

---

# Global Sensitivity Analysis of Predictor Models in Software Engineering

Stefan Wagner  
Software & Systems Engineering  
Technische Universität München  
Germany  
wagnerst@in.tum.de

May 20, 2007

## Introduction

Predictor Models in  
Software

Engineering

Problem

Example:

Fischer-Wagner  
Model

Questions About the  
Model

Overview

Global Sensitivity  
Analysis

---

Approach for SE

---

Summary

---

# Introduction

Introduction

Predictor Models in  
Software  
Engineering

Problem

Example:

Fischer-Wagner

Model

Questions About the

Model

Overview

Global Sensitivity

Analysis

---

Approach for SE

---

Summary

---

- Predictor models describe situations in software engineering
- Various types, aims, ...
- Examples
  - ◆ COCOMO (costs)
  - ◆ Musa-Okumoto (reliability)
  - ◆ Fault-proneness

- [Introduction](#)
- [Predictor Models in Software Engineering](#)
- [Problem](#)**
- [Example: Fischer-Wagner Model](#)
- [Questions About the Model](#)
- [Overview](#)
- [Global Sensitivity Analysis](#)
- [Approach for SE](#)
- [Summary](#)

- The models themselves can be complex
- Their use needs a lot of effort
- How can I analyse those models themselves?
- How can I simplify them?
- How do I improve their predictive power?

- Introduction
- Predictor Models in Software Engineering
- Problem
- Example: Fischer-Wagner Model**
- Questions About the Model
- Overview
- Global Sensitivity Analysis
- Approach for SE
- Summary

- Reliability model used at Siemens COM
- Two parameters estimated by failure data
- Failure probability of a fault:  $p_a = p_1 \cdot d^{(a-1)}$
- Geometrical distribution :  $F_a(t) = 1 - (1 - p_a)^t$
- Cumulated failures up to  $t$ :  

$$\mu(t) = \sum_{a=1}^{\infty} 1 - (1 - p_1 \cdot d^{(a-1)})$$
- Input parameters
  - ◆  $p_1$ : Highest failure probability of a fault
  - ◆  $d$ : Complexity of the system  $]0; 1[$
  - ◆  $t$ : Execution time (incidents)
  - ◆  $inf$ : approximation of  $\infty$
- Output:  $\mu(t)$

More details: S. Wagner, H. Fischer. Ada-Europe, 2006

- [Introduction](#)
- [Predictor Models in Software Engineering](#)
- [Problem Example: Fischer-Wagner Model](#)
- [Questions About the Model](#)**
- [Overview](#)
- [Global Sensitivity Analysis](#)
- [Approach for SE](#)
- [Summary](#)

- How do the input parameters influence the output?
- Which parameter(s) do I have to estimate best to get a good prediction?
- Are there insignificant parameters that I can remove?

Introduction  
Predictor Models in  
Software  
Engineering  
Problem  
Example:  
Fischer-Wagner  
Model  
Questions About the  
Model

**Overview**

Global Sensitivity  
Analysis

Approach for SE

Summary

## Introduction

## Global Sensitivity Analysis

## Approach for SE

## Summary

Introduction

**Global Sensitivity  
Analysis**

Definition

Method

Variance-based

Methods

Main Effect

Higher-Order Effects

Total Effect

Approach for SE

Summary

# Global Sensitivity Analysis

Introduction

Global Sensitivity  
Analysis

**Definition**

Method

Variance-based  
Methods

Main Effect

Higher-Order Effects

Total Effect

Approach for SE

Summary

- Saltelli (2000): “Sensitivity analysis studies the relationships between information flowing in and out of the model.”
- How do the input parameters influence the output?
- How can this influence be quantified?
- Global properties
  - ◆ Inclusion of influence of shape and scale
  - ◆ Multidimensional averaging

[Introduction](#)

[Global Sensitivity Analysis](#)

[Definition](#)

**Method**

[Variance-based Methods](#)

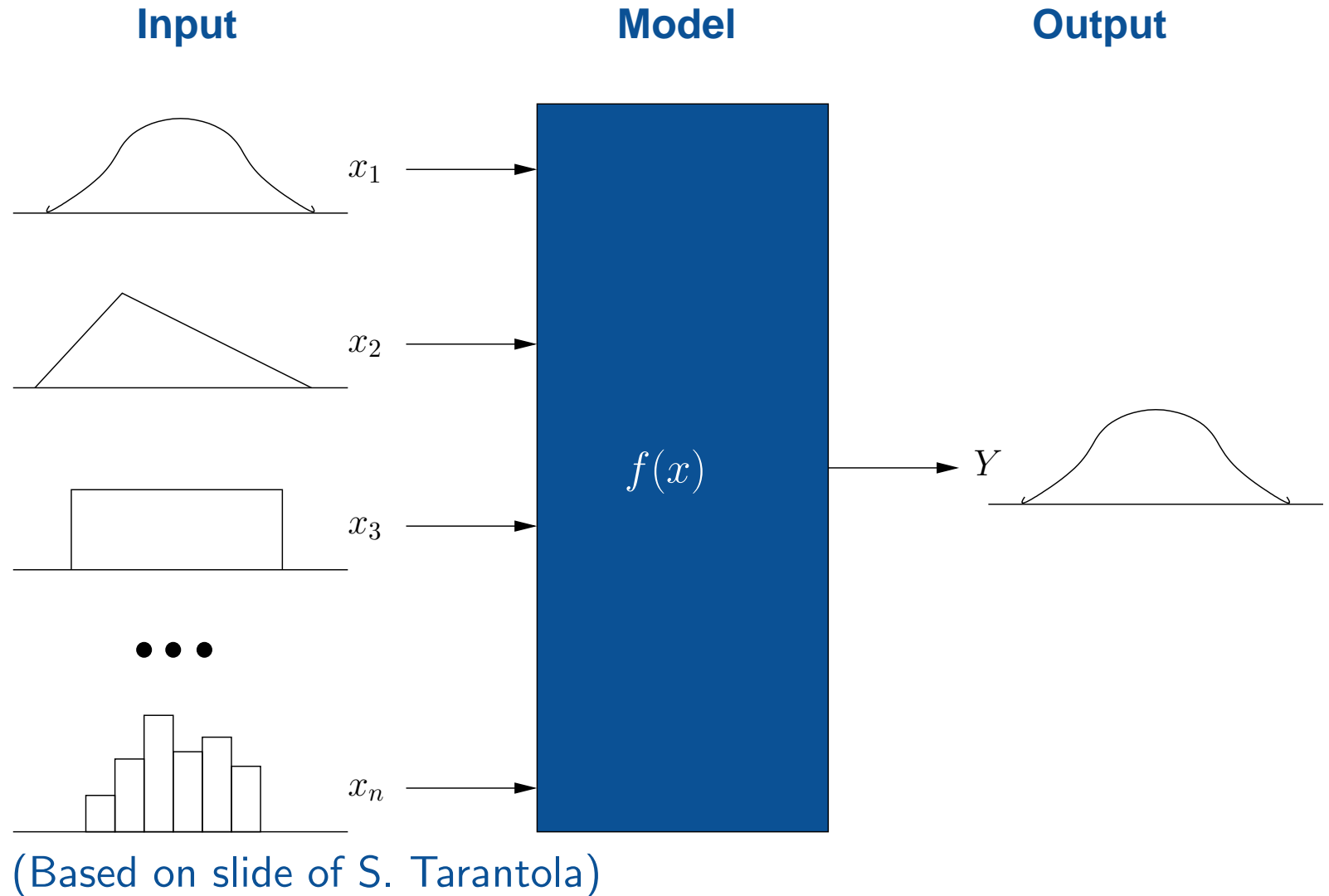
[Main Effect](#)

[Higher-Order Effects](#)

[Total Effect](#)

[Approach for SE](#)

[Summary](#)



[Introduction](#)

[Global Sensitivity Analysis](#)

[Definition](#)

**Method**

[Variance-based Methods](#)

