Sensitive partitioning of the model output space: principle and first results

Sébastien Roux, Patrice Loisel, Samuel Buis

Speaker: Sébastien Roux

Sensitive partitioning of the model output space: principle and first results

Sébastien Roux
MISTEA, Univ Montpellier, INRAE, Montpellier SupAgro, France, sebastien.roux@inrae.fr
Patrice Loisel
MISTEA, Univ Montpellier, INRAE, Montpellier SupAgro, France, patrice.loisel@inrae.fr
Samuel Buis
EMMAH, INRAE, Univ Avignon et Pays de Vaucluse, Avignon, France, samuel.buis@inrae.fr

Context: partitioning the output space

The use of partition of the model output space is a very convenient way to define behaviors of a model which scales to any dimension of the output space and can be powerful to give interpretable characterization of model properties. Considering partitions of the output space in relation to parameter sensitivity comes from the Regional Sensitivity Analysis (RSA) approach of Spear and Hornberger [1]. This subject has lately gained interest through two research directions: i) its application in the context of reliability engineering to characterize parameter sensitivity in relation to a critical domain of the output space and using sensitivity measures compatible with rare events (Target SA, [2]), ii) its application in combination with a clustering procedure in order to characterize parameter sensitivity in relation to the dominant behaviors of the output space (Cluster-based GSA, [3]).

Optimized sensitive partitioning: principle

We introduce here a new perspective on these different approaches. Instead of trying to a priori characterize a target region of the output space, we propose to **optimize** the partitioning in order to reveal the partition of the output space **the most sensitive** to the variations of a given input, i.e. the most explained by the variations of this input. We named this approach **an optimized sensitive partitioning**. It results in associating to each input factor an optimized partition of the output space and a normalized score characterizing the influence of the parameter in driving the output from one region of the partition to the complementary one (see Fig. 1). The optimized sensitive partitioning approach thus aims to find for any model input the two most contrasted model behaviors (defined as regions of the output space) that are influenced by this input.
Two main ingredients are required to define an optimized sensitive partitioning: first a sensitivity measure relating the sensitivity of an input to a partition of the output space, second an optimization procedure that looks for a partition that maximizes the sensitivity score.

Sensitivity measures and optimization algorithms

We propose to use the cluster-based indices defined in [3]. They are defined as Sobol' indices (first order, or total) of the membership functions (MF) associated to a clustering

of the output space. MFs are typically defined in $[0, 1]$ and characterize the degree of membership of any point of the output space to a given cluster. We introduce two optimization algorithms dedicated to the case of a 2-partition problem. In this case, we look for the binarization $(C^*, \bar{C}^*)$ of the output space that maximizes a given cluster-based index.

- The first algorithm is based on a exhaustive search on a set of patches obtained using a first clustering of the output space. It can be applied with any sensitivity measure.
- The second one targets specifically first order indices and uses a property specific to the optimization with this criterion: the fact that two patches having very correlated histograms associated to the distribution of $X_i|Y \in C^*$ belong to the same region of the optimal partition $(C^*, \bar{C}^*)$. Histogram correlation can thus trigger efficiently the clustering of patches and allows to reach finer resolutions in the results.

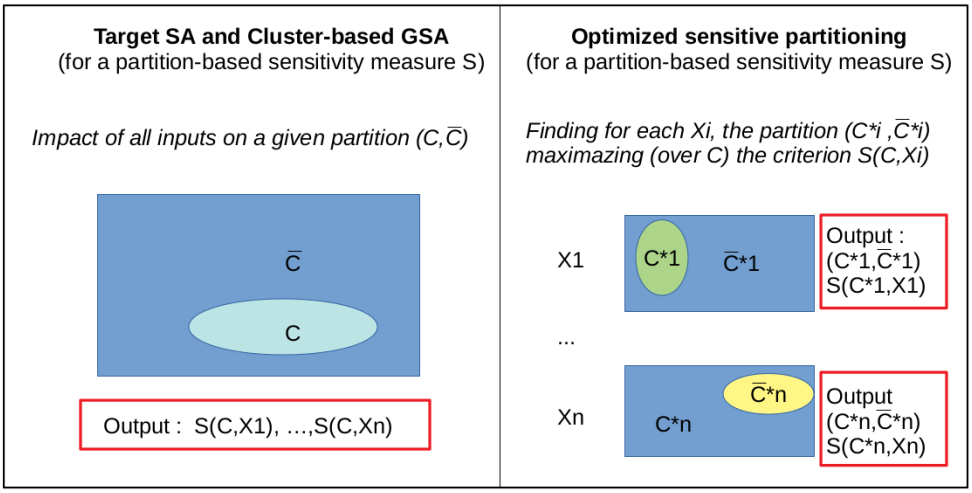

Figure 1: Principle of Optimized Sensitive Partitioning as opposed to Target SA and cluster-based GSA (here in the case of a partitioning into two regions).

First Results

We present the application of the method on different examples.

- First, a 1d example $f(X_1, X_2) = (sign(X_1) \cdot |X_2|)$, with X_1 and X_2 having uniform distributions in $[-1, 1]$ is considered. This example is used for validation purpose as the optimization problem in this case can be solved analytically.
- The two numerical algorithms are then illustrated on a 2d toy model allowing a 2d representation of the optimized partition and on a model with dynamic outputs to show the interest of the approach.

[1] Spear, R. C. and Hornberger, G. M. *Eutrophication in peel inlet—II. Identification of critical uncertainties via generalized sensitivity analysis.* Water research, 14(1), 43-49,1980.

[2] Marrel, A. and Chabridon, V. *Statistical developments for target and conditional sensitivity analysis: application on safety studies for nuclear reactor.* Reliability Engineering & System Safety, 107711, 2021.

[3] Roux, S., Buis, S., Lafolie, F. and Lamboni, M. *Cluster-based GSA: Global sensitivity analysis of models with temporal or spatial outputs using clustering.* Environmental Modelling & Software, vol. 140, p. 105046, 2021.