Article

Likelihood Function through the Delta Approximation in Mixed SDE Models

Nelson T. Jamba [1,2], Gonçalo Jacinto [1,3], Patrícia A. Filipe [1,4] and Carlos A. Braumann [1,3,*]

1 Centro de Investigação em Matemática e Aplicações, Instituto de Investigação e Formação Avançada, Universidade de Évora, 7000-671 Evora, Portugal; d39830@alunos.uevora.pt (N.T.J.); gjcj@uevora.pt (G.J.); patricia.filipe@iscte-iul.pt (P.A.F.)
2 Liceu n° 918 do Município dos Gambos, Chiange, Gambos, Angola
3 Departamento de Matemática, Escola de Ciências e Tecnologia, Universidade de Évora, 7000-671 Evora, Portugal
4 Quantitative Methods for Management and Economics Department, Iscte Business School, Iscte-Instituto Universitário de Lisboa, 1649-026 Lisbon, Portugal
* Correspondence: braumann@uevora.pt

Abstract: Stochastic differential equations (SDE) appropriately describe a variety of phenomena occurring in random environments, such as the growth dynamics of individual animals. Using appropriate weight transformations and a variant of the Ornstein–Uhlenbeck model, one obtains a general model for the evolution of cattle weight. The model parameters are α, the average transformed weight at maturity, β, a growth parameter, and σ, a measure of environmental fluctuations intensity. We briefly review our previous work on estimation and prediction issues for this model and some generalizations, considering fixed parameters. In order to incorporate individual characteristics of the animals, we now consider that the parameters α and β are Gaussian random variables varying from animal to animal, which results in SDE mixed models. We estimate parameters by maximum likelihood, but, since a closed-form expression for the likelihood function is usually not possible, we approximate it using our proposed delta approximation method. Using simulated data, we estimate the model parameters and compare them with existing methodologies, showing that the proposed method is a good alternative. It also overcomes the existing methodologies requirement of having all animals weighed at the same ages; thus, we apply it to real data, where such a requirement fails.

Keywords: delta approximation; maximum likelihood estimation method; mixed models; stochastic differential equations

1. Introduction

In previous works [1–3], we described the individual growth of animals subject to random fluctuations in the environment and study estimation, prediction, and optimization problems with applications to cattle weight data. Considering M animals, we used the following general SDE model:

$$dY_i(t) = \beta(\alpha - Y_i(t))\,dt + \sigma\,dW_i(t), \quad Y_i(t_{i,0}) = y_{i,0}, \quad i = 1,\dots,M, \tag{1}$$

where $Y_i(t) = h(X_i(t))$ is the modified weight by a transformation function h, a known monotonous continuously differentiable function of the real weight $X_i(t)$ of the animal i at age t, and $y_{i,0} = h(x_{i,0})$, where $x_{i,0}$ is the assumed known size of animal i at an initial age of observation $t_{i,0}$. The parameter $\beta > 0$ is the growth coefficient, and α is the mean asymptotic modified size towards which the mean modified size converges as $t \to +\infty$; we denote the corresponding real asymptotic size by $A = h^{-1}(\alpha)$. The intensity of the effect of environmental random fluctuations on growth is measured by the parameter $\sigma > 0$, being $W_i(t)$ $(i = 1,\dots,M)$ independent realizations of the standard Wiener process. A transformation of the size leading to more general models was also suggested by

Reference [4] with applications to tree growth. Adequate choices of the h function lead to stochastic versions of well-known growth models. For instance, the monomolecular model corresponds to $h(x) = x$, the Bertallanfy-Richards model corresponds to $h(x) = x^c$ (with $c > 0$), the Gompertz model corresponds to $h(x) = \ln x$, and the logistic model corresponds to $h(x) = 1/x$, but the theoretical treatment is valid for any monotonous C^1 function h. Here, for comparison purposes, we will illustrate with $h(x) = \ln x$, corresponding to the stochastic Gompertz model.

It is natural to think that the model parameters may vary from animal to animal, and so SDE models with fixed parameters as the one presented in (1) may not be suitable models for these applications. Parameter estimation for models where the parameters are considered random, known as mixed models or mixed-effects models, is presented in References [1,5–10]. A review on the asymptotic inference of SDE mixed models can be found in Reference [11].

For example, in References [1,7], the mixed model considers that different individuals may have different values of A and, consequently, different values of $\alpha = h(A)$, i.e., the case where the average asymptotic weight varies randomly from animal to animal has been considered. In this particular case, in Reference [1], it was considered that α was a random variable, independent of W_t, with a Gaussian distribution with mean μ and variance θ^2, and the maximum likelihood estimation method was applied to estimate the model parameters. In this case, the likelihood function can be explicitly obtained, but it is extremely difficult or impossible to obtain a closed-form expression for the likelihood function in other cases.

More recently, we considered either α or β random, following a Gaussian distribution, and, for these cases, we solved the integral that appears in the likelihood function through approximation methodologies, such as the Laplace and the delta approximation methods (to appear in a forthcoming paper [12]). In References [8,10], for the general case where it is not possible to obtain a closed-form expression for the likelihood function, as is the case of random β, a numerical approximation based on an Hermite expansion was applied, whereas, in Reference [9], in addition to an Hermite expansion, a Gauss-Hermite quadrature was also applied, and the parameters of the SDE mixed model were estimated by the maximum likelihood method. In References [5,6], for mixed-models with linear drift term, when a closed-form expression for the likelihood function is not possible, a different approximation technique is used, based on a discretized version of the continuous-time data likelihood function.

In this work, we consider that both parameters α and β are random variables, and our main contribution is to extend our delta approximation method to this mixed model with two random parameters and compare its performance with other previously proposed estimation methods. The delta approximation method is inspired on the classical statistical delta method, which is properly adapted to serve a quite different purpose. The delta approximation method allows us to approximate the parameter estimates, through numerical maximization of the approximate likelihood function. This approximation methodology, to the best of our knowledge, is the first method to derive simple closed-form expressions for the approximated likelihood function when both α and β are random, allowing anyone to use this method. Notice also that the existing methods for SDE mixed models [5,6,8], when it comes to their numerical implementations, assume that the age vector of the observations is the same for all trajectories and, in some cases, even require equidistant ages of observation. When using real data of cattle weights, these assumptions are not adequate, since the animals are not weighed at the same time instants (ages), nor even with the same elapsed times between weighings. Our proposed methodology allows parameter estimation in scenarios, quite common in applications, where such restrictions on the data structure are not satisfied.

In order to compare our method with existing ones, particularly with the one provided by the *MsdeParEst* R package (see References [5,6,13]), we worked with simulated cattle weight data with 50, 500, and 5000 animals with the same age vector and with consecutive

weights taken at equidistant intervals. We also estimated the model parameters with our method using a real dataset of 16,029 animals with very heterogeneous ages of observations.

This paper is organized as follows. In Section 2, we present the SDE models with fixed parameters, their main properties, and the respective likelihood function, proceeding to the extension to mixed stochastic differential equations models. In Section 3, we develop the delta approximation applied to the likelihood function for the case where both parameters α and β are independent Gaussian distributed random variables. In Section 4, we present the application for both simulated and real datasets and, whenever possible, compare the results with the existing methods. Based on those results, we present practical recommendations on how to deal with real datasets on Section 5 and end with the main conclusions in Section 6.

2. Stochastic Differential Equations Models

Considering data from M individuals, we will denote the size (some measure of weight, volume, height, length, etc.) at age t of the ith individual ($i = 1, \ldots, M$) by $X_i(t)$. If the individual is growing in a randomly fluctuating environment, working with a modified size $Y_i(t) = h(X_i(t))$, where the transformation h is a monotonous continuously differentiable function, we can describe the evolution of individual growth through an SDE of the form (1). This is a variant of the Ornstein–Uhlenbeck model, also called the Vasicek model in the context of interest rate dynamics [14]. The model solution $Y_i(t)$ is a homogeneous diffusion process with drift coefficient $a(y) = \beta(\alpha - y)$ and diffusion coefficient $b(y) = \sigma^2$, given by

$$Y_i(t) = \alpha - (\alpha - y_{i,0})e^{-\beta(t-t_{i,0})} + \sigma e^{-\beta t} \int_{t_{i,0}}^{t} e^{\beta s} dW_i(s), \quad i = 1, \ldots, M, \tag{2}$$

(see, for instance, Reference [15]).

