Transformations and Invariance in Global Sensitivity Analysis

Elmar Plischke
Institute of Disposal Research
Clausthal University of Technology

SAMO Nice, 1.–4. July 2013
Joint Work with

Emanuele Borgonovo  Bocconi University, Milano, Italy
Stefano Tarantola  JRC Ispra, Italy
Max D. Morris  Iowa State University, USA
Contents

Motivation

Global Methods

A Generalized Framework for GSA

Monotonic Invariance

Conclusions
Motivation

- Power transformations, logarithmic transformations, rank transformations are widely used
Motivation

- Power transformations, logarithmic transformations, rank transformations are widely used
- Effective in improving numerical robustness
Motivation

- Power transformations, logarithmic transformations, rank transformations are widely used
- Effective in improving numerical robustness
- Problem of interpreting the results obtained on the transformed data: transfer back to the scale of the original data
Motivation

- Power transformations, logarithmic transformations, rank transformations are widely used
- Effective in improving numerical robustness
- Problem of interpreting the results obtained on the transformed data: transfer back to the scale of the original data
- More severe: estimation of global sensitivity statistics used to identify the key drivers of uncertainty
Motivation

- Power transformations, logarithmic transformations, rank transformations are widely used
- Effective in improving numerical robustness
- Problem of interpreting the results obtained on the transformed data: transfer back to the scale of the original data
- More severe: estimation of global sensitivity statistics used to identify the key drivers of uncertainty
Motivation

- Power transformations, logarithmic transformations, rank transformations are widely used
- Effective in improving numerical robustness
- Problem of interpreting the results obtained on the transformed data: transfer back to the scale of the original data
- More severe: estimation of global sensitivity statistics used to identify the key drivers of uncertainty

*Transformations alter the input-factor model-response mapping*
Setup

Model output of interest $y$ computed through a complex code and is dependent on $k$ uncertain input factors $\mathbf{x}$,

$$g : \mathbf{x} \mapsto y, \quad \Omega_{\mathbf{x}} \subseteq \mathbb{R}^k \to \Omega_\mathit{y} \subseteq \mathbb{R},$$
Setup

Model output of interest $y$ computed through a complex code and is dependent on $k$ uncertain input factors $\mathbf{x}$,

$$g : \mathbf{x} \mapsto y, \quad \Omega_{\mathbf{x}} \subseteq \mathbb{R}^k \to \Omega_{y} \subseteq \mathbb{R},$$

Local methods
Properties around $\mathbf{x}^0 \in \Omega_{\mathbf{x}}$.
- Not responsive to uncertainty in the input factors
- Only a limited exploration of the input factor space
Global Methods: The Golden Standard

Probabilistic formulation
Input $\mathbf{x} = (x_1, \ldots, x_k) \sim \mathbf{X} = (X_1, \ldots, X_k)$, RVec on $(\Omega_{\mathbf{x}}, \mathcal{A}, \mathbb{P}_{\mathbf{x}})$
Output $Y = g(\mathbf{X})$ is RV on $(\Omega_Y, \mathcal{B}, \mathbb{P}_Y)$, $\mathbb{P}_Y(B) = \mathbb{P}_{\mathbf{x}}(g^{-1}[B])$

$$F_{\mathbf{x}}(\mathbf{x}) = \mathbb{P}_{\mathbf{x}}\left( \bigcap_{i=1}^{k} [X_i \leq x_i] \right) \text{ and } F_Y = \mathbb{P}_Y(Y \leq y) \text{ CDFs of } \mathbf{X} \text{ and } Y$$

$$f_{\mathbf{x}}(\mathbf{x}) \text{ and } f_Y(y) \text{ PDFs of } \mathbf{X} \text{ and } Y$$
Global Methods: The Golden Standard

Probabilistic formulation
Input $\mathbf{x} = (x_1, \ldots, x_k) \sim \mathbf{X} = (X_1, \ldots, X_k)$, RVec on $(\Omega_{\mathbf{x}}, \mathcal{A}, P_{\mathbf{x}})$
Output $Y = g(\mathbf{X})$ is RV on $(\Omega_Y, \mathcal{B}, P_Y)$, $P_Y(B) = P_{\mathbf{x}}(g^{-1}[B])$

$F_{\mathbf{x}}(\mathbf{x}) = P_{\mathbf{x}}(\bigcap_{i=1}^{k} [X_i \leq x_i])$ and $F_Y = P_Y(Y \leq y)$ CDFs of $\mathbf{X}$ and $Y$

