



**HAL**  
open science

# Bayesian analysis of growth curves using mixed models defined by stochastic differential equations

Sophie Donnet, Jean-Louis Foulley, Adeline Samson

► **To cite this version:**

Sophie Donnet, Jean-Louis Foulley, Adeline Samson. Bayesian analysis of growth curves using mixed models defined by stochastic differential equations. *Biometrics*, 2010, 66 (3), pp.733-741. 10.1111/j.1541-0420.2009.01342.x . hal-00360111

**HAL Id: hal-00360111**

**<https://hal.science/hal-00360111v1>**

Submitted on 10 Feb 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Bayesian analysis of growth curves using mixed models defined by stochastic differential equations

**Sophie Donnet<sup>1</sup>**

Ceremade, Universite Dauphine, France

\**email*: sophie.donnet@ceremade.dauphine.fr

and

**Jean-Louis Foulley<sup>2</sup>**

INRA, UMR GABI, CR Jouy, France

\**email*: jean-louis.foulley@jouy.inra.fr

and

**Adeline Samson<sup>3</sup>**

Laboratoire MAP5, Universite Paris Descartes, France

\**email*: adeline.samson@parisdescartes.fr

**SUMMARY:** Growth curve data consist of repeated measurements of a continuous growth process over time among a population of individuals. These data are classically analyzed by nonlinear mixed models. However, the standard growth functions used in this context prescribe monotone increasing growth and can fail to model unexpected changes in growth rates. We propose to model these variations using stochastic differential equations (SDEs) that are deduced from the standard deterministic growth function by adding random variations to the growth dynamics. A Bayesian inference of the parameters of these SDE mixed models is developed. In the case when the SDE has an explicit solution, we describe an easily implemented Gibbs algorithm. When the conditional distribution of the diffusion process has no explicit form, we propose to approximate it using the Euler-Maruyama scheme. Finally, we suggest to validate the SDE approach via criteria based on the predictive posterior distribution. We illustrate the efficiency of our method using the Gompertz function to model data on chicken growth, the modeling being improved by the SDE approach.

--

KEY WORDS: Bayesian estimation; Euler-Maruyama scheme; Gompertz model; Growth curves; Mixed models; Predictive posterior distribution; Stochastic differential equation

## 1. Introduction

Growth curve data consist of repeated measurements of a continuous growth process over time among a population of individuals. In agronomy, growth data allow differentiating animal or vegetal phenotypes by characterizing the dynamics of the underlying biological process. In gynecology or pediatrics, height and weight of fetus/children are regularly recorded to control their development. The parametric statistical approach commonly used to analyze these longitudinal data is mixed model methodology (Huggins and Loesch, 1998). The regression function of this mixed model is classically a parametric growth function, such as the Gompertz, logistic, Richards or Weibull functions (Zimmerman and Núñez-Antón, 2001) which prescribe monotone increasing growth, whatever the parameter values. These models have proved their efficiency in animal genetics (Hou et al., 2005; Jaffrézic et al., 2006, *e.g.*) and in pediatrics (Hlaing et al., 2001; Spyrides et al., 2008, *e.g.*). However, as pointed out by Davidian and Giltinan (2003), the used function may not capture the exact process, as responses for some individuals may display some local fluctuations such as weight decreases or growth slow down. These phenomena are not due to error measurements but are induced by an underlying biological process that is still unknown today. In animal genetics, a wrong modeling of these curves could affect the genetic analysis. In fetal growth, the detection of growth slow down is a crucial indicator of fetal development problems.

The aim of this paper is to model these variations in growth rate using a stochastic differential equation (SDE) whose solution is the regression term of the mixed model. More precisely, each growth function is defined as the solution of an ordinary differential equation (ODE); this ODE models the growth rate. We suggest to add a random perturbation to the ODE, resulting in an SDE. Thus, the growth rate varies randomly around the mean dynamics. In this paper, we propose and study Bayesian estimators for mixed models defined by SDEs. Estimator properties are illustrated for simulated and real animal growth data.

Parametric estimation by maximum likelihood of SDE with random parameters (without measurement noise) has been studied by Ditlevsen and De Gaetano (2005). However, estimation of SDE mixed models (including the measurement noise modeling) has received little attention. Overgaard et al. (2005) and Tornøe et al. (2004) proposed estimators based on an extended Kalman filter, but the algorithm convergence was not proved. Donnet and Samson (2008) proposed an EM-based estimator (Dempster et al., 1977) and prove the convergence of their algorithm. Whereas the Bayesian point of view is widely used on standard growth curves, Bayesian estimation of SDE mixed models has not been much investigated. Cano et al. (2006) computed the posterior distribution by approximating the diffusion process by an Euler scheme. Oravec et al. (in press) studied the Bayesian estimation of an Ornstein-Uhlenbeck process with random parameters. In this paper, we propose either to use a judicious transformation of the SDE to compute the exact conditional distribution of the diffusion process, or, if it is not possible, to approximate the diffusion by the Euler-Maruyama scheme. Then we propose a Gibbs algorithm to simulate the exact or the approximate posterior distributions. In the case of approximation by the Euler scheme, we control the error induced by this scheme on the posterior distributions. Finally, we adapt the computation of the posterior predictive distributions to validate the SDE mixed model (Meng, 1994).

Section 2 presents the classical nonlinear mixed model and the mixed model defined by SDEs. We discuss the choice of the volatility term in the SDEs. In Section 3, we suggest some prior specifications and posterior computation. We also present the Euler-Maruyama scheme. Section 4 shows how to validate the SDE mixed model using posterior predictive distributions. In Section 5, the theory is illustrated with the particular case of the Gompertz function applied on chicken growth data.

## 2. Models and notations

### 2.1 Nonlinear mixed models

Let  $\mathbf{y} = (y_i)_{1 \leq i \leq n} = (y_{ij})_{1 \leq i \leq n, 1 \leq j \leq n_i}$  denote the data, where  $y_{ij}$  is the noisy measurement of the observed biological process for individual  $i$  at time  $t_{ij}$ , for  $i = 1, \dots, n$ ,  $j = 0, \dots, n_i$ . In classical mixed models, the evolution of the process is modeled by a deterministic function, depending on individual random parameters. Formally, the classical nonlinear mixed model is defined as:

$$\begin{aligned} y_{ij} &= f(\phi_i, t_{ij}) + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim_{i.i.d.} \mathcal{N}(0, \sigma^2) \\ \phi_i &\sim \mathcal{N}(\mu, \Omega) \end{aligned} \quad (1)$$

with  $f$  being a parametric deterministic function and  $\phi = (\phi_i)_{1 \leq i \leq n}$  the  $p$ -vectors of individual parameter vectors. The  $\phi_i$  are assumed to be independently and identically normally distributed with expectation  $\mu$  and variance  $\Omega$ . The  $\varepsilon_{ij}$  are the residual errors, assumed to be independently and identically normally distributed with null mean and variance equal to  $\sigma^2$ .

