Modeling the biological growth with a random logistic differential equation

Arelly Ornelas
Conacyt - Instituto Politécnico Nacional
Francisco Delgado-Vences (delgado@im.unam.mx)
Conacyt - Universidad Nacional Autónoma de México
Enrique Morales-Bojórquez
Centro de Investigaciones Biológicas del Noroeste SC, B.C.S
Victor Hugo Cruz-Escalona
Instituto Politécnico Nacional
Emigdio Marín-Enríquez
Conacyt - Universidad Autónoma de Sinaloa" Mazatlan
claudia J. Hernández-Camacho
Instituto Politécnico Nacional

Research Article

Keywords: Biological growth, bayesian inference, random logistic differential equation, Data assimilation, simulations

Posted Date: October 12th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1801042/v1

Modeling the biological growth with a random logistic differential equation

Arelly Ornelas[1], Francisco Delgado-Vences[2*], Enrique Morales-Bojórquez[3], Victor Hugo Cruz-Escalona[4], Emigdio Marín-Enríquez[5] and Claudia J. Hernández-Camacho[4]

[1]Centro Interdisciplinario de Ciencias Marinas, Conacyt - Instituto Politécnico Nacional , La Paz, B.C.S., México.
[2*]Instituto de Matemáticas, Conacyt - Universidad Nacional Autónoma de México, Oaxaca de Juarez, Oaxaca, México.
[3]Centro de Investigaciones Biológicas del Noroeste SC., La Paz, B.C.S., México.
[4]Centro Interdisciplinario de Ciencias Marinas, Instituto Politécnico Nacional, , La Paz, B.C.S., México.
[5] Facultad de Ciencias del Mar, Conacyt - Universidad Autónoma de Sinaloa,, Mazatlan, Sinaloa, México.

*Corresponding author(s). E-mail(s): delgado@im.unam.mx;
Contributing authors: aornelasv@ipn.mx; emorales@cibnor.mx; vescalon@ipn.mx; emarin@uas.edu.mx; cjhernandez78@gmail.com;

Abstract

We modeled biological growth using a random differential equation (RDE), where the initial condition is a random variable and the growth rate is a suitable stochastic process. These assumptions let us to obtain a model that represents well the random growth process observed in nature, where only few individuals of the population reach the maximal size of the species and the growth curve for every individual behaves randomly. Since, we assumed that the initial condition is a random variable, we assigned a priori density and we performed Bayesian inference to update the initial condition's density of the RDE. Karhunen-Loeve expansion was then used to approximate the random coefficient of the RDE. Then, using the RDE's approximations, we estimated the density $\boldsymbol{f(p,t)}$. Finally, we fitted this model to the biological growth of

the giant electric ray (or Cortez electric ray) *Narcine entemedor*. Simulations of the solution of the random logistic equation were performed to construct a curve that describes the solutions' mean for each time. As a result, we estimated confidence intervals for the mean growth that described reasonably well the observed data. We fit the proposed model with a training dataset and the model is tested with a different dataset. The model selection is performed with the square of the errors. .

Keywords: Biological growth, bayesian inference, random logistic differential equation, Data assimilation, simulations.

MSC Classification: 34F05 , 62M20 , 62P10 , 62F15 , 92B05

1 Introduction

In this article, we present a stochastic model that takes into account the individual variability of the organisms along the time. The proposed model allows to the individuals to growth at random levels of intensity that changes along time. In addition, in the stochastic model not all the individuals reach the maximal expected growth size, which agrees with what is observed in nature.

The model and methodology studied in this article can be applied to several species of the animal kingdom. Here to illustrate the use of the model, we will focus on marine animal literature and data. Traditionally, the individual growth models that are applied to marine organisms have been based on length-at-age data, describing the average growth of the individuals in a population. However, this approach ignores the intrinsic individual variability of the organisms, which can be influenced by genetic differences in growth rate potential [1], food availability [2], or differences in physiological stress [3] and length dependence of growth rate [4]. This phenomena can be observed when there is differential access to food and dominant organisms interfere with feeding of subordinates, and consequently dominant individuals show better growth rates [5]. Recognizing that the length-at-age data are highly variable, two accepted hypotheses have been proposed: a) growth compensation, which is observed when the length-at-age variability decreases with time or age [2], [3]; and b) growth depensation, which is observed if length-at-age variability increases with age [6], [7]. The parameterization of the individual growth models have used ordinary sum of squares or likelihood functions, assuming different probabilistic density functions, such as: normal, log normal, gamma, mixture distribution or log skew-t distribution [8], [9], [10].

The variability in length-at-age data has also been modeled from stochastic approaches based on transition matrices, where the von Bertalanffy growth model is assumed in the growth matrix, according to: $\bar{\Delta}_t = (L_\infty - l_k)(1 - \exp^{-k})$, where $\bar{\Delta}_t$ represents the mean growth increment,

l_k is the midlength of the length-class l, L_∞ is the asymptotic length, and k is the growth parameter.

A gamma distribution with parameters α_l and β could also describe the variation of the mean growth increment [11, 12], [13]:

$$g(x \mid \alpha_l \beta) = \frac{1}{\beta_l{}^{\alpha_l} \Gamma(\alpha_l)} x^{\alpha_l - 1} \exp \frac{-x}{\beta}, \tag{1}$$

where x represents $\Delta_l = l_{t+1} - l_t$, denoting the growth increment in the length-class l. The mean change in length is $\mathbb{E}(\Delta_l) = \alpha_l \beta$ and the variance is $\sigma_l^2 = \alpha_l \beta^2$. The probability of an individual growing $(P_{l,l'})$ from length-class l to length-class l' can be estimated by integrating over the length-classes, as follows:

$$P_{l|l'} = \int_{l'}^{l'} g(x \mid \alpha_l \beta) dx. \tag{2}$$

The survival matrix is represented as: $S_{(l,t)} = \exp^{-Z_{l,t}}$, where the number of individuals in the length-class l surviving to the start the time t' is reduced by mortality, considering that $Z_{l,t} = F_{l,t} + M_{l,t}$, where M is natural mortality and F is the fishing mortality. Finally, $P_{l,l'}$ matrix can be estimated as: $N_{l,t} = P_{l,l'} \exp^{-Z_{l,t}}$.

The inclusion of this stochastic approach partially improves estimates of individual growth. However, the final estimate requires an estimation of the survival matrix, which is usually not easy to estimate from length-at-age data. Consequently, its use is limited in individual growth modeling. Given the limitations for modeling the individual differences in growth from length-at-age data, a new proposal must be developed from the stochastic approach to improve the estimation of parameters and to increase the realism of the theoretical model.

2 Our contribution

We modeled the biological growth with a random differential equation (RDE). RDEs have been applied in a wide range of disciplines such as: biology, medicine, population dynamics and engineering (see, [14]).

A RDE is essentially an ordinary differential equation where one or more of their coefficients are stochastic processes. Given that differential equations are very dynamic and also given the randomness of their coefficients, RDE's are able to describe the individual differences in growth. We also use Bayesian inference to update the probability density function of the random initial condition of the RDE with new data. With the posterior density, we were able to obtain confidence intervals for the RDE's solution.

More precisely, we have to numerically compute the density probability function $f(p,t)$ of the solution of the random logistic equation with a random initial condition. We used the Wiener process and the Ornstein-Uhlenbeck (OU) process as random coefficients. Therefore, we need to compute or simulate the random coefficients of the RDE. The Wiener process is easier to simulate than the OU process because it is the solution of stochastic differential equations (SDEs) (see, for instance, [15]).

Numerical methods for SDEs (e.g., the Ornstein-Uhlenbeck process, Geometric Brownian motion, etc.) have been developed for the last few decades, such as the Euler-Maruyama method, stochastic Itô-Taylor expansions, Runge-Kutta approximations, Monte Carlo methods and its multiple variation; see for instance the classical book [16].

Recently, spectral methods have been applied to the probability space to solve numerically SDEs. The Wiener-Chaos expansion of the probability space is a spectral method that allows us to write any square-integrable random variable as a Fourier-Hermite series (see, for instance, Chapter 5 in [17] and the references therein). Another option is the Karhunen-Loeve expansion. In this approach, any square-integrable random variable is decomposed as a Fourier expansion using as a basis the eigenfunctions of a suitable operator (usually the covariance operator of a nice stochastic process) via Mercer's Theorem (see, for example, Chapter 4 in [18] or Section 37.5 in [19]). The former approach converges faster than the latter, but is more complicated to compute. In this paper, we use the Karhunen-Loeve expansion to numerically estimate the Wiener process and the Ornstein-Uhlenbeck process. After that, we approximate the solution of the RDE (5) following [20].