$f_{\mathbf{x}}(\mathbf{x})$ and $f_Y(y)$ PDFs of $\mathbf{X}$ and $Y$

Input distributions are known!
Global Methods II: Categories

- Linear regression (also with rank and other transformations)
- Screening methods [5]
- Variance-based methods [8]
- Expected value of information-based methods [9]
- Distribution-based methods [2]

Variance-based methods have been the most widely studied both from the theoretical and numerical viewpoints: ANOVA decomposition
Variance-Based Method

Main effects and total effects

\[ \eta_Y(X_i) = \frac{\text{Var}[E[Y|X_i]]}{\text{Var}[Y]} \]

\[ \eta_T(X_i) = \frac{E[\text{Var}[Y|X_{\sim i}]]}{\text{Var}[Y]} \]

Here \( X_{\sim i} \): Vector \( X \) without \( i \)th component
Variance-Based Method

Main effects and total effects

\[ \eta_Y(X_i) = \frac{\text{Var}[E[Y|X_i]]}{\text{Var}[Y]}, \quad \eta_T(Y_i) = \frac{E[\text{Var}[Y|X_\sim_i]]}{\text{Var}[Y]} \]

Here \( X_\sim_i \): Vector \( X \) without \( i \)th component

Example

\[ Y = \exp(X_1 + 2 \cdot X_2) \quad \text{with} \quad X_1, X_2 \sim \mathcal{N}(0, 1) \quad iid. \]

\( \eta_Y(X_i) \) and \( \eta_T(Y_i) \) \((i = 1, 2)\) computed with [7]’s estimator varying the number of model evaluations \( M \)
Convergence Issues, Interpretation Problems

Main effects

Total effects

Raw Data  Log Trafo  Rank Trafo

Elmar Plischke
IELF, TU Clausthal
Issues and Problems II

- No convergence on the raw data: Analytically

\[ \eta^Y_1(X_1) = 0.012, \quad \eta^Y_1(X_2) = 0.364, \]
\[ \eta^Y_T(X_1) = 0.637, \quad \eta^Y_T(X_2) = 0.988. \]
Issues and Problems II

- No convergence on the raw data: Analytically

\[ \eta_Y(X_1) = 0.012, \quad \eta_Y(X_2) = 0.364, \]
\[ \eta_T(X_1) = 0.637, \quad \eta_T(X_2) = 0.988. \]

- Convergence on the log-transformed data
Issues and Problems II

- No convergence on the raw data: Analytically
  \[ \eta(Y(X_1)) = 0.012, \quad \eta(Y(X_2)) = 0.364, \]
  \[ \eta^T(Y(X_1)) = 0.637, \quad \eta^T(Y(X_2)) = 0.988. \]

- Convergence on the log-transformed data
- Convergence on the rank-scaled data (Copula-related)
Issues and Problems II

- No convergence on the raw data: Analytically
  \[ \eta_Y(X_1) = 0.012, \quad \eta_Y(X_2) = 0.364, \]
  \[ \eta_T(X_1) = 0.637, \quad \eta_T(X_2) = 0.988. \]

- Convergence on the log-transformed data
- Convergence on the rank-scaled data (Copula-related)
- Transformed results not related to analytical values
Issues and Problems II

- No convergence on the raw data: Analytically
  \[
  \eta_Y(X_1) = 0.012, \quad \eta_Y(X_2) = 0.364, \\
  \eta_T(X_1) = 0.637, \quad \eta_T(X_2) = 0.988.
  \]

- Convergence on the log-transformed data
- Convergence on the rank-scaled data (Copula-related)
- Transformed results not related to analytical values
Issues and Problems II

- No convergence on the raw data: Analytically

\[ \eta_Y(X_1) = 0.012, \quad \eta_Y(X_2) = 0.364, \]
\[ \eta_T(X_1) = 0.637, \quad \eta_T(X_2) = 0.988. \]

- Convergence on the log-transformed data
- Convergence on the rank-scaled data (Copula-related)
- Transformed results not related to analytical values

Transformations: Model structure changes from multiplicative to additive:
\[ \hat{\eta}^{\log(Y)}_1(X_i) = \hat{\eta}^{\log(Y)}_T(X_i) \]

Calculations do not readily translate back!
A Generalized Framework for GSA

A decision-maker’s state of knowledge is represented by the probability law $P_Y$ of the random variable of interest $Y$. 
A Generalized Framework for GSA

A decision-maker’s state of knowledge is represented by the probability law $P_Y$ of the random variable of interest $Y$.

