Global Sensitivity Analysis Methods

Which variables matter the most for a model's outcome changes?

Thale Eliassen Fink, Nicole Quattrini, Aleksei Berdiuzhenko, Irene Hofmann, Anette Fagerheim Bjerke, Tore Gude & Martes Zibellina

Content

- Introduction to Sensitivity Analysis
- Types of GSA Methods
- Practical Example - The Sobol Method
- Wrap up

Introduction

Introduction to Sensitivity Analysis

- Sensitivity analysis helps us understand how changes in model inputs influence the outputs. It’s useful for identifying which factors significantly affect a model’s predictions and can guide model improvement, simplification, and decision-making.

- Distinguish between global and local sensitivity analysis

Why Use Global Sensitivity Analysis?

- Look at the impact of each parameter across all possible values
- Find the effects of each parameter and interactive effects
- Suits non-linear problems or complex models ➜ input factors vary simultaneously
- Why do we use it:
  - Model evaluation
  - Model simplification
  - Model refinement
  - Exploratory modelling

Source: Addressing Uncertainty in MultiSector Dynamics Research, Chapter 3

Some foreshadowing

The animal, Sobol

Types of GSA Methods

Elementary Effect (Morris) Method

How it Works:

- Adjusts each input *slightly*, while keeping the others fixed ⚖️
- Gives a set of “elementary effects”, a measure of how output is changed, for each input. 📈
  - Then finds mean and standard deviation of these

When to Use:

- **Speed:** ⚡ Quick, computationally light
- **Model:** Works well with *non-linear* models or those with *interactions*
- **When:** Great for Initial sensitivity analysis 🏁

Source: Chap 3.4.2 Addressing Uncertainty in Multisector Dynamics Research - Book

Regression-based Methods

How it Works:

- Estimates the relationship between input and output using linear or polynomial regression models
- By analysing their coefficients, can identify impact of inputs

When to use:

- **Speed:** ⚡ Quick, Computationally light
- **Model:** Best when relationships can be approximated with regression, complexity is low-moderate
- **When:** Great for Initial sensitivity analysis 🏁

Source: Chap 3.4.3 Addressing Uncertainty in Multisector Dynamics Research - Book

Moment-Independent Methods (e.g., Density-based)

How it Works:

- Calculated how the probability distribution of output is affected by each input change
- Focuses on capturing the full output shape (E.g Spread, skewness, clusters), not just variance and mean alone

When to Use:

- **Speed:** 🐢 Slower
- **Model:** Non-linear, non-normal, complex
- **When:** You want more detailed insight into output distribution 📈🔎

Source: Chap 3.4.7 Addressing Uncertainty in Multisector Dynamics Research - Book

Variance-based Methods (e.g., Sobol)

How it Works:

- Quantifies how much of output variance can be attributed to each input factor
- Measures this in Sobol indices
  - First-order Sobol index: How much each input alone effect output 🎯
  - Total-order Sobol index: How main effect and interactions with other inputs effect output 🌐

When to Use:

- **Speed:** 🐢 Slower, computationally intensive
- **Model:** Complex, non-linear
- **When:** You want detailed SA and capture interactions between factors 🔎

Source: Chap 3.4.5 Addressing Uncertainty in Multisector Dynamics Research - Book

Practical Example - Sobol

Sobol Sensitivity Analysis

- Global Sensitivity Analysis (GSA): Evaluates the influence of inputs across their full range, ideal for non-linear models.

- Variance-Based Decomposition: Decomposes output variance to determine each input's impact on the output.

- Sobol Indices:

First-Order Index (S1)Second-Order Index (S2),*Total-Order Index (ST)*

Ilya M. Sobol'

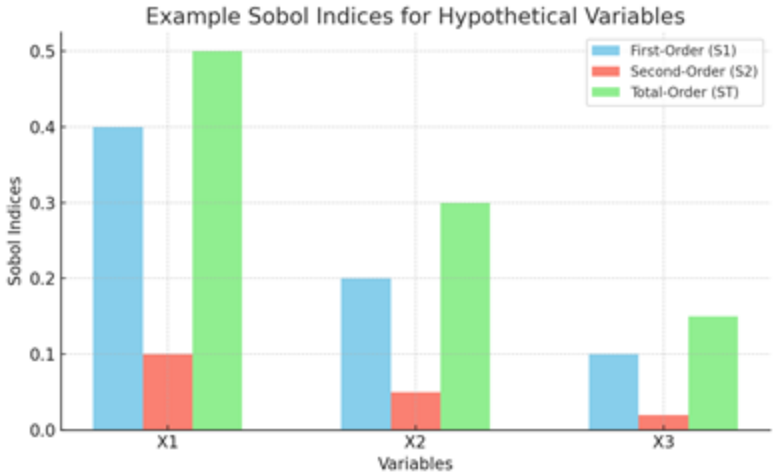
Prof. Sobol' at MCM2001, the third IMACS Seminar on Monte Carlo Methods, September 2001 in Salzburg.

