SAEM methods for statistical PDE parameters estimation and application to biology.

Emmanuel Grenier, Celine Helbert, Violaine Louvet, Adeline Samson, Paul Vigneaux[•]

[•] **ENS de Lyon & Inria NUMED Team**

CIMPA School, Mauritius. Dec. 5-16, 2016

Outline

1. Population approach and PDEs
2. SAEM + Kriging

Introduction

Context : we have a population of “individuals” which evolve in time and discrete values (in time) of these evolutions

Objectives : build a **model** for these dynamics

e.g. one or several ODE’s, SDE’s, PDE’s, etc

containing **parameters** allowing to adjust

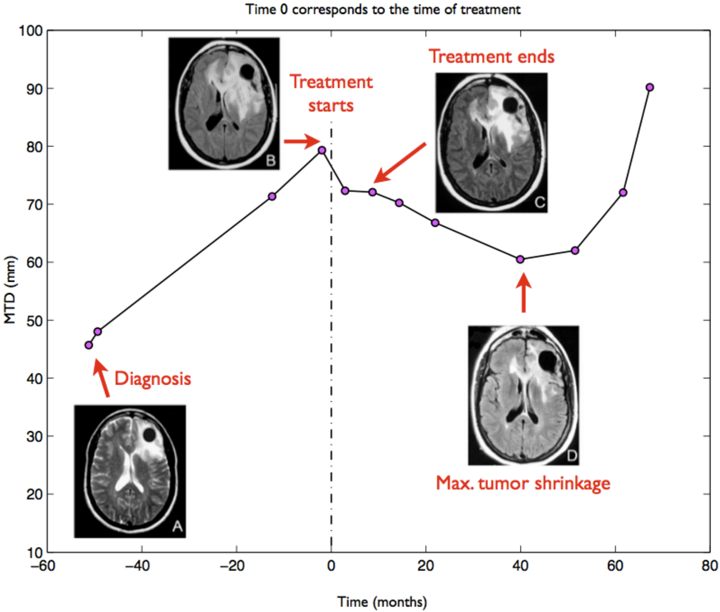
to all individuals’ dynamics

and recover these parameters from time measurements.

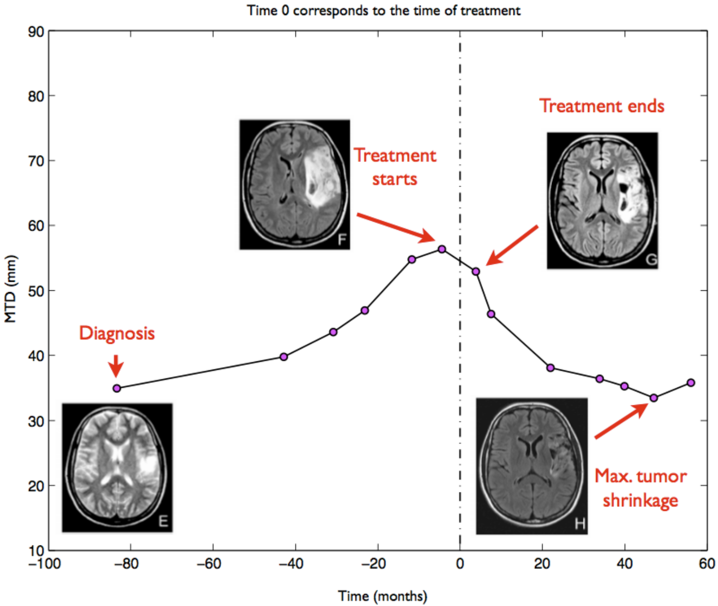
Some existing works

- Inverse problem approaches : huge literature.
  - essentially done indiv. by indiv.
- Another viewpoint : use knowledge from all the population
  - and adopt a statistical approach.
  - Again : huge literature

♣ See Lectures by Marie Doumic and Laurent Dumas ♣



Time 0 corresponds to the time of treatment


Time 0 corresponds to the time of treatment
MTD (mm)
Time (months)
90
80
70
60
50
40
30
20
10
−100
−80
−60
−40
−20
0
20
40
60
Diagnosis
Treatment starts
Treatment ends
Max. tumor shrinkage
E
F
G
H

Previously done in the team

A population study with an **ODE** model : **statistically robust**.

