Estimating Global Sensitivity Measures: Torturing the Data Until They Confess

Elmar Plischke

Institut für Endlagerforschung
TU Clausthal
St. Étienne, MASCOT-NUM, April 10, 2015

Estimating Global Sensitivity Measures: Torturing the Data Until They ~~Confess~~ Prove Innocent

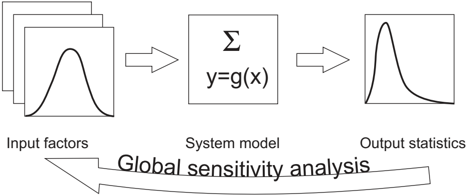
Elmar Plischke

Institut für Endlagerforschung
TU Clausthal
St. Étienne, MASCOT-NUM, April 10, 2015

Contents

Global Sensitivity Measures

[Saltelli et al., 2000]
"Sensitivity analysis is the study of how the variation on the output of a model [. . .] can be apportioned [. . .] to different sources of variation, and of how the given model depends upon the information fed into it."
Saltelli in [Faivre et al., 2013]
"Sensitivity analysis is foremost about exploring the space of input assumptions in such a way as to be able to map the inference to the assumptions in a transparent way."



Many Methods Available for Sensitivity Analysis

- Local Methods: Behaviour of the model about a nominal working/reference point
- Screening Methods: Behaviour of the system within given parameter ranges
- Global Methods: Behaviour under given input parameter distributions

Given Data Methodology

X	k-dimensional random vector
Y	random variable (quantity of interest for time series)

Given Data Methodology

X	k-dimensional random vector
Y	random variable (quantity of interest for time series)

- Physical observations
- Uncertainty propagation through model $Y = g(X)$

Given Data Methodology

X k-dimensional random vector
Y random variable (quantity of interest for time series)

- Physical observations
- Uncertainty propagation through model $Y = g(X)$
  - Simple random sampling of X
  - Latin Hypercube sampling of X
  - Quasi Monte Carlo sampling (Sobol' LPτ, ...) of X
  - But not fast multidimensional/spare grid quadrature designs

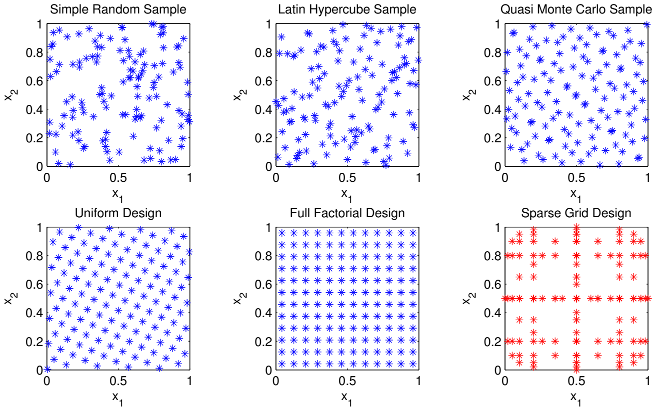
Given Data Methodology

X	k-dimensional random vector
Y	random variable (quantity of interest for time series)

- Physical observations
- Uncertainty propagation through model $Y = g(X)$
  - Simple random sampling of X
  - Latin Hypercube sampling of X
  - Quasi Monte Carlo sampling (Sobol' LPτ, ...) of X
  - But not fast multidimensional/spare grid quadrature designs

Sample must represent the underlying probabilistic framework.
Observations are independent realizations of (X, Y).

Examples for 2D Uniform [0, 1] Input Samples



Red: Bad setup. But fine for a meta-modeling layer

Why Sample-based?

- Minimal requirements for “design of experiments”
- No model-in-the-loop calculations
- Variable sample size: Fit for computational constraints
- Data re-use: Different sensitivity methods may be applied
- Classical statistical estimators: Bootstrap methods are available
- No functional structure necessary: no strict input/output relation

Why Sample-based?

- Minimal requirements for "design of experiments"
- No model-in-the-loop calculations
- Variable sample size: Fit for computational constraints
- Data re-use: Different sensitivity methods may be applied
- Classical statistical estimators: Bootstrap methods are available
- No functional structure necessary: no strict input/output relation

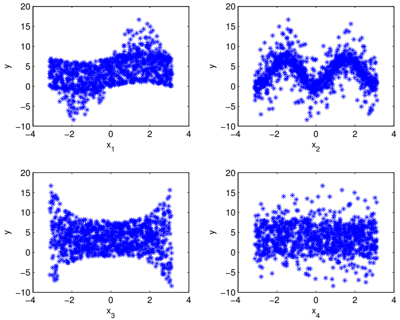
For the rest of the talk

$X = (x_{ij}) \in \mathbb{R}^{n \times k}$ input sample matrix (independent variables),

$Y = (y_i) \in \mathbb{R}^n$ output vector (dependent variable),

Example: The Ishigami test function

$$Y = \sin X_1 + 7\sin^2 X_2 + 0.1 X_3^4 \sin X_1, X_i \sim U(-\pi, \pi) \text{ iid.}$$



Elmar Plischke
IELF, TU Clausthal

Widely used example

- Only first parameter shows up in linear regression
- First order effects explain just 75% of output variance
- Second order interactions between factors 1 and 3
- Dummy serves as a control

Widely used example

- Only first parameter shows up in linear regression
- First order effects explain just 75% of output variance
- Second order interactions between factors 1 and 3
- Dummy serves as a control

Visual interpretation of Sensitivity Analysis: Condense the information of a scatterplot into one number!

