Universidade de Évora - Instituto de Investigação e Formação Avançada

Programa de Doutoramento em Matemática

Tese de Doutoramento

Mixed Models for Individual Growth in a Random Environment

Nelson Tchingui Jamba

Orientador(es) | Gonçalo João Jacinto
Patrícia A. Filipe

Évora 2024

Universidade de Évora - Instituto de Investigação e Formação Avançada

Programa de Doutoramento em Matemática

Tese de Doutoramento

Mixed Models for Individual Growth in a Random Environment

Nelson Tchingui Jamba

Orientador(es) | Gonçalo João Jacinto
Patrícia A. Filipe

Évora 2024

A tese de doutoramento foi objeto de apreciação e discussão pública pelo seguinte júri nomeado pelo Diretor do Instituto de Investigação e Formação Avançada:

Presidente	Carlos Correia Ramos (Universidade de Évora)
Vogais	Clara Sofia Gomes Carlos (Instituto Politécnico de Setúbal - Escola Superior de Tecnologia do Barreiro) Cristiana João Soares da Silva (ISCTE - Instituto Universitário de Lisboa) Célia Maria Pinto Nunes (Universidade da Beira Interior) Dulce Maria de Oliveira Gomes (Universidade de Évora) Patrícia A. Filipe (Universidade de Évora) (Orientador)

Évora 2024

I dedicate this thesis to my loving sons Celson, Neusa and Élio, for the affection and the jokes,
to my mother Regina Chambula (in memory), to my father Fernando Jamba to my uncle
Maurício Cangongo (in memory) and to my uncle Eugénio Sameti.

Acknowledgments

To professors Patrícia Filipe, Gonçalo Jacinto and Carlos Braumann, my advisors, I thank you for your guidance, suggestions, works and rigor, I would never have gone it alone;

To my father Fernando Jamba and my mother Regina Chambula (in memory), many thanks for everything;

To my uncles Maurício Cangongo (in memory) and Eugénio Samety, my thanks for the support they gave me in this academic path, when I needed it most;

To all my family in general, thank you for your support and everyone the values they taught me;

To my great brother that I gained during this important phase of my life, Manuel de Paiva (in memory), thank you very much for help and brotherhood;

To my friends namely: Cesaltina César, Anatilde Timóteo, Augusto Rasga, Emílio Tchikuala, Reginaldo Silva, Robert Sousa, Zacarias Pedro, Abel Zacarias, Nilton Ávido, Armando Coholina, Francisco Chimuco, Francisco Mulangui, Sebastião Mulangui, António Cambanda, Cláudio Jamba, Delfim Jamba, Amauro Chingongo, Higino Chicala, Paulino Bolela, Tiago Nambalo, Longa Paquete, Evaristo Das Mangas, Mwene Vunongue, Cristina Firmino, Inácio Firmino, Érika Silva, Rosária Bunga, Teresa Firmino e Clara Hidipo, thank you very much for the various types of help, from scientific, financial, emotional, hosting, computer loan and dissemination of the work;

To all my friends, a big thank you;

To Lyceum No. 918, Chiange, Gambos, on behalf of the Municipal administration, thank you for service permits for the design of this work, thanks ISCED-Huíla and thanks ISPS-Huíla for support;

To the research center CIMA (Centro de Investigação em Matemática e Aplicações da Universidade de Évora), supported by FCT, project UID/MAT/04674/2020. To Associação de Criadores de Bovinos Mertolengos (ACBM) and José Pais (ACBM head engineer) for providing the data and for continuous support. To RuralBit for the help in extracting the data from Genpro database.

Resumo

Modelos Mistos de Crescimento Individual em Ambiente Aleatório

A utilização de um modelo misto de equações diferenciais estocásticas é uma abordagem adequada para modelar o crescimento individual, sujeito a flutuações aleatórias do ambiente, quando existem evidências de variabilidade nos parâmetros do modelo. Um dos métodos mais utilizado para estimar os parâmetros do modelo é o método da máxima verosimilhança. Na maioria dos casos, ao considerar como aleatórios os parâmetros do modelo, torna-se difícil obter expressões em forma fechada para a função de verosimilhança, o que originou a adoção de métodos de aproximação a esta função. Foi desenvolvido um novo método de aproximação, baseado no conhecido método delta da Estatística clássica, que designamos por método de aproximação de delta. Este método foi desenvolvido para os casos de aleatoriedade no tamanho médio assintótico e/ou na taxa de crescimento e comparado com métodos já existentes, em particular, com o método de Laplace e outros métodos disponíveis em packages do software R. Os métodos foram ilustrados com base em dados reais de peso de bovinos da raça Mertolenga. De forma a poder comparar a performance do método de delta com os outros métodos foi necessário adaptar e desenvolver as expressões do método de Laplace e recorrer a bases de dados simuladas para comparar com métodos já existentes.

Como as características de cada animal podem ser relevantes para estimar a sua curva de crescimento, foi desenvolvido um modelo misto de equações diferenciais estocásticas, para o caso em que o parâmetro que representa o tamanho médio assimptótico do animal, não só é considerado aleatório como pode ser escrito como uma função linear dos valores genéticos do próprio animal, evidenciando desta forma a componente genética de cada animal na sua própria curva de crescimento. Os resultados dos modelos mistos com a incorporação dos valores genéticos revelaram uma melhoria na curva de crescimento do animal estimada.

Abstract

Mixed Models for Individual Growth in a Random Environment

The use of stochastic differential equations mixed models is an appropriate approach to model individual growth, subject to random fluctuations of the environment, when there is evidence of variability in the model parameters. The maximum likelihood method is one of the most used methods to estimate the model parameters. In most cases, when considering the model parameters as random, it becomes difficult to obtain closed-form expressions for the likelihood function, leading to the adoption of approximation methods. A new approximation method was developed, based on the known delta method of classical Statistics, which we call the delta approximation method. This method was developed for the cases where the parameters related to the asymptotic average size and/or growth rate are random and compared with existing methods, in particular, with the Laplace method and other methods available in packages of R software. The methods were illustrated based on real data on the weight of cattle of the Mertolenga breed. To compare the performance of the delta method with the other methods, it was necessary to adapt and develop the expressions of the Laplace method and resort to simulated datasets to compare with already existent methods.

As the characteristics of each animal can be relevant to estimating its growth curve, a stochastic differential equations mixed model was developed, for the case where the parameter that represents the asymptotic average size of the animal, not only is considered random but can be written as a linear function of the genetic values of the animal itself, thus evidencing the genetic component of each animal in its own growth curve. The results of the mixed models with the incorporation of genetic values revealed an improvement in the estimated growth curve of the animal.

Contents

List of Tables

CHAPTER 1

Introduction

Stochastic differential equations (SDE) are mathematical equations used to model systems that involve random or stochastic processes. They combine ordinary differential equations (ODE) with stochastic terms, typically represented as a Wiener process or Brownian motion. SDE can be used to model animal growth because growth processes in living organisms can be influenced by numerous random factors and uncertainties. These factors may include variations in the environment, food availability, genetic traits, and other stochastic influences. Using SDE allows researchers to incorporate randomness and variability into the models, making them more realistic and reflective of the inherent uncertainties observed in biological systems.

When modeling individual growth, many studies use deterministic models and fit them with classic regression methods. These methods assume that the errors are independent, which is only true when the errors are observational. They are inappropriate if the deviations between observations are due to random changes on growth rates resulting from environmental random fluctuations. SDE are built to model growth incorporating the dynamics of the growth process and the effect of the random fluctuations. Our goal is to apply SDE models to describe the dynamics of the evolution of bovine weight. To describe the growth of M animals, a general SDE model can be written as follows,

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

where $Y_i(t) = h(X_i(t))$ is the modified weight by a transformation function h, a known monotonous continuously differentiable function, of the real weight $X_i(t)$ of the animal i at age t, and $y_{i,0} = h(x_{i,0})$, where $x_{i,0}$ is the assumed known size of animal i at an initial age of observation $t_{i,0}$. The parameter $\beta > 0$ is the growth coefficient and α is the mean asymptotic modified size towards which the mean modified size converges as $t \to +\infty$; we denote by $A = h^{-1}(\alpha)$ the corresponding real asymptotic size. The intensity of the effect of environmental random fluctuations on growth is measured by the parameter $\sigma > 0$, being $W_i(t)$ $(i = 1, \ldots, M)$ independent realizations of the standard Wiener process. Specific choices of the function h lead to stochastic versions of some models commonly used to describe growth, such as the monomolecular model for $h(x) = x$, the Bertalanffy-Richards model for $h(x) = x^c$ $(c > 0)$, the Gompertz model for $h(x) = \ln x$ and the logistic model for $h(x) = x^{-1}$, but the theoretical treatment is valid for any monotonous C^1 function h. Here, we will use $h(x) = \ln x$, corresponding to the stochastic Gompertz model. The SDE model described in (1.1) is also called the *Vasicek model* in the context of interest rate dynamics [**1**], was proposed by *Oldrich Vasicek* in 1977 and is widely used in the field of quantitative finance, where in this context, the *Vasicek model* provides a framework for analyzing interest rate derivatives and pricing fixed income securities. It has been used to estimate various interest rate parameters, such as the risk-neutral probability of default and the term structure of interest rates. The field of models for organisms growth has various applications of different models of SDE. For example, [**2**] shows how to model the growth dynamic of trees with fixed parameters that are estimated from data. Another example is [**3**], where they compare evolutionary patterns of a continuous trait on a phylogeny using Brownian motion or Ornstein-Uhlenbeck (OU) processes, among others.

The stochastic differential equation non-mixed (NMSDE) models such as (1.1), are often used to model individual growth, but when using real data an evidence of variability in the fixed effect arise, thus transforming the NMSDE model into a stochastic differential equation mixed (MSDE) model. MSDE models, also known as stochastic mixed-effects models or random effects models, combine the concepts of SDE and mixed-effects models. These models are used to analyze longitudinal or repeated measures data where both fixed effects (population-level effects) and random effects (subject-specific effects) are present, along with stochastic components. In a MSDE the underlying dynamic process is described by a SDE, which captures the evolution of the response variable over time. The model incorporates both fixed effects, which represent the average behavior of the population, and random effects, which capture the individual variations among subjects.

MSDE models find applications in various fields due to their ability to capture both fixed effects and random effects, such as pharmacokinetics and pharmacodynamics modeling [**4**], where they can capture both the individual-level variability and the dynamic behavior of drug concentrations in the body over time, whose extension for biomedical data is presented in [**5**,**6**]; pharmacological studies [**7**] constitute an attractive class of stochastical models for biomedical data; biology and ecology [**8**] models are employed to analyze biological systems with both deterministic growth patterns and random environmental effects; in option pricing [**9**], presents a mixed stochastic differential equation containing both Wiener process and a Hölder continuous process, where a stochastic viability theorem is proven, in [**10**] the asymptotic properties of the estimators of MSDE are studied. This approach was proposed to cope the problems arising when the standard growth functions have a monotone increasing growth and can fail to model unexpected changes in growth rates.

Estimating the parameters in MSDE models typically involves maximizing the likelihood function. However, the estimation process can be complex and in very few cases there exists a closed form expression for the maximum likelihood function. Therefore, estimation and inference for MSDE models involve fitting the model to the observed data using approximation methods or numerical approaches. Some numerical approaches have been implemented in R packages such as *mixedsde* and *MsdeParEst* [**11**, **12**]. These packages have a restriction of using a single time vector for all individuals not being possible to use them when individuals have different time vectors or even in the more complex structure where they have different time vectors and non-equidistant observations. To utilize MSDE in real-world scenarios where different individuals may possess distinct time vectors, it is essential to propose new techniques or adapt existing ones. In this context, we have developed the delta approximation (DA) method, an extension of the classic delta method, tailored for scenarios where the parameters α and/or β can be treated as random variables. Furthermore, we have derived all necessary formulas from the well-established Laplace approximation (LA) method for utilization in real data settings. These approximation approaches allow estimating the parameters of MSDE with vectors of non-equidistant and different observations between animals, the advantage of the DA method over the LA method is its ability to provide a simpler presentation and a closed-form expression for approximating the likelihood function. We present the application to cattle weight real data, comparing the performance of the different methods. We also present the application to simulated data to observe the effect of having diferent datasets sizes and different age vectors on the different estimation methods.

The genetic component appears as a natural addition when we think of the variability between individuals. The inclusion of genetic values in the maximum likelihood function depends on the specific context and model being used, where for our case, we consider the random α

case as a linear function of the genetic values. Thus, the objective is not only to explain part of the variation in the parameters that have not been explained before, but also to identify those characteristics of the animals that contribute most to their growth.

This thesis is organized as follows, in chapter 2 we present the NMSDE models, their main properties and the use of the maximum likelihood estimation method to estimate the models parameters. We also present the extension to MSDE model, for the random α case and/or for the random β case and the methodological developments and expressions for the DA method and the LA method are obtained. In chapter 3 we present the extension to MSDE models, where we focus on the development of the DA method applied to the likelihood function for the case where both parameters α and β are independent Gaussian distributed random variables. In chapter 4, we introduce the particular MSDE models, for the random α case, with the inclusion of the genetics values. For this purpose, we write the random effect of transformed weight at maturity as a linear function of genetic values. In this way it was possible to explain part of the variation of the parameters that were not explained before the inclusion of the genetic values. Chapter 5 presents the general conclusions. The developed R codes can be found in the Appendix.

This thesis is a collection of three research papers published in peer-reviewed indexed scientific journals during my PhD research. The papers forming chapters (2), (3) and (4) of this thesis are mostly presented in the form they were published.

CHAPTER 2

Estimation for stochastic differential equation mixed models using approximation methods

Nelson T. Jamba
Gonçalo Jacinto
Patrícia A. Filipe
Carlos A. Braumann

AIMS Mathematics, 9(4): 7866–7894, 2024

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.

2.1. 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 [**2**, **13**, **14**]. 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 [**1**], as

$$dY_i(t) = \beta\,(\alpha - Y_i(t))\; dt + \sigma\, dW_i(t),\; Y_i(t_{i,0}) = y_{i,0}, \quad i = 1, \ldots, M, \tag{2.1}$$

where $Y_i(t) = h(X_i(t))$ is a modified size 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 [**13**, **15–18**]. 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 ([**19**, **20**]), 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 [**14**] 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 of the 10843 Mertolengo cattle males with recorded genetic information on the database of the Associação de Criadores de Bovinos Mertolengos (ACBM), which registers 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 [**10**] with the discussion of asymptotic properties of estimators and presentation of several applications. For instance, in [**5**] and [**6**] the authors apply SDE mixed effects models to tumor growth dynamics using fully Bayesian inference through the pseudo-marginal methods. In [**7**] 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 [**21**], [**19**], and [**22**]. Some methodologies were already implemented in R packages (see, for example, [**23**], [**24**], and [**25**]) 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 [**24**, **26**].

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 2.2, 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 2.3, we present the SDE mixed model. In Sections 2.4 and 2.5, 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 2.6, the same approach is presented for the case where both α and β are random. In Section 2.7, 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 2.8, where the main conclusions are presented.

2.2. 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 ($i = 1, \dots, M$). 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 (2.1).

It can be seen, for instance in [**27**], that the solution of (2.1) is given, for $t \geq t_{i,0}$, by

$$Y_i(t) = \alpha - (\alpha - y_{i,0})e^{-\beta(t-t_{i,0})} + \sigma e^{-\beta t} \int_{t_{i,0}}^{t} e^{\beta s} dW_i(s), \quad i = 1, \dots, M, \tag{2.2}$$

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 (2.1). For each individual i ($i = 1, \dots, M$), denote by $\mathbf{t_i} = (t_{i,0}, t_{i,1}, \dots, t_{i,n_i})$ 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:

$$Y_{i,j}|Y_{i,j-1} = y_{i,j-1} \sim N\left(\alpha + \left(y_{i,j-1} - \alpha\right) E_{i,j}^{\beta}\, ,\ \frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})\right). \tag{2.3}$$

To estimate the parameter vector $\mathbf{p} = (\alpha, \beta, \sigma)$, the maximum likelihood estimation method is usually applied, see, for instance, [**15**]. From (2.3), using the fact that $Y_i(t)$ is a Markov process, we know that, given $Y_{i,0} = y_{i,0}$ (assumed known), the joint probability density function (p.d.f.)

for animal i is

$$
\begin{aligned}
p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha,\beta,\sigma) &= \prod_{j=1}^{n_i} \frac{\exp\left(-\frac{1}{2}\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}\right)}{\sqrt{2\pi\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}} \qquad (2.4) \\
&= \frac{\exp\left(-\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}\right)}{\prod_{j=1}^{n_i}\sqrt{2\pi\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}}, \quad i=1,\ldots,M,
\end{aligned}
$$

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

$$L(\alpha,\beta,\sigma) = \prod_{i=1}^{M} p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha,\beta,\sigma). \qquad (2.5)$$

The maximum likelihood estimate $\hat{\mathbf{p}}$ of the parameter vector $\mathbf{p}$ is obtained by maximization of (2.5). 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.

2.3. Stochastic differential equation for mixed models

SDE mixed models are used to model growth, with applications in various fields, such as animal growth and pharmacokinetics, among others [**7**, **14**, **21**, **28**, **29**]. For the type of models we present here, in [**14**, **19**, **22**, **30**], 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

$$L(\epsilon,\mathbf{\Psi}) = \prod_{\mathbf{i=1}}^{\mathbf{M}} p_{\mathbf{Y_i}}(\mathbf{y}_i|\epsilon,\mathbf{\Psi}) = \prod_{\mathbf{i=1}}^{\mathbf{M}} \int_{-\infty}^{+\infty} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{b}_i,\epsilon)p(\mathbf{b}_i|\mathbf{\Psi})\,d\mathbf{b}_i, \qquad (2.6)$$

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 [**31**] 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 ($t_j - t_{j-1} \equiv \delta$, $j = 1, \ldots, n$), which considerably simplifies the expressions, and in [**14**] and [**21**] for the general situation.

However, the integral in (2.6) 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, [**14**] proposed the delta approximation method (DA method) to approximate the integral in (2.6). 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 [**14**], 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 [**20**], 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.

2.4. Mixed models with random asymptotic average size

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

$$p_{\mathbf{Y_i}}(\mathbf{y}_i|\mu, \theta, \beta, \sigma) = \int_{-\infty}^{+\infty} p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha_i, \beta, \sigma) p(\alpha_i|\mu, \theta)\, d\alpha_i, \tag{2.7}$$

with $p_{Y_i}(\mathbf{y}_i|\alpha_i, \beta, \sigma)$ obtained by replacing α with α_i in (2.4) and

$$p(\alpha_i|\mu, \theta) = \frac{1}{\sqrt{2\pi\theta^2}} e^{-\frac{(\alpha_i-\mu)^2}{2\theta^2}}. \tag{2.8}$$

Thus, from (2.6) and using (2.7) and (2.8), we get, after some simplifications, the likelihood function

$$\begin{aligned} L(\mu, \theta, \beta, \sigma) \;\; = \;\; & \prod_{i=1}^{M} \left(\frac{1}{\sqrt{2\pi\theta^2}} \prod_{j=1}^{n_i} \left(\frac{1}{\sqrt{\frac{\pi\sigma^2}{\beta}(1 - E_{i,j}^{2\beta})}} \right) \right. \\ & \left. \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2} \sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i) E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})} - \frac{(\alpha_i - \mu)^2}{2\theta^2} \right) d\alpha_i \right). \end{aligned} \tag{2.9}$$

It is shown in [**14**] that it is possible to explicitly compute the integral in (2.9) 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:

$$
\begin{aligned}
LL(\mu,\theta,\beta,\sigma) = \sum_{i=1}^{M}\Bigg(&-\frac{n_i}{2}\ln(2\pi)-\frac{n_i}{2}\ln\left(\frac{\sigma^2}{2\beta}\right)-\frac{1}{2}\sum_{j=1}^{n_i}\ln\left(1-E_{i,j}^{2\beta}\right)-\frac{1}{2}\ln(D_i)\\
&-\tfrac{\beta}{D_i\sigma^2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\mu-\left(y_{i,j-1}-\mu\right)E_{i,j}^{\beta}\right)^2}{1-E_{i,j}^{2\beta}}+\frac{2\beta^2\theta^2}{D_i\sigma^4}\left(\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-y_{i,j-1}E_{i,j}^{\beta}\right)}{1+E_{i,j}^{\beta}}\right)^2\\
&-\tfrac{\beta(D_i-1)}{D_i\sigma^2}\left(\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-y_{i,j-1}E_{i,j}^{\beta}\right)^2}{1-E_{i,j}^{2\beta}}\right)\Bigg),
\end{aligned}
\tag{2.10}
$$

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

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 (2.9) 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.

2.4.1. Likelihood function through the Laplace approximation method

Notice that the integrand function of the integral in (2.9) is the exponential of

$$
f_i(\alpha_i) = -\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}-\frac{(\alpha_i-\mu)^2}{2\theta^2}\,. \tag{2.11}
$$

To use the Laplace approximation method, we first need to obtain the value $\hat{\alpha}_i$ of α_i that maximizes (2.11). The function $f_i(\alpha_i)$ is a real twice continuously differentiable function with first and second order derivatives, respectively

$$
f_i'(\alpha_i) = -\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta}\right)\left(E_{i,j}^{\beta}-1\right)}{\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)}-\frac{\alpha_i-\mu}{\theta^2} \tag{2.12}
$$

and

$$
f_i''(\alpha_i) = -\sum_{j=1}^{n_i}\frac{\left(E_{i,j}^{\beta}-1\right)^2}{\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)}-\frac{1}{\theta^2}\,. \tag{2.13}
$$

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

$$
\hat{\alpha}_i \;=\; \left(\sum_{j=1}^{n_i}\frac{\left(E_{i,j}^{\beta}-1\right)^2}{\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)}+\frac{1}{\theta^2}\right)^{-1}\left(\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-y_{i,j-1}E_{i,j}^{\beta}\right)\left(1-E_{i,j}^{\beta}\right)}{\frac{\sigma^2}{2\beta}\left(1-E_{i,j}^{2\beta}\right)}+\frac{\mu}{\theta^2}\right) \tag{2.14}
$$

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

$$\begin{aligned}
\int_{-\infty}^{+\infty} \exp\left(f_i(\alpha_i)\right) d\alpha_i &\approx \int_{-\infty}^{+\infty} \exp\left(f_i(\hat{\alpha}_i) + \frac{1}{2} f_i''(\hat{\alpha}_i)(\alpha_i - \hat{\alpha}_i)^2\right) d\alpha_i \\
&= \exp\left(f_i(\hat{\alpha}_i)\right) \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2} \frac{(\alpha_i - \hat{\alpha}_i)^2}{-\left(f_i''(\hat{\alpha}_i)\right)^{-1}}\right) d\alpha_i \\
&= \exp\left(f_i(\hat{\alpha}_i)\right) \sqrt{\frac{2\pi}{\left|f_i''(\hat{\alpha}_i)\right|}} \,. \qquad (2.15)
\end{aligned}$$

The usual procedure in applying the Laplace approximation method would consist in replacing expressions (2.13) and (2.14) in (2.15) and then in (2.9) 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 (2.15) 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 (2.15) 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 (2.15) and, therefore, replacing (2.15) in (2.9) provides, not just an approximation of the log-likelihood function but rather the exact log-likelihood function, for which we have already provided expression (2.10). 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 (2.10) w.r.t. the parameters).

2.4.2. Likelihood function through the delta approximation method

A second approximation of the integral in (2.9) is obtained using the delta approximation method. For this approximation we need to rearrange the expression of the likelihood function as

$$\begin{aligned}
L(\mu, \theta, \beta, \sigma) = \prod_{i=1}^{M} \Bigg(\prod_{j=1}^{n_i} \left(\frac{1}{\sqrt{2\pi \frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}} \right) \\
\int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2} \sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i) E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})} \right) \frac{\exp\left(-\frac{(\alpha_i-\mu)^2}{2\theta^2}\right)}{\theta\sqrt{2\pi}} d\alpha_i \Bigg). \qquad (2.16)
\end{aligned}$$

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

$$\begin{aligned}
L(\mu, \theta, \beta, \sigma) &= \prod_{i=1}^{M} \left(\prod_{j=1}^{n_i} \left(\frac{1}{\sqrt{2\pi \frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}} \right) \int_{-\infty}^{+\infty} u_i(\alpha_i) \frac{1}{\theta\sqrt{2\pi}} \exp\left(-\frac{(\alpha_i - \mu)^2}{2\theta^2} \right) d\alpha_i \right) \\
&= \prod_{i=1}^{M} \left(\prod_{j=1}^{n_i} \left(\frac{1}{\sqrt{2\pi \frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}} \right) \mathbf{E}[u_i(\alpha_i)] \right), \qquad (2.17)
\end{aligned}$$

where

$$u_i(\alpha_i) = \exp\left(g_i(\alpha_i)\right) \qquad (2.18)$$

with

$$g_i(\alpha_i) = -\frac{1}{2}\sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})} . \tag{2.19}$$

Using the same idea of the delta method to approximate $\mathbf{E}[u_i(\alpha_i)]$, as in [**14**], 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

$$u_i(\alpha_i) = u_i(\mu + \theta\delta_i) \approx u_i(\mu) + \theta\delta_i u_i'(\mu) + \frac{1}{2}(\theta\delta_i)^2 u_i''(\mu) \tag{2.20}$$

and

$$\mathbf{E}[u_i(\alpha_i)] \approx u_i(\mu) + \frac{1}{2}\theta^2 u_i''(\mu), \tag{2.21}$$

with

$$u_i''(\mu) = u_i(\mu)\left[(g_i'(\mu))^2 + g_i''(\mu)\right], \tag{2.22}$$

where the first and second order derivatives are already presented in [**14**] and so we omit them here.

