Parameter control in the presence of uncertainties

Robust Estimation of Bottom friction

Victor Trappler victor.trappler@univ-grenoble-alpes.fr
É. Arnaud, L. Debreu, A. Vidard
AIRSEA Research team (Inria) team.inria.fr/airsea/en/
Laboratoire Jean Kuntzmann

Applied Inverse Problems, 12/07/2019

Introduction

Uncertainties and errors are introduced at each stage of the modelling, by simplifications, parametrizations…

In the end, we have a set of parameters we want to calibrate, but how can we be sure that this calibration is acting upon the errors of the modelling, and does not compensate the effect of the natural variability of the physical system?

Outline

Deterministic problem

Computer code and inverse problem

Input
- $\boldsymbol{k}$: Control parameter
- $\boldsymbol{u}$: Environmental variables (fixed and known)

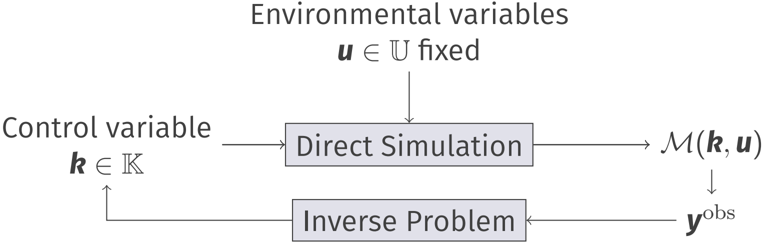
Output
- $\mathcal{M}(\boldsymbol{k}, \boldsymbol{u})$: Quantity to be compared to observations

$$\begin{array}{ccccc} & & \text{Environmental variables} & & \\ & & \boldsymbol{u} \in \mathbb{U} \text{ fixed} & & \\ & & \downarrow & & \\ \text{Control variable} & & & & \\ \boldsymbol{k} \in \mathbb{K} & \longrightarrow & \boxed{\text{Direct Simulation}} & \longrightarrow & \mathcal{M}(\boldsymbol{k}, \boldsymbol{u}) \end{array}$$

Computer code and inverse problem

Input
- $\boldsymbol{k}$: Control parameter
- $\boldsymbol{u}$: Environmental variables (fixed and known)

Output
- $\mathcal{M}(\boldsymbol{k}, \boldsymbol{u})$: Quantity to be compared to observations



We have $\boldsymbol{y}^{\mathrm{obs}} = \mathcal{M}(\boldsymbol{k}_{\mathrm{obs}}, \boldsymbol{u}_{\mathrm{obs}})$ with $\boldsymbol{u}_{\mathrm{obs}} = \boldsymbol{u}$

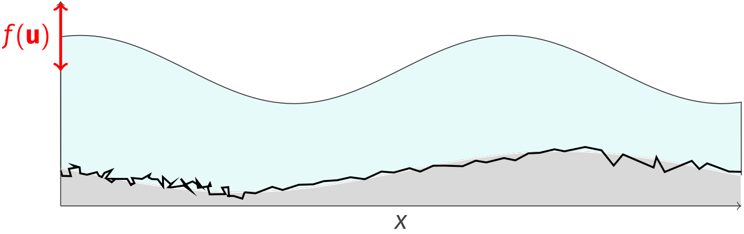
$$\hat{\boldsymbol{k}} = \underset{\boldsymbol{k}\in\mathbb{K}}{\arg\min}\, J(\boldsymbol{k}) = \underset{\boldsymbol{k}\in\mathbb{K}}{\arg\min}\, \frac{1}{2}\|\mathcal{M}(\boldsymbol{k}, \boldsymbol{u}) - \boldsymbol{y}^{\mathrm{obs}}\|^2$$

- $\rightarrow$ Deterministic optimization problem
- $\rightarrow$ Possibly add regularization
- $\rightarrow$ Classical methods: Adjoint gradient and Gradient-descent

BUT

- What if $\boldsymbol{u} \neq \boldsymbol{u}_{\mathrm{obs}}$?
- Does $\hat{\boldsymbol{k}}$ compensate the errors brought by this misspecification?

