# Parametric inference for PK models defined by stochastic differential equations

Erhardt, Elvira M. (1); Leclercq Samson, Adeline (2); Bibbona, Enrico (1)

(1) Politecnico di Torino, Turin, Italy; (2) Université Grenoble Alpes, Grenoble, France

## INTRODUCTION

Pharmacokinetic (PK) studies usually consist in repeated measurements of drug concentration obtained from a population of subjects. $\Rightarrow$ Statistical parametric approach commonly used to analyse this longitudinal data:
**Mixed Effects Model** Pinheiro and Bates (2000): $y_{ij}$ is the observation of subject $i$ at time $t_{ij}$. Same regression function $f$ is used for all the subjects, but regression parameters $\phi_i$ differ between the individuals $i$:

$$y_{ij} = f(t_{ij}; \phi_i) + \epsilon_{ij}, \quad , i = 1, \dots N, \quad j = 1, \dots J.$$
$$\phi_i \sim_{\text{i.i.d}} \pi(\cdot, \beta), \quad \epsilon_{ij} \sim_{\text{i.i.d}} \mathcal{N}(0, \sigma^2)$$

the regression function $f$ **is solution of a system of ordinary differential equations (ODE)**.

- Compartment models are usually deterministic
- Real pharmacological processes are always exposed to influences that are not fully understood or not feasible to model explicitly
- PK processes have both deterministic and stochastic components: drug concentrations follow determinable trends but exact concentration at any given time is not completely determined.
- Ignoring these phenomena in the modelling may affect the estimation of PK parameters and the derived conclusions.
  $\Rightarrow$ *stochastic differential equations (SDEs)* as natural extension of deterministic ODEs

## SDE and BROWNIAN MOTION

- SDE: differential equation in which one or more of the terms is a *stochastic process* (=collection of random variables)

$$dz(t, \phi_i) = F(z, t, \phi_i)dt + \gamma dB_t$$

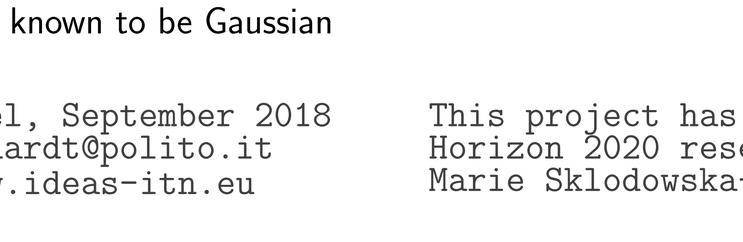
- Its solution, the *diffusion process* $z$, is again a stochastic process
- Typically, SDEs contain a variable which represents random white noise calculated as the "derivative" (in *Itô stochastic calculus* sense) of the *Brownian motion* $B(t)$ $\rightarrow$ requires its own rules of calculus.

### One-compartment SDE model

$$\begin{aligned} y_{ij} &= z(t_{ij}; \phi_i) + \epsilon_{ij}, \\ \phi_i &\sim_{\text{i.i.d}} \mathcal{N}(\mu, \Omega), \qquad \epsilon_{ij} \sim_{\text{i.i.d}} \mathcal{N}(0, \sigma^2), \\ dz(t, \phi_i) &= \left( \frac{\text{Dose } k_e k_a}{Cl} \cdot^{-k_a t} - k_e z(t, \phi) \right) dt + \gamma dB_t, \\ z(t_0) &= 0. \end{aligned}$$

Note: For $\gamma = 0$, the SDE reduces to the ODE.
**Three**, instead of only two, **fundamentally different types of noise** are distinguished:

1. *inter-subject variability* $\omega^2$: variance of the individual parameters $\phi_i$
2. *measurement noise* $\sigma^2$: uncorrelated part of the residual variability associated with dosing and sampling errors
3. *dynamic noise* $\gamma^2$: random fluctuations around the corresponding theoretical dynamic model.




