
RAPPORT SUR LA THÈSE DE SÉBASTIEN MARMIN

E. Le Pennec

La thèse de S. Marmin intitulée *Warping and Sampling Approaches to Non-Stationary Gaussian Process Modeling* a pour thème l'utilisation de processus Gaussien à matrice de covariance non stationnaire pour interpoler des fonctions. Les contributions de S. Marmin s'articulent autour de deux directions : une extension du modèle classique par une déformation (*warping*) de l'espace avant la modélisation Gaussienne et une utilisation des modèles gaussiens (warpés ou non) pour de l'échantillonnage adaptatif.

Le manuscrit de 148 pages rédigées en anglais est très agréable à lire. Il est organisé en cinq chapitres : une courte introduction suivie d'un état de l'art sur les processus Gaussiens (non stationnaires) et les plans d'expériences adaptatifs, un chapitre sur les processus Gaussiens déformés, un autre sur les questions d'échantillonnage et, enfin, un dernier chapitre présentant des résultats numériques. Le manuscrit se termine par une conclusion, deux annexes techniques ainsi qu'une bibliographie.

L'introduction présente de manière très synthétique le contexte de la thèse ainsi que son contenu. S. Marmin justifie à travers un exemple la nécessité de procédure d'interpolation prenant en compte l'hétérogénéité des variations des fonctions. Il explique comment un modèle de déformation combiné à des processus Gaussiens plus classiques peut permettre cela. Il décrit ensuite l'utilisation

qu'il va faire de ces modèles Gaussiens dans un cadre de design adaptatif où l'on choisit les points d'échantillonnage en fonction de l'incertitude de l'estimation courante de la fonction. Il termine cette introduction de 8 pages par le plan de son manuscrit.

Le deuxième chapitre de 33 pages présente un panorama très clair des processus Gaussiens et de leur utilisation pour le design d'expérience. Le ton est volontairement pédagogique et permet de bien se familiariser avec le champ et ses notations. Ce chapitre se termine par le problème de la non stationnarité des variations et de sa prise en compte par paramétrisation des matrices de covariances (y compris par partitionnement) ou par déformations de l'espace.

Cette dernière approche est celle poursuivie par S. Marmin dans son troisième chapitre de 19 pages. Il y présente son modèle de déformation, des transformations unidimensionnelles après application d'une transformée linéaire, ainsi que les propriétés du Processus Gaussien correspondant. Il montre ainsi que sous des hypothèses raisonnables ce processus a des propriétés de régularité suffisantes. Il donne ensuite des exemples de réalisation d'un tel modèle. La dernière section est consacrée à un algorithme numérique d'estimation de déformation basée sur une transformation en ondelettes. On peut regretter que cette section paraisse un peu indépendantes du reste du chapitre : le modèle (non paramétrique) de déformation mais sans application d'une transformée linéaire est très différent du modèle présenté en début de chapitre où des déformations paramétrées par des lois beta sont appliquées à une transformation linéaire des données. Ce léger défaut n'enlève rien au reste du chapitre et la description de l'estimation numérique des paramètres du modèle paramétrique est très bien faite.

Le quatrième chapitre de 25 pages est consacré à l'utilisation des processus Gaussiens pour construire des plans d'expériences adaptatifs. Il débute par la proposition de se concentrer sur les zones d'incertitude sur la norme du gradient plutôt que de la fonction elle-même. S. Marmin propose des critères basés sur des variances de puissance de la norme du gradient et donne des formules explicites permettant de calculer ces quantités quand la puissance vaut 1 ou 2. Ce chapitre se termine par une contribution très fine à un problème de recherche de minimum de fonction à l'aide d'un design adaptatif. Il donne des approximations très efficaces de l'espérance de diminution du minimum vu du processus pour tout choix de nouveaux points. Ces approximations sont suffisamment efficaces pour permettre d'aborder numériquement des problèmes de choix de design intractables autrement. Les résultats les plus techniques concer-

nant cette approximation sont repoussés dans un très intéressant appendice.

Le dernier chapitre de 25 pages correspond à une mise en oeuvre des méthodes proposées. Les nombreuses expériences montrant l'intérêt des outils introduits par S. Marmin sont menées avec soin. La description des cas et des algorithmes utilisés est très précise et les résultats sont analysés avec finesse. Il se poursuit par une conclusion reprenant les points essentiels de la thèse et donnant quelques perspectives.

L'ensemble des travaux présenté par S. Marmin est conséquent et de très bonne qualité. S. Marmin maîtrise les outils (fins) mathématiques nécessaires à ses constructions ainsi que les outils informatiques permettant de les implémenter. Pour toutes ces raisons, il ne fait aucun doute que sa soutenance doit être autorisée.



Erwan LE PENNEC



Department of Statistics (MC0439)
Hutcheson Hall
Virginia Polytechnic & State University
250 Drillfield Drive
Blacksburg, VA 24061

November 13, 2017

Doctoral Committee
Institut de Mathématiques de Marseille
c/o Sarah Mostefa

DOCTORAL THESIS REPORT FOR SÉBASTIAN MARMIN

Dear Committee,

I am pleased to provide this report in support of Sébastien Marmin's Ph.D. dissertation and to recommend that he be allowed to progress to the defense stage. The document was an enjoyable read, and I think represents a substantial body of scholarly work containing valuable and novel contributions to the literatures of nonstationary response surface modeling, modeling of computer simulation experiments, sequential design of such experiments, and their "blackbox" (Bayesian) optimization. I believe the work to be technically sound and well-presented, and will ultimately comprise of some very well-received journal publications. Below I shall outline what I see to be the main contributions, first at a high level and then chapter-by-chapter, with commentary on how the novel methodology described therein augments (and compares favorably to) the current state-of-the art.

