Journal of Theoretical Biology 263 (2010) 59–69

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

A diffusion process to model generalized von Bertalanffy growth patterns: Fitting to real data

Patricia Román-Román *, Desirée Romero, Francisco Torres-Ruiz

Departamento de Estadística e Investigación Operativa, Facultad de Ciencias, Universidad de Granada. Avda. Fuentenueva s/n 18071 Granada, Spain

ARTICLE INFO

Article history:
Received 7 May 2009
Received in revised form
26 October 2009
Accepted 4 December 2009
Available online 16 December 2009

Keywords:
Diffusion process
Stochastic growth
Von Bertalanffy curve

ABSTRACT

The von Bertalanffy growth curve has been commonly used for modeling animal growth (particularly fish). Both deterministic and stochastic models exist in association with this curve, the latter allowing for the inclusion of fluctuations or disturbances that might exist in the system under consideration which are not always quantifiable or may even be unknown. This curve is mainly used for modeling the length variable whereas a generalized version, including a new parameter $b \geq 1$, allows for modeling both length and weight for some animal species in both isometric ($b=3$) and allometric ($b \neq 3$) situations.

In this paper a stochastic model related to the generalized von Bertalanffy growth curve is proposed. This model allows to investigate the time evolution of growth variables associated both with individual behaviors and mean population behavior. Also, with the purpose of fitting the above-mentioned model to real data and so be able to forecast and analyze particular characteristics, we study the maximum likelihood estimation of the parameters of the model. In addition, and regarding the numerical problems posed by solving the likelihood equations, a strategy is developed for obtaining initial solutions for the usual numerical procedures. Such strategy is validated by means of simulated examples. Finally, an application to real data of mean weight of swordfish is presented.

1. Introduction

Von Bertalanffy (1938) introduced an equation to study the growth of individuals belonging to several types of animal populations. As most growth models, it comes from an adaptation of the Verhulst logistic growth by assuming a maximal value of the growth variable (which might eventually be attained), and considering the growth rate as proportional to the difference between maximal and current value. It is currently the most common model used by fishery biologists to study growth in fish and its interpretations, such as fish population dynamics and the effects of fishery regulations on the catch.

The von Bertalanffy growth curve is

$$L(t) = L_\infty[1-e^{-k(t-a)}], \quad t \geq a, \;\; k > 0, \tag{1}$$

where L_∞ is the upper bound for the variable under study, that can only be reached after infinity time and k is the *curvature* parameter, or von Bertalanffy growth rate that determines the speed with which the fish attains L_∞. As regards the parameter a, sometimes named the *initial condition parameter*, it determines the time at which the fish has a size equal to zero and could be negative (in the biological literature is common to note this parameters as t_0, but we have used a in order to avoid confusion with the notation usually employed for the initial instant in the context of stochastic processes). From a biological point of view, this question is meaningless because growth in the embryonic stage usually does not fit the von Bertalanffy growth pattern. Nevertheless, it is important to note that fish sufficiently aged to be exploited, as for consumption, show a trend modeled by the von Bertalanffy curve. The embryonic stage has then no interest in this context.

A general expression for the von Bertalanffy curve, also called "generalized von Bertalanffy growth curve" (see García-Rodríguez et al., 2005 and references therein), is

$$B(t) = B_\infty[1-e^{-k(t-a)}]^b, \quad t \geq a, \;\; k > 0, \;\; b \geq 1, \tag{2}$$

where the parameter b can be known or unknown. For example, the value $b=1$ (curve (1)) is used when the variable under study is the length. However, when focusing on the weight, and taking into account the existing relation between the weight and the length, the value $b=3$ is associated with the *isometric* growth, whereas the case $b \neq 3$ is related to the *allometric* growth.

Curves (1) and (2) provide appropriate deterministic models to describe the growth (in length or weight) of fish and others animals. The study of procedures in order to determine the parameters of such curves for fitting real data has been widely

* Corresponding author. Tel.: +34 958240491; fax: +34 958243267.
E-mail address: proman@ugr.es (P. Román-Román).

0022-5193/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2009.12.009

considered. For example, Rafail (1973) developed a procedure based on a straight-line relationship between the natural logarithms of growth increments per unit of age against age as the independent variable.

Nevertheless, such models do not include variability among individuals of the same age, or environmental variability (fluctuations or disturbances that might exist in the system under consideration).

In order to take these variations into account, several stochastic models have been considered in the literature. The first type of model, according to the classification provided by Russo et al. (2009), are those described by a stochastic differential equation, obtained by including a noise term in the ordinary differential equation of the respective deterministic model.

In this context, and as concerns the Bertalanffy growth model, Lv and Pitchford (2007) have recently considered three stochastic models, associated with the curve (1). Said models are built from stochastic differential equations with identical von Bertalanffy deterministic parts and different stochastic terms, to show, among other points, that stochasticity can have positive impact on fish recruitment. Gudmundsson's (2005) criticism of such models is that the solution for these differential equations may not be strictly increasing. From this point of view, the aforementioned models may not be appropriate for the study of some growth variables, like length, that are not subject to decreasing (Weatherly and Gill, 1987).

Such defect, real from a theoretical point of view, is practically diluted in practice. In the practical application of this type of models, the noise term has a moderate magnitude (see Gutiérrez et al., 2007 in the context of Gompertz-type growth). This feature makes the sample paths of the resulting processes show von Bertalanffy-like behavior patterns, except for the presence of small random fluctuations. In addition, if the resulting stochastic model verifies that its mean function is a von Bertalanffy growth curve, the use of the model for fitting and predicting is fully justified. On the other hand, such defect disappears if the stochastic model is associated with the von Bertalanffy generalized curve. This allows for the study of weight-to-age (generalizing the case for length-to-age) since the evolution of the animal's weight in time is not necessarily a strictly increasing function.

The second type of stochastic models are those that assume that the parameters of the von Bertalanffy curve are different for each member of the population and are thus considered random variables with a certain probability distribution.

Among the authors that deal with this second type of models, we must emphasize the one by Cheng and Kuk (2002), that considered the parameters of the model as random effects following a trivariate normal distribution. More recently, Tovar-Ávila et al. (2009) considered a reparametrization of the von Bertalanffy growth rate using three different probability distribution functions: Weibull, gamma and log-normal. Such models show an increasing behavior and consider individual variations, but are limited by the fact that they do not account for environmental variations, caused by multiple factors which are not always quantifiable or may even be unknown.

Russo et al. (2009), in order to overcome the flaws of the two types of models already mentioned, introduced a type of stochastic models related to the von Bertalanffy curve as a solution for stochastic equations that include a subordinator (a class of strictly increasing stochastic processes which make the solution process of the stochastic equation also increasing). Such models account for both individual and environmental sources of randomness.

In practical situations, the use of a stochastic model for fitting, making forecasts or drawing conclusions about the particular growth of the individuals of a population or about the mean population growth, requires the estimation of the unknown parameters of the model.

As a first approach along this line, some authors, such as Kimura (1980), have studied likelihood methods under the assumption of independent and normally distributed errors using classic nonlinear least square methods. Wang (1999) introduced unbiased estimating functions for a class of growth models that incorporate stochastic components and explanatory variables. More recently, Hart and Chute (2009) introduced a novel linear mixed-effects method for estimating von Bertalanffy growth parameters from growth increment data that lack explicit age information. In Cheng and Kuk (2002) the problem of estimating the parameters of their already mentioned model is considered (see also Laslett et al., 2003 for a discussion of the results).

Russo et al. (2009) also dealt with the problem of parameter estimation based on the probability distributions of the growth variable in each of the observed time instants.

Nevertheless, in the case of models described by stochastic processes that are a solution of stochastic differential equations, an efficient estimation of parameters must be based on data that provide information of the evolution of variables along time (sample paths). Thus such estimation must be based on data which can be of two types:

- Data related to the time evolution of the variable of interest (length, weight, etc.) for each individual of a sample of the population.
- Data relative to the characteristic of interest taken at different time instants from recapture. In this case, data belong to different individuals. This is the most usual case, since measuring growth variables for the same individual at successive time instants can be time-consuming and expensive. For this reason, it is common to consider a single sample path composed of the mean of the recapture data in each time instant.

In the first case, and using several sample paths, it is possible to estimate a model related to the particular growth of individuals, whereas in the second case, and from a single sample path, the estimated model is associated with the mean population growth.

