Statistics and Computing (2025) 35:52
https://doi.org/10.1007/s11222-025-10562-5

ORIGINAL PAPER

Inference on diffusion processes related to a general growth model

Giuseppina Albano[1] · Antonio Barrera[2,5] · Virginia Giorno[3] · Francisco Torres-Ruiz[4,5]

Received: 12 July 2024 / Accepted: 3 January 2025 / Published online: 20 February 2025

Abstract
This paper considers two stochastic diffusion processes associated with a general growth curve that includes a wide family of growth phenomena. The resulting processes are lognormal and Gaussian, and for them inference is addressed by means of the maximum likelihood method. The complexity of the resulting system of equations requires the use of metaheuristic techniques. The limitation of the parameter space, typically required by all metaheuristic techniques, is also provided by means of a suitable strategy. Several simulation studies are performed to evaluate to goodness of the proposed methodology, and an application to real data is described.

Keywords Lognormal process · Gaussian process · Estimation · Moth-flame optimization algorithm

1 Introduction

The construction and development of mathematical models to describe dynamic phenomena associated with growth curves are topics that have caught the eyes of many researchers for a long time. As a result, the literature on this type of model has become so extensive that we can stated that the development of growth theory is, in itself, a growth phenomenon.

It is well known that the origins of such models can be found in the studies on human population growth conducted by Malthus. However, the Malthusian model does not consider restrictions on the evolution of growth, which makes it unsuitable for many practical situations. The introduction of regulatory effects in the model gave rise to alternative formulations, among which are particularly salient the logistic and Gompertz curves. There is no doubt that these models owe their success to their usefulness in a wide variety of fields of application (see, for example, Banks (1994) where many traditional growth curves are presented together with real examples of application).

Despite their applicability, classical models tend to be slightly rigid in their behavior, which has led researchers to study generalizations able to describe growth curves with greater flexibility. Various techniques have been used to provide such generalizations. One of them is based on the inclusion of parameters in classical curves. In this regard, we may cite Von Bertalanffy's curve, in turn generalized by the Richards curve, which is an extension of the logistic one. The introduction of these models made it possible to expand the range of fields of application. For example, the Bertalanffy curve has been applied in ecology to model the growth of fish species (Flinn and Midway 2021), or in paleontology to model the sclerochronological parameters of shell growth (Moss et al. 2021). As for the Richards curve, it has been widely used to study the evolution of pandemics

Giuseppina Albano, Antonio Barrera, Virginia Giorno and Francisco Torres-Ruiz have contributed equally to this work.

✉ Francisco Torres-Ruiz
fdeasis@ugr.es

Giuseppina Albano
pialbano@unisa.it

Antonio Barrera
antonio.barrera@uma.es

Virginia Giorno
giorno@unisa.it

[1] Department of Political and Social Studies, University of Salerno, Via Giovanni Paolo II, 84084 Fisciano, SA, Italy

[2] Department of Mathematical Analysis, Statistics and Operational Research and Applied Mathematics, University of Málaga, Bulevar Louis Pasteur 31, 29010 Málaga, Spain

[3] Department of Computer Science, University of Salerno, Via Giovanni Paolo II, 84084 Fisciano, SA, Italy

[4] Department of Statistics and Operational Research, University of Granada, Avda. Fuentenueva s/n, 18071 Granada, Spain

[5] Institute of Mathematics (IMAG), University of Granada, C/ Ventanilla, 11, 18001 Granada, Spain

(Hsieh et al. 2010) or to describe the growth and development of forest ecosystems (Protazio et al. 2022). The methodology of including new parameters in the classical growth models led Turner et al. (1976) to propose a general theory of growth based on general postulates. Subsequent studies have been carried out by other authors, such as Tsoularis and Wallace (2002), Koya and Goshu (2013) or, more recently, Albano et al. (2022).

It must be acknowledged that the resulting models, while offering greater flexibility, easily become overly complex. For this reason, generalization can otherwise be carried out through the introduction of functions with flexible behaviors. Such is the case of hyperbolastic curves (Tabatatai et al. 2005), oscilobolastic curves (Eby and Tabatabai 2014) or various multisigmoidal type models (Di Crescenzo et al. 2021; Román-Román et al. 2019). A particularly interesting version of this methodology appears when the constant growth rate is replaced by a function of time in the Malthusian equation, which governs the exponential curve. This leads to curves such as the Gompertz (Gutiérrez et al. 2007), the Korf-type curve (Di Crescenzo and Spina 2016) or a mixed version of both of them (Bhowmick and Bhattacharya 2014). On the other hand, convenient reformulations and reparametrizations have allowed this idea to be applied to other models, as in the case of hyperbolastic curves (Barrera et al. 2021).

The fact that deterministic models cannot control the intrinsic variability of phenomena led to the introduction of the randomization into the ordinary differential equations (ODE) that govern such models. This in turn led to considering stochastic differential equations (SDE) whose solutions are, under certain conditions, diffusion processes. Undoubtedly, the extensions of the logistic and Gompertz cases have caught the most attention, being today the base models for the study of current phenomena. In this line we can mention, within the logistic case, the works of Rajasekar and Pitchaimani (2020), Rajasekar et al. (2020) in which the authors analyzed a stochastic version of the SIR models for the spread of the COVID-19 pandemic. Regarding the Gompertz case, the evolution of tumors has been the main object of study. The work of Lo (2010) is a good example, from which several modifications have been proposed with the aim of describing the evolution of tumor growth in the presence of therapeutic treatments (see Albano et al. 2013, 2020 and references therein). Sadly though, the literature on diffusion processes associated with other growth curves is not as extensive. Perhaps the reason why is to be found in the fact that the associated SDEs do not generally have an explicit solution or they give rise to processes that are difficult to handle. However, reformulations of some curves have led to SDEs having solutions that are suitable non-homogeneous lognormal diffusion processes, whose distributions are known (as, for instance, Di Crescenzo et al. (2021), Román-Román and Torres-Ruiz (2015), Albano et al. (2023)).

Based on the fact that the model studied in Tsoularis and Wallace (2002) and Albano et al. (2022) includes a great variety of growth curves, in Albano et al. (2023) two diffusion processes were provided with the goal of having common stochastic models for the study of a wide variety of growth phenomena. Specifically, in this last paper the authors focus on first-passage-time problems through one or two boundaries which are, generally, time dependent.

The main objective of the present paper is to make inference for the two diffusion processes introduced in Albano et al. (2023). Parameter inference for diffusion processes has been solved successfully by using a variety of approaches in several fields of application such as finance, economics and biology. We would like to point out that, even though diffusion processes are continuous-time stochastic models, their observations are only made at discrete time points, for example at n equally spaced time instants (see, for instance, Lo (1988), Aït-Sahalia (2003), Fan (2005) for a discussion and overview of the estimation of diffusion processes based on discrete observations). Parameter inference for diffusion processes is generally based on the method of moments or maximum likelihood approach. The methodology based on the method of moments presents a more flexible framework than maximum likelihood methods, since it does not assume knowledge of the transition density of the process. However, approximations must be derived from discrete schemes such as Euler-Maruyama, which can compromise the quality of the final results. Moreover, the problem of parameter identification based on sample moments is not always satisfactorily resolved. As discussed in López-Pérez et al. (2021), these procedures have poor finite sample properties. In certain cases, estimates are subject to large finite sample biases. This can happen when the functions that match the true parameters with the data provide weak parameter identification. This is usually the case for highly persistent time series, in which the value of the variable at a given date is closely related to the previous value. Concerning the method of maximum likelihood estimation, it provides parameter estimators that have good properties such as consistency, efficiency and asymptotic unbiasedness and normality. (cf. Dacunha-Castelle and Florens-Zmirou 1986).

Starting with a discrete sampling of the process trajectories, the likelihood function is formulated from the initial distribution of the process and the transition probability distributions. When the transition distributions are not explicitly known, approximations of them must be used Pedersen (1995). In any case, determining the maximum of the likelihood function requires solving a system of n equations, where n is the dimension of the parametric space. Such a system can be solved analytically only in a few particular cases, so that it becomes necessary to use numerical methods, for instance Newton–Raphson's. One cannot guarantee convergence of this method, especially when the Hessian matrix is

ill-conditioned, which can lead to unstable solutions (cf., also Román-Román and Torres-Ruiz 2015). Additionally, these methods require initial solutions that are not always easy to obtain. Another way to solve the maximum-likelihood estimation is based on the use of metaheuristic algorithms. Many such algorithms employ heuristic strategies based on evolving random procedures to find the local and global minima (or maxima) of a function within a bounded region of the parametric space.

In this paper, and in order to make inference for our processes, we use a procedure based on the Moth Flame Optimizer (MFO) method. Such procedure is the result of combining the MFO algorithm with the Box-Cox transformation for linear models with the goal of finding the maximum of the loglikelihood function in relation to the parameters that will be estimated.

The present paper is structured as follows: Sect. 2 presents a review of the general growth model treated in Albano et al. (2022); Sect. 3 presents the two diffusion processes associated with the general growth curve. They are essentially obtained by introducing in the ODE a multiplicative and an additive noises; Sect. 4 focuses on the inference of the models under maximum likelihood from discrete sampling. The complexity of the likelihood equations discourages the use of numerical methods for their resolution, therefore the use of metaheuristic optimization methods is proposed. In Sect. 5, several simulation studies for both processes are analyzed in order to evaluate the goodness of the proposed procedure. Finally, an application to real data is provided in Sect. 6.

2 A review of the deterministic general growth model

Turner et al. (1976), established a general theory of growth, based on three rather general postulates, that led to an ODE which solution includes well-known growth curves as exponential, logistic, Gompertz, Richards-Bertalanffy, among others. Later, Tsoularis and Wallace (2002) reconsidered this expression, interpreting it as a generalization of the classical logistic model and analyzing the great variety of curves that this general version includes.

For such a model, the positive real function $x(t)$, describing the population size at time t, is defined as the solution of the ODE

$$\frac{d}{dt}x(t) = \gamma\, k^{n(p-1)} x(t)^{1+n(1-p)} \left(1 - \left(\frac{x(t)}{k}\right)^n\right)^p, \quad t \geq t_0 \geq 0, \tag{1}$$

with initial condition $x(t_0) = x_0$. Here, $k = \lim_{t\to\infty} x(t) > 0$ represents the carrying capacity, while γ, n and p are real and positive shape parameters verifying $0 < p < 1 + 1/n$.

For different values of p and n, Eq. (1) becomes the equation associated to classical growth curves. For instance, for $p = 1$ the Bertalanffy-Richards model is obtained, from which the logistic model follows when $n = 1$. Moreover, if $k \to \infty$, the general growth model goes to the exponential (or Malthusian) one. Generalized versions of Gompertz and logistic models can also be derived from (1); such is the case of the hyper-Gompertz model which is obtained by taking $n \to 0$ and assuming $\lim_{n\to 0} \gamma\, n^p = \gamma'$. In this case it holds

$$\frac{d}{dt}x(t) = \gamma' x(t) \left(\log\frac{k}{x(t)}\right)^p, \quad x(t_0) = x_0. \tag{2}$$

On the other hand, hyper-logistic model is derived by taking $n = 1$ in (1), resulting in

$$\frac{d}{dt}x(t) = \gamma\, k^{p-1} x(t)^{2-p} \left(1 - \frac{x(t)}{k}\right)^p, \quad x(t_0) = x_0. \tag{3}$$

This last expression is a particular case of that introduced originally by Blumberg (1968) (see Tsoularis and Wallace 2002 for details). Note that, by choosing $p = 1$, Eqs. (2) and (3) become the classical Gompertz and logistic curves, respectively.

By considering the initial condition $x(t_0) = x_0$, the solution of (1) is

$$x(t) = k\left(1 + \left(\gamma\, n(p-1)(t - t_0) + A_n^{1-p}\right)^{\frac{1}{1-p}}\right)^{-\frac{1}{n}}, \tag{4}$$

where $A_n = (k/x_0)^n - 1$ is a term depending on the shape parameter n, and the ratio of the carrying capacity and the initial population size.

The stochastic extension of the general growth model is usually established by introducing randomness in (1), thus giving rise to an SDE. Unfortunately, in several cases the explicit solution of this SDE is not available. In other cases, even if the solution exists in closed-form, the probability distribution of the associated stochastic process is not obtainable in terms of finite-dimensional distributions or transition probability distributions. These problems can be solved making use of an adequate reparametrization of (4). Precisely, it can be rewritten as

$$x_{\boldsymbol{\theta}}(t) \equiv x(t) = x_0 \frac{g_{\boldsymbol{\theta}}(t_0)}{g_{\boldsymbol{\theta}}(t)}, \tag{5}$$

for $t \geq t_0 \geq 0$ and where

$$g_{\boldsymbol{\theta}}(t) = \left(\eta + \left(1 + \eta^{1-p}\log\alpha(1-p)t\right)^{\frac{1}{1-p}}\right)^{\frac{1}{\eta}}, \tag{6}$$

defined for a vector of parameters $\boldsymbol{\theta} = (\eta, \alpha, p, n)^T$, with $\eta, n > 0$, $0 < \alpha < 1$ and $0 < p < 1 + 1/n$. Equations (5) and (6) are obtained from (4) by setting $\alpha = e^{-\gamma n}$ and $\eta = \left(A_n^{1-p} + n\,\gamma\,(1-p)\,t_0\right)^{-1/(1-p)}$.

