How can we quantify and reduce the uncertainty of a watershed-scale pesticide transfer model? A comparison of several approaches

Emilie Rouzies, Claire Lauvernet, Bruno Sudret, Arthur Vidard

To cite this version:
Emilie Rouzies, Claire Lauvernet, Bruno Sudret, Arthur Vidard. How can we quantify and reduce the uncertainty of a watershed-scale pesticide transfer model? A comparison of several approaches. UNCECOMP 2021 - 4th International Conference on Uncertainty Quantification in Computational Sciences and Engineering, Jun 2021, Athens, Greece. pp.1-14. hal-03462086
How can we quantify and reduce the uncertainty of a watershed-scale pesticide transfer model? A comparison of several approaches.

Emilie Rouzies\textsuperscript{1}, Claire Lauvernet\textsuperscript{1}, Bruno Sudret\textsuperscript{3}, Arthur Vidard\textsuperscript{2}

\textsuperscript{1}INRAE, UR RiverLy, Villeurbanne, France
\textsuperscript{2}INRIA, Grenoble, France
\textsuperscript{3}ETH Zürich, Institute of Structural Engineering, Zürich
Introduction

Context

Emilie Rouzies (INRAE, France)

UNCECOMP 2021
June 29, 2021
Introduction

The PESHMELBA model

Development of the PESHMELBA model (Rouzies et al. 2019) to simulate pesticide transfers and fate on small agricultural catchments

✓ Simulations of heterogenous landscapes composed of plots, vegetative filter zones, hedges, ditches and rivers

✓ Modular structure to explore landscape management scenarios
The PESHMELBA model

- Process-oriented, fully spatialized model
- Water transfers on surface and subsurface + pesticide advection, adsorption and degradation

⇒ Complex structure may lead to additional difficulties to diagnose model behavior!
Introduction
The PESHMELBA model

✓ Process-oriented, fully spatialized model
✓ Water transfers on surface and subsurface + pesticide advection, adsorption and degradation
✓ One module $\equiv$ one process or ensemble of processes on a landscape element
✓ Coupling of modules within the OpenPAML coupler (Buis, Piacentini, and Déclat 2006) turning the structure flexible

Complex structure may lead to additional difficulties to diagnose model behavior!
Introduction
The PESHMELBA model

✓ Process-oriented, fully spatialized model
✓ Water transfers on surface and subsurface + pesticide advection, adsorption and degradation
✓ One module ≡ one process or ensemble of processes on a landscape element
✓ Coupling of modules within the OpenPALM coupler (Buis, Piacentini, and Déclat 2006) turning the structure flexible

⇒ Complex structure may lead to additional difficulties to diagnose model behavior!
We have a dream that one day PESHMELBA will be used as a decision-making tool to set up management scenarios and to identify an optimal landscape configuration for pesticide transfer mitigation.
Introduction

PhD Objectives

We have a dream that one day PESHMELBA will be used as a decision-making tool to set up management scenarios and to identify an optimal landscape configuration for pesticide transfer mitigation.

This is our objective...but before, it is necessary to **quantify** and **reduce** the uncertainty associated to PESHMELBA output variables.
Introduction

PhD Objectives

We have a dream that one day PESHMELBA will be used as a decision-making tool to set up management scenarios and to identify an optimal landscape configuration for pesticide transfer mitigation.

This is our objective...but before, it is necessary to quantify and reduce the uncertainty associated to PESHMELBA output variables.

PhD objectives

1. **Quantify**: performing an uncertainty analysis and a sensitivity analysis of the model

2. **Reduce**: performing data assimilation to integrate different sources of data: soil moisture images, ERT measurements and in-situ data of pesticide concentration
Case study

First GSA of PESHMELBA: let’s keep it simple…but realistic! (types of landscape elements, number of parameters, climate conditions...)

Emilie Rouzies (INRAE, France)
UNCECOMP 2021
June 29, 2021
Case study

First GSA of PESHMELBA: let’s keep it simple…but realistic! (types of landscape elements, number of parameters, climate conditions...)

Which method to address sensitivity of such a process-oriented, spatialized model?

ghost

3 soil types + 2 vegetation types + ...

⇒ 145 parameters !!!
Case study

First GSA of PESHMELBA: let’s keep it simple...but realistic! (types of landscape elements, number of parameters, climate conditions...)

3 soil types + 2 vegetation types + ... ⇒ 145 parameters !!!
First GSA of PESHMELBA: let’s keep it simple...but realistic! (types of landscape elements, number of parameters, climate conditions...)

Which method to address sensitivity of such a process-oriented, spatialized model?

3 soil types + 2 vegetation types + ...

