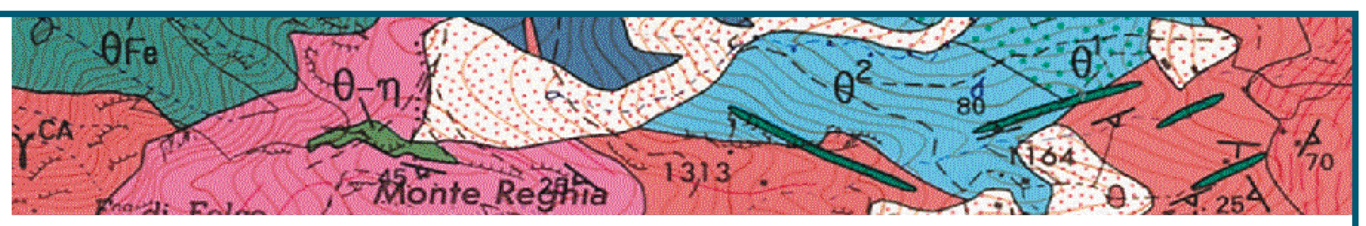


COMPARISON OF STATISTICAL METHODS FOR RANKING THE IMPORTANCE OF MODEL AND PARAMETER UNCERTAINTIES: A CASE STUDY WITH NUMERICAL FLOW SIMULATIONS OF CO2 SEQUESTRATION

General context

- **Uncertainty analysis** is a **key component** of **modelling flow processes**
- Flow modelling:
  - Two types of uncertainties : **parameter** and **model**
  - Potentially **large computation time cost (>hours)**
- **Global sensitivity analysis (GSA) :**
  - what sources of uncertainty **contribute the most** to the uncertainties in the flow simulation results?
  - how to **rank these sources** of uncertainties?
  - how to **set priorities** for future investigations?
- **Observation :** different SA approaches can be used for computing sensitivities, **BUT** they may result in a different importance ranking.

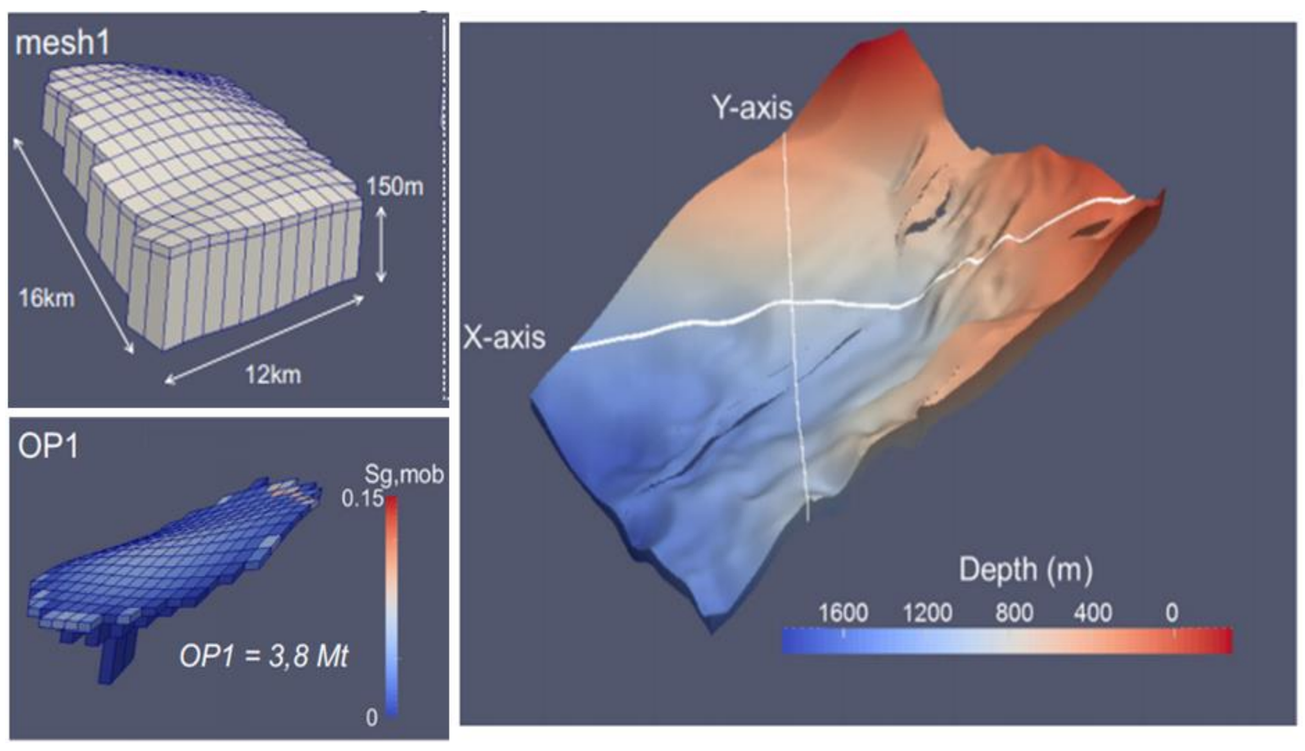
Question : which GSA methods are best applicable regarding the specificities of the situation ?

