



HAL
open science

Processus gaussiens déformés pour l'apprentissage de zones instationnaires

Sébastien Marmin, David Ginsbourger, Jean Baccou, Frédéric Perales,
Jacques Liandrat

► **To cite this version:**

Sébastien Marmin, David Ginsbourger, Jean Baccou, Frédéric Perales, Jacques Liandrat. Processus gaussiens déformés pour l'apprentissage de zones instationnaires. 47èmes Journées de Statistique de la SFdS, Jun 2015, Lille, France. hal-01402259

HAL Id: hal-01402259

<https://hal.science/hal-01402259>

Submitted on 24 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PROCESSUS GAUSSIENS DÉFORMÉS POUR L'APPRENTISSAGE DE ZONES INSTATIONNAIRES

Sébastien Marmin ^{1 2 3} & David Ginsbourger ¹ & Jean Baccou ^{2 4}
& Frédéric Perales ^{2 4} & Jacques Liandrat ³

¹ *Universität Bern, IMSV, Alpeneggstrasse 22, CH-3012 Bern, Switzerland.*

`sebastien.marmin@stat.unibe.ch` `david.ginsbourger@stat.unibe.ch`

² *Institut de Radioprotection et de Sûreté Nucléaire (IRSN), PSN-RES/SEMIA, CE
Cadarache, 13114 Saint-Paul-lez-Durance, France.*

`jean.baccou@irsn.fr` `frederic.perales@irsn.fr`

³ *Centrale Marseille, I2M, UMR 7373, CNRS, Université Aix-Marseille,
13453 Marseille, France.*

`jacques.liandrat@centrale-marseille.fr`

⁴ *Laboratoire de Micromécanique et d'Intégrité des Structures, IRSN-CNRS-UMII, B.P.
3, 13115 Saint-Paul-lez-Durance Cedex, France.*

Résumé. En planification d'expériences simulées, le recours à des méthodes de prédiction issues des statistiques spatiales a fait ses preuves pour raffiner l'exploration de l'espace des variables dans des régions d'intérêt. Les modèles considérés ici substituent la fonction inconnue par un processus gaussien interpolant (krigeage). Dans ce contexte, la covariance des processus gaussiens est souvent supposée stationnaire. Cette hypothèse peut conduire à des prédictions décevantes si le comportement de la réponse est très hétérogène en espace. Cette situation est rencontrée dans beaucoup d'études en sûreté nucléaire, ce qui motive cette présentation. La première partie est consacrée à la comparaison et à la validation de modèles sur des données hautement non-stationnaires, provenant d'un code de calcul développé à l'IRSN simulant la fissuration de matériaux hétérogènes et utilisé dans le cadre d'études relative au vieillissement des centrales nucléaires. Deux approches connues de modélisation par processus gaussien non-stationnaires sont abordées : déformation de l'espace des variables et combinaison localisée de plusieurs processus gaussiens. La deuxième partie est dédiée à la planification adaptative d'expériences pour l'exploration de zones de variations brutales. Pour cela, nous définissons et comparons plusieurs critères qui sélectionnent les points à évaluer dans des zones fortement non-stationnaires. Nous concluons en présentant les avancées obtenues sur un cas test en mécanique de la fissuration d'un bimatériau.

Mots-clés. Processus gaussien, noyau non-stationnaire, plan d'expériences adaptatif.

Abstract. When a computer code is expensive to evaluate, it seems essential for the user to look for a design of experiment adapted to its needs, anticipating the response of

the code. To predict the response, one can build a model extrapolating the initial data in unexplored areas. Here we consider models substituting the unknown function by an interpolating Gaussian process (kriging). The covariance function of the Gaussian process is often assumed stationary. This hypothesis can lead to poor predictions if the response behaves heterogeneously across the input space. This kind of behaviour happens in many nuclear-safety studies conducted by IRSN which motivate this work. The first part of the presentation is dedicated to the comparison and validation of non-stationary Gaussian process models. The data comes from a IRSN computer code simulating the fracture dynamics of heterogeneous materials and used to study the mechanical ageing of nuclear power plants. Several known non-stationary modeling approaches are reviewed : source-space warping and space-dependent combinations of different Gaussian processes. The second part is about adaptive design of experiment for the exploration of high variation regions. We define and compare criteria which detect non-stationary zones. In conclusion, we present progresses made on the nuclear-safety test case.