Let $t_{i,j}$ ($i = 1, \ldots, M, j = 1, \ldots, n_i$) be the age of the jth observation of individual number i and let $Y_{i,j} = Y_i(t_{i,j}) = h(X_i(t_{i,j}))$ be the corresponding modified weight according to model (1). For each individual i ($i = 1, \ldots, M$), denote its age vector of observations (which may differ from individual to individual) by $\mathbf{t}_i = (t_{i,0}, t_{i,1}, \ldots, t_{i,n_i})$, the corresponding vector of modified sizes by $\mathbf{Y}_i = (Y_{i,0}, Y_{i,1}, \ldots, Y_{i,n_i})$, and the observed value of $\mathbf{Y}_i$ by $\mathbf{y}_i = (y_{i,0}, y_{i,1}, \ldots, y_{i,n_i})$. We assume $t_{i,j-1} < t_{i,j}$ and make $E_{i,j} = e^{-(t_{i,j} - t_{i,j-1})}$. We see that, for $Y_{i,j}$ conditioned on $Y_{i,j-1} = y_{i,j-1}$, the transition distribution for animal i is Gaussian:

$$Y_{i,j} | (Y_{i,j-1} = y_{i,j-1}) \sim N\left(\alpha + (y_{i,j-1} - \alpha)E_{i,j}^{\beta}, \frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})\right). \tag{3}$$

In Reference [1], we applied the maximum likelihood estimation method to estimate the parameter vector $\mathbf{p} = (\alpha, \beta, \sigma)$. From (3), using the fact that $Y_i(t)$ is a Markov process, we know that, given $Y_{i,0} = y_{i,0}$ (assumed known), the $\mathbf{Y}_i$ joint probability density function for individual i is given by the product of the transition densities between consecutive observation times of animal i; thus, it takes the form

$$
\begin{aligned}
p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha,\beta,\sigma) &= \prod_{j=1}^{n_i} \frac{\exp\left(-\frac{1}{2}\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)}\right)}{\sqrt{2\pi\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)}} \\
&= \frac{\exp\left(-\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)}\right)}{\prod_{j=1}^{n_i}\sqrt{2\pi\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)}}, \quad i=1,\ldots,M.
\end{aligned}
\tag{4}
$$

By independence among individuals, we obtain the likelihood function for the M animals:

$$
L(\alpha,\beta,\sigma) = \prod_{i=1}^{M} p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha,\beta,\sigma).
\tag{5}
$$

The maximum likelihood estimate of the parameter vector $\mathbf{p}$ is obtained by maximization of (5) or of the log-likelihood function $LL_Y(\alpha,\beta,\sigma) = \ln L(\alpha,\beta,\sigma)$.

$$
\begin{aligned}
LL_Y(\alpha,\beta,\sigma) &= \ln L(\alpha,\beta,\sigma) = \sum_{i=1}^{M} -\frac{n_i}{2}\ln(2\pi) \\
&\quad -\frac{n_i}{2}\sum_{i=1}^{M}\ln\left(\frac{\sigma^2}{2\beta}\right) - \sum_{i=1}^{M}\frac{1}{2}\sum_{j=1}^{n_i}\ln\left(1-E_{i,j}^{2\beta}\right) \\
&\quad -\frac{\beta}{\sigma^2}\sum_{i=1}^{M}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha-\left(y_{i,j-1}-\alpha\right)E_{i,j}^{\beta}\right)^2}{1-E_{i,j}^{2\beta}}.
\end{aligned}
\tag{6}
$$

The maximum likelihood estimators are asymptotically Gaussian with mean vector $\mathbf{p}$ and variance-covariance matrix $\mathbf{V} = \mathbf{F}^{-1}$, where $\mathbf{F}$ is the Fisher information matrix with elements given by $F_{i,j} = -\mathbb{E}\left[\partial^2 L(\mathbf{p})/\partial p_i \partial p_j\right]$. We can estimate $\mathbf{V}$ by the inverse of the empirical information matrix $\hat{\mathbf{F}}$, with elements $\hat{F}_{i,j} = -\partial^2 L(\hat{\mathbf{p}})/\partial p_i \partial p_j$. From these values, we can obtain the approximate confidence bands for the parameters.

For these type of models, in terms of estimation methods, we also developed and applied parametric and non-parametric bootstrap methods [1,3]. Since the asymptotic confidence intervals obtained from the empirical Fisher information matrix may be quite unreliable for small sample sizes, the bootstrap methods can be used in such cases.

In References [1,16], non-parametric estimation methods were developed in order to estimate the drift and diffusion coefficients of a stochastic differential equation model for the case of non-equidistant data. For our application on cattle weight data, we had been working with models with specific functional forms for the drift and the diffusion coefficients, for example, in the case of model (1), with a drift coefficient of linear form $a(y) = \beta(\alpha - y)$ and a constant diffusion coefficient $b(y) = \sigma^2$. These non-parametric methods are useful to assess whether our specific choice of functional forms is appropriate for our data or whether some alternative functional forms for these coefficients are suggested.

Recently, weighted maximum likelihood estimation methods were studied and adapted to overcome one very common limitation in the cattle weight data applications, related to the fact that animals are usually not weighed very frequently, and a scarce number of weight observations exists for older ages. In the weighted maximum likelihood estimation method, the weights are built such that the times elapsed between consecutive observations are considered in the likelihood function [3].

We described the general SDE model (1) for the complete growth curve of the animals where the model's parameters α, β, and σ are assumed common to all individuals. Here, we consider a different generalization to account for the fact that it is natural to think that different animals, due to their specific genetics and other characteristics, may have different values of the parameters. So, in this paper, we will consider the situation where different individuals may have different randomly assigned parameters.

In References [1,5,6,9,10,17], it has been shown that, to consider at least one of the two parameters of the drift term, α or β, as random variables, the likelihood function can be obtained from the transition density function conditioned on the respective random parameter.

We now briefly review how to do it in a more general setting. Let **b** be the d-dimensional vector of parameters that vary randomly among animals and assume that the distribution of **b** among animals has probability density function (p.d.f.) $p_B(\mathbf{b}|\mathbf{\Psi})$, where $\mathbf{\Psi}$ is the parameter vector that characterizes this distribution and needs to be estimated. Assuming independence among the animals, the M parameter vectors $\mathbf{b}_i$ of the different animals i ($i = 1, \ldots, M$) are independent identically distributed random variables with common p.d.f. p_B and assume the $\mathbf{b}_i$ ($i = 1, \ldots, M$) are also independent of the Wiener processes that characterize the environmental conditions under which the animals are growing. Let $\mathbf{\Lambda}$ be the vector of the remaining model parameters (the ones not involved in p_B), assumed to be common to all animals. The likelihood function for M trajectories (animals) is given by

$$
\begin{aligned}
L(\mathbf{\Lambda}, \mathbf{\Psi}) &= \prod_{i=1}^{M} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{\Lambda}, \mathbf{\Psi}) \\
&= \prod_{i=1}^{M} \int_{\mathbb{R}^d} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{b}_i, \mathbf{\Lambda}) p_B(\mathbf{b}_i|\mathbf{\Psi}) \, d\mathbf{b}_i.
\end{aligned}
\tag{7}
$$

The case of a single random parameter $\mathbf{b}_i = (\alpha_i)$ has already been studied for when $\alpha_i \sim N(\mu, \theta^2)$. When we have the special situation of a time vector of observations $\mathbf{t}_i \equiv \mathbf{t} = (t_0, t_1, \ldots, t_n)$, $i = 1, \ldots, M$ common to all animals, we can see it in Reference [8] (for the particular situation of $\mu = 0$ and uniform time spacing $t_j - t_{j-1} \equiv \delta$) and in References [6,7]. The general situation, with no such restrictions, can be seen in Reference [1]. In this case, it is possible to explicitly compute the integral in the likelihood function, resulting in a final closed-form expression for this function. This is shown in Reference [1], where the log-likelihood function for all animals $LL_Y(\mu, \theta, \beta, \sigma) = \ln L(\mu, \theta, \beta, \sigma)$ is given by

$$
\begin{aligned}
LL_Y(\mu, \theta, \beta, \sigma) = \sum_{i=1}^{M} \Bigg(& -\frac{n_i}{2}\ln(2\pi) - \frac{n_i}{2}\ln\left(\frac{\sigma^2}{2\beta}\right) - \frac{1}{2}\sum_{j=1}^{n_i}\ln\left(1 - E_{i,j}^{2\beta}\right) - \frac{1}{2}\ln(D_i) \\
& -\frac{\beta}{D_i\sigma^2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j} - \mu - (y_{i,j-1} - \mu)E_{i,j}^{\beta}\right)^2}{1 - E_{i,j}^{2\beta}} + \frac{2\beta^2\theta^2}{D_i\sigma^4}\left(\sum_{j=1}^{n_i}\frac{\left(y_{i,j} - y_{i,j-1}E_{i,j}^{\beta}\right)}{1 + E_{i,j}^{\beta}}\right)^2 \\
& -\frac{\beta(D_i - 1)}{D_i\sigma^2}\left(\sum_{j=1}^{n_i}\frac{\left(y_{i,j} - y_{i,j-1}E_{i,j}^{\beta}\right)^2}{1 - E_{i,j}^{2\beta}}\right)\Bigg),
\end{aligned}
\tag{8}
$$

with $D_i = \frac{2\beta\theta^2}{\sigma^2}\sum_{j=1}^{n_i}\frac{1 - E_{i,j}^{\beta}}{1 + E_{i,j}^{\beta}} + 1$. However, despite the existence of a closed-form expression for the likelihood function for this particular case, we also applied approximation methods (Laplace and delta approximations—to appear in a forthcoming paper [12]), showing that the approximation methods also provide very good results when compared with the exact method.

