Parameter control in the presence of uncertainties

Victor Trappler
victor.trappler@univ-grenoble-alpes.fr
Élise Arnaud, Laurent Debreu, Arthur Vidard
AIRSEA Research team (Inria) – Laboratoire Jean Kuntzmann

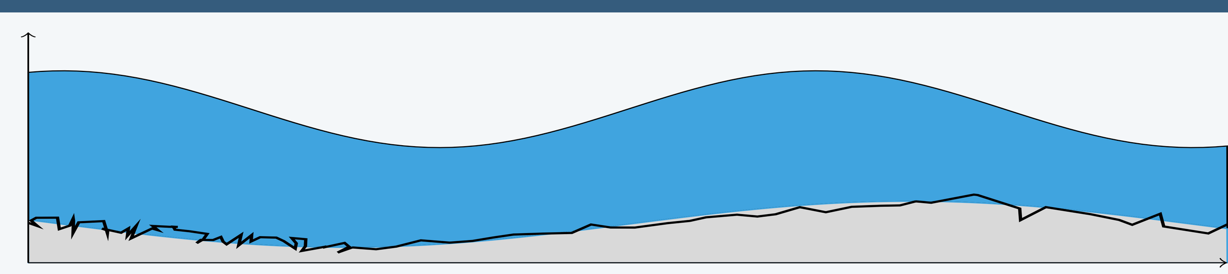
Workshop on Sensitivity Analysis and Data Assimilation in Meteorology and Oceanography, Aveiro, Portugal 2018

How can one calibrate a numerical model so that it performs reasonably well for different random operating conditions ?

Objectives

- Define suitable definitions of robustness in the field of computer code calibration
- Develop efficient techniques and algorithms in order to estimate those parameters
- Deal with the high-dimension of the parameter spaces: Dimension reduction

Background: estimation of the bottom friction in a shallow water model



The calibration problem is to be able to find a value of $\boldsymbol{k} \in \mathcal{K}$ denoted $\hat{\boldsymbol{k}}$ that matches the best the observations $\boldsymbol{y}_{\mathrm{obs}}$. We define a loss function, that is the misfit between the observations to the model.

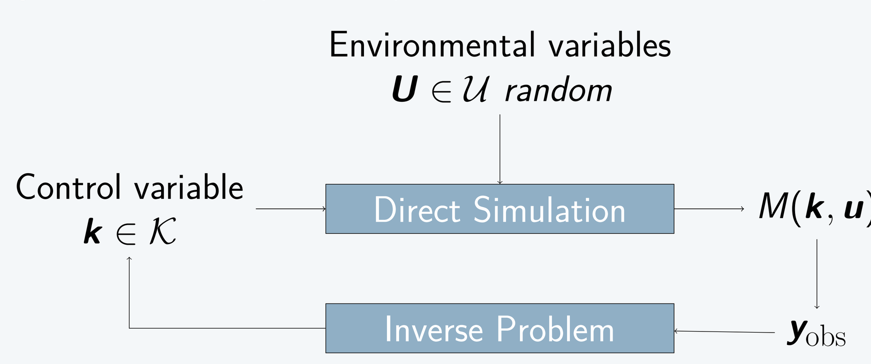
$$J(\boldsymbol{k}) = \frac{1}{2}\|M(\boldsymbol{k}) - \boldsymbol{y}_{\mathrm{obs}}\|^2_{\Sigma^{-1}}$$

and we have to perform the following minimisation problem, usually with the help of the adjoint method

$$\hat{\boldsymbol{k}} = \arg\min_{\boldsymbol{k}\in\mathcal{K}} J(\boldsymbol{k})$$

Stochastic Inverse Problem

Now, $\boldsymbol{u} \in \mathcal{U} \sim \boldsymbol{U}$ of density p_U and $\boldsymbol{y}_{\mathrm{obs}} = M(\boldsymbol{k}_{\mathrm{ref}}, \boldsymbol{u}_{\mathrm{ref}})$



The loss function is now

$$J(\boldsymbol{k}, \boldsymbol{U}) = \underbrace{\frac{1}{2}\|M(\boldsymbol{k}, \boldsymbol{U}) - \boldsymbol{y}_{\mathrm{obs}}\|^2_{\Sigma^{-1}}}_{\text{Random variable}}$$

- What criteria to use to "optimize" in a sense J ?
- Evaluating J is time consuming. How to deal with a limited budget of evaluations ?