By deriving (5) with respect to t, we obtain

$$\frac{dx_{\boldsymbol{\theta}}(t)}{dt} = -x_0 g_{\boldsymbol{\theta}}(t_0)\frac{g'_{\boldsymbol{\theta}}(t)}{g^2_{\boldsymbol{\theta}}(t)} = h_{\boldsymbol{\theta}}(t)x_{\boldsymbol{\theta}}(t), \tag{7}$$

where

$$\begin{aligned} h_{\boldsymbol{\theta}}(t) &= -\frac{d}{dt}\log g_{\boldsymbol{\theta}}(t) \\ &= -\eta^{1-p}\log\alpha\left(1 + (1-p)\,t\,\eta^{1-p}\log\alpha\right)^{p/(1-p)} \\ &\quad \times \left(n\,g_{\boldsymbol{\theta}}(t)^n\right)^{-1}. \end{aligned} \tag{8}$$

We should point out that Eq. (7) represents a Malthusian growth with time dependent fertility. By introducing randomization we are able to present, in the next section, two diffusion processes with means equal to (5) and with manageable probability distribution functions (cf. Albano et al. 2023).

In Albano et al. (2022), an extensive study was carried out to show the behavior of the curve taking into account the values of n and p. This study leads to the consideration of three cases:

1. the case when $1 < p < 1 + 1/n$;
2. the case when $p = 1$;
3. the case when $0 < p < 1$.

The study of monotonicity and inflection points of the curve is important in the context of growth phenomena. Indeed, it is of interest to find the time instants at which the model reaches some proportion of the carrying capacity, or when the model dramatically changes its tendency. In Albano et al. (2022), it is shown that cases 1 and 2 are characterized by a monotonically strictly increasing behavior, exhibiting one inflection point at positions that are proportional to the parameter p. In contrast, case 3 shows different and interesting growth patterns.

In the present paper we will focus the attention on case 1, which is related to a large number of growth phenomena obeying a bounded sigmoidal pattern. For this case, the upper limit of the curve is

$$k_{\boldsymbol{\theta}} = x_0\frac{g_{\boldsymbol{\theta}}(t_0)}{\eta^{1/n}},$$

whereas the time instant in which the inflection is reached is given by

$$t_{\text{inf}} = \frac{\left(\dfrac{np}{1+n(1-p)}\right)^{1-p} - \eta^{p-1}}{(1-p)\log\alpha}, \tag{9}$$

verifying

$$x_{\boldsymbol{\theta}}(t_{\text{inf}}) = k_{\boldsymbol{\theta}}\left(1 - \frac{np}{n+1}\right)^{1/n}.$$

The previous expressions, in particular that of the inflection time, allow to establish -under certain conditions- limitations for some parameters of the models, as we will show later in this paper.

3 Stochastic models derived from the growth curve

The use of deterministic models seems to be appropriate in some practical cases due to their technical simplicity. Nevertheless, they cannot capture all the information provided by the observations. Uncertainties must be considered in order to get a more realistic model. The reasons of such uncertainties are multiple: they could appear, among others, in the form of hidden endogenous of exogenous factors, tiny discrepancies in the measurements or the observations, or even in the pure nature of a random behavior.

The theory of stochastic processes has been developed to a great extent in the last decades, allowing to use sophisticated tools to describe random behavior in a more accurate way. In addition to this, the rapid growth of computational power and network capabilities has broken the barriers between the theory and the practice.

Numerous stochastic models have been introduced as a random extension of well-known deterministic ones, among them diffusion processes. Some of these diffusion models emerge as solutions to a SDE after modifying a deterministic one by introducing in it a term of white noise. Other diffusion processes are constructed in such a way that their mean function is a certain sigmoidal growth curve.

Based on the reparametrization given by (5) and (6), in this section we introduce two diffusion processes with mean functions equal to the general growth curve and with known probability distributions. Said processes are the solution of the SDE resulting from including in the ODE (7) a multiplicative and an additive noise, respectively.

3.1 Multiplicative noise: lognormal-type diffusion model

Let us consider the ODE (7), where the fertility rate $h_{\boldsymbol{\theta}}(t)$ is given by Eq. (8). By including a white noise $\varphi(t)$ with variance σ_{L}^2, and then replacing $h_{\boldsymbol{\theta}}(t)$ with $h_{\boldsymbol{\theta}}(t) + \varphi(t)$, the

population size $x(t)$ is generalized to the one-dimensional stochastic process $X^{\mathrm{L}}(t)$, satisfying the following SDE:

$$dX^{\mathrm{L}}(t) = h_{\boldsymbol{\theta}}(t)X^{\mathrm{L}}(t)\,dt + \sigma_{\mathrm{L}}\,X^{\mathrm{L}}(t)\,dW(t), \quad t \geq t_0.$$

Here, $\{W(t), t \geq t_0\}$ is a Wiener process on the probability space (Ω, F, P), equipped with the natural filtration $(F_t)_{t\geq t_0}$, where $F_t = \sigma(W(s),\ t_0 \leq s \leq t)$ is the sigma-algebra generated by the variables $W(s)$, $t_0 \leq s \leq t$. We assume that the initial condition $X^{\mathrm{L}}(t_0) = X_0^{\mathrm{L}}$ is independent of the Brownian motion $W(t)$ for $t \geq t_0$.

The solution of this equation is a non-homogeneous lognormal diffusion process, characterized by drift $h_{\boldsymbol{\theta}}(t)x$ and diffusion coefficient $\sigma_{\mathrm{L}}^2 x^2$ (see Román-Román et al. 2018 and references therein for details) and with explicit expression given by

$$X^{\mathrm{L}}(t) = X_0^{\mathrm{L}}\,\exp\left[H_{\boldsymbol{\theta}}(t_0, t) + \sigma_{\mathrm{L}}\,(W(t) - W(t_0))\right],$$

where

$$\begin{aligned} H_{\boldsymbol{\xi}_{\mathrm{L}}}(s,t) &= \int_s^t h_{\boldsymbol{\theta}}(u)du - \frac{\sigma_{\mathrm{L}}^2}{2}(t-s) \\ &= \log\frac{g_{\boldsymbol{\theta}}(s)}{g_{\boldsymbol{\theta}}(t)} - \frac{\sigma_{\mathrm{L}}^2}{2}(t-s), \end{aligned}$$

$\boldsymbol{\xi}_{\mathrm{L}} = (\boldsymbol{\theta}^T, \sigma_{\mathrm{L}}^2)^T$ and function $g_{\boldsymbol{\theta}}(t)$ is given in (6).

In addition, if the initial distribution is lognormal, i.e. $X_0^{\mathrm{L}} \sim \Lambda_1\left[\mu_0, \sigma_0^2\right]$ for some parameters μ_0 and σ_0^2, or degenerate at a point x_0, that is $P(X_{\mathrm{L}}(t_0) = x_0) = 1$, the finite-dimensional distributions of the process are lognormal. Concretely, $\forall n \in \mathbb{N}$ and $t_1 < \cdots < t_n$, vector $(X^{\mathrm{L}}(t_1), \ldots, X^{\mathrm{L}}(t_n))^T$ has a n-dimensional lognormal joint distribution $\Lambda_n[\boldsymbol{\varepsilon}, \boldsymbol{\Sigma}]$, where the components of vector $\boldsymbol{\varepsilon}$ and of matrix $\boldsymbol{\Sigma}$ are

$$\varepsilon_i = \mu_0 + H_{\boldsymbol{\xi}_{\mathrm{L}}}(t_0, t_i), \quad i = 1, \ldots, n$$

and

$$\sigma_{ij} = \sigma_0^2 + \sigma_{\mathrm{L}}^2(\min(t_i, t_j) - t_0), \quad i, j = 1, \ldots, n,$$

respectively. From the joint distribution of $(X^{\mathrm{L}}(s), X^{\mathrm{L}}(t))^T$, $s < t$, the transition probability distribution, related to the variable $X^{\mathrm{L}}(t) \mid X^{\mathrm{L}}(s) = y$, $y > 0$, can be obtained. Further, from the properties of the multidimensional lognormal distribution, it follows that $X^{\mathrm{L}}(t) \mid X^{\mathrm{L}}(s) = y$ has a one-dimensional lognormal distribution $\Lambda_1\left[\log y + H_{\boldsymbol{\xi}_{\mathrm{L}}}(s,t),\ \sigma_{\mathrm{L}}^2(t-s)\right]$, so the transition probability density function, denoted by $f_{\boldsymbol{\xi}_{\mathrm{L}}}(x, t|y, s)$, is given by

$$\begin{aligned} f_{\boldsymbol{\xi}_{\mathrm{L}}}(x,t|y,s) &= \frac{1}{x\sqrt{2\pi\sigma_{\mathrm{L}}^2(t-s)}} \\ &\quad \times \exp\left(-\frac{\left[\log(x/y) - H_{\boldsymbol{\xi}_{\mathrm{L}}}(s,t)\right]^2}{2\sigma_{\mathrm{L}}^2(t-s)}\right). \end{aligned} \tag{10}$$

We point out that, from the one-dimensional marginal distributions and from the transition distributions, it is verified that the mean function and the conditional mean function are

$$m_{\boldsymbol{\theta}}^{\mathrm{L}}(t) = E[X^{\mathrm{L}}(t)] = E[X_0^{\mathrm{L}}]\frac{g_{\boldsymbol{\theta}}(t_0)}{g_{\boldsymbol{\theta}}(t)} \quad \text{and}$$

$$m_{\boldsymbol{\theta}}(t|s) = E[X^{\mathrm{L}}(t)|X^{\mathrm{L}}(t_0) = x_s] = x_s\frac{g_{\boldsymbol{\theta}}(t_0)}{g_{\boldsymbol{\theta}}(t)},$$

respectively, and they have the same shape of the deterministic curve given by (5).

3.2 Additive noise: Gaussian diffusion model

In this subsection we start again from the ODE (7) and obtain a stochastic generalization of it by introducing a white noise with variance σ_{G}^2 in additive form, where $\sigma_{\mathrm{G}} > 0$ represents the amplitude of the random fluctuations. In this way, $x(t)$ becomes the stochastic process $X^{\mathrm{G}}(t)$ satisfying the following SDE:

$$dX^{\mathrm{G}}(t) = h_{\boldsymbol{\theta}}(t)X^{\mathrm{G}}(t)\,dt + \sigma_{\mathrm{G}}\,dW(t), \quad t \geq t_0,$$

with initial condition $X^{\mathrm{G}}(t_0) = X_0^{\mathrm{G}}$. This is a linear SDE whose solution is assured by virtue of the continuity of the function $h_{\boldsymbol{\theta}}(t)$, a property that also ensures that the solution is a diffusion process with infinitesimal mean $h_{\boldsymbol{\theta}}(t)x$ and infinitesimal variance σ_{G}^2. Specifically, the solution is (cf. Albano et al. 2023)

$$X^{\mathrm{G}}(t) = \Phi(t)\left[X_0^{\mathrm{G}} + \sigma_{\mathrm{G}}\int_{t_0}^t \Phi(s)^{-1}\,dW(s)\right]$$

where

$$\Phi(t) = \exp\left(\int_{t_0}^t h_{\boldsymbol{\theta}}(s)\,ds\right) = \frac{g_{\boldsymbol{\theta}}(t_0)}{g_{\boldsymbol{\theta}}(t)}.$$

Hence, we have

$$X^{\mathrm{G}}(t) = \frac{g_{\boldsymbol{\theta}}(t_0)}{g_{\boldsymbol{\theta}}(t)}\left(X_0^{\mathrm{G}} + \sigma_{\mathrm{G}}\int_{t_0}^t \frac{g_{\boldsymbol{\theta}}(s)}{g_{\boldsymbol{\theta}}(t_0)}\,dW(s)\right).$$

Regarding the distribution of the process, in the case in which the initial distribution is normal or degenerate, $X^{\mathrm{G}}(t)$ is a Gaussian process with mean function given by

$$m^{\mathrm{G}}_{\boldsymbol{\theta}}(t) = E[X^{\mathrm{G}}_0]\,\frac{g_{\boldsymbol{\theta}}(t_0)}{g_{\boldsymbol{\theta}}(t)},$$

and covariance function

$$K_{\boldsymbol{\xi}_{\mathrm{G}}}(s,t) = \frac{g^2_{\boldsymbol{\theta}}(t_0)}{g_{\boldsymbol{\theta}}(s)g_{\boldsymbol{\theta}}(t)}\left(\mathrm{Var}X^{\mathrm{G}}_0 + \sigma^2_{\mathrm{G}}\int_{t_0}^{s\wedge t}\frac{g^2_{\boldsymbol{\theta}}(u)}{g^2_{\boldsymbol{\theta}}(t_0)}du\right),$$
$$\boldsymbol{\xi}_{\mathrm{G}} = (\boldsymbol{\theta}^T, \sigma^2_{\mathrm{G}})^T$$

where $s\wedge t = \min(s,t)$. Finally, since the finite-dimensional distributions are Gaussian, we have that $X^{\mathrm{G}}(t)|X^{\mathrm{G}}(s) = x_s$ is normally distributed with mean and variance

$$m_{\boldsymbol{\theta}}(t|s) = \frac{g_{\boldsymbol{\theta}}(s)}{g_{\boldsymbol{\theta}}(t)}\,x_s \quad \text{and} \quad \sigma^2_{\boldsymbol{\xi}_{\mathrm{G}}}(t|s) = \frac{\sigma^2_{\mathrm{G}}}{g^2_{\boldsymbol{\theta}}(t)}\int_s^t g^2_{\boldsymbol{\theta}}(u)\,du,$$

respectively. So, the transition probability density function is

$$f_{\boldsymbol{\xi}_{\mathrm{G}}}(x,t|x_s,s) = \frac{1}{\sqrt{2\pi\sigma^2_{\boldsymbol{\xi}_{\mathrm{G}}}(t|s)}}\exp\left(-\frac{[x - m_{\boldsymbol{\theta}}(t|s)]^2}{2\sigma^2_{\boldsymbol{\xi}_{\mathrm{G}}}(t|s)}\right). \tag{11}$$

Again, we can see how both the mean function and the conditional mean function are of the type given in (5).