The main premise motivating the work in this dissertation is that many response surfaces, whether arising from computer simulations or real-data (e.g., spatially referenced) collection, violate the stringent assumptions underling the canonical models in the literature: in particular stationarity. Although remedies abound, e.g., generalizing processes to the nonstationary setting, many exacerbate an already daunting set of computational obstacles present in the simpler (stationary) setup. Marmin's methodological contributions here target a tactic called "warping", which is a sensible idea for obtaining nonstationary flexibility, however at potentially great computational expense. By limiting to an affine structure, via single- (and ultimately multiple-) index modeling from the econometrics literature, he is able to dramatically reduce many of the computational demands while retaining, for the most part, the framework's modeling flexibility. After illustrating these methods, and comparing to some of the other more computationally tractable options from the literature, including divide-and-conquer approaches, he moves on to sequential design and optimization. The thinking here is that a model which adequately captures nonstationary dynamics will be more "data hungry", especially in parts of the input space where the response dynamics are more challenging to model: something a stationary model could not "realize", let alone accommodate. In the case of optimization, Marmin provides some cost-savings techniques for batch- q expected improvement-like methods which are

promising from a computational perspective. These enhancements are more generic, although they also work with the proposed nonstationary modeling methodology.

After a survey of previous work and some of the mathematical and computational tools that are relevant to the proposed methodology and applications (which is very good but I shall not elaborate on further here), Marmin begins outlining his modeling contributions in Chapter 3. These center around a practical approach for nonstationary nonlinear process modeling via Gaussian processes on warped input spaces. These were perhaps first (most successfully) introduced in the 1990s by Sampson and Guttorp. The trouble with the canonical setup here is that computational costs involved render practical application intractable on all but the smallest data sets. Whereas ordinary (stationary) GP inference and prediction typically involves computation in $\mathcal{O}(N^3)$, for N input–output pairs, methods based on warping can be upwards of $\mathcal{O}(N^4)$, although perhaps they not typically characterized in that way. Marmin’s proposed WaMI-GP drastically simplifies the apparatus while retaining many of its essential features, at least in terms of modeling fidelity. After introducing the main ingredients, the chapter discusses a litany of technical results, illustrating desirable theoretical properties, and concludes with a demonstration on synthetic data. That empirical work showcases flexibility in an illustrative context, and moreover quantifies that predictions derived from WaMI-GP predictors outperform both richer (i.e., more cumbersome computationally) and cruder (i.e., computationally thriftier) alternatives. For a special 1- d input case Marmin also provides a nonparametric warping mechanism via Wavelets (Wav-GP), with many attractive technical properties. These are further explored (with potential in higher dimension) within the empirical work of Chapter 5.

Chapter 4 contains methodological contributions on sequential design and optimization. This is in two parts, and both represent substantial (and novel) enhancements to the literature. The first involves the development of a new “active learning” heuristic with the aim of focusing sampling on the part of the input space where the response is changing most quickly, i.e., where the estimated derivative is high (in absolute value). This is in contrast to previous methods which focus on (reducing) variance, possibly integrated over the input space. Marmin provides a variation targeting gradients, utilizing the properties of derivatives of GP predictive surfaces, with point-wise and integrated alternatives. In a number of controlled settings, it would seem that the proposed heuristic has the desired effect. Having such a capability is particularly important for nonstationary models which may have regions in the input space which are changing more rapidly than others. The second part of the chapter involves a new set of calculations for evaluating a recently proposed batch expected improvement criteria called q -EI. Although q -EI has been shown to work well relative to many comparators, calculating and optimizing the criteria has been too cumbersome for many practical applications. Marmin has developed workarounds for both aspects (evaluating and optimizing) which offer orders of magnitude speedup.

Chapter 5 contains a slew of empirical results on synthetic and real data applications, offering a stress test and relative comparison to competing methodology. The real-data both come from engineering contexts: one from nuclear safety (via IRSN) and another from aeronautics (via NASA). These are challenging examples, and the proposed nonstationary modeling methods (WaMI-GP and Wav-GP) and sequential design (and optimization) heuristics therein perform well relative to comparators, including the method of treed Gaussian processes via a popular

R package on CRAN. The results here are promising indeed, and will serve as a nice basis for empirical sections of (perhaps) several papers supporting the methodological contributions introduced in earlier chapters.

In closing, my opinion of this work is strongly favorable. I believe it meets—indeed, well exceeds—the threshold for advancing to defense and ultimately to the awarding the degree of Ph.D. at any of the top academic institutions world wide. The work is innovative, well documented, and will have substantial impact once converted into journal/conference publication form. I congratulate the candidate, and his advising team, and wish him the best of luck going forward in his (surely promising) academic career.

Please do not hesitate to contact me via telephone or email if I can be of any further assistance.

Best regards, Bobby

A handwritten signature in black ink, appearing to read "Bobby B. Gramacy". The signature is fluid and cursive, with the first name "Bobby" and last name "Gramacy" clearly distinguishable.

Robert B. Gramacy
Professor of Statistics
rbg@vt.edu