Which criterion to choose ?

- Global minimum

$$(\boldsymbol{k}^*, \boldsymbol{u}^*) = \arg\min_{(\boldsymbol{k},\boldsymbol{u})} J(\boldsymbol{k}, \boldsymbol{u}) \quad \text{and} \quad \hat{\boldsymbol{k}}_{\mathrm{global}} = \boldsymbol{k}^*$$

- Assuming that the environmental variables have little influence:

$$J_{\mathbb{E}}(\boldsymbol{k}) = J(\boldsymbol{k}, \mathbb{E}[\boldsymbol{U}]) \quad \text{and} \quad \hat{\boldsymbol{k}}_{\mathbb{E}} = \arg\min_{\boldsymbol{k}} J_{\mathbb{E}}(\boldsymbol{k}) \qquad \text{(Classical methods)}$$

$\longrightarrow$ Those approaches are not robust: inherent variability of $\boldsymbol{U}$ not taken into account

- Consider the worst-case scenario

$$J_{\mathrm{w}}(\boldsymbol{k}) = \max_{\boldsymbol{u}\in\mathcal{U}} J(\boldsymbol{k}, \boldsymbol{u}) \quad \text{and} \quad \hat{\boldsymbol{k}}_{\mathrm{wc}} = \arg\min_{\boldsymbol{k}} J_{\mathrm{w}}(\boldsymbol{k}) \qquad \text{(Explorative EGO)}$$

- The solution gives good results on average:

$$\mu(\boldsymbol{k}) = \mathbb{E}_U[J(\boldsymbol{k}, \boldsymbol{U})] \quad \text{and} \quad \hat{\boldsymbol{k}}_{\mu} = \arg\min_{\boldsymbol{k}} \mu(\boldsymbol{k}) \qquad \text{(Iterative EGO)}$$

- The estimate gives steady results:

$$\sigma^2(\boldsymbol{k}) = \mathbb{V}\mathrm{ar}_U[J(\boldsymbol{k}, \boldsymbol{U})] \quad \text{and} \quad \hat{\boldsymbol{k}}_{\sigma^2} = \arg\min_{\boldsymbol{k}} \sigma^2(\boldsymbol{k}) \qquad \text{(PCE gradient)}$$

- Compromise between Mean and Variance $\rightarrow$ multiobjective optimization problem:

$$\text{Pareto front of } (\mu(\boldsymbol{k}), \sigma^2(\boldsymbol{k})) \qquad \text{(Layered kriging)}$$

- Probability of being below threshold $T \in \mathbb{R}$: Reliability analysis

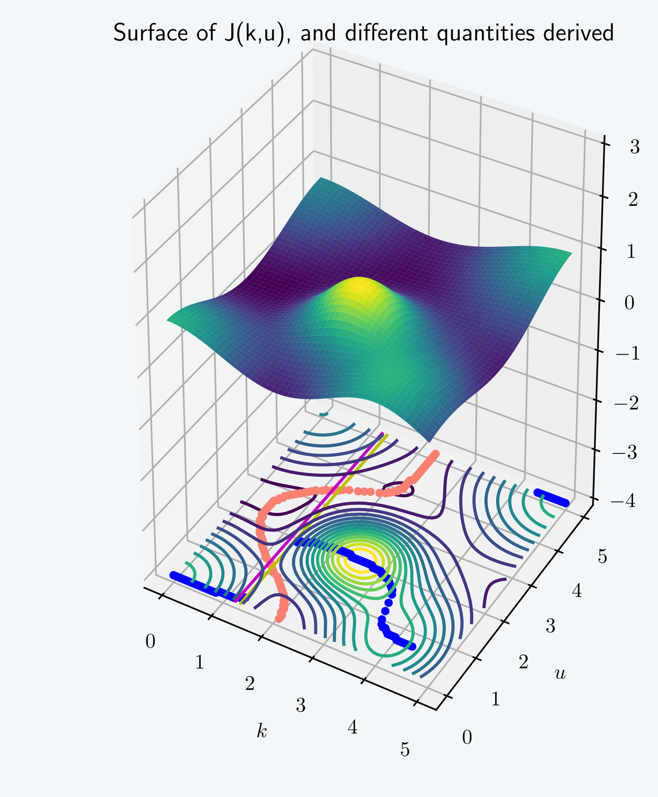
$$R_T(\boldsymbol{k}) = \mathbb{P}[J(\boldsymbol{k}, \boldsymbol{U}) \leq T], \quad \hat{\boldsymbol{k}}_{R_T} = \arg\max R_T(\boldsymbol{k}) \qquad \text{(GP simulations)}$$