4 Inference

One of the main goals of introducing these diffusion processes is to develop a tool able to fit growth patterns to real cases, and to this end we address the estimation of the parameters of the model by means of the maximum likelihood method.

Although the literature in the field of lognormal and Gaussian processes is very extensive (cf., e.g., Gutiérrez et al. 2003; Bibby et al. 2009; Tang and Chen 2009; Albano and Giorno 2020), the classical methods are not applicable for the proposed model because function $h_{\boldsymbol{\theta}}(t)$, given in (8), presents a complex expression involving four non-linearly related parameters. Furthermore, there are restrictions on the parameters (e.g., $0 < p < 1 + 1/n$). To these parameters we must add the one associated with the volatility of the model, which increases the complexity of the problem. However, the fact that the functional form of the transitions is available allows for directly addressing the maximum likelihood estimation of the parameters from the discrete sampling of trajectories.

The starting point is the observation of d sample-paths at discrete time instants. It is noticed that the length of each sample-path may vary, and that the observation time instants of each one may be different. Consequently, let us denote by $t_{i,j}$ the j-th observation time instant for the i-th sample-path, whose length is k_i, $(i = 1,\ldots,d,\ j = 1,\ldots,k_i)$. However, since the first observation of each sample-path must come from a common initial distribution, we will consider that the first time of observation is common for all sample-paths, that is, $t_{i,1} = t_0$, $i = 1,\ldots,d$. Let $X(t_{i,j})$ be the random variable of the process associated with $t_{i,j}$ and let $x_{i,j}$ be the observed sample value. For $i = 1,\ldots,d$, let $\mathbf{X}_i = (X(t_{i,1}),\ldots,X(t_{i,k_i}))^T$ be the vector containing the random variables of the i-th sample-path, and let $\mathbf{X} = \left(\mathbf{X}_1^T|\cdots|\mathbf{X}_d^T\right)^T$.

For the two processes considered, we will assume that the initial distribution is non-degenerate, being lognormal $\Lambda_1\left[\eta_{\mathrm{L}}, \delta^2_{\mathrm{L}}\right]$ for $X^{\mathrm{L}}(t)$ and normal $N_1\left[\eta_{\mathrm{G}}, \delta^2_{\mathrm{G}}\right]$ for $X^{\mathrm{G}}(t)$. Clearly, the degenerate initial distribution is a particular case of both and can be obtained by choosing $\delta^2_{\mathrm{L}} = \delta^2_{\mathrm{G}} = 0$. Let's denote $\boldsymbol{\varsigma}_\nu = (\eta_\nu, \delta^2_\nu)^T$, $\nu = \mathrm{L}, \mathrm{G}$, the parametric vectors of the initial distributions.

For a fixed value $\mathbf{x}$ of $\mathbf{X}$, from (10) and (11), and from the initial distributions, we can write the expression of the logarithm of the likelihood function jointly for both processes, resulting in

$$\log \mathbb{L}^\nu_{\mathbf{x}}\left(\boldsymbol{\varsigma}_\nu, \boldsymbol{\xi}_\nu\right) = \begin{cases} \mathcal{J}^\nu_{\mathbf{x}}(\boldsymbol{\xi}_\nu) & \text{if } P(X_\nu(t_0) = x_0) = 1 \\ \mathcal{H}^\nu_{\mathbf{x}}(\boldsymbol{\varsigma}_\nu) + \mathcal{J}^\nu_{\mathbf{x}}(\boldsymbol{\xi}_\nu) & \text{otherwise}, \end{cases} \qquad \nu = \mathrm{L}, \mathrm{G} \tag{12}$$

being

$$\mathcal{H}^{\nu}_{\mathbf{x}}(\boldsymbol{\varsigma}_{\nu}) = \begin{cases} -\dfrac{k+d}{2}\log(2\pi) - \dfrac{d}{2}\log\delta^2_{\mathrm{L}} - \displaystyle\sum_{i=1}^{d}\log x_{i,1} - \dfrac{1}{2\delta^2_{\mathrm{L}}}\sum_{i=1}^{d}\left(\log x_{i,1} - \eta_{\mathrm{L}}\right)^2 & \text{if } \nu = \mathrm{L}, \\ -\dfrac{k+d}{2}\log(2\pi) - \dfrac{d}{2}\log\delta^2_{\mathrm{G}} - \dfrac{1}{2\delta^2_{\mathrm{G}}}\displaystyle\sum_{i=1}^{d}\left(x_{i,1} - \eta_{\mathrm{G}}\right)^2 & \text{if } \nu = \mathrm{G}, \end{cases} \tag{13}$$

and

$$\mathcal{J}^{\nu}_{\mathbf{x}}(\boldsymbol{\xi}_{\nu}) = \begin{cases} -\dfrac{k}{2}\log\sigma^2_{\mathrm{L}} - \dfrac{1}{2}\displaystyle\sum_{i=1}^{d}\sum_{j=2}^{k_i}\log\Delta_{ij} - \dfrac{Z + \Phi_{\boldsymbol{\xi}_{\mathrm{L}}} - 2\Gamma_{\boldsymbol{\xi}_{\mathrm{L}}}}{2\sigma^2_{\mathrm{L}}} & \text{if } \nu = \mathrm{L}, \\ -\dfrac{k}{2}\log\sigma^2_{\mathrm{G}} - \dfrac{1}{2}\displaystyle\sum_{i=1}^{d}\sum_{j=2}^{k_i}\log\tilde{g}^{ij}_{\boldsymbol{\theta}} - \dfrac{\tilde{Z}_{\boldsymbol{\theta}} + \tilde{\Phi}_{\boldsymbol{\theta}} - 2\tilde{\Gamma}_{\boldsymbol{\theta}}}{2\sigma^2_{\mathrm{G}}} & \text{if } \nu = \mathrm{G}. \end{cases} \tag{14}$$

Here $k = \sum_{i=1}^{d}(k_i - 1)$ and

$$Z = \sum_{i=1}^{d}\sum_{j=2}^{k_i}\frac{\log^2\left(x_{i,j}/x_{i,j-1}\right)}{\Delta_{ij}},$$

$$\Phi_{\boldsymbol{\xi}_{\mathrm{L}}} = \sum_{i=1}^{d}\sum_{j=2}^{k_i}\frac{\left(m^{ij}_{\boldsymbol{\xi}_{\mathrm{L}}}\right)^2}{\Delta_{ij}},$$

$$\Gamma_{\boldsymbol{\xi}_{\mathrm{L}}} = \sum_{i=1}^{d}\sum_{j=2}^{k_i}\frac{\log\left(x_{i,j}/x_{i,j-1}\right)m^{ij}_{\boldsymbol{\xi}_{\mathrm{L}}}}{\Delta_{ij}},$$

$$\tilde{Z}_{\boldsymbol{\theta}} = \sum_{i=1}^{d}\sum_{j=2}^{k_i}\frac{x^2_{i,j}}{\tilde{g}^{ij}_{\boldsymbol{\theta}}},$$

$$\tilde{\Phi}_{\boldsymbol{\theta}} = \sum_{i=1}^{d}\sum_{j=2}^{k_i}\frac{x^2_{i,j-1}\left(g^{ij}_{\boldsymbol{\theta}}\right)^2}{\tilde{g}^{ij}_{\boldsymbol{\theta}}},$$

$$\tilde{\Gamma}_{\boldsymbol{\theta}} = \sum_{i=1}^{d}\sum_{j=2}^{k_i}\frac{x_{i,j}x_{i,j-1}g^{ij}_{\boldsymbol{\theta}}}{\tilde{g}^{ij}_{\boldsymbol{\theta}}},$$

being $\Delta_{ij} = t_{i,j} - t_{i,j-1}$, $m^{ij}_{\boldsymbol{\xi}_L} = H_{\boldsymbol{\xi}_{\mathrm{L}}}(t_{i,j-1}, t_{i,j})$, $g^{ij}_{\boldsymbol{\theta}} = g_{\boldsymbol{\theta}}(t_{i,j-1})/g_{\boldsymbol{\theta}}(t_{i,j})$ and

$$\tilde{g}^{ij}_{\boldsymbol{\theta}} = \frac{1}{g^2_{\boldsymbol{\theta}}(t_{i,j})}\int_{t_{i,j-1}}^{t_{i,j}} g^2_{\boldsymbol{\theta}}(u)\,du.$$

The maximum likelihood estimation requires determining the maximum of the objective function $\log\mathbb{L}^{\nu}_{\mathbf{x}}\left(\boldsymbol{\varsigma}_{\nu}, \boldsymbol{\xi}_{\nu}\right)$. Assuming that $\boldsymbol{\varsigma}_{\nu}$ and $\boldsymbol{\xi}_{\nu}$ are not functionally dependent, from (13) it follows that their estimation can be obtained in an independent form. In this way, we obtain

$$\widehat{\eta}_{\mathrm{L}} = \frac{1}{d}\sum_{i=1}^{d}\log x_{i,1}, \qquad \widehat{\delta}^2_{\mathrm{L}} = \frac{1}{d}\sum_{i=1}^{d}[\log x_{i,1} - \widehat{\eta}_{\mathrm{L}}]^2$$

and

$$\widehat{\eta}_{\mathrm{G}} = \frac{1}{d}\sum_{i=1}^{d}x_{i,1}, \qquad \widehat{\delta}^2_{\mathrm{G}} = \frac{1}{d}\sum_{i=1}^{d}[x_{i,1} - \widehat{\eta}_{\mathrm{G}}]^2.$$

With regard to $\boldsymbol{\xi}_{\nu}$, the resulting system of equations derived from (14) is quite complex, due to the number of parameters and to the random nature of the observations, so analytical and classical numerical methods (such as Newton–Raphson) are not always exploitable. Indeed, it is not possible to carry out a general study of the system of equations in order to check the conditions of convergence of the chosen numerical method, since the system is dependent on sample data and may lead to unforeseeable behavior. Adding to this problem, we must also select an optimal initial solution, which can be quite a complex task.

In this situation, the use of stochastic metaheuristic optimization procedures is highly recommended. These algorithms are designed to solve problems of the type $\min_{\theta\in\Theta} f(\theta)$, and are often more appropriate than classical numerical methods, since they impose fewer restrictions on the space of solutions Θ and on the analytical properties of the function to be optimized. In this paper, we propose to use an ad hoc strategy that, starting from the likelihood function, uses the MFO algorithm.

4.1 Strategy to find bounds for a metaheuristic approach

Metaheuristic models are used in many areas of applied research, and have shown to be very useful in solving optimization problems, such as those related with maximum likelihood estimation. The common pattern in most metaheuristic algorithms involves the parallel use of many agents moving across a region until certain stability conditions are reached. The region could be the parametric space where a function is defined, and the movements are usually guided by random functions of the value of the function at every point of the region. By taking advantage of modern computational power, a high number of agents are able to "find" the points on the region where the objective function reaches local o global optima in a reasonable amount of time.

There are many metaheuristic algorithms, based on a variety of concepts coming from biology, physics, etc. Indeed, this is a very active area of research, where new and sophisticated procedures are currently being developed in numerous published papers.

