AI safety for complex systems – 2-Years Post-doctoral Proposal for 2023

« Deep learning methods with Bayesian-based uncertainty quantification for the emulation of CPU-expensive numerical simulators »

The use of numerical simulators with a high level of resolution, so-called high fidelity simulators, allows modeling that is more and more faithful to reality but also more and more expensive to evaluate, which sometimes limits the direct and intensive use of these simulators, especially for studies of the propagation of uncertainties affecting the model's input variables. To overcome this limitation, the use of substitute mathematical models, called **metamodels or emulators**, has become a valuable and indispensable tool ([1,2]). The objective is to replace a physico-numerical model by a statistical (or machine) learning model that is inexpensive to evaluate and which best approximates the phenomenology of the initial model. This metamodel is trained (i.e. fitted) on a set of available simulations of the model (black-box approach). **The data-driven metamodeling of numerical simulators mainly relies on machine learning (ML) algorithms**, which is a vast and very active field of research, motivated by strong application issues of many industrial sectors (aerospace, transportation, energy, environment, etc.).

Among the usual ML methods commonly used, we can cite chaos polynomials, neural networks, Gaussian process regression or kernel ridge regression ([3,4]). **Gaussian process (GP) metamodels** ([5,6]) have been of particular interest to the computer experiment community since they propose both a prediction and an uncertainty on the output, which is very appealing in a **context of safety studies or risk assessments**. Moreover, they have shown efficient and relevant results on many industrial applications, especially in the case of a few tens of uncertain inputs with a few hundreds of model simulations, or even a few thousands ([7]). However, GP metamodels rely on a *prior* specification of the mean and covariance structure. Parametric models are generally assumed whose parameters have to be estimated on the training data.

However, in some applications, these **GP parametric metamodels can show some limitations**, especially in the case of very irregular functions and/or with strong non-stationarities, or stationary but piecewise. Recent work has focused on the interest of **Bayesian-based deep learning approaches** such as ***Bayesian neural networks*** (**BNN**, see [8,9,10]) or ***deep GP*** (**DGP**, see [11, 12, 13]) to address the applicative limitations of shallow GP metamodels, while providing as output a predictive distribution and not only a simple prediction (predictive distribution resulting from *posterior* inference). **These elements revive the interest of deep learning methods for our applications.**

The objective of the post-doctorate will therefore be **to study the applicability and potential of deep learning approaches for the emulation of expensive numerical simulators** by addressing the following issues:

- **Study the tractability of BNN and DGP-based metamodels**: how to train these models in a robust way when only a few hundred to a few thousand simulations of the model are available (which is the case, for example, for most of our thermal-hydraulic studies for safety studies on PWR or fast RNR-Na reactors)?
- **Evaluate their benefit compared to classical shallow learning methods**, for our applications in nuclear safety studies.
- **Assess the reliability of the uncertainty associated with the deep metamodel predictions.** Is the associated posterior error reliable? We will be able to deploy all our expertise and tools in the validation of the predictive law for the simple GP, to these deep models.

Duration

Post-doctorate of 12 months extendable 12 months (in agreement with all the parties involved).

Location:

CEA Cadarache Center, 13108 Saint-Paul-Lez-Durance, France.

Formation and Skills

PhD in applied mathematics, specialized in probabilities and statistics, with experience in Machine Learning methods and software.
Programming skills: Python, eventually R or Matlab. Ability to publish.

Contacts

Amandine MARREL[1] : amandine.marrel [at] cea.fr
Georis BILLO[1] : georis.billo [at] cea.fr
[1]CEA Cadarache, CEA/DES/IRESNE/DER/SESI/LEMS, 13108 Saint-Paul-lez-Durance

Main references

[1] Cheng et al. *Surrogate-assisted global sensitivity analysis: an overview*. Struct Multidisc Optim 61, 1187–1213, 2020.

[2] Kianifar & Campean. *Performance evaluation of metamodelling methods for engineering problems: towards a practitioner guide*. Struct Multidisc Optim 61, 159–186, 2020.

[3] Fang et al. *Design and modeling for computer experiments.* Computer Science & Data Analysis Series. Chapman& Hall/CRC, 2006.

[4] De Lozzo. *Substitution de modèle et approche multifidélité en expérimentation numérique. Journal de la société française de statistique*, Tome 156 no. 4, pp. 21-55, 2015.

[5] Rasmussen & Williams. *Gaussian processes for machine learning*. MIT Press, 2006.

[6] Gramacy. Surrogates: *Gaussian Process Modeling, Design and Optimization for the Applied Sciences*, Chapman Hall/CRC, 2020.

[7] Marrel et al. *The ICSCREAM Methodology: Identification of Penalizing Configurations in Computer Experiments Using Screening and Metamodel - Applications in Thermal Hydraulics*. Nuclear Science and Engineering, 196(3):301–321, 2022.

[8] Lampinen & Vehtari. *Bayesian approach for neural networks-review and case studies*. Neural Networks, vol. 14, 257 – 274, 2001.

[9] Titterington. *Bayesian methods for neural networks and related models.* Statist. Sci., vol. 19, no. 1, pp. 128–139, 2004.

[10] Goan & Fookes, *Bayesian Neural Networks: An Introduction and Survey. Springer International Publishing*, 45–87, 2020.

[11] Damianou & Lawrence. *Deep Gaussian Processes, in Artificial Intelligence and Stats*, pp. 207–215, 2013.

[12] Sauer et al. *Active Learning for Deep Gaussian Process Surrogates*, Technometrics, 1–15, 2022.

[13] Hebbal et al. N. *Bayesian Optimization using Deep Gaussian Processes with Applications to Aerospace System Design,* Optimization and Engineering, 22, 321–361, 2021.