



HAL
open science

Calibrating a hydrological model robustly to rain perturbations with stochastic surrogates

Katarina Radišić, Claire Lauvernet, Arthur Vidard

► To cite this version:

Katarina Radišić, Claire Lauvernet, Arthur Vidard. Calibrating a hydrological model robustly to rain perturbations with stochastic surrogates. LEFE/MANU 2023 - Colloque sur les Méthodes mathématiques et numériques, Météo France, Jun 2023, Toulouse, France. hal-04383513

HAL Id: hal-04383513

<https://hal.inrae.fr/hal-04383513>

Submitted on 9 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Calibrating a hydrological model robustly to rain perturbations with stochastic surrogates

Katarina Radišić^{1,2}

Claire Lauvernet¹, Arthur Vidard²

¹INRAE, RiverLy, Lyon-Villeurbanne

²Univ. Grenoble-Alpes, Inria, CNRS, Grenoble-INP, LJK

Context

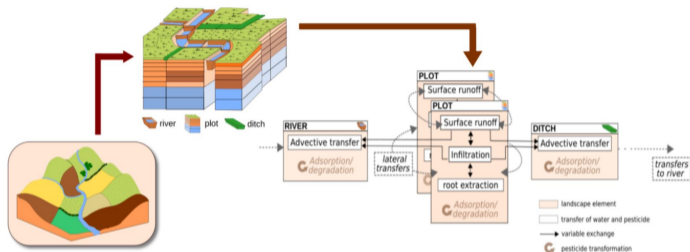


Figure: PESHMELBA^a, a process-based, spatially distributed water and pesticide transfer model, representing dynamical behavior of pesticides in agricultural catchments. Semi-conceptual, semi-physically based model, highly non-linear.

^aEmilie Rouzies et al. (June 2019). "From agricultural catchment to management scenarios: A modular tool to assess effects of landscape features on water and pesticide behavior". en. In: *Science of The Total Environment* 671, pp. 1144–1160.

Context

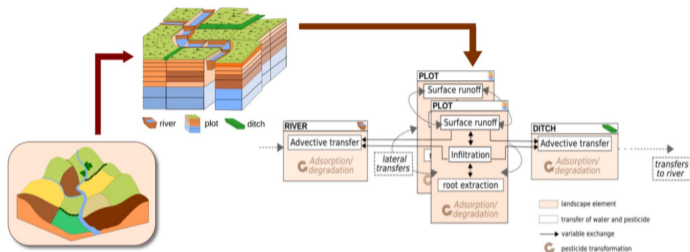


Figure: PESHMELBA^a, a process-based, spatially distributed water and pesticide transfer model, representing dynamical behavior of pesticides in agricultural catchments. Semi-conceptual, semi-physically based model, highly non-linear.

^aEmilie Rouzies et al. (June 2019). "From agricultural catchment to management scenarios: A modular tool to assess effects of landscape features on water and pesticide behavior". en. In: *Science of The Total Environment* 671, pp. 1144–1160.

- Not all model parameters can be measured directly

Context

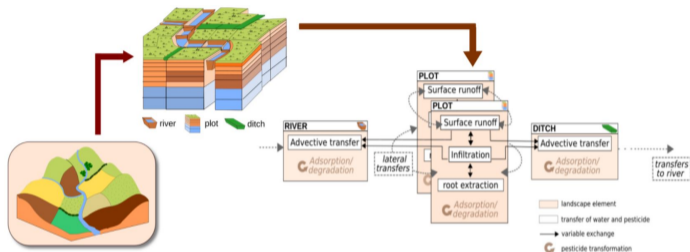


Figure: PESHMELBA^a, a process-based, spatially distributed water and pesticide transfer model, representing dynamical behavior of pesticides in agricultural catchments. Semi-conceptual, semi-physically based model, highly non-linear.

^aEmilie Rouzies et al. (June 2019). "From agricultural catchment to management scenarios: A modular tool to assess effects of landscape features on water and pesticide behavior". en. In: *Science of The Total Environment* 671, pp. 1144–1160.

- Not all model parameters can be measured directly
- → calibrate these model parameters with terrain observations

Context

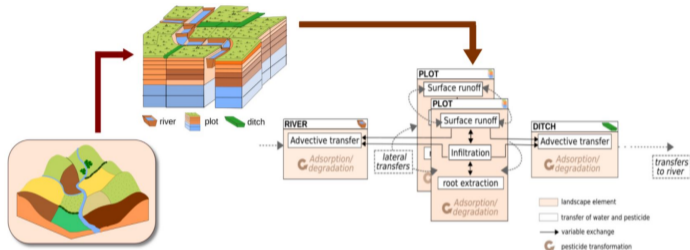


Figure: PESHMELBA^a, a process-based, spatially distributed water and pesticide transfer model, representing dynamical behavior of pesticides in agricultural catchments. Semi-conceptual, semi-physically based model, highly non-linear.

^aEmilie Rouzies et al. (June 2019). "From agricultural catchment to management scenarios: A modular tool to assess effects of landscape features on water and pesticide behavior". en. In: *Science of The Total Environment* 671, pp. 1144–1160.