- The friction $\boldsymbol{k}$ of the ocean bed has an influence on the water circulation
- Depends on the type and/or characteristic length of the asperities
- Subgrid phenomenon
- $\boldsymbol{u}$ parametrizes the BC



Dealing with uncertainties

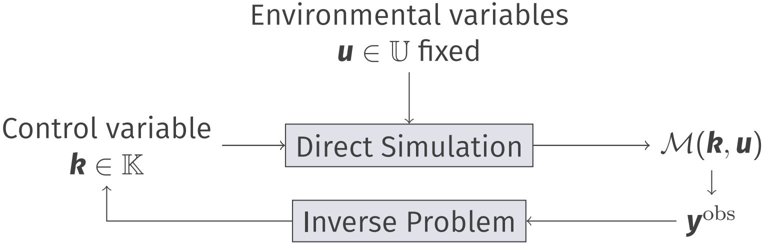
Epistemic or aleatoric uncertainties? [WHR+03]

- Epistemic uncertainties: From a lack of knowledge, that can be reduced with more research/exploration
- Aleatoric uncertainties: From the inherent variability of the system studied, operating conditions

$\rightarrow$ But where to draw the line?
Our goal is to take into account the aleatoric uncertainties in the estimation of our parameter.

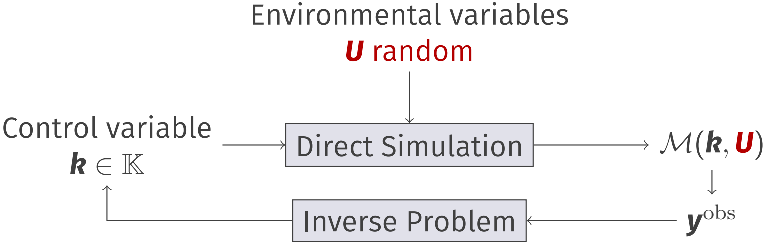
Aleatoric uncertainties

Instead of considering $\boldsymbol{u}$ fixed, we consider that $\boldsymbol{U}$ is a random variable (pdf $\pi(\boldsymbol{u})$), and the output of the model depends on its realization.



Aleatoric uncertainties

Instead of considering $\boldsymbol{u}$ fixed, we consider that $\boldsymbol{U}$ is a random variable (pdf $\pi(\boldsymbol{u})$), and the output of the model depends on its realization.



- Output of the computer code ($\boldsymbol{u}$ is an input):

$$\mathcal{M}(\boldsymbol{k}, \boldsymbol{u})$$

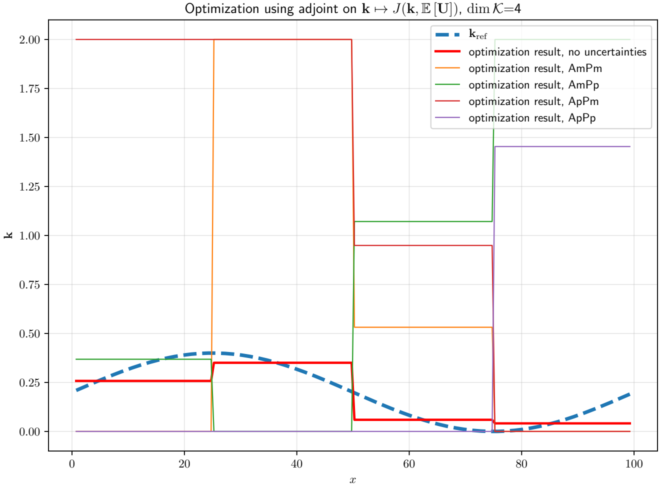
- The (deterministic) quadratic error is now

$$J(\boldsymbol{k}, \boldsymbol{u}) = \frac{1}{2}\|\mathcal{M}(\boldsymbol{k}, \boldsymbol{u}) - \boldsymbol{y}^{\mathrm{obs}}\|^2$$

"$\hat{\boldsymbol{k}} = \arg\min_{\boldsymbol{k}\in\mathbb{K}} J(\boldsymbol{k}, \boldsymbol{u})$" but what can we do about $\boldsymbol{u}$?