Unfortunately, the integral in (7) does not always have a closed-form expression, for example, when the random parameter is β, corresponding to $\mathbf{b}_i = (\beta_i)$, and, in such cases, the approximation methods are good alternatives. In the forthcoming paper [12]), we

used the Laplace and delta approximations to obtain closed-form approximate expressions of the likelihood function for this case, also with very good results.

In the following section, we present the study of the case where both α and β are considered random variables and propose the application of the delta approximation method to the integral in (7).

3. Mixed Model for Two Random Effects

The mixed model where both α and β are random variables can be written as

$$dY_i(t) = \beta_i(\alpha_i - Y_i(t))\,dt + \sigma\,dW_i(t), \quad Y_i(t_{i,0}) = y_{i,0}, \quad i = 1, \ldots, M, \tag{9}$$

where we assume that the $2M$ random variables $\alpha_i \sim \mathcal{N}(\mu, \theta^2)$ $(i = 1, \ldots, M)$ and $\beta_i \sim \mathcal{N}(\lambda, \omega^2)$ $(i = 1, \ldots, M)$ are independent. Now, the random parameter vector is the 2-dimensional vector $\mathbf{b}_i = (\alpha_i, \beta_i)$, the fixed effects parameter vector becomes the 1×1 vector $\mathbf{\Lambda} = (\sigma)$, the random effects parameter vector is $\mathbf{\Psi} = (\mu, \lambda, \theta, \omega)$, and the p.d.f. of the random effects p_B is a bivariate Gaussian distribution.

To obtain the likelihood function in (7) for this situation of two random parameters, notice that, due to the assumed independence structure,

$$
\begin{aligned}
p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{\Lambda}, \mathbf{\Psi}) &= \int_{\mathbb{R}^2} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{b}_i, \mathbf{\Lambda}) p_B(\mathbf{b}_i|\mathbf{\Psi})\,d\mathbf{b}_i \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i, \beta_i, \mathbf{\Lambda}) p(\alpha_i|\mu, \theta) p(\beta_i|\lambda, \omega)\,d\alpha_i\,d\beta_i,
\end{aligned}
\tag{10}
$$

with

$$p(\alpha_i|\mu, \theta) = \frac{1}{\sqrt{2\pi\theta^2}} \exp\left(-\frac{(\alpha_i - \mu)^2}{2\theta^2}\right), \tag{11}$$

$$p(\beta_i|\lambda, \omega) = \frac{1}{\sqrt{2\pi\omega^2}} \exp\left(-\frac{(\beta_i - \lambda)^2}{2\omega^2}\right). \tag{12}$$

To determine $p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i, \beta_i, \mathbf{\Lambda})$ in (10), notice that, due to the Markov property of each individual trajectory i when conditioned on α_i and β_i, we have

$$p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i, \beta_i, \mathbf{\Lambda}) = \prod_{j=1}^{n_i} p_{\mathbf{Y}_i}(y_{i,j}, t_i, t_{i-1}|y_{i,j-1}, \alpha_i, \beta_i, \mathbf{\Lambda}), \tag{13}$$

with the transition densities between consecutive observation ages of the animal i trajectory given by

$$p_{\mathbf{Y}_i}(y_{i,j}, t_i, t_{i-1}|y_{i,j-1}, \alpha_i, \beta_i, \mathbf{\Lambda}) = \frac{\exp\left(-\frac{1}{2}\frac{\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})}\right)}{\sqrt{2\pi\frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})}}. \tag{14}$$

We will apply the delta approximation to the integral in (7), which, in this case, is given by (10) and does not have an explicit solution. This approximation method allows us to obtain simpler and closed-form expressions for the likelihood function. The method is also adapted to general age vectors of observations, allowing for estimation of the parameters of the models even when the different animals have their weights observed at different ages, and those ages may be non-equidistant, which, as far as we are aware, is theoretically possible but is not currently available in the numerical implementations of other existing approximation methods.

Let us denote expression (14) by $u_i(\alpha_i, \beta_i) = \exp(g_i(\alpha_i, \beta_i))$, where

$$g_i(\alpha_i, \beta_i) = \sum_{j=1}^{n_i} \left(-\frac{1}{2} \frac{\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i) E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})} - \frac{1}{2} \ln\left(2\pi \frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})\right) \right). \tag{15}$$

Taking into account the previous expressions, the likelihood function (7) can be writen as an expectation w.r.t. the bivariate Gaussian distribution with density p_B:

$$L(\mu, \theta, \lambda, \omega, \sigma) = \prod_{i=1}^{M} \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} u_i(\alpha_i, \beta_i) p_B(\alpha_i, \beta_i | \mu, \lambda, \theta, \omega) \, d\alpha_i \, d\beta_i \right)$$
$$= \prod_{i=1}^{M} \mathbb{E}[u_i(\alpha_i, \beta_i)]. \tag{16}$$

The mathematical expectation of $u_i(\alpha_i, \beta_i)$ in (16) can be approximately obtained by applying the delta approximation, obtained by adapting the delta method. The delta approximation consists of using the second-order Taylor series expansion about the point (μ, λ)

$$u_i(\alpha_i, \beta_i) \approx u_i(\mu, \lambda) + \left[\frac{\partial u_i}{\partial \alpha_i}(\mu, \lambda) \quad \frac{\partial u_i}{\partial \beta_i}(\mu, \lambda) \right] \begin{bmatrix} \alpha_i - \mu \\ \beta_i - \lambda \end{bmatrix}$$
$$+ \frac{1}{2} \begin{bmatrix} \alpha_i - \mu & \beta_i - \lambda \end{bmatrix} \begin{bmatrix} \frac{\partial^2 u_i}{\partial \alpha_i^2}(\mu, \lambda) & \frac{\partial^2 u_i}{\partial \alpha_i \beta_i}(\mu, \lambda) \\ \frac{\partial^2 u_i}{\partial \beta_i \alpha_i}(\mu, \lambda) & \frac{\partial^2 u_i}{\partial \beta_i^2}(\mu, \lambda) \end{bmatrix} \begin{bmatrix} \alpha_i - \mu \\ \beta_i - \lambda \end{bmatrix}. \tag{17}$$

We now apply mathematical expectations to this expression. Noticing that

$$\mathbb{E}[\alpha_i - \mu] = \mu - \mu = 0, \mathbb{E}[\beta_i - \lambda] = \lambda - \lambda = 0, \mathbb{E}[(\alpha_i - \mu)^2] = \theta^2, \mathbb{E}[(\beta_i - \lambda)^2] = \omega^2,$$

and, due to independence between α_i and β_i, $\mathbb{E}[(\alpha_i - \mu)(\beta_i - \lambda)] = 0$, we obtain

$$\mathbb{E}[u_i(\alpha_i, \beta_i)] \approx u_i(\mu, \lambda) + \frac{\theta^2}{2} \frac{\partial^2 u_i}{\partial \alpha_i^2}(\mu, \lambda) + \frac{\omega^2}{2} \frac{\partial^2 u_i}{\partial \beta_i^2}(\mu, \lambda). \tag{18}$$

Finally, the log-likelihood function, $LL_Y(\mu, \theta, \lambda, \omega, \sigma) = \ln L(\mu, \theta, \lambda, \omega, \sigma)$, can be obtained as

$$LL_Y(\mu, \theta, \lambda, \omega, \sigma) \approx \sum_{i=1}^{M} \ln \left\{ u_i(\mu, \lambda) + \frac{\theta^2}{2} \frac{\partial^2 u_i}{\partial \alpha_i^2}(\mu, \lambda) + \frac{\omega^2}{2} \frac{\partial^2 u_i}{\partial \beta_i^2}(\mu, \lambda) \right\}. \tag{19}$$

By the chain rule, the second-order derivative of $u_i(\alpha_i, \beta_i)$ with respect to α_i is

$$\frac{\partial^2 u_i(\alpha_i, \beta_i)}{\partial \alpha_i^2} = u_i(\alpha_i, \beta_i) \left(\frac{\partial^2 g_i(\alpha_i, \beta_i)}{\partial \alpha_i^2} + \left(\frac{\partial g_i(\alpha_i, \beta_i)}{\partial \alpha_i} \right)^2 \right), \tag{20}$$

where the first and second-order derivatives with respect to α_i are given by

$$\frac{\partial g_i(\alpha_i, \beta_i)}{\partial \alpha_i} = \sum_{j=1}^{n_i} \left(-\frac{2\beta_i(-1 + E_{i,j}^{\beta_i})\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i) E_{i,j}^{\beta_i}\right)}{(1 - E_{i,j}^{2\beta_i})\sigma^2} \right)$$
$$\frac{\partial^2 g_i(\alpha_i, \beta_i)}{\partial \alpha_i^2} = \sum_{j=1}^{n_i} \left(-\frac{2\beta_i(-1 + E_{i,j}^{\beta_i})^2}{(1 - E_{i,j}^{2\beta_i})\sigma^2} \right). \tag{21}$$