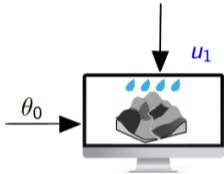
- Not all model parameters can be measured directly
- → calibrate these model parameters with terrain observations

- Impact of external uncertainties on the calibration results

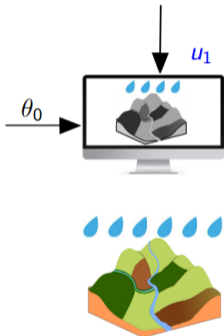
Introduction



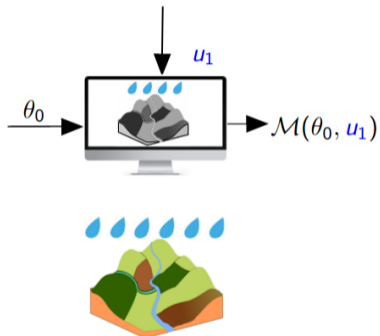
Introduction



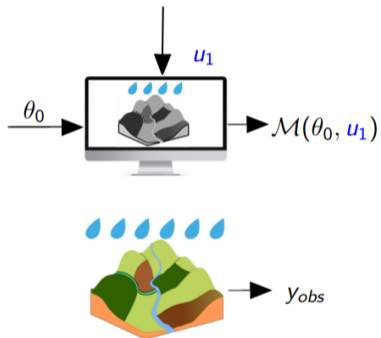
Introduction



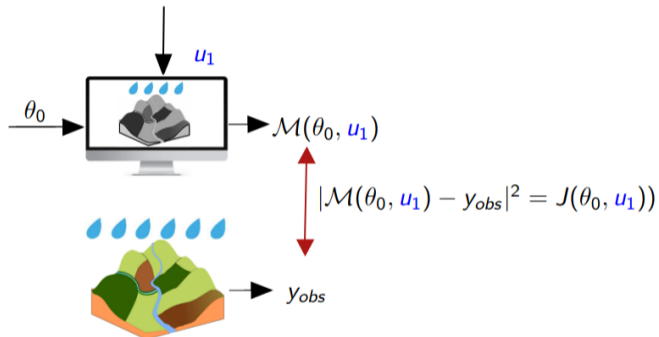
Introduction



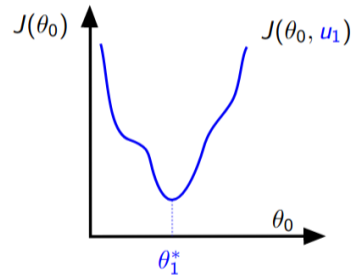
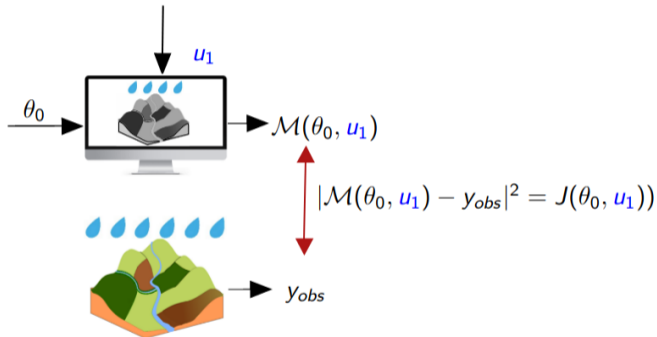
Introduction



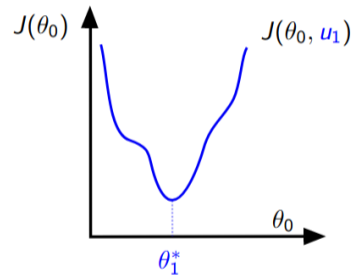
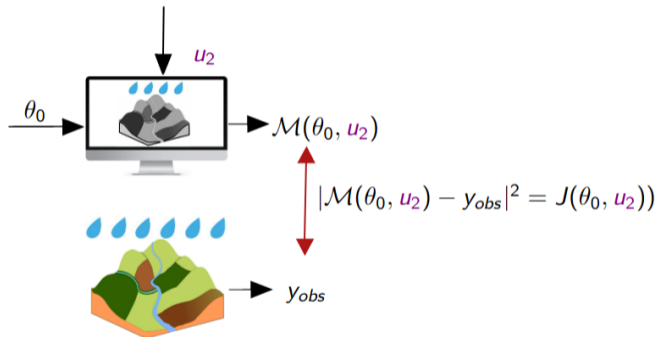
Introduction



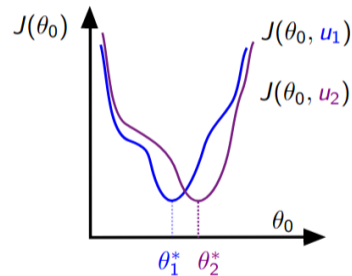
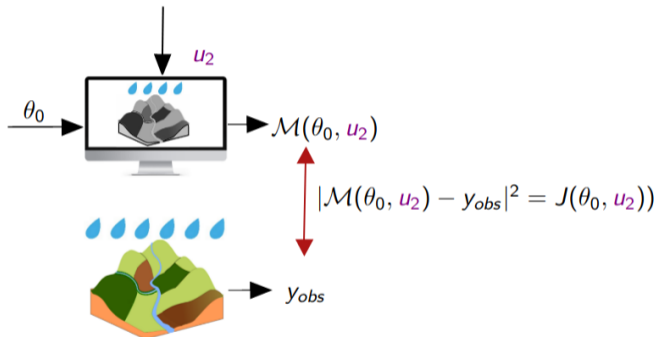
Introduction