Toy Problem: Influence of misspecification of $\boldsymbol{u}_{\mathrm{obs}}$

Minimization performed on $\boldsymbol{k} \mapsto J(\boldsymbol{k}, \mathbb{E}[\boldsymbol{U}])$, for different $\boldsymbol{u}_{\mathrm{obs}}$: Naïve approach



- Main objectives:
  - Define criteria of robustness, based on $J(\boldsymbol{k}, \boldsymbol{u})$, that will depend on the final application
  - For each criterion, be able to compute an estimate $\hat{\boldsymbol{k}}$ in a reasonable time
- Questions to be answered along the way:
  - Good exploration of $\mathbb{U}$, based on the density of $\boldsymbol{U}$ (Design of Experiment: LHS, Monte-Carlo, OA,… ?)
  - Deal with dimension of $\mathbb{K}$?

Robust minimization

Robust minimization

Criteria of robustness

- Worst case [MWP13]:

$$\min_{\boldsymbol{k}\in\mathbb{K}} \left\{ \max_{\boldsymbol{u}\in\mathbb{U}} J(\boldsymbol{k}, \boldsymbol{u}) \right\}$$

- M-robustness [LSN04]:

$$\min_{\boldsymbol{k}\in\mathbb{K}} \mathbb{E}_{\boldsymbol{U}} \left[J(\boldsymbol{k}, \boldsymbol{U}) \right]$$

- V-robustness [LSN04]:

$$\min_{\boldsymbol{k}\in\mathbb{K}} \mathbb{V}\mathrm{ar}_{\boldsymbol{U}} \left[J(\boldsymbol{k}, \boldsymbol{U}) \right]$$

- Multiobjective [Bau12]:

Pareto frontier

- Best performance attainable for each configuration $\boldsymbol{u} \sim \boldsymbol{U}$

Main idea: For each $\boldsymbol{u} \sim \boldsymbol{U}$, compare the value of the cost function to its optimal value $J^*(\boldsymbol{u})$ and define $\boldsymbol{k}^*(\boldsymbol{u}) = \arg\min_{\boldsymbol{k}\in\mathbb{K}} J(\boldsymbol{k},\boldsymbol{u})$

"Most Probable Estimate", and relaxation

Main idea: For each $\boldsymbol{u} \sim \boldsymbol{U}$, compare the value of the cost function to its optimal value $J^*(\boldsymbol{u})$ and define $\boldsymbol{k}^*(\boldsymbol{u}) = \arg\min_{\boldsymbol{k}\in\mathbb{K}} J(\boldsymbol{k}, \boldsymbol{u})$ The minimizer as a random variable:

$$\boldsymbol{K}^* = \underset{\boldsymbol{k}\in\mathbb{K}}{\arg\min}\, J(\boldsymbol{k}, \boldsymbol{U})$$

$\longrightarrow$ estimate its density (how often is the value $\boldsymbol{k}$ a minimizer)

$$p_{\boldsymbol{K}^*}(\boldsymbol{k}) = "\,\mathbb{P}\left[J(\boldsymbol{k}, \boldsymbol{U}) = J^*(\boldsymbol{U})\right]\,"$$

