Parameter control in the presence of uncertainties

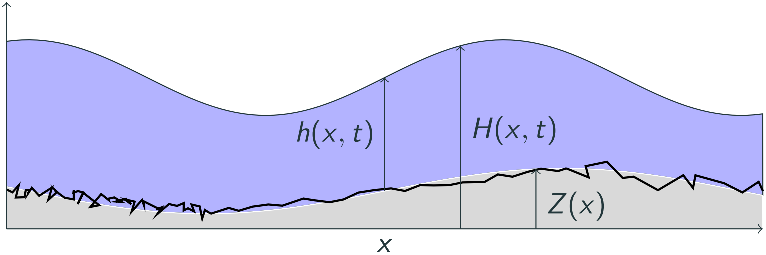
Victor Trappler
Supervisors: Élise Arnaud, Laurent Debreu, Arthur Vidard

June 15, 2018

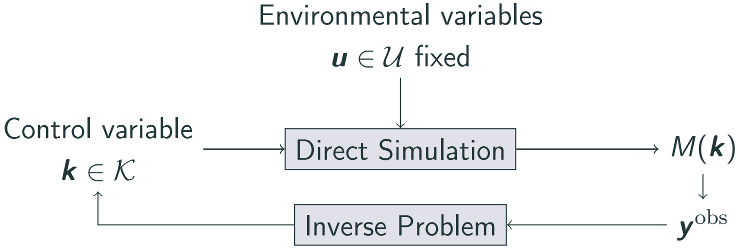
AIRSEA Research team (Inria)– Laboratoire Jean Kuntzmann

Friction de fond

- Friction du fond a une influence sur la circulation de l'eau en surgace
- Dépend de la taille caractéristique des aspérités
- Phénomène sous-maille



Modèle numérique: Shallow Water Equations



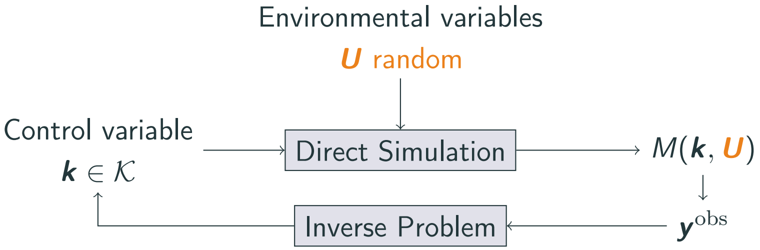
Expériences jumelles: On choisit $\boldsymbol{k}_{\mathrm{ref}}$
On a donc $\boldsymbol{y}^{\mathrm{obs}} = M(\boldsymbol{k}_{\mathrm{ref}})$

- Fonction coût: $J(\boldsymbol{k}) = \frac{1}{2}\|M(\boldsymbol{k}) - \boldsymbol{y}^{\mathrm{obs}}\|^2 +$ Régul. eventuelle
- Estimateur : $\hat{\boldsymbol{k}} = \arg\min_{\boldsymbol{k}\in\mathcal{K}} J(\boldsymbol{k})$

Incertitudes dans un code déterministe

$\boldsymbol{U}$ est maintenant une variable aléatoire (densité $\pi(\boldsymbol{u})$)

$\boldsymbol{y}^{\mathrm{obs}}$ a été générée avec $\boldsymbol{u}_{\mathrm{ref}}$, échantillon de $\boldsymbol{U}$



- $M(\boldsymbol{k})$ devient $M(\boldsymbol{k}, \boldsymbol{u})$ ($\boldsymbol{u}$ est une entrée du modèle)
- Fonction coût: $J(\boldsymbol{k}, \boldsymbol{u}) = \frac{1}{2}\|M(\boldsymbol{k}, \boldsymbol{u}) - \boldsymbol{y}^{\mathrm{obs}}\|^2$ + Régul

Estimateur robuste ?