Thus, the aim of this article is twofold. Firstly, the introduction of a stochastic model (specifically a diffusion process) to model behavior patterns associated with the generalized von Bertalanffy curve (2) that allows the consideration of both length and weight for some animal species. To this end, the model is obtained by applying the methodology developed by Gutiérrez et al. (2007) in the context of gompertzian growth: from the deterministic equation, whose solution is the curve of interest, a stochastic component is introduced as well as the condition that the mean function of the resulting diffusion process be a curve of the type (2) (that fits well the sample data and can be used for forecasting purposes) is imposed. The building of the model as well as the study of some of its characteristics (probability distribution of the process, mean, mode and quantile functions) is presented in Sections 2 and 3. For point predictions the mean and the mode functions can be used, whereas for interval predictions, the quantile functions provide intervals containing the growth variable of the process, for each time, with a specific probability. Furthermore, the consideration of a diffusion process in this context will allow the study of time variables, such as first-passage-times, associated with von Bertalanffy type growth models.

Secondly, since the purpose of model is to fit and forecast real data, the next objective is the estimation of the model. In Section 4 an inferential study of the parameters of the process is carried out (particularly their maximum likelihood estimation) on the basis of discrete sampling (in both cases, by considering one or several sample paths). In addition, and regarding the numerical problems posed by solving the likelihood equations, in Section 4.1 a strategy is developed for obtaining initial solutions in order to apply the usual numerical procedures. Such strategy is validated by means of simulated examples in Section 5. Finally, in Section 6 the possibilities of the new stochastic model, relative to fit, forecasting and first-passage-times, are illustrated by means of an application to real data of mean weight of swordfish in Southeastern Pacific.

2. The generalized von Bertalanffy diffusion process

Before introducing the diffusion process, we will obtain an equivalent expression for the generalized von Bertalanffy curve (2), more suitable for our purpose. We will consider that the observations of the variable under study are made from a time instant $t_0 \geq 0$, with $x_0 > 0$ the observed value at t_0. With this realistic hypothesis in mind, and taking into account (2), we deduce that $a < t_0$, since this function is increasing and verifies $B(a) = 0$. On the other hand, as we suppose $B(t_0) = x_0$, and denoting $c = e^{ka}$, we conclude that

$$B_\infty = \frac{x_0}{(1-ce^{-kt_0})^b}, \tag{3}$$

from where

$$B(t) = x_0\left(\frac{1-ce^{-kt}}{1-ce^{-kt_0}}\right)^b, \quad t \geq t_0 > \frac{\ln c}{k}, \ k > 0, \ b \geq 1. \tag{4}$$

Now we introduce the new diffusion process related to the curve (4). To this end, we follow the methodology considered in Gutiérrez et al. (2007) in which several methods are presented in order to introduce a diffusion process in the context of gompertzian growth. Specifically, we look for a process in which the solution of the Fokker–Planck equation, without noise, is such a curve. In addition, the resulting process must verify that the mean function conditioned on the initial value, $E[X(t)|X(t_0) = x_0]$, coincides with (4), which is specially useful for the purposes of forecasting.

Following the same scheme presented there, let us consider the first order equation verified by the transition probability density of the process, $f \equiv f(x,t|x_0,t_0)$:

$$\frac{\partial f}{\partial t} = -\frac{bck}{e^{kt}-c}\frac{\partial}{\partial x}[xf], \quad x > 0, \ t \geq t_0$$

with the initial condition $\lim_{t \to t_0} f(x,t|x_0,t_0) = \delta(x-x_0)$. Its solution is

$$f(x,t|x_0,t_0) = \delta\left(x - x_0\left(\frac{1-ce^{-kt}}{1-ce^{-kt_0}}\right)^b\right),$$

which implies that the population under consideration increases according to (4). Now, we consider the Fokker–Planck equation of the homogeneous lognormal diffusion process with infinitesimal moments $A_1(x) = mx$ and $A_2(x) = \sigma^2 x^2$ $(m, \sigma > 0)$,

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}[mxf] + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}[x^2 f], \quad x > 0, \ t \geq t_0,$$

and modify the infinitesimal mean by multiplying it by the term $bck/(e^{kt}-c)$. Thus, we obtain

$$\frac{\partial f}{\partial t} = -\frac{bck}{e^{kt}-c}\frac{\partial}{\partial x}[xf] + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}[x^2 f], \quad x > 0, \ t \geq t_0, \tag{5}$$

that is, the forward equation of a new diffusion process with infinitesimal moments

$$A_1(x,t) = \frac{bck}{e^{kt}-c}x,$$

$$A_2(x,t) = \sigma^2 x^2. \tag{6}$$

It is obvious that the solution (5), when σ^2 vanishes, is the curve (4).

Alternatively, the process can be obtained from the Langevin equation

$$\frac{dX(t)}{dt} = \frac{bck}{e^{kt}-c}X(t) + X(t)\sigma W(t), \tag{7}$$

where $W(t)$ denotes the standard Wiener process. The derivation of (7) can be achieved from the deterministic growth equation

$$\frac{dx(t)}{dt} = \frac{bck}{e^{kt}-c}x(t), \quad x(t_0) = x_0, \tag{8}$$

which can be seen as a generalization of the malthusian growth model with a deterministic fertility depending on the time, $r(t) = bck/(e^{kt}-c)$, and replacing this fertility with $r(t) + \sigma W(t)$. Moreover, Eq. (7) has an important meaning in terms of how the population increases. In fact, this equation can be expressed in the form

$$\frac{dX(t)}{dt} = mX(t)[1 - \varphi(X(t),t)],$$

where $\varphi(x,t) = 1 - bck/(e^{kt}-c) - (1/m)\sigma W(t)$ is the well-known *regulation function* (see Capocelli and Ricciardi, 1974) that, over time, introduces several changes into m, the growth rate of the Malthusian model.

By rewriting (7) in the usual form for stochastic differential equations, that is

$$dX(t) = \frac{bck}{e^{kt}-c}X(t)dt + \sigma X(t)dW(t), \tag{9}$$

its solution is a non-homogeneous diffusion process $\{X(t); t \leqslant t_0\}$ taking values on $\mathbb{R}^+$ and with infinitesimal moments

$$A_1(x,t) = h(t)x,$$

$$A_2(x,t) = \sigma^2 x^2, \tag{10}$$

where $h(t) = bck/(e^{kt}-c)$ or $h(t) = bck/(e^{kt}-c) + \sigma^2/2$ according to whether the Itô or the Stratonovich integral is used to solve it, respectively. In addition, it is not difficult to show that the mean function, conditioned on the initial value x_0, is

$$E[X(t)|X(t_0) = x_0] = x_0 \exp\left(\int_{t_0}^t h(s)\,ds\right)$$
$$= \begin{cases} x_0\left(\dfrac{1-ce^{-kt}}{1-ce^{-kt_0}}\right)^b & \text{for the Itô solution,} \\ x_0\left(\dfrac{1-ce^{-kt}}{1-ce^{-kt_0}}\right)^b \exp\left(\dfrac{\sigma^2}{2}(t-t_0)\right) & \text{for the Stratonovich solution.} \end{cases} \tag{11}$$

From (11), and taking into account our objectives, the usefulness of the model to fit and forecasting purposes, the Itô solution must be chosen (note that in the case of the Stratonovich solution, the mean is not even a bounded curve), so we introduce a new von Bertalanffy-type diffusion process associated with the curve (4) as the diffusion process $\{X(t); t \geq t_0\}$ defined on $\mathbb{R}^+$ and with infinitesimal moments given by (6).

Finally, another way to develop this process, in the line shown by Albano and Giorno (2006) in the Gompertzian case, is based on the discretization of (8) and then its randomization (see Appendix A for more details).

In short, the diffusion process presented is a stochastic model associated with the generalized von Bertalanffy curve (2), or its rewriting as (4), which accounts for both individual and environmental variability. Its main advantages are:

- Unlike other existing stochastic models, related to the von Bertalanffy curve (1) and useful to study the length of some animal species, the model under discussion allows to study both length and weight.
- Its mean function is a generalized von Bertalanffy curve, which justifies it use for the fitting and prediction of real data with such a behavior pattern.

In addition, one of the advantages of using this sort of model for the study of dynamic phenomena is that it allows to consider some questions about their evolution through time variables. For instance, the study of time variables as the time an animal takes to reach the minimum size at which it can be sold for consumption, or the time a population takes to reach a certain size, can be of great interest. Problems like these may be solved by obtaining the density function of the first-passage-time through a constant boundary. The application developed in Section 6 deals with a problem of this type, illustrating the model's usage for this end.