Born	Ilya Meyerovich Sobol' 15 August 1926 (age 98) Panevezhas, Lithuania
Known for	Sobol' sequence Sensitivity analysis Optimization Astrophysics
Awards	USSR Medal for Labour Valour and the Order of the Honour Badge
	Scientific career
Fields	Mathematics

In some languages, “sobol” is this beauty (above)

But we are talking about a method, developed renowned Soviet mathematician, Ilya Sobol’ (yes, his last name literally means that furry animal above)

Applying Sobol Analysis and Key Benefits



Steps in Conducting Sobol Analysis

1. Define model and input ranges.
2. Generate quasi-random sample sets (Sobol Sequences)
3. Run model for each sample.
4. Calculate Sobol Indices.

Key Benefits

- **Handles Nonlinearity**: Suitable for models with complex, nonlinear interactions.
- **Quantifies Interactions**: Clearly distinguishes between individual and combined variable effects.
- **Improves Model Interpretability**: Provides insights into the importance of each variable.

Air Quality Dataset

Contains daily air quality measurements in New York from May - Sep 1973

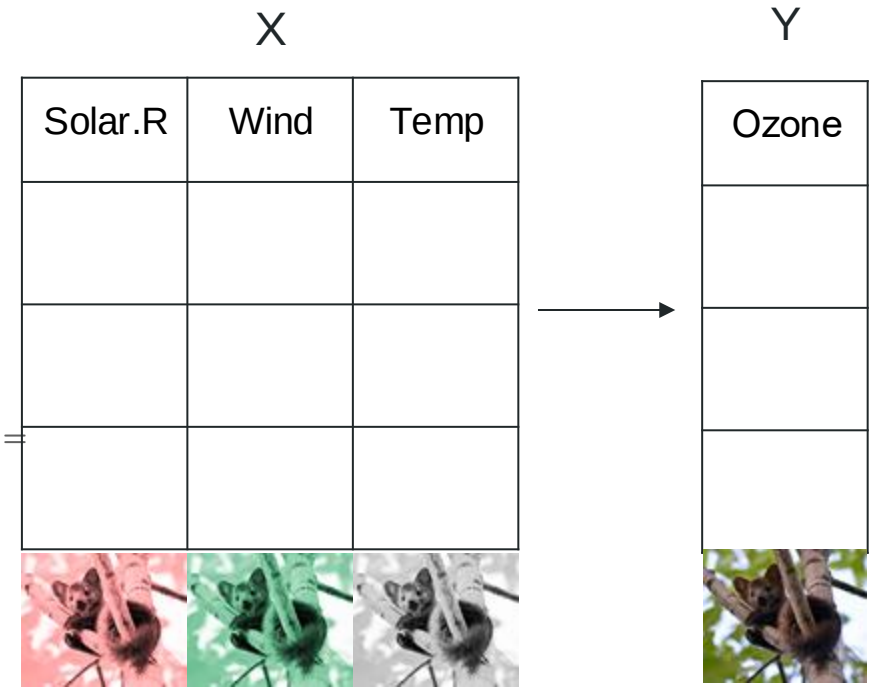
RQ: How do different factors affect the Ozone concentration?

Model: Linear Regression:

```
lm(Ozone ~ Solar.R + Wind + Temp,
                                data =
airquality)
```

Sensitivity Analysis:

Sobol Global Sensitivity



Air Quality Example

Full code located in pdf file:
week11_group4.pdf

Results Interpretation:

1) **First-order Sensitivity Index (S1)**:
   a) S1 for Solar.R = 0.3 → 30% of Ozone variance due to Solar.R alone
   b) S1 for Wind = 0.4
   c) S1 for Temp = 0.2
2) **Total Effect Index (ST)**:
   a) ST for Solar.R = 0.5 → 50% of Ozone variance due to Solar.R and its interaction with other parameters
   b) ST for Wind = 0.6
   c) ST for Temp = 0.3

TK8117_Week11_Group4

N.Quattrini

2024-11-05

SOBOL'S GLOBAL SENSITIVITY ANALYSIS using airquality dataset - contains daily air quality measurements in New York from May to Semptember 1973 - Q: how do different factors affect the Ozone concentration?

