A new method for sensitivity analysis of models with dynamic and/or spatial outputs

Buis Samuel[*1], Ruget Françoise[1], Lafolie François[1], Lamboni Matieyendou[2], Roux Sébastien[3]*

[1] INRA, UMR 1114 EMMAH, 84914, Avignon, France, samuel.buis@inra.fr
[2] University of Guyane, 228-UMR-Espace dev, Cayenne, France
[3] INRA, UMR MISTEA, 34060, Montpellier, France

Introduction

Global Sensitivity Analysis (GSA) is recognized as a powerful tool for measuring the impact of models inputs on simulated outputs under prescribed inputs' variability. Although many simulation models, among which crop models, produce temporal and/or spatial data, extracting relevant information from GSA of such outputs is still challenging. This requires the use of Multivariate Sensitivity Analysis methods (MSA) that are often based on a dimension reduction principle: model outputs are projected onto predefined or data-driven orthogonal bases such as polynomial or eigenvectors (Lamboni et al., 2011). They are however so far limited by the selection of the associated bases which is constrained by orthogonality requirements. Indeed, these bases do not always allow extracting relevant and interpretable information on structural properties of multivariate outputs. More applicable MSA methods are thus expected to be developed (Wei et al., 2015). In this work, we propose a new MSA method combining GSA and clustering.

Cluster-based GSA

Clustering methods have been designed to identify groups of similar objects in multivariate data sets. They may thus be particularly adapted to capture the variability of behaviors of models' temporal and/or spatial outputs. However, while binary clustering has been extensively used in scalar sensitivity analysis to assess the importance of factors leading to a region of interest (Raguet and Marrel, 2018), there is still a lack of quantitative sensitivity analysis methods taking benefit of a clustering of multivariate outputs with any number of clusters.

The main idea of the proposed method is to apply clustering to model outputs simulated on a numerical design-of-experiment generated using a given GSA method, and to compute standard GSA indices (e.g. Sobol' indices) not on the models outputs but on new variables indicating the membership of each output to the different clusters (see Fig. 1). We propose to use a fuzzy clustering method: the new variables are thus the so-called membership functions (MF, valued in [0, 1]) that quantify the degree of membership of any model simulated output to each cluster. The computation of sensitivity indices on either the MF or MF differences allows discussing which parameters influence the membership to a given cluster or drive the output from one cluster to another. A generalized sensitivity index (Lamboni et al, 2011) is also introduced to quantify the overall contribution of the parameters wrt any change of clusters.

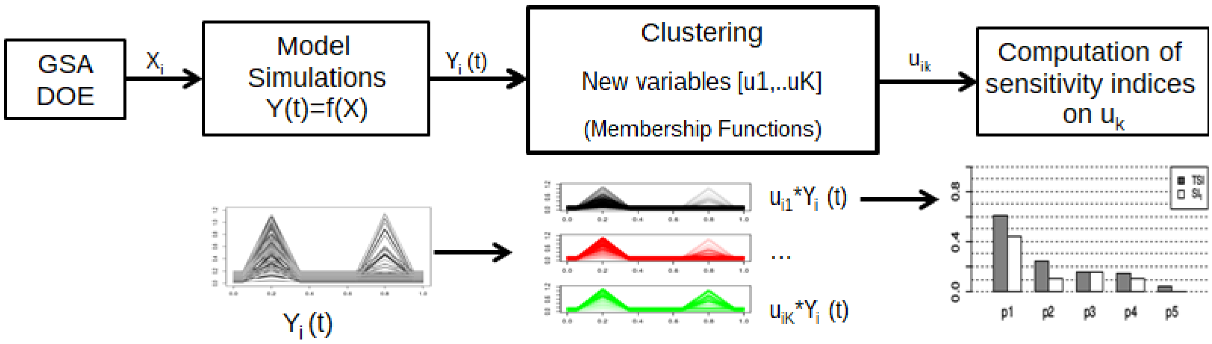

Figure 1. Workflow of the Cluster-based GSA. X represents the vector of model inputs that varies in the sensitivity analysis, Y(t) the (temporal in this case) output simulated by the model. K is the number of clusters, i the index in the design-of-experiment.

Applications

The method has been applied using Sobol' and FAST GSA methods to:

(i) a dedicated toy model producing temporal signals with one or two maxima in response to five parameters,

(ii) the Cantis model (Garnier et al., 2003) simulating the transformations of carbon and nitrogen in soils (10 parameters varying),

(iii) the Stics crop model (Coucheney et al. 2015), on the Multi-Model Ideotyping Agmip 2019 exercise (27 parameters varying).

Results have shown that the model behaviors can be efficiently reported by the newly proposed method.

Conclusions

The proposed method is particularly adapted to models with dynamic and/or spatial outputs that produce distinguishable sets of responses, i.e. when clustering of these outputs lead to well separated and interpretable clusters. In this case, it is particularly powerful for identifying the model inputs that drive these different behaviors. The method is generic wrt clustering and GSA method used.

Keywords
Sensitivity analysis, multivariate outputs, generalized sensitivity indices

References

Coucheney E, Buis S, Launay M, Constantin J, Mary B, de Cortázar-Atauri I G, ..., Léonard J (2015). Accuracy, robustness and behavior of the STICS soil–crop model for plant, water and nitrogen outputs: evaluation over a wide range of agro-environmental conditions in France. Environmental Modelling & Software, 64, 177-190.
Garnier P, Neel C, Aita C, Recous S, Lafolie F, Mary B (2003). Modelling carbon and nitrogen dynamics in a bare soil with and without straw incorporation. European Journal of Soil Science, 54(3), 555-568.
Lamboni M, Monod H, Makowski D (2011) Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. Reliability Engineering & System Safety, 96(4), 450-459.
Raguet H, Marrel A (2018). Target and Conditional Sensitivity Analysis with Emphasis on Dependence Measures. arXiv preprint arXiv:1801.10047.
Wei P, Lu Z, Song J (2015) Variable importance analysis: a comprehensive review, Reliability Engineering & System Safety 142, 399–432.