**Main objective of this work:**
Reproduce and (possibly improve) the MCMC and SAEM based procedures of Donnet and Samson (2008) maximum likelihood method to estimate the parameter vector $\theta = ((\mu, \Omega), \gamma^2, \sigma^2)$ for these mixed models and applying this method to a new dataset with unsatisfactory fit by deterministic ODEs.
**Challenges:**

1. Diffusion is nonlinear with respect to the individual parameters $\phi_i$ (are appearing as exponent) $\rightarrow$ likelihood of the corresponding nonlinear mixed model has no analytical form.
2. Transition density of $z$ (diffusion process) has generally no closed form.

Note: In the PK case, SDE is linear in $z$ and thus, explicitly solvable $\rightarrow$ mean and variance can be calculated and distribution of the diffusion $z$ is analytically known to be Gaussian

## EULER-MARUYAMA (EM) APPROXIMATION OF DIFFUSION MODEL

EM is a scheme to approximate the trajectory of the SDE solution. Here, it is applied to replace the (generally) untractable likelihood of the SDE by its closed form approximation of the EM. In PK-setting, time intervals between observations too large to approximate well $\rightarrow$ auxiliary latent data points $\tau$ between every pair of observations $t_J$ are introduced. By enriching the observed data $y$ with the missing data $(w, \phi)$, and by the Markov property of the diffusion process, the complete data likelihood is *analytically known*:

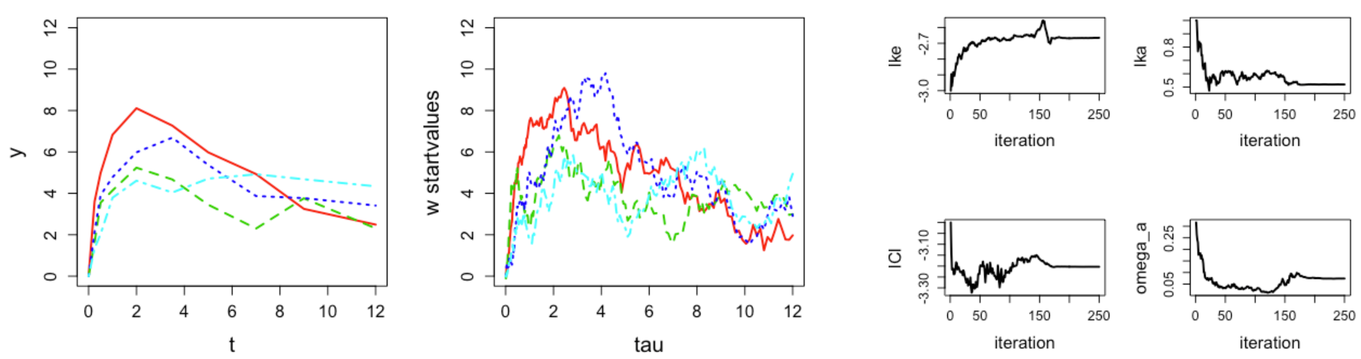
### Approximation $w$ of the diffusion process

$$\begin{aligned} y_{ij} &= w_{i,n_J} + \epsilon_{ij}, \\ \phi_i &\sim_{\text{i.i.d}} \mathcal{N}(\mu, \Omega), \qquad \epsilon_{ij} \sim_{\text{i.i.d}} \mathcal{N}(0, \sigma^2), \\ w_{i,n} &= w_{i,n-1} + h_n F(w_{i,n-1}, \tau_{n-1}, \phi_i) + \gamma \sqrt{h_n} \xi_n, \\ h_n &= \tau_n - \tau_{n-1}, \\ w(t_0, \phi_i) &= z(t_0, \phi_i), \\ \xi_n &\sim_{\text{i.i.d}} \mathcal{N}(0, 1). \end{aligned}$$

## COMBINATION OF MCMC AND SAEM

Proposal of candidates for $\phi$ and $w$ at observation times through random walk Metropolis-Hastings algorithm (a MCMC method) and for the $w$ at latent measurement times using Brownian bridges. Having found candidate values for $w$ and $\phi$ which increase the likelihood, the estimates for $\theta$ are updated using the *Stochastic Approximation Expectation Maximization (SAEM)* algorithm.

