

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

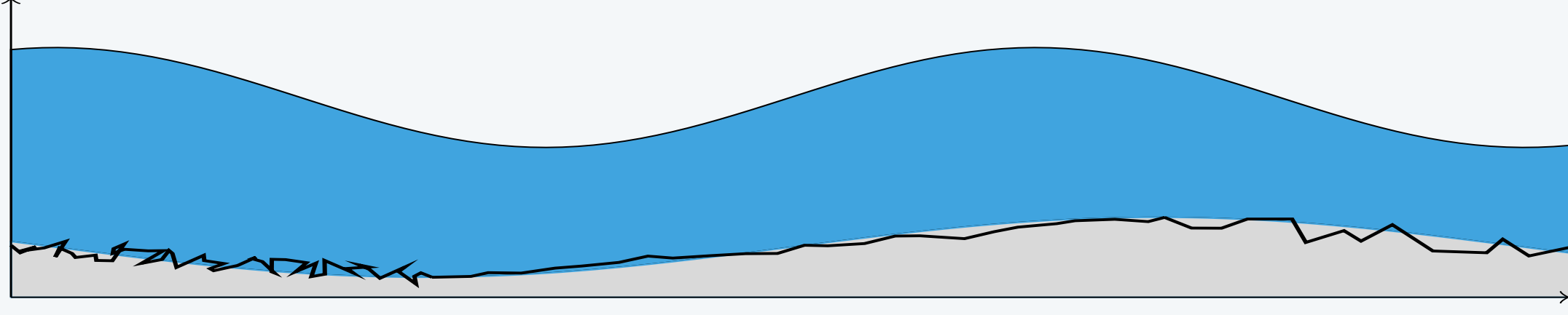
WORKSHOP ON SENSITIVITY ANALYSIS AND DATA ASSIMILATION IN METEOROLOGY AND OCEANOGRAPHY, AVEIRO, PORTUGAL 2018

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 $\mathbf{k} \in \mathcal{K}$ denoted $\hat{\mathbf{k}}$ that matches the best the observations \mathbf{y}_{obs} . We define a loss function, that is the misfit between the observations to the model.

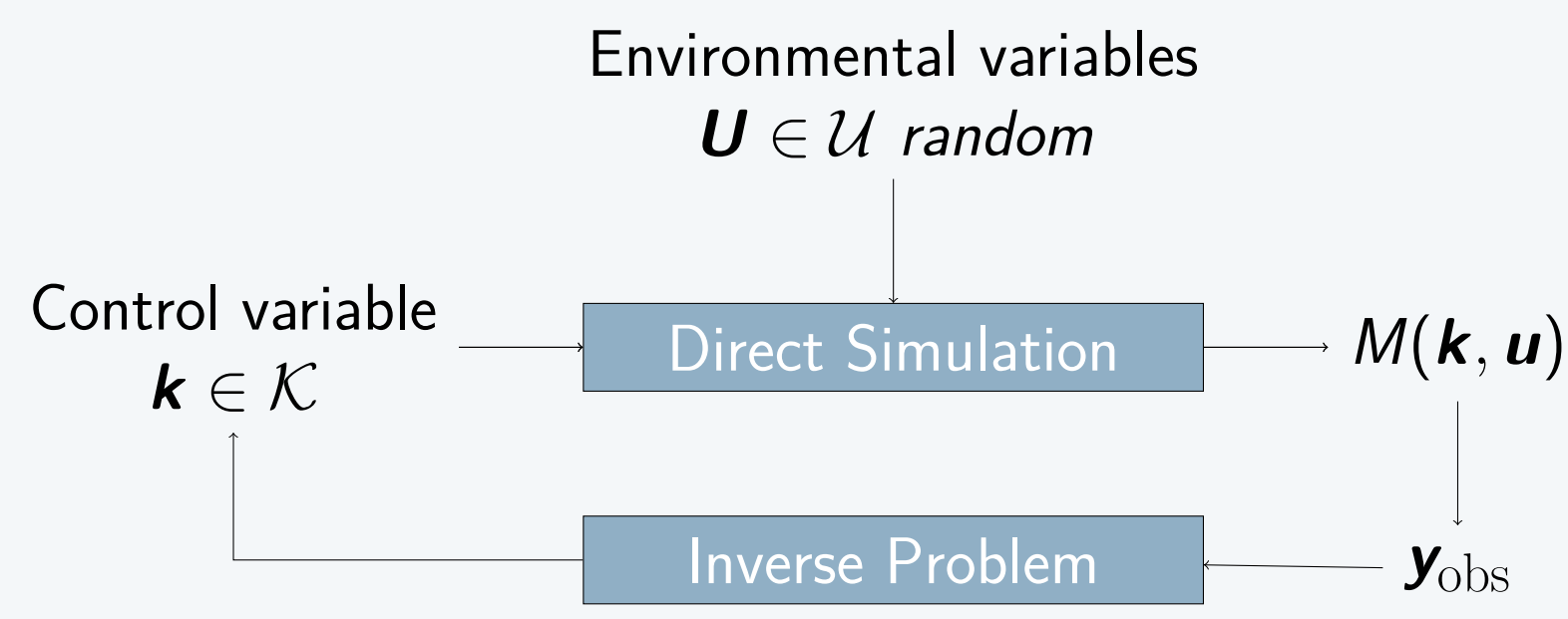
$$J(\mathbf{k}) = \frac{1}{2} \|\mathbf{M}(\mathbf{k}) - \mathbf{y}_{\text{obs}}\|_{\Sigma}^2 - 1$$

