AIMS Mathematics, 9(3): xxx–xxx.
DOI: 10.3934/math.2024xxx
Received: 28 November 2023
Revised: 17 February 2024
Accepted: 18 February 2024
Published: XX February 2024

http://www.aimspress.com/journal/Math

Research article

Estimation for stochastic differential equation mixed models using approximation methods

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, Évora, Portugal

[2] Liceu nº 918 do município dos Gambos, Gambos, Angola

[3] Departamento de Matemática, Escola de Ciência e Tecnologia, Universidade de Évora, Évora, Portugal

[4] Departamento de Métodos Quantitativos para Gestão e Economia, ISCTE Business School, Iscte-Instituto Universitário de Lisboa, Lisboa, Portugal

* **Correspondence:** Email: patricia.filipe@iscte-iul.pt; Tel: +351217903000.

Abstract: We used a class of stochastic differential equations (SDE) to model the evolution of cattle weight that, by an appropriate transformation of the weight, resulted in a variant of the Ornstein-Uhlenbeck model. In previous works, we have dealt with estimation, prediction, and optimization issues for this class of models. However, to incorporate individual characteristics of the animals, the average transformed size at maturity parameter α and/or the growth parameter β may vary randomly from animal to animal, which results in SDE mixed models. Obtaining a closed-form expression for the likelihood function to apply the maximum likelihood estimation method is a difficult, sometimes impossible, task. We compared the known Laplace approximation method with the delta method to approximate the integrals involved in the likelihood function. These approaches were adapted to allow the estimation of the parameters even when the requirement of most existing methods, namely having the same age vector of observations for all trajectories, fails, as it did in our real data example. Simulation studies were also performed to assess the performance of these approximation methods. The results show that the approximation methods under study are a very good alternative for the estimation of SDE mixed models.

Keywords: delta method; Laplace method; maximum likelihood estimation; mixed models; stochastic differential equations
Mathematics Subject Classification: 60H10, 60J70, 62F10, 92D99

1. Introduction

The growth of individual organisms must take into account random environmental fluctuations that affect the growth rate and so we use stochastic differential equation (SDE) models. To describe the growth of an individual in a randomly fluctuating environment, a class of SDE models has been studied, for instance in [1–3]. The models of this class can be written, by using an appropriate transformation of the size, in the form of a mean-reverting *Ornstein-Uhlenbeck model*, also called the *Vasicek model* in financial mathematics [4], as

$$dY_i(t) = \beta\,(\alpha - Y_i(t))\;dt + \sigma\,dW_i(t),\; Y_i(t_{i,0}) = y_{i,0}, \quad i = 1, \ldots, M, \tag{1.1}$$

where $Y_i(t) = h(X_i(t))$ is the size modified by the transformation h, a monotonous C^1 function (which we assume to be known) of the real size $X_i(t)$ at age t of the i^{th} individual ($i = 1, \ldots, M$). We have $Y_i(t_{i,0}) = y_{i,0} = h(x_{i,0})$, where $x_{i,0}$ is the size observed at age $t_{i,0}$ (known initial observation) for individual i, and $\alpha = h(A)$, where A is the asymptotic size or size at maturity. The growth parameter β is the rate of approach to maturity, σ measures the strength of environmental fluctuations on growth, and $W_i(t)$ ($i = 1, \ldots, M$) are independent standard Wiener processes. Specific choices of the function h lead to stochastic versions of some models commonly used to describe growth (e.g., monomolecular, Bertalanffy-Richards, Gompertz, and logistic). The estimation, prediction, and optimization issues for this class of models with applications to cattle weight data can be found in [1, 5–8]. For this type of data, one of the most adequate choices of the transformation function h was the logarithm of the weight, which corresponds to the stochastic Gompertz model. For this reason, when illustrating the results obtained in this paper, this particular model was used in the application.

Since model parameters may differ among animals, we extend the study to mixed SDE models where the parameters α and/or β vary randomly among animals. The maximum likelihood estimation method was applied and we studied the delta approximation method, inspired by the classical delta method, to approximate the integral involved in the likelihood function, and compared it with the available R package for SDE mixed models *mixedsde*. Another already-known approximation method is the Laplace approximation method. Despite some papers having already proposed the Laplace approximation method ([9, 10]), none of them use this approximation method in the cases where the number of observations and/or the ages of such observations vary from animal to animal. The motivation of the paper is to study approximation methods that can be applied to the more realistic SDE mixed models in the context of real cattle growth data, where different animals have different age vectors and different numbers of weighings. In particular, we further develop and assess the performance of an approximation method (the delta method) that is easy to implement and a reliable and fast alternative to other existing approximation methods. In [3] the delta approximation method was proposed for the particular case where both parameters α and β are independent Gaussian distributed random variables.

Our main objective is to compare the delta and Laplace approximation methods. In addition to this methodological comparison, we have two other objectives. The first is to assess the robustness of these approximation methods for estimating model parameters when there is no exact method available as an alternative. These methods can be particularly useful when the integral appearing in the likelihood function cannot be explicitly computed, and no closed-form expression of this function can be obtained, as is the case in the random β scenario and when both parameters α and β are considered

random. The second objective is to use the approximation methods in the particular case where the likelihood function can be explicitly obtained, and therefore an exact method is available (α random), as a benchmark to assess the robustness of the estimates obtained by the approximation methods.

To illustrate the results, we have used the weight in kilograms g of the 10843 Mertolengo cattle males with recorded genetic information on the database of the Associação de Criadores de Bovinos Mertolengos (ACBM), which reports on of the growing and finishing phases of the young Mertolengo males, and its associated breeders in the Alentejo region in Portugal. Each animal has multiple weight records, ranging from 3 to 33, at various ages from birth to a maximum age between 0.2 and 8.1 years, totaling 69782 observations.

A survey of several parametric and nonparametric methods for asymptotic inference in SDE mixed models can be found in [11] with the discussion of asymptotic properties of estimators and presentation of several applications. For instance, in [12] and [13] the authors apply SDE mixed effects models to tumor growth dynamics using fully Bayesian inference through the pseudo-marginal methods. In [14] an application to electroencephalography data is used by discretizing the continuous time likelihood. The Laplace approximation approach and other approximation methodologies have been addressed, for instance, in [15], [9], and [16]. Some methodologies were already implemented in R packages (see, for example, [17], [18], and [19]) and here we will compare the results obtained by the delta and the Laplace approximation methods with those obtained using the R package *mixedsde*. The *mixedsde* R package allows inference on the Ornstein-Uhlenbeck and Cox-Ingersoll-Ross SDE mixed models for random effects on the trend of additive, multiplicative, or jointly additive-multiplicative types. It allows one to obtain the maximum likelihood estimates of the model parameters through approximation methods using a parametric approach, a parametric Bayesian approach, and a non-parametric Bayesian approach. In its parametric approach, it uses the continuous-time likelihood function obtained via the Girsanov formula and approximates the likelihood with a discretized version [18, 20].

The available packages require that the time vector of observations is the same for all trajectories. This is not the case in our application, since in our real data the animals are not weighted at the same ages and, for this reason, the existing packages could not be used. To be able to compare the results of both delta and Laplace approximation methods considered here with the results obtained by the *mixedsde* R package, we have used a dataset of simulated weights at the same equidistant ages. A bias analysis of the parameter estimates and an evaluation of the adequacy of the confidence intervals based on the empirical Fisher information matrix was also applied to both the delta and the Laplace approximation methods.

In summary, in this paper, we present useful contributions not available elsewhere.

- Closed-form expressions for the delta approximation method: This paper offers all closed-form expressions of the delta approximation method, particularly when the parameters α or β are considered random.
- Detailed formulas for the Laplace approximation method: This paper introduces intricate formulas for the Laplace approximation method, accommodating data with varying numbers of observations per subject and non-equidistant observations. These formulas enable a straightforward implementation.
- Comprehensive comparison among methods: A thorough comparison is conducted between the two approximation methods (delta and Laplace) and the *mixedsde* R package. This comprehensive analysis provides a clear understanding of their performance under different scenarios involving

random parameter considerations.

This paper is organized as follows: In Section 2, we present the SDE models, their main properties, and the use of the maximum likelihood estimation method to estimate the model's parameters. In Section 3, we present the SDE mixed model. In Sections 4 and 5, the particular SDE mixed models for the random α case and for the random β case are presented, and the methodological developments and expressions for the delta approximation method and the Laplace approximation method are obtained. In Section 6, the same approach is presented for the case where both α and β are random. In Section 7, we present the application to our real cattle weight data and to simulated data, comparing the performance of the different methods. We end with Section 8, where the main conclusions are presented.

2. Stochastic differential equations for non-mixed models

We consider data from M individuals. Let $X_i(t)$ be the size at time t of the i^{th} individual ($i = 1, \ldots, M$). The corresponding modified size is $Y_i(t) = h(X_i(t))$, where the transformation h is a monotonous C^1 function. If the individual is growing in a randomly fluctuating environment, we can model growth through (1.1).

It can be seen, for instance in [21], that the solution of (1.1) is given, for $t \geq t_{i,0}$, by

$$Y_i(t) = \alpha - (\alpha - y_{i,0})e^{-\beta(t-t_{i,0})} + \sigma e^{-\beta t} \int_{t_{i,0}}^{t} e^{\beta s} dW_i(s), \quad i = 1, \ldots, M, \tag{2.1}$$

and follows a Gaussian distribution with mean $\alpha - (\alpha - y_{i,0})e^{-\beta(t-t_{i,0})}$ and variance $\frac{\sigma^2}{2\beta}(1 - e^{-2\beta(t-t_{i,0})})$. Note that $Y_i(t)$ is a homogeneous diffusion process with drift coefficient $a(y) = \beta(\alpha - y)$ and diffusion coefficient $b(y) = \sigma^2$.

