Kriging-based prediction of probability measures

Application in numerical simulation for nuclear safety studies

Florian Gossard

2nd year PhD student

Institut de Mathématiques de Toulouse

Workshop Avignon 2025

Supervisors : J. Baccou, F. Bachoc, T. Le Gouic, J. Liandrat

Introduction

- The study of complex industrial systems, such as a nuclear reactor, requires the use of complex codes (software)

- These codes are often computationally intensive

- The amount of simulated data may be insufficient to accurately study a phenomenon across its entire domain of interest

Introduction

- The study of complex industrial systems, such as a nuclear reactor, requires the use of complex codes (software)
- These codes are often computationally intensive
- The amount of simulated data may be insufficient to accurately study a phenomenon across its entire domain of interest

Goal : Build a fast-evaluating model (metamodel) to predict the output and thus create new data for the study of the phenomenon

Ordinary kriging in $\mathbb{R}$ [3]

In this case, the computation code can be written as an unknown function:

$$f : \mathscr{D} \subset \mathbb{R}^d \longrightarrow \mathbb{R}$$

and we consider the observations of f as realizations of a **spatially related** random process $\{\mathbf{Y}(x), x \in \mathscr{D}\}$. We denote by $(y(x_1), \ldots, y(x_n))$ the observations.

Assumptions

- isotropic
- stationary
- unknown constant mean

BLUP

Given $(\mathbf{Y}(x_1),\dots,\mathbf{Y}(x_n))$ coming from the stochastic process, the ordinary Kriging estimator of $\mathbf{Y}$ at a new point $x^\star \in \mathscr{D}$ is the Best Linear Unbiased Predictor (BLUP) written as:

$$\hat{\mathbf{Y}}(x^\star) = \sum_{i=1}^{n} \bar{\lambda}_i \mathbf{Y}(x_i) \tag{1}$$

where

$$\bar{\lambda} = \operatorname{argmin}_{\lambda=(\lambda_1,\dots,\lambda_n)} \left\{ \mathbb{E}\left[\left| \mathbf{Y}(x^\star) - \hat{\mathbf{Y}}_\lambda(x^\star) \right|^2 \right], \sum_{i=1}^{n} \lambda_i = 1 \right\}, \tag{2}$$

An estimation of $y(x^\star)$ denoted $\hat{y}(x^\star)$ is therefore given by (1) when replacing $\mathbf{Y}(x_i)$ by $y(x_i)$. This estimation can also be interpreted as a barycenter:

$$\hat{y}(x^\star) = \operatorname{argmin}_{y\in\mathbb{R}} \left\{ \sum_{i=1}^{n} \bar{\lambda}_i |y(x_i) - y|^2 \right\}. \tag{3}$$

Semivariogram

The spatial correlation can be obtained by estimating the semivariogram:

$$\gamma(\|x - x'\|) = \frac{1}{2}\mathbb{E}\left[\left\|\mathbf{Y}(x) - \mathbf{Y}(x')\right\|^2\right]$$

Experimental semivariogram:

$$\gamma_{exp}(h) = \frac{1}{2\,Card(N(h))} \sum_{(m,n)\in N(h)} \|\mathbf{Y}(x_m) - \mathbf{Y}(x_n)\|^2$$

with $N(h) = \{(m,n) \in \{0,\dots,N-1\}^2,\ h-\epsilon \le ||x_m - x_n||_2 \le h+\epsilon\}$ (ϵ depends on the problem)

Matern semivariogram models

The candidates for fitting models are classic semivariogram models like Matern functions:

$$\gamma_{\sigma,l,\nu}(h) = \sigma^2\left(1 - \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\sqrt{2\nu}\frac{h}{l}\right)^{\nu} K_\nu\left(\sqrt{2\nu}\frac{h}{l}\right)\right)$$

where :

- σ^2 is the standard deviation,
- ν is the smoothness parameter,
- l is the length-scale parameter,
- K_ν is a modified Bessel function,
- $\Gamma(\nu)$ is the gamma function.

