Estimating and quantifying uncertainties on level sets using the Vorob'ev expectation and deviation with Gaussian process models

Clément Chevalier, David Ginsbourger, Julien Bect, Ilya Molchanov

Estimating and quantifying uncertainties on level sets using the Vorob'ev expectation and deviation with Gaussian process models

Clément Chevalier, David Ginsbourger, Julien Bect and Ilya Molchanov

Abstract Several methods based on Kriging have been recently proposed for calculating a probability of failure involving costly-to-evaluate functions. A closely related problem is to estimate the set of inputs leading to a response exceeding a given threshold. Now, estimating such level set – and not solely its volume – and quantifying uncertainties on it are not straightforward. Here we use notions from random set theory to obtain an estimate of the level set, together with a quantification of estimation uncertainty. We give explicit formulae in the Gaussian process set-up and provide a consistency result. We then illustrate how space-filling versus adaptive design strategies may sequentially reduce level set estimation uncertainty.

1 Introduction

Reliability studies increasingly rely on complex deterministic simulations. A problem that is often at stake is to identify, from a limited number of evaluations of $f : D \subset \mathbb{R}^d \mapsto \mathbb{R}$, the level set of "dangerous" configurations $\Gamma_f = \{x \in D : f(x) \geq T\}$, where T is a given threshold. In such context, it is commonplace to predict quantities of interest relying on a surrogate model for f. This approach was popularized in the design and analysis of computer experiments [12, 11, 7] . In the Kriging framework, several works have already been proposed for reliability problems (see, e.g., [2, 9, 10, 6] and the references therein). However, the quantity of interest is usually the *volume* of Γ_f, and none of the methods explicitly reconstruct Γ_f itself.

Clément Chevalier, David Ginsbourger and Ilya Molchanov
Institute of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5 and Alpeneggstrasse 22, CH-3012 Bern, Switzerland, e-mail: clement.chevalier@stat.unibe.ch; david.ginsbourger@stat.unibe.ch; ilya.molchanov@stat.unibe.ch

Julien Bect
Supélec Sciences des Systèmes - EA4454 (E3S), SUPELEC, Plateau de Moulon, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette cedex, France, e-mail: julien.bect@supelec.fr

An illustrative example for this issue is given on Figure 1. A Kriging model is built from five evaluations of a $1d$ function (left plot). Three level set realisations (with $T = 0.8$) are obtained from Gaussian process (GP) conditional simulations. The focus here is on summarizing the conditional distribution of excursion sets using ad hoc notions of expectation and deviation from the theory of random sets. We address this issue using an approach based on the Vorob'ev expectation [1, 8].

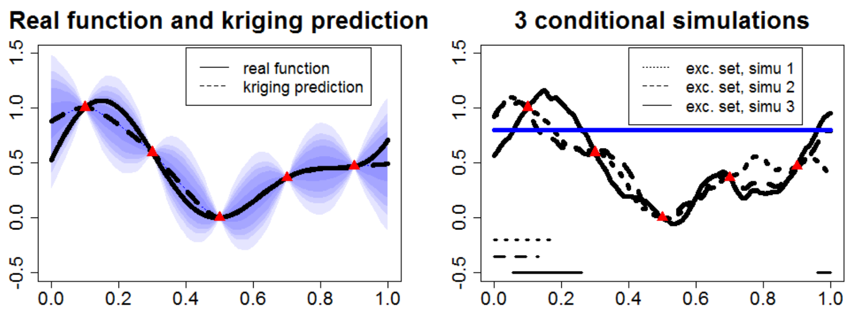

Fig. 1: Conditional simulations of level sets. Left: Kriging model obtained from five evaluations of a $1d$ function. Right: Three GP conditional simulations, leading to three different level sets. Here the threshold is fixed to $T = 0.8$.

In Section 2 we present the Vorob'ev expectation and deviation for a closed random set. In Section 3 we then give analytical expressions for these quantities in the GP framework. In addition we give consistency result regarding the convergence of the Vorob'ev expectation to the actual level set. To the best of our knowledge, this is the first Kriging-based approach focusing on the level set itself, and not solely its volume. Our results are illustrated on a test case in Section 4.