Replacing (2.21) in (2.17), we obtain

$$L(\mu, \theta, \beta, \sigma) \approx \prod_{i=1}^{M}\left(\prod_{j=1}^{n_i}\left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}}\right)\left(u_i(\mu) + \frac{1}{2}\theta^2 u_i''(\mu)\right)\right). \tag{2.23}$$

Now, replacing (2.18) and (2.22) in (2.23), we get the expression for the log-likelihood function

$$\begin{aligned} LL(\mu, \theta, \beta, \sigma) \approx \sum_{i=1}^{M}\Bigg(& -\frac{1}{2}\sum_{j=1}^{n_i} \ln\left(2\pi\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})\right) - \frac{1}{2}\sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \mu - (y_{i,j-1} - \mu)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})} \\ & + \ln\Bigg(1 + \frac{1}{2}\theta^2\Bigg(-\sum_{j=1}^{n_i} \frac{(1 - E_{i,j}^{\beta})^2}{\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})} \\ & + \left(\sum_{j=1}^{n_i} \frac{(1 - E_{i,j}^{\beta})(y_{i,j} - \mu - (y_{i,j-1} - \mu)E_{i,j}^{\beta})}{\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}\right)^2\Bigg)\Bigg)\Bigg). \end{aligned} \tag{2.24}$$

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

2.5. 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

$$p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha, \lambda, \omega, \sigma) = \int_{-\infty}^{+\infty} p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha, \beta_i, \sigma)p(\beta_i|\lambda, \omega)\, d\beta_i, \tag{2.25}$$

where $p_{Y_i}(\mathbf{y}_i|\alpha_i, \beta, \sigma)$ is obtained by replacing β with β_i in (2.4) and the p.d.f. of β_i is

$$p(\beta_i|\lambda, \omega) = \frac{1}{\sqrt{2\pi\omega^2}} e^{-\frac{(\beta_i - \lambda)^2}{2\omega^2}} . \tag{2.26}$$

Thus, from (2.6) and using (2.25), after some simplifications we obtain the likelihood function

$$
\begin{aligned}
L(\alpha, \lambda, \omega, \sigma) = \prod_{i=1}^{M} \Bigg(\frac{1}{\sqrt{2\pi\omega^2}} \prod_{j=1}^{n_i} \left(\frac{1}{\sqrt{2\pi}} \right) \\
\int_{-\infty}^{+\infty} \frac{\exp\left(-\frac{1}{2} \sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \alpha - (y_{i,j-1} - \alpha) E_{i,j}^{\beta_i} \right)^2}{\frac{\sigma^2}{2\beta_i} (1 - E_{i,j}^{2\beta_i})} \right)}{\exp\left(\frac{1}{2} \sum_{j=1}^{n_i} \ln\left(\frac{\sigma^2}{2\beta_i} (1 - E_{i,j}^{2\beta_i}) \right) \right)} \exp\left(-\frac{(\beta_i - \lambda)^2}{2\omega^2} \right) d\beta_i \Bigg).
\end{aligned}
\tag{2.27}
$$

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

2.5.1. Likelihood function through the Laplace approximation method

The integrand function in expression (2.27) can written as the exponential of the function

$$
f_i(\beta_i) = -\frac{1}{2} \sum_{j=1}^{n_i} \left(\frac{\left(y_{i,j} - \alpha - (y_{i,j-1} - \alpha) E_{i,j}^{\beta_i} \right)^2}{\frac{\sigma^2}{2\beta_i} (1 - E_{i,j}^{2\beta_i})} + \frac{1}{n_i} \frac{(\beta_i - \lambda)^2}{\omega^2} + \ln\left(\frac{\sigma^2}{2\beta_i} (1 - E_{i,j}^{2\beta_i}) \right) \right). \tag{2.28}
$$

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 (2.28) 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 (2.15), the Laplace approximation of the integral in (2.27) will be given by

$$
\exp\left(f_i(\hat{\beta}_i) \right) \sqrt{\frac{2\pi}{|f_i''(\hat{\beta}_i)|}}\,, \tag{2.29}
$$

with

$$
\begin{aligned}
f_i'(\beta_i) = \sum_{j=1}^{n_i} \Bigg(&-\frac{E_{i,j}^{2\beta_i} (t_{i,j} - t_{i,j-1})}{\left(1 - E_{i,j}^{2\beta_i} \right)} - \frac{\left(y_{i,j} - \alpha - (y_{i,j-1} - \alpha) E_{i,j}^{\beta_i} \right)^2}{\left(1 - E_{i,j}^{2\beta_i} \right) \sigma^2} \\
&- \frac{\beta_i - \lambda}{n_i \omega^2} + \frac{2\beta_i E_{i,j}^{2\beta_i} (t_{i,j} - t_{i,j-1}) \left(y_{i,j} - \alpha - (y_{i,j-1} - \alpha) E_{i,j}^{\beta_i} \right)^2}{\left(1 - E_{i,j}^{2\beta_i} \right)^2 \sigma^2} \\
&+ \frac{1}{2\beta_i} - \frac{2\beta_i E_{i,j}^{\beta_i} (t_{i,j} - t_{i,j-1}) (y_{i,j-1} - \alpha) \left(y_{i,j} - \alpha - (y_{i,j-1} - \alpha) E_{i,j}^{\beta_i} \right)}{\left(1 - E_{i,j}^{2\beta_i} \right) \sigma^2} \Bigg)
\end{aligned}
\tag{2.30}
$$

and

$$\begin{aligned}
f_i''(\beta_i) = \sum_{j=1}^{n_i} \Bigg(& \frac{2E_{i,j}^{4\beta_i}(t_{i,j}-t_{i,j-1})^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2} - \frac{2\beta_i E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})^2(y_{i,j-1}-\alpha)^2}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \\
& + \frac{2E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})^2}{\left(1-E_{i,j}^{2\beta_i}\right)} + \frac{4E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \\
& - \frac{1}{2\beta_i^2} - \frac{8\beta_i E_{i,j}^{4\beta_i}(t_{i,j}-t_{i,j-1})^2\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^3\sigma^2} \\
& - \frac{1}{n_i\omega^2} - \frac{4\beta_i E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})^2\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \\
& - \frac{4E_{i,j}^{\beta_i}(t_{i,j}-t_{i,j-1})(y_{i,j-1}-\alpha)\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \\
& + \frac{2\beta_i E_{i,j}^{\beta_i}(t_{i,j}-t_{i,j-1})^2(y_{i,j-1}-\alpha)\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \\
& + \frac{8\beta_i E_{i,j}^{3\beta_i}(t_{i,j}-t_{i,j-1})^2(y_{i,j-1}-\alpha)\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \Bigg).
\end{aligned} \tag{2.31}$$

Replacing (2.29) in (2.27), we obtain, after some algebraic manipulations,

$$LL(\alpha, \lambda, \omega, \sigma) \approx \sum_{i=1}^{M} \left(-\ln\omega - \frac{n_i}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\left|f_i''(\hat{\beta}_i)\right|\right) + f_i(\hat{\beta}_i) \right). \tag{2.32}$$

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.

2.5.2. Likelihood function through the delta approximation method

Considering now the likelihood function (2.27) and, proceeding as in subsection 2.4.2 with

$$u_i(\beta_i) = \exp\left(-\frac{1}{2} \sum_{j=1}^{n_i} \left(\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})} + \ln\left(2\pi\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})\right) \right) \right), \tag{2.33}$$

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

$$L(\alpha, \lambda, \omega, \sigma) = \prod_{i=1}^{M} \mathbf{E}[u_i(\beta_i)] \approx \prod_{i=1}^{M} \left(u_i(\lambda) + \frac{1}{2}\omega^2 u_i''(\lambda) \right), \tag{2.34}$$

where we have used the approximation

$$u_i(\beta_i) = u_i(\lambda + \omega\delta_i) \approx u_i(\lambda) + \omega\delta_i u_i'(\lambda) + \frac{1}{2}(\omega\delta_i)^2 u_i''(\lambda) \tag{2.35}$$

and

$$\mathbf{E}[u_i(\beta_i)] \approx u_i(\lambda) + \frac{1}{2}\omega^2 u_i''(\lambda), \tag{2.36}$$

with

$$u_i''(\lambda) = u_i(\lambda)\left[(g_i'(\lambda))^2 + g_i''(\lambda)\right], \tag{2.37}$$

So, we obtain for the log-likelihood function

$$\begin{aligned} LL(\alpha, \lambda, \omega, \sigma) \approx \sum_{i=1}^{M} \Bigg(& -\frac{1}{2}\sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \alpha - (y_{i,j-1} - \alpha)E_{i,j}^{\lambda}\right)^2}{\frac{\sigma^2}{2\lambda}(1 - E_{i,j}^{2\lambda})} \\ & - \frac{1}{2}\sum_{j=1}^{n_i} \ln\left(2\pi\frac{\sigma^2}{2\lambda}(1 - E_{i,j}^{2\lambda})\right) \\ & + \ln\left(1 + \frac{1}{2}\omega^2\left[(g_i'(\lambda))^2 + g_i''(\lambda)\right]\right)\Bigg), \end{aligned} \tag{2.38}$$

where

$$g_i(\lambda) = -\frac{1}{2}\sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \alpha - (y_{i,j-1} - \alpha)E_{i,j}^{\lambda}\right)^2}{\frac{\sigma^2}{2\lambda}(1 - E_{i,j}^{2\lambda})} - \frac{1}{2}\sum_{j=1}^{n_i} \ln\left(2\pi\frac{\sigma^2}{2\lambda}(1 - E_{i,j}^{2\lambda})\right) \tag{2.39}$$

and its first and second order derivatives are presented in [**14**] and omitted here.

2.6. 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

$$\begin{aligned} p(\alpha_i, \beta_i|\mu, \lambda, \theta, \omega) &= p(\alpha_i|\mu, \theta)p(\beta_i|\lambda, \omega) \\ &= \frac{1}{\sqrt{2\pi\theta^2}}\exp\left(-\frac{(\alpha_i - \mu)^2}{2\theta^2}\right)\frac{1}{\sqrt{2\pi\omega^2}}\exp\left(-\frac{(\beta_i - \lambda)^2}{2\omega^2}\right). \end{aligned} \tag{2.40}$$

To obtain the likelihood function in (2.6), the joint p.d.f. for animal i is given by

$$\begin{aligned} p_{\mathbf{Y}_i}(\mathbf{y}_i|\epsilon, \mathbf{\Psi}) &= \int\limits_{\mathbf{R}^2} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{b}_i, \epsilon)p_B(\mathbf{b}_i|\mathbf{\Psi})\,d\mathbf{b}_i \\ &= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i, \beta_i, \epsilon)p(\alpha_i|\mu, \theta)p(\beta_i|\lambda, \omega)\,d\alpha_i\,d\beta_i, \end{aligned} \tag{2.41}$$

To obtain $p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i, \beta_i, \epsilon)$ in (2.41), notice that, due to the Markov property of each individual trajectory i when conditioned on α_i and β_i, we have

$$p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i, \beta_i, \epsilon) = \prod_{j=1}^{n_i} p_{\mathbf{Y}_i}(y_{i,j}, t_i, t_{i-1}|y_{i,j-1}, \alpha_i, \beta_i, \epsilon), \tag{2.42}$$

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

$$p_{\mathbf{Y}_i}(y_{i,j},t_i,t_{i-1}|y_{i,j-1},\alpha_i,\beta_i,\epsilon)=\frac{\exp\left(-\frac{1}{2}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})}\right)}{\sqrt{2\pi\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})}}. \tag{2.43}$$

Thus, replacing (2.40) and (2.43) on (2.6), the likelihood function, after some simplifications, is given by

$$L(\mu,\theta,\lambda,\omega,\sigma)=\prod_{\boldsymbol{i=1}}^{\boldsymbol{M}}\frac{1}{2\pi\theta\omega}\left(\frac{1}{\sqrt{2\pi}}\right)^{n_i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left(-\frac{1}{2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})}\right) \tag{2.44}$$
$$\exp\left(-\frac{1}{2}\sum_{j=1}^{n_i}\ln\left(\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})\right)\right)\exp\left(-\frac{1}{2}\frac{(\alpha_i-\mu)^2}{\theta^2}\right)\exp\left(-\frac{1}{2}\frac{(\beta_i-\lambda)^2}{\omega^2}\right)d\alpha_i\,d\beta_i.$$

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

2.6.1. Likelihood function through the Laplace approximation method

Following the same approach taken in subsections 2.4.1 and 2.5.1, the integrand function in the likelihood function (2.44) can be written as the exponential of the function

$$f_i(\alpha_i,\beta_i)=\sum_{j=1}^{n_i}\left(-\frac{1}{2}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})}\right. \tag{2.45}$$
$$\left.-\frac{1}{2}\ln\left(\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})\right)-\frac{(\alpha_i-\mu)^2}{2n_i\theta^2}-\frac{(\beta_i-\lambda)^2}{2n_i\omega^2}\right).$$

Rewriting expression (2.44) we obtain

$$L(\mu,\theta,\lambda,\omega,\sigma)=\prod_{\boldsymbol{i=1}}^{\boldsymbol{M}}\left(\frac{1}{2\pi\theta\omega}\left(\frac{1}{\sqrt{2\pi}}\right)^{n_i}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left(f_i(\alpha_i,\beta_i)\right)\,d\alpha_i\,d\beta_i\right). \tag{2.46}$$

As shown in subsection 2.4.1, when it is not possible to obtain an explicit expression for the vector $(\hat{\alpha}_i,\hat{\beta}_i)$ that maximizes the function (2.46), 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 $\nabla f_i(\hat{\alpha}_i,\hat{\beta}_i)=0$, we can write

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left(f_i(\alpha_i,\beta_i)\right)\,d\alpha_i\,d\beta_i$$
$$\approx\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left(f_i(\hat{\alpha}_i,\hat{\beta}_i)+\frac{1}{2}\begin{bmatrix}\alpha_i-\hat{\alpha}_i & \beta_i-\hat{\beta}_i\end{bmatrix}\mathbf{H}\Big[f_i(\hat{\alpha}_i,\hat{\beta}_i)\Big]\begin{bmatrix}\alpha_i-\hat{\alpha}_i\\ \beta_i-\hat{\beta}_i\end{bmatrix}\right)d\alpha_i\,d\beta_i$$
$$=2\pi\left|\det\mathbf{\Sigma}_{(\hat{\alpha}_i,\hat{\beta}_i)}\right|^{1/2}\exp(f_i(\hat{\alpha}_i,\hat{\beta}_i)), \tag{2.47}$$

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 $\mathbf{\Sigma}_{(\hat{\alpha}_i,\hat{\beta}_i)}=-\mathbf{H}^{-1}[f_i(\hat{\alpha}_i,\hat{\beta}_i)]$.

The first and second partial derivatives with respect to α_i are given in (2.12) and (2.13) with β replaced by β_i. Therefore, an exact expression for $\hat{\alpha}_i$ is given by (2.14) with β replaced by $\hat{\beta}_i$.

The first and second partial derivatives with respect to β_i were also already given in subsection 2.5.1, in expressions (2.30) and (2.31), 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

$$\begin{aligned}\frac{\partial^2 f_i(\alpha_i,\beta_i)}{\partial\alpha_i\partial\beta_i} &= \frac{4\beta_i}{\sigma^2}\sum_{j=1}^{n_i}\frac{(t_j-t_{j-1})E_{i,j}^{2\beta_i}\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)\left(E_{i,j}^{\beta_i}-1\right)}{(1-E_{i,j}^{2\beta_i})^2}\\ &- \frac{2\beta_i}{\sigma^2}\sum_{j=1}^{n_i}\frac{(t_j-t_{j-1})(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\left(E_{i,j}^{\beta_i}-1\right)}{(1-E_{i,j}^{2\beta_i})}\\ &+ \frac{2\beta_i}{\sigma^2}\sum_{j=1}^{n_i}\frac{(t_j-t_{j-1})E_{i,j}^{\beta_i}\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)}{(1-E_{i,j}^{2\beta_i})}\\ &- \frac{2}{\sigma^2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)\left(E_{i,j}^{\beta_i}-1\right)}{(1-E_{i,j}^{2\beta_i})},\end{aligned} \tag{2.48}$$

Replacing (2.47) in (2.46), we obtain, after some algebraic manipulation

$$LL(\mu,\theta,\lambda,\omega,\sigma) \approx \sum_{i=1}^{M}\left\{-\ln(\theta\omega)-\frac{n_i}{2}\ln(2\pi)-\frac{1}{2}\ln(|\det\mathbf{H}[f_i(\hat{\alpha}_i,\hat{\beta}_i)]|)+f_i(\hat{\alpha}_i,\hat{\beta}_i)\right\}. \tag{2.49}$$

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.

2.6.2. Likelihood function through the delta approximation method

In [**14**] the delta approximation method was proposed to approximate the integral in (2.6), which, in this case, is given by (2.41) 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 [**14**], denoting expression (2.44) by $u_i(\alpha_i,\beta_i) = \exp(g_i(\alpha_i,\beta_i))$, where

$$g_i(\alpha_i,\beta_i) = \sum_{j=1}^{n_i}\left(-\tfrac{1}{2}\,\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{-\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})}-\tfrac{1}{2}\ln\left(2\pi\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})\right)\right), \tag{2.50}$$

the likelihood function (2.44) can be written as

$$\begin{aligned}L(\mu,\theta,\lambda,\omega,\sigma) &= \prod_{i=1}^{M}\left(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}u_i(\alpha_i,\beta_i)p_B(\alpha_i,\beta_i|\mu,\lambda,\theta,\omega)\,d\alpha_i\,d\beta_i\right)\\ &= \prod_{i=1}^{M}\mathbf{E}[u_i(\alpha_i,\beta_i)].\end{aligned} \tag{2.51}$$

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

$$\mathbf{E}[u_i(\alpha_i,\beta_i)] \approx u_i(\mu,\lambda)+\frac{\theta^2}{2}\frac{\partial^2 u_i}{\partial\alpha_i^2}(\mu,\lambda)+\frac{\omega^2}{2}\frac{\partial^2 u_i}{\partial\beta_i^2}(\mu,\lambda). \tag{2.52}$$

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

$$LL_Y(\mu, \theta, \lambda, \omega, \sigma) \approx \sum_{i=1}^{M} \ln \left\{ u_i(\mu, \lambda) + \frac{\theta^2}{2} \frac{\partial^2 u_i}{\partial \alpha_i^2}(\mu, \lambda) + \frac{\omega^2}{2} \frac{\partial^2 u_i}{\partial \beta_i^2}(\mu, \lambda) \right\}. \tag{2.53}$$

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 [**14**].

2.7. 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 (2.3) 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 (2.3) 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 2.7.1, 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 2.7.2, 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 2.7.2, 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 2.7.4, 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).

2.7.1. 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 θ, we obtain the estimates of the parameters, μ, θ, β, and σ, by maximization of the exact log-likelihood function (2.10) ("Exact method"). We will also estimate the parameters by the DA method, i.e., by maximization of the approximate log-likelihood function (2.24) using the DA method. The maximum likelihood estimates based on the exact log-likelihood function (2.10) ("Exact method") will be the same estimates obtained by the LA method because, in this case and as shown on subsection 2.4.1, 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 (2.10), 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 2.7);
- 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.

2.7.2. 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 2.7, 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 2.7.1, 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 2.7). 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.

2.7.3. 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 2.7, 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 2.7.1 and 2.7.2, 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.

2.7.4. 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 [**23–25**]) 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 performs 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 2.1, 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 2.4, 2.5, and 2.6 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 ± 0.0108	$<$ 0.0001 ± 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 animals 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 of α being 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.

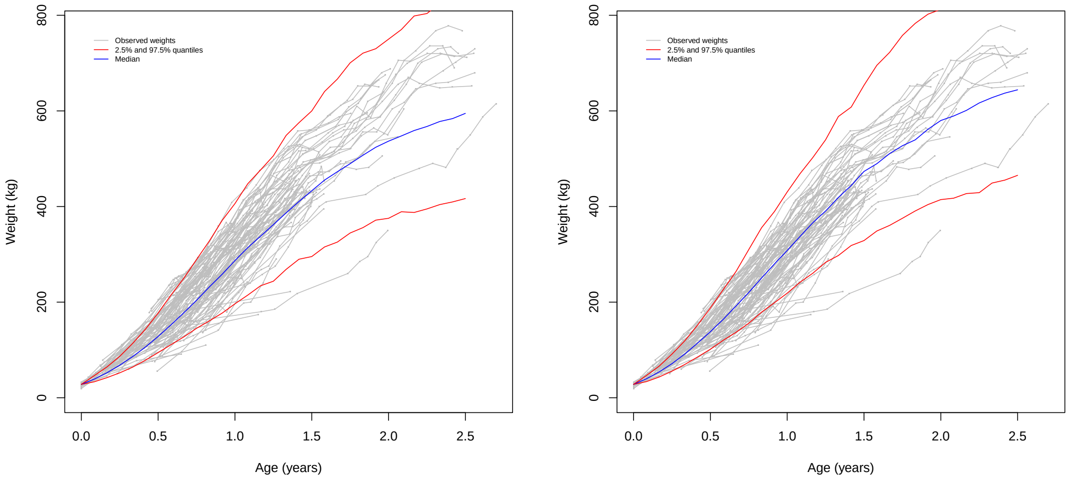

FIGURE 1. Real observed weight trajectories (in grey) of 100 animals (randomly chosen among the 10843 available real trajectories) and estimated 2.5%, 50% (median), and 97.5% quantile curves for the predictions of the Gompertz SDE mixed model with random α (on the left) or with random β (on the right).

The findings presented in Section 2.7 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.

2.8. 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 of 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.

CHAPTER 3

Likelihood Function through the delta approximation in Mixed SDE models

Nelson T. Jamba
Gonçalo Jacinto
Patrícia A. Filipe
Carlos A. Braumann

MDPI Mathematics 10 (3), 385, 2022

Abstract

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

Keywords delta approximation, maximum likelihood estimation method, mixed models, stochastic differential equations.

3.1. Introduction

In previous works [**13**, **15**, **32–35**], we describe the individual growth of animals subject to random fluctuations in the environment and study estimation, prediction and optimization problems with applications to cattle weight data. Considering M animals, we have used the following general SDE model:

$$dY_i(t) = \beta\,(\alpha - Y_i(t))\;dt + \sigma\,dW_i(t), \quad Y_i(t_{i,0}) = y_{i,0}, \quad i = 1, \ldots, M, \tag{3.1}$$

where $Y_i(t) = h(X_i(t))$ is the modified weight by a transformation function h, a known monotonous continuously differentiable function of the real weight $X_i(t)$ of the animal i at age t, and $y_{i,0} = h(x_{i,0})$, where $x_{i,0}$ is the assumed known size of animal i at an initial age of observation $t_{i,0}$. The parameter $\beta > 0$ is the growth coefficient and α is the mean asymptotic modified size towards which the mean modified size converges as $t \to +\infty$; we denote by $A = h^{-1}(\alpha)$ the corresponding real asymptotic size. The intensity of the effect of environmental random fluctuations on growth is measured by the parameter $\sigma > 0$, being $W_i(t)$ $(i = 1, \ldots, M)$ independent realizations of the standard Wiener process. Adequate choices of the h function lead to stochastic versions of well-known growth models. For instance, the monomolecular model corrsponds to $h(x) = x$, the Bertallanfy-Richards model corresponds to $h(x) = x^c$ (with $c > 0$), the Gompertz model corresponds to $h(x) = \ln x$, and the logistic model corresponds to $h(x) = 1/x$, but the theoretical treatment is valid for any monotonous C^1 function h. Here, for comparison purposes, we will illustrate with $h(x) = \ln x$, corresponding to the stochastic Gompertz model.

It is natural to think that the model parameters may vary from animal to animal, and so SDE models with fixed parameters as the one presented in (3.1) may not be a suitable model for these applications. Parameter estimation for models where the parameters are considered random, known as mixed models or mixed-effects models, are presented in [**19**,**22**,**26**,**31**,**33**,**34**,**36**].

For example, in [**33**, **34**] the mixed model considers that different individuals may have different values of A and consequently different values of $\alpha = h(A)$, i.e., the case where the average asymptotic weight varies randomly from animal to animal has been considered. In this particular case, it was considered that α was a random variable, independent of W_t, with a Gaussian distribution with mean μ and variance θ^2, and in [**34**] the maximum likelihood estimation method was applied to estimate the model parameters. In this case, the likelihood function can be explicitly obtained, but it is extremely difficult or impossible to obtain a closed-form expression for the likelihood function in other cases.

More recently, we have considered either α or β random, following a Gaussian distribution, and for these cases we have solved the integral that appears in the likelihood function through approximation methodologies, such as the Laplace and the delta approximation methods (to appear in a forthcomming paper [**37**]). In [**19**, **31**] for the general case where it is not possible to obtain a closed-form expression for the likelihood function, as is the case of random β, a numerical approximation based on an Hermite expansion was applied, whereas in [**22**], in addition to an Hermite expansion, a Gauss-Hermite quadrature was also applied and the parameters of the SDE mixed model were estimated by the maximum likelihood method. In [**26**, **36**], for mixed-models with linear drift term, when a closed-form expression for the likelihood function is not possible a different approximation technique is used, based on a discretized version of the continuous-time data likelihood function.

In this work we consider that both parameters α and β are random variables and apply an adaptation of the delta method to approximate the likelihood function, which we will call delta approximation method. This method allowed us to approximate the parameter estimates, through numerical maximization of the approximate likelihood function. Notice also that the existing methods for SDE mixed models [**26**, **31**, **36**] assume that the age vector of the observations is the same for all trajectories and in some cases even require equidistant ages of observation. When using real data of cattle weights, these assumptions are not adequate, since the animals are not weighted at the same time instants (ages), not even with the same elapsed times between weighings.

To compare our method with existing ones, like the one considered here and provided by the *MsdeParEst* R package (see [**23**, **26**, **36**]), we have worked with simulated cattle weight data with 50, 500 and 5000 animals with the same age vector and with consecutive weights taken at equidistant intervals. We have also estimated the model parameters with our method using a real dataset of 16029 animals with very heterogeneous ages of observations.

This paper is organized as follows. In Section 3.2 we present the SDE models with fixed parameters, their main properties and the respective likelihood function, proceeding to the extension to mixed stochastic diferential equations models. In Section 3.3 we develop the delta approximation applied to the likelihood function for the case where both parameters α and β are independent Gaussian distributed random variables. In Section 3.4 we present the application for both simulated and real datasets and, whenever possible, compare the results with the existing methods. Based on those results, we present practical recommendations on how to deal with real datasets on Section 3.5 and end up with the main conclusions on Section 3.6.

3.2. Stochastic differential equations models

Considering data from M individuals, we will denote by $X_i(t)$ the size (some measure of weight, volume, height, length,...) at age t of the i^{th} individual ($i = 1, \ldots, M$). If the individual is growing in a randomly fluctuating environment, working with a modified size $Y_i(t) = h(X_i(t))$, where the transformation h is a monotonous continuously differentiable function, we can describe the evolution of individual growth through an SDE of the form (3.1). The model solution $Y_i(t)$ is an homogeneous diffusion process with drift coefficient $a(y) = \beta(\alpha - y)$ and diffusion coefficient

$b(y) = \sigma^2$, given by

$$Y_i(t) = \alpha - (\alpha - y_{i,0})e^{-\beta(t-t_{i,0})} + \sigma e^{-\beta t} \int_{t_{i,0}}^{t} e^{\beta s} dW_i(s), \quad i = 1, \ldots, M, \tag{3.2}$$

(see, for instance [**27**]).