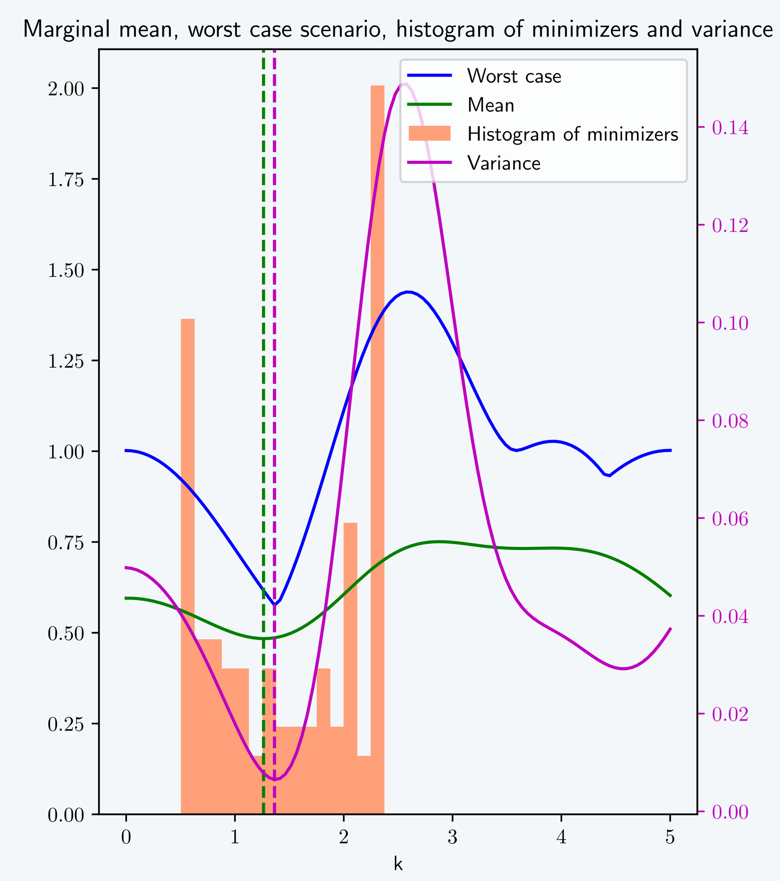
- Distribution of minimizers: $T_{\min} = T(\boldsymbol{U}) = \min_{\boldsymbol{k}} J(\boldsymbol{k}, \boldsymbol{U})$

$$R_{\min}(\boldsymbol{k}) = \mathbb{P}[J(\boldsymbol{k}, \boldsymbol{U}) \leq T_{\min}] = \mathbb{P}\left[\boldsymbol{k} = \arg\min_{\tilde{\boldsymbol{k}}} J(\tilde{\boldsymbol{k}}, \boldsymbol{U})\right] \qquad \text{(Estimation and maximization of density)}$$

- Relaxation of the constraint: we define $T_\alpha(\boldsymbol{U}) = \alpha \min_{\boldsymbol{k}} J(\boldsymbol{k}, \boldsymbol{U})$, for $\alpha \geq 1$, and $R_\alpha = R_{T_\alpha}$