For growth curve data,  $f$  is classically one of the four most famous parametric functions modeling growth curves, namely the logistic, the Gompertz, the Richards and the Weibull functions. Each of them can be written as the solution of an ordinary differential equation (ODE) describing the evolution of growth rate, which are respectively:

$$f'(t) = Cf(t) \left[ 1 - \frac{1}{A}f(t) \right], \quad f(0) = \frac{A}{1+B} \quad (\text{Logistic}) \quad (2)$$

$$f'(t) = BCe^{-Ct}f(t), \quad f(0) = Ae^{-B} \quad (\text{Gompertz}) \quad (3)$$

$$f'(t) = \frac{BCDe^{-Ct}}{1 + Be^{-Ct}}f(t), \quad f(0) = \frac{A}{(1+B)^D} \quad (\text{Richards}) \quad (4)$$

$$f'(t) = DCt^{D-1}(A - f(t)), \quad f(0) = A - B \quad (\text{Weibull}) \quad (5)$$

where  $A, B, C, D$  are non-negative parameters.  $A$  is the upper asymptote,  $C$  and  $D$  are growth rate parameters. All four models prescribe monotone increasing curves. More gen-

erally, if  $\phi$  denotes the parameter vector (either  $(A, B, C)$ ,  $(A, B, C, D)$  or a well-chosen parametrization),  $f$  is the solution of the following ODE:

$$\frac{\partial f(\phi, t)}{\partial t} = F(f, t, \phi), \quad f(\phi, 0) = f_0(\phi) \quad (6)$$

## 2.2 Nonlinear mixed models defined by stochastic differential equations

In this section, we extend the classical nonlinear mixed model by replacing the regression function by a stochastic process. In order to take into account individuals whose growth curve suffers from an unexpected growth rate change, we propose to introduce a stochastic term in the ODE (6). Growth curve is thus described by a random process, denoted  $(Z_t)$ , which is the solution of the following SDE:

$$dZ_t = F(Z_t, t, \phi)dt + \Gamma(Z_t, \phi, \gamma^2)dW_t, \quad Z(t=0) = Z_0(\phi) \quad (7)$$

where  $W_t$  is a Brownian motion.  $\Gamma(Z_t, \phi, \gamma^2)$  is the volatility function depending on the unknown parameter  $\gamma^2$ .

The nonlinear mixed model defined by an SDE is thus:

$$\begin{aligned} y_{ij} &= Z_{t_{ij}}(\phi_i) + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim_{i.i.d.} \mathcal{N}(0, \sigma^2) \\ dZ_t(\phi_i) &= F(Z_t, t, \phi_i)dt + \Gamma(Z_t, \phi, \gamma^2)dW_t \\ \phi_i &\sim \mathcal{N}(\mu, \Omega) \end{aligned} \quad (8)$$

In model (8), three fundamentally different noises are distinguished: the inter-subject variability  $\Omega$ , which is the individual parameter variance, the dynamic noise  $\gamma^2$ , reflecting the random fluctuations around the corresponding theoretical dynamic model, and the measurement noise  $\sigma^2$  representing the uncorrelated part of the residual variability associated with assay or sampling errors.

Many types of volatility functions can be proposed to extend an ODE into an SDE (*e.g.* constant, square root or polynomial volatility). For a given equation, the choice depends on several considerations. If the observed biological process is non-negative, a volatility function

that ensures the positivity of  $(Z_t)$  will be chosen. If biological reasons imply that a model parameter fluctuates along the experiment record, then the volatility can be derived by adding a random perturbation to this parameter. If heteroscedastic variances have been used in an ODE modeling approach, a polynomial volatility can be chosen. Finally, algorithmic and computational constraints have to be considered: an SDE with explicit solution implies a simpler estimation scheme leading to good estimation properties (convergence of the algorithm to the true posterior distribution) whereas an SDE without explicit solution implies additional computational difficulties (use of an approximation scheme). As an example, we propose to use an affine volatility function  $\Gamma(Z_t, \phi, \gamma^2) = \gamma Z_t$ , for the logistic (2), Gompertz (3) and Richards (4) models: the process  $(\log Z_t)$  is then a Gaussian process (see Section 5.2 for more details).

### 3. Bayesian estimation

#### 3.1 Prior specification

The Bayesian approach consists in the evaluation of the posterior distribution of the population parameters  $\mu, \Omega, \sigma^2$  and the volatility  $\gamma^2$  for the SDE model. The first step is thus the choice of the prior distribution. Usual diffuse prior distributions can be chosen but the resulting posterior distributions may not be proper. Therefore, we suggest to use standard prior distributions proposed, among others, by De la Cruz-Mesia and Marshall (2006) for expectation or variance parameters in hierarchical models:

$$\begin{aligned}
 \mu_k &\sim \mathcal{N}(m_k^{prior}, v_k^{prior}), \quad k = 1, \dots, p \\
 \Omega^{-1} &\sim W(R, p + 1) \\
 1/\sigma^2 &\sim \Gamma(\alpha_\sigma^{prior}, \beta_\sigma^{prior})
 \end{aligned} \tag{9}$$



The  $\gamma^2$  parameter controls the variance of the random perturbation. Therefore it seems reasonable to choose as prior for  $\gamma^2$  an inverse-Gamma distribution:

$$1/\gamma^2 \sim \Gamma(\alpha_\gamma^{prior}, \beta_\gamma^{prior})$$

In practice the specification of hyperparameters  $m_k^{prior}, v_k^{prior}, R, \alpha_\sigma^{prior}, \beta_\sigma^{prior}, \alpha_\gamma^{prior}, \beta_\gamma^{prior}$  may be difficult. Therefore we can choose the values of hyperparameters to obtain non-informative priors.

### 3.2 Posterior computation

Since models (1) and (8) are non-linear, posterior distributions are not explicit and iterative estimation procedures have to be used. For the ODE model (1), Gibbs sampling algorithms including the sampling of the auxiliary random variables  $\phi_i$  under their conditional distributions have been proposed in the literature (Carlin and Louis, 2000, *e.g.*). These algorithms do not present any particular difficulties and are not detailed here. For the SDE model (8), we propose to use a Gibbs algorithm, including the sampling of the auxiliary random variables  $\phi_i$  and the vectors  $Z_i$  of realizations of process  $(Z_t)$  for each individual at each observation time. Let  $\mathbf{Z} = (Z_1, \dots, Z_n) \in \mathbf{R}^{(n_1+1)+\dots+(n_n+1)}$  denote the vector of the  $n$  realizations. Hence the Gibbs sampling algorithm for the SDE model is outlined as follows:

- STEP 1: initialize the iteration counter of the chain  $k = 1$  and start with initial values  $\sigma^{-2(0)}, \gamma^{2(0)}, \mu^{(0)}, \phi^{(0)}, \mathbf{Z}^{(0)}$ .
- STEP 2: obtain  $\sigma^{-2(k)}, \gamma^{2(k)}, \mu^{(k)}, \phi^{(k)}, \mathbf{Z}^{(k)}$  from  $\sigma^{-2(k-1)}, \gamma^{2(k-1)}, \mu^{(k-1)}, \phi^{(k-1)}, \mathbf{Z}^{(k-1)}$  through successive generations of

- (1)  $\mathbf{Z}^{(k)} \sim p(\mathbf{Z}|\phi^{(k-1)}, \gamma^{-2(k-1)}, \sigma^{-2(k-1)}, \mathbf{y})$
- (2)  $\phi^{(k)} \sim p(\phi|\sigma^{-2(k-1)}, \gamma^{-2(k-1)}, \mu^{(k-1)}, \Omega^{(k-1)}, \mathbf{Z}^{(k)}, \mathbf{y}_0)$  where  $\mathbf{y}_0 = (y_{i0})_{i=1\dots n}$
- (3)  $\mu^{(k)} \sim p(\mu|\phi^{(k)})$  and  $\Omega^{(k)} \sim p(\Omega|\phi^{(k)})$
- (4)  $\sigma^{-2(k)} \sim p(\sigma^{-2}|\mathbf{Z}^{(k)}, \phi^{(k)}, \mathbf{y})$  and  $\gamma^{-2(k)} \sim p(\gamma^{-2}|\mathbf{Z}^{(k)}, \phi^{(k)})$

- **STEP 3:** change  $k$  to  $k + 1$  and return to **STEP 2** until convergence is reached.

Some conditional distributions are explicit. A Gamma prior distribution on  $\sigma^{-2}$  implies that  $p(\sigma^{-2} | \mathbf{Z}^{(k)}, \boldsymbol{\phi}^{(k)}, \mathbf{y})$  is a Gamma density. The prior distribution of  $p(\boldsymbol{\phi} | \mu, \Omega)$  being Gaussian, then the conditional distribution of  $\mu$  is Gaussian and the conditional distribution of  $\Omega$  is inverse Wishart. The conditional distributions on  $\boldsymbol{\phi}$ ,  $\mathbf{Z}$  and  $\gamma^2$  depend on the specific form of the SDE and will be detailed in the particular example of the Gompertz model in Section 5. Depending on the complexity of the model, we may have to resort to Metropolis-Hastings algorithms. Moreover, for SDEs without explicit solution, the conditional distribution on  $\mathbf{Z}$  has generally not a closed form. In this case, we suggest to approximate the diffusion by the Euler-Maruyama scheme, which leads to Gaussian approximations of the transition densities. We then introduce an approximate statistical model on which the posterior distributions are computed. Details are given thereafter.

