ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Universität Potsdam

EGU General Assembly 2020

Session HS-8.1.6: Handling Uncertainties in Model Concepts, Parameters, Forcings and Forecasts: Diagnostics, Sensitivity, Inversion and Uncertainty Analysis

A comprehensive global sensitivity analysis using generic sampling designs by means of a combination of variance- and distribution-based approaches

Gabriele Baroni[1] and Till Francke[2]

(1) Università di Bologna (Italy), g.baroni@unibo.it

(2) University of Potsdam (Germany)

CC BY

Outline

A) **Motivations:** why and how to improve current best practices in global sensitivity analysis (GSA)?

B) A new **C**ombined **V**ariance- and **D**istribution-based global sensitivity analysis **CVD - GSA**: how does it work?

C) **Tests** to three analytic functions and one hydrological model

D) **Conclusions** and Outlook

cc BY

A) Motivation

1. **global sensitivity analysis (GSA) is an important tool for**
   - supporting model developments
   - processes understanding

2. State-of-the-art **variance-based approach** (Saltelli et al., 2010)
   - Identify important parameters (main effect) and interactions (total effect – main effect)
   - It works also on non-scalar factors (e.g., Baroni and Tarantola, 2014)

3. **Limitations**
   - Specific sampling design
   - Relatively high number of simulation runs
   - Issues with non-gaussian distributions (distribution-based approaches)
   - how to improve?

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)} \qquad T_i = \frac{E[V(Y|X_{\sim i})]}{V(Y)}$$

Where:

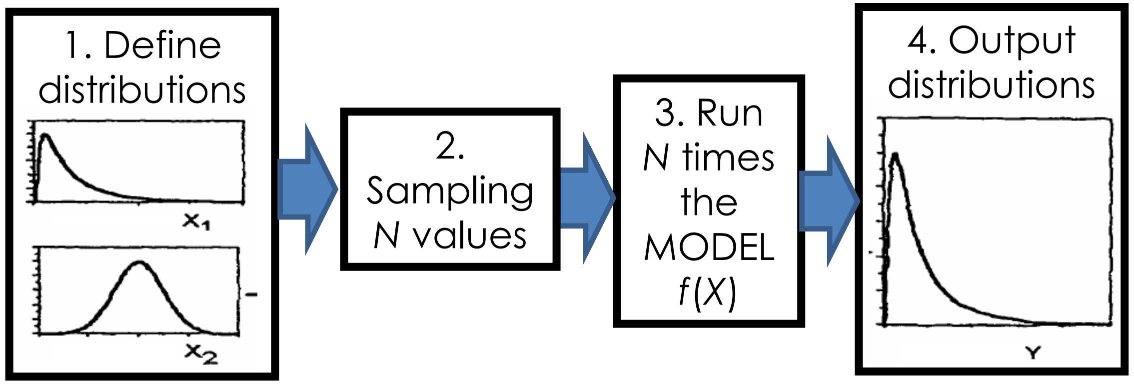
S_i main effect of factor i
T_i total effect of factor i
E mean operator
V variance operator
$Y\,|\,X_i$ output Y conditioned to X_i
$Y\,|\,X_{\sim i}$ output Y conditioned to all but not X_i

Interaction $I_i = T_i - S_i$

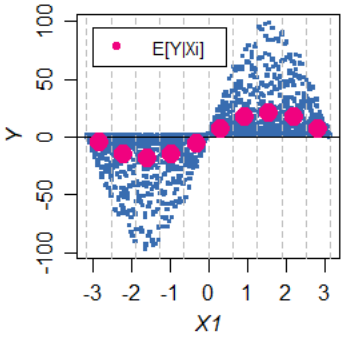
CC BY

B) An effective Combined Variance- and Distribution-based strategy (CVD)

1. Run generic Monte Carlo simulations



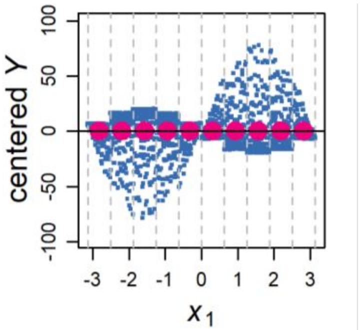
2. Main effect is estimated based on a filtering approach (Kucherenko et al., 2017)



- X_i is divided in m intervals
- $E[Y|X_i]$ is calculated in each m
- Main effect S_i is estimated