PACKAGES

```
library(sensitivity)
```

```
## Registered S3 method overwritten by 'sensitivity':
##   method    from
##   print.src dplyr
```

```
library(boot)
library(ggplot2)
```

STEP 1: LOADING & PREPARING DATA

- importing dataset: 153 rows, 6 cols
- removing rows containing NAs (42) –> 111 rows, 6 cols remain

```
data(airquality)
airquality <- na.omit(airquality)
```

STEP 2 DEFINE THE MODEL

Creating simple linear model to predict Ozone concentration based on Solar.R, Wind, and Temp

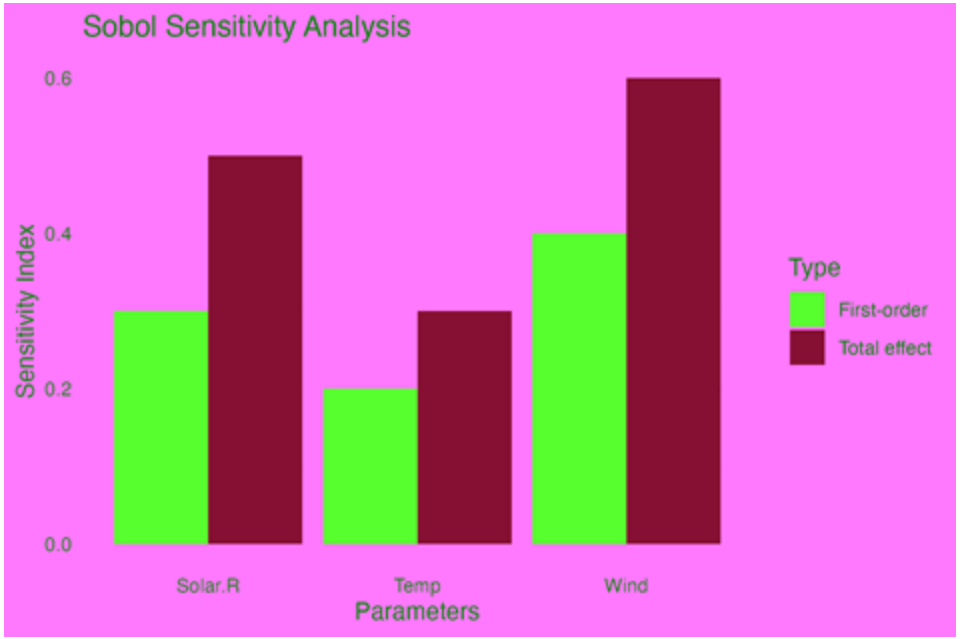
```
model <- function(X){
  fit <- lm(Ozone ~ Solar.R + Wind + Temp, data = airquality)
  predict(fit, newdata = data.frame(Solar.R = X[,1], Wind = X[,2],
                                    Temp = X[,3]))
}
```

Air Quality Dataset - Conclusions

Solar.R:
- 30% contribution to Ozone variation individually;
- 50% contribution when considering interactions.

Temp:
- 20% contribution to Ozone variation individually;
- 30% contribution when considering interactions.

Wind:
- 40% contribution individually;
- 60% contribution w/ interactions.



