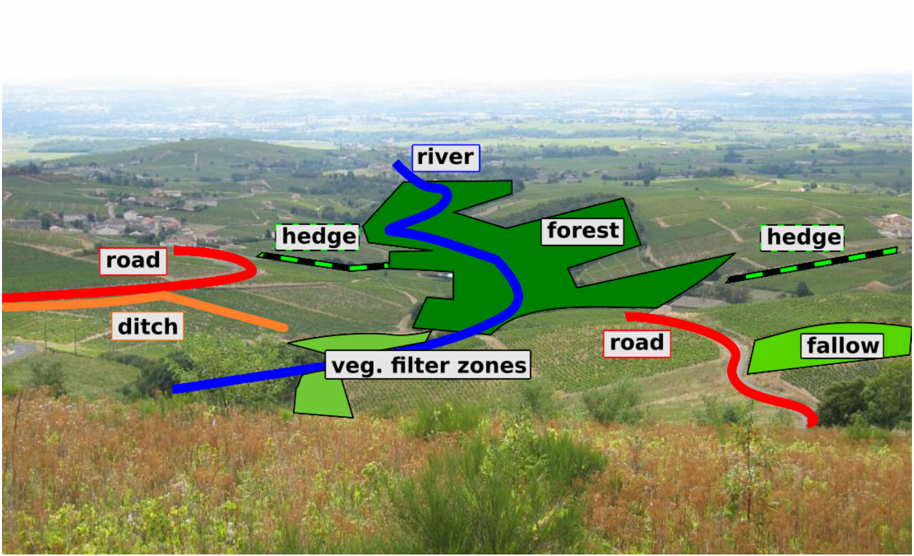
How can we quantify and reduce the uncertainty of a watershed-scale pesticide transfer model? Application to the PESHMELBA model.

Emilie Rouzies (INRAE, France)

PhD supervised by Arthur Vidard (LJK/Inria, France) and Claire Lauvernet (Inrae, France)

Introduction

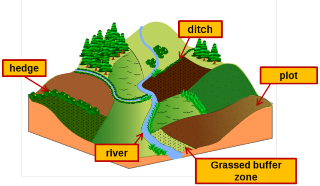
Context



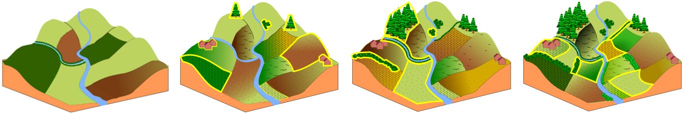
river
hedge
forest
hedge
road
ditch
road
fallow
veg. filter zones

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



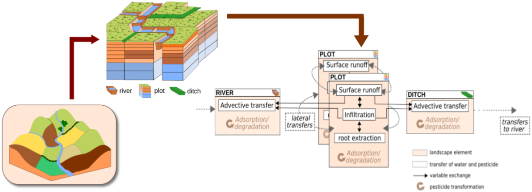
✓ Modular structure to explore landscape management scenarios



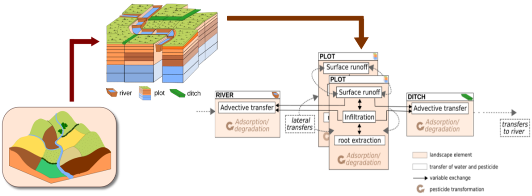
Introduction

The PESHMELBA model

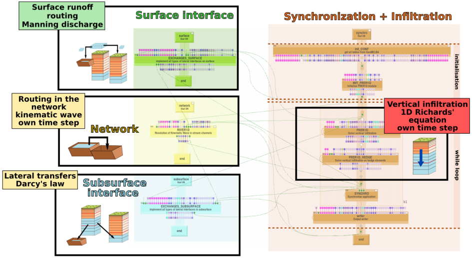
- ✓ Process-oriented, fully spatialized model
- ✓ Water transfers on surface and subsurface + pesticide advection, adsorption and degradation



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



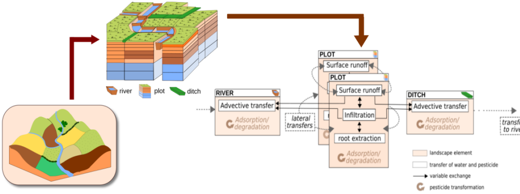
- ✓ One module ≡ one process or ensemble of processes on a landscape element
- ✓ Coupling of modules within the OpenPALM coupler (Buis et al. 2006) turning the structure flexible