⇒ 145 parameters !!!
GSA methods

Notations \( Y = f(X_1, X_2, \ldots, X_k) \)

Variance-based Sobol method (Sobol 1993)
Decomposition of the output variance in conditional variances.

\[
S_i = \frac{\nabla_i}{\nabla(Y)} \quad \text{main effect of } i^{th} \text{ parameter}
\]

\[
S_{ij} = \frac{\nabla_{ij}}{\nabla(Y)} \quad \text{interaction effect due to the } i^{th} \text{ and the } j^{th} \text{ factors}
\]

\[
S_{Ti} = S_i + \sum S_{ij} + \ldots + \sum S_{1,\ldots,k} \quad \text{overall output sensitivity}
\]
GSA methods

Notations $Y = f(X_1, X_2, ..., X_k)$

Variance-based Sobol method (Sobol 1993)
Decomposition of the output variance in conditional variances.

$$S_i = \frac{\text{Var}(i)}{\text{Var}(Y)}$$ main effect of $i^{th}$ parameter

$$S_{ij} = \frac{\text{Var}(ij)}{\text{Var}(Y)}$$ interaction effect due to the $i^{th}$ and the $j^{th}$ factors

$$S_{Ti} = S_i + \sum S_{ij} + ... + \sum S_{1,...,k}$$ overall output sensitivity

Classical Sobol sampling $>$ 75000 model runs, impossible!
⇒ Sobol indices obtained with Polynomial Chaos Expansion surrogate model (Wiener 1938) from 4000 simulation runs using UQLab (Marelli and Sudret 2014).
GSA methods

Alternative methods

- **HSIC dependence measure** (Da Veiga 2015)
  Main idea: describe the similarity between $P_Y$ and $P_{Y|X_i}$ by using a dependence measure $d$

$$S_i^d = \mathbb{E}_{X_i}(d(P_Y, P_{Y|X_i}))$$

Chosen dependence measure: Hilbert-Schmidt independence criterion (HSIC) (Gretton et al. 2005)

$\Rightarrow$ **Screening method** (De Lozzo and Marrel 2014)
GSA methods

Alternative methods

- **HSIC dependence measure** (Da Veiga 2015)
  Main idea: describe the similarity between $P_Y$ and $P_{Y|X_i}$ by using a dependence measure $d$

  $$S^d_i = \mathbb{E}_{X_i}(d(P_Y, P_{Y|X_i}))$$

  Chosen dependence measure: Hilbert-Schmidt independence criterion (HSIC) (Gretton et al. 2005)

  ⇒ **Screening method** (De Lozzo and Marrel 2014)

- **Random Forest**
Scalar variables: informative variables: cumulated water volume and pesticide mass transferred from each HU by subsurface lateral transfers and by surface runoff.
✓ **Scalar variables**: informative variables: cumulated water volume and pesticide mass transferred from each HU by subsurface lateral transfers and by surface runoff.

![Diagram showing water and pesticide flow](image)

✓ **Temporal series**: target variables for DA: surface moisture, mean moisture in first 100 cm, water table pest. conc., water flow and pest. conc. at the outlet

![Graphs showing temporal series](image)
Results
Scalar variables - screening

Screening: independance test based on HSIC measure (power of the test $\alpha=1\%$)

After screening:
- Water surface runoff: 43 parameters
- Pesticide surface runoff: 45 parameters

- High number of influential parameters remaining after screening: method not discriminant enough? Many physical processes at stake?
- Spatial heterogeneities consistent with heterogeneities in physical processes activation
Results
Scalar variables - ranking

Ranking for cumulated pesticide mass transferred in surface runoff
Results
Scalar variables - ranking

Ranking for cumulated pesticide mass transferred in surface runoff

✓ Discrepancies in ranking between the 3 methods
Results
Scalar variables - ranking

Ranking for cumulated pesticide mass transferred in surface runoff

- Discrepancies in ranking between the 3 methods
- Pesticide transfers at surface result from the interaction of several physical processes.
Results
Surface moisture time series - ranking

Random Forest feature importance for surface moisture on HU 4

✓ Uncertainty on influential parameters will be reduced during the DA process
✓ Variable mainly Gaussian along the simulation: valuable info to choose DA method
Random Forest feature importance for surface moisture on HU 4

✓ Uncertainty on influential parameters will be reduced during the DA process
Results

Surface moisture time serie - ranking

Random Forest feature importance for surface moisture on HU 4

✓ Uncertainty on influential parameters will be reduced during the DA process
✓ Variable mainly gaussian along the simulation: valuable info to choose DA method
✓ PESHMELBA specificities turn sensitivity analysis a challenging task ⇒ need for adapted tools: Sobol’ indices from PCE, HSIC, Random Forest

✓ Sensitivity analysis provides valuable information about hydrological processes activation and interaction for a given scenario

✓ Uncertainty/Sensitivity analysis are a necessary preliminary task for data assimilation (which parameters could be estimated? which method can be used?)
✓ PESHMELBA specificities turn sensitivity analysis a challenging task ⇒ need for adapted tools: Sobol’ indices from PCE, HSIC, Random Forest

✓ Sensitivity analysis provides valuable information about hydrological processes activation and interaction for a given scenario

✓ Uncertainty/Sensitivity analysis are a necessary preliminary task for data assimilation (which parameters could be estimated? which method can be used?)

To be further explored:
Sensitivity analysis on temporal series may be improved, especially for pesticide concentration series ⇒ On-going test of PCA-PCE analysis.
Thanks for your attention