Problem Settings & Overview of GSA Methods

Thierry A. Mara
thierry.mara@univ-reunion.fr

University of Reunion

12$^{\text{th}}$ SAMO Summer School
June 24-28 2024
University of Parma

Outline

Notations & Assumptions

- $y = f(\boldsymbol{x}) \sim p_y$ the scalar QoI
- $\boldsymbol{x} = (x_1, x_2, \ldots, x_d) \sim p_x$
- $x_i \sim p_{x_i} = \dfrac{\mathrm{d}F_{x_i}}{\mathrm{d}x_i}$, the marginal pdf ($F_{x_i}$ cdf) of x_i
- $\boldsymbol{X}$, $\boldsymbol{y}$, samples of $\boldsymbol{x}$ and y respectively
- $\mathcal{D} = \{0, 1, \ldots, d\}$, $\mathcal{D}_{-i} \bigcap \{i\} = \emptyset$, $\mathcal{D}_{+i} \bigcap \{i\} \neq \emptyset$
- $\boldsymbol{x} = (\boldsymbol{x}_\alpha, \boldsymbol{x}_{-\alpha})$ with $\boldsymbol{x}_\alpha \bigcap \boldsymbol{x}_{-\alpha} = \emptyset$
- $\alpha = \{i_1, \ldots, i_k\} \subseteq \mathcal{D} \Leftrightarrow \boldsymbol{x}_\alpha = \boldsymbol{x}_{i_1,\ldots,i_k} = (x_{i_1}, \ldots, x_{i_k})$

Outlook of SS2024

When conducting Global Sensitivity Analysis (GSA), two situatons are to be considered:

- Design-driven methods: $(\boldsymbol{X}, \boldsymbol{y})$ to be generated by the analyst/modeller (Mon-Tues)
- Data-driven methods: $\boldsymbol{X}$ or $(\boldsymbol{X}, \boldsymbol{y})$ given (Wed-Thu)

Outlook of SS2024

When conducting Global Sensitivity Analysis (GSA), two situatons are to be considered:

- Design-driven methods: $(\boldsymbol{X}, \boldsymbol{y})$ to be generated by the analyst/modeller (Mon-Tues)
- Data-driven methods: $\boldsymbol{X}$ or $(\boldsymbol{X}, \boldsymbol{y})$ given (Wed-Thu)

and two cases:

- $p_x = \prod_{i=1}^{d} p_{x_i}$, independence (Mon-Wed)
- $p_x \neq \prod_{i=1}^{d} p_{x_i}$, dependence (Thu)

Steps of GSA

One question to be asked: What is the goal of the GSA?

According to Saltelli et al.(2008): "Experience shows that a poor definition of the objective of the sensitivity analysis ... can lead to confused or inconclusive results."

They argue that:

- There are different sensitivity indices proposed in the literature
- They all provide different rankings
- It is recommended to define the objective of the analysis beforehand

Steps of GSA

One question to be asked: What is the goal of the GSA?

According to Saltelli et al.(2008): "Experience shows that a poor definition of the objective of the sensitivity analysis . . . can lead to confused or inconclusive results."

They argue that:

- There are different sensitivity indices proposed in the literature
- They all provide different rankings
- It is recommended to define the objective of the analysis beforehand

GSA can be conducted following different steps:

1. Define the goal of GSA
2. Choose the adequate QoI and the adequate sensitivity indices (SI)
3. (Design-driven): Define p_x
4. (Design-driven): Choose the adequate DOE
5. (Design-driven): Sample $\boldsymbol{x}$ (get $\boldsymbol{X}$) and Run the model (get $\boldsymbol{y}$)
6. Compute the SI and Conclude

GSA settings

Possible goals of the analysis (Saltelli et al. (2008)):

- Model corroboration. Is the inference robust? Is the model overly dependent on fragile assumptions ?
- Factor prioritization: Which factor is most deserving of further analysis or measurement ?
- Factor fixing: Can some factors or compartments of the model be fixed ?
- Regional SA: Which factors are mostly responsible for producing results in critical region of the output space ? (design space, failure detection, ..)
- Model interpretation: How the input factors combine to produce the model response ?
- ...