3. Probability distribution and some characteristics of the process

The probability distribution of the process, determined by the finite-dimensional distributions, can be obtained from the theory of stochastic differential equations, as solution of (9), or employing the theory of partial differential equations from the forward equation, (5), and the Kolmogorov or backward equation

$$\frac{\partial f}{\partial t_0}+\frac{bck}{e^{kt_0}-c}x_0\frac{\partial f}{\partial x_0}+\frac{\sigma^2}{2}x_0^2\frac{\partial^2 f}{\partial x_0^2}=0,\quad x_0>0,\ \ t\geq t_0.$$

In Gutiérrez et al. (2006) both approaches have been developed in the context of the lognormal diffusion process with exogenous factors, of which the von Bertalanffy process is a particular case. In this case, the transition probability density function of the process is

$$f(x,t|y,s)=\frac{1}{x\sqrt{2\pi\sigma^2(t-s)}}\times\exp\left(-\frac{1}{2}\frac{\left[\ln\left(\frac{x}{y}\right)-b\ln\left(\frac{1-ce^{-kt}}{1-ce^{-ks}}\right)+\frac{\sigma^2}{2}(t-s)\right]^2}{(t-s)\sigma^2}\right),\quad t>s,\tag{12}$$

which corresponds to a lognormal distribution, that is,

$$[X(t)|X(s)=y]\sim\Lambda\left[\ln y+b\ln\left(\frac{1-ce^{-kt}}{1-ce^{-ks}}\right)-\frac{\sigma^2}{2}(t-s);(t-s)\sigma^2\right].\tag{13}$$

Because of the Markovian property of the process, from (13) and the initial distribution, we can calculate the finite-dimensional distributions. In this case, we consider two initial distributions: a degenerate distribution, that is, $P[X(t_0)=x_0]=1$, and a lognormal distribution, $X(t_0)\sim\Lambda(\mu_0,\sigma_0^2)$, these choices ensuring that the finite-dimensional distributions are lognormal. We must remark that the former choice can be seen as a particular case of the second considering $\sigma_0=0$ and $\mu_0=\ln(x_0)$. Moreover, the degenerate initial distribution is the real situation when only a sample path is available, whereas the lognormal case requires several trajectories. In any case, the random vector $(X(t_1),\ldots,X(t_n))$ follows an $n-$dimensional lognormal distribution $\Lambda_n(\mu,\Sigma)$, where

$$\mu_i=\mu_0+b\ln\left(\frac{1-ce^{-kt_i}}{1-ce^{-kt_0}}\right)-\frac{\sigma^2}{2}(t_i-t_0),\quad i=1,\ldots,n$$

and

$$\Sigma_{ij}=\sigma_0^2+\sigma^2(\mathrm{Min}\{t_i,t_j\}-t_0),\quad i,j=1,\ldots,n.$$

We now describe the main characteristics of the process, focussing particularly on the three most commonly employed in practice, especially for forecasting purposes. These characteristics are the mean function (which by the structure of the model is a type (2) curve, and thus particularly appropriate for fitting and predicting), the mode function (that provides, for each time, the most probable value of the growth variable) and the quantile functions (which allow to make predictions through intervals that contain the growth variable for each time, with a specific probability). Its expressions can be formulated jointly for the two initial distributions under consideration:

- Mean function:

$$m(t)=E[X(t)]=E[X(t_0)]\left(\frac{1-ce^{-kt}}{1-ce^{-kt_0}}\right)^b,\quad t\geq t_0.$$

- Mode function:

$$\mathrm{Mo}(t)=\mathrm{Mode}[X(t)]=\mathrm{Mode}[X(t_0)]\left(\frac{1-ce^{-kt}}{1-ce^{-kt_0}}\right)^b\times\exp\left(-\frac{3}{2}\sigma^2(t-t_0)\right),\quad t\geq t_0.$$

- Quantile function:

$$C_\alpha(t)=\alpha\text{-}quantile[X(t)]=\alpha\text{-}quantile[X(t_0)]\left(\frac{1-ce^{-kt}}{1-ce^{-kt_0}}\right)^b\times\exp\left(-\frac{\sigma^2}{2}(t-t_0)+z_{1-\alpha}\left[\sqrt{\sigma^2(t-t_0)+Var[\ln(X(t_0))]}-\sqrt{Var[\ln(X(t_0))]}\right]\right),t\geq t_0,$$

where $z_{1-\alpha}$ is the $\alpha-$th quantile of a standard normal distribution.

Furthermore, and from (13), the conditioned versions of these functions can be also obtained (see Gutiérrez et al., 2006 in the general case).

- Conditional mean function:

$$m(t|s)=E[X(t)|X(s)=x_s]=x_s\left(\frac{1-ce^{-kt}}{1-ce^{-ks}}\right)^b,\quad t>s.$$

- Conditional mode function:

$$\mathrm{Mo}(t|s)=\mathrm{Mode}[X(t)|X(s)=x_s]=x_s\left(\frac{1-ce^{-kt}}{1-ce^{-ks}}\right)^b\exp\left(-\frac{3}{2}\sigma^2(t-s)\right),\quad t>s.$$

- Conditional quantile function:

$$C_\alpha(t|s)=\alpha\text{-}quantile[X(t)|X(s)=x_s]=x_s\left(\frac{1-ce^{-kt}}{1-ce^{-ks}}\right)^b\times\exp\left(-\frac{\sigma^2}{2}(t-s)+z_{1-\alpha}\sqrt{\sigma^2(t-s)}\right),\quad t>s.$$

4. Inference on the model

As we have noted above, the mean function of the process is a von Bertalanffy curve of the type (4). Therefore the mean function, as well as its conditional version, $E[X(t)|X(s)=x_s]$, $t>s$, can be useful for making predictions with this model. For this reason, let us examine in this section its maximum likelihood (ML) estimation. First, we obtain the ML estimators of the parameters of the model and then that of the mean function, as well as the other parametric functions above-mentioned.

Let us consider a discrete sampling of the process, based on d sample paths, for times t_{ij}, $(i=1,\ldots,d, j=1,\ldots,n_i)$ with $t_{i1}=t_1$ $i=1,\ldots,d$. That is, we observe variables $X(t_{ij})$, the values of which, $\{x_{ij}\}_{i=1,\ldots,d; j=1,\ldots,n_i}$, make up the sample for the inferential study.

The likelihood function depends on the choice of the initial distribution. When $P[X(t_1)=x_1]=1$, this function is $L_{x_{ij}}(b,c,k,\sigma^2)=\prod_{i=1}^{d}\prod_{j=2}^{n_i}f(x_{ij},t_{ij}|x_{i,j-1},t_{i,j-1})$, where b, c, k and σ^2 are the parameters to be estimated. If $X(t_1)\sim\Lambda(\mu_1,\sigma_1^2)$ the likelihood is $L_{x_{ij}}(\mu_1,\sigma_1^2,b,c,k,\sigma^2)=\prod_{i=1}^{d}f_{X(t_1)}(x_{i1})\prod_{j=2}^{n_i}f(x_{ij},t_{ij}|x_{i,j-1},t_{i,j-1})$.

In the second case there are two additional parameters that must be included in the estimation procedure. Nevertheless, the estimations of μ_1 and σ_1^2 depend only on the initial values and do not influence the estimation of the other parameters. Hence, the ML estimators of b, c, k and σ^2 are the same in both cases.

Henceforth, we will consider the case when the initial distribution is lognormal, because this situation gives full meaning to the model in relation to the comments made above.

In this case (see Appendix B for a more detailed development) by maximizing the likelihood function it can be derived that the ML estimates of μ_1 and σ_1^2 are

$$\hat{\mu}_1=\frac{1}{d}\sum_{i=1}^{d}\ln x_{i1} \quad \text{and} \quad \hat{\sigma}_1^2=\frac{1}{d}\sum_{i=1}^{d}(\ln x_{i1}-\hat{\mu}_1)^2,$$

while the ML estimates for $D=1/c$ and $A=e^{-k}$ result of the solution of the system of Eq. (B.11) for the particular case of equally spaced data, that is $t_{ij}-t_{i,j-1}=h$, $i=1,\ldots,d$; $j=2,\ldots,n_i$. This system has no explicit solution and must be dealt with by numerical methods. Once these estimates are obtained, the corresponding for b and σ^2 are $\hat{b}=b^{\hat{A},\hat{D}}$ and $\hat{\sigma}^2=\sigma^2_{\hat{A},\hat{D}}$ from (B.9). Finally, the ML estimate of any parametric function expressed in terms of b, c, k and σ^2, as for example the mean function, is calculated by applying Zehna's theorem.