Similarly, the second-order derivative of $u_i(\alpha_i, \beta_i)$ with respect to β_i is given by

$$\frac{\partial^2 u_i(\alpha_i, \beta_i)}{\partial \beta_i^2} = u_i(\alpha_i, \beta_i)\left(\frac{\partial^2 g_i(\alpha_i, \beta_i)}{\partial \beta_i^2} + \left(\frac{\partial g_i(\alpha_i, \beta_i)}{\partial \beta_i}\right)^2\right), \tag{22}$$

where the first two derivatives of $g_i(\alpha_i, \beta_i)$ with respect to β_i are given by

$$\frac{\partial g_i(\alpha_i, \beta_i)}{\partial \beta_i} = \sum_{j=1}^{n_i}\left(\frac{1}{2\beta_i} + \frac{2\beta_i E_{i,j}^{2\beta_i}(t_{i,j} - t_{i,j-1})\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)^2}{\left(1 - E_{i,j}^{2\beta_i}\right)^2 \sigma^2}\right.$$

$$-\frac{E_{i,j}^{2\beta_i}(t_{i,j} - t_{i,j-1})}{\left(1 - E_{i,j}^{2\beta_i}\right)} - \frac{\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)^2}{\left(1 - E_{i,j}^{2\beta_i}\right)\sigma^2} \tag{23}$$

$$\left.-\frac{2\beta_i E_{i,j}^{\beta_i}(t_{i,j} - t_{i,j-1})(y_{i,j-1} - \alpha_i)\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)}{\left(1 - E_{i,j}^{2\beta_i}\right)\sigma^2}\right)$$

and

$$\frac{\partial^2 g_i(\alpha_i, \beta_i)}{\partial \beta_i^2} = \sum_{j=1}^{n_i}\left(-\frac{4\beta_i E_{i,j}^{2\beta_i}(t_{i,j} - t_{i,j-1})^2\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)^2}{\left(1 - E_{i,j}^{2\beta_i}\right)^2 \sigma^2}\right.$$

$$-\frac{1}{2\beta_i^2} - \frac{4E_{i,j}^{\beta_i}(t_{i,j} - t_{i,j-1})(y_{i,j-1} - \alpha_i)\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)}{\left(1 - E_{i,j}^{2\beta_i}\right)\sigma^2}$$

$$+\frac{2E_{i,j}^{2\beta_i}(t_{i,j} - t_{i,j-1})^2}{\left(1 - E_{i,j}^{2\beta_i}\right)} + \frac{2\beta_i E_{i,j}^{\beta_i}(t_{i,j} - t_{i,j-1})^2(y_{i,j-1} - \alpha_i)\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)}{\left(1 - E_{i,j}^{2\beta_i}\right)\sigma^2}$$

$$+\frac{2E_{i,j}^{4\beta_i}(t_{i,j} - t_{i,j-1})^2}{\left(1 - E_{i,j}^{2\beta_i}\right)^2} + \frac{8\beta_i E_{i,j}^{3\beta_i}(t_{i,j} - t_{i,j-1})^2(y_{i,j-1} - \alpha_i)\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)}{\left(1 - E_{i,j}^{2\beta_i}\right)^2 \sigma^2} \tag{24}$$

$$+\frac{4E_{i,j}^{2\beta_i}(t_{i,j} - t_{i,j-1})\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)^2}{\left(1 - E_{i,j}^{2\beta_i}\right)^2 \sigma^2}$$

$$-\frac{8\beta_i E_{i,j}^{4\beta_i}(t_{i,j} - t_{i,j-1})^2\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)^2}{\left(1 - E_{i,j}^{2\beta_i}\right)^3 \sigma^2}$$

$$\left.-\frac{2\beta_i E_{i,j}^{2\beta_i}(t_{i,j} - t_{i,j-1})^2(y_{i,j-1} - \alpha_i)^2}{\left(1 - E_{i,j}^{2\beta_i}\right)\sigma^2}\right).$$

Replacing these derivatives in the expression (19), we obtain an approximate expression for the log-likelihood function $LL_Y(\mu, \theta, \lambda, \omega, \sigma, \rho) = \ln L(\mu, \theta, \lambda, \omega, \sigma, \rho)$ for the case of two random Gaussian parameters. Using it, in the next section, we are going to apply this approach to obtain approximate maximum likelihood parameter estimates for the mixed stochastic Gompertz model with two random parameters using cattle weight data.

4. Results

The main interest with the methodology proposed in the previous sections is to have a reliable estimation method that can be applied to real animal weight data, where the animals' weights are often not taken at the same age instants.

In our application, we worked with real cattle weight data provided by the Associação de Criadores de Bovinos Mertolengos (ACBM), which performs the growing and finishing phases of young Mertolengo males, and by associated breeders, from the Alentejo region in Portugal. The available dataset contains a total of 96,204 observations of the weight (in kg) of 16,029 Mertolengo cattle males, where each animal has several observations with a minimum of 3 and a maximum of 33 weights at ages varying from birth until a maximum age that ranges between 0.2 and 16 years old. This is a case where, indeed, each animal has its weight measurements taken at different (and even non-equidistant) age instants, varying from animal to animal.

We obtained the estimation results implementing the proposed delta approximation method (DA) in the software R project [18]. However, since it is a new method, we first will compare the results obtained with this DA method with the estimation methods available in the R package *MsdeParEst* described in Reference [13], which we will call MPE methods. The package includes, when closed-form likelihood function expressions are not possible, techniques based on a discretized version of the continuous-time data maximum likelihood developed in References [5,6]. The main drawback of this package (and, to the best of our knowledge, of other similar available R packages [19,20]) is that all observations must be measured at the same age instants for all animals, and, in some cases, the age instants are even required to be equidistant. Therefore, in order to compare the performance of the proposed approximation method DA with the methods available in the literature (in particular, the MPE method used here), we would also need datasets where weights are observed at the same equidistant age instants for all the animals. Whenever possible, we will also compare the DA and MPE methods, which consider approximations for the likelihood function, with existing estimation methods that use exact closed-form expressions for the likelihood function, in particular, when the likelihood function assumes both parameters fixed or Non-Mixed (which we will call the NMSDE method, presented in Reference [1]) or, if appropriate, when the likelihood function assumes just the parameter α as random (which we will call the Exact(α) method, presented in Reference [1]).

For this purpose, we worked with simulated datasets of equidistant monthly weights for *M* animals, since birth until four years of age, totaling *M**49 observations.

The animal's weights were simulated based on the stochastic Gompertz model ($Y(t) = h(X(t)) = \ln(X(t))$). We simulated four datasets:

- $D_M(\alpha, \beta)$: Mixed SDE model with random independent α_i and β_i, where $\alpha_i \sim \mathcal{N}(\mu, \theta^2)$ and $\beta_i \sim \mathcal{N}(\lambda, \omega^2)$ with $\mu = 6.45$ ($h^{-1}(\mu) = 632.70$ kg), $\theta = 0.15$, $\lambda = 1.43$ year^{-1}, $\omega = 0.30$ year^{-1}, and with fixed parameter $\sigma = 0.33$ year$^{-1/2}$;

- $D_M(\alpha)$: Mixed SDE model with random $\alpha_i \sim \mathcal{N}(\mu, \theta^2)$ with $\mu = 6.45$ ($h^{-1}(\mu) = 632.70$ kg) and $\theta = 0.15$, and with fixed parameters $\beta = 1.43$ year^{-1} and $\sigma = 0.33$ year$^{-1/2}$;

- $D_M(\beta)$: Mixed SDE model with random $\beta \sim \mathcal{N}(\lambda, \omega^2)$ with $\lambda = 1.43$ year^{-1} and $\omega = 0.30$ year^{-1}, and with fixed parameters $\alpha = 6.45$ ($h^{-1}(\alpha) = 632.70$ kg) and $\sigma = 0.33$ year$^{-1/2}$;

- D_M: Non-mixed SDE model (1) with fixed parameters $\alpha = 6.45$ ($h^{-1}(\alpha) = 632.70$ kg), $\beta = 1.43$ year^{-1} and $\sigma = 0.33$ year$^{-1/2}$.

Each of the simulated datasets based on a Mixed SDE model were obtained in the following way. First, the *M* random parameter values $\mathbf{b}_i$ ($i = 1, \ldots, M$) of the *M* animals were simulated based on the Gaussian distribution p_B; then, for each animal i ($i = 1, \ldots, M$), the simulated values were incorporated in the transition density (3) as the true values of α_i and/or β_i for that animal. Then, we simulated the weights of each animal based on the Markov property and the use of the transition densities between consecutive observation ages. For the non-mixed or fixed parameters model, the simulation of the dataset is obtained without simulating the parameter values of α_i and β_i since they are fixed.

The advantage of using simulated data, besides the possibility of comparing the different methods due to its common age vector of observations, is the knowledge of the true parameter values, which allows us to compare the performances of the different

methods when they can be applied. We first consider simulated datasets for weights of $M = 5000$ animals to evaluate the behavior of the methods for a large number of trajectories close to an asymptotic regimen. However, in applications, we usually do not have datasets with so many animals available, and it is important to evaluate the performance of the different methods for smaller datasets. For this reason, we also present a comparison of the main methods, considering datasets with $M = 500$ and $M = 50$ animals. To distinguish them, we index the simulated datasets by the number of animals M.