**Sensitivity summarizes two states of knowledge**
Before and after knowing that $X_i = x_i$
More on this idea: [3].
A Generalized Framework

Definition (Sensitivity Measure)
Let $d(\cdot, \cdot)$ measure a shift (discrepancy, distance) between the unconditional probability $\mathbb{P}_Y$ and the probability $\mathbb{P}_{Y|X_i = x_i}$ conditional to a realization $X_i = x_i$.

The associated sensitivity measure is defined as an expected shift of the conditional probabilities,

$$
\gamma^Y_{d}(X_i) = \mathbb{E} \left[ d(\mathbb{P}_Y, \mathbb{P}_{Y|X_i}) \right] \quad (1)
$$
A Generalized Framework

Definition (Sensitivity Measure)

Let $d(\cdot, \cdot)$ measure a shift (discrepancy, distance) between the unconditional probability $P_Y$ and the probability $P_{Y|X_i=x_i}$ conditional to a realization $X_i = x_i$. The associated sensitivity measure is defined as an expected shift of the conditional probabilities,

$$
\gamma_d^Y(X_i) = \mathbb{E} \left[ d(P_Y, P_{Y|X_i}) \right]
$$

(1)

First order effect: $d(P_Y, P_{Y|X_i=x_i}) = \nabla[\gamma]^{-1} (\nabla[\gamma] - \nabla[Y|X_i = x_i])$
A Generalized Framework

Definition (Sensitivity Measure)
Let $d(\cdot, \cdot)$ measure a shift (discrepancy, distance) between the unconditional probability $\mathbb{P}_Y$ and the probability $\mathbb{P}_{Y|X_i=x_i}$ conditional to a realization $X_i = x_i$.

The associated sensitivity measure is defined as an expected shift of the conditional probabilities,

$$\gamma_d(X_i) = \mathbb{E}\left[ d(\mathbb{P}_Y, \mathbb{P}_{Y|X_i}) \right] \tag{1}$$

First order effect: $d(\mathbb{P}_Y, \mathbb{P}_{Y|X_i=x_i}) = \nabla[Y]^{-1} (\nabla[Y] - \nabla[Y|X_i=x_i])$

Approach allows for estimates: Analogously to Correlation Ratio estimation of first order effects
Monotonic Transformations

Consider a monotonic transformation of $y$,

$$u : \Omega_Y \to \Omega_U \subseteq \mathbb{R}, \ y \mapsto u(y) \text{ and } u \circ g : \Omega_X \to \mathbb{R}, \ x \mapsto u(g(x))$$
Monotonic Transformations

Consider a monotonic transformation of $y,$

$$u : \Omega_Y \to \Omega_U \subseteq \mathbb{R}, \ y \mapsto u(y)$$

and

$$u \circ g : \Omega_X \to \mathbb{R}, \ x \mapsto u(g(x))$$

When $x$ is uncertain, $u$ becomes a random variable (denoted by $U$) induced by $\mathbb{P}_X$ through the composition of $u$ with $g$.

$\gamma^U(X_i)$: sensitivity statistics of $X_i$ with respect to $U = u(Y)$. 
Monotonic Invariance

A sensitivity measure is monotonic invariant, if $\gamma^U(X_i) = \gamma^Y(X_i)$ for all suitable models $Y = g(X)$ and transformations $U = u: g(X)$
Monotonic Invariance

A sensitivity measure is monotonic invariant, if \( \gamma^U(X_i) = \gamma^Y(X_i) \) for all suitable models \( Y = g(X) \) and transformations \( U = u: g(X) \).

**Theorem**

A sensitivity measure \( \gamma_d \) is monotonic invariant if its shift \( d \) is monotonic invariant.