### 3.3 Posterior distribution using Euler-Maruyama approximation

Let us briefly recall the Euler-Maruyama scheme for subject  $i$ . If the time intervals between the observation instants are too great to obtain a good approximation of the transition density, a natural approach is to introduce a set of auxiliary latent data points between every pair of observations, as first proposed by Pedersen (1995). Let  $t_{i0} = \tau_0 < \tau_1 < \dots < \tau_m < \dots < \tau_{M_i} = t_{i,n_i}$  denote the equally spaced discretization of the time interval  $[t_{i0}, t_{i,n_i}]$ . Let  $h$  be the step size of the discretization. Let us assume that, for all  $j = 0 \dots n_i$ , there exists an integer  $m_j$  verifying  $t_{ij} = \tau_{m_j}$ , with  $m_0 = 0$  by definition. Then the diffusion process denoted  $\tilde{Z}^h$  and supplied by the Euler-Maruyama approximation of the SDE is described by the following iterative scheme: for a fixed  $\phi_i$ ,  $\tilde{Z}_0^h = Z_0(\phi_i)$ , and for  $m = 1 \dots M_i$ ,

$$\tilde{Z}_m^h = \tilde{Z}_{m-1}^h + h F(\tilde{Z}_{m-1}^h, \tau_{m-1}, \phi_i) + \Gamma(\tilde{Z}_{m-1}^h, \phi_i, \gamma^2) \sqrt{h} \xi_m, \xi_m \sim_{i.i.d} \mathcal{N}(0, 1)$$

Consequently,  $(\tilde{Z}_{m_0}^h, \dots, \tilde{Z}_{m_{n_i}}^h)$  is an approximation of the original diffusion process at time points  $(t_{i0}, \dots, t_{in_i})$ .

Using this approximation of the diffusion process provided by the Euler-Maruyama scheme of step size  $h$ , an approximate statistical model is defined as:

$$y_{ij} = \tilde{Z}_{m_j}^h(\phi_i) + \varepsilon_{ij}, \varepsilon_{ij} \sim_{i.i.d.} \mathcal{N}(0, \sigma^2) \quad (10)$$

$$\tilde{Z}_m^h(\phi_i) = \tilde{Z}_{m-1}^h(\phi_i) + h F(\tilde{Z}_{m-1}^h, \tau_{m-1}, \phi_i) + \Gamma(\tilde{Z}_{m-1}^h, \phi_i, \gamma^2) \sqrt{h} \xi_m, 1 \leq m \leq M_i, \quad (11)$$

$$\xi_m \sim_{i.i.d.} \mathcal{N}(0, 1)$$

$$\phi_i \sim \mathcal{N}(\mu, \Omega)$$

For model (10), the conditional distribution of the approximate diffusion  $\tilde{Z}^h$  is Gaussian, allowing to implement the previously presented Gibbs algorithm. The convergence of this Gibbs algorithm is ensured by classical results (Carlin and Louis, 2000). However, this Gibbs algorithm is performed on the approximate model (10), and computes the posterior distribution  $p^h(\theta|\mathbf{y})$  of model (10), with  $\theta = (\mu, \Omega, \sigma^2, \gamma^2)$ , instead of the original posterior distribution  $p(\theta|\mathbf{y})$ . But, the error induced by the Euler scheme on the posterior distributions can be controlled, as shown in the Appendix.

#### 4. Model validation

The goal in model checking is to monitor the quality of the proposed model, i.e. to determine whether the observed data are representative of the type of data we might expect under this model. Posterior predictive checks set this up by generating replicated data sets from the estimated posterior distribution of the model. These replicated data sets are then compared with the observed data. The function used to compare observed and replicated datasets is the discrepancy function; it depends on data and model parameters and is denoted  $T(y, \eta)$ ,  $\eta$  being used as generic notation for a function of the model parameters. It quantifies incompatibility of the model with the observed data. In our case, we consider for  $T$  the

$\chi^2$  discrepancy:

$$T(\mathbf{y}, \eta) = \frac{(\mathbf{y} - \eta)^2}{\text{Var}(\mathbf{y} - \eta)}$$

For the observation at time  $t_{ij}$ , we choose  $\eta = f(\phi, t_{ij})$  for the ODE model and  $\eta = Z_{ij}(\phi_i)$  for the SDE model. Consequently, for both models,  $\text{Var}(\mathbf{y} - \eta) = \sigma^2$ .

We aim at comparing the posterior distribution  $p(T(\mathbf{y}, \eta)|\mathbf{y})$  of the observed data  $\mathbf{y}$  with the posterior distribution  $p(T(\mathbf{y}_{rep}, \eta)|\mathbf{y})$  where  $\mathbf{y}_{rep}$  denotes the replicated data drawn from the posterior predictive distribution  $p(\mathbf{y}_{rep}|\mathbf{y})$ . A short version of that posterior predictive distribution is the posterior predictive  $p$ -value:

$$\begin{aligned} p_{pp} &= P \left[ T(\mathbf{y}_{rep}, \eta) \geq T(\mathbf{y}, \eta) | \mathbf{y} \right] \\ &= \int P \left[ T(\mathbf{y}_{rep}, \eta) > T(\mathbf{y}, \eta) | \mathbf{y}, \eta \right] p(\eta | \mathbf{y}) d\eta \end{aligned} \quad (12)$$

Since this quantity has no closed form, the idea is to approximate it by the Monte Carlo method. For each estimated model (ODE and SDE), the Gibbs algorithm used to estimate the posterior distribution provides a set of vectors  $\eta^l$  ( $l = 1 \dots L$ ) drawn from the posterior distribution  $p(\eta|\mathbf{y})$ . For each of this draw, a replicated data set  $\mathbf{y}_{rep}^l$  is simulated from the posterior predictive distribution of the data  $p(\mathbf{y}_{rep}|\eta^l)$ . Finally, the posterior predictive  $p$ -value (12) is estimated by the Monte Carlo method as  $\frac{1}{L} \sum_{l=1}^L 1_{T(\mathbf{y}_{rep}^l, \eta^l) > T(\mathbf{y}, \eta^l)}$ . By definition of  $T$  and  $\eta$  and for both models (ODE and SDE), we can remark that  $T(\mathbf{y}_{rep}^l, \eta^l)$  is simulated under a  $\chi^2(1)$  distribution.

## 5. An example: chicken growth modeling with the Gompertz function

We focus on the modeling of chicken growth previously analyzed by Jaffrézic et al. (2006). Data  $\mathbf{y}$  are noisy weight measurements of  $n = 50$  chickens at days  $t = 0, 4, 6, 8, 12, 16, 20, 24, 28, 32, 36, 40$  after birth: see the corresponding curves on Figure 1.

### 5.1 The classical Gompertz nonlinear mixed model

Jaffrézic et al. (2006) propose to model these growth data by a Gompertz function (3) and a mixed model. Jaffrézic and Foulley (2006) underline that a heteroscedastic error model is required to obtain satisfactory results. For simplicity's sake, we consider modeling the logarithm of the data  $\mathbf{y}$  by adding an additive measurement error with a constant variance:

$$\begin{cases} \log y_{ij} &= \log A_i - B_i e^{-C_i t_{ij}} + \varepsilon_{ij}, \varepsilon_{ij} \sim_{i.i.d.} \mathcal{N}(0, \sigma^2), \forall i = 1, \dots, n, j = 0, \dots, n_i \\ \phi_i &= (\log A_i, B_i, \log C_i) \sim_{i.i.d.} \mathcal{N}(\mu, \Omega), \forall i = 1, \dots, n \end{cases} \quad (13)$$

We use the log-parametrization for parameters  $A_i$  and  $C_i$ . This parametrization has two advantages: it simplifies the computation of the posterior distributions and it ensures the positivity of the parameters. We set  $\mu = (\log(a), b, \log(c))$ .