Widely used example

- Only first parameter shows up in linear regression
- First order effects explain just 75% of output variance
- Second order interactions between factors 1 and 3
- Dummy serves as a control

Visual interpretation of Sensitivity Analysis: Condense the information of a scatterplot into one number!
In the following: Input factor *j* of interest is fixed – Group effects only touched marginally

Torture Chamber

Painful Least Squares Interrogations

Detecting linear dependence

Correlation between X_j and Y

Sample size			n
x-values			$x_1, \dots, x_i, \dots, x_n$
y-values			$y_1, \dots, y_i, \dots, y_n$
Sums			
x-values		$\Sigma x =$	$x_1 + \dots + x_i \dots + x_n$
y-values		$\Sigma y =$	$y_1 + \dots + y_i \dots + y_n$
x-squares		$\Sigma x^2 =$	$x_1^2 + \dots + x_i^2 \dots + x_n^2$
y-squares		$\Sigma y^2 =$	$y_1^2 + \dots + y_i^2 \dots + y_n^2$
xy-products		$\Sigma xy =$	$x_1 y_1 + \dots + x_i y_i \dots + x_n y_n$
Mean values		$\bar{x} = \frac{1}{n}(\Sigma x)$	$\bar{y} = \frac{1}{n}(\Sigma y)$
Deviances			
	$Sxx =$	$\sum_i (x_i - \bar{x})^2$	$= (\Sigma x^2) - \frac{1}{n}(\Sigma x)^2$
	$Syy =$	$\sum_i (y_i - \bar{y})^2$	$= (\Sigma y^2) - \frac{1}{n}(\Sigma y)^2$
	$Sxy =$	$\sum_i (x_i - \bar{x})(y_i - \bar{y})$	$= (\Sigma xy) - \frac{1}{n}(\Sigma x)(\Sigma y)$
Regression line		$y = \hat{a} + \hat{b}x :$	$\hat{b} = Sxy/Sxx,\ \hat{a} = \bar{y} - \hat{b}\bar{x}$
Correlation coefficient			$\rho = Sxy/\sqrt{SxxSxy}$
Coefficient of determination			$R^2 = (Sxy)^2/(SxxSyy)$

R^2 measures goodness of fit for a linear regression model!

Goodness of Fit – Coefficient of Determination

If $\hat{\varphi}_j(x) = \hat{a}x + \hat{b}$ is the linear regression model for factor j, then

$$R_j^2 = \frac{\mathbb{V}[\hat{\varphi}_j(X_j)]}{\mathbb{V}[Y]} = \frac{\mathbb{E}[(\hat{\varphi}(X_j) - \mathbb{E}[Y])^2]}{\mathbb{E}[(Y - \mathbb{E}[Y])^2]}$$

$\mathbb{V}[\cdot]$	variance
$\mathbb{E}[\cdot]$	expectation

Goodness of Fit – Coefficient of Determination

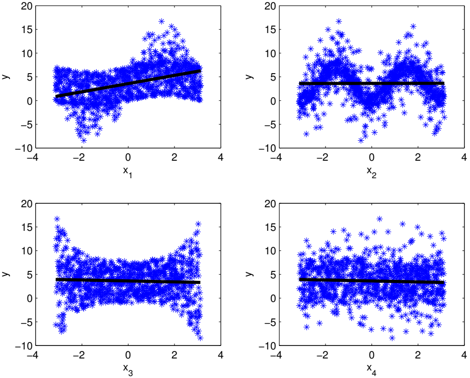
If $\hat{\varphi}_j(x) = \hat{a}x + \hat{b}$ is the linear regression model for factor j, then

$$R_j^2 = \frac{\mathbb{V}[\hat{\varphi}_j(X_j)]}{\mathbb{V}[Y]} = \frac{\mathbb{E}[(\hat{\varphi}(X_j) - \mathbb{E}[Y])^2]}{\mathbb{E}[(Y - \mathbb{E}[Y])^2]}$$

$\mathbb{V}[\cdot]$	variance
$\mathbb{E}[\cdot]$	expectation

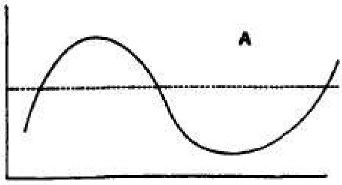
Fraction of the output variance explained by a linear dependence on input factor X_j

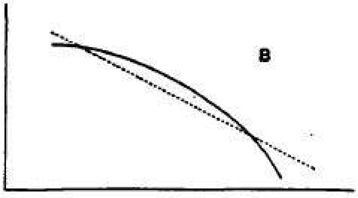
Example: Linear Regression



Parameter	1	2	3	4
R^2	0.1936	0.0000	0.0026	0.0026

But...





[Pearson, 1912]: "Nothing can be learnt of association by assuming linearity in a case with a regression line (plane, etc.) like A, much in a case like B."

Torture Chamber

Stretching and Squeezing with Rank Transformations

Detecting monotonic dependence

Spearman Rank–Correlation

Replace each value with its rank before computing the regression line

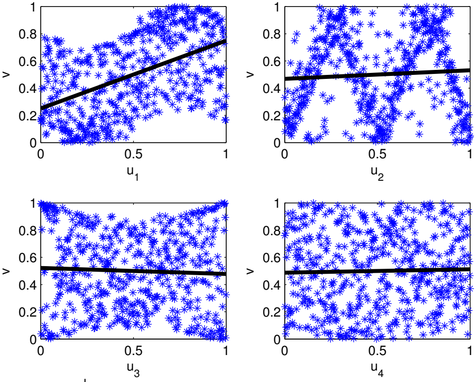
$$Rk(z) = \sum_{i=1}^{n} \mathbf{1}\{z_i \leq z\}$$

Almost the same as the empirical cdf

$$\hat{F}_Z(z) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\{z_i \leq z\}$$

Counting the number of realisations smaller or equal to a given value

Example: Rank Regression



Parameter	1	2	3	4
R^{*2}	0.2493	0.0040	0.0019	6.e-04

A Look through the Hidden Backdoor Advanced Instruments

Copula Theory

Study of properties invariant under strictly monotonic transformations [Nelsen, 2006]

A Look through the Hidden Backdoor
Advanced Instruments

Copula Theory

Study of properties invariant under strictly monotonic transformations [Nelsen, 2006]

Spearman R^{*2} is just one prominent example

A Look through the Hidden Backdoor
Advanced Instruments

Copula Theory

Study of properties invariant under strictly monotonic transformations [Nelsen, 2006]

Spearman R^{*2} is just one prominent example

[Borgonovo et al., 2014]: Many moment-independent sensitivity measures are transformation-invariant (Secret passage to Chamber 4)

Empirical Copula

Already used:

$$(x,y) \mapsto (u,v) \in [0,1]^2, u = \widehat{F}_X(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}\{x_j \leq x\}, v = \dots$$

