

Article

# T-Growth Stochastic Model: Simulation and Inference via Metaheuristic Algorithms

Antonio Barrera <sup>1,2,\*,†</sup> , Patricia Román-Román <sup>2,3,†</sup>  and Francisco Torres-Ruiz <sup>2,3,†</sup> 

- <sup>1</sup> Departamento de Análisis Matemático, Estadística e Investigación Operativa y Matemática Aplicada, Facultad de Ciencias, Universidad de Málaga, Bulevar Louis Pasteur, 31, 29010 Málaga, Spain
- <sup>2</sup> Instituto de Matemáticas de la Universidad de Granada (IMAG), Calle Ventanilla 11, 18001 Granada, Spain; proman@ugr.es (P.R.-R.); fdeasis@ugr.es (F.T.-R.)
- <sup>3</sup> Departamento de Estadística e Investigación Operativa, Facultad de Ciencias, Universidad de Granada, Avenida Fuente Nueva s/n, 18071 Granada, Spain
- \* Correspondence: antonio.barrera@uma.es; Tel.: +34-952131896
- † These authors contributed equally to this work.

**Abstract:** The main objective of this work is to introduce a stochastic model associated with the one described by the T-growth curve, which is in turn a modification of the logistic curve. By conveniently reformulating the T curve, it may be obtained as a solution to a linear differential equation. This greatly simplifies the mathematical treatment of the model and allows a diffusion process to be defined, which is derived from the non-homogeneous lognormal diffusion process, whose mean function is a T curve. This allows the phenomenon under study to be viewed in a dynamic way. In these pages, the distribution of the process is obtained, as are its main characteristics. The maximum likelihood estimation procedure is carried out by optimization via metaheuristic algorithms. Thanks to an exhaustive study of the curve, a strategy is obtained to bound the parametric space, which is a requirement for the application of various swarm-based metaheuristic algorithms. A simulation study is presented to show the validity of the bounding procedure and an example based on real data is provided.

**Keywords:** growth curves; diffusion processes; inference on diffusion processes; metaheuristic algorithms



**Citation:** Barrera, A.; Román-Román, P.; Torres-Ruiz, F. T-Growth Stochastic Model: Simulation and Inference via Metaheuristic Algorithms. *Mathematics* **2021**, *9*, 959. <https://doi.org/10.3390/math9090959>

Academic Editor: Elvira Di Nardo

Received: 2 March 2021

Accepted: 22 April 2021

Published: 25 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In recent years, the mathematical description of growth-related phenomena has become a burgeoning field of research. Today, the formal analysis of such phenomena and the prediction of future behavior from past observed data make good use of the so-called growth models. The availability of large datasets and the increasing computational capacity for the analysis and development of new statistical techniques have propelled the traditional field of population dynamics towards the center stage, where new, increasingly sophisticated models are required. Taking growth curves as a basis, researchers work on models which are applied to fields as diverse as cell growth (whether tumoral, stem, or other tissue), the exploitation of energy resources, or infection transmission and the spread of pandemics. In this latter regard, it is obvious that the SARS-CoV-2 virus has further fueled interest in the development and application of new dynamic models.

For the reasons outlined above, the use of growth curves in the study of biological phenomena has become more widespread than ever before. Recent studies have applied growth curves to the description of a variety of phenomena. For example, Ram et al. [1] employed logistic, exponential, and Gompertz curves to the study of microbial growth in a mixed culture. Spiess et al. [2,3] studied the methods used in the analysis of PCR techniques (currently applied to the detection of the new coronavirus), whose observations can be modeled through growth curves. Additionally, and also concerning pandemic dynamics, Rajasekar et al. [4] obtained some results about a SIRS model with logistic growth. Several

fields of study can be formally approached by means of deterministic growth models, such as the analysis of fish growth [5], which is carried out by considering a new framework related with the von Bertalanffy curve. However, such tools lend themselves not only to the description of phenomena: their behavior in context itself has become the object of study. In this sense, one very successful approach is to extend classic models by introducing functions that modify their original behavior and allow them to reproduce real phenomena of greater complexity. Concerning this approach, the next section makes a brief review of models that have emerged recently thanks to the incorporation of hyperbolic functions to classic models such as the logistic and the Weibull.

Despite the high degree of effectiveness shown by many of these deterministic models, many of them lack the ability to account for random factors affecting growth dynamics. To improve a model's ability to fit a phenomenon, one must take into consideration the uncertainty generated by factors both internal and external to the process, which are unknown and therefore commonly disregarded. The same applies to errors generated by the observation and measurement of the phenomenon and to the random nature of certain dynamics. For this reason, stochastic models must be introduced which can be built on the basis of said growth curves as a basis. As a matter of fact, including certain noise factors in the ordinary differential equations which generate said curves leads to stochastic differential equations, which have, under certain conditions, diffusion processes as a solution. The literature is extensive in this regard and continues to grow, as exemplified by Hening et al. [6], who developed a model characterized by a density-dependent growth rate and applies it to structured populations under the conditions of a heterogeneous environment. Additionally, Parsons [7] considered some stochastic processes and their effects on populations with a logistic behavior and Román-Román and Torres-Ruiz [8] defined a stochastic model based on a Richards-type growth curve. On the other hand, concerning birth-and-death processes, Asadi et al. [9] dealt with a generalization of the Gompertz model showing good fit to real data from COVID-19 outbreak. In connection with the growth curves, widely used to model tumor dynamics, other studies such as the one carried out by Ascione et al. [10] also build a fractional version of stochastic models. Furthermore, and as mentioned above, stochastic models can also be developed by generalizing classic ones (see the next section for some examples linked with the objective of the present paper).

However, to obtain stochastic models derived from growth curves, one must face certain problems concerning inference, as the very nature of said processes commonly prevents the explicit knowledge of certain features such as the transition density functions. Similarly, analytical complexity thwarts the estimation of parameters from mere observation. In this regard, non-homogeneous lognormal diffusion processes allow for many of the characteristics of the process to be known and render the process in general much more manageable. To build these processes from growth curves, the original model must usually be reformulated. The convenience of such an approach is complemented by the possibility of carrying out estimations by means of maximum likelihood procedures. Along these lines, Román-Román et al. [11] carried out a study regarding the estimation by maximum likelihood in the lognormal process with a generic exogenous factor. This enables the estimation of various processes associated with growth curves, such as multisigmoidal models, to be addressed [12]. Moreover, maximum-likelihood equations can also be dealt with by means of metaheuristic algorithms, as the complexity of the model may hamper efforts to solve the equations. Several algorithms exist which can be used to optimize the likelihood function, among them Simulated Annealing or Variable Neighbourhood Search, which is applied in [8] to the case of a Richards-type model. More successful recent algorithms are, for example, the Whale Optimization or the Grey Wolf and Moth Flame optimizers algorithms (see [13–15] for more details), all of them based on animal social behavior. In addition, the Clonal Selection algorithm [16] is based on cell behavior and has been applied, for example, to optimization problems related with electromagnetics [17]. Another

procedure is the Harmony Search algorithm, widely extended to many applications and subject to several extensions (see [18–20]).

To a large extent, the procedure used to develop stochastic growth models takes into account the factors stated above. However, the sheer complexity of the initial deterministic model plays a central role in the development and application of any stochastic model. In this regard, growth curves must be used that allow for a better fit to the data without needlessly increasing complexity. The present work deals with building a model based on the T-growth curve, which has been shown to have great potential in modeling tumor growth dynamics in the presence of an immune response. Given its inherent analytical simplicity, this model, which is derived from the classic logistic one, may be successfully combined with equations associated with dynamics involving the effector cells of the immune system.

The structure of the present work is as follows. Section 2 provides a brief summary of the evolution followed by the study of hyperbolastic and derived models, both deterministic and stochastic. Section 3 presents the T-growth curve, analyzes its main properties, and introduces a reformulation that enables asymptotic dependence on the initial conditions. Thanks to this, a stochastic model associated with the curve based on the non-homogeneous lognormal diffusion process is introduced, which is the objective of Section 3. Inference in this process is addressed in Section 4, which considers the use of five metaheuristic algorithms (“Clonal Selection”, “Harmony Search”, “Grey Wolf Optimizer”, “Moth Flame Optimizer”, and “Whale Optimization Algorithm”) with the goal of optimizing the maximum-likelihood function. In addition, a proposal is made concerning the strategy employed in order to bound the parametric space associated with the process. Section 5 presents a simulation study in order to compare the performance of the aforementioned metaheuristic algorithms regarding the inference of the new process. Additionally, it shows the validity of the parametric bounding procedure proposed in these pages. Finally, Section 6 provides an application of the proposed model and methodology to real data.

## 2. Brief Overview on Hyperbolastic and Associated Models

The origins of the T-type growth model can be traced back to the methodology suggested by Tabatabai et al. [21] and Bursac et al. [22] for generating the so-called hyperbolastic models. Their starting concept was using classic growth curves as the basis to develop other, more flexible models which do not require a large increase in the number of parameters. By virtue of this flexibility, such models can be used to represent certain real phenomena which classic models are unable to adequately explain.

Following these premises, in the aforementioned work, the authors introduced Type I, II, and III hyperbolastic curves (usually denoted as H1, H2 and H3) as the solution of certain ordinary differential equations. In essence, curves H1 and H3 are obtained by solving certain generalizations of the logistic and Weibull differential equations, while curve H2 is a non-homogeneous quadratic equation. All such differential equations have in common the presence of hyperbolic functions (or derivatives of them). The resulting models have been applied to several fields, mostly dealing with bio-sanitary phenomena such as cell proliferation. Thus, Eby et al. [23] used hyperbolastic curves to analyze the growth of solid Ehrlich carcinoma using a variety of treatments. Tabatabai et al. [24] used them as a model for the study of stem cell proliferation, showing how some of them, specifically H3, improve the results obtained with other models such as Deasy’s [25]. In other fields, the models have been used to represent wound healing [26], specifically in the research of variations in the healing speed of different types of wounds at several stages of the healing process. Mixed-effects models based on the H2 curve have also been considered (see [27]).