Let $t_{i,j}$ ($i = 1, ..., M$, $j = 1, ..., n_i$) be the age of the j^{th} 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 (1.1). For each individual i ($i = 1, \ldots, M$), denote by $\mathbf{t_i} = (t_{i,0}, t_{i,1}, \ldots, t_{i,n_i})$ its time vector of observations (which may differ from individual to individual), by $\mathbf{Y_i} = (Y_{i,0}, Y_{i,1}, ..., Y_{i,n_i})$ its random vector of observables, and by $\mathbf{y_i} = (y_{i,0}, y_{i,1}, ..., y_{i,n_i})$ the observed value of $\mathbf{Y_i}$. 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 $Y_{i,j} = \alpha + E_{i,j}^{\beta}(Y_{i,j-1} - \alpha) + \sigma e^{-\beta t_{i,j}} \int_{t_{i,j-1}}^{t_{i,j}} e^{\beta s} dW_s$, and so, conditioned on $Y_{i,j-1} = y_{i,j-1}$, the transition distribution for animal i is Gaussian:

$$Y_{i,j}|Y_{i,j-1} = y_{i,j-1} \sim N\left(\alpha + \left(y_{i,j-1} - \alpha\right)E_{i,j}^{\beta}\,,\ \frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})\right). \tag{2.2}$$

To estimate the parameter vector $\mathbf{p} = (\alpha, \beta, \sigma)$, the maximum likelihood estimation method is usually applied, see, for instance, [5]. From (2.2), using the fact that $Y_i(t)$ is a Markov process, we know that, given $Y_{i,0} = y_{i,0}$ (assumed known), the joint probability density function (p.d.f.) for animal i is

$$p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha, \beta, \sigma) = \prod_{j=1}^{n_i} \frac{\exp\left(-\frac{1}{2}\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}\right)}{\sqrt{2\pi\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}} \tag{2.3}$$

$$= \frac{\exp\left(-\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}\right)}{\prod_{j=1}^{n_i}\sqrt{2\pi\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}}, \quad i=1,\ldots,M,$$

and, by independence among individuals, we obtain the likelihood function for the M animals

$$L(\alpha,\beta,\sigma) = \prod_{i=1}^{M} p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha,\beta,\sigma). \tag{2.4}$$

The maximum likelihood estimate $\hat{\mathbf{p}}$ of the parameter vector $\mathbf{p}$ is obtained by maximization of (2.4). The maximum likelihood estimators are asymptotically Gaussian with mean $\mathbf{p} = (\alpha,\beta,\sigma)$ and variance-covariance matrix $\mathbf{V}$. We can estimate $\mathbf{V}$ by the inverse of the empirical Fisher information matrix.
We have described a quite general SDE animal growth model and the maximum likelihood estimation method to estimate the model's parameters α, β, and σ, that are assumed to be the same for all individuals. However, in cattle weight data, it is natural to think that different animals may differ in some parameter value. So, we will consider this situation using mixed SDE models.

3. Stochastic differential equation mixed models

SDE mixed models are used to model growth, with applications in various fields, such as animal growth and pharmacokinetics, among others [3, 14, 15, 22, 23]. For the type of models we present here, in [3,9,16,24], it can be seen that, when one or both drift parameters α and β may vary randomly among animals, the likelihood function can be obtained from the transition density function conditioned on the respective random parameter(s). Assume that $\mathbf{b}$ is a d-dimensional vector of parameters that vary randomly among animals and let $\mathbf{b}_i$ be its random effect on animal i. Let ϵ be the vector of the remaining model parameters, assumed to be common to all animals. The distribution of $\mathbf{b}$ among animals has p.d.f. $p(\mathbf{b}|\mathbf{\Psi})$, where $\mathbf{\Psi}$ is the parameter vector that characterizes that 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. Assume also that the $\mathbf{b}_i$ ($i = 1,\ldots,M$) are independent of the Wiener processes that characterize the environmental conditions under which the animals are growing. The likelihood function for M trajectories (animals) is given by

$$L(\epsilon,\mathbf{\Psi}) = \prod_{\mathbf{i=1}}^{\mathbf{M}} p_{\mathbf{Y_i}}(\mathbf{y}_i|\epsilon,\mathbf{\Psi}) = \prod_{\mathbf{i=1}}^{\mathbf{M}} \int_{-\infty}^{+\infty} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{b}_i,\epsilon)p(\mathbf{b}_i|\mathbf{\Psi})\,d\mathbf{b}_i, \tag{3.1}$$

where $p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{b}_i,\epsilon)$ is the joint p.d.f. for animal i conditional on $\mathbf{b}_i$. The case of a single parameter $\mathbf{b}_i = (\alpha_i)$ has already been studied for $\alpha_i \sim N(\mu,\theta^2)$ and allows a closed-form computation of the integral in the likelihood function, resulting in a final closed-form (heavy) expression for this function. One can see it, for example, in [25] for the particular situation of $\mu = 0$ and 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 and having uniform time spacing ($t_j - t_{j-1} \equiv$

δ , $j = 1, \ldots, n$), which considerably simplifies the expressions, and in [3] and [15] for the general situation.

However, the integral in (3.1) does not always have a closed-form solution, like when the random parameter is β, corresponding to $\mathbf{b}_i = (\beta_i)$, and when both parameters are considered random, corresponding to $\mathbf{b}_i = (\alpha_i, \beta_i)$. For the case when both parameters are considered random, [3] proposed the delta approximation method (DA method) to approximate the integral in (3.1). This approximation method allows us to obtain simpler expressions for the likelihood function, usually in closed-form. Although inspired by the classical statistical delta method, it differs from it in purpose and in the methodology used. In the following subsections, we are going to detail the same DA method for the cases not shown in [3], namely for the cases where just one parameter, α or β, is random. This method can be useful if we are interested in considering just the case of a single random parameter, for instance, when it is the most reasonable assumption for a certain real application. So, in this paper, we will also compare the DA method with other methods for the specific cases where only one of the parameters, α or β, is considered random.

This method is also suitable to use in the situation of a general age vector of observations, allowing us to estimate the parameters of the models even when the different animals have their weights observed at different ages and those ages may be non-equidistant.

We are also going to develop the numerical implementation of the Laplace approximation method (LA method) since, despite having already been proposed in [10], we could not find available packages enabling us to test the performance of the Laplace approximation method and compare it with the performance of the DA method. On the numerical implementation of the LA method, we have generalized the known existing references and applications, so that the method can be used to estimate the parameters when the number and ages of observation fail to be equidistant and/or common to all observed animals. Finally, the DA method and the LA method will be compared with the existing implemented methodologies that require an age vector common to all animals.

4. Mixed models with random asymptotic average size

If we assume, for each animal, a random asymptotic size α_i, with Gaussian distribution, mean μ, and variance θ^2, the joint p.d.f. for animal i is given by

$$p_{\mathbf{Y_i}}(\mathbf{y}_i|\mu,\theta,\beta,\sigma) = \int_{-\infty}^{+\infty} p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha_i,\beta,\sigma)p(\alpha_i|\mu,\theta)\,d\alpha_i, \tag{4.1}$$

with $p_{Y_i}(\mathbf{y}_i|\alpha_i,\beta,\sigma)$ obtained by replacing α with α_i in (2.3) and

$$p(\alpha_i|\mu,\theta) = \frac{1}{\sqrt{2\pi\theta^2}} e^{-\frac{(\alpha_i-\mu)^2}{2\theta^2}}. \tag{4.2}$$

Thus, from (3.1) and using (4.1) and (4.2), we get, after some simplifications, the likelihood function

$$L(\mu,\theta,\beta,\sigma) = \prod_{i=1}^{M}\left(\frac{1}{\sqrt{2\pi\theta^2}}\prod_{j=1}^{n_i}\left(\frac{1}{\sqrt{\frac{\pi\sigma^2}{\beta}(1-E_{i,j}^{2\beta})}}\right)\right. \tag{4.3}$$

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}-\frac{(\alpha_i-\mu)^2}{2\theta^2}\right)d\alpha_i\Bigg).$$

It is shown in [3] that it is possible to explicitly compute the integral in (4.3) and to obtain the log-likelihood function for all animals $LL(\mu,\theta,\beta,\sigma)=\ln L(\mu,\theta,\beta,\sigma)$ as the closed-form expression:

$$\begin{aligned} LL(\mu,\theta,\beta,\sigma) = &\sum_{i=1}^{M}\Bigg(-\frac{n_i}{2}\ln(2\pi)-\frac{n_i}{2}\ln\left(\frac{\sigma^2}{2\beta}\right)-\frac{1}{2}\sum_{j=1}^{n_i}\ln\left(1-E_{i,j}^{2\beta}\right)-\frac{1}{2}\ln(D_i)\\ &-\frac{\beta}{D_i\sigma^2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\mu-\left(y_{i,j-1}-\mu\right)E_{i,j}^{\beta}\right)^2}{1-E_{i,j}^{2\beta}}+\frac{2\beta^2\theta^2}{D_i\sigma^4}\left(\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-y_{i,j-1}E_{i,j}^{\beta}\right)}{1+E_{i,j}^{\beta}}\right)^2\\ &-\frac{\beta(D_i-1)}{D_i\sigma^2}\left(\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-y_{i,j-1}E_{i,j}^{\beta}\right)^2}{1-E_{i,j}^{2\beta}}\right)\Bigg), \end{aligned} \tag{4.4}$$

with $D_i=\frac{2\beta\theta^2}{\sigma^2}\sum_{j=1}^{n_i}\frac{1-E_{i,j}^{\beta}}{1+E_{i,j}^{\beta}}+1$.

This is not always possible for other situations and approximation methodologies are very useful to overcome this problem.

In the following subsections we are going to approximate the integral in (4.3) using the Laplace and delta approximation methods, allowing to obtain closed-form approximate expressions for the likelihood function. This might seem unnecessary since, in this case, we have an exact explicit expression, but it allows us to use this case as a benchmark to evaluate the quality of the approximations.

4.1. Likelihood function through the Laplace approximation method

Notice that the integrand function of the integral in (4.3) is the exponential of

$$f_i(\alpha_i)=-\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}-\frac{(\alpha_i-\mu)^2}{2\theta^2}\,. \tag{4.5}$$

To use the Laplace approximation method, we first need to obtain the value $\hat{\alpha}_i$ of α_i that maximizes (4.5). The function $f_i(\alpha_i)$ is a real twice continuously differentiable function with first and second order derivatives, respectively

$$f_i'(\alpha_i)=-\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta}\right)\left(E_{i,j}^{\beta}-1\right)}{\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)}-\frac{\alpha_i-\mu}{\theta^2} \tag{4.6}$$

and

$$f_i''(\alpha_i)=-\sum_{j=1}^{n_i}\frac{\left(E_{i,j}^{\beta}-1\right)^2}{\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)}-\frac{1}{\theta^2}\,. \tag{4.7}$$

In this case, due to the quadratic form of the function to be maximized, an exact expression for $\hat{\alpha}_i$ can be easily obtained and is given by

$$\hat{\alpha}_i = \left(\sum_{j=1}^{n_i} \frac{\left(E_{i,j}^{\beta}-1\right)^2}{\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)} + \frac{1}{\theta^2}\right)^{-1} \left(\sum_{j=1}^{n_i} \frac{\left(y_{i,j}-y_{i,j-1}E_{i,j}^{\beta}\right)\left(1-E_{i,j}^{\beta}\right)}{\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)} + \frac{\mu}{\theta^2}\right) \tag{4.8}$$

In other cases, where it is not possible to obtain an explicit expression, a numerical method must be used. By taking the second order Taylor series expansion of $f_i(\alpha_i)$ at $\hat{\alpha}_i$, and since $\hat{\alpha}_i$ is a maximum of f_i and therefore $f'(\hat{\alpha}_i) = 0$, we can write

$$\begin{aligned}
\int_{-\infty}^{+\infty} \exp\left(f_i(\alpha_i)\right) d\alpha_i &\approx \int_{-\infty}^{+\infty} \exp\left(f_i(\hat{\alpha}_i) + \frac{1}{2} f_i''(\hat{\alpha}_i)(\alpha_i-\hat{\alpha}_i)^2\right) d\alpha_i \\
&= \exp\left(f_i(\hat{\alpha}_i)\right) \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}\frac{(\alpha_i-\hat{\alpha}_i)^2}{-\left(f_i''(\hat{\alpha}_i)\right)^{-1}}\right) d\alpha_i \\
&= \exp\left(f_i(\hat{\alpha}_i)\right) \sqrt{\frac{2\pi}{\left|f_i''(\hat{\alpha}_i)\right|}}\,.
\end{aligned} \tag{4.9}$$

The usual procedure in applying the Laplace approximation method would consist in replacing expressions (4.7) and (4.8) in (4.9) and then in (4.3) in order to obtain an approximate expression for $LL(\mu,\theta,\beta,\alpha)$. Maximizing such expression w.r.t. the parameters, we would obtain approximately their maximum likelihood estimates.