One such procedure is the MFO algorithm. This is a population-based metaheuristic algorithm inspired on the navigation method of moths (see Mirjalili 2015). Indeed, such insects fly following what is called transverse orientation, a method based on a fixed angle flight with respect to the Moon. The high rate of success of MFO and other related algorithms based on this natural behavior is illustrated by Kotary and Nanda (2020) and Korashy et al. (2020). The algorithm takes two populations, one of moths and one of flames, which act as flags and represent the matrix of best solutions according to the fitness of the objective function. The moths then search across a parametric space delimited by the boundaries imposed by the researcher.

Mirjalili (2015) provides a wide-ranging presentation and analysis of the method, and supports its use in optimization problems. In particular, a comparative study is made involving some of the most well-known and recent algorithms (Simulated Annealing, Genetic Algorithms, Firefly Algorithm, Particle swarm optimization,...) The results of the study illustrate how MFO provides better results than a wide variety of metaheuristic methods, with the results being

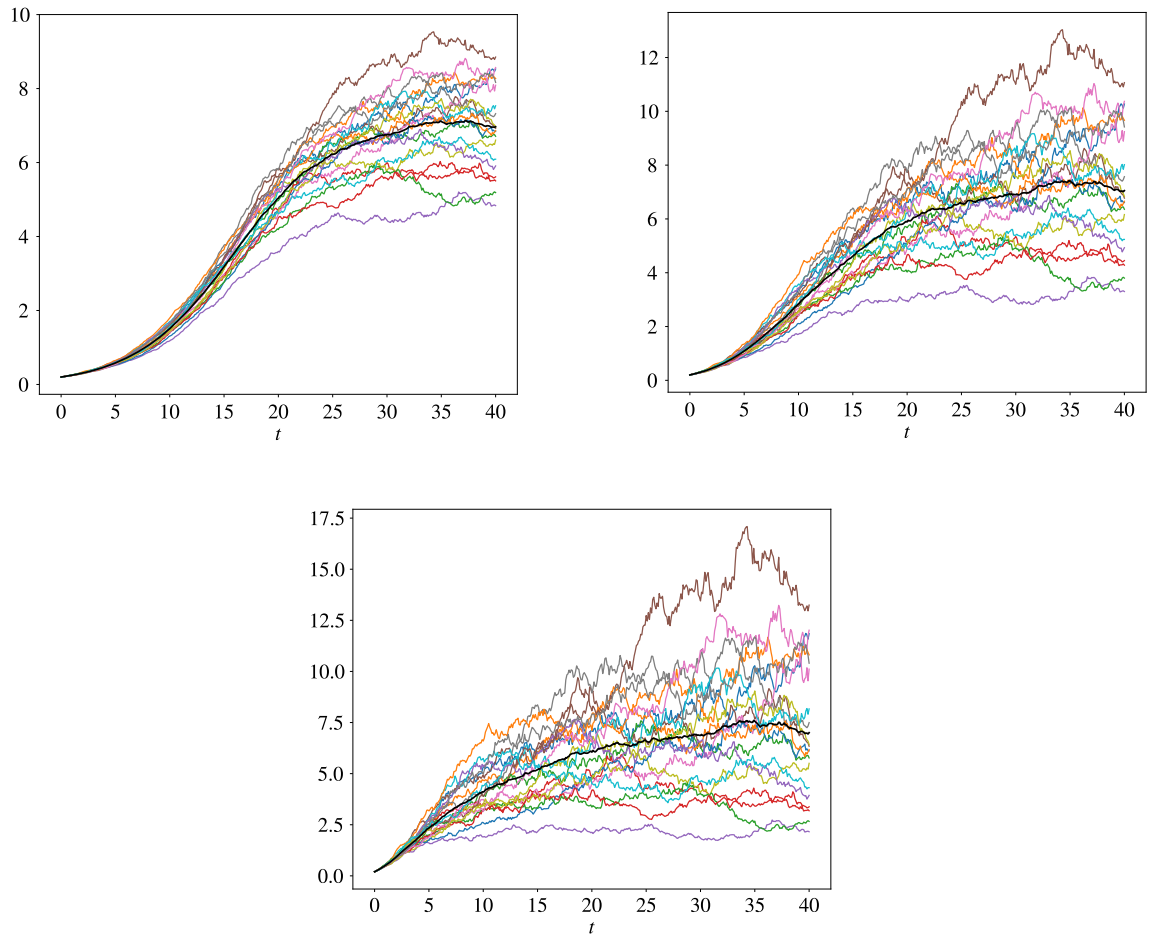

Fig. 1 Simulated paths for lognormal case with $n = 1$ known. Up (left to right): $p = 1.01, \sigma = 0.025$, $p = 1.2, \sigma = 0.05$. Down: $p = 1.5, \sigma = 0.075$

Table 1 Lognormal case with $n = 1$ known: estimations of the parameters and RAE for $\eta = 0.03$, $\alpha = 0.8$ and for different combinations of p, σ

p	σ	$\widehat{\eta}$	$\widehat{\alpha}$	$\widehat{p}$	$\widehat{\sigma}$	RAE
1.01	0.025	0.0289	0.8034	1.0084	0.0252	0.0049
		($< 10^{-4}$)	(0.0002)	(0.0005)	($< 10^{-4}$)	
	0.050	0.0279	0.8043	1.0054	0.0503 ($< 10^{-4}$)	0.0100
		($< 10^{-4}$)	(0.0002)	(0.0004)	($< 10^{-4}$)	
	0.075	0.0269	0.8047	1.0014	0.0755 ($< 10^{-4}$)	0.0152
		($< 10^{-4}$)	(0.0002)	(0.0005)	($< 10^{-4}$)	
1.20	0.025	0.0287	0.8114	1.2172	0.0251	0.0047
		($< 10^{-4}$)	(0.0002)	(0.0004)	($< 10^{-4}$)	
	0.050	0.0273	0.8223	1.2338	0.0503	0.0092
		($< 10^{-4}$)	(0.0002)	(0.0004)	($< 10^{-4}$)	
	0.075	0.0260	0.8325	1.2492	0.0755	0.0141
		($< 10^{-4}$)	(0.0003)	(0.0006)	($< 10^{-4}$)	
1.50	0.025	0.0297	0.8071	1.5127	0.0252	0.0171
		(0.0005)	(0.0040)	(0.0050)	($< 10^{-4}$)	
	0.050	0.0291	0.8168	1.5287	0.0503	0.0302
		(0.0005)	(0.0035)	(0.0045)	($< 10^{-4}$)	
	0.075	0.0284	0.8265	1.5449	0.0755	0.0420
		(0.0005)	(0.0038)	(0.0049)	($< 10^{-4}$)	

Standard errors in parentheses

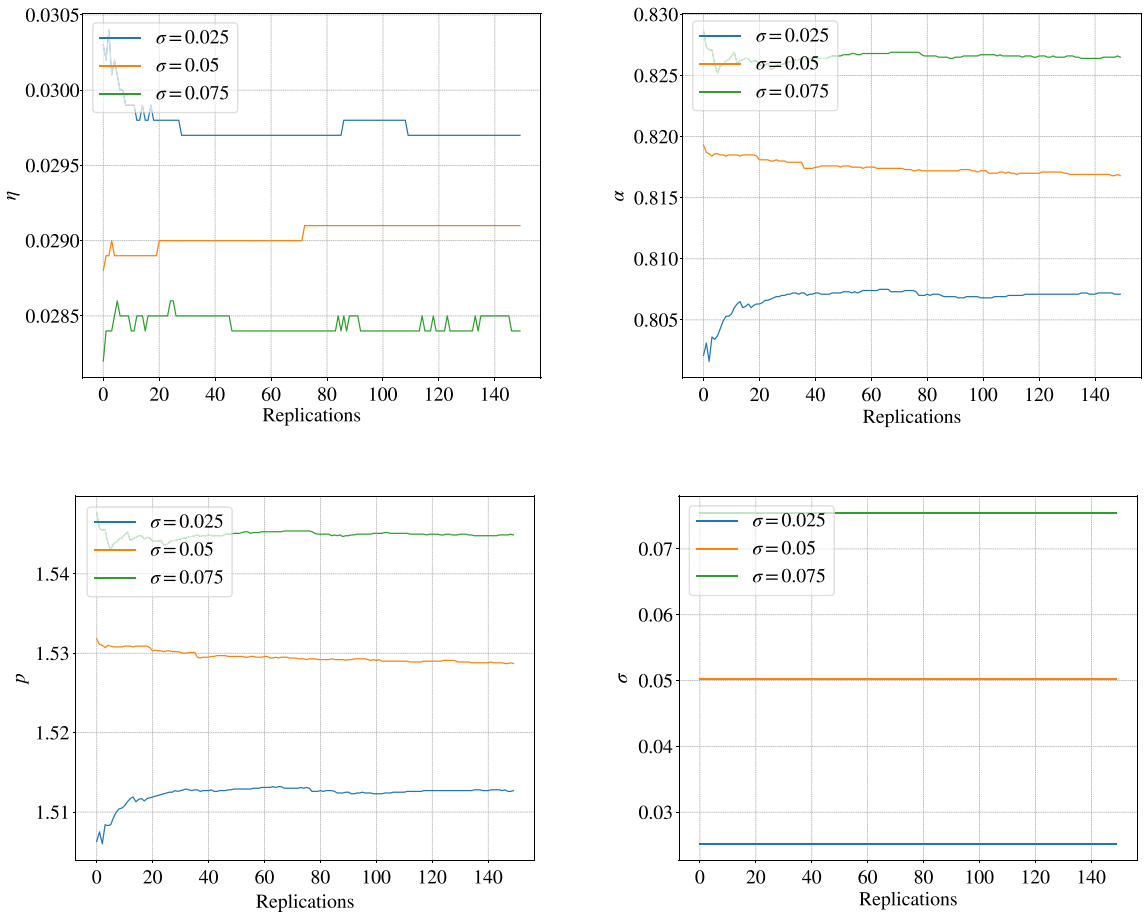

Fig. 2 Lognormal case with $n = 1$ known: evolution of the sequence of the mean of the estimates across replications for $p = 1.5$ and for $\sigma = 0.025, 0.05, 0.075$

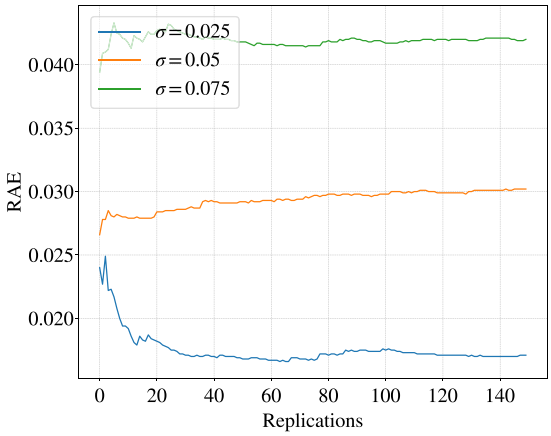

Fig. 3 Lognormal case with $n = 1$ known and $p = 1.5$: RAE across replications for $\sigma = 0.025, 0.05, 0.075$

quite competitive with the rest of the procedures compared. This has motivated us to choose this procedure for the optimization of the objective function.

Concretely, in the MFO algorithm, the configuration of moths can be represented by a matrix $Q = (q_{ij})$, with $i = 1, \ldots, n$ being the number of moths and $j = 1, \ldots, d$ the number of variables. Essentially, the MFO algorithm is determined by three key elements. The first one is a function that provides the random population of moths and the corresponding fitness (a vector in $\mathbb{R}^n$ associated to each configuration matrix Q). Such function generates the random configuration by using any probability distribution and associates fitness by calculating the values of the objective function at the points set by the configuration. The second element of the algorithm is the function in charge of moving the moths around the space. Formally, this function updates matrix Q at each step. Finally, a termination criterion function determines if the algorithm stops or not, according to the state of the matrix Q.

The function for the movement of moths is defined by considering a logarithmic spiral function. As a matter of fact, for a moth i at position q_i and a flame f_j, between which there is a distance d_i, the new position $\tilde{q}_i$ is computed according to the formula $\tilde{q}_i = d_i\, e^{bt} \cos(2\pi t) + f_j$, where b is a constant related with the shape of the spiral and t is taken uniformly at random in $[-1, 1]$.