2 The Vorob'ev expectation and deviation in Random Set theory

Random variables are usually defined as measurable maps from a probability space $(\Omega, \mathscr{G}, \mathbb{P})$ to some measurable space, such as $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ or $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. However there has been a growing interest during the last decades for set-valued random elements, and in particular for *random closed sets* [8].

Definition 1. Let $\mathscr{F}$ be the family of all closed subsets of D. A map $X : \Omega \mapsto \mathscr{F}$ is called a *random closed set* if, for every compact set K in D,

$$\{\omega : X(\omega) \cap K \neq \emptyset\} \in \mathscr{G}. \tag{1}$$

As mentioned in [8], this definition basically means that for any compact K, one can always say when observing X if it hits K or not. Defining the expectation of a random set is far from being straightforward. Different candidate notions of expectation from the random set literature are documented in [8] (Chapter 2), with a major development on the *selection expectation*. Some alternative expectations mentioned in [8] include the *linearisation approach*, the *Vorob'ev expectation*, the *distance average*, the *Fréchet expectation*, and the *Doss and Herer expectations*.

In the present work we focus on the Vorob'ev expectation, which is based on the intuitive notion of coverage probability function. Given a random closed set X over a space D with σ-finite measure μ (say $D \subset \mathbb{R}^d$ and $\mu = \text{Leb}_d$), then X is associated with a random field $(\mathbb{1}_X(x))_{x \in D}$. The coverage function is defined as the expectation of this binary random field:

Definition 2 (coverage function and α-quantiles of a random set). The function

$$p_X : x \in D \mapsto \mathbb{P}(x \in X) = \mathbb{E}(\mathbb{1}_X(x)) \tag{2}$$

is called the *coverage function* of X. The α*-quantiles* of X are the level sets of p_X,

$$Q_\alpha := \{x \in D : p_X(x) \geq \alpha\},\ \alpha \in (0,1]. \tag{3}$$

Note that in Equation 2, the expectation is taken with respect to the set X and not to the point x. In Figure 1 (right) we plotted three conditional realizations of the random set $X := \{x \in [0,1], \xi(x) \geq T\}$, where ξ is a GP. The α-quantile defined in Definition 2 can be seen as the set of points having a (conditional, in Figure 1) probability of belonging to X greater or equal than α. This definition is particularly useful here as, now, the so-called Vorob'ev expectation of the random set X will be defined as a "well-chosen" α-quantile of X.

Definition 3 (Vorob'ev expectation). Assuming $\mathbb{E}(\mu(X)) < \infty$, the *Vorob'ev expectation* of X is defined as the α^*-quantile of X, where α^* is determined from

$$\mathbb{E}(\mu(X)) = \mu(Q_{\alpha^*}) \tag{4}$$

if this equation has a solution, or in general, from the condition

$$\mu(Q_\beta) \leq \mathbb{E}(\mu(X)) \leq \mu(Q_{\alpha^*}) \text{ for all } \beta > \alpha^*. \tag{5}$$

Throughout this paper, an α^* satisfying the condition of Definition 3 will be referred to as a *Vorob'ev threshold*.

Property 1. For any measurable set M with $\mu(M) = \mathbb{E}(\mu(X))$, we have:

$$\mathbb{E}(\mu(Q_{\alpha^*} \Delta X)) \leq \mathbb{E}(\mu(M \Delta X)), \tag{6}$$

where $A \Delta B$ denotes the symmetric difference between two sets A and B. The quantity $\mathbb{E}(\mu(Q_{\alpha^*} \Delta X))$ is called *Vorob'ev deviation*.

The Vorob'ev expectation thus appears as a global minimizer of the deviation, among all closed sets with volume equal to the average volume of X. A proof can be found in [8], p. 193. In the next section, we will use these definitions and properties for our concrete problem, where the considered random set is a level set of a GP.

3 Conditional Vorob'ev expectation for level sets of a GP

In this section, we focus on the particular case where the random set (denoted by X in the previous section) is a level set

$$\Gamma := \{x \in D : \xi(x) \geq T\} \tag{7}$$

of a GP ξ above a fixed threshold $T \in \mathbb{R}$. Once n evaluation results $\mathscr{A}_n := ((x_1, \xi(x_1)), \dots, (x_n, \xi(x_n)))$ are known, the main object of interest is then the conditional distribution of the level set Γ given $\mathscr{A}_n$. We propose to use the Vorob'ev expectation and deviation to capture and quantify the variability of the level set Γ conditionally on the available observations $\mathscr{A}_n$.