2D Illustration

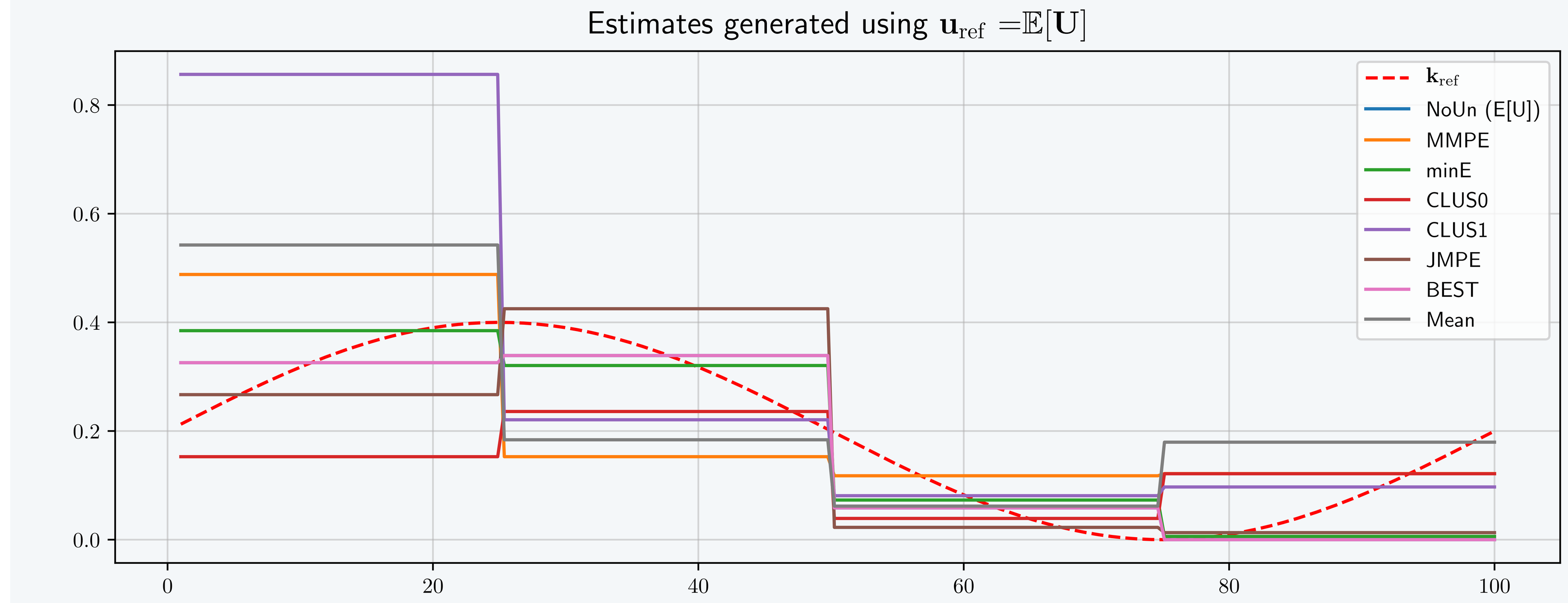


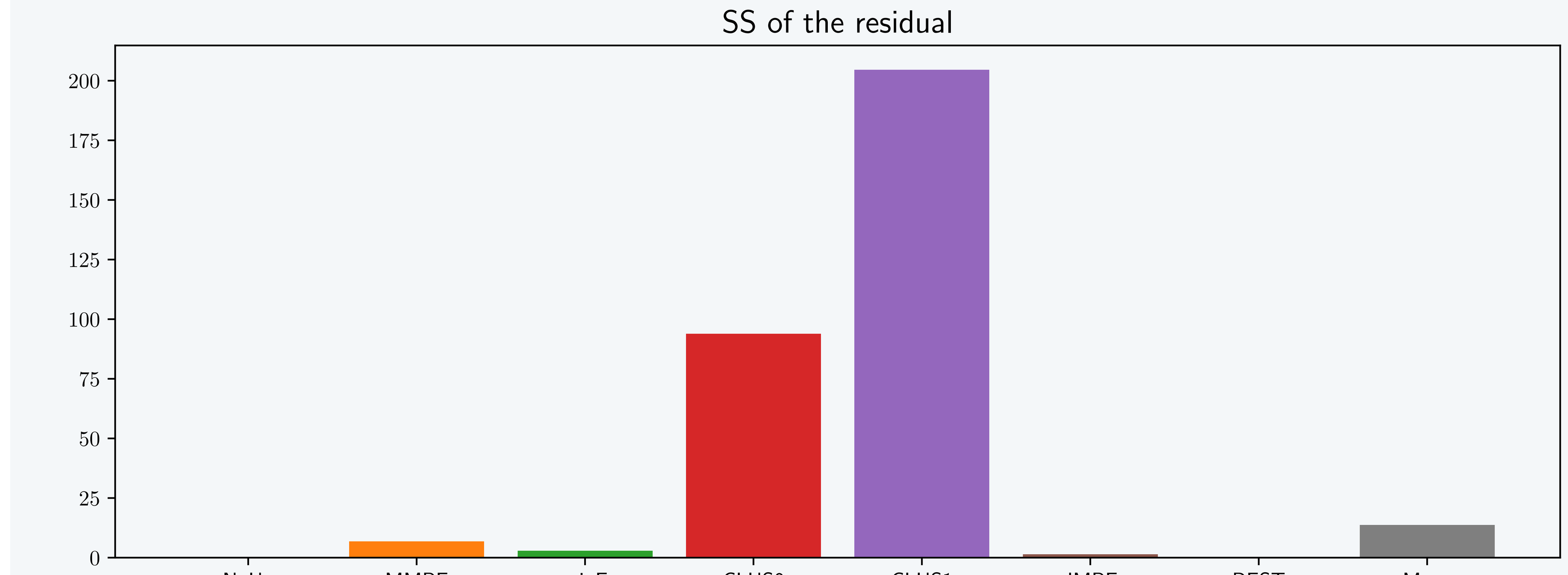


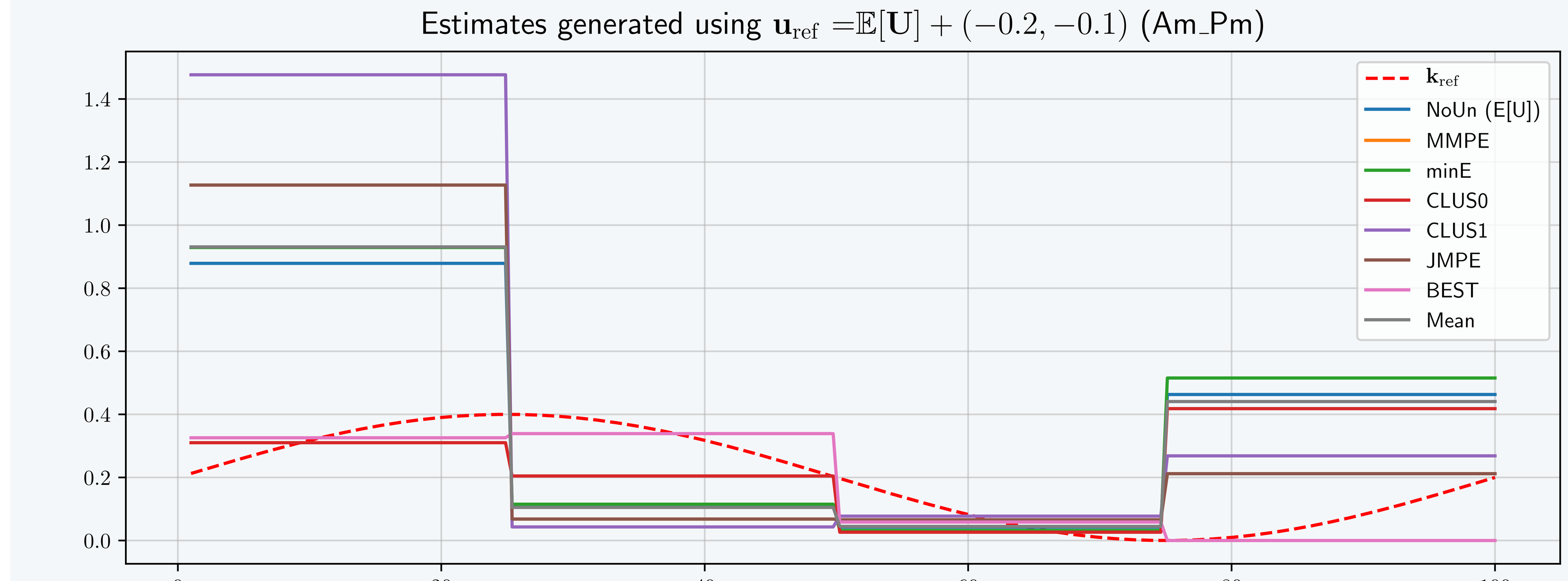
General methods

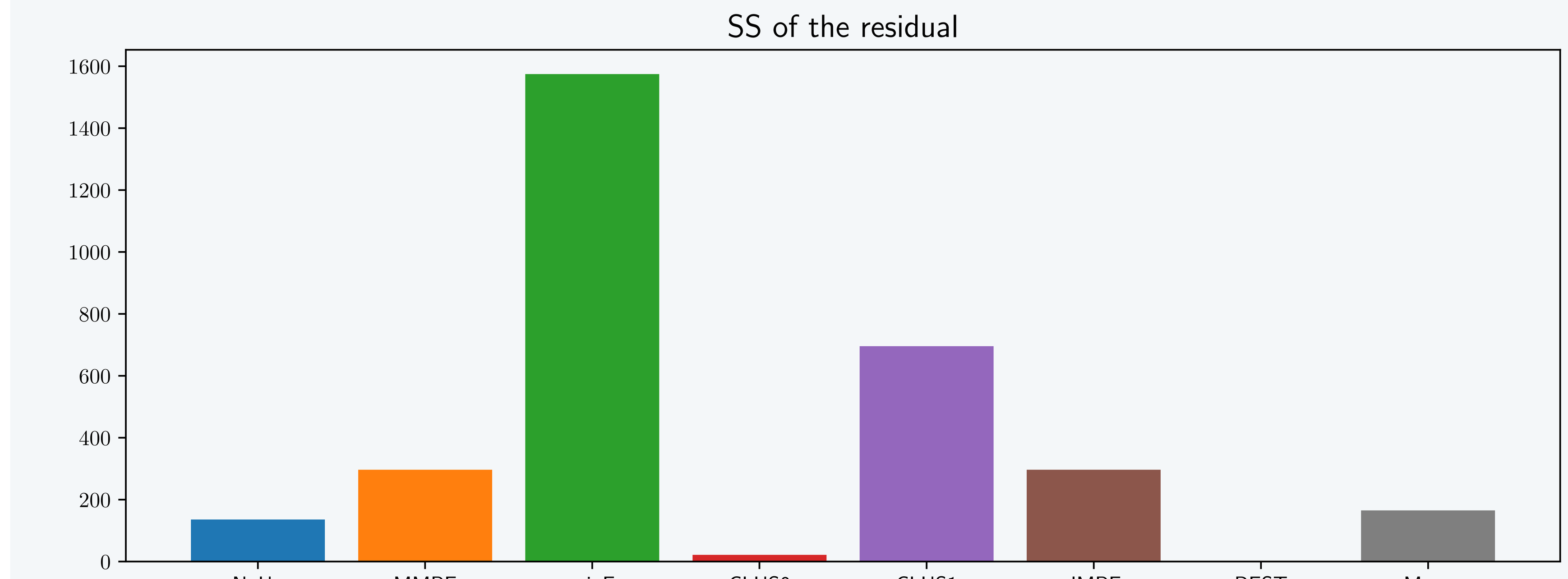
- Design of Experiment
  - Efficient exploration of the input space: LHS, space filling designs
- Statistical/Probabilistic aspects
  - Bayesian/Frequentist approach: Markov-chain based methods, study of the posterior distribution
  - Choice of prior on $\boldsymbol{K}$ to take into account specific information on spatial variation of the friction
  - Marginalization with respect to $\boldsymbol{U}$
- Surrogate modelling
  - Kriging (Gaussian Process Regression)
  - Polynomial Chaos Expansion