Transformation of marginal distributions to uniform

Empirical Copula

Already used:

$$(x, y) \mapsto (u, v) \in [0, 1]^2, u = \widehat{F}_X(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}\{x_j \leq x\}, v = \dots$$

Transformation of marginal distributions to uniform
Same idea in 2D: Empirical bivariate copula

$$\widehat{C}(u, v) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}\{u_j \leq u\}\mathbf{1}\{v_j \leq v\}$$

Sensitivity from the Empirical Copula

Under Innocence: $\hat{C}(u, v)$ is a MC integral (relative frequency of hits) of the box area $u \cdot v$

Sensitivity from the Empirical Copula

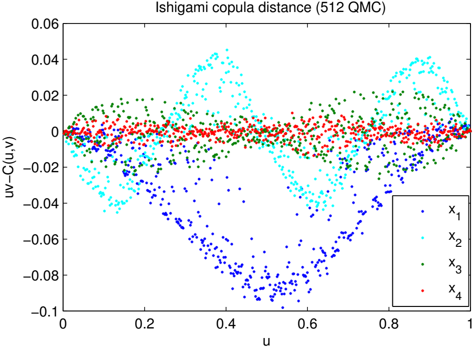
Under Innocence: $\hat{C}(u, v)$ is a MC integral (relative frequency of hits) of the box area $u \cdot v$

Hence: Compare the independent (product) copula with the empirical copula [Plischke and Borgonovo, 2015]

$$\varphi : u \mapsto uv - \widehat{C}(u, v)$$

Visual tool – Copula distance plots: u vs. $\varphi(u)$

Ishigami Copula Distance: Quasi MC, 512 runs



Spearman ϱ^*	$-12\int_u \varphi(u)du$	$(0.44, 0.00, 0.00, -0.00)$
Schweizer Wolff	$12\int_u \lvert\varphi(u)\rvert\, du$	$(0.44, 0.22, 0.09, 0.03)$
Discrepancy	$4\max_u \lvert\varphi(u)\rvert$	$(0.38, 0.18, 0.10, 0.04)$

Torture Chamber

Slicing into Functional ANOVA Pieces

Detecting functional dependence

Functional ANOVA

$g(\cdot)$ square integrable, $\mathbf{X}$ independent with $f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{k} f_i(x_i)$.
Unique representation:

$$g(\mathbf{x}) = \sum_{r=0}^{k} \sum_{\alpha:|\alpha|=r} g_\alpha(\mathbf{x}_\alpha)$$

where $\sum_\alpha$: sum over all subsets of indices of cardinality r.
$g_\alpha(\mathbf{x}_\alpha)$ determined by

$$g_0 = \int_{\mathcal{X}} g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

$$g_\alpha(\mathbf{x}_\alpha) = \int_{\mathcal{X}_{\sim\alpha}} \left(g(\mathbf{x}_\alpha, \mathbf{x}_{\sim\alpha}) - \sum_{\beta \subsetneq \alpha} g_\beta(\mathbf{x}_\beta) \right) f_{\sim\alpha}(\mathbf{x}_{\sim\alpha})\,\mathrm{d}\mathbf{x}_{\sim\alpha}$$

where $\mathcal{X}_\beta = \bigotimes_{j\in\beta} \mathcal{X}_j$ and $f_\beta(\mathbf{x}_\beta) = \prod_{j\in\beta} f_j(x_j)$.

Variance Decomposition

For $Y = g(\mathbf{X})$

$$\mathbb{V}[Y] = \sum_{r=1}^{m} \sum_{|\alpha|=r} \mathbb{V}(\alpha) \text{ where } \mathbb{V}(\alpha) = \int_{\mathcal{X}_\alpha} [g_\alpha(\mathbf{x}_\alpha)]^2 f_\alpha(\mathbf{x}_\alpha)\,\mathrm{d}\mathbf{x}_\alpha$$

First order effects	$S_j = \eta_j^2 = \frac{\mathbb{V}(\{j\})}{\mathbb{V}[Y]}$
Total effects	$S_j^T = \sum_{j\in\alpha} \frac{\mathbb{V}(\alpha)}{\mathbb{V}[Y]} = 1 - \sum_{j\notin\alpha} \frac{\mathbb{V}(\alpha)}{\mathbb{V}[Y]}$

Fraction of output variance explained by functional dependence on input factors or groups of input factors.

Sobol' Method

Two independent input samples are mixed to form new samples

$$x_A = \begin{pmatrix} .47 & .09 & .84 & .28 \\ .97 & .59 & .34 & .78 \\ .72 & .34 & .59 & .03 \\ .22 & .84 & .09 & .53 \\ .16 & .16 & .53 & .84 \end{pmatrix} \qquad x_B = \begin{pmatrix} .16 & .84 & .91 & .22 \\ .66 & .34 & .41 & .72 \\ .91 & .09 & .66 & .97 \\ .41 & .59 & .16 & .47 \\ .84 & .66 & .97 & .16 \end{pmatrix}$$

$$x_i = \begin{pmatrix} .47 & .09 & .91 & .28 \\ .97 & .59 & .41 & .78 \\ .72 & .34 & .66 & .03 \\ .22 & .84 & .16 & .53 \\ .16 & .16 & .97 & .84 \end{pmatrix}$$

Associated model output vectors: y_A, y_B and y_i

Sobol' Method II

Correlation coefficient between y_B and y_i:
Sobol' main effect $S_i = \eta_i^2$

Sobol' Method II

Correlation coefficient between y_B and y_i:
Sobol' main effect $S_i = \eta_i^2$

1− Correlation coefficient between y_A and y_i:
Sobol' total effect S_i^T

Sobol' Method II

Correlation coefficient between y_B and y_i:
Sobol' main effect $S_i = \eta_i^2$

$1-$ Correlation coefficient between y_A and y_i:
Sobol' total effect S_i^T

Computational Costs
Each sample block n simulations, given a model with k factors:
$(k+2)n$