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 (3.1). 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 age vector of observations (which may differ from individual to individual), by $\mathbf{Y}_i = (Y_{i,0}, Y_{i,1}, ..., Y_{i,n_i})$ the corresponding vector of modified sizes 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 for $Y_{i,j}$ conditioned on $Y_{i,j-1} = y_{i,j-1}$, the transition distribution for animal i is Gaussian:

$$Y_{i,j} | \left(Y_{i,j-1} = y_{i,j-1}\right) \sim N\left(\alpha + \left(y_{i,j-1} - \alpha\right) E_{i,j}^{\beta}, \frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})\right). \tag{3.3}$$

In [**15**, **33**, **34**] we have applied the maximum likelihood estimation method to estimate the parameter vector $\mathbf{p} = (\alpha, \beta, \sigma)$. From (3.3), using the fact that $Y_i(t)$ is a Markov process, we know that, given $Y_{i,0} = y_{i,0}$ (assumed known), the $\mathbf{Y_i}$ joint probability density function for indivifual i takes the form

$$\begin{aligned} p_{\mathbf{Y_i}}(\mathbf{y}_i | \alpha, \beta, \sigma) &= \prod_{j=1}^{n_i} \frac{\exp\left(-\frac{1}{2} \frac{\left(y_{i,j} - \alpha - (y_{i,j-1} - \alpha) E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}\right)}{\sqrt{2\pi \frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}} \\ &= \frac{\exp\left(-\frac{1}{2} \sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \alpha - (y_{i,j-1} - \alpha) E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}\right)}{\prod_{j=1}^{n_i} \sqrt{2\pi \frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}}, \quad i = 1, \ldots, M, \end{aligned} \tag{3.4}$$

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

$$L(\alpha, \beta, \sigma) = \prod_{i=1}^{M} p_{\mathbf{Y}_i}(\mathbf{y}_i | \alpha, \beta, \sigma). \tag{3.5}$$

The maximum likelihood estimate of the parameter vector $\mathbf{p}$ is obtained by maximization of (3.5) or of the log-likelihood function $LL_Y(\alpha, \beta, \sigma) = \ln L(\alpha, \beta, \sigma)$.

$$\begin{aligned} LL_Y(\alpha, \beta, \sigma) &= \ln L(\alpha, \beta, \sigma) = \sum_{i=1}^{M} - \frac{n_i}{2} \ln(2\pi) \\ &- \frac{n_i}{2} \sum_{i=1}^{M} \ln\left(\frac{\sigma^2}{2\beta}\right) - \sum_{i=1}^{M} \frac{1}{2} \sum_{j=1}^{n_i} \ln\left(1 - E_{i,j}^{2\beta}\right) \\ &- \frac{\beta}{\sigma^2} \sum_{i=1}^{M} \sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \alpha - \left(y_{i,j-1} - \alpha\right) E_{i,j}^{\beta}\right)^2}{1 - E_{i,j}^{2\beta}}. \end{aligned} \tag{3.6}$$

The maximum likelihood estimators are asymptotically Gaussian with mean $\mathbf{p}$ and variance-covariance matrix $\mathbf{V} = \mathbf{F}^{-1}$, where $\mathbf{F}$ is the Fisher information matrix with elements given by

$F_{i,j} = -\mathbb{E}\left[\partial^2 L(\mathbf{p})/\partial p_i \partial p_j\right]$. We can estimate $\mathbf{V}$ by the inverse of the empirical information matrix $\hat{\mathbf{F}}$, with elements $\hat{F}_{i,j} = -\partial^2 L(\hat{\mathbf{p}})/\partial p_i \partial p_j$. From these values we can obtain the approximate confidence bands for the parameters.

For these type of models, in terms of estimation methods, we have also developed and applied parametric and nonparametric boostrap methods ([**34**, **38**]). Since the asymptotic confidence intervals obtained from the empirical Fisher information matrix may be quite unreliable for small sample sizes, the bootstrap methods can be used in such cases.

In [**15**, **34**], nonparametric estimation methods were developed in order to estimate the drift and difusion coefficients of a stochastic diferential equation model for the case of nonequidistant data. For our application on cattle weight data, we had been working with models with specific functional forms for the drift and the diffusion coefficients, for example, in the case of model (3.1), with a drift coefficient of linear form $a(y) = \beta(\alpha - y)$ and a constant diffusion coefficient $b(y) = \sigma^2$. These nonparametric methods are useful to assess whether our specific choice of functional forms is appropriate for our data or whether some alternative functional forms for these coefficients are suggested.

Recently, weighted maximum likelihood estimation methods were studied and adapted to overcome one very common limitation in the cattle weight data applications, that it is related to the fact that usually animals are not weighted very frequently and there exists scarce number of weight observations for older ages. In the weighted maximum likelihood estimation method, the weights are built such that the times elapsed between consecutive observations are considered in the likelihood function ([**38**]).

We have described the general SDE model (3.1) for the complete growth curve of the animals where the model's parameters α, β and σ are assumed common to all individuals. In [**16**], we present a generalization of (3.1) to a multiphasic model, where it is considered that the growth coefficient β can assume different values for different phases of the animals' growth curve.

Here, we consider a different generalization to account for the fact that it is natural to think that different animals, due to their specific genetical and other characteristics, may have different values of the parameters. So, in this paper we will consider the situation where different individuals may have different randomly assigned parameters.

In [**19**, **22**, **26**, **30**, **34**, **36**], it has been shown that, to consider at least one of the two parameters of the drift term, α or β, as random variables, the likelihood function can be obtained from the transition density function conditioned on the respective random parameter.

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

$$\begin{aligned} L(\mathbf{\Lambda}, \mathbf{\Psi}) &= \prod_{i=1}^{M} p_{\mathbf{Y}_i}(\mathbf{y}_i | \mathbf{\Lambda}, \mathbf{\Psi}) \\ &= \prod_{i=1}^{M} \int_{\mathbb{R}^d} p_{\mathbf{Y}_i}(\mathbf{y}_i | \mathbf{b}_i, \mathbf{\Lambda}) p_B(\mathbf{b}_i | \mathbf{\Psi})\, d\mathbf{b}_i. \end{aligned} \tag{3.7}$$

The case of a single random parameter $\mathbf{b}_i = (\alpha_i)$ has already been studied with $\alpha_i \sim N(\mu, \theta^2)$. When we have the special situation of a time vector of observations $\mathbf{t_i} \equiv \mathbf{t} = (t_0, t_1, \ldots, t_n)$, $i = 1, \ldots, M$ common to all animals, we can see it in [**31**] (for the particular situation of $\mu = 0$ and uniform time spacings $t_j - t_{j-1} \equiv \delta$) and in [**26**]. For the general situation, with no such restrictions, you can see it in [**34**]. In this case, it is possible to explicitly compute the integral in the likelihood function, resulting in a final closed-form expression for this function. This is shown in [**34**], where the log-likelihood function for all animals $LL_Y(\mu, \theta, \beta, \sigma) = \ln L(\mu, \theta, \beta, \sigma)$ is given by

$$\begin{aligned} LL_Y(\mu, \theta, \beta, \sigma) = \sum_{i=1}^{M} \Bigg(& -\frac{n_i}{2} \ln(2\pi) - \frac{n_i}{2} \ln\left(\frac{\sigma^2}{2\beta}\right) - \frac{1}{2} \sum_{j=1}^{n_i} \ln\left(1 - E_{i,j}^{2\beta}\right) - \frac{1}{2} \ln(D_i) \\ & - \tfrac{\beta}{D_i \sigma^2} \sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \mu - \left(y_{i,j-1} - \mu\right) E_{i,j}^{\beta}\right)^2}{1 - E_{i,j}^{2\beta}} + \frac{2\beta^2\theta^2}{D_i \sigma^4} \left(\sum_{j=1}^{n_i} \frac{\left(y_{i,j} - y_{i,j-1} E_{i,j}^{\beta}\right)}{1 + E_{i,j}^{\beta}} \right)^2 \\ & - \tfrac{\beta(D_i - 1)}{D_i \sigma^2} \left(\sum_{j=1}^{n_i} \frac{\left(y_{i,j} - y_{i,j-1} E_{i,j}^{\beta}\right)^2}{1 - E_{i,j}^{2\beta}} \right) \Bigg), \end{aligned} \tag{3.8}$$

with $D_i = \frac{2\beta\theta^2}{\sigma^2} \sum_{j=1}^{n_i} \dfrac{1 - E_{i,j}^{\beta}}{1 + E_{i,j}^{\beta}} + 1$. However, despite the existence of a closed-form expression for the likelihood function for this particular case, we have also applied approximation methods (Laplace and delta approximations - to appear in a forthcomming paper [**37**]), showing that the approximation methods also provide very good results when compared with the exact method.

Unfortunately, the integral in (3.7) does not always have a closed-form expression, for example when the random parameter is β, corresponding to $\mathbf{b}_i = (\beta_i)$, and in such cases the approximation methods are good alternatives.

In the following section, we present the study of the case where both α and β are considered random variables and propose the aplication of the delta approximation to the integral in (3.7).

3.3. Mixed model for two random effects

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

$$dY_i(t) = \beta_i \left(\alpha_i - Y_i(t)\right) dt + \sigma\, dW_i(t), \qquad Y_i(t_{i,0}) = y_{i,0}, \quad i = 1, \ldots, M, \tag{3.9}$$

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

To obtain the likelihood function in (3.7) for this two random parameters situation, notice that, due to the assumed independence structure,

$$\begin{aligned} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{\Lambda},\mathbf{\Psi}) &= \int\limits_{\mathbb{R}^2} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{b}_i,\mathbf{\Lambda})p_B(\mathbf{b}_i|\mathbf{\Psi})\,d\mathbf{b}_i \\ &= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i,\beta_i,\mathbf{\Lambda})p(\alpha_i|\mu,\theta)p(\beta_i|\lambda,\omega)\,d\alpha_i\,d\beta_i, \qquad (3.10) \end{aligned}$$

with

$$p(\alpha_i|\mu,\theta) = \frac{1}{\sqrt{2\pi\theta^2}}\exp\left(-\frac{(\alpha_i-\mu)^2}{2\theta^2}\right), \qquad (3.11)$$

$$p(\beta_i|\lambda,\omega) = \frac{1}{\sqrt{2\pi\omega^2}}\exp\left(-\frac{(\beta_i-\lambda)^2}{2\omega^2}\right). \qquad (3.12)$$

To determine $p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i,\beta_i,\mathbf{\Lambda})$ in (3.10), notice that, due to the Markov property of each individual trajectory i when conditioned on α_i and β_i, we have

$$p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha_i,\beta_i,\mathbf{\Lambda}) = \prod_{j=1}^{n_i} p_{\mathbf{Y}_i}(y_{i,j},t_i,t_{i-1}|y_{i,j-1},\alpha_i,\beta_i,\mathbf{\Lambda}), \qquad (3.13)$$

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

$$p_{\mathbf{Y}_i}(y_{i,j},t_i,t_{i-1}|y_{i,j-1},\alpha_i,\beta_i,\mathbf{\Lambda}) = \frac{\exp\left(-\frac{1}{2}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})}\right)}{\sqrt{2\pi\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})}}. \qquad (3.14)$$

We will apply the delta approximation to the integral in (3.7), which in this case is given by (3.10) and does not have an explicit solution. This approximation method allows us to obtain simpler expressions for the likelihood function. The method is also adapted to general age vectors of observations, allowing 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.

Let $u_i(\alpha_i,\beta_i) = \exp\left(g_i(\alpha_i,\beta_i)\right)$, where

$$\begin{aligned} g_i(\alpha_i,\beta_i) &= \sum_{j=1}^{n_i}\left(-\frac{1}{2}\frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)^2}{\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})}\right. \\ &\qquad \left. -\frac{1}{2}\ln\left(2\pi\frac{\sigma^2}{2\beta_i}(1-E_{i,j}^{2\beta_i})\right)\right). \qquad (3.15) \end{aligned}$$

Taken into account the previous expressions, the likelihood function (3.7) can be writen as an expectation w.r.t. the bivariate Gaussian distribution with density p_B:

$$\begin{aligned} L(\mu, \theta, \lambda, \omega, \sigma) &= \prod_{i=1}^{M} \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} u_i(\alpha_i, \beta_i) p_B(\alpha_i, \beta_i | \mu, \lambda, \theta, \omega) \, d\alpha_i \, d\beta_i \right) \\ &= \prod_{i=1}^{M} \mathbb{E}[u_i(\alpha_i, \beta_i)]. \end{aligned} \tag{3.16}$$

The mathematical expectation of $u_i(\alpha_i, \beta_i)$ in (3.16) can be obtained approximately by applying the delta approximation, which we obtain by adapting the delta method. The delta approximation consists in using the second-order Taylor series expansion about the point (μ, λ)

$$\begin{aligned} u_i(\alpha_i, \beta_i) \approx u_i(\mu, \lambda) &+ \begin{bmatrix} \frac{\partial u_i}{\partial \alpha_i}(\mu, \lambda) & \frac{\partial u_i}{\partial \beta_i}(\mu, \lambda) \end{bmatrix} \begin{bmatrix} \alpha_i - \mu \\ \beta_i - \lambda \end{bmatrix} \\ &+ \frac{1}{2} \begin{bmatrix} \alpha_i - \mu & \beta_i - \lambda \end{bmatrix} \begin{bmatrix} \frac{\partial^2 u_i}{\partial \alpha_i^2}(\mu, \lambda) & \frac{\partial^2 u_i}{\partial \alpha_i \beta_i}(\mu, \lambda) \\ \frac{\partial^2 u_i}{\partial \beta_i \alpha_i}(\mu, \lambda) & \frac{\partial^2 u_i}{\partial \beta_i^2}(\mu, \lambda) \end{bmatrix} \begin{bmatrix} \alpha_i - \mu \\ \beta_i - \lambda \end{bmatrix} \end{aligned} \tag{3.17}$$

and applying mathematical expectations to obtain

$$\mathbb{E}[u_i(\alpha_i, \beta_i)] \approx u_i(\mu, \lambda) + \frac{\theta^2}{2} \frac{\partial^2 u_i}{\partial \alpha_i^2}(\mu, \lambda) + \frac{\omega^2}{2} \frac{\partial^2 u_i}{\partial \beta_i^2}(\mu, \lambda). \tag{3.18}$$

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

$$LL_Y(\mu, \theta, \lambda, \omega, \sigma) \approx \sum_{i=1}^{M} \ln \left\{ u_i(\mu, \lambda) + \frac{\theta^2}{2} \frac{\partial^2 u_i}{\partial \alpha_i^2}(\mu, \lambda) + \frac{\omega^2}{2} \frac{\partial^2 u_i}{\partial \beta_i^2}(\mu, \lambda) \right\}. \tag{3.19}$$

The first and second-order derivatives with respect to α_i are

$$\begin{aligned} \frac{\partial g_i(\alpha_i, \beta_i)}{\partial \alpha_i} &= \sum_{j=1}^{n_i} \left(-\frac{2\beta_i(-1 + E_{i,j}^{\beta_i}) \left(y_{i,j} - \alpha_i - (y_{i,j-1} - \alpha_i) E_{i,j}^{\beta_i} \right)}{(1 - E_{i,j}^{2\beta_i}) \sigma^2} \right) \\ \frac{\partial^2 g_i(\alpha_i, \beta_i)}{\partial \alpha_i^2} &= \sum_{j=1}^{n_i} \left(-\frac{2\beta_i(-1 + E_{i,j}^{\beta_i})^2}{(1 - E_{i,j}^{2\beta_i}) \sigma^2} \right) \end{aligned} \tag{3.20}$$

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

$$\frac{\partial^2 u_i(\alpha_i, \beta_i)}{\partial \alpha_i^2} = u_i(\alpha_i, \beta_i) \left(\frac{\partial^2 g_i(\alpha_i, \beta_i)}{\partial \alpha_i^2} + \left(\frac{\partial g_i(\alpha_i, \beta_i)}{\partial \alpha_i} \right)^2 \right). \tag{3.21}$$

The first two derivatives of $g_i(\alpha_i,\beta_i)$ with respect to β_i are

$$\begin{aligned}\frac{\partial g_i(\alpha_i,\beta_i)}{\partial\beta_i} = \sum_{j=1}^{n_i}\Bigg(& \frac{1}{2\beta_i} + \frac{2\beta_i E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \\ & -\frac{E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})}{\left(1-E_{i,j}^{2\beta_i}\right)} - \frac{\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \\ & -\frac{2\beta_i E_{i,j}^{\beta_i}(t_{i,j}-t_{i,j-1})(y_{i,j-1}-\alpha_i)\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2}\Bigg)\end{aligned} \tag{3.22}$$

and

$$\begin{aligned}\frac{\partial^2 g_i(\alpha_i,\beta_i)}{\partial\beta_i^2} = \sum_{j=1}^{n_i}\Bigg(& -\frac{4\beta_i E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})^2\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \\ & -\frac{1}{2\beta_i^2} - \frac{4E_{i,j}^{\beta_i}(t_{i,j}-t_{i,j-1})(y_{i,j-1}-\alpha_i)\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \\ & +\frac{2E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})^2}{\left(1-E_{i,j}^{2\beta_i}\right)} + \frac{2\beta_i E_{i,j}^{\beta_i}(t_{i,j}-t_{i,j-1})^2(y_{i,j-1}-\alpha_i)\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2} \\ & +\frac{2E_{i,j}^{4\beta_i}(t_{i,j}-t_{i,j-1})^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2} + \frac{8\beta_i E_{i,j}^{3\beta_i}(t_{i,j}-t_{i,j-1})^2(y_{i,j-1}-\alpha_i)\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \\ & +\frac{4E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^2\sigma^2} \\ & -\frac{8\beta_i E_{i,j}^{4\beta_i}(t_{i,j}-t_{i,j-1})^2\left(y_{i,j}-\alpha_i-(y_{i,j-1}-\alpha_i)E_{i,j}^{\beta_i}\right)^2}{\left(1-E_{i,j}^{2\beta_i}\right)^3\sigma^2} \\ & -\frac{2\beta_i E_{i,j}^{2\beta_i}(t_{i,j}-t_{i,j-1})^2(y_{i,j-1}-\alpha_i)^2}{\left(1-E_{i,j}^{2\beta_i}\right)\sigma^2}\Bigg).\end{aligned} \tag{3.23}$$

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

$$\frac{\partial^2 u_i(\alpha_i,\beta_i)}{\partial\beta_i^2} = u_i\left(\alpha_i,\beta_i\right)\left(\frac{\partial^2 g_i(\alpha_i,\beta_i)}{\partial\beta_i^2} + \left(\frac{\partial g_i(\alpha_i,\beta_i)}{\partial\beta_i}\right)^2\right). \tag{3.24}$$

Replacing these derivatives in the expression (3.19), we obtain an approximate expression for the log-likelihood function $LL_Y(\mu,\theta,\lambda,\omega,\sigma,\rho) = \ln L(\mu,\theta,\lambda,\omega,\sigma,\rho)$ for the case of two random Gaussian parameters. Using it, we are going in the next Section to apply this approach to obtain approximate maximum likehood parameter estimates for the mixed stochastic Gompertz model with two random parameters using cattle weight data.

3.4. Results

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

In our application we have worked with real cattle weight data provided by the Associação de Criadores de Bovinos Mertolengos (ACBM), which performs the growing and finishing phases of young Mertolengo males, and by associated breeders, from the Alentejo region in Portugal. The available dataset contains a total of 96204 observations of the weight (in kg) of 16029 Mertolengo cattle males, where each animal has several observations with a minimum of 3 and a maximum of 33 weights at ages varying from birth until a maximum age that ranges between 0.2 and 16 years old. This is a case where indeed each animal has its weight measurements taken at different (and even non-equidistant) age instants, varying from animal to animal.

We have obtained the estimation results implementing the proposed delta approximation method (DA method) in the software R project [**39**]. However, since it is a new method, we first will compare the results obtained with this DA method with the estimation methods available in the R package *MsdeParEst* described in [**23**], which we will call the MPE methods. The package includes, when closed-form likelihhod function expressions are not possible, tecnhiques based on a discretized version of the continuous-time data maximum likelihood developed in [**26**, **36**]. The main drawback of this package (and, to the best of our knowledge, of other similar available R packages [**24**,**25**]) is that all observations must be measured at the same age instants for all animals and, in some cases, the age instants are even required to be equidistant. Therefore, in order to compare the performance of the proposed approximation method DA with the methods available in the literature (in particular, the MPE method used here), we should also need datasets where weights are observed at the same equidistant age instants for all the animals. Whenever possible, we will also compare the DA and MPE methods, which consider approximations for the likelihood function, with existing estimation methods that use exact closed-form expressions for the likelihood function, in particular, when the likelihood function assumes both parameters fixed or Non-Mixed (which we will call the NMSDE method, presented in [**13**, **15**, **34**]) or, if appropriate, when the likelihood function assumes just the parameter α as random (which we will call the Exact(α) method, presented in [**34**]).

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

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

- $D_M(\alpha,\beta)$: Mixed SDE model with random independent α_i and β_i, where $\alpha_i \sim \mathcal{N}(\mu, \theta^2)$ and $\beta_i \sim \mathcal{N}(\lambda, \omega^2)$ with $\mu = 6.45$ $(h^{-1}(\mu) = 632.70$ kg), $\theta = 0.15$, $\lambda = 1.43$ $year^{-1}$, $\omega = 0.30$ $year^{-1}$, and with fixed parameter $\sigma = 0.33$ $year^{-1/2}$;
- $D_M(\alpha)$: Mixed SDE model with random $\alpha_i \sim \mathcal{N}(\mu, \theta^2)$ with $\mu = 6.45$ $(h^{-1}(\mu) = 632.70$ kg) and $\theta = 0.15$, and with fixed parameters $\beta = 1.43$ $year^{-1}$ and $\sigma = 0.33$ $year^{-1/2}$;
- $D_M(\beta)$: Mixed SDE model with random $\beta \sim \mathcal{N}(\lambda, \omega^2)$ with $\lambda = 1.43$ $year^{-1}$ and $\omega = 0.30$ $year^{-1}$, and with fixed parameters $\alpha = 6.45$ $(h^{-1}(\alpha) = 632.70$ kg) and $\sigma = 0.33$ $year^{-1/2}$;
- D_M: Non-mixed SDE model (3.1) with fixed parameters $\alpha = 6.45$ $(h^{-1}(\alpha) = 632.70$ kg), $\beta = 1.43$ $year^{-1}$ and $\sigma = 0.33$ $year^{-1/2}$.

Each of the simulated datasets based on a Mixed SDE model were obtained in the following way. First, the M random parameter values $\mathbf{b}_i$ ($i = 1, \ldots, M$) of the M animals were simulated based on the Gaussian distribution p_B, and then, for each animal i ($i = 1, \ldots, M$), the simulated values were incorporated in the transition density (3.3) as the true values of α_i and/or β_i for that animal. Then, we have simulated the weights of each animal based on the Markov property and the use of the transition densities between consecutive observation ages. For the non-mixed or fixed parameters model, the simulation of the dataset is obtained without simulating the parameter values of α_i and β_i since they are fixed.

The advantage of using simulated data, besides the possibility of comparing the different methods due to its common age vector of observations, is the knowledge of the true parameter values, which allows us to compare the performances of the different methods when they can be applied. We first consider a simulated dataset for weights of $M = 5000$ animals to evaluate the behaviour of the methods for a large number of trajectories close to an asymptotic regimen. However, in applications we usually do not have datasets with so many animals available and it is important to evaluate the performance of the different methods for smaller data sets. For this reason, we also present a comparison of the main methods considering datasets with $M = 500$ and $M = 50$ animals. To distinguish them we index the simulated datasets by the number of animals M.

Tables 1-4 presents the results for each of the datasets $D_{5000}(\alpha,\beta)$, $D_{5000}(\alpha)$, $D_{5000}(\beta)$ and D_{5000} simulated under the four stochastic Gompertz mixed and fixed models. In each table, we present the results obtained using the different DA methods, DA(α,β) (the one that assumes both parameters α and β random and is described in Section 3.3), DA(α) and DA(β) (the ones that assume just α random or just β random as described in [**37**]), as well as the estimates obtained by the different MPE methods, MPE(α,β), MPE(α) and MPE(β) (again when we consider random both or just one of the parameters). We also present, for comparison purposes, the results obtained when using the NMSDE estimation method (which assumes both parameters to be fixed). So, we can assess what happens when an appropriate estimation method is applied [for instance, when applied to the $D_{5000}(\alpha)$ dataset, the estimation method MPE(α) and the estimation method DA(α,β) are appropiate] and when an inappropriate estimation method is applied [for instance, when applied to the $D_{5000}(\alpha,\beta)$ dataset, the MPE(α) method is inappropriate since it cannot estimate the parameters involved in the random β]; this may be important since, in real non-simulated data, we may not know which parameters are random or not. Of course, DA(α,β) and MPE(α,β) are appropriate for all datasets but, when the data has one or both parameters non-random, they are overparametrized, and therefore, may not be so accurate as non-overparametrized methods.

The maximum likelihood estimates and the approximate 95% confidence bands based on the inverse of the empirical Fisher information matrix are presented only for the DA methods and the NMSDE method, but not for the MPE methods since the confidence intervals of the estimates are not provided by the R package *MsdeParEst*. Instead of presenting the results for the overall mean of the modified asymptotic weight μ [which is $= \alpha$ when α is a fixed parameter], we present them in terms of the corresponding actual weight value $h^{-1}(\mu)$. On Tables 1-4, we have also underlined the headings of the estimation methods that are appropriate to analyse the corresponding dataset, in the sense of being able to estimate all the parameters of the model used to simulate the dataset. Appropriate methods include the ones that use in the likelihood function or its approximation the same random parameters that were used in the dataset generation (these would be the most appropriate methods when we know beforehand which parameteres are random), but include as well overparametrized methods that also use additional random parameters in the likelihood function or its approximation.

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

When considering on Table 1 the $D_{5000}(\alpha,\beta)$ dataset, the appropriate estimation methods are the ones assuming α and β random, the DA(α,β) method we have proposed and the MPE(α,β) method. We can observe that the DA(α,β) method provides a lower bias for all parameters than the MPE(α,β) method, noticing that in the MPE(α,β) method the estimate of ω is extremely biased, close to zero. The use of an inappropriate method should obviously be avoided, but when using real data we often do not know beforehand exactly which are the random parameters and may make a wrong assumption, leading to the use of an inappropriate estimation method. These simulated datasets in which we know exactly which parameters are random, give us the opportunity of assessing the consequences of a wrong assumption on the random parameters when dealing with real data. Let us look at this issue, beyond the obvious fact that an inappropriate method cannot estimate some of the parameters. When comparing the appropriate DA(α,β) method with the inappropriate (for this dataset) NMSDE method, the former method leads to less biased estimates. When comparing the appropriate DA(α,β) with the inappropriate (for this dataset) DA(α) and DA(β) methods, the inappropriate methods estimate the standard deviation of their own random parameter (resp. the standard deviation θ of the random α and the standard deviation ω of the random β) better than the appropriate method, which underestimates both standard deviations. On the estimation of the noise intensity parameter σ, the appropriate DA(α,β) method shows the better performance and it also outperforms the inappropriate DA(β) method on the estimation of the means $h^{-1}(\mu)$ and λ, but these means are curiously better estimated by the inappropriate DA(α) method. As for the inappropriate MPE(α) and MPE(β) methods, besides the obvious impossibility of estimating one parameter, have a behaviour not too different from the MPE(α,β) appropriate method.