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

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

Stochastic Inverse Problem

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



The loss function is now

$$J(\mathbf{k}, \mathbf{U}) = \frac{1}{2} \|\mathbf{M}(\mathbf{k}, \mathbf{U}) - \mathbf{y}_{\text{obs}}\|_{\Sigma}^2 - 1$$

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

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

- Assuming that the environmental variables have little influence:

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

→ Those approaches are not robust: inherent variability of \mathbf{U} not taken into account

- Consider the **worst-case scenario**

$$J_w(\mathbf{k}) = \max_{\mathbf{u} \in \mathcal{U}} J(\mathbf{k}, \mathbf{u}) \quad \text{and} \quad \hat{\mathbf{k}}_{\text{wc}} = \arg \min_{\mathbf{k}} J_w(\mathbf{k}) \quad (\text{Explorative EGO})$$

- The solution gives **good results on average**:

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

- The estimate gives **steady results**:

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

- Compromise** between Mean and Variance → multiobjective optimization problem:

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

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

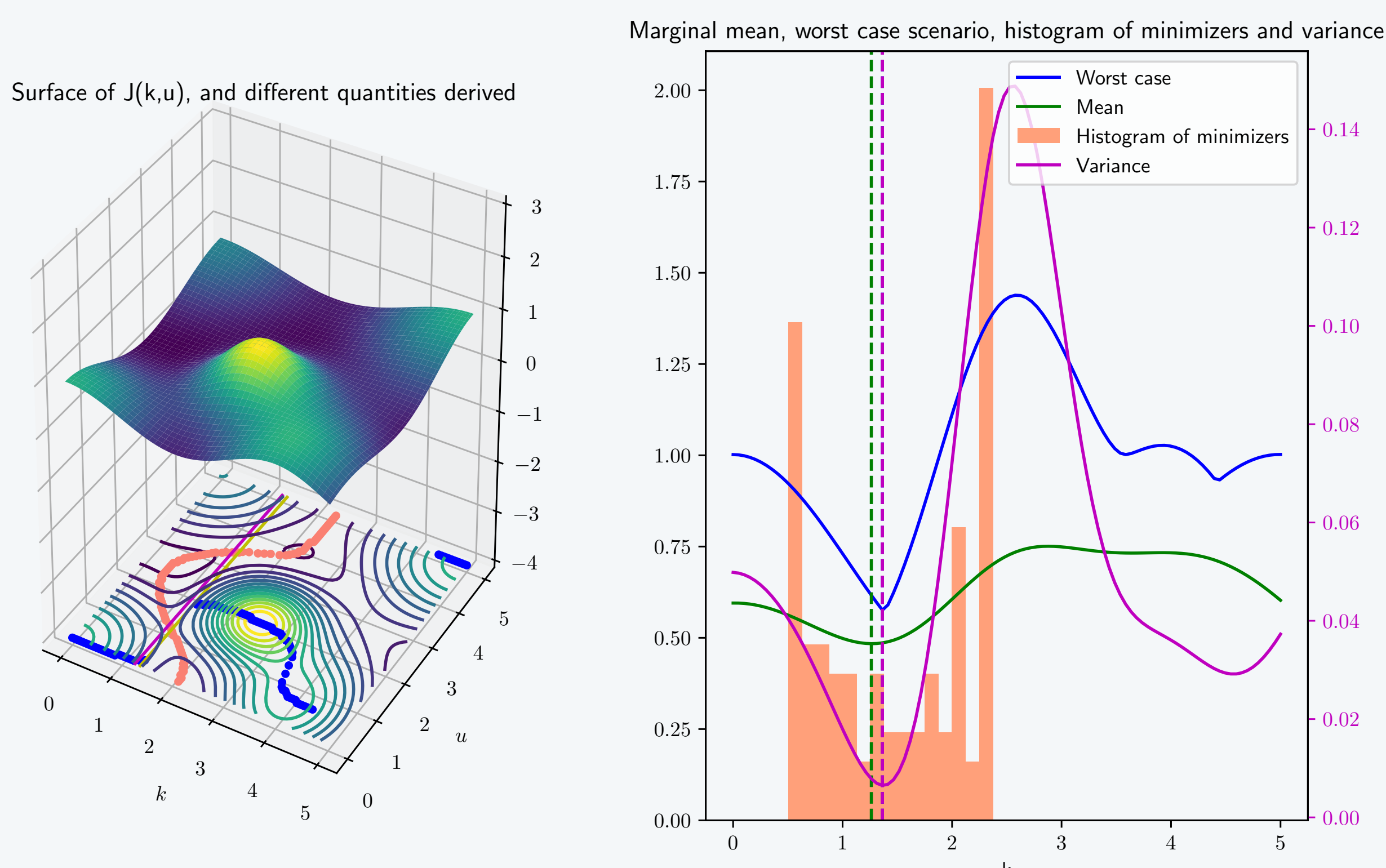
$$R_T(\mathbf{k}) = \mathbb{P}[J(\mathbf{k}, \mathbf{U}) \leq T], \quad \hat{\mathbf{k}}_{R_T} = \arg \max_{\mathbf{k}} R_T(\mathbf{k}) \quad (\text{GP simulations})$$