Sobol' Method II

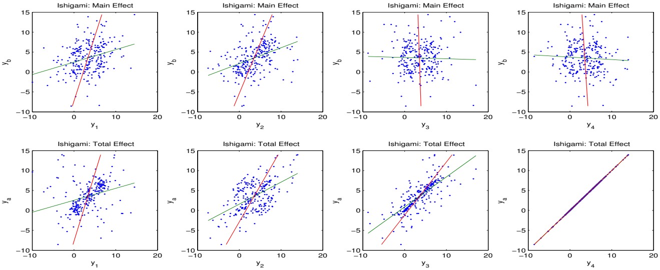
Correlation coefficient between y_B and y_i:
Sobol' main effect $S_i = \eta_i^2$

$1-$ Correlation coefficient between y_A and y_i:
Sobol' total effect S_i^T

Computational Costs

Each sample block n simulations, given a model with k factors:
$(k+2)n$
Model in the loop
Sophisticated sample design

Sobol' Method: Correlations for Ishigami



Factor	1	2	3	4
main effect	0.3223	0.4429	-0.0060	-0.0003
total effect	0.5644	0.4427	0.2399	0

Regression Lines: Acute angle $\Longrightarrow$ important main effect or unimportant total effect

First Order/Main Effects from Given Data

Given a nonlinear regression curve for factor j:
$\varphi_j(x) = \hat{\mathbb{E}}[Y|X_j = x]$
Estimation of first order effect $\eta_j^2 = \frac{\mathbb{V}[\mathbb{E}[Y|X_j]]}{\mathbb{V}[Y]}$

$$\hat{\eta}_j^2 = \frac{\sum_{i=1}^n (\varphi_i(x_{ij}) - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Goodness-of-fit or coefficient of determination, nonlinear R^2
Cost reduction for main effects: From $(k+1)n$ to n
For dependent input samples: no functional ANOVA needed

How to choose φ?

- Affine linear regression models
- Piecewise constant step functions (Pearson, 1905)
- (Piecewise) polynomials, Splines
- Orthogonal polynomial bases (Rabitz & Alış, 1999)
- Harmonic functions (Plischke, 2010)
- Wavelet decompositions
- Moving averages (Doksum & Samarov, 1995)
- LOESS/LASSO/ACOSSO Data Smoothing
- PCA of the output autocovariance, Karhunen-Loéve
- . . .

Cosine Transformation for Sensitivity Indicators [Plischke, 2012b]

For a given sample $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$

- Let ψ denote the order permutation of x, $x_{\psi(i)} \leq x_{\psi(i+1)}$.

Cosine Transformation for Sensitivity Indicators [Plischke, 2012b]

For a given sample $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$

- Let ψ denote the order permutation of x, $x_{\psi(i)} \leq x_{\psi(i+1)}$.
- Reorder y accordingly, $(y_\psi(i))_{i=1,\ldots,n}$

Cosine Transformation for Sensitivity Indicators [Plischke, 2012b]

For a given sample $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$

- Let ψ denote the order permutation of x, $x_{\psi(i)} \leq x_{\psi(i+1)}$.
- Reorder y accordingly, $(y_\psi(i))_{i=1,\dots,n}$
- Compute cosine transformation frequencies,

$$c_i = \sqrt{\frac{2}{n}} \sum_{j=1}^{n} \cos\left(\frac{\pi(2j-1)i}{2n}\right) y_{\psi(j)}.$$

Cosine Transformation for Sensitivity Indicators [Plischke, 2012b]

For a given sample $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$

- Let ψ denote the order permutation of x, $x_{\psi(i)} \leq x_{\psi(i+1)}$.
- Reorder y accordingly, $(y_\psi(i))_{i=1,\dots,n}$
- Compute cosine transformation frequencies,

$$c_i = \sqrt{\frac{2}{n}} \sum_{j=1}^{n} \cos\left(\frac{\pi(2j-1)i}{2n}\right) y_{\psi(j)}.$$

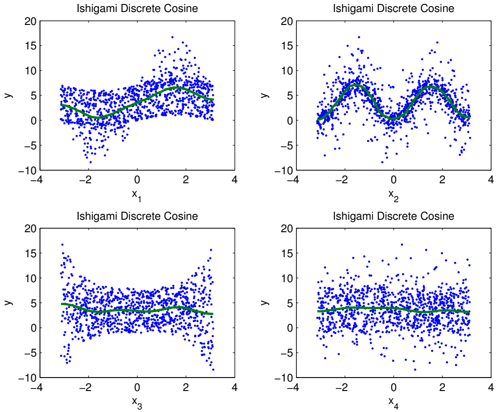
- Gather resonating frequencies,

$$\hat{\eta}^2 = \frac{\sum_{i=1}^{M} c_i^2}{\sum_{i=1}^{n-1} c_i^2}.$$

MATLAB implementation computing first order sensitivity indices

```
 [xr,index]=sort(x);
 yr=y(index); % Reorder output
 allcoeff=dct(yr); % Compute transformation
% Unconditional variance
 V = sum(allcoeff(2:end,:).^2);
% Conditional variance with M resonating harmonics
 Vi= sum(allcoeff(1+(1:M),:).^2);  Si= Vi./V;
```

Cosine sensitivity for the Ishigami example



Parameter	1	2	3	4
η^2	0.3121	0.4333	0.0162	0.0079

Condensing several scatterplots into one figure

Contribution of the input factors to the output mean: Absolute concentration curve, contribution to the sample mean, [Plischke, 2012a]

Cumulative Sum of Normalised Reordered Output (CUSUNORO): Plot the index i against

$$z(i) = \frac{\sum_{j=1}^{i}(y_{\psi(j)} - \bar{y})}{\sqrt{\sum_{j=1}^{n}(y_j - \bar{y})^2}}$$

where $\psi(\cdot)$ is the order permutation for (x_j), $(x_{\psi(j)}) = (x_{(j)})$, $x_{(j)} \leq x_{(j+1)}$.

