Some properties of variance-based sensitivity indices for spatially distributed models

Nathalie Saint-Geours[1] **Christian Lavergne**[2]
Jean-Stéphane Bailly[1] **Frédéric Grelot**[3]

[1]AgroParisTech, UMR TETIS, Montpellier.
nathalie.saint-geours@teledetection.fr

[2]Institut de Mathématiques et de Modélisation de Montpellier, UM2.

[3]Cemagref, UMR G-EAU, Montpellier.

GDR MASCOT NUM
Villard de Lans - March, 23^{th}, 2011

General framework

General framework

- Numerical model $Y = f(X_1, \dots X_d)$:
  - black box
  - deterministic
  - uncertain inputs and outputs

General framework

- Numerical model $Y = f(X_1, \ldots X_d)$:
  - black box
  - deterministic
  - uncertain inputs and outputs
- Variance-based sensitivity analysis
  - Contribution of X_i to the variance of model output Y
  - Sobol' sensitivity indices S_i :

$$S_i = \frac{Var\left(E\left[Y \mid X_i\right]\right)}{\mathrm{Var}\left(Y\right)}$$

General framework

- Numerical model $Y = f(X_1, \ldots X_d)$:
  - black box
  - deterministic
  - uncertain inputs and outputs
- Variance-based sensitivity analysis
  - Contribution of X_i to the variance of model output Y
  - Sobol' sensitivity indices S_i :

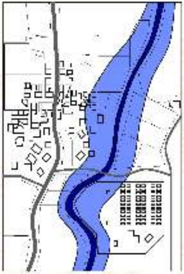
$$S_i = \frac{Var\left(E\left[Y \mid X_i\right]\right)}{\mathrm{Var}\left(Y\right)}$$

- Properties of S_i if X_i or Y is spatially distributed ?

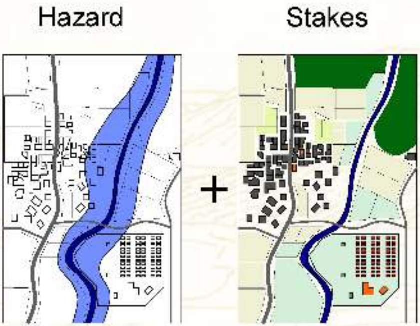
Model ACB-DE : assessment of flood damages

Model ACB-DE : assessment of flood damages

Hazard

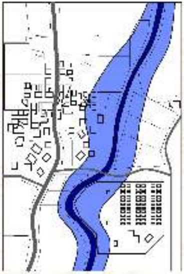


Model ACB-DE : assessment of flood damages



Model ACB-DE : assessment of flood damages

Hazard

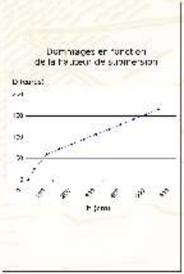


\+

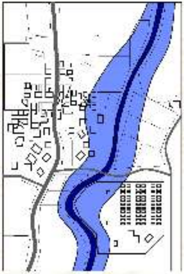
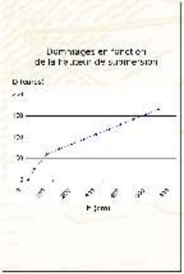
Stakes

\+

Vulnerability



Model ACB-DE : assessment of flood damages

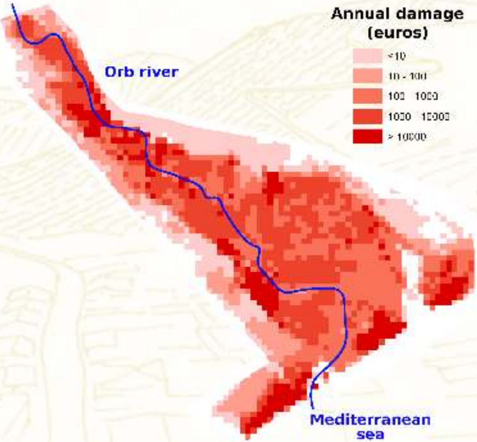
Hazard		Stakes		Vulnerability		Damage
	+		+		➤	1,3 M€

Model ACB-DE : inputs

	Nature	***Notation***
Water levels	Map	$Z_0(u) \in \mathbb{R}, \quad \forall u \in \mathbb{R}^2$
Landuse	Map	$Z_1(u) \in \mathbb{N}, \quad \forall u \in \mathbb{R}^2$
Terrain elevation	Map	$Z_2(u) \in \mathbb{R}, \quad \forall u \in \mathbb{R}^2$
Flood return period	Vector	$X_3 \in \mathbb{R}^5$
Infinite flood coefficient	Vector	$X_4 \in \mathbb{R}$
Damage curves	Vector	$X_5 \in \mathbb{R}^{100}$

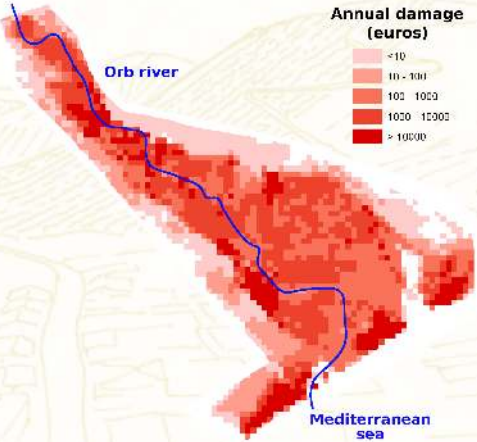
Model ACB-DE : output

Map of annual flood damages



Uncertainty of model output ?

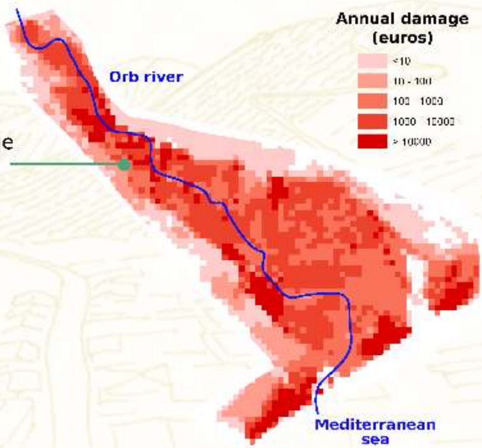
Different questions for different stakeholders !



Uncertainty of model output ?

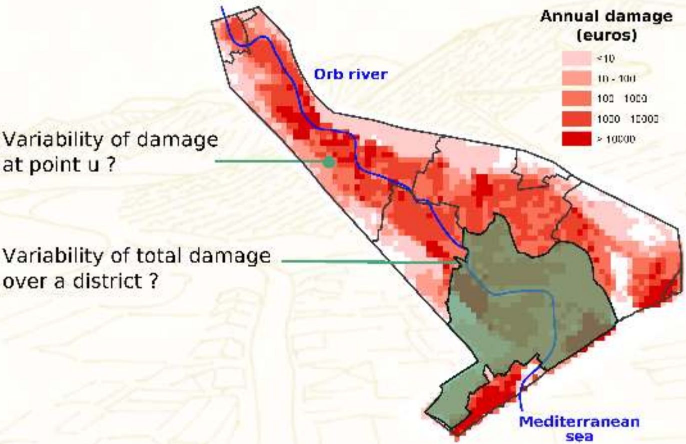
Different questions for different stakeholders !

Variability of damage at point u ?



Uncertainty of model output ?

Different questions for different stakeholders !



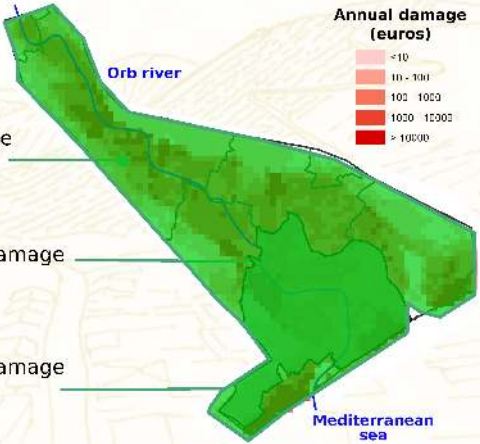
Uncertainty of model output ?

Different questions for different stakeholders !

Variability of damage at point u ?

Variability of total damage over a district ?

Variability of total damage over the territory ?



What the modeler wants to know

What the modeler wants to know

- For a chosen spatial support *(point, district, territory. . .)*

What the modeler wants to know

- For a chosen spatial support *(point, district, territory. . .)*
  - what's the uncertainty of total damage over the support ?

What the modeler wants to know

- For a chosen spatial support *(point, district, territory. . .)*
  - what's the uncertainty of total damage over the support ?
  - what model inputs explain this uncertainty ?

What the modeler wants to know

- For a chosen spatial support *(point, district, territory...)*
  - what's the uncertainty of total damage over the support ?
  - what model inputs explain this uncertainty ?
  - which input data should I put more money in ?

What the modeler wants to know

- For a chosen spatial support *(point, district, territory. . .)*
  - what's the uncertainty of total damage over the support ?
  - what model inputs explain this uncertainty ?
  - which input data should I put more money in ?