For more details, see : B. Ribba et al. *A Tumor Growth Inhibition Model for Low-Grade Glioma Treated with Chemotherapy or Radiotherapy.* Clin Cancer Res. Sep 15 ;18(18) :5071-80 (2012)

One of the next questions is : **can we extend this kind of population approach to a PDE model ?**

In particular, for gliomas, numerous previous studies are based on a **KPP model** (reaction-diffusion). See, *e.g.* :

- Mandonnet et al. *Computational modeling of the WHO grade II glioma dynamics : principles and applications to management paradigm.* Neurosurg Rev 2008 ;31 :263-9.
- Murray JD. Mathematical biology. 3rd ed. New York : Springer ; 2002.
- Swanson et al. *Virtual and real brain tumors : using mathematical modeling to quantify glioma growth and invasion.* J Neurol Sci 2003 ;216 :1-10.

Outline

1. Population approach and PDEs

2. SAEM + Kriging

A closer look at the population algorithm

$$y_{ij} = f(x_{ij}, \psi_i) + \varepsilon_{ij}, 1 \leq i \leq N, 1 \leq j \leq n_i \qquad (1)$$

- $y_{ij} \in \mathbb{R}$: j^{th} observation of individual i
- N : number of individuals
- n_i : number of observations of individual i
- $x_{ij} \in \mathbb{R}^{n_x}$: **known** design variables
- ψ_i : vector of the n_ψ **unknown individual parameters**

A closer look – Non linear mixed effects models

$$y_{ij} = f(x_{ij}, \psi_i) + \varepsilon_{ij}, 1 \leq i \leq N, 1 \leq j \leq n_i$$

$$\psi_i = h(c_i, \mu, \eta_i), \quad \text{e.g.} \quad \psi_{ik} = \mu_k + \eta_{ik} = \psi_k^{pop} + \eta_{ik} \quad (2)$$

- c_i : known vector of M covariates
- μ : unknown vector of **fixed effects** (size p)
- $\eta_i \sim_{i.i.d.} \mathcal{N}(0, \Omega)$: unkn. vect. of **random effects** (size q)

  Ω is the $q \times q$ var.– covariance matrix of the rand. eff.
- $\varepsilon_{ij} \sim_{i.i.d.} \mathcal{N}(0, \sigma^2)$: residual errors

Parameters of the model to be determined : $\theta = (\mu, \Omega, \sigma^2)$

Stochastic Approximation of EM + Estimation/Maximization of the conditional distributions with MCMC

NLMEM and SAEM : what's done ?

To our knowledge, the following is working with MONOLIX :

- ODE's
- Systems of ODE's and Chains of ODE's
- Stochastic DE's
- Numerous validation on real applications :
  - PK/PD (1 or more compart.), viral dynamics models
  - oncology, etc

but the **integration of PDE's remains an open problem**.

Some attempts here and there but essentially done by transforming the PDE into a set of ODE's.

Why ? 'Cause of the computational cost and polymorphism

The primitive idea ...

You want to keep the PDE to have the solution :

- decouple PDE resolution and SAEM evaluation :
- **precompute** solutions (as functions of parameters)
- store them and call them when SAEM needs them

This is the classical **Offline/Online** concept

- Offline step : very long computational time
- Online step : “instantaneous” $\Rightarrow$ SAEM doable

Rk : there is still the problem of storage ... (balance v.s. cpu)

Precomputation

To evaluate quickly a function f, ...

