Parameter control in the presence of uncertainties

PhD Defense, June 11th 2021

Victor Trappler

Advisors:	Élise Arnaud	Univ. Grenoble Alpes
	Laurent Debreu	Inria Grenoble
	Arthur Vidard	Inria Grenoble
Jury:	Youssef Marzouk	MIT
	Pietro Marco Congedo	Inria Paris Saclay
	Olivier Roustant	INSA Toulouse
	Rémy Baraille	SHOM

Why do we need models ?

The ability to understand is essential in order to forecast, and to take decisions.

- Natural phenomena are often very complex to understand in their entirety
- Mathematical models are simplified versions, which allow to study the relationships between some observed (or not) quantities
- Mathematical models can be used to construct numerical models, which are used for forecasts

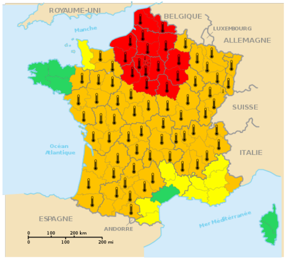
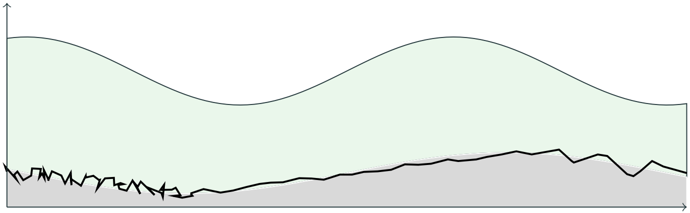



Figure 1: From the reality to predictions

Modelling of the ocean, bottom friction

Example: modelling of the ocean

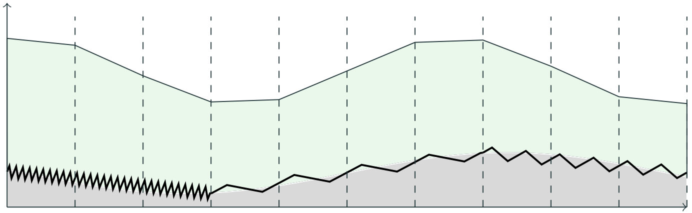
- Ocean domain discretized
- External tidal forcing
- Energy is dissipated through turbulences caused by the asperities at the bottom
  - The water currents at the surface are affected
  - The friction is taken into account at a cell level



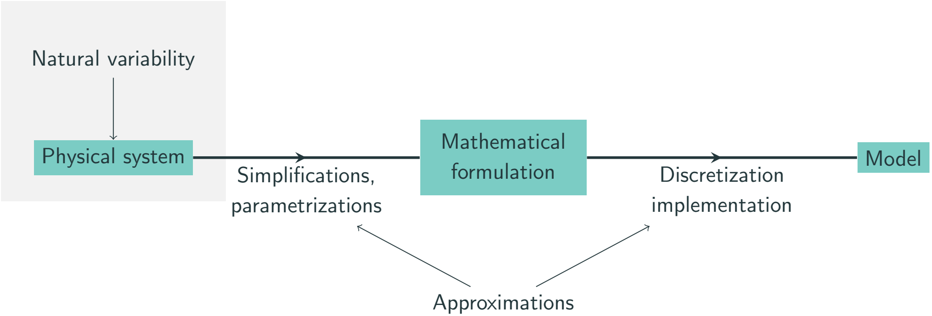
Modelling of the ocean, bottom friction

Example: modelling of the ocean

- Ocean domain discretized
- External tidal forcing
- Energy is dissipated through turbulences caused by the asperities at the bottom
    - The water currents at the surface are affected
    - The friction is taken into account at a cell level



The modelling process

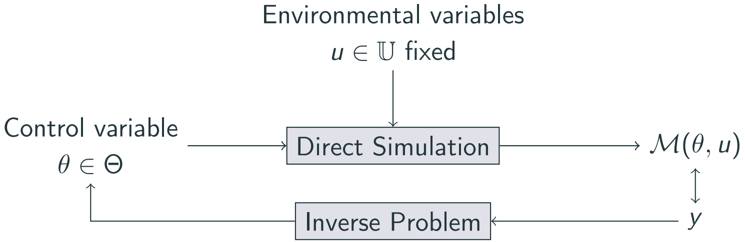


$\rightarrow$ How well can we calibrate the model, so that it depicts *accurately* the reality ?

Computer code and inverse problem

Input
- θ: Control parameter
- u: Environmental variables (fixed and known)

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



Data assimilation framework

Let $u \in \mathbb{U}$, assumed fixed and known

Objective function

We define J as the squared difference between the output of the model and the observations

$$J(\theta, u) = \frac{1}{2}\|\mathcal{M}(\theta, u) - y\|^2 \tag{1}$$

$\rightarrow$ the smaller J is, the better the fit is

Data assimilation framework

Let $u \in \mathbb{U}$, assumed fixed and known

Objective function

We define J as the squared difference between the output of the model and the observations

$$J(\theta, u) = \frac{1}{2}\|\mathcal{M}(\theta, u) - y\|^2 \tag{1}$$

$\rightarrow$ the smaller J is, the better the fit is

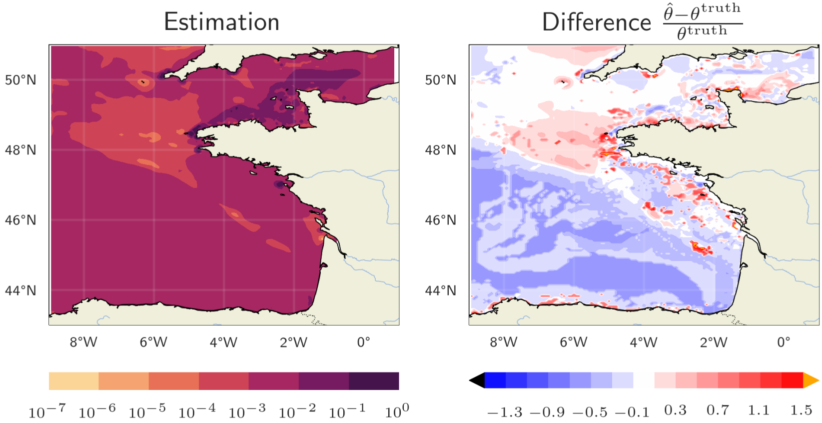
We can get an estimate by solving an optimisation problem:

$$\min_{\theta \in \Theta} J(\theta, u) = J(\hat{\theta}, u) \tag{2}$$

- $\hat{\theta}$ depends inherently on u
- What if u is uncertain by nature ?
- Does $\hat{\theta}$ compensate the errors brought by variability? ($\sim$ overfitting)

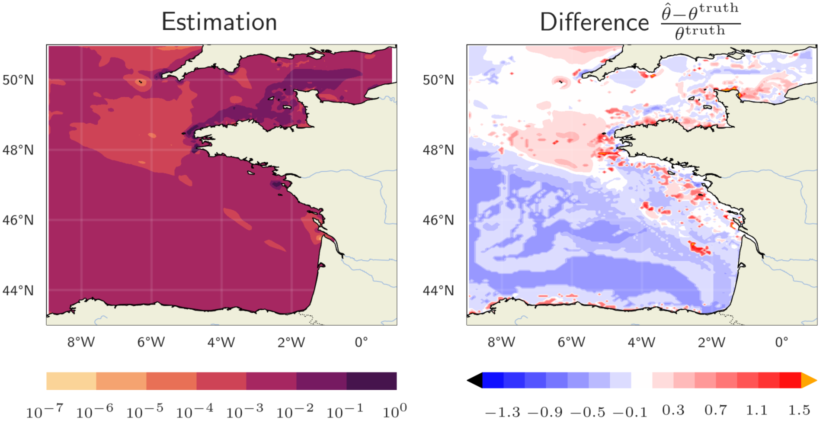
An example: Misspecification of u

Minimization of $\theta \mapsto J(\theta, u)$, for different u, which parametrizes some boundary conditions:
$u = (0.0, 0.0)$



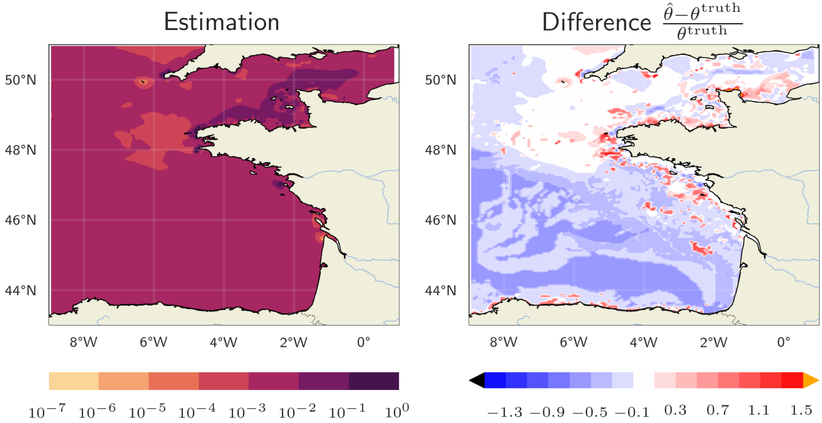
An example: Misspecification of u

Minimization of $\theta \mapsto J(\theta, u)$, for different u, which parametrizes some boundary conditions:
$u = (0.0, 0.5)$



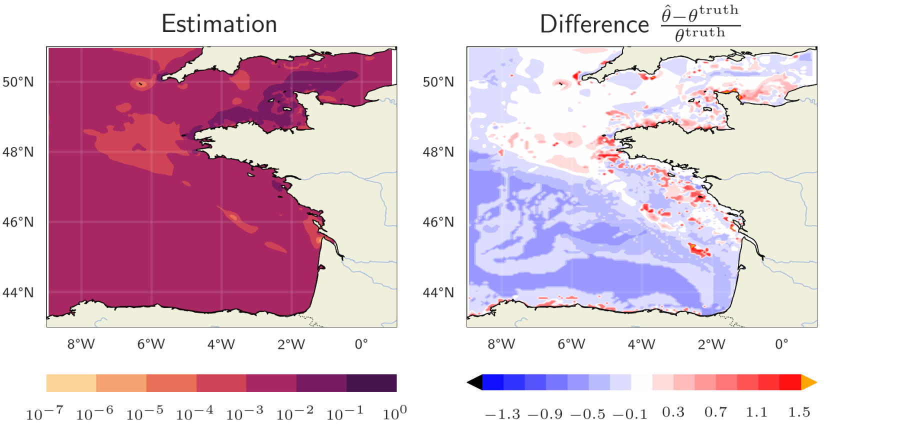
An example: Misspecification of u

Minimization of $\theta \mapsto J(\theta, u)$, for different u, which parametrizes some boundary conditions:
$u = (0.0, 1.0)$