→ Sensitivity indices of model inputs on various spatial support

Spatially distributed model $\mathcal{M}$

Definition (Spatially distributed model $\mathcal{M}$)

$$Y = \mathcal{M}(X, Z)$$

Spatially distributed model $\mathcal{M}$

Definition (Spatially distributed model $\mathcal{M}$)

$$Y = \mathcal{M}(X, Z)$$

with :

- $X = (X_1, \ldots, X_d) \in \mathbb{R}^d$, joint pdf $p(X)$ *(economic parameters)*

Spatially distributed model $\mathcal{M}$

Definition (Spatially distributed model $\mathcal{M}$)

$$Y = \mathcal{M}(X, Z)$$

with :

- $X = (X_1, \dots, X_d) \in \mathbb{R}^d$, joint pdf $p(X)$ *(economic parameters)*
- $Z(u)$ Gaussian Random Field *(water levels)*

Spatially distributed model $\mathcal{M}$

Definition (Spatially distributed model $\mathcal{M}$)

$$Y = \mathcal{M}(X, Z)$$

with :

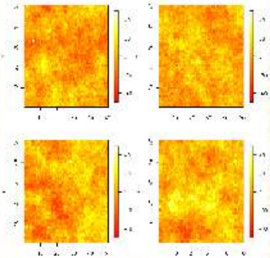
- $X = (X_1, \ldots, X_d) \in \mathbb{R}^d$, joint pdf $p(X)$ *(economic parameters)*
- $Z(u)$ Gaussian Random Field *(water levels)*
- $Y(u)$ a random field *(map of flood damage)* :

$$\forall u \in \mathcal{D}, \quad Y(u) = \psi\left[X, Z(u)\right]$$

with $\psi(.,.) : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$

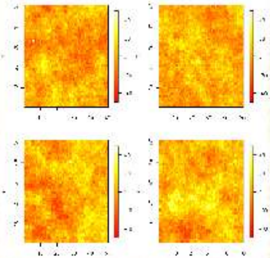
Model input : Gaussian random field $Z(u)$

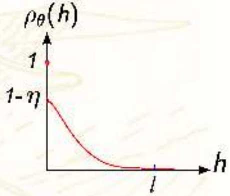
- Gaussian Random Field
- Domain $\mathcal{D} \subset \mathbb{R}^2$
- Order 2 stationary



Model input : Gaussian random field $Z(u)$

- Gaussian Random Field
- Domain $\mathcal{D} \subset \mathbb{R}^2$
- Order 2 stationary



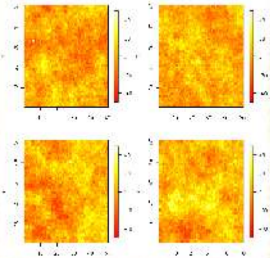


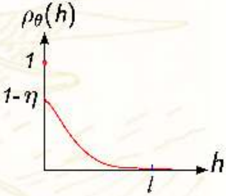
$$\forall u \in \mathcal{D}, \quad E\left[Z(u)\right] = \mu$$

$$\text{Cov}\left[\,Z(u), Z(u+h)\,\right] = \sigma_Z^2 \cdot \rho_\theta(h)$$

Model input : Gaussian random field $Z(u)$

- Gaussian Random Field
- Domain $\mathcal{D} \subset \mathbb{R}^2$
- Order 2 stationary





$$\forall u \in \mathcal{D}, \quad E\,[Z(u)] = \mu$$

$$\mathrm{Cov}\,[\,Z(u), Z(u+h)\,] = \sigma_Z^2 \cdot \rho_\theta(h)$$

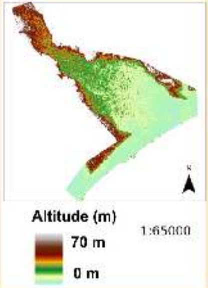
Parameter $\theta = (\eta$: nugget, l : range$)$

Example : terrain elevation

Example : terrain elevation

Spatial model input : Digital Elevation Model

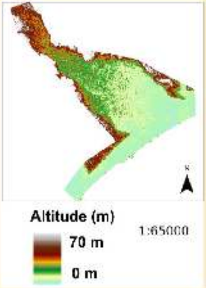
- Grid of 5m resolution
- Created from aerial photography
- Uncertainty : measure errors + interpolation errors



Example : terrain elevation

Spatial model input : Digital Elevation Model

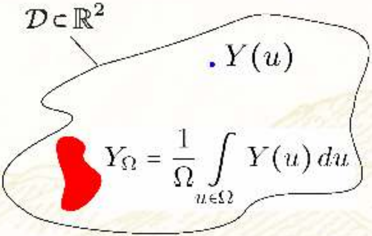
- Grid of 5m resolution
- Created from aerial photography
- Uncertainty : measure errors + interpolation errors



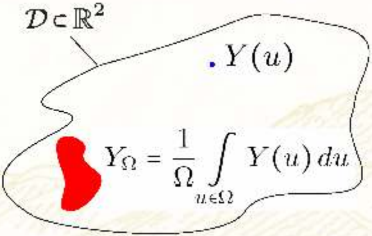
Modelling uncertainty on spatial model input

- 500 ground-validation points
- Estimation of mean error
- Estimation of covariance function

Model output : random field $Y(u)$



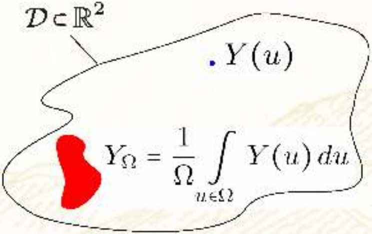
Model output : random field $Y(u)$



- Variance at point u :

$$\sigma_Y^2(u) = \mathrm{Var}\left[Y(u)\right]$$

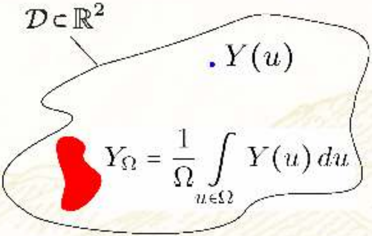
Model output : random field $Y(u)$



- Variance at point u :

$$\sigma_Y^2(u) = \mathrm{Var}[Y(u)]$$
$$= \sigma_Y^2 \qquad (Z \text{ stationary GRF})$$

Model output : random field $Y(u)$



- Variance at point u :

$$\sigma_Y^2(u) = \mathrm{Var}\left[Y(u)\right]$$
$$= \sigma_Y^2 \qquad \text{(Z stationary GRF)}$$

- Block variance :

$$\sigma_Y^2(\Omega) = \mathrm{Var}\left[Y_\Omega\right]$$

Point-based and zonal sensitivity indices

Definition (Point-based sensitivity indices)

Let $u \in \mathcal{D}$.
Output of interest : $Y(u)$

Point-based and zonal sensitivity indices

Definition (Point-based sensitivity indices)

Let $u \in \mathcal{D}$.
Output of interest : $Y(u)$

$$S_Z(u) = \frac{\mathrm{var}\left[E(Y(u) \mid Z)\right]}{\mathrm{var}\left[Y(u)\right]}$$

$$S_X(u) = \frac{\mathrm{var}\left[E(Y(u) \mid X)\right]}{\mathrm{var}\left[Y(u)\right]}$$

Point-based and zonal sensitivity indices

Definition (Zonal sensitivity indices)

Let $\Omega \subset \mathcal{D}$.

Output of interest : $Y_\Omega = \frac{1}{\Omega} \int\limits_{u \in \Omega} Y(u)\, du$

Point-based and zonal sensitivity indices

Definition (Zonal sensitivity indices)

Let $\Omega \subset \mathcal{D}$.

Output of interest : $Y_\Omega = \frac{1}{\Omega} \int_{u \in \Omega} Y(u)\, du$

$$S_Z(\Omega) = \frac{\text{var}\left[E(Y_\Omega \mid Z)\right]}{\text{var}\left[Y_\Omega\right]}$$

$$S_X(\Omega) = \frac{\text{var}\left[E(Y_\Omega \mid X)\right]}{\text{var}\left[Y_\Omega\right]}$$

Random field $E_Z Y(u)$

Definition

Random field $E_Z Y(u)$

Definition

Let $E_Z Y(u)$ be defined by :

$$\forall u \in \mathcal{D}, \quad E_Z Y(u) = E\left[Y(u) \,|\, Z(u)\right]$$

Random field $E_Z Y(u)$

Definition

Let $E_Z Y(u)$ be defined by :

$$\forall u \in \mathcal{D}, \quad E_Z Y(u) = E\left[Y(u) \,|\, Z(u)\right] = \bar{\psi}\left[Z(u)\right]$$

Random field $E_Z Y(u)$

Definition

Let $E_Z Y(u)$ be defined by :

$$\forall u \in \mathcal{D}, \quad E_Z Y(u) = E\left[Y(u) \,|\, Z(u)\right] = \bar{\psi}\left[Z(u)\right]$$

with

$$\bar{\psi}(z) = \int_{x \in \mathbb{R}^d} \psi(x, z) \cdot p(x)\, dx$$

Random field $E_Z Y(u)$

Definition

Let $E_Z Y(u)$ be defined by :

$$\forall u \in \mathcal{D}, \quad E_Z Y(u) = E\left[Y(u) \mid Z(u)\right] = \bar{\psi}\left[Z(u)\right]$$