... interpolate from precomputed values on a grid

For efficiency :

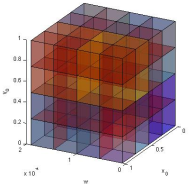
- Interpolation should be easy/fast $\rightarrow$ quadtree/octree
- Mesh refined in "high variations" zones of f ...
  ... in a sense to be defined

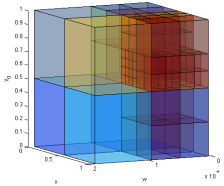
Precomputation algorithm

Start with an hyper-rectangle (let’s say a “cube”) :

$$C_{init} = \Pi_{i=1}^{N} [x_{min,i}, x_{max,i}]$$

- Divide the “cube” and compute weigths of children
- Choose a child (e.g. highest weight) and divide it
- Iterate as needed





Can be donne in parallel. As such, doable for ~ 5 parameters.
→ **for more param, additional ideas are needed, cf next part**

Reconnecting with PDE's

- MONOLIX is not able to deal with PDE output
- Whole solution needs to be "reduced"
- → transformation into scalar time series
- → Rely on the knowledge about the PDE
- → Difficult to expect generalization for all PDE's
- → Identifiability becomes even more crucial

Rk : reduction/transformation is also good for storage ...

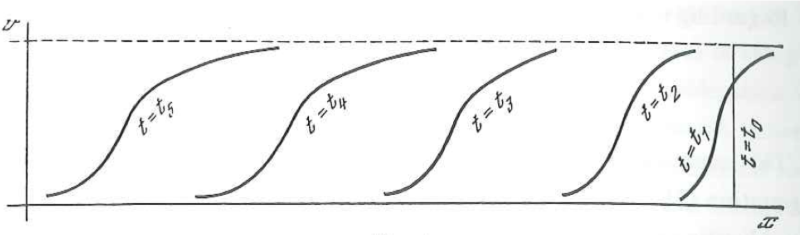
But, still, some things are doable :o)

Description of the KPP model

We consider the classical reaction-diffusion PDE named after Kolmogoroff, Petrovsky and Piscounoff (1937)[1].

$$\partial_t u - \nabla.(D\nabla u) = Ru(1-u), \forall t > 0, \forall x \in \Delta \tag{3}$$

$$u(T_0, x) = \alpha 1_{|x-x_0|\leq\varepsilon}, \text{ and Neumann B.C. on } \Delta \tag{4}$$



---

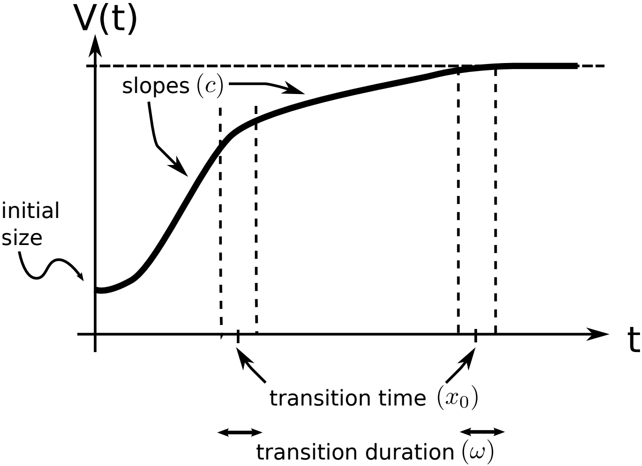
[1] Etude de l'equation de la diffusion avec croissance de la quantite de matiere et son application a un probleme biologique. Bulletin de l'universite d'Etat a Moscou. Section A, I(6) :1-26, 1937.

Properties of the KPP model

- Maximum principle : $\forall t > 0, \quad 0 \le u(t,.) \le 1$
- Good model for front propagation
- Speed $:= c = 2\sqrt{RD}$, Front width $:= \omega \propto \sqrt{\frac{D}{R}}$
- Define the “volume” of the invaded zone :

$$V(t) := \int_{\Delta} u(t,x)dx \tag{5}$$

Consequences on the volume



- Pract. identif. : Yes. 1D (up to symetry in x_0), 2D (up to Δ)
- **PDE-SAEM algo used on time series of the volumes of a population of individuals**