This paper is organized as follows. In section 3, we introduce the random logistic differential equation and its Karhunen-Loeve expansion. In section 4, we briefly review Bayesian theory, particularly the Beta-Binomial model. A model description for the biological growth is discussed in section 5, we also describe the data used in this paper in this section. In sections 6 and 7, we apply the RDE model with Wiener and Ornstein-Uhlenbeck processes as its random coefficient and with a Beta random variable as prior distribution of the initial condition. Our conclusions and a discussion of the model and its numerical applications are given in section 9. Finally, in the two appendices we provide a proof of the solution of the RDE and a description of the heuristic method that we use to fix the parameters of the Ornstein-Uhlenbeck process.

The procedure described above can accomplish the characteristics of data assimilation definition. In our case, we applied to RDEs and we updated the state of the system with new observations just one time. Data assimilation is the set of techniques that allow us to optimally combine observations of a physical system with numerical models and prior information of that system.

Finally, an updated estimate of the state of the system is obtained *see, e.g.* [21].

3 A random logistic differential equation

The classical logistic differential equation has been applied to model bounded growth, for instance, tumor growth in medicine, reaction models in chemistry, population dynamics and Fermi distribution in physics, and so on. These problems are formulated via the Initial Value Problems (IVP) [20, *for instance.*]

$$\left.\begin{array}{ll} p'(t) = rp(t)\big[1 - p(t)\big], & t_0 < t \\ p(t_0) = p_0, & p_0 \in (0,1)' \end{array}\right\} \tag{3}$$

with solution:

$$p(t) = \frac{p_0}{p_0 + (1 - p_0)e^{-r(t-t_0)}}. \tag{4}$$

Where r is the intrinsic growth rate and p_0 and $p(t)$, in our case, denote the size proportion at time instant t_0 and t.

We are interested in a model where both parameters r and p_0 are stochastic and therefore $p(t)$ is also stochastic. Therefore, this model will take into account the intrinsic variability of the growth rate r and the value of p_0. We introduce a modification of the logistic model, as follows:

$$\left.\begin{array}{ll} P'(t,\omega) = A(t,\omega)P(t,\omega)\Big[1 - P(t,\omega)\Big], & t_0 < t < T, \quad \omega \in \Omega \\ P(t_0,\omega) = p_0(\omega). & \end{array}\right\} \tag{5}$$

$A(t,\omega)$ is a stochastic process and $p_0(\omega)$ is a bounded absolutely continuous random variable $p_0(\omega) : \Omega \to [a_1, a_2] \subset (0,1)$. We assume that both, $A(\cdot,\omega)$ and $p_0(\omega)$, are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $T > 0$ is fixed. We called equation(5) the random logistic equation, which is a Random differential equation (RDE). We observe that the parameters in the proposed model are still interpretable; meaning that, in our case, they correspond to the random size at birth and to the stochastic growth rate at each time t.

The solution to equation (5) is given by (a proof is presented in appendix A)

$$P(t,\omega) = \frac{p_0(\omega)}{p_0(\omega) + \big(1 - p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t} A(s,\omega)ds\Big)} \tag{6}$$

Note that the $\lim_{t\to\infty} P(t,\omega) = 1$ is satisfied only when the stochastic process $A(s,\omega)$ is strictly positive. If $A(s,\omega)$ does not ensure positivity, then

with some strictly positive probability the $\lim_{t\to\infty} P(t,\omega) = 0$, so that the IVP stochastic solution (5) could not behave as its deterministic counterpart (3). However, in nature some species could decrease their sizes under some conditions; then, even in these situations the stochastic model considered in this manuscript could be an appropriate model. To illustrate the suitability of the model we choose as examples two stochastic processes with reflection at zero. However, other possible choices could be done acordingly to the phenomena/data to study.

The fact that $P(t)$ reaches an asymptotic value of 1 does not imply that all individuals are expected to growth to the species' maximum observed size. Instead, this means that with strictly positive probability some individuals could reach the asymptotic size and some will not.

To approximate the coefficient $A(s,\omega)$, we will use the Karhunen-Loève expansion. This approximation is used for those cases where it is highly difficult to obtain a closed form of the stochastic process or when performing simulations in an efficient way is desired or when statistical properties of the referred stochastic process will be studied.

Therefore, we present a theorem on the L^2 convergence of the Karhunen-Loève expansion of a mean square continuous process. The proof of this theorem can be consulted in [22], Theorem 5.28.

Theorem 1 *Consider a mean square integrable continuous time stochastic process* $X \equiv \{X(t,\omega) : t \in \mathcal{T}, \omega \in \Omega\}$, *i.e.,* $X \in L^2\big(\Omega, L^2(\mathcal{T})\big)$ *with* $\mu_X(t)$ *and* $c_X(s,t)$ *as its mean and covariance functions, respectively. Then,*

$$X(t,\omega) = \mu_X(t) + \sum_{j=1}^{\infty} \sqrt{\nu_j}\phi_j(t)\xi_j(\omega), \qquad \omega \in \Omega, \tag{7}$$

where, this sum converges in $L^2\big(\Omega, L^2(\mathcal{T})\big)$ *and*

$$\xi_j(\omega) := \sqrt{\nu_j}\big\langle X(t,\omega) - \mu_X(t), \phi_j(t)\big\rangle_{L^2(\mathcal{T})}, \tag{8}$$

associated to the covariance function $c_X(s,t)$. *The random variables* $\xi_j(\omega)$ *have zero mean* $\mathbb{E}[\xi_j(\omega)] = 0$, *unit variance* $Var[\xi_j(\omega)] = 1$ *and are pairwise uncorrelated* $\mathbb{E}[\xi_j(\omega)\xi_k(\omega)] = \delta_{jk}$. *Furthermore, if* $X(t,\omega)$ *is Gaussian, then* $\xi_j(\omega) \sim N(0,1)$ *are independent and identically distributed.*

This theorem guarantees that each process with finite second moment can be written as a series of the product between random variables and a sequence of functions that are some suitable basis. By using this theorem, we will approximate $A(s,\omega)$. In this paper, we will use $\mathcal{T} = [t_0, T]$ for some $T > t_0$.

Note that it is necessary to truncate the series because it is impossible to take the sum to infinite in equation (7). Then, we work with a finite approximation to the process $X(t,\omega)$.

In the following paragraphs, we apply a truncation of the Karhunen-Loève expansion to $A(s,\omega)$ and then we will get an approximation for the density $f(p,t)$ of the RDE's solution (5) at time t. Notice that $f(p,t)$ is not a transition density but is a true fixed-time density for the process $P(t,\omega)$. More precisely, given a fixed t the function $(0,1) \ni p \to f(p,t)$ is the density probability function of the random variable $P(t,\omega)$ at every fixed time t.

Set $\overline{\boldsymbol{\xi}_N}(\omega) := (\xi_1, \ldots, \xi_N)$ and denote by

$$Q_N\big(t, \overline{\boldsymbol{\xi}_N}(\omega)\big) = \int_{t_0}^{t} \Big(\mu_A(s) + \sum_{j=1}^{N} \sqrt{\nu_j}\phi_j(s)\xi_j(\omega)\Big) ds. \tag{9}$$

Some hypotheses on p_0, $A(t,\omega)$, $\{\xi_j\}_{j\geq 1}$ and $f(p,t)$ are required:

(H1) $p_0 : \Omega \to [b_1, b_2] \subset (0,1)$ and $A(t,\omega) \in L^2\big(\Omega, L^2(\mathcal{T})\big)$.

(H2) $p_0(\omega)$, $\xi_j(\omega)$, $1 \leq j \leq N$ are absolutely continuous random variables. Moreover, $p_0(\omega)$, $\overline{\boldsymbol{\xi}_N}(\omega)$ are independent with densities $f_{p_0}(p_0)$ and $f_{\overline{\boldsymbol{\xi}_N}(\omega)}(\xi_1, \ldots, \xi_N)$, respectively.