On veut pouvoir trouver une valeur $\hat{\boldsymbol{k}}$, tel que $M(\hat{\boldsymbol{k}}, \boldsymbol{U})$ soit relativement semblable à $\boldsymbol{y}^{\mathrm{obs}}$

Idéalement, $M(\hat{\boldsymbol{k}}, \cdot)$

- reste assez performant pour fournir des prédictions acceptables
- ne varie pas trop avec $\boldsymbol{U}$

Approche variationnelle ou Bayésienne ?

- **Variationnelle**: Variable aléatoire indexée par $\boldsymbol{k}$:
  $\boldsymbol{k} \mapsto J(\boldsymbol{k}, \boldsymbol{U})$,
  Extrema des moments ? $\mathbb{E}[J(\boldsymbol{K}, \boldsymbol{U}) | \boldsymbol{K} = \boldsymbol{k}] \dots$
- **Bayésienne**: $e^{-J(\boldsymbol{k},\boldsymbol{u})} \propto p(\boldsymbol{y}^{\mathrm{obs}} | \boldsymbol{k}, \boldsymbol{u})$= Vraisemblance
  Inférence bayésienne, marginalisation, Estimation bayésienne

Mais

- Estimer efficacement moments ?
- Quelle connaissance de $\boldsymbol{U}$?
- Gérer le coût de calcul du modèle

Différents critères

- Optimum global: $\min_{(\boldsymbol{k},\boldsymbol{u})} J(\boldsymbol{k},\boldsymbol{u}) \longrightarrow$ EGO
- Pire des cas: $\min_{\boldsymbol{k}} \max_{\boldsymbol{u}} J(\boldsymbol{k},\boldsymbol{u}) \longrightarrow$ Explorative EGO

$\pi(\boldsymbol{u})$?

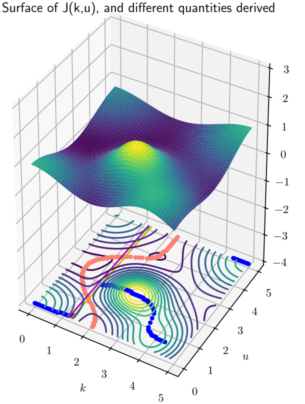
- M-robustesse: $\min_{\boldsymbol{k}} \mathbb{E}\left[J(\boldsymbol{k},\boldsymbol{U})\right] \longrightarrow$ iterated LHS
- V-robustesse: $\min_{\boldsymbol{k}} \mathbb{V}\mathrm{ar}\left[J(\boldsymbol{k},\boldsymbol{U})\right] \longrightarrow$ gradient-descent with PCE
- ρ-robustesse: $\min \rho(J(\boldsymbol{k},\boldsymbol{U})) \longrightarrow$ gradient-descent with PCE
- Multiobjectifs: estimer le front de Pareto $\longrightarrow$ 1L/2L kriging

$\mathbb{E}$, $\mathbb{V}\mathrm{ar}$

- Minimiser probabilité de défaillance: $\min_{\boldsymbol{k}} \mathbb{P}\left[J(\boldsymbol{k},\boldsymbol{U}) > T\right]$
- $\mathbb{P}\left[J(\boldsymbol{k},\boldsymbol{U}) \leq \min_{\tilde{\boldsymbol{k}}} J(\tilde{\boldsymbol{k}},\boldsymbol{U})\right] = \mathbb{P}\left[\boldsymbol{k} = \arg\min_{\tilde{\boldsymbol{k}}} J(\tilde{\boldsymbol{k}},\boldsymbol{U})\right]$
  $\rightarrow$ Distribution et mode de $\arg\min_{\tilde{\boldsymbol{k}}} J(\tilde{\boldsymbol{k}},\boldsymbol{U})$: MPE
- Relaxation de la contrainte:
  $R_\alpha(\boldsymbol{k}) = \mathbb{P}\left[J(\boldsymbol{k},\boldsymbol{U}) \leq \alpha \min_{\tilde{\boldsymbol{k}}} J(\tilde{\boldsymbol{k}},\boldsymbol{U})\right]$

