Contribution ID: **50** Type: **not specified**

# Multi-fidelity Gaussian processes for noisy outputs and non-nested experiment designs: a comparison between the recursive and non-recursive formulations

Surrogate models provide fast approximations of computationally expensive simulations (or experiments) and are trained using a limited set of observations generated by these codes. In the multi-fidelity framework, we assume the availability of two computer models with different levels of cost and accuracy. The high-fidelity model $z_H$ provides the most accurate predictions but is also the most expensive to evaluate. In contrast, the low-fidelity model $z_L$ is significantly cheaper to compute but offers lower accuracy. We focus in this work on the auto-regressive model which supposes a linear relation between the two codes: $\forall \mathbf{x} \in \mathbb{R}^D$, $z_H(\mathbf{x}) = \rho(\mathbf{x}) \cdot z_L(\mathbf{x}) + \delta_H(\mathbf{x})$, where $\rho$ is the scaling factor and the functions $z_L$ and $\delta_H$ are approximated with Gaussian processes.

This model was initially developed by Kennedy and O'Hagan and improved afterwards by Le Gratiet and Garnier with the more computationally efficient recursive formulation. However, these works rely on two important assumptions: first, that the observed outputs are deterministic; and second, that the experimental designs are nested, meaning that each high-fidelity input point coincides with a low-fidelity input point. Under these assumptions, the optimization of the model parameters is significantly simplified.

We generalize the recursive formulation to the case of noisy outputs and non-nested designs. An alternative optimization approach based on the expectation-maximization algorithm is compared to the direct maximum likelihood estimation for the initial, non-recursive formulation. We apply both formulations of the auto-regressive model to several cases of varying difficulty and show that the proposed approach achieves faster training times for large low-fidelity datasets.

## Special/ Invited session

## Classification

Both methodology and application

## Keywords

Surrogate modeling, Auto-regressive co-kriging, Expectation-Maximization

**Primary author:** BAILLIE, Nils (Université Paris-Saclay, CEA, Service d'Etudes Mécaniques et Thermiques, 91191 Gif-sur-Yvette, France)

**Co-authors:** Dr KERLEGUER, Baptiste (CEA, DAM, DIF, F-91297, Arpajon, France); Dr FEAU, Cyril (Université Paris-Saclay, CEA, Service d'Etudes Mécaniques et Thermiques, 91191 Gif-sur-Yvette, France); Prof. GARNIER, Josselin (CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France)

**Presenter:** BAILLIE, Nils (Université Paris-Saclay, CEA, Service d'Etudes Mécaniques et Thermiques, 91191 Gif-sur-Yvette, France)

**Track Classification:** Uncertainty quantification and computer experiments