4.1. Numerical aspects

The system of Eq. (B.11) is quite complex, which makes its resolution difficult, especially when the sample is large. For this reason, it is necessary to make use of numerical procedures, most of which need an initial solution to be applied.

In order to obtain a good initial solution, we propose some alternatives based on the sample information provided by the observed trajectories of the process. In this line, we will distinguish several cases depending on the parameter b and will use some expressions derived from the curve.

Firstly, and taking into account that $A=e^{-k}$ and $D=1/c$, from (4) we have

$$\frac{x_0}{B_\infty}=\left(1-\frac{A^{t_0}}{D}\right)^b,$$

from where

$$D=\frac{A^{t_0}}{1-\left(\frac{x_0}{B_\infty}\right)^{1/b}}. \tag{14}$$

On the other hand, the inflection occurs at

$$t_I=\frac{\ln(bc)}{k}=\frac{\ln(D/b)}{\ln A},$$

so

$$A=\left(b\left(1-\left(\frac{x_0}{B_\infty}\right)^{1/b}\right)\right)^{1/(t_0-t_I)}. \tag{15}$$

Now, we will distinguish the cases in which b is known and unknown. In each case we will consider two situations according to the inflection time being visualized or not.

- *b known.*
  - Firstly, let us suppose that $t_I>t_0$ (i.e. $b>e^{k\,t_0}/c$). In such a case, and since $(1-1/b)^b$ is the quotient between the value of the curve at t_I and the upper bound, we can approximate t_I, for each trajectory, taking the first time instant at which the sample path exceeds $x_{i,\infty}(1-1/b)^b$, where $x_{i,\infty}$ is the upper bound for the i-th trajectory ($i=1,\ldots,d$), and, finally, consider the mean of these values. Obviously, this procedure can be applied if the limit value is known, at least in a approximate form. In this sense, usually the last value x_{i,n_i} of each trajectory is taken as $x_{i,\infty}$.
    Once t_I is approached, the initial values for A and D are obtained from (14) and (15), respectively. To this end, the value x_0/B_∞ is approached by considering the mean of the values $x_{i1}/x_{i,n_i}$, $i=1,\ldots,d$.
  - If $b=1$ or $\ln(c)/k<t_I\le t_0$ (i.e., $1<b\le e^{k\,t_0}/c$), the inflection time cannot be guessed from the observation of the sample paths. Nevertheless, expression (14) provides a relationship between A and D, or equivalently between k and c, concretely $c=\alpha e^{kt_0}$ where $\alpha=1-(x_0/B_\infty)^{1/b}$. From this expression the curve remains

    $$B(t)=B_\infty(1-\alpha e^{-k(t-t_0)})^b$$

    and presents only an unknown parameter, k. We propose, for each sample path, to calculate k from a least square fit to the previous curve. For this, the value x_0/B_∞ is approached in the same way described in the previous case, whereas B_∞ is approached by means of the values x_{i,n_i}, $i=1,\ldots,d$. Finally, the initial value for A is the exponential of the mean of these estimations.
- *b unknown.*
  - If for each sample path the inflection time can be visualized, or guessed, we propose to find an initial value for b as follows: firstly, we calculate an approximate value for the time at which the inflection occurs, for example by examining the sample paths, and taking the mean of the values $x_{i,t_{I_i}}/x_{i,\infty}$, namely $\overline{x}_I$, where t_{I_i} and $x_{i,\infty}$ are, respectively, the inflection time and the upper bound for the i-th sample path (approached by the last value of the trajectory). Once this value is obtained, the initial one for b is the solution of the equation

    $$\overline{x}_I=\left(1-\frac{1}{b}\right)^b.$$

    After calculating the initial value for b, the corresponding for A and D are obtained from (14) and (15), respectively, taking into account the same remarks made before.
  - If the inflection point cannot be visualized from the sample paths, that is $\ln(c)/k\le t_I\le t_0$, we propose a slight modification of the previous procedure which is based on the increasing behavior of the von Bertalanffy curve. In fact, in such a case, and for $b>1$, the following relation holds:

    $$\left(1-\frac{1}{b}\right)^b=\frac{f(t_I)}{B_\infty}\le\frac{x_0}{B_\infty}.$$

    Thus, and in an approximate form, we can consider that $(1-1/b)^b<x_b$, where $x_b=\text{Mean}_{i=1,\ldots,d}\,x_{i,1}/x_{i,n_i}$. Furthermore,

$(1-1/b)^b$ is a strictly increasing function verifying $(1-1/b)^b < e^{-1}$, $\forall b \geq 1$, from which we can establish the following strategy in order to take an initial value for b in this situation:

- If $x_b < e^{-1}$, then the solution of the equation $\overline{x}_b = (1-1/b)^b$ provides an upper bound for b, namely b_1. In such a case, we propose to consider an initial value for b randomly chosen from an uniform distribution in the interval $(1, b_1)$.
- If $x_b \geq e^{-1}$, then the previous equation has no solution, so we propose to consider $b=1$ as initial value.

Once the value of b is obtained, we proceed as in the case when b is known.

5. Simulation study

In this section we present several examples in order to validate the estimation procedure previously developed, together with the strategy showed for establishing the initial solution of the system of equations that must be solved. To this end, we have simulated sample paths from the von Bertalanffy diffusion process following the algorithm derived from the numerical solution of the stochastic differential equation associated with the process (see Kloeden et al., 1994). In our case, the algorithm is

$$x_{n+1} = x_n \left[1 + h\left(\frac{bck}{e^{k\,t_n} - c} - \frac{\sigma^2}{2} \right) + \frac{h^2}{2} \left(\frac{(bck^2)(bc - e^{k\,t_n})}{(e^{k\,t_n} - c)^2} - \frac{\sigma^2 bck}{e^{k\,t_n} - c} + \frac{\sigma^4}{4} \right) \right.$$
$$\left. + \sigma \left(1 + \frac{hbck}{e^{k\,t_n} - c} - \frac{h\sigma^2}{2} \right) \left(1 + \frac{\sigma}{2} Z_{1n} \right) Z_{1n} + \frac{\sigma^3}{6} Z_{1n}^3 + \frac{\sigma^4}{24} Z_{1n}^4 \right],$$

for $n = 0, \ldots, N-1$, where N is the number of simulated values, h is the integration step, $x_n = X(t_n)$ and Z_{1n} is a normal variable of zero mean and variance h.

We have considered four cases in which $d=100$ sample paths have been simulated. Each trajectory has been simulated with 301 data starting at $t_0 = 0$, taking $h = 0.1$ as the integration step and an initial lognormal distribution $\Lambda(3, 0.2)$. Table 1 shows the values of the others parameters of the process, together with the theoretical inflection time instant. Note that in only in the first example the inflection time instant can be guessed.

In order to make the subsequent inference we have considered, in each case, 31 data with $t_{ij} - t_{i,j-1} = 1$, $i = 1, \ldots, 100$, $j = 1, \ldots, 31$. The graphics in Fig. 1 show the mean of the simulated sample paths for each case.

As concerns the estimation procedure, each of the examples has been treated considering the case b known and unknown.

- For example 1 (in which the inflection occurs inside the considered time interval), if b is known, the procedure aforementioned leads to 2.5 as approximate value or the inflection time (the true value is 2.35). We must note that the time instants considered are discrete; for this reason, when a time instant is found in which the value $x_{i,\infty}(1-1/b)^b$ is exceeded, we have considered as the inflection time the mean between this time and the previous. Finally, from (14) and (15) we have obtained the initial values for A and D and, therefore,

Table 1
Parameters chosen for the simulation of the sample paths.