Table 2 shows the results for the $D_{5000}(\alpha)$ dataset. For this case, we have included the Exact(α) method, where the maximum likelihood is obtained exactly in closed-form, and that is certainly the best choice method for this dataset. The methods DA(α,β) and DA(α) are both appropriate for this dataset, but the first is overparametrized, while the latter is the most appropriate of the two and we expect it to be more accurate. Still, it is interesting to notice that these two methods provide exactly the same parameter estimates, except naturally for $\omega = 0$ ("parameter" obviously out of the DA(α) method). Comparing the MPE(α,β) with the MPE(α) methods, they do not give exactly the same parameter estimates but their estimates are very close. As expected, the most appropriate method for this dataset, the Exact(α) method, provides slightly better estimates than the DA(α) and the DA(α,β) methods and than the MPE(α) and the MPE(α,β) methods. The estimates for $h^{-1}(\mu)$ and θ are better when using the appropriate MPE methods than the ones obtained with the corresponding DA methods, but the reverse happens for the estimates of the mean λ and the standard deviations $\omega = 0$ of parameter β. The

Table 1. Results for the simulated dataset of 5000 animals when both α and β are random - $D_{5000}(\alpha,\beta)$. Besides the true parameter values used in the simulations, the Table shows their maximum likelihood estimates and, when available, approximate 95% confidence bands using the estimation methods: delta approximation methods DA(α,β), DA(α) and DA(β), NMSDE method (using the exact likelihood when both α and β are fixed), and MPE methods MPE(α,β), MPE(α) and MPE(β), where the likelihoods or approximate likelihoods work with the indicated parameters taken as random. Appropriate methods are underlined. Notice that the LL_X values of approximate methods may differ from the true values of the log-likelihood function at the parameter estimated values.

	True	**DA(α,β)**	**DA(α)**	**DA(β)**
$h^{-1}(\mu)$	632.70	640.87 ± 3.17	629.77 ± 3.16	651.31 ± 2.83
θ	0.15	0.1011 ± 0.0040	0.1397 ± 0.0024	–
λ	1.43	1.3316 ± 0.0067	1.3548 ± 0.0060	1.3036 ± 0.0065
ω	0.30	0.1786 ± 0.0042	–	0.2002 ± 0.0029
σ	0.33	0.3399 ± 0.0010	0.3414 ± 0.0010	0.3405 ± 0.0001
LL_X		−1191084	−1192146	−1191418
	NMSDE	**MPE(α,β)**	**MPE(α)**	**MPE(β)**
$h^{-1}(\mu)$	646.29 ± 2.92	623.77	624.34	622.40
θ	–	0.2158	0.2127	–
λ	1.3208 ± 0.0061	1.3272	1.3249	1.3182
ω	–	0.0001	–	0.0444
σ	0.3452 ± 0.0010	0.2983	0.3140	0.3155
LL_X	−1194145	−1209481	−1208249	−1208631

inappropriate methods DA(β) and MPE(β) give worse estimates than the appropriate methods of the same type. The use of the inappropriate (for this data) NMSDE method performs quite badly in the estimation of $h^{-1}(\mu)$. Notice that the dataset was simulated for fixed β, i.e, $\omega = 0$, and this is better captured by the DA method.

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

For the case of the $D_{5000}(\beta)$ dataset, Table 3 shows that again the appropriate (for this dataset) DA(α,β) (which is overparametrized) and DA(β) methods provide very similar parameter estimates. Curiously, the inappropriate NMSDE model performs surprisingly well. In this case, the MPE methods, both appropriate and inappropriate ones, present very poor estimates. We also highlight that the estimates of the standard deviations $\theta = 0$ and $\omega = 0.30$ using the appropriate MPE methods severely fail, suggesting even that α is random and β is not when the reverse is the true situation, while the DA(α,β) method captures very well the non-random nature of parameter α and both DA(α,β) and DA(β) methods capture the random nature of β, although

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

	True	**DA(α,β)**	**DA(α)**	**DA(β)**
$h^{-1}(\mu)$	632.70	637.79 ± 3.01	637.79 ± 3.01	648.73 ± 2.70
θ	0.15	0.1180 ± 0.0027	0.1180 ± 0.0027	–
λ	1.43	1.4158 ± 0.0060	1.4158 ± 0.0060	1.3838 ± 0.0069
ω	0	< 0.0001	–	0.1190 ± 0.0059
σ	0.33	0.3328 ± 0.0009	0.3328 ± 0.0009	0.3343 ± 0.0010
LL_X		−1191370	−1191370	−1192336

	NMSDE	**MPE(α,β)**	**MPE(α)**	**MPE(β)**
$h^{-1}(\mu)$	644.13 ± 2.66	632.81	633.34	630.96
θ	–	0.1439	0.1532	–
λ	1.3906 ± 0.0060	1.3517	1.3515	1.3494
ω	–	0.0124	–	0.0042
σ	0.3359 ± 0.0010	0.3074	0.3085	0.3075
LL_X	−1192590	−1207378	−1207350	−1202663

	Exact(α)
$h^{-1}(\mu)$	633.53 ± 3.66
θ	0.1523 ± 0.0047
λ	1.4332 ± 0.0061
ω	–
σ	0.3307 ± 0.0009
LL_X	−1190931

somewhat underestimating its standard deviation ω.

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

In this case, not only the NMSDE method and the DA methods give almost coincidental parameter estimates, but also the LL_X value of the the NMSDE method (which is the exact maximum of the exact log-likelihood function) is almost coincidental with the LL_X values for the DA methods, so these LL_X DA values, which are approximations, are likely to be good approximations. If these values were exact instead of just approximations, a likelihood ratio test

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

	True	**DA(α,β)**	**DA(α)**	**DA(β)**
$h^{-1}(\mu)$	632.70	630.64 ± 2.60	611.04 ± 3.3176	630.70 ± 2.60
θ	0	< 0.0001	0.1105 ± 0.0035	–
λ	1.43	1.3559 ± 0.0068	1.3852 ± 0.0061	1.3482 ± 0.0068
ω	0.30	0.1977 ± 0.0033	–	0.1972 ± 0.0031
σ	0.33	0.3362 ± 0.0010	0.3379 ± 0.0010	0.3360 ± 0.0009
LL_X		−1186396	−1187992	−1186334
	NMSDE	**MPE(α,β)**	**MPE(α)**	**MPE(β)**
$h^{-1}(\mu)$	630.12 ± 2.7063	620.22	619.63	620.66
θ	–	0.1413	0.1458	–
λ	1.3624 ± 0.0061	1.3229	1.3252	1.3155
ω	–	0.0001	–	0.0275
σ	0.3403 ± 0.0010	0.3271	0.3138	0.3153
LL_X	−1188686	−1203829	−1203643	−1203829

for $\theta = 0$ (i.e, for the randomness of α) and/or the significance of $\omega = 0$ (i.e., for the randomness of β), would obtain non-significant results.

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

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

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

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

	True	**DA**(α,β)	**DA**(α)	**DA**(β)
$h^{-1}(\mu)$	632.70	631.17 ± 2.49	631.16 ± 2.49	631.19 ± 2.47
θ	0	0.0067 ± 0.0407	0.0070 ± 0.0375	–
λ	1.43	1.4310 ± 0.0061	1.4310 ± 0.0061	1.4309 ± 0.0060
ω	0	0.0047 ± 0.1298	–	0.0074 ± 0.0796
σ	0.33	0.3295 ± 0.0009	0.3295 ± 0.0009	0.3295 ± 0.0009
LL_X		−1186643	−1186643	−1186643

	NMSDE	**MPE**(α,β)	**MPE**(α)	**MPE**(β)
$h^{-1}(\mu)$	631.19 ± 2.4747	630.31	631.07	630.96
θ	–	0.0262	0.0186	–
λ	1.4309 ± 0.0060	1.3518	1.3493	1.3494
ω	–	0.0071	–	0.0042
σ	0.3295 ± 0.0009	0.2920	0.3075	0.3075
LL_X	−1186643	−1202431	−1201744	−1201744

clearly, when both effects are fixed).

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

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

For the fixed-effects datasets D_{50} and D_{500}, the estimates of the DA(α,β) method replicate the same results as the NMSDE method, which, for these datasets, is the most appropriate method. The LL_X values of the NMSDE method are exact maximum log-likelihood values and they are again pratically equal to the LL_X approximate values of the DA(α,β) method, suggesting again that, for fixed-effects datasets, these are good approximations, in which case a likelihood ratio test of $\theta = \omega = 0$ would accept this null hypothesis.

Finally, in Table 9 we used the real cattle weight dataset of 16029 Mertolengo cattle males (totalling 96204 observations) and present the estimates obtained by the DA(α,β) method and by the NMSDE method. The true parameter values are not known and we cannot use the R package *MsdeParEst* since the real weight data from the 16029 animals presents a different age

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

	True	**DA(α,β)**	**MPE(α,β)**	**NMSDE**
			50 animals	
$h^{-1}(\mu)$	632.70	648.87 ± 33.04	612.27	646.62 ± 32.22
θ	0.15	0.0392 ± 0.1214	0.2594	–
λ	1.43	1.2265 ± 0.0600	1.2703	1.2259 ± 0.0599
ω	0.30	0.2164 ± 0.0297	< 0.0001	–
σ	0.33	0.3418 ± 0.0096	0.3059	0.3474 ± 0.0099
LL_X		−11819.3	−11713.8	−11868.8
			500 animals	
$h^{-1}(\mu)$	632.70	651.52 ± 10.06	634.72	655.81 ± 9.52
θ	0.15	0.0944 ± 0.0130	0.2087	–
λ	1.43	1.3216 ± 0.0209	1.3061	1.3056 ± 0.0191
ω	0.30	0.1809 ± 0.0124	0.0002	–
σ	0.33	0.3407 ± 0.0030	0.3114	0.3458 ± 0.0031
LL_X		−119331.8	−120800.7	−119620.8

vector of observations for each animal. Analyzing the results of Table 9, we can conclude that, despite the database being large and heterogeneous, the two estimation methods provide very similar estimates. According to our findings obtained above, it is interesting to note that the DA(α,β) method presents clear evidence of random effects on the α parameter and of strong random effects on the β parameter. Although we cannot perform a likelihood ratio test to reject the null hypothesis $\theta = \omega = 0$ because the displayed LL_X value for the DA(α,β) method is an approximation, the difference of LL_X values between this method and the NMSDE method (which displays an exact LL_X value) is so striking to leave doubts about the random nature of α and β.

3.5. Recommendations

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

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

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

	True	**DA**(α,β)	**MPE**(α,β)	**NMSDE**	**Exact**(α)
			50 animals		
$h^{-1}(\mu)$	632.70	641.87 ± 30.74	639.16	648.94 ± 26.51	640.20 ± 36.01
θ	0.15	0.1152 ± 0.0284	0.1446	–	0.1437 ± 0.0464
λ	1.43	1.4173 ± 0.0627	1.3501	1.3939 ± 0.0596	1.4385 ± 0.0608
ω	0	0.0637 ± 0.1213	< 0.0001	–	–
σ	0.33	0.3284 ± 0.0093	0.3051	0.3318 ± 0.0094	0.3268 ± 0.0095
LL_X		−11892.8	−11767.1	−11905.6	−11659.9
			500 animals		
$h^{-1}(\mu)$	632.70	635.46 ± 9.56	635.62	636.59 ± 8.35	644.44 ± 37.87
θ	0.15	0.1197 ± 0.0085	0.1537	–	0.1493 ± 0.0484
λ	1.43	1.4037 ± 0.0188	1.3417	1.3781 ± 0.0.0189	1.3957 ± 0.0610
ω	0	< 0.0001	0.0064	–	–
σ	0.33	0.3312 ± 0.0030	0.3067	0.3343 ± 0.0030	0.3308 ± 0.0096
LL_X		−118752.5	−120099.7	−118878.5	−11667.9

if one just uses the DA(α,β) method.

In the case the preliminary analysis indicates that $\omega = 0$ (by having a very small estimate of this parameter), we should then apply the Exact(α) method ([**34**]), which is designed for the case of α being the only random parameter and uses the exact log-likelikelihood function (3.8), to get better estimates of all parameters.

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

Finally, in the case where the preliminary analysis indicates $\omega = 0$ and $\theta = 0$ (i.e., the dataset only has fixed parameters), we can apply either the most appropriate and classical NMSDE method ([**13**, **15**, **34**]), designed for the fixed-effects model and using the exact log-likelihood function (3.6), or the DA(α,β) method (which gives basically the same estimates).

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

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

	True	**DA**(α,β)	**MPE**(α,β)	**NMSDE**
			50 animals	
$h^{-1}(\mu)$	632.70	619.32 ± 23.98	613.94	618.43 ± 24.13
θ	0	0.0260 ± 0.1038	0.1026	–
λ	1.43	1.4725 ± 0.0730	1.3944	1.4591 ± 0.0618
ω	0.30	0.2043 ± 0.0359	0.0002	–
σ	0.33	0.3323 ± 0.0094	0.3036	0.3365 ± 0.0096
LL_X		−11881.2	−11765.4	−11902.1
			500 animals	
$h^{-1}(\mu)$	632.70	640.71 ± 8.45	626.32	639.20 ± 8.77
θ	0	< 0.0001	0.1602	–
λ	1.43	1.3451 ± 0.0215	1.3213	1.3497 ± 0.0191
ω	0.30	0.1923 ± 0.01051	< 0.0001	–
σ	0.33	0.3360 ± 0.0030	0.2947	0.3399 ± 0.0031
LL_X		−118814.8	−120401.4	−119038.4

3.6. Conclusion

To describe the individual growth of animals in a randomly varying environment, a general stochastic differential equation (SDE) model was used and, in order to take into account that the model parameters may vary from animal to animal, which, for instance, occurs due to different individual genetical and other characteristics of the animals, we have consider SDE mixed models. Here we have studied the SDE mixed model where both parameters included in the drift term, the parameter α (asymptotic modified weight) and the growth parameter β, were assumed to be Gaussian distributed.

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

To evaluate the performance of the proposed method on parameter estimation we have used simulated datasets from different mixed-effects models (with only α random, with only β random, and with both parameters random) and from a fixed-effects model (both α and β fixed), with different numbers of animals (5000, 500 and 50), datasets in which all animals were weighted at the same ages so that we could use all the estimation methods considered here

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

	True	**DA**(α,β)	**MPE**(α,β)	**NMSDE**
			50 animals	
$h^{-1}(\mu)$	632.70	624.23 ± 24.67	624.32	624.23 ± 24.13
θ	0	< 0.001	0.0013	–
λ	1.43	1.4465 ± 0.0616	1.3623	1.4465 ± 0.0616
ω	0.30	< 0.001	< 0.0001	–
σ	0.33	0.3362 ± 0.0096	0.3136	0.3362 ± 0.0096
LL_X		−11901.2	−11762.8	−11901.0
			500 animals	
$h^{-1}(\mu)$	632.70	628.22 ± 7.81	628.15	628.21 ± 7.81
θ	0	< 0.0001	0.0196	–
λ	1.43	1.4331 ± 0.0193	1.3511	1.4330 ± 0.0191
ω	0	0.0135 ± 0.1381	< 0.0001	–
σ	0.33	0.3309 ± 0.0030	0.2996	0.3309 ± 0.0030
LL_X		−118686.3	−119963.8	−118686.3

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

	DA(α,β)	**NMSDE**
$h^{-1}(\mu)$	629.70 ± 6.30	631.27 ± 6.38
θ	0.0003 ± 0.0104	–
λ	1.4261 ± 0.0122	1.4198 ± 0.0116
ω	0.1998 ± 0.0050	–
σ	0.3273 ± 0.0016	0.3398 ± 0.0017
LL_X	−368732.10	−369753.20

and compare them. Since in real life we usually do not know which are the random parameters, wrong assumptions may be made on that issue. In order to evaluate the consequences of wrong assumptions on what parameters are random, we have included in the comparisons also the methods designed under such assumptions. We have compared the DA method (in its variants of assuming both or just one parameter random) with an existing method specifically designed for mixed-effects models (also with the same variants), referred here as the MPE method, which is provided by the R package *MsdeParEst*. We have also included in the comparisons the estimation methods for which we have exact log-likelihood expressions in closed-form, namely the NMSDE method and, in some cases, the Exact(α) method, which are designed, respectively, for the fixed-effects model and the mixed-effects model with only α random.

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

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

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

CHAPTER 4

Stochastic differential equations mixed model for individual growth with the inclusion of genetic characteristics

Nelson T. Jamba
Patrícia A. Filipe
Gonçalo Jacinto
Carlos A. Braumann

Statistics, Optimization and Information Computing, 12(2), 298–309, 2024

Abstract

In early work we have studied a class of stochastic differential equation (SDE) models, for which the Gompertz and the Bertalanffy-Richards stochastic models are particular cases, to describe individual growth in random environments, and applied it to model cattle weight evolution using real data. We have started to work on these type of models considering that the model parameters are fixed, i.e. the same for all the animals. Aiming to incorporate variability among individuals, we consider that the model parameters can be random variables, resulting in SDE mixed models. In additon, here we consider SDE mixed models, allowing the parameters to be random and propose to incorporate each animal's genetic characteristics considering the transformed animal's weight at maturity to be a function of its genetic values. The main objective is to extend the SDE mixed model to the more realistic case where the individual genetic value becomes an important component in the estimated growth curve. For the estimation of the model parameters we have used maximum likelihood estimation theory. Estimates and asymptotic confidence intervals of the parameters are presented. A comparison with SDE non-mixed model and SDE mixed model without the inclusion of genetic characteristics is shown with the conclusion that the incorporation of some genetic characteristics in the model parameters improves the estimation of the animal's growth curve.

Keywords genetic traits; individual growth; mixed models; stochastic differential equations.

4.1. Introduction

In the last four decades, Stochastic Differential Equations (SDE) have been studied and applied to individual growth modeling by several authors using SDE non-mixed models [**2**, **13**, **15**, **35**, **40–43**] where, for instance, authors have modeled the growth dynamic of animals, trees or glucose, and where the parameters of the SDE have fixed, although unknown, values that need to be estimated from data. When the model parameters are considered random, SDE mixed models are used to model individual growth, with applications in various fields, such as animal growth and pharmacokinetics, among others [**14**, **19**, **21**, **22**, **26**, **36**, **44**].

Considering M animals, we have used the following general SDE model

$$dY_i(t) = \beta\,(\alpha - Y_i(t))\; dt + \sigma\, dW_i(t), \quad Y_i(t_{i,0}) = y_{i,0}, \quad i = 1, \ldots, M, \tag{4.1}$$

where $Y_i(t) = h(X_i(t))$ is the modified size by the transformation h, a monotonous continuously differentiable function (which we assume 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 $t_{i,0}$ (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, such as the monomolecular model for $h(x) = x$, the Bertalanffy-Richards model for $h(x) = x^c$ ($c > 0$), the Gompertz model for $h(x) = \ln x$ and the logistic model for $h(x) = x^{-1}$ [**2**, **15**]. We refer to model (4.1) as SDE non-mixed model, when the parameters α, β and σ are fixed values common to all individuals. However, it is natural to think that different individuals might have different values for these parameters, due to their individual characteristics. To take this into account we have studied SDE mixed models, where in in model (4.1) we have considered that the parameters α and/or β were random variables [**14**,**37**].

The maximum likelihood methods are the commonly used methods to estimate the model parameters. Unlike the SDE non-mixed models, where it is usually possible to obtain a closed-form expression for the likelihood function, that is not always the case when SDE mixed models are used. Recently we have made some developments in estimation methods for SDE mixed models, when trend parameters can be considered random, [**14**, **37**], using approximation methods to solve estimation difficulties in SDE mixed models [**22**, **26**, **36**]. In our recent work, we have introduced the delta approximation method (DA method), a new approach inspired by the classical statistical delta method, and evaluated its performance with existing approximation methods, such as the Laplace approximation method (LA method) or other approximation methods available in *MsdeParEst* and *mixedsde R* packages. The DA method has an advantage over the available *R* packages by not requiring a common time vector of observations for all individuals.

The inclusion of genetic values in animal growth allows breeders to develop targeted breeding objectives and selection strategies. By prioritizing growth-related traits based on their estimated genetic values, breeders can focus on enhancing specific aspects of animal growth. This approach ensures that breeding efforts are directed towards maximizing growth potential, efficiency, and overall productivity.

In this work we will include a genetic value of the animals in the mixed model, allowing that each animal has a different growth curve influenced by its genetic information. Here, we will consider the case where the asymptotic weight at maturity is random, allowing us to study the influence of genetic values on this parameter of the SDE mixed model, particularly whether they provide a significant contribution to the precision of the growth curve. Comparing the parameters estimation using the SDE non-mixed model (parameters are fixed values, common to all individuals), the SDE mixed model without the inclusion of genetic values (parameters, in this case α, vary randomly from animal to animal) and the SDE mixed model with inclusion of the genetic values (parameter α is considered random and writen as a linear function of the animal's genetic values), we can evaluate if a random effect on the asymptotic weight at maturity is needed, and if so if the animal's specific genetic information would improve the parameters estimation. A more precise parameters estimation, and as a consequence, a more precise growth curve, will improve the farmers profit in livestock production.

We illustrate the results using real cattle weight (in kg) data of 10843 animals provided by the Associação de Criadores de Bovinos Mertolengos (ACBM), which performs the growing and finishing phases of young Mertolenga breed males, and by associated breeders, whose agricultural holdings are located in the Alentejo region in Portugal. Each animal has weight measurements taken at different age instants, varying from animal to animal, contrary to what is usually assumed in estimation methodologies.

The data regarding the genetic information of the animals results from the genetic evaluation of the Mertolenga cattle breed that was carried out at the National Institute of Agricultural and Veterinary Research (INIAV). Seven genetic traits were considered: growth capacity, maternal capacity, calving interval, average daily weight gain, carcass weight per day of age, feed conversion index and reproductive longevity. Seven models were fitted, each including a different genetic trait. We then evaluate which genetic trait results in an increasing performance on the model parameter estimates of the individuals.

We begin by describing the SDE mixed models with random α, in section 4.2, and discuss their main properties. Then, in section 4.3, we extend the SDE mixed models to incorporate the genetic values. In section 4.4, we apply our model to real cattle weight data and compare its performance with other SDE models. Finally, in section 4.5, we summarize the main findings and conclusions of our study.

4.2. Stochastic differential equations models

Let us consider M individuals, where $X_i(t)$ denotes the size of the i^{th} individual ($i = 1, \ldots, M$) at age t. In our application, the size refers to the animal's weight, but in other cases, it can denote another growth measure of the individual. We will consider a modified size $Y_i(t) = h(X_i(t))$, $(i = 1, \cdots, M)$, and describe the evolution of individual growth through an SDE of the form (4.1). For the real cattle weight data we used here for illustration, the transformation $h(x) = \ln x$, that leads to the Stochastic Gompertz model. revealed to be the most appropriate and so we will use it in the application. The solution of this SDE is an ergodic diffusion process with drift and diffusion coefficients $\beta\,(\alpha - y_i)$ and σ^2, respectively. The solution of (4.1), given by

$$Y_i(t) = \alpha + e^{-\beta(t-t_{i,0})}(y_{i,0} - \alpha) + \sigma e^{-\beta t} \int_{t_{i,0}}^{t} e^{\beta s} dW(s), \tag{4.2}$$

follows a Gaussian distribution with mean $\alpha + e^{-\beta(t-t_{i,0})}(y_{i,0} - \alpha)$ and variance $\frac{\sigma^2}{2\beta}(1 - e^{-2\beta(t-t_{i,0})})$ (see, for instance, [**15**, **27**]).

Let $t_{i,j}$ ($i = 1, ..., M$, $j = 1, ..., n_i$) be the age of the j^{th} weight measurement of the i^{th} animal and let $Y_{i,j} = Y_i(t_{i,j}) = h(X_i(t_{i,j}))$ be the corresponding modified weight according to model (4.1). For each animal i ($i = 1, \ldots, M$), denote by $\mathbf{t}_i = (t_{i,0}, t_{i,1}, \ldots, t_{i,n_i})$ its age vector of observations (which may differ from individual to individual), by $\mathbf{Y}_i = (Y_{i,0}, Y_{i,1}, ..., Y_{i,n_i})$ the corresponding vector of modified sizes 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 for $Y_{i,j}$ conditioned on $Y_{i,j-1} = y_{i,j-1}$, the transition distribution for animal i is Gaussian:

$$Y_{i,j}|\,(Y_{i,j-1} = y_{i,j-1}) \;\sim\; N\left(\alpha + \left(y_{i,j-1} - \alpha\right) E_{i,j}^{\beta}\,,\; \frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})\right), \tag{4.3}$$

with $i = 1, \cdots, M$, $j = 1, \cdots, n_i$. We have applied the maximum likelihood estimation method, [**14**, **15**, **37**], to estimate the parameter vector $\mathbf{p} = (\alpha, \beta, \sigma)$. From (4.3), using the fact that $Y_i(t)$ is a Markov process, we know that, given $Y_{i,0} = y_{i,0}$ (assumed known), the $\mathbf{Y_i}$ joint probability density function for individual i takes the form

$$\begin{aligned} p_{\mathbf{Y_i}}(\mathbf{y}_i|\alpha, \beta, \sigma) &= \prod_{j=1}^{n_i} \frac{\exp\left(-\frac{1}{2}\frac{\left(y_{i,j}-\alpha-(y_{i,j-1}-\alpha)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1-E_{i,j}^{2\beta})}\right)}{\sqrt{2\pi\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}} \qquad (4.4) \\ &= \frac{\exp\left(-\frac{1}{2}\sum_{j=1}^{n_i} \frac{\left(y_{i,j} - \alpha - (y_{i,j-1} - \alpha)E_{i,j}^{\beta}\right)^2}{\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}\right)}{\prod_{j=1}^{n_i} \sqrt{2\pi\frac{\sigma^2}{2\beta}(1 - E_{i,j}^{2\beta})}}, \quad i = 1, \ldots, M, \end{aligned}$$

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

$$L(\alpha, \beta, \sigma) = \prod_{i=1}^{M} p_{\mathbf{Y}_i}(\mathbf{y}_i|\alpha, \beta, \sigma). \tag{4.5}$$

The maximum likelihood estimate of the parameter vector **p** is obtained by maximization of (4.5) or of the log-likelihood function

$$\begin{aligned} LL_Y(\alpha,\beta,\sigma) &= \ln L(\alpha,\beta,\sigma) = \sum_{i=1}^{M} -\frac{n_i}{2}\ln(2\pi) \\ &- \frac{n_i}{2}\sum_{i=1}^{M}\ln\left(\frac{\sigma^2}{2\beta}\right) - \sum_{i=1}^{M}\frac{1}{2}\sum_{j=1}^{n_i}\ln\left(1 - E_{i,j}^{2\beta}\right) \\ &- \frac{\beta}{\sigma^2}\sum_{i=1}^{M}\sum_{j=1}^{n_i}\frac{\left(y_{i,j} - \alpha - \left(y_{i,j-1} - \alpha\right)E_{i,j}^{\beta}\right)^2}{1 - E_{i,j}^{2\beta}}. \end{aligned} \tag{4.6}$$

We have described the general SDE model (4.1) for the complete growth curve of the animals where the model's parameters α, β, and σ are assumed common to all individuals. In [**14**], we consider a generalization to SDE mixed models to take into account that different animals, due to their individual characteristics, may have different values of the parameters.