Hyperbolastic curves are not the only product of the methodology outlined above. Tabatabai et al. [28] considered the so-called oscillabolastic model, which is used to explain the behavior of dynamic biomedical data in which an oscillatory behavior is observed. Such model was applied to the study of Ehrlich ascites tumor growth and to the modeling

of signal intensity in the gene expression of the Hes1 gene. Another curve of interest is the hypertabastic curve, introduced in Eby and Tabatabai [29] within a context of survival models. In that work, the authors recommended its use to model stem cell differentiation time and suggested the inclusion in the curve of covariates that enable the study of the influence of various factors. Finally, the main subject of the present work, the T-type growth curve was introduced by Tabatabai et al. [30] from a modified differential equation of a logistic one. As the authors mentioned, the model shows great versatility in the representation of dynamic growth phenomena, while enabling models of biphasic growth patterns to be built. Such patterns appear, for instance, in the development of tumors that display temporary decreases in their growth.

Although all these models have been shown to be useful in the description and adjustment of phenomena such as those mentioned above, they are in essence deterministic models since they do not take into account the random fluctuations that may exist in the system under study. In light of that, these curves would be well served by associating them with dynamic models with the power to represent such random information. Some such models are those representing stochastic processes and, more specifically, diffusion processes. Such processes arise when white noise is introduced into the ordinary differential equation that produces the curve, giving way to the consideration of stochastic differential equations whose solution are, under certain conditions of regularity, diffusion processes. These processes are advantageously applied to the computation of multiple statistics aiming to describe the dynamics under study, such as the mean, mode or quantile functions. These statistical functions could also be used for prediction purposes. At the same time, this stochastic-type approach makes it possible to address another series of problems associated with the study of these types of dynamic variables. Among them, the determination of temporal variables that represent the moment in which the variable under study verifies a certain condition stands out, such as the moment of inflection or the moment in which a certain threshold is exceeded. These questions can be posed as first-passage time problems for which there is a well-established methodology within the field of diffusion processes, from theoretical and applied perspectives (see [31–33], respectively).

Within this context, Barrera et al. [34,35] established two different diffusion processes associated with curves H1 and H3. In both cases, they analyzed their properties and addressed the estimation problem. Both processes were used for the determination of fluorescence, as a function of the cycle, in the application of the qPCR technique of molecular biology. On the other hand, Barrera et al. [36] presented several alternatives for modeling oscillabolastic-type behavior, more specifically a Gaussian-type process and another derived from the inhomogeneous lognormal process.

### 3. The T-Growth Curve

The T-growth curve is defined as the solution of the ordinary differential equation

$$\frac{d}{dt}S(t) = S(t)\left(1 - \frac{S(t)}{M}\right)\beta \cosh(\beta t), \quad t \geq t_0, \quad (1)$$

with initial condition  $S(t_0) = S_0$ . Parameter  $\beta > 0$  is the growth rate and  $M > 0$  is the asymptotic value of the quantity  $S(t)$ , known as the *carrying capacity*. Its solution is (see [30]),

$$S(t) = \frac{M}{1 + \alpha \exp(-\sinh(\beta t))}, \quad t \geq t_0, \quad (2)$$

where  $\alpha$  is a parameter depending on the initial values,

$$\alpha = \frac{M - S_0}{S_0} \exp(\sinh(\beta t_0)).$$

Note that Equation (1) can be viewed as the classic logistic differential equation with a multiplicative factor  $\beta \cosh(\beta t)$ . Solving  $\beta \cosh(\beta t) > 1$  for  $t$ , it follows that the T-growth model grows faster than the logistic model when  $t > \operatorname{arccosh}(1/\beta)/\beta$ .

As is well known, in this type of sigmoidal increasing curves, the inflection point of the curve determines the time instant of maximum growth velocity. Thus, this point is of special interest in any study about population growth. As a matter of fact, the T-growth model may have different inflection points, which increases the flexibility of the curve and its ability to adapt to different behaviours.

The time  $t^*$  at which inflection occurs verifies the equation

$$\frac{\exp(\sinh(\beta t^*)) - \alpha}{\exp(\sinh(\beta t^*)) + \alpha} = \tanh(\beta t^*) \operatorname{sech}(\beta t^*). \tag{3}$$

After some algebra, and from Expression (2), it follows that the value of the curve at the inflection time is

$$S(t^*) = \frac{M}{2} (\tanh(\beta t^*) \operatorname{sech}(\beta t^*) + 1).$$

This guarantees that, even if it depends on growth rate  $\beta$ , at inflection  $t^*$ , the value of the curve is greater than half of its maximum capacity, that is  $S(t^*) > M/2$ .

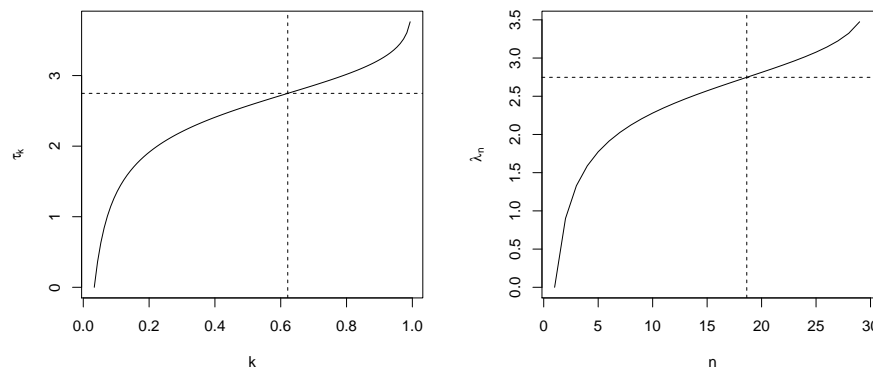
Another remarkable time is that at which the curve reaches a particular value. The first time the curve crosses a threshold  $T$  such that  $S_0 < T < M$  is important in terms of growth analysis. Indeed, such threshold can be expressed as  $T = kM$ , where  $S_0/M \leq k < 1$  is a positive number representing 100k% of the total carrying capacity. If such *crossing-time* is denoted by  $\tau_k$ , it means that  $S(\tau_k) = kM$ . Solving this expression, it follows that

$$\tau_k = \frac{1}{\beta} \operatorname{arcsinh}\left(\log\left(\frac{\alpha k}{1 - k}\right)\right).$$

A different approach can be taken by considering the threshold as a quantity  $T = nS_0$  consisting of  $n > 0$  times the initial value of the curve. As far as the study of growth phenomena is concerned, this approach is particularly useful in fields such as the analysis of tumor growth, as it can be used to estimate the moment at which the initial volume of the tumor reaches a particular size (expressed in percentages). This expression could be also included in other dynamical systems, e.g., those taking into account the dynamics of the tumor growth under certain treatments. Now, being  $\lambda_n$  the time at which the curve reaches  $n$  times its initial value, then  $S(\lambda_n) = nS_0$ , and finally

$$\lambda_n = \frac{1}{\beta} \operatorname{arcsinh}\left(\log\left(\frac{\alpha n S_0}{M - n S_0}\right)\right).$$

Please note that  $M > nS_0$ , provided that  $M$  is an asymptotic value. Then, it follows that  $0 < n < M/S_0$ . Figure 1 shows the evolution of times  $\tau_k$  and  $\lambda_n$  for  $k$  when the carrying capacity  $M = 3$ , and for  $n$  when the initial volume is  $S_0 = 0.1$ , respectively. The growth rate is  $\beta = 0.75$ . On the one hand,  $0.1/3 \leq k < 1$  and, on the other,  $0 < n < 30$ . The effect of  $\beta$  is shown in the scale of the vertical axis. Actually, as the growth rate increases, the range of each time becomes smaller.



**Figure 1.** From left to right: times  $\tau_k$  and  $\lambda_n$  for different magnitudes  $k$  of carrying capacity  $M$  and  $n$  of initial volume  $S_0 = 0.1$ , in the case of parameters  $\beta = 0.75$  and  $M = 3$ . Horizontal dashed lines indicate inflection time. Vertical dashed lines show the values of  $k$  and  $n$  at inflection time, respectively.

*Reformulation of the T-Growth Model*

The original formulation of the T-growth model implies that the carrying capacity will be independent of the initial value. Nevertheless, in many practical applications, such independence would not adequately describe the dynamic phenomena under study. In such cases, the carrying capacity must be made to depend on the initial value. For example, concerning the study of livestock, the weight or size of individuals is studied to determine the moment at which they become marketable. Similarly, in fruit farms, knowing the distribution of fruit size at harvest is of great importance for fruit producers. An increase in benefits may result from proper forecasting of this distribution and from the prediction of calibers (an application in this sense was shown by Román-Román and Torres-Ruiz [37]). In both cases, it is logical to think that the limit values reached by individuals are not the same, and that they depend on the starting values reached at the first moment of observation. For that reason, a reformulation of the T-growth curve is appropriate. To deal with this issue, from initial condition  $S(t_0) = S_0$  it follows that  $S(t)$  can be reformulated as

$$S(t) = S_0 \frac{\eta + \exp(-\sinh(\beta t_0))}{\eta + \exp(-\sinh(\beta t))}, \tag{4}$$

where  $\eta = 1/\alpha$  is a parameter. Note that  $\lim_{t \rightarrow \infty} S(t)$  now depends on  $S_0$  and  $t_0$ . Note also that  $\alpha$  (and hence  $\eta$ ) becomes a parameter, instead of a function of the initial values. With this formulation, the expression for the inflection time becomes

$$\frac{\eta - \exp(-\sinh(\beta t^*))}{\eta + \exp(-\sinh(\beta t^*))} = \tanh(\beta t^*) \operatorname{sech}(\beta t^*), \tag{5}$$

similar to (3) with new parameter  $\eta$ . The value of the curve at the inflection is the same as in the original curve, but replacing parameter  $M$  with  $S_0 \frac{\eta + \exp(-\sinh(\beta t_0))}{\eta}$ .