Main idea: For each $\boldsymbol{u} \sim \boldsymbol{U}$, compare the value of the cost function to its optimal value $J^*(\boldsymbol{u})$ and define $\boldsymbol{k}^*(\boldsymbol{u}) = \arg\min_{\boldsymbol{k}\in\mathbb{K}} J(\boldsymbol{k}, \boldsymbol{u})$ The minimizer as a random variable:

$$\boldsymbol{K}^* = \arg\min_{\boldsymbol{k}\in\mathbb{K}} J(\boldsymbol{k}, \boldsymbol{U})$$

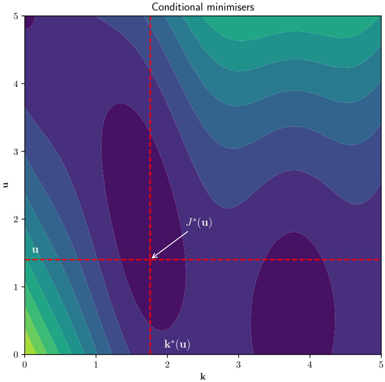
$\longrightarrow$ estimate its density (how often is the value $\boldsymbol{k}$ a minimizer)

$$p_{\boldsymbol{K}^*}(\boldsymbol{k}) = {}''\,\mathbb{P}\left[J(\boldsymbol{k}, \boldsymbol{U}) = J^*(\boldsymbol{U})\right]''$$

How to take into account values not optimal, but not too far either $\longrightarrow$ relaxation of the equality with $\alpha > 1$:

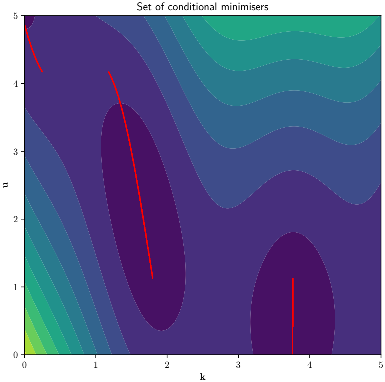
$$\Gamma_\alpha(\boldsymbol{k}) = \mathbb{P}_{\boldsymbol{U}}\left[J(\boldsymbol{k}, \boldsymbol{U}) \leq \alpha J^*(\boldsymbol{U})\right]$$

Illustration

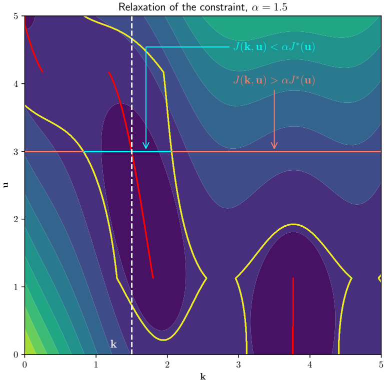


- Sample $\boldsymbol{u} \sim \boldsymbol{U}$, and solve $\boldsymbol{k}^*(\boldsymbol{u}) = \arg\min_{\boldsymbol{k}\in\mathbb{K}} J(\boldsymbol{k}, \boldsymbol{u})$

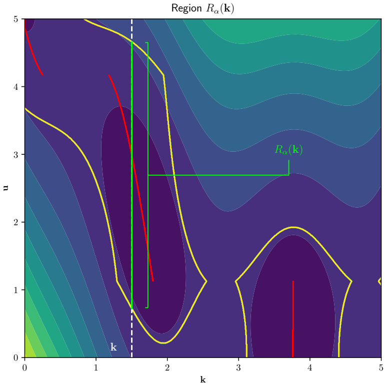
Illustration



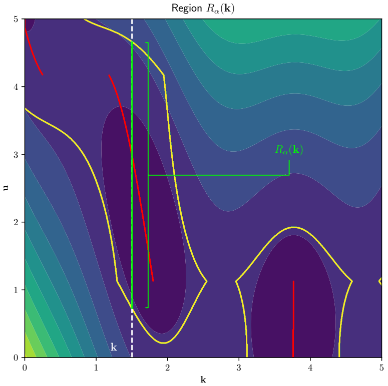
- Sample $\boldsymbol{u} \sim \boldsymbol{U}$, and solve $\boldsymbol{k}^*(\boldsymbol{u}) = \arg\min_{\boldsymbol{k} \in \mathbb{K}} J(\boldsymbol{k}, \boldsymbol{u})$
- Set of conditional minimisers: $\{(\boldsymbol{k}^*(\boldsymbol{u}), \boldsymbol{u}) \mid \boldsymbol{u} \in \mathbb{U}\}$