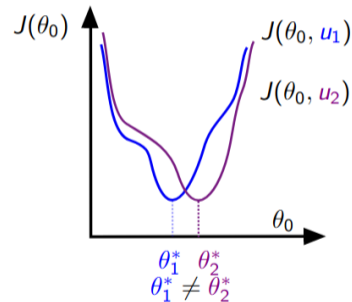
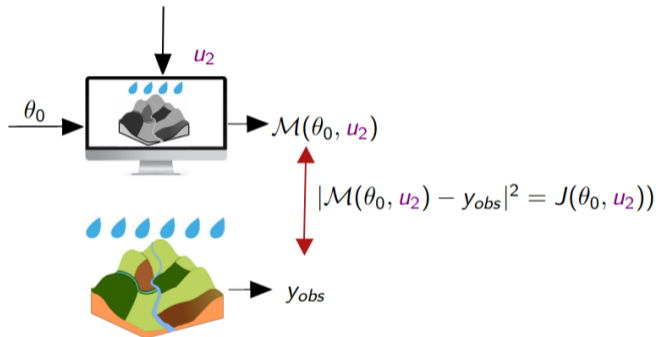
Introduction



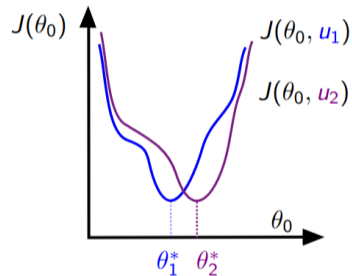
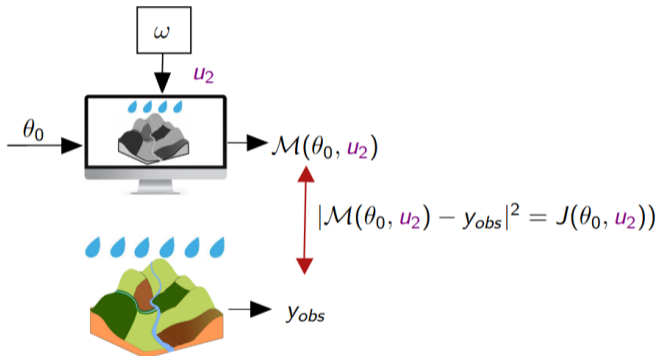
Introduction