Crossing times  $\tau_k$  and  $\lambda_n$  can be obtained in a similar way as in the original case, resulting in

$$\tau_k = \frac{1}{\beta} \operatorname{arcsinh} \log \frac{kM}{\eta(S_0 - kM) + S_0 e^{-\sinh(\beta t_0)}}, \quad \frac{\eta}{\eta + e^{-\sinh(\beta t_0)}} \leq k < 1,$$

$$\lambda_n = \frac{1}{\beta} \operatorname{arcsinh} \log \frac{n}{\eta(1 - n) + e^{-\sinh(\beta t_0)}}, \quad 0 < n < \frac{\eta + e^{-\sinh(\beta t_0)}}{\eta}.$$

It is important to note that, following this reformulation, the curve can now be obtained as solution of the linear differential equation

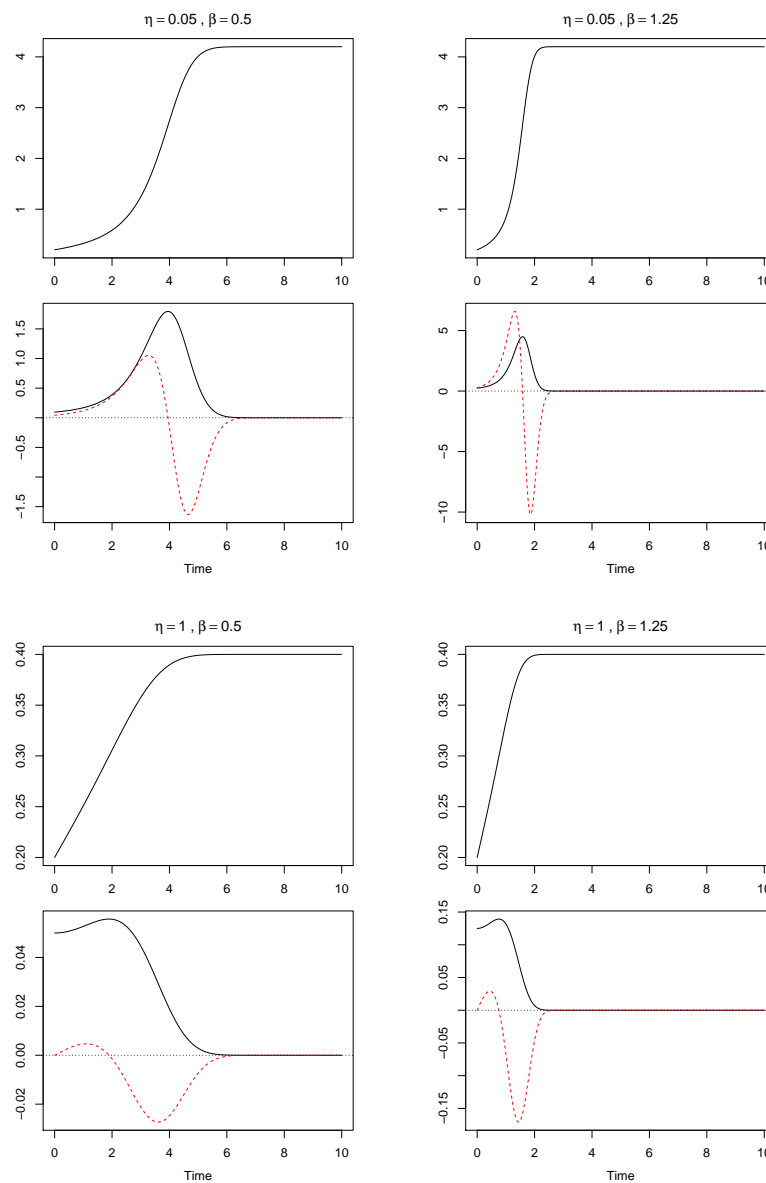
$$\frac{dS(t)}{dt} = h_{\xi}(t)S(t),$$

where

$$h_{\xi}(t) = \frac{e^{-\sinh(\beta t)}}{\eta + e^{-\sinh(\beta t)}} \cosh(\beta t)\beta, \quad \xi = (\eta, \beta)^T. \tag{6}$$

Indeed, its logarithmic derivative is independent of the curve itself, and mathematical treatment is therefore greatly simplified. At the same time, this approach allows the use of stochastic differential equation with explicit solutions, as demonstrated in the following section.

Figure 2 shows the T-growth curve for different values of  $\eta$  and  $\beta$ . The presence of inflections is determined mainly by  $\eta$ , whereas  $\beta$ , being the growth rate parameter, affects the position of such inflection. Beside each plot of the curve, the first and second derivatives are shown.



**Figure 2.** (Top) T-growth curve for different values of  $\eta$  and  $\beta$ , starting at  $S_0 = 0.2$ ; and (Bottom) first (black line) and second (red dashed line) derivative.

#### 4. Stochastic Model

Deterministic models such as the T-growth curve may be useful to predict and describe the behavior of dynamic phenomena, as they take into account known variables and the relations between them. Nevertheless, in real life, factors remain that are unknown to the researcher. In addition, errors or misleading interpretations of the measurements (caused sometimes from the instruments themselves) may arise during the observation phase and the collection of data. Finally, the random nature of many processes and affecting factors may have an influence on the dynamics under study. For all of these reasons, the use of deterministic models in such studies is not advised. Stochastic models, on the contrary, are able to address these issues and provide an effective way to study, analyze, and describe dynamic phenomena such as the ones concerning growth. For this reason, this section proposes a stochastic counterpart to the T-growth model.

There are different ways to construct such model, but many of them result in equations and processes that are, due to their complexity, useless from a theoretical or even practical point of view. Here, a lognormal diffusion process with exogenous factors, based on the T-growth curve, is proposed. This model can be viewed as a *nice* one, meaning that many useful properties and explicit expressions are satisfied.

The construction of the model relies on the property, satisfied by the reformulated curve (4),  $\frac{d}{dt}S(t) = h_{\zeta}(t)S(t)$  where  $h_{\zeta}(t)$  is the real continuous function given in (6). Such function can be viewed as the time-dependent fertility rate of a Malthusian model. A stochastic version follows from the perturbation of such fertility, by replacing  $h_{\zeta}(t)$  by  $h_{\zeta}(t) + \zeta(t)$ , where  $\zeta(t)$  is a white noise with variance  $\sigma^2$ . As a result, the stochastic differential equation

$$dX_{\zeta'}(t) = h_{\zeta}(t)X_{\zeta'}(t)dt + \sigma X_{\zeta'}(t)dW(t), \quad \zeta' = (\eta, \beta, \sigma)^T,$$

follows, where  $W(t)$  is the standard Wiener process, independent on initial value  $X_0 := X_{\zeta'}(t_0)$  for every  $t \geq t_0$ . The continuity of function  $h_{\zeta}(t)$  implies the existence and uniqueness of a solution. In this case, such solution  $X_{\zeta'}(t)$  is meant to be a diffusion process for which an explicit formula is given by

$$X_{\zeta'}(t) = X_0 \exp\left(\int_{t_0}^t h_{\zeta}(u)du - \frac{\sigma^2}{2}(t - t_0) + \sigma(W(t) - W(t_0))\right).$$

Its finite-dimensional distributions, provided  $X_0$  is either a degenerate variable (i.e.,  $P[X_0 = x_0] = 1$ ) or lognormal distributed  $\Lambda_1[\mu_0, \sigma_0]$  are lognormal. That is, for  $n \in \mathbb{N}$  and time instants  $t_1 < \dots < t_n$ , the random vector  $(X_{\zeta'}(t_1), \dots, X_{\zeta'}(t_n))^T$  follows an  $n$ -dimensional distribution  $\Lambda_n[\varepsilon, \Sigma]$ , where

$$\begin{aligned} \varepsilon_i &= \mu_0 + \int_{t_0}^{t_i} h_{\zeta}(u)du - \frac{\sigma^2}{2}(t_i - t_0), \\ \sigma_{ij} &= \sigma_0^2 + \sigma^2(\min(t_i, t_j) - t_0) \end{aligned}$$

are the components of vector  $\varepsilon$ , for  $i, j = 1, \dots, n$ , and matrix  $\Sigma$ , respectively.

Then, it follows that the conditional distribution is also lognormal, that is, for  $s < t$  and state  $y$ ,

$$X_{\zeta'}(t)|X_{\zeta'}(s) = y \sim \Lambda_1\left[\log y + \int_s^t h_{\zeta}(u)du - \frac{\sigma^2}{2}(t - s); \sigma^2(t - s)\right].$$

This leads to the explicit form of the probability transition density function, allowing future inference procedures such as maximum-likelihood estimation with the exact density and metaheuristic optimization of an objective function derived from such density.



