We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Simulation and Parametric Inference of a Mixed Effects Model with Stochastic Differential Equations Using the Fokker-Planck Equation Solution

Bakrim Fadwa, Hamid El Maroufy and Hassan Ait Mousse

Abstract

This chapter is concerned with estimation method for multidimensional and nonlinear dynamical models including stochastic differential equations containing random effects (random parameters). This type of model has proved useful for describing continuous random processes, for distinguishing intra- and interindividual variability as well as for accounting for uncertainty in the dynamic model itself. Pharmacokinetic/pharmacodynamic modeling often involves repeated measurements on a series of experimental units, and random effects are incorporated into the model to simulate the individual behavior in the entire population. Unfortunately, the estimation of this kind of models could involve some difficulties, because in most cases, the transition density of the diffusion process given the random effects is not available. In this work, we focus on the approximation of the transition density of a such process in a closed form in order to obtain parameter estimates in this kind of model, using the Fokker-Planck equation and the Risken approximation. In addition, the chapter discusses a simulation study using Markov Chain Monte Carlo simulation, to provide results of the proposed methodology and to illustrate an application of mixed effects models with SDEs in the epidemiology using the minimal model describing glucose-insulin kinetics.

Keywords: stochastic differential equations, mixed effects model, Fokker-Planck equation, transition density, maximum likelihood estimators, genetic algorithm, Markov Chain Monte Carlo simulation

1. Introduction

In pharmacokinetic/pharmacodynamic studies, repeated measurements are performed on a sample of individuals (units/subjects), and responses for all experimental subjects are assumed to be described by a common structural model. This model contains both fixed and random effects to distinguish between individuals in a population, leading to a mixed effects model in which fixed effects represent fixed

parameters for all individuals in the population and random effects account for individual differences. Moreover, mixed effects modeling has been shown to be useful in pharmacokinetic/pharmacodynamic studies; particularly for modeling total variation in and between individuals, the models used in pharmacokinetic/pharmacodynamic analysis are often presented by a system of deterministic differential equations (ordinary, partial, or delay). However, the pharmacological processes in reality are always exposed to incomprehensible effects that are difficult to model, and ignoring these effects can affect the parameter estimation results and their interpretations. So, the introduction of stochastic components to deterministic models is an important tool of analysis [1] and is more appropriate to model the intra-individual variations rather than ODEs. In addition, the extension of ODEs to the SDEs makes it possible to explain the differences between the observations and the predictions by two types of noise: dynamic noise that enters through the dynamics of the system and that can result from its random fluctuations or from the shortages of model and the measurement error which is added in the case of an indirectly observed process, which may be due to a test error or to the existence of a disturbance and represent the uncorrelated part of the residual variability. In the theory, there are rich and developed resources for mixed effects models whether deterministic [2–6] or stochastic and linear or nonlinear models (see many applications of stochastic NLME models in biomedical [7–9] and in pharmacokinetic [10–12] fields).

Parameter estimation in mixed effects models with SDEs, known by stochastic differential mixed effects models, is not an obvious procedure except in some cases simpler [14] because it is often difficult to write the likelihood function in its closed form. In this context, we propose a review on estimation methods of SDME models in [13, 15] and an example case that treats a generalized linear mixed models in [16] and also an example to approximate the likelihood function of an NLME with the likelihood of a linear mixed effects model in [17]. Moreover, to strengthen knowledge on estimation methods of SDME, we refer to [18, 19] that propose an example of stochastic mixed effects model with random effects log-normally distributed with a constant diffusion term.

In general, it is difficult to obtain an explicit likelihood function because the transition density of the stochastic process is often unknown or that the integral in the likelihood given the random effects cannot be computed analytically, and although the size of the random effects increases, the complexity of the problem increases also rapidly. Therefore, this requires a significant need for approximate methods to compute the transition density in an approximate closed form and for effective numerical integration methods to compute or approximate the integral in the likelihood function, using, for example, the Laplacian and Gaussian quadrature approximation [3, 8, 20] or other approaches [21, 22]. In the literature, several solutions have been proposed to approximate the transition density and have shown their effectiveness despite certain limitations. For example, the transition density could be approximated by the solution of the partial differential equations of Kolmogorov [23]; by the derivation of a Hermite expansion of closed form at the transition density [24–26] (this method has been reviewed and applied for many known stochastic processes for one-dimensional [8] and multidimensional [20] time-homogeneous SDME model); or by simulating the process to Monte-Carlo-integrate the transition density [27–29]. These techniques are very useful and can solve the problem, but unfortunately, they involve intense calculations which make the problem always complicated.

In this work, we focus on two fundamental issues concerning the implementation of SDEs in NLME models. The first is how the transition density of an SDME model can be approximated when it is not known, and the second is about

approximating methods of the likelihood function when the integral given the random effects has no analytic solution. Then, we propose an optimization algorithm to obtain maximum likelihood estimators in order to facilitate the estimation procedure for these models. Finally, the methodology is evaluated by simulation studies on the bidimensional Ornstein-Uhlenbeck model, and then it is implemented on the minimal model describing the glucose and insulin kinetics.

2. Theoretical tools

Consider an N-dimensional continuous and stochastic process Y_t in the state space $E \subset \mathbb{R}^N$ described by the general first-order nonlinear stochastic differential equations of the Itô type [30]:

$$dY_t^i = \mu(Y_t^i, t, \theta, b^i)dt + \Sigma(Y_t^i, \theta, b^i)dW_t^i, \quad Y_0^i = y_0^i, \quad i = 1, ..., M, \qquad (1)$$

where Y_t^i is defined as the solution of the SDME model Eq. (1) that exists under some conditions that we suppose satisfied [31–33] and represents the observation of individual i from M different experimental units, $(i = 1, .., M)$, at the moment $t \geq t_0^i$, and $Y_0^i = Y_{t_0}^i$ is the initial state of Y_t for each subject. The process $\left\{ (Y_t^i)_{t \geq 0}, i = 1..M \right\}$ is assumed to verify the same model structure Eq. (1) according to the individual parameters b^i; $\theta \in \Theta \subset \mathbb{R}^p$ is a p-dimensional fixed effects parameter which represents the same and common characteristics for all subjects, and $b^i \in B \subseteq \mathbb{R}^q$ are the q-dimensional individual random parameters assumed mutually independent, also called random effects because they are not the same for all individuals; they change between them according to a distribution of density $P_B(b^i|\Psi)$ depending on a population parameter Ψ; in the population approach, this parameter vector allows for a data from several subjects to be considered simultaneously. Each component b_l^i may follow a different distribution, $(l = 1, ..., q)$, and a standard choice for the joint density function $P_B(b^i|\Psi)$ of the vector b^i could be the Gaussian distribution; however, any other distributions may be considered continuous or discrete:

$$b^i \sim i.i.d \ N(\vartheta, \phi). \qquad (2)$$

The joint density function of the vector b^i is parameterized by a q-dimensional parameter $\vartheta \in \upsilon \subset \mathbb{R}^q$ and a $q \times q$-dimensional matrix $\phi \in \Phi \subset \mathbb{R}^{q \times q}$ representing the covariance matrix of b^i and specifying the parameters of the marginal distributions of the components b_l^i, $(1 \leq l \leq q)$; the components of ϕ and ϑ represent the population parameters Ψ. For Y_0^i, it is not necessarily known, and when its components are unknown, they must be considered as random effects since they change between individuals; but in some cases it can be known and assumed equal to a real constant. Also, we assume that the distribution of $Y^i(t)$ given (b^i, θ) and $Y^i(t') = y_{t'}, t' < t$, has a strictly positive density with regard to the Lebesgue measure on E:

$$y \rightarrow P_Y(y, t - t'|y_{t'}, b^i, \theta) > 0, \ y \in E. \qquad (3)$$

$W^i(t)$ are the standard Brownian motions, and they are assumed mutually independent with b^j for all $1 \leq i, j \leq M$. The functions $\mu(\cdot) : E \times \mathbb{R} \times \Theta \times B \rightarrow \mathbb{R}$ and

$\Sigma(\cdot) : E \times \Theta \times B \to \mathbb{R}^+$ represent, respectively, the drift and the diffusion term of the model and are assumed to have some properties sufficiently regular to ensure a unique solution to the model [33].

According to the model Eq. (1), the process Y is the same and follows the same form for each individual of the population, and the SDME model of type Itô expresses the dynamic of the individual i perturbed by the Brownian motion. So, the differences between the subjects are modeled on the one hand by the different realizations of the Brownian motion paths $\{W_t^i\}_{t \geq t_0^i}$ and, on the other hand, by the incorporation of the random parameters b^i in the model. Therefore, the introduction of parameters varying randomly between subjects allows the model Eq. (1) to explain the variability between individuals.