Classical Matern functions are given for $\nu \in \{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}\}$

Matern semivariogram models

The candidates for fitting models are classic semivariogram models like Matern functions:

$$\gamma_{\sigma,l,\nu}(h) = \sigma^2\left(1 - \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\sqrt{2\nu}\frac{h}{l}\right)^{\nu} K_\nu\left(\sqrt{2\nu}\frac{h}{l}\right)\right)$$

where :

- σ^2 is the standard deviation,
- ν is the smoothness parameter,
- l is the length-scale parameter,
- K_ν is a modified Bessel function,
- $\Gamma(\nu)$ is the gamma function.

Classical Matern functions are given for $\nu \in \{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}\}$

Optimisation methods

Least squared, Cross Validation,...

By introducing a Lagrange multiplier α, we can show that $\bar{\lambda}$ is the solution of the following system:

$$\begin{bmatrix} \Sigma & \mathbb{1}_n \\ \mathbb{1}_n^T & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \alpha \end{bmatrix} = \begin{bmatrix} k^\star \\ 1 \end{bmatrix}$$

where Σ is the $n \times n$ matrix with $\Sigma_{i,j} = \gamma(\|x_i - x_j\|)$, $k^\star$ is the column vector of size n with $(k^\star)_i = \gamma(\|x_i - x^\star\|)$ and $\mathbb{1}_n$ is the column vector of ones.

Ordinary Kriging in $\mathscr{P}_2(\mathbb{R})$ [6]

$$\mathscr{P}_2(\mathbb{R}) = \left\{ \mu \text{ proba measure} \mid \int_{\mathbb{R}} x^2 \, d\mu(x) < \infty \right\}$$

Consider

$$f : \mathscr{D} \subset \mathbb{R}^d \to \mathscr{P}_2(\mathbb{R})$$

We extend the barycenter construction from the real case

$$\hat{\mu}(x^\star) = \operatorname{argmin}_{\mu \in \mathscr{P}_2(\mathbb{R})} \left\{ \sum_{i=1}^{n} \bar{\lambda}_i W_2^2(\mu(x_i), \mu) \right\}, \tag{4}$$

where:

$$\bar{\lambda} = \operatorname{argmin}_{\lambda=(\lambda_1,\ldots,\lambda_n)} \left\{ \mathbb{E}\left[W_2\left(\mu(x^\star), \hat{\mu}(x^\star)\right)^2 \right], \sum_{i=0}^{N-1} \lambda_i = 1 \right\}. \tag{5}$$

Wasserstein distance[7]

Second order Wasserstein distance:

$$W_2(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \left(\int_{\mathbb{R}\times\mathbb{R}} \|x-y\|^2 \, d\pi(x,y) \right)^{\frac{1}{2}},$$

In 1D ($\mathscr{P}_2(\mathbb{R})$), the Wasserstein distance of order 2 can be written:

$$W_2(\mu,\nu) = \left(\int_0^1 \left| F^{-1}(t) - G^{-1}(t) \right|^2 dt \right)^{\frac{1}{2}},$$

where F^{-1} and G^{-1} are the quantile functions of μ and ν

Predictor [2]

These statements allow us to introduce a new linear predictor based on quantile functions.

$$\hat{\mathbf{Q}}_{\mu(x^\star)} = \sum_{i=1}^{n} \bar{\lambda}_i \mathbf{Q}_{\mu(x_i)} \tag{6}$$

where

$$\bar{\lambda} = \text{argmin}_{\lambda=(\lambda_1,\dots,\lambda_n)} \left\{ \mathbb{E}\left[\int_0^1 \left(\mathbf{Q}_{\mu(x^\star)}(\xi) - \hat{\mathbf{Q}}_{\lambda,\mu(x^\star)}(\xi) \right)^2 d\xi \right], \sum_{i=1}^{n} \lambda_i = 1 \right\}, \tag{7}$$

Experimental semivariogram for probability measures:

$$\gamma_{exp}^{W}(h) = \frac{1}{2Card(N(h))} \sum_{(i,j)\in N(h)} \left[\int_0^1 \left(Q_{\mu(x_i)}(\xi) - Q_{\mu(x_j)}(\xi) \right)^2 d\xi \right], \tag{8}$$

where $N(h)$ is as in the real case.