3.1 Conditional Vorob'ev expectation and deviation

In the simple Kriging GP set-up (see, e.g., [5]), we know the marginal conditional distributions of $\xi(x)|\mathscr{A}_n$:

$$\mathscr{L}(\xi(x)|\mathscr{A}_n) = \mathscr{N}(m_n(x), s_n^2(x)), \tag{8}$$

where $m_n(x) = \mathbb{E}(\xi(x)|\mathscr{A}_n)$ and $s_n^2(x) = Var(\xi(x)|\mathscr{A}_n)$ are respectively the *simple Kriging* mean and variance functions. The coverage probability function and any α-quantile of Γ can be straightforwardly calculated (given $\mathscr{A}_n$) as follows.

Property 2. **(i)** The coverage probability function of Γ has the following expression:

$$p_n(x) := \mathbb{P}(x \in \Gamma|\mathscr{A}_n) = \mathbb{P}(\xi(x) \geq T|\mathscr{A}_n) = \Phi\left(\frac{m_n(x) - T}{s_n(x)}\right), \tag{9}$$

where $\Phi(.)$ denotes the c.d.f. of the standard Gaussian distribution.
(ii) For any $\alpha \in (0,1]$, the α-quantile of Γ (conditional on $\mathscr{A}_n$) is

$$Q_{n,\alpha} = \{x \in D : \ m_n(x) - \Phi^{-1}(\alpha) s_n(x) \geq T\}. \tag{10}$$

(iii) For any $\alpha \in (0,1]$, the α-quantile of Γ can also be seen as the excursion set above T of the Kriging quantile with level $1-\alpha$.

From Property 2, one can see hat the Vorob'ev expectation is in fact the excursion set above T of a certain Kriging quantile. In applications, an adequate Vorob'ev threshold value can be determined by tuning α to a level α_n^* such that $\mu(Q_{n,\alpha_n^*}) = \mathbb{E}(\mu(\Gamma)|\mathscr{A}_n) = \int_D p_n(x)\mu(\mathrm{dx})$. This can be done through a simple dichotomy.

Once the Vorob'ev expectation is calculated, the computation of the Vorob'ev deviation $\mathbb{E}(\mu(Q_{n,\alpha_n^*}\Delta\Gamma)|\mathscr{A}_n)$ does not require to simulate Γ. Indeed,

$$\begin{aligned}
\mathbb{E}(\mu(Q_{n,\alpha_n^*}\Delta\Gamma)|\mathscr{A}_n) &= \mathbb{E}\left(\int_D (\mathbb{1}_{x\in Q_{n,\alpha_n^*},x\notin\Gamma} + \mathbb{1}_{x\notin Q_{n,\alpha_n^*},x\in\Gamma})\mu(\mathrm{dx})\middle|\mathscr{A}_\mathrm{n}\right) \\
&= \int_{Q_{n,\alpha_n^*}} \mathbb{E}(\mathbb{1}_{x\notin\Gamma}|\mathscr{A}_n)\mu(\mathrm{dx}) + \int_{\mathrm{Q}^\mathrm{c}_{\mathrm{n},\alpha_\mathrm{n}^*}} \mathbb{E}(\mathbb{1}_{\mathrm{x}\in\Gamma}|\mathscr{A}_\mathrm{n})\mu(\mathrm{dx}) \\
&= \int_{Q_{n,\alpha_n^*}} (1-p_n(x))\mu(\mathrm{dx}) + \int_{\mathrm{Q}^\mathrm{c}_{\mathrm{n},\alpha_\mathrm{n}^*}} \mathrm{p_n(x)}\mu(\mathrm{dx}). \qquad (11)
\end{aligned}$$

We will present in Section 4 an example of computation of Vorob'ev expectation and deviation. Before that, we give in the next subsection a consistency result for the case where observations of ξ progressively fill the space D.