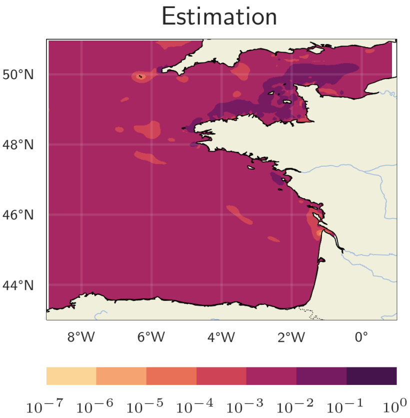
An example: Misspecification of u

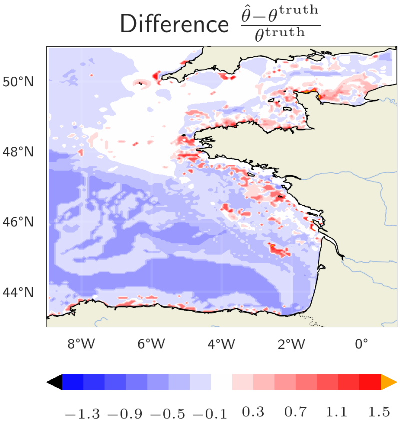
Minimization of $\theta \mapsto J(\theta, u)$, for different u, which parametrizes some boundary conditions:
$u = (0.5, 0.0)$



An example: Misspecification of u

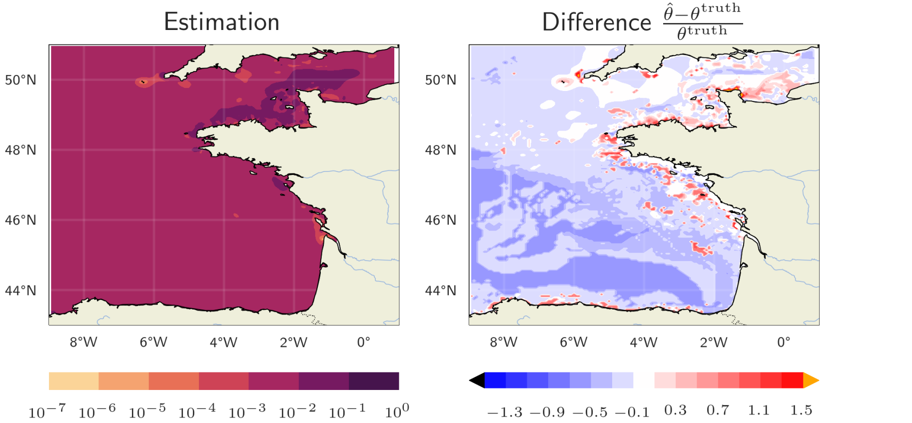
Minimization of $\theta \mapsto J(\theta, u)$, for different u, which parametrizes some boundary conditions: $u = (0.5, 0.5)$





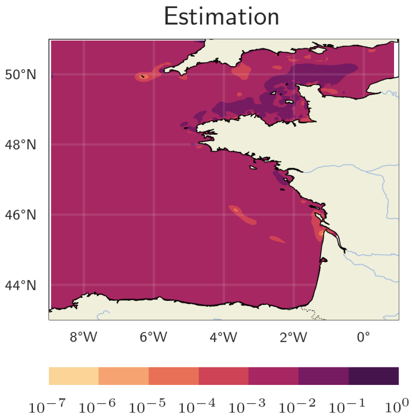
An example: Misspecification of u

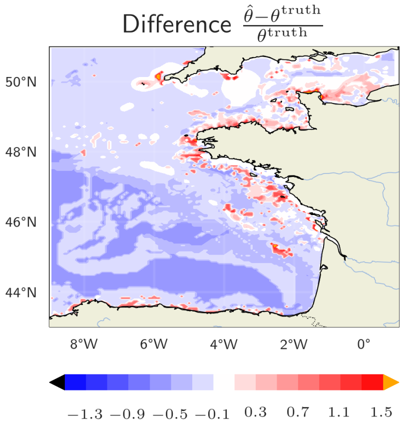
Minimization of $\theta \mapsto J(\theta, u)$, for different u, which parametrizes some boundary conditions: $u = (0.5, 1.0)$



An example: Misspecification of u

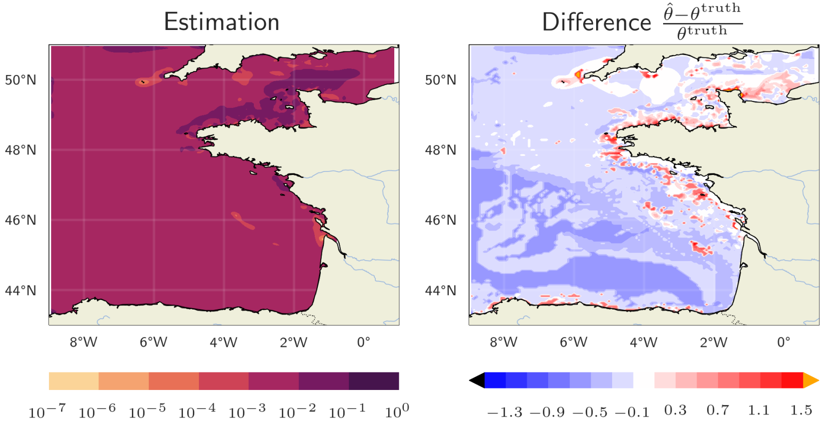
Minimization of $\theta \mapsto J(\theta, u)$, for different u, which parametrizes some boundary conditions:
$u = (1.0, 0.0)$



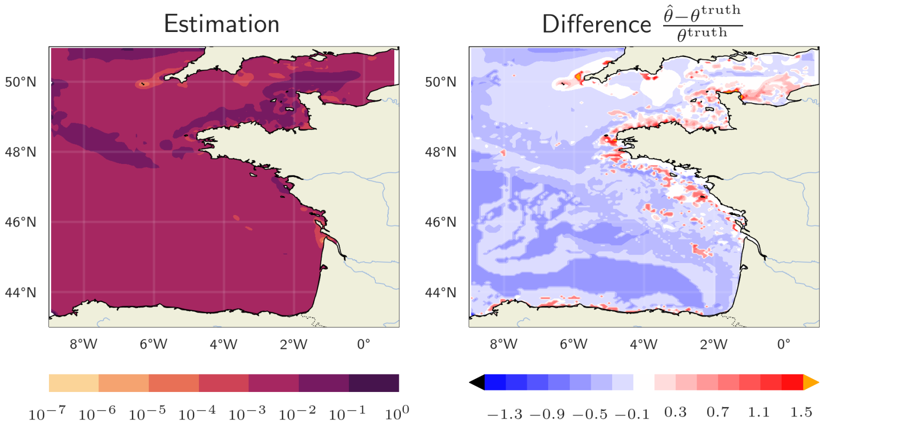


An example: Misspecification of u

Minimization of $\theta \mapsto J(\theta, u)$, for different u, which parametrizes some boundary conditions:
$u = (1.0, 0.5)$



Minimization of $\theta \mapsto J(\theta, u)$, for different u, which parametrizes some boundary conditions: $u = (1.0, 1.0)$



A definition of robustness ?

Robustness

$\hat{\theta}$ can be considered "robust" if $J(\hat{\theta}, u)$ gives "good enough" performances when u varies

Main objectives:

- Define quantitative criteria of robustness
- Develop methods in order to compute robust estimates efficiently
- Apply those methods to the robust calibration of CROCO

Outline

Robustness in calibration

Different types of uncertainties

Epistemic or aleatoric uncertainties? [Walker *et al.* , 2003]

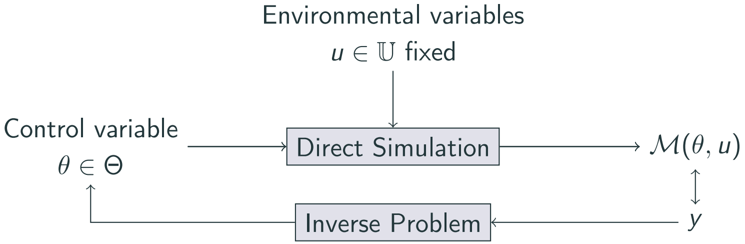
- Epistemic uncertainties: From a lack of knowledge, that can be reduced with more research/exploration
- Aleatoric uncertainties: Inherent variability of the system studied, operating conditions that we cannot afford to research further

- θ: Control parameter, needs to be tuned
- u: Environmental variable, subject to natural variability

Our goal is to take into account the aleatoric uncertainties (=the assumed variability of the environmental conditions) in the estimation of the control parameter (=reduction of the epistemic uncertainties).

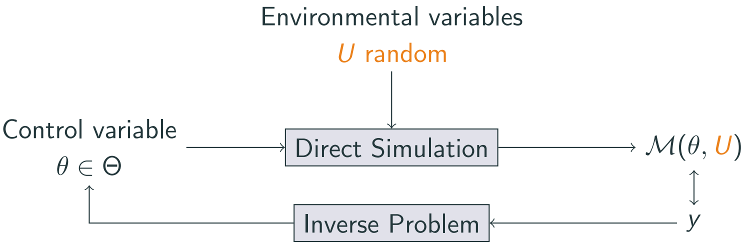
Aleatoric uncertainty as a random variable

- U: random variable of known distribution, with support $\mathbb{U}$
- u is a sample of U



Aleatoric uncertainty as a random variable

- U: random variable of known distribution, with support $\mathbb{U}$
- u is a sample of U



The objective function as a random variable

- The computer code is *still* deterministic, and takes θ and u as inputs (from the user):

$$\mathcal{M}(\theta, u)$$

- Due to the previous assumptions, the quadratic error J is now considered as a random variable, indexed by θ

$$J(\theta, U) = \frac{1}{2}\|\mathcal{M}(\theta, U) - y\|^2$$

Robust objectives of the objective function

We are looking for $\hat{\theta}$, such that $J(\hat{\theta}, U)$ gives “good performances”

- Worst case [Marzat *et al.* , 2013]:

$$\min_{\theta \in \Theta} \left\{ \max_{u \in \mathbb{U}} J(\theta, u) \right\}$$

- M-robustness [Lehman *et al.* , 2004]:

$$\min_{\theta \in \Theta} \mathbb{E}_U \left[J(\theta, U) \right]$$

- Multiobjective [Baudoui, 2012]:

$$\text{Pareto frontier of } \left(\mathbb{E}_U \left[J(\theta, U) \right], \mathbb{V}\mathrm{ar}_U \left[J(\theta, U) \right] \right)$$

- Reliability

$$\min_{\theta \in \Theta} Q_U(J(\theta, U); p)$$

- Regret-based estimates

The notion of relative-regret

Given $u \in \mathbb{U}$, how well can we calibrate the model ?