### 5.2 Extension to the Gompertz stochastic nonlinear mixed model

We now deduce the SDE model from the Gompertz equation (3). Given the heteroscedasticity of the process, the volatility function is set to be equal to  $\Gamma(Z_t, \phi, \gamma^2) = \gamma Z_t$ :

$$dZ_t = BCe^{-Ct} Z_t dt + \gamma Z_t dW_t, \quad Z_0 = Ae^{-B} \quad (14)$$

This means that the standard error of the random perturbations of the growth rate is proportional to weight. This choice of volatility has two main advantages. First, SDE (14) has an explicit solution. Indeed, set  $X_t = \log(Z_t)$ . By the Ito's formula, for  $h > 0$ , the conditional distribution of  $X_{t+h}$  given  $(X_s), s \leq t$  is:

$$X_{t+h} | (X_s)_{s \leq t} \sim \mathcal{N}\left(X_t - Be^{-Ct}(e^{-Ch} - 1) - \frac{1}{2}\gamma^2 h, \gamma^2 h\right), \quad X_0 = \log(A) - B$$

Thus,  $\forall t > 0$ , we have:

$$Z_t = Ae^{-Be^{-Ct}} e^{-\frac{1}{2}\gamma^2 t + \eta_t} = f(t) e^{-\frac{1}{2}\gamma^2 t + \eta_t}, \quad \eta_t \sim \mathcal{N}(0, \gamma^2 t)$$

with  $Z_0 = Ae^{-B}$ . As a consequence,  $Z_t$  is a multiplicative random perturbation of the solution of the Gompertz model. Second, due to the assumption of the non-negativity of  $A$ ,  $Z_t$  is almost surely non-negative, which is a natural constraint to model weight records.

We then discretize the SDE. The discrete realization  $(X_{t_{ij}})$  of the SDE is Markovian:

$$X_{i,t_{ij}} | X_{i,t_{ij-1}} \sim \mathcal{N} \left( X_{i,t_{ij-1}} - B_i e^{-C_i t_{ij-1}} (e^{-C_i(t_{ij}-t_{ij-1})} - 1) - \frac{1}{2} \gamma^2 (t_{ij} - t_{ij-1}), \gamma^2 (t_{ij} - t_{ij-1}) \right)$$

with  $X_{i,0} = \log(A_i) - B_i$ . The SDE model (8) on the logarithm of data is thus defined as:

$$\begin{cases} \log y_{ij} & = X_{t_{ij}} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim_{i.i.d.} \mathcal{N}(0, \sigma^2), \quad \forall i = 1, \dots, n, j = 0, \dots, n_i \\ \phi_i = (\log A_i, B_i, \log C_i) & \sim_{i.i.d.} \mathcal{N}(\mu, \Omega), \quad \forall i = 1, \dots, n \end{cases} \quad (15)$$

which can be written under the following matrix form:

$$\begin{cases} (\log y_{i0}, \log y_{i1}, \dots, \log y_{in_i})' = (\log(A_i) - B_i, X_{t_{i1}}, \dots, X_{t_{in_i}})' + \varepsilon_i, & \varepsilon_i \sim_{i.i.d.} \mathcal{N}(0, \sigma^2 \mathbf{I}_{n_i+1}) \\ (X_{t_{i1}}, \dots, X_{t_{in_i}})' = (\log(A_i) - B_i e^{-C_i t_{i1}}, \dots, \log(A_i) - B_i e^{-C_i t_{in_i}})' - \gamma^2 (t_{i1}, \dots, t_{in_i})' + \eta_i \\ \eta_i \sim_{i.i.d.} \mathcal{N}(0_J, \gamma^2 T_i), \quad T_i = (\min(t_{ij}, t_{ij'}))_{1 \leq j, j' \leq n_i} \\ (\log A_i, B_i, \log C_i) \sim_{i.i.d.} \mathcal{N}(\mu, \Omega) \end{cases} \quad (16)$$

### 5.3 Posterior computation and inference in the Gompertz model

Conditional distribution computation for the ODE mixed model is standard. We detail the computation under the SDE mixed model. Let  $m_a^{prior}, m_b^{prior}, m_c^{prior}, v_a^{prior}, v_b^{prior}, v_c^{prior}$  denote the parameters of the three priors of the components of  $\mu$ .

The conditional distribution of  $\mathbf{X}_i = (X_{ij})_{1 \leq j \leq n_i}$  given  $(\phi_i, \gamma^{-2}, \mathbf{y}_i, \sigma^2)$  is Gaussian:

$$\begin{aligned} \mathbf{X}_i | \mathbf{y}_i, \sigma^2, \gamma^2, \phi_i &\sim \mathcal{N}(m_{\mathbf{X}_i}^{post}, V_{\mathbf{X}_i}^{post}), \\ V_{\mathbf{X}_i}^{post} &= (\sigma^{-2} \mathbf{I}_{n_i-1} + \gamma^{-2} T_i^{-1})^{-1}, \quad m_{\mathbf{X}_i}^{post} = V_{\mathbf{X}_i}^{post} \left[ \sigma^{-2} (\log y_{i1} \dots \log y_{in_i})' + \gamma^{-2} T_i^{-1} u_{\mathbf{X}_i} \right] \\ u_{\mathbf{X}_i} &= \log A_i - B_i \left( e^{-C_i t_{i1}} \dots e^{-C_i t_{in_i}} \right)' - \frac{1}{2} \gamma^2 (t_{i1} \dots t_{in_i})' \end{aligned}$$

Let us introduce the  $(n_i + 1) \times (n_i + 1)$ -matrix:

$$G_i = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \gamma^2 T_i \end{pmatrix} \quad (17)$$

Let  $\omega_{\log A}^2, \omega_B^2, \omega_{\log C}^2$  denote the three diagonal elements of  $\Omega$ . Let  $\Omega_{k,(j,j')}$  denote the two-vector composed of the elements on the  $k$ -th row and  $(j, j')$  columns of  $\Omega$  and  $\Omega_{(j,j'),(j,j')}$  the two-symmetric-matrix composed of the elements on the  $(j, j')$ -th rows and  $(j, j')$ -th columns of  $\Omega$ . We have the following conditional distributions for the individual parameters  $\log A_i$

and  $B_i$ :

$$\begin{aligned} \log A_i \mid \mathbf{y}_i, \mathbf{X}_i, \mu, \Omega, B_i, \log C_i, \sigma^2, \gamma^2 &\sim \mathcal{N}(m_{1i}^{post}, V_{1i}^{post}) \\ B_i \mid \mathbf{y}_i, \mathbf{X}_i, \mu, \Omega, \log A_i, \log C_i, \sigma^2, \gamma^2 &\sim \mathcal{N}(m_{2i}^{post}, V_{2i}^{post}) \end{aligned}$$

where

$$\begin{aligned} V_{1i}^{post} &= \left[ (1 \dots 1)G_i^{-1}(1 \dots 1)' + \frac{1}{\omega_{\log A \mid B, \log C}^2} \right]^{-1} \\ m_{1i}^{post} &= V_{1i}^{post} \left[ (1 \dots 1)G_i^{-1}u_{Ai} + \frac{\mu_{\log A \mid B, \log C}}{\omega_{\log A \mid B, \log C}^2} \right] \\ u_{Ai} &= (\log y_{i0} \ X_{i1} \dots X_{in_i})' + B_i (e^{-C_i t_{i0}} \dots e^{-C_i t_{in_i}})' - \frac{1}{2}\gamma^2 (t_{i0} \dots t_{in_i})' \\ \omega_{\log A \mid B, \log C}^2 &= \omega_{\log A}^2 - \Omega_{\log A, (B, \log C)} \Omega_{(B, \log C), (B, \log C)}^{-1} \Omega'_{\log A, (B, \log C)} \\ \mu_{\log A \mid B, \log C} &= \log a + \Omega_{\log A, (B, \log C)} \Omega_{(B, \log C), (B, \log C)}^{-1} ((B_i, \log C_i)' - (b, \log c)') \end{aligned}$$