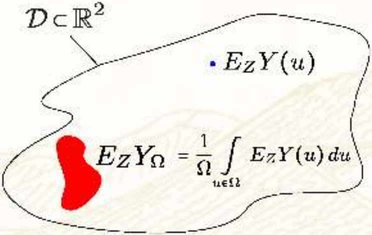
with

$$\bar{\psi}(z) = \int_{x \in \mathbb{R}^d} \psi(x, z) \cdot p(x)\, dx$$

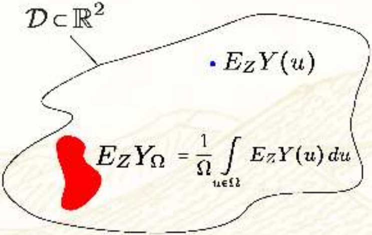
$$E_Z Y = \bar{\psi}(Z)$$

→ transformation of a Gaussian random field

Random field $E_Z Y(u)$



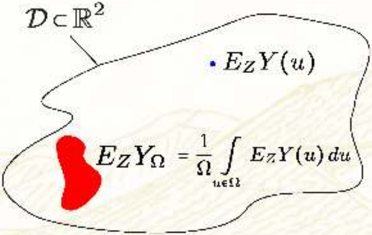
Random field $E_Z Y(u)$



- Variance at point u :

$$\sigma^2_{E_Z Y}(u) = \mathrm{Var}\left[E\left(Y(u) | Z(u)\right)\right]$$

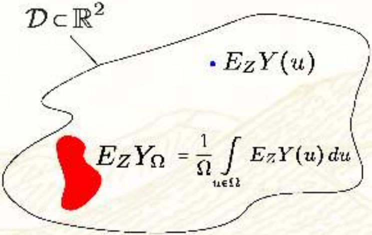
Random field $E_Z Y(u)$



- Variance at point u :

$$\sigma^2_{E_Z Y}(u) = \text{Var}\left[E\left(Y(u) | Z(u)\right)\right]$$
$$= \sigma^2_{E_Z Y} \qquad \text{(Z stationary GRF)}$$

Random field $E_Z Y(u)$



- Variance at point u :

$$\sigma^2_{E_Z Y}(u) = \text{Var}\left[E\left(Y(u) | Z(u)\right)\right]$$
$$= \sigma^2_{E_Z Y} \qquad \text{(Z stationary GRF)}$$

- Block variance :

$$\sigma^2_{E_Z Y}(\Omega) = \text{Var}\left[(E_Z Y)_\Omega\right] = \text{Var}\left[E_Z(Y_\Omega)\right]$$

Point-based and zonal sensitivity indices

Definition (Point-based sensitivity indices)

Let $u \in \mathcal{D}$.
Output of interest : $Y(u)$

$$S_Z(u) = \frac{\mathrm{var}\left[E(Y(u) \mid Z)\right]}{\mathrm{var}\left[Y(u)\right]}$$

$$S_X(u) = \frac{\mathrm{var}\left[E(Y(u) \mid X)\right]}{\mathrm{var}\left[Y(u)\right]}$$

Point-based and zonal sensitivity indices

Definition (Point-based sensitivity indices)

Let $u \in \mathcal{D}$.

Output of interest : $Y(u)$

$$S_Z(u) = \frac{\text{var}\left[E(Y(u) \mid Z)\right]}{\text{var}\left[Y(u)\right]} = \frac{\sigma^2_{E_Z Y}}{\sigma^2_Y}$$

$$S_X(u) = \frac{\text{var}\left[E(Y(u) \mid X)\right]}{\text{var}\left[Y(u)\right]}$$

Point-based and zonal sensitivity indices

Definition (Point-based sensitivity indices)

Let $u \in \mathcal{D}$.
Output of interest : $Y(u)$

$$S_Z(u) = \frac{\text{var}\left[E(Y(u) \mid Z)\right]}{\text{var}\left[Y(u)\right]} = \frac{\sigma^2_{E_Z Y}}{\sigma^2_Y}$$

$$S_X(u) = \frac{\text{var}\left[E(Y(u) \mid X)\right]}{\text{var}\left[Y(u)\right]} = \frac{\text{var}\left[E(Y(.) \mid X)\right]}{\sigma^2_Y}$$

Point-based and zonal sensitivity indices

Definition (Point-based sensitivity indices)

Let $u \in \mathcal{D}$.
Output of interest : $Y(u)$

$$S_Z(u) = \frac{\operatorname{var}[E(Y(u) \mid Z)]}{\operatorname{var}[Y(u)]} = \frac{\sigma^2_{E_Z Y}}{\sigma^2_Y} = S_Z$$

$$S_X(u) = \frac{\operatorname{var}[E(Y(u) \mid X)]}{\operatorname{var}[Y(u)]} = \frac{\operatorname{var}[E(Y(.) \mid X)]}{\sigma^2_Y} = S_X$$

Point-based and zonal sensitivity indices

Definition (Zonal sensitivity indices)

Let $\Omega \subset \mathcal{D}$.

Output of interest : $Y_\Omega = \frac{1}{\Omega} \int_{u \in \Omega} Y(u)\, du$

$$S_Z(\Omega) = \frac{\mathrm{var}\left[E(Y_\Omega \mid Z)\right]}{\mathrm{var}\left[Y_\Omega\right]}$$

$$S_X(\Omega) = \frac{\mathrm{var}\left[E(Y_\Omega \mid X)\right]}{\mathrm{var}\left[Y_\Omega\right]}$$

Point-based and zonal sensitivity indices

Definition (Zonal sensitivity indices)

Let $\Omega \subset \mathcal{D}$.

Output of interest : $Y_\Omega = \frac{1}{\Omega} \int\limits_{u \in \Omega} Y(u)\, du$

$$S_Z(\Omega) = \frac{\mathrm{var}\left[E(Y_\Omega \mid Z)\right]}{\mathrm{var}\left[Y_\Omega\right]} = \frac{\sigma^2_{E_Z Y}(\Omega)}{\sigma^2_Y(\Omega)}$$

$$S_X(\Omega) = \frac{\mathrm{var}\left[E(Y_\Omega \mid X)\right]}{\mathrm{var}\left[Y_\Omega\right]}$$

Point-based and zonal sensitivity indices

Definition (Zonal sensitivity indices)

Let $\Omega \subset \mathcal{D}$.

Output of interest : $Y_\Omega = \frac{1}{\Omega} \int_{u \in \Omega} Y(u)\, du$

$$S_Z(\Omega) = \frac{\text{var}\left[E(Y_\Omega \mid Z)\right]}{\text{var}\left[Y_\Omega\right]} = \frac{\sigma^2_{E_Z Y}(\Omega)}{\sigma^2_Y(\Omega)}$$

$$S_X(\Omega) = \frac{\text{var}\left[E(Y_\Omega \mid X)\right]}{\text{var}\left[Y_\Omega\right]} = \frac{\text{var}\left[E(Y(.) \mid X)\right]}{\sigma^2_Y(\Omega)}$$

Point-based and zonal sensitivity indices

Property

Using previous notations :

$$\frac{S_Z(\Omega)}{S_X(\Omega)} = \frac{S_Z}{S_X} \cdot \frac{\sigma^2_{E_Z Y}(\Omega)}{\sigma^2_{E_Z Y}}$$

Properties of zonal sensitivity indices

Properties of zonal sensitivity indices

- Zonal sensitivity indices depend on $\sigma^2_{E_Z Y}(\Omega)$

Properties of zonal sensitivity indices

- Zonal sensitivity indices depend on $\sigma^2_{E_Z Y}(\Omega)$
- $\sigma^2_{E_Z Y}(\Omega)$ depends on the covariance function of RF $E_Z Y(u)$

Properties of zonal sensitivity indices

- Zonal sensitivity indices depend on $\sigma^2_{E_Z Y}(\Omega)$
- $\sigma^2_{E_Z Y}(\Omega)$ depends on the covariance function of RF $E_Z Y(u)$
- We have $E_Z Y = \bar{\psi}(Z)$

Properties of zonal sensitivity indices

- Zonal sensitivity indices depend on $\sigma^2_{E_Z Y}(\Omega)$
- $\sigma^2_{E_Z Y}(\Omega)$ depends on the covariance function of RF $E_Z Y(u)$
- We have $E_Z Y = \bar{\psi}(Z)$
- Condition on $\bar{\psi}$:

$$\int_{-\infty}^{\infty} \bar{\psi}^2(z) \cdot n(z)\, dz < \infty$$

where $n(.)$ is the $\mathcal{N}(0,1)$ pdf.