However, in the random α case, due to the quadratic nature of the function $f_i(\alpha_i)$, it turns out that the approximate expression (4.9) is in fact exact. Indeed, it is a trivial exercise to show that, for any quadratic function, the second order Taylor expansion is exactly equal to the function and, therefore, in the random α case, the second-order Taylor expansion $f_i(\hat{\alpha}_i) + \frac{1}{2} f_i''(\hat{\alpha}_i)(\alpha_i-\hat{\alpha}_i)^2$ of $f_i(\alpha_i)$, on which (4.9) is based, is an exact expression for $f_i(\alpha_i)$. So, we can in this case write the equal sign instead of the approximation sign in expression (4.9) and, therefore, replacing (4.9) in (4.3) provides, not just an approximation of the log-likelihood function but rather the exact log-likelihood function, for which we have already provided expression (4.4). In summary, in the random α case, the Laplace approximation method provides exactly the same results as the exact method (i.e., the method of direct maximization of expression (4.4) w.r.t. the parameters).

4.2. Likelihood function through the delta approximation method

A second approximation of the integral in (4.3) is obtained using the delta approximation method. For this approximation we need to rearrange the expression of the likelihood function as

$$\begin{aligned}
L(\mu,\theta,\beta,\sigma) = \prod_{i=1}^{M} \left(\prod_{j=1}^{n_i} \left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}} \right) \right. \\
\left. \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2} \sum_{j=1}^{n_i} \frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})} \right) \frac{\exp\left(-\frac{(\alpha_i-\mu)^2}{2\theta^2}\right)}{\theta\sqrt{2\pi}}\, d\alpha_i \right).
\end{aligned} \tag{4.10}$$

Isolating, in the integrand function, the p.d.f. of the Gaussian distribution of the α_i and denoting by $\mathbf{E}$ the expectation with respect to that distribution, we get

$$\begin{aligned} L(\mu,\theta,\beta,\sigma) &= \prod_{i=1}^{M}\left(\prod_{j=1}^{n_i}\left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}}\right)\int_{-\infty}^{+\infty} u_i(\alpha_i)\frac{1}{\theta\sqrt{2\pi}}\exp\left(-\frac{(\alpha_i-\mu)^2}{2\theta^2}\right)d\alpha_i\right) \\ &= \prod_{i=1}^{M}\left(\prod_{j=1}^{n_i}\left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}}\right)\mathbf{E}[u_i(\alpha_i)]\right), \end{aligned} \tag{4.11}$$

where

$$u_i(\alpha_i) = \exp\left(g_i(\alpha_i)\right) \tag{4.12}$$

with

$$g_i(\alpha_i) = -\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}\,. \tag{4.13}$$

Using the same idea of the delta method to approximate $\mathbf{E}[u_i(\alpha_i)]$, as in [3], where $\alpha_i = \mu + \theta\delta_i$ with δ_i $(i = 1,\ldots,M)$ i.i.d. standard Gaussian random variables, and noting that $u_i(\alpha_i)$ is a real twice continuously differentiable function, we obtain

$$u_i(\alpha_i) = u_i(\mu+\theta\delta_i) \approx u_i(\mu) + \theta\delta_i u_i'(\mu) + \frac{1}{2}(\theta\delta_i)^2 u_i''(\mu) \tag{4.14}$$

and

$$\mathbf{E}[u_i(\alpha_i)] \approx u_i(\mu) + \frac{1}{2}\theta^2 u_i''(\mu), \tag{4.15}$$

with

$$u_i''(\mu) = u_i(\mu)\left[\left(g_i'(\mu)\right)^2 + g_i''(\mu)\right], \tag{4.16}$$

where the first and second order derivatives are already presented in [3] and so we omit them here.

Replacing (4.15) in (4.11), we obtain

$$L(\mu,\theta,\beta,\sigma) \approx \prod_{i=1}^{M}\left(\prod_{j=1}^{n_i}\left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}}\right)\left(u_i(\mu)+\frac{1}{2}\theta^2 u_i''(\mu)\right)\right). \tag{4.17}$$

Now, replacing (4.12) and (4.16) in (4.17), we get the expression for the log-likelihood function

$$\begin{aligned} LL(\mu,\theta,\beta,\sigma) \approx \sum_{i=1}^{M}\Bigg(&-\frac{1}{2}\sum_{j=1}^{n_i}\ln\left(2\pi\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})\right) - \frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\mu-(y_{i,j-1}-\mu)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})} \\ &+\ln\Bigg(1+\frac{1}{2}\theta^2\Bigg(-\sum_{j=1}^{n_i}\frac{(1-E_{i,j}^{\beta})^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})} \\ &+\left(\sum_{j=1}^{n_i}\frac{(1-E_{i,j}^{\beta})(y_{i,j}-\mu-(y_{i,j-1}-\mu)E_{i,j}^{\beta})}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}\right)^2\Bigg)\Bigg)\Bigg). \end{aligned} \tag{4.18}$$

As we can observe, the approximation of the log-likelihood function through the delta method provides a very simple closed-form expression.

5. Mixed models with random growth parameter

A similar procedure as the one presented in the previous section can be used in the case where the growth parameter β_i is considered a random variable following a Gaussian distribution with mean λ and variance ω^2. The joint p.d.f. for animal i is now given by

$$p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha,\lambda,\omega,\sigma) = \int_{-\infty}^{+\infty} p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha,\beta_i,\sigma)p(\beta_i|\lambda,\omega)\,d\beta_i, \tag{5.1}$$

where $p_{Y_i}(\mathbf{y}_i|\alpha_i,\beta,\sigma)$ is obtained by replacing β with β_i in (2.3) and the p.d.f. of β_i is

$$p(\beta_i|\lambda,\omega) = \frac{1}{\sqrt{2\pi\omega^2}} e^{-\frac{(\beta_i-\lambda)^2}{2\omega^2}}. \tag{5.2}$$

Thus, from (3.1) and using (5.1), after some simplifications we obtain the likelihood function

$$\begin{aligned} L(\alpha,\lambda,\omega,\sigma) = \prod_{i=1}^{M}\Bigg(&\frac{1}{\sqrt{2\pi\omega^2}}\prod_{j=1}^{n_i}\left(\frac{1}{\sqrt{2\pi}}\right) \\ &\int_{-\infty}^{+\infty}\frac{\exp\left(-\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})}\right)}{\exp\left(\frac{1}{2}\sum_{j=1}^{n_i}\ln\left(\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})\right)\right)}\exp\left(-\frac{(\beta_i-\lambda)^2}{2\omega^2}\right)d\beta_i\Bigg). \end{aligned} \tag{5.3}$$

The use of the delta and the Laplace approximation methods allows us to obtain approximate expressions for the likelihood function.

5.1. Likelihood function through the Laplace approximation method

The integrand function in expression (5.3) can written as the exponential of the function

$$f_i(\beta_i) = -\frac{1}{2}\sum_{j=1}^{n_i}\left(\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})} + \frac{1}{n_i}\frac{(\beta_i-\lambda)^2}{\omega^2} + \ln\left(\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})\right)\right). \tag{5.4}$$

The function $f_i(\beta_i)$ is a real twice continuously differentiable function. First, we need to obtain the value $\hat{\beta}_i$ that maximizes (5.4) and then, in order to obtain a Laplace approximation, we need to apply to $f_i(\beta_i)$ a second order Taylor expansion about $\hat{\beta}_i$.
Since in this case we cannot obtain a closed-form expression for the maximum $\hat{\beta}_i$, we can apply a numerical method. Having the numerical maximum, $\hat{\beta}_i$, and resorting to the same approach used in (4.9), the Laplace approximation of the integral in (5.3) will be given by

$$\exp\left(f_i(\hat{\beta}_i)\right)\sqrt{\frac{2\pi}{|f_i''(\hat{\beta}_i)|}}, \tag{5.5}$$

with

$$
\begin{aligned}
f_i'(\beta_i) = \sum_{j=1}^{n_i} \Bigg(& -\frac{E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})}{\left(1-E_{i,j}^{2\beta_i}\right)} - \frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \\
& -\frac{\beta_i-\lambda}{n_i\omega^2} + \frac{2\beta_i E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \\
& +\frac{1}{2\beta_i} - \frac{2\beta_i E_{i,j}^{\beta_i}(t_{i,j}-t_{i,j-1})(y_{i,j-1}-\alpha)\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \Bigg)
\end{aligned} \tag{5.6}
$$

and

$$
\begin{aligned}
f_i''(\beta_i) = \sum_{j=1}^{n_i} \Bigg(& \frac{2E_{i,j}^{4\beta_i}(t_{i,j}-t_{i,j-1})^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2} - \frac{2\beta_i E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})^2(y_{i,j-1}-\alpha)^2}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \\
& +\frac{2E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})^2}{\left(1-E_{i,j}^{2\beta_i}\right)} + \frac{4E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \\
& -\frac{1}{2\beta_i^2} - \frac{8\beta_i E_{i,j}^{4\beta_i}(t_{i,j}-t_{i,j-1})^2\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^3\sigma^2} \\
& -\frac{1}{n_i\omega^2} - \frac{4\beta_i E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})^2\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \\
& -\frac{4E_{i,j}^{\beta_i}(t_{i,j}-t_{i,j-1})(y_{i,j-1}-\alpha)\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \\
& +\frac{2\beta_i E_{i,j}^{\beta_i}(t_{i,j}-t_{i,j-1})^2(y_{i,j-1}-\alpha)\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \\
& +\frac{8\beta_i E_{i,j}^{3\beta_i}(t_{i,j}-t_{i,j-1})^2(y_{i,j-1}-\alpha)\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \Bigg).
\end{aligned} \tag{5.7}
$$

Replacing (5.5) in (5.3), we obtain, after some algebraic manipulations,

$$
LL(\alpha,\lambda,\omega,\sigma) \approx \sum_{i=1}^{M}\left(-\ln\omega - \frac{n_i}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\left|f_i''(\hat{\beta}_i)\right|\right) + f_i(\hat{\beta}_i)\right). \tag{5.8}
$$

For the Laplace approximation we can observe that it is possible to obtain a closed-form approximation formula of the log-likelihood, but for that we need to obtain the maximum value of β_i by numerical methods. This slows down the maximization of the log-likelihood function, which uses an iterative procedure, since, at each step of that procedure, the numerical determination of the maximum values of the β_i needs to be repeated. These repeated numerical determinations are not required in the delta approximation method.