Example	b	c	k	σ	t_I
1	2	0.8	0.2	0.01	2.35
2	1	0.6	0.5	0.01	−1.02
3	1.3	0.7	0.4	0.01	−0.23
4	3	0.2	0.3	0.01	−1.70

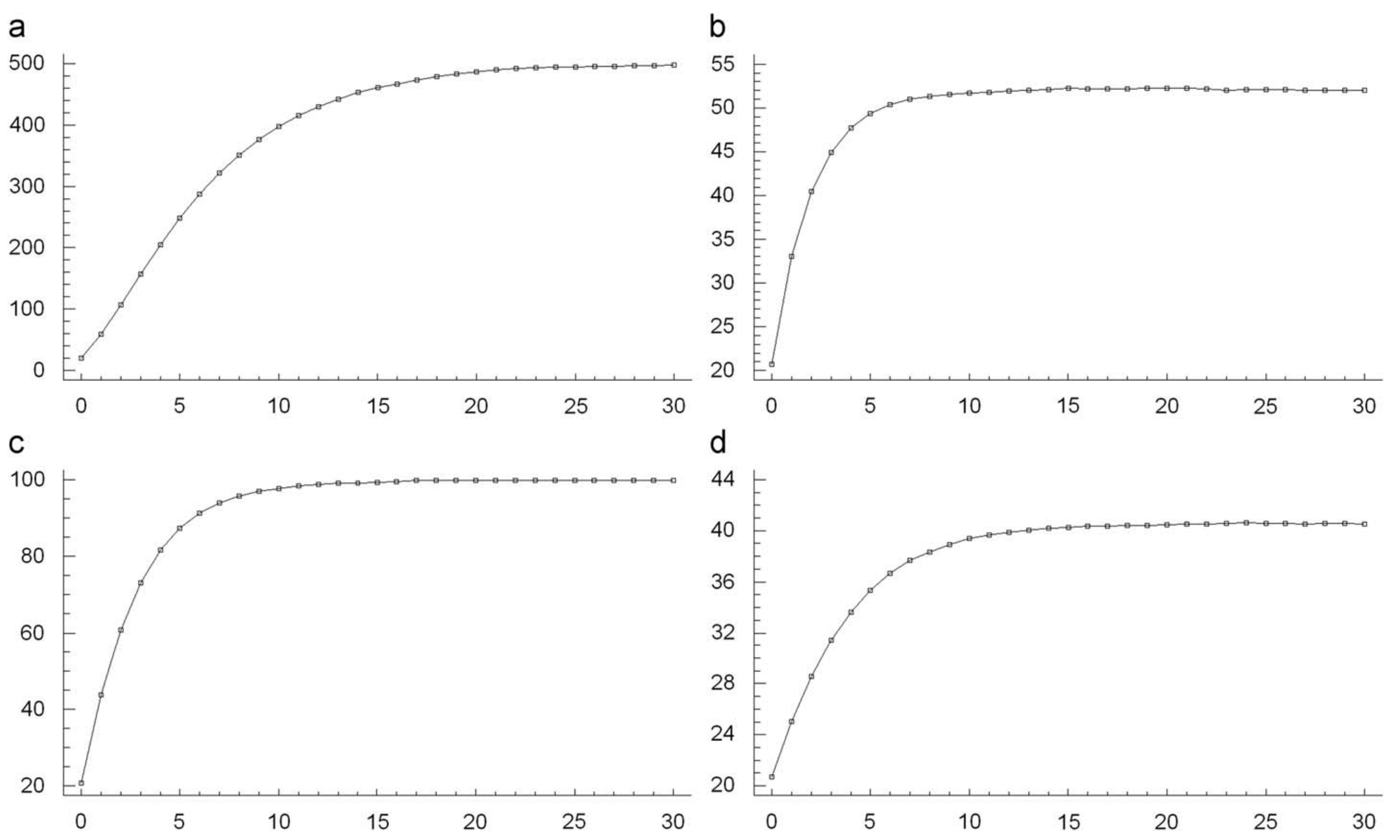

Fig. 1. Mean of the simulated sample paths.

those for c and k. In addition, the value x_0/B_∞ has been approximated by 0.04042, whereas the true value is 0.04.
In the case b unknown, we have considered $t_I = 3$ as the inflection time. From this value, and following the procedure described above, we have obtained $b_0 = 1.22789$, and then the others initial values as we have previously remarked. Table 2 summarizes the results obtained for this example. In the last case, the choice of t_I is arbitrary and it is based on the visualization of the sample paths. For this reason we have considered other possible values for t_I and compared the results with the previous one. Specifically, the values $t_I = 2, 4$ have been also considered without changes in the estimated values (see Table 3).

- For example 2 we must remark that the inflection cannot be guessed (in fact, there is not inflection in the model since $b = 1$). In this example, the initial solution coincides in both cases, that is, when b is known and unknown because of the value obtained for $\bar{x}_b$ is $0.40121 > e^{-1}$. Therefore, when b is unknown, the procedure described leads to consider $b = 1$, from which follows that the initial solution for k (and then for A) and D is the same. Obviously, the final estimation of the parameters does not coincide in both cases since the likelihood equations differ. Table 4 contain the results for this example.
- As example 3 concerns, we have another situation in which the inflection exists but cannot be seen. In this case we must remark that when b is unknown, its initial value is calculated by solving the equation $\bar{x}_b = (1-1/b)^b$ since $\bar{x}_b$ is now $0.20807 < e^{-1}$. Thus, this equation has only one solution, $b_1 = 1.59925$, which is an upper bound for b. Later, we have taken a random value in the interval $(1, b_1)$ as the initial value for b (Table 5 includes the results). Finally, the procedure for taking the initial value for b leads to consider a sensibility analysis of the results. To this end, we have considered other random values in the interval in order to validate the procedure followed (see Table 6). However, no differences have been found, so the estimation of the parameters is robust to the choice of b.
- The last example shows another situation in which the inflection exists but occurs before t_0, so it cannot be seen. The difference between this example and the previous is that, in this case, when b is considered unknown $\bar{x}_b = 0.51208 > e^{-1}$ and then we take $b_0 = 1$ as the initial value for b. Observe that, unlike in example 2, the initial values are different according to b being known or unknown. The results are included in Table 7.

Table 2
Initial and estimated values for example 1.

Parameter	True value	b known		b unknown	
		Initial solution	Estimated value	Initial solution	Estimated value
b	2			1.22789	1.97997
c	0.8	0.79894	0.79994	0.92668	0.80339
k	0.2	0.18747	0.20024	0.04305	0.19847
σ	0.01		0.00991		0.00991

Table 3
Sensibility analysis for example 1 (b unknown).

t_I	b_0	c_0	k_0	b	c	k	σ
2	1.09777	0.94621	0.01899	1.97997	0.80339	0.19847	0.00991
4	1.49086	0.88375	0.06894	1.97997	0.80339	0.19847	0.00991

Table 4
Initial and estimated values for example 2.

Parameter	True value	b known		b unknown	
		Initial solution	Estimated value	Initial solution	Estimated value
b	1			1	1.00201
c	0.6	0.59878	0.60125	0.59878	0.60428
k	0.5	0.50850	0.49515	0.50850	0.49354
σ	0.01		0.00996		0.01011

Table 5
Initial and estimated values for example 3.

Parameter	True value	b known		b unknown	
		Initial solution	Estimated value	Initial solution	Estimated value
b	1.3			1.37891	1.30028
c	0.7	0.70107	0.70075	0.67968	0.70067
k	0.4	0.39513	0.39842	0.40180	0.39847
σ	0.01		0.01017		0.01017

Table 6
Sensibility analysis for example 3 (b unknown).

b_0	c_0	k_0	b	c	k	σ
1.11985	0.75385	0.85843	1.30028	0.70067	0.39847	0.01017
1.17977	0.73568	0.73702	1.30028	0.70067	0.39847	0.01017
1.2397	0.71813	0.62710	1.30028	0.70067	0.39847	0.01017
1.29962	0.70118	0.52597	1.30028	0.70067	0.39847	0.01017
1.35955	0.68484	0.43135	1.30028	0.70067	0.39847	0.01017
1.41947	0.66909	0.34097	1.30028	0.70067	0.39847	0.01017
1.4794	0.65393	0.25291	1.30028	0.70067	0.39847	0.01017
1.53932	0.63934	0.17442	1.30028	0.70067	0.39847	0.01017

Table 7
Initial and estimated values for example 4.

Parameter	True value	b known		b unknown	
		Initial solution	Estimated value	Initial solution	Estimated value
b	3			1	3.10916
c	0.2	0.19995	0.19998	0.40073	0.19361
k	0.3	0.29958	0.30198	0.26639	0.30305
σ	0.01		0.00980		0.00980

6. Application to real data

The following application is based on the studies developed by Chong and Aguayo (2009) on some aspects related to the swordfish age and growth in the Southeastern Pacific. Specifically, we have considered data about the mean weight to age in both sexes, as shown in Fig. 2.