Hermite polynomials

Definition (Sequence of normalized Hermite polynomials)

Hermite polynomials

Definition (Sequence of normalized Hermite polynomials)

Consider sequence $(\chi_k)_{k\in\mathbb{N}}$ such that :

$$\forall k \in \mathbb{N}, \quad \chi_k(z) = \frac{1}{\sqrt{k!}} \cdot \frac{1}{n(z)} \cdot \frac{\partial^k}{\partial z^k} n(z)$$

Hermite polynomials

Definition (Sequence of normalized Hermite polynomials)

Consider sequence $(\chi_k)_{k\in\mathbb{N}}$ such that :

$$\forall k \in \mathbb{N}, \quad \chi_k(z) = \frac{1}{\sqrt{k!}} \cdot \frac{1}{n(z)} \cdot \frac{\partial^k}{\partial z^k} n(z)$$

Property

The sequence $(\chi_k)_{k\in\mathbb{N}}$ forms an orthonormal basis of Hilbert space $L^2(\mathcal{N})$:

$$L^2(\mathcal{N}) = \left\{ f : \mathbb{R} \to \mathbb{R} \quad tq. \int_{-\infty}^{\infty} f^2(z) \cdot n(z)\, dz < \infty \right\}$$

Properties of random field $E_Z Y(u)$

Properties of random field $E_Z Y(u)$

Property (Hermite expansion)

$\bar{\psi} \in L^2(\mathcal{N})$ *and* $E_Z Y = \bar{\psi}(Z)$, *thus :*

Properties of random field $E_Z Y(u)$

Property (Hermite expansion)

$\bar{\psi} \in L^2(\mathcal{N})$ and $E_Z Y = \bar{\psi}(Z)$, thus :

$$\bar{\psi} = \sum_{k=0}^{\infty} \lambda_k \cdot \chi_k \qquad \text{with } \lambda_k = E\left[\chi_k \cdot \bar{\psi}\right]$$

Properties of random field $E_Z Y(u)$

Property (Hermite expansion)

$\bar{\psi} \in L^2(\mathcal{N})$ *and* $E_Z Y = \bar{\psi}(Z)$, *thus :*

$$\bar{\psi} = \sum_{k=0}^{\infty} \lambda_k \cdot \chi_k \qquad \text{with } \lambda_k = E\left[\chi_k \cdot \bar{\psi}\right]$$

$$E_Z Y(u) = \sum_{k=0}^{\infty} \lambda_k \cdot \chi_k \left[Z(u)\right] \qquad \forall u \in \mathcal{D}$$

Properties of random field $E_Z Y(u)$

Property (Hermite expansion)

$\bar{\psi} \in L^2(\mathcal{N})$ *and* $E_Z Y = \bar{\psi}(Z)$, *thus :*

$$\bar{\psi} = \sum_{k=0}^{\infty} \lambda_k \cdot \chi_k \qquad \text{with } \lambda_k = E\left[\chi_k \cdot \bar{\psi}\right]$$

$$E_Z Y(u) = \sum_{k=0}^{\infty} \lambda_k \cdot \chi_k\left[Z(u)\right] \qquad \forall u \in \mathcal{D}$$

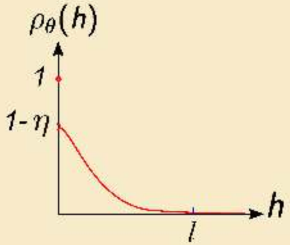
Property (Covariance function)

$$\forall u \in \mathcal{D}, \quad Cov\left[E_Z Y(u), E_Z Y(u+h)\right] = \sum_{k=0}^{\infty} \lambda_k^2 \cdot \sigma_Z^{2k} \cdot \rho_0^k(h)$$

Influence of range l

Property

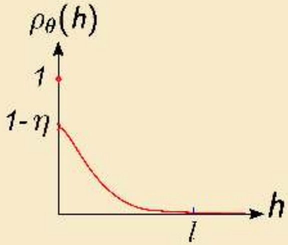
- Assume : $\forall h > 0, \quad \frac{\partial \rho_\theta}{\partial l}(h) > 0$



Influence of range l

Property

- Assume : $\forall h > 0, \quad \dfrac{\partial \rho_\theta}{\partial l}(h) > 0$

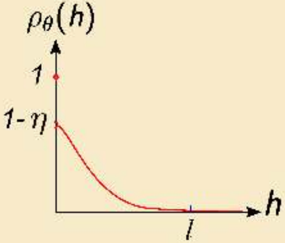


Influence of range l

Property

- Assume : $\forall h > 0, \quad \frac{\partial \rho_\theta}{\partial l}(h) > 0$
- then $\frac{\partial \sigma^2_{E_Z Y}(\Omega)}{\partial l} > 0$, thus :

$$\frac{\partial}{\partial l}\left[\frac{S_Z(\Omega)}{S_X(\Omega)}\right] > 0$$

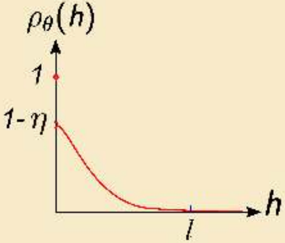


Influence of range l

Property

- Assume : $\forall h > 0, \quad \frac{\partial \rho_\theta}{\partial l}(h) > 0$
- then $\frac{\partial \sigma^2_{E_Z Y}(\Omega)}{\partial l} > 0$, thus :

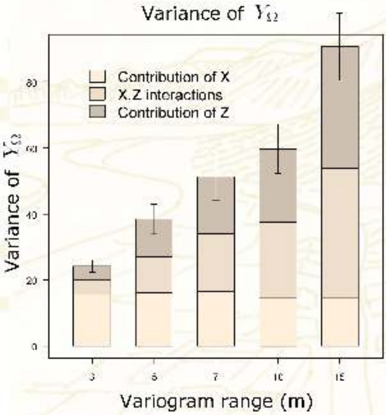
$$\frac{\partial}{\partial l}\left[\frac{S_Z(\Omega)}{S_X(\Omega)}\right] > 0$$

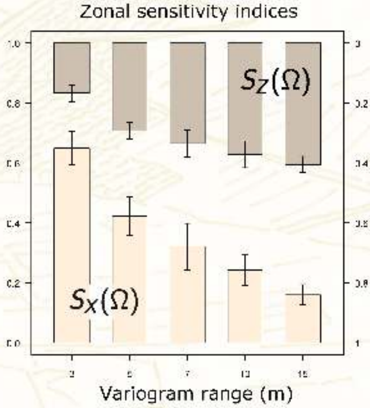


- Short range $l \rightarrow$ averaging of local errors $\rightarrow$ low $S_Z(\Omega)$
- Long range $l \rightarrow$ no averaging of local errors $\rightarrow$ high $S_Z(\Omega)$

Influence of range l - numerical example

- exponential covariance $\rho(h)$
- support $\Omega = \mathcal{D}$





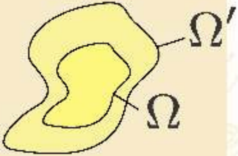
TETIS

Change of support

Change of support

Property

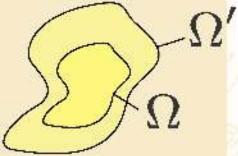
- Let $\Omega \subset \mathcal{D}$ and Ω' homothetic transformation of Ω of center O and ratio $\tau > 1$.



Change of support

Property

- Let $\Omega \subset \mathcal{D}$ and Ω' homothetic transformation of Ω of center O and ratio $\tau > 1$.
- Assume that : $\forall h > 0, \quad \frac{\partial \rho_0}{\partial h}(h) < 0$

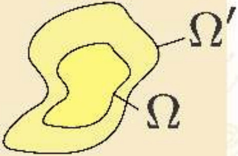


Change of support

Property

- Let $\Omega \subset \mathcal{D}$ and Ω' homothetic transformation of Ω of center O and ratio $\tau > 1$.
- Assume that : $\forall h > 0, \quad \dfrac{\partial \rho_0}{\partial h}(h) < 0$
- Then $\sigma^2_{E_{\angle}Y}(\Omega') < \sigma^2_{E_{\angle}Y}(\Omega)$, thus :

$$\frac{S_Z(\Omega')}{S_X(\Omega')} < \frac{S_Z(\Omega)}{S_X(\Omega)}$$

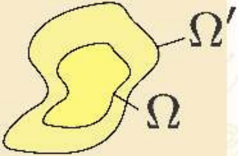


Change of support

Property

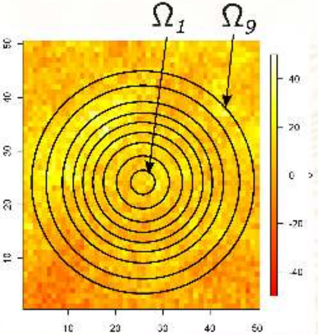
- Let $\Omega \subset \mathcal{D}$ and Ω' homothetic transformation of Ω of center O and ratio $\tau > 1$.
- Assume that : $\forall h > 0, \quad \frac{\partial \rho_0}{\partial h}(h) < 0$
- Then $\sigma^2_{E_Z Y}(\Omega') < \sigma^2_{E_Z Y}(\Omega)$, thus :

$$\frac{S_Z(\Omega')}{S_X(\Omega')} < \frac{S_Z(\Omega)}{S_X(\Omega)}$$

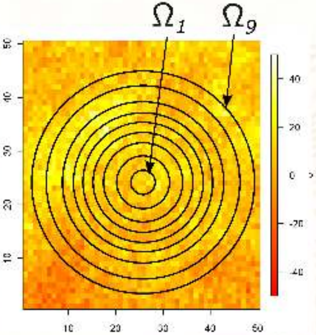


Larger zone $\Omega \rightarrow$ averaging of local errors $\rightarrow$ lower $S_Z(\Omega)$