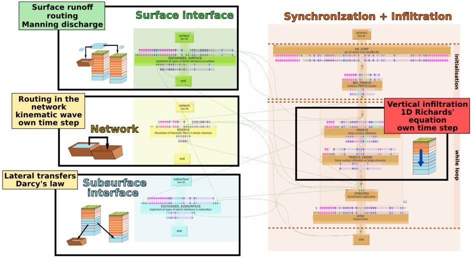
Introduction

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 et al. 2006) turning the structure flexible



⇒ **Complex structure may lead to additionnal 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.

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.

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 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

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.

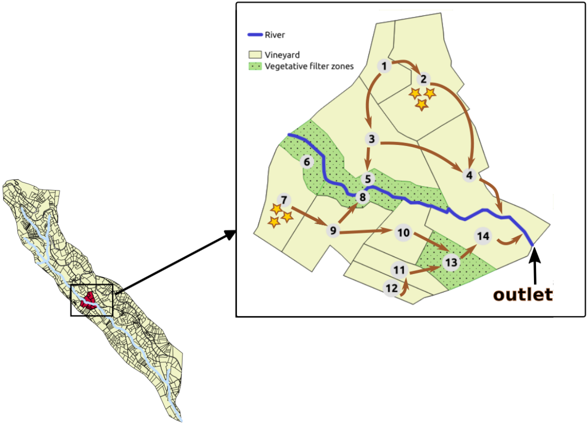
This is our objective...but before, it is necessary to **quantify** and **reduce** the uncertainty associated to PESHMELBA output variables.

PhD objectives

1. Quantify: performing 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

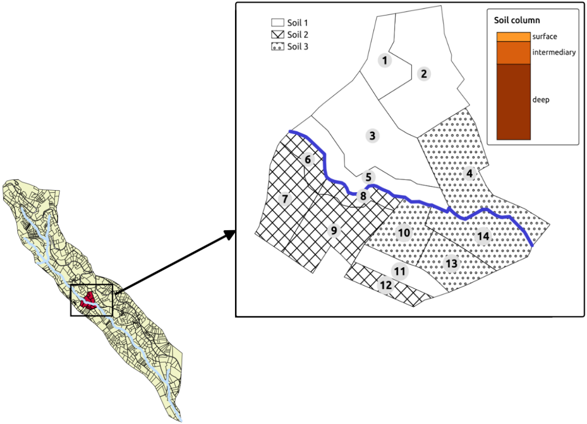
Case study

First attempt of GSA and DA in the PESHMELBA model: let's keep it simple...but realistic! (types of landscape elements, number of parameters, climate conditions...)



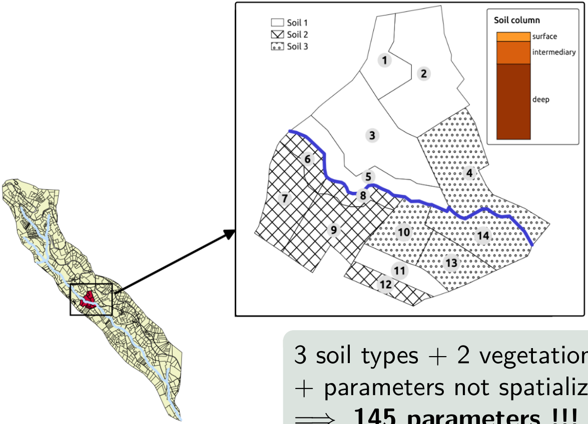
Case study

First attempt of GSA and DA in the PESHMELBA model: let's keep it simple...but realistic! (types of landscape elements, number of parameters, climate conditions...)



Case study

First attempt of GSA and DA in the PESHMELBA model: let's keep it simple...but realistic! (types of landscape elements, number of parameters, climate conditions...)