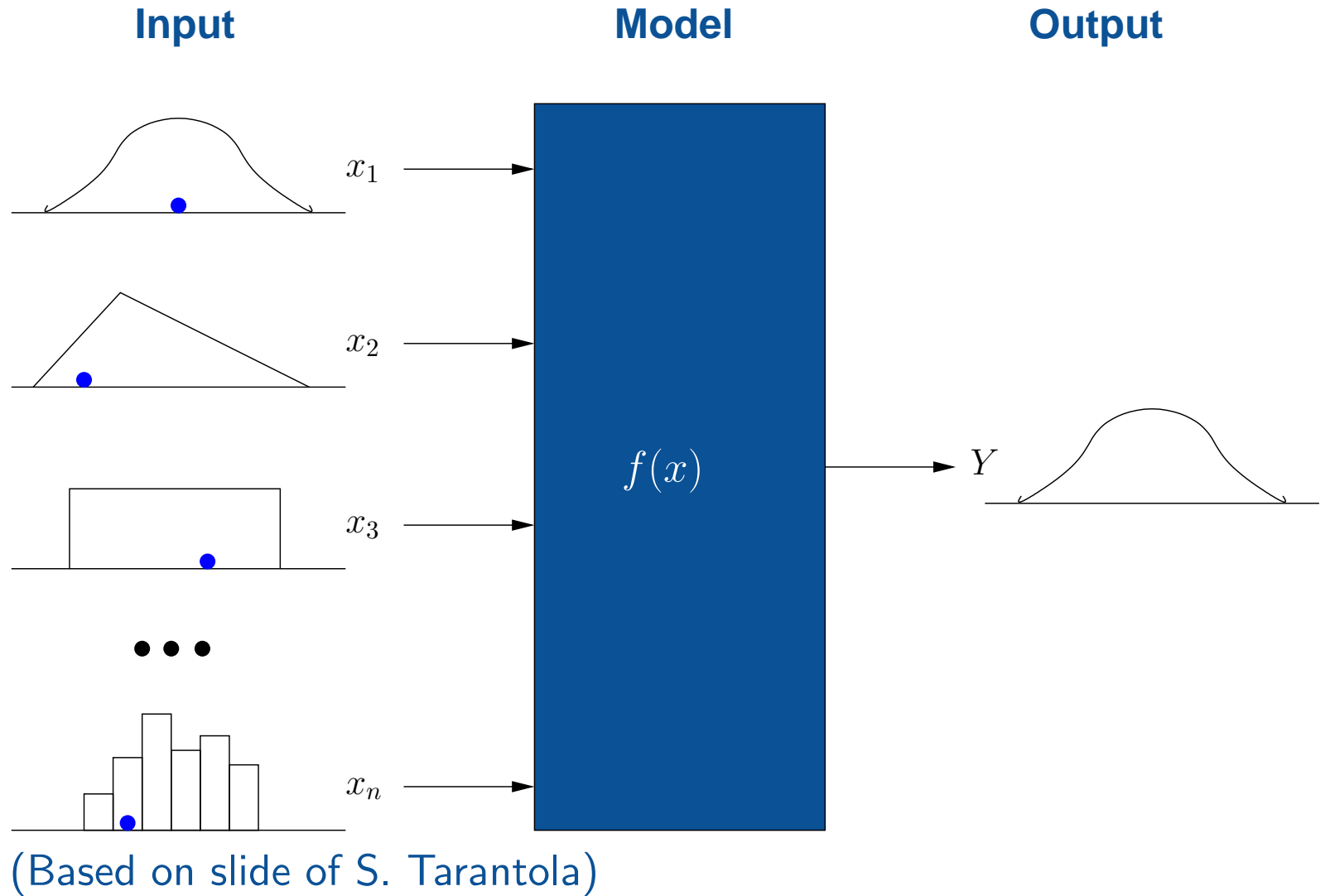
[Main Effect](#)

[Higher-Order Effects](#)

[Total Effect](#)

[Approach for SE](#)

[Summary](#)



[Introduction](#)

[Global Sensitivity Analysis](#)

[Definition](#)

**Method**

[Variance-based Methods](#)

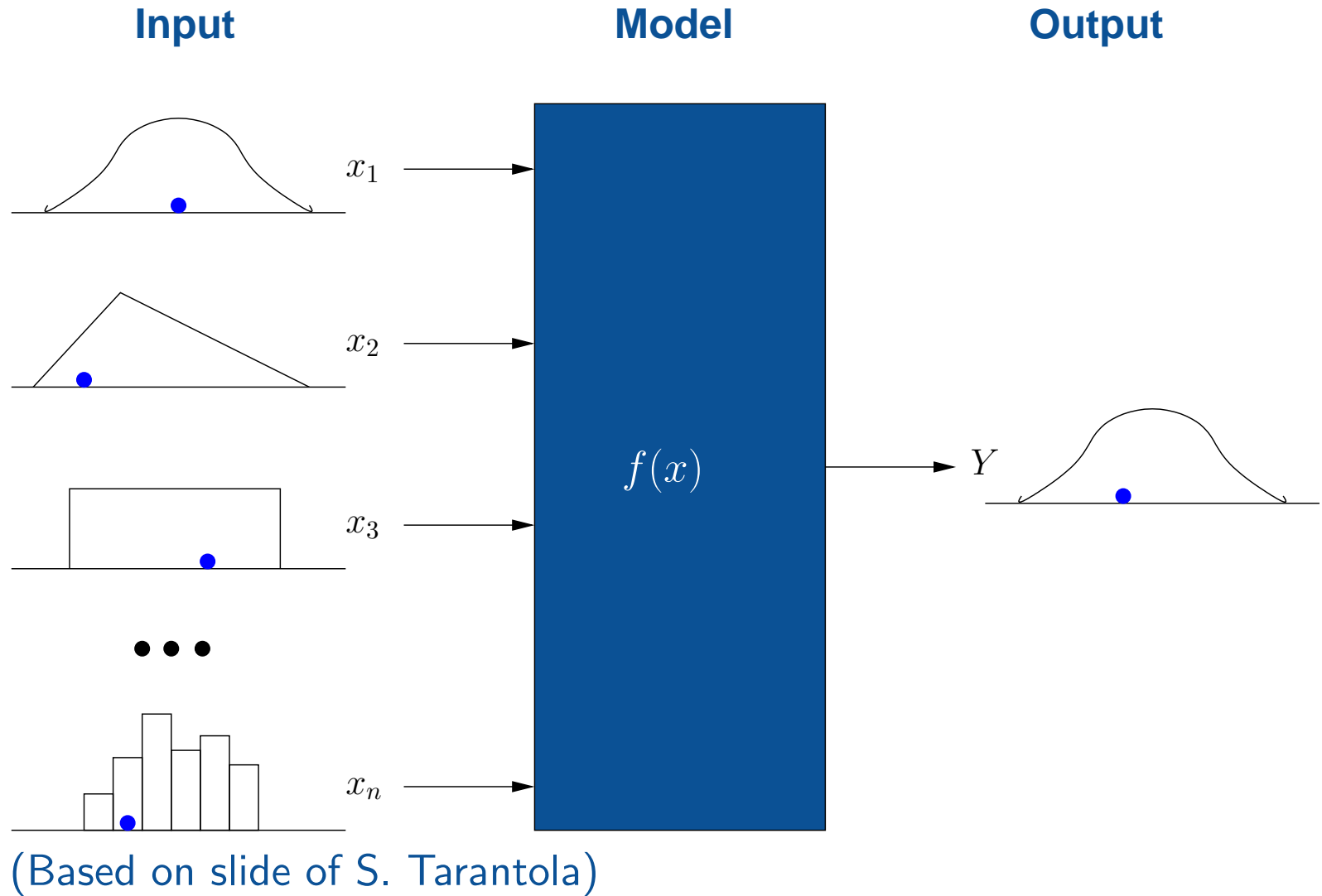
[Main Effect](#)

[Higher-Order Effects](#)

[Total Effect](#)

[Approach for SE](#)

[Summary](#)



[Introduction](#)

[Global Sensitivity Analysis](#)

[Definition](#)

**Method**

[Variance-based Methods](#)

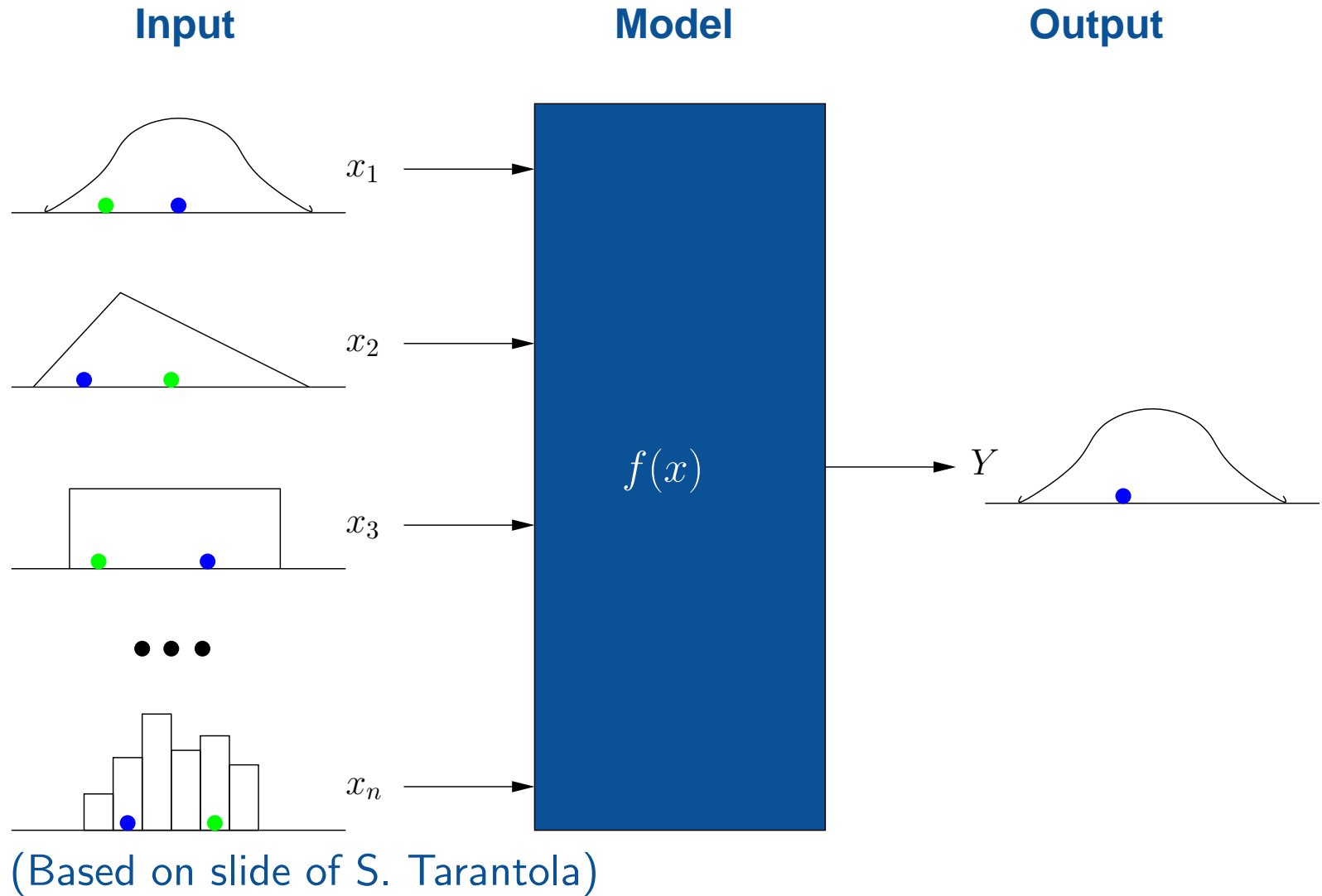
[Main Effect](#)