(H3) $f_{p_0}(p_0)$ is Lipschitz continuous in its domain; that is,

$$\exists L = L_{f_{p_0}} \text{ such that } \quad |\, f_{p_0}(p) - f_{p_0}(q)\,| \leq L\,|\,p-q\,| \qquad \forall p, q \in [b_1, b_2].$$

(H4) $A(t,\omega)$ admits a Karhunen-Loève expansion of type (7) such that there exists a constant $C > 0$ that satisfies

$$\mathbb{E}\big[e^{2Q_N(t, \boldsymbol{\xi}_N(\omega))}\big] \leq C, \quad \text{for all positive integers } N.$$

Note that because $A(t,\omega) \in L^2\big(\Omega, L^2(\mathcal{T})\big)$, we can apply Theorem 1. Then, the stochastic process (6) solution to (5) can be written as

$$P(t,\omega) = \frac{p_0(\omega)}{p_0(\omega) + \big(1 - p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t}\Big[\mu_A(s) + \sum_{j=1}^{\infty}\sqrt{\nu_j}\phi_j(t)\xi_j(\omega)\Big]ds\Big)}. \tag{10}$$

We define the N-truncated solution of $P(t,\omega)$ as

$$P^N(t,\omega) = \frac{p_0(\omega)}{p_0(\omega) + \big(1 - p_0(\omega)\big)\exp\Big(-Q_N\big(t, \overline{\boldsymbol{\xi}_N}(\omega)\big)\Big)}. \tag{11}$$

Following the procedure used in [20, *see, eq. (11)*], we obtain the probability density function of the truncated solution $P^N(t,\omega)$

$$f^N(p,t) = \int_{\mathbb{R}^N} f_{p_0}\left(\frac{pe^{-Q_N(t,\overline{\boldsymbol{\xi}_N})}}{1+p(-1+\exp\left(-Q_N(t,\overline{\boldsymbol{\xi}_N})\right)}\right) f_{\overline{\boldsymbol{\xi}_N}}(\xi_1,\ldots,\xi_N) \times \frac{pe^{-Q_N(t,\overline{\boldsymbol{\xi}_N})}}{\left[1+p(-1+\exp\left(-Q_N(t,\overline{\boldsymbol{\xi}_N})\right)\right]^2} d\xi_N \cdots d\xi_1. \tag{12}$$

Note that in the previous expression, even when it is not explicitly pointed out, the density f_{p_0} depends on time.

The following theorem ensures the convergence of (12) to the exact density $f(p,t)$ [20, Th. 3].

Theorem 2 *Under Hypotheses* **(H1)**-**(H4)**, *the sequence of densities* $\{f^N(p,t)\}_{N\geq 1}$ *defined by* (12) *converges for every* $J\times[t_0,T]$, *for all* $J\subset\mathbb{R}$, *to the exact density* $f(p,t)$ *of the solution of the random Initial Value Problem* (5).

4 Bayesian Inference

In this section, for the sake of completeness, we illustrate the use of Bayesian inference to update the density $f_{p_0}(p_0)$. The main idea of Bayesian inference is to update the probability of occurrence of some event as more evidence or data becomes available. We refer to [23] for further reading.

Suppose that the unknown parameter θ of a given model is regarded as a random variable. Prior information is expressed using probability density $f(\theta)$, which is called *prior probability*. The Bayesian approach connects data and parameter through the *likelihood function* $L(x \mid \theta)$ and is based on the evaluation of the *posterior probability*, which by Bayes theorem is

$$f(\theta \mid x) \propto f(\theta)L(x \mid \theta).$$

Note that parameter values where the likelihood is high are those that have a high probability of producing the observed data. This agrees with the maximum likelihood approach, where the best estimate of θ is the one that maximizes the likelihood function.

We want to update f_{p_0} by applying Bayesian inference. Suppose that, at times $\tau_1,\tau_2,\ldots,\tau_m$ we obtain new information to calculate the likelihood function and consequently update the posterior probability. We denote by $L_i(x \mid \theta)$ and $f_{P_i}(\theta \mid x)$ as the likelihood function and the *posterior probability* at time

τ_i $(i = 1, 2, \ldots, m)$, respectively. Then,

$$f_{P_i}(\theta \mid x) \propto f_{P_{i-1}}(\theta \mid x)\, L(x \mid \theta), \qquad i = 1, \ldots, m,$$

where $f_{P_0}(\theta \mid x) = f_{p_0}(\theta)$. Note that the last expression is applied in the case when the SDE (5) has a closed solution. In those cases where there is an approximate solution,

$$f^N_{P_i}(\theta \mid x) \propto f^N_{P_{i-1}}(\theta \mid x)\, L(x \mid \theta), \qquad i = 1, \ldots, m,$$

where $f^N_{P_0}(\theta \mid x) = f_{p_0}(\theta)$ and N are the number of terms that we take in the approximation of $A(t, \omega)$ (compare equations (10) and (11) taking into account (9)).

We use the well-known Beta-Binomial model to our data. For the sake of completeness we present a brief review of this model in appendix.

5 Biological growth of the Narcine entemedor

We are interested in modeling the biological growth of a giant electric ray (or Cortez electric ray) *Narcine entemedor* using the RDE (5).

Narcine entemedor specimens were collected between October 2013 through December 2015 in the south of Bahia de La Paz, which is located in the southern portion of the Gulf of California (24" 25' N, 110" 18' W). The organisms were captured by artisanal fishers using monofilament gill nets (200–300 m long, 1.5 m high, 20–25 cm stretch mesh), traditionally called "chinchorros," which are set in the afternoon at depths between 10 and 30 m over sandy bottoms and recovered the next morning. Individuals were measured for total length (TL, cm) and the sex was determined by the presence of copulatory organs in males. Vertebrae were collected from the abdominal region of each specimen. The radius of each vertebrae was measured on the corpus calcareum along a straight line through the focus of each vertebra with SigmaScan Pro 5.0.0 Software (SPSS Inc). Vertebral radius (VR) was plotted against TL and tested for a linear relationship to determine if these vertebrae were a suitable structure for age determination and for back-calculated estimation of length at previous ages. The number of the sample is $N = 244$.

We have split the whole sample into two disjoint datasets. The biggest dataset contains around 70% of the observations and it was used for learning—that is, to fit the parameters of the model—which is usually called the training dataset. The fitted model was used to run simulations. We then construct confidence intervals, which we use to compare the observations of a second dataset—called the validation dataset—, which is the complement of the training data. The sample size for training data was $N_1 = 170$ and for the test data it was $N_2 = 74$.

5.1 Bayesian inference for the initial random variable

In this subsection, we applied the Beta-Binomial model to the random variable p_0. We noted two limitations about this approach. Fistly, there was not available data that help us to set up a density function for the variable p_0. Secondly, the random variable p_0 has as a support an interval strictly inside the $[a, b] \subsetneq [0, 1]$. Thus, we used the Beta-Binomial model, because it allowed us to approximate a beta distribution for p_0 in a simple but efficient approach. Another possible choice could be to take into account the deviations from the predictions, that means try to model with a continuous likelihood instead of a discrete as in the Beta-Binomial model. However, we constrained the density to be in an interval $[a, b] \subsetneq [0, 1]$ which implies we have to truncate the continuous likelihood, and consequently the posterior density as well. Moreover, the fact that a Binomial converges to a Gaussian for a large sample, indicate that the election of the Beta-Binomial model is suitable to the data.

The strategy to construct a density for p_0 using the the Beta-Binomial model is as follows. First, we select a noninformative or weakly informative prior density, meaning a beta density that reflects our ignorance on the initial value p_0. Afterward, we use the available data to construct a Binomial density that incorporates information about the initial density. This last part will we done fitting the logistic model (4).

We focused on two random variables

- *Estimated-aged*, which is an age that is assigned to each individual according to its largest vertebra.
- *Proportion size* of each individual, we define a new variable by dividing "individual size" by the maximum size of the specie observed in the sample.

Observe that in opposition of the definition of the already defined two random variables, when we use explicit data we get the sampled aged and the sampled proportion size, which is no longer a random variable but an actual number.

Using these variables, we fitted a logistic curve according to equation (4), which is the solution of equation (3), with the variables *Estimated-aged* and *Proportion size*. p_0 and r were fitted using a optimization procedure. To estimate those values, the sum of the square error between the observed data and the curve (4) was calculated,

$$S = \sum_{j=1}^{N_1} \left[\hat{p}(t_j;\ r, p_0) - p_j(t) \right]^2, \tag{13}$$

where $\hat{p}(t_j;\ r, p_0)$ is the value of the curve at the time of the $p_j(t)$. $\hat{r}$ and $\hat{p}_0$ were the values that minimize S. By using the `R-` function `optim`, $\hat{r} = 0.2266$ and $\hat{p}_0 = 0.4414$ were obtained; see figure 1. Note that this prior information is reflected on the prior distribution of p_0.