GSA settings

Several global sensitivity measures proposed in the litterature:

- Variance-based sensitivity indices: (Mon-Tue,Thu)
- Moment-independent sensitivity indices: (Wed-Thu)
  1. Distribution-based
  2. Entropy-based

Global Sensitivity Indices

Several global sensitivity measures proposed in the litterature:

- Variance-based sensitivity indices: (Mon-Tue,Thu)

$$S_i = \frac{\mathbb{V}\left[\mathbb{E}\left[y|x_i\right]\right]}{\mathbb{V}\left[y\right]} \quad (1)$$

$$T_i = \frac{\mathbb{E}\left[\mathbb{V}\left[y|\boldsymbol{x}_{-i}\right]\right]}{\mathbb{V}\left[y\right]} \quad (2)$$

$$Sh_i = \sum_{\alpha \subseteq \mathcal{D}_{-i}} \frac{|\alpha|!\,(d-1-|\alpha|)!}{d!} \left(S^{clo}_{\alpha\cup i} - S^{clo}_{\alpha}\right) \quad (3)$$

- Moment-independent sensitivity indices: (Wed-Thu)

$$\alpha_i = A\left[d(y, y|x_i)\right],\ (A : \mathbb{E}\left[\cdot\right] \mathit{or} \sup)$$

1. Distribution-based
2. Entropy-based

The sensitivity indices should preferably be scaled within (0,1).

Global Sensitivity Indices

Several global sensitivity measures proposed in the litterature:

- Variance-based sensitivity indices: (Mon-Tue,Thu)

$$S_i = \frac{\mathbb{V}\left[\mathbb{E}\left[y|x_i\right]\right]}{\mathbb{V}\left[y\right]} \quad (4)$$

$$T_i = \frac{\mathbb{E}\left[\mathbb{V}\left[y|\boldsymbol{x}_{-i}\right]\right]}{\mathbb{V}\left[y\right]} \quad (5)$$

$$Sh_i = \sum_{\alpha \subseteq \mathcal{D}_{-i}} \frac{|\alpha|!\,(d-1-|\alpha|)!}{d!}\left(S^{clo}_{\alpha \cup i} - S^{clo}_{\alpha}\right) \quad (6)$$

- Moment-independent sensitivity indices: (Wed-Thu)

$$\alpha_i = A\left[d(y, y|x_i)\right],\ (A, d : \mathbb{E}\left[\cdot\right] \textit{ or} \sup)$$

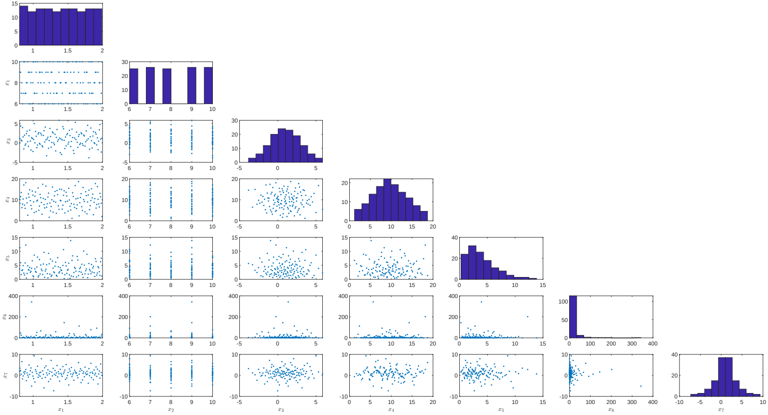
1. Distribution-based

$$\delta_i = \frac{1}{2}\int_{\mathbb{R}^2} |p_y - p_{y|x_i}| \mathrm{d}y \mathrm{d}x_i \quad (7)$$