Keywords. Gaussian process, non-stationary kernel, adaptive design of experiment.

1 Modélisation par processus gaussien

Les modèles considérés ici utilisent des processus gaussiens pour approcher une fonction objectif de \mathbb{R}^d dans \mathbb{R} à partir d'un nombre limité d'observations. La fonction objectif est alors prédite par la moyenne conditionnelle du processus gaussien et l'incertitude de prédiction est quantifiable via sa covariance conditionnelle. Dans le cas d'observations non bruitées, la moyenne interpole les observations. Les formules permettant d'obtenir la moyenne et la covariance conditionnelle du modèle sont souvent appelées « équations de krigeage », en référence à la méthode éponyme venue de la géostatistique, et sont rappelées notamment dans Roustant et al. (2012). Dans notre cas test, la fonction objectif est la réponse d'un code de calcul en mécanique de la fissuration, Xper (Perales et al., 2010), utilisé à l'IRSN dans le cadre d'études relatives au vieillissement des centrales nucléaires. Plus précisément, on s'intéresse à la prise en compte dans les simulations Xper des incertitudes de deux paramètres d'entrée du code relatifs à des caractéristiques géométriques et mécaniques d'un matériau hétérogène (typiquement du béton). Dans la suite, ces deux paramètres sont notés x_1 et x_2 et sont redimensionnées entre 0 et 1. 216 observations réparties sur une grille irrégulière sont initialement fournies et représentées sur la figure 1. On observe une zone rectiligne de forte variation. Une première modélisation est faite par un processus gaussien avec une covariance stationnaire. La figure 1 compare la fonction objectif avec la fonction prédite calculée à partir d'un sous-ensemble d'observations appelées données d'apprentissage. Le modèle obtenu reproduit bien la zone de forte variation. Pourtant on observe que la prédiction est éloignée de la véritable fonction dans cette région de transition. On est donc amené à considérer des processus gaussiens non-stationnaires pour la modélisation de notre fonction objectif.

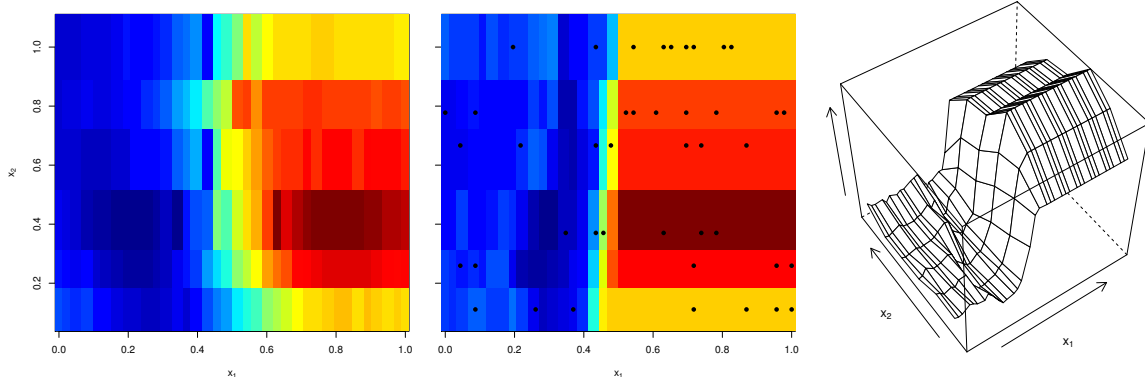


Figure 1 : Comparaison de la prédiction (à gauche) avec la vraie fonction (au milieu et à droite en trois dimensions). Les données d'apprentissage sont marquées d'un point noir.