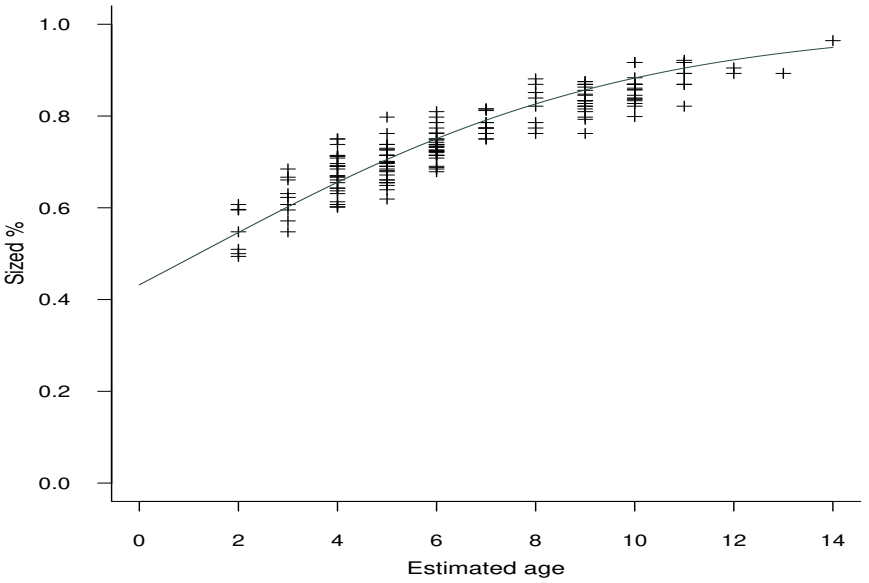
A set of random variables $\{Y_j\}_{j=1}^{N_1}$ is then defined: $Y_j = 1$ if the j observation of the variable "estimated-aged" is below or on the logistic curve, otherwise $Y_j = 0$. The sum over all the random variables Y_j:

$$Z := \sum_{j=1}^{N_1} Y_j,$$

follows a $Z \sim Bin(N_1, p)$, with p unknown. The random variable Z describes how many of the observed data are less than or equal to the expected size according to age. It is important to note that p is the initial condition of the model equation (5) which we are assuming is random.

From our training data, we obtain that $z = 62$ and because $N_1 = 170$, then $n - z = 108$. Moreover, data goes from 0 to 14 years in the estimated aged variable, then the interval $[0, T] = [0, 14]$ in the RDE equation (5).

Figure 1 Logistic curve for the the variables estimated-aged and proportion-size.



We consider parameter p as a random variable that follows a Beta distribution: $Beta_{[0.1,0.9]}(\alpha, \beta)$, which is the prior distribution. We denote by z the value of the random variable Z. We use the Beta-Binomial model to obtain

the posterior distribution accordingly to equation (21), then the posterior distribution is $Beta_{[0.1,0.9]}(\alpha + z, n - z + \beta)$. Prior and posterior distributions need to be standardized to be a true density.

We consider two Beta distributions with parameters (α, β):

- $(\alpha, \beta) = (1, 1)$
- $(\alpha, \beta) = (2, 2)$,

in the former case, each segment with same length, included in the range of the random variable, has equal probability; while in the latter case, the probability is high around the middle point of the segment $[0.1, 0.9]$. These choices could be considered as noninformative distributions and, as we mention before, they reflect our ignorance on the true value of the random initial condition.

We use a mesh for the interval $[0.1, 0.9]$ with $\Delta_x = 1/100$.

The density of the RDE's solution with Beta as prior distribution for p_0, with two different Betas is presented in figure 2. This means that the use of the posterior distribution is not included. As we mention before, they are noninformative and therefore they are among the best choices as prior distribution when we do not know the value of the parameter (or any data) (see for instance [24]).

In the following two sections, we simulate the posterior distribution with two examples of stochastic processes as random coefficients of the RDE (5). The stochastic processes are the Wiener and the Ornstein-Uhlenbeck processes.

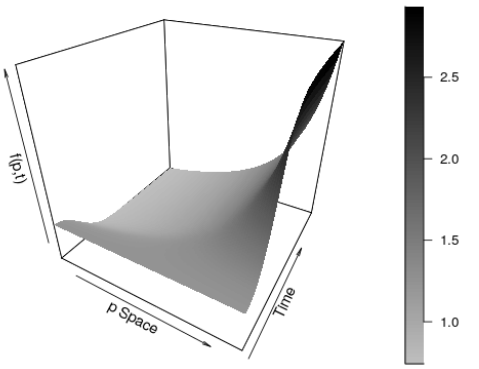
We use the Karhunen-Loeve expansion with $N = 1$, as in [20] to approximate the posterior distribution. We apply the Simpson's rule *see, e.g.*, [25] to perform the numerical integration of the expression (12).

6 Wiener process as random coefficient.

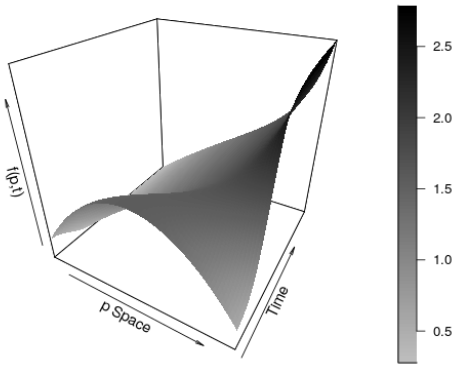
In this section, we consider $A(t, \omega)$ to be the Wiener process (also known as Brownian motion). We recall some basic facts of the Karhunen-Loeve expansion for the Wiener process $W(t, \omega)$, $t \in [0, T]$ [22] . $W(t, \omega)$ has mean function $\mu_W(t) = 0$ and covariance function given by $C_W(s, t) = \min(s, t)$, for all $s, t \in [0, T]$. The Karhunen-Loeve expansion, equation (7), is performed with a set of uncorrelated standard Gaussian random variables $\{\xi_j(\omega)\}_{j \geq 1}$; that is, $\xi_j(\omega) \sim N(0, 1)$ and with the set of functions $\phi_j(t)$ and eigenvalues ν_j given by

$$\nu_j = \frac{4T^2}{(2j-1)^2\pi^2}, \qquad \phi_j(t) = \sqrt{\frac{2}{T}} \sin\Big(\frac{(2j-1)\pi t}{2T}\Big), \qquad j = 1, 2, \ldots$$

Figure 2 Density of the solution of the random logistic equation using the *prior distribution* $Beta(\alpha, \beta)$.



(a) $Beta(1, 1)$



(b) $Beta(2, 2)$

We choose $P_0(\omega)$ independent of the random vector $\overline{\boldsymbol{\xi}_N} = (\xi_1, \ldots, \xi_N)$.

With this expansion, we simulate the solution of the RDE by using the equation (12). The posterior distribution of the solution when the prior density is $Beta(1, 1)$ and $Beta(2, 2)$ is presented in figure 3. Notice that we get almost the same shape of the density for both choices of prior distribution.

This is most probably due to the fact that the sample number is greater than 50. We use a mesh for the interval $[0.1, 0.9]$ with $\Delta_x = 1/100$ and a time step of $\Delta_t = 1/100$.

Note that the variability increases when time increases. This can be explained by the fact that during the first period of life, the availability of food is essential in the growth of each organism. If the food is very abundant, then their size will be greater than expected. This fact generates variability on the data. Meanwhile, when the organisms achieve the maximum age, the variability decreases because they approach the maximum expected size. Therefore, the model reproduces the biological growth dynamic well.

We then run 1000 simulations of the Wiener process. We use the function `snssde1d` from the `R-library Sim.DiffProc` [26, 27] to simulate the Wiener process.

From the posterior density at time $t = 0$, as shown in figure 3, we take 1000 samples. We make that for each prior distribution $Beta(1, 1)$ and $Beta(2, 2)$.

Subsequently, we run 1000 simulations of the solution of the equation (6), each of them combining the posterior density $Beta(\alpha, \beta)$ and the previous Wiener processes simulations. Figure 4 show the simulations of the solutions of the equation (5).

In these figures, every single line is a realization of the process, that means that every line represents an estimated growth along time. Note that not all the individuals reach the maximal expected growth size, this agrees with what is observed in nature. Therefore, the proposed model adequately reproduces this phenomenon.

We also observe that the dynamic of the differential equation governs the behavior of the solution because each of the solutions stays close to the solution mean. In other words, the variations of the Wiener processes are not strong.

Note that when the posterior distribution is updated with new data, we obtain a $Beta(21, 32)$ and a $Beta(22, 33)$, and these distributions are quite similar. Therefore, as in the case of the densities (figure 3), there is no substantial difference between $Beta(1, 1)$ and $Beta(2, 2)$. Thus, if the sample size is large enough ($n \geq 20$), then we could use as a prior distributions either Beta and the result will be very similar.

One important fact about this model is that it is able to provide a *credible interval* at certain level α (analogous to confidence intervals in classical statistics). However, because we are using a very low number of terms in the Karhunen-Loeve expansion, the credible intervals are wide and meaningful. Consequently, they are not included in the results. To improve this part it is

necessary to include more terms in the Karhunen-Loeve expansion. Another idea could be to use the Wiener-chaos expansion to improve the numerical simulations.

Figures 5 show the mean, confidence intervals at 95% for the simulations; black points represent observed data, dotted lines represent confidence intervals (CI) of the simulation. We presented two graphs, one for each prior distribution $Beta(1,1)$ and $Beta(2,2)$. Bayesian inference and the numerical approximation to the solution given by the Karhunen-Loeve expansion are well behaved because all of the observed data are inside the confidence interval.