and

$$\begin{aligned} V_{2i}^{post} &= \left[ (e^{-C_i t_{i0}} \dots e^{-C_i t_{in_i}})G_i^{-1}(e^{-C_i t_{i0}} \dots e^{-C_i t_{in_i}})' + \frac{1}{\omega_{B \mid \log A, \log C}^2} \right]^{-1} \\ m_{2i}^{post} &= V_{2i}^{post} \left[ (e^{-C_i t_{i0}} \dots e^{-C_i t_{in_i}})G_i^{-1}u_{Bi} + \frac{\mu_{B \mid \log A, \log C}}{\omega_{B \mid \log A, \log C}^2} \right] \\ u_{Bi} &= (\log y_{i0}, X_{i1} \dots X_{in_i})' + \log A_i - \frac{1}{2}\gamma^2 (t_{i0} \dots t_{in_i}) \\ \omega_{B \mid \log A, \log C}^2 &= \omega_B^2 - \Omega_{B, (\log A, \log C)} \Omega_{(\log A, \log C), (\log A, \log C)}^{-1} \Omega'_{B, (\log A, \log C)} \\ \mu_{B \mid \log A, \log C} &= b + \Omega_{B, (\log A, \log C)} \Omega_{(\log A, \log C), (\log A, \log C)}^{-1} ((\log A_i, \log C_i)' - (\log a, \log c)') \end{aligned}$$

The conditional distribution of  $\log(a)$  is:

$$\begin{aligned} \log a \mid (\log A_i)_{i=1 \dots n}, \omega_{\log A}^2, m_a^{prior}, v_a^{prior} &\sim \mathcal{N}(m_a^{post}, V_a^{post}) \\ V_a^{post} &= [n\omega_{\log A}^{-2} + (v_a^{prior})^{-1}]^{-1} \text{ and } m_a^{post} = V_a^{post} \left[ \omega_{\log A}^{-2} \frac{1}{n} \sum_{i=1}^n \log A_i + \frac{m_a^{prior}}{v_a^{prior}} \right] \end{aligned}$$

Similarly, we have:

$$\begin{aligned} b \mid (B_i)_{i=1 \dots n}, \omega_B^2, m_b^{prior}, v_b^{prior} &\sim \mathcal{N}(m_b^{post}, V_b^{post}) \\ V_b^{post} &= [n\omega_B^{-2} + (v_b^{prior})^{-1}]^{-1} \text{ and } m_b^{post} = V_b^{post} \left[ \omega_B^{-2} \frac{1}{n} \sum_{i=1}^n B_i + \frac{m_b^{prior}}{v_b^{prior}} \right] \end{aligned}$$

The conditional distribution of  $\Omega^{-1}$  is:

$$\Omega^{-1} | (\phi_i)_{i=1\dots n}, \mu, R \sim W(R + (\boldsymbol{\phi} - \mu)(\boldsymbol{\phi} - \mu)', n + p + 1)$$

where  $\boldsymbol{\phi} - \mu = [(\phi_1 - \mu) \dots (\phi_n - \mu)] \in \mathbf{R}^{3 \times n}$ .

The conditional distribution of  $\sigma^2$  is:

$$\begin{aligned} \sigma^{-2} | y, \mathbf{X}, \boldsymbol{\phi}, \alpha_\sigma^{prior}, \beta_\sigma^{prior} &\sim \Gamma(\alpha_\sigma^{post}, \beta_\sigma^{post}) \\ \alpha_\sigma^{post} &= \alpha_\sigma^{prior} + \sum_{i=1} n \frac{n_i + 1}{2} \quad \text{and} \quad \beta_\sigma^{post} = \left[ \frac{1}{\beta_\sigma^{prior}} + \frac{1}{2} \sum_{i=1, j=0}^{n, n_i} (\log y_{ij} - X_{ij})^2 \right]^{-1} \end{aligned}$$

The posterior distributions of  $\log C_i$ ,  $\log c$  and  $\gamma^2$  have no explicit form and we use the Metropolis-Hastings random-walks.

The convergence of this Gibbs algorithm is ensured by the classical convergence theorem proposed by Carlin and Louis (2000), the convergence of the Metropolis-Hastings algorithm is ensured by the theorem proposed by Mengersen and Tweedie (1996).

#### 5.4 Simulations

We simulate datasets mimicking chicken growth with  $n = 50$  individuals and  $n_i = 9$  measurements obtained every 5 days after birth. The population parameters are  $\log(a) = \log(3000)$ ,  $b = 5$ ,  $\log(c) = \log(14)$ ,  $\Omega$  is assumed diagonal with diagonal elements equal to 100 and  $\sigma^{-2} = 5$ . A 100 datasets are simulated via the mixed model defined by the Gompertz model (13) and a 100 datasets with the mixed model defined by the Gompertz SDE (15), with  $\gamma^2 = 1$ . We estimate all the parameters under the ODE mixed model (13) and the SDE mixed model (15), successively. Estimates are obtained as the expectation of the parameter posterior distribution. Bias and root mean square error (RMSE) obtained with both algorithms are presented in Table 1.

[Table 1 about here.]

When data are simulated under the ODE model, estimates obtained with the Bayesian ODE algorithm are very satisfactory. Those obtained by the Bayesian SDE algorithm are also



satisfactory although the bias for the variance parameters is larger. Note that, as expected, the volatility parameter  $\gamma^2$  is estimated to be close to zero. When data are simulated under the SDE model, estimates obtained with the Bayesian SDE model are very satisfactory, with small bias and RMSE. On the contrary, the estimates obtained with the Bayesian ODE algorithm have larger bias and RMSE, especially  $\omega_{\log A}^{-2}$  (bias around 91%) and  $\sigma^{-2}$  (bias equal to 27%).

### 5.5 *Application on chicken growth data*

In this section, we apply the proposed models on the real data of chicken growth. The ODE and SDE models (13) and (15) are used to model the logarithm of the data. Posterior expectations of the parameters are presented in Table 2. Diagnostic tools to validate the models are applied to both ODE and SDE models: Table 3 presents the posterior predictive p-values of both models computed for each time point and Figure 2 presents the corresponding boxplots of the posterior predictive distributions. The estimates of the volatility parameter  $\gamma^2$  is strictly positive: this means that the dynamical process that most likely represents the growth is a stochastic process with non-negligible noise. Furthermore, the diagnostic tools show a clear improvement from the ODE model to SDE model, both at early and late ages. Figure 3 reports, for four subjects, the observed weights, the ODE prediction, the empirical mean of the last 1000 simulated trajectories of the SDE (15) generated during the Gibbs algorithm, their empirical 95 % confidence limits (from the 2.5th percentile to the 97.5th percentile) and one simulated trajectory. Subjects 4 and 13 are examples of subjects with no growth slow down. Both ODE and SDE models satisfactorily fit the observations. Subject 14 has a small observed weight decrease. For subject 1, the weight decrease is more important. For both subjects, the ODE model fails to capture this phenomenon while the SDE model does.

[Figure 1 about here.]

[Table 2 about here.]

[Figure 2 about here.]

[Table 3 about here.]

[Figure 3 about here.]

## 6. Conclusion and discussion

We propose a Bayesian approach to nonlinear mixed models defined by stochastic differential equations. These models are an alternative to classical nonlinear mixed models whose deterministic regression function is too restrictive to model some unexplained biological processes such as growth rate changes. We detail the case where the diffusion process has no explicit distribution by proposing to use the Euler-Maruyama scheme to approximate the diffusion: the conditional distribution is then Gaussian, implying an easy Bayesian implementation. We control the error induced by this Euler approximate scheme on the posterior distribution. In this context, auxiliary latent points are introduced to obtain a better approximation of the diffusion. The choice of the discrete grids  $(\tau_0, \dots, \tau_{M_i})$  is complex and has been evoked by Pedersen (1995) and Donnet and Samson (2008).