Objective

- **Test the feasibility of the available methods/approaches** for dealing with GSA with respect to **parameter and model uncertainties**
- **4 types of approaches tested:**
  - **DGSA:** distance-based generalized sensitivity analysis developed by Fenwick et al. (2014) and further extended by Park et al. (2016) based on the Regionalized sensitivity analysis **RSA** method (Spear and Hornberger 1980);
  - **PAWN:** a density-based GSA (aka moment-independent) developed by Pianosi & Wagener (2015);
  - **M-VBSA:** combination of variance-based GSA (Saltelli et al. 2008) and metamodeling techniques adapted to situations using continuous and categorical variables (Storlie et al. 2013);
  - **RF:** a machine learning approach based on the random forest technique (e.g., Wei et al. 2015).

Case study

- Same as in Manceau and Rohmer (2016) :**dataset of 1000 simulations available**
- **Injection of 30 Mt of CO_2 during 30 years** in the lower Triassic sandstone formation based on a potential project in the Paris basin (France).
- Sensitivity analysis on the quantity of mobile CO_2 150y. after the injection stops



Case study geological model [Manceau and Rohmer, 2016]

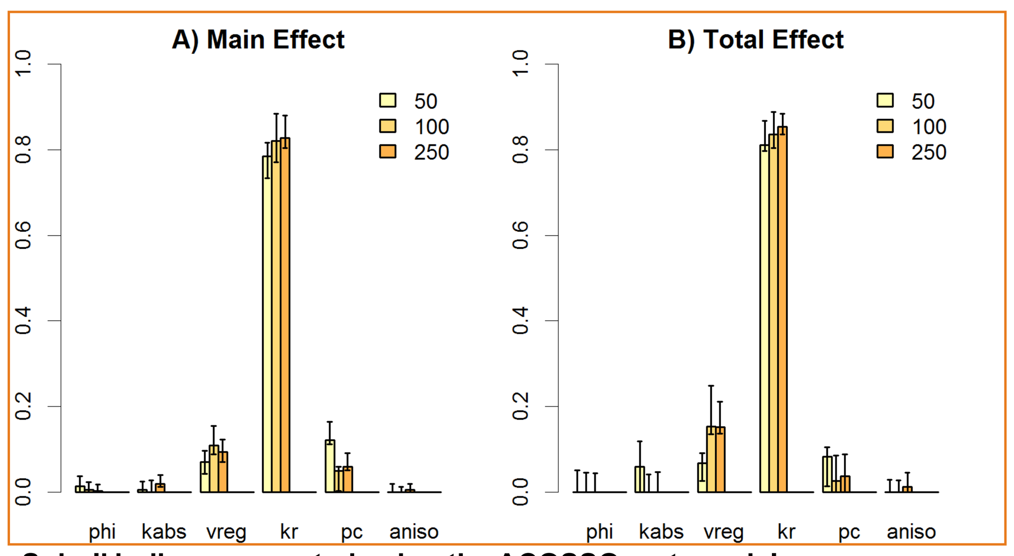
- Input variables and **associated uncertainties**:

Parameter	Uncertainty type	Representation
Porosity	Parametric	Probability density
Permeability	Parametric	Probability density
Permeability anisotropy	Model	3 scenarios (k_v/k_h = 0.1; 0.5 and 1)
Regional hydraulic gradient	Model	2 scenarios (hydrostatic and 0.01 m/m)
Relative permeability	Model	10 scenarios (10 different relative permeability datasets)
Capillary pressure	Model	2 scenarios (no-capillary pressure and “strong” capillary pressure)

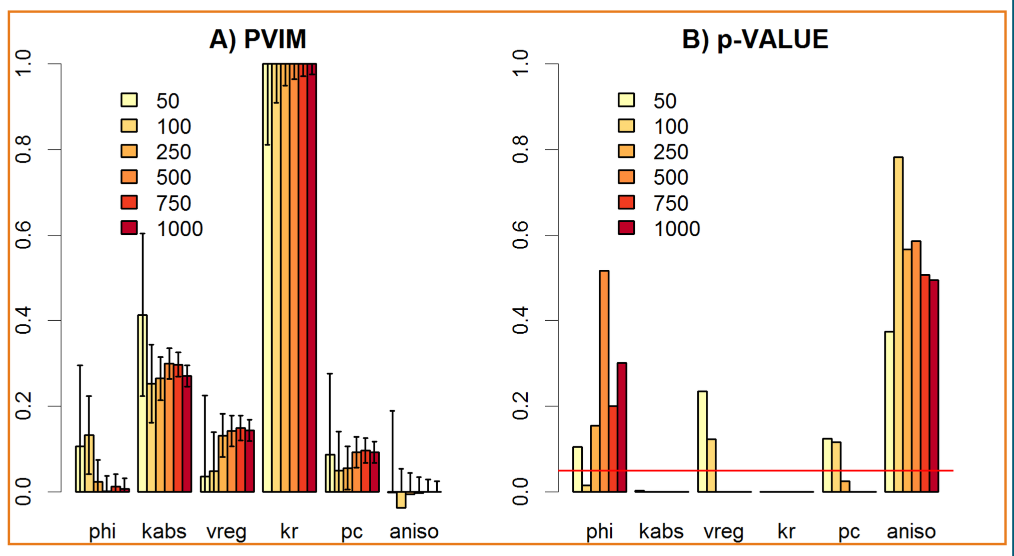
Within-method analysis

- Three criteria for assessing **the impact of the number of simulation results N and the robustness to the parametrization** of each method:

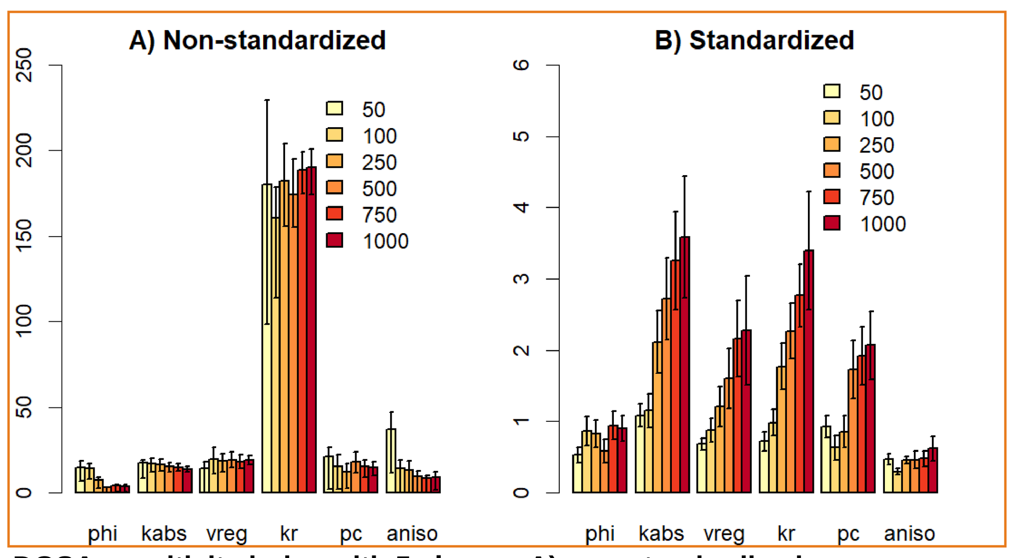
1) **Convergence of the sensitivity indices**: reached if the values of the indices remain stable;
2) **Convergence of ranking:** achieved if the ordering between the parameters remains stable;
3) **Convergence of screening:** reached if the partitioning between non- and -influential parameters remains stable.