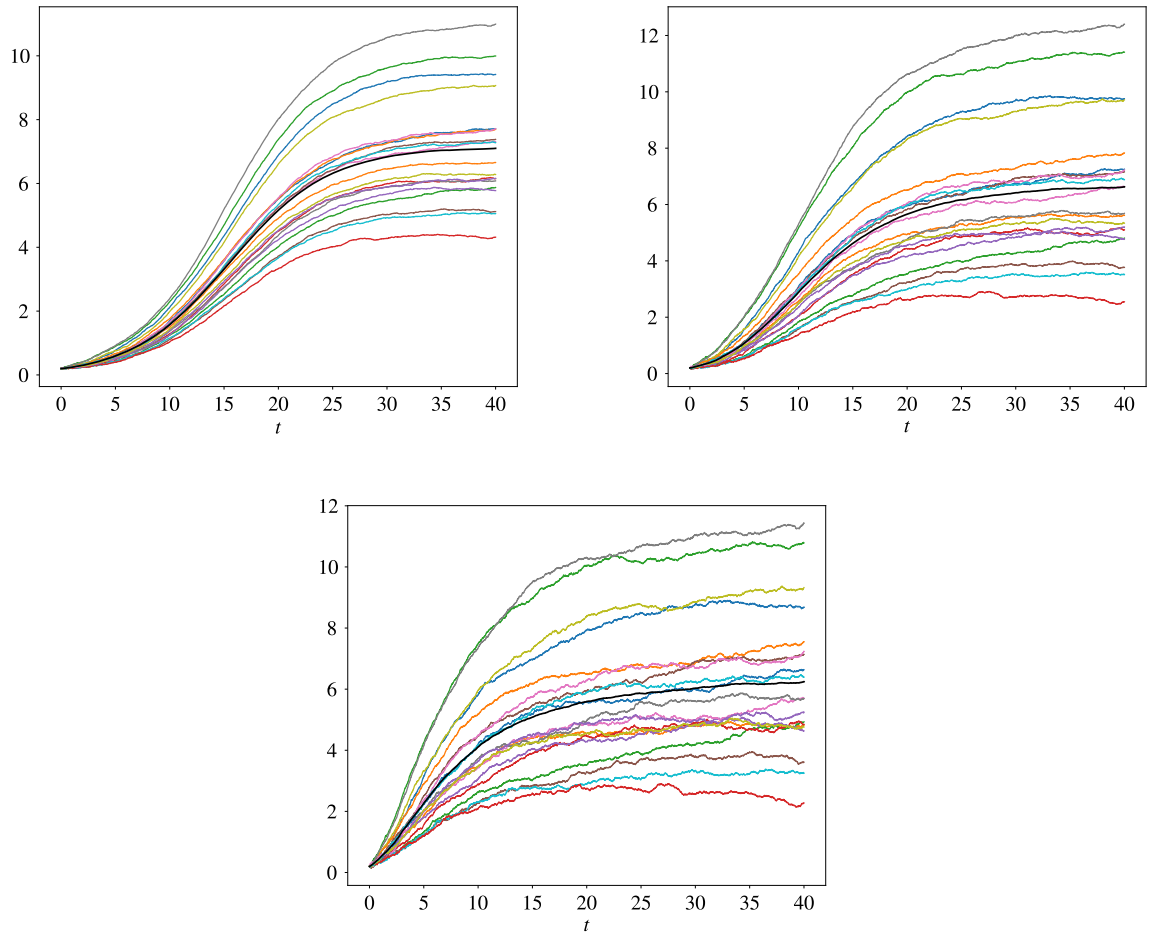

Fig. 4 Simulated paths for the Gaussian case with $n = 1$ known. Up (left to right): $p = 1.01, \sigma = 0.025$, $p = 1.2, \sigma = 0.05$. Down: $p = 1.5, \sigma = 0.075$

Consequently, this approach results in a combination of exploration and exploitation. A moth is exploring when it moves outside the space between itself and the flame. On the contrary, exploitation occurs when the moth moves inside the space between itself and the flame.

In order to avoid local optima that might restrict movement, the position of each moth must be updated (according to the logarithmic spiral) taking only one fixed flame (the list of flames is sorted, by means of the Quicksort algorithm, from high to low fitness at each step).

By considering the sorting algorithm and the function of movement, the computational complexity of the MFO algorithm is $\mathcal{O}(rn^2 + rnd)$, where n is the number of moths, r the maximum number of iterations and d the number of variables, i.e. the dimension of the search space. This has been computed by considering the worst case of the Quicksort algorithm, that is $\mathcal{O}(n^2)$.

In the following, bounds for the parameters of processes $X^{\mathrm{L}}(t)$ and $X^{\mathrm{G}}(t)$ are obtained according to some characteristics of the deterministic curve $x_{\boldsymbol{\theta}}(t)$ given by (5).

Regarding parameter σ_ν, $\nu = \mathrm{L}, \mathrm{G}$, it is known that when it has high values it leads to sample paths with great variability around the mean of the process. Thus, excessive variability in available paths would make a sigmoidal-type modeling unadvisable. Some simulations performed for several values of σ_ν have led us to consider that $0 < \sigma_\nu < 0.1$, so that we may have paths compatible with a sigmoidal-type growth.

On the other hand, whereas $0 < \alpha < 1$, there does not seem to be an upper bound for η and n. From (9), it follows that the growth curve considered shows an inflection point which is higher than t_0, i.e. it is observable, if and only if $\eta \leq b(n, p)$, where

$$b(n, p) = \frac{n(1-p)+1}{np}, \quad 1 < p < 1 + 1/n.$$

Function b is decreasing in p for fixed n. In this case, since $\lim_{p\to 1} b(n, p) = 1/n$, it is deduced that $0 < \eta \leq 1/n$. Consequently, when n is known (namely n_0), it is verified that $1 < p < 1 + 1/n_0$ and $0 < \eta < 1/n_0$. On the contrary, we suggest a procedure to find the bounds for parameter n. This procedure is based on the one used to find a range of values for the parameter of the Box-Cox transformation in linear models. To this end, if we call $L^*(n) = \sup_{\eta,\alpha,p,\sigma_\nu^2} \mathcal{J}_{\mathbf{x}}^{\nu}(\eta, \alpha, p, n, \sigma_\nu^2)$ and consider $L^*(\widehat{n}) = \sup_n L^*(n)$, then $2\,(L^*(\widehat{n}) - L^*(n)) \rightsquigarrow \chi_1^2$, from which

$$P\left[L^*(n) \geq L^*(\widehat{n}) - \frac{1}{2}\chi^2_{1,\widetilde{\alpha}}\right] = 1 - \widetilde{\alpha},$$

being $\chi^2_{1,\widetilde{\alpha}}$ the $(1-\widetilde{\alpha})$-th percentile of a χ^2 distribution with one degree of freedom. Splitting $L^*(n)$ with ordinate $L^*(\widehat{n}) - \frac{1}{2}\chi^2_{1,\widetilde{\alpha}}$, we find two values for n which determine the interval, namely (n_1, n_2). Once this interval is calculated, the corresponding interval for η is $(0, 1/n_1)$ whereas for p is $(1, 1 + 1/n_1)$.

5 Simulation study

In order to study the behavior of the proposed models and the use of metaheuristic approaches for the estimation of the parameters, simulation studies have been carried out for both processes. Moreover, taking into account the arguments provided at the end of the previous section, we consider the cases for n known and unknown separately.

5.1 The case for *n* known

When the value of n is known, the general growth model relies on four parameters, namely η, α, p and σ (throughout this section we will denote the volatility coefficient by σ, without distinguishing the process being used). In this case, upper bounds for η and p are then fixed, as it has been shown in the previous section. In the following, we consider $n = 1$, which is related to the case of the hyperlogistic curve.

The general setup for the simulations consists of 20 paths of a process with initial distribution degenerate at $x_0 = 0.2$. Initial and final times are, respectively, $t_0 = 0$ and 40. Each sample path consists of $N = 401$ values obtained at instants $t_i = 0.1 \times (i-1)$. Parameters η and α are fixed, with values 0.03 and 0.8, respectively. The choice of such values guarantees the existence of one observed inflection point in the curve for values of p between 1 and 2, according to function $b(n, p)$ given in Sect. 4.1.

The bounded parametric space for parametric vector $\boldsymbol{\theta}_1 = (\eta, \alpha, p)^T$ (the vector $\boldsymbol{\theta}$ for $n = 1$) is $[0, 1] \times [0, 1] \times [1, 2]$. Once paths were simulated, the estimation of the parameters was performed by means of the metaheuristic algorithm MFO with 30 agents working across 100 generations, making several replications and taking mean values. The computer implementation was made in Python using the MEALPY package (Van Thieu and Mirjalili 2023).

For the evaluation of the results, the relative absolute error (RAE) of the mean is computed according to the expression

$$\mathrm{RAE} = \frac{1}{N}\sum_{i=1}^{N} \frac{|m_i - x_{\widehat{\boldsymbol{\theta}}_1}(t_i)|}{m_i},$$

where m_i is the value of the observed mean at time t_i and $x_{\widehat{\boldsymbol{\theta}}_1}(t_i)$ is the value of the estimated mean function of the process at the same time instant.

Table 2 Gaussian case with $n = 1$ known: estimations of the parameters and RAE for $\eta = 0.03$, $\alpha = 0.8$ and for different combinations of p, σ

p	σ	$\widehat{\eta}$	$\widehat{\alpha}$	$\widehat{p}$	$\widehat{\sigma}$	RAE
1.01	0.025	0.0284	0.8013	1.0157	0.0251	0.0152
		(0.0007)	(0.0006)	(0.0050)	($< 10^{-4}$)	
	0.050	0.0276	0.8013	1.0212	0.0501	0.0382
		(0.0007)	(0.0006)	(0.0054)	($< 10^{-4}$)	
	0.075	0.0272	0.8012	1.0248	0.0755	0.0528
		(0.0009)	(0.0008)	(0.0067)	(0.0028)	
1.20	0.025	0.0295	0.7988	1.2001	0.0251	0.0306
		(0.0001)	(0.0001)	(0.0003)	($< 10^{-4}$)	
	0.050	0.0289	0.7975	1.2007	0.0501	0.0604
		(0.0026)	(0.0024)	(0.0164)	(0.0001)	
	0.075	0.0279	0.7969	1.2050	0.0752	0.1058
		($< 10^{-4}$)	(0.0001)	(0.0003)	($< 10^{-4}$)	
1.50	0.025	0.0308	0.7966	1.4894	0.0251	0.0053
		($< 10^{-4}$)	($< 10^{-4}$)	(0.0001)	($< 10^{-4}$)	
	0.050	0.0309	0.7929	1.4795	0.0501	0.0186
		($< 10^{-4}$)	($< 10^{-4}$)	(0.0001)	($< 10^{-4}$)	
	0.075	0.0307	0.7890	1.4706	0.0752	0.0419
		($< 10^{-4}$)	($< 10^{-4}$)	($< 10^{-4}$)	($< 10^{-4}$)	

Standard errors in parentheses

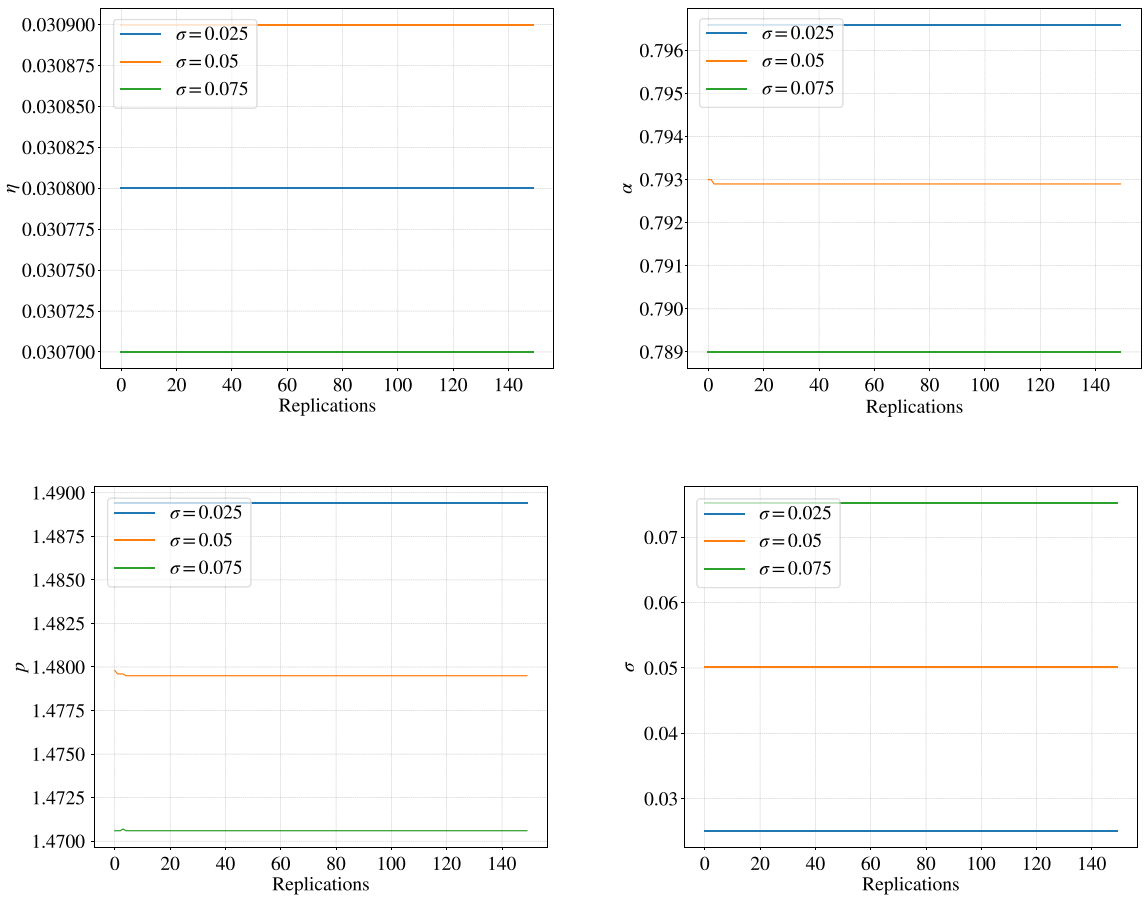

Fig. 5 Gaussian case with $n = 1$ known: evolution of the sequence of the mean of the estimates across replications for $p = 1.5$ and for $\sigma = 0.025, 0.05, 0.075$