## RESULTING ESTIMATIONS ON SIMULATED DATA



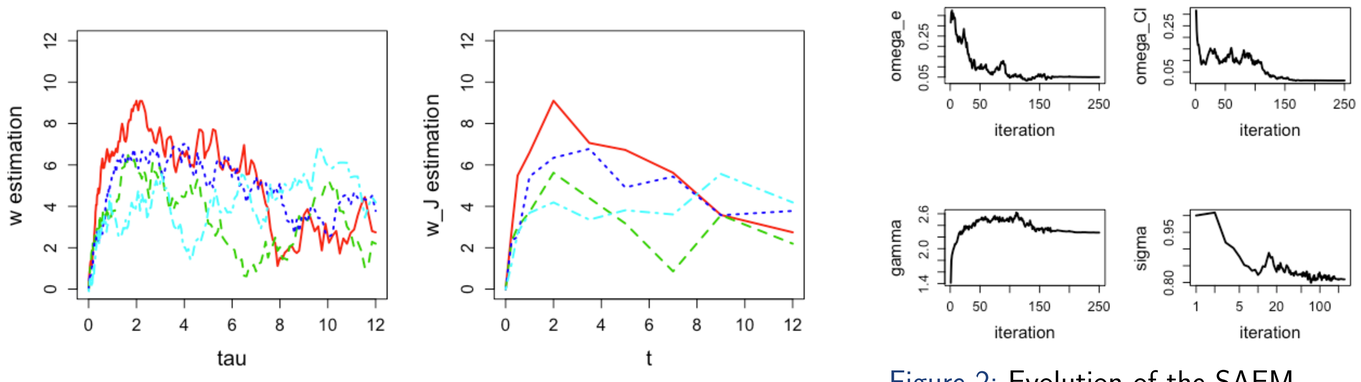




Figure 1: Individual concentrations for the pharmacokinetics of the simulated data, 4 out of 36 patients.

Figure 2: Evolution of the SAEM parameter estimates function of the iteration number.

| **Parameter** | $\log(k_e)$ | $\log(k_a)$ | $\log(Cl)$ | $\omega_e$ | $\omega_a$ | $\omega_{Cl}$ | $\gamma$ | $\sigma$ |
|---|---|---|---|---|---|---|---|---|
| true value | -2.52 | 0.40 | -3.22 | 0.10 | 0.10 | 0.10 | 0.45 | 0.32 |
| initial value | -3.00 | 1.00 | -3.00 | 0.32 | 0.32 | 0.32 | 1.41 | 1.00 |
| estimations (mean) | -2.76 | 0.59 | -3.26 | 0.02 | 0.06 | 0.02 | 2.28 | 0.78 |
| estimations (st. dev.) | 0.13 | 0.14 | 0.07 | 0.02 | 0.03 | 0.01 | 0.05 | 0.04 |

Table 1: Mean values and standard deviations of 5 PK studies simulated based on the given true values as in Donnet and Samson (2008).

## DISCUSSION

Our reproduction of the combined estimation method proposed by Donnet and Samson (2008) estimates the PK parameters correctly. However, it needs further improvement since the estimation of the different types of noise is not satisfactory yet and the overall performance of our code is greatly dependent on the chosen starting values. Possibilities are the replacement of Brownian bridges by *diffusion bridges*, to create dependency on the process' distribution, or a *Particle MCMC* algorithm which is expected to separate the diffusion noise better.

## REFERENCES

**Delyon, B., M. Lavielle and Moulines, E. (1999).** Convergence of a stochastic approximation version of the EM algorithm. *Ann. Statist., 27; 94–128*
**Dempster, A. P., Laird, N.M. and Rubin, D.B. (1977).** Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B, 39; 1-38*
**Donnet, S., Samson, A. (2013).** A review on estimation of stochastic differential equations for pharmacokinetic/pharmacodynamic model. *Advanced Drug Delivery Reviews, Elsevier, pp.1.*
**Donnet, S. Samson, A. (2008)** Parametric inference for mixed models defined by stochastic differential equations
**Pinheiro, J.C. and Bates, D.M. (2000)** Mixed-Effects Models in S and S-PLUS. *Springer.*
**Kuhn, E. and Lavielle, M., (2004)** Coupling a stochastic approximation version of EM with a MCMC procedure. *ESAIM: PS 8; 115–131.*

EFSPI Basel, September 2018
elvira.erhardt@polito.it
http://www.ideas-itn.eu

This project has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 633567.