Decision Theory: Sensitivity measure independent under choice of utility function [1]
Invariant shifts

**Definition**
Let \( P \) and \( Q \) be probability measures on \((\Omega, \mathcal{A})\). Let \( h : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \) be a continuous and non-decreasing function with \( h(0) = 0 \) and \( \sup_t \frac{h(2t)}{h(t)} < \infty \) (Orlicz condition for generalized triangle inequality). Then, we define

\[
d(P, Q) = \sup_{A \in \mathcal{A}} h \left( |P(A) - Q(A)| \right)
\]

(2)

**Theorem**
Such a \( d(\cdot, \cdot) \) is monotonic invariant.
Invariant shifts: PDF based

Example
For \( h(t) = t \) and \( f \) and \( g \) PDFs of \( \mathbb{P} \) and \( \mathbb{Q} \): From (2)

\[
\sup_{A \in \mathcal{A}} |\mathbb{P}(A) - \mathbb{Q}(A)| = \sup_{A \in \mathcal{A}} \left| \int_A (f(y) - g(y)) dy \right|
\]

\[
= \frac{1}{2} \int_{\mathcal{Y}} |f(y) - g(y)| dy
\]

(Consider \( A = \{ y : f(y) \geq g(y) \} \))

Hence: Borgonovo importance measure is transformation invariant,

\[
\delta^Y(x_i) = \frac{1}{2} \int_{\mathcal{X}_i} f_{X_i}(x_i) \int_{\mathcal{Y}} |f_Y(y) - f_{Y|X_i=x_i}(y)| dy dx_i
\]
Monotonic invariant shifts: CDF based

Consider only half-rays \( A(y) = \{ z \leq y \} \) in (2):

\[
d(\mathbb{P}, \mathbb{Q}) = \sup_{y \in \mathbb{R}} h(|F(y) - G(y)|)
\]

(3)

where \( F(y) = \mathbb{P}(z \leq y) \) and \( G(y) = \mathbb{Q}(z \leq y) \) are the CDFs.

Transformation invariance: Birnbaum-Orlicz family of metrics

\[
h(t) = t: \text{Kolmogorov-Smirnov distance},
\]

\[
\beta^Y(X_i) = \int_{\mathcal{X}} f_{X_i}(x_i) \sup_{y \in \mathcal{Y}} \left| F_Y(y) - F_Y|_{X_i=x_i}(y) \right| \, dx_i
\]
Example revisited

\[ Y = \exp(X_1 + 2 \cdot X_2) \quad \text{with} \quad X_1, X_2 \sim \mathcal{N}(0, 1) \quad \text{iid}. \]
Example revisited

\[ Y = \exp(X_1 + 2 \cdot X_2) \quad \text{with} \quad X_1, X_2 \sim \mathcal{N}(0, 1) \quad \text{iid.} \]
Level E Geosphere Transport Model

Sensitivity of total dose over time: QMC sample, size 8192:

Main Effects

Main Effects on Log

Kolmogorov Smirnov Sensitivity

Borgonovo Sensitivity

Elmar Plischke
IELF, TU Clausthal

Invariance GSA
Conclusions

- Uncertainty is coded in the probability in general, not only in the variance: Need for stronger sensitivity measures
Conclusions

- Uncertainty is coded in the probability in general, not only in the variance: Need for stronger sensitivity measures
- General Framework: Estimators from given data are available
Conclusions

- Uncertainty is coded in the probability in general, not only in the variance: Need for stronger sensitivity measures
- General Framework: Estimators from given data are available
- Transformation invariance: no change of interpretation
Conclusions

- Uncertainty is coded in the probability in general, not only in the variance: Need for stronger sensitivity measures
- General Framework: Estimators from given data are available
- Transformation invariance: no change of interpretation
- Suitable domain for estimation: gain in numerical precision
Thank you!

Any questions?

More details: [4]
Estimation from given data: [6]

Financial support from German Ministry of Education and Research, Grant BMBF, 02S9082A is greatly acknowledged
References I

M. Baucells and E. Borgonovo.
Invariant probabilistic sensitivity analysis.
In Press.

E. Borgonovo.
A new uncertainty importance measure.

E. Borgonovo, G. B. Hazen, and E. Plischke.
2013.
In Preparation.

E. Borgonovo, S. Tarantola, E. Plischke, and M. D. Morris.
Transformations and invariance in the sensitivity analysis of computer experiments.
Submitted.

M. D. Morris.
Factorial sampling plans for preliminary computational experiments.
References II

E. Plischke, E. Borgonovo, and C. L. Smith.
Global sensitivity measures from given data.

Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index.

*Global Sensitivity Analysis – The Primer.*

M. Strong and J. E. Oakley.
An efficient method for computing single-parameter partial expected value of perfect information.
*Medical Decision Making*, 2013.
DOI:10.1177/0272989X12465123.