Change of support - numerical example

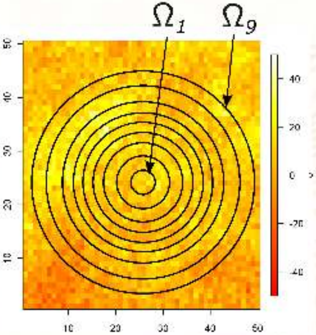


Change of support - numerical example



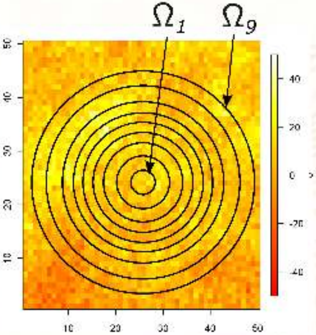
- Exponential covariance $\rho(h)$

Change of support - numerical example



- Exponential covariance $\rho(h)$
- Zones Ω_1 to Ω_9 of increasing size

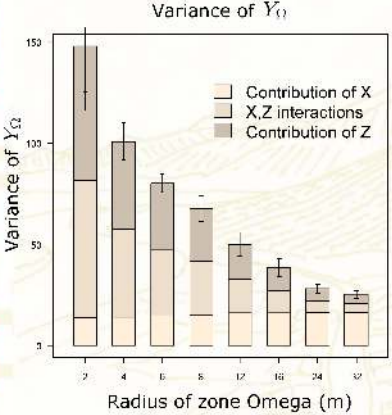
Change of support - numerical example

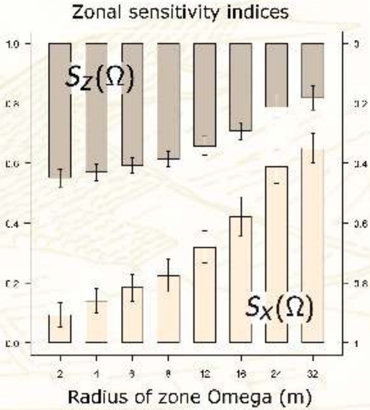


- Exponential covariance $\rho(h)$
- Zones Ω_1 to Ω_9 of increasing size
- Outputs of interest Y_{Ω_1} to Y_{Ω_9} :

$$Y_{\Omega_i} = \frac{1}{\Omega_i} \int_{u \in \Omega_i} Y(u)\, du$$

Change of support - numerical example

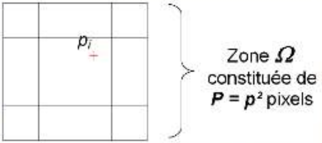




Approximation of zonal indices on a grid

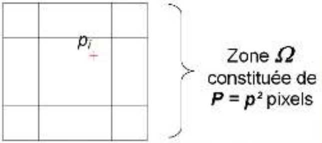
Approximation of zonal indices on a grid

- Zone Ω represented on a linear grid



Approximation of zonal indices on a grid

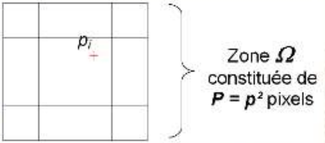
- Zone Ω represented on a linear grid



- Approximation of output of interest : $\tilde{Y}_\Omega = \frac{1}{P}\sum_{i=1}^{P} Y(p_i)$

Approximation of zonal indices on a grid

- Zone Ω represented on a linear grid



- Approximation of output of interest : $\tilde{Y}_\Omega = \frac{1}{P} \sum_{i=1}^{P} Y(p_i)$
- Approximation of zonal sensitivity indices : $\tilde{S}_X^\Omega = S_X(\tilde{Y}_\Omega)$
- Bias ε (for ψ linear) :

$$\varepsilon \underset{p \to \infty}{\sim} \frac{K}{p}$$

Other properties (ongoing work)

Other properties (ongoing work)

- Fields of point-based sensitivity indices
  - linking vocabulary of geostatistics and sensitivity analysis
  - properties of $S_X(u)$ and $S_Z(u)$?

Other properties (ongoing work)

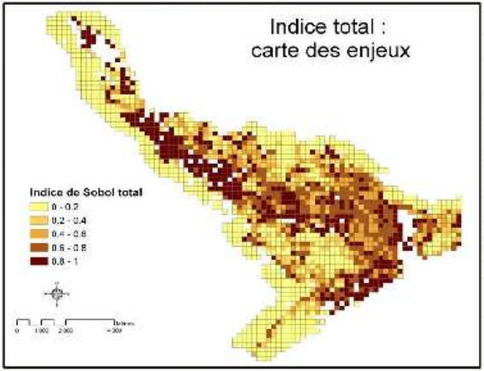
- Fields of point-based sensitivity indices
  - linking vocabulary of geostatistics and sensitivity analysis
  - properties of $S_X(u)$ and $S_Z(u)$?
- Estimation of zonal sensitivity indices
  - various techniques for spatial inputs
  - sampling random field $Z(u)$?

1. Motivations and notations
2. Point-based and zonal sensitivity indices
3. Some properties of zonal sensitivity indices
4. Conclusion

First answers to the modeler

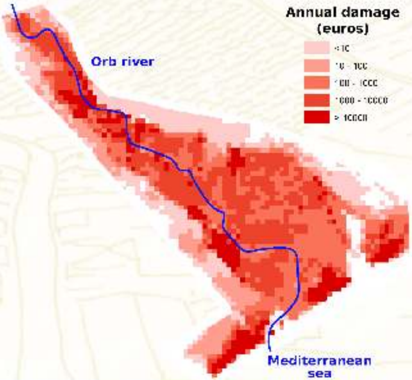
First answers to the modeler

1. Maps of point-based sensitivity indices



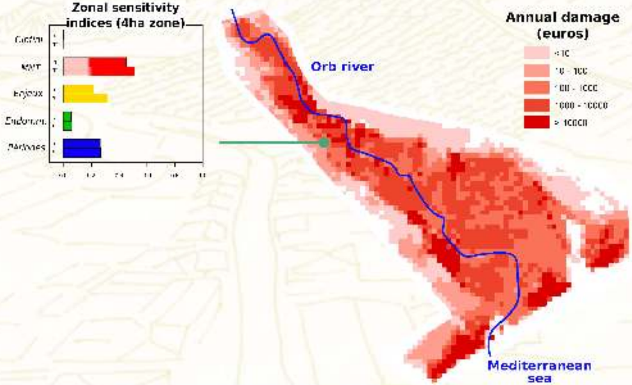
First answers to the modeler

1. Maps of point-based sensitivity indices
2. Zonal sensitivity indices depend on the chosen spatial support Ω



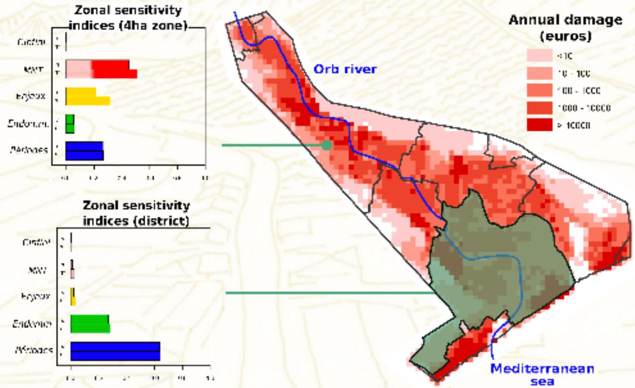
First answers to the modeler

1. Maps of point-based sensitivity indices
2. Zonal sensitivity indices depend on the chosen spatial support Ω
   - Individual stake : most important model input = water level



First answers to the modeler

1. Maps of point-based sensitivity indices
2. Zonal sensitivity indices depend on the chosen spatial support Ω
   - Individual stake : most important model input = water level
   - District : most important model input = flood return period



Variance-based spatial sensitivity analysis ?

Variance-based spatial sensitivity analysis ?

1. Need for UA/SA tools in environmental modeling
   - Spatial uncertainties are everywhere
   - Models used for decision-making

Variance-based spatial sensitivity analysis ?

1. Need for UA/SA tools in environmental modeling
   - Spatial uncertainties are everywhere
   - Models used for decision-making

2. Results
   - Ranking of model inputs at different spatial scale
   - Interactions between model inputs
   - Spatial structure of sensitivities, change of support. . .

Variance-based spatial sensitivity analysis ?

1. Need for UA/SA tools in environmental modeling
   - Spatial uncertainties are everywhere
   - Models used for decision-making
2. Results
   - Ranking of model inputs at different spatial scale
   - Interactions between model inputs
   - Spatial structure of sensitivities, change of support. . .
3. Limits
   - High CPU cost
   - « *Point-based* » model only

References

Grelot, F. ; Bailly, J.-S. ; Blanc, C. ; Erdlenbruch, K. ; Meriaux, P. ; Saint-Geours, N. ; Tourment, R.
Sensibilité d'une analyse coût-bénéfice - Enseignements pour l'évaluation des projets d'atténuation des inondations
Ingénieries Eau-Agriculture-Territoires, Numéro spécial inondations :95-108, 2009.