The goal is to estimate the vector of fixed parameters θ and the parameter vector Ψ characterizing the distribution of random parameters b^is, but the statistical inference for such models is a difficult issue, and the level of difficulty is not the same whether the transition density is explicit or not and whether the process is observed directly or with measurement noise; in this work we assume that the process was exactly observed and no observation noise was considered.

2.1 Maximum likelihood estimation in SDME model

The likelihood function of an SDME model is expressed as follows:

$$L(\theta, \Psi) = \prod_{i=1}^{M} P\left(\underline{y}^i | \theta, \Psi\right) = \prod_{i=1}^{M} \int P_{\underline{Y}}\left(\underline{y}^i | b^i, \theta\right) P_B(b^i | \Psi) db^i \qquad (4)$$

with

$$P_{\underline{Y}}\left(\underline{y}^i | b^i, \theta\right) = \prod_{j=1}^{n_i} P_Y\left(y_j^i, \Delta_j^i | y_{j-1}^i, b^i, \theta\right), \qquad (5)$$

where n_i is the number of observations for the subject i at discrete points of time $\left\{t_0^i, t_1^i, \ldots, t_{n_i}^i\right\}$, $i = 1, \ldots, M$ and $\Delta_j^i = t_j^i - t_{j-1}^i$, $j = 1, \ldots, n_i$. The conditional density $P_Y\left(y^i | \cdot\right)$ is equal to the product of the transition densities Eq. (5) for given random effects and θ, but the availability of the explicit transition density is the second constraint for the statistical issue of model Eq. (1) to obtain an exact likelihood function and exact estimators, since computing the transition density is not always obvious and requires approximation methods. However, there are some cases where the exact likelihood function is known, and the exact MLEs of θ are obtained (see references in the introduction). In fact, to compute the likelihood function in a closed form of an SDME model, we can encounter two types of problems that require approximate methods to overcome them: First, when the transition density $P_Y\left(y_j^i, \Delta_j^i | y_{j-1}^i, b^i, \theta\right)$ is known but the integral in Eq. (4) has no solution, in this case, the numerical methods of approximation of the integral are required. Or, second, when $P_Y\left(y_j^i, \Delta_j^i | y_{j-1}^i, b^i, \theta\right)$ cannot even be expressed explicitly and must also be approximated, see next paragraphs. Usually, in realistic examples, we have both an unknown transition density and an integral that is difficult to solve analytically. In theory, several methods for approximating transition densities and integrals have been proposed (see references cited in the introduction).

In this work, we propose an approximate method to estimate the transition density for a time-inhomogeneous NLME model with SDEs in a closed approximate form. So, we suggest to derive the transition density by solving the motion equations of the process using the Fokker-Planck equation, and we deal with the use of its solution given in [34]. Then, using the expression obtained, we get a closed form approximation of the likelihood function that we maximize to obtain the approximated MLEs $\hat{\theta}$ and $\hat{\Psi}$. The approximated transition density obtained from the proposed method is denoted by $P_Y^{(a)}\left(Y_j^i, \Delta_j^i | Y_{j-1}^i, b^i, \theta\right)$ that we substitute in Eq. (4) to obtain the following approximated likelihood function for the SDME model Eq. (1):

$$L^{(a)}(\theta, \Psi) = \prod_{i=1}^{M} \int \left(\prod_{j=1}^{n_i} P_Y^{(a)}\left(y_j^i, \Delta_j^i | y_{j-1}^i, b^i, \theta\right) \right) P_B(b^i | \Psi) db^i. \qquad (6)$$

2.1.1 Laplace approximation

For a multidimensional vector of random parameters, if the exact transition density or its closed form approximation is available, we can use the Laplace approximation method [16, 20, 35] to obtain an explicit expression of the likelihood function to maximize, despite the fact that the integral in Eq. (4) has no closed solution. So, for a q-dimensional random vector b^i, the likelihood function Eq. (4) can be approximated as:

$$logL(\theta, \Psi) \simeq \sum_{i=1}^{M} \left[logP_{\underline{Y}}\left(\underline{y}^i | \tilde{b}^i, \theta\right) + logP_B\left(\tilde{b}^i | \Psi\right) + \frac{q}{2} log(2\pi) - \frac{1}{2} log\left| -H\left(\tilde{b}^i | \theta, \Psi\right) \right| \right],$$

$$(7)$$

where

$$\tilde{b}^i = argmax_{b^i}\left(f(b^i | \theta, \Psi)\right) \quad and \quad f(b^i | \theta, \Psi) = logP_{\underline{Y}}\left(\underline{y}^i | \tilde{b}^i, \theta\right) + logP_B\left(\tilde{b}^i | \Psi\right) \quad (8)$$

and $| \cdot |$ denotes the determinant of the Hessian matrix $H(b^i | \theta, \Psi)$:

$$H\left(\tilde{b}^i | \theta, \Psi\right) = \frac{\partial^2 \left[logP_{\underline{Y}}\left(\underline{y}^i | \tilde{b}^i, \theta\right) + logP_B\left(\tilde{b}^i | \Psi\right) \right]}{\partial \tilde{b}^i \partial \tilde{b}^{iT}} \qquad (9)$$

2.2 Approximate transition density for multidimensional and nonlinear SDME model

Usually, a formal general solution of the stochastic differential equations as in Eq. (1) cannot be given, which makes the calculation of the transition density for this process more complicated. Moreover, this process has a lot of fluctuations that its exact position cannot be determined but can be known given a region by its probability density; with the FP equation, such a probability density can be determined. The FP equation is a differential equation for the distribution function describing a Brownian motion by which the probability density of the stochastic process can be calculated in a much simpler way by solving this equation. This motion equation is usually used for variables describing a macroscopic but small

system, where the fluctuations are important as for some cases in physics, e.g., the position and the speed of the Brownian motion of a small particle. However, it can be also used for the larger system where, in spite of their small fluctuations, the stochastic description remains necessary when the deterministic equations may not be stable for this type of system.

So, under some assumptions, the transition density $P_Y\left(y_j^i, \Delta_j^i | y_{j-1}^i, b^i, \theta\right)$ of the process Y in the SDME model Eq. (1) is the solution of the following functional partial differential Equation [23, 34]:

$$\frac{\partial P_Y\left(y_j^i, \Delta_j^i | y_{j-1}^i, b^i, \theta\right)}{\partial t} = L_{FP} P_Y\left(y_j^i, \Delta_j^i | y_{j-1}^i, b^i, \theta\right), \tag{10}$$

where L_{FP} is the FP operator and $L_{FP} = -\sum_{k=1}^{N} \frac{\partial}{\partial Y_k} \left[\mu_k\left(Y^i, t, \theta, b^i\right)\right] + \frac{1}{2}\sum_{k=1}^{N}\sum_{l=1}^{N} \frac{\partial^2}{\partial Y_k \partial Y_l} \left[\Sigma_{kl}\right]$ and μ_k is the k-th element of the drift vector and Σ_{kl} is the kl-th element of the diffusion matrix. We assume that the diffusion matrix is positive definite so that its inverse exists $Det(\Sigma) \neq 0$.

In [34], H. Risken deals in this book with the derivation of the FP equation and its solution methods with some of its applications especially for problems of Brownian motion. So, here, we propose to characterize the transition density of the N-dimensional process Y_t^i in Eq. (1) using the Risken approximation based on the formal solution of Eq. (10) proposed in [34], p: 4.109, denoted by $P_Y^{(a)}\left(y_j^i, \Delta_j^i | y_{j-1}^i, b^i, \theta\right)$:

$$P_Y^{(a)}\left(Y_j^i, \Delta_j^i | Y_{j-1}^i, b^i, \theta\right) = \left(2\sqrt{\Pi\Delta_j}\right)^{-N} [Det\Sigma]^{-\frac{1}{2}} * exp\left(-\frac{1}{4\Delta_j}\sum_{l=1}^{N}\sum_{k=1}^{N} [\Sigma^{-1}]_{lk}\right.$$
$$\left.\left[\left(Y_j^i\right)_l - \left(Y_{j-1}^i\right)_l - \mu_l\left(Y_{j-1}^i, t, \theta, b^i\right)\Delta_j^i\right]\left[\left(Y_j^i\right)_k - \left(Y_{j-1}^i\right)_k - \mu_k\left(Y_{j-1}^i, t_{j-1}, \theta, b^i\right)\Delta_j^i\right]\right) \tag{11}$$

To test the effectiveness of this approach, we will guide our statistical methodology by simulation studies in order to examine the flexibility of its application to deduce its advantages and disadvantages. So, we substitute $P_Y^{(a)}\left(Y_j^i, \Delta_j^i | Y_{j-1}^i, b^i, \theta\right)$ in Eq. (6), and by solving the integral with respect to the random effects density and maximizing the obtained likelihood function, we get the approximated estimators $\hat{\theta}$ and $\hat{\Psi}$.

