Sequential Design of Computer Experiments for Numerical Dosimetry

Marjorie Jala
(Télécom ParisTech/Orange Labs)

Ph.D. supervisors : Céline Lévy-Leduc, Éric Moulines, Emmanuelle Conil and Joe Wiart

MASCOT-SAMO 2013 * Ph.D. Students Day * July 1st 2013

Motivation

Sequential Design of Computer Experiments ...

Estimation of the α-quantile q_α of the distribution of $Y = f(\mathbf{X})$, for a given α in $(0, 1)$,

$$q_\alpha = \inf \{q : \mathbb{P}(Y \leqslant q) > \alpha\} \ .$$

- f is an unknown, expensive-to-evaluate real-valued function
- $\mathbf{X}$ is a random vector having a known distribution on a compact subset $\mathcal{A} \subseteq \mathbb{R}^d$.

We aim at estimating q_α by using as few evaluations of f as possible

... for Numerical Dosimetry

At wich level are fetuses exposed to Radio Frequency Electromagnetic Fields ?

Background: Gaussian Process Modelling

Assume that f is a sample of a zero-mean Gaussian process (GP) having a covariance function k: $\mathrm{GP}(0, k(.,.))$

Conditionally to $\mathbf{y}_t = (y_1, \ldots, y_t)'$, the mean $\mu_t(u)$ and covariance $k_t(u, v)$ are given by

$$\mu_t(u) = \mathbf{k}_t(u)'\mathbf{K}_t^{-1}\mathbf{y}_t \ ,$$
$$k_t(u, v) = k(u, v) - \mathbf{k}_t(u)'\mathbf{K}_t^{-1}\mathbf{k}_t(v) \ ,$$

where $\mathbf{k}_t(u) = [k(x_1, u) \ldots k(x_t, u)]'$, $'$ denotes the matrix transposition, $\mathbf{K}_t = [k(x_i, x_j)]_{1 \leqslant i,j \leqslant t}$, u and v and the x_i's are in $\mathcal{A}$.

Covariance function

Since the SAR is supposed to be smooth, we shall use the square exponential covariance function

$$k_{\mathrm{SE}}(u, v) = \exp\left(-\frac{\|u - v\|^2}{2\ell^2}\right) \ , u, v \in \mathcal{A} \ , \ell > 0 \ ,$$

where $\|u\|$ denotes the euclidean norm of u in $\mathbb{R}^d$.

Background: Methodologies

Sequential strategies

Bayesian optimization: find the maximum of f, optimizing an acquisition function

- Expected Improvement [Vazquez et al., 2010]
- Confidence Bound Criteria (GP-UCB [Srinivas et al., 2010], Branch and Bounds [De Freitas et al., 2012])

EI has been adapted for

- Contour estimation [Ranjan et al., 2009]
- Estimation of $\mathbb{P}(Y \geqslant s)$ where s is a given threshold (SUR) [Bect et al., 2012]

Quantile estimation

- Non sequential approach [Oakley, 2004]
- Extension of the SUR criterion [Arnaud et al., 2010]

We really need a sequential strategy, but improvement based criteria demand Monte Carlo samplings of the GP and the conditional GPs, which made them difficult to use for $d > 2$

Quantile estimation

We shall compare the quantile estimators with $\tilde{q}_{\alpha,m}$ defined by

$$\tilde{q}_{\alpha,m} = \inf\left\{ q : \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{\{f(\mathrm{x}_i)\leqslant q\}} > \alpha \right\} ,$$

where $\mathrm{x}_1, \ldots, \mathrm{x}_m$ are m fixed points in $\mathcal{A}$.
Let $\mathrm{A} = \{\mathrm{x}_1, \ldots, \mathrm{x}_m\} \subset \mathcal{A}$.

Pure exploration criterion

- Minimizes the global uncertainty on the estimation of f
- New point x_{t+1} to add to the set of t observations:

$$x_{t+1} \in \arg\max_{x \in \mathrm{A}} \sigma_t(x) .$$

- Propose methodologies more adapted to our quantile estimation issue to realize the exploration-exploitation trade-off

GPS

- Let $\mu_t^U(x) = \mu_t(x) + \sqrt{\beta_t}\sigma_t(x)$ and $\mu_t^L(x) = \mu_t(x) - \sqrt{\beta_t}\sigma_t(x)$ with $\beta_t = 2\ln\left(\frac{\pi^2 t^2}{6}\right) + 2\ln\left(\frac{m}{\delta}\right)$ where m is the cardinal of $\mathbb{A}$

- Let $\hat{q}^U_{\alpha,t}$ and $\hat{q}^L_{\alpha,t}$ be the estimators of the α-quantile of μ_t^U and μ_t^L

$$\hat{q}^U_{\alpha,t} = \inf\left\{q : \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{\{\mu_t^U(\mathrm{x}_i) \leqslant q\}} > \alpha\right\}$$

$$\hat{q}^L_{\alpha,t} = \inf\left\{q : \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{\{\mu_t^L(\mathrm{x}_i) \leqslant q\}} > \alpha\right\}$$

Proposition

For all δ in $(0, 1)$, for all $t \geqslant 1$, with probability greater than $(1 - \delta)$,

$$\tilde{q}_{\alpha,m} \in [\hat{q}^L_{\alpha,t}, \hat{q}^U_{\alpha,t}] \, .$$

GPS (cont.)

