Non-stationary Gaussian process modelling and sequential design of experiments for exploration of high variation regions

S. Marmin
University of Bern, IRSN Cadarache and École Centrale Marseille

Supervisor(s): Dr. David Ginsbourger (Idiap Research Institute and University of Bern), Prof. Jacques Liandrat (École Centrale Marseille), Dr Jean Baccou (Institut de Radioprotection et de Sûreté Nucléaire)

Ph.D. expected duration: 2014-2017

Adress: IRSN/PSN-RES/SEMIA/LIMAR, Centre de Cadarache, 13115 St Paul-Lez-Durance, France

Email: sebastien.marmin@irsn.fr

Abstract:

In the context of expensive deterministic simulations, Gaussian process (GP) models surrogating the response of the code have become a standard device for sequential design of experiments. The objective function f is seen as a black box and assumed to be one realization of a GP with given mean and covariance. Based on such a GP model, sampling criteria dedicated to various aims (approximation, optimization, set estimation, etc.) involving f, can be worked out and used to appropriately choose next evaluation points. State of the art criteria for approximation notably include MSE and IMSE [5]. Further criteria dedicated to localizing optima, and also to the estimation of target regions have been proposed [2]. Here we focus on the case of functions possessing heterogeneous variations across the input space. First we propose a sequential design approach dedicated to learning high variation regions: the Gaussian process model is initially assumed stationary, and non-stationary is progressively introduced through warpings built upon prediction errors (see figure 1 for an example of a warped GP model).

Second we compare the use of different infill criteria for this method, including recently introduced IWS derivative-based infill criteria. They are inspired by MSE and IMSE, but they focus on gradients under the GP model and attempt to balance the exploration of unknown regions with the refinement of high variation regions.

The non-stationary modelling is motivated by observations on data simulating the fracture dynamics of two-phase materials, from a IRSN computer code XPER [3]: different cracks propagation types are determined by the inputs, implying heterogeneous behaviours in the parameter space (see figure 2). We finally present the progresses made on this test case.

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

References

[1] Sébastien Marmin, David Ginsbourger, Jean Baccou, Frédéric Perales, and Jacques Liandrat. Processus gaussiens déformés pour l'apprentissage de zones instationnaires. In *47èmes journées de la Société Française de Statistique, Lille, France*, 2015. http://papersjds15.sfds.asso.fr/submission_125.pdf.

[2] J. Mockus. *Bayesian Approach to Global Optimization. Theory and Applications.* Kluwer Academic Publisher, Dordrecht, 1989.

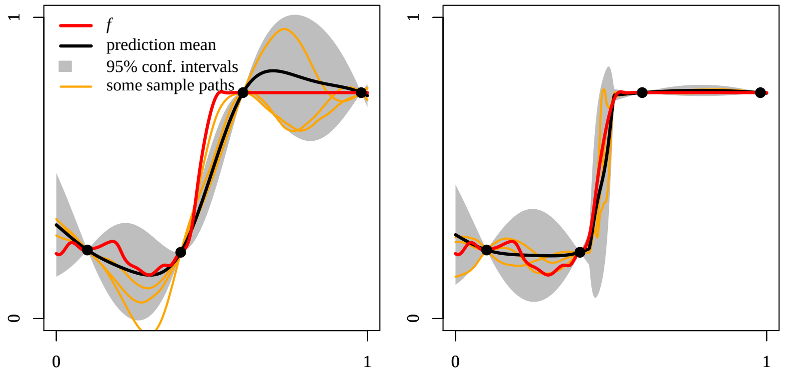

Figure 1: Example of stationary (left) and non-stationary (right) modelling.

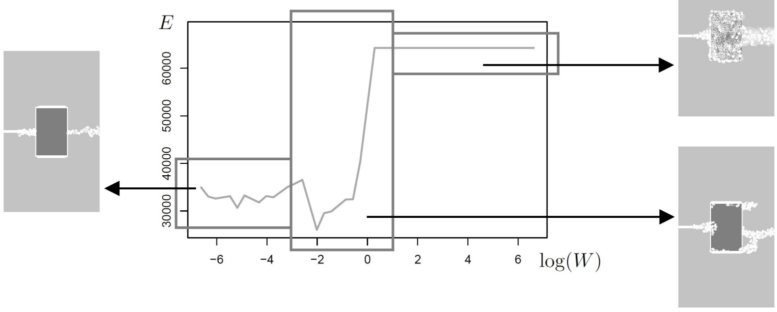

Figure 2: One dimensional section of a response surface (cracking energy over material parameters). The input space is divided into regions where the material cracking behaves differently.

[3] F Perales, F Dubois, Y Monerie, B Piar, and L Stainier. A NonSmooth Contact Dynamics-based Multi-domain Solver. Code coupling (Xper) and application to fracture. *European Journal of Computational Mechanics*, 19:389–417, 2010.

[4] O. Roustant, D. Ginsbourger, and Y. Deville. DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by Kriging-Based Metamodelling and Optimization. *Journal of Statistical Software*, 51 (1):1–55, 2012.

[5] Jerome Sacks, William J Welch, Toby J Mitchell, and Henry P Wynn. Design and analysis of computer experiments. *Statistical science*, pages 409–423, 1989.

[6] PD Sampson and P Guttorp. Nonparametric Estimation of Nonstationary Spatial Covariance Structure. *Journal of the American Statistical Association*, 87(417):108–119, 1992.

[7] Michael Scheuerer. *A comparison of models and methods for spatial interpolation in statistics and numerical analysis*. PhD thesis, Georg-August-Universität Göttingen, 2009.

Short biography – I have studied engineering and applied mathematics at École des Mines de Saint-Étienne. I am doing a joint PhD between University of Bern and École Centrale Marseille. The PhD is funded by the French Nuclear Safety Institute (IRSN). After a year at the Institute of Mathematical Statistics and Actuarial Science of Bern University (IMSV), I am now continuing my research activities in LIMAR, a laboratory of the IRSN (Cadarache, France).