2 Processus gaussiens déformés (warping)

L'idée de déformer un processus gaussien stationnaire pour le rendre non-stationnaire (e.g. l'approche de Sampson et al. (1992)) est séduisante, parce qu'elle s'interprète intuitivement en terme de transformation de l'espace des variables. En effet, le noyau de covariance $k^{\text{NS}}(\mathbf{x}_1, \mathbf{x}_2)$ du modèle choisi provient du processus $Y_{\mathbf{x}}^{\text{NS}} = Y_{T(\mathbf{x})}^{\text{S}}$, avec $Y_{\mathbf{x}}^{\text{S}}$ un processus stationnaire de covariance $k^{\text{S}}(\mathbf{x}_1, \mathbf{x}_2)$, et T une fonction de déformation d'un domaine de \mathbb{R}^n . La covariance de $Y_{\mathbf{x}}^{\text{NS}}$ est la composée de la covariance stationnaire avec T : $k^{\text{NS}}(\mathbf{x}_1, \mathbf{x}_2) = k^{\text{S}}(T(\mathbf{x}_1), T(\mathbf{x}_2))$. Dans le cas de la *non-linear map method* (méthode de transformation non-linéaire introduite par Gibbs (1997)), la fonction de transformation T est l'intégrale vectorielle d'une fonction de densité $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_n(\mathbf{x}))$. Chaque fonction $g_k(\mathbf{x})$ détermine la déformation locale autour du point \mathbf{x} de l'espace des variables et dirigée par l'axe k : si $g_k(\mathbf{x})$ est inférieur à 1, le processus gaussien est dilaté autour de \mathbf{x} dans la direction k , et si $g_k(\mathbf{x})$ supérieur à 1 il y a un contraction du processus au points \mathbf{x} dans la direction k (voir l'exemple unidimensionnel de la figure 2). Un possible inconvénient de cette méthode est d'avoir besoin d'un nombre important de paramètres à estimer, T étant une fonction multivariée. Une manière de réduire le nombre de paramètres proposée par Ying et al. (2007) est de traiter indépendamment la non-stationnarité suivant chaque coordonnée. Mais cette hypothèse est souvent trop forte. En particulier, elle n'est pas justifiée pour notre cas test de mécanique de la fissuration. L'observation des données (fig 1) nous a donc conduit à considérer une déformation adaptée illustrée par la figure 3. D'une part, pour tenir compte du fait que la forte variation se fait autour d'une droite, nous proposons de faire une transformation localisée autour d'un hyperplan quelconque à estimer. D'autre part, l'alignement des lignes de niveau dans l'espace $0,5 \leq x_1$ nous amène à diriger la déformation par un vecteur fixé ou à estimer.

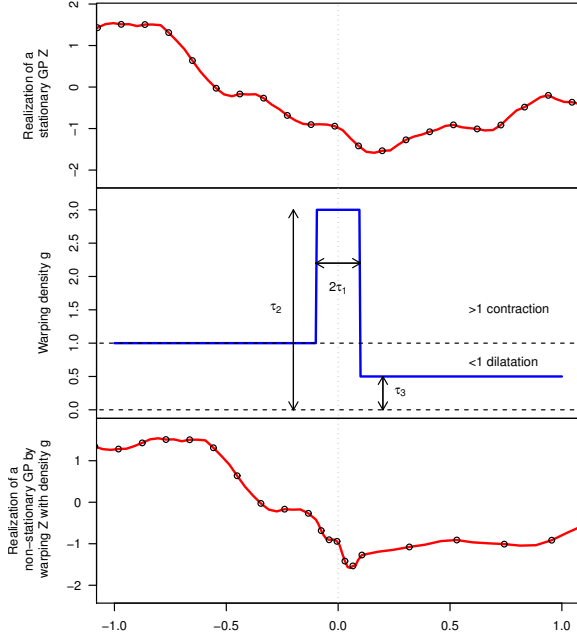


Figure 2 : Exemple unidimensionnel d'une déformation avec sa densité.

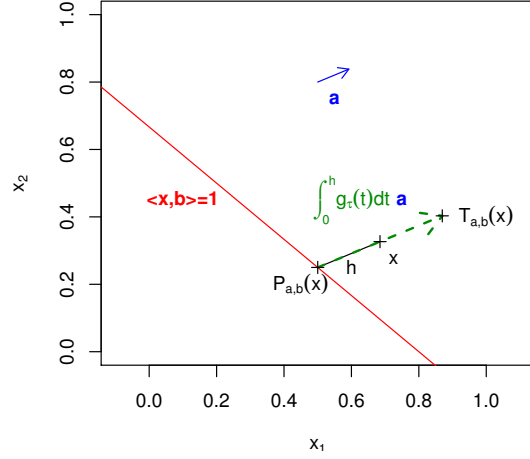


Figure 3 : Schéma de la transformation de l'espace des variables. L'image de \mathbf{x} par T est obtenue par modification de la distance h entre \mathbf{x} et sa projection $P_{\mathbf{a},\mathbf{b}}(\mathbf{x})$ sur l'hyperplan $\langle \mathbf{b}, \mathbf{x} \rangle = 1$ et dirigée par \mathbf{a} . La nouvelle distance est donnée par l'intégrale sur $[0, h]$ d'une fonction g_τ .