- The best performance is $\min_{\theta \in \Theta} J(\theta, u) = J^*(u) > 0$
- How does $J^*(u)$ compare to $J(\theta, u)$?
- What is the cost of choosing θ instead of $\theta^*(u) = \arg\min_\theta J(\theta, u)$?

The notion of relative-regret

Given $u \in \mathbb{U}$, how well can we calibrate the model ?

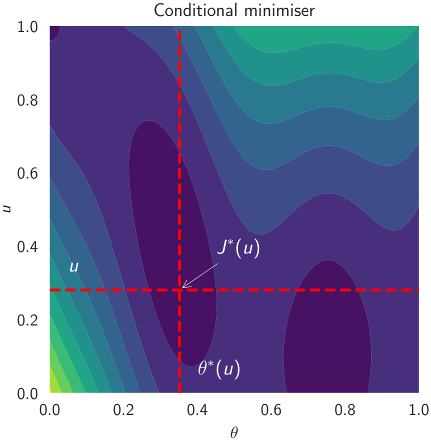
- The best performance is $\min_{\theta \in \Theta} J(\theta, u) = J^*(u) > 0$
- How does $J^*(u)$ compare to $J(\theta, u)$?
- What is the cost of choosing θ instead of $\theta^*(u) = \arg\min_\theta J(\theta, u)$?

The relative-regret, α-acceptability

- We define the relative-regret as the ratio $\frac{J(\theta,u)}{J^*(u)}$
- (θ, u) said α-acceptable if $\frac{J(\theta,u)}{J^*(u)} \leq \alpha \Rightarrow J(\theta, u) \leq \alpha J^*(u)$

Construction of regions of acceptability

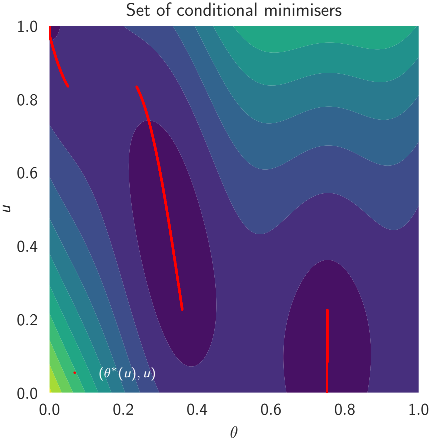
Analytical objective function $J(\theta, u) = \frac{1}{51.95}\left[\left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + (10 - \frac{10}{8\pi})\cos(x_1) - 44.81\right] + 2, \quad x_1 = 15\theta - 5, \quad x_2 = 15u$



- Sample u from U, and solve $\theta^*(u) = \arg\min_{\theta\in\Theta} J(\theta, u)$

Construction of regions of acceptability

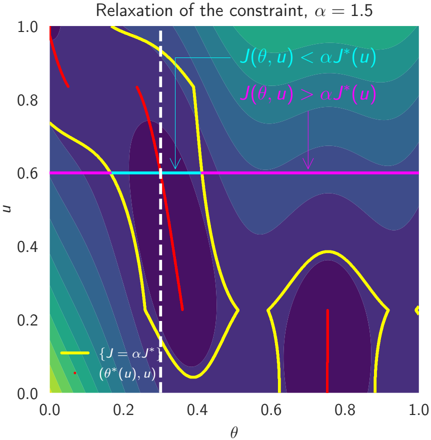
Analytical objective function $J(\theta, u) = \frac{1}{51.95}\left[\left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + \left(10 - \frac{10}{8\pi}\right)\cos(x_1) - 44.81\right] + 2, \quad x_1 = 15\theta - 5, \quad x_2 = 15u$



- Sample u from U, and solve $\theta^*(u) = \arg\min_{\theta\in\Theta} J(\theta, u)$
- Set of conditional minimisers: $\{(\theta^*(u), u) \mid u \in \mathbb{U}\}$

Construction of regions of acceptability

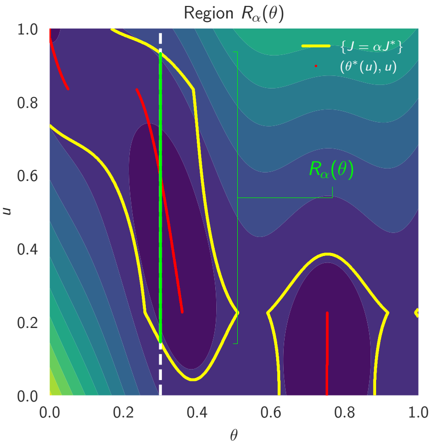
Analytical objective function $J(\theta, u) = \frac{1}{51.95}\left[\left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + \left(10 - \frac{10}{8\pi}\right)\cos(x_1) - 44.81\right] + 2, \quad x_1 = 15\theta - 5, \quad x_2 = 15u$



- Sample u from U, and solve $\theta^*(u) = \arg\min_{\theta \in \Theta} J(\theta, u)$
- Set of conditional minimisers: $\{(\theta^*(u), u) \mid u \in \mathbb{U}\}$
- Set $\alpha > 1$

Construction of regions of acceptability

Analytical objective function $J(\theta, u) = \frac{1}{51.95}\left[\left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + \left(10 - \frac{10}{8\pi}\right)\cos(x_1) - 44.81\right] + 2, \quad x_1 = 15\theta - 5, \quad x_2 = 15u$



- Sample u from U, and solve $\theta^*(u) = \arg\min_{\theta\in\Theta} J(\theta, u)$
- Set of conditional minimisers: $\{(\theta^*(u), u) \mid u \in \mathbb{U}\}$
- Set $\alpha > 1$
- $R_\alpha(\theta) = \{u \mid J(\theta, u) \leq \alpha J^*(u)\}$
- $\Gamma_\alpha(\theta) = \mathbb{P}_U\left[U \in R_\alpha(\theta)\right]$

Γ_α as a measure of robustness

- probability that θ is close enough to the optimal value
- probability that the relative-regret is less than α

Regret-based estimates

Relative-regret family of estimators [Trappler *et al.* , 2020]

Let $J > 0 \quad \Gamma_\alpha(\theta) = \mathbb{P}_U[J(\theta, U) \leq \alpha J^*(U)] = \mathbb{P}_U[\frac{J(\theta,U)}{J^*(U)} \leq \alpha]$

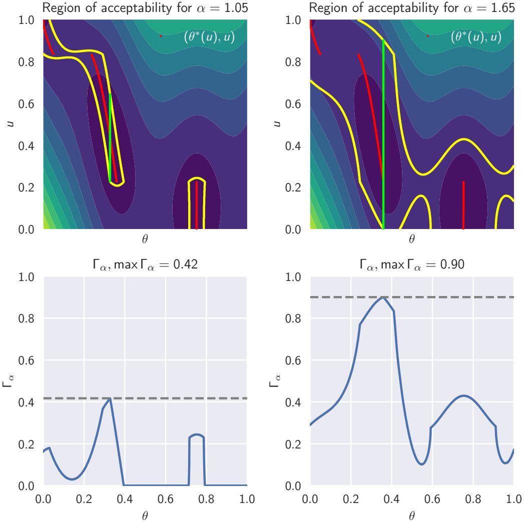
Relative-regret family of estimators:

$$\left\{\hat{\theta} \mid \hat{\theta} = \operatorname*{arg\,max}_{\theta\in\Theta} \Gamma_\alpha(\theta), \alpha > 1\right\} \tag{3}$$

- $\hat{\theta}$: calibrated value of the control parameter
- α: range of variation from the optimal value
- $p = \Gamma_\alpha(\hat{\theta})$: probability of being within this range

$\rightarrow$ the relative-regret is bounded by α with probability p

Conservative or optimistic estimate



By acting either on α or on p, we can choose a level of robustness

- α small: Optimistic, good performances but maybe not often
- α large: Conservative, performances controlled with high probability
- Choose p, in order to ensure a certain level of confidence

Relative-regret estimations as optimisation problems

- If α known, maximize the probability that θ gives acceptable values:

$$\max_{\theta \in \Theta} \Gamma_\alpha(\theta) = \max_{\theta \in \Theta} \mathbb{P}_U \left[J(\theta, U) \leq \alpha J^*(U) \right] \tag{4}$$

$\rightarrow$ Probability maximization

- Set a target probability p, find the smallest α such the probability p is reached

$$\inf \{ \alpha \mid \max_{\theta \in \Theta} \Gamma_\alpha(\theta) \geq p \} \tag{5}$$

$\rightarrow$ Quantile minimization

Relative-regret estimations as optimisation problems

- If α known, maximize the probability that θ gives acceptable values:

$$\max_{\theta \in \Theta} \Gamma_\alpha(\theta) = \max_{\theta \in \Theta} \mathbb{P}_U \left[J(\theta, U) \leq \alpha J^*(U) \right] \tag{4}$$

$\rightarrow$ Probability maximization

- Set a target probability p, find the smallest α such the probability p is reached

$$\inf\{\alpha \mid \max_{\theta \in \Theta} \Gamma_\alpha(\theta) \geq p\} \tag{5}$$

$\rightarrow$ Quantile minimization

Adaptive strategies using Gaussian Processes

The computational bottleneck

$$\max_{\theta\in\Theta} \Gamma_\alpha(\theta) = \underbrace{\max_{\theta\in\Theta}}_{\text{expensive}} \overbrace{\mathbb{P}_U}^{\text{expensive}} \left[\underbrace{J(\theta, U)}_{\text{expensive}} \leq \alpha \overbrace{J^*(U)}^{\text{expensive}}\right] \tag{6}$$

In general, getting estimates can be very expensive:

- Estimate statistical quantities ($\mathbb{E}_U$, $\mathbb{P}_U$)
  - $\rightarrow$ Sufficient exploration of $\mathbb{U}$ with respect to $\mathbb{P}_U$ (Monte-Carlo methods, numerical integration)
- Optimize those quantities with respect to $\theta \in \Theta$
  - $\rightarrow$ Focus on regions of interest of Θ
  - $\rightarrow$ Take into account the uncertainty on the estimation

$\Rightarrow$ requires a lot of computational effort (*i.e.* extensive number of calls to J)

How to make the best use of a specific budget of evaluations ?