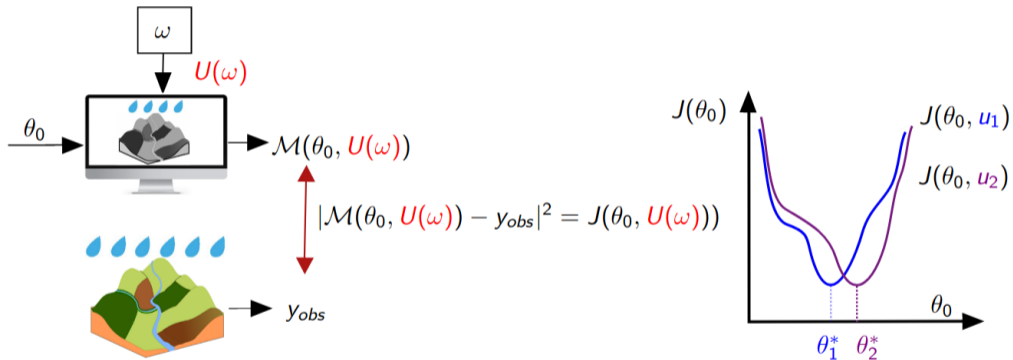
Introduction



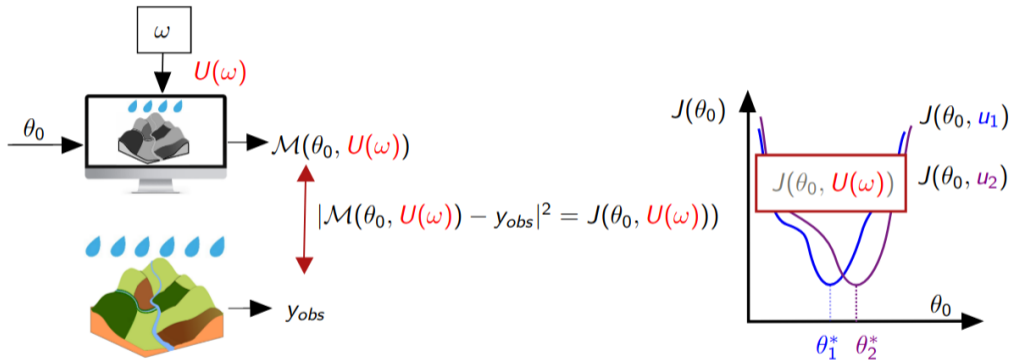
Introduction



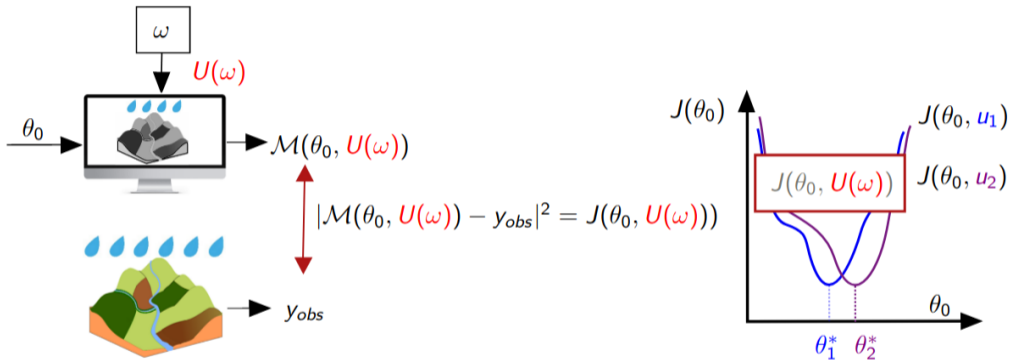
Introduction



Introduction

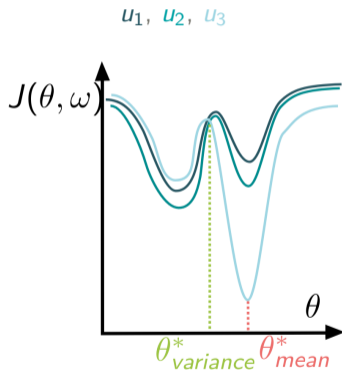


Introduction



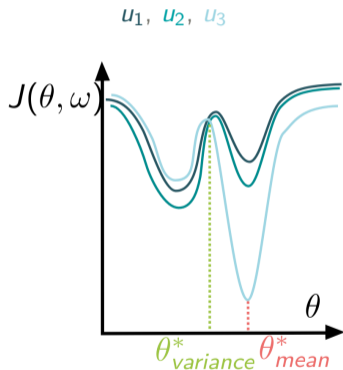
What does it mean to find a *robust* minimizer ?

Methods: Robust estimators



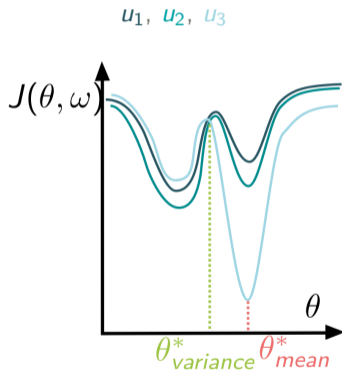
- $\theta_{\mathbb{E}}^* = \underset{\theta}{\operatorname{argmin}} \mathbb{E}_U [J(\theta, U)],$

Methods: Robust estimators



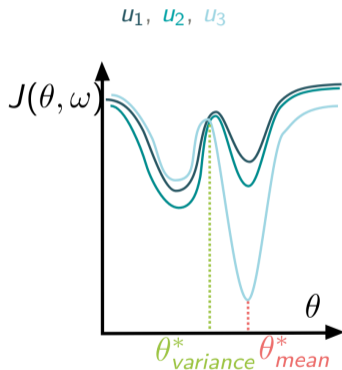
1. $\theta_{\mathbb{E}}^* = \underset{\theta}{\operatorname{argmin}} \mathbb{E}_U[J(\theta, U)],$
2. $\theta_{\text{Var}}^* = \underset{\theta}{\operatorname{argmin}} \text{Var}_U[J(\theta, U)],$

Methods: Robust estimators



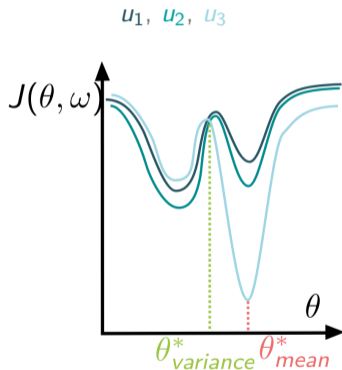
1. $\theta_{\mathbb{E}}^* = \underset{\theta}{\operatorname{argmin}} \mathbb{E}_U[J(\theta, U)],$
2. $\theta_{\mathbb{V}ar}^* = \underset{\theta}{\operatorname{argmin}} \mathbb{V}ar_U[J(\theta, U)],$
3. Pareto of $\theta_{\mathbb{E}}^*, \theta_{\mathbb{V}ar}^*$

Methods: Robust estimators



1. $\theta_{\mathbb{E}}^* = \underset{\theta}{\operatorname{argmin}} \mathbb{E}_U[J(\theta, U)],$
2. $\theta_{\text{Var}}^* = \underset{\theta}{\operatorname{argmin}} \text{Var}_U[J(\theta, U)],$
3. Pareto of $\theta_{\mathbb{E}}^*, \theta_{\text{Var}}^*$
4. ...

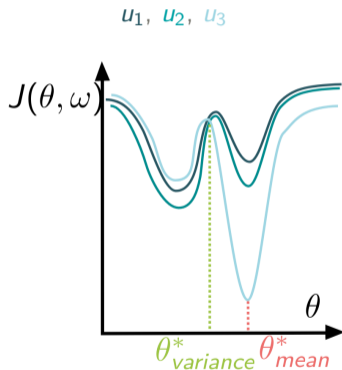
Methods: Robust estimators



1. $\theta_{\mathbb{E}}^* = \underset{\theta}{\operatorname{argmin}} \mathbb{E}_U[J(\theta, U)],$
2. $\theta_{\mathbb{V}ar}^* = \underset{\theta}{\operatorname{argmin}} \mathbb{V}ar_U[J(\theta, U)],$
3. Pareto of $\theta_{\mathbb{E}}^*$, $\theta_{\mathbb{V}ar}^*$
4. ...

- How to estimate the robust parameters from a limited number of model simulations?

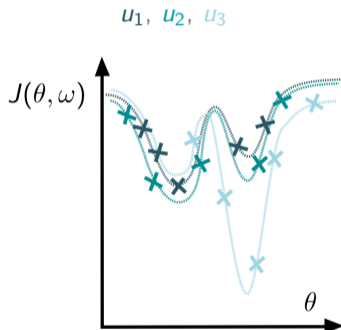
Methods: Robust estimators



1. $\theta_{\mathbb{E}}^* = \underset{\theta}{\operatorname{argmin}} \mathbb{E}_U[J(\theta, U)],$
2. $\theta_{\mathbb{V}ar}^* = \underset{\theta}{\operatorname{argmin}} \mathbb{V}ar_U[J(\theta, U)],$
3. Pareto of $\theta_{\mathbb{E}}^*, \theta_{\mathbb{V}ar}^*$
4. ...