Condensing several scatterplots into one figure

Contribution of the input factors to the output mean: Absolute concentration curve, contribution to the sample mean, [Plischke, 2012a]

Cumulative Sum of Normalised Reordered Output (CUSUNORO): Plot the index i against

$$z(i) = \frac{\sum_{j=1}^{i}(y_{\psi(j)} - \bar{y})}{\sqrt{\sum_{j=1}^{n}(y_j - \bar{y})^2}}$$

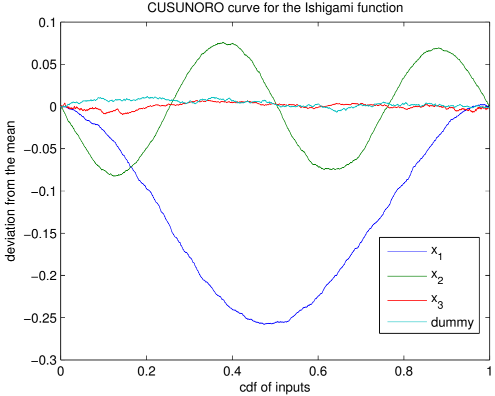
where $\psi(\cdot)$ is the order permutation for (x_j), $(x_{\psi(j)}) = (x_{(j)})$, $x_{(j)} \le x_{(j+1)}$.

Sample-based version of

$$u \mapsto \frac{1}{\sqrt{\mathbb{V}[Y]}}\mathbb{E}\left[(Y - \mathbb{E}[Y])\mathbf{1}\{X \le F_X^{-1}(u)\}\right]$$

Example: Again Ishigami

Instead of *k* scatterplots: *k* curves in a CUSUNORO plot



Link to First Order Effects

Derivative of CUSUNORO is conditional expectation!
Parameters are sensitive if CUSUNORO deviates from 0-line:
With slight abuse of notation,

$$\hat{\eta}^2 = \int_0^1 (\nabla z(t))^2 \, dt$$

Main effect estimate: Riemann sum of the square of the derivate of the CUSUNORO curve

Sensitivity and Interactions

So far: Methods for estimating first order effects from given data.

Sensitivity and Interactions

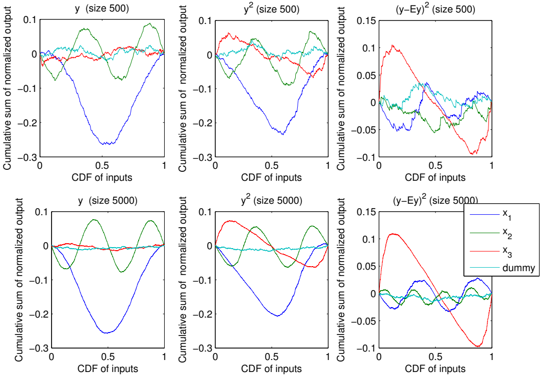
So far: Methods for estimating first order effects from given data.
But: No estimate of total effects
Total effects are important in Factor Fixing Setting -
Identification of input factors without influence on the output

Sensitivity and Interactions

So far: Methods for estimating first order effects from given data.
But: No estimate of total effects
Total effects are important in Factor Fixing Setting -
Identification of input factors without influence on the output
Alternatives

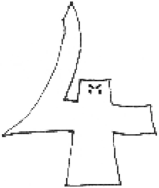
- Higher order effects using generalized regression with feature maps from an orthogonal function basis (HDMR, Polynomial Chaos approach)
- Main effects of higher moments - Replace $y \to (y - \bar{y})^2$
- Compare the probability of Y with the probability of Y conditional to $X_i = x$

CUSUNORO for the second moment



x_3 is sensitive!

Torture Chamber



Iron Maiden of Conditional Piercings

Detecting statistical dependence

General Framework for Sensitivity

A general sensitivity measure

$$\gamma_j(X_j, Y) = \mathbb{E}[\zeta(\mathbb{P}_Y, \mathbb{P}_{Y|X_j})]$$

Suitable distance measure $\zeta(\cdot,\cdot)$: shift/separation/contrast between total and conditional probability measures.

General Framework for Sensitivity

A general sensitivity measure

$$\gamma_j(X_j, Y) = \mathbb{E}[\zeta(\mathbb{P}_Y, \mathbb{P}_{Y|X_j})]$$

Suitable distance measure $\zeta(\cdot, \cdot)$: shift/separation/contrast between total and conditional probability measures.

Aswers the question:
What is the (average) value of getting to know that $X_j = x$?

Examples for Shift/Separation/Contrast Functions

$$\zeta_{\mathrm{EI}}(\mu_Y, \mu_{Y|X=x}) = \max\{\mu_{Y|X=x}, 0\} - \max\{\mu_Y, 0\}$$
EVPI, null alternative

$$\zeta_{\mathrm{SI}}(\mu_Y, \mu_{Y|X=x}) = \sigma_Y^{-2}(\mu_Y - \mu_{Y|X=x})^2$$
Main Effect

$$\zeta_{\mathrm{KS}}(F_Y, F_{Y|X=x}) = \sup \left|F_Y - F_{Y|X=x}\right|$$
Kolmogorov-Smirnov

$$\zeta_{\mathrm{Ku}}(F_Y, F_{Y|X=x}) = \sup \left(F_Y - F_{Y|X=x}\right) - \inf \left(F_Y - F_{Y|X=x}\right)$$
Kuiper

$$\zeta_{\mathrm{CvM}}(F_Y, F_{Y|X=x}) = \frac{1}{2}\int \left(F_{Y|X=x}(y) - F_Y(y)\right)^2 dy$$
Cramér, L^2 (cdf)

$$\zeta_{\mathrm{Bo}}(f_Y, f_{Y|X=x}) = \frac{1}{2}\int \left|f_{Y|X=x}(y) - f_Y(y)\right| dy$$
Borgonovo, L^1 (pdf)

$$\zeta_{\mathrm{KL}}(f_Y, f_{Y|X=x}) = \int f_{Y|X=x}(y) \log \frac{f_{Y|X=x}(y)}{f_Y(y)} dy$$
Kullback-Leibler

$$\zeta_{\mathrm{He}}(f_Y, f_{Y|X=x}) = 1 - \int \sqrt{f_Y(y) \cdot f_{Y|X=x}(y)}dy$$
Hellinger

Moment-Independent Sensitivity

The variance of the conditional expectation is is problematic when the output distribution is multi-modal or highly skewed or in case of heteroscedastic data.

Variance works as uncertainty measure only for distributions close to normal.