Surrogates and cost function

Given a set comprising points and their evaluations $\mathcal{X} = \{(x_i, J(x_i))\}_{1 \leq i \leq n}$ (*training set*), we can construct an approximation of the expensive function J

$\rightarrow$ Polynomial interpolation, Gaussian Process Regression (Kriging), Polynomial Chaos Expansion

Surrogates and cost function

Given a set comprising points and their evaluations $\mathcal{X} = \{(x_i, J(x_i))\}_{1 \leq i \leq n}$ (*training set*), we can construct an approximation of the expensive function J

$\rightarrow$ Polynomial interpolation, Gaussian Process Regression (Kriging), Polynomial Chaos Expansion

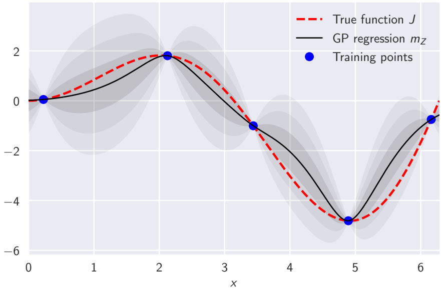
- It replaces the expensive original function J by a computationally cheap surrogate ($\sim$ plug-in approach)
- It can be adapted for sequential strategies

Let $x = (\theta, u) \in \Theta \times \mathbb{U} = \mathbb{X}$, $J(x) = J(\theta, u)$
$\mathcal{X} = \{(x_i, J(x_i))\}_{1 \leq i \leq N}$ initial design of experiments ($\sim$ training points)

GP regression [Matheron, 1962; Krige, 1951]

$Z \sim \mathrm{GP}\left(m_Z, C_Z\right)$ is the GP constructed on $\mathcal{X}$ with $m_Z : \mathbb{X} \to \mathbb{R}$ and $C_Z : \mathbb{X}^2 \to \mathbb{R}$

- m_Z: GP (or kriging) regression
- C_Z: covariance function
- $\sigma_Z^2 : x \mapsto C_Z(x, x)$ variance function
- $Z(x) \sim \mathcal{N}\left(m_Z(x), \sigma_Z^2(x)\right)$



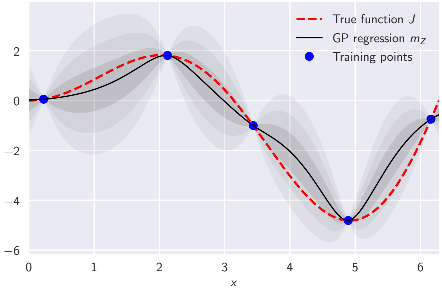
Gaussian Process Regression

Let $x = (\theta, u) \in \Theta \times \mathbb{U} = \mathbb{X}$, $J(x) = J(\theta, u)$
$\mathcal{X} = \{(x_i, J(x_i))\}_{1 \leq i \leq N}$ initial design of experiments ($\sim$ training points)

GP regression [Matheron, 1962; Krige, 1951]

$Z \sim \mathrm{GP}(m_Z, C_Z)$ is the GP constructed on $\mathcal{X}$ with $m_Z : \mathbb{X} \to \mathbb{R}$ and $C_Z : \mathbb{X}^2 \to \mathbb{R}$

- m_Z: GP (or kriging) regression
- C_Z: covariance function
- $\sigma_Z^2 : x \mapsto C_Z(x, x)$ variance function
- $Z(x) \sim \mathcal{N}\left(m_Z(x), \sigma_Z^2(x)\right)$



$\rightarrow$ m_Z is an approximation of J

$\rightarrow$ Information on prediction error with σ_Z^2

Adaptive strategies

$Z \sim \mathrm{GP}(m_Z, C_Z)$ constructed using $\mathcal{X}_n = \{(x_i, J(x_i))\}_{1\leq i\leq n}$

As a surrogate, m_Z is a cheap alternative to J in order to compute Γ_α,

- still an approximation based on $\mathcal{X}_n$
- is it accurate enough for this purpose ?

Adaptive methods: enrichment of the design

- Choosing iteratively the most relevant point(s) to add to the design
- ... until the budget of evaluations runs out

$\Rightarrow$ Definition of a criterion κ, that measures the relevancy

$\Rightarrow$ How to exploit such criterion ?

Objectives of adaptive strategies

Recalling that $\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$

- m_Z should be a good approximation of J
- m_Z should be strictly positive
- Good approximation of J^* ?

Objectives of adaptive strategies

Recalling that $\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$

- m_Z should be a good approximation of J
- m_Z should be strictly positive
- Good approximation of J^* ?

Objectives of adaptive strategies

Recalling that $\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$

- m_Z should be a good approximation of J
- m_Z should be strictly positive
- Good approximation of J^* ?

Objectives of adaptive strategies

Recalling that $\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$

- m_Z should be a good approximation of J
- m_Z should be strictly positive
- Good approximation of J^* ?

- Enrich the design for the estimation of J^* and ensure positivity as well
  $\rightarrow$ PEI criterion
- Improve the estimation of $J - \alpha J^*$ globally
  $\rightarrow$ Reduction of the IMSE using a 1-step criterion
- Improve the estimation of the set $\{J - \alpha J^* \leq 0\}$
  $\rightarrow$ AK-MCS: enrichment using batches of points

1-step criteria

Let κ be such a criterion measuring the relevancy of points:

- κ is constructed using the properties of the GP
- $\kappa(x)$ measures how *interesting* would be the evaluation of x

Selection of the next point and update of the design

Given a GP Z, constructed using $\mathcal{X}_n = \{(x_i, J(x_i))\}_{1 \leq i \leq n}$

$$x_{n+1} = \arg\max_{x \in \mathbb{X}} \kappa(x) = \arg\max_{x \in \mathbb{X}} \kappa(x; Z; \mathcal{X}_n) \tag{7}$$

$$\mathcal{X}_{n+1} = \mathcal{X}_n \cup \{(x_{n+1}, J(x_{n+1}))\} \tag{8}$$

1-step criteria

Let κ be such a criterion measuring the relevancy of points:

- κ is constructed using the properties of the GP
- $\kappa(x)$ measures how *interesting* would be the evaluation of x

Selection of the next point and update of the design

Given a GP Z, constructed using $\mathcal{X}_n = \{(x_i, J(x_i))\}_{1\leq i\leq n}$

$$x_{n+1} = \arg\max_{x\in\mathbb{X}} \kappa(x) = \arg\max_{x\in\mathbb{X}} \kappa(x; Z; \mathcal{X}_n) \tag{7}$$

$$\mathcal{X}_{n+1} = \mathcal{X}_n \cup \{(x_{n+1}, J(x_{n+1}))\} \tag{8}$$

Different criteria caters to different objectives

- Optimization: PI, EGO [Jones *et al.* , 1998; Hernández-Lobato *et al.* , 2014]
- Exploration: prediction variance, aIMSE
- Contour/levelsets estimation: reliability index [Bect *et al.* , 2012; Picheny *et al.* , 2010]

Estimation of $\theta^*(u)$, $J^*(u)$ with the PEI

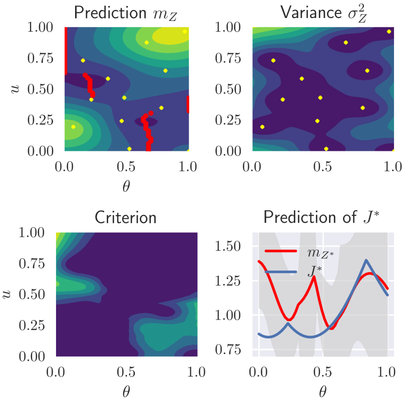
$$\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$$

$\rightarrow$ Need to have a good approximation of the conditional minimum and minimisers: $m_{Z^*}(u) = \min_\theta m_Z(\theta, u)$, and $\theta^*_Z(u) = \arg\min_\theta m_Z(\theta, u)$

PEI [Ginsbourger *et al.* , 2014]

Profile Expected Improvement: $\kappa(\theta, u) = \mathbb{E}_{Z(\theta,u)}\left[(f_{\min}(u) - Z(\theta, u))_+\right]$ and close form available due to the GP properties of Z

Design enriched iteratively using the PEI criterion, one point at a time: m_{Z^*} becomes a more accurate approximation of J^*



Estimation of $\theta^*(u)$, $J^*(u)$ with the PEI

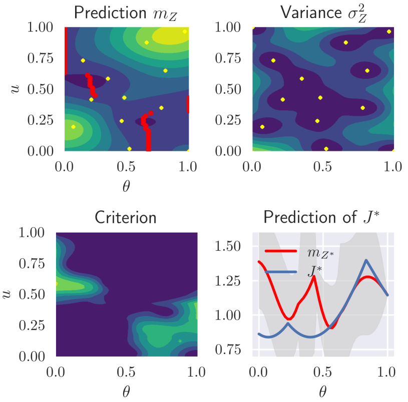
$\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$

$\rightarrow$ Need to have a good approximation of the conditional minimum and minimisers:
$m_{Z^*}(u) = \min_\theta m_Z(\theta, u)$, and $\theta^*_Z(u) = \arg\min_\theta m_Z(\theta, u)$

PEI [Ginsbourger *et al.* , 2014]

Profile Expected Improvement:
$\kappa(\theta, u) = \mathbb{E}_{Z(\theta,u)}\left[(f_{\min}(u) - Z(\theta, u))_+\right]$ and close form available due to the GP properties of Z

Design enriched iteratively using the PEI criterion, one point at a time:
m_{Z^*} becomes a more accurate approximation of J^*



Estimation of $\theta^*(u)$, $J^*(u)$ with the PEI

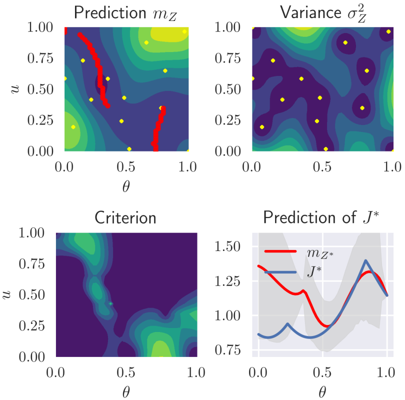
$$\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$$

$\rightarrow$ Need to have a good approximation of the conditional minimum and minimisers:
$m_{Z^*}(u) = \min_\theta m_Z(\theta, u)$, and $\theta^*_Z(u) = \arg\min_\theta m_Z(\theta, u)$

PEI [Ginsbourger *et al.* , 2014]

