MASCOT-NUM 2022
Clermont Ferrand, France
7-9 June 2022

Relaxed Gaussian process interpolation: a goal-oriented approach to Bayesian optimization

S. J. Petit
Safran Aircraft Engines (SAE)
Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes (L2S)

Supervisor(s): J. Bect (L2S), J. Demange (SAE), S. Da Veiga (Safran Tech), E. Vazquez (L2S)

PhD expected duration: Apr. 2019 - Mar. 2022

Address: 3 rue Joliot-Curie, 91190, Gif-sur-Yvette

E-mail: sebastien.julien.petit@gmail.com

Abstract:

Gaussian process (GP) interpolation and regression [see, e.g., 1, 2] is a very classical method for predicting an unknown function from data. It has found applications in Bayesian optimization, a popular derivative-free global optimization technique for "expensive-to-evaluate" functions.

A GP model is defined by a mean and a covariance functions, which are generally selected from data within parametric families. The most popular models assume stationarity and rely on standard covariance functions such as the Matérn covariance. The assumption of stationarity yields models with relatively low-dimensional parameters. However, such a hypothesis can sometimes result in poor modeling when functions to be predicted have different scales of variation or different local regularities across the domain.

This is the case, for instance, in the motivating example given by [3] or in the even simpler toy example shown in Figure 1. This function, which we shall call the Steep function, has a clear global minimum around the point $x = 8$, which is overshadowed by the variations on the left. A stationary GP fit is shown in Figure 2 (a): observe that the confidence bands are too large and that the conditional mean varies too much in the neighborhood of the global minimum. One expects Bayesian optimization techniques to be somehow inefficient on this problem.

Going beyond the stationary hypothesis has been a prevailing direction of research. One can distinguish two categories of approaches that all use stationary Gaussian processes as a core building block: local models [see, e.g., 3, 4] and composition of models [see, e.g., 5, 6]. The gain in expressiveness has usually the counterpart of an increasing number of parameters to be estimated from data, relying on advanced inference techniques, or leaving to the user the choice of key parameters.

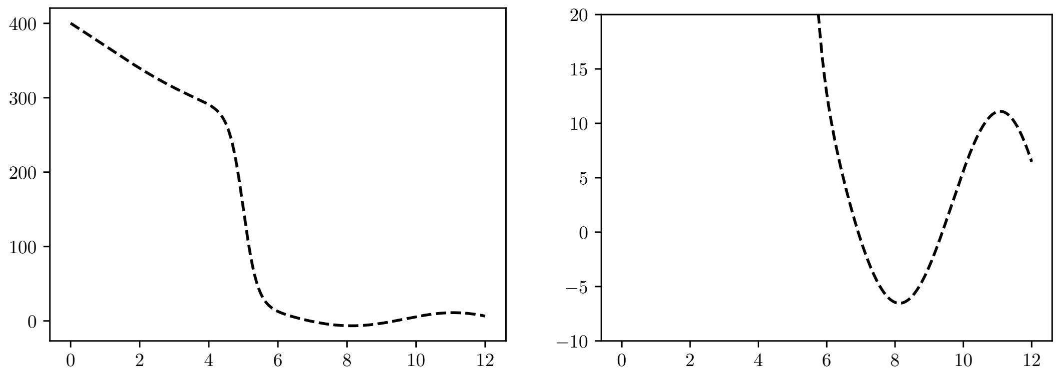
In this work, we suppose that there are regions of interest in the output range. We propose to explore goal-oriented GP modeling, where we want predictive models in regions of interest, even if it means being less predictive elsewhere. The proposed method is based on the relaxation of the interpolation constraint outside the region of interest. This yields goal-oriented interpolation models designed to be accurate for a subset of the range of the function, which are very beneficial for Bayesian optimization, as in Figure 2 (b).

This work presents three main contributions. First, we propose new models called *Relaxed Gaussian processes* (GP-R) for goal-oriented modeling with noiseless observations. Secondly, we give theoretical and empirical results justifying the method and its use for Bayesian optimization. Finally, we propose a new scoring rule called *truncated continuous ranked probability score* (tCRPS) designed for the problem of goal-oriented modeling.

Joint work with J. Bect and E. Vazquez.

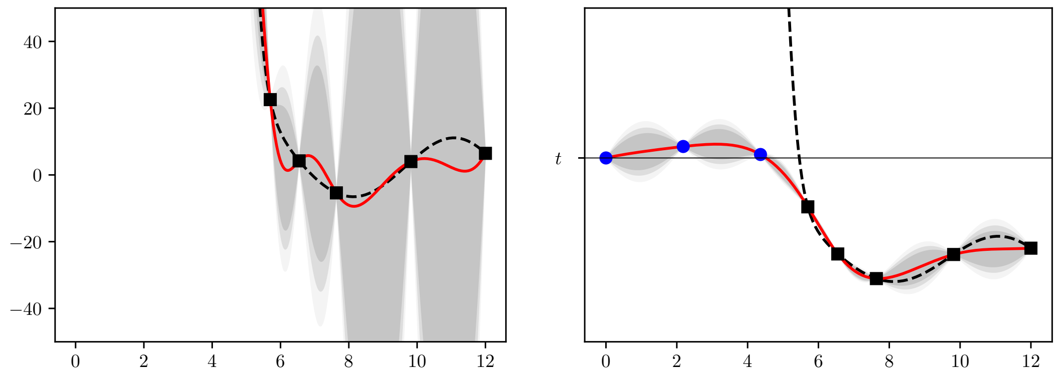
References

[1] M. L. Stein. *Interpolation of Spatial Data: Some Theory for kriging.* Springer New York, 1999.



(a) (b)

Figure 1: We show (a): the Steep function, and b): also the Steep function, but with a restriction on the y-axis for illustration purposes.



(a) A stationary GP fit (b) A GP-R fit

Figure 2: We show (a): a vanilla stationary Gaussian process fit on the Steep function based on some observations (black squares), and (b): a GP-R fit; the quantity t is a relaxation threshold that is chosen automatically. The red lines are the conditional means and the gray bands represent the uncertainty estimates.

[2] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, USA, 2006.

[3] R. B. Gramacy and H. K. H. Lee. Bayesian treed Gaussian process models with an application to computer modeling. *Journal of the American Statistical Association*, 103(483):1119–1130, 2008.

[4] C. E. Rasmussen and Z. Ghahramani. Infinite mixtures of Gaussian process experts. *Advances in neural information processing systems*, 2:881–888, 2002.

[5] I. Rychlik, P. Johannesson, and M. R. Leadbetter. Modelling and statistical analysis of ocean-wave data using transformed Gaussian processes. *Marine Structures*, 10(1):13–47, 1997.

[6] A. Damianou and N. D. Lawrence. Deep Gaussian processes. In *Artificial intelligence and statistics*, pages 207–215. PMLR, 2013.

Short biography – With a background in applied mathematics, Sébastien began his PhD in April 2019 with L2S and Safran Aircraft Engines. This PhD thesis is funded by Safran Aircraft Engines, under the scope of a CIFRE agreement, and aims at further developing Safran Aircraft Engines *Bayesian optimization* tools.