Table 8 shows the estimated values of the parameters of the model, together with the initial solution obtained by means of the strategy proposed before.

Since the model considered verifies that $E[X(t)|X(t_0) = x_0]$ is a generalized von Bertalanffy curve, it is obvious that a good fit to

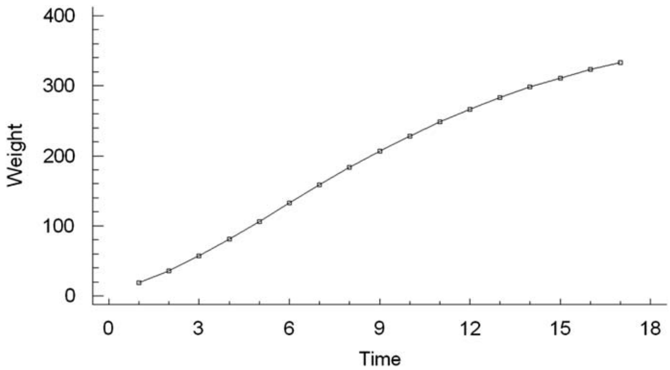

Fig. 2. Mean weight to age for swordfish in both sexes.

Table 8
Estimation of the parameters of the model using all swordfish data.

Parameter	Initial solution	Estimated value
b	4.2379	3.0631
c	0.5901	0.7305
k	0.1833	0.1451
σ		0.0003

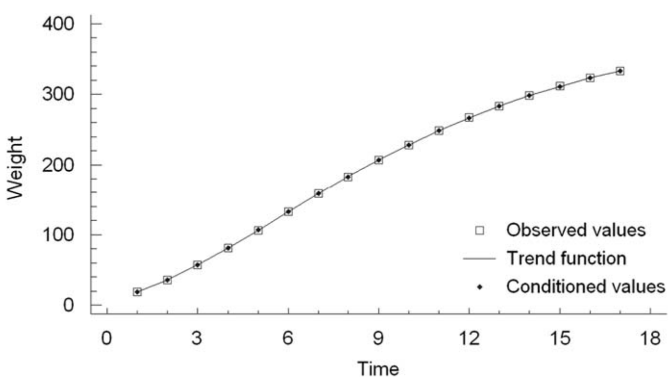

Fig. 3. Observed values and estimations of the trend function and the $E[X(t_i)|X(t_{i-1})=x_{i-1}]$ values, $i=1,\ldots,17$.

the data will be provided by the function $E[X(t)|X(t_1)=x_1]$, where t_1 is the first observation time instant and x_1 is the initial value of the sample path. Nevertheless, and with the aim of obtaining a best fit to the observed data, in practice it is usual to consider the values $E[X(t_i)|X(t_{i-1})=x_{i-1}]$, $i=1,\ldots,17$, where x_{i-1} is the observed value at t_{i-1}.

Fig. 3 shows the observed values, the estimation of the $E[X(t)|X(t_1)=x_1]$ function and the estimation of the $E[X(t_i)|X(t_{i-1})=x_{i-1}]$ values (joined by a solid line for better visualization), $i=1,\ldots,17$.

To illustrate the predictive capability of the model, maximum likelihood estimates for the parameters using data for the age 1–16 were calculated, and we predicted the value for 17. The estimates of the parameters, and the initial solution, are summarized in Table 9, whereas Table 10 shows the observed values at age 17, point prediction given by the estimation of the mean and mode of $X(17)|X(16)=x_{16}$, as well as an interval prediction provided by the estimation of 0.025 and 0.975 quantiles of the variable above.

One advantage of using growth stochastic models, like the one shown here, is the possibility of studying some questions related to their evolution, such as the achievement of a certain value or a change in a behavior pattern. For example, the time an animal takes to reach the minimum size at which it can be sold for consumption, or the time a population takes to reach a certain size, can be of great interest and can be formulated as a problem of first-passage-time through a constant boundary S, that is the time variable

$$T_{S,x_1} = \operatorname*{Inf}_{t \geq t_1} \{t : X(t) > S | X(t_1) = x_1\}$$

for which its probability density function must be calculated. In order to illustrate this question, we have considered two boundaries, $S=300$ and 350. For the first case the value S is achieved inside the time interval considered, whereas the second

Table 9
Estimation of the parameters of the model without the last data.

Parameter	Initial solution	Estimated value
b	5.2573	3.0589
c	0.5067	0.7310
k	0.1959	0.1450
σ		0.0003

Table 10
Observed and predicted values for the mean weight of swordfish at 17-th year.

Observed value	Estimated conditional mean	Estimated conditional mode	Estimated conditional quantiles
333.2	333.331	333.445	(333.092,333.571)

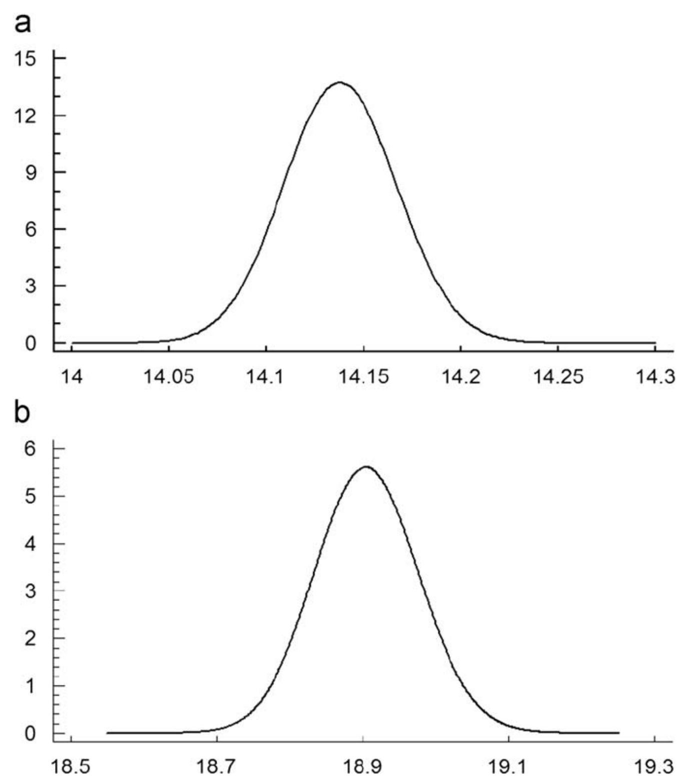

Fig. 4. First-passage probability densities for the weight of the swordfish through a constant boundary: (a) $S=300$; (b) $S=350$.

will be attained later (the estimated upper bound for the model is 405.755).

As regards the obtaining of the probability density function, it is the solution of a Volterra integral equation of the second kind that must be solved by numerical procedures since a closed-form solution is not available in this case (see Gutiérrez et al., 1997). To this end we have employed the methodology developed in Román et al. (2008). Fig. 4 shows the probability density functions for the two cases considered for which, moreover, we can point out that the means of the variables first-passage-time through the constant S are 14.138 and 18.903, respectively.

7. Conclusions

A new stochastic diffusion model has been presented for the modeling of animal growth with a behavior that fits a von Bertalanffy growth pattern. The main innovation from other existing models (see Gudmundsson, 2005; Lv and Pitchford, 2007; Russo et al., 2009) is that it allows for modeling both length and weight for some animal species in isometric and allometric situations. This is done by considering a generalized von Bertalanffy growth curve including a parameter $b \geq 1$, and whose particular case with $b = 1$ is the one considered in previous stochastic models.

The new model is a diffusion process determined by the solution of a differential stochastic equation, which can be obtained by including a noise term in the ordinary differential equation associated with the respective deterministic model. Thus, and in the absence of noise, the growth variable increases according to (2) and the sample paths of the process show such behavior patterns (according to a generalized von Bertalanffy curve), save for the presence of random disturbances (whose magnitude depends on that of the noise). In addition, the fact that the mean function of the process is also a type (2) curve, makes the model particularly interesting for real data applications with fitting and forecasting purposes.

In addition to the mean function, other parametric functions prove useful for fitting and forecasting aims. In the case of point fitting and prediction, the mode function provides, for each time, the most probable value for the growth variable. For interval predictions, through consideration of the percentile function, intervals containing the growth variable for each time (with a specific probability), can be obtained.

