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

# Multi-fidelity surrogate modelling for time-series output

B. Kerleguer
*CEA, DAM, DIF and Ecole Polytechnique*

**Supervisor(s):** C. Cannamela (CEA, DAM, DIF) and J. Garnier (Ecole Polytechnique)

**PhD expected duration:** Sep. 2019 - Sep. 2022

**Address:** Bruyères-le-Châtel, 91297 Arpajon Cedex

**E-mail:** baptiste.kerleguer@polytechnique.edu

**Abstract:**

We focus on multi-fidelity hierarchical computer codes, which are codes modelling the same phenomenon, but which can be hierarchically sorted according to their accuracy and numerical cost. To improve the high-fidelity surrogate model prediction we want using information from low-fidelity data. Standard techniques for surrogate modelling have been extended to multi-fidelity framework via co-kriging models. Some numerical codes may, however, have high-dimension outputs, such as time series, and this is the topic of this talk.

We have first proposed a method that deals with dimension reduction and time-dependent kernel in co-kriging [2]. The code output is expanded onto a basis built from the experimental design. The first coefficients of the expansion of the code output are processed by a co-kriging approach. The last coefficients are collectively processed by a kriging approach with covariance tensorization. The prediction uncertainty is fully taken into account with this method. This method is more efficient than simple dimension reduction on an example set.

Second we have developed our own regression method using Gaussian processes and Bayesian neural networks. Indeed, methods based on neural networks have proven to be effective even in the context of small amounts of data [3]. However, the goal is to have a quantification of prediction uncertainty. We have therefore proposed an approach combining Bayesian neural networks (BNN) and Gaussian processes, called GP-BNN method. The Gaussian process emulates the low-fidelity code. The BNN emulates the high-fidelity code with inputs from the code and the low-fidelity Gaussian process. We have investigated several methods to transfer the output information of the low-fidelity surrogate model to the high-fidelity BNN emulator. It turns out that the best method is based on a quasi Monte Carlo sample. The method is found to be effective for non-linear interactions between codes, and reproduces the results of the AR(1) auto-regressive method [1] for the linear case.

Finally, we have constructed a method based on wavelet decomposition and Gaussian process regression. We use the correlation properties of the wavelet coefficients under prior assumptions about the output regularity. We have to deal with a large number of points in the space of wavelet coefficients. We condition the Gaussian process in the wavelet space only to those points whose realisations are significantly Larger from the prior. We use the AR(1) model with a Gaussian process in a multi-scale space for multi-fidelity. An advantage of this method is that it is adaptable to several types of high-dimensional outputs. Moreover, wavelets are suitable for dimension reduction in many applications, which is our goal here.

## References

[1] Loic Le Gratiet and Josselin Garnier. Recursive co-kriging model for design of computer experiments with multiple levels of fidelity. *International Journal for Uncertainty Quantification*, 4(5):365–386, 2014.

[2] Baptiste Kerleguer. Multi-fidelity surrogate modeling for time-series outputs. *SIAM/ASA Journal on Uncertainty Quantification*, Submitted December 2021.

[3] Xuhui Meng and George Em Karniadakis. A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse pde problems. *Journal of Computational Physics*, 401:109020, 2020.

**Short biography** – Baptiste Kerleguer got a Master's Degree in applied mathematics from Ecole Normale Supérieure Paris-Saclay. His thesis, funded by CEA DAM, focuses on surrogate models with functional inputs and outputs for the analysis and quantification of uncertainties in complex models.