Our model differs from mixed models with continuous time autoregressive measurement errors, as proposed by De la Cruz-Mesia and Marshall (2006) or others. These authors assume that measurement errors have an auto-regressive structure. We assume that the auto-regressive structure observed in residuals of classical nonlinear mixed models comes from a model failure: the regression function is too restrictive and rigid to model random variations of the biological process. Therefore, in our model, it is the regression process that has an auto-regressive structure, while the observation measurements are assumed to be independant and identically distributed.

The proposed model should prove to be useful for other applications in which deterministic

models are too restrictive to take into account the variabilities that exist in real life. For example, Picchini et al. (2006) propose a stochastic differential equation to model glucose/insulin dynamics, where sources of variability are various (anxiety, rest, etc). The extension of this work to mixed models using our approach should be of great interest.

An interesting area for future research is the development of model selection tools in this context. Indeed, the analysis of covariate effects or the comparison between the ODE and the SDE models require specific selection tools. Indeed, the method of pseudo-priors proposed by Carlin and Chib (1995) and developed by others, which is very sensitive to the choice of priors and pseudo-priors, would be difficult to use in practice in our context. Bayes factors are complex to compute in these models but could be an interesting alternative. Finally, the extension of this work to multidimensional SDEs would be of great interest in several biological applications.

#### ACKNOWLEDGEMENTS

The authors are grateful to Dr Florence Jaffrzic for helpful discussions on mixed model methodology for longitudinal data and to Mrs Wendy Brand Williams and Mrs Genevieve Vincent-Arnaud for their English revision of the manuscript. This research is partly supported by a grant (2005-2008) from the Agence Nationale de La Recherche for the "Monolix" project.

#### REFERENCES

- Cano, J., Kessler, M., and Salmern, D. (2006). Approximation of the posterior density for diffusion processes. *Statistics & Probability Letters* **76**, 39–44.
- Carlin, B. P. and Chib, S. (1995). Bayesian model choice via Markov Chain Monte Carlo methods. *Journal of the Royal Statistical Society. Series B. Statistical Methodology* **57**, 473–484.

- Carlin, B. P. and Louis, T. A. (2000). *Bayes and empirical Bayes methods for data analysis*, volume 69 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.
- Davidian, M. and Giltinan, D. (2003). Nonlinear models for repeated measurements: An overview and update. *Journal of Agricultural, Biological, and Environmental Statistics* **8**, 387–419.
- De la Cruz-Mesia, R. and Marshall, G. (2006). Non-linear random effects models with continuous time autoregressive errors: a Bayesian approach. *Statistics in Medicine* **25**, 1471–1484.
- Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B. Statistical Methodology* **39**, 1–38.
- Ditlevsen, S. and De Gaetano, A. (2005). Stochastic vs. deterministic uptake of dodecanedioic acid by isolated rat livers. *Bulletin of Mathematical Biology*, **67**, 547–561.
- Donnet, S. and Samson, A. (2008). Parametric inference for mixed models defined by stochastic differential equations. *ESAIM Probability & Statistics* **12**, 196–218.
- Hlaing, W., Prineas, R., Zhu, Y., and Leaverton, P. (2001). Body mass index growth in a sample of U.S. children: repeated measures data analysis of the Minneapolis children’s blood pressure study. *American Journal of Human Biology* **13**, 821–831.
- Hou, W., Garvan, C., Zhao, W., Behnke, M., Eyster, F., and Wu, R. (2005). A general model for detecting genetic determinants underlying longitudinal traits with unequally spaced measurements and nonstationary covariance structure. *Biostatistics* **6**, 420–433.
- Huggins, R. and Loesch, D. (1998). On the analysis of mixed longitudinal growth data. *Biometrics* **54**, 583–595.
- Jaffrézic, F. and Foulley, J. (2006). Modelling variances with random effects in non linear

- mixed models with an example in growth curve analysis. *XXIII International Biometric Conference, Montreal, 16-21 juillet* .
- Jaffrézic, F., Meza, C., Lavielle, M., and Foulley, J. (2006). Genetic analysis of growth curves using the SAEM algorithm. *Genetics Selection Evolution* **38**, 583–600.
- Meng, X. (1994). Posterior predictive p-values. *Annals of Statistics* **22**, 1142–1160.
- Mengersen, K. L. and Tweedie, R. L. (1996). Rates of convergence of the Hastings and Metropolis algorithms. *Annals of Statistics* **24**, 101–121.
- Oravecz, Z., Tuerlinckx, F., and Vandekerckhove, J. (in press). A hierarchical Ornstein-Uhlenbeck model for continuous repeated measurement data. *Psychometrika* .
- Overgaard, R., Jonsson, N., Tornøe, C., and Madsen, H. (2005). Non-linear mixed-effects models with stochastic differential equations: Implementation of an estimation algorithm. *Journal of Pharmacokinetics and Pharmacodynamics* **32**, 85–107.
- Pedersen, A. (1995). A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scandinavian Journal of Statistics* **22**, 55–71.
- Picchini, U., Ditlevsen, S., and De Gaetano, A. (2006). Modeling the euglycemic hyperinsulinemic clamp by stochastic differential equations. *Journal of Mathematical Biology* **53**, 771–796.
- Spyrides, M., Struchiner, C., Barbosa, M., and Kac, G. (2008). Effect of predominant breastfeeding duration on infant growth: a prospective study using nonlinear mixed effect models. *Journal of Pediatrics* **84**, 237–243.
- Tornøe, C., Agerso, H., Jonsson, E., Madsen, H., and Nielsen, H. (2004). Non-linear mixed-effects pharmacokinetic/pharmacodynamic modelling in nlme using differential equations. *Computer Methods and Programs in Biomedicine* **76**, 31–40.
- Zimmerman, D. and Núñez-Antón, V. (2001). Parametric modelling of growth curve data:

an overview. *Test* **10**, 1–73.

#### APPENDIX

For SDE mixed models (10) where the Euler-Maruyama scheme has to be used to implement the Gibbs algorithm, the approximate posterior distribution  $p^h(\theta|\mathbf{y})$  of the approximate model (10) is estimated instead of the original posterior distribution  $p(\theta|\mathbf{y})$ . The error induced by the Euler scheme on the posterior distributions may be controlled for restrictive volatility functions.

We either assume a volatility function proportionnal to  $\gamma$  ( $\Gamma(Z_t, \phi, \gamma^2) = \gamma g(\phi)$ , with  $g(\phi)$  a function of  $\phi$ ) or an affine function of  $Z_t$  ( $\Gamma(Z_t, \phi, \gamma^2) = \gamma(g_1(\phi)Z_t + g_2(\phi))$ , with  $g_1(\phi)$  and  $g_2(\phi)$  two functions of  $\phi$ ).

**PROPOSITION 1:** Let us assume that the drift function  $F : \mathbf{R} \times [t_0, T] \times \mathbf{R}^p \rightarrow \mathbf{R}$  is infinitely differentiable in the variable space and its partial derivatives of any order are uniformly bounded with respect to  $z$  and  $\phi$ .

Let  $p(\theta|y)$  and  $p^h(\theta|y)$  be the posterior distributions of the original (8) and the approximated (10) SDE mixed models, respectively. There exists a constant  $C(y)$  such that for all  $0 < h < H_0$ :

$$\|p^h(\theta|y) - p(\theta|y)\|_{TV} \leq C(y)h.$$

where  $\|\cdot\|_{TV}$  denotes the total variation distance.