Once the diffusion process has been defined, the maximum likelihood estimation of its parameters is done (and, consequently, that of the parametrical functions of interest). However, the likelihood equations obtained do not have an explicit solution and classical numerical procedures are required. The convergence of such procedures may depend on the choice of good initial solutions. For this reason, one of the contributions of the present paper is a proposal for a strategy for the search of such initial solutions in several potential situations. The cases of b known and unknown have been considered and also, in order to use the information on the characteristics of the model provided by the sample data, we have distinguished two possibilities regarding the visualization of the inflection point in the sample paths. Such strategy is validated by means of simulated examples that consider all the situations described.

Finally, one of the main advantages of the use of stochastic models is, in addition to its fitting and forecasting capabilities, the chance to study time variables that affect the growth process. In an application to real data of mean weight of swordfish, the capability of the model for fitting and predicting is shown, as well as for studying time variables of interest like the first-passage-time through constant boundaries.

Acknowledgments

The authors are very grateful to the anonymous referees whose comments and suggestions have contributed to improve this paper. This work was supported in part by the Ministerio de Educación y Ciencia, Spain, under Grants MTM2008-05785/MTM and HI2007-0034, and by the Consejería de Innovación, Ciencia y Empresa (Junta de Andalucía), Spain, under Grant P06-FQM-02271.

Appendix A

Another way to develop the diffusion process proposed in this paper is in the line shown by Albano and Giorno (2006). The procedure is based on the discrete version of (8),

$$x_{(n+1)\tau} - x_{n\tau} = \frac{bck}{e^{kn\tau} - c}\tau x_{n\tau}, \quad n = 0, 1, \ldots \tag{A.1}$$

that approaches (8) when $n \to \infty$ and $\tau \to 0$, with $n\tau = t$.

Now we introduce *random environment* as follows: the relative intrinsic change $bck\tau/(e^{kn\tau} - c)$ in the interval $[n\tau, (n+1)\tau)$, $n = 0, 1, \ldots$ can bee seen as the mean value of a sequence of independent (not identically distributed) Bernoulli variables, $Z_{n\tau}$, $n = 0, 1, \ldots$, verifying

$$P(Z_{n\tau} = \sigma\sqrt{\tau}) = \frac{1}{2} + \frac{bck\sqrt{\tau}}{2(e^{kn\tau} - c)\sigma},$$

$$P(Z_{n\tau} = -\sigma\sqrt{\tau}) = \frac{1}{2} - \frac{bck\sqrt{\tau}}{2(e^{kn\tau} - c)\sigma},$$

where $\sigma > 0$ is a constant that evaluates the width of the environmental fluctuations. Furthermore, the moments of $Z_{n\tau}$ are

$$E[Z_{n\tau}] = \frac{bck}{e^{kn\tau} - c}\tau, \quad E[Z_{n\tau}^2] = \sigma^2\tau, \quad E[Z_{n\tau}^{2+p}] = o(\tau), \quad p \in \mathbb{N}.$$

In this way, we randomize the model by replacing the term $bck\tau/(e^{kn\tau} - c)$ by $Z_{n\tau}$ in (A.1), thus obtaining

$$X_{(n+1)\tau} - X_{n\tau} = Z_{n\tau}X_{n\tau}, \quad n = 0, 1, \ldots.$$

The increments $X_{(n+1)\tau} - X_{n\tau}$ conditional upon $X_{n\tau} = x$ are

$$\frac{1}{\tau}E[X_{(n+1)\tau} - X_{n\tau} | X_{n\tau} = x] = \frac{bck}{e^{kn\tau} - c}x,$$

$$\frac{1}{\tau}E[(X_{(n+1)\tau} - X_{n\tau})^2 | X_{n\tau} = x] = \sigma^2 x^2,$$

$$\frac{1}{\tau}E[(X_{(n+1)\tau} - X_{n\tau})^{2+p} | X_{n\tau} = x] = \frac{x^2}{\tau}o(\tau), \quad \forall p \in \mathbb{N}, \tag{A.2}$$

in such a way that when τ approaches zero and $n \to \infty$, when the condition $n\tau = t$, then $X_{n\tau}$ converges to the diffusion process with infinitesimal moments provided by (6).

Appendix B

The transition p.d.f. (12) can be rewritten as

$$f(x_{ij}, t_{ij} | x_{i,j-1}, t_{i,j-1}) = \frac{1}{x_{ij}\sqrt{2\pi\sigma^2(t_{ij} - t_{i,j-1})}} \times \exp\left(-\frac{1}{2}\frac{\left[\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right) - b\ln\left(\frac{D - A^{t_{ij}}}{D - A^{t_{i,j-1}}}\right) + \frac{\sigma^2}{2}(t_{ij} - t_{i,j-1})\right]^2}{(t_{ij} - t_{i,j-1})\sigma^2}\right),$$

where $D = 1/c$ and $A = e^{-k}$. Denoting $n = \sum_{i=1}^{d} n_i$, the log-likelihood function of the sample is

$$\ln L_{x_{ij}}(\mu_1, \sigma_1^2, A, D, b, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{d}{2}\ln\sigma_1^2 - \frac{n-d}{2}\ln\sigma^2 - \sum_{i=1}^{d}\ln x_{i1}$$

$$-\frac{1}{2\sigma_1^2}\sum_{i=1}^{d}[\ln x_{i1}-\mu_1]^2-\sum_{i=1}^{d}\sum_{j=2}^{n_i}\ln x_{ij}-\frac{1}{2}\sum_{i=1}^{d}\sum_{j=2}^{n_i}\ln(t_{ij}-t_{i,j-1})$$
$$-\frac{1}{2}\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{[\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right)-b\ln\left(\frac{D-A^{t_{ij}}}{D-A^{t_{i,j-1}}}\right)+\frac{\sigma^2}{2}(t_{ij}-t_{i,j-1})]^2}{(t_{ij}-t_{i,j-1})\sigma^2},$$

from which the ML estimates of μ_1 and σ_1^2 are

$$\hat{\mu}_1=\frac{1}{d}\sum_{i=1}^{d}\ln x_{i1}\quad\text{and}\quad\hat{\sigma}_1^2=\frac{1}{d}\sum_{i=1}^{d}(\ln x_{i1}-\hat{\mu}_1)^2,$$

whereas the corresponding to A, b, D and σ^2 follow from the solution of the system of equations:

$$\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right)-b\ln\left(\frac{D-A^{t_{ij}}}{D-A^{t_{i,j-1}}}\right)+\frac{\sigma^2}{2}(t_{ij}-t_{i,j-1})}{t_{ij}-t_{i,j-1}}\times\frac{t_{i,j-1}A^{t_{i,j-1}-1}(D-A^{t_{ij}})-t_{i,j}A^{t_{ij}-1}(D-A^{t_{i,j-1}})}{(D-A^{t_{ij}})(D-A^{t_{i,j-1}})}=0, \tag{B.1}$$

$$\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right)-b\ln\left(\frac{D-A^{t_{ij}}}{D-A^{t_{i,j-1}}}\right)+\frac{\sigma^2}{2}(t_{ij}-t_{i,j-1})}{t_{ij}-t_{i,j-1}}\times\ln\left(\frac{D-A^{t_{ij}}}{D-A^{t_{i,j-1}}}\right)=0, \tag{B.2}$$

$$\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right)-b\ln\left(\frac{D-A^{t_{ij}}}{D-A^{t_{i,j-1}}}\right)+\frac{\sigma^2}{2}(t_{ij}-t_{i,j-1})}{t_{ij}-t_{i,j-1}}\times\frac{A^{t_{ij}}-A^{t_{i,j-1}}}{(D-A^{t_{ij}})(D-A^{t_{i,j-1}})}=0, \tag{B.3}$$

$$\sigma^4\sum_{i=1}^{d}\sum_{j=2}^{n_i}(t_{ij}-t_{i,j-1})+4\sigma^2(n-d)-4\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\ln^2\left(\frac{x_{ij}}{x_{i,j-1}}\right)}{t_{ij}-t_{i,j-1}}-4b^2\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\ln^2\left(\frac{D-A^{t_{ij}}}{D-A^{t_{i,j-1}}}\right)}{t_{ij}-t_{i,j-1}}+8b\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right)\ln\left(\frac{D-A^{t_{ij}}}{D-A^{t_{i,j-1}}}\right)}{t_{ij}-t_{i,j-1}}=0. \tag{B.4}$$