Finally, the theory about lognormal processes allows us to find explicit expressions for many useful functions. Following Román-Román et al. [11], many of these characteristics can be expressed in a unique way by the function

$$G_{\xi'}^{\rho}(t|y, s) = \left( \frac{\eta + e^{-\sinh(\beta s)}}{\eta + e^{-\sinh(\beta t)}} \right)^{\rho_1} \exp\left( \rho_1 \left( y - \frac{\sigma^2}{2}(t - s) \right) + \rho_2 \left( \rho_3 \sigma_0^2 + \sigma^2(t - s) \right)^{\rho_4} \right),$$

for a vector  $\rho = (\rho_1, \rho_2, \rho_3, \rho_4)^T$  and a generic state  $y$ , being  $s < t$ . Different values of  $\rho$ ,  $s$  and  $y$  lead to some interesting characteristics, as summarized in Table 1.

**Table 1.** Some characteristics of the T-growth process, according to the values of  $G_{\xi'}^{\rho}(t|z, s)$ . Here,  $Q_{\alpha}$  is the quantile function and  $z_{\alpha}$  is the  $\alpha$ -quantile of the standard normal distribution.

Function of the Process	$z$	$s$	$\rho$
$E[X_{\xi'}(t)^n]$	$\mu_0$	$t_0$	$(n, n^2/2, 1, 1)^T$
$E[X_{\xi'}(t)^n   X_{\xi'}(s) = z]$	$\log y$	$s$	$(n, n^2/2, 0, 1)^T$
$\text{Mode}[X_{\xi'}(t)]$	$\mu_0$	$t_0$	$(1, -1, 1, 1)^T$
$\text{Mode}[X_{\xi'}(t)   X_{\xi'}(s) = z]$	$\log y$	$s$	$(1, -1, 0, 1)^T$
$Q_{\alpha}[X_{\xi'}(t)]$	$\mu_0$	$t_0$	$(1, z_{\alpha}, 1, 1/2)^T$
$Q_{\alpha}[X_{\xi'}(t)   X_{\xi'}(s) = z]$	$\log y$	$s$	$(1, z_{\alpha}, 0, 1/2)^T$

Please note that, in the case of the mean function of the process, it follows

$$E[X_{\xi'}(t)] = E[X_0] e^{\int_{t_0}^t h_{\xi}(u) du} = E[X_0] \frac{\eta + e^{-\sinh(\beta t_0)}}{\eta + e^{-\sinh(\beta t)}},$$

which is the T-growth curve starting at  $E[X_0]$ .

### 5. Inference on the Model

The present section addresses the estimation of the parameters by means of maximum-likelihood procedure. In the particular case of the T-growth model, there are only two parameters, namely  $\eta$  and  $\beta$ , in addition to the diffusion coefficient  $\sigma^2$ .

Let us consider the observation of  $d$  sample paths at time instants  $t_{ij}$  for  $i = 1, \dots, d$  and  $j = 1, \dots, n_i$ . For the sake of expositional clarity, we suppose that  $t_{i1}$ , i.e., the first observation time, is the same for all sample paths, that is,  $t_{i1} = t_0 \geq 0$  for all  $i$ . Now, if  $\mathbf{X}_i^T = (X(t_{i1}), \dots, X(t_{in_i}))^T$ , we denote  $\mathbf{X} = (\mathbf{X}_1^T | \dots | \mathbf{X}_d^T)^T$  as the matrix of observations. Let us denote  $x_{ij}$  as the observation of variable  $X(t_{ij})$  at instant  $t_{ij}$ . Such observations form a vector  $\mathbf{x}$ . On the other hand, let us consider an initial lognormal distribution, i.e.,  $X_0 \sim \Lambda_1[\mu_1, \sigma_1^2]$  of parameters  $\mu_1$  and  $\sigma_1^2 > 0$ .

To address the maximum-likelihood estimation, and regarding the conditional distribution of the process (and hence the knowledge of its probability transition density function), we deal with the corresponding log-likelihood function:

$$\begin{aligned} \log L(\mathbf{x}; \mu_1, \sigma_1^2, \xi') &= -\frac{1}{2} \left( n \log(2\pi) + d \log \sigma_1^2 + (n - d) \log \sigma^2 \right) \\ &\quad - \sum_{i=1}^d \left( \log x_{i1} + \frac{1}{2\sigma^2} (\log x_{i1} - \mu_1)^2 \right) \\ &\quad - \sum_{i=1}^d \sum_{j=2}^{n_i} \left( \log x_{ij} + \frac{1}{2} \log \Delta_{ij} + \frac{1}{2\sigma^2 \Delta_{ij}} \left( Q_{ij} - T_{ij}^{\xi'} + \frac{\sigma^2}{2} \Delta_{ij} \right)^2 \right), \end{aligned} \tag{7}$$

for any observation set  $\mathbf{x} = (x_{ij}) \in \mathbf{X}$ , where  $n = \sum_{i=1}^d n_i$  and, for every  $i = 1, \dots, d$  and  $j = 1, \dots, n_i$ , the following grouped variables have been defined:

$$Q_{ij} = \log \frac{x_{ij}}{x_{ij-1}},$$

$$T_{ij}^{\xi} = \log \frac{\eta + \exp(-\sinh(\beta t_{ij-1}))}{\eta + \exp(-\sinh(\beta t_{ij}))},$$

$$\Delta_{ij} = t_{ij} - t_{ij-1}.$$

The optimization of this function can be carried out by different methods. The direct one is to solve the maximum-likelihood equations. To this end, numerical methods might be required. Nevertheless, the convergence conditions imposed by numerical methods such as Newton–Raphson and its derivatives cannot always be guaranteed. As a matter of fact, such methods require calculating (or at least approximating) the inverse of the Hessian matrix. The existence of this inverse cannot always be guaranteed given the dependence of the system of likelihood equations on the observed data. On other occasions, this inverse may be badly conditioned and give rise to unstable solutions. Additionally, these methods require initial solutions whose determination is not always easy to carry out both due to the difficulty in calculating them and because of their optimality to favor the convergence of the numerical method. Another way to solve the maximum-likelihood estimation is based on the use of metaheuristic algorithms. Such algorithms have the ability to find the local and global minima (or maxima) of a function within a bounded region of the parametric space. These boundaries depend on the properties of the parameters and the original model.

### 5.1. Bounded Parametric Space

Many metaheuristic algorithms, including those used in this work, require a bounded region within the parametric space. The Grey Wolf Optimizer, for example, will not work properly if the wolves are unable to meet each other while moving without limitations. For that reason, an initial region to search maxima or minima could avoid issues related with the convergence of the algorithm.

The stochastic T-growth model has two parameters, named  $\eta, \beta > 0$ , which do not have, a priori, upper bounds. However, in the context of growth phenomena, some assumptions could lead to these parameters being totally bound.

Indeed, we suppose that there exists at least one inflection point in our model. This is due to the required flexibility of the model and its relation with inflections. If observations could be explained by a curve with no inflections, many other simpler models would be preferable over the T-growth.

#### 5.1.1. Case $t_0 \neq 0$

By assuming that there exists an inflection at time  $t^*$ , inflection condition (5) holds for every  $\eta$  and  $\beta$ . Let us then define auxiliary functions  $v_1(t) = \frac{\eta - \exp(-\sinh(\beta t))}{\eta + \exp(-\sinh(\beta t))}$  and  $v_2(t) = \tanh(\beta t) \operatorname{sech}(\beta t)$ . Then (5) could be expressed as

$$w(t^*) = 0,$$

where  $w(t) = v_1(t) - v_2(t)$  for all  $t$ . After some algebra, it follows that  $v_1$  is monotonically increasing and  $\lim_{t \rightarrow \infty} v_1(t) = 1$ . Since  $v_1(0) = \frac{\eta - 1}{\eta + 1}$ , it follows that the range of  $v_1$  is the interval  $\left[\frac{\eta - 1}{\eta + 1}, 1\right)$ . Analogously, it can be shown that  $\lim_{t \rightarrow \infty} v_2(t) = 0$ , with  $v_2(0) = 0$  with a maximum at  $t = \frac{1}{2}$ . It follows that the range of  $v_2$  is  $\left[0, \frac{1}{2}\right]$ . Given this, we can conclude that  $w(t) = 0$  has no solutions when the ranges of  $v_1$  and  $v_2$  do not intersect, that is, when  $\frac{\eta - 1}{\eta + 1} > \frac{1}{2}$ . Solving for  $\eta$  shows that, if inflections do exist, then  $\eta \leq 3$ . Furthermore, a more

precise upper bound can be obtained by solving (5) for  $\eta$ . Indeed, for an inflection time  $t^*$ , it follows that

$$\eta = \frac{\cosh^2(\beta t^*) + \sinh(\beta t^*)}{\cosh^2(\beta t^*) - \sinh(\beta t^*)} e^{-\sinh(\beta t^*)}.$$

Then,  $\eta$  can be viewed as a function  $f_\eta(\beta)$  of  $\beta$ . Then,  $f_\eta(0) = 1$  and  $\lim_{\beta \rightarrow \infty} f_\eta(\beta) = 0$ , i.e., for high values of the growth rate, and when there exists an inflection, parameter  $\eta$  must decrease in order to maintain such flexibility. The study of  $\frac{d}{d\beta} f_\eta(\beta)$  leads to the existence of a maximum at  $\tilde{\beta} = \frac{1}{t^*} \operatorname{arcsinh} \sqrt{z}$ , where  $z = (\sqrt{3} - 3)/2 \approx 0.526$ . Then, the maximum value of  $\eta$  is  $f_\eta(\tilde{\beta}) \approx 1.42$ . With this, it can be concluded that, for a T-growth model with at least one inflection point, it holds that  $\eta \in [0, 1.42]$ .