Cross validation[1]

- Limitation of the semivariogram: its estimation becomes unreliable when based on a limited number of observations.

Cross validation[1]

- Limitation of the semivariogram: its estimation becomes unreliable when based on a limited number of observations.

- We can estimate the model parameters by cross validation with the LOO MSE criterion :

$$\mathrm{MSE}_{\mathrm{LOO}} = \frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left(Q_{\mu(x_i)}(\xi) - \hat{Q}^{(-i)}_{\mu(x_i)}(\xi)\right)^2 d\xi. \tag{9}$$

.

- Extension of virtual cross validation formulas for quantile functions.

Extension of virtual cross validation formulas [4]

Proposition

Let $\mathbf{Y}$ *be a stochastic process with values in* $\mathscr{P}_2(\mathbb{R})$ *with unknown constant mean, its semivariogram is denoted by* γ^W. *Let* $\mathbf{Y}(x_1), \dots, \mathbf{Y}(x_n)$ *be observations of the process,* $(Q_1, \cdots, Q_n)$ *the quantile functions associated with the observation and* $(\Sigma_W)_{i,j} = \gamma^W(\|x_i - x_j\|)$. *Then* $\forall \xi \in [0,1]$ *and* $\forall i \in \{1, \cdots, n\}$ *we have:*

$$Q_i(\xi) - \hat{Q}_i(\xi) = \sum_{j=1}^{n} \frac{\tilde{\Sigma}_{ij}}{\tilde{\Sigma}_{i,i}} Q_j(\xi) \tag{10}$$

with:

- $\tilde{\Sigma} = \Sigma_W^{-1} - \Sigma_W^{-1} \mathbb{1}_n (\mathbb{1}_n^t \Sigma_W^{-1} \mathbb{1}_n)^{-1} \mathbb{1}_n^t \Sigma_W^{-1}$
- $\hat{Q}_i$ *is the estimator of* Q_i *based on all the other observations.*

Application in reflooding studies for nuclear safety

Loss of primary coolant accident

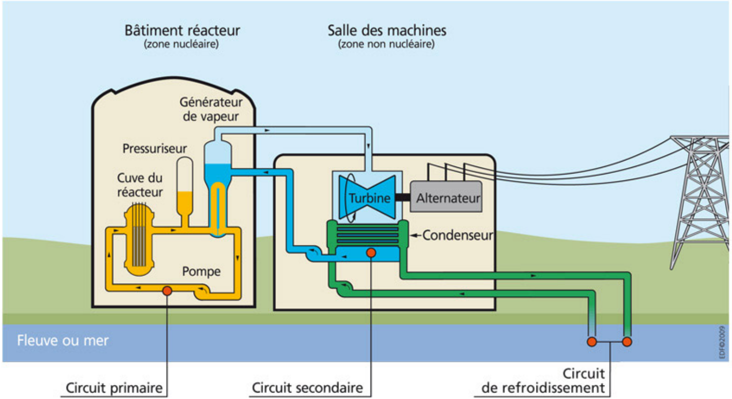

Figure 1: Scheme of a nuclear plant

Application in reflooding studies for nuclear safety

DRACCAR : Déformation et Renoyage d'un Assemblage de Crayon de Combustibles pendant un Accident de Refroidissement (ASNR software)

Figure 2: DRACCAR Process

Quantities of interest

- Average temperature
- 95% quantile
- Entire distribution (error measured with the wasserstein distance)

Models

Model based on ordinary kriging in $\mathbb{R}$ (prediction of the map then computation of the quantities of interest)

- Principal Component Analysis over the 100 discretization points of the maps [5]
- Kriging on the first three components
- Method for semivariogram estimation: Max likelihood under Gaussian assumption (model 1)

Models

Models based on ordinary kriging in $\mathscr{P}_2(\mathbb{R})$ (prediction of the distribution)

- Transform temperature maps into histograms
- Kriging based on quantile functions
- Method for semivariogram estimation:
  - Least squared with empirical semivariogram and positives weights (model 3)
  - Least squared with empirical semivariogram and no constrains on the weights (model 4)
  - Cross validation (model 5)