Profile Expected Improvement:
$\kappa(\theta, u) = \mathbb{E}_{Z(\theta,u)}\left[(f_{\min}(u) - Z(\theta, u))_+\right]$ and close form available due to the GP properties of Z

Design enriched iteratively using the PEI criterion, one point at a time:
m_{Z^*} becomes a more accurate approximation of J^*



Estimation of $\theta^*(u)$, $J^*(u)$ with the PEI

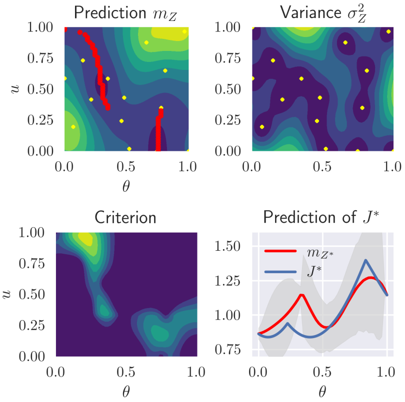
$$\Gamma_\alpha(\theta) = \mathbb{P}_U \left[J(\theta, U) \leq \alpha J^*(U) \right]$$

$\rightarrow$ Need to have a good approximation of the conditional minimum and minimisers: $m_{Z^*}(u) = \min_\theta m_Z(\theta, u)$, and $\theta^*_Z(u) = \arg\min_\theta m_Z(\theta, u)$

PEI [Ginsbourger *et al.* , 2014]

Profile Expected Improvement:
$\kappa(\theta, u) = \mathbb{E}_{Z(\theta,u)} \left[(f_{\min}(u) - Z(\theta, u))_+ \right]$ and close form available due to the GP properties of Z

Design enriched iteratively using the PEI criterion, one point at a time:
m_{Z^*} becomes a more accurate approximation of J^*



Estimation of $\theta^*(u)$, $J^*(u)$ with the PEI

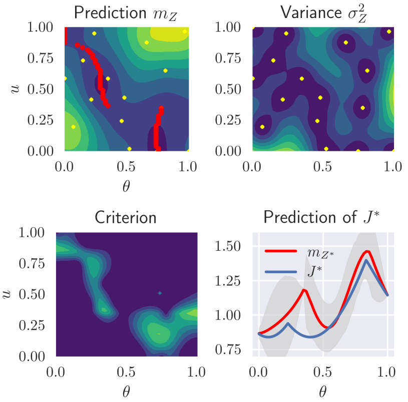
$$\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$$

$\rightarrow$ Need to have a good approximation of the conditional minimum and minimisers: $m_{Z^*}(u) = \min_\theta m_Z(\theta, u)$, and $\theta^*_Z(u) = \arg\min_\theta m_Z(\theta, u)$

PEI [Ginsbourger *et al.* , 2014]

Profile Expected Improvement:
$\kappa(\theta, u) = \mathbb{E}_{Z(\theta,u)}\left[(f_{\min}(u) - Z(\theta, u))_+\right]$ and close form available due to the GP properties of Z

Design enriched iteratively using the PEI criterion, one point at a time:
m_{Z^*} becomes a more accurate approximation of J^*



Estimation of $\theta^*(u)$, $J^*(u)$ with the PEI

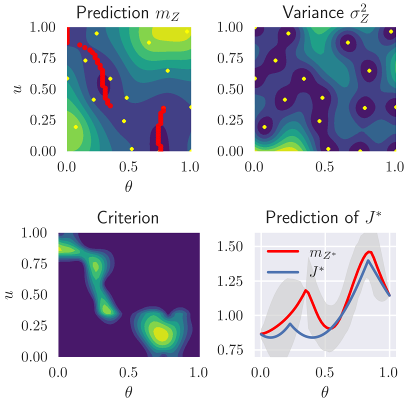
$$\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$$

$\rightarrow$ Need to have a good approximation of the conditional minimum and minimisers:
$m_{Z^*}(u) = \min_\theta m_Z(\theta, u)$, and $\theta^*_Z(u) = \arg\min_\theta m_Z(\theta, u)$

PEI [Ginsbourger *et al.* , 2014]

Profile Expected Improvement:
$\kappa(\theta, u) = \mathbb{E}_{Z(\theta,u)}\left[(f_{\min}(u) - Z(\theta, u))_+\right]$ and close form available due to the GP properties of Z

Design enriched iteratively using the PEI criterion, one point at a time:
m_{Z^*} becomes a more accurate approximation of J^*



Estimation of $\theta^*(u)$, $J^*(u)$ with the PEI

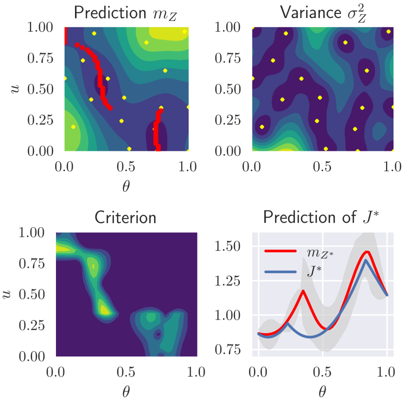
$\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$

$\rightarrow$ Need to have a good approximation of the conditional minimum and minimisers:
$m_{Z^*}(u) = \min_\theta m_Z(\theta, u)$, and $\theta^*_Z(u) = \arg\min_\theta m_Z(\theta, u)$

PEI [Ginsbourger *et al.* , 2014]

Profile Expected Improvement:
$\kappa(\theta, u) = \mathbb{E}_{Z(\theta,u)}\left[(f_{\min}(u) - Z(\theta, u))_+\right]$ and close form available due to the GP properties of Z

Design enriched iteratively using the PEI criterion, one point at a time:
m_{Z^*} becomes a more accurate approximation of J^*



Estimation of $\theta^*(u)$, $J^*(u)$ with the PEI

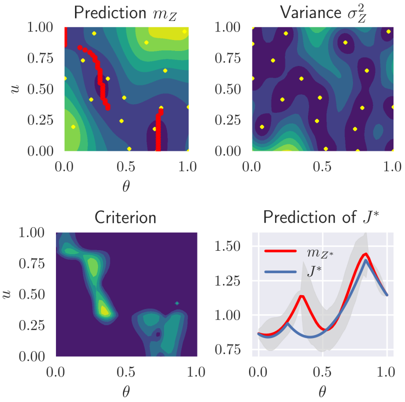
$$\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$$

$\rightarrow$ Need to have a good approximation of the conditional minimum and minimisers:
$m_{Z^*}(u) = \min_\theta m_Z(\theta, u)$, and $\theta^*_Z(u) = \arg\min_\theta m_Z(\theta, u)$

PEI [Ginsbourger *et al.* , 2014]

Profile Expected Improvement:
$\kappa(\theta, u) = \mathbb{E}_{Z(\theta,u)}\left[(f_{\min}(u) - Z(\theta, u))_+\right]$ and close form available due to the GP properties of Z

Design enriched iteratively using the PEI criterion, one point at a time:
m_{Z^*} becomes a more accurate approximation of J^*



Estimation of $\theta^*(u)$, $J^*(u)$ with the PEI

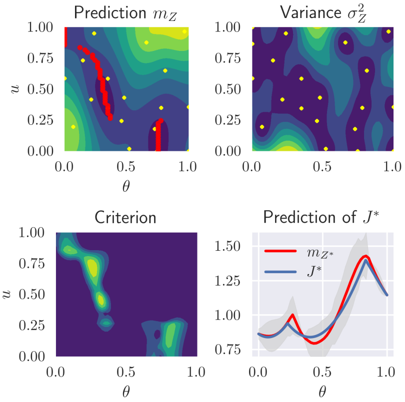
$$\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$$

$\rightarrow$ Need to have a good approximation of the conditional minimum and minimisers:
$m_{Z^*}(u) = \min_\theta m_Z(\theta, u)$, and $\theta^*_Z(u) = \arg\min_\theta m_Z(\theta, u)$

PEI [Ginsbourger *et al.* , 2014]

Profile Expected Improvement:
$\kappa(\theta, u) = \mathbb{E}_{Z(\theta,u)}\left[(f_{\min}(u) - Z(\theta, u))_+\right]$ and close form available due to the GP properties of Z

Design enriched iteratively using the PEI criterion, one point at a time:
m_{Z^*} becomes a more accurate approximation of J^*



Estimation of $\theta^*(u)$, $J^*(u)$ with the PEI

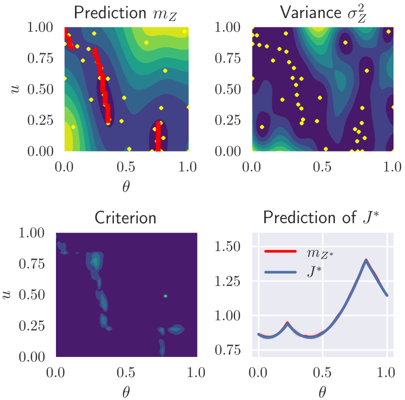
$$\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) \leq \alpha J^*(U)\right]$$

$\rightarrow$ Need to have a good approximation of the conditional minimum and minimisers:
$m_{Z^*}(u) = \min_\theta m_Z(\theta, u)$, and $\theta^*_Z(u) = \arg\min_\theta m_Z(\theta, u)$

PEI [Ginsbourger *et al.* , 2014]

Profile Expected Improvement:
$\kappa(\theta, u) = \mathbb{E}_{Z(\theta,u)}\left[(f_{\min}(u) - Z(\theta, u))_+\right]$ and close form available due to the GP properties of Z

Design enriched iteratively using the PEI criterion, one point at a time:
m_{Z^*} becomes a more accurate approximation of J^*



Taking into account J and J^*

$\rightarrow \Gamma_\alpha$ involves interaction between J and J^*

$$\Gamma_\alpha(\theta) = \mathbb{P}_U\Big[\underbrace{J(\theta, u) \leq \alpha J^*(U)}_{\text{approximation ?}}\Big] \tag{9}$$

$$= \mathbb{P}_U\left[\underbrace{\frac{J(\theta, U)}{J^*(U)}}_{\text{approximation ?}} \leq \alpha\right] \tag{10}$$

GP formulation of the difference and of the ratio

- $\Delta_\alpha = Z - \alpha Z^*$ is a GP
  $\Rightarrow \Delta_\alpha \sim \mathrm{GP}(m_{\Delta_\alpha}, C_{\Delta_\alpha})$
- $\Xi = \log Z/Z^*$ is approximately normal if $Z^* > 0$ with high enough probability
  $\Rightarrow \Xi \sim \mathrm{GP}(m_\Xi, C_\Xi)$