3.2 Consistency result

Let us consider a (zero-mean, stationary) GP Z and a deterministic sequence of sampling points $x_1, x_2, \ldots$, such that $s_n^{\max} \triangleq \sup_{x\in D} s_n \to 0$ (this holds, e.g., for any space-filling sequence, assuming that the covariance function is merely continuous). We denote by α_n^* the Vorob'ev threshold selected for the first n sampling points, and by $\kappa_n = \Phi^{-1}(\alpha_n^*)$ and $Q_{n,\alpha_n^*} \subset D$ the corresponding quantile and Vorob'ev expectation. Our goal here is to prove that the Vorob'ev expectation is a consistent estimator of the true excursion set Γ, in the sense that $\mu\left(Q_{n,\alpha_n^*}\,\Delta\,\Gamma\right) \to 0$ for some appropriate convergence mode. To do so, we shall consider a slightly modified estimator Q_{n,α_n^*}, where the choice of the Vorob'ev threshold α_n^* is constrained in such a way that $|\kappa_n| \leq \kappa_n^{\max}$, for some deterministic sequence of positive constants $\kappa_n^{\max}$.

Proposition 1 *Assume that $\mu(D) < +\infty$ and $\kappa_n^{max} = O\left(\sqrt{|\log s_n^{max}|}\right)$. Then*

$$\mathbb{E}\left(\mu\left(Q_{n,\alpha_n^*}\,\Delta\,\Gamma\right)\right) \;=\; O\left(s_n^{max}\sqrt{|\log s_n^{max}|}\right).$$

As a consequence, $\mu\left(Q_{n,\alpha_n^*}\,\Delta\,\Gamma\right) \to 0$ for the convergence in mean.

Proof. The result has been proven in [13, 14] in the special case $\kappa_n^{\max} = 0$ (i.e., with $\alpha_n^* = 1/2$). We follow their proof very closely.

Let us first rewrite the probability of misclassification at $x \in D$ as

$$\mathbb{E}\left(\mathbb{1}_{Q_{n,\alpha_n^*}\,\Delta\,\Gamma}(x)\right) \;=\; \mathbb{E}\left(\mathbb{1}_{p_n(x)\geq\alpha_n^*}\,(1-p_n(x)) + \mathbb{1}_{p_n(x)<\alpha_n^*}\,p_n(x)\right), \qquad (12)$$

and consider the events

$$E_n^+ = \{m_n(x) \geq T + w_n(x)\}\,, \quad E_n^- = \{m_n(x) \geq T - w_n(x)\}\,,$$

where $w_n(x)$ is a deterministic sequence that will be specified later. Let us assume that $\kappa_n^{\max} s_n(x) = O(w_n(x))$, uniformly in x. Then we have

$$|\kappa_n|\, s_n(x) \;\leq\; \kappa_n^{\max} s_n(x) \;\leq\; C\, w_n(x)$$

for some $C > 1$ (without loss of generality), and thus

$$\mathbb{1}_{p_n(x) \geq \alpha_n^*} \;=\; \mathbb{1}_{m_n(x) \geq T + \kappa_n s_n(x)} \;\leq\; \mathbb{1}_{|m_n(x)-T| \leq C w_n(x)} + \mathbb{1}_{E_n^+}.$$

As a consequence, noting that $\frac{m_n(x)-T}{s_n(x)} \geq \frac{w_n(x)}{s_n(x)}$ on E_n^+, we obtain:

$$\begin{aligned}\mathbb{1}_{p_n(x) \geq \alpha_n^*}\,(1-p_n(x)) \;&\leq\; \mathbb{1}_{|m_n(x)-T| \leq C w_n(x)} + \mathbb{1}_{E_n^+}\,(1-p_n(x)) \\ &\leq\; \mathbb{1}_{|m_n(x)-T| \leq C w_n(x)} + \Psi\left(\frac{w_n(x)}{s_n(x)}\right),\end{aligned}$$

where Ψ denotes the standard normal complementary cdf. Proceeding similarly with the second term in (12), it follows that

$$\mathbb{E}\left(\mathbb{1}_{Q_{n,\alpha_n^*}\,\Delta\,\Gamma}(x)\right) \;\leq\; 2\left[\Psi\left(\frac{w_n(x)}{s_n(x)}\right) + \mathbb{P}\left(|m_n(x)-T| \leq C w_n(x)\right)\right].$$