Saint-Geours, N. ; Lavergne, C. ; Bailly, J.-S. ; Grelot, F.
Analyse de sensibilité de Sobol d'un modèle spatialisé pour l'évaluation économique du risque d'inondation
Journal de la Société Française de Statistique, 2011, 152, 24-46.

Saint-Geours, N. ; Lavergne, C. ; Bailly, J.-S. ; Grelot, F.
Is there room to optimise the use of geostatistical simulations for sensitivity analysis of spatially distributed models ?
Accuracy2010, Leicester, UK 20-23 July 2010

Merci pour votre attention

Indices de sensibilité (Sobol) (1/5)

Soit le modèle $Y = f(X_1, \ldots, X_p)$ avec $X_i \in \mathbb{R}$ et $X_i \perp X_j$.

Indices de sensibilité (Sobol) (1/5)

Soit le modèle $Y = f(X_1, \ldots, X_p)$ avec $X_i \in \mathbb{R}$ et $X_i \perp X_j$.

Si f est de carré intégrable, elle peut se décomposer (Hoeffding, 1948) en :

$$f(x_1, \ldots, x_p) = f_0 + \sum_{i=1}^{p} f_i(x_i) + \sum_{1 \leq i < j \leq p} f_{ij}(x_i, x_j) + \ldots + f_{1 \ldots p}(x_1, \ldots, x_p)$$

Indices de sensibilité (Sobol) (2/5)

La décomposition est unique sous la condition :

$$\forall u \subset \{1, \ldots, p\}, \quad \forall i \in u, \quad E_{X_i}[\, f_u \,] = 0$$

Indices de sensibilité (Sobol) (2/5)

La décomposition est unique sous la condition :

$$\forall u \subset \{1, \ldots, p\}, \quad \forall i \in u, \quad E_{X_i}[\, f_u \,] = 0$$

On a alors $f_u \perp f_v$ et :

$$\begin{aligned} f_0 &= E[Y] \\ f_i(X_i) &= E[Y|X_i] - f_0 \\ f_{ij}(X_i, X_j) &= E[Y|X_i, X_j] - f_i - f_j - f_0 \end{aligned} \qquad (1)$$

Indices de sensibilité (Sobol) (2/5)

La décomposition est unique sous la condition :

$$\forall u \subset \{1, \ldots, p\}, \quad \forall i \in u, \quad E_{X_i}[\, f_u \,] = 0$$

On a alors $f_u \perp f_v$ et :

$$\begin{aligned} f_0 &= E[Y] \\ f_i(X_i) &= E[Y \,|\, X_i] - f_0 \\ f_{ij}(X_i, X_j) &= E[Y \,|\, X_i, X_j] - f_i - f_j - f_0 \end{aligned} \tag{1}$$

La variance de Y se décompose alors en :

$$V(Y) = \sum_{i=1}^{p} V(f_i) + \sum_{1 \leq i < j \leq p} V(f_{ij}) + \ldots + V(f_{1 \ldots p}) \tag{2}$$

Indices de sensibilité (Sobol) (3/5)

Pour $u \subset \{1, \ldots, p\}$, on définit

l'indice de sensibilité S_u du groupe de variables $(X_i)_{i \in u}$:

Indices de sensibilité (Sobol) (3/5)

Pour $u \subset \{1, \ldots, p\}$, on définit

l'indice de sensibilité S_u du groupe de variables $(X_i)_{i \in u}$:

$$S_u = \frac{\text{var}(f_u)}{\text{var}(Y)} \quad \forall u \subset \{1, \ldots, p\}$$

Indices de sensibilité (Sobol) (4/5)

Pour un modèle à p variables d'entrées indépendantes :

$$1 = S_1 + \cdots + S_d + \cdots + S_{1,\cdots,d}$$

Indices de sensibilité (Sobol) (4/5)

Pour un modèle à p variables d'entrées indépendantes :

$$1 = S_1 + \cdots + S_d + \cdots + S_{1,\cdots,d}$$

plus l'indice est grand, plus la variable ou le groupe de variables est important vis à vis de la variance de Y

Indices de sensibilité (Sobol) (5/5)

Indices de sensibilité (Sobol) (5/5)

Indice de sensibilité de premier ordre

$$S_i = \frac{V(f_i)}{V(Y)} = \frac{\mathrm{Var}\left[E\left(Y \mid X_i\right)\right]}{\mathrm{Var}(Y)}$$

$\Rightarrow$ réduction espérée de la variance si l'on fixe X_1

Indices de sensibilité (Sobol) (5/5)

Indice de sensibilité de premier ordre

$$S_i = \frac{V(f_i)}{V(Y)} = \frac{\text{Var}\left[E\left(Y \mid X_i\right)\right]}{\text{Var}(Y)}$$

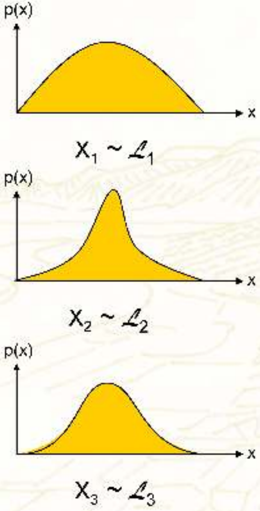
$\Rightarrow$ réduction espérée de la variance si l'on fixe X_1

Indice de sensibilité total

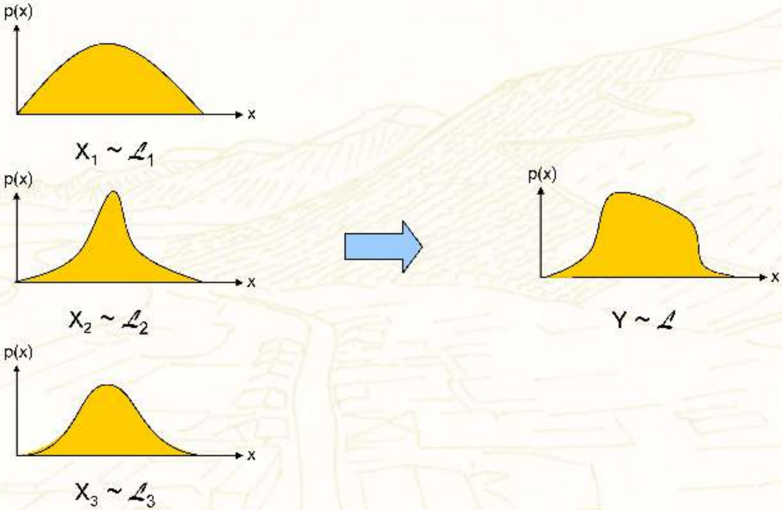
$$ST_i = S_i + S_{i,1} + S_{i,2} + \cdots + S_{1,\ldots,d}$$