[Higher-Order Effects](#)

[Total Effect](#)

[Approach for SE](#)

[Summary](#)



[Introduction](#)

[Global Sensitivity Analysis](#)

[Definition](#)

**Method**

[Variance-based Methods](#)

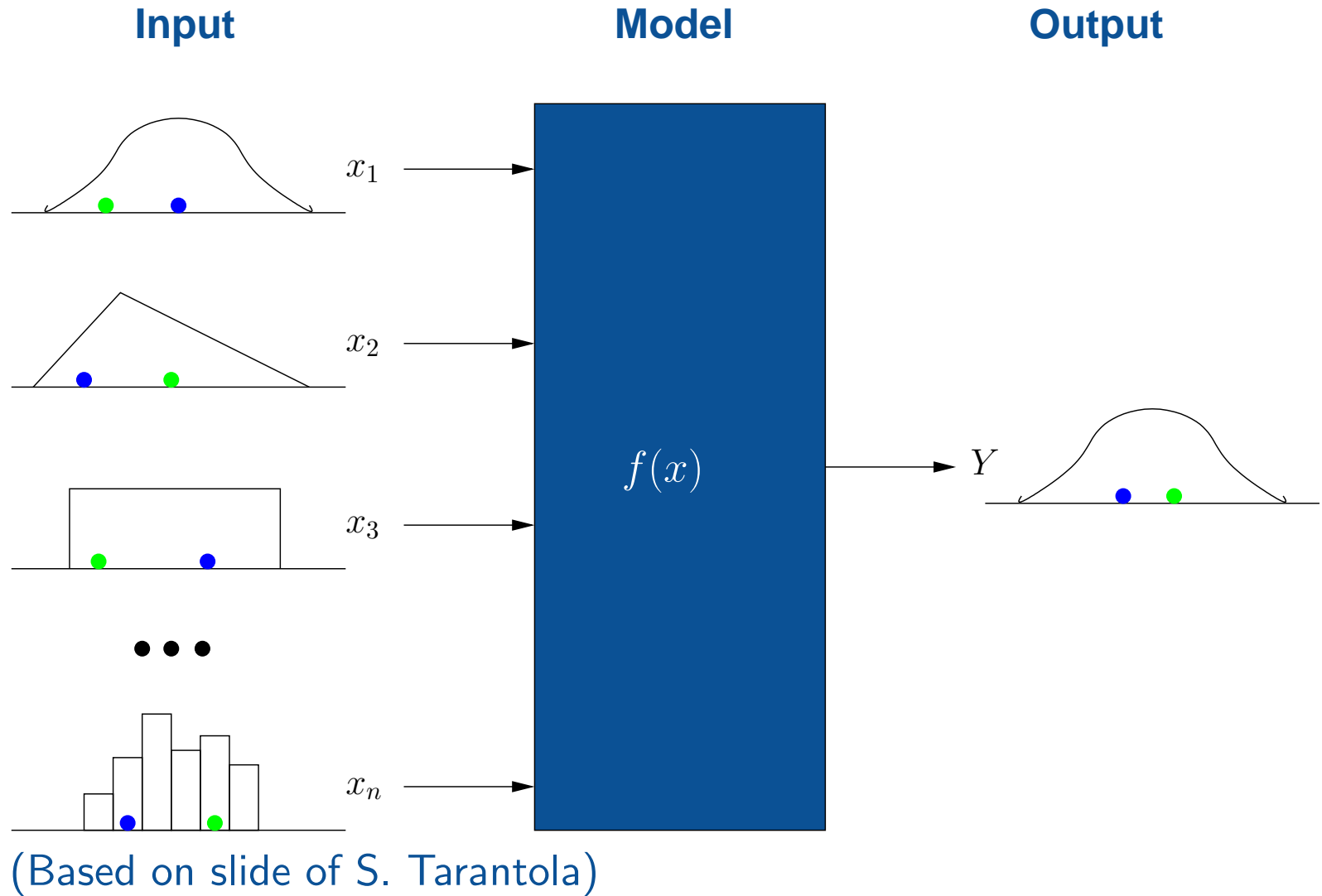
[Main Effect](#)

[Higher-Order Effects](#)

[Total Effect](#)

[Approach for SE](#)

[Summary](#)



[Introduction](#)

[Global Sensitivity Analysis](#)

[Definition](#)

**Method**

[Variance-based Methods](#)

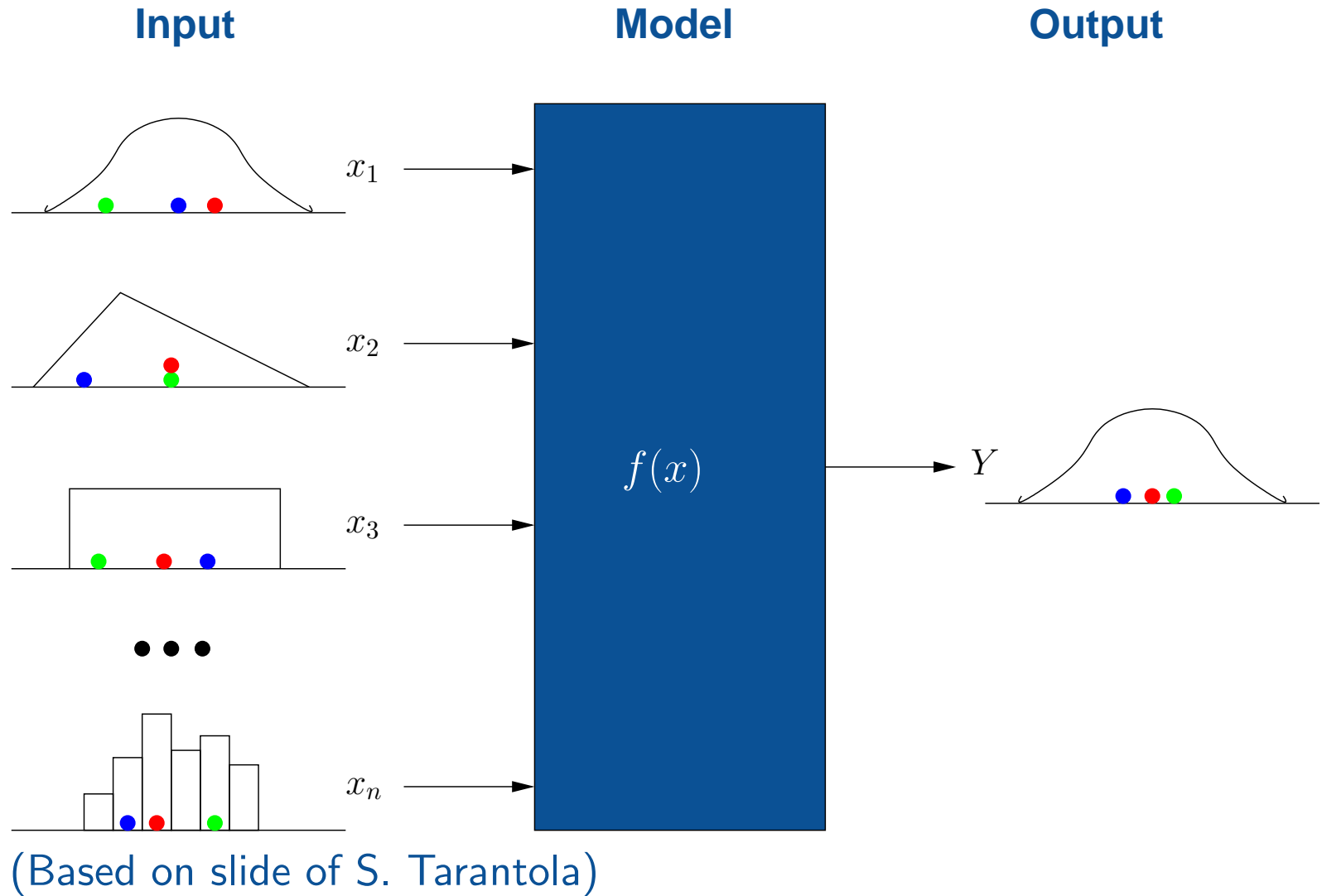
[Main Effect](#)