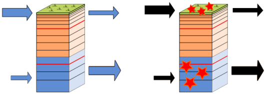
3 soil types + 2 vegetation types + ...
+ parameters not spatialized
$\Longrightarrow$ **145 parameters !!!**

⚠ parameters are assumed independent

Uncertainty quantification

Objectives

✓ Get insight into the (recent) model functioning

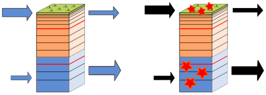


- ✓ Get insight into the (recent) model functioning **GSA on integrated variables**

Integrated variables: cumulated water volume and pesticide mass transferred from each HU by subsurface lateral transfers and by surface runoff.

Uncertainty quantification

Objectives



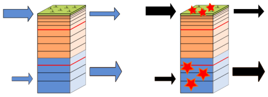
Integrated variables: cumulated water volume and pesticide mass transferred from each HU by subsurface lateral transfers and by surface runoff.

- ✓ Get insight into the (recent) model functioning **GSA on integrated variables**
- ✓ Identify the parameters that may be estimated by data assimilation (most influential)

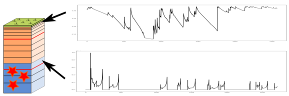
Uncertainty quantification

Objectives

- ✓ Get insight into the (recent) model functioning **GSA on integrated variables**
- ✓ Identify the parameters that may be estimated by data assimilation (most influential) **GSA on time series ⇒ see next talk (Katarina)**



Integrated variables: cumulated water volume and pesticide mass transferred from each HU by subsurface lateral transfers and by surface runoff.



Target variables for DA: surface moisture,mean moisture in first 100 cm, water table pest. conc., water flow and pest. conc. at the outlet

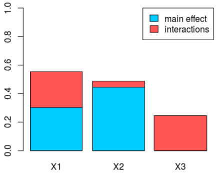
Notations $Y = f(X_1, X_2, ..., X_k)$

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

$S_i = \frac{\mathbb{V}_i}{\mathbb{V}(Y)}$ main effect of i^{th} parameter

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

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



Sobol indices for Ishigami function

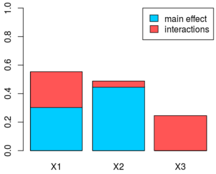
Notations $Y = f(X_1, X_2, ..., X_k)$

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

$S_i = \frac{\mathbb{V}_i}{\mathbb{V}(Y)}$ main effect of i^{th} parameter

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

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



Sobol indices for Ishigami function

Classical Sobol sampling $>$ 75,000 model runs, impossible!
$\Rightarrow$ Sobol indices obtained with Polynomial Chaos Expansion surrogate model (Wiener 1938) using UQLab (Marelli and Sudret 2014).

Uncertainty quantification

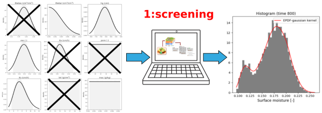
GSA methods

- **HSIC dependence measure** **(Da Veiga 2015)**
  Dependence measures: aim at quantifying, from a probabilistic point of view, the dependence between X_i and Y with the property that the measure equals zero if and only if X_i and Y are independent.

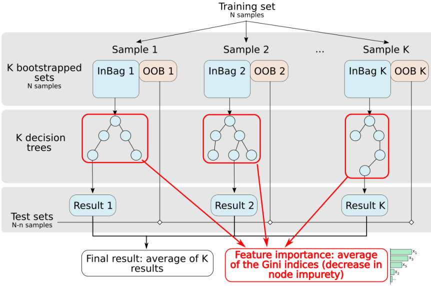
  $\Rightarrow$ Chosen dependence measure: **Hilbert-Schmidt independence criterion (HSIC)** (Gretton et al. 2005): calculate the cross-correlation between any non-linear transformations of some input factor X_i and the output Y.

$$HSIC(\mathrm{X}_i, \mathrm{Y})_{\mathcal{F}_i,\mathcal{G}} = ||C[\mathcal{G}\mathcal{F}_i]||^2_{HS}$$

Also used as a **screening** method based on an independence test (De Lozzo and Marrel 2014)



Feature importance from Random Forest (Breiman 2001)



Feature importance measures: an input parameter X_i is considered important if when breaking the link between X_i and the output Y by permutation, the RF prediction error increases.