On the other hand, a bounded interval of  $\beta$  can be obtained from the expression of the carrying capacity. Indeed, from the original expression of such asymptotic value  $M$  and its expression in the reformulation, it follows that

$$M = s_0 \frac{\eta + \exp(-\sinh(\beta t_0))}{\eta}.$$

Solving for  $\beta$  leads to a function of  $\eta$ ,  $f_\beta(\eta)$ ,

$$f_\beta(\eta) = \frac{1}{t_0} \operatorname{arcsinh} \left( -\log \left( \left( \frac{M}{S_0} - 1 \right) \eta \right) \right).$$

This function verifies that  $\lim_{\eta \rightarrow 0} f_\beta(\eta) = \infty$ , but from  $\frac{1}{f_\beta(\eta)} \frac{d}{d\eta} f_\beta(\eta)$  it follows that the growth is slow for lower values of  $\eta$ . This implies that, for a positive lower bound of  $\eta$ , and even of high orders such as  $10^{-10}$ , a bounded interval for  $\beta$  follows. Furthermore, being  $\beta$  the growth rate in the context of growth phenomena, it would be reasonable to consider  $\beta \geq 0$ . That is, if  $\eta \in [\epsilon_\eta, 1.42]$ , for  $\epsilon_\eta \leq 10^{-10}$ , then  $\beta \in [0, f_\beta(\epsilon_\eta)]$ .

### 5.1.2. Case $t_0 = 0$

The methodology described above requires that  $t_0 \neq 0$  in order to get the expression of  $f_\beta(\eta)$  and hence the interval of  $\beta$ . However, when  $t_0 = 0$ , this approach cannot be accomplished. In such case, provided the relation  $\eta = \frac{S_0}{M - S_0}$  between  $\eta$  and the carrying capacity, and where  $x_{in_i}$  is the last observed value of the  $i$ th path, we propose that  $\eta$  be bound as follows

$$\frac{x_0}{\max_{i=1, \dots, d} (x_{in_i}) - x_0} \leq \eta \leq \frac{x_0}{\min_{i=1, \dots, d} (x_{in_i}) - x_0}.$$

Then, solving (5) at each one of those bounds of  $\eta$  and inflection time  $t^*$ , corresponding bounds for  $\beta$  follow. Such equation might be solved, for example, by applying the bisection method. Even in such case, the quest for an interval where the root (here, a bound of  $\beta$ ) lies, may lead to two minimal intervals where the two bounds lie. Another possibility here is to consider a mixture of those two intervals, by taking the minimum and the maximum of the four values as the lower and upper ends, respectively, of the interval of  $\beta$ .

### 5.2. Metaheuristic Approaches

In this subsection, the metaheuristic algorithms used to estimate the maximum-likelihood parameters of the T-growth model are introduced. The common framework for these procedures comprises a bounded parametric space and an objective function to optimize. The parametric space can be bounded according to the aforementioned methodology.

On the other hand, the objective function is obtained from the log-likelihood function (7) by getting rid of constants and initial-value terms. Indeed, the function to be optimized is

$$f_o(\eta, \beta, \sigma^2) = -\frac{d(n-1)}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^d \sum_{j=2}^{n_i} \frac{1}{\Delta_{ij}} \left( Q_{ij} - T_{ij}^{\xi} + \frac{\sigma^2}{2} \Delta_{ij} \right)^2.$$

### 5.2.1. Clonal Selection Algorithm

The “Clonal Selection” algorithm (CLONALG) was introduced by Castro and Zuben [16], and it is based on the clonal selection principle, which studies the immune response to antigenic stimuli. Related to evolutionary algorithms, clonal selection has shown great performance in engineering, as illustrated in [17,38]. The algorithm describes two populations, one of antigens and one of antibodies, and the relations of affinity between them. The basic steps can be summarized as follows:

1. Initialize random population of antigens.
2. Determine the affinity to all antibodies.
3. Select the antibodies showing the highest affinity and clone them according to their antigenic affinities.
4. Mature generated clones inversely to the antigenic affinity (high affinity, small mutation rate).
5. Select those with the highest affinity to antigens and replace if it is better.
6. Check the stopping criteria (maximum number of iterations or good fitness).

### 5.2.2. Harmony Search Algorithm

The “Harmony Search” algorithm (HS) was introduced by Geem et al. [18] and improved in many recent works (e.g., [19]). This algorithm tries to mimic the process of improvisation performed by music players. With fewer mathematical requirements than other heuristic algorithms, it has been successful in its application to several optimization problems (see [20]). Its steps are as follows:

1. Initialize a matrix of random values, sorted by the value of the objective function, called “Harmony Memory” (HM).
2. Improvise a new harmony (generated according to the parameters) after considerations of HM.
3. Update HM by including the new harmony (according to fitness, and replacing if better) and excluding the worst harmony.
4. Check the stopping criteria (maximum number of improvisations).

### 5.2.3. Grey Wolf Optimizer

This algorithm, inspired by nature, was introduced by Mirjalili et al. [14] and attempts the leadership structure of wolf packs. Some of its results have been successfully applied, and this in turn has produced some interesting modifications (see [39,40]). The algorithm is based on the hierarchy of wolf populations, in which individuals are classified as leaders (both male and female) or  $\alpha$ , followed by  $\beta$  and  $\delta$  in decreasing order or leadership. The rest of the wolves are tagged as  $\omega$ . The iterations consist of updating the leadership according to fitness. The rest of the wolves move towards their leaders. Its steps can be summarized as:

1. Initialize the random population of wolves and established the hierarchy according to the fitness coming from the objective function.
2. Update the positions of wolves according to the position of leaders. This updating is performed in several stages, such as hunting, searching, or attacking the prey.
3. Replace the leaders according to better fitness.
4. Check the stopping criteria and return  $\alpha$  as the optimal value.

#### 5.2.4. Moth Flame Optimizer

The “Moth Flame Optimizer” (MFO) (see [15]) is a population-based algorithm inspired by the particular navigation method exhibited by moths. These insects fly with a fixed angle with respect to the Moon. This technique, called transverse orientation, allows them to fly long distances efficiently. Some works about the applications and modifications of this highly successful optimizer are found in [41,42]. The algorithm considers two populations, one of moths and one of flames (flashes). The former act as search agents across the parametric space, whereas the latter represent the matrix of best solutions according to the fitness of the objective function. By this procedure, moths are ensured to never discard the best solution. Its steps are as follows:

1. Initialize random population of moths and calculate fitness according to objective function. This leads to a flame as the best position so far.
2. Update position of the moths moving around the flames, which decrease in number with every iteration.
3. Update flame position according to better solutions.
4. Check the stopping criteria (maximum number of iterations).

#### 5.2.5. Whale Optimization Algorithm

The “Whale Optimization Algorithm” (WOA) was introduced by Mirjalili and Lewis [13] and has shown very good results in, among others, the study of combinatorial problems. This algorithm draws from nature to mimic the social behavior of humpback whales. This species has a particular way of hunting called “bubble-net”. It creates bubbles in a circular pattern in order to hunt their prey by driving it closer to the surface of water. The algorithm models such ascending spirals of bubbles in order to perform the optimization process. Its basic steps can be summarized as follows:

1. Initialize a population of random whales and calculate fitness according to the objective function.
2. Update whale positions randomly or according to the bubble-net technique.
3. Update the best search agent (whale), which is the one approaching the best fitness.
4. Check the stopping criteria (maximum number of iterations).

## 6. Simulation Study

The metaheuristic algorithms described above were applied to simulated data from the T-growth stochastic process in order to study the robustness of the proposed methodology concerning the bounded parametric space. Specifically, 25 sample paths were simulated in the same time interval,  $[t_0, 30]$ , for  $t_0 \geq 0$ . The time step was calculated for a sample size  $n$ , which takes values 11, 31, and 51. The initial variable here was degenerate at  $x_0 = 9$ , meaning that  $P[X_0 = x_0] = 1$ . Parameters  $\eta$  and  $\beta$  were fixed at values 0.25 and 0.1, respectively. Furthermore, diffusion coefficient  $\sigma$  took values 0.01, 0.05, and 0.1 in order to illustrate the performance of the algorithms in the parametric space under conditions of low as well as high variability. For each algorithm and every combination of parameters, 50 replications were made, and their mean was taken as the final value. Regarding the metaheuristic algorithms, none required parameter tuning, and the default values recommended by their respective authors were used.

The results were determined by considering the relative absolute error between the observed mean and the mean of the simulated process with the estimated values. The expression of such measure is

$$\text{RAE} = \frac{1}{N} \sum_{i=1}^N \frac{|m_i - \hat{E}[X(t_i)]|}{m_i},$$

where  $m_i$  is the value of the observed mean at time  $t_i$  and  $\hat{E}[X(t_i)]$  is the value of the mean of the estimated process at the same time.  $N$  is the corresponding sample size in every case.

The results are summarized in Table 2, for all combinations between  $\sigma$  and  $n$ . It is shown that all the metaheuristics perform very well despite the sample size or the variability. The parametric space was bounded according to the proposed methodology, and it was the same for all five algorithms.

Estimated values for parameters  $\eta$ ,  $\beta$ , and  $\sigma$  are shown for each algorithm in the rest of the tables. For each estimation, the corresponding mean relative error is shown, which was calculated according to

$$\text{MRE} = \frac{1}{N} \sum_{i=1}^N \frac{|\alpha - \hat{\alpha}|}{\alpha},$$

where  $\alpha$  is the corresponding parameter and  $\hat{\alpha}$  its estimated value.