2.2.1 Approximated estimators

For a nonlinear SDME model with Gaussian random effects using Eqs. (6), (7), and (11), we obtain the following approximated likelihood function:

$$logL^{(a)}(\theta, \Psi) \simeq \sum_{i=1}^{M}\left[-\frac{q}{2}log((2\pi)) - \frac{1}{2}log(det(\phi)) - \frac{n_i}{2}log(det(\Sigma)) + \left(\sum_{j=1}^{n_i}log\left(\left(2\sqrt{\Pi\Delta_j}\right)^{-N}\right) - \frac{1}{4\Delta_j}\right.\right.$$
$$\left.[\Sigma^{-1}]_{lk}\left[\left(Y_j^i\right)_l - \left(Y_{j-1}^i\right)_l - \mu_l\left(Y_{j-1}^i, t, \theta, \tilde{b}^i\right)\Delta_j^i\right]\left[\left(Y_j^i\right)_k - \left(Y_{j-1}^i\right)_k - \mu_k\left(Y_{j-1}^i, t_{j-1}, \theta, \tilde{b}^i\right)\Delta_j^i\right]\right)$$
$$\left. -\frac{1}{2}\left(\tilde{b}^i - \nu\right)'\phi^{-1}\left(\tilde{b}^i - \nu\right) + \frac{q}{2}log(2\pi) - \frac{1}{2}log\left(det\left(-H\left(\tilde{b}^i | \theta, \Psi\right)\right)\right)\right]. \tag{12}$$

The MLEs of (θ, Ψ) can be obtained using one of the optimization tools and numerical computation software, especially when it is complicated to compute the

gradients of the likelihood analytically. Here, we propose to use the genetic algorithm as an optimization tool to maximize the approximate likelihood function Eq. (12) using MATLAB software:

$$\left(\hat{\theta}, \hat{\Psi}\right) = argmin_{GA}\left(-log\left(L^{(a)}(\theta, \Psi)\right)\right). \tag{13}$$

2.3 Genetic algorithm

The genetic algorithm is a random search technique to look for an exact or approximated optimum points for optimization problems [36–38]. It is based on the concepts of natural genetic evolution which contains the following stages: the reproduction, the crossing, and the mutation of a constantly evolving population. It sets up the evolution of a random population of potential solutions of the N cardinal; then, the N simultaneous iterative trajectories interact with each other by following or imitating the biological evolution, for a convergence of certain elements of the population towards an optimal point of the fitness function.

The GA can search in multiple directions to explore all the search space by the possibility of jumping across them, so that the seeds spread uniformly over the whole search space. In this algorithm, we have a diversity of initial populations which gives the global optimum faster than other algorithms, where the initial value is very important and should be enough close to the global optimum. All of these features allows the GA to be regarded as a driving tool of evolution giving good results for optimization processes [37, 39]. In the literature, there were many works about the application of GA in optimizing problems specially for likelihood function [40, 41]. For the use of GA, we must first define some parameters of the algorithm: Population size N, EN, SR, CP, MP, fitness function, and convergence criteria. In the following we present the GA steps:

Steps of GA:

1. Generate initial population $\left\{\beta_1^{(0)}, \beta_2^{(0)}, ..., \beta_N^{(0)}\right\}, m = 0$ via an initialization strategy (random generation), in our case $\beta = (\theta, \Psi)$.

For $m = 0$:

2. Evaluate the Fitness function $log\left(-L^{(a)}\left(\theta^{(m)}, \Psi^{(m)}\right)\right)$.

3. While (convergence criteria are not satisfied):
Do:

4. Replacement step (by using SR and EN): At the SR rate, individuals with the worst results in step 2 of fitness function are replaced by new ones randomly generated, and a number EN of individuals is selected and accepted for the next step.

5. Selection operator by using roulette wheel method, based on the fact that the more the individual has a good result of fitness function, the more likely he will be selected.

6. Crossover operator by using CP and mutation operator by using MP: It is a mechanism of perturbation on the candidate individuals (parents) according to CP and MP to generate new groups of individuals, and we obtain a new $(m + 1)nd$ population $\left\{\beta_1^{(m+1)}, \beta_2^{(m+1)}, ..., \beta_N^{(m+1)}\right\}$.

Else:

7. Evolution stops; get GA output.

8. $m = m + 1$.

End For.

In the simulation studies paragraphs, the GA is implemented using the MATLAB software, where the function "ga," to generate the genetic algorithm, requires inputs that are chosen according to the constraints of each example (see the help window in MATLAB). Moreover, the algorithm parameters are chosen according to what is early used in the literature [39], $EN = 4$, $MP = 0.2$, $CP = 0.8$ *and* $SR = 1/3$, and the search spaces are around the confidence interval of the minimal model parameters (see [42] and references therein).

3. Simulation studies for SDME model

3.1 The two-dimensional Ornstein-Uhlenbeck process

To apply the proposed methodology and evaluate its effectiveness, we consider the two-dimensional OU process that is very useful in pharmacokinetic/pharmaco-dynamic studies and in biology [43], physics, engineering, finance, and neuroscience applications [14, 44]. Indeed, the choice of this process is due to the fact that it is one of the few known multivariate SDME models with known transition density. For this reason, we choose the OU process to evaluate the methodology presented above, and we perform a comparison study between the results obtained using the proposed transition density in Eq. (11) and those obtained using the exact density. The model is defined as follows:

$$dY^{(1)i}(t) = -\left(\beta_{11}b^i_{11}\left(Y^{(1)i}(t) - \alpha_1\right) + \beta_{12}b^i_{12}\left(Y^{(2)i}(t) - \alpha_2\right)\right)dt + \Sigma_{11}dW^{(1)i}(t), \ Y^i_0 = y^{(1)i}_0, i = 1, \ldots, M$$

$$dY^{(2)i}(t) = -\left(\beta_{21}b^i_{21}\left(Y^{(1)i}(t) - \alpha_1\right) + \beta_{22}b^i_{22}\left(Y^{(2)i}(t) - \alpha_2\right)\right)dt + \Sigma_{22}dW^{(2)i}(t), \ Y^i_0 = y^{(2)i}_0, i = 1, \ldots, M$$

$$(14)$$

With $Y^i(t) = \begin{pmatrix} Y^{(1)i}(t) \\ Y^{(2)i}(t) \end{pmatrix}$; $\beta = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}$; $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$; $\Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$;

$W^i(t) = \begin{pmatrix} W^{(1)i}(t) \\ W^{(2)i}(t) \end{pmatrix}$; $Y^i(0) = \begin{pmatrix} Y^{(1)i}(0) \\ Y^{(1)i}(0) \end{pmatrix}$ and $b^i = \begin{pmatrix} b^i_{11} & b^i_{12} \\ b^i_{21} & b^i_{22} \end{pmatrix}$ where $b^i_{ll'}$ *i.i.d* $\sim$

$\Gamma\left(r_{ll'}, r^{-1}_{ll'}\right)$, $l, l' = 1, 2$; $i = 1, \ldots, M$.

We rewrite the system in matrix notation under the Itô formula; we denote by $(\cdot)$ the elementwise multiplication:

$$dY^i(t) = \beta \cdot b^i\left(\alpha - Y^i(t)\right)dt + \Sigma dW^i(t), \ Y^i_0 = y^i_0, \ i = 1, \ldots, M. \quad (15)$$

Here, the random effects b^i are a matrix and not a vector in order to have a uniform dimension writing of the Eq. (15) and are assumed mutually independent and independent of Y^i_0 and W^i. The fixed parameter vector is $\theta = (\beta_{11}, \ \beta_{12}, \beta_{21}, \ \beta_{22}, \ \alpha_1, \ \alpha_2, \ \Sigma_{11}, \ \Sigma_{22})$, and the population parameter vector is $\Psi = (r_{11}, \ r_{12}, \ r_{21}, \ r_{22})$. The exact transition density of model Eq. (15) for a given realization of the random effects is a bivariate normal:

$$P_Y\left(Y^i_{t_j}, \Delta^i_j | Y^i_{t_{j-1}}, b^i, \theta\right) = (2\pi)^{-1}|\varsigma|^{\frac{-1}{2}} exp\left(\frac{-\left(Y^i_{t_j} - \mu\right)' \varsigma^{-1}\left(Y^i_{t_j} - \mu\right)}{2}\right), \quad (16)$$

with mean vector $\mu = \alpha + \left(Y^i_{t_{j-1}} - \alpha\right) exp\left(-(\beta.b^i)\Delta^i_j\right)$ and covariance matrix $\varsigma = \tau - \left(exp\left(-(\beta.b^i)\Delta^i_j\right) \tau \; exp\left(-(\beta.b^i)'\Delta^i_j\right)\right)$, where

$$\tau = \frac{1}{2tr(\beta.b^i)|\beta.b^i|}\left(|\beta.b^i|\Sigma\Sigma' + (\beta.b^i - tr(\beta.b^i).I)\Sigma\Sigma'(\beta.b^i - tr(\beta.b^i).I)'\right). \tag{17}$$