A moment-independent measure should compare conditional densities/distributions with the unconditional one without recurring to special moments.

Density-Based Sensitivity Measue

Use mean distance between conditional and unconditional densities [Borgonovo, 2007].

$$\delta(X,Y) = \tfrac{1}{2}\,\mathbb{E}[S(X)] = \frac{1}{2}\int_{\mathcal{X}} f_X(x)\int_{\mathcal{Y}} \left| f_Y(y) - f_{Y|X}(y|x) \right| \, dy\, dx$$

δ: L^1-distance of the product of the marginal distributions and the joint distribution

Estimating δ

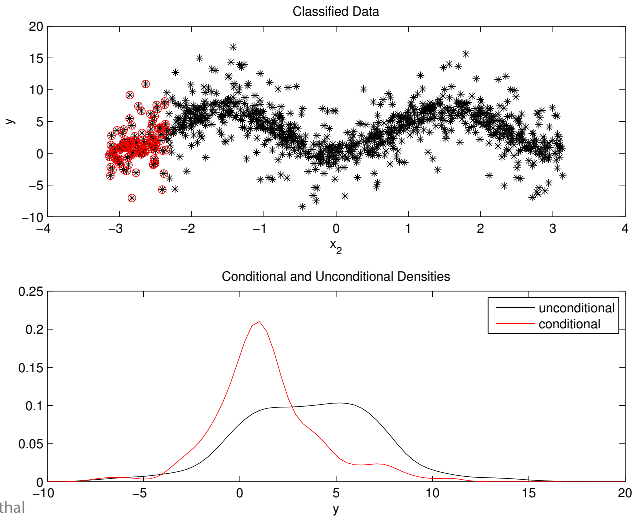
Comparison with variance-based methods:
Instead of conditional expectation $\mathbb{E}[Y|X = x]$: conditional density $f_{Y|X}(y|x)$.

Estimating δ

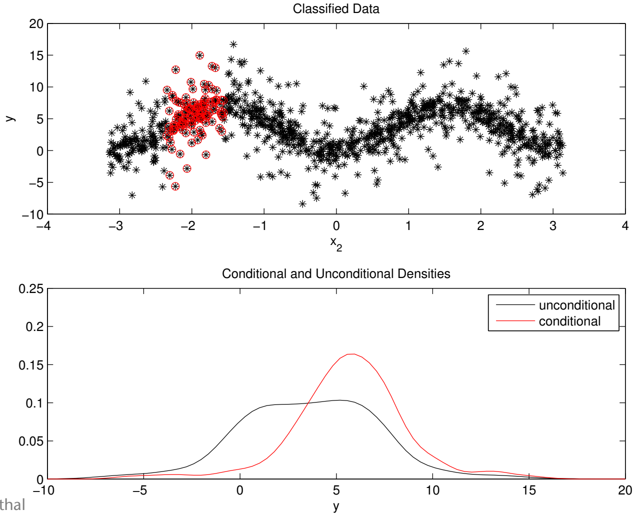
Comparison with variance-based methods:
Instead of conditional expectation $\mathbb{E}[Y|X = x]$: conditional density $f_{Y|X}(y|x)$.

Replace $f_{Y|X=x}(y)$ by $\hat{f}_{Y|X\in\mathcal{C}_r}(y)$ for suitable classes $\mathcal{C}_r$, $r = 1, \ldots, s$, and use kernel density estimation on the classified data $\{y_i|\, x_i \in \mathcal{C}_r\}$ [Plischke et al., 2013]