Uncertainty quantification

Workflow for scalar variables

Case study

PESHMELBA
=
vertical infiltration
+
surface runoff
+
lateral exchanges
+
pest. degradation
+
pest. adsorption

→ 145 input parameters

Spatialized scalar output:

- ✓ water/pest. surface runoff
- ✓ water/pest. subsurface flow

1 Screening (4,000-pts LHS)

Independence test based on HSIC measure

$$p\text{-val}_B = \frac{1}{B}\sum_{b=1}^{B} \mathbb{1}_{\widehat{HSIC}^{[b]}(X_i,Y) > \widehat{HSIC}(X_i,Y)}$$

2 Ranking (1,000-pts LHS)

a. Site indices: one GSA per HU

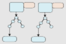
- ✓ Sobol' indices from PCE

$$S_{i_1,\ldots,i_s} = \frac{V_{i_1,\ldots,i_s}}{\mathrm{Var}[Y]}$$

- ✓ HSIC dependence measures

$$HSIC(X_i, Y)_{\mathcal{F}_i,\mathcal{G}} = ||C[\mathcal{G}\mathcal{F}_i]||^2_{HS}$$

- ✓ RF feature importance measures



b. Aggregated indices: one GSA for the whole catchment

- ✓ Aggregated Sobol indices (Gamboa 2014)

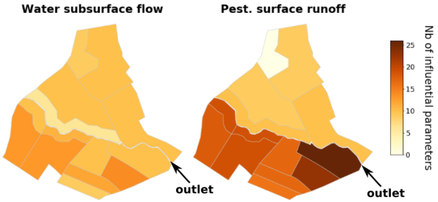
$$S_u = \frac{\sum_{j=1}^{d} \mathrm{Var}[Y_j] S_{u,j}}{\sum_{j=1}^{d} \mathrm{Var}[Y_j]}$$

- ✓ Aggregated HSIC indices
- ✓ Aggregated RF indices

Screening: independence test based on HSIC measure (power of the test α=1%)

After screening:

- Water lateral transfer: 42 parameters
- Pesticide surface runoff: 45 parameters



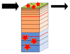
Number of parameters per HU after screening

✓ High number of influential parameters remaining after screening: method not discriminant enough? Many physical processes at stake?

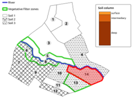
✓ Spatial heterogeneities consistent with heterogeneities in physical processes activation

Uncertainty quantification

Results - ranking (scalar variables)

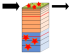


Ranking for cumulated pesticide mass transferred in surface runoff

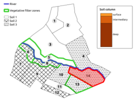


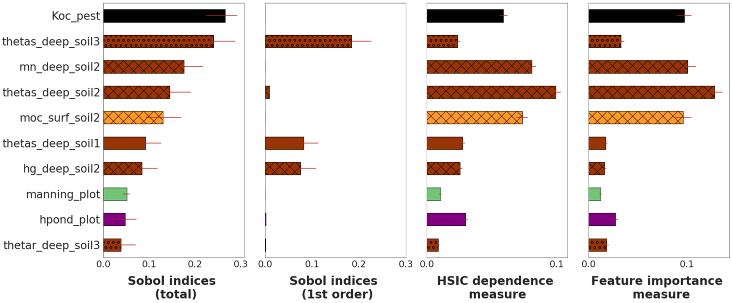
Uncertainty quantification

Results - ranking (scalar variables)

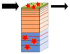


Ranking for cumulated pesticide mass transferred in surface runoff

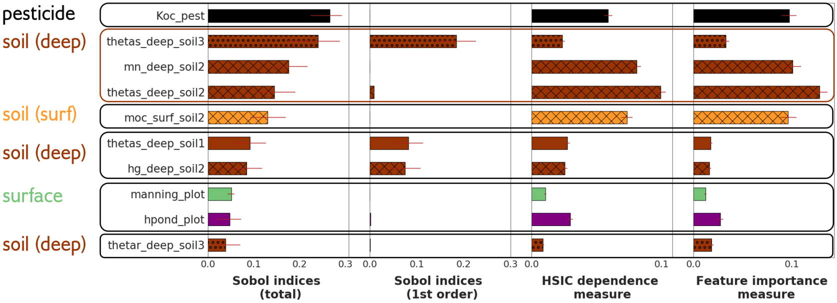




- ✓ Rankings from Sobol' **total** indices, HSIC and RF measures are mainly consistent
- ✓ Quantitative differences due to the contrasts in *Sensitivity* definition
- ✓ Uncertainty + on Sobol' indices: PCE estimation quality ?