We assume that the matrices $\beta.b^i$ and Σ have full rank and the real parts of the eigenvalues of $\beta.b^i$ are positive definite in order that a stationary solution to Eq. (15) exists. Under these assumptions, we derive from Eqs. (14) and (11) the following approximated transition density of Y:

$$P^{(a)}_Y\left(Y^i_{t_j}, \Delta^i_j|Y^i_{t_{j-1}}, b^i, \theta\right) = (2\sqrt{\Pi\Delta_j})^{-2}(\Sigma_{11}\Sigma_{22})^{-\frac{1}{2}} * exp\left(-\frac{1}{4\Delta_j}\left(\Sigma^{-1}_{11}\right)\left[Y^{(1)i}_{t_j} - Y^{(1)i}_{t_{j-1}} + \left(\beta_{11}b^i_{11}\left(Y^{(1)i}_{t_{j-1}} - \alpha_1\right)\right.\right.\right.$$
$$\left.\left.\left.+ \beta_{12}b^i_{12}\left(Y^{(2)i}_{t_{j-1}} - \alpha_2\right)\right)\Delta^i_j\right]^2 + \left(\Sigma^{-1}_{22}\right)\left[Y^{(2)i}_{t_j} - Y^{(2)i}_{t_{j-1}} + \left(\beta_{21}b^i_{21}\left(Y^{(1)i}_{t_{j-1}} - \alpha_1\right) + \beta_{22}b^i_{22}\left(Y^{(2)i}_{t_{j-1}} - \alpha_2\right)\right)\Delta^i_j\right]^2\right). \tag{18}$$

In **Figure 1**, we illustrate in (a) the simulation of the OU process using the Euler scheme [45] with the following set of parameters:
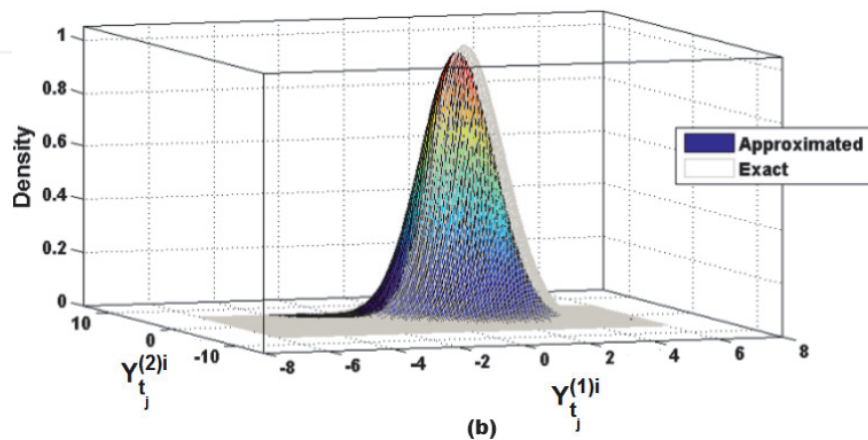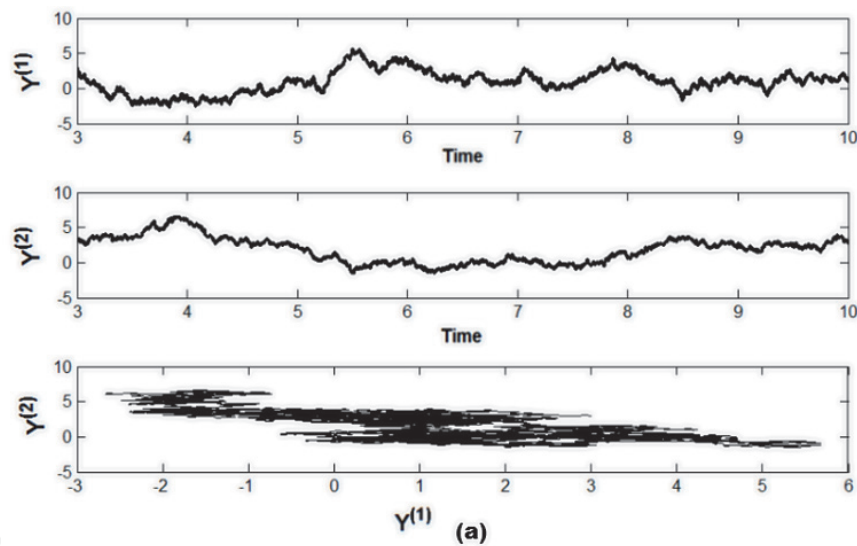
Figure 1.
A sample path of the OU process in the third graph of (a) for the given parameters set with the initial condition:
$Y_0 = (3,3)$ *and time interval* $[3,10]$; *and the transition density for a transition from* Y_j *to* Y_{j+1} *in (b).*

$(\beta_{11} = 2.8, \beta_{12} = 2.5, \beta_{21} = 1.8, \beta_{22} = 2, \alpha_1 = 0.8, \alpha_2 = 1.5, \Sigma_{11} = 0.3, \Sigma_{22} = 0.5, r_{11} = 45, r_{12} = 100, r_{21} = 100, \ r_{22} = 125)$, and a time step of $\Delta_j = 0.001$ and represent in (b) the graph of the two transition densities given by Eqs. (16) and (18) for Y_j to Y_{j+1} using the same set of parameters and time step.

3.1.1 Simulation results

In this simulation study, we generate 1000 artificial datasets of dimension $2 \times (n + 1) \times M$ from Eq. (15), where M is the number of subjects and n presents the number of repetitions of the experiment for each subject; then we estimate the parameters using the proposed approximated method, and we obtain 1000 sets of parameter estimates. The observations are obtained by linear interpolation from simulated trajectories using the Euler-Maruyama scheme with step size equal to 10^{-3}.

By plugging Eq. (16) in Eq. (4) and Eq. (18) in Eq. (6), we obtain a huge expression of the likelihood function but without a closed solution of integrals, so the exact estimators of θ and Ψ are unavailable. Therefore, in both cases, either using the exact or the approximated transition density, the Hessian matrix in Laplace approximation can be obtained analytically after a tedious calculation; then we apply the GA to obtain parameter estimates. But we cannot ignore the time consumed by this algorithm to get the results because of the long and complex expressions.

Table 1 shows that, for the given sample size, the results can be correctly identified using this estimated approach; even if some of parameters are overestimated or underestimated, the results remain acceptable because the results belong to the confidence interval. However, we believe that these results could be further proven by using other sample sizes and by adding alternative assumptions for the model that we did not consider in the methodology proposed in this chapter, which could further complicate the problem and be more time-consuming, as well as the present methodology suffers from some limitations. For the random parameters, the estimates can be provided using the optimization algorithm on Eq. (8) using the obtained estimate results of the parameters vector Ψ. Moreover, we conduct this simulation study using **Figure 2**, which shows that the empirical

True values				Mean and (Std) (M = 40, n = 10)			
β_{11}	β_{12}	β_{21}	β_{22}	$\hat{\beta}_{11}$	$\hat{\beta}_{12}$	$\hat{\beta}_{21}$	$\hat{\beta}_{22}$
2.8	2.5	1.8	2	3.10–3.25	2.48–2.56	1.72–1.65	2.11–2.37
				(0.164)–(0.283)	(0.015)–(0.283)	(0.095)–(0.189)	(0.153)–(0.255)
α_1	α_2	Σ_{11}	Σ_{22}	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\Sigma}_{11}$	$\hat{\Sigma}_{22}$
0.8	1.5	0.3	0.5	1.06–1.13	1.57–1.64	0.28–0.37	0.56–0.45
				(0.071)–(0.102)	(0.109)–(0.158)	(0.026)–(0.073)	(0.023)–(0.061)
r_{11}	r_{12}	r_{21}	r_{22}	$\hat{r}_{11}$	$\hat{r}_{12}$	$\hat{r}_{21}$	$\hat{r}_{22}$
45	100	100	25	44.75–52.43	100–112.75	102.35–89.64	24.72–31.02
				(0.523)–(9.372)	(01.166)–(28.113)	(01.04)–(22.05)	(2.297)–(11.20)