Tables 1–4 presents the results for each of the datasets, $D_{5000}(\alpha, \beta)$, $D_{5000}(\alpha)$, $D_{5000}(\beta)$, and D_{5000}, simulated under the four stochastic Gompertz mixed and fixed models. In each table, we present the results obtained using the different DA methods, DA(α, β) (the one that assumes both parameters α and β random and is described in Section 3), and DA(α) and DA(β) (the ones that assume just α random or just β random as described in Reference [12]), as well as the estimates obtained by the different MPE methods, MPE(α, β), MPE(α), and MPE(β) (again, when we consider random both or just one of the parameters). We also present, for comparison purposes, the results obtained when using the NMSDE estimation method (which assumes both parameters to be fixed).

In this way, we can assess what happens when an appropriate estimation method is applied (for instance, when applied to the $D_{5000}(\alpha)$ dataset, the estimation method MPE(α) and the estimation method DA(α, β) are appropriate) and when an inappropriate estimation method is applied (for instance, when applied to the $D_{5000}(\alpha, \beta)$ dataset, the MPE(α) method is inappropriate since it cannot estimate the parameters involved in the random β); this may be important since, in real non-simulated data, we may not know which parameters are random or not. Of course, DA(α, β) and MPE(α, β) are appropriate for all datasets, but, when the data has one or both parameters as non-random, they are overparametrized and, therefore, may not be as accurate as non-overparametrized methods.

In Tables 1–4, we have also underlined the headings of the estimation methods that are appropriate to analyze the corresponding dataset, in the sense of being able to estimate all the parameters of the model used to simulate the dataset. Appropriate methods include the ones that use, in the likelihood function or its approximation, the same random parameters that were used in the dataset generation (these would be the most appropriate methods when we know beforehand which parameters are random) but include, as well, overparametrized methods that also use additional random parameters in the likelihood function or its approximation.

The maximum likelihood estimates and the approximate 95% confidence bands based on the inverse of the empirical Fisher information matrix are presented only for the DA methods and the NMSDE method but not for the MPE methods since the confidence intervals of the estimates are not provided by the R package *MsdeParEst*. Instead of presenting the results for the overall mean of the modified asymptotic weight μ (which is $=\alpha$ when α is a fixed parameter), we present them in terms of the corresponding actual weight value $h^{-1}(\mu)$.

The tables also display the log-likelihood function values computed at the estimated parameter values. We display the log-likelihood values LL_X in terms of the actual weights X, which can be easily obtained by a simple conversion using the function h from the corresponding log-likelihood values LL_Y (expressed in terms of the modified weights Y we have been working with so far in all computations). However, although the displayed log-likelihood values use the exact log-likelihood known expressions for the NMSDE and the Exact(α) methods, for the DA methods, they use the approximate log-likelihood expressions of those methods, and the same is possibly happening with the values delivered by the *MsdeParEst* Package for the MPE methods. Since these approximations involve some incomparable errors, the displayed log-likelihood values LL_X should, with rare exceptions, not be used for comparisons in which an approximate method is involved.

When considering, in Table 1, the $D_{5000}(\alpha, \beta)$ dataset, the appropriate estimation methods are the ones assuming α and β random, the DA(α, β) method we have proposed, and the MPE(α, β) method. We can observe that the DA(α, β) method provides a lower bias for all parameters than the MPE(α, β) method, noticing that, in the MPE(α, β) method,

the estimate of ω is extremely biased, close to zero. The use of an inappropriate method should obviously be avoided, but, when using real data, we often do not know beforehand exactly which are the random parameters and may make a wrong assumption, leading to the use of an inappropriate estimation method. These simulated datasets, in which we know exactly which parameters are random, give us the opportunity to assess the consequences of a wrong assumption on the random parameters when dealing with real data. Let us look at this issue, beyond the obvious fact that an inappropriate method cannot estimate some of the parameters. When comparing the appropriate DA(α, β) method with the inappropriate (for this dataset) NMSDE method, the former method leads to less biased estimates. When comparing the appropriate DA(α, β) with the inappropriate (for this dataset) DA(α) and DA(β) methods, the inappropriate methods estimate the standard deviation of their own random parameter (respectively, the standard deviation θ of the random α and the standard deviation ω of the random β) better than the appropriate method, which underestimates both standard deviations. On the estimation of the noise intensity parameter σ, the appropriate DA(α, β) method shows the better performance, and it also outperforms the inappropriate DA(β) method on the estimation of the means $h^{-1}(\mu)$ and λ, but these means are curiously better estimated by the inappropriate DA(α) method. As for the inappropriate MPE(α) and MPE(β) methods, besides the obvious impossibility of estimating one parameter, they have a behavior that is not too different from the MPE(α, β) appropriate method.

Table 1. Results for the simulated dataset of 5000 animals when both α and β are random-$D_{5000}(\alpha, \beta)$. Besides the true parameter values used in the simulations, the table shows their maximum likelihood estimates and, when available, approximate 95% confidence bands for the different estimation methods: delta approximation methods, DA(α, β), DA(α), and DA(β); NMSDE method (which wrongly assumes α and β fixed and works with the corresponding exact likelihood); and MPE methods, MPE(α, β), MPE(α), and MPE(β). For each method, the indicated parameters are the ones taken as random in its likelihoods or approximate likelihoods. Appropriate methods (i.e., those that allow the estimation of all the dataset parameters) are underlined. Notice that the LL_X values of approximate methods may differ from the true values of the log-likelihood function at the parameter estimated values.

	True	**DA(α, β)**	**DA(α)**	**DA(β)**
$h^{-1}(\mu)$	632.70	640.87 ± 3.17	629.77 ± 3.16	651.31 ± 2.83
θ	0.15	0.1011 ± 0.0040	0.1397 ± 0.0024	$-$
λ	1.43	1.3316 ± 0.0067	1.3548 ± 0.0060	1.3036 ± 0.0065
ω	0.30	0.1786 ± 0.0042	$-$	0.2002 ± 0.0029
σ	0.33	0.3399 ± 0.0010	0.3414 ± 0.0010	0.3405 ± 0.0001
LL_X		$-1,191,084$	$-1,192,146$	$-1,191,418$
	NMSDE	**MPE(α, β)**	**MPE(α)**	**MPE(β)**
$h^{-1}(\mu)$	646.29 ± 2.92	623.77	624.34	622.40
θ	$-$	0.2158	0.2127	$-$
λ	1.3208 ± 0.0061	1.3272	1.3249	1.3182
ω	$-$	0.0001	$-$	0.0444
σ	0.3452 ± 0.0010	0.2983	0.3140	0.3155
LL_X	$-1,194,145$	$-1,209,481$	$-1,208,249$	$-1,208,631$

Table 2 shows the results for the $D_{5000}(\alpha)$ dataset. For this case, we included the Exact(α) method, where the maximum likelihood is obtained exactly by the closed-form expression (8), and that is certainly the best choice method for this dataset. The methods DA(α, β) and DA(α) are both appropriate for this dataset, but the first is overparametrized, while the latter is the most appropriate of the two, and we expect it to be more accurate. Still, it is interesting to notice that these two methods provide exactly the same parameter estimates, except, naturally, for $\omega = 0$ ("parameter" obviously out of the DA(α) method).

Comparing the MPE(α, β) with the MPE(α) methods, they do not give exactly the same parameter estimates, but their estimates are very close. As expected, the most appropriate method for this dataset, the Exact(α) method, provides slightly better estimates than the DA(α) and the DA(α, β) methods, and than the MPE(α) and the MPE(α, β) methods. The estimates for $h^{-1}(\mu)$ and θ are better when using the appropriate MPE methods than the ones obtained with the corresponding DA methods, but the reverse happens for the estimates of the mean λ and the standard deviations $\omega = 0$ of parameter β. The inappropriate methods DA(β) and MPE(β) give worse estimates than the appropriate methods of the same type. The use of the inappropriate (for this data) NMSDE method performs quite badly in the estimation of $h^{-1}(\mu)$. Notice that the dataset was simulated for fixed β, i.e, $\omega = 0$, and this is better captured by the DA method.

Table 2. Results for the simulated dataset of 5000 animals when α is random, but β is fixed ($\omega = 0$) – $D_{5000}(\alpha)$. Besides the true parameter values used in the simulations, the table shows their maximum likelihood estimates and, when available, approximate 95% confidence bands, for the different estimation methods: DA(α, β), DA(α), DA(β), NMSDE, MPE(α, β), MPE(α), MPE(β), and Exact(α) (the most appropriate since it uses the exact log-likelihood function when only α is random). Appropriate methods are underlined. Notice that the LL_X values of approximate methods may differ from the true values of the log-likelihood function at the parameter estimated values.