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

$$\begin{aligned} L(\mathbf{\Lambda}, \mathbf{\Psi}) &= \prod_{i=1}^{M} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{\Lambda}, \mathbf{\Psi}) \\ &= \prod_{i=1}^{M} \int_{\mathbb{R}^d} p_{\mathbf{Y}_i}(\mathbf{y}_i|\mathbf{b}_i, \mathbf{\Lambda}) p_B(\mathbf{b}_i|\mathbf{\Psi})\, d\mathbf{b}_i. \end{aligned} \tag{4.7}$$

The case of a single random parameter $\mathbf{b}_i = (\alpha_i)$ with $\alpha_i \sim N(\mu, \theta^2)$ has already been studied. When we have the special situation of an age vector of observations $\mathbf{t_i} \equiv \mathbf{t} = (t_0, t_1, \ldots, t_n)$, $i = 1, \ldots, M$ common to all animals, we can see it in [**21**]. For the general situation, with no such restrictions, we can see it in [**37**] and [**14**]. In this case, it is possible to explicitly compute the integral in the likelihood function, resulting in a final closed-form expression for this function. This is shown in [**14, 37**], where the log-likelihood function for all animals is given by

$$\begin{aligned} LL_Y(\mu,\theta,\beta,\sigma) = &\sum_{i=1}^{M}\left(-\frac{n_i}{2}\ln(2\pi) - \frac{n_i}{2}\ln\left(\frac{\sigma^2}{2\beta}\right) - \frac{1}{2}\sum_{j=1}^{n_i}\ln\left(1 - E_{i,j}^{2\beta}\right) - \frac{1}{2}\ln(D_i)\right. \\ &- \tfrac{\beta}{D_i\sigma^2}\sum_{j=1}^{n_i}\frac{\left(y_{i,j} - \mu - \left(y_{i,j-1} - \mu\right)E_{i,j}^{\beta}\right)^2}{1 - E_{i,j}^{2\beta}} + \frac{2\beta^2\theta^2}{D_i\sigma^4}\left(\sum_{j=1}^{n_i}\frac{\left(y_{i,j} - y_{i,j-1}E_{i,j}^{\beta}\right)}{1 + E_{i,j}^{\beta}}\right)^2 \\ &\left.- \tfrac{\beta(D_i-1)}{D_i\sigma^2}\left(\sum_{j=1}^{n_i}\frac{\left(y_{i,j} - y_{i,j-1}E_{i,j}^{\beta}\right)^2}{1 - E_{i,j}^{2\beta}}\right)\right), \end{aligned} \tag{4.8}$$

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

For the random β case, $\mathbf{b}_i = (\beta_i)$, it is not possible to obtain a closed-form expression for the likelihood function, and we need to solve the integral that appears in the likelihood function (4.7) through approximation methods. In [**14**] we propose a method to approximate the likelihood function when both α and β are random variables, the delta method. This approach gives simple and closed-form approximation formulas, making it easy to apply. This method also has the advantage of not being restricted to the cases where the age vector of the observations is the same for all trajectories nor does it require equidistant ages of observation, assumptions quite common in most estimation techniques. In [**37**] the same method is used but for the particular cases where just one of the parameters α or β is considered random, and a comparison of our new proposed method with the known Laplace approximation method is performed.

Here, we consider a generalization of SDE mixed models that allows one to incorporate each animal's genetic value with the aim of obtaining a more realistic model and improve the estimates of the model parameters at the individual level.

4.3. Mixed models with genetic values

It is natural to assume that the growth of an animal can depend on its individual characteristics, in particular on the animal's genetic information. We will develop the particular case where the parameter α, which represents the transformed asymptotic maturity weight of the animal, is considered random and can incorporate the animals genetic values. Consider that the genetic characterization of each animal is determined by K genetic values. In the first stage, we are interested in identifying the genetic values that can improve the estimation of the animal's weight at maturity and can lead to a significantly better model. So, for simplicity, we will consider incorporating one genetic value at a time ($K = 1$).

Denoting by g_i the genetic value of animal $i, (i = 1, \ldots, M)$, let us consider α_i to be defined as a linear function of the genetic value in the form of

$$\alpha_i = c_0 + c_1 g_i + \delta_i, \tag{4.9}$$

where δ_i are independent and identically distributed with $\delta_i \sim \mathcal{N}(0, \sigma_\delta^2)$, $(i = 1, \ldots, M)$. Then α_i follows a Gaussian distribution with mean $c_0 + c_1 g_i$ and variance σ_δ^2.

Using the parameter α_i written as a linear function of the genetic value, we can adapt the model presented in the previous section, but instead of $\alpha_i \sim N(\mu, \theta^2)$, we will have $\alpha_i \sim N(c_0 + c_1 g_i, \sigma_\delta^2), (i = 1, \ldots, M)$. In fact, using the same simple math used to derive expression (4.8), i.e. replacing in (4.7) the Gaussian density with mean $\mu_i = c_0 + c_1 g_i$ and standard deviation σ_δ

for animal i and solving the resulting integral, we end up obtaining a closed-form expression for the log-likelihood for all animals, similar to the expression (4.8), which will be given by

$$\begin{aligned} LL_Y(c_0, c_1, \sigma_\delta, \beta, \sigma) = \sum_{i=1}^{M} \Bigg(& -\frac{n_i}{2} \ln (2\pi) - \frac{n_i}{2} \ln\left(\frac{\sigma^2}{2\beta}\right) - \frac{1}{2} \sum_{j=1}^{n_i} \ln\left(1 - E_{i,j}^{2\beta}\right) \\ & -\tfrac{1}{2} \ln (D_i) - \tfrac{\beta}{D_i\sigma^2} \sum_{j=1}^{n_i} \frac{\left(y_{i,j} - (c_0 + c_1 g_i) - \left(y_{i,j-1} - (c_0 + c_1 g_i)\right) E_{i,j}^{\beta}\right)^2}{1 - E_{i,j}^{2\beta}} \\ & + \tfrac{2\beta^2\sigma_\delta^2}{D_i\sigma^4} \left(\sum_{j=1}^{n_i} \frac{\left(y_{i,j} - y_{i,j-1} E_{i,j}^{\beta}\right)}{1 + E_{i,j}^{\beta}} \right)^2 - \tfrac{\beta(D_i-1)}{D_i\sigma^2} \left(\sum_{j=1}^{n_i} \frac{\left(y_{i,j} - y_{i,j-1} E_{i,j}^{\beta}\right)^2}{1 - E_{i,j}^{2\beta}} \right) \Bigg), \end{aligned} \tag{4.10}$$

with $D_i = \frac{2\beta\sigma_\delta^2}{\sigma^2} \sum_{j=1}^{n_i} \dfrac{1 - E_{i,j}^{\beta}}{1 + E_{i,j}^{\beta}} + 1$, where the models parameters are now c_0, c_1, σ_δ, β, σ.

In order to effectively verify if a genetic value can improve the estimation of the animal's weight, a comparison between the SDE mixed model and the SDE mixed model with the inclusion of the genetic value is made (using a likelihood ratio and also the Bayesian information criterion (BIC) and the Akaike information criterion (AIC)). If the two models show no significant differences, the genetic value does not have a significant effect and we may conclude that c_1 is not significantly different from 0, i.e. the mean value of the parameter α_i is common to all animals and equal to $\mu = c_0$. On the other hand, if a significant difference between models is detected, it reveals that the inclusion of the genetic value has a significant effect on the estimation of the animal's weight at maturity, i.e., c_1 is significantly different from 0 and the mean value of α_i is different for each animal and given as a linear function of the specific genetic value g_i as $\mu_i = c_0 + c_1\, g_i$.

4.4. Real data application

In this section we use the maximum likelihood method to estimate the parameters for the stochastic Gompertz model using real cattle weight data, for the case where the parameter α is considered random and a function of the genetic values. We compare the results with the ones obtained when no genetic value is included, and using the likelihood ratio test (LRT) we compare both models. In this way, we infer if the inclusion of the genetic values improves the estimation of the model parameters. Since we have a different set of animals for each genetic value, we also present the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) to compare different models. The database provided by ACBM has information on the weights and ages of 10843 animals, totaling 69782 observations, with a minimum of 3 weight observations and a maximum of 33 observations.

4.4.1. The genetic information

Genetic evaluation of the Mertolenga cattle breed was carried out at the National Institute of Agricultural and Veterinary Research, from all the data collected by ACBM, namely, records of genealogies, births, weights and carcass information, taking into account the following characteristics: growth capacity until weaning (GC), maternal capacity until weaning (MC), calving interval (CI), average daily weight gain (WG), carcass weight per day of age (CW), feed conversion index (FC) and reproductive longevity (RL). The estimation of the genetic values of each animal for these seven types of traits, takes into account the animal's performance, if known, and the performances of all its relatives (ascendants, descendants and collateral's), taking into consideration the various environmental effects that affect the respective trait. Currently, at an international level and in several livestock species (cattle, sheep, pigs poultry, goats, horses, etc.) the use of BLUP - Animal Model for the genetic evaluation is widespread [**45–47**].

Not all animals have the same age at the available genetic evaluations. Most of the animals had genetic evaluations taken at least at two different instants: at birth and at entrance on ACBM for the finishing phase. We have compared the estimation results considering, for each animal and each genetic characteristic, the first genetic value available, the genetic value closest to 7 months of age (mean age of entrance in ACBM for the growing and finishing phases), and also the mean of the mother and father genetic values. Comparing the estimates obtained using these three different genetic information, we have concluded that, using the values from the genetic evaluation closest and prior to the time of entrance in ACBM, results in the best choice in terms of the likelihood of the mixed stochastic differential equations model. That is also the most coherent decision, since the final objective is to give a tool to ACBM to better estimate the animal's growth and accurately estimate the profit obtained by raising and selling the animal. Using the most recent animal's genetic information available by the time of entrance on ACBM will accomplish better that purpose.

Therefore, in our analysis, we will use for each animal and genetic trait, the animal's available genetic value closest and prior to the mean entrance age at ACBM (approximately 7 months). Not all animals had the seven genetic values available. For that reason, depending on the genetic value used in the model, the number of observations used to estimate the parameters vary.

The genetic value of an animal for a given trait represents the value of that animal as a breeder (expressed in the respective units of measurement, i.e. kg, days, %, etc.) and should be interpreted as the genetic superiority or inferiority for the trait relative to the population average. The accuracy of the genetic value estimate is also available and gives us an idea of how confidently the genetic value of the animal for a given characteristic was estimated. However, it is not an indicator of the genetic potential of the animal. The more information about the animal (e.g., various records of birth intervals) and about their relatives (mother, sisters, daughters, grandparents, etc.) there are, the more accurate the estimate of their breeding value will be.

Regarding the seven genetic values in terms of genetic potential:

- For MC, RL, GC, WG and CW, the higher their values are, preferably positive, the better.
- For CI and FC, the lower their values are, preferably negative, the better.

In Table 1 we present the minimum, the 25^{th} quantile (Q1), the median, the mean, the 75^{th} quantile (Q3), the maximum, and the standard deviation as well as the number of animals with the respective genetic value available and the total number of weight observations available for that group of animals. Since, to achieve a high genetic potential, some of the genetic values should be high and positive while others should be small and negative, and all genetic values have a very high range (positive and negative) values, it is clear that some animals would have a greater genetic potential that others.

The first five genetic values presented in Table 1 are the ones that should be as high as possible and positive. The last two genetic values should be as small as possible and negative. At the moment of the evaluation of the genetic values, they are usually centered to obtain an evaluated population with a zero mean and making it easy to determine which animals have higher potential and which have lower potential. In the results presented in Table 1, the mean is not 0 because the animals used in the parameter estimation had their genetic values evaluated at a wide variety of evaluation times and so belong to different evaluated populations. The CW, WG and CI values have higher variability.

TABLE 1. Descriptive statistics for the seven genetic values of the genetic evaluation: minimum (min), 25^{th} quantile (Q_1), median, mean, 75^{th} quantile (Q_3), maximum (max), standard deviation (s.d.), number of animals (M) and total number of weight observations (n).

Genetic value	min	Q_1	median	mean	Q_3	max	s.d.	M	n
MC	-22.82	-3.45	0.63	0.24	4.56	19.92	6.17	7981	45014
GC	-39.44	-5.07	-0.02	-0.17	5.08	24.56	7.78	5186	34526
RL	-11.66	-1.52	0.99	1.04	3.32	19.88	3.82	4234	21317
WG	-45.00	-1.00	8.63	11.54	21.51	106.50	17.30	8302	45424
CW	-99.27	-18.94	-6.18	-6.48	6.53	69.77	21.23	5184	26742
CI	-67.15	-9.27	-0.75	-0.28	8.89	78.48	16.14	7960	44860
FC	-0.73	-0.18	-0.07	-0.10	0.00	0.47	0.15	6952	36415

4.4.2. Parameters estimates

In Table 2 we can observe the parameter estimates using the stochastic differential equations non-mixed model (NMSDE) - obtained by maximization of (4.6) - and the mixed model with random α without the genetic characters (MSDE) - obtained by maximization of (4.8). We can observe that the estimates of all parameters are quite similar in both models. The approximate 95% confidence bands based on the inverse of the empirical Fisher information matrix in both models are also very similar. We present the maximum values of the log-likelihood in terms of the actual weight, LL_X. Comparing the two models through the LRT, we can conclude that they differ significantly ($p \leq 0.001$). This result shows that assuming the parameter α as a random parameter should be considered as a better model for the growth data of the Mertolenga cattle breed.

TABLE 2. Estimation results for the NMSDE and for the MSDE models assuming random α. The table shows the values of the maximum likelihood estimates and corresponding approximate 95% asymptotic confidence bands, as well as the corresponding value of the log-likelihood. Besides showing the results for the modified weight α, we present them for the weight $A = h^{-1}(\mu)$ (for the NMSDE model $\mu = \alpha$). The AIC and BIC values are also presented. The magnitude of the p-values of the LRT test between the NMSDE and MSDE are shown: ^{ns} ($p > 0.05$); * ($0.01 < p \leq 0.05$); ** ($0.001 < p \leq 0.01$); *** ($p \leq 0.001$).

	NMSDE	**MSDE**
$\boldsymbol{\mu}$	6.5867 ± 0.0142	6.5889 ± 0.0148
A	725.36 ± 10.28	726.98 ± 10.74
$\boldsymbol{\theta}$	$-$	0.1428 ± 0.0092
$\boldsymbol{\beta}$	1.2914 ± 0.0138	1.3002 ± 0.0143
$\boldsymbol{\sigma}$	0.3078 ± 0.0018	0.3010 ± 0.0019
$\mathbf{LL_X}$	-264443.2	-264281.0^{***}
AIC	528892.4	528570.0
BIC	528919.4	528606.0

Let us now explore the mixed model with random α with the inclusion of genetic values (MSDEG). We fit seven MSDEG models, one for each genetic trait, and compared them with the MSDE and NMSDE models using the same datasets (for each trait, the number of animals M and the number of weight observations n of the corresponding dataset can be seen on Table 1). The parameter estimates were obtained by maximization of (4.10), (4.8) and (4.6), respectively. First, we will compare the NMSDE models with the equivalent MSDE models, to evaluate the relevance of considering a random effect on the parameter α, and then we compare the MSDE models with the MSDEG models, to evaluate the effect of adding that specific genetic trait. Table 3 and 4 present the estimates obtained through the maximum likelihood estimation method and the 95% approximate confidence bands based on the inverse of the empirical Fisher information matrix. The AIC and BIC are also presented to compare the models.

In Table 3 we can observe that, for all the genetic traits, the differences between NMSDE and MSDE were significant at $p \leq 0.01$ or $p \leq 0.001$ and the AIC and BIC values were lower on these MSDE models when compared to the NMSDE models. This shows that a random effect should be present in all groups of animals.

When comparing the MSDEG and MSDE models of Table 3 for the genetic trait MC, we can observe that c_1 (the parameter that indicates the importance of the genetic trait on explaining the variability in α) has a small estimate not significantly different from 0. But we should be aware that such lack of significance was based on an approximation of the confidence interval of c_1 obtained from the inverse of empirical Fisher information matrix. The LRT, however, confirms the lack of significance of the maternal capacity until weaning. For the cases of the reproductive longevity and calvin interval, a significance exists with a p-value of the LRT lower than 0.1% (the AIC values are also lower for these two genetic traits). The maximum likelihood estimates from the NMSDE, the MSDE and the MSDEG models are very similar for all common parameters. For the models with random α, it is interesting to compare the standard deviation of the random effect, i.e. to compare the θ of the MSDE model with the σ_δ of the MSDEG model. With the exception of the non-significant MC trait, we see that the incorporation of the genetic trait, RL or CI, reduces the variability of the random effect.

In Table 4 for all genetic traits, the differences between NMSDE and MSDE were significant at $p \leq 0.001$. Also the AIC and BIC values were lower on the MSDE models, allowing us to conclude that for these groups of animals we should consider a random asymptotic weight at maturity. The maximum likelihood estimates from the NMSDE, the MSDE and the MSDEG models are very similar for all parameters. For the models with random α, it is interesting to compare the standard deviation of the random effect, i.e. to compare the θ of the MSDE model with the σ_δ of the MSDEG model. With the exception of the non-significant GC trait, we see that the incorporation of the genetic trait, Wg, CW or FC, reduces the variability of the random effect. Regarding the AIC values we can see that the MSDE models provides lower values than the NMSDE equivalent models, which in turn provides lower values that the equivalent MSDEG models.

In summary, comparing models with the same number of observations, the comparison between NMSDE and MSDE models shows that considering the parameter α to randomly vary among animals provides a very significant improvement, and the comparison between MSDE and MSDEG models shows a very significant improvement for the incorporation of some genetic traits (which help explain part of the variability in α), but not for others (which seems to have no or low influence on α). The genetic traits that are not significant are the maternal capacity until weaning (MC) and the growth capacity until weaning (GC). The genetic traits

Table 3. Estimation results for the NMSDE, MSDE and MSDEG models. The table shows the values of the maximum likelihood estimates and corresponding approximate 95% asymptotic confidence bands, as well as the corresponding value of the log-likelihood. Estimates for the actual asymptotic weight (in kg) are shown with $A = h^{-1}(\alpha)$, $A = h^{-1}(\mu)$ and $A = h^{-1}(c_0 + c_1 g_1)$, for NMSDE, MSDE and MSDEG, respectively. The AIC and BIC values are also presented. The magnitude of the p-values of the LRT test between the NMSDE and MSDE and between the MSDE and MSDEG are shown: ^{ns} ($p > 0.05$); * ($0.01 < p \leq 0.05$); ** ($0.001 < p \leq 0.01$); *** ($p \leq 0.001$).

	MC	RL	CI
	NMSDE		
$\boldsymbol{\alpha}$	6.5826 ± 0.0157	6.5892 ± 0.0281	6.2829 ± 0.0157
A	722.40 ± 11.32	727.17 ± 20.43	722.68 ± 11.36
$\boldsymbol{\beta}$	1.2938 ± 0.0154	1.3099 ± 0.0272	1.2930 ± 0.0154
$\boldsymbol{\sigma}$	0.3060 ± 0.0020	0.3395 ± 0.0033	0.3062 ± 0.0020
$\mathbf{LL_X}$	−202064.0	−96107.3	−201397.3
AIC	404134.0	192220.5	402800.7
BIC	404160.1	192244.4	402826.8
	MSDE		
$\boldsymbol{\mu}$	6.5814 ± 0.0163	6.5872 ± 0.0284	6.5816 ± 0.0164
A	721.55 ± 11.78	725.75 ± 20.61	721.69 ± 11.82
$\boldsymbol{\theta}$	0.1452 ± 0.0102	0.0707 ± 0.0313	0.1450 ± 0.0102
$\boldsymbol{\beta}$	1.3052 ± 0.0160	1.3143 ± 0.0277	1.3046 ± 0.0160
$\boldsymbol{\sigma}$	0.2989 ± 0.0021	0.3381 ± 0.0035	0.2991 ± 0.0021
$\mathbf{LL_X}$	-201924.8^{***}	-96104.7^{**}	-201259.7^{***}
AIC	403857.6	192217.3	402527.3
BIC	403889.2	192230.2	402514.0
	MSDEG		
$\mathbf{c_0}$	6.5826 ± 0.0164	6.5951 ± 0.0287	6.5799 ± 0.0163
$\mathbf{c_1}$	0.0010 ± 0.0011	−0.0059 ± 0.0026	0.0016 ± 0.0004
A	722.24 ± 11.839	727.18 ± 20.68	720.14 ± 11.74
$\boldsymbol{\sigma_\delta}$	0.1453 ± 0.0102	0.0679 ± 0.0325	0.1410 ± 0.0104
$\boldsymbol{\beta}$	1.3045 ± 0.0160	1.3134 ± 0.0277	1.3076 ± 0.01610
$\boldsymbol{\sigma}$	0.2989 ± 0.0021	0.3381 ± 0.0035	0.2993 ± 0.0021
$\mathbf{LL_X}$	-201923.2^{ns}	-96095.2^{***}	-201235.6^{***}
AIC	403856.3	192200.3	402481.1
BIC	403899.9	192240.2	402524.7

reproductive longevity (RL), calving interval (CI), average daily weight gain (WG), carcass weight per day of age (CW), and feed conversion index (FC) do have a very significant effect, RL and FC with a negative effect $c_1 < 0$ (i.e., α is high for low genetic values, preferably negative) and CI, WG and CW with a positive effect $c_1 > 0$ (i.e., α is high for high genetic values, preferably positive).

TABLE 4. Estimation results for the NMSDE, MSDE and MSDEG models. The table shows the values of the maximum likelihood estimates and corresponding approximate 95% asymptotic confidence bands, as well as the corresponding value of the log-likelihood. Estimates for the actual asymptotic weight (in kg) are shown with $A = h^{-1}(\alpha)$, $A = h^{-1}(\mu)$ and $A = h^{-1}(c_0 + c_1 g_1)$, for NMSDE, MSDE and MSDEG, respectively. The AIC and BIC values are also presented. The magnitude of the p-values of the LRT test between the NMSDE and MSDE and between the MSDE and MSDEG are shown: ^{ns} $(p > 0.05)$; * $(0.01 < p \leq 0.05)$; ** $(0.001 < p \leq 0.01)$; *** $(p \leq 0.001)$.

	NMSDE			
	GC	WG	CW	FC
α	6.5826 ± 0.0157	6.5884 ± 0.0165	6.6007 ± 0.0240	6.6241 ± 0.0210
A	722.40 ± 11.32	726.61 ± 12.01	735.60 ± 17.65	753.00 ± 15.84
β	1.2938 ± 0.0154	1.2998 ± 0.0161	1.2835 ± 0.0226	1.2553 ± 0.0192
σ	0.3060 ± 0.0020	0.3156 ± 0.0021	0.3308 ± 0.0028	0.3254 ± 0.0024
$\mathbf{LL_X}$	−202050.7	−204171.3	−120058.1	−163493.7
AIC	404107.3	408348.7	240122.2	326993.5
BIC	404133.4	408374.8	240146.8	327019.0
	MSDE			
μ	6.5814 ± 0.0163	6.5846 ± 0.0170	6.5932 ± 0.0247	6.6216 ± 0.0216
A	721.55 ± 11.78	723.86 ± 12.33	730.11 ± 18.01	751.15 ± 16.23
θ	0.1452 ± 0.0102	0.1256 ± 0.0117	0.1241 ± 0.0175	0.1181 ± 0.0157
β	1.3052 ± 0.0160	1.3107 ± 0.0167	1.2988 ± 0.0237	1.2644 ± 0.0198
σ	0.2989 ± 0.0021	0.3106 ± 0.0022	0.3265 ± 0.0030	0.3216 ± 0.0025
$\mathbf{LL_X}$	-201911.4^{***}	-204100.8^{***}	-120028.9^{***}	-163461.3^{***}
AIC	403830.9	408209.6	240065.7	326930.6
BIC	403864.9	408189.3	240081.6	326897.9
	MSDEG			
$\mathbf{c_0}$	6.5818 ± 0.0164	6.5667 ± 0.0175	6.5987 ± 0.0248	6.6040 ± 0.0218
$\mathbf{c_1}$	0.0004 ± 0.0008	0.0015 ± 0.0004	0.0010 ± 0.0005	−0.2223 ± 0.0534
A	721.64 ± 11.79	723.44 ± 12.31	729.4 ± 17.95	753.89 ± 16.34
σ_δ	0.1450 ± 0.0102	0.1215 ± 0.0120	0.1217 ± 0.0178	0.1131 ± 0.0163
β	1.3051 ± 0.0160	1.3087 ± 0.0166	1.2991 ± 0.0236	1.2608 ± 0.0198
σ	0.2989 ± 0.0021	0.3106 ± 0.0022	0.3265 ± 0.0030	0.3215 ± 0.0025
$\mathbf{LL_X}$	-201911.0^{ns}	-204073.1^{***}	-120020.4^{***}	-163427.9^{***}
AIC	403832.1	408156.1	240050.8	326865.9
BIC	403875.6	408199.7	240091.8	326908.4

4.5. Conclusion

Incorporating genetic values in the animal growth process is a valuable approach that can lead to significant improvements in livestock production. By considering the genetic potential of animals, breeders can select individuals with desirable traits, resulting in enhanced growth, productivity, and overall performance of the population.

We have studied a general class of SDE mixed models with the inclusion of genetic values 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) is assumed to be random and where the random effect is written as a linear function of the genetic values of the animal. We have used, in the numerical application to Mertolengo cattle data, the weight of the animals and the logarithmic transformation, which reveals to be the most appropriate to that data, leading to the stochastic Gompertz model. Our interest in using SDE mixed models with genetic values 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 with inclusion of genetic values (MSDEG) and compare each model with the correspondent ones without the inclusion of the genetic values (MSDE) and with the fixed effects model (NMSDE). We compare the models using the likelihood ratio test and also present the AIC and BIC values. The comparison between the SDE mixed model and the SDE non-mixed model clearly shows that, just by assuming the asymptotic weight at maturity of each animal to be a random variable, results in a highly significant improvement. The further importance of considering a genetic value was evaluated by comparing the corresponding MSDEG model with the MSDE model. The results revealed that the incorporation of the genetic values of reproductive longevity, calving interval, average daily weight gain, carcass weight per day of age or feed conversion index, provides a highly significant improvement, while the incorporation of the maternal capacity until weaning and the growth capacity until weaning genetic values shows non-significant improvement.

This study enables the identification of genetic values that contribute the most to describe the animal growth asymptotic size. By incorporating these genetic values, breeders can improve the efficiency and productivity of animal production systems, offering significant benefits to livestock production.

In future work, we intend to implement MSDE and MSDEG models to improve the optimization of the breeder's profit by raising an animal. We should estimate the optimal selling age and optimize the profit using SDE mixed models instead of SDE non-mixed models. We also intend to extend the SDE mixed model by letting the random growth parameter β depend on the genetic values. A combination of several genetic values simultaneously instead of just one at a time should also be addressed, either for the SDE mixed effect model with a random α and the SDE mixed model with a random β.