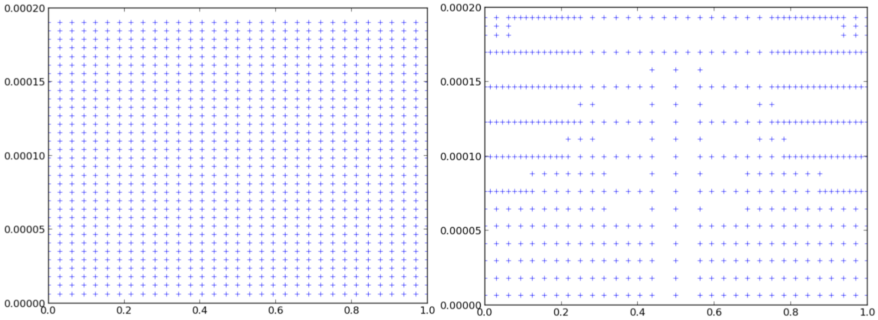
Technical details

- I.C. $\varepsilon = 0.03$ and $\alpha = 1$.
- Define the parameters' space (medical appl.) : x_0, R, D
- $0 \leq x_0 \leq 1$
- $7.2 \times 10^{-3} \leq R \leq 4.0 \times 10^{-2}$
- $2.5 \times 10^{-7} \leq D \leq 13.9 \times 10^{-7}$

Build 2 databases (see next slide) :

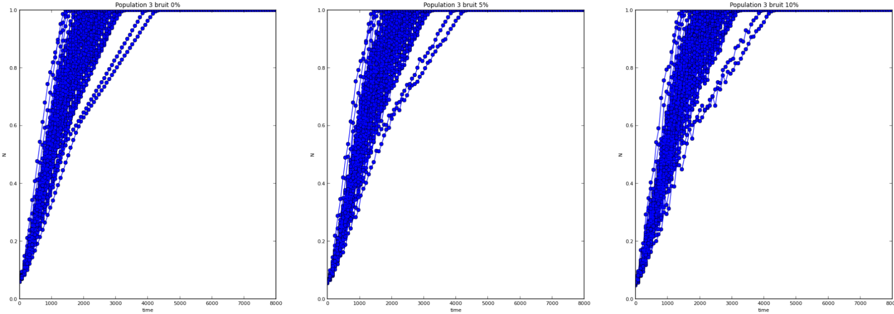
- homogeneous : 1089 summits
- heterogeneous : 500 summits

Technical details – Databases



Note the finer zones (compared to Left) on the Right.

Technical details - Populations



100 individuals in each population. Noise : 0%, 5%, 10%

Lognormal distribution of parameters.

101 points in time.

Results : Case (x_0, R, D) – population errors

	Theor	E1		E2		E3	
			error		error		error
R	0.0245	0.0237	-3.3%	0.0234	-4.5%	0.0231	-5.7%
D	$8.64e^{-7}$	$8.67e^{-7}$	0.3%	$8.79e^{-7}$	1.7%	$9.62e^{-7}$	11%
x_0	0.415	0.399	-3.9%	0.393	-5.3%	0.37	-11%
ω_R	0.201	0.196	-2.5%	0.263	31%	0.253	26%
ω_D	0.205	0.188	-8.3%	0.247	20%	0.395	93%
ω_{x_0}	0.254	0.244	-3.9%	0.241	-5%	0.616	143%

TABLE – Homogeneous grid : Column E1 : 0% noise. E2 (resp. E3) refers to a population with a 5% (resp. 10%) noise.

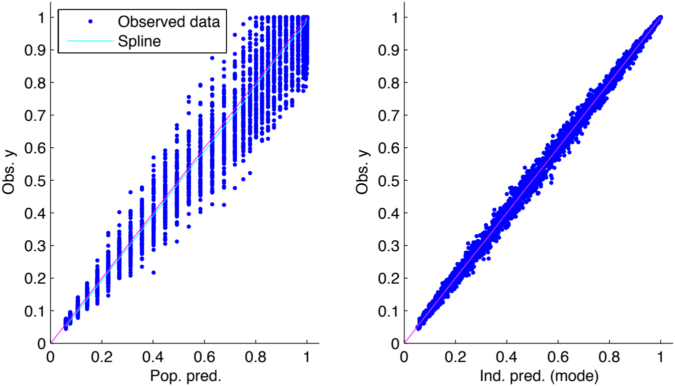
Results : Case (x_0, R, D) – population errors