Random sets

Illustration en 2D: $\mathcal{K} \times \mathcal{U} = [0; 5]^2$



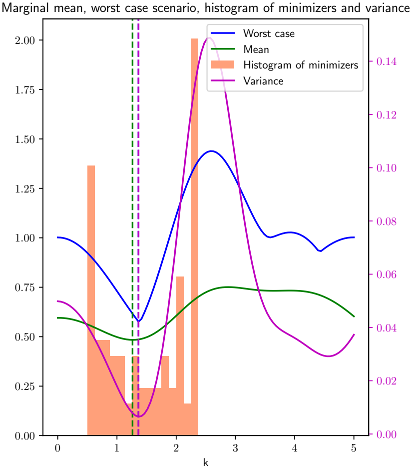
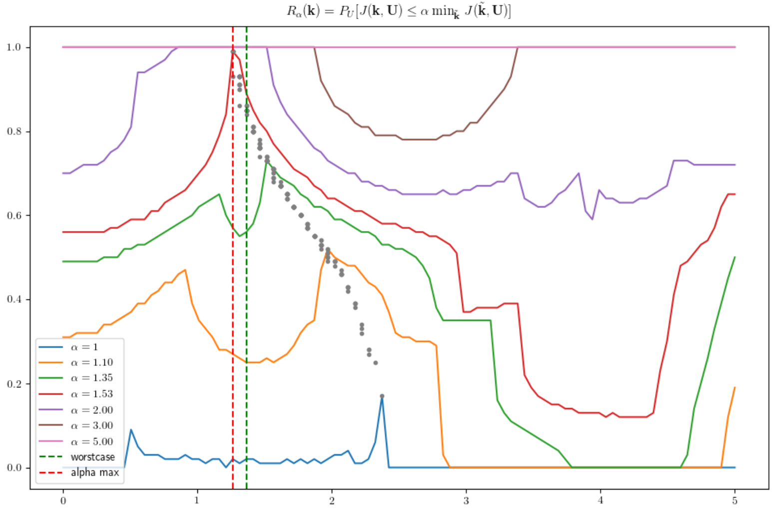
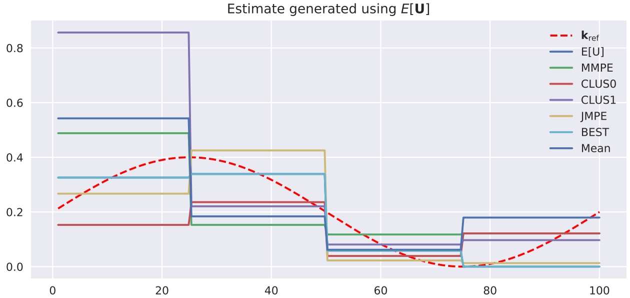


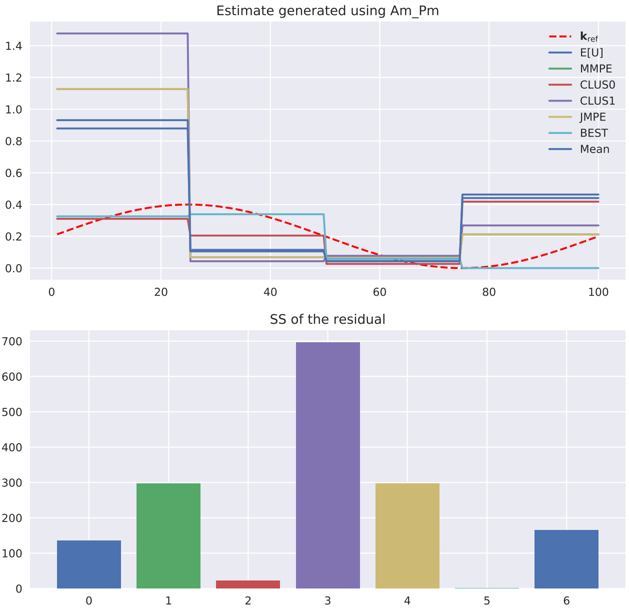
Illustration en 2D: relaxation de la contrainte avec α



MPE et cluster, u_{ref} centré



MPE et cluster, u_{ref} décalé



Théorème de Bayes

On définit la distribution a priori de $\boldsymbol{K} \sim \pi(\boldsymbol{k})$.