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

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

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

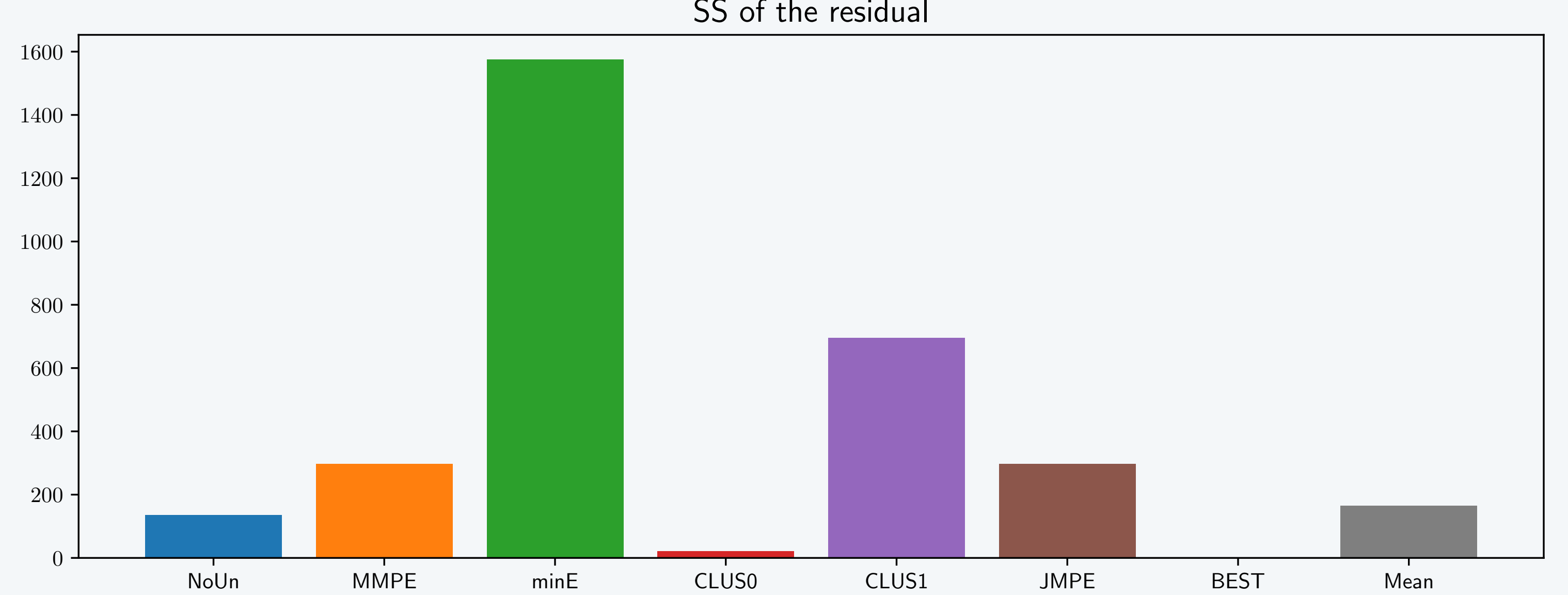
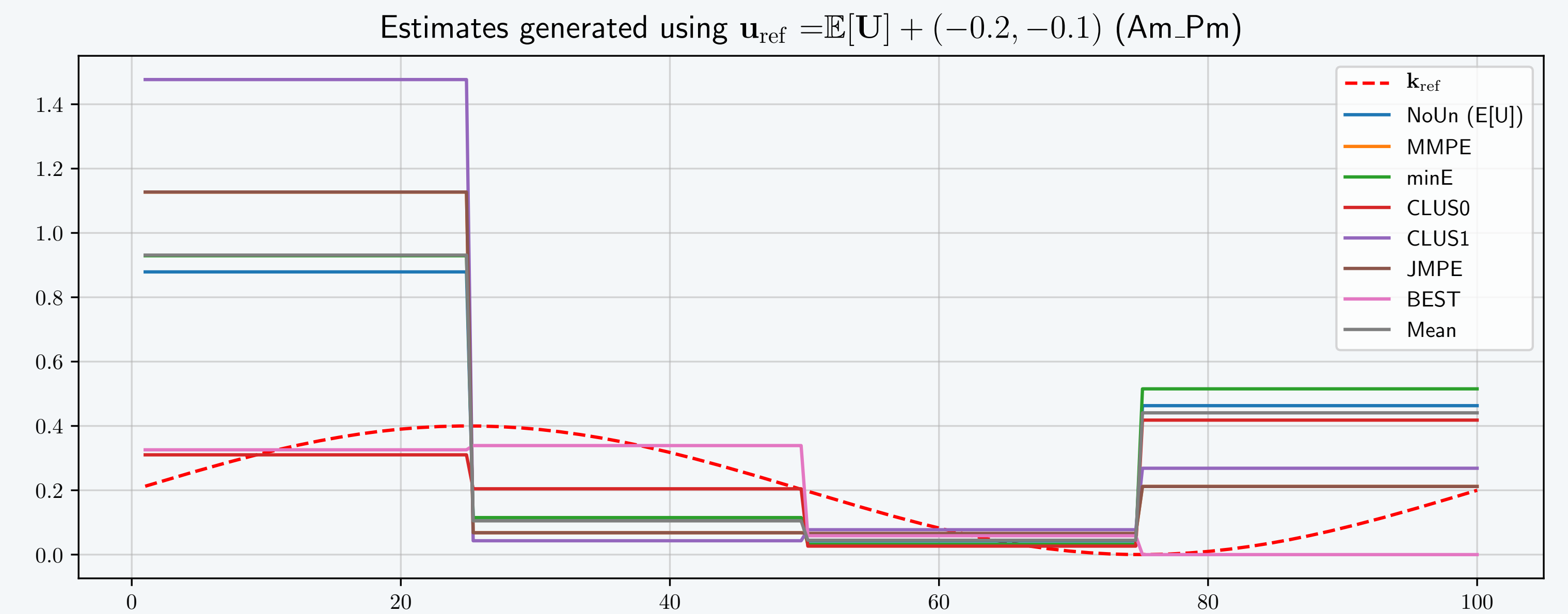
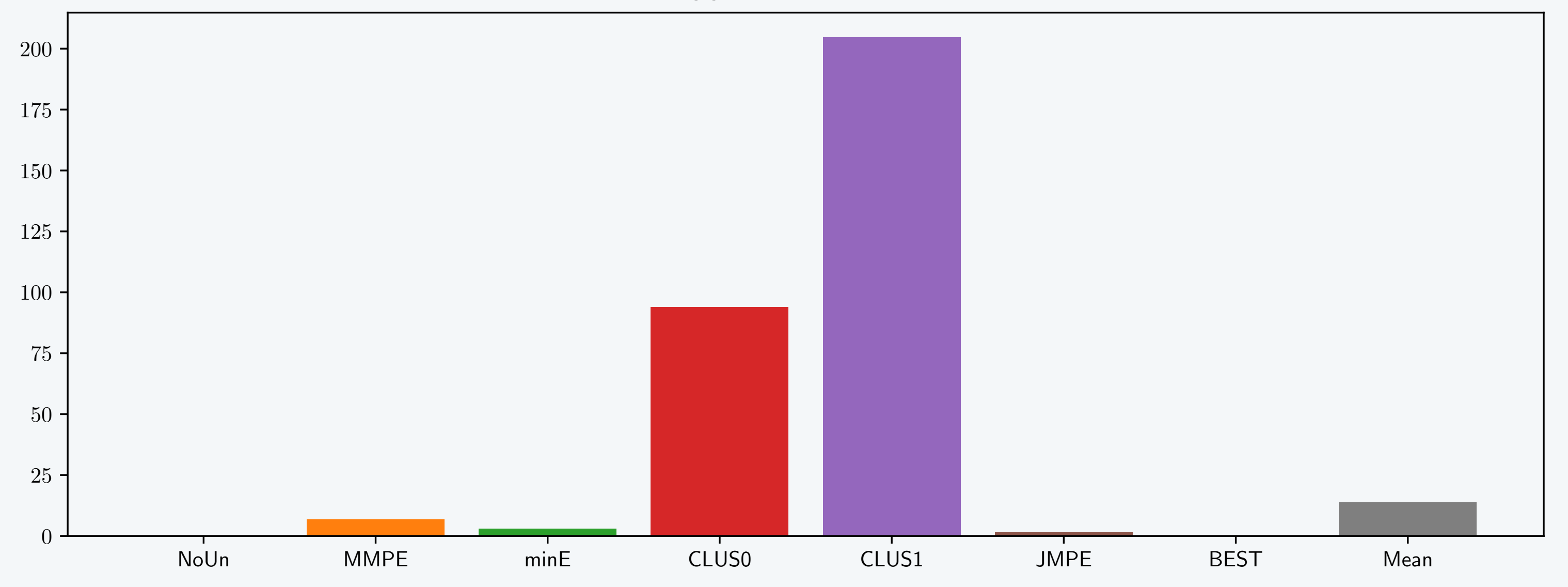