7 Ornstein-Uhlenbeck process as random coefficient.

We will now address the case in which the random coefficient $A(t,\omega)$ is the Ornstein-Uhlenbeck process.

The Ornstein-Uhlenbeck (OU) process was introduced by Leonard Ornstein and George Eugene Uhlenbeck in 1930. This is a mean reverting process (i.e., the process tends to drift towards its long-term mean). This process has been studied deeply by several authors, *see, e.g.* [15]

Consider the following stochastic differential equation (SDE)

$$\begin{aligned} dX_t &= -\theta(X_t - \mu)dt + \sigma dW_t \\ X_0 &= x_0 \end{aligned} \tag{14}$$

where θ, μ, σ are the parameters of the SDE and W_t is a Wiener process:

- μ represents the equilibrium or mean value.
- σ is the degree of errors.
- θ is the rate by which the errors dissipate and the variable reverts towards the mean.
- x_0 is the initial value.

As usual, the OU process is defined as the unique strong solution:

$$X_t = e^{-\theta t}\left(x_0 - \mu(1 - e^{\theta t}) + \sigma \int_0^t e^{\theta s} dW_s\right), \tag{15}$$

when $\mu = 0$ and $x_0 \sim N(0, \sigma^2/2_s)$, it is possible to prove that the covariance function is $cov(X_t, X_s) = 2\theta e^{-\theta t\text{-}s}$.

The Karhunen-Loeve expansion, equation (7), is performed with the set of functions $\phi_j(t)$ and eigenvalues ν_j given by

$$\nu_j = \frac{\sigma^2}{w_j + \theta^2}, \qquad \phi_j(t) = \left(\frac{T}{2} - \frac{\sin\left(2w_j T\right)}{4w_j} \right)^{-1/2} \sin(w_j t), \qquad j = 1, 2, \ldots$$

where w_j are the strictly positive solutions of the following equation

$$\theta \sin(w_j T) + w_j \cos(w_j T) = 0 \tag{16}$$

for $j \geq 1$.

Moreover, ξ_j is given by

$$\xi_j = \frac{1}{\nu_j} \left\langle X_t - \mu_{X_t}, \phi_j \right\rangle_{L^2([0,T])} \tag{17}$$

Note that $\{\xi_j\}$ is a sequence of independent Gaussian random variables. For a detailed derivation of the Karhunen-Loeve expansion for the OU process, see for instance Section 2.3 of [28] or [29].

We choose $P_0(\omega)$ independent of the random vector $\overline{\boldsymbol{\xi}_N} = (\xi_1, \ldots, \xi_N)$.

To simulate solutions of the random logistic equation with the OU process and beta posterior, we need to fix the values of the parameters in the SDE (14). However, estimating these parameters formally and rigorously seems complicated and is beyond the scope of this work.

Therefore, we choose these values heuristically without the use of any inference tools (see appendix for the procedure to fit these values). We fixed the values as $\theta = 1/2$, $\sigma = 1.5$ and $x_0 = 0$.

We repeat the procedure described in the previous section: We run 1000 simulations of the OU process using the function `snssde1d`. We take 1000 samples from the posterior density at time $t = 0$, as shown in figure 6. Finally, we run 1000 simulations of the solution of the equation (6), each of them combining the posterior density $Beta(\alpha, \beta)$ and the previous OU processes simulations.

Figure 7 show the simulations of the solutions of the equation (5). As before, we observe that the dynamic of the differential equation governs the behavior of the solution. In fact, the behavior of the OU case is very similar to the Wiener process. The only difference that we observe is that the confidence intervals are smaller than for the Wiener process.

As in previous section, we now present figure 8 with the mean, confidence intervals at 95% for the simulations and we compare with the real data. From here, we observe that almost all of the experimental data are inside the interval, meaning that the Bayesian inference and the numerical approximation to the solution given by the Karhunen-Loeve expansion are well behaved.

8 Testing the model with new data

In this section we test the model with a different dataset. The observations are stored as a test dataset, which is approximately 30% of the whole dataset. We use the mean of simulations and confidence intervals, and compare with these new data. We observe in figure 9 that the new data is inside the confidence intervals.

We have fitted four models, with different choices of initial condition and different stochastic process $A(t, \omega)$.

To evaluate the goodness of fit the adjusted values of each model where compared, we calculated the square of the errors for each simulation,

$$Sq = \sum_{j=1}^{N} \left[\hat{p}(t_j) - p_j(t) \right]^2, \tag{18}$$

where $\hat{p}(t_j)$ is the predicted curve at the time t_j of the $p_j(t)$ observation. Since $\hat{p}(t_j)$ is random, then we run 1000 simulations and we take the value $\hat{p}(t_j)$ equal to the mean of such simulations, in this way we stabilize the mean curve. We use $\hat{p}(t_j)$ as the predicted curve. We observe that for each set of 1000 simulations, the values of Sq have very small variations. Consequently, we decided to repeat this procedure 10 times, and with these values take the average to provide a better approximation to Sq. We have also calculated the standard deviation of Sq.

The model with the OU process and prior Beta$(2, 2)$ has the smallest sum of squares, which means that the predicted values are closest to the real data and are consequently a better goodness of fit. Moreover, when comparing goodness of fit of the OU process with the Wiener processes, when they are used as stochastic growth rate, we conclude that the the OU process has a better performance no matter the choice of Beta as prior distribution. We also conclude that the election of the prior distributions is not as important as was expected. This means that we could start with a total absence of knowledge and if the data is large enough, then it is possible to get a very similar result if we have a little knowledge about the initial condition. We observe that the confidence interval for the random initial condition of the solution covers the

estimated p_0 in section 5.

We present the results of this procedure in the table 1.

(α, β)	**Wiener process**	**OU process**
(1,1)	0.7104 (0.0221)	0.5398 (0.0226)
(2,2)	0.7081 (0.0216)	0.5238 (0.0182)

Table 1: Mean and Standard deviation of Sq

9 Conclusions

In this paper, we modeled the biological growth of Narcine entemedor with the use of a random logistic differential equation (RLDE). Because we allow the initial condition to be a a random variable, we used the classical Beta-Binomial model from Bayesian inference to incorporate real data in the calculation of the distribution of the mentioned initial condition. Our second main assumption is that the growth rate r is not a deterministic but is a stochastic process. We then used two examples for the stochastic growth rate: the Brownian motion and the Ornstein-Uhlenbeck process. For these examples, the density of the solution by using the Karhunen-Loeve expansion of the solution of the RLDE was constructed numerically. Observe that we have calculated a probability density function, and thus it is possible to calculate directly credible intervals.

The proposed model allows us to include the inherent randomness of the variables in the model, hence the within-subject variation is well represented. In the deterministic model, the parameters are usually estimated by minimizing the square of the difference between the value of the function and the observed data. In contrast, in the stochastic model it is assumed that the initial condition is stochastic. Therefore, it is assumed that the value of the parameter may vary depending on the circumstances. Assuming that the parameter is a random variable, under certain conditions, its probability density function can be obtained and therefore confidence intervals can be calculated. This is an advantage because it is not clear how to build confidence intervals in the deterministic model.

We have also performed several simulations for the RLDE. With these simulations, we construct confidence intervals of their mean. We simulated the initial condition accordingly to the density obtained by the Karhunen-Loeve expansion, and we simulated the stochastic process $A(t, \omega)$ using simulations of the solution of the RLDE. We find that the mean of the simulations fits well the real data. Moreover, the simulated trajectories of the solution reaches maximum sizes at different times, which is consistent with the observation in

nature where different individuals of the same species reach a maximum size at different ages; so the model reproduces this phenomena well.

The method introduced in this paper could be applied to different growth equations; for instance, to a random Von Bertalanffy equation. A natural extension of this research could be to increase the number of term in the Karhunen-Loeve expansion. Another option is to use Wiener-Chaos expansion instead of using KL expansion to numerically construct the density of the solution of the RLDE and compare their performance. Another avenue of future research is to construct a model with the use of a Stochastic Differential Equation, instead of a Random Differential Equation, and then study if it could produce better results than those presented here.