Using the tail inequality $\Psi(u) \leq \frac{1}{u\sqrt{2\pi}} \exp(-\frac{1}{2}u^2)$, and observing that $Var\,(m_n(x)) \geq s_0^2 - (s_n^{\max})^2 \geq s_0^2/4$ for n larger than some n_0 that does not depend on x, we have:

$$\mathbb{E}\left(\mathbb{1}_{Q_{n,\alpha_n^*}\,\Delta\,\Gamma}(x)\right) \;\leq\; \sqrt{\frac{2}{\pi}}\left[\frac{s_n(x)}{w_n(x)} \exp\left(-\frac{1}{2}\frac{w_n^2(x)}{s_n^2(x)}\right) + 4C\,\frac{w_n(x)}{s_0}\right]. \tag{13}$$

Finally, taking $w_n(x) = \sqrt{2}\, s_n(x)\, \sqrt{|\log s_n(x)|}$ as in [13], we have indeed $\kappa_n^{\max} s_n(x) = O(w_n(x))$ uniformly in x, and from (13) we deduce that

$$\mathbb{E}\left(\mathbb{1}_{Q_{n,\alpha_n^*}\,\Delta\,\Gamma}(x)\right) \;=\; O\left(s_n^{\max}\sqrt{|\log s_n^{\max}|}\right)$$

uniformly in x. The result follows by integrating with respect to μ over D.

4 Application to adaptive design for level set estimation

Here we present a 2-dimensional example on the notions and results previously detailed. We consider the Branin-Hoo function, with variables normalised so that the domain D is $[0,1]^2$. We multiply the function by a factor -1 and we are interested in the set $\{x \in D : f(x) \geq -10\}$. Figure 2 (top) gives the real level set and the coverage probability function obtained from $n = 10$ observations. The covariance parameters of the Gaussian process used for Kriging are assumed to be known. The measure μ is the uniform measure on $D = [0,1]^2$ and the current Vorob’ev deviation is $\mathbb{E}(\mu(Q_{n,\alpha_n^*} \Delta \Gamma)|\mathscr{A}_n) \approx 0.148$. All the integrals are calculated using the KrigInv R package [4] with a Sobol’ Quasi Monte Carlo sequence of 10000 points.

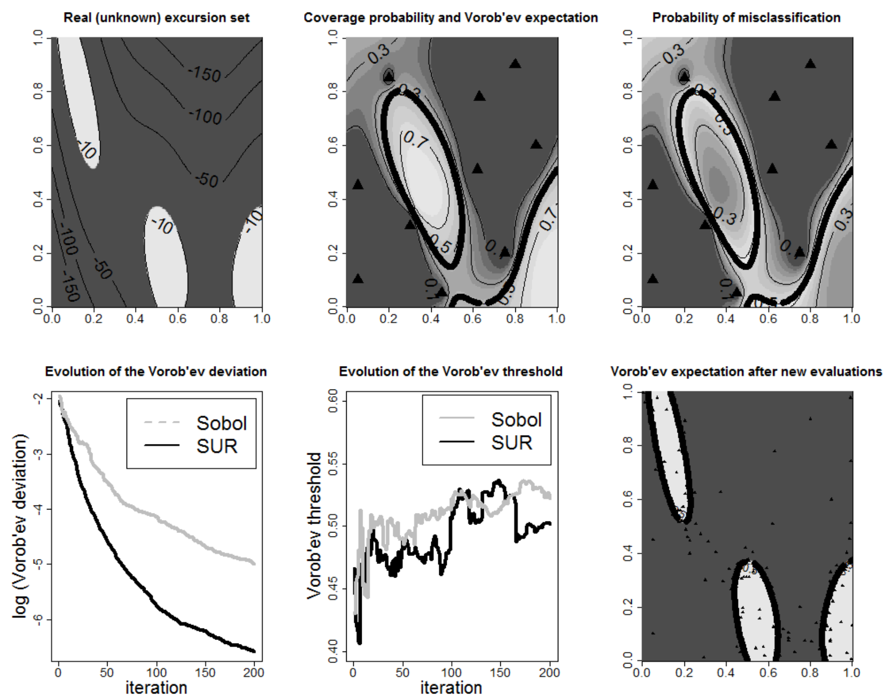

Fig. 2: Top left: Level set of a $2d$ function. Middle: Coverage probability function after 10 evaluations of f. Top right: $\mathbb{E}\left(\mathbb{1}_{Q_{n,\alpha_n^*} \Delta \Gamma}(\cdot)\right)$ function. Bottom left: Decrease of the Vorob’ev deviation when new points are added (2 strategies). Middle: Evolution of α^*. Bottom right: New Vorob’ev expectation (SUR strategy).