- How to estimate the robust parameters from a limited number of model simulations?

Methods: Robust estimators



1. $\theta_{\mathbb{E}}^* = \operatorname{argmin}_{\theta} \mathbb{E}_U[J(\theta, U)],$
2. $\theta_{\mathbb{V}ar}^* = \operatorname{argmin}_{\theta} \mathbb{V}ar_U[J(\theta, U)],$
3. Pareto of $\theta_{\mathbb{E}}^*, \theta_{\mathbb{V}ar}^*$
4. ...

- How to estimate the robust parameters from a limited number of model simulations?

→ Stochastic metamodel^a

^aNora Lüthen, Stefano Marelli, and Bruno Sudret (Mar. 2023). “A spectral surrogate model for stochastic simulators computed from trajectory samples”. en. In: *Computer Methods in Applied Mechanics and Engineering* 406, p. 115875.

Methods: Stochastic metamodel

- Polynomial chaos expansion (useful for models with sparse observations)

Methods: Stochastic metamodel

- Polynomial chaos expansion (useful for models with sparse observations)
- → ok, for when a rain realization is fixed

Methods: Stochastic metamodel

- Polynomial chaos expansion (useful for models with sparse observations)
- → ok, for when a rain realization is fixed

$$J(\boldsymbol{\theta}) = \sum_{\alpha \in \mathbb{N}^K} c_{\alpha} \psi_{\alpha}(\boldsymbol{\theta}) \approx \sum_{\alpha \in \mathcal{A}_q^{K,p}} c_{\alpha} \psi_{\alpha}(\boldsymbol{\theta})$$

Methods: Stochastic metamodel

- Polynomial chaos expansion (useful for models with sparse observations)
- → ok, for when a rain realization is fixed

$$J(\theta) = \sum_{\alpha \in \mathbb{N}^K} c_{\alpha} \psi_{\alpha}(\theta) \approx \sum_{\alpha \in \mathcal{A}_q^{K,p}} c_{\alpha} \psi_{\alpha}(\theta)$$

- ...but when the rain is an external, unknow, unparametrized uncertainty ?

Methods: Stochastic metamodel

- Polynomial chaos expansion (useful for models with sparse observations)
- → ok, for when a rain realization is fixed

$$J(\boldsymbol{\theta}) = \sum_{\alpha \in \mathbb{N}^K} c_{\alpha} \psi_{\alpha}(\boldsymbol{\theta}) \approx \sum_{\alpha \in \mathcal{A}_q^{K,p}} c_{\alpha} \psi_{\alpha}(\boldsymbol{\theta})$$

- ...but when the rain is an external, unknown, unparametrized uncertainty ?

$$\hat{J}(\boldsymbol{\theta}, \omega) = \hat{\mu}(\boldsymbol{\theta}) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\boldsymbol{\theta}) \right)$$

Methods: Stochastic metamodel

- Polynomial chaos expansion (useful for models with sparse observations)
- → ok, for when a rain realization is fixed

$$J(\boldsymbol{\theta}) = \sum_{\alpha \in \mathbb{N}^K} c_{\alpha} \psi_{\alpha}(\boldsymbol{\theta}) \approx \sum_{\alpha \in \mathcal{A}_q^{K,p}} c_{\alpha} \psi_{\alpha}(\boldsymbol{\theta})$$

- ...but when the rain is an external, unknown, unparametrized uncertainty ?

$$\hat{J}(\boldsymbol{\theta}, \omega) = \hat{\mu}(\boldsymbol{\theta}) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\boldsymbol{\theta}) \right)$$

- PCE on each rain realization, then Principal component analysis on the PCE coefficients

Methods: Stochastic metamodel

- Polynomial chaos expansion (useful for models with sparse observations)
- → ok, for when a rain realization is fixed

$$J(\boldsymbol{\theta}) = \sum_{\alpha \in \mathbb{N}^K} c_{\alpha} \psi_{\alpha}(\boldsymbol{\theta}) \approx \sum_{\alpha \in \mathcal{A}_q^{K,p}} c_{\alpha} \psi_{\alpha}(\boldsymbol{\theta})$$

- ...but when the rain is an external, unknown, unparametrized uncertainty ?

$$\hat{J}(\boldsymbol{\theta}, \omega) = \hat{\mu}(\boldsymbol{\theta}) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\boldsymbol{\theta}) \right)$$

- PCE on each rain realization, then Principal component analysis on the PCE coefficients
- representation of the impact of the variability of the rain on the cost function through a random variable in the latent space

Methods: Stochastic metamodel

- Polynomial chaos expansion (useful for models with sparse observations)
- → ok, for when a rain realization is fixed

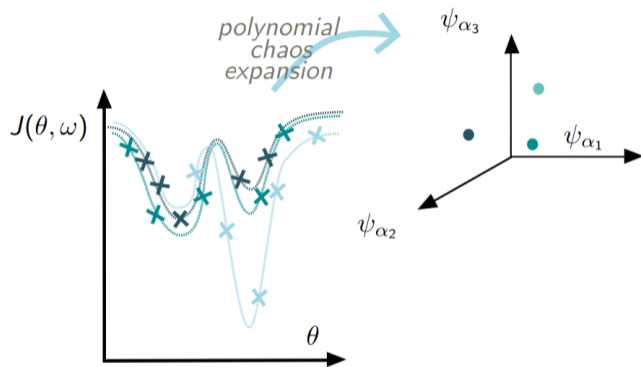
$$J(\boldsymbol{\theta}) = \sum_{\alpha \in \mathbb{N}^K} c_{\alpha} \psi_{\alpha}(\boldsymbol{\theta}) \approx \sum_{\alpha \in \mathcal{A}_q^{K,p}} c_{\alpha} \psi_{\alpha}(\boldsymbol{\theta})$$