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)}$$

3. The main effect is removed from the input-output space



By removing the conditional mean $E(Y|X_i)$, we obtain m centralized conditional distributions

$$P(Y|X_i) = P(Y|X_i) - E(Y|X_i)$$

4. Distribution-based measure (e.g., Kolmogorov-Smirnov test - KS) is used to estimate the interaction term I_i by comparing the m centralized conditional distributions

$$I_i = median\{KS[P(Y|X_{i=j}), P(Y|X_{i=j+1})]\}$$

With $j = [1\ldots m]$

cc BY

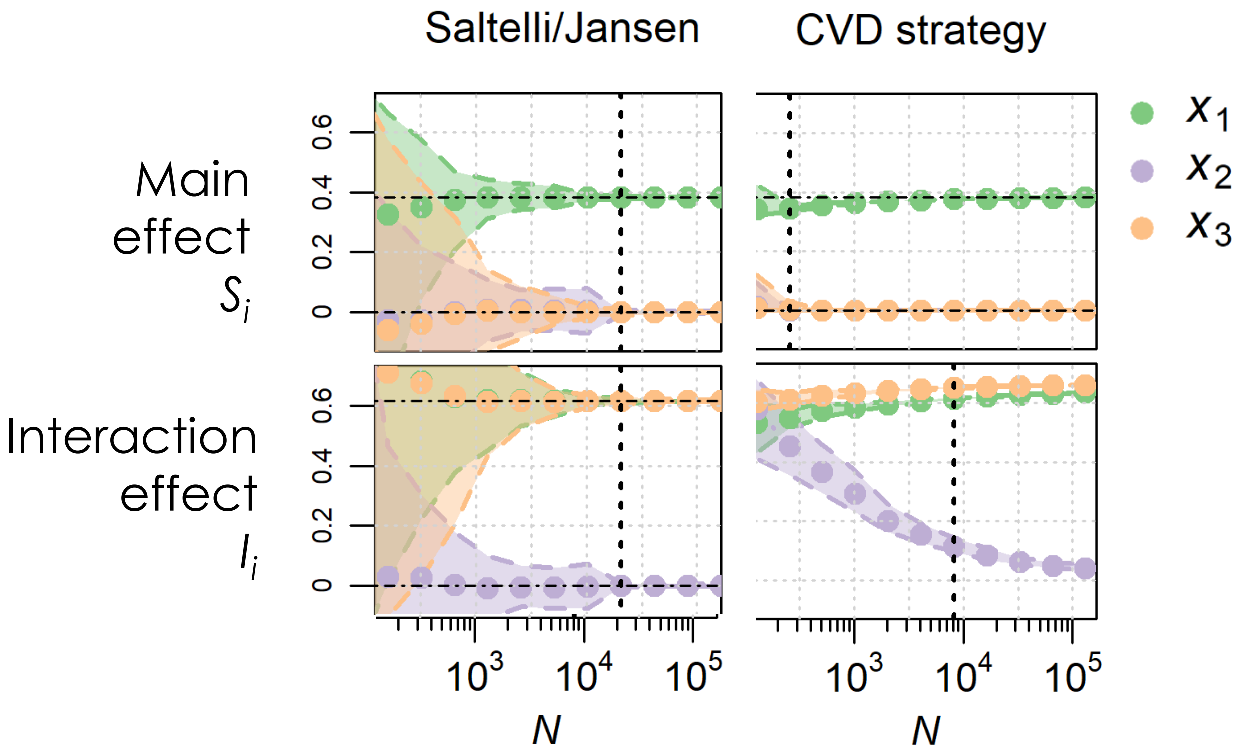
C) Tests to three analytic functions and one hydrological model