Taking into account J and J^*

$\rightarrow$ Γ_α involves interaction between J and J^*

$$\Gamma_\alpha(\theta) = \mathbb{P}_U\Big[\underbrace{J(\theta, u) \leq \alpha J^*(U)}_{\text{approximation ?}}\Big] \tag{9}$$

$$= \mathbb{P}_U\left[\underbrace{\frac{J(\theta, U)}{J^*(U)}}_{\text{approximation ?}} \leq \alpha\right] \tag{10}$$

GP formulation of the difference and of the ratio

- $\Delta_\alpha = Z - \alpha Z^*$ is a GP
  $\Rightarrow$ $\Delta_\alpha \sim \mathrm{GP}(m_{\Delta_\alpha}, C_{\Delta_\alpha})$
- $\Xi = \log Z/Z^*$ is approximately normal if $Z^* > 0$ with high enough probability
  $\Rightarrow$ $\Xi \sim \mathrm{GP}(m_\Xi, C_\Xi)$

Improving the plug-in estimation of Δ_α

$\rightarrow$ We want to define a criterion κ in order to reduce the global uncertainty

Integrated Mean Square Error

Uncertainty associated with the GP Δ_α using the design $\mathcal{X}_n$

$$\mathrm{IMSE}(\mathcal{X}_n) = \int_{\mathbb{X}} \sigma^2_{\Delta_\alpha}(x)\,\mathrm{d}x \tag{11}$$

After having chosen *and evaluated* the next point x_+, we want the IMSE to be as small as possible

$$\min_{x_+} \overbrace{\mathrm{IMSE}\left(\mathcal{X}_n \cup \{(x_+, \underbrace{J(x_+)}_{\text{unknown}})\}\right)}^{\kappa} \tag{12}$$

Improving the plug-in estimation of Δ_α

$\rightarrow$ We want to define a criterion κ in order to reduce the global uncertainty

Integrated Mean Square Error

Uncertainty associated with the GP Δ_α using the design $\mathcal{X}_n$

$$\mathrm{IMSE}(\mathcal{X}_n) = \int_{\mathbb{X}} \sigma^2_{\Delta_\alpha}(x)\,\mathrm{d}x \tag{11}$$

After having chosen *and evaluated* the next point x_+, we want the IMSE to be as small as possible

$$\min_{x_+} \overbrace{\mathrm{IMSE}\left(\mathcal{X}_n \cup \{(x_+, \underbrace{Z(x_+)}_{\text{r.v.}})\}\right)}^{\kappa} \tag{12}$$

Improving the plug-in estimation of Δ_α

$\rightarrow$ We want to define a criterion κ in order to reduce the global uncertainty

Integrated Mean Square Error

Uncertainty associated with the GP Δ_α using the design $\mathcal{X}_n$

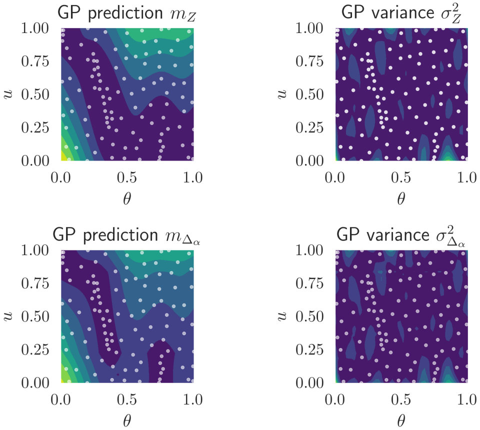
$$\mathrm{IMSE}(\mathcal{X}_n) = \int_{\mathbb{X}} \sigma^2_{\Delta_\alpha}(x)\,\mathrm{d}x \tag{11}$$

After having chosen *and evaluated* the next point x_+, we want the IMSE to be as small as possible

$$\min_{x_+} \overbrace{\mathbb{E}_{Z(x_+)}\left[\mathrm{IMSE}\left(\mathcal{X}_n \cup \{(x_+, \underbrace{Z(x_+)}_{\text{r.v.}})\}\right)\right]}^{\kappa} = \min_{x_+} \kappa(x_+) \tag{12}$$

x_{n+1} presents the smallest IMSE on average once evaluated

Numerical illustration for Δ_α



- Exploration of the whole input space $\mathbb{X} = \Theta \times \mathbb{U}$, but intensification also near the conditional minimisers
- only 1 point added at a time

$\rightarrow$ Adding multiple points every iteration ?

Sampling-based methods: AK-MCS

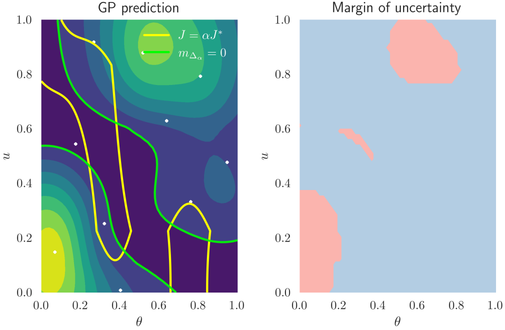
Example: Improve estimation of set

$\{m_{\Delta_\alpha} < 0\} = \{m_Z - \alpha m_{Z^*} < 0\}$

Selection of a batch of K points and update of the design [Dubourg *et al.* , 2011]

Given $\mathbb{M} \subset \mathbb{X}$, margin of uncertainty

$$g(x) = \begin{cases} 1 \text{ if } x \in \mathbb{M} \\ 0 \text{ elsewhere} \end{cases}$$



Sampling-based methods: AK-MCS

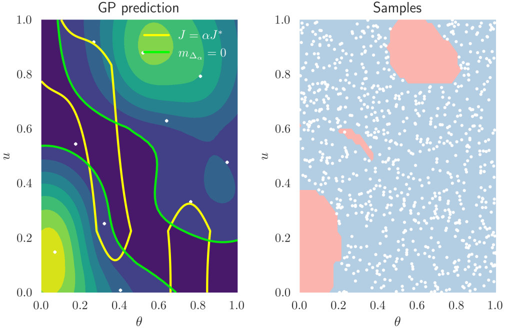
Example: Improve estimation of set
$\{m_{\Delta_\alpha} < 0\} = \{m_Z - \alpha m_{Z^*} < 0\}$

Selection of a batch of K points and update of the design [Dubourg *et al.* , 2011]

- Sample points according to the sampling density g

Given $\mathbb{M} \subset \mathbb{X}$, margin of uncertainty

$$g(x) = \begin{cases} 1 \text{ if } x \in \mathbb{M} \\ 0 \text{ elsewhere} \end{cases}$$



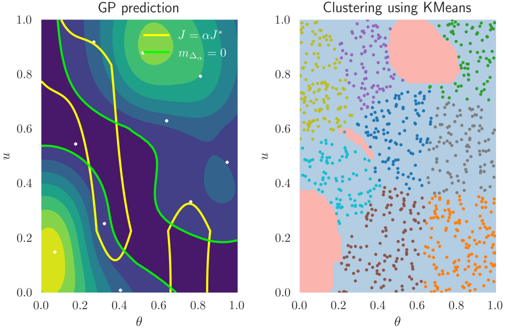
Example: Improve estimation of set
$\{m_{\Delta_\alpha} < 0\} = \{m_Z - \alpha m_{Z^*} < 0\}$

Selection of a batch of K points and update of the design [Dubourg *et al.* , 2011]

- Sample points according to the sampling density g
- Cluster the samples into K clusters (KMeans)

Given $\mathbb{M} \subset \mathbb{X}$, margin of uncertainty

$$g(x) = \begin{cases} 1 \text{ if } x \in \mathbb{M} \\ 0 \text{ elsewhere} \end{cases}$$



Sampling-based methods: AK-MCS

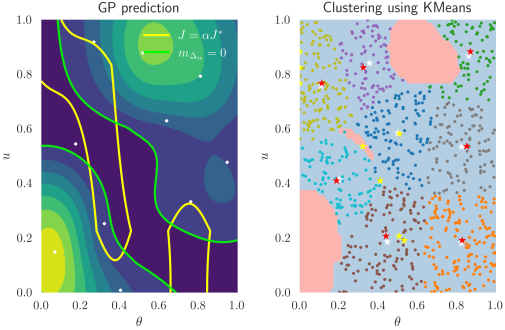
Example: Improve estimation of set
$\{m_{\Delta_\alpha} < 0\} = \{m_Z - \alpha m_{Z^*} < 0\}$

Selection of a batch of K points and update of the design [Dubourg *et al.* , 2011]

- Sample points according to the sampling density g
- Cluster the samples into K clusters (KMeans)
- For each cluster, select the sample which is closest to the cluster center
- Eventually postprocessing

Given $\mathbb{M} \subset \mathbb{X}$, margin of uncertainty

$$g(x) = \begin{cases} 1 \text{ if } x \in \mathbb{M} \\ 0 \text{ elsewhere} \end{cases}$$



Sampling-based methods: AK-MCS

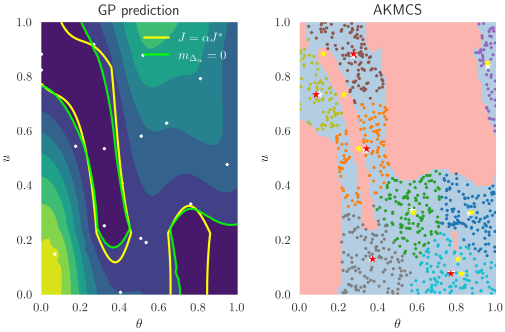
Example: Improve estimation of set
$\{m_{\Delta_\alpha} < 0\} = \{m_Z - \alpha m_{Z^*} < 0\}$

Selection of a batch of K points and update of the design [Dubourg *et al.* , 2011]

- Sample points according to the sampling density g
- Cluster the samples into K clusters (KMeans)
- For each cluster, select the sample which is closest to the cluster center
- Eventually postprocessing
- Evaluate the K chosen samples and update the design accordingly

Given $\mathbb{M} \subset \mathbb{X}$, margin of uncertainty

$$g(x) = \begin{cases} 1 \text{ if } x \in \mathbb{M} \\ 0 \text{ elsewhere} \end{cases}$$