- ...but when the rain is an external, unknown, unparametrized uncertainty ?

$$\hat{J}(\boldsymbol{\theta}, \omega) = \hat{\mu}(\boldsymbol{\theta}) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\boldsymbol{\theta}) \right)$$

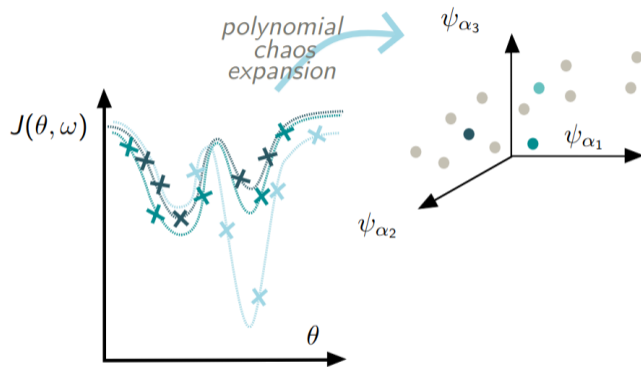
- PCE on each rain realization, then Principal component analysis on the PCE coefficients
- representation of the impact of the variability of the rain on the cost function through a random variable in the latent space
- the mean and variance are analytically available and independent of the distribution of Z

Methods: Stochastic metamodel



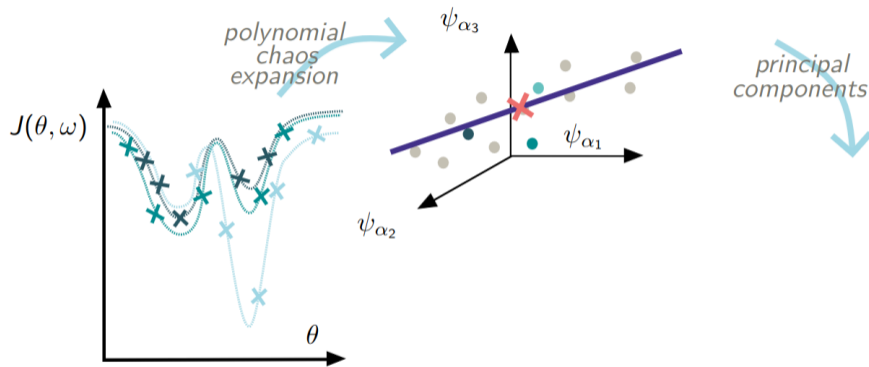
$$\hat{J}(\theta, U(\omega)) = \hat{\mu}(\theta) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\theta) \right)$$

Methods: Stochastic metamodel



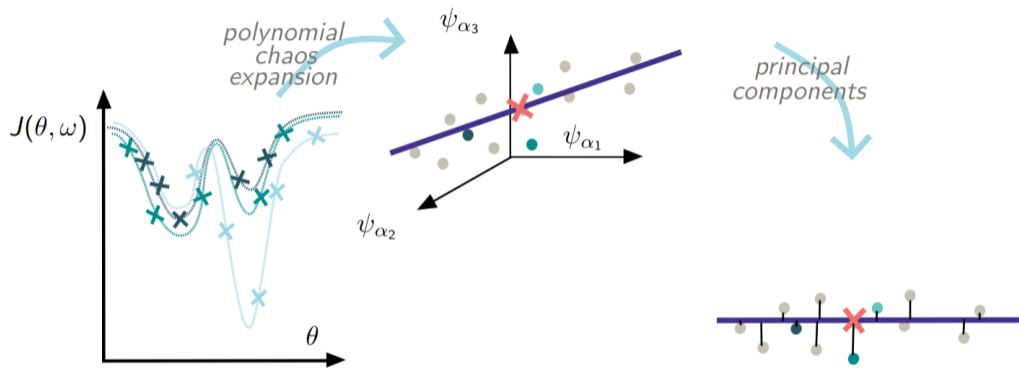
$$\hat{J}(\theta, U(\omega)) = \hat{\mu}(\theta) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\theta) \right)$$

Methods: Stochastic metamodel



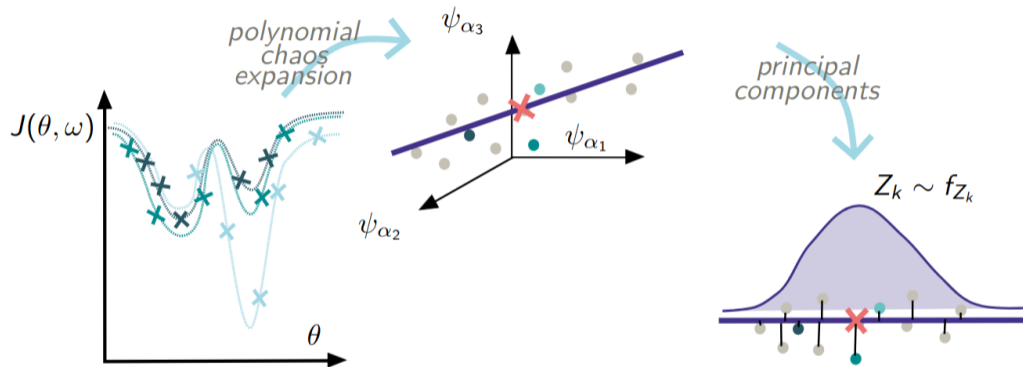
$$\hat{J}(\theta, U(\omega)) = \hat{\mu}(\theta) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\theta) \right)$$