Wrap Up

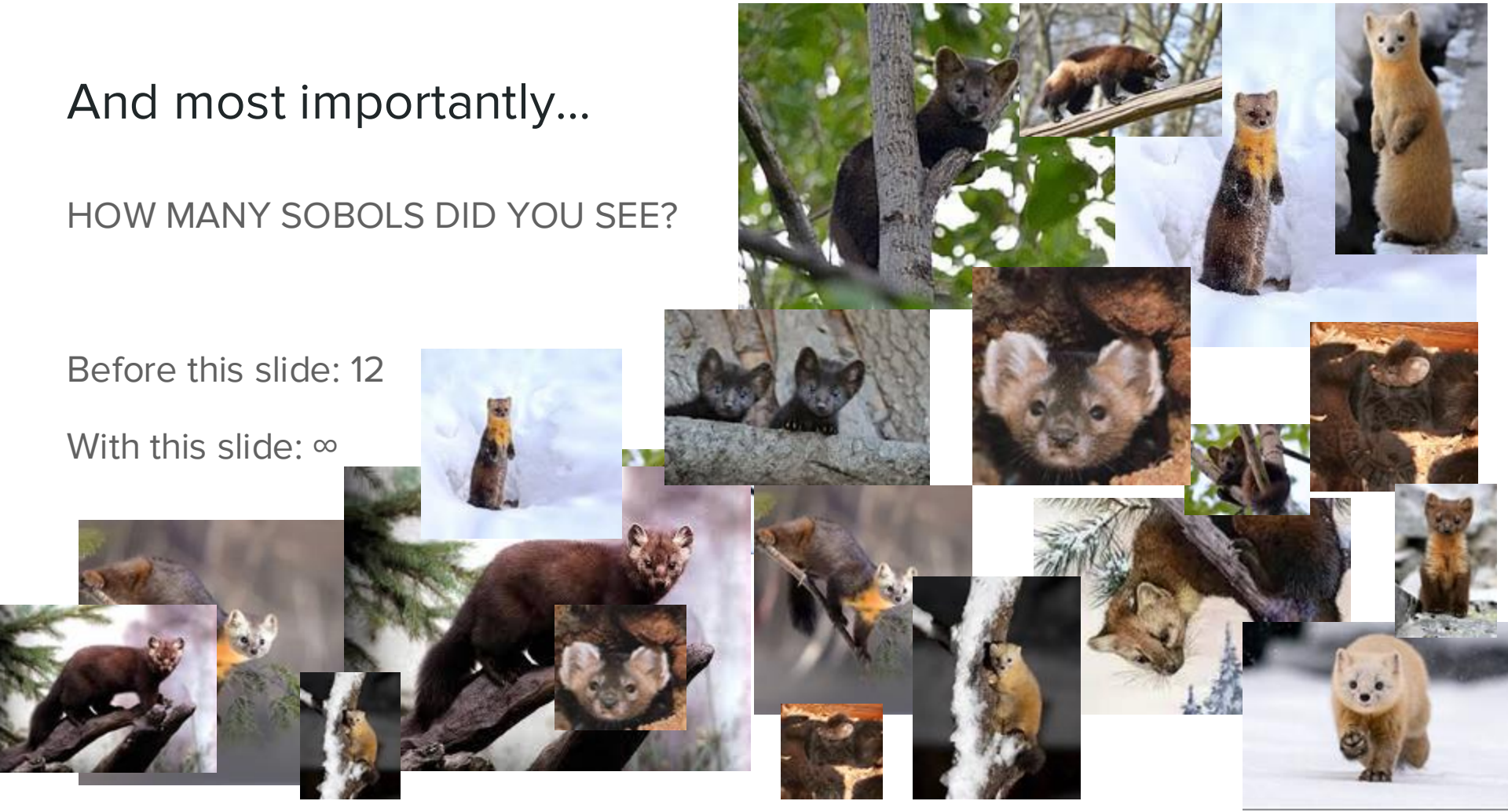
Name (e.g.)	Sobol 	Morris	Regression based	Density-based
Type	Variance-based, quantitative	Screening, qualitative	Regression based	Moment-Independent Methods
Computational cost	High	Low	Low	High
Detail	High	Low	Low	High
Best for	Precise sensitivity ranking	Quick screening of high dimensions	Initial sensitivity analysis	Detailed insight into output distribution
Use cases (examples)	Environmental modeling, engineering design	Manufacturing and production optimization, economic modeling	Risk assessment, climate modelling	Financial modeling, risk analysis

And most importantly...

HOW MANY SOBOLS DID YOU SEE?

Before this slide: 12

With this slide: ∞



Sources

- Reed, P.M., Hadjimichael, A., Malek, K., Karimi, T., Vernon, C.R., Srikrishnan, V., Gupta, R.S., Gold, D.F., Lee, B., Keller, K., Thurber, T.B, & Rice, J.S. (2022). Addressing Uncertainty in Multisector Dynamics Research [Book]. Zenodo. https://doi.org/10.5281/zenodo.6110623 https://uc-ebook.org/docs/html/index.html