Let $U_{\alpha,t}$ and $L_{\alpha,t}$ be the following sets :

$$U_{\alpha,t} = \left\{ x \in \mathrm{A} : \mu_t^U(x) \geqslant \hat{q}^L_{\alpha,t} \right\} \text{ and } L_{\alpha,t} = \left\{ x \in \mathrm{A} : \mu_t^L(x) \leqslant \hat{q}^U_{\alpha,t} \right\} \ , \ t \geqslant 1.$$

Proposition

With probability greater than $(1 - \delta)$, for all $t \geqslant 1$,

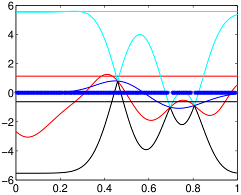
$$|\hat{q}_{\alpha,t} - \tilde{q}_{\alpha,m}| \leqslant \sqrt{\beta_t} \sup_{\mathrm{x} \in U_{\alpha,t}} \sigma_t(\mathrm{x}) \ .$$

Criterion

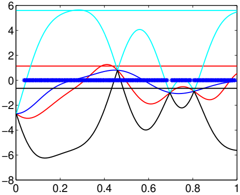
x_{t+1} to add to the set of t observations is such that:

$$x_{t+1} \in \arg\max_{x \in U_{\alpha,t}} \sigma_t(x) \ .$$

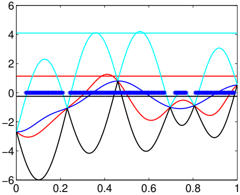
Illustration : 1D Gaussian Process sample path



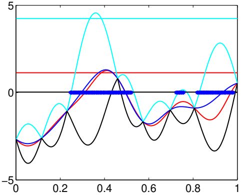
$t = 3$



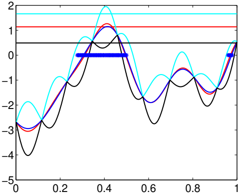
$t = 4$



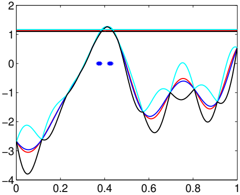
$t = 6$



$t = 8$



$t = 10$



$t = 12$

GPS+

- Let $\mathcal{S}_{\alpha,t} \subseteq \mathcal{A}$ be the compact subset such that $\mathcal{S}_{\alpha,t} = \prod_{i=1}^{d} [x_{\min,t}^{(i)}, x_{\max,t}^{(i)}] \times \cdots \times [x_{\min}^{(d)}, x_{\max}^{(d)}]$. Here $x_{\min,t}^{(i)}$ and $x_{\max,t}^{(i)}$ denote the smallest (resp. the largest) ith component of the points in $\bar{U}_{\alpha,t}$
- where $\bar{U}_{\alpha,t} = \left\{ x \in \mathrm{S}_{\alpha,t-1} : \mu_t^U(x) \geqslant \hat{q}_{\alpha,t}^L \right\}$
- where $\mathrm{S}_{\alpha,t} = \{\mathrm{x}_{t,1}, \ldots, \mathrm{x}_{t,m_t}\} \cup \bar{U}_{\alpha,t}$,
- where $\{\mathrm{x}_{t,1}, \ldots, \mathrm{x}_{t,m_t}\}$ are m_t points randomly chosen in $\mathcal{S}_{\alpha,t}$
- By convention $\mathrm{S}_{\alpha,0} = \mathrm{A}$.

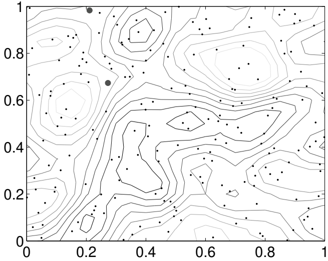
Criterion

x_{t+1} to add to the set of t observations is such that:

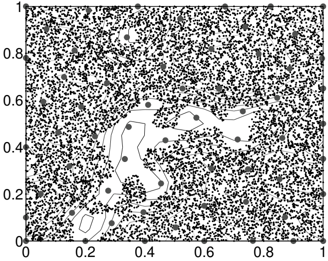
$$x_{t+1} \in \arg\max_{x \in \bar{U}_{\alpha,t}} \sigma_t(x) \ .$$

Note: Since the size of the grid varies at each iteration of the process, we use $\beta_t = 2\ln\left(\frac{\pi^2 t^2}{6}\right) + 2\ln\left(\frac{|\mathrm{S}_{\alpha,t-1}|}{\delta}\right)$.