Methods: Stochastic metamodel



$$\hat{J}(\theta, U(\omega)) = \hat{\mu}(\theta) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\theta) \right)$$

Methods: Stochastic metamodel



$$\hat{J}(\theta, U(\omega)) = \hat{\mu}(\theta) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\theta) \right)$$

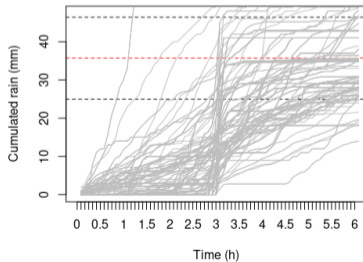
Case study: Moisture profile observations, twin experience

$$J(\theta, u) = (\mathcal{M}(\theta, u) - y_{true})^2$$

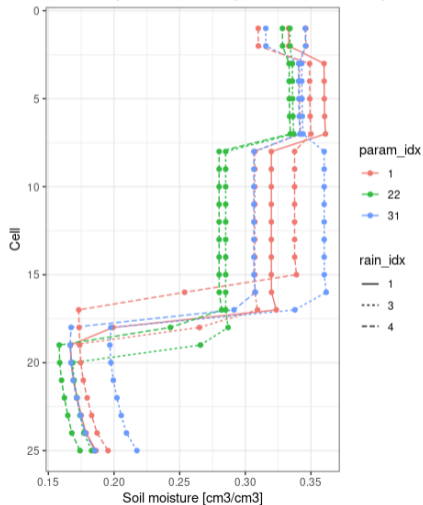
$$\theta \in \mathbb{R}^6$$

$$y_{true} \in \mathbb{R}^{25}$$

Season = S - Duration = 6h - Intensity = 35.7mm



Rain and parameter changes on the moisture profile



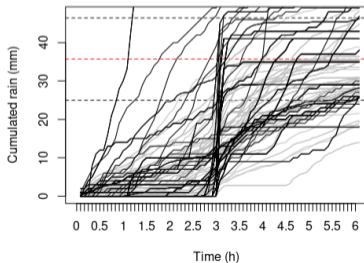
Case study: Moisture profile observations, twin experience

$$J(\theta, u) = (\mathcal{M}(\theta, u) - y_{true})^2$$

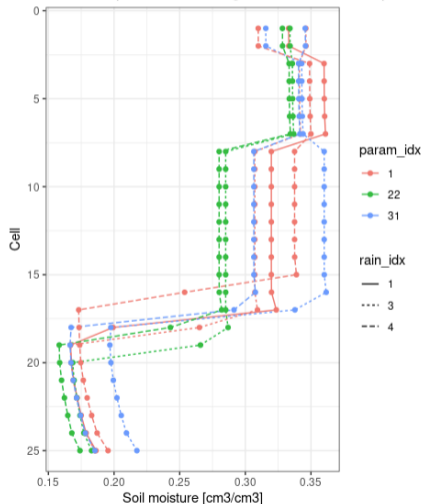
$$\theta \in \mathbb{R}^6$$

$$y_{true} \in \mathbb{R}^{25}$$

Season = S - Duration = 6h - Intensity = 35.7mm



Rain and parameter changes on the moisture profile



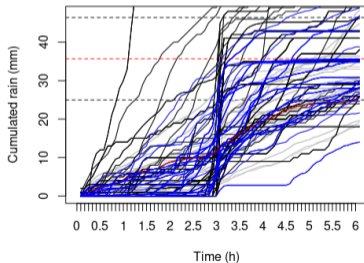
Case study: Moisture profile observations, twin experience

$$J(\theta, u) = (\mathcal{M}(\theta, u) - y_{true})^2$$

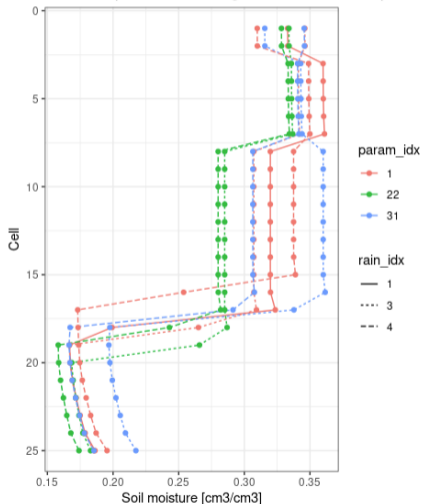
$$\theta \in \mathbb{R}^6$$

$$y_{true} \in \mathbb{R}^{25}$$

Season = S - Duration = 6h - Intensity = 35.7mm



Rain and parameter changes on the moisture profile



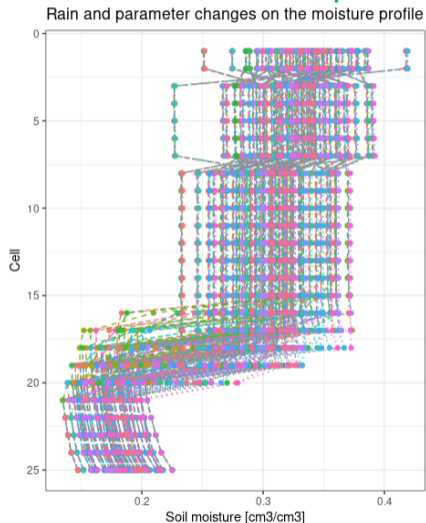
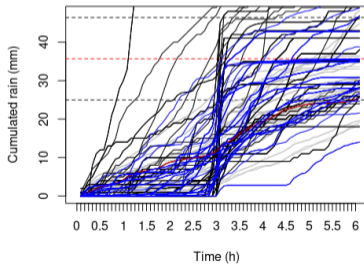
Case study: Moisture profile observations, twin experience

$$J(\theta, u) = (\mathcal{M}(\theta, u) - y_{true})^2$$

$$\theta \in \mathbb{R}^6$$

$$y_{true} \in \mathbb{R}^{25}$$

Season = S - Duration = 6h - Intensity = 35.7mm



Results: Conditional minimization

- PCE constructed conditionally to each train rain realization.
- Determination coefficients $R^2 > 0.95$
- Minimization BFGS conditionally to each rain realization.