[Higher-Order Effects](#)

[Total Effect](#)

[Approach for SE](#)

[Summary](#)



[Introduction](#)

[Global Sensitivity Analysis](#)

[Definition](#)

**Method**

[Variance-based Methods](#)

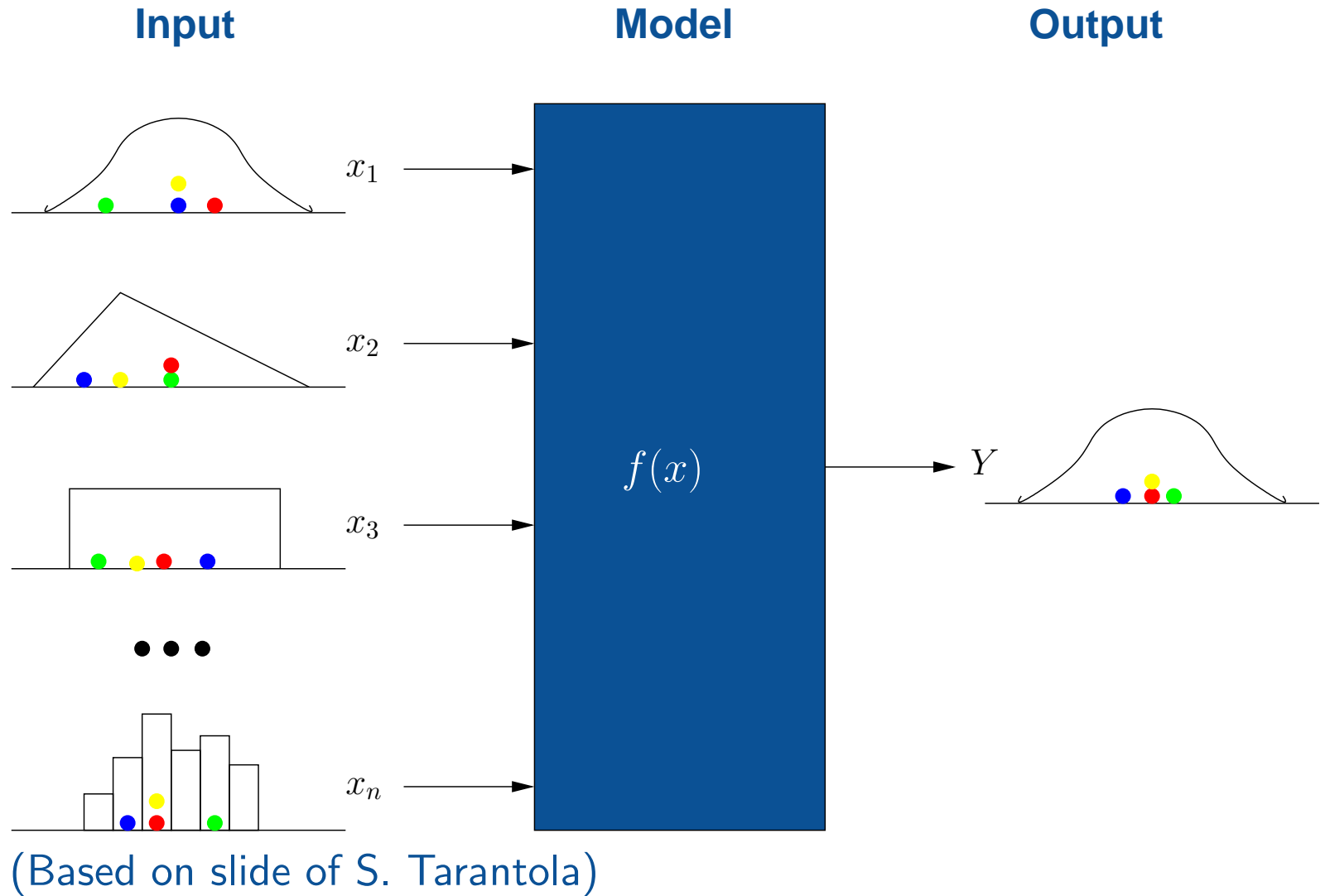
[Main Effect](#)

[Higher-Order Effects](#)

[Total Effect](#)

[Approach for SE](#)

[Summary](#)



[Introduction](#)

[Global Sensitivity Analysis](#)

[Definition](#)

**[Method](#)**

[Variance-based Methods](#)

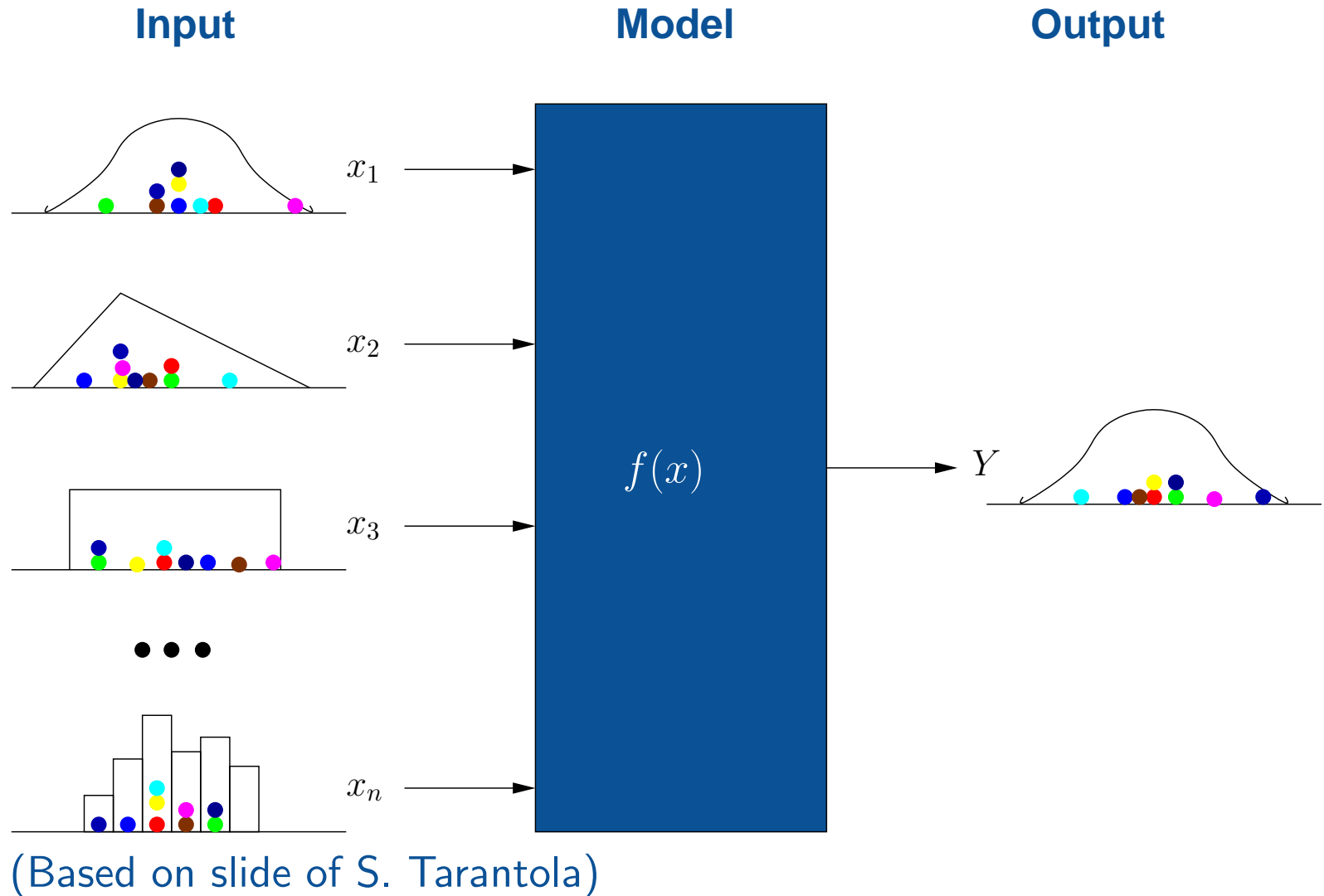
[Main Effect](#)