*Proof.* Let  $p(\theta)$  denote the prior distribution. With the Bayes theorem, we have  $p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$ . Donnet and Samson (2008) prove that there exists a constant  $C_1$ , independent of  $\theta$  such that  $|p(y|\theta) - p^h(y|\theta)| \leq hC_1$ . Consequently  $|p(y) - p^h(y)| \leq C_1h$  and

$$\begin{aligned} |p(\theta|y) - p^h(\theta|y)| &\leq \frac{p(\theta)}{p(y)} \left| |p(y|\theta) - p^h(y|\theta)| + \frac{p^h(y|\theta)}{p^h(y)} |p^h(y) - p(y)| \right| \\ &\leq \frac{C_1h}{p(y)} p(\theta) \left| 1 + \frac{p^h(y|\theta)}{p(y)} \right| = C_2(y) h [p(\theta) + p^h(\theta|y)]. \end{aligned}$$

The result can be directly deduced:

$$\begin{aligned} \|p^h(\theta|y) - p(\theta|y)\|_{TV} &= \int |p^h(\theta|y) - p(\theta|y)| d\theta \\ &\leq C_2(y) h \int (p(\theta) + p^h(\theta|y)) d\theta \leq 2 C_2(y) h \end{aligned}$$

As a principal consequence of the proposition, the bias on the posterior mean is controlled: under hypotheses on the moments of  $p_\theta(\theta)$  and  $p^h(\theta|y)$ , there exists a constant  $C'_y$  such that  $|E_{\theta|y}[\theta] - E_{\theta|y}^h[\theta]| = |\int \theta p^h(\theta|y) d\theta - \int \theta p(\theta|y) d\theta| \leq C'_y h$  where  $E_{\theta|y}[\theta]$  and  $E_{\theta|y}^h[\theta]$  are the expectation under the posterior distributions  $p(\theta|y)$  and  $p^h(\theta|y)$ , respectively. A similar result can be obtained for the bias of the posterior mode.

*Received. Revised.*

*Accepted.*

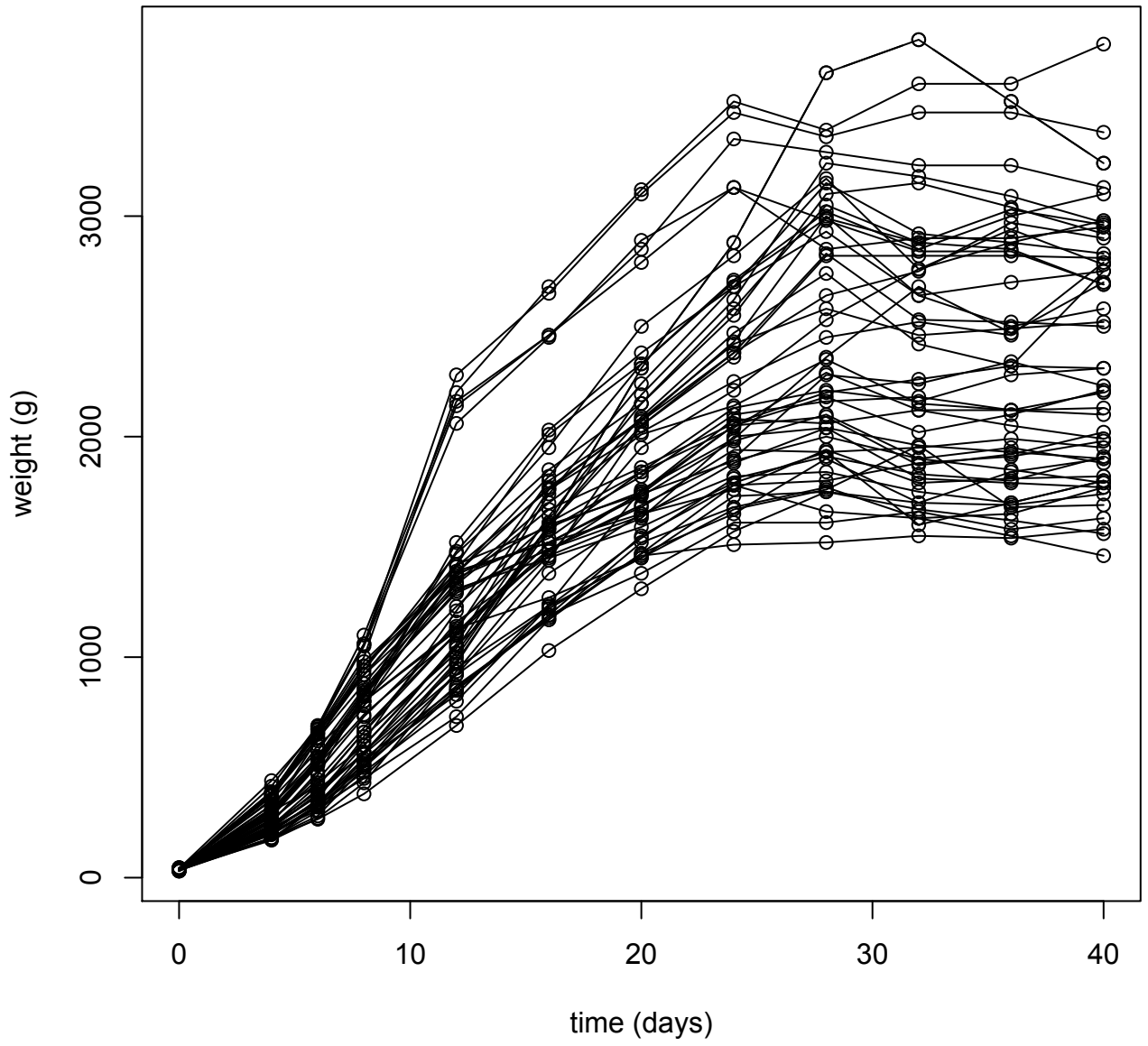
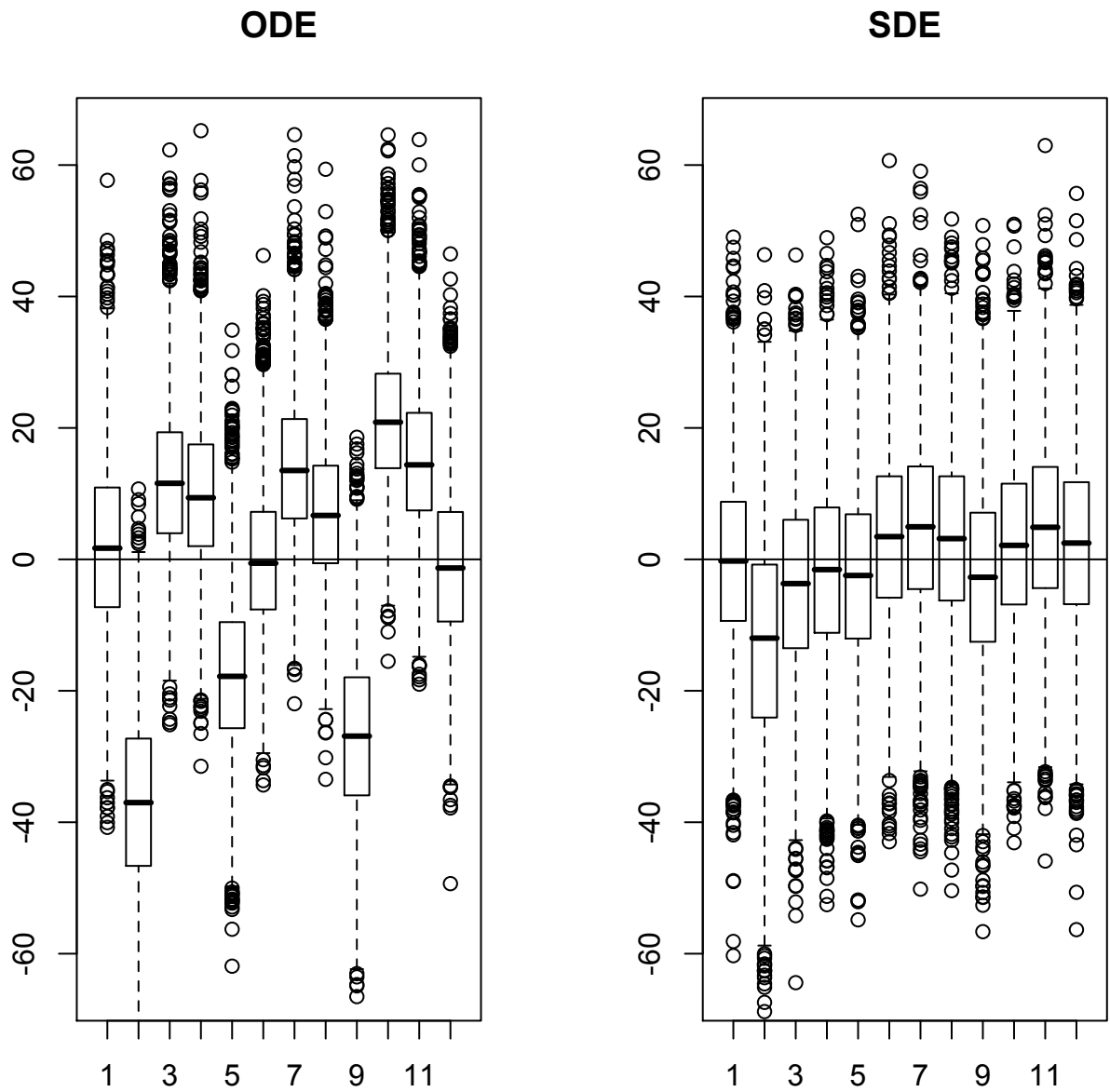
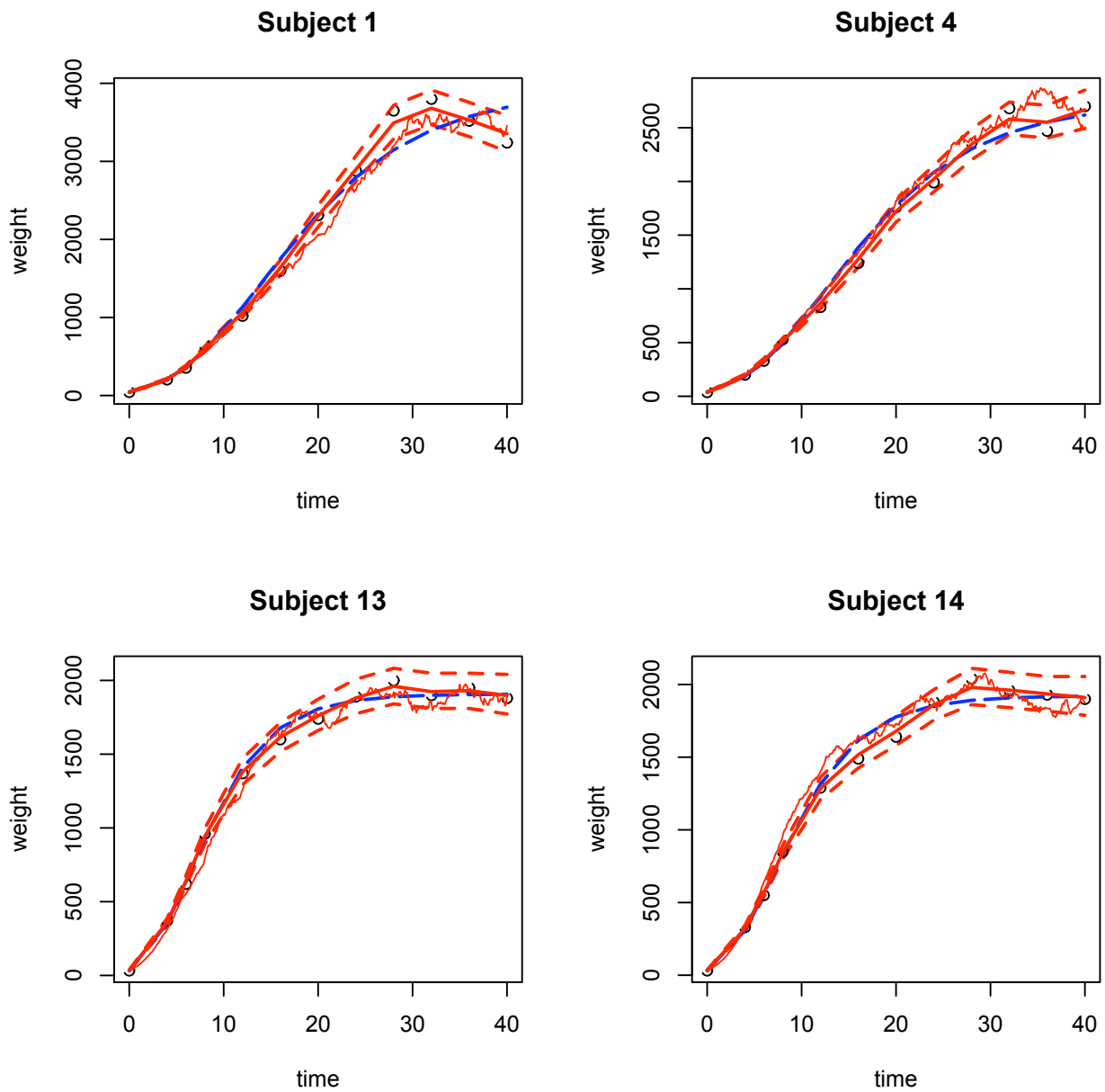