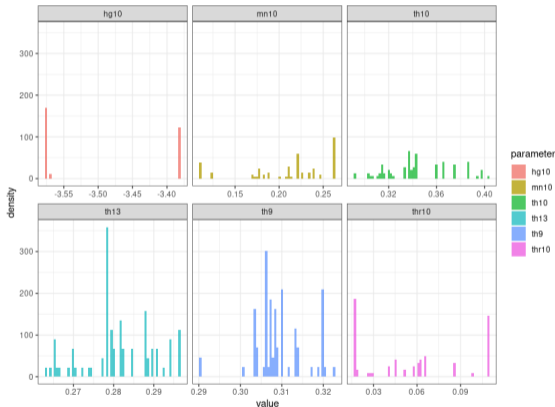


Figure: Histogram of conditional minimizers to each train rain realization.

Results: Validation of the emulator

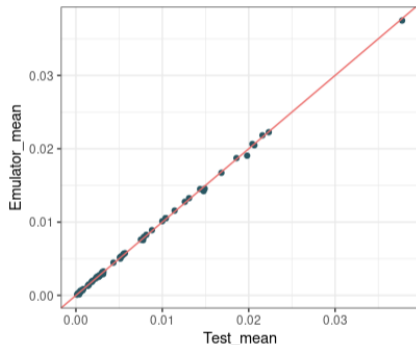


Figure: Comparison of the **train** and the **test** means $\mathbb{E}_U[J(\theta, U)]$.

$$\hat{J}(\theta, U(\omega)) = \hat{\mu}(\theta) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\theta) \right)$$

- The emulator mean is a good estimator of the true mean.

Results: Validation of the emulator

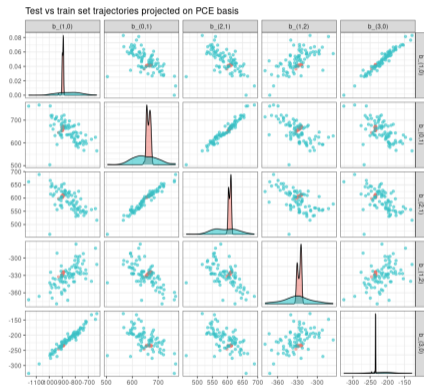


Figure: Comparison of the **train** and the **test** trajectories of PESHMELBA simulations, projected on the PCE basis.

$$\hat{J}(\theta, U(\omega)) = \hat{\mu}(\theta) + \sum_{k=1}^K \sqrt{\lambda_k} Z_k(\omega) \left(\sum_{\alpha \in \mathcal{A}} b_{\alpha}^{(k)} \psi_{\alpha}(\theta) \right)$$

- The emulator mean is a good estimator of the true mean.
- The **train** set and the **test** set do not present the same variabilities, thus the emulator does not reproduce correctly the impact of rain perturbations on the cost function.

Results: Validation of the emulator

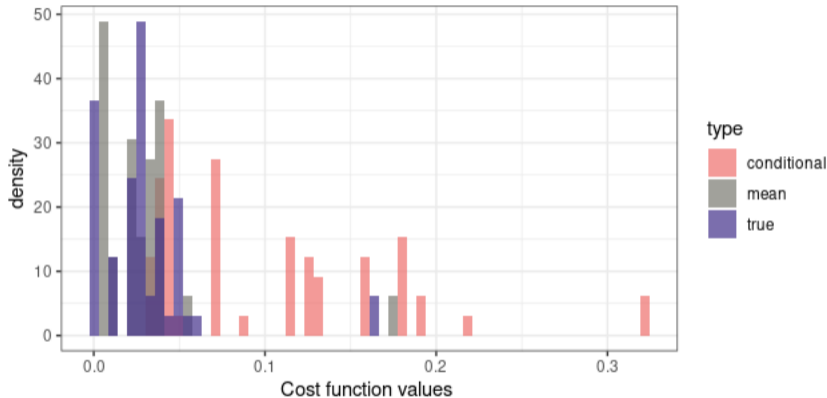


Figure: Cost function evaluated at the test rains, the simulations performed with the minimizer of the mean have a lower mean and variance than the one obtained with the minimizer conditioned on one rain realization.

Conclusion

- The PCE approximations on each conditioned rains have a determination coefficient $R^2 > 0.95$ and the approximation of the metamodel mean is good.
- The minimizer of the mean has lower cost function on unseen rains than a randomly chosen conditional minimizer.
- The variance of the metamodel is not a good estimate

What next ?

- work on the estimation of the variance
- study the Pareto-optimal minimizers
- augment rain perturbations
- minimize mean and variance with Gaussian Process based Efficient Global Optimization, compare the precision and number of model evaluations needed.
- observe another output, for example pesticide quantities in the river, interaction will be added.
- up the domain to the catchment scale

Thank you for your attention !