Sampling-based methods: AK-MCS

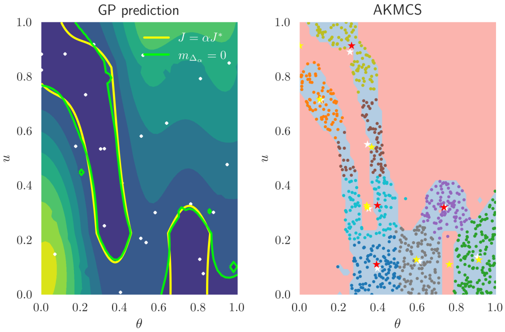
Example: Improve estimation of set
$\{m_{\Delta_\alpha} < 0\} = \{m_Z - \alpha m_{Z^*} < 0\}$

Selection of a batch of K points and update of the design [Dubourg *et al.* , 2011]

- Sample points according to the sampling density g
- Cluster the samples into K clusters (KMeans)
- For each cluster, select the sample which is closest to the cluster center
- Eventually postprocessing
- Evaluate the K chosen samples and update the design accordingly

Given $\mathbb{M} \subset \mathbb{X}$, margin of uncertainty

$$g(x) = \begin{cases} 1 \text{ if } x \in \mathbb{M} \\ 0 \text{ elsewhere} \end{cases}$$



Sampling-based methods: AK-MCS

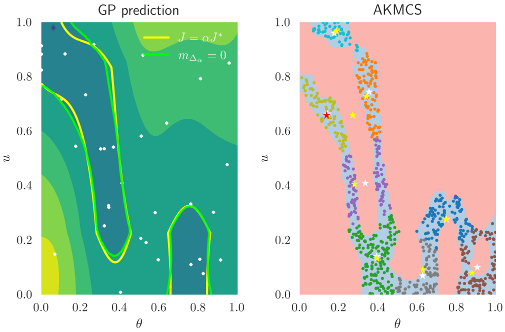
Example: Improve estimation of set
$\{m_{\Delta_\alpha} < 0\} = \{m_Z - \alpha m_{Z^*} < 0\}$

Selection of a batch of K points and update of the design [Dubourg *et al.* , 2011]

- Sample points according to the sampling density g
- Cluster the samples into K clusters (KMeans)
- For each cluster, select the sample which is closest to the cluster center
- Eventually postprocessing
- Evaluate the K chosen samples and update the design accordingly

Given $\mathbb{M} \subset \mathbb{X}$, margin of uncertainty

$$g(x) = \begin{cases} 1 \text{ if } x \in \mathbb{M} \\ 0 \text{ elsewhere} \end{cases}$$



Sampling-based methods: AK-MCS

Example: Improve estimation of set

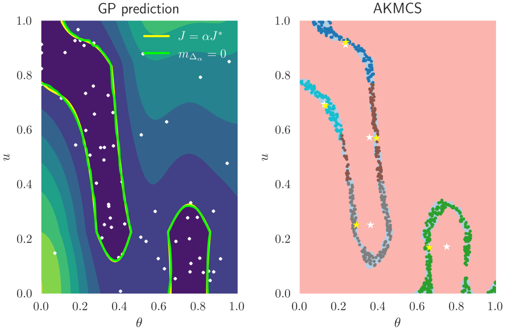
$\{m_{\Delta_\alpha} < 0\} = \{m_Z - \alpha m_{Z^*} < 0\}$

Selection of a batch of K points and update of the design [Dubourg *et al.* , 2011]

- Sample points according to the sampling density g
- Cluster the samples into K clusters (KMeans)
- For each cluster, select the sample which is closest to the cluster center
- Eventually postprocessing
- Evaluate the K chosen samples and update the design accordingly

Given $\mathbb{M} \subset \mathbb{X}$, margin of uncertainty

$$g(x) = \begin{cases} 1 \text{ if } x \in \mathbb{M} \\ 0 \text{ elsewhere} \end{cases}$$



Computations using Δ_α or Ξ

Recalling that $\Gamma_\alpha(\theta) = \mathbb{P}_U\left[J(\theta, U) - \alpha J^*(U) \leq 0\right] = \mathbb{P}_U\left[\frac{J(\theta,U)}{J^*(U)} \leq \alpha\right]$

Plug-in and sample average approximation

- J is replaced by m_Z
- $J - \alpha J^*$ is replaced by $m_{\Delta_\alpha} \to \alpha$ fixed, estimation of probability
- $\frac{J}{J^*}$ is replaced by $m_\Xi \to p$ fixed, estimation of quantile

$$\Gamma_\alpha(\theta) \approx \mathbb{P}_U\left[m_{\Delta_\alpha}(\theta, U) \leq 0\right] \tag{13}$$

$$\approx \mathbb{P}_U\left[m_\Xi(\theta, U) \leq \log \alpha\right] \tag{14}$$

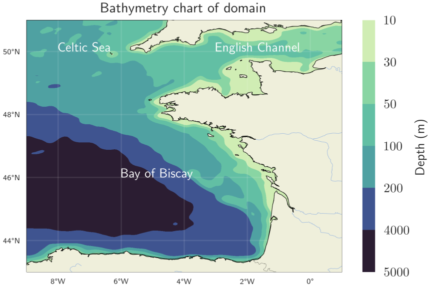
Outer probability approximated using Monte-Carlo since m_{Δ_α} and m_Ξ are cheaper to evaluate than J, and optimized

Robust calibration of CROCO

Coastal and **R**egional **O**cean **CO**mmunity model (here in a 2D configuration)

- Solves the shallow-water equations
- Grid resolution of $1/14^{\circ}$ (5.5 km)
- 15 684 cells located in the ocean

$\rightarrow$ Academic toy problem



The shallow-water equations

ζ: sea-water height, $\mathbf{v}$: velocity vector

$$\mathcal{M} : \begin{cases} \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} + 2\boldsymbol{\Omega} \wedge \mathbf{v} & = & -g\nabla H + \frac{\boldsymbol{\tau}_b}{\rho H} + F \\ \frac{\partial \zeta}{\partial t} + \nabla (H \cdot \mathbf{v}) & = & 0 \\ \zeta & = & \mathrm{BC}(u) \text{ at the open boundaries} \end{cases} \tag{15}$$

- $\boldsymbol{\tau}_b = \boldsymbol{\tau}_b(\theta)$: bottom shear stress $\rightarrow$ depends on the control parameter θ
- $\mathrm{BC}(u)$: tidal boundary conditions $\rightarrow$ subject to aleatoric uncertainties (u)

Model and objective function

- Output of the model: $\mathcal{M}(\theta, u) = (\zeta_{i,t}(\theta, u))_{\substack{1 \leq i \leq N_{\textbf{Mesh}} \\ 1 \leq t \leq N_{\textbf{time}}}}$
- Observations: $y = \mathcal{M}(\theta^{\text{truth}}, u^{\text{truth}})$, truth values defined later
- Objective function: $J(\theta, u) = \|\mathcal{M}(\theta, u) - y\|^2 = \sum_{i,t} (\zeta_{i,t}(\theta, u) - y_{i,t})^2$

Modelling of the bottom friction

Quadratic friction coefficient

Let $\boldsymbol{\tau}_b$ be the shear stress at the bottom, and v_b the velocity vector at the bottom

$$\boldsymbol{\tau}_b = -C_d \|v_b\| v_b, \quad \text{with} \quad C_d = \left(\frac{k}{\log\left(\frac{H}{z_b}\right) - 1} \right)^2 \tag{16}$$

and we define the control parameter as $\theta = \log z_b$

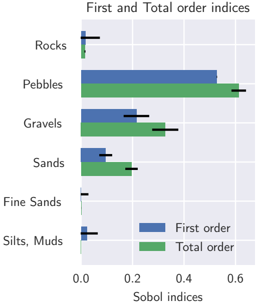
- We assume that θ is uniform and constant for each type of sediment
- θ^{truth}: constructed using the 6 types of sediments present

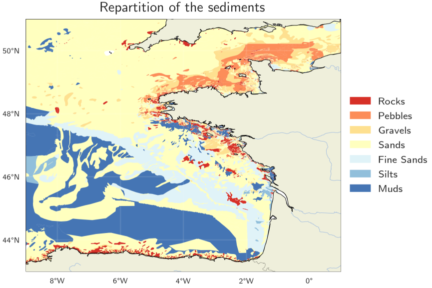
Due to the water height H, influence of each sediment type depends on how deep it is

$\Rightarrow$ Sensitivity analysis

Sensitivity analysis on the friction associated with the sediments

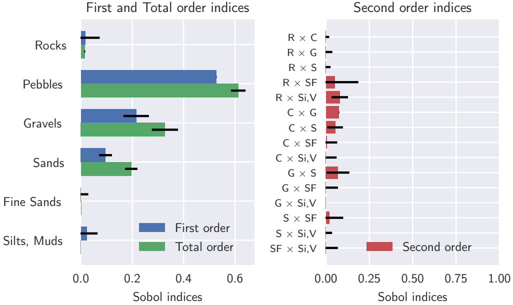
Global Sensitivity Analysis with Sobol' indices [Sobol, 2001, 1993]: quantify the influence of each input variable on J (computed using Gilquin *et al.* [2019])

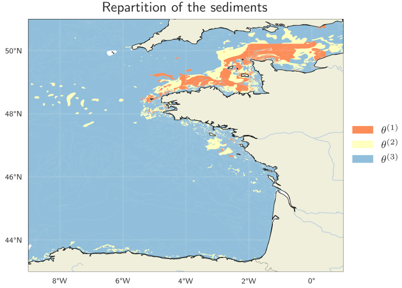




Sensitivity analysis on the friction associated with the sediments

Global Sensitivity Analysis with Sobol' indices [Sobol, 2001, 1993]: quantify the influence of each input variable on J (computed using Gilquin *et al.* [2019])





The control variable is then $\theta = (\theta^{(1)}, \theta^{(2)}, \theta^{(3)}) \in \Theta$

- $\theta^{(1)}$: Pebbles
- $\theta^{(2)}$: Gravels
- $\theta^{(3)}$: Other sediments (in deeper water)

Control parameter and environmental variable

- Aleatoric uncertainty on boundary conditions
- u parametrizes an error on the amplitude of the M_2 and S_2 tide components.
- We assume that $U \sim \mathcal{U}(\mathbb{U})$, with $\mathbb{U} = [0,1]^2$

Summary of the problem

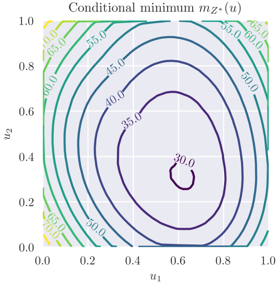
- $\theta = (\theta^{(1)}, \theta^{(2)}, \theta^{(3)}) \in \Theta \subset \mathbb{R}^3$
- $u \in \mathbb{U} = [0,1]^2$
- $U \sim \mathcal{U}(\mathbb{U})$
- $J(\theta, u) = \|\mathcal{M}(\theta, u) - y\|^2$
  - $y = \mathcal{M}(\theta^{\text{truth}}, u^{\text{truth}})$
  - $\theta^{\text{truth}} \in \mathbb{R}^6 \neq \Theta \Rightarrow$ The model is not able to replicate the observations: $J > 0$
  - $u^{\text{truth}} = (0.5, 0.5)$

Initial design, preliminary analysis

- Initial design evaluated $\mathcal{X}_{\mathrm{LHS}} = \{(x_i, J(x_i))\}_{1\leq i\leq n}$, LHS on $\Theta \times \mathbb{U}$
- $J^* > 0$ by definition, but m_{Z^*} ?