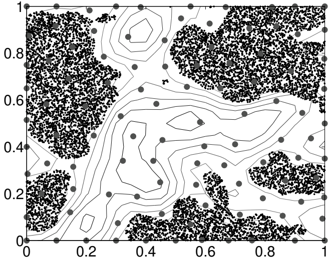
Illustration : 2D Gaussian Process sample path



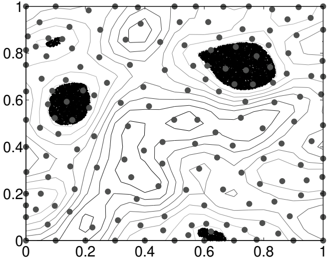
$t = 2$



$t = 62$



$t = 102$



$t = 162$

Numerical Dosimetry ?

In general

Virtually expose human 3D-models to one source of EMF in order to evaluate the Specific Absorption Rate (the SAR, in $W.kg^{-1}$)
SAR computation in our case is done through Finite Difference in Time Domain (FDTD) method
The SAR depends on

- the geometry of the models
- the dielectric properties of the tissues
- the type and position of the EMF source

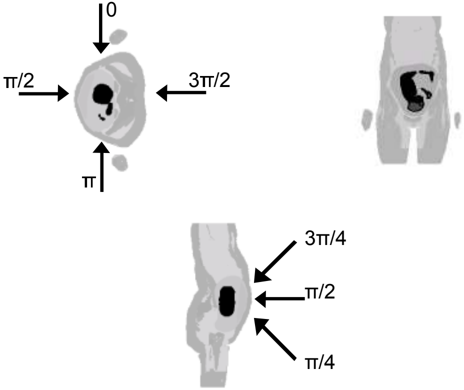
Fetus exposure

- Very few models are available
- The simulations are expensive in terms of computational load
- The preparation of the simulations is very complex

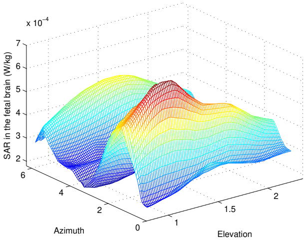
We focus on the fetal brain exposure

Application I: GPS, Japanese model and plane wave

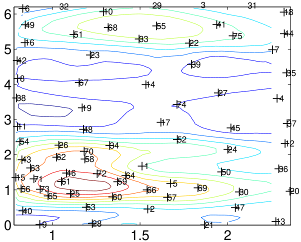
- Plane wave exposure: far field sources (base stations antennas, WiFi boxes)
- 900 MHz vertically polarized electromagnetic plane waves with a 1 Volt per meter amplitude
- Start by performing 5 randomly chosen evaluations of the SAR in order to have an estimation of I



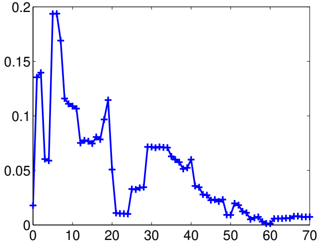
Application I: GPS, Japanese model and plane wave (cont.)



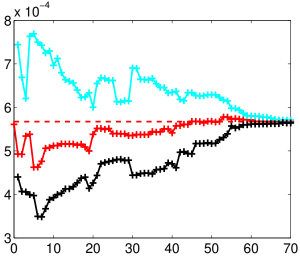
3D-plot



Contour plot and observations



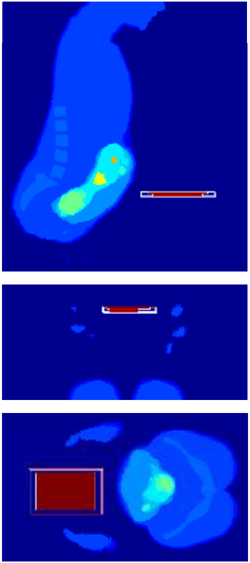
Relative errors



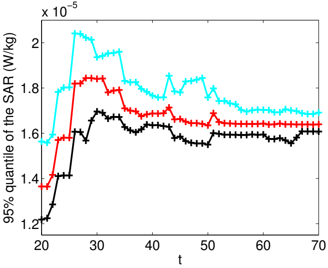
Quantile convergence

Application II: GPS+, Victoria and Samsung Galaxy Tab

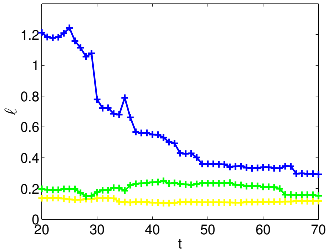
- Model Victoria is sitting working on her Samsung Galaxy Tab at 3G frequency (1940 MHz)
- 3 parameters: height, nearness and slope of the tablet
- Start by performing 20 evaluations of the SAR from a LHS in order to have an estimation of I



Application II: GPS+, Victoria and Samsung Galaxy Tab (cont.)



Quantile convergence



ℓ convergence

Conclusion

- We propose two novel sequential approaches for quantile estimation
- Successfully applied to real data coming from numerical dosimetry