Distribution jointe de $(\boldsymbol{K}, \boldsymbol{U})$ ayant observé $\boldsymbol{y}^{\mathrm{obs}}$: $p(\boldsymbol{k}, \boldsymbol{u} | \boldsymbol{y}^{\mathrm{obs}})$?

Théorème de Bayes

$$\begin{aligned} p(\boldsymbol{k}, \boldsymbol{u} | \boldsymbol{y}^{\mathrm{obs}}) &\propto p(\boldsymbol{y}^{\mathrm{obs}} | \boldsymbol{k}, \boldsymbol{u}) \pi(\boldsymbol{k}, \boldsymbol{u}) \\ &\propto L(\boldsymbol{k}, \boldsymbol{u}; \boldsymbol{y}^{\mathrm{obs}}) \pi(\boldsymbol{k}) \pi(\boldsymbol{u}) \end{aligned}$$

Lien avec J ?: Erreur quadratique $\leftrightarrow$ Erreur gaussienne

$$L(\boldsymbol{k}, \boldsymbol{u}; \boldsymbol{y}^{\mathrm{obs}}) \propto \exp\left[-\frac{1}{2} \| M(\boldsymbol{k}, \boldsymbol{u}) - \boldsymbol{y}^{\mathrm{obs}} \|^2_{\boldsymbol{\Sigma}^{-1}}\right] = \exp\left[-J(\boldsymbol{k}, \boldsymbol{u})\right]$$

$$\begin{aligned}
\text{ML :}\quad & \operatorname*{arg\,max}_{(\boldsymbol{k},\boldsymbol{u})} L(\boldsymbol{k}, \boldsymbol{u}; \boldsymbol{y}^{\mathrm{obs}}) \\
\text{MAP :}\quad & \operatorname*{arg\,max}_{(\boldsymbol{k},\boldsymbol{u})} p(\boldsymbol{k}, \boldsymbol{u} | \boldsymbol{y}^{\mathrm{obs}}) = L(\boldsymbol{k}, \boldsymbol{u}; \boldsymbol{y}^{\mathrm{obs}})\pi(\boldsymbol{k})\pi(\boldsymbol{u}) \\
\text{MMAP :}\quad & \operatorname*{arg\,max}_{\boldsymbol{k}} p(\boldsymbol{k} | \boldsymbol{y}^{\mathrm{obs}}) = \int_{\mathcal{U}} p(\boldsymbol{k}, \boldsymbol{u} | \boldsymbol{y}^{\mathrm{obs}}) \,\mathrm{d}\boldsymbol{u} \\
\text{Min variance :}\quad & \operatorname*{arg\,min}_{\boldsymbol{k}} \mathbb{V}\mathrm{ar}_{U} \left[p(\boldsymbol{k} | \boldsymbol{y}^{\mathrm{obs}}, \boldsymbol{U}) \right] \\
\text{Pire des cas:}\quad & \operatorname*{arg\,max}_{\boldsymbol{k}} \{ \min_{\boldsymbol{u}} p(\boldsymbol{k} | \boldsymbol{y}^{\mathrm{obs}}, \boldsymbol{u}) \} \\
\text{MPE :}\quad & \underbrace{\text{Mode of } \boldsymbol{K}_{\operatorname{arg\,max}} = \operatorname*{arg\,max}_{\boldsymbol{k}} p(\boldsymbol{k} | \boldsymbol{y}^{\mathrm{obs}}, \boldsymbol{U})}_{\text{mieux} \rightarrow \text{Cluster analysis}}
\end{aligned}$$

Conclusion

Interrogations

- Fonction coût ou vraisemblance/postérieur ? "Mêmes" quantités, méthodes différentes pour les manipuler ?
- Quels critères sont vraiment pertinents ? liens entre eux ?
- À partir d'un ensemble de candidats: lequel choisir ?
- Quel budget d'évaluations en pratique ?

À faire

- Utilisation des métamodèles (sur quelles fonctions ?)
- Ensembles aléatoires pour probabilité de défaillance (voir avec Reda)
- $\boldsymbol{y}^{\mathrm{obs}}$?