5.2. Likelihood function through the delta approximation method

Considering now the likelihood function (5.3) and, proceeding as in subsection 4.2 with

$$u_i(\beta_i) = \exp\left(-\frac{1}{2}\sum_{j=1}^{n_i}\left(\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})}+\ln\left(2\pi\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})\right)\right)\right), \tag{5.9}$$

we see that the likelihood function can be approximated, using the delta method, by

$$L(\alpha,\lambda,\omega,\sigma) = \prod_{i=1}^{M}\mathbf{E}[u_i(\beta_i)] \approx \prod_{i=1}^{M}\left(u_i(\lambda)+\frac{1}{2}\omega^2 u_i''(\lambda)\right), \tag{5.10}$$

where we have used the approximation

$$u_i(\beta_i) = u_i(\lambda+\omega\delta_i) \approx u_i(\lambda)+\omega\delta_i u_i'(\lambda)+\frac{1}{2}(\omega\delta_i)^2 u_i''(\lambda) \tag{5.11}$$

and

$$\mathbf{E}[u_i(\beta_i)] \approx u_i(\lambda)+\frac{1}{2}\omega^2 u_i''(\lambda), \tag{5.12}$$

with

$$u_i''(\lambda) = u_i(\lambda)\left[(g_i'(\lambda))^2+g_i''(\lambda)\right], \tag{5.13}$$

So, we obtain for the log-likelihood function

$$\begin{aligned} LL(\alpha,\lambda,\omega,\sigma) \approx \sum_{i=1}^{M}\Bigg(&-\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\lambda}\right)^2}{\frac{\sigma^2}{2\lambda}(1-E_{i,j}^{2\lambda})} \\ &-\frac{1}{2}\sum_{j=1}^{n_i}\ln\left(2\pi\frac{\sigma^2}{2\lambda}(1-E_{i,j}^{2\lambda})\right) \\ &+\ln\left(1+\frac{1}{2}\omega^2\left[(g_i'(\lambda))^2+g_i''(\lambda)\right]\right)\Bigg), \end{aligned} \tag{5.14}$$

where

$$g_i(\lambda) = -\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\lambda}\right)^2}{\frac{\sigma^2}{2\lambda}(1-E_{i,j}^{2\lambda})}-\frac{1}{2}\sum_{j=1}^{n_i}\ln\left(2\pi\frac{\sigma^2}{2\lambda}(1-E_{i,j}^{2\lambda})\right) \tag{5.15}$$

and its first and second order derivatives are presented in [3] and omitted here.

6. Mixed models with random asymptotic average size and random growth parameter

In this section, we consider the mixed model where $\alpha_i\,(i=1,\ldots,M)$ are i.i.d. random variables following a Gaussian distribution with mean μ and variance θ^2 and $\beta_i\,(i=1,\ldots,M)$ are i.i.d. random variables following a Gaussian distribution with mean λ and variance ω^2 and are independent of the

α_i variables. 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 $\epsilon = (\sigma)$, the random effects parameter vector is $\mathbf{\Psi} = (\mu, \lambda, \theta, \omega)$, and the p.d.f. of the random effects p is a bivariate Gaussian distribution

$$\begin{aligned} p(\alpha_i, \beta_i|\mu, \lambda, \theta, \omega) &= p(\alpha_i|\mu, \theta)p(\beta_i|\lambda, \omega) \\ &= \frac{1}{\sqrt{2\pi\theta^2}} \exp\left(-\frac{(\alpha_i - \mu)^2}{2\theta^2}\right) \frac{1}{\sqrt{2\pi\omega^2}} \exp\left(-\frac{(\beta_i - \lambda)^2}{2\omega^2}\right). \end{aligned} \tag{6.1}$$

To obtain the likelihood function in (3.1), the joint p.d.f. for animal i is given by

$$\begin{aligned} p_{\mathbf{Y}_i}(\mathbf{y}_i|\epsilon, \mathbf{\Psi}) &= \int\limits_{\mathbf{R}^2} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{b}_i, \epsilon)p_B(\mathbf{b}_i|\mathbf{\Psi})\, d\mathbf{b}_i \\ &= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i, \beta_i, \epsilon)p(\alpha_i|\mu, \theta)p(\beta_i|\lambda, \omega)\, d\alpha_i\, d\beta_i, \end{aligned} \tag{6.2}$$

To obtain $p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i, \beta_i, \epsilon)$ in (6.2), notice that, due to the Markov property of each individual trajectory i when conditioned on α_i and β_i, we have

$$p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i, \beta_i, \epsilon) = \prod_{j=1}^{n_i} p_{\mathbf{Y}_i}(y_{i,j}, t_i, t_{i-1}|y_{i,j-1}, \alpha_i, \beta_i, \epsilon), \tag{6.3}$$

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

$$p_{\mathbf{Y}_i}(y_{i,j}, t_i, t_{i-1}|y_{i,j-1}, \alpha_i, \beta_i, \epsilon) = \frac{\exp\left(-\frac{1}{2}\frac{\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})}\right)}{\sqrt{2\pi\frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})}}. \tag{6.4}$$

Thus, replacing (6.1) and (6.4) on (3.1), the likelihood function, after some simplifications, is given by

$$\begin{aligned} L(\mu, \theta, \lambda, \omega, \sigma) = \prod_{\boldsymbol{i=1}}^{\boldsymbol{M}} \frac{1}{2\pi\theta\omega} \left(\frac{1}{\sqrt{2\pi}}\right)^{n_i} \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}\sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})}\right) \\ \exp\left(-\frac{1}{2}\sum_{j=1}^{n_i} \ln\left(\frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})\right)\right) \exp\left(-\frac{1}{2}\frac{(\alpha_i - \mu)^2}{\theta^2}\right) \exp\left(-\frac{1}{2}\frac{(\beta_i - \lambda)^2}{\omega^2}\right) d\alpha_i\, d\beta_i. \end{aligned} \tag{6.5}$$

In the next sections, we will use the Laplace and the delta approximation methods to obtain approximate expressions to the likelihood function.

6.1. Likelihood function through the Laplace approximation method

Following the same approach taken in subsections 4.1 and 5.1, the integrand function in the likelihood function (6.5) can be written as the exponential of the function

$$f_i(\alpha_i, \beta_i) = \sum_{j=1}^{n_i} \left(-\frac{1}{2}\frac{\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})}\right. \tag{6.6}$$

$$-\frac{1}{2}\ln\left(\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})\right)-\frac{(\alpha_i-\mu)^2}{2n_i\theta^2}-\frac{(\beta_i-\lambda)^2}{2n_i\omega^2}\Bigg).$$

Rewriting expression (6.5) we obtain

$$L(\mu,\theta,\lambda,\omega,\sigma)=\prod_{i=1}^{M}\left(\frac{1}{2\pi\theta\omega}\left(\frac{1}{\sqrt{2\pi}}\right)^{n_i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left(f_i(\alpha_i,\beta_i)\right)\,d\alpha_i\,d\beta_i\right). \tag{6.7}$$

As shown in subsection 4.1, when it is not possible to obtain an explicit expression for the vector $(\hat{\alpha}_i,\hat{\beta}_i)$ that maximizes the function (6.7), a numerical method must be used. By taking the second order Taylor series expansion of (α_i,β_i) at $(\hat{\alpha}_i,\hat{\beta}_i)$, and since $(\hat{\alpha}_i,\hat{\beta}_i)$ is a maximum of $f_i(\hat{\alpha}_i,\hat{\beta}_i)$ and therefore $\triangledown f_i(\hat{\alpha}_i,\hat{\beta}_i)=0$, we can write

$$\begin{aligned}
&\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left(f_i(\alpha_i,\beta_i)\right)\,d\alpha_i\,d\beta_i\\
&\approx\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left(f_i(\hat{\alpha}_i,\hat{\beta}_i)+\frac{1}{2}\begin{bmatrix}\alpha_i-\hat{\alpha}_i & \beta_i-\hat{\beta}_i\end{bmatrix}\mathbf{H}\Big[f_i(\hat{\alpha}_i,\hat{\beta}_i)\Big]\begin{bmatrix}\alpha_i-\hat{\alpha}_i\\ \beta_i-\hat{\beta}_i\end{bmatrix}\right)d\alpha_i\,d\beta_i\\
&=2\pi\left|\det\boldsymbol{\Sigma}_{(\hat{\alpha}_i,\hat{\beta}_i)}\right|^{1/2}\exp(f_i(\hat{\alpha}_i,\hat{\beta}_i)),
\end{aligned} \tag{6.8}$$

where $\mathbf{H}[f_i(\hat{\alpha}_i,\hat{\beta}_i)]$ is the Hessian matrix of $f_i(\alpha_i,\beta_i)$ computed at $(\hat{\alpha}_i,\hat{\beta}_i)$, assumed to be negative definite, and $\boldsymbol{\Sigma}_{(\hat{\alpha}_i,\hat{\beta}_i)}=-\mathbf{H}^{-1}[f_i(\hat{\alpha}_i,\hat{\beta}_i)]$.
The first and second partial derivatives with respect to α_i are given in (4.6) and (4.7) with β replaced by β_i. Therefore, an exact expression for $\hat{\alpha}_i$ is given by (4.8) with β replaced by $\hat{\beta}_i$. The first and second partial derivatives with respect to β_i were also already given in subsection 5.1, in expressions (5.6) and (5.7), but we must replace in these expressions α with $\hat{\alpha}_i$. The expression for $\frac{\partial^2 f_i(\alpha_i,\beta_i)}{\partial\alpha_i\partial\beta_i}=\frac{\partial^2 f_i(\alpha_i,\beta_i)}{\partial\beta_i\partial\alpha_i}$ is given by

$$\begin{aligned}
\frac{\partial^2 f_i(\alpha_i,\beta_i)}{\partial\alpha_i\partial\beta_i} &= \frac{4\beta_i}{\sigma^2}\sum_{j=1}^{n_i}\frac{(t_j-t_{j-1})E_{i,j}^{2\beta_i}\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)\left(E_{i,j}^{\beta_i}-1\right)}{(1-E_{i,j}^{2\beta_i})^2}\\
&- \frac{2\beta_i}{\sigma^2}\sum_{j=1}^{n_i}\frac{(t_j-t_{j-1})(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\left(E_{i,j}^{\beta_i}-1\right)}{(1-E_{i,j}^{2\beta_i})}\\
&+ \frac{2\beta_i}{\sigma^2}\sum_{j=1}^{n_i}\frac{(t_j-t_{j-1})E_{i,j}^{\beta_i}\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)}{(1-E_{i,j}^{2\beta_i})}\\
&- \frac{2}{\sigma^2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)\left(E_{i,j}^{\beta_i}-1\right)}{(1-E_{i,j}^{2\beta_i})},
\end{aligned} \tag{6.9}$$

Replacing (6.8) in (6.7), we obtain, after some algebraic manipulation

$$LL(\mu,\theta,\lambda,\omega,\sigma)\approx\sum_{i=1}^{M}\left\{-\ln(\theta\omega)-\frac{n_i}{2}\ln(2\pi)-\frac{1}{2}\ln(|\det\mathbf{H}[f_i(\hat{\alpha}_i,\hat{\beta}_i)]|)+f_i(\hat{\alpha}_i,\hat{\beta}_i)\right\}. \tag{6.10}$$

For the Laplace approximation it is possible to obtain a closed-form approximation formula of the log-likelihood, but, like in the random β case, that formula involves now the maximum value of (α_i, β_i), which has to be determined numerically and repeatedly at each step of the iterative procedure of the log-likelihood maximization. In the next section we present the delta approximation method and we can see that, for such approximation method, it is possible to obtain a closed-form approximation formula without having to resort to such repeated numerical determinations.