- Sample $\boldsymbol{u} \sim \boldsymbol{U}$, and solve $\boldsymbol{k}^*(\boldsymbol{u}) = \arg\min_{\boldsymbol{k}\in\mathbb{K}} J(\boldsymbol{k},\boldsymbol{u})$
- Set of conditional minimisers: $\{(\boldsymbol{k}^*(\boldsymbol{u}),\boldsymbol{u}) \mid \boldsymbol{u}\in\mathbb{U}\}$
- Set $\alpha \geq 1$

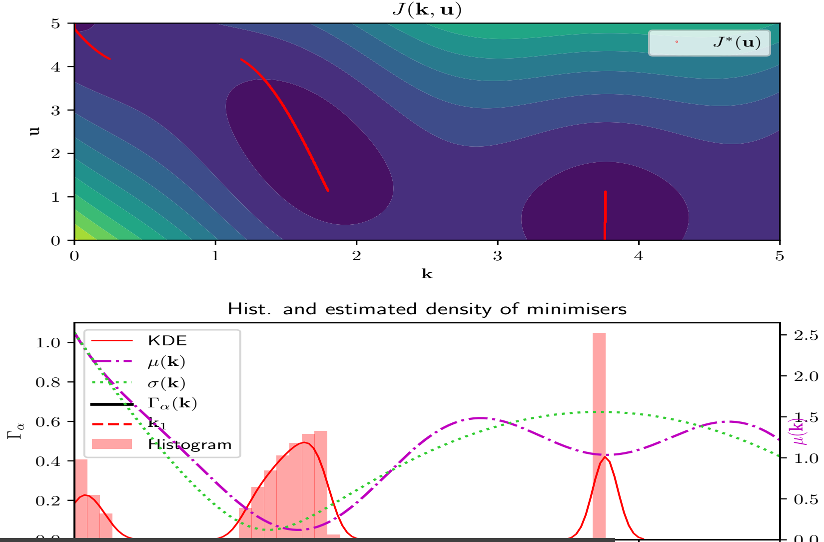


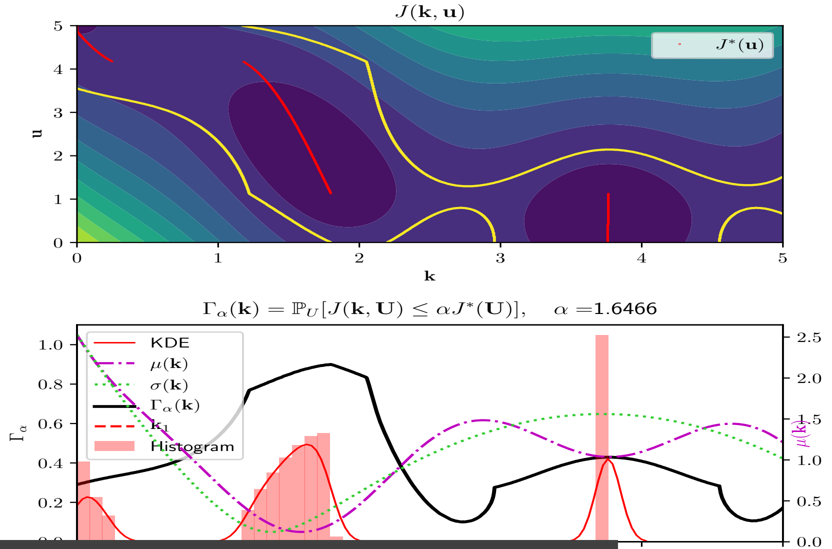
- Sample $\boldsymbol{u} \sim \boldsymbol{U}$, and solve $\boldsymbol{k}^*(\boldsymbol{u}) = \arg\min_{\boldsymbol{k}\in\mathbb{K}} J(\boldsymbol{k}, \boldsymbol{u})$
- Set of conditional minimisers: $\{(\boldsymbol{k}^*(\boldsymbol{u}), \boldsymbol{u}) \mid \boldsymbol{u} \in \mathbb{U}\}$
- Set $\alpha \geq 1$
- $R_\alpha(\boldsymbol{k}) = \{\boldsymbol{u} \mid J(\boldsymbol{k}, \boldsymbol{u}) < \alpha J^*(\boldsymbol{u})\}$
- $\Gamma_\alpha(\boldsymbol{k}) = \mathbb{P}_{\boldsymbol{U}}\left[\boldsymbol{U} \in R_\alpha(\boldsymbol{k})\right]$