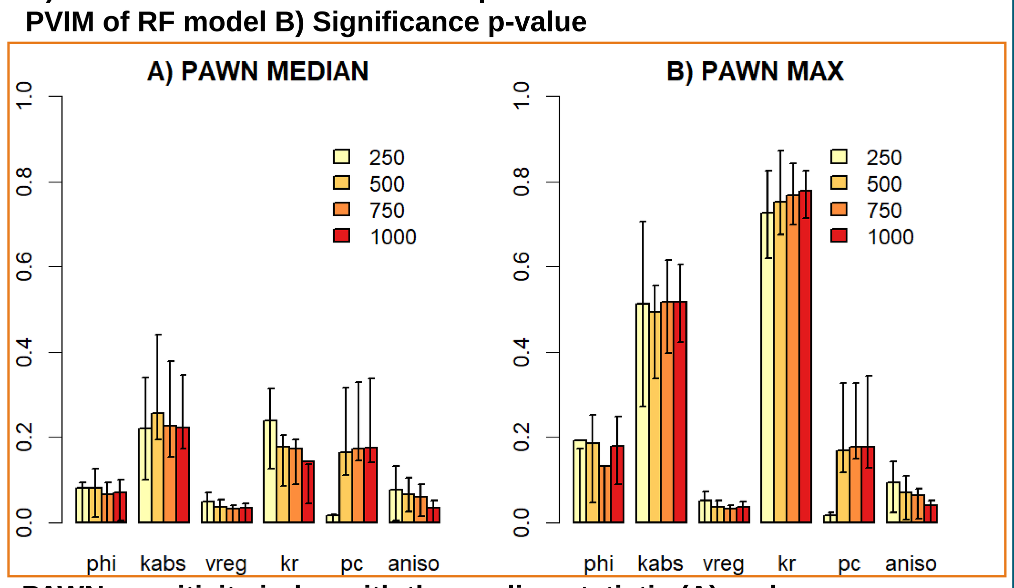
Sobol' indices computed using the ACOSSO metamodel



A) Permutation-based Variable Importance measure
PVIM of RF model B) Significance p-value



DGSA sensitivity index with 5 classes A) non-standardized procedure and B) standardized procedure



PAWN sensitivity index with the median statistic (A) and with the max statistic (B)