This system of equations does not have an explicit solution, so numerical procedures are required to find it. In the particular case $t_{ij}-t_{i,j-1}=h$, $i=1,\ldots,d$; $j=2,\ldots,n_i$, Eqs. (B.1) to (B.4) remain

$$D(1-A^h)(2X_{2,*}^{A,D}-2bX_{3,*}^{A,D}+\sigma^2\,h\,X_{1,*}^{A,D})+hA^{h-1}(2W_2^{A,D}-2bW_3^{A,D}+\sigma^2\,h\,W_1^{A,D})=0, \tag{B.5}$$

$$\sigma^2hY_1^{A,D}-2bY_3^{A,D}+2Y_2^{A,D}=0, \tag{B.6}$$

$$\sigma^2hX_1^{A,D}-2bX_3^{A,D}+2X_2^{A,D}=0, \tag{B.7}$$

$$\sigma^4h^2(n-d)+4\sigma^2h(n-d)-4b^2Y_3^{A,D}+8bY_2^{A,D}-4Z=0, \tag{B.8}$$

where, and denoting $S_{ij}^{A,D}=(D-A^{t_{ij}})(D-A^{t_{i,j-1}})$ and $T_{ij}^{A,D}=\ln((D-A^{t_{ij}})/(D-A^{t_{i,j-1}}))$,

$$X_1^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{A^{t_{i,j-1}}}{S_{ij}^{A,D}},\quad X_2^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{A^{t_{i,j-1}}}{S_{ij}^{A,D}}\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right),$$

$$X_3^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{A^{t_{i,j-1}}}{S_{ij}^{A,D}}T_{ij}^{A,D},\quad Y_1^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}T_{ij}^{A,D},$$

$$Y_2^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}T_{ij}^{A,D}\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right),\quad Y_3^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}(T_{ij}^{A,D})^2,$$

$$X_{1,*}^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{t_{i,j-1}A^{t_{i,j-1}}}{S_{ij}^{A,D}},\quad X_{2,*}^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{t_{i,j-1}A^{t_{i,j-1}}}{S_{ij}^{A,D}}\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right),$$

$$X_{3,*}^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{t_{i,j-1}A^{t_{i,j-1}}}{S_{ij}^{A,D}}T_{ij}^{A,D},\quad W_1^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{A^{2t_{i,j-1}}}{S_{ij}^{A,D}},$$

$$W_2^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{A^{2t_{i,j-1}}}{S_{ij}^{A,D}}\ln\left(\frac{x_{ij}}{x_{i,j-1}}\right),\quad W_3^{A,D}=\sum_{i=1}^{d}\sum_{j=2}^{n_i}\frac{A^{2t_{i,j-1}}}{S_{ij}^{A,D}}T_{ij}^{A,D}$$

and $Z=\sum_{i=1}^{d}\sum_{j=2}^{n_i}\ln^2(x_{ij}/x_{i,j-1})$.

After some algebra, from Eqs. (B.6) and (B.7), we obtain

$$b^{A,D}=\frac{X_2^{A,D}Y_1^{A,D}-X_1^{A,D}Y_2^{A,D}}{X_3^{A,D}Y_1^{A,D}-X_1^{A,D}Y_3^{A,D}}\quad\text{and}\quad\sigma_{A,D}^2=\frac{2C^{A,D}}{h}, \tag{B.9}$$

where

$$C^{A,D}=\frac{X_2^{A,D}Y_3^{A,D}-X_3^{A,D}Y_2^{A,D}}{X_3^{A,D}Y_1^{A,D}-X_1^{A,D}Y_3^{A,D}}. \tag{B.10}$$

Replacing these expressions in (B.5) and (B.8) the following system of equations appears:

$$D(1-A^h)(X_{2,*}^{A,D}-b^{A,D}X_{3,*}^{A,D}+C^{A,D}X_{1,*}^{A,D})+hA^{h-1}(W_2^{A,D}-b^{A,D}W_3^{A,D}+C^{A,D}W_1^{A,D})=0,\ (n-d)(C^{A,D})^2+2(n-d)C^{A,D}-(b^{A,D})^2Y_3^{A,D}+2b^{A,D}Y_2^{A,D}-Z=0. \tag{B.11}$$

As we pointed out in the introduction, in some cases the parameter b is known. For example, in the study of the length of fish b is equal to one, whereas in the case of weight in *isometric* growth $b=3$. In such a case, that is, when b is known, the Eq. (B.6) disappears, (B.10) turns into

$$C^{A,D}=\frac{bX_3^{A,D}-X_2^{A,D}}{X_1^{A,D}},$$

whereas in the system of Eqs. (B.11) the expression $b^{A,D}$ must be changed to b.

References

Albano, G., Giorno, V., 2006. A stochastic model in tumor growth. Journal of Theoretical Biology 242, 329–336.

Capocelli, R.M., Ricciardi, L.M., 1974. Growth with regulation in random environment. Kybernetik 15, 147–157.

Cheng, Y.W., Kuk, A., 2002. Determination of the unknown age at first capture of western rock lobsters (panulirus cygnus) by random effects model. Biometrics 58, 459–462.

Chong, J., Aguayo, M., 2009. Edad y crecimiento del pez espada (*Xiphias gladius* Linnaeus 1758) en el Pacífico suroriental (Diciembre 1994-Septiembre 1996). Latin American Journal of Aquatic Research 37 (1), 1–15.

García-Rodríguez, M., Pereda, P., Landa, J., Esteban, A., 2005. On the biology growth of the anglerfish lophius Budegasa Spinola, 1807 in the Spanish Mediterranean: a preliminary approach. Fisheries Research 71 (2), 197–208.

Gudmundsson, G., 2005. Stochastic growth. Canadian Journal of Fisheries and Aquatic Sciences 62 (8), 1746–1755.

Gutiérrez, R., Ricciardi, L.M., Román, P., Torres, F., 1997. First-passage-time densities for time-non-homogeneous diffusion processes. Journal of Applied Probability 34 (3), 623–631.

Gutiérrez, R., Rico, N., Román, P., Torres, F., 2006. Approximate and generalized confidence bands for some parametric functions of the lognormal diffusion process with exogenous factors. Scientiae Mathematicae Japonicae 64, 843–859.

Gutiérrez, R., Román, P., Romero, D., Serrano, J.J., Torres, F., 2007. A new gompertz-type diffusion process with application to random growth. Mathematical Biosciences 208, 147–165.

Hart, D.R., Chute, A.S., 2009. Estimating von bertalanffy growth parameters from growth increment data using a linear mixed-effects model with an application

to the sea scallop *Placopecten magellanicus*. ICES Journal of Marine Science Advances: Journal du Conseil 66 (10), 2165-2175.
Kimura, D.K., 1980. Likelihood methods for the Von Bertalanffy growth curve. Fishery Bulletin 77 (4), 765–776.
Kloeden, P.E., Platen, E., Schurz, H., 1994. Numerical Solution of SDE Through Computer Experiments. Springer, Berlin.
Laslett, G.M., Eveson, J.P., Polacheck, T., 2003. Biological and statistical issues in fitting growth curves to capture-recapture data. Biometrics 59, 1186–1189.
Lv, Q., Pitchford, J.W., 2007. Stochastic Von Bertalanffy models, with applications to fish recruitment. Journal of Theoretical Biology 244, 640–655.
Rafail, S.Z., 1973. A simple and precise method for fitting a Von Bertalannfy growth curve. Marine Biology 19, 354–358.
Román, P., Serrano, J.J., Torres, F., 2008. First-passage-time location function: application to determine first-passage-time densities in diffusion processes. Computational Statistics and Data Analysis 52, 4132–4146.
Russo, T., Baldi, P., Parisi, A., Magnifico, G., Mariani, S., Cataudella, S., 2009. Lévy processes and stochastic von Bertalanffy models of growth, with application to fish population analysis. Journal of Theoretical Biology 258, 521–529.
Tovar-Ávila, J., Troynikov, V.S., Walter, T.I., Day, R.W., 2009. Use of stochastic models to estimate the growth of the Port–Jackson shark, *Heterodontus portusjacksoni*, off eastern Vistoria, Australia. Fisheries Research 95, 230–235.
Von Bertalanffy, L., 1938. A quantitative theory of organic growth. Human Biology 0 (2), 181–213.
Wang, Y., 1999. Estimating equations for parameters in stochastic growth model from tag-recapture data. Biometrics 55, 900–903.
Weatherly, A., Gill, H., 1987. The Biology of Fish Growth. Academic Press Inc. Harcourt Brace Jovanovich Publishers, New York.