- Sample $\boldsymbol{u} \sim \boldsymbol{U}$, and solve $\boldsymbol{k}^*(\boldsymbol{u}) = \arg\min_{\boldsymbol{k}\in\mathbb{K}} J(\boldsymbol{k}, \boldsymbol{u})$
- Set of conditional minimisers: $\{(\boldsymbol{k}^*(\boldsymbol{u}), \boldsymbol{u}) \mid \boldsymbol{u} \in \mathbb{U}\}$
- Set $\alpha \geq 1$
- $R_\alpha(\boldsymbol{k}) = \{\boldsymbol{u} \mid J(\boldsymbol{k}, \boldsymbol{u}) < \alpha J^*(\boldsymbol{u})\}$
- $\Gamma_\alpha(\boldsymbol{k}) = \mathbb{P}_{\boldsymbol{U}}\left[\boldsymbol{U} \in R_\alpha(\boldsymbol{k})\right]$
- How to choose α? When $\max_{\boldsymbol{k}} \Gamma_\alpha(\boldsymbol{k})$ reaches fixed levels

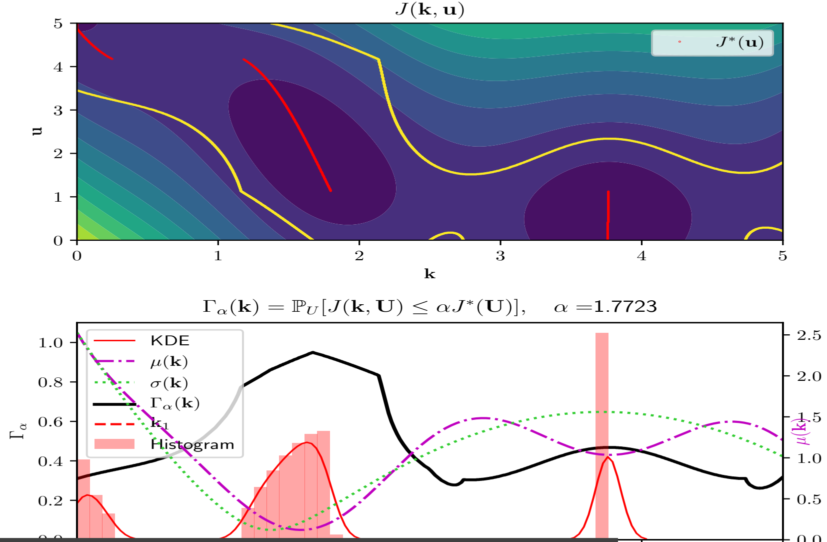
Choosing a α



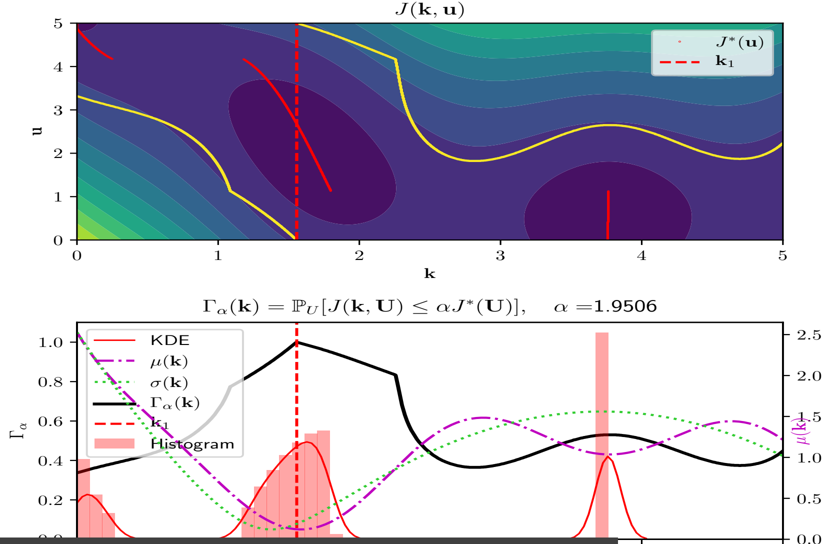


$J(\mathbf{k},\mathbf{u})$
$J^*(\mathbf{u})$
u
k
0
1
2
3
4
5
$\Gamma_\alpha(\mathbf{k}) = \mathbb{P}_U[J(\mathbf{k},\mathbf{U}) \leq \alpha J^*(\mathbf{U})], \quad \alpha = 1.6466$
KDE
$\mu(\mathbf{k})$
$\sigma(\mathbf{k})$
$\Gamma_\alpha(\mathbf{k})$
$\mathbf{k}_1$
Histogram