Initial design, preliminary analysis

- Initial design evaluated $\mathcal{X}_{\mathrm{LHS}} = \{(x_i, J(x_i))\}_{1\leq i\leq n}$, LHS on $\Theta \times \mathbb{U}$
- $J^* > 0$ by definition, but m_{Z^*} ?
  - Improve first m_{Z^*} using PEI criterion to ensure $m_{Z^*} > 0$



Global minimum not attained at the truth value of the environmental parameter

$\rightarrow$ Compensation of errors due to the dimension reduction

Robust calibration

With this preliminary analysis, we can choose α

- we set $\alpha = 1.3$: with which probability can we stay within 30% of the optimal value ?
- we enriched the design by looking to reduce globally the IMSE using κ

	$\theta^{(1)}$	$\theta^{(2)}$	$\theta^{(3)}$	
$\hat{\theta}_\alpha$, $\alpha = 1.3$	-3.43	-5.20	-6.48	$\Gamma_\alpha(\hat{\theta}_\alpha) = 0.93$
$\hat{\theta}_{\text{global}}$	-3.516	-5.078	-6.346	
θ^{truth}	-3.689	-4.962	n.a.	

In this case:

- $\hat{\theta}_\alpha^{(1)} > \hat{\theta}_{\text{global}}^{(1)} > \theta^{\text{truth},(1)}$
- $\hat{\theta}_\alpha^{(2)} < \hat{\theta}_{\text{global}}^{(2)} < \theta^{\text{truth},(2)}$
- $\hat{\theta}_\alpha^{(3)} < \hat{\theta}_{\text{global}}^{(3)}$

Computational overview

- Total of 500 runs of the numerical model
  - 100 for the initial LHS
  - 200 with the PEI criterion
  - 200 for the reduction of the IMSE
- Each iteration for the reduction of the IMSE requires
  - Evaluations of an integral of dimension $1 + \dim(\Theta \times \mathbb{U})$
  - Optimization of this integral in a space of dimension $\dim \Theta$

  $\rightarrow$ Dependent on the ability to compute m_{Z^*}, m_{Δ_α}, $\sigma^2_{\Delta_\alpha}$

  $\rightarrow$ As is, limits severely the possibility of increased dimension
- For sampling-based methods:
  - Size of margin of uncertainty decreases

  $\rightarrow$ Sampling becomes increasingly difficult

Conclusion

Conclusion: overview

Notion of robustness

- Notion of robustness is context dependent
- Relative-regret estimates control the deviation from the *optimal value*
- α can reflect risk-adverse or risk-seeking preferences

Adaptive methods for computations

- GP used as tools for plug-in estimation
- Allow to derive sequential strategies for enrichment of the design
- Can be adapted for batch evaluations

Conclusion: perspectives

Computational improvements

- Dimension of the input space $\Theta \times \mathbb{U}$ can be limiting
  - $\rightarrow$ Reduction of the input space required
- Adaptive methods still require expensive or difficult tasks (optimization, sampling, integration)
  - $\rightarrow$ 2-stages methods
  - $\rightarrow$ Focus computational effort for optimization
- Include gradient information in the procedures ?

Perspectives

- Discrepancy between $J^*(U)$ and $J(\hat{\theta}, U)$ in terms of r.v.
- Study of CROCO with a more complex configuration
- Study and compare robust estimates in predictions

Baudoui, Vincent. 2012. *Optimisation Robuste Multiobjectifs Par Modèles de Substitution*. Ph.D. thesis, Toulouse, ISAE.

Bect, Julien, Ginsbourger, David, Li, Ling, Picheny, Victor, & Vazquez, Emmanuel. 2012. Sequential Design of Computer Experiments for the Estimation of a Probability of Failure. *Statistics and Computing*, **22**(3), 773–793.

Dubourg, V., Sudret, B., & Bourinet, J.-M. 2011. Reliability-Based Design Optimization Using Kriging Surrogates and Subset Simulation. *Structural and Multidisciplinary Optimization*, **44**(5), 673–690.

Gilquin, Laurent, Arnaud, Elise, Prieur, Clémentine, & Janon, Alexandre. 2019. Making Best Use of Permutations to Compute Sensitivity Indices with Replicated Orthogonal Arrays. *Reliability Engineering and System Safety*, **187**(July), 28–39.

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

Hernández-Lobato, José Miguel, Hoffman, Matthew W., & Ghahramani, Zoubin. 2014. Predictive Entropy Search for Efficient Global Optimization of Black-Box Functions. *arXiv:1406.2541 [cs, stat]*, June.

Jones, Donald R., Schonlau, Matthias, & Welch, William J. 1998. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global optimization*, **13**(4), 455–492.

Krige, Daniel G. 1951. A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, **52**(6), 119–139.

Lehman, Jeffrey S., Santner, Thomas J., & Notz, William I. 2004. Designing Computer Experiments to Determine Robust Control Variables. *Statistica Sinica*, 571–590.

Marzat, Julien, Walter, Eric, & Piet-Lahanier, Hélène. 2013. Worst-Case Global Optimization of Black-Box Functions through Kriging and Relaxation. *Journal of Global Optimization*, **55**(4), 707–727.

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

Picheny, Victor, Ginsbourger, David, Roustant, Olivier, Haftka, Raphael T., & Kim, Nam-Ho. 2010. Adaptive Designs of Experiments for Accurate Approximation of a Target Region. *Journal of Mechanical Design*, **132**(7), 071008.

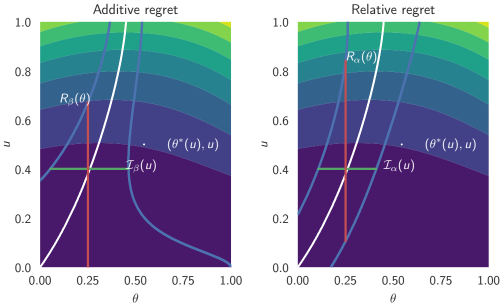
Sobol, Ilya M. 1993. Sensitivity Analysis for Non-Linear Mathematical Models. *Mathematical modelling and computational experiment*, **1**, 407–414.

Sobol, Ilya M. 2001. Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates. *Mathematics and computers in simulation*, **55**(1-3), 271–280.

Trappler, Victor, Arnaud, Élise, Vidard, Arthur, & Debreu, Laurent. 2020. Robust Calibration of Numerical Models Based on Relative Regret. *Journal of Computational Physics*, Nov., 109952.

Walker, Warren E., Harremoës, Poul, Rotmans, Jan, van der Sluijs, Jeroen P., van Asselt, Marjolein BA, Janssen, Peter, & Krayer von Krauss, Martin P. 2003. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated assessment*, **4**(1), 5–17.

Relative or additive regret



- Relative regret
  - α-acceptability regions large for flat and bad situations ($J^*(u)$ large)
  - Conversely, puts high confidence when $J^*(u)$ is small
  - No units $\rightarrow$ ratio of costs

Modelling of $J(\theta, u) - \alpha J^*(u)$ using Δ_α

Let $Z \sim \mathrm{GP}\left(m_Z; C_Z\right)$ on $\Theta \times \mathbb{U}$, constructed using $\{(\theta_i, u_i), J(\theta_i, u_i))\}$

We define $Z^*(u) = Z(\theta^*_Z(u), u)$, with $\theta^*_Z(u) = \min_\theta m_Z(\theta, u)$, and $\Delta_\alpha = Z - \alpha Z^*$:

The GP Δ_α

As a linear combination of GP, Δ_α is a GP as well:

$$\Delta_\alpha \sim \mathrm{GP}\left(m_{\Delta_\alpha}; C_{\Delta_\alpha}\right) \tag{17}$$

$$m_{\Delta_\alpha}(\theta, u) = m_Z(\theta, u) - \alpha m^*_Z(u) \tag{18}$$

$$\sigma^2_{\Delta_\alpha}(\theta, u) = \sigma^2_Z(\theta, u) + \alpha^2 \sigma^2_{Z^*}(u) - 2\alpha C_Z\left((\theta, u), (\theta^*_Z(u), u)\right) \tag{19}$$

Approximation of the ratio

Let us assume that $Z^* > 0$ with high enough probability: $\Xi(\theta, u) = \log \frac{Z(\theta,u)}{Z^*(u)}$ is approximately normal

Log-normal approximation of the ratio of GP

$$\Xi(\theta, u) \sim \mathcal{N}\left(m_\Xi(\theta, u), \sigma^2_\Xi(\theta, u)\right) \tag{20}$$

$$m_\Xi(\theta, u) = \log \frac{m_Z(\theta, u)}{m^*_Z(u)} \tag{21}$$

$$\sigma^2_\Xi(\theta, u) = \frac{\sigma^2_Z(\theta, u)}{m_Z(\theta, u)^2} + \frac{\sigma^2_{Z^*}(u)}{m_{Z^*}(u)^2} - 2\frac{\mathrm{Cov}[Z(\theta, u), Z^*(u)]}{m_Z(\theta, u)m_{Z^*}(u)} \tag{22}$$

Objective-oriented exploration: 2-stage methods

Instead of reducing *globally* the uncertainty, we can look directly to optimize Γ_α

- Select a candidate $\tilde{\theta}$ with "high potential" to optimize Γ_α
- Find the point x_{n+1} which reduces the most a measure of uncertainty on $\{\tilde{\theta}\} \times \mathbb{U}$

IMSE given a candidate $\tilde{\theta}$

$$\mathrm{IMSE}(\mathcal{X}_n; \tilde{\theta}) = \int_{\{\tilde{\theta}\} \times \mathbb{U}} \sigma_\Phi^2(x)\, \mathrm{d}x \tag{23}$$

$x_{n+1} = (\theta_{n+1}, u_{n+1})$ presents the smallest IMSE (given $\tilde{\theta}$) on average once evaluated.

$\tilde{\theta} \neq \theta_{n+1}$ in general