Table 1.
Ornstein-Uhlenbeck model: maximum likelihood estimates from 1000 simulations of model Eq. (14), using the exact and the approximated transition density.
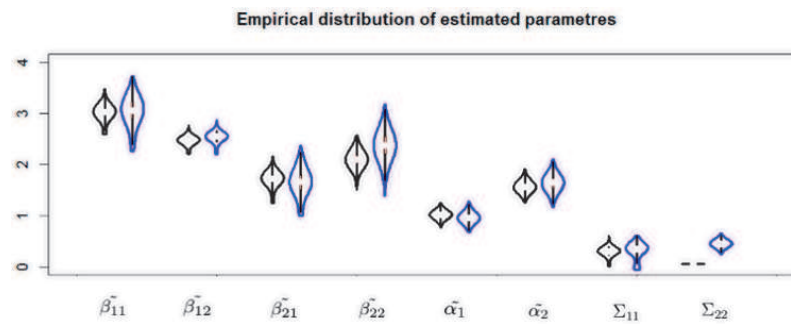
Figure 2.
Empirical distribution of estimates obtained using the exact and approximated transition density.

distribution of the most approximated estimators seems to be reasonably close to a normal distribution:

3.2 Stochastic minimal model

The minimal model describes the glucose-insulin kinetics and the dynamics of these processes illustrating the diabetes disease mechanisms. The diabetes is one of the most prevalent diseases in individuals and the nature and degree of its assignment changes from an individual to another and depends on certain individual characteristics, which implies that the concepts of stochastic modeling with random effects could be a good approach to modeling this disease. The Diabetes may be due to the insufficient insulin production (type 1 diabetes) or to the fact that the cells do not respond to the secreted insulin (type 2 diabetes) and the T2D patients tend to have substantially lower insulin sensitivity than healthy individuals, so that the T2D can be characterized by the level of insulin sensitivity for each individual. Therefore, to model the T2D, we observe how a person's body responds to insulin in the process of transporting glucose to tissues by measuring his insulin sensitivity. So, we present in this section the estimation of the minimal model which represents a powerful model describing the glucose-insulin kinetics for an individual's body in three differential equations, see the mathematical formulation of the model in [46–49]. So, it is already clear that the model will contain both fixed and random effects, because the study of diabetes disease takes into account the response of each individual according to his own parameters and other common parameters that describe the process of glucose-insulin for the entire population. See the description of the glucose-insulin kinetics in **Figure 3**.

From the mentioned literature and **Figure 3**, the glucose-insulin disposal can be represented, with respect to time, by the following nonlinear stochastic differential equations, perturbed by the stochastic terms $\sigma_1 dw_1(t)$, $\sigma_2 dw_2(t)$, and $\sigma_3 dw_3(t)$:

$$dG(t) = \left[-\left(p_1 + X(t)\right)G(t) + p_1 G_b \right]dt + \sigma_1 dw_1(t), \quad G(0) = G_0$$
$$dX(t) = \left[-p_2 X(t) + p_3 (I(t) - I_b) \right]dt + \sigma_2 dw_2(t), \quad X(0) = 0 \qquad (19)$$
$$dI(t) = \left[-n(I(t) - I_b) + \gamma(G(t) - h)t \right]dt + \sigma_3 dw_3(t), \quad I(0) = I_0,$$

where $G(t)$ and $I(t)$ are, respectively, the concentration of glucose and insulin at time t in the blood. G_b and I_b indicate the basal level of glucose and insulin concentration before the glucose injection, this injection will cause a disturbance of the concentrations according to the mechanism described in these equations, and these values are assumed known for each individual. And, G_0 and I_0 are the theoretical measure of the concentrations at glucose injection moment at the beginning of the experiment.
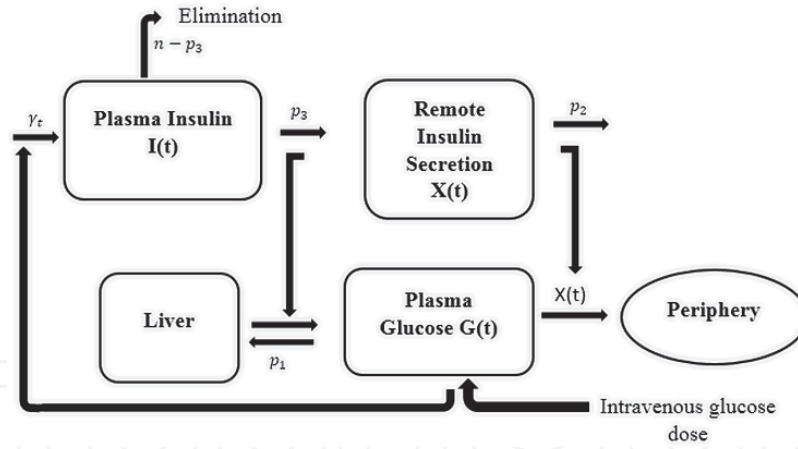
Figure 3.
At first, glucose and insulin concentrations in the blood are described by two sets of differential equations (see [50]); at a rate p_1, glucose leaves and enters the glucose space in proportion to the difference between the plasma glucose concentration $G(t)$ and the basal plasma concentration G_b that represent the known pre-injection glucose level for each individual. Therefore, the parameter p_1 represents the glucose's own ability to be eliminated in muscles, liver, and tissues independently of insulin which is called glucose efficiency and denoted by S_G. Then, the glucose disappears from the glucose space at a rate proportional to insulin concentration in the insulin compartment $X(t)$ which represents the dynamic of insulin response according to the two rates p_2 and p_3. These two parameters represent, respectively, the decreased glucose absorption capacity in tissues and its increased insulin dependency. For insulin secretion $I(t)$, it is secreted by the pancreas independently of the glucose concentration, and proportional to a rate n to its own level already in the body and to the glucose level deferred from a threshold h at a rate γ when $G(t)$ is above h, the insulin secretion does not only depend on the hyperglycemia level but also to the time spent since glucose injection.

The insulin sensitivity is defined by combining the two rates p_3 and p_2 as $S_I = p_3/p_2$, representing insulin's ability to increase the net glucose utilization [51].

Finally, the stochastic minimal model under the Itô sense re-parameterized by S_G and S_I can be rewritten as:

$$dY^i(t) = \begin{pmatrix} -\left(S_G^i + X(t)^i\right)G(t)^i + S_G^i G_b^i \\ -p_2\left(X(t)^i + S_I^i\left(I(t)^i - I_b^i\right)\right) \\ -n\left(I(t)^i - I_b^i\right) + \gamma\left(G(t)^i - h\right)t \end{pmatrix} dt + \Sigma dW(t); \quad Y_0^i = \begin{pmatrix} G_0^i \\ 0 \\ I_0^i \end{pmatrix},$$

(20)

where $Y^i(t) = \begin{pmatrix} G(t)^i \\ X(t)^i \\ I(t)^i \end{pmatrix}$ and Σ is the diagonal diffusion matrix which contains

the unknown constants σ_1, σ_2, and σ_3. The parameters $S_G^i, p_2, S_I^i, n, \gamma, h, G_0^i$ and I_0^i are unknown in the model and will be estimated. The parameters S_G^i, S_I^i, I_0^i and G_0^i are assumed random, because they represent individual parameters that change from an individual to another. Each person has its own insulin sensitivity S_I^i which allows to know if the cells of his body react correctly or not to the insulin and if the insulin produced by the pancreas is sufficient or not, which can make some people with T2D and others without diabetes. Also, for glucose effectiveness S_G^i, which represents the glucose's own ability to be eliminated independently of insulin, it is unique to each individual and changes from a person to another, as well as for the measurement of glucose and insulin concentration. For the rest of the parameters, we consider them fixed since they describe the common side of the glucose-insulin

process for the entire population (all individuals). So, we have the following random-dom effect vector $b^i = \left(S_G^i, \quad S_I^i, \quad I_0^i, \quad G_0^i \right)$, and we assume that:

$$S_G^i \sim N\left(\mu_{S_G}, \sigma_{S_G}\right), \quad S_I^i \sim N\left(\mu_{S_I}, \sigma_{S_I}\right), \quad I_0^i \sim N\left(\mu_{I_0}, \sigma_{I_0}\right), \quad G_0^i \sim N\left(\mu_{G_0}, \sigma_{G_0}\right). \quad (21)$$

Random effects are assumed to be independent with a multinormal joint density function, with the mean $\vartheta = \left(\mu_{S_G}, \mu_{S_I}, \mu_{I_0}, \mu_{G_0}\right)$ and the covariance matrix $\phi = diag(\sigma_{S_G}, \sigma_{S_I}, \sigma_{I_0}, \sigma_{G_0})$, so we have $\Psi = \left(\mu_{S_G}, \mu_{S_I}, \mu_{I_0}, \mu_{G_0}, \sigma_{S_G}, \sigma_{S_I}, \sigma_{I_0}, \sigma_{G_0}\right)$ and $\theta = (p2, n, \gamma, h, \sigma_1, \sigma_2, \sigma_3)$.