	Theor	E1		E2		E3	
			error		error		error
R	0.0245	0.0245	0%	0.0241	-1.6%	0.0239	-2.4%
D	$8.64e^{-7}$	$8.31e^{-7}$	-3.8%	$8.47e^{-7}$	-1.9%	$8.66e^{-7}$	0.2%
x_0	0.415	0.414	-0.2%	0.406	-2.1%	0.436	5%
ω_R	0.201	0.197	-1.9%	0.238	18.4%	0.257	27.8%
ω_D	0.205	0.191	-6.8%	0.238	16%	0.299	45.8%
ω_{x_0}	0.254	0.262	3.1%	0.247	-2.7%	0.290	14.1%

TABLE – Inhomogeneous grid : Column E1 : 0% noise. E2 (resp. E3) refers to a population with a 5% (resp. 10%) noise.

Same quality with lower cost

Results : pred vs obs indiv params (100 ind), 10%



Results : computational cost

	"Exact" case	Interpolation with homogeneous mesh	Interpolation with heterogeneous mesh
Offline	No offline computation	Mesh with n segmentations, $(2^n+1)^2$ points. For 5 segmentations, 1089 points	Mesh with n points. Example with 500 points
Unit average CPU	-	$2.12s$	$2.12s$
Offline total CPU	-	$38mn28s$	$17mn40s$
Online	SAEM, 10^6 KPP evaluations	SAEM, 10^6 interpolations	SAEM, 10^6 interpolations
Unit average CPU	$2s$	$4.5 \times 10^{-4}s$	$5.1 \times 10^{-4}s$
Online total Cost	$\sim 23\ days\ 3\ h$	$7mn30s$	$8mn30s$
Total cost	$\sim 23\ days\ 3\ h$	$45mn58s$	$26mn10s$

TABLE – The number of calls of the solver in SAEM is about 10^6 for this case. Note that this is sequential CPU time. The mesh generation can be easily parallelized on many cores with an excellent scalability.

All details and results of this part in : **Parameter estimation in non-linear mixed effects models with SAEM algorithm : extension from ODE to PDE**. E. Grenier, V. Louvet & PV. M2AN, 48(5), pp. 1303-1329, 2014. HAL link.

Another test on a few real data w/ Pierre Gabriel

Starting point :
The contribution of age structure to cell population responses to targeted therapeutics. P. Gabriel, S.P. Garbett, V. Quaranta, D. R. Tyson and G. F. Webb. J. of Theor. Biology. 311(19). (2012)
Model :

$$\forall t > 0, \forall a > 0, \quad \partial_t p(t,a) + \partial_a p(t,a) + \beta(a) p(t,a) = 0 \tag{6}$$

$$\text{Initial condition : } \forall a \geq 0, \quad p(0,a) = p_0(a) \tag{7}$$

$$\text{B. C. : } \forall t \geq 0, \quad p(t,0) = 2(1-f) \int_0^\infty \beta(a) p(t,a)\, da \tag{8}$$

$$\frac{dQ(t)}{dt} = 2f \int_0^\infty \beta(a) p(t,a)\, da \tag{9}$$

$$P(t) = \int_0^\infty p(t,a)\, da \tag{10}$$

$$N(t) = P(t) + Q(t) \tag{11}$$

The data

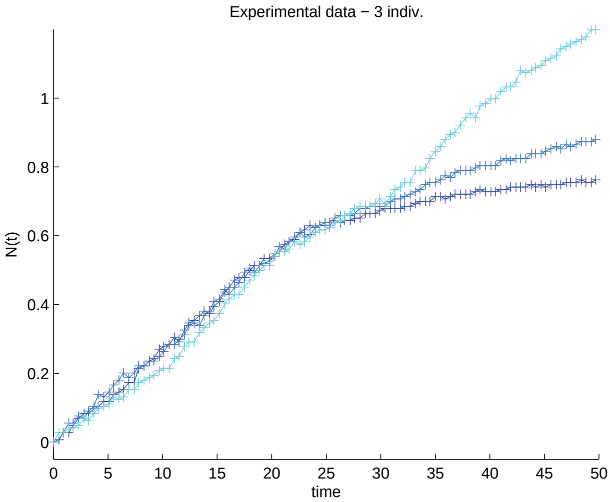


FIGURE – Typical profiles for a "population" of ... 3 individuals.