Γ_α
$\mu(\mathbf{k})$
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5

CHOOSING A α



CHOOSING A α



Surrogates

SURROGATES

HOW TO COMPUTE $\hat{k}$ IN A REASONABLE TIME?

WHY SURROGATES?

- Computer model: expensive to run
- $\dim \mathbb{K}$, $\dim \mathbb{U}$ can be very large: curse of dimensionality
- Uncertainties upon $\boldsymbol{u}$ may be incorporated directly in the surrogate

Environmental variables
$\boldsymbol{U}$ random

Control variable
$\boldsymbol{k} \in \mathcal{K}$

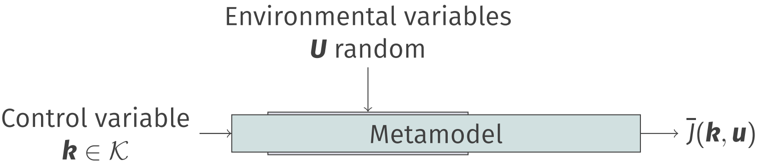
$$\xrightarrow{} \boxed{\text{Direct Simulation}} \longrightarrow \mathcal{M}(\boldsymbol{k}, \boldsymbol{u}) \longrightarrow J(\boldsymbol{k}, \boldsymbol{u})$$

Two main forms:

- Kriging (Gaussian Process Regression) [Mat62]
- Polynomial Chaos Expansion [XK02, Sud15]

Why surrogates?

- Computer model: expensive to run
- $\dim \mathbb{K}$, $\dim \mathbb{U}$ can be very large: curse of dimensionality
- Uncertainties upon $\boldsymbol{u}$ may be incorporated directly in the surrogate

