Gaussian process metamodeling for functional-input coastal flooding code

J. Betancourt
Toulouse Mathematics Institute, University Paul Sabatier, France

Supervisor(s): Prof. Thierry Klein (Toulouse Mathematics Institute), Prof. François Bachoc (Toulouse Mathematics Institute)

Ph.D. expected duration: Aug. 2017 - Sep. 2020

Adress: Toulouse Mathematics Institute, 118 Route de Narbonne, 31400 Toulouse

Email: jbetanco@math.univ-toulouse.fr

Abstract:

Floods in general affect more people than any other hazard with 1.5 billion people affected in the last decade of the 20th century. Many recent events (e.g.: Katrina, USA 2005; Xynthia, France 2010) illustrate the complexity of coastal systems and the limits of traditional forecast and early warning systems and flood risk analysis. Recent scientific progresses now allow properly modeling coastal flooding events. Such models are nevertheless very expensive in terms of computation time (multiple hours) which prevents any use for forecast and warning. A widely used method to approach this type of limitation is to build a reduced model (often called surrogate model or metamodel), able to provide high precision estimates of the response surface at an acceptable computation time.

This study was developed in the frame of the ANR RISCOPE project, which studies the development of metamodels for coastal flooding early warning [1]. In the coastal flooding context, the metamodel should be able to deal with functional inputs associated to time varying maritime conditions such as the tide and surge. Among all the types of metamodels available (polynomials, splines, neural networks, etc.), we focus on Kriging (also called Gaussian process model), characterized by its mean and covariance functions [6, 5]. The main advantage of Kriging is its ability to provide both a prediction of the computer code and the uncertainty attached to this prediction [3]. Kriging metamodels were originally developed for scalar inputs, however, they can be also built for functional inputs. To this end, we project each functional input on a functional basis, and then we use the projection as inputs of the metamodel. In this study we compare results of metamodels based on B-splines [2] and PCA [4] projection methods, as well as two different forms to meassure the distance among functional basis within the covariance function: (i) taking the coefficients of the functional decomposition as independent scalar inputs; and (ii) using an adapted distance for functional decompositions. We illustrate a procedure for identification of relevant functional inputs for the metamodel. We further discuss two approaches to tune the dimension of the projections: i) based on the error of the projection; ii) based on the performance of the metamodel. Our results show that the approach based on the error of the projection, being the most widely used in the literature, may lead to unnecessarily large projection dimensions. In contrast, the approach based on metamodel performance presents the virtue of directly pointing to the final objective of building a fast and accurate metamodel. All codes were implemented in R and the metamodel was validated through a case study based on real data gathered at Gâvres coast in France.

References

[1] Anr riscope project. `https://perso.math.univ-toulouse.fr/riscope/`. Accessed: 2018-12-04.

[2] Thomas Muehlenstaedt, Jana Fruth, and Olivier Roustant. Computer experiments with functional inputs and scalar outputs by a norm-based approach. *Statistics and Computing*, 27(4):1083–1097, 2017.

[3] Simon Nanty, Céline Helbert, Amandine Marrel, Nadia Pérot, and Clémentine Prieur. Sampling, metamodeling, and sensitivity analysis of numerical simulators with functional stochastic inputs. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):636–659, 2016.

[4] Simon Nanty, Céline Helbert, Amandine Marrel, Nadia Pérot, and Clémentine Prieur. Uncertainty quantification for functional dependent random variables. *Computational Statistics*, 32(2):559–583, 2017.

[5] Carl Edward Rasmussen and Christopher KI Williams. Gaussian processes for machine learning. 2006. *The MIT Press, Cambridge, MA, USA*, 38:715–719, 2006.

[6] Thomas J Santner, Brian J Williams, and William I Notz. *The design and analysis of computer experiments*. Springer Science & Business Media, 2013.

Short biography – José Betancourt is currently a PhD student at the Toulouse Mathematics Institute. His doctoral thesis is part of the ongoing ANR project RISCOPE, which focuses on the development of metamodeling techniques for coastal flooding early warning. His responsibility in the project lies in the frame of objective-based dimension reduction & meta-modeling for functional-input hydrodynamic codes.