Parameters estimation with SAEM

Key ingredient from P. Gabriel et al. : one can take

$$\beta(a) \quad = \quad \beta_0 \, Erfc\left(\frac{m-a}{\sigma}\right), \tag{12}$$

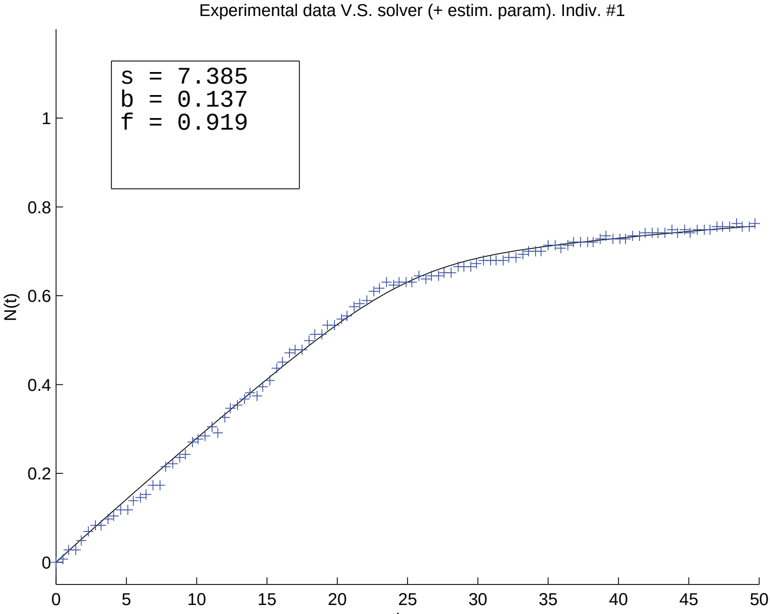
Parameter m can be fixed to 25.

Unknow parameters studied via SAEM : f ; β_0, σ

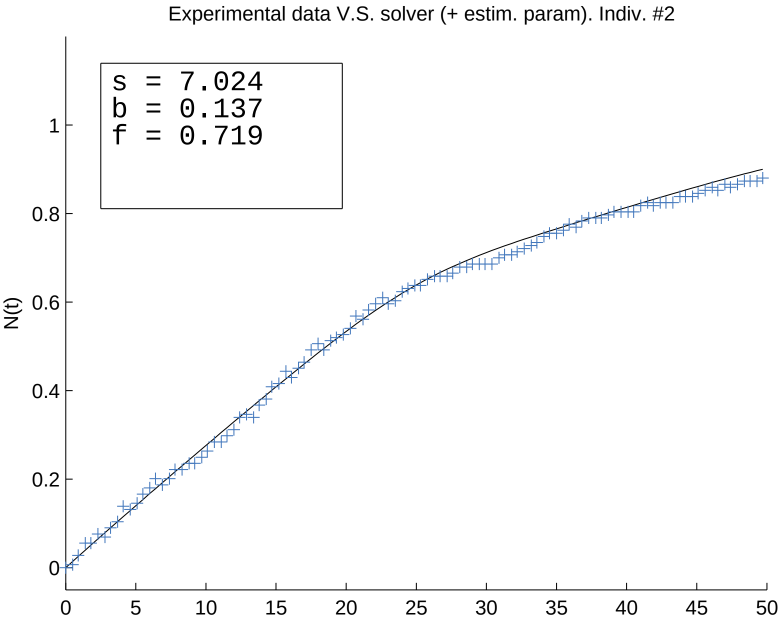
We run SAEM only with these 3 individuals, even if this is not statistically relevant.