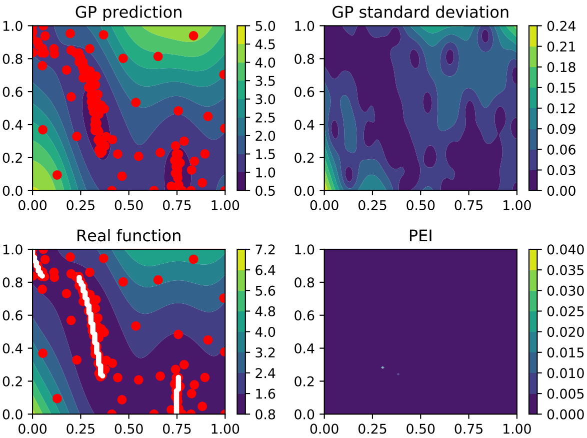


Two main forms:

- Kriging (Gaussian Process Regression) [Mat62]
- Polynomial Chaos Expansion [XK02, Sud15]

Estimation of $\boldsymbol{K}^*$, $J^*(\boldsymbol{U})$

Iterative procedures to estimate set of conditional minimum/minimisers [GBC+14]



- Sensitivity analysis [Sud08, LGMS16]: Based on intensive computation of the metamodel, or analytic computation based on coefficients of the expansion computed
- Isotropic by groups kernels [BHRV17, Rib18]: Group variables to have a few isotropic kernels

Conclusion

Conclusion

Wrapping up

- Problem of a *good* definition of robustness
- Strategies rely heavily on surrogate models, to embed aleatoric uncertainties directly in the modelling

Perspective and future work

- Cost of computer evaluations $\rightarrow$ limited number of runs?
- Dimensionality of the input space $\rightarrow$ reduction of the input space?
- How to deal with uncontrollable errors $\rightarrow$ realism of the model?

References

Vincent Baudoui.
Optimisation Robuste Multiobjectifs Par Modèles de Substitution.
PhD thesis, Toulouse, ISAE, 2012.

Christophette Blanchet-Scalliet, Céline Helbert, Mélina Ribaud, and Céline Vial.
A specific kriging kernel for dimensionality reduction: Isotropic by group kernel, March 2017.

David Ginsbourger, Jean Baccou, Clément Chevalier, Frédéric Perales, Nicolas Garland, and Yann Monerie.
Bayesian Adaptive Reconstruction of Profile Optima and Optimizers.
SIAM/ASA Journal on Uncertainty Quantification, 2(1):490–510, January 2014.

Loic Le Gratiet, Stefano Marelli, and Bruno Sudret.
Metamodel-based sensitivity analysis: Polynomial chaos expansions and Gaussian processes.
In *Handbook of Uncertainty Quantification - Part III: Sensitivity Analysis*. 2016.

Jeffrey S. Lehman, Thomas J. Santner, and William I. Notz.
Designing computer experiments to determine robust control variables.
Statistica Sinica, pages 571–590, 2004.

Georges Matheron.
***Traité de Géostatistique Appliquée. 1 (1962)*, volume 1.**
Editions Technip, 1962.

Julien Marzat, Eric Walter, and Hélène Piet-Lahanier.
Worst-case global optimization of black-box functions through Kriging and relaxation.
Journal of Global Optimization, 55(4):707–727, April 2013.

Mélina Ribaud.
Krigeage Pour La Conception de Turbomachines : Grande Dimension et Optimisation Multi-Objectif Robuste.
Thesis, Lyon, October 2018.

Bruno Sudret.
Global sensitivity analysis using polynomial chaos expansion.
Reliability Engineering & System Safety, 93:964–979, July 2008.

Bruno Sudret.
Polynomial chaos expansions and stochastic finite element methods.
In Jianye Ching Kok-Kwang Phoon, editor, *Risk and Reliability in Geotechnical Engineering*, pages 265–300. CRC Press, 2015.

Warren E. Walker, Poul Harremoës, Jan Rotmans, Jeroen P. van der Sluijs, Marjolein BA van Asselt, Peter Janssen, and Martin P. Krayer von Krauss.
Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support.
Integrated assessment, 4(1):5–17, 2003.

D. Xiu and G. Karniadakis.
The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations.
SIAM Journal on Scientific Computing, 24(2):619–644, January 2002.