$\Rightarrow$ variance résiduelle de Y lorsque tous les X_j sauf X_i sont fixés

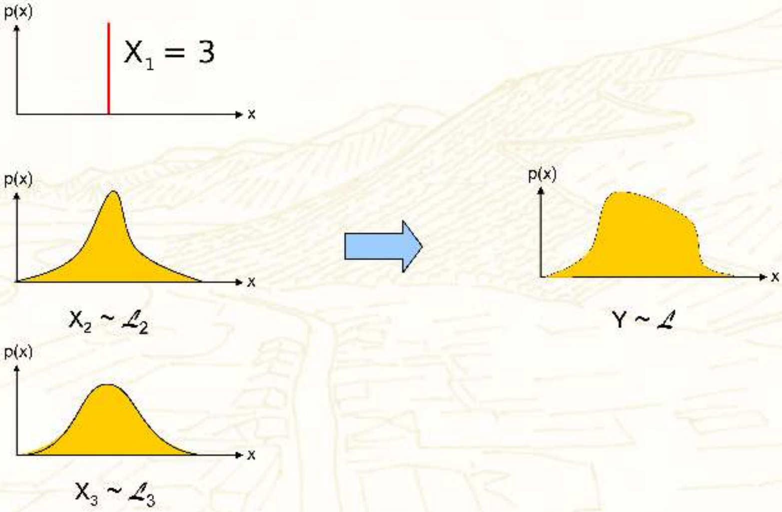
Indices de sensibilité : intuition



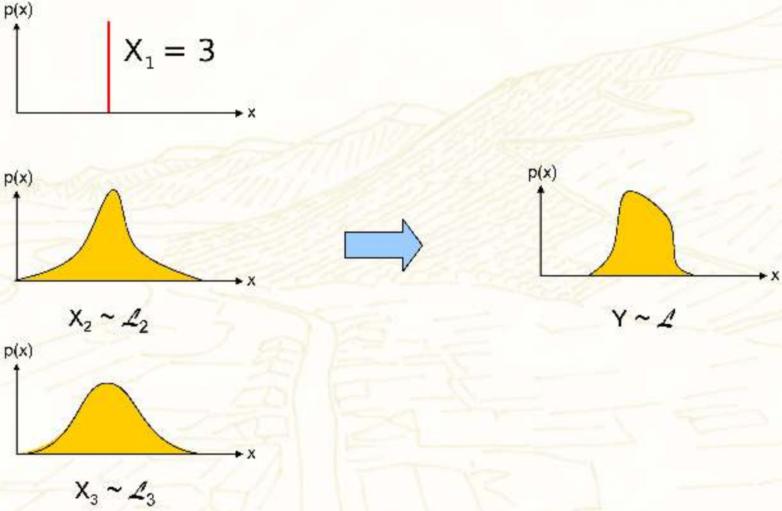
Indices de sensibilité : intuition



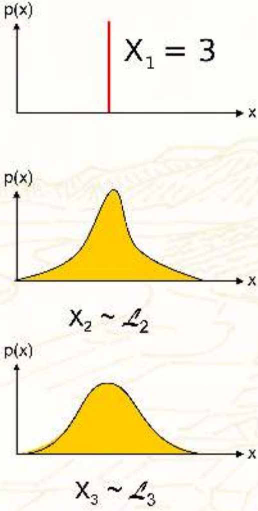
Indices de sensibilité : intuition



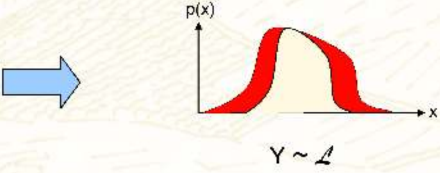
Indices de sensibilité : intuition



Indices de sensibilité : intuition

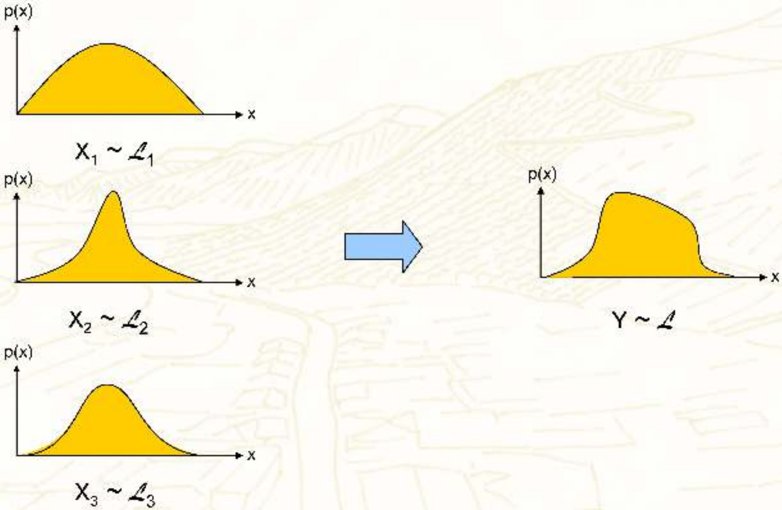


Indice de sensibilité
de premier ordre S_1

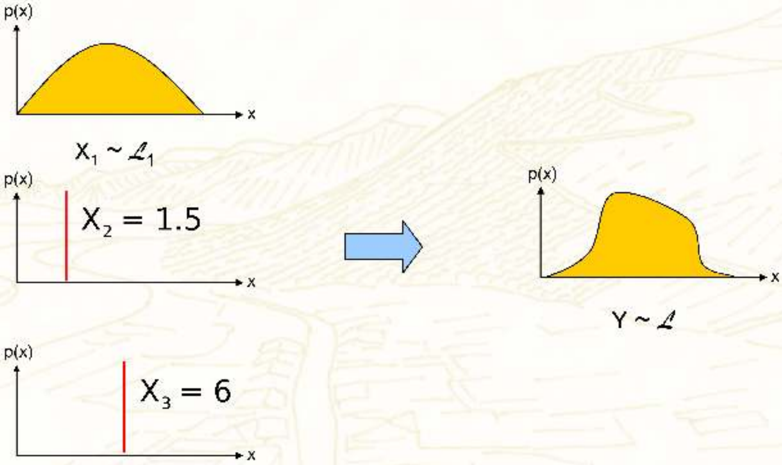


Réduction espérée de la variance de Y lorsque X1 est fixé

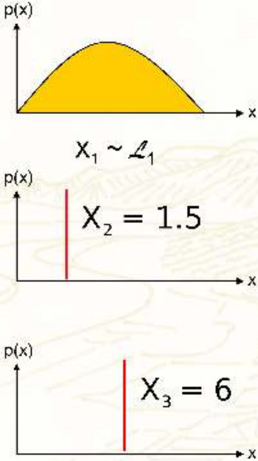
Indices de sensibilité : intuition



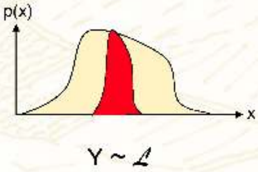
Indices de sensibilité : intuition



Indices de sensibilité : intuition



Indice de sensibilité
total ST_1

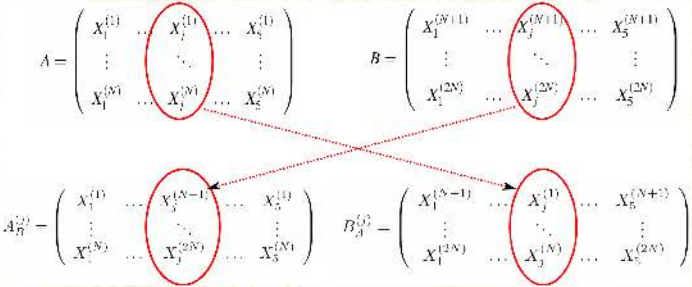


Variance résiduelle de Y lorsque X2 et X3 sont fixés

Indices de sensibilité : estimation

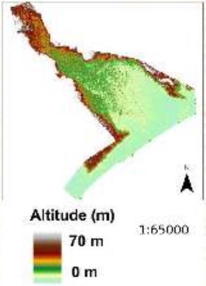
Indices de sensibilité : estimation

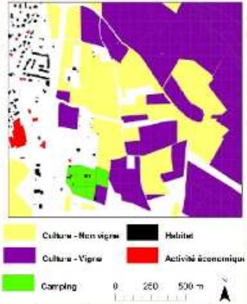
- estimation de type Monte-Carlo proposée par Sobol (1993)
- taille de l'échantillon nécessaire : $N \sim 1000 \cdot p$



Estimation de l'indice de sensibilité S_Z

- Rappel : indices de sensibilité estimés par une procédure de type Monte-Carlo (Sobol)

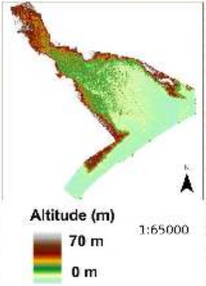


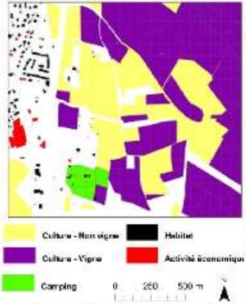


Estimation de l'indice de sensibilité S_Z

- Rappel : indices de sensibilité estimés par une procédure de type Monte-Carlo (Sobol)

- Approche 1 : considérer Z comme un groupe de variables

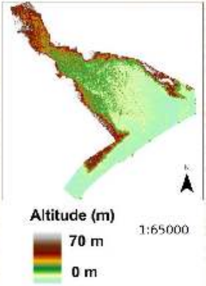


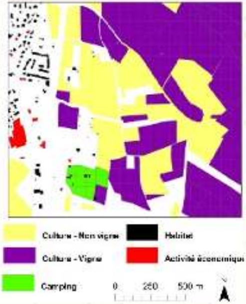


Estimation de l'indice de sensibilité S_Z

- Rappel : indices de sensibilité estimés par une procédure de type Monte-Carlo (Sobol)

- Approche 1 : considérer Z comme un groupe de variables
- Approche 2 : variable fonctionnelle, « *non contrôlable* »





Estimation de l'indice de sensibilité S_Z

Approche 1 : considérer Z comme un groupe de variables :

$$Z = (Z_1, \ldots, Z_M)$$

Estimation de l'indice de sensibilité S_Z

Approche 1 : considérer Z comme un groupe de variables :

$$Z = (Z_1, \dots, Z_M)$$

- méthode brute : Z_i valeur du champ en un pixel, variables non indépendantes (Jacques et al.) $\rightarrow M$ très grand

Estimation de l'indice de sensibilité S_Z

Approche 1 : considérer Z comme un groupe de variables :

$$Z = (Z_1, \dots, Z_M)$$

- méthode brute : Z_i valeur du champ en un pixel, variables non indépendantes (Jacques et al.) $\rightarrow$ M très grand
- Z_i valeurs en différentes zones, indépendantes (Volkova et al.)

Estimation de l'indice de sensibilité S_Z

Approche 1 : considérer Z comme un groupe de variables :

$$Z = (Z_1, \dots, Z_M)$$

- méthode brute : Z_i valeur du champ en un pixel, variables non indépendantes (Jacques et al.) $\rightarrow$ M très grand
- Z_i valeurs en différentes zones, indépendantes (Volkova et al.)
- projection sur une base : Z_i indépendantes (Busby et al.)