Figure 1. Growth curves of the 50 chickens.





**Figure 2.** Posterior predictive distributions for the ODE and SDE models on chicken growth data.



**Figure 3.** Observations (circles), predictions obtained with the ODE mixed model (long dashed line), mean SDE prediction (smooth solid line), 95% credibility interval obtained with the SDE mixed model (dotted line) and one SDE realization (solid line), for subjects 1, 4 13 and 14.

**Table 1**

*Relative bias (RMSE) (%) obtained from the ODE and the SDE mixed models when datasets are simulated with the ODE or the SDE mixed model.*

Simulation model	ODE		SDE	
Estimation model	ODE	SDE	ODE	SDE
$\mu_{lnA}$	-0.06 (0.49)	0.37 (0.72)	-2.22 (2.44)	-0.05 (1.36)
$\mu_B$	-0.21 (1.25)	0.55 (1.41)	-3.34 (3.88)	-0.02 (2.47)
$\mu_{lnC}$	-0.11 (1.35)	-0.19 (1.57)	1.93 (2.59)	-0.02 (1.77)
$\omega_{lnA}^{-2}$	31.69 (45.83)	65.06 (68.60)	-91.55 (91.58)	17.92 (35.27)
$\omega_B^{-2}$	3.32 (25.32)	2.56 (25.70)	-18.19 (35.88)	7.94 (23.38)
$\omega_{lnC}^{-2}$	22.19 (46.69)	36.59 (55.36)	20.69 (45.15)	22.54 (43.83)
$\gamma^{-2}$	-	5.03 (18.27)	-	8.28 (33.08)
$\sigma^{-2}$	1.18 (7.52)	7.45 (10.98)	-27.27 (27.92)	1.17 (8.87)

**Table 2**  
*Posterior distributions for the ODE and SDE models on chicken growth data.*

	ODE			SDE		
	mean	sd	median	mean	sd	median
$\log a$	7.77	0.03	7.77	7.75	0.04	7.75
$b$	4.17	0.03	4.17	4.15	0.04	4.15
$\log c$	2.75	0.03	2.75	2.78	0.03	2.78
$\Omega_{1,1}^{-1}$	117.30	31.74	113.60	93.89	20.45	92.10
$\Omega_{1,2}^{-1}$	-128.50	38.13	-123.70	-90.04	24.68	-88.46
$\Omega_{1,3}^{-1}$	-4.57	0.22	-4.02	-4.40	9.96	-4.13
$\Omega_{2,2}^{-1}$	172.10	49.95	165.40	146.10	37.45	141.70
$\Omega_{2,3}^{-1}$	22.64	15.87	21.16	23.86	15.55	22.01
$\Omega_{3,3}^{-1}$	36.68	8.14	35.89	38.04	10.07	36.57
$\sigma^{-2}$	225.5	14.74	225.10	630.22	83.44	623.92
$\gamma^{-2}$				9.07	1.02	9.00

**Table 3**  
*Posterior predictive p-values for the ODE and SDE mixed models on chicken growth data.*

time points	0	4	6	8	12	16	20	24	28	32	36	40
ODE model	0.55	0.00	0.86	0.80	0.08	0.48	0.90	0.73	0.02	0.99	0.91	0.46
SDE model	0.49	0.23	0.40	0.45	0.43	0.60	0.64	0.60	0.43	0.57	0.64	0.56