However, note that there is a lot of points in time so we get something :

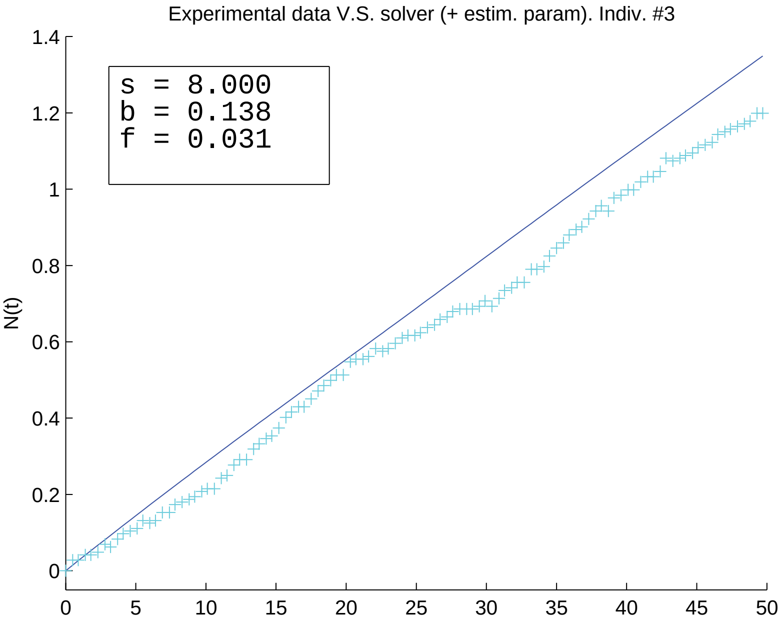
Erlotinib 5000nM



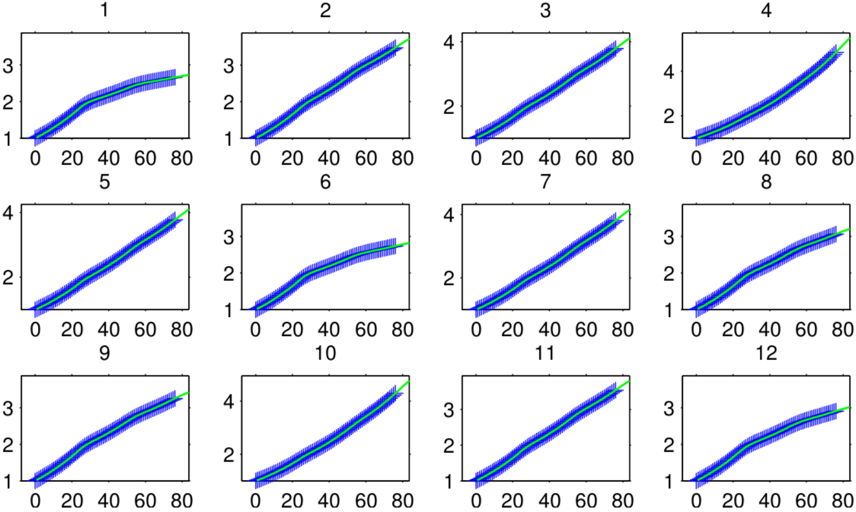
Erlotinib 50nM



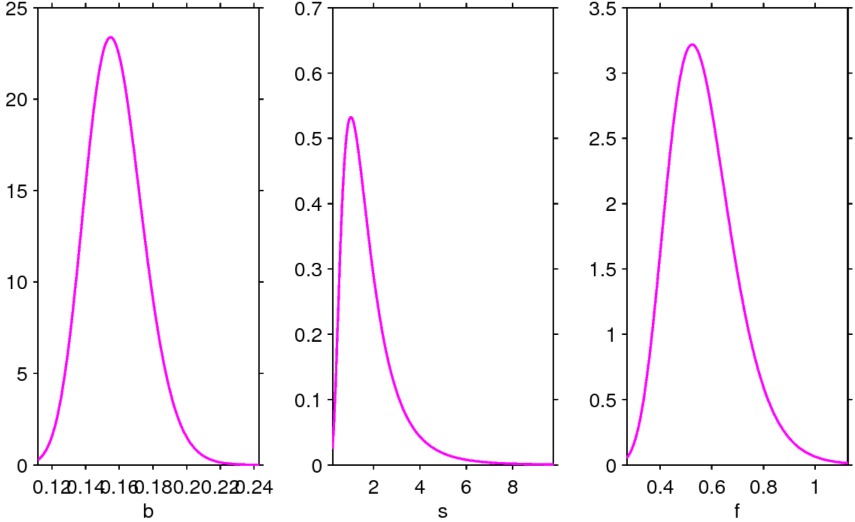
Erlotinib 5nM



Statistical check with 100 digital individuals



Statistical check with 100 digital individuals



Outline

1. Population approach and PDEs

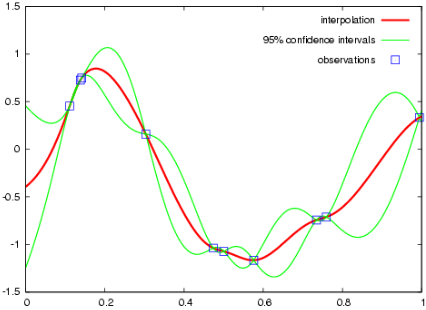
2. SAEM + Kriging

Trying to diminish the computational cost

Recall : previous approach limited to 5-6 parameters.
= "**dimensional curse**"

A possible choice is to make a SAEM method with an approximate model $f_{app}(k)$ which evolves during the SAEM iterations k :

- use **kriging** as the interpolation (has a *variance* !)
- start from **coarse** kriging and **refine** $f_{app}(k)$ only when needed along the SAEM "exploration" of the parameters' space, informed by the kriging **variance**



KSAEM = SAEM + Kriging

- Cannot be compared directly to the Part 1 Method :
    - if the SAEM algo is used many times on the same model : can be better to pay for a costly offline step, once for all
    - if SAEM algo is used a few times : better to use KSAEM
- By property of kriging :
  very few points added in the basis/grid along KSAEM
- For PDE models : can expect a drastic ↘ of costly eval of f
- Rigorous proof of convergence of KSAEM : open problem.
  At least, numerical experiments so far show quite good computations.

KSAEM illustrated on KPP

Same model as presented in Part 1.

- SAEM : 9×10^5 calls of f, 3 days on a laptop (C++ code)