6.2. Likelihood function through the delta approximation method

In [3] the delta approximation method was proposed to approximate the integral in (3.1), which, in this case, is given by (6.2) and does not have an explicit solution. This approximation method allows us to obtain simpler and closed-form expressions for the likelihood function. As shown in [3], denoting expression (6.5) by $u_i(\alpha_i, \beta_i) = \exp(g_i(\alpha_i, \beta_i))$, where

$$g_i(\alpha_i, \beta_i) = \sum_{j=1}^{n_i} \left(-\frac{1}{2} \frac{\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i) E_{i,j}^{-\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})} - \frac{1}{2} \ln\left(2\pi \frac{\sigma^2}{2\beta_i}(1 - E_{i,j}^{2\beta_i})\right) \right), \tag{6.11}$$

the likelihood function (6.5) can be written as

$$\begin{aligned} L(\mu, \theta, \lambda, \omega, \sigma) &= \prod_{i=1}^{M} \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} u_i(\alpha_i, \beta_i) p_B(\alpha_i, \beta_i | \mu, \lambda, \theta, \omega) \, d\alpha_i \, d\beta_i \right) \\ &= \prod_{i=1}^{M} \mathbf{E}[u_i(\alpha_i, \beta_i)]. \end{aligned} \tag{6.12}$$

The mathematical expectation of $u_i(\alpha_i, \beta_i)$ can be approximately obtained by applying the delta approximation, consisting in a second-order Taylor series expansion and resulting in

$$\mathbf{E}[u_i(\alpha_i, \beta_i)] \approx u_i(\mu, \lambda) + \frac{\theta^2}{2} \frac{\partial^2 u_i}{\partial \alpha_i^2}(\mu, \lambda) + \frac{\omega^2}{2} \frac{\partial^2 u_i}{\partial \beta_i^2}(\mu, \lambda). \tag{6.13}$$

The log-likelihood function, $LL_Y(\mu, \theta, \lambda, \omega, \sigma) = \ln L(\mu, \theta, \lambda, \omega, \sigma)$, will be then given by

$$LL_Y(\mu, \theta, \lambda, \omega, \sigma) \approx \sum_{i=1}^{M} \ln \left\{ u_i(\mu, \lambda) + \frac{\theta^2}{2} \frac{\partial^2 u_i}{\partial \alpha_i^2}(\mu, \lambda) + \frac{\omega^2}{2} \frac{\partial^2 u_i}{\partial \beta_i^2}(\mu, \lambda) \right\}. \tag{6.14}$$

The closed-form expressions for the first and second order derivatives of $u_i(\alpha_i, \beta_i)$ with respect to α_i and β_i are presented in [3].

7. Performance comparison of the methods

In this section, for the SDE mixed cattle growth models, our focus is on the parameters estimation and the approximation methods to achieve it when no exact method is available. Initially, we will evaluate the performance of the DA method and compare it with the LA method. Additionally, to benchmark these two approximation methods against numerically implemented alternatives, we will compare the estimates of DA and LA methods with those obtained using the *mixedsde* R package. However, it is important to note that, unlike the DA and LA approximation methods, the *mixedsde*

package requires uniformity in the time vector of observations across all trajectories. This constraint prevents us from using the real cattle dataset for a comprehensive comparison of the three methods. For this purpose, we generated simulated datasets containing monthly weights for 500 animals, covering the period from birth to four years of age and resulting in a total of 24,500 observations. The trajectories were simulated using the stochastic Gompertz model (using the logarithm transformation of the weights). We used the following methodology:

- The 500 random parameter values of the $M = 500$ animals were simulated based on a Gaussian distribution.
  - For the case of random α, we have used a mean $\mu = 6.45$ (corresponding to the asymptotic weight $h^{-1}(\mu) = 632.7$ kg) and a standard deviation $\theta = 0.15$, thus obtaining, for each animal i ($i = 1, \ldots, 500$), the corresponding α_i value, to be incorporated in the transition density (2.2) as the true value of α for that animal.
  - For the random β case, we have used a mean $\lambda = 1.43\ year^{-1}$ and a standard deviation $\omega = 0.30\ year^{-1}$, thus obtaining, for each animal i ($i = 1, \ldots, 500$), the corresponding β_i value, to be incorporated in the transition density (2.2) as the true value of β for that animal.
  - For the case where both α and β are considered random, we have used the same values for the parameters as used in the separated α and β cases.
  - As for the remaining non-random parameters, we use $\sigma = 0.33\ year^{-1/2}$.
  - We have used $\beta = 1.43\ year^{-1}$ and $\alpha = 6.45$, respectively for the random α case and the random β case.
- Having the parameters defined, we have simulated the trajectory of each animal based on the Markov property and the use of the transition densities between consecutive observation times.
- We have obtained one simulated dataset for the case where the parameter α is considered random, one for the case where the parameter β is considered random, and another for the case where both parameters are considered random.

In Section 7.1, we will analyze the simulated dataset where the parameter α is considered random, evaluating the performance of the estimation methods with a random average asymptotic size. Following this, in Section 7.2, we will examine the simulated dataset with the parameter β considered as random, assessing the performance of the estimation methods with a random growth parameter. Subsequently, in Section 7.2, we will explore the simulated dataset when both parameters are considered random, evaluating the performance of the estimation methods with two random effects. In this scenario, we will also assess the performance when only the parameter α is random or when only the parameter β is random. Finally, in Section 7.4, we will consider the real cattle weight dataset, comprising 10843 animals. For this real data, we will compare the DA and LA estimation methods when only the parameter α is random, when only the parameter β is random, and when both parameters are considered random. The comparative performance study cannot in this case use the results of the *mixedsde* package.

To assess the performance of the estimates, within each subsection, we conduct a bias analysis and evaluate the precision of the standard deviation estimates obtained from the empirical Fisher information matrix. The ultimate goal is to assess the performance of the DA method. First, for benchmarking purposes, when an exact method is available (the case when the parameter α is considered random) and, at last, in the absence of an exact method, to investigate whether the DA

method can be used for parameter estimation of the model (cases when the parameter β is considered random or both parameters are considered random).

7.1. Model with a random average asymptotic size

In this subsection, using the simulated dataset with α a Gaussian random variable with mean μ and standard deviation ω, we obtain the estimates of the parameters, μ, θ, β, and σ, by maximization of the exact log-likelihood function (4.4) ("Exact method"). We will also estimate the parameters by the DA method, i.e., by maximization of the approximate log-likelihood function (4.18) using the DA method. The maximum likelihood estimates based on the exact log-likelihood function (4.4) ("Exact method") will be the same estimates obtained by the LA method because, in this case and as shown on subsection 4.1, the approximation of the log-likelihood function obtained by the LA method is in fact exact and coincides with the exact log-likelihood function. These "Exact method" estimates allow us to evaluate the performance of the DA method for the case where the parameter α is random. In this case, the "Exact method" estimates are available and can be used to assess, by comparison, the quality of the DA method estimates.

We also compare with the estimates obtained by the R package *mixedsde*. Table 1 presents the estimation results for the stochastic Gompertz model. The maximum likelihood estimates and the 95% approximate confidence bands based on the inverse of the empirical Fisher information matrix are presented for the DA method and also, using the exact log-likelihood function (4.4), for the "Exact method". Note that the *mixedsde* package does not provide confidence intervals for the estimates.

Table 1. Results for the simulated dataset assuming a random α. The table shows the true parameter values used in the simulations, the maximum likelihood estimates and corresponding 95% asymptotic confidence bands using the "Exact method"/LA method and the DA method. Estimates obtained by the *mixedsde* R package are shown. Instead of showing the results for the modified weight μ, we present them for the real weight $A = h^{-1}(\mu)$.

	True	**DA method**	**Exact/LA method**	**mixedsde**
A	632.70	635.47 ± 9.56	625.71 ± 11.56	633.79
θ	0.15	0.1197 ± 0.0085	0.1555 ± 0.0150	0.0787
β	1.43	1.4037 ± 0.0188	1.4221 ± 0.0191	1.3117
σ	0.33	0.3312 ± 0.0030	0.3289 ± 0.0030	0.4671

The DA method and the *mixedsde* method present an estimate of A closer to the true value than the "Exact method". The estimate of the standard deviation θ of the random parameter α presents a greater bias in the *mixedsde* method and in the DA method than in the "Exact method". We must refer that for the DA method, the asymptotic confidence interval does not include the true value for θ. For the fixed parameters β and σ, the estimates obtained from the DA method are quite close to the true value and similar to those obtained by the "Exact method". The *mixedsde* underestimates or overestimates these two parameters.

Since for this simulated dataset, we know the true value of the parameters, we can compare them with the estimates obtained by the different methods and confirm if they are included in the asymptotic confidence intervals for the DA method and the "Exact method". We can observe that only for the

DA method the asymptotic confidence interval does not include the true value for θ. Since these asymptotic confidence intervals based on the Fisher information matrix may be quite unreliable for small sample sizes, we proceed with a bias analysis and an evaluation of the quality of the standard deviation estimates obtained from the Fisher information matrix. With this purpose:

- We have obtained $K = 1000$ replications of the simulated dataset assuming random α with the same parameter vector $\mathbf{p} = (\mu = 6.45, \theta = 0.15, \beta = 1.43, \sigma = 0.33)$ (using the same procedure detailed at the beginning of Section 7);
- For each simulated dataset, we obtain the estimates through the DA method for random α and the LA method for random α (which in this case coincides with the "Exact method"), denoting them by $\hat{\mathbf{p}}_k$, $k = 1, \ldots, K$;
- The estimated standard deviations of the parameters are obtained by taking the square roots of the diagonal elements of the inverted empirical Fisher information matrix;
- The sample means and the sample standard deviations of the 1000 estimates $\hat{\mathbf{p}}_k$ were obtained;
- The sample means of the $K = 1000$ estimates of the parameter variances (obtained approximately for each replication by the inverse of the empirical Fisher information matrix) are also computed and their square roots will serve as estimates of the approximated standard deviations based on the empirical Fisher information matrix.

Following the procedure described above, in Table 2 we present, for the case of random α and this sample of 1000 parameter estimates, their sample mean values and standard deviations (which are very precise approximations of the true expected values and standard deviations of the parameter estimators), as well as the approximated standard deviations based on the empirical Fisher information matrix.