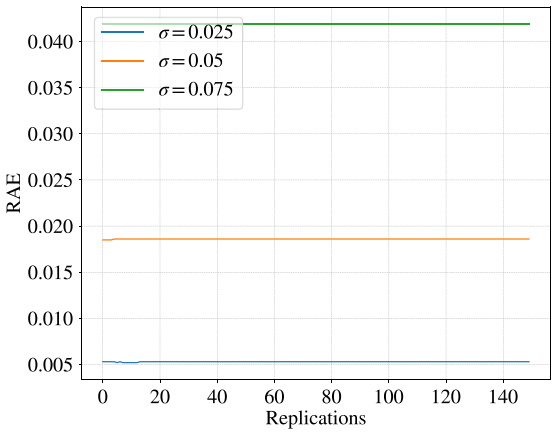

Fig. 6 Gaussian case with $n = 1$ known and $p = 1.5$: RAE across replications for $\sigma = 0.025, 0.05, 0.075$

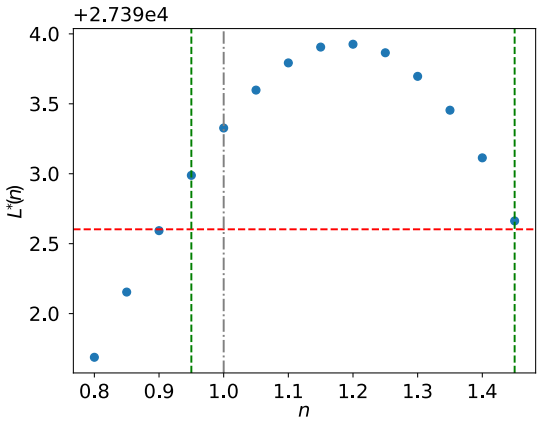

Fig. 7 Lognormal case with $n = 1$ unknown: bounds for n in the general growth process in Table 3 for $p = 1.5$. Blue dots represent values of $L(n_i^*)$ for every n_i^*. The red horizontal line is the threshold that cuts the interpolated function. The vertical lines are the real value of n (grey) and the bounds n_0 and n_1 (green) for n

Under these settings, 150 replications are performed for combination of values $p = 1.01, 1.2, 1.5$ and $\sigma = 0.025, 0.05, 0.075$. The bounds of σ are chosen to be 0.001 and 0.1, a natural choice according to experience, as previously noted.

For the lognormal process described in Sect. 3.1, in Fig. 1 the plots of some simulated paths are shown for several values of p and σ. Precisely, we have chosen $p = 1.01, \sigma = 0.025$, $p = 1.2, \sigma = 0.05$ and $p = 1.5, \sigma = 0.075$ (from left to right). The trajectories grow around the mean function (black curve) and the variability around it increases as σ grows, as expected. Furthermore, we can also observe that the inflection time decreases as the parameter p increases, as it is analytically deduced by Eq. (9).

In Table 1 we list the means and the standard deviation (in parentheses) of the estimates obtained in each replication. The procedure seems to work properly and is able to find estimates that are very close to the true parameters, keeping errors small even as the variability linked to parameter σ increases. This is further made clear by looking at the individual errors given by the standard errors as well as the global errors described by the RAEs in the last column.

For parameters η, α, p and σ, Fig. 2 shows the plots of the evolution of the estimated values found by the algorithm across 150 replications. Precisely, for each of the plots, each line represents the sequence of the mean of the estimates. For example, for parameter η, the green line is obtained at step i as the mean of estimates $\widehat{\eta}_j$ for $j = 1, \ldots, i$, being $\widehat{\eta}_j$ the estimate at the j-th replication for $\sigma = 0.075$.

Finally, in Fig. 3, evolution of RAE across replications can be seen for $p = 1.5$ and for $\sigma = 0.025, 0.05, 0.075$. The error tends to stabilize after just a few replications. Clearly, increasing σ increases the value of the RAE in each replication.

For the Gaussian process described in Sect. 3.2, in Fig. 4 the plots of some simulated paths are shown for several values of p and σ. Also in this case, the trajectories grow around the mean function (black curve) and the variability around it increases as σ grows and the inflection time decreases as the parameter p increases, as it is analytically deduced by Eq. (9). Clearly, since in this case the process is homoscedastic, the oscillations in each trajectory are less sharp.

In Table 2 we list the means and the standard deviation (in parentheses) of the obtained estimates in each replication. Also in this case, estimates very closed to the parameters are found keeping the errors small even as σ increases both in terms of global and individual errors.

For parameters η, α, p and σ, Fig. 5 shows the plots of the evolution of the estimated values found by the algorithm. We

Table 3 Lognormal case with $n = 1$ unknown: estimations of the parameters and RAE for $\eta = 0.03, \alpha = 0.8$ and for $p = 1.01, 1.2, 1.5$ and $\sigma = 0.025$

p	$\hat{\eta}$	$\hat{\alpha}$	$\hat{p}$	$\hat{n}$	$\hat{\sigma}$	RAE	n_0	n_1
1.01	0.0261	0.7975	1.0161	1.0595	0.0251	0.0192	0.99	1.30
	(0.0059)	(0.0045)	(0.0098)	(0.0785)	($< 10^{-4}$)			
1.20	0.0287	0.8003	1.1978	1.0190	0.0248	0.0371	0.95	1.45
	(0.0075)	(0.0045)	(0.0017)	(0.1048)	($< 10^{-4}$)			
1.50	0.0364	0.8018	1.5157	0.9445	0.0249	0.0250	0.90	1.10
	(0.0069)	(0.0053)	(0.0169)	(0.0623)	($< 10^{-4}$)			

The last two columns show bounds of n. Standard errors in parentheses

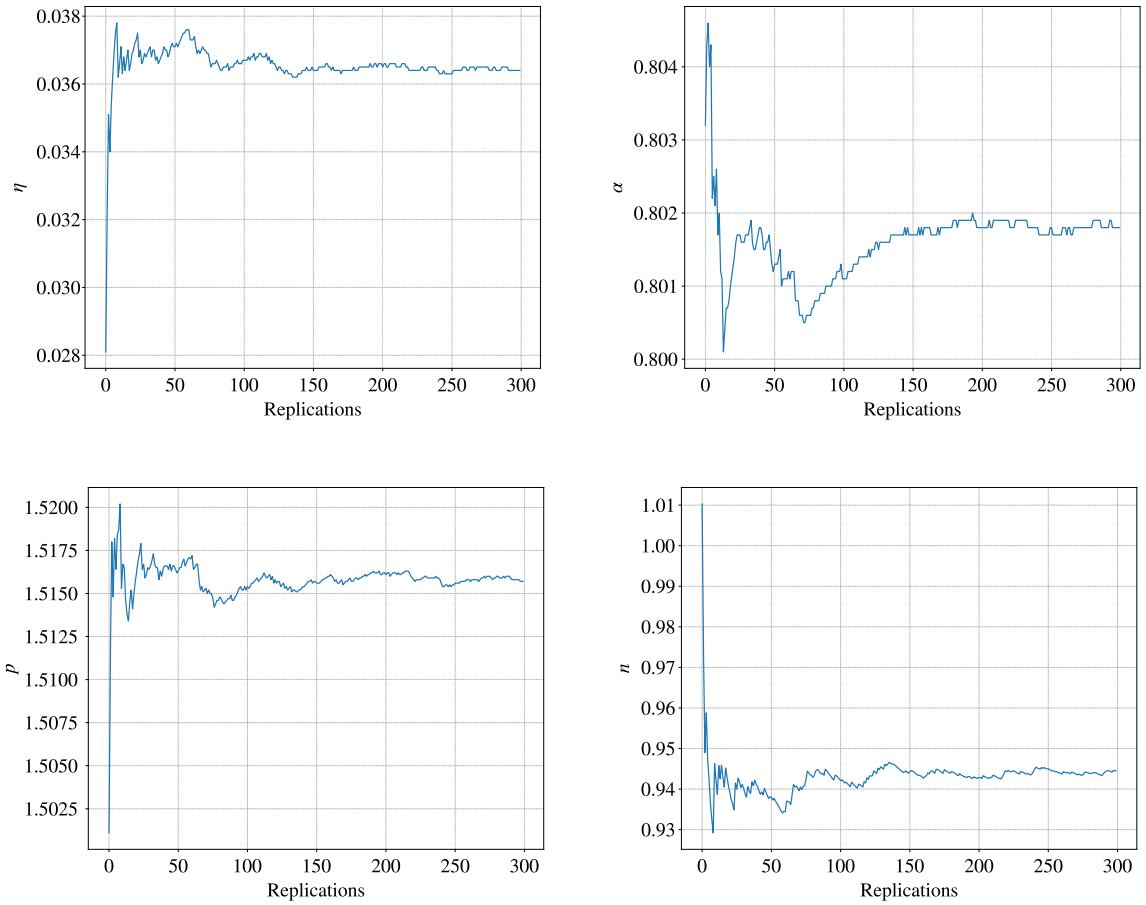

Fig. 8 Lognormal case with $n = 1$ unknown: evolution of the sequence of the mean of the estimates across replications for $p = 1.5$ and $\sigma = 0.025$

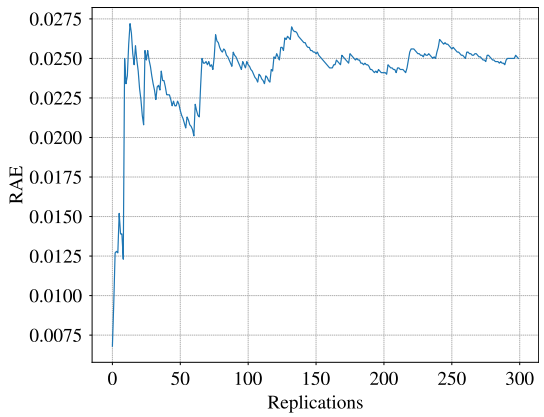

Fig. 9 Lognormal case with $n = 1$ unknown: RAE across replications for $p = 1.5$ and $\sigma = 0.025$

observe that the sequence of the estimates draws a constant line showing that the convergence is reached immediately; this is in accordance with the results of the Table 2 in which the standard errors are lower than 10^{-4}. Finally, in Fig. 6, evolution of RAE across replications can be seen for $p = 1.5$ and for $\sigma = 0.025, 0.05, 0.075$.

5.2 The case for *n* unknown

A simulation study for the general growth processes for n unknown is carried out in this section. In this case, lower and upper bounds for the parameter n can be found according to the discussion done at the end of Sect. 4.1. We must bear in mind that the metaheuristic approach presupposes an initial limitation of the parameter space. Since $\alpha \in (0, 1)$, $\eta \in (0, 1/n)$, $p \in (1, 1 + 1/n)$ and $\sigma \in (0, 0.1)$, when n is unknown we need bounds for n. To obtain these bounds we propose the following two-step procedure: in the first step the lower and upper bounds of n are obtained; in the second step we implement an MFO algorithm to estimate parameters $\eta, \alpha, p, n, \sigma$.