e.g., Ishigami-Homma function

$$y = \sin(x_1) + a \cdot \sin(x_2)^2 + b \cdot x_3^4 \cdot \sin(x_1)$$

with x_i in [-π, +π]
and a = 2 and b = 1

Number of simulations N ~ 10^5
Number of repetitions r = 100



CVD well estimates main and interaction with lower sample size

CC BY

C) Tests to three analytic functions and one hydrological model

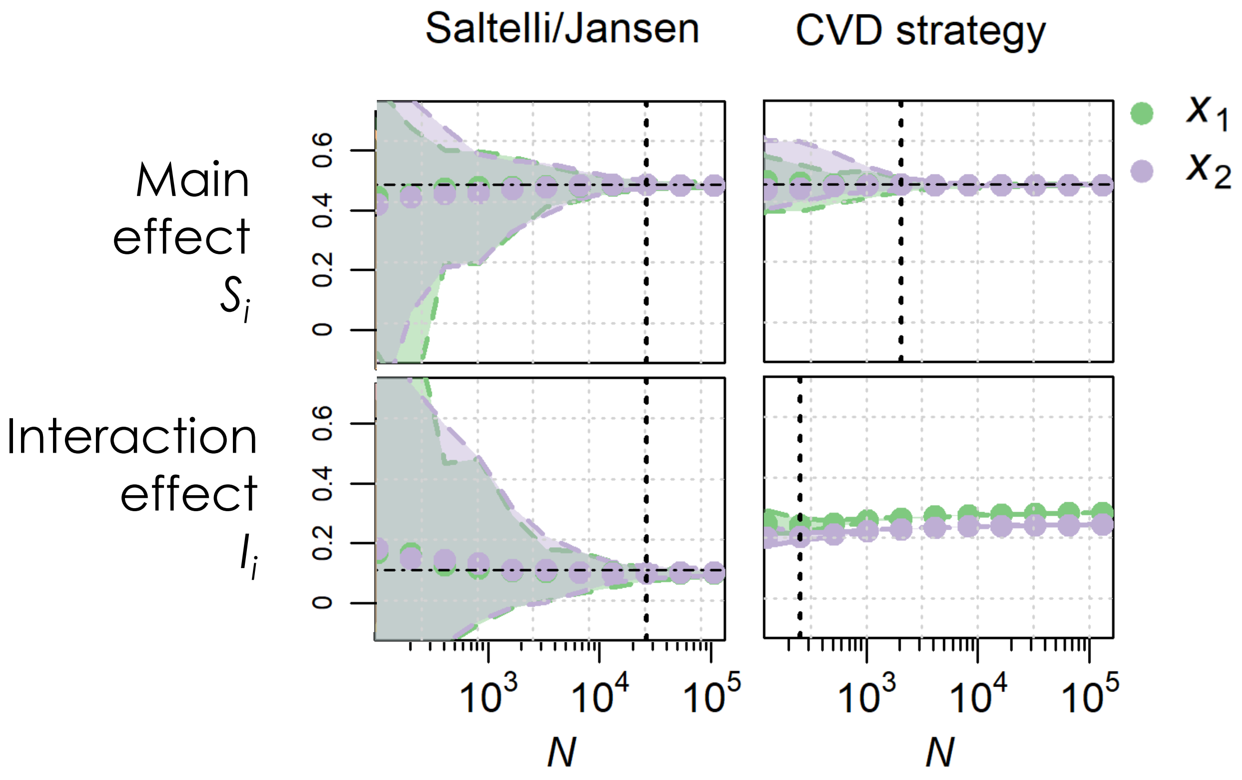
e.g., Skewed function

$$y = \frac{x_1}{x_2}$$

With x_1 and x_2 both follow χ2 distributions with degrees of freedom of 10 and 13.978, respectively

Number of simulations N ~ 10^5

Number of repetitions r = 100



Variance-based measures do not identify differences between the two parameters. The interaction terms *I* based on CVD identify differences between the two factors

cc BY

D) Conclusions and Outlook

Conclusions

- Main and interactions effects are estimated from a generic sampling design. CVD strategy can be easily integrated in any modelling framework
- The new approach converges faster than Saltelli/Jansen formula and combines the strength of variance and distribution-based approaches in exploring input-output space

Repository and document

- GitHub: https://github.com/baronig/GSA-cvd
- Manuscript under review: Baroni and Francke, An effective strategy for combining variance- and distribution-based global sensitivity analysis

Outlook

- Alternatives estimation of main and interaction term can be tested e.g., spline interpolation, δ-measure (Borgonovo et al., 2007)
- Testing on highly skewed modelling output

CC BY

References

Baroni, G., Tarantola, S., 2014. A General Probabilistic Framework for uncertainty and global sensitivity analysis of deterministic models: A hydrological case study. Environ. Model. Softw. 51, 26–34. https://doi.org/10.1016/j.envsoft.2013.09.022

Borgonovo, E., 2007. A new uncertainty importance measure. Reliab. Eng. Syst. Saf. 92, 771–784. https://doi.org/10.1016/j.ress.2006.04.015

Kucherenko, S., Song, S., 2017. Different numerical estimators for main effect global sensitivity indices. Reliab. Eng. Syst. Saf. 165, 222–238. https://doi.org/10.1016/j.ress.2017.04.003

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput. Phys. Commun. 181, 259–270. https://doi.org/10.1016/j.cpc.2009.09.018

Thank you for the attention

Gabriele Baroni

Department of Food and Agricultural sciences (DISTAL)

University of Bologna (Italy)

Till Francke

Institute of Environmental Science and Geography

University of Potsdam (Germany)

g.baroni@unibo.it

CC BY