[Higher-Order Effects](#)

[Total Effect](#)

[Approach for SE](#)

[Summary](#)



[Introduction](#)

[Global Sensitivity Analysis](#)

[Definition](#)

**Method**

[Variance-based Methods](#)

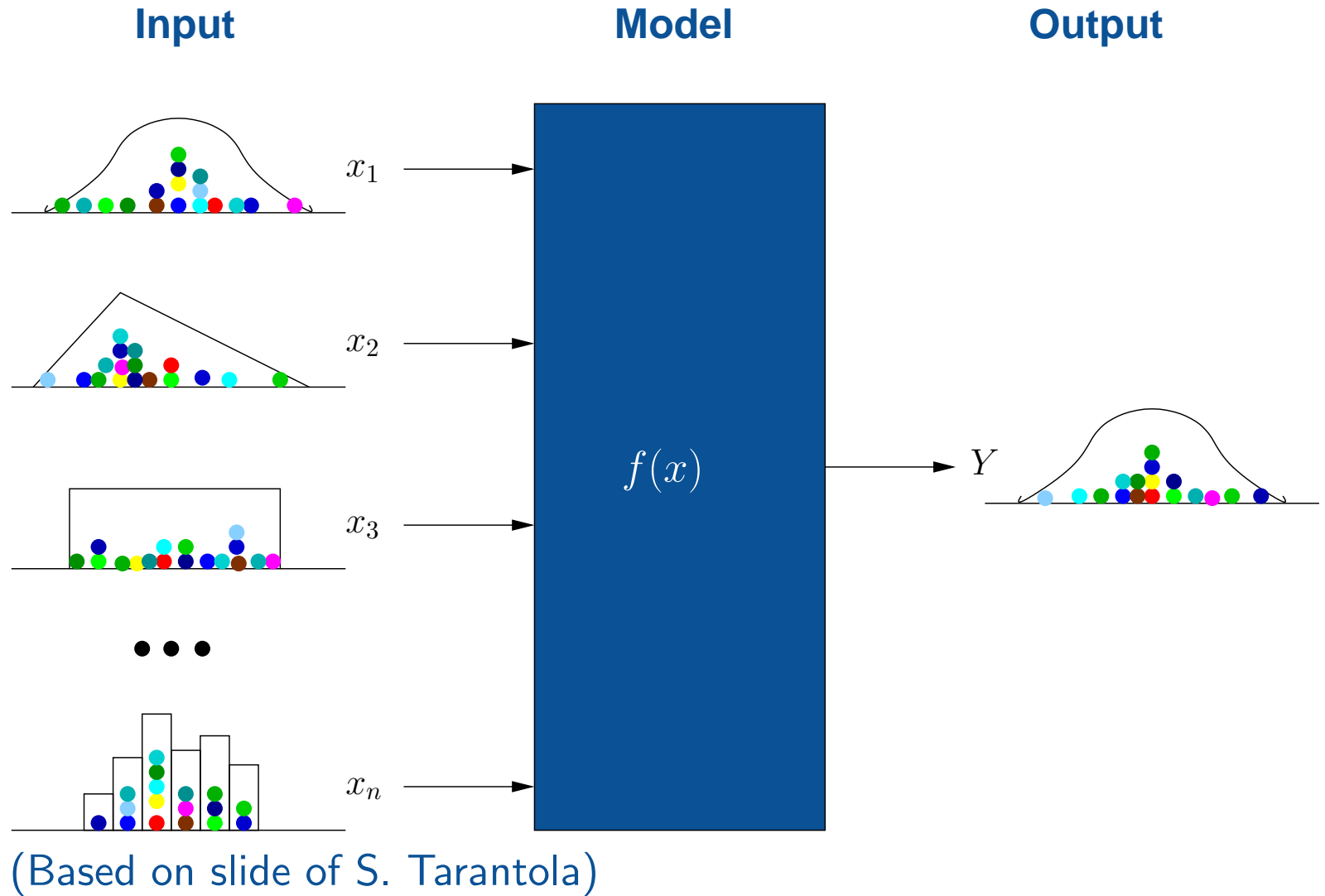
[Main Effect](#)

[Higher-Order Effects](#)

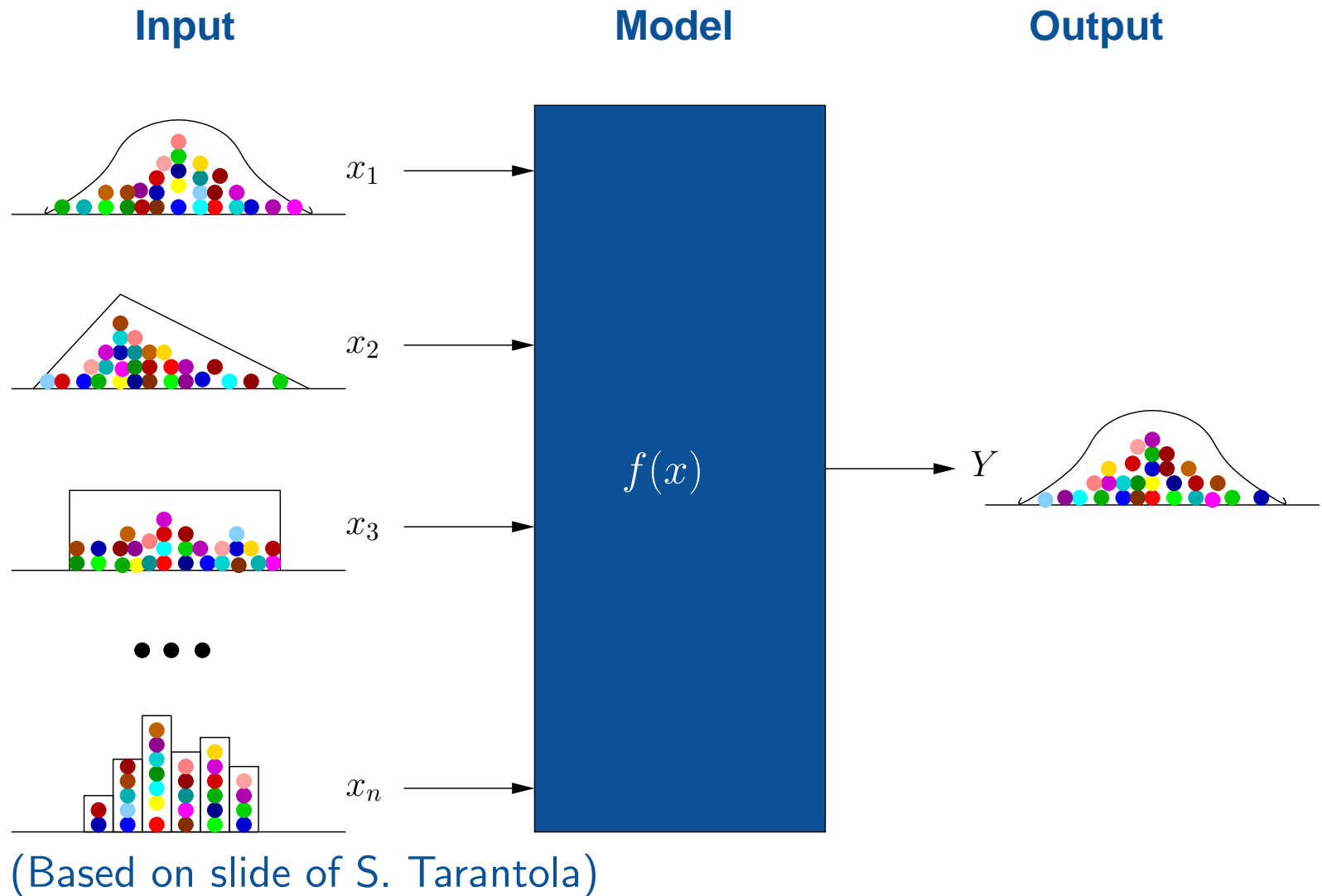
[Total Effect](#)

[Approach for SE](#)

[Summary](#)



- [Introduction](#)
- [Global Sensitivity Analysis](#)
- [Definition](#)
- [Method](#)**
- [Variance-based Methods](#)
- [Main Effect](#)
- [Higher-Order Effects](#)
- [Total Effect](#)
- [Approach for SE](#)
- [Summary](#)



[Introduction](#)

[Global Sensitivity Analysis](#)

[Definition](#)

**Method**

[Variance-based Methods](#)