	True	**DA(α, β)**	**DA(α)**	**DA(β)**
$h^{-1}(\mu)$	632.70	637.79 ± 3.01	637.79 ± 3.01	648.73 ± 2.70
θ	0.15	0.1180 ± 0.0027	0.1180 ± 0.0027	–
λ	1.43	1.4158 ± 0.0060	1.4158 ± 0.0060	1.3838 ± 0.0069
ω	0	<0.0001	–	0.1190 ± 0.0059
σ	0.33	0.3328 ± 0.0009	0.3328 ± 0.0009	0.3343 ± 0.0010
LL_X		$-1{,}191{,}370$	$-1{,}191{,}370$	$-1{,}192{,}336$
	NMSDE	**MPE(α, β)**	**MPE(α)**	**MPE(β)**
$h^{-1}(\mu)$	644.13 ± 2.66	632.81	633.34	630.96
θ	–	0.1439	0.1532	–
λ	1.3906 ± 0.0060	1.3517	1.3515	1.3494
ω	–	0.0124	–	0.0042
σ	0.3359 ± 0.0010	0.3074	0.3085	0.3075
LL_X	$-1{,}192{,}590$	$-1{,}207{,}378$	$-1{,}207{,}350$	$-1{,}202{,}663$
			Exact(α)	
$h^{-1}(\mu)$			633.53 ± 3.66	
θ			0.1523 ± 0.0047	
λ			1.4332 ± 0.0061	
ω			–	
σ			0.3307 ± 0.0009	
LL_X			$-1{,}190{,}931$	

Since the LL_X values for the approximate methods are only approximations, we cannot perform a likelihood ratio test for the randomness of the β parameter (null hypothesis $\omega = 0$ corresponding to a non-random β). However, if one knows (or assumes) that β is fixed, one can perform a likelihood-ratio test on whether α is fixed or random (null hypothesis $\theta = 0$ corresponding to a fixed α) by using the log-likelihood LL_X values of the NMSDE method and the Exact(α) method, which are exact values; the result is the rejection of the null hypothesis at the usual significance levels (p-value < 0.001).

For the case of the $D_{5000}(\beta)$ dataset, Table 3 shows that, again, the appropriate (for this dataset) DA(α, β) (which is overparametrized) and DA(β) methods provide very similar parameter estimates. Curiously, the inappropriate NMSDE model performs surprisingly well. In this case, the MPE methods, both appropriate and inappropriate ones, present very poor estimates. We also highlight that the estimates of the standard deviations $\theta = 0$

and $\omega = 0.30$ using the appropriate MPE methods severely fail, suggesting even that α is random and β is not when the reverse is the true situation, while the DA(α, β) method very well captures the non-random nature of parameter α, and both DA(α, β) and DA(β) methods capture the random nature of β, although somewhat underestimating its standard deviation ω.

Table 3. Results for the simulated dataset of 5000 animals when β is random, but α is fixed ($\theta = 0$) – $D_{5000}(\beta)$. Besides the true parameter values used in the simulations, the table shows their maximum likelihood estimates and, when available, approximate 95% confidence bands, for the different estimation methods: DA(α, β), DA(α), DA(β), NMSDE, MPE(α, β), MPE(α), and MPE(β). Appropriate methods are underlined. Notice that the LL_X values of approximate methods may differ from the true values of the log-likelihood function at the parameter estimated values.

	True	**DA(α, β)**	**DA(α)**	**DA(β)**
$h^{-1}(\mu)$	632.70	630.64 ± 2.60	611.04 ± 3.3176	630.70 ± 2.60
θ	0	<0.0001	0.1105 ± 0.0035	$-$
λ	1.43	1.3559 ± 0.0068	1.3852 ± 0.0061	1.3482 ± 0.0068
ω	0.30	0.1977 ± 0.0033	$-$	0.1972 ± 0.0031
σ	0.33	0.3362 ± 0.0010	0.3379 ± 0.0010	0.3360 ± 0.0009
LL_X		$-1,186,396$	$-1,187,992$	$-1,186,334$
	NMSDE	**MPE(α, β)**	**MPE(α)**	**MPE(β)**
$h^{-1}(\mu)$	630.12 ± 2.7063	620.22	619.63	620.66
θ	$-$	0.1413	0.1458	$-$
λ	1.3624 ± 0.0061	1.3229	1.3252	1.3155
ω	$-$	0.0001	$-$	0.0275
σ	0.3403 ± 0.0010	0.3271	0.3138	0.3153
LL_X	$-1,188,686$	$-1,203,829$	$-1,203,643$	$-1,203,829$

In Table 4, the estimates using the D_{5000} dataset, with both parameters α and β fixed, i.e., with $\theta = 0$ and $\omega = 0$, are presented. Now, all the methods are appropriate. However, with the exception of the NMSDE method, which, in this case, is the most appropriate one, they are overparametrized. It is quite interesting to note that the different DA methods give practically coincidental parameter estimates among them and are practically coincidental with the parameter estimates of the NMSDE model. This reveals robustness of the proposed DA method, which also captures the non-random character of α and β very well. The MPE methods give reasonable results, but they are not so good.

In this case, not only the NMSDE method and the DA methods give almost coincidental parameter estimates, but the LL_X value of the the NMSDE method (which is the exact maximum of the exact log-likelihood function) is also almost coincidental with the LL_X values for the DA methods; so, these LL_X DA values, which are approximations, are likely to be good approximations. If these values were exact instead of just approximations, a likelihood ratio test for $\theta = 0$ (i.e, for the randomness of α) and/or the significance of $\omega = 0$ (i.e., for the randomness of β) would obtain non-significant results.

Table 4. Results for the simulated dataset of 5000 animals when α and β are both fixed ($\theta = \omega = 0$) – D_{5000}. Besides the true parameter values used in the simulations, the table shows their maximum likelihood estimates and, when available, approximate 95% confidence bands, for the different estimation methods: DA(α, β), DA(α), DA(β), NMSDE, MPE(α, β), MPE(α), and MPE(β). Appropriate methods are underlined. Notice that the LL_X values of approximate methods may differ from the true values of the log-likelihood function at the parameter estimated values.

	True	**DA(α, β)**	**DA(α)**	**DA(β)**
$h^{-1}(\mu)$	632.70	631.17 ± 2.49	631.16 ± 2.49	631.19 ± 2.47
θ	0	0.0067 ± 0.0407	0.0070 ± 0.0375	$-$
λ	1.43	1.4310 ± 0.0061	1.4310 ± 0.0061	1.4309 ± 0.0060
ω	0	0.0047 ± 0.1298	$-$	0.0074 ± 0.0796
σ	0.33	0.3295 ± 0.0009	0.3295 ± 0.0009	0.3295 ± 0.0009
LL_X		$-1,186,643$	$-1,186,643$	$-1,186,643$
	NMSDE	**MPE(α, β)**	**MPE(α)**	**MPE(β)**
$h^{-1}(\mu)$	631.19 ± 2.4747	630.31	631.07	630.96
θ	$-$	0.0262	0.0186	$-$
λ	1.4309 ± 0.0060	1.3518	1.3493	1.3494
ω	$-$	0.0071	$-$	0.0042
σ	0.3295 ± 0.0009	0.2920	0.3075	0.3075
LL_X	$-1,186,643$	$-1,202,431$	$-1,201,744$	$-1,201,744$

Since the typical situation when dealing with real data is not knowing beforehand whether α and/or β are random or fixed, and, since we have seen that, among the approximate methods DA, the method DA(α, β) performs as well or almost as well as other appropriate DA methods (in what concerns the estimation of the common parameters), even when it is overparametrized with respect to the dataset, we will consider from now on, among the three delta approximation methods, only the DA(α, β) method. For similar reasons, among the MPE methods, we will consider from now on just the MPE(α, β) method. For comparison purposes, we also consider the usual NMSDE method, which is the most appropriate method for the fixed-effects model considered in most applications, but it is, of course, inappropriate when our datasets have α random and/or β random.

With the purpose of analyzing the influence of smaller sample sizes on the estimates, Tables 5–8 present the results for the four mixed- and fixed-effects stochastic Gompertz models for the datasets with smaller samples sizes, 50 and 500 animals. We will use the DA(α, β), the MPE(α, β), and the NMSDE estimation methods. For comparison purposes, for the datasets $D_{50}(\alpha)$ and $D_{500}(\alpha)$, we also present the results of the Exact(α) method.

We can conclude that the same main characteristic features observed in the estimates for the large datasets with 5000 animals still hold. The DA(α, β) method is able to accurately identify the fixed and the random effects, while the MPE(α, β) method usually fails (except, and not clearly, when both effects are fixed).

As expected, when comparing the confidence intervals of the parameter estimates for different number of animals in the datasets (see Tables 5–8 for 50 and 500 animals, and Tables 1–4 for 5000 animals), their amplitudes decrease as the number of animals in the dataset increases. In general, the MPE(α, β) method have worst estimates than the DA(α, β) method, both for small and large samples, except in estimating the standard deviation θ of the parameter α when this parameter is random.

For the datasets $D_{50}(\alpha)$ and $D_{500}(\alpha)$, the results obtained with the DA(α, β) method are not exactly the same as the ones obtained by the Exact(α) method, but they are quite close, even very close for the 500 animals dataset $D_{500}(\alpha)$.

For the fixed-effects datasets D_{50} and D_{500}, the estimates of the DA(α, β) method replicate the same results as the NMSDE method, which, for these datasets, is the most appropriate method. The LL_X values of the NMSDE method are exact maximum log-likelihood values, and they are again practically equal to the LL_X approximate values

of the DA(α, β) method, again suggesting that, for fixed-effects datasets, these are good approximations, in which case a likelihood ratio test of $\theta = \omega = 0$ would accept this null hypothesis.