- Optimization
  - Adjoint method provides the gradient of the cost function $\rightarrow$ Adapt principles of gradient descent on specific objectives
  - Adaptative sampling: based on surrogate, choose the next point to be evaluated based on a specific criterion: EGO, IAGO and more general *Stepwise Uncertainty Reduction* strategies

Numerical Results: toy model of Shallow Waters









Conclusion and perspectives

- Several objectives can be defined, often concurrent
- Choice of criterion of robustness is application-dependent
- Scalability of methods in high dimension ? Need to perform Dimension reduction on $\mathcal{K}$ and $\mathcal{U}$

References

- Maria J. Bayarri, James O. Berger, Rui Paulo, Jerry Sacks, John A. Cafeo, James Cavendish, Chin-Hsu Lin, and Jian Tu. A framework for validation of computer models. *Technometrics*, 49(2):138–154, 2007.
- Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Society for Industrial and Applied Mathematics, Philadelphia, Pa, 2005. OCLC: 265659758.
- Jeffrey S. Lehman, Thomas J. Santner, and William I. Notz. Designing computer experiments to determine robust control variables. *Statistica Sinica*, pages 571–590, 2004.
- Abraham Wald. Statistical Decision Functions Which Minimize the Maximum Risk. *Annals of Mathematics*, 46(2):265–280, 1945.
- Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.