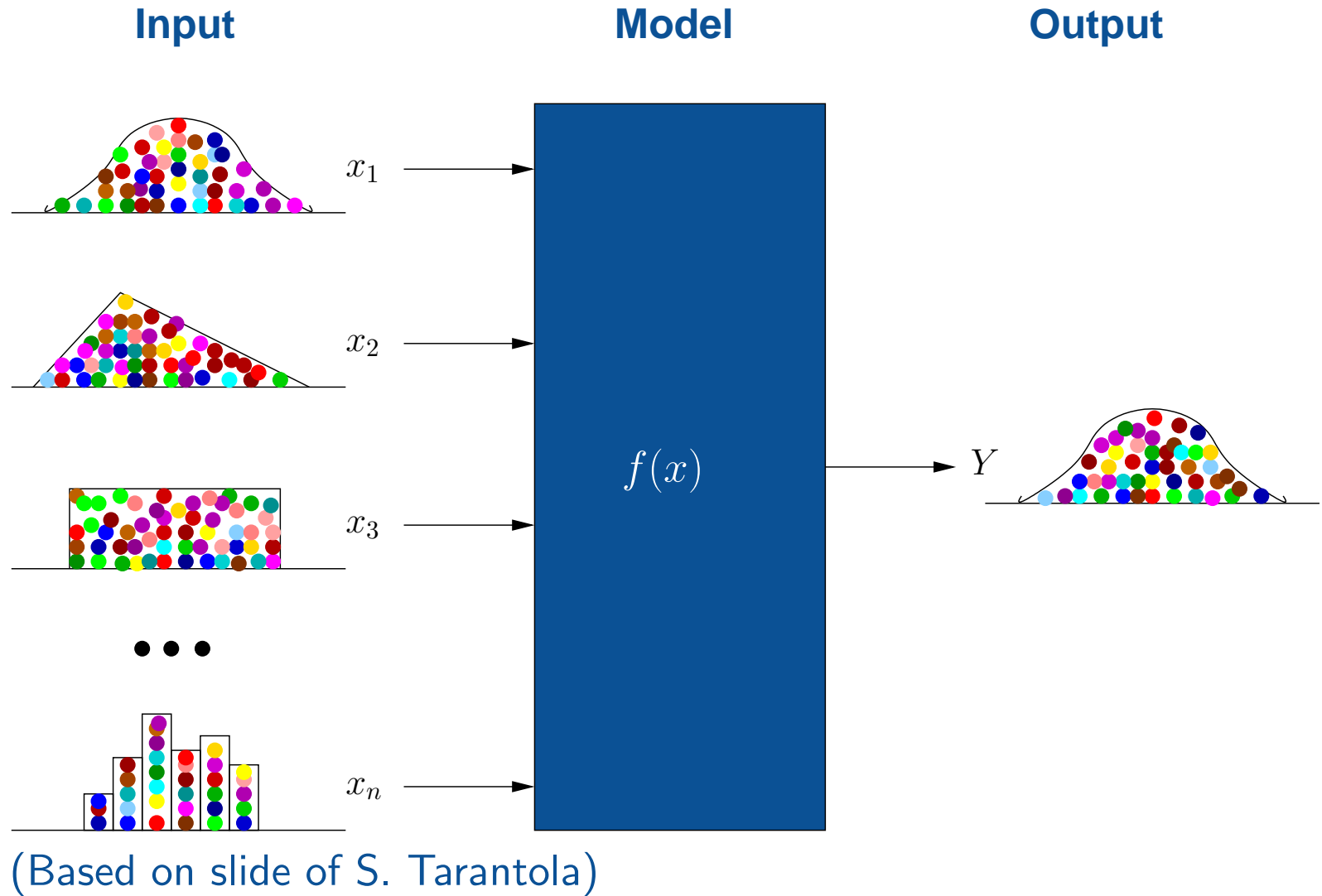
[Main Effect](#)

[Higher-Order Effects](#)

[Total Effect](#)

[Approach for SE](#)

[Summary](#)



- Global analysis usually variance-based
- “model-free”
- Analysis on the basis of **sensitivity indices**
  - ◆ *Main or first-order effect*
  - ◆ *Higher-order effects*
  - ◆ *Total effect*
- Decomposition in *main effects* and *interactions*

$$\begin{aligned}
 y = f(x) &= f_0 + \sum_{i=1}^k f_i(x_i) + \sum_i \sum_{j>i} f_{ij}(x_i, x_j) \\
 &+ \dots + f_{1,2,\dots,k}(x_1, x_2, \dots, x_k)
 \end{aligned}$$

- Global analysis usually variance-based
- “model-free”
- Analysis on the basis of **sensitivity indices**
  - ◆ *Main or first-order effect*
  - ◆ *Higher-order effects*
  - ◆ *Total effect*
- Decomposition in *main effects* and *interactions*

For example  $k = 3$ :

$$\begin{aligned}
 f(x) = & f_0 + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_{12}(x_1, x_2) \\
 & + f_{13}(x_1, x_3) + f_{23}(x_2, x_3) + f_{123}(x_1, x_2, x_3)
 \end{aligned}$$

Decomposition of the variance  $V$  of  $f(x)$

$$\text{Var}(y) = V = \sum_{i=1}^k V_i + \sum_i \sum_j V_{ij} + \sum_i \sum_j \sum_k V_{ijk} + \dots + V_{1,2,\dots,k}$$

First-order sensitivity coefficient

$$S_i = \frac{V_i}{V}$$

(Factors Priorisation)

Introduction

Global Sensitivity  
Analysis

Definition

Method  
Variance-based  
Methods

Main Effect

**Higher-Order Effects**

Total Effect

Approach for SE

Summary

## Higher-order indices analog

$$S_{i_1 \dots i_o} = V_{i_1 \dots i_o} / V$$

With  $k = 4$ :

2. order:  $S_{12}, S_{13}, S_{14}, S_{23}, S_{24}, S_{34}$

3. order:  $S_{123}, S_{124}, S_{134}, S_{234}$

4. order:  $S_{1234}$

- [Introduction](#)
- [Global Sensitivity Analysis](#)
- [Definition](#)
- [Method](#)
- [Variance-based Methods](#)
- [Main Effect](#)
- [Higher-Order Effects](#)
- [Total Effect](#)**
- [Approach for SE](#)
- [Summary](#)

- Sum of the first- and higher-order effects of  $x_i$
- Removal of the non-relevant parts

$$S_1 + S_2 + S_3 + S_4 + S_{12} + S_{13} + S_{14} + S_{23} + S_{24} + S_{34} + S_{123} + S_{124} + S_{134} + S_{234} + S_{1234} = 1$$

(Factors Fixing)

- [Introduction](#)
- [Global Sensitivity Analysis](#)
- [Definition](#)
- [Method](#)
- [Variance-based Methods](#)
- [Main Effect](#)
- [Higher-Order Effects](#)
- [Total Effect](#)**
- [Approach for SE](#)
- [Summary](#)

- Sum of the first- and higher-order effects of  $x_i$
- Removal of the non-relevant parts

$$S_1 + \cancel{S_2} + S_3 + S_4 + S_{12} + S_{13} + S_{14} + S_{23} + S_{24} + S_{34} + S_{123} + S_{124} + S_{134} + S_{234} + S_{1234} = 1$$

(Factors Fixing)

Introduction

Global Sensitivity Analysis

Definition

Method

Variance-based

Methods

Main Effect

Higher-Order Effects

**Total Effect**

Approach for SE

Summary

- Sum of the first- and higher-order effects of  $x_i$
- Removal of the non-relevant parts

$$S_1 + \cancel{S_2} + \cancel{S_3} + S_4 + S_{12} + S_{13} + S_{14} + S_{23} + S_{24} + S_{34} + S_{123} + S_{124} + S_{134} + S_{234} + S_{1234} = 1$$

(Factors Fixing)

- Sum of the first- and higher-order effects of  $x_i$
- Removal of the non-relevant parts

$$S_1 + \cancel{S_2} + \cancel{S_3} + \cancel{S_4} + S_{12} + S_{13} + S_{14} + S_{23} + S_{24} + S_{34} + S_{123} + S_{124} + S_{134} + S_{234} + S_{1234} = 1$$

(Factors Fixing)

- Sum of the first- and higher-order effects of  $x_i$
- Removal of the non-relevant parts

$$S_1 + \cancel{S_2} + \cancel{S_3} + \cancel{S_4} + S_{12} + S_{13} + S_{14} + \cancel{S_{23}} + \cancel{S_{24}} + \cancel{S_{34}} + S_{123} + S_{124} + S_{134} + S_{234} + S_{1234} = 1$$

(Factors Fixing)

- [Introduction](#)
- [Global Sensitivity Analysis](#)
- [Definition](#)
- [Method](#)
- [Variance-based Methods](#)
- [Main Effect](#)
- [Higher-Order Effects](#)
- [Total Effect](#)**
- [Approach for SE](#)
- [Summary](#)

- Sum of the first- and higher-order effects of  $x_i$
- Removal of the non-relevant parts

$$S_1 + \cancel{S_2} + \cancel{S_3} + \cancel{S_4} + S_{12} + S_{13} + S_{14} + \cancel{S_{23}} + \cancel{S_{24}} + \cancel{S_{34}} + S_{123} + S_{124} + S_{134} + \cancel{S_{234}} + S_{1234} = 1$$

(Factors Fixing)