Table 5. Results for the simulated datasets of 50 and 500 animals when both α and β are random–$D_{50}(\alpha, \beta)$ and $D_{500}(\alpha, \beta)$. Besides the true parameter values used in the simulations, the table shows their maximum likelihood estimates and, when available, approximate 95% confidence bands, for the appropriate estimation methods DA(α, β) and MPE(α, β) and the inappropriate estimation method NMSDE. Notice that the LL_X values of approximate methods may differ from the true values of the log-likelihood function at the parameter estimated values.

	True	<u>DA(α, β)</u>	<u>MPE(α, β)</u>	NMSDE
		50 animals		
$h^{-1}(\mu)$	632.70	648.87 ± 33.04	612.27	646.62 ± 32.22
θ	0.15	0.0392 ± 0.1214	0.2594	–
λ	1.43	1.2265 ± 0.0600	1.2703	1.2259 ± 0.0599
ω	0.30	0.2164 ± 0.0297	<0.0001	–
σ	0.33	0.3418 ± 0.0096	0.3059	0.3474 ± 0.0099
LL_X		$-11{,}819.3$	$-11{,}713.8$	$-11{,}868.8$
		500 animals		
$h^{-1}(\mu)$	632.70	651.52 ± 10.06	634.72	655.81 ± 9.52
θ	0.15	0.0944 ± 0.0130	0.2087	–
λ	1.43	1.3216 ± 0.0209	1.3061	1.3056 ± 0.0191
ω	0.30	0.1809 ± 0.0124	0.0002	–
σ	0.33	0.3407 ± 0.0030	0.3114	0.3458 ± 0.0031
LL_X		$-119{,}331.8$	$-120{,}800.7$	$-119{,}620.8$

Table 6. Results for the simulated dataset of 50 and 500 animals when α is random, but β is fixed ($\omega = 0$) – $D_{50}(\alpha)$ and $D_{500}(\alpha)$. Besides the true parameter values used in the simulations, the table shows their maximum likelihood estimates and, when available, approximate 95% confidence bands, for the appropriate estimation methods DA(α, β), MPE(α, β) and Exact(α) and the inappropriate estimation method NMSDE. Notice that the LL_X values of approximate methods may differ from the true values of the log-likelihood function at the parameter estimated values.

	True	<u>DA(α, β)</u>	<u>MPE(α, β)</u>	NMSDE	<u>Exact(α)</u>
		50 animals			
$h^{-1}(\mu)$	632.70	641.87 ± 30.74	639.16	648.94 ± 26.51	640.20 ± 36.01
θ	0.15	0.1152 ± 0.0284	0.1446	–	0.1437 ± 0.0464
λ	1.43	1.4173 ± 0.0627	1.3501	1.3939 ± 0.0596	1.4385 ± 0.0608
ω	0	0.0637 ± 0.1213	<0.0001	–	–
σ	0.33	0.3284 ± 0.0093	0.3051	0.3318 ± 0.0094	0.3268 ± 0.0095
LL_X		$-11{,}892.8$	$-11{,}767.1$	$-11{,}905.6$	$-11{,}659.9$
		500 animals			
$h^{-1}(\mu)$	632.70	635.46 ± 9.56	635.62	636.59 ± 8.35	644.44 ± 37.87
θ	0.15	0.1197 ± 0.0085	0.1537	–	0.1493 ± 0.0484
λ	1.43	1.4037 ± 0.0188	1.3417	$1.3781 \pm 0.0.0189$	1.3957 ± 0.0610
ω	0	<0.0001	0.0064	–	–
σ	0.33	0.3312 ± 0.0030	0.3067	0.3343 ± 0.0030	0.3308 ± 0.0096
LL_X		$-118{,}752.5$	$-120{,}099.7$	$-118{,}878.5$	$-11{,}667.9$

Table 7. Results for the simulated datasets of 50 and 500 animals when β is random, but α is fixed ($\theta = 0$) – $D_{50}(\beta)$ and $D_{500}(\beta)$. Besides the true parameter values used in the simulations, the table shows their maximum likelihood estimates and, when available, approximate 95% confidence bands, for the appropriate estimation methods DA(α, β) and MPE(α, β) and the inappropriate estimation method NMSDE. Notice that the LL_X values of approximate methods may differ from the true values of the log-likelihood function at the parameter estimated values.

	True	DA(α, β)	MPE(α, β)	NMSDE
		50 animals		
$h^{-1}(\mu)$	632.70	619.32 ± 23.98	613.94	618.43 ± 24.13
θ	0	0.0260 ± 0.1038	0.1026	–
λ	1.43	1.4725 ± 0.0730	1.3944	1.4591 ± 0.0618
ω	0.30	0.2043 ± 0.0359	0.0002	–
σ	0.33	0.3323 ± 0.0094	0.3036	0.3365 ± 0.0096
LL_X		$-11{,}881.2$	$-11{,}765.4$	$-11{,}902.1$
		500 animals		
$h^{-1}(\mu)$	632.70	640.71 ± 8.45	626.32	639.20 ± 8.77
θ	0	<0.0001	0.1602	–
λ	1.43	1.3451 ± 0.0215	1.3213	1.3497 ± 0.0191
ω	0.30	0.1923 ± 0.01051	<0.0001	–
σ	0.33	0.3360 ± 0.0030	0.2947	0.3399 ± 0.0031
LL_X		$-118{,}814.8$	$-120{,}401.4$	$-119{,}038.4$

Table 8. Results for the simulated datasets of 50 and 500 animals when α and β are both fixed ($\theta = \omega = 0$) – D_{50} and D_{500}. Besides the true parameter values used in the simulations, the table shows their maximum likelihood estimates and, when available, approximate 95% confidence bands, for the appropriate estimation methods DA(α, β) and MPE(α, β) and the estimation method NMSDE, which is the most appropriate for this dataset. Notice that the LL_X values of approximate methods may differ from the true values of the log-likelihood function at the parameter estimated values.

	True	DA(α, β)	MPE(α, β)	NMSDE
		50 animals		
$h^{-1}(\mu)$	632.70	624.23 ± 24.67	624.32	624.23 ± 24.13
θ	0	<0.001	0.0013	–
λ	1.43	1.4465 ± 0.0616	1.3623	1.4465 ± 0.0616
ω	0.30	<0.001	<0.0001	–
σ	0.33	0.3362 ± 0.0096	0.3136	0.3362 ± 0.0096
LL_X		$-11{,}901.2$	$-11{,}762.8$	$-11{,}901.0$
		500 animals		
$h^{-1}(\mu)$	632.70	628.22 ± 7.81	628.15	628.21 ± 7.81
θ	0	<0.0001	0.0196	–
λ	1.43	1.4331 ± 0.0193	1.3511	1.4330 ± 0.0191
ω	0	0.0135 ± 0.1381	<0.0001	–
σ	0.33	0.3309 ± 0.0030	0.2996	0.3309 ± 0.0030
LL_X		$-118{,}686.3$	$-119{,}963.8$	$-118{,}686.3$

Finally, in Table 9, we used the real cattle weight dataset of 16,029 Mertolengo cattle males (totaling 96,204 observations) and present the estimates obtained by the DA(α, β) method and by the NMSDE method. The true parameter values are not known, and we cannot use the R package *MsdeParEst* since the real weight data from the 16,029 animals presents a different age vector of observations for each animal. Analyzing the results of Table 9, we can conclude that, despite the database being large and heterogeneous, the two estimation methods provide very similar estimates. According to our findings obtained above, it is interesting to note that the DA(α, β) method presents clear evidence of random effects on the α parameter and of strong random effects on the β parameter. Although we

cannot perform a likelihood ratio test to reject the null hypothesis $\theta = \omega = 0$ because the displayed LL_X value for the DA(α, β) method is an approximation, the difference of LL_X values between this method and the NMSDE method (which displays an exact LL_X value) is striking enough to leave doubts about the random nature of α and β.

Table 9. Results for the real cattle weight dataset. The maximum likelihood estimates and approximate 95% confidence bands are shown when assuming a mixed model with both α and β random and using the DA(α, β) estimation method and when assuming a non-mixed model (α and β fixed) and using the NMSDE method (which is the classical exact maximum likelihood method under that assumption). Notice that the LL_X values of the approximate method DA(α, β) method may differ from the true value of the log-likelihood function at the parameter estimated values.

	DA(α, β)	**NMSDE**
$h^{-1}(\mu)$	629.70 ± 6.30	631.27 ± 6.38
θ	0.0909 ± 0.0083	−
λ	1.4261 ± 0.0122	1.4198 ± 0.0116
ω	0.1998 ± 0.0050	−
σ	0.3273 ± 0.0016	0.3398 ± 0.0017
LL_X	$-368{,}732.10$	$-369{,}753.20$

5. Recommendations

Taking into account all the analyzed scenarios, it makes sense to present a summary of the advantages and disadvantages of each method and recommendations for its use with a real dataset depending on the number of animals. If one has no information or strong reasons to consider a given parameter as random, one should apply the DA(α, β) method, which, for all datasets considered here, was able to correctly identify which parameters are fixed and which are random.