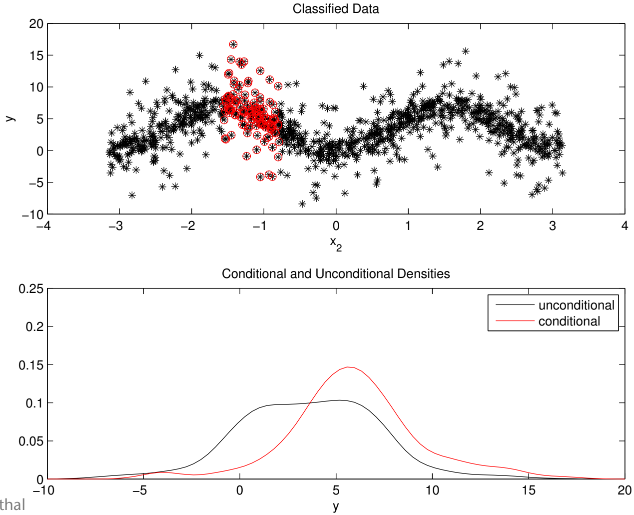
Example: Ishigami Function, Parameter 2



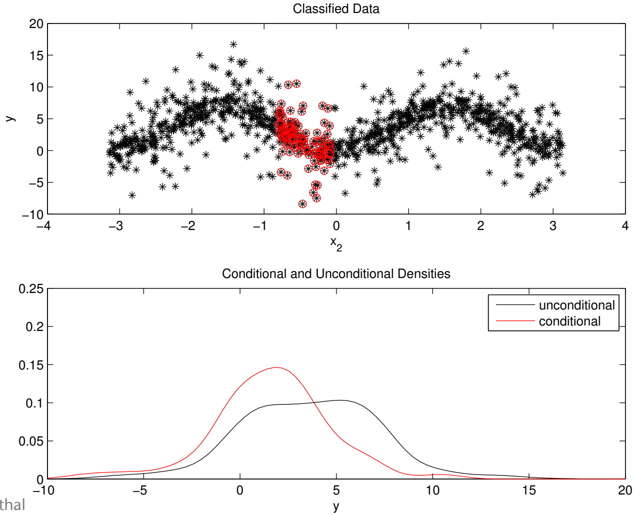
Example: Ishigami Function, Parameter 2



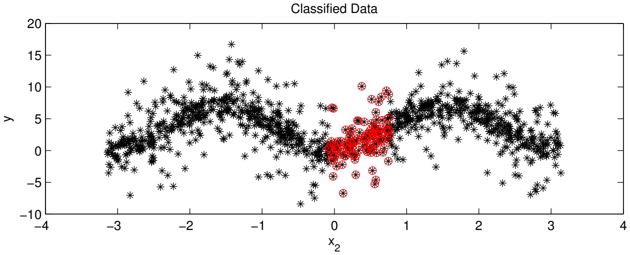
Example: Ishigami Function, Parameter 2

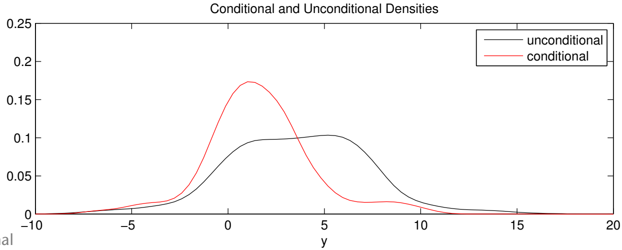


Example: Ishigami Function, Parameter 2

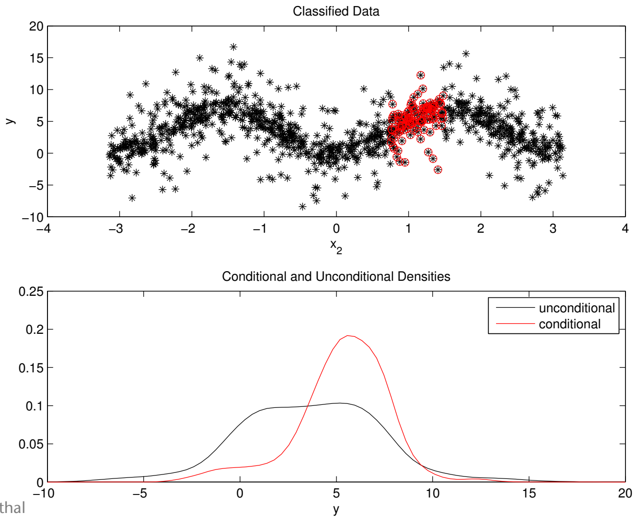


Example: Ishigami Function, Parameter 2

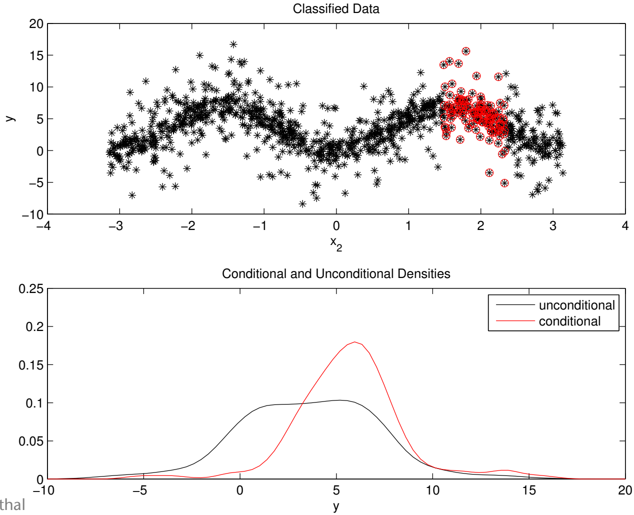




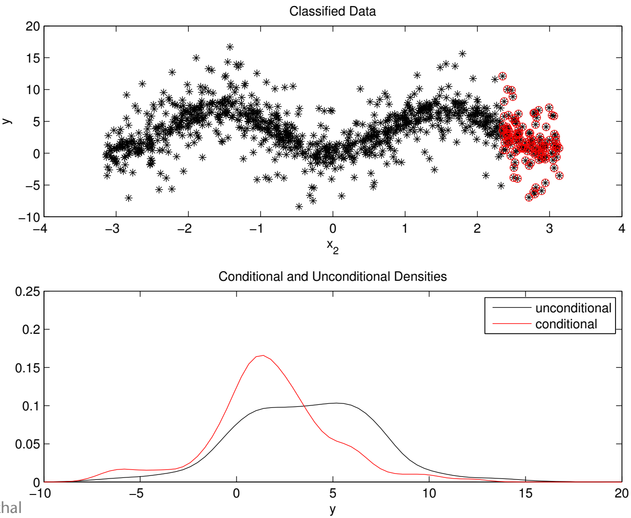
Example: Ishigami Function, Parameter 2



Example: Ishigami Function, Parameter 2



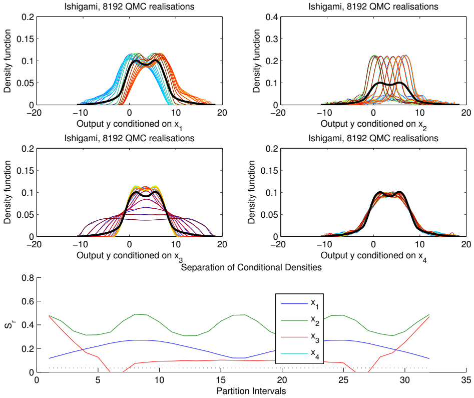
Example: Ishigami Function, Parameter 2



MATLAB Implementation for δ estimation

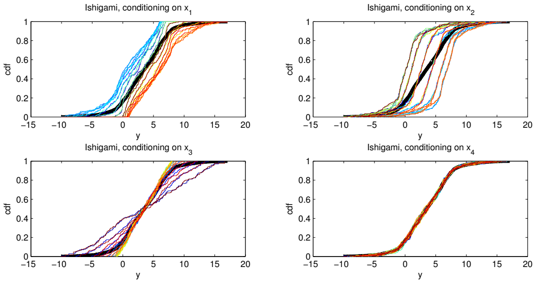
```
Kernel=@(x)3/(4*sqrt(5))*max(1-(x.^2/5),0);[n,k]=size(x);
iqry=median(abs(median(y)-y));
% bandwidth estimate (rule of thumb)
stdy=min(std(y),iqry/(2*0.675)); h=stdy*((4/(3*n))^(1/5));
z=linspace(min(y),max(y),100); % quadrature points
W=Kernel(bsxfun(@minus,z,y)/h)/h; densy=mean(W); %KDE
[xr,indxx]=sort(x);
for i=1:k;   xr(indxx(:,i),i)=1:n; end % ranks
for j=1:M
   indx=((j-1)*n/M<xr) & (xr <= j*n/M);   nm(:,j)=sum(indx);
   for i=1:k
   densc=mean(W(indx(:,i),:)); % conditional density
   Sm(i,j)=trapz(z,max(densy-densc,0)); %only positive part
end
Sm(Sm<Cutoff.*sqrt(1/n+1./nm))=0; d=sum(Sm.*nm,2)'/n;
```