- [Introduction](#)
- [Global Sensitivity Analysis](#)
- [Definition](#)
- [Method](#)
- [Variance-based Methods](#)
- [Main Effect](#)
- [Higher-Order Effects](#)
- [Total Effect](#)**
- [Approach for SE](#)
- [Summary](#)

- Sum of the first- and higher-order effects of  $x_i$
- Removal of the non-relevant parts

$$S_1 + \cancel{S_2} + \cancel{S_3} + \cancel{S_4} + S_{12} + S_{13} + S_{14} + \cancel{S_{23}} + \cancel{S_{24}} + \cancel{S_{34}} + S_{123} + S_{124} + S_{134} + \cancel{S_{234}} + S_{1234} = 1$$

$$S_{T1} = S_1 + S_{12} + S_{13} + S_{14} + S_{123} + S_{124} + S_{134} + S_{1234}$$

(Factors Fixing)

- Introduction
- Global Sensitivity Analysis
- Definition
- Method
- Variance-based Methods
- Main Effect
- Higher-Order Effects
- Total Effect**
- Approach for SE
- Summary

- Sum of the first- and higher-order effects of  $x_i$
- Removal of the non-relevant parts

$$S_1 + \cancel{S_2} + \cancel{S_3} + \cancel{S_4} + S_{12} + S_{13} + S_{14} + \cancel{S_{23}} + \cancel{S_{24}} + \cancel{S_{34}} + S_{123} + S_{124} + S_{134} + \cancel{S_{234}} + S_{1234} = 1$$

$$S_{T1} = S_1 + S_{12} + S_{13} + S_{14} + S_{123} + S_{124} + S_{134} + S_{1234}$$

$$S_{T2} = S_2 + S_{12} + S_{23} + S_{24} + S_{123} + S_{124} + S_{234} + S_{1234}$$

(Factors Fixing)

Introduction

Global Sensitivity  
Analysis

**Approach for SE**

Overview of the  
Approach

Determining  
Distributions

Example:  
Distributions  
Visualising Using  
Scatterplots

$p_1$  and  $\mu$

$d$  and  $\mu$

$t$  and  $\mu$

$inf$  and  $\mu$

Applying Global SA  
Indices

Summary

# Approach for SE

Introduction

Global Sensitivity Analysis

Approach for SE

Overview of the Approach

Determining Distributions

Example: Distributions Visualising Using Scatterplots

$p_1$  and  $\mu$

$d$  and  $\mu$

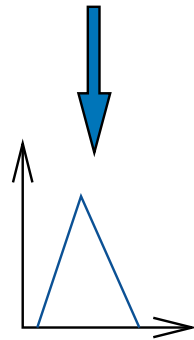
$t$  and  $\mu$

$inf$  and  $\mu$

Applying Global SA Indices

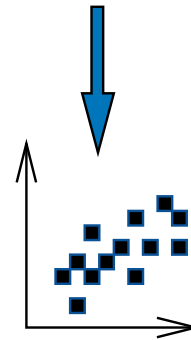
Summary

Determining distributions



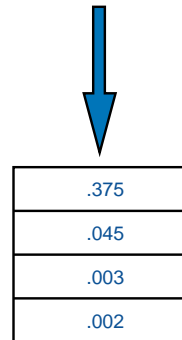
Input distributions

Visualising using scatterplots



Scatterplots

Global sensitivity analyses



Sensitivity indices

Introduction

Global Sensitivity  
Analysis

Approach for SE  
Overview of the  
Approach

**Determining  
Distributions**

Example:  
Distributions  
Visualising Using  
Scatterplots

$p_1$  and  $\mu$

$d$  and  $\mu$

$t$  and  $\mu$

$inf$  and  $\mu$

Applying Global SA  
Indices

Summary

- Distribution of the values of the input parameters
- Derived from
  - ◆ scientific literature
  - ◆ physical boundaries
  - ◆ expert opinion
  - ◆ surveys
  - ◆ experiments

Introduction

Global Sensitivity Analysis

Approach for SE

Overview of the Approach

Determining Distributions

**Example: Distributions**

Visualising Using Scatterplots

$p_1$  and  $\mu$

$d$  and  $\mu$

$t$  and  $\mu$

$inf$  and  $\mu$

Applying Global SA Indices

Summary

- Based on expert opinion
- General parameter
  - ◆  $p_1$  uniformly distributed between 0 and 1
  - ◆  $d$  uniformly distributed between 0.9 and 1
- Dependent on application
  - ◆  $inf$  e.g.. uniformly distributed between 50 and 500
  - ◆ For  $t$  e.g.  $t \sim \mathcal{N}(1000, 200)$

Introduction

Global Sensitivity  
Analysis

Approach for SE

Overview of the  
Approach

Determining  
Distributions  
Example:  
Distributions

Visualising Using  
Scatterplots

$p_1$  and  $\mu$

$d$  and  $\mu$

$t$  and  $\mu$

$inf$  and  $\mu$

Applying Global SA  
Indices

Summary

- Sampling using the distributions
- Pairwise relationship between factors
- Detection of errors in the model
- Indications of influences