$$\tau_i = \int_{\mathbb{R}} \sup |F_y - F_{y|x_i}| \mathrm{d}x_i \quad (8)$$

Sampling from p_x

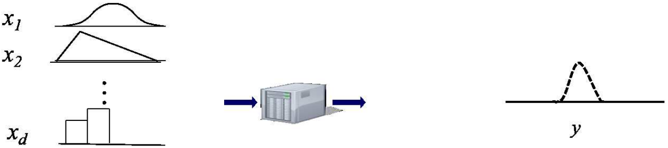
Sampling is an important issue (discussed on Mon. afternoon)



The Sampler must be preferably space-filling.

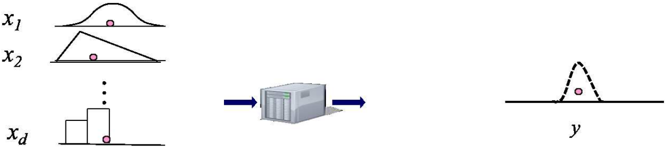
Run the model

Monte Carlo simulations are a numerical way to propagate the input uncertainty into the model.



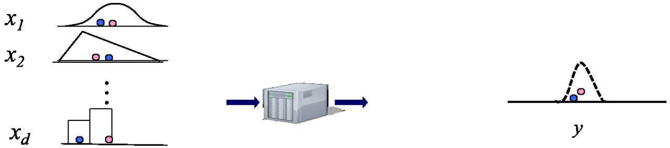
Run the model

Monte Carlo simulations are a numerical way to propagate the input uncertainty into the model.



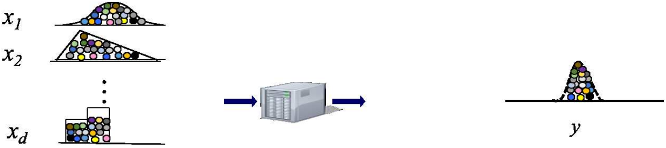
Run the model

Monte Carlo simulations are a numerical way to propagate the input uncertainty into the model.



Run the model

Monte Carlo simulations are a numerical way to propagate the input uncertainty into the model.



Then, p_y and other statistics of y can be inferred from the Monte Carlo sample.

Compute the SI's and conclude

For this job, use Siml@b a recent online tool developed by the european Commission to perform GSA (more this afternoon)

Legal Notice Cookies policy

European Commission

SIML@B: an EC tool for sensitivity analysis

Siml@b GSA Variance-Based Moment-Free Screening Terms & Conditions Thanks

Introduction

The Competence Centre on Modelling (CC-MOD) of the Joint Research Centre (Directorate General of the European Commission) has been developing the present toolbox in order to help modellers perform uncertainty analysis and sensitivity analysis of their models.

The issue of uncertainty: Computer models can be highly parameterised and subject to different plausible assumptions. Therefore, defining the precise value of each input variable can be challenging. Meanwhile, good modelling practice requires that uncertainty in model inputs be accounted for in simulation studies. GSA (Global Sensitivity Analysis) can facilitate this task by pointing out those uncertain input variables (denoted $\mathbf{x} = (x_1, \ldots, x_d)$) mostly responsible for the model output (denoted y) uncertainty. Knowing this, modellers can focus their attention on the sensitive variables.

About `SIML@B`

`SIML@B` provides a set of online tools to perform uncertainty analysis and sensitivity analysis (UASA) of model output. The diagram below provides a classification of some existing methods to perform global sensitivity analysis. **For the time being, only the method in the green boxes are implemented**. While the `data-driven` methods only require an input/output Monte Carlo sample, `design-driven` methods require specific design of experiments which are provided by the module `Sampler`. See tabs, `Moment-Free`, `Variance-based` or `Screening` for a short description of the methods and the definition of the associated sensitivity indices. By clicking on a green box, you will be redirected to the corresponding R-Shiny online tool from which you can upload your dataset and perform the analysis.

Thank You!