Between-method analysis

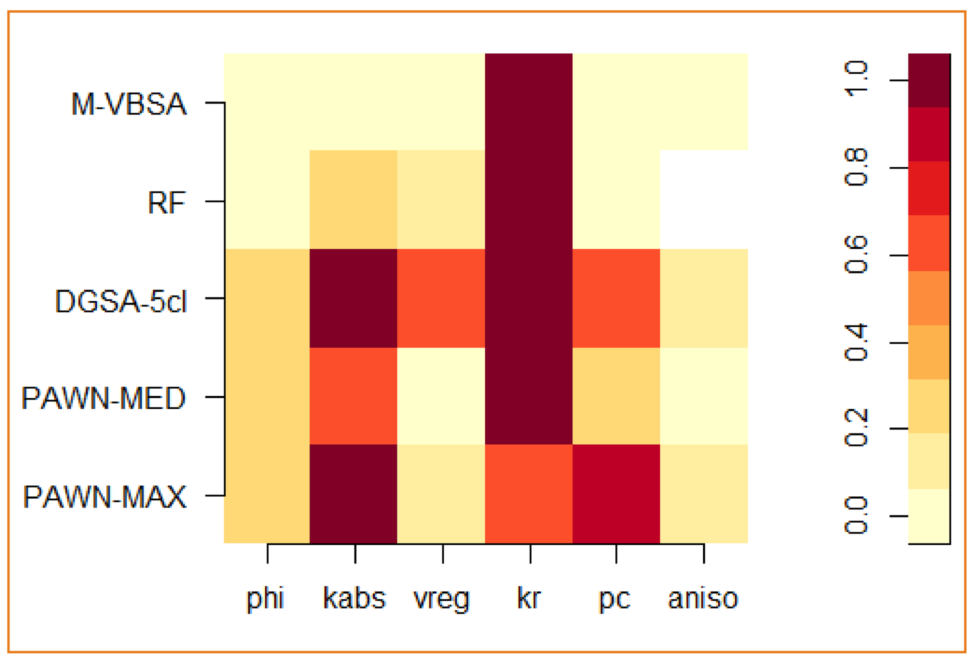
- **Normalization of the sensitivity measures** with respect to the maximum value reached for each method considering N=250

⇒**Clear similarity** among most methods regarding ***kr* largest importance**

BUT:

⇒The importance ranking of the **other parameters is less straightforward**

⇒The normalized sensitivity measures **cannot be compared**



- Comparison of the methods form a **practical point of view**

	PROS	CONS
M-VBSA	**Intuitive and rigorous** interpretation as a proportion of variance; **Feedbacks** in a large variety of domains	**Sensitivity to the number of simulations**, which imposes a careful examination of the predictability of the metamodel. **Convergence analysis** when using Monte-Carlo algorithm.
RF	**Little influence** of the RF parameters (mtry, ntree, nodesize and even split rule); **Robustness** to the number of simulations	**Sensitivity of the p-value** algorithm to the number of permutations. **Difficulties in the interpretation** of the sensitivity measure (here as a decrease in predictability)
RSA	**Easy to compute**, and possible even for low number of samples and for categorical inputs. Adapted when the outputs can be naturally divided into two different groups, and useful for **factor mapping.**	**Do not account for interactions** among input parameters. **No procedure for factor fixing.** **Cannot handle more than two groups,** and very much influenced by the choice of these two groups: may lead to difficulties in interpretation.
DGSA	Can handle **multiple groups of outputs**. Provide a lot of information on **interactions** among parameters (two-way interactions) Can help **fixing parameters** to the less influential value/range	The proposed statistical test might lead to a strong **sensitivity to the number of simulations.** The test for statistically significant interactions require a **high number of simulations.** Very much **influenced by the output classification**: : may lead to difficulties in interpretation..
PAWN	Relatively **good convergence** for sensitivity analysis, ranking and fixing with the number of simulations. Compared to other density-based GSA, rely on **CDF** whose approximation is easier than PDF	Dependent on the **choice of the statistic**: may lead to difficulties in interpretation. A procedure is proposed for factor fixing but **interactions** are not accounted for.

References

Fenwick et al., 2014, Math. Geosci., 46(4), 493-511 / Manceau and Rohmer, 2016, Comput. Geosci., 20(6), 1251-1267 / Park et al., 2016, Comput. Geosci., 97, 15-29 / Pianosi and Wagener, 2015, Environ. Modell. Softw., 67, 1-11 / Saltelli et al., 2008, The Primer. John Wiley & Sons, Chichester. / Spear and Hornberger, 1980, Water Res., 14, 43-49 / Storlie et al., 2013, Reliab. Eng. Syst. Safe., 113, 30-41 / Wei et al., 2015, Reliab. Eng. Syst. Safe., 142, 399-432

Authors

Jean-Charles Manceau, Jérémy Rohmer and Pascal Audigane
Contact: jc.manceau@brgm.fr