As displayed in Table 2, the “Moth Flame Optimizer” algorithm performs very well in all possible scenarios, even when sample size is small and variability is high. The estimated parameters obtained by this metaheuristic are shown in Table 3, where the maximum MRE is just 0.0067, corresponding to the case  $\sigma = 0.1$  and  $n = 11$ . The best estimation obtained with this algorithm corresponds to the case  $\sigma = 0.01$  and  $n = 51$ . This result is expected due to the low variability and large sample size. Nevertheless, as illustrated by Table 2, the best estimation of MFO occurs when  $\sigma = 0.01$  and  $n = 11$ . This is explained by measure MRE being affected by the deviation of any parameter, whereas the RAE takes the mean of the deviations with respect to the estimated mean (taking into account the estimation of  $\sigma$ ).

Besides MFO, the other algorithms also perform well, although they present higher errors in cases of high variability. This is shown in detail in Tables 4–7 for the WOA, GWO, CLONALG, and HS algorithms, respectively.

In particular, the aforementioned tables show that, in general, MRE decreases as the number of points  $n$  increases, except maybe for the case  $\sigma = 0.05$ , where it remains the same or even increases slightly. In any case, values are low and acceptable. The final results for the parameters are essentially the same despite the algorithm employed. Focusing on each parameter, a maximum variation of approximately 0.04 is apparent for values of  $\eta$  across all algorithms. This variation occurs when  $n = 31$  and  $n = 51$  for  $\sigma = 0.1$  (the maximum and minimum estimates of  $\eta$ , respectively). Of course, such high variability has an impact on the solutions and appears to be particularly sensitive to variations of  $n$ . Nevertheless, MRE decreases across all algorithms for such cases. On the other hand, results for parameter  $\beta$ , mainly involved in the flexibility of the curve, show a maximum variation below 0.004. In particular, the cases  $\sigma = 0.01$  with  $n = 11$  and  $\sigma = 0.05$  with  $n = 31$  present the largest differences, although these remain very small.

All of these results show the robustness of the strategy proposed to bound the parametric space. As a matter of fact, and despite their disparate behaviors, all the algorithms were able to find optimal values with low errors. A particular case of interest is that of low sample sizes. This is a very common case in many practical applications, which makes algorithm performance within the bounded parametric space all the more important. The tables mentioned above show very good performance for low sample sizes. Additionally, even when errors naturally decrease as sample size increases in nearly every algorithm, the estimations when  $n = 11$  remain quite good. Once again, MFO shows lower errors for small sample sizes even when variability is high. Nevertheless, the values yielded by the rest of the algorithms remain good, to an extent that guarantees the robustness of the procedure and the applicability of this model and strategy to the description of growth phenomena in which few data observations are available.

Finally, Figure 3 shows a comparison between the observed and the estimated means for every algorithm in every case. The conclusions drawn are the same, with variations being somewhat higher in the case  $\sigma = 0.1$ . In particular, simulated mean paths for the results of every algorithm are shown against the sample mean observation. Cases are sorted by increasing values of  $\sigma$  from top to bottom and increasing values of  $n$  from left to right. As can be expected, figures in the first row show small differences between the paths. On the contrary, the last row plots show wider variability due to value  $\sigma = 0.1$ . With

respect to  $n$ , such differences are less noticeable in the plots, but  $\sigma = 0.05$  illustrates that high values of  $n$  lead to slightly larger distances between paths.

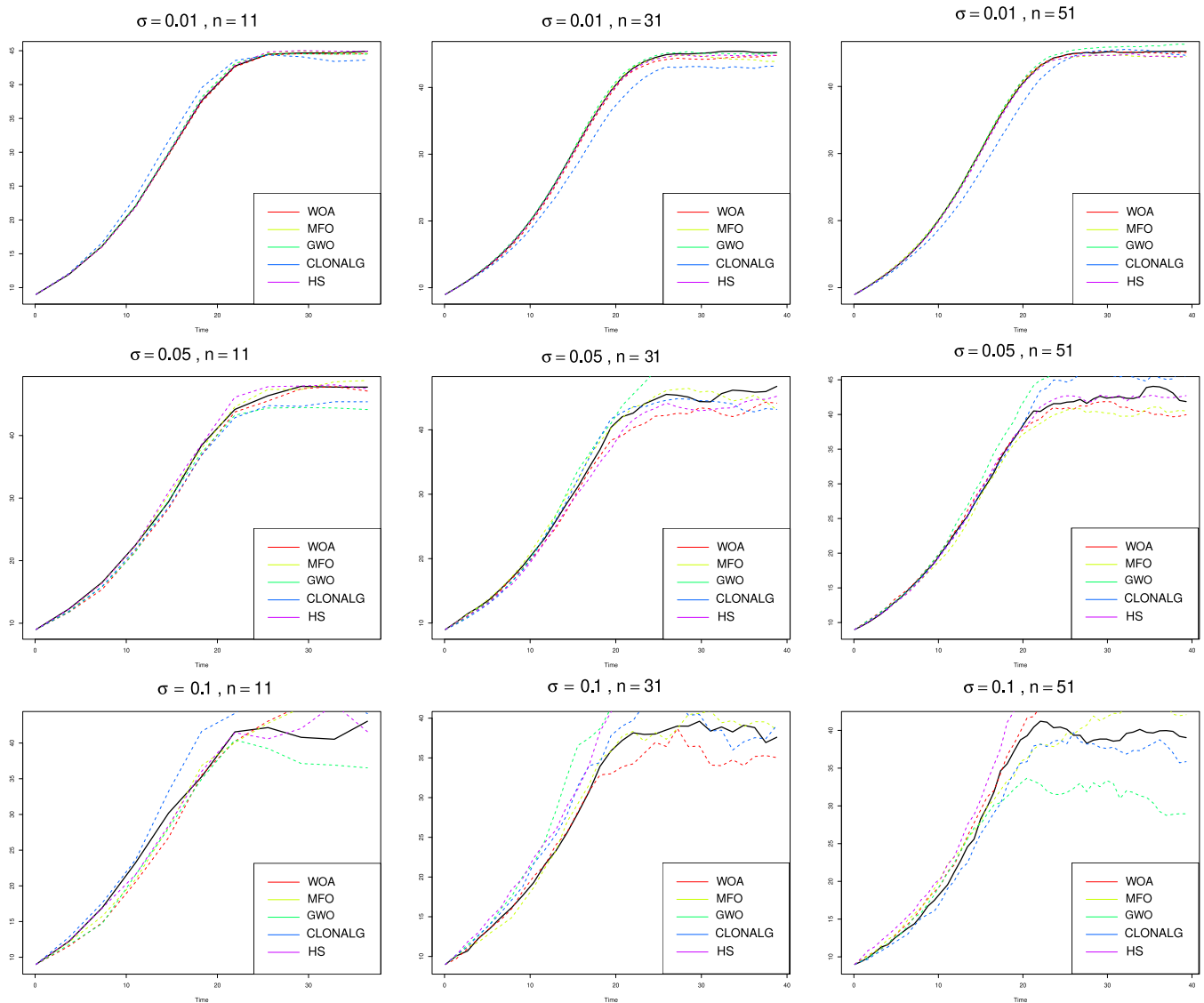


Figure 3. Comparison of observed mean (solid black line) and estimated means for every algorithm.

Table 2. Relative absolute errors for every metaheuristic algorithm in the case of simulated T-growth process with parameters  $\eta = 0.25$  and  $\beta = 0.1$  and different values of  $\sigma$  and  $n$ .

		WOA	MFO	GWO	CLONALG	HS
$\sigma = 0.01$	$n = 11$	0.004596	0.004203	0.005010	0.027998	0.004583
	$n = 31$	0.013944	0.008499	0.007087	0.051982	0.005861
	$n = 51$	0.003792	0.009805	0.009740	0.037694	0.008499
$\sigma = 0.05$	$n = 11$	0.025560	0.015626	0.040875	0.036209	0.014687
	$n = 31$	0.039790	0.022904	0.062858	0.030954	0.036023
	$n = 51$	0.027456	0.036456	0.082239	0.036688	0.013407
$\sigma = 0.1$	$n = 11$	0.074296	0.038146	0.092243	0.037846	0.047159
	$n = 31$	0.047714	0.02956951	0.1718793	0.054826	0.1773981
	$n = 51$	0.107721	0.062996	0.123859	0.050686	0.215345

**Table 3.** Estimated parameters and mean relative error of the MFO algorithm in the case of  $\eta = 0.25$  and  $\beta = 0.1$ , for different values of  $\sigma$  and  $n$ .

		$\hat{\eta}$	$\hat{\beta}$	$\hat{\sigma}$	MRE
$\sigma = 0.01$	$n = 11$	0.2509756	0.1009127	0.01064425	0.02581848
	$n = 31$	0.2478262	0.09944957	0.009785756	0.01187461
	$n = 51$	0.2471421	0.09995375	0.01005502	0.005798523
$\sigma = 0.05$	$n = 11$	0.2409002	0.09842545	0.05054152	0.02099172
	$n = 31$	0.2455311	0.09914826	0.05329188	0.03074345
	$n = 51$	0.26337	0.09822727	0.05227099	0.03887568
$\sigma = 0.1$	$n = 11$	0.2681699	0.101778	0.1109459	0.0666395
	$n = 31$	0.2710887	0.09861253	0.09997612	0.03282277
	$n = 51$	0.2353388	0.09912472	0.09934843	0.0246378

**Table 4.** Estimated parameters and mean relative error of the WOA algorithm in the case of  $\eta = 0.25$  and  $\beta = 0.1$ , for different values of  $\sigma$  and  $n$ .

