# Link between the sensitivity indices at different scales

Bruchou, C. * and Saint-Geours, N. †

*Biosp INRA
84914 Avignon, France
† itk
CEEI CAP ALPHA - Avenue de l'Europe, 34830 Clapiers, France

## 1 Objectives

It is well known that importance of the impact of input factors of a numerical model may depends on the range of variation of factors. That importance can also vary by regions in the domain of definition. Local and Global sensitivity analyses evaluate factor impacts at a point or along the whole range of the domain respectively (Saltelli et al, 2004). Sobol and Kucherenko (2009) defined the derivative based indices and showed their link with global ones. The first objective of that work in progress is to show the link between variance based sensitivity indices at different scales. The scales are defined by cutting the domain of definition. The second objective, that is a practical one, aims to estimate the terms of the relationship linking the indices. A method of estimation is proposed and evaluated on an analytical deterministic model.

## 2 Notations and definitions

- $f$ a real square-integrable function defined on hypercube $\Omega = [0,1]^K$, $\mathcal{F} = \{1,\dots,K\}$:
  - $\mathbf{X}^{(\mathcal{F})} = (X^{(1)},\dots,X^{(K)})$ the vector of the variates or factors of $f$, $\mathbf{x}$ a value of $\mathbf{X}^{(\mathcal{F})}$,
  - $\mathcal{F} = \mathcal{I} \cup \{\sim\mathcal{I}\}$, with $\mathcal{I} = \{1,\dots,M\}$ and $\sim\mathcal{I}$ complement of $\mathcal{I}$ in $\mathcal{F}$,
  - $\mathbf{X}^{(\mathcal{F})} = (\mathbf{X}^{(\mathcal{I})}, \mathbf{X}^{(\sim\mathcal{I})})$, $\mathbf{x} = (\mathbf{x_i}, \mathbf{x_j})$ with $\mathbf{x_i} \in \Omega_{\mathcal{I}} = [0,1]^M$ and $\mathbf{x_j} \in \Omega_{\sim\mathcal{I}} = [0,1]^{K-M}$,
- a tiling of $\Omega_{\mathcal{F}}$ :
  - $\forall k \in \mathcal{F}$, $[0,1] = \bigcup\limits_{q\in\mathcal{Q}} I_q^{(k)}$, $\mathcal{Q} = \{1,\dots,Q\}$, with $|I_q^{(k)}| = \delta_q^{(k)}$, a partition of $[0,1]$,
  - $\omega_{\mathbf{i}}^{(\mathcal{I})} = \prod\limits_{l=1}^{M} I_{i_l}^{(l)}$ cuboid in $\Omega_{\mathcal{I}}$ , having volume $\delta_{\mathbf{i}} = \prod\limits_{l=1}^{M} \delta_{i_l}^{(l)}$,
  - $\omega_{\mathbf{i},\mathbf{j}} = \omega_{\mathbf{i}}^{(\mathcal{I})} \times \omega_{\mathbf{j}}^{(\sim\mathcal{I})} \in \Omega$, $\mathbf{i} \in \mathcal{Q}^M$, $\mathbf{j} \in \mathcal{Q}^{K-M}$ cuboid in $\Omega$, $\Omega = \bigcup\limits_{\mathbf{i}\in\mathcal{Q}^K} \omega_{\mathbf{i}}^{(\mathcal{F})}$.
- statistics:
  - $\mu_{\mathbf{i},\mathbf{j}} = \dfrac{1}{\delta_{\mathbf{i}}\delta_{\mathbf{j}}} \displaystyle\int_{\omega_{\mathbf{i}}^{(\mathcal{I})}} \int_{\omega_{\mathbf{j}}^{(\sim\mathcal{I})}} f(\mathbf{x_i}, \mathbf{x_j})\, d\mathbf{x_j}\, d\mathbf{x_i}$, mean of $f$ in $\omega_{\mathbf{i},\mathbf{j}}$, $\mu_\Omega$ mean of $f$ in $\Omega$,
  - $\sigma^2_{\mathbf{i},\mathbf{j}} = \dfrac{1}{\delta_{\mathbf{i}}\delta_{\mathbf{j}}} \displaystyle\int_{\omega_{\mathbf{i}}^{(\mathcal{I})}} \int_{\omega_{\mathbf{j}}^{(\mathcal{J})}} \left(f(\mathbf{x_i}, \mathbf{x_j}) - \mu_{\mathbf{i},\mathbf{j}}\right)^2 d\mathbf{x_j}\, d\mathbf{x_i}$ variance of $f$ in $\omega_{\mathbf{i},\mathbf{j}}$, $\sigma^2_\Omega$ variance of $f$ in $\Omega$,
  - part of the variance of the regression of $f(\mathbf{X})$ to $\mathbf{X}^{(\mathcal{I})}$ in $\Omega$ (resp. $\omega_{\mathbf{i},\mathbf{j}}$) to $\sigma^2_\Omega$ (resp. $\sigma^2_{\mathbf{i},\mathbf{j}}$) :
    $\mathbf{P}_\Omega^{(\mathcal{I})} = \dfrac{\mathbb{V}_{\Omega_{\mathcal{I}}}\left(\mathbb{E}_{\Omega_{\sim\mathcal{I}}}(f(\mathbf{X})|\mathbf{X}^{(\mathcal{I})})\right)}{\sigma^2_\Omega}$ (resp. $\mathbf{P}_{\omega_{\mathbf{i},\mathbf{j}}}^{(\mathcal{I})} = \mathbb{V}_{\omega_{\mathbf{i}}^{(\mathcal{I})}}\left(\mathbb{E}_{\omega_{\mathbf{j}}^{(\sim\mathcal{I})}}(f(\mathbf{X})|\mathbf{X}^{(\mathcal{I})})\right)$ ),
  - $\bar{f}$ the function defined on grid $\mathcal{Q}^K$ : $\bar{f}(\mathbf{u}) = \frac{1}{\delta_{\mathbf{u}}} \int_{\omega_{\mathbf{u}}} f(\mathbf{x}) d\mathbf{x}$, $\mathbf{u} \in \mathcal{Q}^K$,
  - $\tilde{\mathbf{P}}_{\mathcal{Q}^K}^{(\mathcal{I})} = \dfrac{\mathbb{V}_{\mathcal{Q}^M}\left(\mathbb{E}_{\mathcal{Q}^{K-M}}(\bar{f}(\mathbf{U})|\mathbf{U}^{(\mathcal{I})})\right)}{\sigma^2_B}$, where $\sigma^2_B = \sum\limits_{\mathbf{u}\in\mathcal{Q}^K} \delta_{\mathbf{u}} \left(\bar{f}(\mathbf{u}) - \mu_\Omega\right)^2$.