Plus précisément, nous écrivons :

$$T(\mathbf{x}) = P_{\mathbf{a},\mathbf{b}}(\mathbf{x}) + \int_0^{(\mathbf{x}-P_{\mathbf{a},\mathbf{b}}(\mathbf{x})) \cdot \mathbf{a}} g_\tau(t) dt \mathbf{a}, \text{ où :} \quad (1)$$

- $P_{\mathbf{a},\mathbf{b}}$ est la projection de direction \mathbf{a} sur l'hyperplan $\langle \mathbf{b}, \mathbf{x} \rangle = 1$, \mathbf{a} et \mathbf{b} vecteurs de \mathbb{R}^d (N.B. : ce n'est en général pas une projection orthogonale).
- g_τ est une fonction positive appelée densité de déformation. Elle est paramétrée par $\tau \in \mathbb{R}^p$, $p \in \mathbb{N}$.

Pendant la phase d'apprentissage, cette déformation permet de capter avec un nombre limité de paramètres des effets de non-stationnarité qui ne sont pas simplement tensoriel. La prédiction de la fonction avec cette méthode est meilleure que la prédiction obtenue avec un processus de covariance stationnaire comme l'illustre la figure 4. On veut maintenant confronter la méthode avec d'autres modèles basés sur des processus gaussiens non-stationnaires.

3 Comparaison de modèles

On compare 7 modèles de trois types différents :

- **Modèle stationnaire** : il correspond à ce qui se fait couramment en modélisation par processus gaussien. La moyenne est supposée constante et la covariance est choisie du type « Matérn-3/2 ». On le note **modèle 1**.
- **Modèle par déformation autour d'un hyperplan** : présenté dans la section précédente, on le décline en quatre variantes selon les choix de la fonction de densité g_τ et du vecteur \mathbf{a} qui sont résumées par ce tableau :

	modèle 2	modèle 3	modèle 4	modèle 5
g_τ	type exponentielle	const. par morceau	type exponentielle	const. par morceau
\mathbf{a}	ortho. à $\langle \mathbf{x}, \mathbf{b} \rangle = 1$	ortho. à $\langle \mathbf{x}, \mathbf{b} \rangle = 1$	ortho. à l'axe 2	ortho. à l'axe 2

- **Modèle par combinaison de processus stationnaire indépendants** : le processus de modélisation $Y(\mathbf{x})$ s'écrit comme la somme de deux processus gaussiens $Z_{\mathbf{x}}^{(1)}$ et $Z_{\mathbf{x}}^{(2)}$ pondérée par une fonction de localisation $p(\mathbf{x})$ qui vaut 1 dans un domaine et 0 dans son complémentaire : $Y_{\mathbf{x}} = \mu + p(\mathbf{x})Z_{\mathbf{x}}^{(1)} + (1 - p(\mathbf{x}))Z_{\mathbf{x}}^{(2)}$. Le **modèle 6** suppose $Z_{\mathbf{x}}^{(1)}$ et $Z_{\mathbf{x}}^{(2)}$ centrés alors que le **modèle 7** autorise $Z_{\mathbf{x}}^{(1)}$ à avoir une moyenne constante non-nulle.

Les performances de chaque modèle sont évaluées sur le cas test IRSN à partir d'un ensemble d'apprentissage comportant 43 points. Afin de garantir la robustesse des conclusions, l'étude numérique est répétée 10 fois en faisant varier aléatoirement l'ensemble d'apprentissage. Il apparaît alors que les écart-types des estimateurs des paramètres des modèles représentent entre 15 et 30 % de leurs valeurs moyennes. De plus, le graphe 5 donne, pour les 7 modèles, l'erreur quadratique entre les prédictions et les données réelles issues du code de calcul. Il montre, pour notre cas test, que les modèles avec déformation de processus gaussien sont les plus performants.