Table 2. Mean values and standard deviations ("SD") of the estimates obtained through the DA method and the "Exact method"/LA method for random α case, based on the 1000 simulated datasets, as well as the standard deviation estimate based on the empirical Fisher information matrix ("SD. Fisher"). Instead of showing the results for the modified weight μ, we show them for the weight $A = h^{-1}(\mu)$.

		DA method			**Exact/LA method**		
	True	**Mean estim.**	**SD**	**SD. Fisher**	**Mean estim.**	**SD**	**SD. Fisher**
A	632.70	637.17	6.4986	4.8735	632.69	5.9932	5.9059
θ	0.15	0.1179	0.0044	0.0043	0.1525	0.0075	0.0076
β	1.43	1.4118	0.0098	0.0097	1.4301	0.0096	0.0098
σ	0.33	0.3322	0.0015	0.0015	0.3300	0.0015	0.0015

Observing the results of Table 2, we can conclude that, on average, the parameter standard deviation estimates (SD. Fisher) obtained approximately by the empirical information matrix do not differ much (and with the exception of the parameter $A = h^{-1}(\mu)$, practically coincide) with the more precise standard deviations (SD) obtained from this extremely computer intensive and time-consuming procedure. This shows that the empirical Fisher information matrix is a relatively reliable method to estimate the standard deviation of the parameters and therefore to obtain approximate confidence

intervals for them. Comparing the mean estimates of the DA method with the true values, we can conclude that the DA method estimates very well all the parameters, except for the parameter θ, which is slightly underestimated. For the LA method, we can conclude that the standard deviations obtained from the empirical Fisher information matrix are very reliable for all parameters. Looking at the mean estimates, we can see that the LA method, which in this random α case coincides with the "Exact method", behaves, as expected, very well for all parameters.

7.2. Model with random growth parameter

Here we repeat the approach taken in the previous subsection, but now for the model where the growth parameter varies from animal to animal, considering β to be a Gaussian random variable with mean λ and standard deviation ω. For the simulated data assuming β is a random parameter, obtained as described at the beginning of Section 7, we present in Table 3 the results based on the DA and LA methods for the random parameter β, as well as the estimates obtained through the *mixedsde* package. Here, we do not have an exact closed-form expression for the log-likelihood function, so we can not present a comparison with an "Exact method". The maximum likelihood estimates and the 95% approximate confidence bands based on the inverse of the empirical Fisher information matrix were computed for the DA method and the LA method.

Table 3. Results for the simulated dataset assuming a random β. The table shows the true parameter values used in the simulations, the maximum likelihood estimates, and corresponding 95% asymptotic confidence bands using the DA and the LA methods for a random β. Estimates obtained for the *mixedsde* R package are shown. Instead of showing the results for the modified weight α, we present for the weight $A = h^{-1}(\alpha)$.

	True	**DA method**	**LA method**	**mixedsde**
A	632.70	640.70 ± 8.45	640.87 ± 7.88	639.32
λ	1.43	1.3451 ± 0.0216	1.4181 ± 0.0327	1.2765
ω	0.30	0.1925 ± 0.0105	0.3051 ± 0.0250	< 0.0001
σ	0.33	0.3360 ± 0.0030	0.3290 ± 0.0030	0.4661

From Table 3 we can conclude that the estimates obtained through the DA method and the LA method for the mean of the random growth parameter λ are closer to the true value than the estimate obtained from the *mixedsde*. For the standard deviation of the random growth parameter ω, the estimates given by the LA method (the best one) and the DA method are much better than those obtained with the *mixedsde* package. This last method was unable to detect the random effect of the parameter β. For the estimate of the measure of the random fluctuations of the environment, σ, the estimates obtained through the DA and LA methods are the closest to the true value.

We must refer that also in this case, for the DA method the asymptotic confidence intervals failed to include the true values for parameters λ and ω. Proceeding similarly as in subsection 7.1, but now for the case where the parameter β is considered random, we performed a bias analysis and an evaluation of the quality of the standard deviation estimates obtained from the empirical Fisher information matrix. We have simulated $K = 1000$ datasets assuming that the parameter β follows a Gaussian distribution

with mean $\lambda = 1.43\ year^{-1}$ and a standard deviation $\omega = 0.30\ year^{-1}$. For the fixed parameters we have considered $\alpha = 6.45$ and $\sigma = 0.33\ year^{-1/2}$ (using the same procedure detailed in Section 7). In Table 4 we present, for the DA method and the LA method with random β, the means and standard deviations of the parameter estimates of the 1000 simulated datasets.

Table 4. Mean values and standard deviations (SD) of the estimates obtained through the DA and the LA methods for random β, based on the 1000 simulated datasets and the standard deviation estimate based on the empirical Fisher information matrix ("SD. Fisher"). Instead of showing the results for the modified weight α, we show them for the weight $A = h^{-1}(\alpha)$.

		DA method			**LA method**		
	True	**Mean estim.**	**SD**	**SD. Fisher**	**Mean estim.**	**SD**	**SD. Fisher**
A	632.70	632.77	4.1616	4.1913	632.71	3.9484	3.9504
λ	1.43	1.3608	0.0218	0.0109	1.4301	0.0164	0.0165
ω	0.3	0.1979	0.0054	0.0053	0.2990	0.0123	0.0127
σ	0.33	0.3360	0.0017	0.0015	0.3301	0.0015	0.0015

Observing the results of Table 4, both for the DA method and for the LA method, we can conclude that the estimates of the standard deviations obtained from the empirical Fisher information matrix are practically the same as the more reliable (but extremely computer-intensive and time consuming) standard deviation estimates obtained from the 1000 simulated datasets showing that the empirical Fisher information matrix is a reliable method to estimate the variance of the parameters for the DA method and for the LA method with random β. Only for the DA method and the parameter λ the empirical Fisher information matrix seems to underestimate its standard deviation. Comparing the mean estimates of both methods with the true values, we can see that the DA method presents estimates that are very close to the true values of the fixed parameters α (A) and σ and underestimates the mean λ and the standard deviation ω of the random parameter, while for the LA method all the parameter estimates are very close to the true values.

7.3. Model with two random effects

We repeat the approach used in previous sections but now for the model where both the average asymptotic weight and the growth parameter are varying from animal to animal, with α being a Gaussian random variable with parameters vector (μ, θ) and β being a Gaussian random variable with parameters vector (λ, ω). A simulated dataset was obtained as described at the beginning of section 7, but assuming that both α and β are random and independent. In Table 5, the results based on the DA and LA methods are shown, as well as the estimates obtained through the *mixedsde* package when both α and β are random. The goal is to evaluate if the estimates obtained using the different methods are aligned. Here, we do not have an exact closed-form expression for the log-likelihood function, so we cannot present a comparison with an "Exact method". The maximum likelihood estimates and the 95% approximate confidence bands based on the inverse of the empirical Fisher information matrix were computed for the DA and the LA methods.

To check for the consequences of ignoring that both parameters are random, Table 5 also presents the parameter estimates under the DA and LA methods when we assume that only the parameter α is

random or when we assume that only the parameter β is random.

Table 5. Results for the simulated dataset assuming both α and β are random. The table shows the true parameter values used in the simulations, the maximum likelihood estimates, and corresponding 95% asymptotic confidence bands using the DA and the LA methods. Estimates obtained for the *mixedsde* R packages are shown. Instead of showing the results for the modified weight α (or, when α is random, for its mean μ), we present for the weight $A = h^{-1}(\alpha)$ (or $A = h^{-1}(\mu)$).

	True	**DA method**	**LA method**	**mixedsde**
		random (α, β)		
A	632.70	651.52 ± 10.06	661.02 ± 12.97	643.82
θ	0.15	0.0944 ± 0.0130	0.0316 ± 0.1026	0.1556
λ	1.43	1.3216 ± 0.0209	1.3515 ± 0.0272	1.2745
ω	0.30	0.1809 ± 0.0124	0.3100 ± 0.0366	< 0.0001
σ	0.33	0.3407 ± 0.0030	0.3344 ± 0.0030	0.4690
		random α		
A	632.70	643.91 ± 10.86	634.71 ± 14.27	643.82
θ	0.15	0.1330 ± 0.0083	0.2068 ± 0.0173	0.1557
λ	1.43	1.3361 ± 0.0190	1.3816 ± 0.0195	1.2745
ω	1.43	–	–	–
σ	0.33	0.3425 ± 0.0031	0.3374 ± 0.0031	0.4690
		random β		
A	632.70	659.58 ± 9.26	665.06 ± 8.80	645.38
θ	0.15	–	–	–
λ	1.43	1.3034 ± 0.0211	1.3693 ± 0.0336	1.2644
ω	0.30	0.1969 ± 0.0098	0.3187 ± 0.0258	0.0276
σ	0.33	0.3416 ± 0.0030	0.3341 ± 0.0030	0.4690

From Table 5 we can conclude that, for the case of random (α, β), the *mixedsde* method was unable to identify that the parameter β is a random parameter since the estimate of the parameter ω is practically null (the same happens in last section for the case of the random β). The LA and the DA methods estimate more accurately the parameters related to β with a slight underestimation of the mean value λ. The standard deviation of the β distribution, ω, was best estimated through the LA method. Regarding the estimation of the parameters related to α (A), the DA and the LA methods underestimate the parameter θ and slightly overestimate the weight at maturity A. The LA method even allows the value $\theta = 0$ inside the 95% confidence interval for θ, but one should keep in mind that this is an approximate confidence interval based on the inverse of the empirical Fisher information matrix. The *mixedsde* was the best method to estimate the parameter θ. All methods were able to correctly estimate the environmental random fluctuations σ. Observing the estimates of the DA method when using the random (α, β), the random α, and the random β, we can see that the estimates are very close. When

in the presence of a dataset with both parameters random and independent, the use of the DA method with only one parameter considered random can correctly identify as random that same parameter, and when using the DA method with both parameters considered random can correctly identify both parameters as random. For the LA method only the parameter θ was not correctly estimated when both parameters were random.

Proceeding similarly to subsections 7.1 and 7.2, but now for the case where both parameters α and β are considered random, we performed a bias analysis and an evaluation of the quality of the standard deviation estimates obtained from the empirical Fisher information matrix. We have simulated $K = 1000$ datasets assuming that the parameter α follows a Gaussian distribution with mean $\mu = 6.45$ and a standard deviation $\theta = 0.15$ and the parameter β follows an independent Gaussian distribution with mean $\lambda = 1.43\ year^{-1}$ and a standard deviation $\omega = 0.30\ year^{-1}$. For the fixed parameter we have considered $\sigma = 0.33\ year^{-1/2}$. In Table 6 we present, for the DA method and the LA method with both α and β random, the means and standard deviations of the parameter estimates of the 1000 simulated datasets.

Table 6. Mean values and standard deviations (SD) of the estimates obtained through the DA and the LA methods for both random α and β, based on the 1000 simulated datasets and the standard deviation estimate based on the empirical Fisher information matrix ("SD. Fisher"). Instead of showing the results for the modified weight α (or, when α is random, for its mean μ), we present for the weight $A = h^{-1}(\alpha)$ (or $A = h^{-1}(\mu)$).