10 Appendix A

In this appendix, it is proved that the function (6) is in fact a solution of the Initial value problem equation (5). Assume that the stochastic process $A(t,\omega)$ is continuous in time. Notice that most of the stochastic process that are continuous functions of the Wiener process, such as the Ornstein-Uhlenbeck process, Geometric Brownian motion, and so on are in fact continuous in time, which implies that the Riemann integrals in time of each of these processes are differentiable functions in time and that their derivatives are continuous in time. Note that,

$$\begin{aligned} 1-P(t,\omega) &= 1-\frac{p_0(\omega)}{p_0(\omega)+\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t}A(s,\omega)ds\Big)} \\ &= \frac{p_0(\omega)+\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t}A(s,\omega)ds\Big)-p_0(\omega)}{p_0(\omega)+\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t}A(s,\omega)ds\Big)} \\ &= \frac{\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t}A(s,\omega)ds\Big)}{p_0(\omega)+\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t}A(s,\omega)ds\Big)}. \end{aligned} \tag{19}$$

The exponential of the integral in time of the stochastic process $A(t,\omega)$ is continuous, then the stochastic process $P(t,\omega)$ has a classical continuous derivative in time. Therefore, taking the derivative in time

$$P'(t,\omega)=\frac{p_0(\omega)\times(-1)\left[p_0(\omega)+\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t}A(s,\omega)ds\Big)\right]'}{\left[p_0(\omega)+\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t}A(s,\omega)ds\Big)\right]^2}$$

$$
\begin{aligned}
&= \frac{p_0(\omega) \times (-1)\left[\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t} A(s,\omega)ds\Big)\right] \times \left[-\int_{t_0}^{t} A(s,\omega)ds\right]'}{\left[p_0(\omega)+\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t} A(s,\omega)ds\Big)\right]^2} \\
&= \frac{p_0(\omega) \times \left[\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t} A(s,\omega)ds\Big)\right] \times A(t,\omega)}{\left[p_0(\omega)+\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t} A(s,\omega)ds\Big)\right]^2} \\
&= A(t,\omega)\frac{p_0(\omega)}{\left[p_0(\omega)+\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t} A(s,\omega)ds\Big)\right]} \\
&\quad \times \frac{\left[\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t} A(s,\omega)ds\Big)\right]}{\left[p_0(\omega)+\big(1-p_0(\omega)\big)\exp\Big(-\int_{t_0}^{t} A(s,\omega)ds\Big)\right]} \\
&= A(t,\omega)P(t,\omega)\Big[1-P(t,\omega)\Big].
\end{aligned}
$$

The last equality comes from equation (19).

Then, the solution of the Initial Value Problem equation (5) is (6).

11 Appendix B

In this appendix, we briefly describe how to estimate the SDE's parameters in (14).

After applying some algebra to equation (6), we obtain,

$$
\ln\left(\frac{P_0\left(\frac{1}{P_t(\omega)}-1\right)}{1-P_0(\omega)}\right) = -\int_0^t A(s,\omega)ds, \quad \text{for } 0<t\leq T. \tag{20}
$$

$A(s,\omega) = X_t(\omega)$ is the solution of the SDE Eq. (14). $x_0 = 0$ was fixed. Note that in (20) the stochastic process $A(s,\omega)$ is behind a Riemann time integral, therefore it is difficult to estimate the parameters θ and σ of the SDE (14). See, [30] for a method to estimate these parameters using observations of the integral through a discretization with time step Δ, which in our case does not apply. We now describe the method to estimate the parameters.

Denote by Y_t as

$$Y_t = \ln\left(\frac{P_0\left(\frac{1}{P_t(\omega)} - 1\right)}{1 - P_0(\omega)}\right).$$

We look for parameters that generate solutions whose integrals "cover" the data. We have performed 1000 simulations of the OU process equation (14), we have integrated them in time, and the confidence intervals and mean of these simulations were obtained. The Function `snssde1d` from the `R-library Sim.DiffProc` [26, 27] was used.

We have chosen these parameters such that the variance of the simulations were minimal. We fixed the parameters $\theta = 1/2$ and $\sigma = 1.5$.

12 Appendix C: Beta-binomial model

We now describe shortly the well-known Beta-Binomial model

Suppose $y_1, \ldots, y_n$ follow independent binomial distributions:

$$Y \sim Bin(n, \theta),$$

where n is known and θ is an unknown parameter that will be assumed as a random variable, thus we assign it to be a prior distribution.

The prior distribution should reflect our ignorance about the unknown parameter θ and it should not have any special properties. Then, we suppose that $\theta \sim Beta(\alpha, \beta)$ with $0 < \alpha < \beta < 1$.

Beta distribution is a conjugate of binomial distribution [24, *see example 5.4]* or [31, *see example 2.2]*, this implies that the posterior distribution is also a Beta. Assuming that the random variable X is Binomial(n, x), the posterior distribution $p(\theta \mid x)$ is also Beta:

$$\theta \mid x \sim Beta(x + \alpha, n - x + \beta). \tag{21}$$

References

[1] Gurney, W.S.C., Tyldesley, G., Wood, S.N., Bacon, P.J., Heath, M.R., Youngson, A., Ibbotson, A.: Modelling length-at-age variability under irreversible growth. Canadian Journal of Fisheries and Aquatic Sciences **64**(4), 638–653 (2007). https://doi.org/10.1139/F07-039

[2] Fujiwara, M., Kendall, B.E., Nisbet, R.M.: Growth autocorrelation and animal size variation. Ecology Letters **7**(2), 106–113 (2004)

[3] Benhaïm, D., Péan, S., Brisset, B., Leguay, D., Bégout, M.-L., Chatain, B.: Effect of size grading on sea bass (dicentrarchus labrax) juvenile self-feeding behaviour, social structure and culture performance. Aquatic Living Resources **24**(4), 391–402 (2011)

[4] Gurney, W.S.C., Veitch, A.R.: The dynamics of size-at-age variability. Bulletin of Mathematical Biology **69**(3), 861–885 (2007). https://doi.org/10.1007/s11538-006-9167-8

[5] Abbott, J.C., Dill, L.M.: The relative growth of dominant and subordinate juvenile steelhead trout (salmo gairdneri) fed equal rations. Behaviour **108**(1-2), 104–113 (1989)

[6] Luquin-Covarrubias, M.A., Morales-Bojórquez, E., González-Peláez, S.S., Hidalgo-De-La-Toba, J., Lluch-Cota, D.B.: Modeling of growth depensation of geoduck clam panopea globosa based on a multimodel inference approach. Journal of Shellfish Research **35**(2), 379–387 (2016). https://doi.org/10.2983/035.035.0212

[7] Luquin-Covarrubias, M.A., Morales-Bojórquez, E., González-Peláez, S.S., Lluch-Cota, D.B.: Joint likelihood function based on multinomial and normal distributions for analyzing the phenotypic growth variability of geoduck clam panopea globosa. California Cooperative Oceanic Fisheries Investigations Reports **57**, 151–162 (2016)

[8] Chen, Y., Fournier, D.: Impacts of atypical data on bayesian inference and robust bayesian approach in fisheries. Canadian Journal of Fisheries and Aquatic Sciences **56**(9), 1525–1533 (1999). https://doi.org/10.1139/f99-076

[9] Tang, M., Jiao, Y., Jones, J.W.: A hierarchical bayesian approach for estimating freshwater mussel growth based on tag-recapture data. Fisheries Research **149**, 24–32 (2014). https://doi.org/10.1016/j.fishres.2013.09.005

[10] López Quintero, F.O., Contreras-Reyes, J.E., Wiff, R., Arellano-Valle, R.B.: Flexible bayesian analysis of the von bertalanffy growth function with the use of a log-skew-t distribution. Fishery Bulletin **115**(1), 13–26 (2017). https://doi.org/10.7755/FB.115.1.2

[11] Haddon, M.: Modelling and Quantitative Methods in Fisheries. CRC press, Boca Raton, Fl (2010)

[12] Helidoniotis, F., Haddon, M., Tuck, G., Tarbath, D.: The relative suitability of the von bertalanffy, gompertz and inverse logistic models for describing growth in blacklip abalone populations (haliotis rubra) in tasmania, australia. Fisheries Research **112**(1-2), 13–21 (2011). https:

//doi.org/10.1016/j.fishres.2011.08.005

[13] Cao, J., Chen, Y., Richards, R.A.: Improving assessment of pandalus stocks using a seasonal, size-structured assessment model with environmental variables. part i: Model description and application. Canadian Journal of Fisheries and Aquatic Sciences **74**(3), 349–362 (2017). https://doi.org/10.1139/cjfas-2016-0020

[14] Han, X., Kloeden, P.E.: Random Ordinary Differential Equations and Their Numerical Solution. Springer, Singapore (2017)

[15] Øksendal, B.: Stochastic differential equations. In: Stochastic Differential Equations, pp. 65–84. Springer, Berlin, Heidelberg (2003)

[16] Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations vol. 23. Springer, Springer Berlin (2013)

[17] Lototsky, S.V., Rozovsky, B.L.: Stochastic Partial Differential Equations. Springer, Springer International Publishing (2017)

[18] Xiu, D.: Numerical Methods for Stochastic Computations: a Spectral Method Approach. Princeton university press, Princeton, New Jersey (2010)

[19] Loeve, M.: Probability Theory. II vol. 46. Graduate Texts in Mathematics, Springer New York, NY (1978)

[20] Cortés, J.-C., Navarro-Quiles, A., Romero, J.-V., Roselló, M.-D.: Analysis of random non-autonomous logistic-type differential equations via the karhunen–loève expansion and the random variable transformation technique. Communications in Nonlinear Science and Numerical Simulation **72**, 121–138 (2019)

[21] Reich, S., Cotter, C.: Probabilistic Forecasting and Bayesian Data Assimilation. Cambridge University Press, England (2015)

[22] Lord, G.J., Powell, C.E., Shardlow, T.: An Introduction to Computational Stochastic PDEs vol. 50. Cambridge Texts in Applied Mathematics, New York (2014)

[23] Gelman, A., Stern, H.S., Carlin, J.B., Dunson, D.B., Vehtari, A., Rubin, D.B.: Bayesian Data Analysis. Chapman and Hall/CRC, Boca Raton, FL (2013)

[24] Bernardo, J.M., Smith, A.F.: Bayesian Theory. John Wiley & Sons, New York (2006)

[25] Burden, R.L., Faires, J.D., Reynolds, A.C.: Numerical Analysis, 7th edn.