CHAPTER 5

Conclusions

The study of MSDE models has provided valuable insights into the analysis and modeling of dynamic systems subject to both deterministic and stochastic components. Throughout this research, we have explored the unique characteristics and applications of these models in animal growth, allowing a more realistic representation of growth dynamics accounting for unobserved factors, measurement errors, or random fluctuations in the system.

While fixed effects describe the average or general trend of the system, random effects capture the individual variations or heterogeneity among different units or subjects. This integration of fixed and random effects allows for a more comprehensive understanding of the underlying dynamics and enables the estimation of both population-level and individual-level effects.

We have studied a general class of MSDE 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 (OU) model, where the parameter α (asymptotic modified weight) and/or the growth parameter β (the rate of approach to maturity) are assumed to be random. We have worked with the logarithm of the weight, the transformation that reveals to be the most appropriate for our type of data, leading to the stochastic Gompertz model.

We apply the maximum likelihood estimation method to obtain the parameter estimates of the MSDE models. For this type of models, 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 issue. We have developed approximation methods used to obtain simpler approximate expressions for the integral involved in the likelihood function. We propose the DA method, a new approximation method we have developed, based in the classical delta method used in Statistics. We compare this new method with an already known approach, the LA method, for which we have explicitly derived the formulas to meet the needs of our real application (not available in the literature) and we have implemented numerically and also compared with available R packages (*mixedsde* and *MsdeParEst*). However these packages assume that the animals should have the weight taken at the same ages. Our newly proposed DA method and the derived LA method were implemented for datasets where animals can have different age vectors of observations, and simulated datasets with animals with the same age vector were created to compare the DA and LA methods with the methods proposed in the *mixedsde* and *MsdeParEst* packages.

These approximation methods have provided robust and efficient means for estimating the unknown parameters and evaluating model fit, enabling researchers to make more accurate inferences and predictions. We have developed MSDE models for three specific cases, namely, the random α case, the random β case and also the case where both α and β are random. The random parameters were assumed to follow a Gaussian distribution.

For the random α case, we have proven that the LA method leads to an exact closed form expression for the likelihood function. We also concluded that the DA method provides similar estimates of the parameters to those obtained by the R packages. For the random beta case, the results obtained by the two methods (LA and DA) are very similar and better than the obtained by the R packages. When both (α and β) are random, the DA method, for all simulated scenarios, with several numbers of observations (50, 500 and 5000), was able to identify randomness in the parameters, while the approach provided by the R packages has difficulty identifying the variability of parameter β. When applied to real cattle weight data, the DA method was also able to identify both parameters as random. The results showed good performance of the DA method with both random parameters compared to the methods available through the R packages. The DA and LA approximation have the advantage over the methods presented in the R packages for their ability to be used with different age vectors for each animal.

Finally in livestock production the genetic values of animals are very important elements to take into account in their growth process due to their genetic potential that allows the breeder to select animals with the characteristics desirable for productivity. We use real data of genetic value, in the MSDE models, in the case where the α parameter is considered random, and α is written as a linear function of the genetic values. The importance of considering a genetic value was evaluated by comparing the MSDEG model and the MSDE model. The results revealed that the incorporation of the genetic values of reproductive longevity, calving interval, average daily weight gain, carcass weight per day of age or feed conversion index, provides a highly significant improvement, while the incorporation of the maternal capacity until weaning and the growth capacity until weaning genetic values shows non-significant improvement.

In future work, we intend to implement MSDE and MSDEG models to improve the optimization of the breeder's profit by raising an animal. We intend estimate the optimal selling age and optimize the profit using MSDE models instead of NMSDE models. We also intend to extend the MSDE model by letting the random growth parameter β depend on the genetic values. Another research topic is to develop MSDE models, where the choice of the distribution for the positive random parameter β is e.g., log-normal. A combination of several genetic values simultaneously instead of just one at a time should also be addressed, either for the MSDE model with a random α and the MSDE model with a random β.

Appendix A: R codes for data Simulation

```
##---------------------------------##---------------------------------##
##  R Code to simulate Log of animal’s weights at constant instants   ##
##   from birth (with 28 kgs) until  X years old - monthly weights    ##
##---------------------------------##---------------------------------##

#################
set.seed(123)
X<-4
# montly weights
pm<-1/12
# 15 days weights
pq<-1/24
# ages at montly weights
Im<-seq(0,X, by=pm)
# ages at every 15 days weights
Iq<-seq(0,X, by=pq)


n<-length(Im)
nq<-length(Iq)

# birth weight
P00m<-28

m<-500 # Define de number of animals


#  random ALPHA
#alfa= 6.45;  theta<-0.15; beta= 1.43; sigma^2=0.33^2
a<-rnorm(m,6.45,0.1531)
exp(6.45)
b<-1.43
s2<-0.33^2

#Log weights - random alpha
PGm<-matrix(nrow=m,ncol=n)
PGm[,1]<-log(P00m)
for (i in 1:m){
for (j in 2:n){
PGm[i,j]<-rnorm(1,a[i]+(PGm[i,j-1]-a[i])*exp(-b*(Im[j]-Im[j-1])),
sqrt(((s2/(2*b))*(1-exp(-2*b*(Im[j]-Im[j-1]))))))
}}
```

```
randomalpha1<-rbind(Im,PGm)
#write.csv2(randomalpha1,file = "Data_randalpha_50anim.csv",row.names = FALSE)


# random  BETA
#alfa= 6.45; lambda= 1.43; omega<-0.30; sigma^2=0.33^2
a<-6.45
exp(6.45)
b<-rnorm(m,1.43,0.30)
s2<-0.33^2

#Log weights
PGm<-matrix(nrow=m,ncol=n)
PGm[,1]<-log(P00m)
for (i in 1:m){
for (j in 2:n){
PGm[i,j]<-rnorm(1,a+(PGm[i,j-1]-a)*exp(-b[i]*(Im[j]-Im[j-1])),
sqrt(((s2/(2*b[i]))*(1-exp(-2*b[i]*(Im[j]-Im[j-1]))))))
}}

randombeta1<-rbind(Im,PGm)
# write.csv2(randombeta1,file = "Data_randbeta_50anim.csv",row.names = FALSE)


#  Random ALPHA and BETA independents
#alfa= 6.45;  theta<-0.15; lambda= 1.43; omega<-0.30; sigma^2=0.33^2
a<-rnorm(m,6.45,0.1531)
exp(6.45)
b<-rnorm(m,1.43,0.30)
s2<-0.33^2

#Log weights
PGm<-matrix(nrow=m,ncol=n)
PGm[,1]<-log(P00m)
for (i in 1:m){
for (j in 2:n){
PGm[i,j]<-rnorm(1,a[i]+(PGm[i,j-1]-a[i])*exp(-b[i]*(Im[j]-Im[j-1])),
sqrt(((s2/(2*b[i]))*(1-exp(-2*b[i]*(Im[j]-Im[j-1]))))))
}}

randomAlphaBeta<-rbind(Im,PGm)
#write.csv2(randomAlphaBeta,file = "Data_rndAlphaBeta_50anim.csv",
row.names = FALSE)


#  Random ALPHA and BETA dependents
#alfa= 6.45;  theta<-0.15; lambda= 1.43; omega<-0.30; sigma^2=0.33^2
rho <- 0.88
mu1 <- 6.45; s1 <- 0.15
```

```
mu2 <- 1.43; s2 <- 0.3
s2<-0.33^2

# Parameters for bivariate normal distribution
mu <- c(mu1,mu2) # Mean
sigma <- matrix(c(s1^2, s1*s2*rho, s1*s2*rho, s2^2),2) # Covariance matrix

library(MASS)
alfa_beta <- mvrnorm(m, mu = mu, Sigma = sigma)
a<-alfa_beta[,1]
b<-alfa_beta[,2]

#Log weights
PGm<-matrix(nrow=m,ncol=n)
PGm[,1]<-log(P00m)
for (i in 1:m){
for (j in 2:n){
PGm[i,j]<-rnorm(1,a[i]+(PGm[i,j-1]-a[i])*exp(-b[i]*(Im[j]-Im[j-1])),
sqrt(((s2/(2*b[i]))*(1-exp(-2*b[i]*(Im[j]-Im[j-1]))))))
}}

randomAlphaBetaCorr<-rbind(Im,PGm)
#write.csv2(randomAlphaBetaCorr,file = "Data_rndAlphaBetaCorr.csv",
row.names = FALSE)

############################################################################
# No random effect simulation
#  ALPHA e BETA fixos
#alfa= 6.45;  theta<-0; lambda= 1.43; omega<-00; sigma^2=0.33^2
a<-rnorm(m,6.45,0)
exp(6.45)
b<-rnorm(m,1.43,0)
s2<-0.33^2

# Log weights
PGm<-matrix(nrow=m,ncol=n)
PGm[,1]<-log(P00m)
for (i in 1:m){
for (j in 2:n){
PGm[i,j]<-rnorm(1,a[i]+(PGm[i,j-1]-a[i])*exp(-b[i]*(Im[j]-Im[j-1])),
sqrt(((s2/(2*b[i]))*(1-exp(-2*b[i]*(Im[j]-Im[j-1]))))))
}}

randomAlphaBeta<-rbind(Im,PGm)
#write.csv2(randomAlphaBeta,file = "Data_fixAlphaBeta_50anim.csv",
row.names = FALSE)

############################################################################
# Simulation with estimates of exact  MCV
# Estimates from 16029 animals for SGM
```

```
mu<-6.45
exp(6.45)
theta<-0.1531
b<-1.432
s2<-0.33^2

# Log weights
PGam<-matrix(nrow=m,ncol=n)
PGam[,1]<-log(P00m)
for (i in 1:m){
for (j in 2:n){
PGam[i,j]<-rnorm(1,mu+(PGam[i,j-1]-mu)*exp(-b*(Im[j]-Im[j-1])),
sqrt(((s2/(2*b))*(1-exp(-2*b*(Im[j]-Im[j-1]))))+(theta^2)*
(1-exp(-b*(Im[j]-Im[j-1])))^2))
}}

DaAlea_m<-rbind(Im,PGam)

#write.csv2(DaAlea_m,file = "DadosG2_aAlea_mensais.csv", row.names = FALSE)

###LogPesos quinzenais
PGaq<-matrix(nrow=m,ncol=nq)
PGaq[,1]<-log(P00m)
for (i in 1:m){
for (j in 2:nq){
PGaq[i,j]<-rnorm(1,mu+(PGaq[i,j-1]-mu)*exp(-b*(Iq[j]-Iq[j-1])),
sqrt(((s2/(2*b))*(1-exp(-2*b*(Iq[j]-Iq[j-1]))))+(theta^2)*
(1-exp(-b*(Iq[j]-Iq[j-1])))^2))
}}

DaAlea_q<-rbind(Iq,PGaq)

write.csv2(DaAlea_q,file = "DadosG2_aAlea_quinzenais.csv", row.names = FALSE)

##---------------------------------##---------------------------------##
##                    Code to call the DA or LA methods              ##
##---------------------------------##---------------------------------##
##                                Data                                ##
##---------------------------------##---------------------------------##

## Dataset Simulated data
I <-read.csv("Mat_Time_500.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
P <-read.csv("Sim_Obs_500_alpha.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
P <-read.csv("Sim_Obs_500_beta.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
P <-read.csv("Sim_Obs_500_alphabeta.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
```

```
## Datasets real data paper Mathematics
I <-read.csv("Idades16029.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
P <-read.csv("Pesos16029.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data

## Datasets real data
I <-read.csv("Idades10843.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
P <-read.csv("Pesos10843.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
dim(I)
dim(P)
```

Appendix B: Estimates throught the delta approximation (DA) method

```
## Estimates throught the DA method
start_time <- Sys.time()
DA.fit(I, P,  0, c(6.3,1.3,0.3))
end_time <- Sys.time()
end_time - start_time

start_time <- Sys.time()
DA.fit(I, P,  1, c(6.5,0.13,1.11,0.28))
end_time <- Sys.time()
end_time - start_time

start_time <- Sys.time()
DA.fit(I, P,  11, c(6.59205,1.30,0.1835, 0.30))
end_time <- Sys.time()
end_time - start_time

start_time <- Sys.time()
DA.fit(I, P,  2, c(6.592,0.1,1.30,0.183, 0.3))
end_time <- Sys.time()
end_time - start_time

## For real Dataset with 10843 animals
start_time <- Sys.time()
DA.fit(I, P,  2, c(6.4,0.03,1.36,0.21,0.29))
end_time <- Sys.time()
end_time - start_time
```

Appendix C: Estimates throught the Laplace approximation (LA) method

```
## Estimates throught the LA method
start_time <- Sys.time()
LA.fit(I, P,  1, c(6.5,0.13,1.11,0.28))
end_time <- Sys.time()
end_time - start_time

start_time <- Sys.time()
LA.fit(I, P,  11, c(6.6,1.25,0.16, 0.28))
end_time <- Sys.time()
end_time - start_time

start_time <- Sys.time()
LA.fit(I, P,  2, c(6.5002,0.0001,1.2042,0.3288, 0.2503))
end_time <- Sys.time()
end_time - start_time

## For real Dataset with 10843 animals
start_time <- Sys.time()
LA.fit(I, P,  2, c(6.4,0.01,1.3,0.3,0.33))
end_time <- Sys.time()
end_time - start_time



## Simulated datasets
I <-read.csv("Mat_Time_100_sim.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
P <-read.csv("Mat_Obs_100_sim.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data

I <-read.csv("Mat_Time_100_real.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
P <-read.csv("Mat_Obs_100_real.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
dim(I)
dim(P)

## Estimates throught the DA method
start_time <- Sys.time()
DA.fit(I, P,  0, c(6.3,1.3,0.3))
end_time <- Sys.time()
end_time - start_time
```

```
start_time <- Sys.time()
DA.fit(I, P,  1, c(6.5,0.13,1.11,0.28))
end_time <- Sys.time()
end_time - start_time

start_time <- Sys.time()
DA.fit(I, P,  11, c(6.59205,1.30,0.1835, 0.30))
end_time <- Sys.time()
end_time - start_time

start_time <- Sys.time()
DA.fit(I, P,  2, c(6.4,0.03,1.36,0.21,0.29))
end_time <- Sys.time()
end_time - start_time



## Estimates throught the LA method
start_time <- Sys.time()
LA.fit(I, P,  1, c(6.5,0.13,1.11,0.28))
end_time <- Sys.time()
end_time - start_time

start_time <- Sys.time()
LA.fit(I, P,  11, c(6.4,1.30,0.01, 0.30))
end_time <- Sys.time()
end_time - start_time

start_time <- Sys.time()
LA.fit(I, P,  2, c(6.4,0.03,1.36,0.21,0.29))
end_time <- Sys.time()
end_time - start_time
```

Appendix D: delta approximation (DA) codes

```
##---------------------------------##---------------------------------##
##    Inference on SDE Stochastic Gompertz model:                     ##
## 1) The fixed effects closed formula                                ##
## 2) The delta approximation method for random alpha case            ##
## 3) The delta approximation method for random beta case             ##
## 4) The delta approximation method both alpha and beta random       ##
##---------------------------------##---------------------------------##
##                                Data                                ##
##---------------------------------##---------------------------------##


## Usage
# DA.fit(Mat_Time, Mat_Obs,  random, c(a,t,b,o,s))

## Mat_Time - matrix of m observation times. Each column represents the
observation time of an individual.
## Mat_Obs - matrix of m trajectories. Each column represents  the
observations of an individual.
## Notes: Mat_Time and Mat_Obs must  have the same dimensions. If different
animals had different observation times, the remain values should have NA

## random  - the random effects in the drift. If random=0, no random effects. If
random=1,  a random effect on alpha
## If random=11,  a random effect on beta. If random=2,  a random effect on both
alpha and beta
## starting parameter values for the minimization
## c(a,t,b,o,s): initial values for a - alpha, t - theta (give the value 0 if no
random effect is used)
## b - beta, o - omega (give the value 0 if no random effect is used)
## s - sigma (the diffusion coefficient assumed fixed)
## If random=0 use c(a,b,s). if random=1 use c(a,t,b,s). if random=11 use
## c(a,b,o,s). if random=2 use c(a,t,b,o,s).

DA.fit <- function(I,P,rnd, y) {
library("numDeriv")

d <- dim(I)  # Matrix dimension 'D';
n <- d[2]  # Number of animals;
m <- d[1]    # Number of weightings per animal;

Nc<-vector(length=n)
for (j in 1:n){
```

```
for (i in 1:m){
ifelse(I[i,j] == 'NA',Nc,Nc[j]<-Nc[j]+1)}}
N<-Nc-1


##---------------------------------##---------------------------------##
##                         Fixed effects case                          ##
##---------------------------------##---------------------------------##
if (rnd==0) {

if (length(y)!=3) {
print(paste("Error: wrong number of parameter"))
}
else {
# - Likelihood function
L<-function(I1,I2,P1,P2,N,x){
-(-(N/2)*log(2*pi)-(N/2)*log(x[3]^2/(2*x[2]))-
(1/2)*(sum(log(1-(exp(-x[2]*(I2-I1)))^2)))-
(x[2]/(x[3]^2))*(sum((P2-x[1]-(P1-x[1])*(exp(-x[2]*(I2-I1))))^2/
(1-(exp(-x[2]*(I2-I1)))^2))))
}

LTG<-function(x){
LTG<-0
for (i in 1:n) {
LTG<-LTG+L(I[1:(Nc[i]-1),i],I[2:Nc[i],i],log(P[1:(Nc[i]-1),i]),
log(P[2:Nc[i],i]),N[i],c(x[1],x[2],x[3]))
}
return(LTG)
}

#LTG is the  symmetric of the likelihood function and nlm (or nlminb)
minimizes symmetric of the likelihood function.
mLTG<-nlm(LTG,c(y[1],y[2],y[3]),hessian=TRUE)
uG <- mLTG$estimate[1]
UG <- exp(uG)
lG <- mLTG$estimate[2]
sG <- abs(mLTG$estimate[3])

# - Empirical Fisher information matrix; Inverse of the Hessian LT at
point (a,b1,b2,s)
vhG<-solve(mLTG$hessian)
MEuG <- sqrt(vhG[1,1])
MEUG <- sqrt(vhG[1,1])*exp(uG)
MElG <- sqrt(vhG[2,2])
MEsG <- sqrt(vhG[3,3])

print(paste("SDE with fixed effects estimates"))
print(paste("alpha", round(uG,5)))
print(paste("beta", round(lG,5)))
```

```
print(paste("sigma", round(sG,5)))
print(paste("____"))
print(paste("Standard deviation estimate from the empirical Fisher
information matrix"))
print(paste("alpha", round(MEuG,5)))
print(paste("beta", round(MElG,5)))
print(paste("sigma",round(MEsG,5)))

}


}

##---------------------------------##----------------------------------##
##                 DA method for random alpha case                    ##
##---------------------------------##----------------------------------##

else if (rnd==1) {

if (length(y)!=4) {
print(paste("Error: wrong number of parameter"))
}
else {
# - Likelihood function
LG <- function(I1,I2,P1,P2,N,x){
f <- (-1/2)*sum((P2 - x[1] - (P1 - x[1])*(exp(- x[3]*(I2 - I1))))^2/
((x[4]^2/(2*x[3]))*(1 - (exp(- x[3]*(I2 - I1)))^2)))
f1 <- (sum((1 - (exp(- x[3]*(I2 - I1))))*(P2 - x[1] - (P1 - x[1])*
(exp(- x[3]*(I2 - I1))))/((x[4]^2/(2*x[3]))*(1 - (exp(- x[3]*(I2 - I1)))^2))))
f2 <- (- sum((1 - (exp(- x[3]*(I2 - I1))))^2/((x[4]^2/(2*x[3]))*
(1 - (exp(- x[3]*(I2 - I1)))^2))))
MV <- -((-1/2)*sum(log(2*pi*((x[4]^2/(2*x[3]))*
(1 - (exp(- x[3]*(I2 - I1)))^2)))) + f + log(1 + (1/2)*(x[2]^2*(f2 + (f1)^2))))
MV}


LTG <- function(x){
LTG <- 0
for (k in 1:n) {
LTG<-LTG+LG(I[1:(Nc[k]-1),k],I[2:Nc[k],k],log(P[1:(Nc[k]-1),k]),
log(P[2:Nc[k],k]),N[k],c(x[1],x[2],x[3],x[4]))
}
return(LTG)
}
#LTG is the  symmetric of the likelihood function and nlm (or nlminb)
minimizes symmetric of the likelihood function.
mLTG<-nlm(LTG,c(y[1],y[2],y[3], y[4]),hessian=TRUE)

uG <- mLTG$estimate[1]
UG <- exp(uG)
```

```
tG <- mLTG$estimate[2]
lG <- mLTG$estimate[3]
sG <- abs(mLTG$estimate[4])

# - Empirical Fisher information matrix; Inverse of the Hessian
LT at point (a,b1,b2,s)
vhG<-solve(mLTG$hessian)
MEuG <- sqrt(vhG[1,1])
MEUG <- sqrt(vhG[1,1])*exp(uG)
MEtG <- sqrt(vhG[2,2])
MElG <- sqrt(vhG[3,3])
MEsG <- sqrt(vhG[4,4])

print(paste("SDE with random effect on alpha estimates"))
print(paste("mu", round(uG,5), "theta", round(tG,5)))
print(paste("beta", round(lG,5)))
print(paste("sigma", round(sG,5)))
print(paste("___"))
print(paste("Standard deviation estimate from the empirical Fisher
information matrix"))
print(paste("mu", round(MEuG,5), "theta", round(MEtG,5)))
print(paste("beta", round(MElG,5)))
print(paste("sigma",round(MEsG,5)))
}

}

##---------------------------------##---------------------------------##
##                    DA method for random beta case                  ##
##---------------------------------##---------------------------------##

else if (rnd==11) {
if (length(y)!=4) {
print(paste("Error: wrong number of parameter"))
}
else {
# - Likelihood function
LG<-function(I1,I2,P1,P2,N,x){

f <- (-1/2)*sum((P2 - x[1] - (P1 - x[1])*(exp(- x[2]*(I2 - I1))))^2/
(((x[4]^2)/(2*x[2]))*(1 - (exp(- x[2]*(I2 - I1)))^2)) +
log(2*pi*(((x[4]^2)/(2*x[2]))*(1 - (exp(- x[2]*(I2 - I1)))^2))))
f1 <- (sum(((I1 - I2)*(P1 - x[1])*(exp(- x[2]*(I2 - I1)))*(P2 - x[1] -
(P1 - x[1])*(exp(- x[2]*(I2 - I1)))))/(((x[4]^2)/(2*x[2]))*(1 - (exp(- x[2]*
(I2 - I1)))^2))) - sum((P2 - x[1] - (P1 - x[1])*(exp(- x[2]*(I2 - I1))))^2/
(x[4]^2*(1 - (exp(- x[2]*(I2 - I1)))^2))) + 2*sum(((P2 - x[1] - (P1 - x[1])*
(exp(- x[2]*(I2 - I1))))^2*(I2 - I1)*(exp(- x[2]*(I2 - I1)))^2)/(((x[4]^2)/
(x[2]))*(1 - (exp(- x[2]*(I2 - I1)))^2)^2)) + N/(2*x[2]) - sum(((I2 - I1)*
(exp(- x[2]*(I2 - I1)))^2)/(1 - (exp(- x[2]*(I2 - I1)))^2)))
```

```
f2 <- (2*sum(((I2 - I1)^2*(P1 - x[1])*(exp(- x[2]*(I2 - I1)))*(P2 - x[1] -
(P1 - x[1])*(exp(- x[2]*(I2 - I1)))))/(((x[4]^2)/(x[2]))*(1 - (exp(- x[2]*
(I2 - I1)))^2))) - 2*sum(((I2 - I1)^2*(P1 - x[1])^2*(exp(- x[2]*(I2 - I1)))^2)/
(((x[4]^2)/(x[2]))*(1 - (exp(- x[2]*(I2 - I1)))^2))) -  4*sum(((I2 - I1)*
(P1 - x[1])*(exp(- x[2]*(I2 - I1)))*(P2 - x[1] - (P1 - x[1])*(exp(- x[2]*
(I2 - I1)))))/(x[4]^2*(1 - (exp(- x[2]*(I2 - I1)))^2))) +  8*sum(((I2 - I1)^2*
(P1 - x[1])*(exp(- x[2]*(I2 - I1)))^3*(P2 - x[1] - (P1 - x[1])*(exp(- x[2]*
(I2 - I1)))))/(((x[4]^2)/(x[2]))*(1 - (exp(- x[2]*(I2 - I1)))^2)^2)) +
4*sum(((I2 - I1)*(exp(- x[2]*(I2 - I1)))^2*(P2 - x[1] - (P1 - x[1])*(exp(- x[2]*
(I2 - I1))))^2)/(x[4]^2*(1 - (exp(- x[2]*(I2 - I1)))^2)^2)) -
4*sum(((I2 - I1)^2*(exp(- x[2]*(I2 - I1)))^2*(P2 - x[1] - (P1 - x[1])*
(exp(- x[2]*(I2 - I1))))^2)/((x[4]^2/x[2])*(1 - (exp(- x[2]*(I2 - I1)))^2)^2)) -
8*sum(((I2 - I1)^2*(exp(- x[2]*(I2 - I1)))^4*(P2 - x[1] - (P1 - x[1])*
(exp(- x[2]*(I2 - I1))))^2)/((x[4]^2/x[2])*(1 - (exp(- x[2]*(I2 - I1)))^2)^3)) +
2*sum(((I2 - I1)^2*(exp(- x[2]*(I2 - I1)))^4)/(1 - (exp(- x[2]*(I2 - I1)))^2)^2) +
2*sum(((I2 - I1)^2*(exp(- x[2]*(I2 - I1)))^2)/(1 - (exp(- x[2]*(I2 - I1)))^2)) -
N/(2*x[2]^2))

MV<- - ( f + log(1 + (1/2)*(x[3]^2)*(f2 + (f1)^2)))
MV}

LTG <- function(x){
LTG <- 0
for (k in 1:n) {
LTG <- LTG + LG(I[1:(Nc[k]-1),k],I[2:Nc[k],k],log(P[1:(Nc[k]-1),k]),
log(P[2:Nc[k],k]),N[k],c(x[1],x[2],x[3],x[4]))
}
return(LTG)
}

#LTG is the  symmetric of the likelihood function and nlm (or nlminb)
minimizes symmetric of the likelihood function.
mLTG <- nlm(LTG,c(y[1],y[2],y[3], y[4]),hessian=TRUE)

uG <- mLTG$estimate[1]
UG <- exp(uG)
lG <- mLTG$estimate[2]
oG <- mLTG$estimate[3]
sG <- abs(mLTG$estimate[4])

# - Empirical Fisher information matrix; Inverse of the Hessian LT at
point (a,b1,b2,s)
vhG<-solve(mLTG$hessian)

MEuG <- sqrt(vhG[1,1])
MEUG <- sqrt(vhG[1,1])*exp(uG)
MElG <- sqrt(vhG[2,2])
MEoG <- sqrt(vhG[3,3])
MEsG <- sqrt(vhG[4,4])
```

```
print(paste("SDE with random effect on beta estimates"))
print(paste("alpha", round(uG,5) ))
print(paste("lambda", round(lG,5), "omega", round(oG,5)))
print(paste("sigma", round(sG,5)))
print(paste("___"))
print(paste("Standard deviation estimate from the empirical Fisher
information matrix"))
print(paste("alpha", round(MEuG,5)))
print(paste("lambda", round(MElG,5), "omega", round(MEoG,5)))
print(paste("sigma",round(MEsG,5)))
}}

##---------------------------------##---------------------------------##
##               DA method for both alpha and beta random             ##
##---------------------------------##---------------------------------##

else if (rnd==2) {

if (length(y)!=5) {
print(paste("Error: wrong number of parameter"))
}
else {
# - Likelihood function
LG <- function(I1,I2,P1,P2,N,x){
f <- ((-x[3]/x[5]^2)*sum((P2-x[1]-(P1-x[1])*(exp(-x[3]*(I2-I1))))^2/
(1-(exp(-x[3]*(I2-I1)))^2))-(1/2)*sum(log(((pi*x[5]^2)/x[3])*
(1-(exp(-x[3]*(I2-I1)))^2))))
fa1 <- (((-2*x[3])/x[5]^2)*sum(((P2-x[1]-(P1-x[1])*
(exp(-x[3]*(I2-I1))))*((exp(-x[3]*(I2-I1)))-1))/(1-(exp(-x[3]*(I2-I1)))^2)))
fa2 <- (((-2*x[3])/x[5]^2)*sum(((exp(-x[3]*(I2-I1)))-1)^2/
(1-(exp(-x[3]*(I2-I1)))^2)))
fb1 <- (((-2*x[3])/x[5]^2)*sum(((I2-I1)*(P1-x[1])*(exp(-x[3]*(I2-I1)))*
(P2-x[1]-(P1-x[1])*(exp(-x[3]*(I2-I1)))))/(1-(exp(-x[3]*(I2-I1)))^2)) -
(1/x[5]^2)*sum((P2-x[1]-(P1-x[1])*(exp(-x[3]*(I2-I1))))^2/
(1-(exp(-x[3]*(I2-I1)))^2)) +
((2*x[3])/x[5]^2)*sum(((I2-I1)*(exp(-x[3]*(I2-I1)))^2*(P2-x[1]-(P1-x[1])*
(exp(-x[3]*(I2-I1))))^2)/(1-(exp(-x[3]*(I2-I1)))^2)^2) -
sum(((I2-I1)*(exp(-x[3]*(I2-I1)))^2)/(1-(exp(-x[3]*(I2-I1)))^2)) + N/(2*x[3]))
fb2 <- (((2*x[3])/x[5]^2)*sum(((I2-I1)^2*(P1-x[1])*(exp(-x[3]*(I2-I1)))*
(P2-x[1]-(P1-x[1])*(exp(-x[3]*(I2-I1)))))/(1-(exp(-x[3]*(I2-I1)))^2)) -
((2*x[3])/x[5]^2)*sum(((I2-I1)^2*(P1-x[1])^2*(exp(-x[3]*(I2-I1)))^2)/
(1-(exp(-x[3]*(I2-I1)))^2)) -  (4/x[5]^2)*sum(((I2-I1)*(P1-x[1])*
(exp(-x[3]*(I2-I1)))*(P2-x[1]-(P1-x[1])*(exp(-x[3]*(I2-I1)))))/
(1-(exp(-x[3]*(I2-I1)))^2)) + ((8*x[3])/x[5]^2)*sum(((I2-I1)^2*(P1-x[1])*
(exp(-x[3]*(I2-I1)))^3*(P2-x[1]-(P1-x[1])*(exp(-x[3]*(I2-I1)))))/
(1-(exp(-x[3]*(I2-I1)))^2)^2) + (4/x[5]^2)*sum(((I2-I1)*(exp(-x[3]*(I2-I1)))^2*
(P2-x[1]-(P1-x[1])*(exp(-x[3]*(I2-I1))))^2)/(1-(exp(-x[3]*(I2-I1)))^2)^2) -
((4*x[3])/x[5]^2)*sum(((I2-I1)^2*(exp(-x[3]*(I2-I1)))^2*(P2-x[1]-(P1-x[1])*
(exp(-x[3]*(I2-I1))))^2)/(1-(exp(-x[3]*(I2-I1)))^2)^2) -
```

```
((8*x[3])/x[5]^2)*sum(((I2-I1)^2*(exp(-x[3]*(I2-I1)))^4*(P2-x[1]-(P1-x[1])*
(exp(-x[3]*(I2-I1))))^2)/(1-(exp(-x[3]*(I2-I1)))^2)^3) +
2*sum(((I2-I1)^2*(exp(-x[3]*(I2-I1)))^4)/(1-(exp(-x[3]*(I2-I1)))^2)^2) +
2*sum(((I2-I1)^2*(exp(-x[3]*(I2-I1)))^2)/(1-(exp(-x[3]*(I2-I1)))^2)) -
N/(2*x[3]^2))
MV <- -(-sum(P2)+f+log(1+(x[2]^2/2)*(fa2+(fa1)^2)+(x[4]^2/2)*(fb2+(fb1)^2)))
MV}
# - Sum of likelihood functions
LTG<-function(x){
LTG<-0
for (k in 1:n) {
LTG <- LTG + LG(I[1:(Nc[k]-1),k],I[2:Nc[k],k],log(P[1:(Nc[k]-1),k]),
log(P[2:Nc[k],k]),N[k],c(x[1],x[2],x[3],x[4],x[5]))
}
return(LTG)
}

#LTG is the  symmetric of the likelihood function and nlm (or nlminb)
minimizes symmetric of the likelihood function.
mLTG <- nlm(LTG,c(y[1],y[2],y[3], y[4],y[5]),hessian=TRUE)
uG <- mLTG$estimate[1]
UG <- exp(uG)
tG <- mLTG$estimate[2]
lG <- mLTG$estimate[3]
oG <- mLTG$estimate[4]
sG <- abs(mLTG$estimate[5])

# - Empirical Fisher information matrix; Inverse of the Hessian LT at
point (a,b1,b2,s)
vhG<-solve(mLTG$hessian)
MEuG <- sqrt(vhG[1,1])
MEUG <- sqrt(vhG[1,1])*exp(uG)
MEtG <- sqrt(vhG[2,2])
MElG <- sqrt(vhG[3,3])
MEoG <- sqrt(vhG[4,4])
MEsG <- sqrt(vhG[5,5])

print(paste("SDE with random effect on alpha and beta estimates"))
print(paste("mu", round(uG,5), "theta", round(tG,5)))
print(paste("lambda", round(lG,5), "omega", round(oG,5)))
print(paste("sigma", round(sG,5)))
print(paste("___"))
print(paste("Standard deviation estimate from the empirical Fisher
information matrix"))
print(paste("mu", round(MEuG,5), "theta", round(MEtG,5)))
print(paste("lambda", round(MElG,5), "omega", round(MEoG,5)))
print(paste("sigma",round(MEsG,5)))

}}
```

```
else {
print("Choice of parameter rnd incorrect")
}}
```