		DA method			**LA method**		
	True	**Mean estim.**	**SD**	**SD. Fisher**	**Mean estim.**	**SD**	**SD. Fisher**
A	632.70	638.56	6.6072	5.0429	653.23	6.9991	4.3928
θ	0.15	0.0967	0.0067	0.0067	0.0059	0.0332	0.0075
λ	1.43	1.3387	0.0213	0.0110	1.3855	0.0171	0.0169
ω	0.30	0.1785	0.0073	0.0068	0.3085	0.0238	0.0129
σ	0.33	0.3394	0.0018	0.0015	0.3333	0.0015	0.0015

Observing the results of Table 6, we can conclude that the mean estimates under both methods are, in general, aligned with the true values. The DA method underestimates the parameters related to β, λ and ω, and the parameter θ. The LA method slightly overestimates the parameter A and the observation made above that the LA method allows the value $\theta = 0$ inside the 95% confidence interval for θ is now confirmed, and even reinforced when using a more reliable estimate (SD) of the standard deviation of the θ estimator. Both methods estimate very well the fixed parameter σ.

The estimates of the standard deviations obtained from the empirical Fisher information matrix are practically the same as the more reliable standard deviation estimates obtained from the 1000 simulated datasets for the DA method. However, for the LA method, the standard deviations obtained from the empirical Fisher information matrix for the parameters A, θ, and ω, are not close to the ones obtained from the standard deviation estimates of 1000 simulated datasets. This shows that, in the scenario of random and independent α and β, the empirical Fisher information matrix is a reliable method to estimate the variance of the parameters for the DA method, but not so robust when considering the LA method.

7.4. Real data application

In this section we are going to use the DA method and the LA method to estimate the parameters for the stochastic Gompertz model using real cattle weight data, where each animal has its weight measurements taken at different age instants, varying from animal to animal. For that reason, other available R packages (for example [17–19]) cannot be used since they require the same age vector for all the animals. Our real data was provided by the Associação de Criadores de Bovinos Mertolengos (ACBM), which preforms the growing and finishing phases of young Mertolengo males, and by associated breeders, whose agricultural holdings are located in the Alentejo region in Portugal. As mentioned in Section 1, the database contains information on the weight of 10843 Mertolengo cattle males taken at heterogeneous ages and totalling 69782 weight observations. We apply to this dataset the DA and the LA methods presented in Sections 4, 5, and 6 to the cases of random α, random β, and random both α and β, respectively.

Table 7 presents, for the real dataset, the estimates obtained through the LA and the DA methods for the case where both parameters are random. The true parameter values are not known for this dataset.

Table 7. Results for the real dataset. Maximum likelihood estimates and corresponding 95% asymptotic confidence bands for the DA and LA methods for the case where both parameters α and β are assumed random. Instead of showing the results for the modified weight μ, we show them for the weight $A = h^{-1}(\mu)$.

	random (α, β)	
	DA method	**LA method**
A	709.76 ± 9.70	688.70 ± 9.23
θ	$< 0.0001 \pm 0.0108$	$< 0.0001 \pm 0.0045$
λ	1.3141 ± 0.0142	1.3741 ± 0.0163
ω	0.1769 ± 0.0055	0.2264 ± 0.0082
σ	0.2977 ± 0.0017	0.2910 ± 0.0018

Observing the results of Table 7, both methods present similar estimates of the parameters θ and σ, allowing us to conclude that when both parameters are considered random the methods were unable to identify a random effect on the weight at maturity of the animal's, identifying only a random effect on the animal's growth rate. The LA method shows a higher estimate of the animals growth rate mean value (λ) and standard deviation (ω). The estimate of the weight at maturity of the animal is higher for the DA method. The 95% asymptotic confidence bands are wider in the LA method for the parameters related to animal growth rate and to σ, and are wider in the DA method for the parameters related to the weight at maturity of the animal.

Table 8 presents, for the real dataset, the estimates obtained for the random α case through the LA method and the DA method, and the estimates obtained for the random β case through the LA and the DA methods. Observing Table 8 for the random α case, we can conclude that, despite having a large dataset with very heterogeneous data, the estimates of all parameters through the "Exact method"/LA method and the DA method are similar, except for a slight difference observed in θ, the standard deviation of the random parameter α. The approximate confidence intervals based on the empirical Fisher's information matrix are also very similar.

Table 8. Results for the real dataset. Maximum likelihood estimates and corresponding 95% asymptotic confidence bands using the "Exact method"/LA method and the DA method for the case where only α is assumed random, and the LA and DA methods for the case where only β is assumed random. For the cases of random α, instead of showing the results for the modified weight μ, we show them for the weight $A = h^{-1}(\mu)$. For the case of random β, instead of showing the results for the modified weight α, we show them for the weight $A = h^{-1}(\alpha)$.

	random α		**random β**	
	DA method	**Exact/LA method**	**DA method**	**LA method**
A	724.97 ± 10.54	726.99 ± 10.74	709.76 ±9.70	687.24 ± 9.16
θ	0.1091 ± 0.0089	0.1428 ± 0.0092	–	–
λ	1.2959 ± 0.0141	1.3002 ± 0.0143	1.3141 ± 0.0142	1.3764 ± 0.0162
ω	–	–	0.1769 ± 0.0055	0.2259 ± 0.0082
σ	0.3039 ± 0.0018	0.3010 ± 0.0019	0.2977 ± 0.0017	0.2911 ± 0.0018

Since the DA and LA methods with α and β both random indicate that α should not be considered random because the estimate of θ is almost zero, the appropriate action to improve the estimates of the other parameters is to apply the DA and LA methods with only β considered to vary randomly. This is done on the right side of Table 8 and the results are consistent with those of Table 7 since the parameter estimates and their approximate confidence intervals are indistinguishable from the ones obtained in Table 7.

However, for the case where only α is assumed random (left side of Table 8), the estimates of common parameters for the DA method are practically the same as the ones obtained in Table 7, except for the mean and standard deviation of the random parameter α. Surprisingly, when both parameters are assumed random, the θ estimate is close to zero, while, when just the parameter α is assumed random, the estimate is quite different from zero. However, for the LA method the difference in estimates is presented in all parameters. This effect may be due to parameters α and β being correlated in the real life dataset, which would impact the correctness of the estimation of their standard deviations θ and ω when one forces α and β to be independent as assumed on Table 7. The effect may also be due to the fact that on the left side of Table 8 we may be wrongly forcing α to be random and β to be fixed, thus forcing the methods to compensate the variability of β with a false variability on α. Whatever the reason, further research is needed to extend the DA method to the case where the random parameters are correlated.

The possible correlation explanation for this effect is worth considering since, in the case of the simulated dataset with α and β both really random and really independent, the same exercise (not shown) of using the DA and LA methods with the wrong assumption, that α is the only random parameter, shows no such effect, i.e., the estimates of the common parameters are in line with those obtained in Table 5 (which is based on the now correct assumption that α and β are both random and independent). The same is true if we use the DA and LA methods with the wrong assumption that β is the only random parameter. However, this is not the case if we use the available R package for the same simulated dataset (we cannot use it for the real dataset) since, when using the package with the

wrong assumption of only one parameter being random, parameter estimates differ noticeably.

For the real dataset we also present visual diagnostics. To obtain information on the distribution of model predictions, we have simulated 1000 trajectories (1000 simulated animals), with monthly observations from birth to 2.5 years of age, using the Gompertz SDE mixed model with random α. We have used as "true" parameter values the ones obtained in Table 8 through the DA method with random α. For each monthly age, we therefore have a sample of 1000 values for which we have obtained the empirical 2.5%, 50% (median), and 97.5% quantiles. The left side of Figure 1 shows the corresponding quantile curves as a function of age. For a visual check, we would like to superimpose on the same Figure the real observed trajectories of the real 10843 animals, but that produces a confusing cloud where individual trajectories are indistinguishable. So, we display instead just 100 real animal trajectories chosen at random from the 10843 available trajectories. A similar procedure, now for the Gompertz SDE mixed model with random β, was also performed and shown on the right side of Figure 1. We can observe that for both cases (random α and random β) the real observed weights are almost all contained in the empirical 95% interval. The plot for the case of both α and β random is not shown since for our application to real cattle weight data the random effect on α was not identified as significant and so the case where both parameters are considered random is reduced to the β random case.

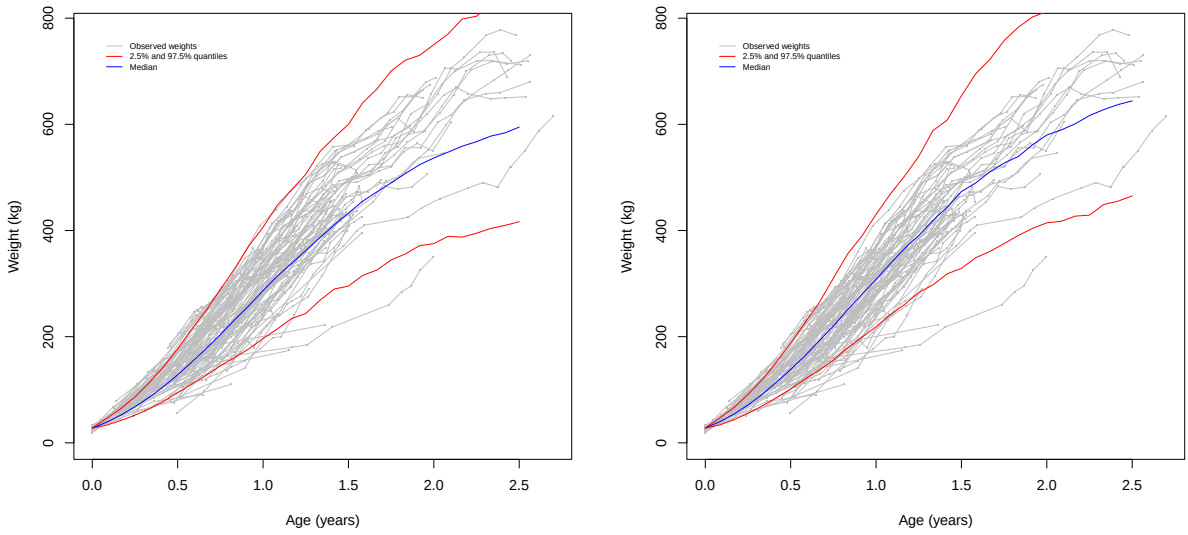

Figure 1. Real observed weight trajectories (in grey) of 100 animals (randomly chosen among the 10843 available real trajectories) and estimated 2.5%, 50% (median), and 97.5% quantile curves for the predictions of the Gompertz SDE mixed model with random α (on the left) or with random β (on the right).

The findings presented in Section 7 concerning both simulated and real datasets highlight the good performance of the DA method. Notably, the DA method exhibits versatility by effortlessly handling large datasets with heterogeneous ages of observation. Furthermore, its efficiency surpasses (except obviously for the random α case) that of the LA method. This is due to the fact that, at each step of the numerical maximization of the likelihood, we use in the DA method a closed-form expression

for the approximate likelihood, while in the LA method the approximate likelihood depends on an intermediate numerical maximization procedure of one or two variables.