Brooks/cole Pacific Grove, CA, Boston (2000)

[26] R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2018). R Foundation for Statistical Computing. https://www.R-project.org/

[27] Guidoum, A.C., Boukhetala, K.: Sim.DiffProc: Simulation of Diffusion Processes. (2018). R package version 4.3. https://cran.r-project.org/package=Sim.DiffProc

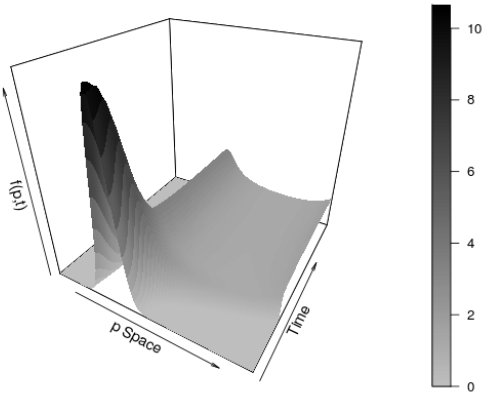
[28] Ghanem, R.G., Spanos, P.D.: Stochastic Finite Elements: A Spectral Approach. Springer, Berlin, Heidelberg (1991)

[29] Corlay, S., Pagés, G.: Functional quantization-based stratified sampling methods. Monte Carlo Methods and Applications **21**(1), 1–32 (2015)

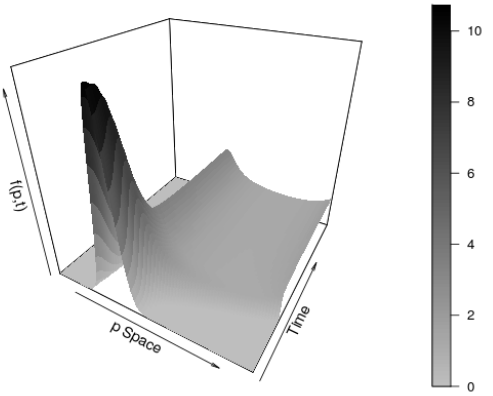
[30] Gloter, A.: Parameter estimation for a discrete sampling of an intergrated ornstein-uhlenbeck process. Statistics **35**(3), 225–243 (2001)

[31] Ghosh, J.K., Delampady, M., Samanta, T.: An Introduction to Bayesian Analysis: Theory and Methods. Springer, New York (2006)

Figure 3 Density of the solution of the random logistic equation using the Wiener process as random coefficient and the *posterior distribution*, where the prior distribution is $Beta(1,1)$ or $Beta(2,2)$.

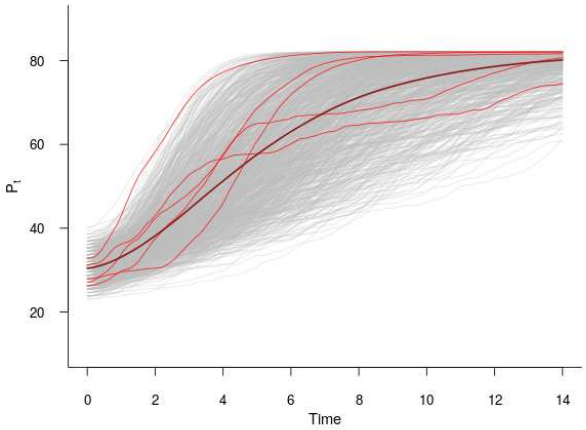


(a) $Beta(1,1)$

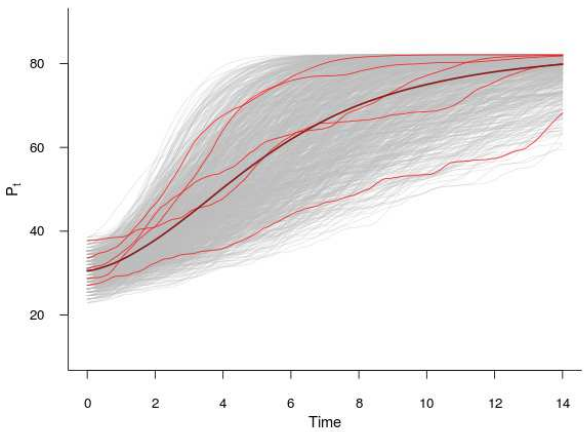


(b) $Beta(2,2)$

Figure 4 Simulation with the data obtained from the logistic regression using the Wiener process as random coefficient and the *posterior distribution*, where the prior distribution is $Beta(1,1)$ or $Beta(2,2)$. Mean of the solution in dark. Each line is a simulated's individual growth and some of them are highlighted in red.

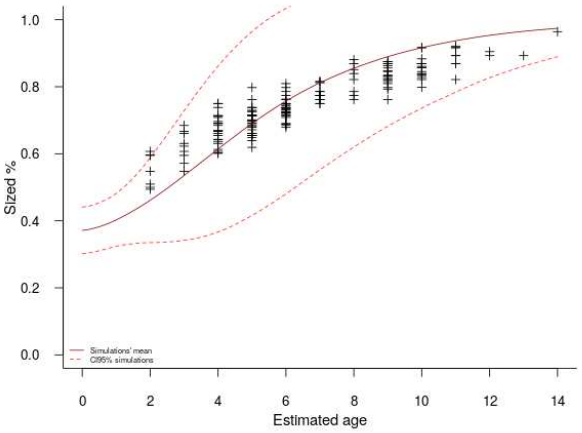


(a) $Beta(1,1)$

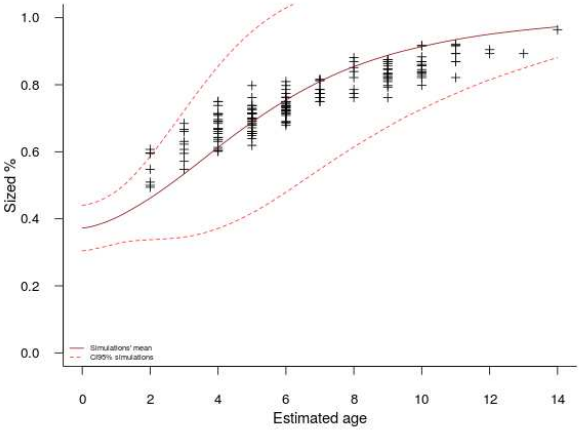


(b) $Beta(2,2)$

Figure 5 Confidence intervals at 95% level using the Wiener process as random coefficient and the *posterior distribution*, where the prior distribution is $Beta(1,1)$ or $Beta(2,2)$.

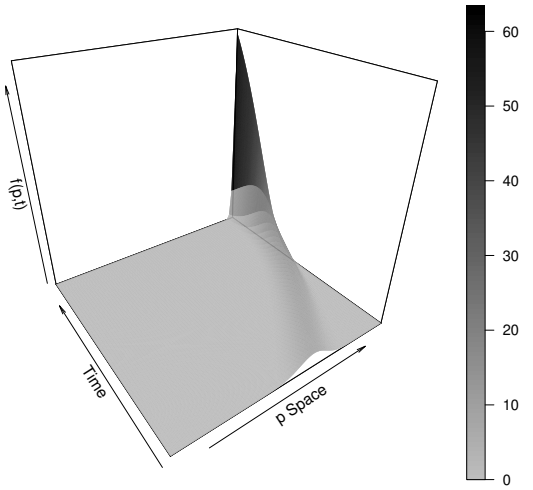


(a) $Beta(1,1)$

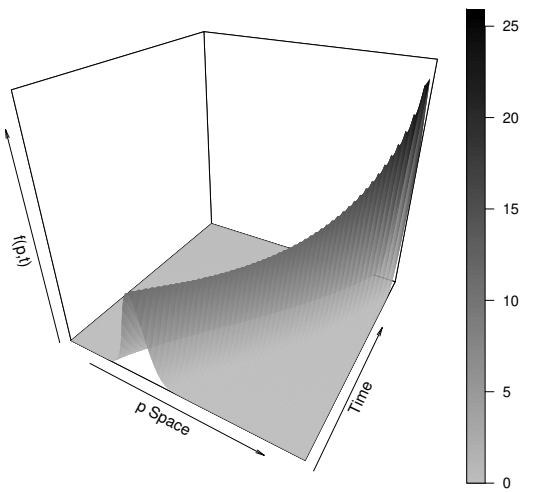


(b) $Beta(2,2)$

Figure 6 Density of the solution of the random logistic equation using the OU process as random coefficient and the *posterior distribution*, where the prior distribution is $Beta(1,1)$ or $Beta(2,2)$.

