TU Clausthal

# 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

- 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 $\mathbf{x}$,

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

# Setup

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

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

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

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

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

# Global Methods: The Golden Standard

## Probabilistic formulation

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

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

$$\eta_1^Y(X_i) = \frac{\mathrm{Var}[\mathrm{E}[Y|X_i]]}{\mathrm{Var}[Y]}, \quad \eta_T^Y(X_i) = \frac{\mathrm{E}[\mathrm{Var}[Y|\mathbf{X}_{\sim i}]]}{\mathrm{Var}[Y]}$$

Here $\mathbf{X}_{\sim i}$: Vector $\mathbf{X}$ without *i*th component

# Variance-Based Method

Main effects and total effects

$$\eta_1^Y(X_i) = \frac{\mathrm{Var}[\mathrm{E}[Y|X_i]]}{\mathrm{Var}[Y]}, \quad \eta_T^Y(X_i) = \frac{\mathrm{E}[\mathrm{Var}[Y|\mathbf{X}_{\sim i}]]}{\mathrm{Var}[Y]}$$

Here $\mathbf{X}_{\sim i}$: Vector $\mathbf{X}$ without $i$th component

## Example

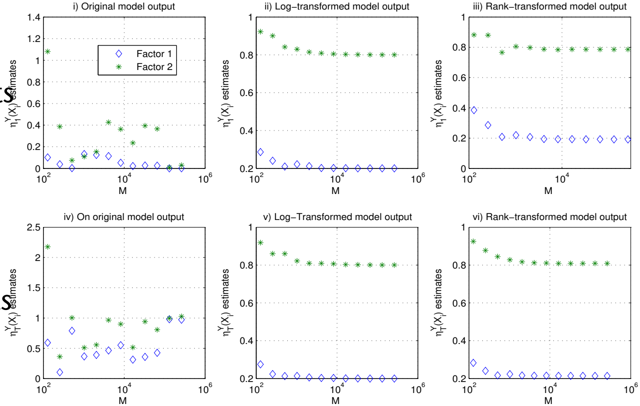
$$Y = \exp(X_1 + 2 \cdot X_2) \quad \text{with} \quad X_1, X_2 \sim \mathcal{N}(0,1) \quad iid.$$

$\eta_1^Y(X_i)$ and $\eta_T^Y(X_i)$ ($i = 1,2$) 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

# Issues and Problems II

- No convergence on the raw data: Analytically

$$\eta_1^Y(X_1) = 0.012, \quad \eta_1^Y(X_2) = 0.364,$$
$$\eta_T^Y(X_1) = 0.637, \quad \eta_T^Y(X_2) = 0.988.$$

# Issues and Problems II

- No convergence on the raw data: Analytically

$$\eta_1^Y(X_1) = 0.012, \quad \eta_1^Y(X_2) = 0.364,$$
$$\eta_T^Y(X_1) = 0.637, \quad \eta_T^Y(X_2) = 0.988.$$

- Convergence on the log-transformed data

# Issues and Problems II

- No convergence on the raw data: Analytically

$$\eta_1^Y(X_1) = 0.012, \quad \eta_1^Y(X_2) = 0.364,$$
$$\eta_T^Y(X_1) = 0.637, \quad \eta_T^Y(X_2) = 0.988.$$

- 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

$$\eta_1^Y(X_1) = 0.012, \quad \eta_1^Y(X_2) = 0.364,$$
$$\eta_T^Y(X_1) = 0.637, \quad \eta_T^Y(X_2) = 0.988.$$

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

# Issues and Problems II

- No convergence on the raw data: Analytically

$$\eta_1^Y(X_1) = 0.012, \quad \eta_1^Y(X_2) = 0.364,$$
$$\eta_T^Y(X_1) = 0.637, \quad \eta_T^Y(X_2) = 0.988.$$

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

# Issues and Problems II

- No convergence on the raw data: Analytically

$$\eta_1^Y(X_1) = 0.012, \quad \eta_1^Y(X_2) = 0.364,$$
$$\eta_T^Y(X_1) = 0.637, \quad \eta_T^Y(X_2) = 0.988.$$

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

Transformations: Model structure changes from multiplicative to additive: $\hat{\eta}_1^{\log(Y)}(X_i) = \hat{\eta}_T^{\log(Y)}(X_i)$
*Calculations do not readily translate back!*

# A Generalized Framework for GSA

A decision-maker's state of knowledge is represented by the probability law $\mathbb{P}_Y$ 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 $\mathbb{P}_Y$ 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,

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

# 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,

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

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

# 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,

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

First order effect: $d(\mathbb{P}_Y, \mathbb{P}_{Y|X_i=x_i}) = \mathbb{V}[Y]^{-1}\left(\mathbb{V}[Y] - \mathbb{V}[Y|X_i = x_i]\right)$
Approach allows for estimates: Analogously to Correlation Ratio estimation of first order effects

# Monotonic Transformations

Consider a monotonic transformation of $y$,

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

# Monotonic Transformations

Consider a monotonic transformation of $y$,

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

When $\mathbf{x}$ 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 $Y = g(\mathbf{X})$ and transformations $U = u \colon g(\mathbf{X})$