*Corresp. author: bruchou@avignon.inra.fr
†nathalie.saint-geours@itk.fr

## 3 Relationship between the sensitivity indices at two scales

Let $m^{(\mathcal{I})}_{\mathbf{i},\mathbf{j}'}(\mathbf{x_i}) = \frac{1}{\delta_{\mathbf{j}'}} \int_{\omega^{(\mathcal{J})}_{\mathbf{j}'}} f(\mathbf{x_i}, \mathbf{x_j})\, d\mathbf{x_j}$ (The exponent $(\mathcal{I})$ highlights the space of definition $\Omega_{\mathcal{I}}$ of $m$). The covariance of $m^{(\mathcal{I})}_{\mathbf{i},\mathbf{j}'}$ and $m^{(\mathcal{I})}_{\mathbf{i},\mathbf{j}''}$ is : $\mathbb{C}(m^{(\mathcal{I})}_{\mathbf{i},\mathbf{j}'}, m^{(\mathcal{I})}_{\mathbf{i},\mathbf{j}''}) = \dfrac{1}{\delta_{\mathbf{i}}} \displaystyle\int_{\omega^{(\mathcal{I})}_{\mathbf{i}}} \left(m^{(\mathcal{I})}_{\mathbf{i},\mathbf{j}'}(\mathbf{x_i}) - \mu_{\mathbf{i},\mathbf{j}'}\right)\left(m^{(\mathcal{I})}_{\mathbf{i},\mathbf{j}''}(\mathbf{x_i}) - \mu_{\mathbf{i},\mathbf{j}''}\right) d\mathbf{x_i}$.

The index $\mathbf{P}^{(\mathcal{I})}_{\Omega}$ verifies:

$$\mathbf{P}^{(\mathcal{I})}_{\Omega} = \underbrace{\sum_{\substack{\mathbf{i}\in\mathcal{Q}^M \\ \mathbf{j}\in\mathcal{Q}^{K-M}}} \delta_{\mathbf{i}}\,(\delta_{\mathbf{j}})^2\, \frac{\sigma^2_{\mathbf{i},\mathbf{j}}}{\sigma^2_{\Omega}}\, \mathbf{P}^{(\mathcal{I})}_{\omega_{\mathbf{i},\mathbf{j}}}}_{A} + \underbrace{\frac{\sigma^2_B}{\sigma^2_{\Omega}}\, \tilde{\mathbf{P}}^{(\mathcal{I})}_{\mathcal{Q}^K}}_{B} + \underbrace{\sum_{\substack{\mathbf{i}\in\mathcal{Q}^M \\ \mathbf{j}',\mathbf{j}''\in\mathcal{Q}^{K-M} \\ \mathbf{j}'\neq\mathbf{j}''}} \delta_{\mathbf{i}}\,\delta_{\mathbf{j}'}\delta_{\mathbf{j}''}\, \frac{\mathbb{C}(m^{(\mathcal{I})}_{\mathbf{i},\mathbf{j}'}, m^{(\mathcal{I})}_{\mathbf{i},\mathbf{j}''})}{\sigma^2_{\Omega}}}_{C} \tag{1}$$

.