		$\hat{\eta}$	$\hat{\beta}$	$\hat{\sigma}$	MRE
$\sigma = 0.01$	$n = 11$	0.2509266	0.100913	0.01066472	0.02643622
	$n = 31$	0.2469859	0.09944247	0.009953973	0.007411521
	$n = 51$	0.2471157	0.09995652	0.01005198	0.005723296
$\sigma = 0.05$	$n = 11$	0.2403702	0.09844525	0.05055765	0.0217399
	$n = 31$	0.2452724	0.09915115	0.05329859	0.03112355
	$n = 51$	0.263635	0.09819216	0.05228017	0.03940729
$\sigma = 0.1$	$n = 11$	0.2672859	0.1017511	0.1110155	0.06560311
	$n = 31$	0.2720849	0.09859512	0.09996334	0.0342517
	$n = 51$	0.2380585	0.09985639	0.09908654	0.01944554

**Table 5.** Estimated parameters and mean relative error of the GWO algorithm in the case of  $\eta = 0.25$  and  $\beta = 0.1$ , for different values of  $\sigma$  and  $n$ .

		$\hat{\eta}$	$\hat{\beta}$	$\hat{\sigma}$	MRE
$\sigma = 0.01$	$n = 11$	0.2509764	0.1009246	0.01064277	0.02580951
	$n = 31$	0.2478205	0.09944689	0.009782874	0.01198724
	$n = 51$	0.2471215	0.09996527	0.01004665	0.005508826
$\sigma = 0.05$	$n = 11$	0.2408722	0.09842666	0.050547	0.02106147
	$n = 31$	0.2454393	0.0991375	0.05328924	0.03088416
	$n = 51$	0.2633211	0.09822432	0.0522742	0.03884178
$\sigma = 0.1$	$n = 11$	0.2681796	0.1017863	0.110939	0.06665699
	$n = 31$	0.2709916	0.09860349	0.09996272	0.03276809
	$n = 51$	0.2354084	0.09913233	0.09935355	0.02450257

**Table 6.** Estimated parameters and mean relative error of the CLONALG algorithm in the case of  $\eta = 0.25$  and  $\beta = 0.1$ , for different values of  $\sigma$  and  $n$ .

		$\hat{\eta}$	$\hat{\beta}$	$\hat{\sigma}$	MRE
$\sigma = 0.01$	$n = 11$	0.252627	0.107977	0.02083098	0.3911254
	$n = 31$	0.2479919	0.09313908	0.01602853	0.2264982
	$n = 51$	0.2473206	0.0902483	0.01440435	0.1828897
$\sigma = 0.05$	$n = 11$	0.2401952	0.09537204	0.05049653	0.03180981
	$n = 31$	0.2486393	0.1027989	0.05409315	0.03843151
	$n = 51$	0.2610024	0.09328716	0.05313177	0.05792447
$\sigma = 0.1$	$n = 11$	0.2689292	0.102819	0.1094837	0.0662479
	$n = 31$	0.2692818	0.09915501	0.1035572	0.04038295
	$n = 51$	0.2342171	0.09737465	0.09926846	0.03223356



**Table 7.** Estimated parameters and mean relative error of the HS algorithm in the case of  $\eta = 0.25$  and  $\beta = 0.1$ , for different values of  $\sigma$  and  $n$ .

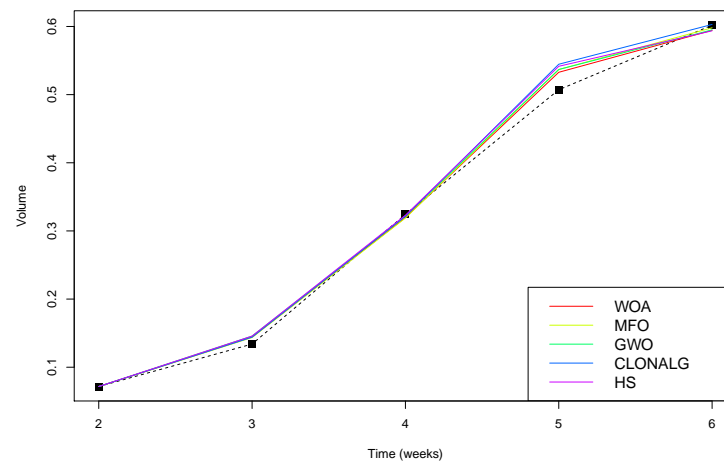
		$\hat{\eta}$	$\hat{\beta}$	$\hat{\sigma}$	MRE
$\sigma = 0.01$	$n = 11$	0.2507133	0.1004446	0.01060876	0.02272496
	$n = 31$	0.2477269	0.09915076	0.009438699	0.02457161
	$n = 51$	0.2469593	0.09944281	0.01004033	0.007255965
$\sigma = 0.05$	$n = 11$	0.240688	0.09844413	0.05050159	0.02094608
	$n = 31$	0.2455214	0.09907743	0.05316942	0.0301762
	$n = 51$	0.2633633	0.09814548	0.05208191	0.03787881
$\sigma = 0.1$	$n = 11$	0.2680154	0.1016684	0.1109767	0.06617093
	$n = 31$	0.271004	0.0984852	0.09997166	0.03314918
	$n = 51$	0.2364888	0.09758188	0.0982698	0.03184272

### 7. Application to Real Data

In this section, the proposed methodology is applied to real data extracted from the work of Song et al. [43]. The dataset concerns the volume of tumors in nasopharyngeal cancer, a disease with high incidence in Southern China. Despite classic treatments such as radiotherapy, chemotherapy, or immunotherapy, local recurrence and distant metastasis are frequent among patients. One technology, called RNA interference (RNAi), uses post-transcriptive gene silencing and has shown good performance inhibiting tumor growth by silencing one single gene involved in tumor development. In [43], it is shown that the effect of RNAi in the reduction of the tumor volume is greater when it is applied to multiple involved genes at once than when applied to a single gene. Nasopharyngeal tumor volume is measured for a number of weeks in several groups according to the affected gene. The involved genes are VEGF, C-myc, Survivin, and hTERT. In the work introducing the T-growth curve, Tabatabai et al. [30] fitted the model to data related to the VEGF gene, showing better results than other classic models such as the logistic or Gompertz curves. In this section, the proposed T-growth stochastic model is fitted to data for all four genes. Following the methodology, parameter estimation was performed considering the five different metaheuristic algorithms with  $t_0 = 2$ ,  $n = 5$  and a time step of 1 (weeks). The results are shown in Table 8. Given these estimations, five T-growth processes were simulated in order to compare their mean paths with the original data (see Figure 4). Again, Table 8 shows that relative absolute errors are low and very similar, with MFO and GWO yielding the best results. In particular, the MFO algorithm boasts great performance, as it did with simulated data. It is worth noting that the process and the proposed methodology are here applied to a small dataset,  $n = 5$ . The results are consistent with others in simulations with few data points like  $n = 11$ . The T-growth stochastic model could thus be used in this context to describe the evolution of tumor volume for the RNAi technique acting over a variety of genetic origins. Furthermore, in [30], the T-growth curve was introduced in a system of equations concerning the dynamics of effector and tumor cells. Such proposal opens possibilities for using the T-growth stochastic model within systems of stochastic differential equations describing more complex behaviors in the dynamics of the RNAi technique and tumor volume.

**Table 8.** Application to real data: estimated parameters and RAE for every metaheuristic algorithm in the case of tumor volume data under RNAi (data from [43]).

	$\hat{\eta}$	$\hat{\beta}$	$\hat{\sigma}$	RAE
WOA	0.04543004	0.4721142	0.08919208	0.03083607
MFO	0.04543903	0.4720851	0.09028120	0.03051089
GWO	0.04545478	0.4720773	0.08920575	0.03291900
CLONALG	0.04551454	0.4718910	0.08924891	0.03187722
HS	0.04539460	0.4718072	0.08921148	0.03449431



**Figure 4.** Application to real data: nasopharyngeal mean tumor volume (black dashed line with squares) and the mean of T-growth model simulated paths according to the estimated parameters followed from every metaheuristic algorithm (data from [43]).

## 8. Conclusions

In the present work, a new stochastic growth model is proposed. It is based on the T-growth curve, which is an extension of the classic logistic growth curve. This new model has some interesting features which make it a suitable choice to describe growth phenomena without compromising efficiency in the estimation of the parameters. As a matter of fact, the T-growth model only has two parameters, in addition to the diffusion coefficient and a simple analytic expression. The use of the hyperbolic sine inside the exponential function leads to a significant increase in the flexibility of the curve, and the presence of inflections makes the model a good fit for a wide range of growth data.

In many practical applications, parameter estimation is a major issue. In this case, the low number of parameters and simplicity of the analytical formulation make this easier to address, although the estimation may require the use of metaheuristic algorithms.

Parameter estimation is particularly interesting given the model's possibilities for real-world applications. Maximum likelihood estimation leads to a system of equations that cannot be solved explicitly, and therefore numerical resolution methods are required. The fact that the verification of the applicability hypotheses of common numerical methods cannot always be guaranteed makes it advisable to study alternatives, such as metaheuristic algorithms.

As a general rule, these procedures require a bounded parametric space to work properly, which in the previous pages is obtained by means of a strategy that employs the T-growth model.

The robustness of such strategy was observed after studying the performance of five different well-known metaheuristic algorithms. All such algorithms were able to estimate parameters  $\eta$ ,  $\beta$ , and  $\sigma$  with low relative absolute errors. Simulations were carried out under different values of parameter  $\sigma$  and sample size  $n$ . The reason these were chosen is that, in practical applications, it must be known whether the process and the estimation are able to perform well with few data values and in conditions of great variability.

The possible limitations of this model become more evident in contexts which require a high number of parameters. However, given its simplicity, its low number of parameters, and the robustness of the methodology developed for bounding the parametric space when it is applied to several metaheuristic algorithms, the T-growth stochastic model appears as a good choice to describe, analyze, and make predictions within many different fields of practical application.