In scenarios where the estimation of the parameter α is considered a random variable, the computational times of the LA method were comparable to those of the DA method. However, when focusing on estimating only the parameter β as a random variable, the LA method required 50% more time than the DA method. Notably, when both α and β were treated as random variables, the LA method needed ten times more computational time than the DA method.

8. Conclusions

We have studied a general class of SDE mixed models to describe individual growth in a randomly varying environment with a real application to the weight of male Mertolengo cattle. A general model can be written as a variant of the Ornstein-Uhlenbeck model, where the parameter α (asymptotic modified weight) and/or the growth parameter β are assumed to be random. For illustration purposes, we have worked with the logarithm of the weight, the transformation that is revealed to be the most appropriate for the cattle weight data, leading to the stochastic Gompertz model. Our interest in using SDE mixed models comes from the reasonable idea that model parameters may vary from animal to animal, which, for instance, occurs due to different individual characteristics of the animals such as genetic differences.

We apply the maximum likelihood estimation method to obtain the parameter estimates of the SDE mixed models. For this type of model, in most cases it is not possible to obtain a closed-form expression for the likelihood function and approximation methods are used to overcome this problem. In this paper, we compare approximation methods used to obtain simpler approximate expressions for the integral involved in the likelihood function. We study the DA method and compare it with the LA method and with the *mixedsde* R package. The DA method is a new approximation method we have developed, while the LA method is an already known approach that we have implemented numerically, including for the case of previously unconsidered datasets. where animals have different age vectors of observations. To illustrate and compare the performances of the DA and LA methods and of the existing *mixedsde* R package, which is based on different numerical approximation methods for the likelihood function, we have to use data compatible with the *mixedsde* package requirement of having observation times equidistant and common to all animals. So, for that comparative study, we have used simulated datasets satisfying that requirement. The DA and LA methods are presented for the case where only α is considered random, for the case where only β is considered random, and for the case where both parameters α and β are considered random and independent. The Gaussian distribution was assumed for the three situations. For the specific case where only the asymptotic modified size α is assumed random, it is possible to obtain an exact expression for the likelihood function and we have used the exact maximum likelihood estimates as a benchmark to evaluate the results obtained by the DA method with random α and by the *mixedsde* R package. We have also proven that, for this case, the LA method leads to the exact expression of the likelihood function.

The results for the simulated datasets show a very good performance of the proposed approximate methods, having similar estimates for some parameters and often outperforming the existing methods for the remaining parameters.

Since the DA method, a faster method than the LA method, is a new approximation method, we have

conducted a bias analysis and an evaluation of the quality of the standard deviation estimates obtained from the empirical Fisher information matrix (which is not provided by the *mixedsde* R package). We show that for all three cases, not only the estimates have a low bias but also the estimated standard deviations obtained by the empirical Fisher information matrix are quite reliable.

In the literature, for this type of models, it is typical to see the methods developed under the assumption of having a unique (sometimes evenly spaced) age vector of observations common to all individuals and this is also a restriction in the R package available for the estimation of the parameters of SDE mixed models. The DA method and the LA method expressions presented in this paper have the advantage of not requiring such restrictions and so they can be used in real situations where each animals size (weight, volume, height, or length) is measured at unique age instants that need not to be evenly spaced. We illustrate the results also with real data, the weight of Mertolengo cattle males from a large and heterogeneous sample, and concluded that either the DA method or the LA method are very good alternatives that have the advantage to always result in simpler and closed-form approximate expressions for the likelihood function.

However, further research is needed to allow for consideration the existence of correlation in the case of both parameters being random, since the correlation effect can impact the estimates of the standard deviation of the distribution of the animal's weight at maturity. Besides this issue, as future work, we also intend to consider distributions other than the Gaussian distribution for the random effects α and β.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors belong to the research center CIMA (Centro de Investigação em Matemática e Aplicações, Universidade de Évora), supported by FCT (Fundação para a Ciência e a Tecnologia, Portugal), project UID/MAT/04674/2020, https://doi.org/10.54499/UIDB/04674/2020. This work was developed within the Operational Group PDR2020-1.0.1-FEADER-031130 - Go BovMais - Productivity improvement in the system of bovine raising for meat, funded by PDR 2020. We are grateful to ACBM and José Pais (ACBM head engineer) for providing the data and for continuous support. We thank to RuralBit for the help in extracting the data from Genpro database.

Data and code availability

The real dataset of Mertolengo cattle weight data was extracted from the proprietary database GENPRO and was gently provided under confidentiality by ACBM, a Mertolengo cattle producers association. Simulated datasets, as well as the R code for the recently proposed delta approximation method are available at https://github.com/patfilipe/DAmethodSDEM

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. P.A. Filipe, C.A. Braumann, N.M. Brites, C.J. Roquete, Prediction for individual growth in a random environment, in *Recent Developments in Modeling and Applications in Statistics* (eds. P. Oliveira, M. da Graça Temido, C. Henriques, M. Vichi), Springer, Berlin, Heidelberg, (2013), 193–201. https://doi.org/10.1007/978-3-642-32419-2_20
2. O. Garcia, A stochastic differential equation model for the height growth of forest stands, *Biometrics*, **39(4)** (1983), 1059–1072. https://doi.org/10.2307/2531339
3. N.T. Jamba, G. Jacinto, P.A. Filipe, C.A. Braumann, Likelihood function through the delta approximation in mixed sde models, *Mathematics*, **10(3)**, 385 (2022). https://doi.org/10.3390/math10030385
4. O. Vasicek, An equilibrium characterization of the term structure, *Journal of Financial Economics*, **5(2)** (1977), 177–188. https://doi.org/10.1016/0304-405X(77)90016-2
5. P.A. Filipe, C.A. Braumann, N.M. Brites, C.J. Roquete, Modelling animal growth in random environments: an application using nonparametric estimation, *Biometrical Journal*, **52(5)** (2010), 653–666. https://doi.org/10.1002/bimj.200900273
6. P.A. Filipe, C.A. Braumann, C.J. Roquete, Multiphasic individual growth models in random environments, *Methodology and Computing in Applied Probability*, **14** (2012), 49–56. https://doi.org/10.1007/s11009-010-9172-0
7. G. Jacinto, P.A. Filipe, C.A. Braumann, Profit optimization of cattle growth with variable prices, *Methodology and Computing in Applied Probability*, **24** (2022a), 1917–1952. https://doi.org/10.1007/s11009-021-09889-z
8. G. Jacinto, P.A. Filipe, C.A. Braumann, Weighted maximum likelihood estimation for individual growth models, *Optimization*, **71(11)** (2022b), 3295–3311. https://doi.org/10.1080/02331934.2022.2075745
9. U. Picchini, S. Ditlevsen, Practical estimation of high dimensional stochastic differential mixed-effects models, *Computational Statistics & Data Analysis*, **55(3)** (2011), 1426–1444. https://doi.org/10.1016/j.csda.2010.10.003
10. R. Wolfinger, Laplace's approximation for nonlinear mixed models, *Biometrika*, **80(4)** (1993), 791–795. https://doi.org/10.2307/2336870
11. M. Delattre, A review on asymptotic inference in stochastic differential equations with mixed effects, *Japanese Journal of Statistics and Data Science*, **4** (2021), 543–575. https://doi.org/10.1007/s42081-021-00105-3
12. I. Botha, R. Kohn, C. Drovandi, Particle methods for stochastic differential equation mixed effects models, *Bayesian Analysis*, **16(2)** (2021), 575–609. https://doi.org/10.1214/20-BA1216

13. S. Wiqvist, A. Golightly, A.T. McLean, U. Picchini, Efficient inference for stochastic differential equation mixed-effects models using correlated particle pseudo-marginal algorithms, *Computational Statistics & Data Analysis*, **157** (2021), 107151. https://doi.org/10.1016/j.csda.2020.107151

14. M.G. Ruse, A. Samson, S. Ditlevsen, Inference for biomedical data by using diffusion models with covariates and mixed effects, *Journal of the Royal Statistical Society Series C: Applied Statistics*, **69(1)** (2020), 167–193. https://doi.org/10.1111/rssc.12386

15. R.V. Overgaard, N. Jonsson, C.W. Tornøe, H. Madsen, Non-linear mixed-effects models with stochastic differential equations: Implementation of an estimation algorithm, *Journal of Pharmacokinetics and Pharmacodynamics*, **32** (2005), 85–107. https://doi.org/10.1007/s10928-005-2104-x

16. U. Picchini, A.D. Gaetano, S. Ditlevsen, Stochastic differential mixed-effects models, *Scandinavian Journal of Statistics*, **37(1)** (2010), 67–90. http://www.jstor.org/stable/41000916

17. M. Delattre, C. Dion, Msdeparest: Parametric estimation in mixed-effects stochastic differential equations, *R package version 1.7*, (2017). https://CRAN.R-project.org/package=MsdeParEst

18. C. Dion, A. Samson A, Hermann S (2018) mixedsde: Estimation methods for stochastic differential mixed effects models, *R package version 5.0*, (2018). https://CRAN.R-project.org/package=mixedsde

19. S. Klim, S.B. Mortensen, N.R. Kristensen, R.V. Overgaard, H. Madsen, Population stochastic modelling (psm)-an r package for mixed-effects models based on stochastic differential equations. *Computer Methods and Programs in Biomedicine*, **94(3)** (2009), 279–289. https://doi.org/10.1016/j.cmpb.2009.02.001

20. M. Delattre, V. Genon-Catalot, A. Samson, Maximum likelihood estimation for stochastic differential equations with random effects, *Scandinavian Journal of Statistics*, **40(2)** (2013), 322–343. https://doi.org/10.1111/j.1467-9469.2012.00813.x

21. C.A. Braumann, *Introduction to Stochastic Differential Equations with Applications to Modelling in Biology and Finance*, John Wiley & Sons, (2019).

22. J. Leander, M. Jirstrand, U.G. Eriksson, R. Palmér, A stochastic mixed effects model to assess treatment effects and fluctuations in home-measured peak expiratory flow and the association with exacerbation risk in asthma, *CPT: Pharmacometrics & Systems Pharmacology*, **11(2)** (2022), 212–224. https://doi.org/10.1002/PSP4.12748

23. K. Wang, L. Marciani, G.L. Amidon, D.E. Smith, D. Sun, Stochastic Differential Equation-based Mixed Effects Model of the Fluid Volume in the Fasted Stomach in Healthy Adult Human, *The AAPS J*, **25(5)** (2023), 76. https://doi.org/10.1208/s12248-023-00840-3

24. U. Picchini, S. Ditlevsen, A. De Gaetano, Maximum likelihood estimation of a time-inhomogeneous stochastic differential model of glucose dynamics, *Mathematical Medicine and Biology: A Journal of the IMA*, **25(2)** (2008), 141–155. https://doi.org/10.1093/imammb/dqn011

25. U. Picchini, *Stochastic Differential Models with Applications to Physiology*, PhD thesis, University of Rome, Rome, Italy, (2006).

AIMS Press

© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0)