Results

- Model 1 : Ordinary kriging in $\mathbb{R}$
  variogram parameters optimisation : Max likelihood
- Model 2 : Ordinary kriging in $\mathscr{P}_2(\mathbb{R})$
  variogram parameters optimisation : Least squared + $\lambda > 0$
- Model 3 : Same + no constraints on λ
- Model 4: Odinary kriging + Cross validation

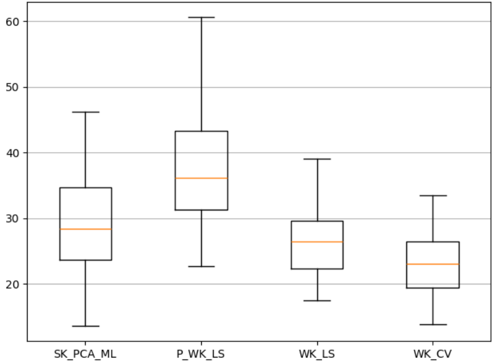

Figure 3: Boxplots of $\text{RMSE}_{\text{mean}}$ for each model.

Results

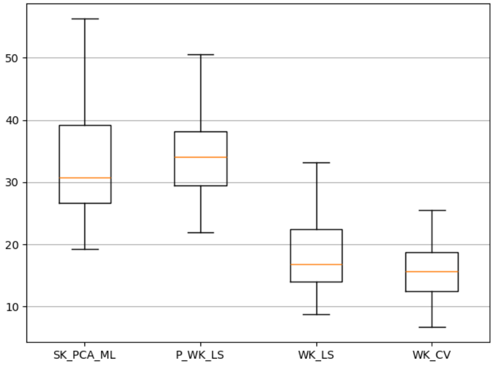

Figure 4: Boxplots of RMSE_{q95} for each model.

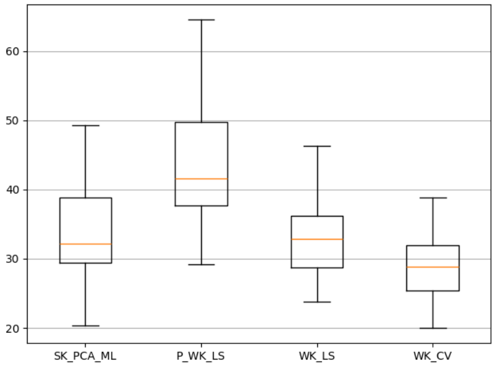

Figure 5: Boxplots of RMSE_W for each model.

Conclusion

- We proposed an extension of kriging for probability measures
- We also extended the virtual cross validation formulas for quantile functions
- These methods produce better results on the prediction of statistical parameters in thermohydraulic studies

Perspectives

- Consider anisotropic models
- Work on a new set of data
- Implement 2D Wasserstein Barycenters

References

[1] François Bachoc. Cross validation and maximum likelihood estimations of hyper-parameters of gaussian processes with model misspecification. *Computational Statistics & Data Analysis*, 66:55–69, 2013.

[2] Antonio Balzanella and Antonio Irpino. Spatial prediction and spatial dependence monitoring on georeferenced data streams. *Statistical Methods and Applications*, 29:101–128, 3 2020.

[3] Noel A.C. Cressie. Statistics for spatial data revised edition. *Statistics for Spatial Data*, pages 1–900, 4 2015.

[4] Florian Gossard, François Bachoc, Jean Baccou, Thibaut Le Gouic, Jacques Liandrat, and Tony Glantz. Kriging measure-valued data with sparse observations: application to nuclear safety studies. working paper or preprint, October 2025.

[5] Amandine Marrel, Bertrand Iooss, Michel Jullien, Béatrice Laurent, and Elena Volkova. Global sensitivity analysis for models with spatially dependent outputs global sensitivity analysis for models with spatially dependent outputs global sensitivity analysis for models with spatially dependent outputs. *Environmetrics*, 22:383–397, 2011.

[6] Alessandra Menafoglio and Piercesare Secchi. Statistical analysis of complex and spatially dependent data: A review of object oriented spatial statistics. *European Journal of*