Future research lines concerning this model could rely on the study of asymptotic distributions of the estimators among other statistics properties derived from the diffusion process. In addition, a more practically oriented approach may follow from the study of dynamic statistical variables such as stopping times or first-passage times. Finally, the introduction of the model into dynamic systems (by taking advantage of its simplicity) may lead to more sophisticated ways to analyze and predict complex behaviors.

**Author Contributions:** Investigation, A.B., P.R.-R. and F.T.-R.; Writing original draft, A.B., P.R.-R. and F.T.-R.; Writing–review and editing, A.B., P.R.-R. and F.T.-R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the Ministerio de Economía, Industria y Competitividad, Spain, under Grant MTM2017-85568-P and FEDER/Junta de Andalucía-Consejería de Economía y Conocimiento, Spain, Grant A-FQM-456-UGR18.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank the three anonymous reviewers for their suggestions that have improved the content of the paper.

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

## References

- Ram, Y.; Dellus-Gur, E.; Bibi, M.; Karkare, K.; Obolski, U.; Feldman, M.W.; Cooper, T.F.; Berman, J.; Hadany, L. Predicting microbial growth in a mixed culture from growth curve data. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 14698–14707. [\[CrossRef\]](#)
- Spiess, A.N.; Rödiger, S.; Burdukiewicz, M.; Volksdorf, T.; Tellinghuisen, J. System-specific periodicity in quantitative real-time polymerase chain reaction data questions threshold-based quantitation. *Sci. Rep.* **2016**, *6*, 1–10. [\[CrossRef\]](#)
- Spiess, A.N.; Feig, C.; Ritz, C. Highly accurate sigmoidal fitting of real-time PCR data by introducing a parameter for asymmetry. *BMC Bioinform.* **2008**, *9*, 221. [\[CrossRef\]](#)
- 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.* **2020**, *377*, 125143. [\[CrossRef\]](#)
- Lee, L.; Atkinson, D.; Hirst, A.G.; Cornell, S.J. A new framework for growth curve fitting based on the von Bertalanffy Growth Function *Sci. Rep.* **2020**, *10*, 7953. [\[CrossRef\]](#)
- Hening, A.; Nguyen, D.H.; Yin, G. Stochastic population growth in spatially heterogeneous environments: The density-dependent case. *J. Math. Biol.* **2018**, *76*, 697–754. [\[CrossRef\]](#)
- Parsons, T.L. Invasion probabilities, hitting times, and some fluctuation theory for the stochastic logistic process. *J. Math. Biol.* **2018**, *77*, 1193–1231. [\[CrossRef\]](#) [\[PubMed\]](#)
- 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.* **2015**, *266*, 579–598. [\[CrossRef\]](#)
- Asadi, M.; Di Crescenzo, A.; Sajadi, F.A.; Spina, S. A generalized Gompertz growth model with applications and related birth-death processes. *Ric. Mat.* **2020**. [\[CrossRef\]](#)
- Ascione, G.; Mishura, Y.; Pirozzi, E. Fractional Ornstein-Uhlenbeck Process with Stochastic Forcing, and its Applications. *Methodol. Comput. Appl. Probab.* **2021**, *23*, 53–84. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- 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.* **2021**, *92*, 884–904. [\[CrossRef\]](#)
- Mirjalili, S.; Lewis, A. The Whale Optimization Algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [\[CrossRef\]](#)
- Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey Wolf Optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [\[CrossRef\]](#)
- Mirjalili, S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowl.-Based Syst.* **2015**, *89*, 228–249. [\[CrossRef\]](#)
- Castro, L.N.d.; Zuben, F.J.V. Learning and optimization using the clonal selection principle. *IEEE Trans. Evol. Comput.* **2002**, *6*, 239–251. [\[CrossRef\]](#)
- Campelo, F.; Guimaraes, F.G.; Igarashi, H.; Ramirez, J.A. A clonal selection algorithm for optimization in electromagnetics. *IEEE Trans. Magn.* **2005**, *41*, 1736–1739. [\[CrossRef\]](#)
- Geem, Z.W.; Kim, J.H.; Loganathan, G. A New Heuristic Optimization Algorithm: Harmony Search. *Simulation* **2001**, *76*, 60–68. [\[CrossRef\]](#)

19. Mahdavi, M.; Fesanghary, M.; Damangir, E. An improved harmony search algorithm for solving optimization problems. *Appl. Math. Comput.* **2007**, *188*, 1567–1579. [[CrossRef](#)]
20. Wang, L.; Hu, H.; Liu, R.; Zhou, X. An improved differential harmony search algorithm for function optimization problems. *Soft Comput.* **2019**, *23*, 4827–4852. [[CrossRef](#)]
21. Tabatabai, M.; Williams, D.; Bursac, Z. Hyperbolastic growth models: Theory and application. *Theor. Biol. Med. Model.* **2005**, *2*, 14. [[CrossRef](#)] [[PubMed](#)]
22. Bursac, Z.; Tabatabai, M.; Williams, D.K. Nonlinear Hyperbolastic Growth Models and Applications in Cranofacial and Stem Cell Growth. In Proceedings of the American Statistical Association Biometrics Section. American Statistical Association, Alexandria, VA, USA, 16–21 July 2005; pp. 190–197.
23. 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](#)]
24. 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](#)]
25. Deasy, B.M.; Jankowski, R.J.; Payne, T.R.; Cao, B.; Goff, J.P.; Green-Berger, J.S.; Huard, J. Modeling stem cell population growth: Incorporating terms for proliferative heterogeneity. *Stem Cells* **2003**, *21*, 536–545. [[CrossRef](#)] [[PubMed](#)]
26. Tabatabai, M.A.; Eby, W.M.; Singh, K.P. Hyperbolastic modeling of wound healing. *Math. Comput. Model.* **2011**, *53*, 755–768. [[CrossRef](#)]
27. Tabatabai, M.A.; Kengwoung-Keumo, J.J.; Eby, W.M.; Bae, S.; Guemmegne, J.T.; Manne, U.; Fouad, M.; Partridge, E.E.; Singh, K.P. Disparities in Cervical Cancer Mortality Rates as Determined by the Longitudinal Hyperbolastic Mixed-Effects Type II Model. *PLoS ONE* **2014**, *9*, e107242. [[CrossRef](#)]
28. 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](#)]
29. Eby, W.M.; Tabatabai, M.A. Methods in Mathematical Modeling for Stem Cells. In *Stem Cells and Cancer Stem Cells*; Hayat M., Ed.; Springer: Dordrecht, The Netherlands, 2014; Volume 12, pp. 201–217. [[CrossRef](#)]
30. 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](#)]
31. Gutiérrez, R.; Román, P.; Torres, F. A note on the Volterra integral equation for the first-passage-time probability density. *J. Appl. Probab.* **1995**, *32*, 635–648. [[CrossRef](#)]
32. Gutiérrez, R.; Ricciardi, L.; Román, P.; Torres, F. First-passage-time densities for time-non-homogeneous diffusion processes. *J. Appl. Probab.* **1997**, *34*, 623–631. [[CrossRef](#)]
33. Luz-Sant’Ana, I.; Román-Román, P.; Torres-Ruiz, F. Modeling oil production and its peak by means of a stochastic diffusion process based on the Hubbert curve. *Energy* **2017**, *133*, 455–470. [[CrossRef](#)]
34. Barrera, A.; Román-Román, P.; Torres-Ruiz, F. A hyperbolastic type-I diffusion process: Parameter estimation by means of the firefly algorithm. *Biosystems* **2018**, *163*, 11–22. [[CrossRef](#)] [[PubMed](#)]
35. 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](#)]
36. Barrera, A.; Román-Román, P.; Torres-Ruiz, F. Two stochastic differential equations for modeling oscillabolastic-type behavior. *Mathematics* **2020**, *8*, 155. [[CrossRef](#)]
37. Román-Román, P.; Torres-Ruiz, F. Forecasting fruit size and caliber by means of diffusion processes. Application to “Valencia late” oranges. *J. Agric. Biol. Environ. Stat.* **2014**, *19*, 292–313. [[CrossRef](#)]
38. Gálvez, A.; Iglesias, A.; Ávila, A.; Otero, C.; Arias, R.; Machado, C. Elitist clonal selection algorithm for optimal choice of free knots in B-spline data fitting. *Appl. Soft Comput.* **2015**, *26*, 90–106. [[CrossRef](#)]
39. Rodríguez, A.; Camarena, O.; Cuevas, E.; Aranguren, I.; Valdivia-G, A.; Morales-Castañeda, B.; Zaldívar, D.; Pérez-Cisneros, M. Group-based synchronous-asynchronous Grey Wolf Optimizer. *Appl. Math. Model.* **2021**, *93*, 226–243. [[CrossRef](#)]
40. Rebello, G.; de Oliveira, E.J. *Modified Binary Grey Wolf Optimizer*; Springer Tracts in Nature-Inspired Computing; Springer: Singapore, 2020; pp. 148–179. [[CrossRef](#)]
41. Kotary, D.K.; Nanda, S.J. Distributed robust data clustering in wireless sensor networks using diffusion moth flame optimization. *Eng. Appl. Artif. Intell.* **2020**, *87*, 103342. [[CrossRef](#)]
42. 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* **2020**, *8*, 87378–87392. [[CrossRef](#)]
43. Song, Y.; Dong, M.M.; Yang, H.F. Effects of RNA interference targeting four different genes on the growth and proliferation of nasopharyngeal carcinoma CNE-2Z cells. *Cancer Gene Ther.* **2011**, *18*, 297–304. [[CrossRef](#)] [[PubMed](#)]