Appendix E: Laplace approximation (LA) codess

```
##---------------------------------##---------------------------------##
##    Inference on SDE Stochastic Gompertz model:                    ##
## 2) The Laplace approximation method for random alpha case         ##
## 3) The Laplace approximation method for random beta case          ##
## 4) The Laplace approximation method both alpha and beta random    ##
##---------------------------------##---------------------------------##
##                               Data                                ##
##---------------------------------##---------------------------------##


## Usage
# LA.fit(Mat_Time, Mat_Obs,  random, c(a,t,b,o,s))

## Mat_Time - matrix of m observation times. Each column represents the
##observation time of an individual.
## Mat_Obs - matrix of m trajectories. Each column represents  the
##observations of an individual.
## Notes: Mat_Time and Mat_Obs must  have the same dimensions. If different
##animals had different observation times, the remain values should have NA

## random  - the random effects in the drift. If random=1,  a random effect on
##alpha
##If random=11,  a random effect on beta. If random=2,  a random effect on
##both alpha and beta
## starting parameter values for the minimization
## c(a,t,b,o,s): initial values for a - alpha, t - theta (give the value 0 if
##no random effect is used)
## b - beta, o - omega (give the value 0 if no random effect is used)
## s - sigma (the diffusion coefficient assumed fixed
## If random=0 use c(a,b,s). if random=1 use
## c(a,t,b,s). if random=11 use c(a,b,o,s). if random=2 use c(a,t,b,o,s).


LA.fit <- function(I,P,rnd, y) {
library('maxLik')
library("numDeriv")
d <- dim(I)  # Matrix dimension 'D';
n <- d[2]  # Number of animals;
m <- d[1]    # Number of weightings per animal;

Nc<-vector(length=n)
for (j in 1:n){
```

```
for (i in 1:m){
ifelse(I[i,j] == ’NA’,Nc,Nc[j]<-Nc[j]+1)}}
N<-Nc-1

##---------------------------------##---------------------------------##
##                 LA method for random alpha case                   ##
##---------------------------------##---------------------------------##

if (rnd==1) {

if (length(y)!=4) {
print(paste("Error: wrong number of parameter"))
}
else {
LG <- function(I1,I2,P1,P2,N,x){
x[5] <- ((2*x[3]/x[4]^2)*sum(((exp(- x[3]*(I2-I1))) - 1)^2/
(1 - (exp(- x[3]*(I2-I1)))^2)) + 1/x[2]^2)^{-1}*
((2*x[3]/x[4]^2)*sum((P2-P1*(exp(- x[3]*(I2-I1))))*
(1 - (exp(- x[3]*(I2-I1))))/(1 - (exp(- x[3]*(I2-I1)))^2)) + x[1]/x[2]^2)

MV <- -(-sum(P2)- log((x[2])) - (1/2)*sum(log(((pi*x[4]^2)/x[3])*
(1 - (exp(- x[3]*(I2 - I1)))^2))) -
(1/2)*sum((P2 - x[5] -(P1 - x[5])*(exp(- x[3]*(I2 - I1))))^2/
((x[4]^2/(2*x[3]))*(1 - (exp(- x[3]*(I2 - I1)))^2))) -
(x[5] - x[1])^2/(2*(x[2])^2) - (1/2)*log(abs(sum((1 - (exp(- x[3]*(I2 - I1))))^2/
((x[4]^2/(2*x[3]))*(1 - (exp(- x[3]*(I2 - I1)))^2))) + 1/(x[2])^2)))
MV}

LTG<-function(x){
LTG<-0
for (i in 1:n) {
LTG<-LTG+LG(I[1:(Nc[i]-1),i],I[2:Nc[i],i],log(P[1:(Nc[i]-1),i]),
log(P[2:Nc[i],i]),N[i],c(x[1],x[2],x[3],x[4]))
}
return(LTG)
}

#LTG is the  symmetric of the likelihood function and nlm (or nlminb)
minimizes symmetric of the likelihood function.
mLTG<-nlm(LTG,c(y[1],y[2],y[3], y[4]),hessian=TRUE)

uG <- mLTG$estimate[1]
UG <- exp(uG)
tG <- mLTG$estimate[2]
lG <- mLTG$estimate[3]
sG <- abs(mLTG$estimate[4])

# - Empirical Fisher information matrix; Inverse of the Hessian LT at
#point (a,b1,b2,s)
vhG<-solve(mLTG$hessian)
```

```
MEuG <- sqrt(vhG[1,1])
MEUG <- sqrt(vhG[1,1])*exp(uG)
MEtG <- sqrt(vhG[2,2])
MElG <- sqrt(vhG[3,3])
MEsG <- sqrt(vhG[4,4])

print(paste("SDE with random effect on alpha estimates"))
print(paste("mu", round(uG,5), "theta", round(tG,5)))
print(paste("beta", round(lG,5)))
print(paste("sigma", round(sG,5)))
print(paste("___"))
print(paste("Standard deviation estimate from the empirical Fisher
#information matrix"))
print(paste("mu", round(MEuG,5), "theta", round(MEtG,5)))
print(paste("beta", round(MElG,5)))
print(paste("sigma",round(MEsG,5)))

}}

##-----------------------------------##-----------------------------------##
##                   LA method for random beta case                      ##
##-----------------------------------##-----------------------------------##

else if (rnd==11) {
if (length(y)!=4) {
print(paste("Error: wrong number of parameter"))
}
else {
A <- matrix(c(0, 0, 0, 0, 1), 1, 5)
B <- 0
betaP<-vector(length=n)
for (k in 1:n) {
beta <- function(param){
I1 <- I[1:(Nc[k]-1),k]
I2 <- I[2:Nc[k],k]
P1 <- P[1:(Nc[k]-1),k]
P2 <- P[2:Nc[k],k]
alpha <- param[1]
beta <- param[5]
lambda <- param[2]
omega <- param[3]
sigma <- param[4]
ll <- - (1/2)*sum((P2 - alpha - (P1 - alpha)*(exp(- beta*(I2 - I1))))^2/
((sigma^2/(2*beta))*(1 - (exp(- beta*(I2 - I1)))^2))) - (1/2)*
(beta - lambda)^2/omega^2
- (1/2)*sum(log((sigma^2/(2*beta))*(1 - (exp(- beta*(I2 - I1)))^2)))
ll
}
resFix <- maxLik(beta,grad = NULL,hess = NULL,  method = "CG",
start=c(alpha = 6.5, lambda = 1.3, omega =0.00001, sigma = 0.5^2, beta = 1.4),
```

```
fixed = c("alpha", "lambda", "omega", "sigma"),
constraints=list(ineqA=A, ineqB=B))
suppressWarnings(summary(resFix))
suppressWarnings(betaP[k] <- resFix$estimate[5])
}


# - Likelihood function
LG <- function(I1,I2,P1,P2,betaP,N,x){
x[5] <- (betaP-(((-2*betaP)/x[4]^2)*sum(((I2-I1)*(P1-x[1])*
(exp(-betaP*(I2-I1)))*(P2-x[1]-(P1-x[1])*(exp(-betaP*(I2-I1)))))/
(1-(exp(-betaP*(I2-I1)))^2)) -  (1/x[4]^2)*sum((P2-x[1]-(P1-x[1])*
(exp(-betaP*(I2-I1))))^2/(1-(exp(-betaP*(I2-I1)))^2)) + ((2*betaP)/x[4]^2)*
sum(((I2-I1)*(exp(-betaP*(I2-I1)))^2*(P2-x[1]-(P1-x[1])*(exp(-betaP*
(I2-I1))))^2)/(1-(exp(-betaP*(I2-I1)))^2)^2) - sum(((I2-I1)*(exp(-betaP*
(I2-I1)))^2)/(1-(exp(-betaP*(I2-I1)))^2)) - (betaP-x[2])/x[3]^2 +
N/(2*betaP))/(((2*betaP)/x[4]^2)*sum(((I2-I1)^2*(P1-x[1])*
(exp(-betaP*(I2-I1)))*(P2-x[1]-(P1-x[1])*(exp(-betaP*(I2-I1)))))/
(1-(exp(-betaP*(I2-I1)))^2)) -  ((2*betaP)/x[4]^2)*sum(((I2-I1)^2*(P1-x[1])^2*
(exp(-betaP*(I2-I1)))^2)/(1-(exp(-betaP*(I2-I1)))^2)) -
(4/x[4]^2)*sum(((I2-I1)*(P1-x[1])*(exp(-betaP*(I2-I1)))*(P2-x[1]-(P1-x[1])*
(exp(-betaP*(I2-I1)))))/(1-(exp(-betaP*(I2-I1)))^2)) +
((8*betaP)/x[4]^2)*sum(((I2-I1)^2*(P1-x[1])*(exp(-betaP*(I2-I1)))^3*
(P2-x[1]-(P1-x[1])*(exp(-betaP*(I2-I1)))))/(1-(exp(-betaP*(I2-I1)))^2)^2) +
(4/x[4]^2)*sum(((I2-I1)*(exp(-betaP*(I2-I1)))^2*(P2-x[1]-(P1-x[1])*
(exp(-betaP*(I2-I1))))^2)/(1-(exp(-betaP*(I2-I1)))^2)^2) -
((4*betaP)/x[4]^2)*sum(((I2-I1)^2*(exp(-betaP*(I2-I1)))^2*(P2-x[1]-(P1-x[1])*
(exp(-betaP*(I2-I1))))^2)/(1-(exp(-betaP*(I2-I1)))^2)^2) -
((8*betaP)/x[4]^2)*sum(((I2-I1)^2*(exp(-betaP*(I2-I1)))^4*(P2-x[1]-(P1-x[1])*
(exp(-betaP*(I2-I1))))^2)/(1-(exp(-betaP*(I2-I1)))^2)^3) +
2*sum(((I2-I1)^2*(exp(-betaP*(I2-I1)))^2)/(1-(exp(-betaP*(I2-I1)))^2)) +
2*sum(((I2-I1)^2*(exp(-betaP*(I2-I1)))^4)/(1-(exp(-betaP*(I2-I1)))^2)^2) -
1/x[3]^2 - N/(2*betaP^2)))

MV <- -(-sum(P2)- log(x[3]) - (N/2)*log(2*pi) -
0.5*log(abs(((2*x[5])/x[4]^2)*sum(((I2-I1)^2*(P1-x[1])*(exp(-x[5]*(I2-I1)))*
(P2-x[1]-(P1-x[1])*(exp(-x[5]*(I2-I1)))))/(1-(exp(-x[5]*(I2-I1)))^2)) -
((2*x[5])/x[4]^2)*sum(((I2-I1)^2*(P1-x[1])^2*(exp(-x[5]*(I2-I1)))^2)/
(1-(exp(-x[5]*(I2-I1)))^2)) -
(4/x[4]^2)*sum(((I2-I1)*(P1-x[1])*(exp(-x[5]*(I2-I1)))*(P2-x[1]-(P1-x[1])*
(exp(-x[5]*(I2-I1)))))/(1-(exp(-x[5]*(I2-I1)))^2)) +
((8*x[5])/x[4]^2)*sum(((I2-I1)^2*(P1-x[1])*(exp(-x[5]*(I2-I1)))^3*
(P2-x[1]-(P1-x[1])*(exp(-x[5]*(I2-I1)))))/(1-(exp(-x[5]*(I2-I1)))^2)^2) +
(4/x[4]^2)*sum(((I2-I1)*(exp(-x[5]*(I2-I1)))^2*(P2-x[1]-(P1-x[1])*(exp(-x[5]*
(I2-I1))))^2)/(1-(exp(-x[5]*(I2-I1)))^2)^2) -
((4*x[5])/x[4]^2)*sum(((I2-I1)^2*(exp(-x[5]*(I2-I1)))^2*(P2-x[1]-(P1-x[1])*
(exp(-x[5]*(I2-I1))))^2)/(1-(exp(-x[5]*(I2-I1)))^2)^2) -
((8*x[5])/x[4]^2)*sum(((I2-I1)^2*(exp(-x[5]*(I2-I1)))^4*(P2-x[1]-(P1-x[1])*
(exp(-x[5]*(I2-I1))))^2)/(1-(exp(-x[5]*(I2-I1)))^2)^3) +
2*sum(((I2-I1)^2*(exp(-x[5]*(I2-I1)))^2)/(1-(exp(-x[5]*(I2-I1)))^2)) +
```

```
2*sum(((I2-I1)^2*(exp(-x[5]*(I2-I1)))^4)/(1-(exp(-x[5]*(I2-I1)))^2)^2) -
1/x[3]^2 - N/(2*x[5]^2))) -
0.5*sum((P2-x[1]-(P1-x[1])*(exp(-x[5]*(I2-I1))))^2/((x[4]^2/(2*x[5]))*
(1-(exp(-x[5]*(I2-I1)))^2))) - 0.5*(x[5]-x[2])^2/x[3]^2 - 0.5*sum(log((x[4]^2/
(2*x[5]))*(1-(exp(-x[5]*(I2-I1)))^2))))
MV}

LTG <- function(x){
LTG <- 0
for (k in 1:n) {
LTG <- LTG + LG(I[1:(Nc[k]-1),k],I[2:Nc[k],k],log(P[1:(Nc[k]-1),k]),
log(P[2:Nc[k],k]),betaP[k],N[k],c(x[1],x[2],x[3],x[4]))
}
return(LTG)
}

#LTG is the  symmetric of the likelihood function and nlm (or nlminb) minimizes
symmetric of the likelihood function.
mLTG <- nlm(LTG,c(y[1],y[2],y[3], y[4]),hessian=TRUE)

uG <- mLTG$estimate[1]
UG <- exp(uG)
lG <- mLTG$estimate[2]
oG <- mLTG$estimate[3]
sG <- abs(mLTG$estimate[4])

# - Empirical Fisher information matrix; Inverse of the Hessian LT at
point (a,b1,b2,s)
vhG<-solve(mLTG$hessian)
MEuG <- sqrt(vhG[1,1])
MEUG <- sqrt(vhG[1,1])*UG
MElG <- sqrt(vhG[2,2])
MEoG <- sqrt(vhG[3,3])
MEsG <- sqrt(vhG[4,4])


print(paste("SDE with random effect on beta estimates"))
print(paste("alpha", round(uG,5) ))
print(paste("lambda", round(lG,5), "omega", round(oG,5)))
print(paste("sigma", round(sG,5)))
print(paste("___"))
print(paste("Standard deviation estimate from the empirical Fisher
information matrix"))
print(paste("alpha", round(MEuG,5)))
print(paste("lambda", round(MElG,5), "omega", round(MEoG,5)))
print(paste("sigma",round(MEsG,5)))
}}


##---------------------------------##---------------------------------##
```

```
##               LA method for both alpha and beta random             ##
##---------------------------------##---------------------------------##

else if (rnd==2) {

if (length(y)!=5) {
print(paste("Error: wrong number of parameter"))
}
else {
A <- matrix(c(0, 0, 0, 0, 1), 1, 5)
B <- 0
alphaP <- vector(length=n)
for (k in 1:n) {
llalpha <- function(parama){
I1 <- I[1:(Nc[k]-1),k]
I2 <- I[2:Nc[k],k]
P1 <- P[1:(Nc[k]-1),k]
P2 <- P[2:Nc[k],k]
mu <- parama[1]
alpha <- parama[5]
theta <- parama[2]
beta <- parama[3]
sigma <- parama[4]
ll <- - (1/2)*sum((P2 - alpha - (P1 - alpha)*(exp(- beta*(I2 - I1))))^2/
((sigma^2/(2*beta))*(1 - (exp(- beta*(I2 - I1)))^2))) - (1/2)*(alpha - mu)^2/
theta^2
ll
}
resFix1 <- maxLik(llalpha,grad = NULL,hess = NULL,  method = "CG",
start=c(mu = 6.45, theta = 0.15,beta =1.43, sigma = 0.33, alpha = 6.45),
fixed = c("mu", "theta", "beta", "sigma"),
constraints=list(ineqA=A, ineqB=B))
suppressWarnings(summary(resFix1))
suppressWarnings( alphaP[k] <- log(resFix1$estimate[5]))
}

E <- matrix(c(0, 0, 0, 0, 1), 1, 5)
F <- 0
betaP<-vector(length=n)
for (k in 1:n) {
LLbeta <- function(paramb){
I1 <- I[1:(Nc[k]-1),k]
I2 <- I[2:Nc[k],k]
P1 <- P[1:(Nc[k]-1),k]
P2 <- P[2:Nc[k],k]
alpha <- paramb[1]
beta <- paramb[5]
lambda <- paramb[2]
omega <- paramb[3]
sigma <- paramb[4]
```

```
LL <- - (1/2)*sum((P2 - alpha - (P1 - alpha)*(exp(- beta*(I2 - I1))))^2/
((sigma^2/(2*beta))*(1 - (exp(- beta*(I2 - I1)))^2))) - (1/2)*
(beta - lambda)^2/omega^2 - (1/2)*sum(log((sigma^2/(2*beta))*
(1 - (exp(- beta*(I2 - I1)))^2)))
LL
}
resFix2 <- maxLik(LLbeta,grad = NULL,hess = NULL,  method = "CG",
start=c(alpha = 6.5, lambda = 1.3, omega =0.00001, sigma = 0.5^2, beta = 1.4),
fixed = c("alpha", "lambda", "omega", "sigma"),
constraints=list(ineqA=E, ineqB=F))
suppressWarnings(summary(resFix2))
suppressWarnings(betaP[k] <- resFix2$estimate[5])
}

# - Likelihood function
LG <- function(I1,I2,P1,P2,alphaP,betaP,N,x){
x[6] <- (alphaP - (2*x[3]/x[5]^2*sum(((P2 - alphaP - (P1 - alphaP)*
(exp(- x[3]*(I2-I1))))*((exp(- x[3]*(I2-I1))) - 1))/(1 - (exp(- x[3]*
(I2-I1)))^2)) +
(alphaP - x[1])/(x[2]^2))/(2*x[3]/x[5]^2*sum(((exp(- x[3]*(I2-I1))) - 1)^2/
(1 - (exp(- x[3]*(I2-I1)))^2)) + 1/(x[2]^2)))
x[7] <- (betaP-((-2*betaP/x[5]^2)*sum(((I2-I1)*(P1-x[1])*(exp(-betaP*(I2-I1)))*
(P2-x[1]-(P1-x[1])*(exp(-betaP*(I2-I1)))))/(1-(exp(-betaP*(I2-I1)))^2)) -
(1/x[5]^2)*sum((P2-x[1]-(P1-x[1])*(exp(-betaP*(I2-I1))))^2/(1-(exp(-betaP*
(I2-I1)))^2)) +
(2*betaP/x[5]^2)*sum(((I2-I1)*(exp(-betaP*(I2-I1)))^2*(P2-x[1]-(P1-x[1])*
(exp(-betaP*(I2-I1))))^2)/(1-(exp(-betaP*(I2-I1)))^2)^2) -
sum(((I2-I1)*(exp(-betaP*(I2-I1)))^2)/(1-(exp(-betaP*(I2-I1)))^2)) +
N/(2*betaP) - (betaP-x[3])/x[4]^2)/
((2*betaP/x[5]^2)*sum(((I2-I1)^2*(P1-x[1])*(exp(-betaP*(I2-I1)))*(P2-x[1] -
(P1-x[1])*(exp(-betaP*(I2-I1)))))/(1-(exp(-betaP*(I2-I1)))^2)) -
(2*betaP/x[5]^2)*sum(((I2-I1)^2*(P1-x[1])^2*(exp(-betaP*(I2-I1)))^2)/
(1-(exp(-betaP*(I2-I1)))^2)) -
(4/x[5]^2)*sum(((I2-I1)*(P1-x[1])*(exp(-betaP*(I2-I1)))*(P2-x[1]-(P1-x[1])*
(exp(-betaP*(I2-I1)))))/(1-(exp(-betaP*(I2-I1)))^2)) +
(8*betaP/x[5]^2)*sum(((I2-I1)^2*(P1-x[1])*(exp(-betaP*(I2-I1)))^3*(P2-x[1]-
(P1-x[1])*(exp(-betaP*(I2-I1)))))/(1-(exp(-betaP*(I2-I1)))^2)^2) +
(4/x[5]^2)*sum(((I2-I1)*(exp(-betaP*(I2-I1)))^2*(P2-x[1]-(P1-x[1])*(exp(-betaP*
(I2-I1))))^2)/(1-(exp(-betaP*(I2-I1)))^2)^2) -
(4*betaP/x[5]^2)*sum(((I2-I1)^2*(exp(-betaP*(I2-I1)))^2*(P2-x[1]-(P1-x[1])*
(exp(-betaP*(I2-I1))))^2)/(1-(exp(-betaP*(I2-I1)))^2)^2) -
(8*betaP/x[5]^2)*sum(((I2-I1)^2*(exp(-betaP*(I2-I1)))^4*(P2-x[1]-(P1-x[1])*
(exp(-betaP*(I2-I1))))^2)/(1-(exp(-betaP*(I2-I1)))^2)^3) +
2*sum(((I2-I1)^2*(exp(-betaP*(I2-I1)))^4)/(1-(exp(-betaP*(I2-I1)))^2)^2) +
2*sum(((I2-I1)^2*(exp(-betaP*(I2-I1)))^2)/(1-(exp(-betaP*(I2-I1)))^2)) -
1/x[4]^2 - N/(2*betaP^2)))
dfAA <- abs(-(2*x[7]/x[5]^2)*sum(((exp(-x[7]*(I2-I1)))-1)^2/(1-(exp(-x[7]*
(I2-I1)))^2)) - 1/(x[2]^2))
dfBB <-  abs((2*x[7]/x[5]^2)*sum(((I2-I1)^2*(P1-x[6])*(exp(-x[7]*(I2-I1)))*
(P2-x[6]-(P1-x[6])*(exp(-x[7]*(I2-I1)))))/(1-(exp(-x[7]*(I2-I1)))^2)) -
```

```
(2*x[7]/x[5]^2)*sum(((I2-I1)^2*(P1-x[6])^2*(exp(-x[7]*(I2-I1)))^2)/
(1-(exp(-x[7]*(I2-I1)))^2)) -
(4/x[5]^2)*sum(((I2-I1)*(P1-x[6])*(exp(-x[7]*(I2-I1)))*(P2-x[6]-(P1-x[6])*
(exp(-x[7]*(I2-I1)))))/(1-(exp(-x[7]*(I2-I1)))^2)) +
(8*x[7]/x[5]^2)*sum(((I2-I1)^2*(P1-x[6])*(exp(-x[7]*(I2-I1)))^3*(P2-x[6]-
(P1-x[6])*(exp(-x[7]*(I2-I1)))))/(1-(exp(-x[7]*(I2-I1)))^2)^2) +
(4/x[5]^2)*sum(((I2-I1)*(exp(-x[7]*(I2-I1)))^2*(P2-x[6]-(P1-x[6])*(exp(-x[7]*
(I2-I1))))^2)/(1-(exp(-x[7]*(I2-I1)))^2)^2) -
(4*x[7]/x[5]^2)*sum(((I2-I1)^2*(exp(-x[7]*(I2-I1)))^2*(P2-x[6]-(P1-x[6])*
(exp(-x[7]*(I2-I1))))^2)/(1-(exp(-x[7]*(I2-I1)))^2)^2) -
(8*x[7]/x[5]^2)*sum(((I2-I1)^2*(exp(-x[7]*(I2-I1)))^4*(P2-x[6]-(P1-x[6])*
(exp(-x[7]*(I2-I1))))^2)/(1-(exp(-x[7]*(I2-I1)))^2)^3) +
2*sum(((I2-I1)^2*(exp(-x[7]*(I2-I1)))^4)/(1-(exp(-x[7]*(I2-I1)))^2)^2) +
2*sum(((I2-I1)^2*(exp(-x[7]*(I2-I1)))^2)/(1-(exp(-x[7]*(I2-I1)))^2)) -
1/x[4]^2 - N/(2*x[7]^2))
dfAB <- 0
HESSIANA <- -matrix(c(dfAA,dfAB,dfAB,dfBB),nrow = 2)
SIGMA  <- solve(HESSIANA)
MV <- -(-sum(P2)-(1/2)*log((x[2]^2))-(1/2)*log(x[4]^2)-(N/2)*log(2*pi)+(1/2)*
log(det(SIGMA)) -
(x[7]/x[5]^2)*sum((P2-x[6]-(P1-x[6])*(exp(-x[7]*(I2-I1))))^2/(1-(exp(-x[7]*
(I2-I1)))^2))-(1/2)*sum(log((x[5]^2/(2*x[7]))*(1-(exp(-x[7]*(I2-I1)))^2)))-
(x[6]-x[1])^2/(2*(x[2]^2))-(x[7]-x[3])^2/(2*x[4]^2))
MV}
# - Sum of likelihood functions;
LTG <- function(x){
LTG <- 0
for (k in 1:n) {
LTG <- LTG + LG(I[1:(Nc[k]-1),k],I[2:Nc[k],k],log(P[1:(Nc[k]-1),k]),
log(P[2:Nc[k],k]),alphaP[k],betaP[k],N[k],c(x[1],x[2],x[3],x[4],x[5]))
}
return(LTG)
}

#LTG is the  symmetric of the likelihood function and nlm (or nlminb)
#minimizes symmetric of the likelihood function.
mLTG <- nlm(LTG,c(y[1],y[2],y[3], y[4],y[5]),hessian=TRUE)
uG <- mLTG$estimate[1]
UG <- exp(uG)
tG <- mLTG$estimate[2]
lG <- mLTG$estimate[3]
oG <- mLTG$estimate[4]
sG <- abs(mLTG$estimate[5])

# - Empirical Fisher information matrix; Inverse of the Hessian LT at point
#(a,b1,b2,s)
vhG<-solve(mLTG$hessian)
MEuG <- sqrt(vhG[1,1])
MEUG <- sqrt(vhG[1,1])*exp(uG)
MEtG <- sqrt(vhG[2,2])
```

```
MElG <- sqrt(vhG[3,3])
MEoG <- sqrt(vhG[4,4])
MEsG <- sqrt(vhG[5,5])

print(paste("SDE with random effect on alpha and beta estimates"))
print(paste("mu", round(uG,5), "theta", round(tG,5)))
print(paste("lambda", round(lG,5), "omega", round(oG,5)))
print(paste("sigma", round(sG,5)))
print(paste("___"))
print(paste("Standard deviation estimate from the empirical Fisher information
matrix"))
print(paste("mu", round(MEuG,5), "theta", round(MEtG,5)))
print(paste("lambda", round(MElG,5), "omega", round(MEoG,5)))
print(paste("sigma",round(MEsG,5)))
}}

else {
print("Choice of parameter rnd incorrect")
}}
```