Estimation de l'indice de sensibilité S_Z

Approche 2 : variable fonctionnelle, « *non contrôlable* »

Estimation de l'indice de sensibilité S_Z

Approche 2 : variable fonctionnelle, « *non contrôlable* »

- variable « *interrupteur* » (Crosetto et al., 2001)

Estimation de l'indice de sensibilité S_Z

Approche 2 : variable fonctionnelle, « *non contrôlable* »

- variable « *interrupteur* » (Crosetto et al., 2001)
- méthode par scénario (Ruffo et al., 2004)

Estimation de l'indice de sensibilité S_Z

Approche 2 : variable fonctionnelle, « *non contrôlable* »

- variable « *interrupteur* » (Crosetto et al., 2001)
- méthode par scénario (Ruffo et al., 2004)
- jeu de n cartes équiprobables (Lilburne & Tarantola, 2009)

Estimation de l'indice de sensibilité S_Z

Approche 2 : variable fonctionnelle, « *non contrôlable* »

- variable « *interrupteur* » (Crosetto et al., 2001)
- méthode par scénario (Ruffo et al., 2004)
- jeu de n cartes équiprobables (Lilburne & Tarantola, 2009)
- méta-modèles joints (Iooss & Ribatet, 2009)

Indices de sensibilité généralisés (Lamboni, 2010)

Indices de sensibilité généralisés (Lamboni, 2010)

- Analyse en Composantes Principales sur Y multivarié
  - P variables : $Y(u_1), \ldots, Y(u_P)$ (P pixels)
  - N individus : les N runs du modèle
  - K composantes $Y_k = (Y_{k,1}, \ldots, Y_{k,N})$ de poids p_k

Indices de sensibilité généralisés (Lamboni, 2010)

- Analyse en Composantes Principales sur Y multivarié
  - P variables : $Y(u_1), \ldots, Y(u_P)$ (P pixels)
  - N individus : les N runs du modèle
  - K composantes $Y_k = (Y_{k,1}, \ldots, Y_{k,N})$ de poids p_k

- Indices de sensibilité $S_i^{(k)} = S_i(Y_k)$

Indices de sensibilité généralisés (Lamboni, 2010)

- Analyse en Composantes Principales sur Y multivarié
  - P variables : $Y(u_1), \ldots, Y(u_P)$ (P pixels)
  - N individus : les N runs du modèle
  - K composantes $Y_k = (Y_{k,1}, \ldots, Y_{k,N})$ de poids p_k
- Indices de sensibilité $S_i^{(k)} = S_i(Y_k)$

Indices de sensibilité généralisés (Lamboni, 2010)

$$S_i^{(\text{gal})} = \sum p_k S_i^{(k)}$$

Changements d'échelle

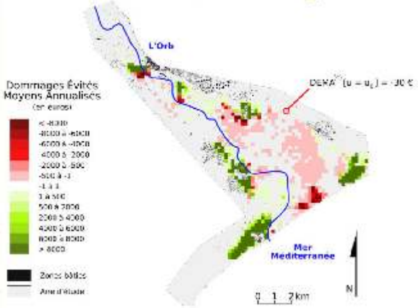
Le champ Z est d'autant moins influent que la zone Ω est grande.

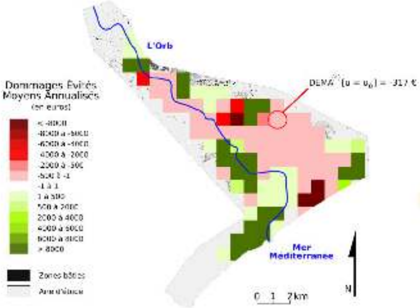
Changements d'échelle

Le champ Z est d'autant moins influent que la zone Ω est grande.

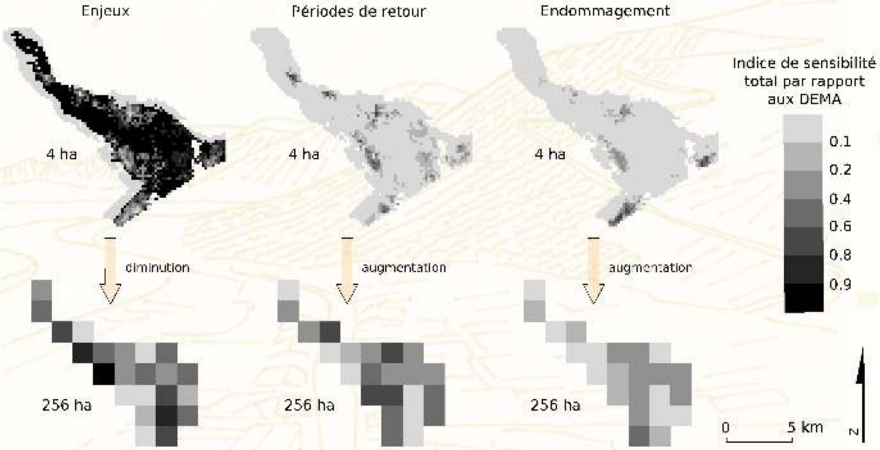
- Zone Ω grande
  $\rightarrow$ compensation des erreurs locales
  $\rightarrow$ faible influence de Z

Cartes des Dommages Evités à différentes échelles

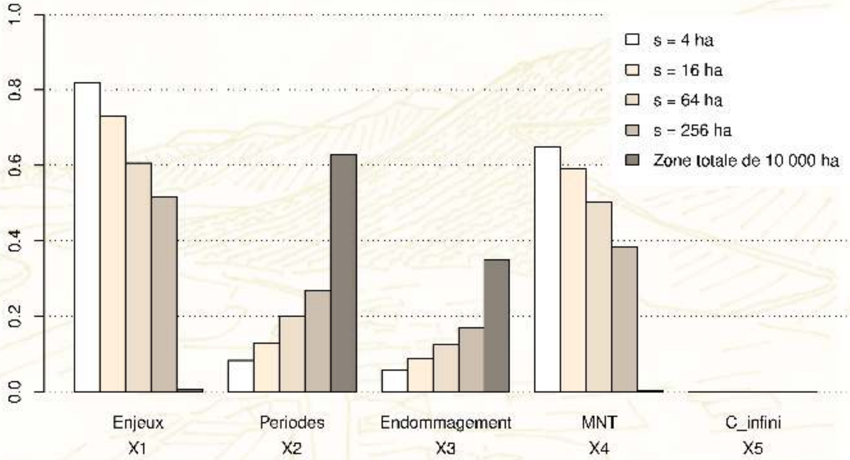




Cartes d'indices de Sobol à différentes échelles

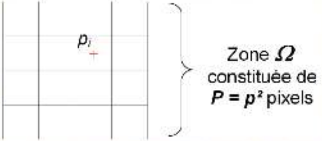


Indices de sensibilité moyens



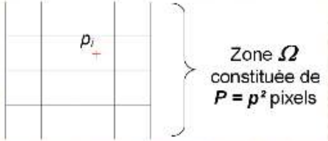
Influence de la discrétisation du champ Z

- Discrétisation de la zone Ω supposée carrée



Influence de la discrétisation du champ Z

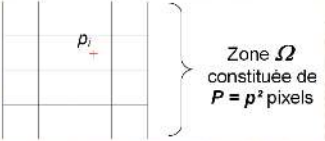
- Discrétisation de la zone Ω supposée carrée



- Dommage moyen approché : $\tilde{Y}_\Omega = \frac{1}{P} \sum_{i=1}^{P} Y(p_i)$

Influence de la discrétisation du champ Z

- Discrétisation de la zone Ω supposée carrée

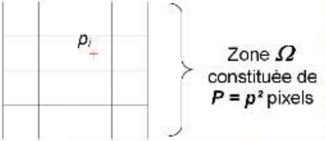


- Dommage moyen approché : $\tilde{Y}_\Omega = \frac{1}{P} \sum_{i=1}^{P} Y(p_i)$

- Hauteur d'eau moyenne approchée : $\tilde{Z}_\Omega = \frac{1}{P} \sum_{i=1}^{P} Z(p_i)$

Influence de la discrétisation du champ Z

- Discrétisation de la zone Ω supposée carrée



- Dommage moyen approché : $\tilde{Y}_\Omega = \frac{1}{P} \sum_{i=1}^{P} Y(p_i)$
- Hauteur d'eau moyenne approchée : $\tilde{Z}_\Omega = \frac{1}{P} \sum_{i=1}^{P} Z(p_i)$
- Indices de sensibilité approchés : $\tilde{S}_X^\Omega = S_X(\tilde{Y}_\Omega)$

Influence de la discrétisation du champ Z

Ecart entre indices approchés et valeurs exactes

- $\epsilon = \tilde{S}_X^\Omega - S_X^\Omega$
- ϵ linéaire en : $\frac{\eta\sigma_Z^2}{P} + \left[\bar{\gamma}(\nu,\nu) - \frac{1}{P^2}\sum_{i,j=1}^{P}\gamma^{-}(d_{p_i,p_j})\right]$
- $\lim_{p\to\infty} \epsilon = 0$, vitesse de convergence en $\frac{1}{p}$