Specifically, the proposed strategy is the following

1. Step 1:
   (a) Let $I = [a, b]$ an interval for n and $0 < \Delta < (b - a)/2$. Set $i = 0, n_i^* = n^* = a$
   (b) **while** $(n^* \leq b)$

Table 4 Gaussian case with $n = 1$ unknown: estimations of the parameters and RAE for $\eta = 0.03, \alpha = 0.8$ and for $p = 1.01, 1.2, 1.5$ and $\sigma = 0.025$

p	$\hat{\eta}$	$\hat{\alpha}$	$\hat{p}$	$\hat{n}$	$\hat{\sigma}$	RAE	n_0	n_1
1.01	0.0331	0.8036	1.0059	0.9855	0.0251	0.0190	0.90	1.40
	(0.0070)	(0.0040)	(0.0131)	(0.0842)	($< 10^{-4}$)			
1.20	0.0377	0.8004	1.1926	0.9540	0.0253	0.0527	0.90	1.20
	(0.0085)	(0.0006)	(0.0074)	(0.0788)	($< 10^{-4}$)			
1.50	0.0369	0.8017	1.5085	0.9528	0.0246	0.0240	0.90	1.10
	(0.0060)	(0.0024)	(0.0054)	(0.0489)	($< 10^{-4}$)			

The last two columns show bounds of n. Standard errors in parentheses

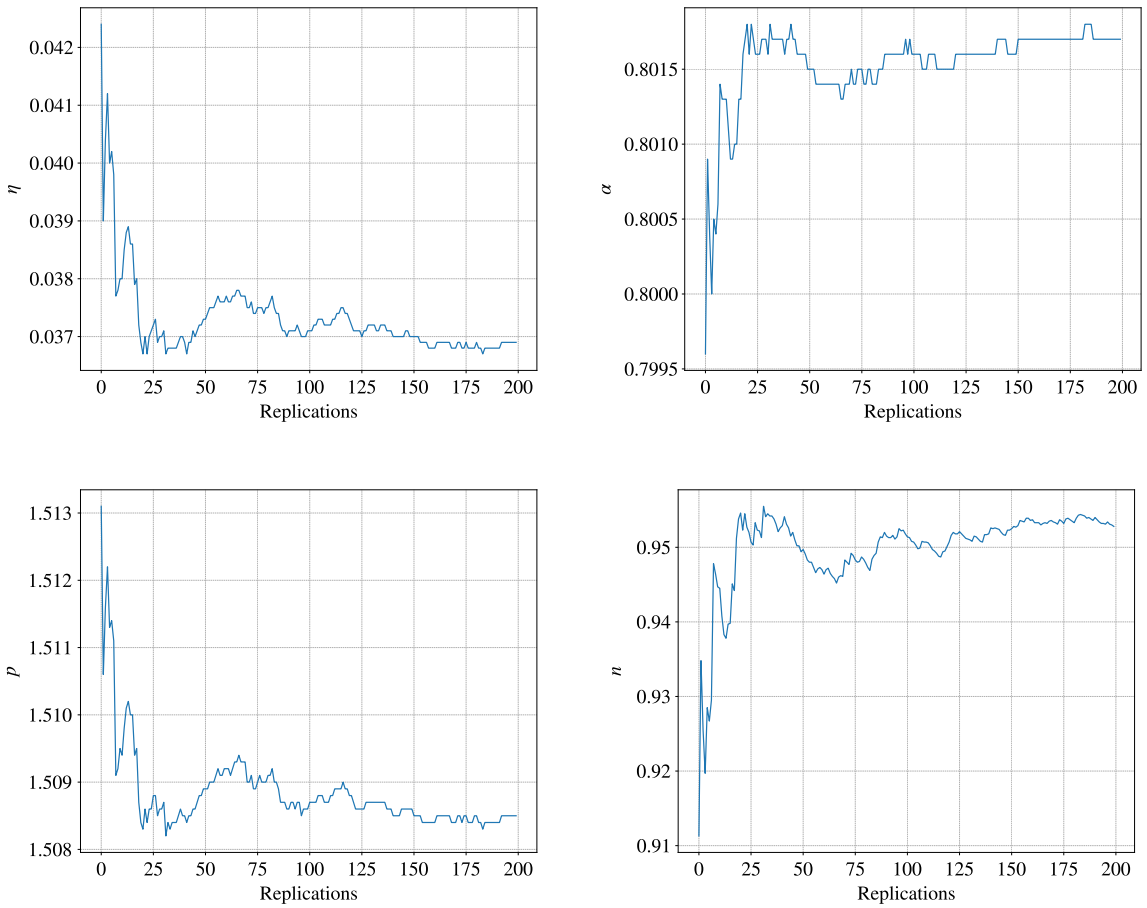

Fig. 10 Gaussian case with $n = 1$ unknown: evolution of the sequence of the mean of the estimates across replications for $p = 1.5$ and $\sigma = 0.025$

- Obtain estimates $\widehat{\eta}, \widehat{\alpha}, \widehat{p}$ and $\widehat{\sigma}$ for η, α, p and σ by using MFO algorithm by assuming n known and equal to n^*
- Evaluate the function $L(n_i^*) = \mathcal{J}_{\mathbf{x}}^v(\widehat{\eta}, \widehat{\alpha}, \widehat{p}, n_i^*, \widehat{\sigma}^2)$
- $i = i + 1, n_i^* = a + i\Delta, n^* = n_i^*$

(c) Plot the values of the vector $L(n_i^*)$ versus n_i^*

(d) Interpolate the points $L(n_i^*)$ in I and obtain the maximum of the interpolated function, namely M

(e) Fixed a confidence level $1 - \widetilde{\alpha}$, calculate the values n_0 and n_1 such that the interpolated function in the interval $[n_0, n_1]$ is greater than $M - \frac{1}{2}\chi^2_{1,\widetilde{\alpha}}$

(1.) Step 2:

- Assume $n \in [n_0, n_1]$, $\alpha \in (0, 1)$, $\eta \in (0, 1/n_0)$, $p \in (1, 1 + 1/n_0)$ and $\sigma \in (0, 0.1)$
- Obtain estimates $\widehat{\eta}, \widehat{\alpha}, \widehat{p}, \widehat{n}$ and $\widehat{\sigma}$ for η, α, p, n and σ by using MFO algorithm in 5 dimensions.

In our study, we simulate 20 sample paths for the lognormal version of the general growth process with parameters $\eta = 0.03, \alpha = 0.8, n = 1$ and $p = 1.01, 1.2, 1.5$, being $\sigma = 0.025$ the diffusion parameter. We must point out that, since we expect that adding parameter n leads to a slower

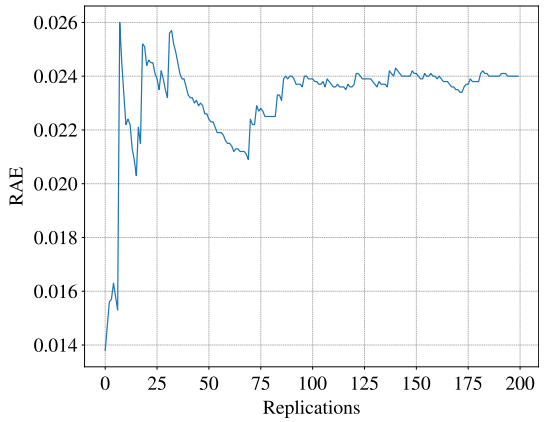

Fig. 11 Gaussian case with $n = 1$ unknown: RAE across replications for $p = 1.5$ and $\sigma = 0.025$

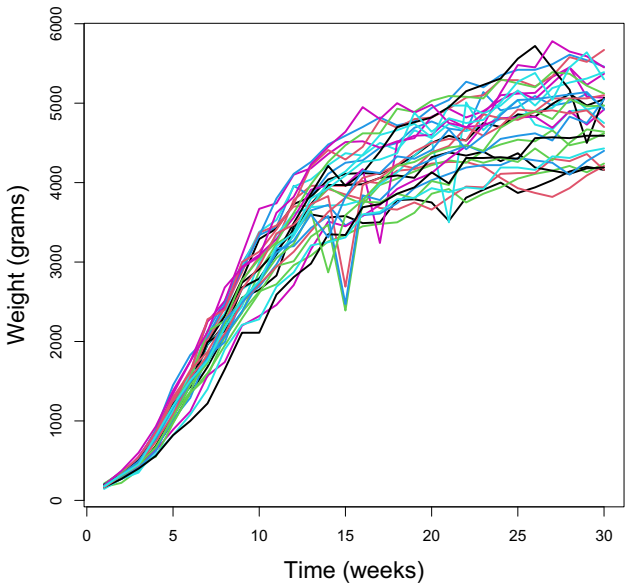

Fig. 12 Weight of 29 rabbits over 30 weeks

convergence of the estimates, for this study we considered 200 replications for the Gaussian case and 300 for the lognormal process because of the heteroscedasticity of the latter. For the lognormal case, in Table 3 we list the mean and the standard deviation (in parentheses) of the estimates obtained, the RAE and the bounds for n obtained by using the proposed methodology. Table 3 considers the same setup parameters for η, α, p and σ of Table 1 but the new parameter n is inclosed. The comparison between two Tables shows that the estimates are very similar in the two cases. However, when n is unknown the estimation errors are bigger, as is as global error RAE.

For illustrative purpose, Fig. 7 shows graphically the computation of the interval $[n_0, n_1]$ for the case $p = 1.5$. Here the confidence level $1 - \widehat{\alpha}$ is fixed to 0.80.

Table 5 Real application: estimates of the parameters, bounds for n and RAE

$\hat{\eta}$	$\hat{\alpha}$	$\hat{p}$	$\hat{n}$	$\hat{\sigma}$	n_0	n_1	RAE
0.1043	0.7225	1.1148	0.6167	0.0691	0.55	0.84	0.0313

For parameters η, α, p and n, Fig. 8 shows the plots of the evolution of the estimated values found by the algorithm. The plot of σ is omitted since its estimated value remains the same across the replications. Figure 9 shows the stabilization of RAE across the replications.

Finally, results for the Gaussian case are shown in Table 4 in which the estimates and errors illustrate the good performance of the procedure, with a similar behavior to the lognormal case comparing it to Table 2 in which parameter n is assumed known. Also the convergence of the estimates is reached as one can see in Figs. 10 and 11.

6 Application to real data

In the following section, we consider data provided from the paper of Blasco et al. (2003), in which a study concerning some aspects related to the growth of a population of rabbits was developed. Figure 12 shows the weight (in grams) of 29 rabbits over 30 weeks. The sample paths start at different initial values, thus exhibiting a sigmoidal behavior, and being their bounds dependent on the initial values. Furthermore, the graph shows how the variability observed between trajectories is greater as the weight of the rabbits increases. These aspects suggest that using the lognormal-type model proposed above would be appropriate with respect to a Gaussian model in which the variability is constant.

To obtain the parameter estimates, the MFO algorithm, with 150 generations, 40 agents and 1000 replications, has been considered. Parameter n is unknown, so the procedure outlined in Sect. 5.2 has been used in the estimation. Final results are listed in Table 5, in which the estimates of the parameters of the model are shown, together with the confidence interval for n with $1 - \alpha = 0.95$ and the value of the related error (RAE), showing that the estimated mean is close to the observed mean function.

In particular, the estimates in Table 5 can be used to forecast the weight of a rabbit given its initial weight x_0. Specifically, thanks to the invariance property of the MLE, we have $\widehat{m_{\boldsymbol{\theta}}(t|t_0)} = m_{\widehat{\boldsymbol{\theta}}}(t|t_0)$. In order to show the goodness of fit of the estimated mean function with respect the weight of rabbits, knowing its initial value, in Fig. 13 we show, for four rabbits, the observed weights along the weeks and the estimated mean function.