# Monotonic Invariance

A sensitivity measure is monotonic invariant, if $\gamma^U(X_i) = \gamma^Y(X_i)$ for all suitable models $Y = g(\mathbf{X})$ and transformations $U = u \colon g(\mathbf{X})$

**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 $\mathbb{P}$ and $\mathbb{Q}$ 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

$$d(\mathbb{P}, \mathbb{Q}) = \sup_{A \in \mathcal{A}} h\left(|\mathbb{P}(A) - \mathbb{Q}(A)|\right) \tag{2}$$

## 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,

$$\delta^Y(X_i) = \frac{1}{2} \int_{\mathcal{X}_\rangle} f_{X_i}(x_i) \int_{\mathcal{Y}} \left| f_Y(y) - f_{Y|X_i=x_i}(y) \right| dydx_i$$

# Monotonic invariant shifts: CDF based

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

$$d(\mathbb{P}, \mathbb{Q}) = \sup_{y \in \mathbb{R}} h(|F(y) - G(y)|) \qquad (3)$$

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

$h(t) = t$: Kolmogorov-Smirnov distance,

$\beta^Y(X_i) = \int_{\mathcal{X}} f_{X_i}(x_i) \sup_{y \in \mathcal{Y}} \left| F_Y(y) - F_{Y|X_i=x_i}(y) \right| dx_i$

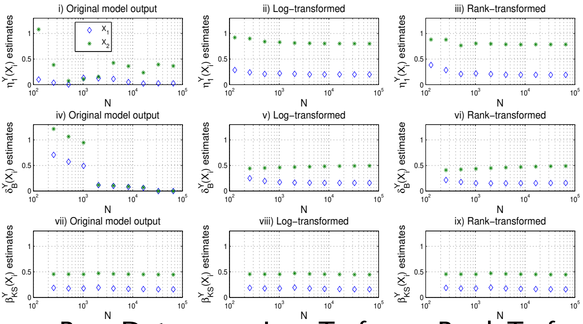
# Example revisited

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

Main effects

Borgonovo

Kolmogorov-Smirnov




Raw Data Log Trafo Rank Trafo

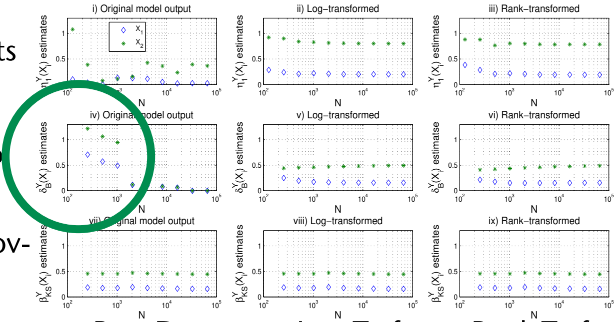
# Example revisited

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

Main effects

Borgonovo

Kolmogorov-Smirnov

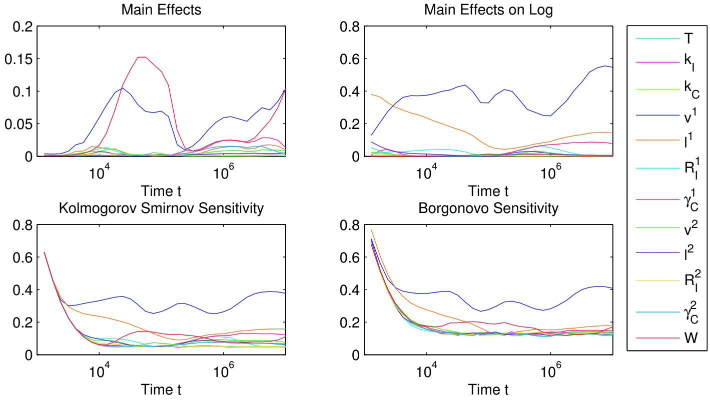



Raw Data | Log Trafo | Rank Trafo

Kernel density estimator fails!

# Level E Geosphere Transport Model

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



# 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.
*Management Science*, 2013.
In Press.

E. Borgonovo.
A new uncertainty importance measure.
*Reliability Engineering&System Safety*, 92(6):771–784, 2007.

E. Borgonovo, G. B. Hazen, and E. Plischke.
Probabilistic sensitivity measures: Foundations and estimation.
2013.
In Preparation.

E. Borgonovo, S. Tarantola, E. Plischke, and M. D. Morris.
Transformations and invariance in the sensitivity analysis of computer experiments.
*Journal of the Royal Statistical Society, Series B*, 2013.
Submitted.

M. D. Morris.
Factorial sampling plans for preliminary computational experiments.
*Technometrics*, 33(2):161–174, 1991.

# References II

E. Plischke, E. Borgonovo, and C. L. Smith.
Global sensitivity measures from given data.
*European Journal of Operational Research*, 226(3):536–550, 2013.

A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola.
Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index.
*Comput. Phys. Commun.*, 181(2):259–270, 2010.

A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola.
*Global Sensitivity Analysis – The Primer*.
John Wiley&Sons, Chichester, 2008.

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.