So, here we deal with a time-inhomogeneous NLME model with SDE describing the glucose-insulin kinetics (see [52] for the implementation of SDE time-inhomogeneous model) and [53] where the maximum likelihood estimation for a time-inhomogeneous stochastic differential model of glucose dynamics was treated. The measurements in the model Eq. (20) are assumed directly observed without measurement errors as well as in the theoretical approach presented above.

So, we wish to estimate (θ, Ψ) given the observations $\underline{y} = \left(\underline{y}^1, ..., \underline{y}^M\right)$ from model Eq. (20). By using the approximated transition density (Eq. (11)), we get the following approximated likelihood function for model Eq. (20):

$$L^{(a)}(\theta, \Psi) = \prod_{i=1}^{M} (2\pi)^{-2} (\sigma_{SG}\sigma_{SI}\sigma_{I0}\sigma_{G0})^{\frac{-1}{2}} \prod_{j=1}^{n_i} \left(\left(2\sqrt{\Pi\Delta_j}\right)^{-3} \right) [\sigma_1\sigma_2\sigma_3]^{-\frac{1}{2}}$$

$$\int_{R^4} exp \left(\sum_{j=1}^{n_i} \left[\frac{-1}{4\Delta_j} \left[\frac{1}{\sigma_1} \left(A_{1j}^i\right)^2 + \frac{1}{\sigma_2} \left(A_{2j}^i\right)^2 + \frac{1}{\sigma_3} \left(A_{3j}^i\right)^2 \right] \right] \right.$$

$$\left. - \frac{1}{2} \left(\sigma_{S_G}^{-1} \left(S_G^i - \mu_{S_G}\right)^2 - \sigma_{S_I}^{-1} \left(S_I^i - \mu_{S_I}\right)^2 - \sigma_{I_0}^{-1} \left(I_0^i - \mu_{I_0}\right)^2 - \sigma_{G_0}^{-1} \left(G_0^i - \mu_{G_0}\right)^2 \right) \right) dS_G^i dS_I^i dI_0^i dG_0^i$$

$$(22)$$

with

$$A_{lj}^i = \left[\left(Y_j^i\right)_l - \left(Y_{j-1}^i\right)_l - \mu_l\left(Y_{j-1}^i, t_{j-1}, \theta, b^i\right)\Delta_j^i \right]. \quad (23)$$

We have no closed form solution to this integral, so exact estimators of θ and Ψ are unavailable. So, we use the Laplace approximation method described in Section 2 to obtain a closed form approximation to the log-likelihood function $log\left(L^{(a)}(\theta, \Psi)\right)$ for the model Eq. (20); then by applying the GA, we get the approximate estimates $\hat{\theta}$ and $\hat{\Psi}$.

3.2.1 Simulation results

We start this study with an application on artificial data generated on the intra-venous glucose tolerance test principle (see [53] for a mathematical modeling of the test where glucose and insulin concentrations in plasma are subsequently sampled after an intravenous glucose injection). We generate 5000 sets of simulated artificial data of dimension $(n_i + 1) \times M$ from Eq. (20) using the Euler-Maruyama scheme [45] with a step size of 10^{-3} and a set of true parameters that are chosen according to [53, 54] representing the normal range of parameters values to simulate healthy subjects (without diabetes), with M being the number of units and n_i being the number of observations or repeated measurements collected on each unit i.

For each data from 5000 generated data sets, we estimate (θ, Ψ) by applying the proposed method. We first assume that the number of repeated measurements

True values				(M,m)	Mean and (Std)			
p_2	n	γ	h		$\hat{p}_2$	$\hat{n}$	$\hat{\gamma}$	$\hat{h}$
0.074	0.10	0.0007	90	(40,60)	0.0737 (0.0014)	0.100 (0.0013)	0.00073 (2.02×10^{-4})	89.08 (0.031)
				(40,10)	0.0784 (0.0059)	0.209 (0.0295)	0.00068 (2.54×10^{-4})	91.34 (0.712)
				(10,20)	0.0794 (0.0402)	0.156 (0.046)	0.00054 (0.0016)	62.11 (1.13)
σ_1	σ_2	σ_3	σ_{S_G}		$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\sigma}_3$	$\hat{\sigma}_{S_G}$
0.01	0.06	0.03	0.006	(40,60)	0.009 (0.0012)	0.0616 (0.0011)	0.0343 (0.0027)	0.0061 (0.00010)
				(40,10)	0.014 (0.0023)	0.0708 (0.0060)	0.0340 (0.0035)	0.0073 (0.00030)
				(10,20)	0.015 (0.0031)	0.0938 (1.0196)	0.0480 (0.0108)	0.0067 (0.0006)
σ_{S_I}		σ_{I_0}	σ_{G_0}	μ_{S_G}	$\hat{\sigma}_{S_I}$	$\hat{\sigma}_{I_0}$	$\hat{\sigma}_{G_0}$	$\hat{\mu}_{S_G}$
0.000025	46	50	0.03	(40,60)	0.000021 (3×10^{-6})	44.98 (1.25)	45.96 (2.11)	0.0315 (0.0007)
				(40,10)	0.000029 (4.4×10^{-6})	43.94 (1.75)	45.16 (2.53)	0.0349 (0.0036)
				(10,20)	0.000016 (0.8×10^{-5})	41.09 (2.82)	44.64 (3.07)	0.0178 (0.0051)
μ_{S_I}	μ_{I_0}	μ_{G_0}			$\hat{\mu}_{S_I}$	$\hat{\mu}_{I_0}$	$\hat{\mu}_{G_0}$	
0.0002	95	320		(40,60)	0.00021 (1.2×10^{-6})	94.20 (1.12)	321.2 (1.07)	
				(40,10)	0.00025 (1.6×10^{-6})	92.05 (2.25)	318.6 (3.11)	
				(10,20)	0.00037 (1.35×10^{-5})	122.51 (2.51)	281.5 (4.88)	

Table 2.
Approximated maximum likelihood estimates and standard deviation from simulations of model Eq. (20), using the approximated transition density Eq. (11) with large and small DATA.

collected on each unit is constant $n_i = m$ and $\Delta_j^i = \Delta$ for all $1 \leq i \leq M$ and $1 \leq j \leq n_i$; then we apply the GA after choosing the good algorithm parameters (N, EN, SR, CP, MP), and we get 5000 sets of parameters estimates. We repeat this for large and small data with different possibilities of repetition of the experiment $(m, M) = (40, 60); (40, 10)$ and $(10, 20)$; for each parameter the sample mean and standard deviation are reported in **Table 2**. The simulation study on small data is treated in order to see if the size of the sample influences on the results and if the number of measures taken over time has a negligible effect or not, in other words, to see if it is possible to select only the essential measuring moments without repeating the measurements several times to well simulate a subject. We treat this issue in relation to our model and its study context, since in epidemiology the availability of data (measurements) at any point of time is an interesting constraint. We note that quantities of G_b and I_b are randomly simulated from the normal range of healthy subjects.