The index $\mathbf{T}^{(\mathcal{I})}_{\Omega} = 1 - \mathbf{P}^{(\sim\mathcal{I})}_{\Omega}$ verifies:

$$\mathbf{T}^{(\mathcal{I})}_{\Omega} = \underbrace{\sum_{\substack{\mathbf{i}\in\mathcal{Q}^M \\ \mathbf{j}\in\mathcal{Q}^{K-M}}} \delta_{\mathbf{j}}\,\delta^2_{\mathbf{i}}\, \frac{\sigma^2_{\mathbf{i},\mathbf{j}}}{\sigma^2_{\Omega}}\, \mathbf{T}^{(\mathcal{I})}_{\omega_{\mathbf{i},\mathbf{j}}}}_{D} + \underbrace{\frac{\sigma^2_B}{\sigma^2_{\Omega}}\, \tilde{\mathbf{T}}^{(\mathcal{I})}_{\mathcal{Q}^K}}_{E} - \underbrace{\sum_{\substack{\mathbf{j}\in\mathcal{Q}^{K-M} \\ \mathbf{i}',\mathbf{i}''\in\mathcal{Q}^{M} \\ \mathbf{i}'\neq\mathbf{i}''}} \delta_{\mathbf{j}}\,\delta_{\mathbf{i}'}\delta_{\mathbf{i}''}\, \frac{\mathbb{C}(m^{(\sim\mathcal{I})}_{\mathbf{i}',\mathbf{j}}, m^{(\sim\mathcal{I})}_{\mathbf{i}'',\mathbf{j}})}{\sigma^2_{\Omega}}}_{F} + \underbrace{\sum_{\substack{\mathbf{i}\in\mathcal{Q}^M \\ \mathbf{j}\in\mathcal{Q}^{K-M}}} \delta_{\mathbf{j}}\,\delta_{\mathbf{i}}\,(1-\delta_{\mathbf{i}})\, \frac{\sigma^2_{\mathbf{i},\mathbf{j}}}{\sigma^2_{\Omega}}}_{G} \tag{2}$$

- $A$ and $D$ are the contribution of indices computed *within* cuboids,
- $B$ and $E$ are the contribution of indices of the mean of $f$ *between* cuboids,
- $C$ quantifies the similarity of the means of $f$ on the margin $\omega_{\mathcal{I}}$ between the cuboids having the face $\omega_{\mathcal{I}}$ in common,
- $F$ quantifies the similarity of the means of $f$ on the margin $\omega_{\sim\mathcal{I}}$ between the cuboids having the face $\omega_{\sim\mathcal{I}}$ in common,
- $G$ is a constant that depends on variance of factors in cuboids and an expression of their volumes,
- main and total sensitivity indices of a factor $X^{(k)}$ are easily obtained with $\mathcal{I} = \{k\}$.

## 4 Example

A polynomial test model with five variables is defined on hypercube $[0,1]^5$. The edges of the hypercube are uniformly cut into $Q$ intervals. That cutting defines $Q^5$ cuboids. Real sensitivity indices are computed at global and cuboids scale. A nested lhs design is proposed to sample cuboids and points inside them to estimate the terms of the relationships (1) and (2). Different sample size and value of $Q$ are used to evaluate the design.

## 5 Perspectives

The relationship between sensitivity indices at different scales being establised, that work will be continued in the following directions:

- trying to improve the estimation by means of a metamodel,
- finding the relationship between the indices with a hierarchical cutting of the domain,
- determining the relevant scale in environmental studies using a numerical model.

## 6 References


Saltelli et al (2004) *Sensitivity Analysis*, Wiley, New York.

I.M. Sobol, , S. Kucherenko, Derivative based global sensitivity measures and their link with global sensitivity indices, *Mathematics and Computers in Simulation*, Volume 79, Issue 10, June 2009, Pages 3009–3017.



**Acknowledgments**

This work was financially supported by the ANR project SimTraces (ANR-2011-CESA-008-01).