- KSAEM :
  - Start with 20 points in the kriging grid.
  - 22 points then added during the KSAEM loop ...
  - ... and 9.3×10^5 fast interpolations via f_{app}
- So : 9×10^5 to be compared with 44 costly eval. of f !
  CPU times : 3 days v.s. 1 minute !
- What about the estimated parameters ? See next :

KSAEM illustrated on KPP

Parameters	True values	SAEM	KSAEM
μ_λ	0.0236	0.0229 (0.009)	0.0259 (0.013)
μ_ν ($\times 10^7$)	8.195	8.6327 (4.058)	8.3869 (4.390)
μ_{x_0}	0.4	0.4024 (0.107)	0.5615 (0.200)
ω_λ^2	0.04	0.0391 (0.035)	0.1382 (0.143)
ω_ν^2	0.04	0.0451 (0.050)	0.1688 (0.173)
$\omega_{x_0}^2$	0.04	0.0426 (0.030)	0.0905 (0.113)
σ_ε^2	0.05	0.0391 (0.035)	0.1383 (0.143)

TABLE – Simulation study with KPP model : results obtained from 100 repetitions, with $N = 100$ individuals with the exact SAEM and KSAEM. Results are presented in means and standard deviation in brackets.

All (tricky) details on KSAEM in : E. Grenier, C. Helbert, V. Louvet, A. Samson, P.V. **Population parametrization of costly black box models using iterations between SAEM algorithm and kriging. Computational and Applied Mathematics, March 2016.** HAL link.

Conclusions

Summary

- Use of SAEM population approaches ...
- in the context of PDE (not just ODE).
- Two methods : one pragmatical (quite easy implementation) – offline grid
- another a bit more tricky, but can be usefull – KSAEM
- Illustrated and encouraging on KPP and Renewal equation

Perspectives

- Applications to other models
- Extensions to take into account full images
- Rigorous proof of convergence of KSAEM