Ishigami Example, 8192 Quasi-MC Samples



Parameter	1	2	3	4
δ	0.2037	0.3918	0.1392	0
η^2	0.3132	0.4365	0.0000	0.0000

Further Moment-Independent Measures

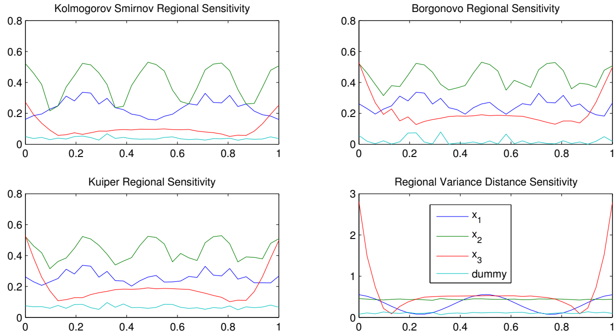


Parameter	1	2	3	4
First Effect η^2	0.3122	0.4321	0.0001	0.0005
Kolmogorov Smirnov β	0.2307	0.3929	0.1018	0.0403
Borgonovo δ	0.2460	0.4213	0.2105	0.0419
Kuiper	0.2525	0.4205	0.1950	0.0694
Wald Wolfowitz	1.6218	2.0463	0.9198	-0.2885

Wald Wolfowitz: Number of conditional subsample runs, normalized

Regional Sensitivity

Plot the separation: $x \mapsto \max_y \left| F_Y - F_{Y|X=x} \right|$ etc.



When to use what?

Sometimes a linear regression provides enough information for a better model understanding.
However, always interpret a sensitivity measure of 0 as inconclusive.

Advantages of linear methods

Small sample of (X, Y) provides info, widely available in software packages

Advantages of variance-based methods

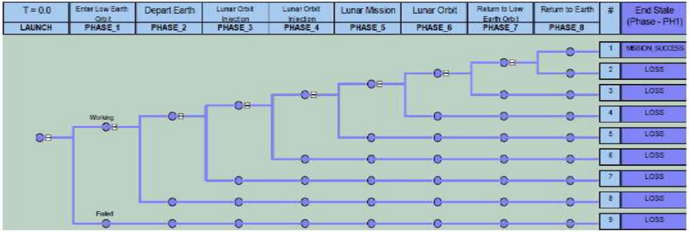
Functional depedence of Y on X detectable

Advantages of moment-independent methods

Indepedence of X and Y $\leftrightarrow$ vanishing indicator

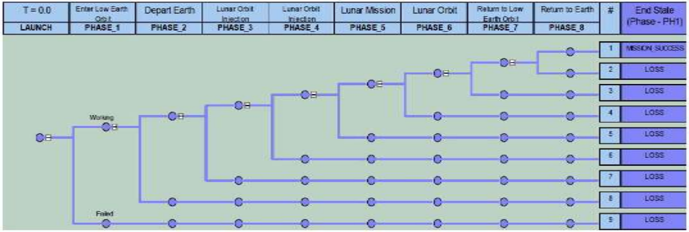
Decision support: NASA lunar mission

Risk assessment of lunar space missions
8-phase process from launch to moon orbit, to return to earth



Decision support: NASA lunar mission

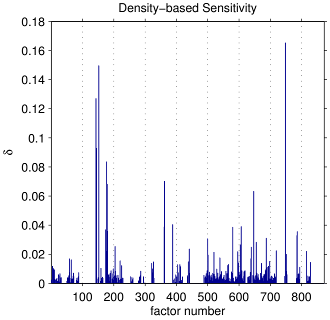
Risk assessment of lunar space missions
8-phase process from launch to moon orbit, to return to earth

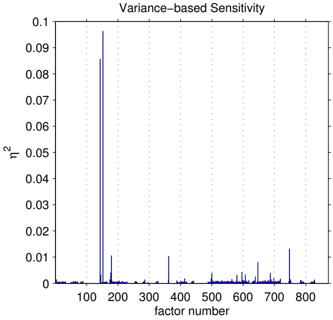


Model is a blackbox processing $k = 872$ uncertain input factors
Crucial: Factors to focus resources in data collection and further modelling efforts

Lunar Mission: Results

QMC sample, 65536×872





Only a few key drivers of uncertainty

Thank You!

Questions, Comments

mailto:elmar.plischke@tu-clausthal.de

Preprints, Scripts, Stuff

http://www.immr.tu-clausthal.de/~epl/

References I

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

Borgonovo, E., Tarantola, S., Plischke, E., and Morris, M. D. (2014).
Transformations and invariance in the sensitivity analysis of computer experiments.
Journal of the Royal Statistical Society, Series B, 76:925–947.

Faivre, R., Iooss, B., Mahévas, S., Makowski, D., and Monod, H., editors (2013).
Analyse de sensibilité et exploration de modèles.
Éditions Quæ, Versailles, France.

Nelsen, R. B. (2006).
An Introduction to Copulas.
Springer Series in Statistics. Springer Verlag, 2nd edition.

Pearson, K. (1912).
On the general theory of the influence of selection on correlation and variation.
Biometrika, 8(3–4):437–443.

Plischke, E. (2012a).
An adaptive correlation ratio method using the cumulative sum of the reordered output.
Reliability Engineering&System Safety, 107:149–156.

References II

Plischke, E. (2012b).
How to compute variance-based sensitivity indicators with your spreadsheet software.
Environmental Modelling&Software, 35:188–191.

Plischke, E. and Borgonovo, E. (2015).
Copula-based sensitivity measures of computer experiments.
Technical report, TU Clausthal.
Submitted to ESREL 2015, ETH Zürich.

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

Saltelli, A., Chan, K., and Scott, E. M. (2000).
Sensitivity Analysis.
John Wiley&Sons, Chichester.