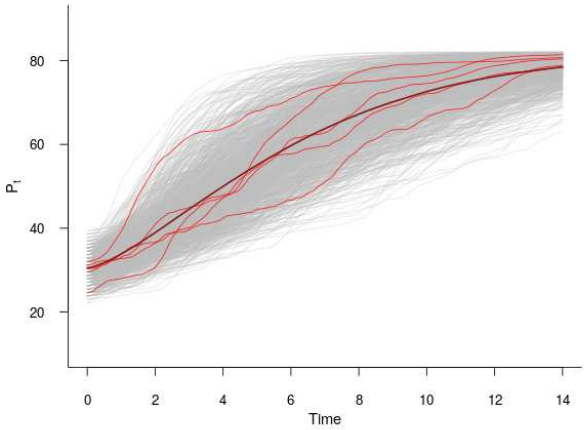


(a) $Beta(1,1)$

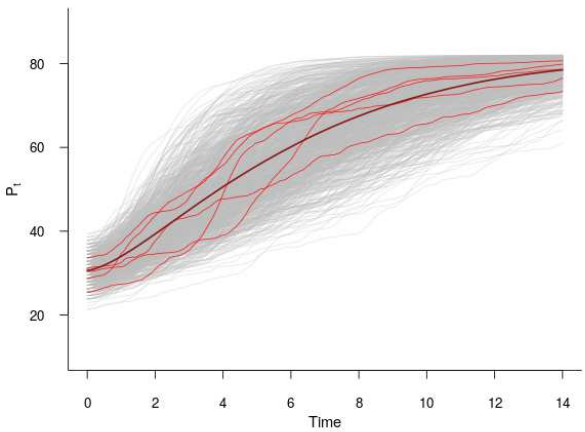


(b) $Beta(2,2)$

Figure 7 Simulation with the data obtained from the logistic regression using the OU process as random coefficient and the *posterior distribution*, where the prior distribution is $Beta(1,1)$ or $Beta(2,2)$. Mean of the solution in dark red. Each line is a simulated's individual growth and some of them are highlighted in red.

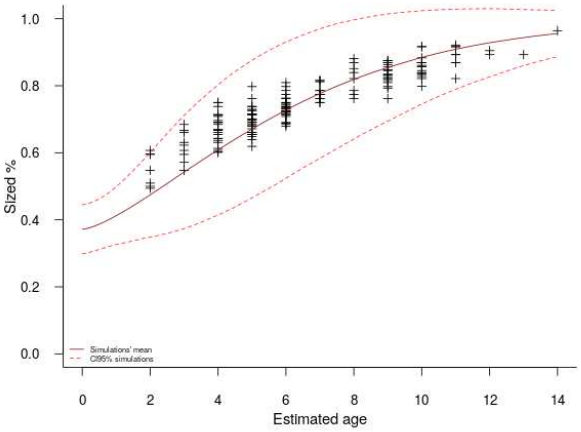


(a) $Beta(1,1)$

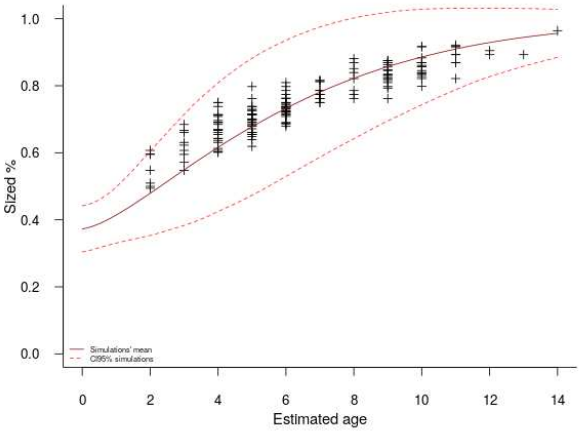


(b) $Beta(2,2)$

Figure 8 Comparison of credibel intervals using the OU process as random coefficient and the *posterior distribution*, where the prior distribution is $Beta(1,1)$ or $Beta(2,2)$..

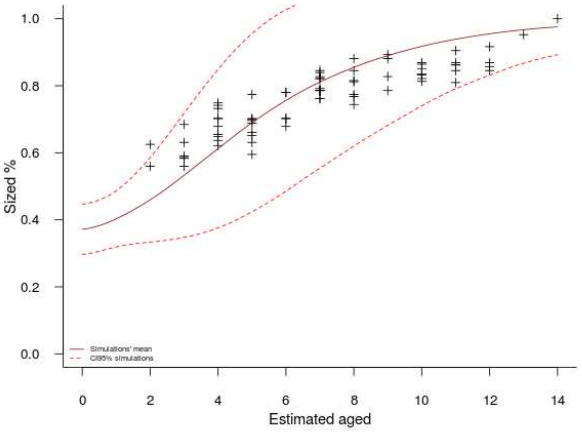


(a) $Beta(1,1)$

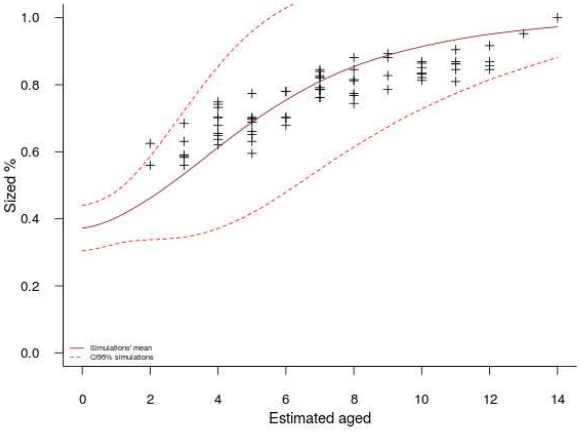


(b) $Beta(2,2)$

Figure 9 Test data in points, confidence intervals in dots. Mean of the solutions of the RDEs in solid line. Here we are using the Wiener process as random coefficient and the *posterior distribution*, where the prior distribution is $Beta(1,1)$ or $Beta(2,2)$.

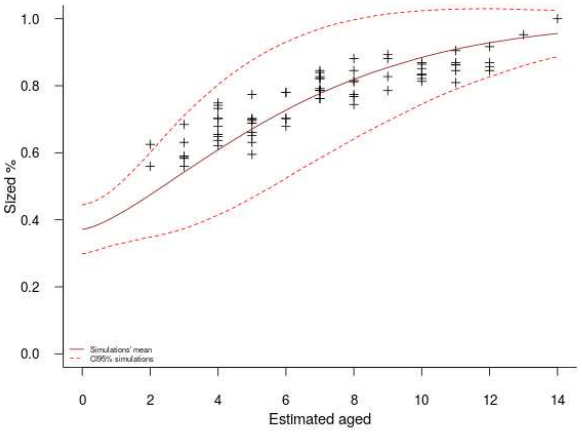


(a) $Beta(1,1)$

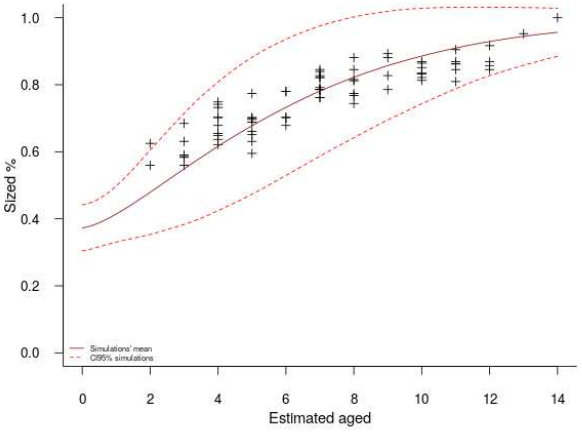


(b) $Beta(2,2)$

Figure 10 Test data in points, confidence intervals in dots. Mean of the solutions of the RDEs in solid line. using the OU process as random coefficient and the *posterior distribution*, where the prior distribution is $Beta(1,1)$ or $Beta(2,2)$.



(a) $Beta(1,1)$



(b) $Beta(2,2)$