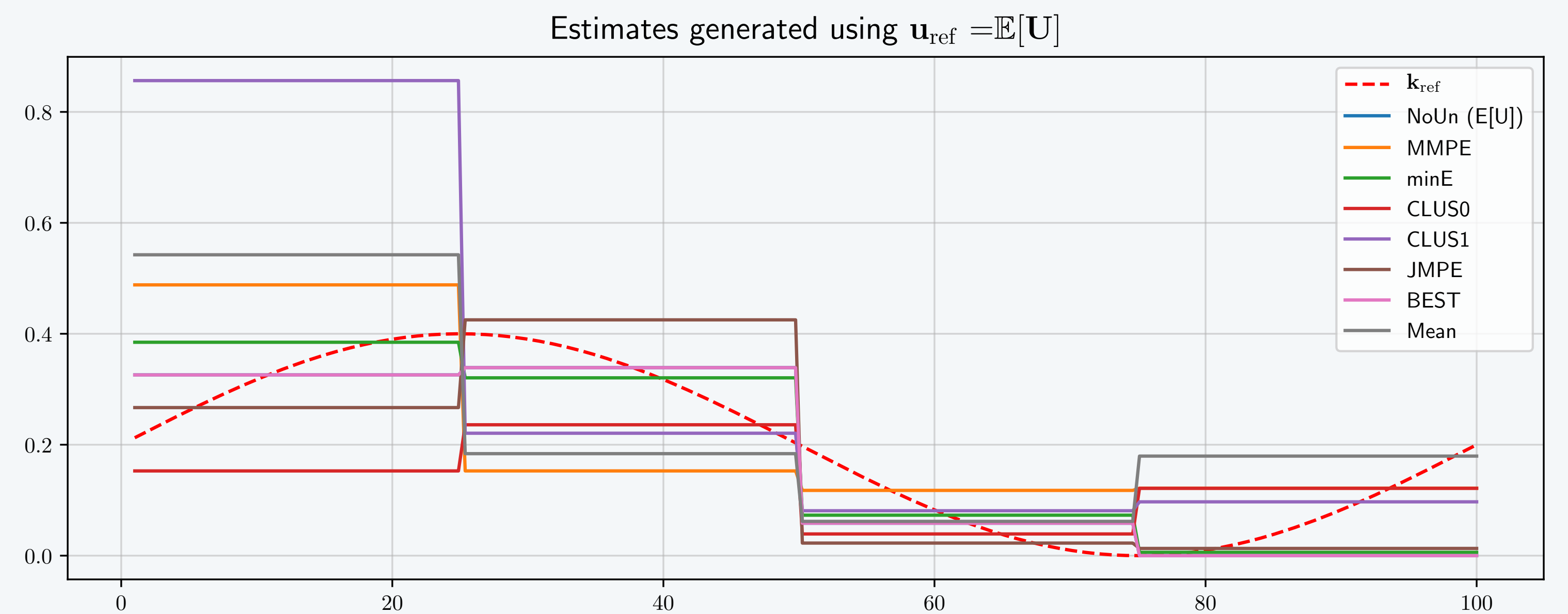
2D Illustration



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 \mathbf{K} to take into account specific information on spatial variation of the friction
 - Marginalization with respect to \mathbf{U}
- Surrogate modelling
 - Kriging (Gaussian Process Regression)
 - Polynomial Chaos Expansion
- Optimization
 - Adjoint method provides the gradient of the cost function → 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. Cafoe, 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.
- 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.