In **Table 2**, we report the results obtained on large and small data by maximizing Eq. (22) using the GA. We notice that, for the case of the large data, the true values
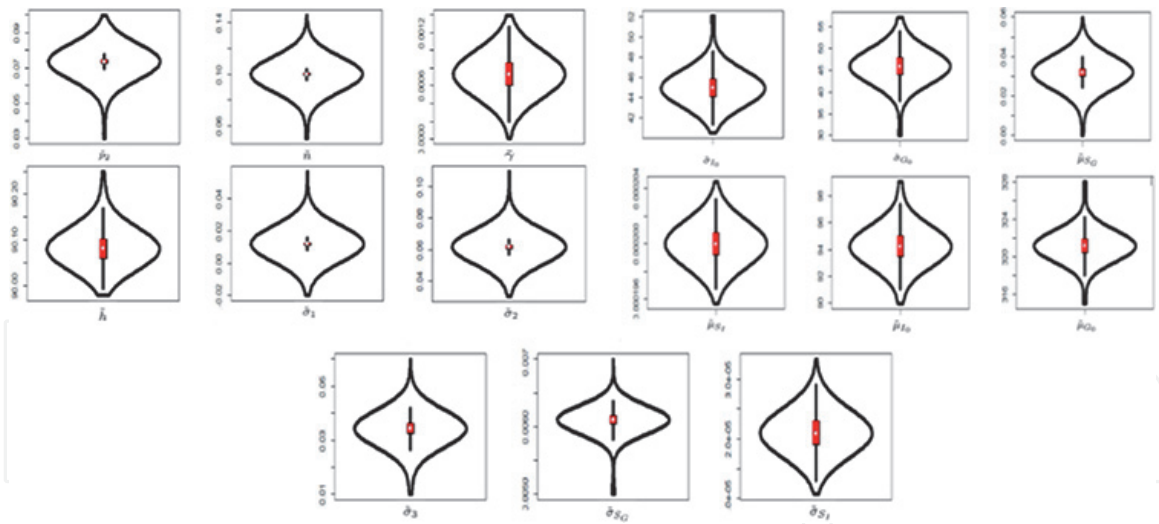
Figure 4.
SMM: empirical distribution of population parameter estimates obtained using (18) for (M, m) = (40,60).

of the parameters are well identified with the exception of some in which, in all the simulations and samples, the true value does not belong to the estimated confidence interval, such as σ_{S_I}, σ_{I_0}, and σ_{G_0}; nevertheless, the results are more satisfactory when the sample size M is large for all parameters. As mentioned before, we cannot ignore the limitations of this method, and the results could be improved by eliminating these limits and improving this approach, but for the given assumptions and tools, we can say that the results are still satisfactory even for a sample of $M = 10$ with 20 measures taken on each subject. So, we can conclude that we can rely on the results obtained from small samples with a small number of measurement repetitions of at least 20 observations. In addition, although we can use a relatively small number of measurements, we need to know how to choose the time points to perform the measurements in the blood, as this could, physiologically, affect selected observations and results. Thus, it is specified here that the essential task is to know how to choose the good moments of measurement after the injection, even in small numbers, chosen according to medical knowledge. **Figure 4** shows that, in the case of $(M; m) = (40; 60)$, the empirical distribution of the most approximated estimates seems to be reasonably close to a normal distribution.

Finally, from this simulation study where we have considered two SDME models, we can conclude that the parameters values of the models appear to be correctly identified using the proposed approach based on the Risken approximation to approximate the transition density of a stochastic process.

4. Conclusions

In this work, we have proposed a procedure to estimate the parameters of a mixed effects model containing stochastic differential equations, known by the SDME models, by proposing an approach to approximate its likelihood function to obtain the MLEs. This method has been evaluated by simulation studies on two SDME models in epidemiology: the two-dimensional Ornstein-Uhlenbeck process and stochastic minimal model. In fact, in models with SDEs instead of ODEs with random effects, the estimation of parameters is still not obvious even for one individual (one trajectory) because of the difficulties in deriving the transition densities, and difficulties become more interesting in using the population approach that treats the entire population simultaneously. The derivation of the exact density

is not always possible for a stochastic and continuous process in an SDME model, so the search for an approximation of this density is an important step and requires an expensive calculation. This task is very interesting to give good results with good statistical properties of the estimators obtained by maximizing the likelihood function. In this work, the proposed estimation method has been applied to multidimensional and nonlinear SDME models with many random parameters normally distributed that can be extended to random parameters of any distribution.

So, an approximation of the transition density $P_Y^{(a)}$ is obtained in a closed form using the Risken approximation for the formal solution of the Fokker-Planck equation proposed by Risken [34], and then the approximated likelihood is obtained using the Laplace approximation method and optimized using the genetic algorithm; these calculation procedures can be obtained using any numerical calculation software or with symbolic computing capabilities.

The classical inference of SDME models implies the problem of the numerical evaluation of the integral for the given random effects in the likelihood function, which becomes complicated especially when the model contains more than two random parameters. In the literature, several methods have been proposed and tested for the approximation of the integral (see references in the introduction) and the following examples: [8] which proposes the Gaussian quadrature method to solve the integrals for the case of an SDME models with a single random effect and [20] where the study was revised for a general case with several random parameters using the Laplace approximation to compute the integral in Eq. (4) and Eq. (6) numerically. For the mixed effects framework, see [3, 16, 55, 56]. In the case of using the Laplace method, as in this chapter, the calculation of the Hessian matrix can be done analytically when it is possible, as the examples in Section 2, or with the help of a symbolic calculus software or the automatic differentiation (AD) tools [57].

The results of simulation studies are satisfactory and can be obtained either by using moderate values for the number of experimental units M and of observations n taken for each experimental unit or by using a small sample size but with a number of measurements taken for each subject of 10 at least; this is relevant for applications where large sets of data are not available, such as biomedical applications where the mixed effect theory is widely applied.

The advantages of this approach, compared to those proposed in the literature for multidimensional SDME models with more than one random parameter [26], are that the computation of the approximate density is very easy and does not require a lot of time to calculate it or to program it in a software; the only task that can be time-consuming is in the optimization step to search for the optimum solution of the likelihood function, and also the proposed method is effective even with large data with a MATLAB program on a common PC (Intel Pentium IV 3.0 GHz with 512 MB of RAM). Nevertheless, the method suffers some limitations, for example, when the conditions to use Eq. (11) are not verified when, e.g., the inverse of the diffusion term does not exist and when, in certain cases, it is not obvious to derive the gradients and Hessians terms. Another limitation is that measurement errors are not considered in this work, and for a good stochastic version, it will be better to include noise on process increments and noise on observations that may be significant compared to system noise. These limitations may provide a perspective towards a more elaborate extension of the statistical study for SDME models, particularly in the field of epidemiology.

In conclusion, in this work, we proposed a method of parameter estimation for a mixed effects models with SDEs proposing an approximation method for the transition density in the case when it cannot be obtained in a closed form, with an

approximation method of integral computation for the case of an SDME model with several random parameters. We have treated two examples to illustrate the effectiveness of this approach using computer tools. Indeed, we believe that this type of models is very interesting and provides a powerful and flexible modeling approach for repeated measurement studies such as biological and pharmacokinetic/pharmacodynamic and financial studies, as they combine the good characteristics of mixed effects and stochastic increments in intra-subject dynamics for a good modeling of a phenomenon.

Acknowledgements

Conflict of interest

The author(s) declare no competing interests.

Abbreviations

SDEs	stochastic differential equations
MCMC	Markov Chain Monte Carlo
FP	Fokker-Planck equation
MLEs	maximum likelihood estimators
ODE	ordinary differential equations
NLME	nonlinear mixed effects
SDME	stochastic differential mixed effects
OU	Ornstein-Uhlenbeck
GA	genetic algorithm
EN	elite number
SR	selection rate
CP	crossover probability
MP	mutation probability
SMM	stochastic minimal model
T2D	type 2 diabetes

Author details

Bakrim Fadwa[1*†], Hamid El Maroufy[1†] and Hassan Ait Mousse[2]

1 Faculty of Science and Techniques, Department of Applied Mathematics, Sultan Moulay Slimane University, Beni-Mellal, Morocco

2 Faculty of Sciences and Techniques, Sultan Moulay Slimane University, Beni Mellal, Morocco

*Address all correspondence to: bakrim.fadwa@gmail.com

† These authors are contributed equally.

IntechOpen

References

[1] Allen E. Modeling with Itô Stochastic Differential Equations. Springer Science and Business Media, Springer; 2007

[2] Vonesh E, Chinchilli V. Linear and Nonlinear Models for the Analysis of Repeated Measurements. New York: Marcel Dekker; 1997

[3] McCulloch CE, Searle SR. Generalized, Linear and Mixed Models. New York: Wiley; 2001

[4] Kuhn E, Lavielle M. Maximum likelihood estimation in nonlinear mixed effects models. Computational Statistics & Data Analysis. 2005;**49**:1020-1038

[5] Guedj J, Thiébaut R, Commenges D. Maximum likelihood estimation in dynamical models of HIV. Biometrics. 2007;**63**:1198-1206

[6] Wang J. EM algorithms for nonlinear mixed effects models. Computational Statistics & Data Analysis. 2007;**51**: 3244-3256

[7] Picchini U, Ditlevsen S, De Gaetano A. Modeling the euglycemic hyperinsulinemic clamp by stochastic differential equations. Journal of Mathematical Biology. 2006;**53**:771-796

[8] Picchini U, De Gaetano A, Ditlevsen S. Stochastic differential mixed-effects models. Scandinavian Journal of Statistics. 2010;**37**:67-90

[9] Ditlevsen S, Yip K-P, Marsh D, Holstein-Rathlou N-H. Parameter estimation of the feedback gain in a stochastic model of renal hemodynamics: Differences between spontaneously hypertensive rats and Sprague-Dawley rats. American Journal of Physiology. Renal Physiology. 2007;**292**:607-616

[10] Sheiner L, Beal S. Evaluation of methods for estimating population pharmacokinetic parameters. I.