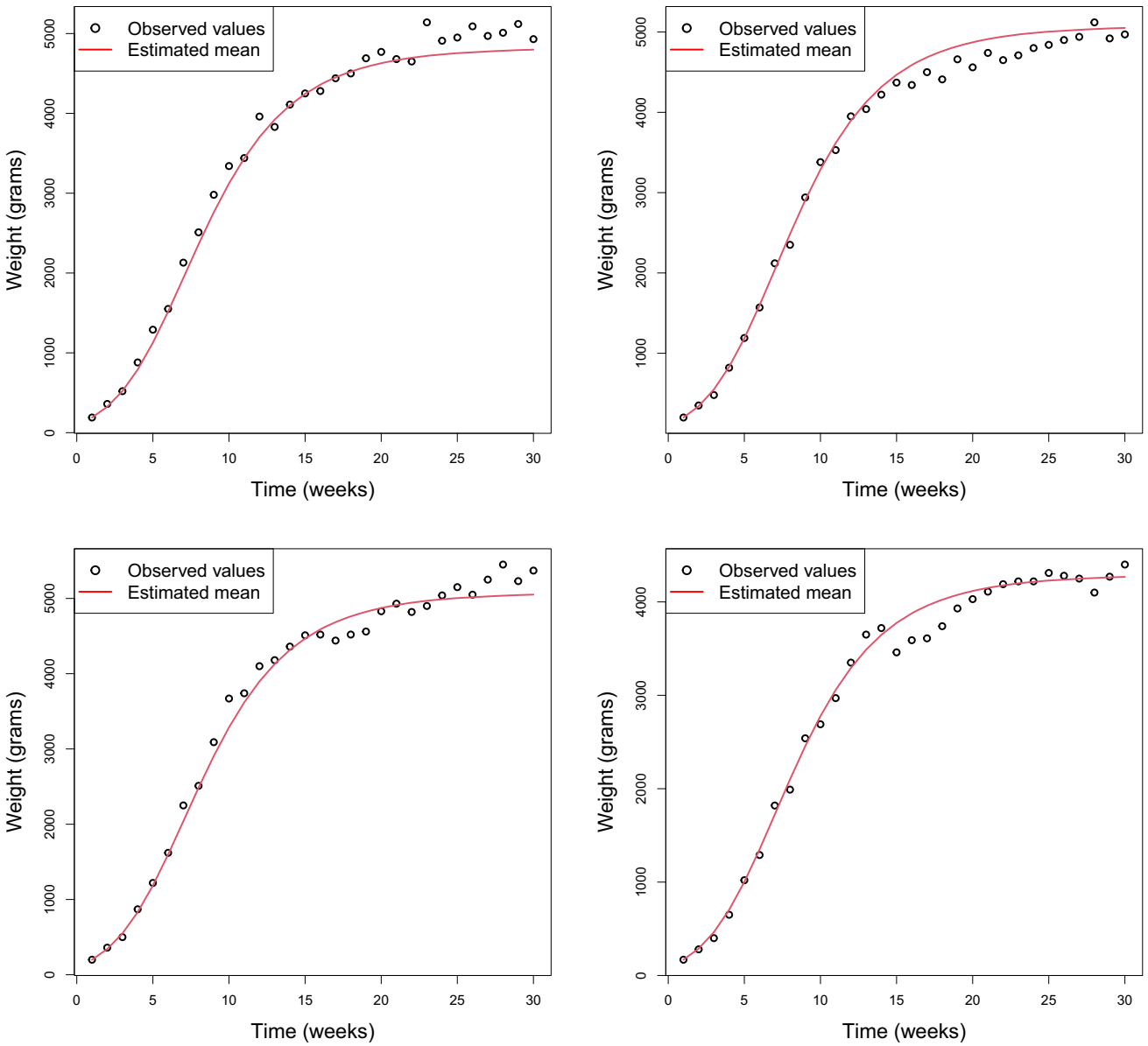

Fig. 13 Observed values and estimated mean function for a choice of rabbits

7 Conclusion

In the previous pages we have considered two stochastic diffusion processes, starting from a general growth curve given by Eq. (1). Such equation includes the well known models of hyper-logistic and hyper-Gompertz growth, so it is therefore able to describe many dynamic systems. Specifically, the diffusion processes are obtained by considering in the deterministic Eq. (1) a multiplicative and an additive noise, leading to a Lognormal and a Gaussian process, respectively. After studying the probability distribution of the two processes, the inference is addressed by means of the maximum likelihood method. Due to the complexity of the resulting system of equations, we have used metaheuristic techniques (MFO algorithm) in order to maximize the likelihood function. In particular, these methods require delimiting the parameter space, and we have therefore considered two cases: n known and n unknown. In the first case, parameter space is in $\mathbb{R}^4$ and the bounds of the parameters are obtained starting from similar information related to the inflection point. In the second case, parameter space is in $\mathbb{R}^5$ and the bounds of the parameters are obtained by making use of a procedure based on the one used to find a range of values for the parameter of the Box-Cox transformation in linear models. Several simulation studies have shown that the estimates produced are very close to the real values of the parameters and provide small errors both for the estimate of each parameter and for the estimate of the mean function of the process.

Acknowledgements The authors are very grateful to the anonymous referees whose comments and suggestions have contributed to improving this paper. This research is partially supported by PID2020-1187879GB-100 and CEX2020-001105-M grants, funded by MCIN/AEI/10.13039/501100011033 (Spain). The authors G. Albano and V. Giorno acknowledge support received from the " European Union – Next Generation EU" through MUR-PRIN 2022, project 2022XZSAFN "Anomalous Phenomena on Regular and Irregular Domains: Approximating Complexity for the Applied Sciences," and MUR-PRIN 2022 PNRR, project P2022XSF5H " Stochastic Models in Biomathematics and Applications.". They are members of the GNCS-INdAM. Funding for open access charge: University of Granada / CBUA.

Declarations

Conflict of interest The authors declare no Conflict of interest.

References

Aït-Sahalia, Y.: Maximum likelihood estimation of discretely sample diffusion: a close form approximation approach. Econometrica **70**, 223–262 (2003). https://doi.org/10.1111/1468-0262.00274

Albano, G., Giorno, V.: Inferring time non-homogeneous Ornstein Uhlenbeck type stochastic process. Comput. Stat. Data Anal. **150**, 107008 (2020). https://doi.org/10.1016/j.csda.2020.107008

Albano, G., Giorno, V., Román-Román, P., Torres-Ruiz, F.: On the effect of a therapy able to modify both the growth rates in a Gompertz stochastic model. Math. Biosci. **245**, 12–21 (2013). https://doi.org/10.1016/j.mbs.2013.01.001

Albano, G., Giorno, V., Román-Román, P., Román-Román, S., Torres-Ruiz, F.: Inference on an heteroscedastic Gompertz tumor growth model. Math. Biosci. **328**, 108428 (2020). https://doi.org/10.1016/j.mbs.2020.108428

Albano, G., Giorno, V., Román-Román, P., Torres-Ruiz, F.: Study of a general growth model. Commun. Nonlinear Sci. Numer. Simul. **107**, 106100 (2022). https://doi.org/10.1016/j.cnsns.2021.106100

Albano, G., Barrera, A., Giorno, V., Román-Román, P., Torres-Ruiz, F.: First passage and first exit times for diffusion processes related to a general growth curve. Commun. Nonlinear Sci. Numer. Simul. **126**, 107494 (2023). https://doi.org/10.1016/j.cnsns.2023.107494

Banks, R.B.: Growth and Diffusion Phenomena. Springer, Berlin (1994)

Barrera, A., Román-Román, P., Torres-Ruiz, F.: Hyperbolastic models from a stochastic differential equation point of view. Mathematics **9**, 1835 (2021). https://doi.org/10.3390/math9161835

Bhowmick, A.R., Bhattacharya, S.: A new growth curve model for biological growth: some inferential studies on the growth of Cirrhinus mrigala. Math. Biosci. **254**, 28–41 (2014). https://doi.org/10.1016/j.mbs.2014.06.004

Bibby, B., Jacobsen, M., Sørensen, M.: Estimating functions for discretely sampled diffusion type models. In: Aït-Sahalia, Y., Hansen, L.P. (eds.) Handbook of Financial Econometrics: Tools and Techniques, pp. 203–268. North-Holland, Amsterdam (2009). https://doi.org/10.1016/B978-0-444-50897-3.50007-9

Blasco, A., Piles, M., Varona, L.A.: A Bayesian analysis of the effect of selection for growth rate on growth curves in rabbits. Genet. Select. Evol. **35**, 21–41 (2003). https://doi.org/10.1051/gse:2002034

Blumberg, A.A.: Logistic growth rate functions. J. Theor. Biol. **21**, 42–44 (1968). https://doi.org/10.1016/0022-5193(68)90058-1

Dacunha-Castelle, D., Florens-Zmirou, D.: Estimation of the coefficients of a diffusion from discrete observations. Stochastics **19**, 263–284 (1986). https://doi.org/10.1080/17442508608833428

Di Crescenzo, A., Spina, S.: Analysis of a growth model inspired by Gompertz and Korf laws, and an analogous birth-death process. Math. Biosci. **282**, 121–134 (2016). https://doi.org/10.1016/j.mbs.2016.10.005

Di Crescenzo, A., Paraggio, P., Román-Román, P., Torres-Ruiz, F.: Applications of the multi-sigmoidal deterministic and stochastic logistic models for plant dynamics. Appl. Math. Model. **92**, 884–904 (2021). https://doi.org/10.1016/j.apm.2020.11.046

Eby, W., Tabatabai, M.: Methods in mathematical modeling for stem cells. In: Hayasl, M.A. (ed.) Stem Cells and Cancer Stem Cells, pp. 201–217. Springer, Berlin (2014). https://doi.org/10.1007/978-94-017-8032-2

Fan, J.: A selective overview of nonparametric methods in financial econometrics. Stat. Sci. **20**, 317–357 (2005). https://doi.org/10.1214/088342305000000412

Flinn, S.A., Midway, S.R.: Trends in growth modeling in fisheries science. Fishes **6**, 1 (2021). https://doi.org/10.3390/fishes6010001

Gutiérrez, R., Román, P., Romero, D., Torres, F.: Forecasting for the univariate lognormal diffusion process with exogenous factors. Cybern. Syst. **34**, 709–724 (2003). https://doi.org/10.1080/716100279

Gutiérrez, R., Román, P., Romero, D., Serrano, J.J., Torres, F.: A new Gompertz-type diffusion process with application to random growth. Math. Biosci. **208**, 147–165 (2007). https://doi.org/10.1016/j.mbs.2006.09.020

Hsieh, Y.H., Fisman, D.N., Wu, J.: On epidemic modeling in real time: an application to the 2009 Novel A (H1N1) influenza outbreak in Canada. BMC Res. Notes **3**, 283 (2010). https://doi.org/10.1186/1756-0500-3-283

Korashy, A., Kamel, S., Alquthami, T., Jurado, F.: Optimal coordination of standard and non-standard direction overcurrent relays using an improved Moth-flame optimization. IEEE Access **8**, 87378–87392 (2020). https://doi.org/10.1109/access.2020.2992566

Kotary, D.K., Nanda, S.J.: Distributed robust data clustering in wireless sensor networks using diffusion moth flame optimization. Eng. Appl. Artif. Intell. **87**, 103342 (2020). https://doi.org/10.1016/j.engappai.2019.103342

Koya, P.R., Goshu, A.T.: Generalized mathematical model for biological growths. Open J. Model. Simul. **1**, 42–53 (2013). https://doi.org/10.4236/ojmsi.2013.14008

Lo, A.: Maximum likelihood estimation of generalized Ito processes with discretely sampled data. Econom. Theory **4**, 231–247 (1988). https://doi.org/10.1017/S0266466600012044

Lo, C.F.: A modified stochastic Gompertz model for tumour cell growth. Comput. Math. Methods Med. **11**, 3–11 (2010). https://doi.org/10.1080/17486700802545543

López-Pérez, A., Febrero-Bande, M., González-Manteiga, W.: Parametric estimation of diffusion processes: a review and comparative study. Mathematics **9**, 859 (2021). https://doi.org/10.3390/math9080859

Mirjalili, S.: Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm. Knowl.-Based Syst. **89**, 228–249 (2015). https://doi.org/10.1016/j.knosys.2015.07.006

Moss, D., Ivany, L., Jones, D.: Fossil bivalves and the sclerochronological reawakening. Paleobiology **47**, 551–573 (2021). https://doi.org/10.1017/pab.2021.16

Pedersen, A.R.: Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes. Bernoulli **1**, 257–279 (1995). https://doi.org/10.2307/3318480

Protazio, J.M.B., Souza, M.A., Hernández-Díaz, J.C., Escobar-Flores, J.G., López-Sánchez, C.A., et al.: A dynamical model based on the Chapman-Richards growth equation for fitting growth curves for four pine species in Northern Mexico. Forests **13**, 1866 (2022). https://doi.org/10.3390/f13111866

Rajasekar, S.P., Pitchaimani, M.: Ergodic stationary distribution and extinction of a stochastic SIRS epidemic model with logistic growth and nonlinear incidence. Appl. Math. Comput. **377**, 1–15 (2020). https://doi.org/10.1016/j.amc.2020.125143

Rajasekar, S.P., Pitchaiman, M., Zhu, Q.: Progressive dynamics of a stochastic epidemic model with logistic growth and saturated treatment. Physica A **538**, 1–20 (2020). https://doi.org/10.1016/j.physa.2019.122649

Román-Román, P., Torres-Ruiz, F.: A stochastic model related to the Richards-type growth curve. Estimation by means of simulated annealing and variable neighborhood search. Appl. Math. Comput. **266**, 579–598 (2015). https://doi.org/10.1016/j.amc.2015.05.096

Román-Román, P., Serrano-Pérez, J.J., Torres-Ruiz, F.: Some notes about inference for the lognormal diffusion process with exogenous factors. Mathematics **6**, 85 (2018). https://doi.org/10.3390/math6050085

Román-Román, P., Serrano-Pérez, J.J., Torres-Ruiz, F.: A note on estimation of multi-sigmoidal Gompertz functions with random noise. Mathematics **7**, 541 (2019). https://doi.org/10.3390/math7060541

Tabatatai, M., Williams, D.K., Bursac, Z.: Hyperbolastic growth models: theory and application. Theor. Biol. Med. Model. **2**, 1–13 (2005). https://doi.org/10.1186/1742-4682-2-14

Tang, C.Y., Chen, S.X.: Parameter estimation and bias correction for diffusion processes. J. Econom. **149**, 65–81 (2009). https://doi.org/10.1016/j.jeconom.2008.11.001

Tsoularis, A., Wallace, J.: Analysis of logistic growth models. Math. Biosci. **179**, 21–55 (2002). https://doi.org/10.1016/S0025-5564(02)00096-2

Turner, M.E., Bradley, E., Kirk, K., Pruit, K.: A theory of growth. Math. Biosci. **29**, 367–373 (1976). https://doi.org/10.1016/0025-5564(76)90112-7

Van Thieu, N., Mirjalili, S.: MEALPY: an open-source library for latest meta-heuristic algorithms in Python. J. Syst. Archit. **139**, 102871 (2023). https://doi.org/10.1016/j.sysarc.2023.102871