If one concludes from this preliminary analysis that none of the parameters are fixed (by not having small estimates of their standard deviations θ and ω, complemented by the information given by the approximate confidence intervals of these parameters), the DA(α, β) method is recommended to estimate the mean and standard deviation of the random β, whereas the MPE(α, β) method is recommended to estimate the mean and standard deviation of the random α, either for small and large samples. However, a good compromise will be achieved if one just uses the DA(α, β) method.

In the case the preliminary analysis indicates that $\omega = 0$ (by having a very small estimate of this parameter), we should then apply the Exact(α) method [1], which is designed for the case of α being the only random parameter and uses the exact log-likelihood function (8), to get better estimates of all parameters.

If the preliminary analysis indicates $\theta = 0$ (by having a very small estimate of this parameter), we better remain with the DA(α, β) method (since its estimates are very similar with those obtained under the DA(β) method) and provides better estimates of the mean and standard deviation for the β random effect and also of the fixed effect α than the MPE method or the NMSDE method.

Finally, in the case where the preliminary analysis indicates $\omega = 0$ and $\theta = 0$ (i.e., the dataset only has fixed parameters), we can apply either the most appropriate and classical NMSDE method [1], designed for the fixed-effects model and using the exact log-likelihood function (6), or the DA(α, β) method (which gives basically the same estimates).

Furthermore, the DA method has the big advantage of being usable for all real datasets, with or without equidistant observations and whether or not the animal's sizes have been obtained at the same age instants for all animals.

6. Conclusions

To describe the individual growth of animals in a randomly varying environment, a general stochastic differential equation (SDE) model was used, and, in order to take into account that the model parameters may vary from animal to animal, which, for instance,

occurs due to different individual genetics and other characteristics of the animals, we considered SDE mixed models. Here, we studied the SDE mixed model where both parameters included in the drift term, the parameter α (asymptotic modified weight) and the growth parameter β, were assumed to be Gaussian distributed.

We applied the maximum likelihood estimation method to obtain the estimates for the parameters of the SDE mixed models. In most cases, for this type of model, it is difficult or impossible to obtain a closed-form expression for the likelihood function and approximation methods are used to cope with this issue. In this paper, we propose to use the delta approximation method (DA method) with both α and β random, in which we adapted the delta method, a classic in Statistics, to obtain approximate closed-form expressions for the likelihood function when both parameters α and β are random. Of course, the DA method can also be applied when only one of the two parameters is assumed random.

To evaluate the performance of the proposed method on parameter estimation, we used simulated datasets from different mixed-effects models (with only α random, with only β random, and with both parameters random) and from a fixed-effects model (both α and β fixed), with different numbers of animals (5000, 500, and 50), datasets in which all animals were weighed at the same ages so that we could use all the estimation methods considered here and compare them. Since, in real life, we usually do not know which are the random parameters, wrong assumptions may be made on that issue. In order to evaluate the consequences of wrong assumptions on what parameters are random, we included in the comparisons also the methods designed under such assumptions. We compared the DA method (in its variants of assuming both or just one parameter random) with an existing method specifically designed for mixed-effects models (also with the same variants), referred to here as the MPE method, which is provided by the R package *MsdeParEst*. We also included in the comparisons the estimation methods for which log-likelihood expressions in closed-form are available, namely the NMSDE method and, in some cases, the Exact(α) method, which are designed, respectively, for the fixed-effects model and the mixed-effects model with only α random.

The results of these comparisons show a very good performance of the proposed DA method with both α and β random, being globally the best method for all the simulated scenarios. This method, unlike the MPE methods, was able to correctly identify, in each of the settings, the fixed and the random parameters. It gives generally better parameter estimates than the MPE method and the estimates are quite close to the true parameter values, with the exception of the standard deviations of random effects, which were somewhat underestimated. The performance of the proposed DA method was confirmed when using a simulated dataset with both parameters fixed, since it provided the same results as the ones obtained when using the exact NMSDE method.

For this type of SDE mixed models, it is usual to find in the literature estimation methods developed under the assumption of having a unique (often also evenly spaced) age vector of observations common to all individuals, and this is also required when using available R packages, such as the *MsdeParEst* package used in the MPE method. The delta approximation (DA) method has the advantage of not requiring such restrictions, so it can be used in real situations where such restrictions usually do not hold. In our real data application, we are precisely in this situation, and we applied the estimation methods to real weights of Mertolengo cattle males from a large and heterogeneous dataset, reaching the conclusion that the proposed DA method identifies both parameters α and (more strikingly) β as being random.

This approach revealed to be a very interesting alternative to the available estimation methods for SDE mixed models.

As future work, we are undergoing the study of the case where we will incorporate the genetic factors of the animals into the model to explain part of the variation in the random parameters, and we intend to implement the current methods in an R package.

References

1. Filipe, P.A. Equações Diferenciais Estocásticas na Modelação do Crescimento Individual em Ambiente Aleatório. Ph.D. Thesis, Universidade de Évora, Evora, Portugal, 2011.
2. Jacinto, G.; Filipe, P.A.; Braumann, C.A. Profit optimization of cattle growth with variable prices. *Methodol. Comput. Appl. Probab.* **2021**. [CrossRef]
3. Jacinto, G.; Filipe, P.A.; Braumann, C.A. Weighted maximum likelihood estimation for individual growth models. *Optimization under review*.
4. Garcia, O. A stochastic differential equation model for the height of forest stands. *Biometrics* **1983**, *39*, 1059–1072. [CrossRef]
5. Delattre, M.; Genon-Catalot, V.; Larédo, C. Parametric inference for discrete observations of diffusion processes with mixed effects. *Stoch. Process. Their Appl.* **2016**, *128*, 1929–1957. [CrossRef]
6. Delattre, M.; Genon-Catalot, V.; Samson, A. Maximum likelihood estimation for stochastic differential equations with random effects. *Scand. J. Stat.* **2013**, *40*, 322–343. [CrossRef]
7. Overgaard, R.V.; Jonsson, N.; Tornøe, C.W.; Madsen, H. Non-Linear Mixed-Effects Models with stochastic differential equations: Implementation of an estimation algorithm. *J. Pharmacokinet. Pharmacodyn.* **2005**, *32*, 85–107. [CrossRef] [PubMed]
8. Picchini, U. Stochastic Differential Models with Applications to Physiology. Ph.D. Thesis, University of Rome, Rome, Italy, 2006.
9. Picchini, U.; De Gaetano, A.; Ditlevsen, S. Stochastic differential mixed-effects models. *Scand. J. Stat.* **2010**, *37*, 67–90. [CrossRef]
10. Picchini, U.; Ditlevsen, S. Practical estimation of high dimensional stochastic differential mixed-effects models. *Comput. Stat. Data Anal.* **2011**, *55*, 1426–1444. [CrossRef]
11. Delattre, M. A review on asymptotic inference in stochastic differential equations with mixed effects. *Jpn. J. Stat. Data Sci.* **2021**, *4*, 543–575. [CrossRef]
12. Jamba, N.T.; Filipe, P.A.; Jacinto, G.; Braumann, C.A. Estimation for SDE mixed models using approximation methods. *Comput. Stat. under review*.
13. Delattre, M.; Dion, C. MsdeParEst: Parametric Estimation in Mixed-Effects Stochastic Differential Equations. R Package Version 1.7. 2017. Available online: https://cran.r-project.org/web/packages/MsdeParEst/index.html (accessed on 25 January 2022).
14. Vasicek, O. An equilibrium characterization of the term structure. *J. Financ. Econ.* **1977**, *5*, 177–188. [CrossRef]
15. Braumann, C.A. *Introduction to Stochastic Differential Equations with Applications to Modelling in Biology and Finance*; John Wiley & Sons: Hoboken, NJ, USA, 2019.
16. Filipe, P.A.; Braumann, C.A.; Brites, N.M.; Roquete, C.J. Modelling animal growth in random environments: An application using nonparametric estimation. *Biom. J.* **2010**, *52*, 653–666. [CrossRef] [PubMed]
17. Picchini, U.; Ditlevsen, S.; De Gaetano, A. Maximum likelihood estimation of a time—Inhomogeneous stochastic differential model of glucose dynamics. *Math. Med. Biol. J. IMA* **2008**, *25*, 141–155. [CrossRef] [PubMed]

18. R Core Development Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.

19. Klim, S.; Mortensen, S.B.; Kristensen, N.R.; Overgaard, R.V.; Madsen, H. Population Stochastic Modelling (PSM)—An R Package for mixed-effects models based on stochastic differential equations. *Comput. Methods Programs Biomed.* **2009**, *94*, 279–289. [CrossRef] [PubMed]

20. Dion, C.; Sansom, A.; Hermann, S. mixedsde: Estimation Methods for Stochastic Differential Mixed Effects Models. R Package Version 5.0. 2018. Available online: https://cran.r-project.org/web/packages/mixedsde/index.html (accessed on 25 January 2022).