Michaelis-Menten model: Routine clinical pharmacokinetic data. Journal of Pharmacokinetics and Biopharmaceutics. 1980;**8**:553-571

[11] Sheiner L, Beal S. Evaluation of methods for estimating population pharmacokinetic parameters. II. Biexponential model and experimental pharmacokinetic data. Journal of Pharmacokinetics and Biopharmaceutics. 1981;**9**:635-651

[12] Donnet S, Samson A. A review on estimation of stochastic differential equations for pharmacokinetic/ pharmacodynamic models. Advanced Drug Delivery Reviews. 2013;**65**(7): 929-939

[13] Donnet S, Samson A. Parametric inference for mixed models defined by stochastic differential equations. European Series in Applied and Industrial Mathematics. 2008;**12**: 196-218

[14] Ditlevsen S, De Gaetano A. Mixed effects in stochastic differential equations models. REVSTAT. 2005;**3**:137-153

[15] Donnet S, Samson A. Parametric inference for mixed models defined by stochastic differential equations. European Series in Applied and Industrial Mathematics. 2008;**12**: 196-218

[16] Pinheiro J, Chao E. Efficient Laplacian and adaptive Gaussian quadrature algorithms for multilevel generalized linear mixed models. Journal of Computational and Graphical Statistics. 2006;**15**:58-81

[17] Lindstrom M, Bates D. Nonlinear mixed-effects models for repeated measures data. Biometrics. 1990;**46**: 673-687

[18] Overgaard R, Jonsson N, Tornoe C, Madsen H. Nonlinear mixed-effects

models with stochastic differential equations: Implementation of an estimation algorithm. Journal of Pharmacokinetics and Pharmacodynamics. 2005;**32**:85-107

[19] Tornøe C, Overgaard R, Agerso H, Nielsen HA, Madsen H, Jonsson EN. Stochastic differential equations in NONMEM: Implementation, application, and comparison with ordinary differential equations. Pharmaceutical Research. 2005;**22**: 1247-1258

[20] Picchini U, Ditlevsen S. Practical estimation of high dimensional stochastic differential mixed-effects models. Computational Statistics & Data Analysis. 2011;**55**: 1426-1444

[21] Fröberg C. Numerical Mathematics: Theory and Computer Applications. Redwood City, CA: Benjamin-Cummings Publishing Company; 1985

[22] Krommer AR, Ueberhuber CW. Computational integration. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1998

[23] Lo A. Maximum likelihood estimation of generalized Ito processes with discretely-sample data. Econometric Theory. 1988;**4**:231-247

[24] Aït-Sahalia Y. Comment on numerical techniques for maximum likelihood estimation of continuous-time diffusion processes by G. Durham and A. Gallant. Journal of Business & Economic Statistics. 2002;**20**:317-321

[25] Aït-Sahalia Y. Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. Econometrica. 2002b;**70**:223-262

[26] Aït-Sahalia Y. Closed-form likelihood expansion for multivariate

diffusions. Ann. Statist. 2008;**36**: 906-937

[27] Nicolau J. A new technique for simulating the likelihood of stochastic differential equations. The Econometrics Journal. 2002;**5**:91-103

[28] Hurn A, Lindsay K, Martin V. On the efficacy of simulated maximum likelihood for estimating the parameters of stochastic differential equations. Journal of Time Series Analysis. 2003; **24**:45-63

[29] Ripley B. Stochastic Simulation (Paperback Edition, 2006). New York: Wiley; 1987

[30] Ramanathan M. An application of ito's lemma in population pharmacokinetics and pharmacodynamics. Pharmaceutical Research. 1999;**16**:584-586

[31] Andersson H, Britton T. Stochastic Epidemic Models and their Statistical Analysis. New York: Springer Lecture Notes in Statistics; 2000

[32] Becker NG. On a general stochastic epidemic model. Theoretical Population Biology. 1976;**11**:23-26

[33] Øksendal B. Stochastic Differential Equations: An Introduction with Applications. 6th ed. Berlin-Heidelberg: Springer-Verlag; 2007

[34] Risken H. The Fokker-Plank Equation, Method of Solution and Application. 2nd ed. New York: Springer-Verlag; 1989

[35] Shun Z, McCullagh P. Laplace approximation of high dimensional integrals. Journal of the Royal Statistical Society B. 1995;**57**(4):749-760

[36] Goldberg DE. Genetic Algorithms in Search. Addison-Wesley Reading Mass: Optimization and Machine Learning; 1989

[37] Mickalewicz Z. Genetic Algorithms + Data Structure = Evolution Programs. Artificial Intelligence, Berlin: Springer-Verlag; 1992

[38] Sivanandam SN, Deepa SN. Introduction to Genetic Algorithms. Berlin: Springer-Verlag; 2008

[39] Jong D, Kenneth A. Evolutionary Computation: A Unified Approach. Cambridge, MA, United States: MIT Press; 2006

[40] Yalçinkaya A, Şenoglŭ B, Yolcu U. Maximum likelihood estimation for the parameters of skew normal distribution using genetic algorithm. Swarm and Evolutionary Computation. 2017;**38**: 127-138. DOI: 10.1016/j.swevo.2017. 07.007

[41] Petrovski A, Wilson A, McCall J. Statistical Analysis of Genetic Algorithms and Inference about Optimal Factors. SCMS Technical Report 1998/2. Aberdeen, Scotland: The Robert Gordon University; 1998

[42] Andersen KE, Malene H. A population-based Bayesian approach to the minimal model of glucose and insulin homeostasis. Statistics in Medicine. 2005;**24**:2381-2400

[43] Favetto B, Samson A. Parameter estimation for a bidimensional partially observed OrnsteinUhlenbeck process with biological application. Scandinavian Journal of Statistics. 2010; **37**:200-220

[44] Picchini U, Ditlevsen S, De Gaetano A. Maximum likelihood estimation of a time-inhomogeneous stochastic differential model of glucose dynamics. Mathematical Medicine and Biology: A Journal of the IMA. 2008; **25**(2):141-155

[45] Kloeden P, Platen E. Numerical Solution of Stochastic Differential Equations. Berlin-Heidelberg: Springer-Verlag; 1992

[46] Bergman RN, Bortolan G, Cobelli C, Toffolo G. Identification of a minimal model of glucose disappearance for estimating insulin sensitivity. Proceedings of IFAC Symposium Identification and System Parameter Estimation. 1979;**2**:883-890

[47] Bergman RN, Phillips LS, Cobelli C. Physiologic evaluation of factors controlling glucose tolerance in man. Journal of Clinical Investigation. 1981; **68**:1456-1467

[48] Toffolo G, De Grandi F, Cobelli C. Estimation of beta-cell sensitivity from intravenous glucose tolerance test C-peptide data. Knowledge of the kinetics avoids errors in modeling the secretion. Diabetes. 1995;**44**:845-854

[49] Cobelli C, Man CD, Sparacino G, Magni L, De Nicolao G, Boris Kovatchev P. Diabetes: Models, signals, and control. IEEE Reviews in Biomedical Engineering. 2009;**2**:54-96

[50] Pacini G, Bergman RN. MINMOD: A computer program to calculate insulin sensitivity and pancreatic responsivity from the frequently sampled intravenous glucose tolerance test. Computer Methods and Programs in Biomedicine. 1986;**23**:113-122

[51] Bergman RN, Ider YZ, Bowden CR, Cobelli C. Quantitative estimation of insulin sensitivity. American Journal of Physiology-Endocrinology And Metabolism. 1979;**236**(6):E667

[52] Egorov A, Li H, Xu Y. Maximum likelihood estimation of time-inhomogeneous diffusions. Journal of Econometrics. 2003;**114**:107-139

[53] Gaetano A, Arino O. Mathematical modelling of the intravenous glucose tolerance test. Journal of Mathematical Biology. 2000;**40**:136-168

[54] Makroglou A, Li J, Kuang Y. Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: An overview. Applied Numerical Mathematics. 2006; **56**(3):559-573

[55] Davidian M, Giltinan D. Nonlinear models for repeated measurements: An overview and update. Journal of Agricultural, Biological, and Environmental Statistics. 2003;**8**: 387-419

[56] Pinheiro J, Bates D. Approximations of the log-likelihood function in the nonlinear mixedeffects model. Journal of Computational and Graphical Statistics. 1995;**4**:12-35

[57] Griewank A. Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation. Philadelphia, PA: SIAM; 2000