On Figure 2 (bottom plots) one can see the evolution of the Vorob’ev deviation and threshold when new points are added. Two different strategies are tested: a simple space filling strategy (with, again, the Sobol’ sequence) and a so-called Stepwise Uncertainty Reduction (SUR) strategy, aiming at reducing the variance of $\mu(\Gamma)$ (see, [2], criterion $J_{4,n}^{\text{SUR}}$, or [3] for more details). We observe that the SUR strategy manages to quickly reduce the Vorob’ev deviation (bottom right plot) and that the Vorob’ev expectation obtained after the new evaluations matches with the true level set. However, note that the consistency of the adaptive approach is not guaranteed by Proposition 1 as the latter only holds for a deterministic space filling sequence. Further research is needed to establish an extension of Proposition 1 to adaptive settings.

5 Conclusion

In this paper we proposed to use random set theory notions, the Vorob'ev expectation and deviation, to estimate and quantify uncertainties on a level set of a real-valued function. This approach has the originality of focusing on the set itself rather than solely on its volume. When the function is actually a GP realization, we proved that the Vorob'ev deviation converges to zero in infill asymptotics, under some mild conditions. However, the final example illustrates that a space-filling approach based on a Sobol' sequence may not be optimal for level set estimation, as it clearly was outperformed by an adaptive strategy dedicated to volume of excursion estimation. In future works, we may investigate sampling criteria and adaptive strategies dedicated to uncertainty reduction in the particular context of set estimation.

References

1. Baddeley, A., Molchanov, I.: Averaging of random sets based on their distance functions. Journal of Mathematical Imaging and Vision **8**(1), 79–92 (1998)
2. Bect, J., Ginsbourger, D., Li, L., Picheny, V., Vazquez, E.: Sequential design of computer experiments for the estimation of a probability of failure. Statistics and Computing **22(3)**, 773–793 (2012)
3. Chevalier, C., Bect, J., Ginsbourger, D., Vazquez, E., Picheny, V., Richet, Y.: Fast parallel kriging-based stepwise uncertainty reduction with application to the identification of an excursion set (2012). URL http://hal.inria.fr/hal-00641108/en
4. Chevalier, C., Picheny, V., Ginsbourger, D.: KrigInv: Kriging-based Inversion for Deterministic and Noisy Computer Experiments (2012). URL http://CRAN.R-project.org/package=KrigInv. R package version 1.3
5. Chilès, J.P., Delfiner, P.: Geostatistics: Modeling Spatial Uncertainty. Wiley, New York (1999)
6. Dubourg, V.: Métamodèles adaptatifs pour l'analyse de fiabilité et l'optimisation sous contrainte fiabiliste. Ph.D. thesis, IFMA (2011)
7. Forrester, A.I.J., Sóbester, A., Keane, A.J.: Engineering design via surrogate modelling: a practical guide. Wiley (2008)
8. Molchanov, I.: Theory of Random Sets. Springer (2005)
9. Picheny, V., Ginsbourger, D., Roustant, O., Haftka, R.T., Kim, N.H.: Adaptive designs of experiments for accurate approximation of target regions. Journal of Mechanical Design **132**(7) (2010)
10. Ranjan, P., Bingham, D., Michailidis, G.: Sequential experiment design for contour estimation from complex computer codes. Technometrics **50**(4), 527–541 (2008)
11. Rasmussen, C.R., Williams, C.K.I.: Gaussian Processes for Machine Learning. MIT Press (2006)
12. Santner, T.J., Williams, B.J., Notz, W.: The Design and Analysis of Computer Experiments. Springer Verlag (2003)
13. Vazquez, E., Piera-Martinez, M.: Estimation of the volume of an excursion set of a Gaussian process using intrinsic Kriging (2006). URL http://arxiv.org/abs/math/0611273
14. Vazquez, E., Piera-Martinez, M.: Estimation du volume des ensembles d'excursion d'un processus gaussien par krigeage intrinsèque. In: 39ème Journées de Statistiques. Angers France (2007)