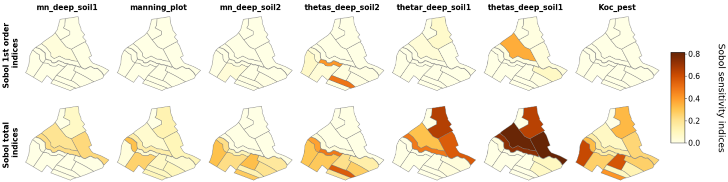
Ranking for cumulated pesticide mass transferred in surface runoff

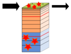


- ✓ Influential parameters relate to various physical processes of transfers **and** transformation: adsorption, overland flow, vertical infiltration...
- ✓ This ranking reflects the interactions of physical processes in PESHMELBA

Ranking for cumulated pesticide mass transferred in surface runoff

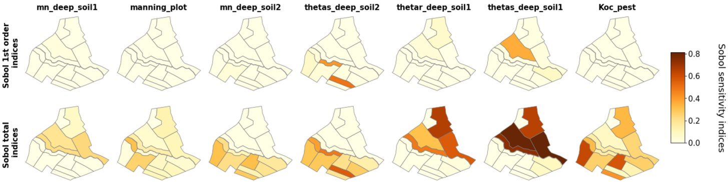
Site sensitivity indices





Ranking for cumulated pesticide mass transferred in surface runoff

Site sensitivity indices



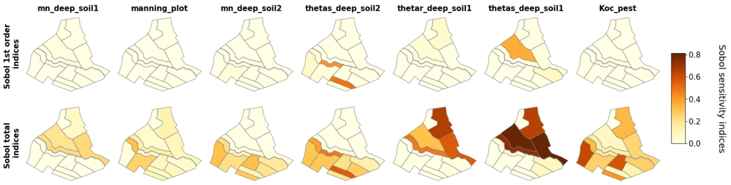
✓Detailed insights into the model sensitivity but computationally costly

Uncertainty quantification

Results - landscape analysis (scalar variables)

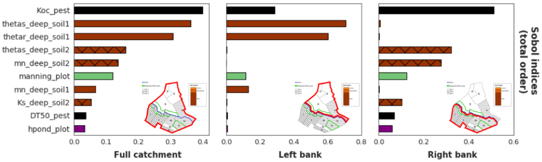
Ranking for cumulated pesticide mass transferred in surface runoff

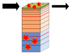
Site sensitivity indices



✓Detailed insights into the model sensitivity but computationally costly

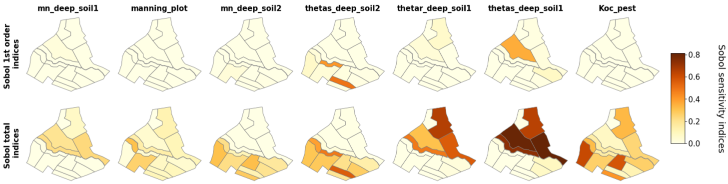
Aggregated sensitivity indices





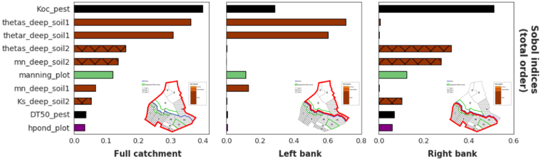
Ranking for cumulated pesticide mass transferred in surface runoff

Site sensitivity indices



✓Detailed insights into the model sensitivity but computationally costly

Aggregated sensitivity indices



- ✓ Summary of overall sensitivity
- ✓ Hillslope scale used as an intermediary scale to mantain physical interpretation of aggregated indices

- Rankings mainly consistent for hydro. variables
- Differences for more complex pest. variables due to differences in "*sensitivity*" definitions
- Sobol aggregated indices at intermediary scale provide valuable information about the physics + overall summary on sensitivity
- We choose Sobol indices as they capture interactions but HSIC and RF should not be discarded for "simple" variables (many advantages).
- How could we transpose the methodology to real catchments ? ⇒ New challenges: spatialized, dependent input parameters ? choice of a relevant intermediary scale to guarantee physical interpretability of results ?