4 Planification d'expériences séquentielle

Une planification d'expériences séquentielle est une méthode d'évaluation de la fonction objectif qui dépense le budget d'observations de manière itérative. On utilise alors la moyenne et la covariance du modèle dans la formulation d'un critère qui juge les choix des prochains points d'évaluation. Nombreux sont les critères utilisés en plan d'expérience séquentielle, le plus connu étant celui qui revient à réduire l'intégrale de la variance du modèle, l'IMSE (Sacks (1989)). Dans le cas d'un modèle stationnaire, le critère de l'IMSE considère toutes les zones inexplorées indépendamment des valeurs de la fonction objectif. Cependant dans notre cas test, les zones d'intérêt sont celles qui sont à la fois inexplorées (qui présentent une variance importante) et fortement non-stationnaires/variationnelles (e.g. qui présentent un fort gradient). Nous montrons comment les différents modèles non-stationnaires permettent de privilégier l'exploration de zones non-stationnaires. La comparaison de critères faisant un compromis entre exploration globale de la fonction et exploration des zones non-stationnaires sera également abordée.

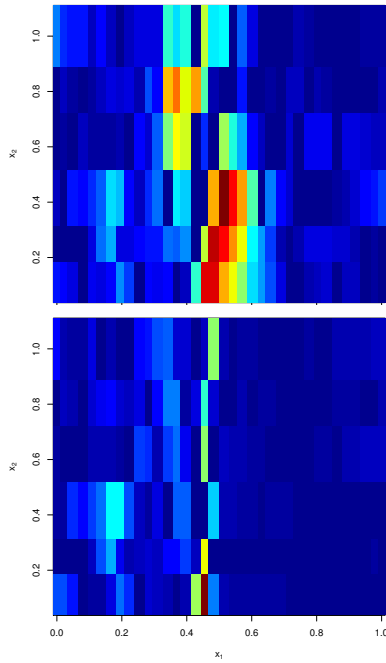


Figure 4 : Erreur de prédiction par le modèle stationnaire (haut) versus par le modèle par déformation (bas).

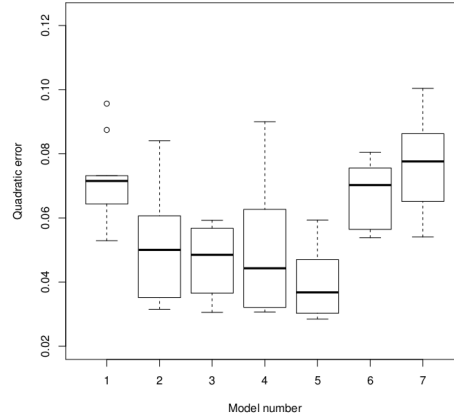


Figure 5 : Erreurs quadratiques des 7 modèles sur 10 ensembles d'apprentissage. 1 : modèle stationnaire, 2 à 5 : différents modèles avec déformations (voir section 2), 6 à 7 : modèles par combinaison de processus gaussien.

Bibliographie

- [1] Roustant O, Ginsbourger D, Deville Y (2012), DiceKriging, DiceOptim : Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization, *Journal of Statistical Software*, vol. 51, issue 1.
- [2] Perales F, Dubois F, Monerie Y, Piar B, Stainier L (2010). A NonSmooth Contact Dynamics-based multi-domain solver. Code coupling (Xper) and application to fracture. *Eur. J. Comp. Mech.* 19, 389–417.
- [3] Sampson PD, Guttorp P (1992), Nonparametric Estimation of Nonstationary Spatial Covariance Structure, *Journal of the American Statistical Association*, 87(417), 108–119.
- [4] Gibbs M (1997), Bayesian Gaussian processes for regression and classification, Ph.D. thesis, University of Cambridge.
- [5] Ying X, Wei C, Daniel A, and Xuru D (2007), A non-stationary covariance-based Kriging method for metamodeling in engineering design, *International Journal for Numerical Methods in Engineering*, (71), 733–756.
- [6] Sacks J, Schiller S, and Welch W (1989), *Designs for Computer Experiments*, *Technometrics*, 31, 41–47.