Appendix F: R code for results in stochastic differential equation (SDE) with genetic values

```
##---------------------------------##---------------------------------##
##             Code of results presented in chapter 4                ##
##---------------------------------##---------------------------------##

I <-read.csv("Idades10843.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data
P <-read.csv("Pesos10843.csv",header=FALSE,sep=";",
dec=",",fill=TRUE)  #reading the data

D <-read.csv("Genetic_values.csv",header=TRUE,sep=";",
dec=",",fill=TRUE)  #reading the data


I<-as.data.frame(I)
P<-as.data.frame(P)
Nc<-vector(length=dim(P)[2])
for (j in 1:dim(P)[2]){
for (i in 1:dim(P)[1]){
ifelse(P[i,j] == 'NA',Nc,Nc[j]<-Nc[j]+1)}}
N<-Nc


### Choice of the Genetic to be tested (Uncomment the GV of interest)

## The D$X1VG_xxxx    - represents the first genetic value obtained
## The D$UltimVG_xxxx - represents the genetic value closest to 7months
## The D$Medvg_xxxx   - represents the mean value of the mother and father
#genetic values

#VGen<-D$X1VG_cap_crescimento
#VGen<-D$X1VG_capacidade_maternal
#VGen<-D$X1VG_gmd_estacao
#VGen<-D$X1VG_carcaca_dia_idade
#VGen<-D$X1VG_Intervalo_Entre_partos
#VGen<-D$X1VG_indice_conversao
#VGen<-D$X1VG_longevidade_produtiva
#VGen<-D$X1VG_consumo_alim_residual


#VGen<-D$UltimVG_cap_crescimento
#VGen<-D$UltimVG_capacidade_maternal
```

```
# VGen<-D$UltimVG_gmd_estacao
#VGen<-D$UltimVG_carcaca_dia_idade
#VGen<-D$UltimVG_Intervalo_Entre_partos
#VGen<-D$UltimVG_indice_conversao
#VGen<-D$UltimVG_longevidade_produtiva
VGen<-D$UltimVG_consumo_alim_residual


#VGen<-D$Medvg_cap_crec_directo
#VGen<-D$Medvg_cap_cresc_materno
#VGen<-D$Medvg_gmd
#VGen<-D$Medvg_carcaca_dia_idade
#VGen<-D$Medvg_intervalo_ent_partos
#VGen<-D$Medvg_Ind_conversao
#VGen<-D$Medvg_longevidade_prod
#VGen<-D$Medvg_consumo_alim_residual


#To clean for possible missing values and to center the GV
MVGen<-round(mean(VGen,na.rm=T),4)


IdxNA<-which(is.na(VGen))
VGenf<-na.omit(VGen)
Pf<-P[,-IdxNA]
If<-I[,-IdxNA]
Ncf<-Nc[-IdxNA]
Nf<-Ncf-1
summary(VGen)
sum(Nf)
sd(VGen, na.rm=T)
length(Nf)

VGen<-VGen-MVGen



##---------------------------------##---------------------------------##
##                          SDE mixed Model with VGs                  ##
##---------------------------------##---------------------------------##


###################################################
#####GOMPERTZ    - Exact Maximum likelihood with random alpha

#- maximum likelihood
LG_VG<-function(If1,If2,Pf1,Pf2,VGenf,Nf,x){
MV<- -(-sum(Pf2)-(Nf/2)*log(2*pi)-(Nf/2)*log((x[5]^2)/(2*x[4]))-(1/2)*
sum(log(1-(exp(-x[4]*(If2-If1)))^2))-(1/2)*
log((((2*x[4]*x[3]^2)/(x[5]^2))*(sum((1-(exp(-x[4]*(If2-If1))))/(1+exp(-x[4]*
```

```
(If2-If1)))))+1))-
((x[4])/((((2*x[4]*x[3]^2)/(x[5]^2))*(sum((1-(exp(-x[4]*(If2-If1))))/
(1+exp(-x[4]*(If2-If1)))))+1)*
(x[5]^2)))*sum((Pf2-(x[1]+x[2]*VGenf)-(Pf1-(x[1]+x[2]*VGenf))*(exp(-x[4]*
(If2-If1))))^2/(1-(exp(-x[4]*(If2-If1)))^2))+
((2*(x[4]^2)*x[3]^2)/((x[5]^4)*(((2*x[4]*x[3]^2)/(x[5]^2))*(sum((1-(exp(-x[4]*
(If2-If1))))/(1+exp(-x[4]*
(If2-If1)))))+1)))*((sum((Pf2-Pf1*(exp(-x[4]*(If2-If1))))/(1+(exp(-x[4]*
(If2-If1))))))^2)-
((x[4]*((((2*x[4]*x[3]^2)/(x[5]^2))*(sum((1-(exp(-x[4]*(If2-If1))))/
(1+exp(-x[4]*(If2-If1)))))+1)-1))/((((2*x[4]*x[3]^2)/
(x[5]^2))*(sum((1-(exp(-x[4]*(If2-If1))))/(1+exp(-x[4]*(If2-If1)))))+1)*
(x[5]^2)))*(sum((Pf2-Pf1*(exp(-x[4]*(If2-If1))))^2/
(1-(exp(-x[4]*(If2-If1)))^2))))
MV}


#-Sum of the Likelihoods
LTG_VG<-function(x){
LTG_VG<-0
for (k in 1:dim(Pf)[2]) {
LTG_VG<-LTG_VG+LG_VG(If[1:(Ncf[k]-1),k],If[2:Ncf[k],k],
log(Pf[1:(Ncf[k]-1),k]),log(Pf[2:Ncf[k],k]),VGenf[k],Nf[k],
c(x[1],x[2],x[3],x[4],x[5]))}
return(LTG_VG)}

# Parameters estimates
c0G<-round(mLTG$estimate[1],4)
c1G<-round(mLTG$estimate[2],4)
c0G_c<-c0G+c1G*MVGen #c0 centrado
AG<-round(exp(c0G+c1G*MVGen),2) #h^-1(c0G+c1G*MVen)
seG<-mLTG$estimate[3]
bG<-round(mLTG$estimate[4],4)
sG<-round(mLTG$estimate[5],4)

# Fisher empirical matrix
vhG<-solve(mLTG$hessian)
MEc0G<-round(1.96*sqrt(vhG[1,1]),4)
MEc1G<-round(1.96*sqrt(vhG[2,2]),4)
MEAG<-round(1.96*sqrt(vhG[1,1]+(MVGen^2)*vhG[2,2]+2*MVGen*vhG[1,2])*AG,3)
MEseG<-round(1.96*sqrt(vhG[3,3]),4)
MEbG<-round(1.96*sqrt(vhG[4,4]),4)
MEsG<-round(1.96*sqrt(vhG[5,5]),4)
c0G;c1G;AG;seG;bG;sG
MEc0G;MEc1G;MEAG;MEseG;MEbG;MEsG

- mLTG$minimum ## ML value
-2*(- mLTG$minimum)+2*5 ## AIC value
-2*(- mLTG$minimum)+5*log(sum(Nf)) ## BIC value
```

```
##---------------------------------##---------------------------------##
##                        SDE Mixed without VGs                       ##
##---------------------------------##---------------------------------##


###################################################
#####GOMPERTZ    MSDE PAPER
N <- Nf;Nc <- Ncf;I <- If;P <- Pf

# maximum likelihood
LG<-function(I1,I2,P1,P2,N,x){
MV<- -(-sum(P2)-(N/2)*log(2*pi)-(N/2)*log((x[4]^2)/(2*x[3]))-(1/2)*
sum(log(1-(exp(-x[3]*(I2-I1)))^2))-(1/2)*
log((((2*x[3]*x[2]^2)/(x[4]^2))*(sum((1-(exp(-x[3]*(I2-I1))))/(1+exp(-x[3]*
(I2-I1)))))+1))- ((x[3])/((((2*x[3]*x[2]^2)/(x[4]^2))*(sum((1-(exp(-x[3]*
(I2-I1))))/(1+exp(-x[3]*(I2-I1)))))+1)* (x[4]^2)))*sum((P2-x[1]-(P1-x[1])*
(exp(-x[3]*(I2-I1))))^2/(1-(exp(-x[3]*(I2-I1)))^2))+ ((2*(x[3]^2)*x[2]^2)/
((x[4]^4)*(((2*x[3]*x[2]^2)/(x[4]^2))*(sum((1-(exp(-x[3]*(I2-I1))))
/(1+exp(-x[3]*(I2-I1)))))+1)))*((sum((P2-P1*(exp(-x[3]*(I2-I1))))/
(1+(exp(-x[3]*(I2-I1))))))^2)-
((x[3]*((((2*x[3]*x[2]^2)/(x[4]^2))*(sum((1-(exp(-x[3]*(I2-I1))))/
(1+exp(-x[3]*(I2-I1)))))+1)-1))/((((2*x[3]*x[2]^2)/
(x[4]^2))*(sum((1-(exp(-x[3]*(I2-I1))))/(1+exp(-x[3]*(I2-I1)))))+1)*(x[4]^2)))*
(sum((P2-P1*(exp(-x[3]*(I2-I1))))^2/(1-(exp(-x[3]*(I2-I1)))^2))))
MV}

LTG<-function(x){
LTG<-0
for (k in 1:dim(P)[2]) {
LTG<-LTG+LG(I[1:(Nc[k]-1),k],I[2:Nc[k],k],log(P[1:(Nc[k]-1),k]),
log(P[2:Nc[k],k]),N[k],c(x[1],x[2],x[3],x[4]))}
return(LTG)}

mLTG<-nlm(LTG,c(6.1,0.09,1.747,0.33),hessian=TRUE)

# parameters estimates
aG<-round(mLTG$estimate[1],4)
AG<-round(exp(aG),2)
tG<-mLTG$estimate[2]
bG<-round(mLTG$estimate[3],4)
sG<-round(mLTG$estimate[4],4)

vhG<-solve(mLTG$hessian)
MEaG<-round(1.96*sqrt(vhG[1,1]),4)
MEAG<-round(1.96*sqrt(vhG[1,1])*AG,3)
MEtG<-round(1.96*sqrt(vhG[2,2]),4)
MEbG<-round(1.96*sqrt(vhG[3,3]),4)
MEsG<-round(1.96*sqrt(vhG[4,4]),4)
aG;AG;tG;bG;sG
MEaG;MEAG;MEtG;MEbG;MEsG
```

```
- mLTG$minimum ## ML value
-2*(- mLTG$minimum)+2*4 ## AIC value
-2*(- mLTG$minimum)+4*log(sum(N)) ## BIC value


############################################ GOMPERTZ NMSDE PAPER

#-Log-likelihood function for a trajectry :
L<-function(I1,I2,P1,P2,N,x){
MV<- +sum(P2)+(N*log(2*pi)/2)+((N/2)*log((x[3]^2)/(2*x[2])))+
(sum(log(1-exp(-2*x[2]*(I2-I1))))/2)+(x[2]/(x[3]^2))*
(sum(((P2-x[1]-(P1-x[1])*exp(-x[2]*(I2-I1)))^2)/
(1-exp(-2*x[2]*(I2-I1)))))
MV}

LTG<-function(x){
LTG<-0
for (k in 1:dim(P)[2]) {
LTG<-LTG+L(I[1:(Nc[k]-1),k],I[2:Nc[k],k],log(P[1:(Nc[k]-1),k]),
log(P[2:Nc[k],k]),N[k],c(x[1],x[2],x[3]))
}
return(LTG)}

minLTG<-nlm(LTG,c(6.6,1.2837,0.400),hessian=TRUE)

aG<-minLTG$estimate[1]
AG<-exp(aG)
bG<-minLTG$estimate[2]
sG<-minLTG$estimate[3]
aG;AG;bG;sG

vhG<-solve(minLTG$hessian)
MEaG<-1.96*sqrt(vhG[1,1])
MEAG<-1.96*sqrt(vhG[1,1])*exp(aG)
MEbG<-1.96*sqrt(vhG[2,2])
MEsG<-1.96*sqrt(vhG[3,3])
MEaG;MEAG;MEbG;MEsG

- minLTG$minimum  ## ML value
-2*(- minLTG$minimum)+2*3 ## AIC value
-2*(- minLTG$minimum)+3*log(sum(N)) ## BIC value
```

Bibliography

[1] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188, 1977.

[2] O. Garcia. A stochastic differential equation model for the height growth of forest stands. *Biometrics*, pages 1059–1072, 1983.

[3] J. M. Beaulieu, D. C. Jhwueng, C. Boettiger, and B. C. O'Meara. Modeling stabilizing selection: expanding the ornstein–uhlenbeck model of adaptive evolution. *Evolution*, 66(8):2369–2383, 2012.

[4] S. Donnet and A. Samson. A review on estimation of stochastic differential equations for pharmacokinetic/pharmacodynamic models. *Advanced drug delivery reviews*, 65(7):929–939, 2013.

[5] I. Botha, R. Kohn, and C. Drovandi. Particle methods for stochastic differential equation mixed effects models. *Bayesian Analysis*, 16(2):575–609, 2021.

[6] S. Wiqvist, A. Golightly, A. T. McLean, and U. Picchini. Efficient inference for stochastic differential equation mixed-effects models using correlated particle pseudo-marginal algorithms. *Computational Statistics & Data Analysis*, 157:107151, 2021.

[7] M. G. Ruse, A. Samson, and 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):167–193, 2020.

[8] P. Rupšys. Height–diameter models with stochastic differential equations and mixed-effects parameters. *Journal of Forest Research*, 20:9–17, 2015.

[9] A. Melnikov, Y. Mishura, and G. Shevchenko. Stochastic viability and comparison theorems for mixed stochastic differential equations. *Methodology and computing in applied probability*, 17(1):169–188, 2015.

[10] M. Delattre. A review on asymptotic inference in stochastic differential equations with mixed effects. *Japanese Journal of Statistics and Data Science*, 4(1):543–575, 2021.

[11] M. Delattre and C. Dion. *MsdeParEst: Parametric Estimation in Mixed-Effects Stochastic Differential Equations*, 2017. R package version 1.7.

[12] C. Dion, A. Sansom, and S. Hermann. *mixedsde: Estimation Methods for Stochastic Differential Mixed Effects Models*, 2018. R package version 5.0.

[13] P. A. Filipe, C. A. Braumann, N. M. Brites, and C. J. Roquete. Prediction for individual growth in a random environment. *In: Recent Developments in Modeling and Applications in Statistics*, pages 193–201, 2013.

[14] N. T. Jamba, G. Jacinto, P. A. Filipe, and C. A. Braumann. Likelihood function through the delta approximation in mixed sde models. *Mathematics*, 10(3):385, 2022.

[15] P. A. Filipe, C. A. Braumann, N. M. Brites, and C. J. Roquete. Modelling animal growth in random environments: an application using nonparametric estimation. *Biometrical Journal*, 52(5):653–666, 2010.

[16] P. A. Filipe, C. A. Braumann, and C. J. Roquete. Multiphasic individual growth models in random environments. *Methodology and Computing in Applied Probability*, 14:49–56, 2012.

[17] G. Jacinto, P. A. Filipe, and C. A. Braumann. Profit optimization of cattle growth with variable prices. *Methodology and Computing in Applied Probability*, 24:1917–1952, 2022.

[18] G. Jacinto, P. A. Filipe, and C. A. Braumann. Weighted maximum likelihood estimation for individual growth models. *Optimization*, 71(11):3295–3311, 2022.

[19] U. Picchini and S. Ditlevsen. Practical estimation of high dimensional stochastic differential mixed-effects models. *Computational Statistics & Data Analysis*, 55(3):1426–1444, 2011.

[20] R. Wolfinger. Laplace's approximation for nonlinear mixed models. *Biometrika*, 80(4):791–795, 1993.

[21] R. V. Overgaard, N. Jonsson, C. W. Tornøe, and H. Madsen. Non-linear mixed-effects models with stochastic differential equations: Implementation of an estimation algorithm. *Journal of Pharmacokinetics and Pharmacodynamics*, 32:85–107, 2005.

[22] U. Picchini, A. D. Gaetano, and S. Ditlevsen. Stochastic differential mixed-effects models. *Scandinavian Journal of Statistics*, 37(1):67–90, 2010.

[23] M. Delattre and C. Dion. Msdeparest: Parametric estimation in mixed-effects stochastic differential equations. 2017. R package version 1.7.

[24] C. Dion, A. Sansom, and S. Hermann. mixedsde: Estimation methods for stochastic differential mixed effects models. 2018. R package version 5.0.

[25] S. Klim, S. B. Mortensen, N. R. Kristensen, R. V. Overgaard, and 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):279–289, 2009.

[26] M. Delattre, V. Genon-Catalot, and A. Samson. Maximum likelihood estimation for stochastic differential equations with random effects. *Scandinavian Journal of Statistics*, 40(2):322–343, 2013.

[27] C. A. Braumann. *Introduction to Stochastic Differential Equations with Applications to Modelling in Biology and Finance*. John Wiley & Sons, 2019.

[28] J. Leander, U. G. Jirstrand, R. Eriksson, and 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*, page 76, 2022.

[29] K. Wang, L. Marciani, G. L. Amidon, D. E. Smith, and D. Sun. Stochastic differential equation-based mixed effects model of the fluid volume in the fasted stomach in healthy adult human. *The AAPS J*, pages 193–201, 2023.

[30] U. Picchini, S. Ditlevsen, and A. D. 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):141–155, 2008.

[31] U. Picchini. *Stochastic Differential Models with Applications to Physiology*. PhD thesis, University of Rome Rome, Italy, 2006.

[32] P. A. Filipe, C. A. Braumann, and C. Carlos. Profit optimization for cattle growing in a randomly fluctuating environment. *Optimization*, 64(6):1393–1407, 2015.

[33] C. A. Braumann, P. A. Filipe, C. Carlos, and C. J. Roquete. Growth of individuals in randomly fluctuating environments. *In Proc. Intern. Conf. in Computational and Mathematical Methods in Science and Engineering*, pages 201–212, 2009.

[34] P. A. Filipe. *Equações diferenciais estocásticas na modelação do crescimento individual em ambiente aleatório*. PhD thesis, University of Évora, 2011.

[35] G. Jacinto, P. A. Filipe, and C. A. Braumann. Profit optimization of cattle growth with variable prices. *Methodology and Computing in Applied Probability*, pages 1–36, 2021.

[36] M. Delattre, V. Genon-Catalot, and C. Larédo. Parametric inference for discrete observations of diffusion processes with mixed effects. *Stochastic Processes and their Applications*, 128(6):1929–1957, 2018.

[37] N. T. Jamba, G. Jacinto, P. A. Filipe, and C. A. Braumann. Estimation for stochastic differential equations mixed models using approximation methods. *AIMS Mathematics*, 9(4):7866–7894, 2024.

[38] P. A. Filipe and C. A. Braumann. Modelling individual animal growth in random environments. *In* Proceedings of the 23rd International Workshop on Statistical Modelling*; Eilers, P.H.C., Ed., Utrecht, Netherlands*, pages 232–237, 2009.

[39] R. C. Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.

[40] Q. Lv and J. W. Pitchford. Stochastic von bertalanffy models, with applications to fish recruitment. *Journal of Theoretical Biology*, 244(4):640–655, 2007.

[41] U. Picchini, S. Ditlevsen, and A. D. Gaetano. Modeling the euglycemic hyperinsulinemic clamp by stochastic differential equations. *Journal of Mathematical Biology*, 53:771 – 796, 2006.

[42] U. Picchini, S. Ditlevsen, and A. D. Gaetano. Maximum likelihood estimation of a time-inhomogeneous stochastic differential model of glucose dynamics. *Mathematical medicine and biology : a journal of the IMA*, 25 2:141–55, 2008.

[43] K. Rennolls. Forest height growth modelling. *Forest Ecology and Management*, 71(3):217–225, 1995. Growth and yield estimation from successive forest inventories.

[44] M. Delattre, V. Genon-Catalot, and A. Samson. Estimation of population parameters in stochastic differential equations with random effects in the diffusion coefficient. *ESAIM: Probability and Statistics*, 19:671–688, 2015.

[45] N. Carolino and L. T. Gama. Indicators of genetic erosion in an endangered population: The Alentejana cattle breed in Portugal1. *Journal of Animal Science*, 86(1):47–56, 2008.

[46] N. Carolino, A. Vitorino, I. Carolino, J. Pais, N. Henriques, M. Silveira, and A. Vicente. Genetic diversity in the Portuguese Mertolenga cattle breed assessed by pedigree analysis. *Animals*, 10(11):1990, 2020.

[47] J. Ruane and J. J. Colleau. Marker assisted selection for genetic improvement of animal populations when a single qtl is marked. *Genetics Research*, 66(1):71–83, 1995.