Introduction

Global Sensitivity Analysis

Approach for SE

Overview of the Approach

Determining Distributions

Example: Distributions Visualising Using Scatterplots

**$p_1$  and  $\mu$**

$d$  and  $\mu$

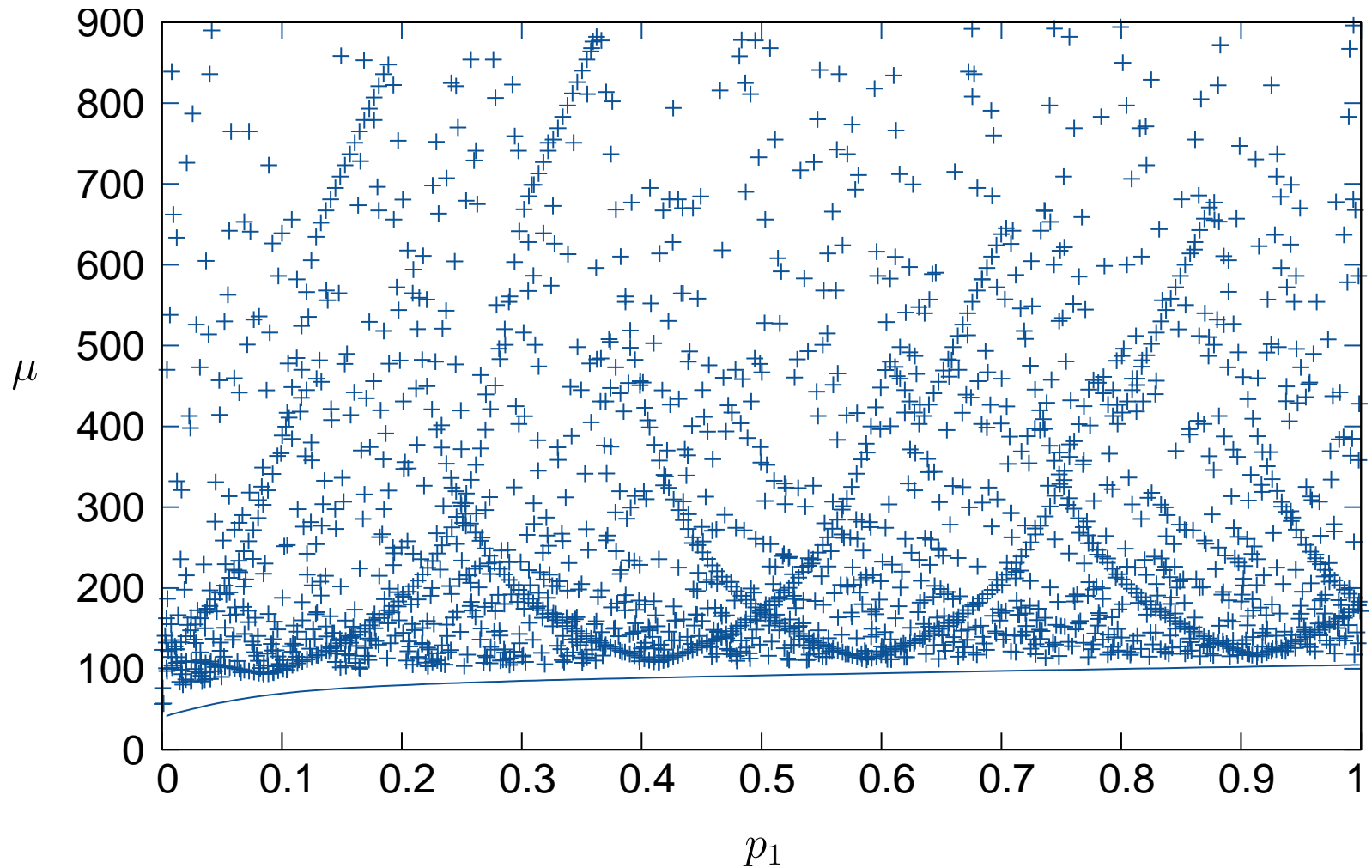
$t$  and  $\mu$

$inf$  and  $\mu$

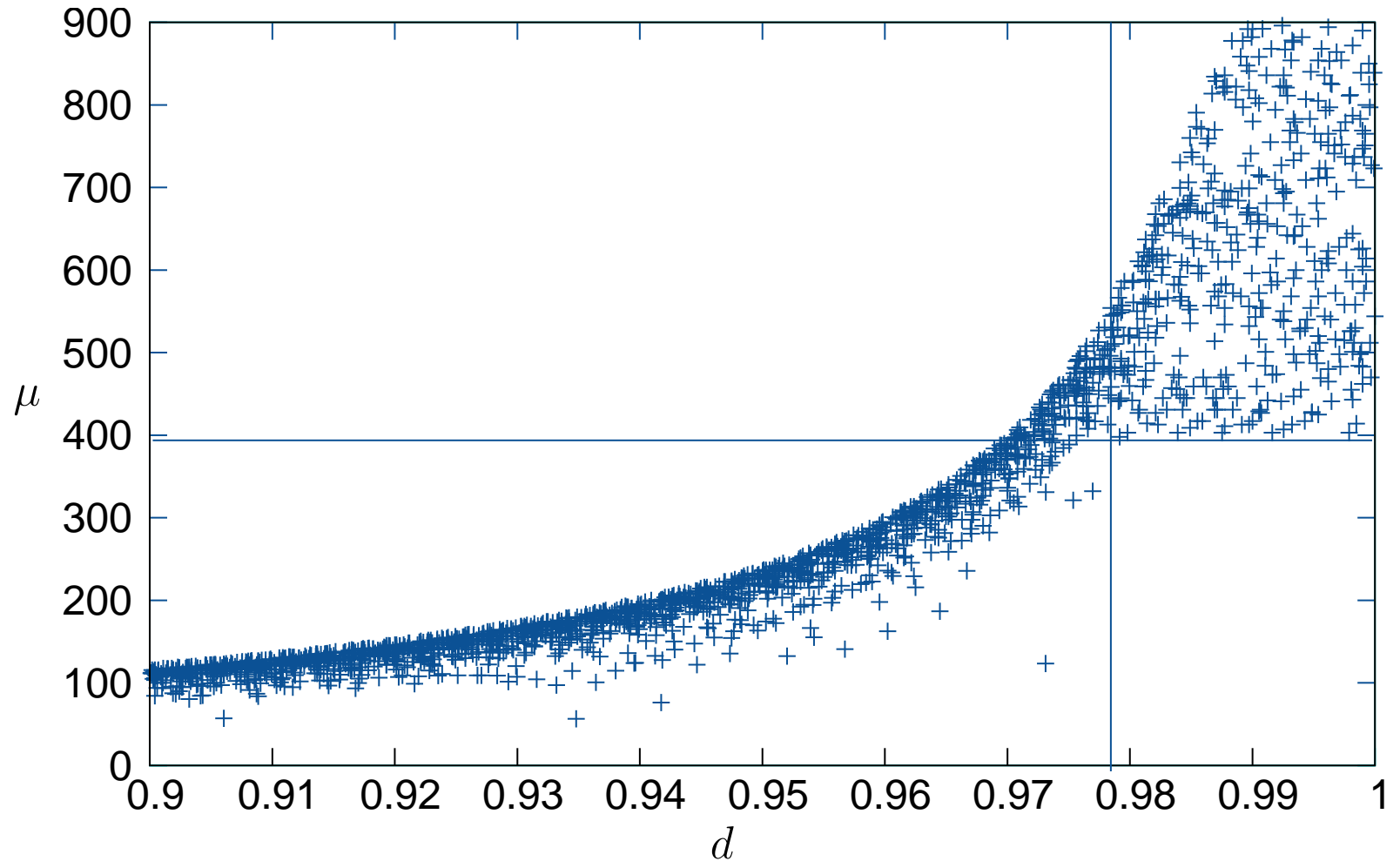
Applying Global SA

Indices

Summary



- [Introduction](#)
- [Global Sensitivity Analysis](#)
- [Approach for SE](#)
- [Overview of the Approach](#)
- [Determining Distributions](#)
- [Example: Distributions](#)
- [Visualising Using Scatterplots](#)
- [p<sub>1</sub> and  \$\mu\$](#)
- [d and  \$\mu\$](#)
- [t and  \$\mu\$](#)
- [inf and  \$\mu\$](#)
- [Applying Global SA Indices](#)
- [Summary](#)



[Introduction](#)

[Global Sensitivity Analysis](#)

[Approach for SE](#)

[Overview of the Approach](#)

[Determining Distributions](#)

[Example: Distributions Visualising Using Scatterplots](#)

[p<sub>1</sub> and  \$\mu\$](#)

[d and  \$\mu\$](#)

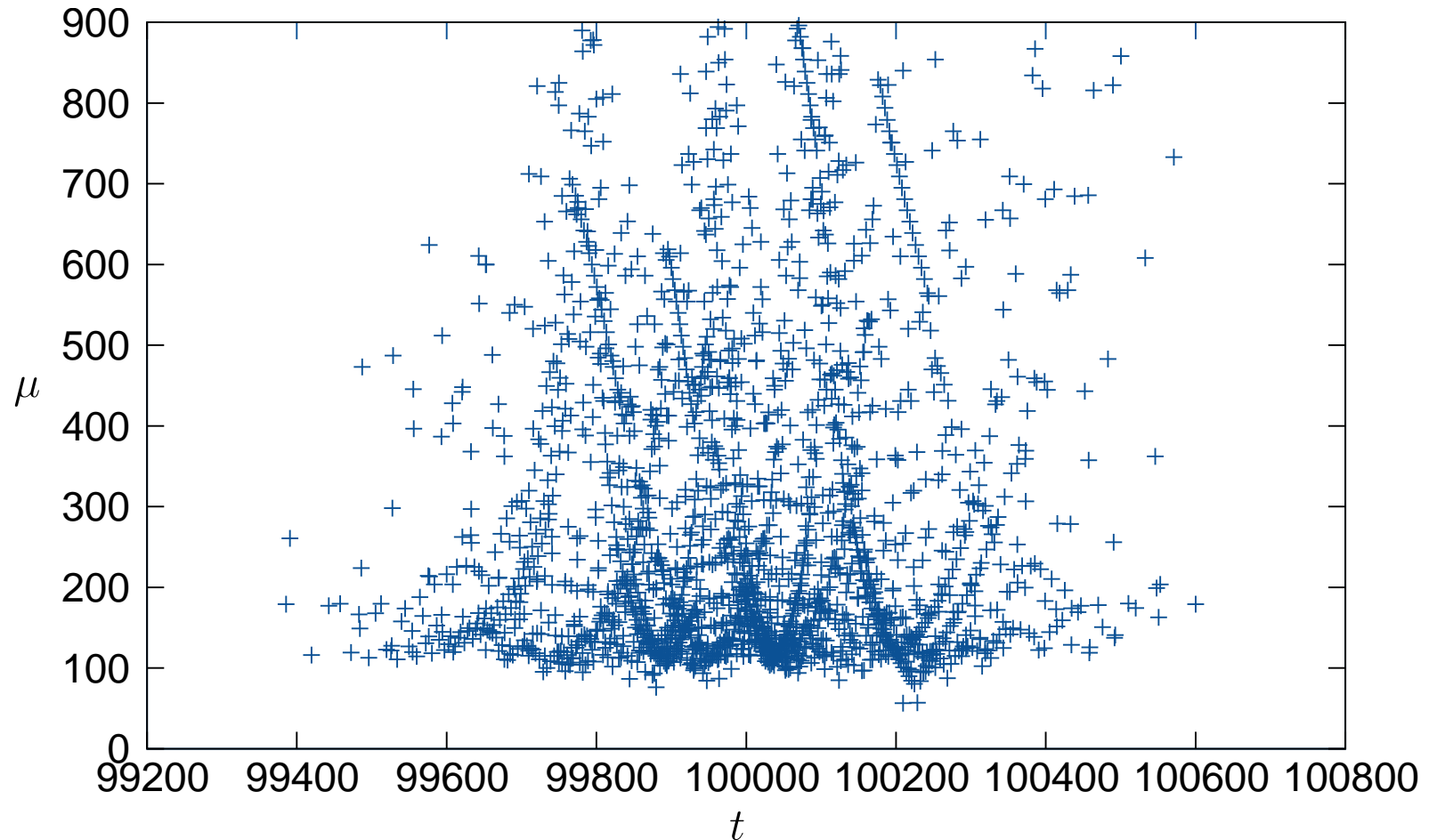
**[t and  \$\mu\$](#)**

[inf and  \$\mu\$](#)

[Applying Global SA](#)

[Indices](#)

[Summary](#)



Introduction

Global Sensitivity Analysis

Approach for SE

Overview of the Approach

Determining Distributions

Example: Distributions Visualising Using Scatterplots

$p_1$  and  $\mu$

$d$  and  $\mu$

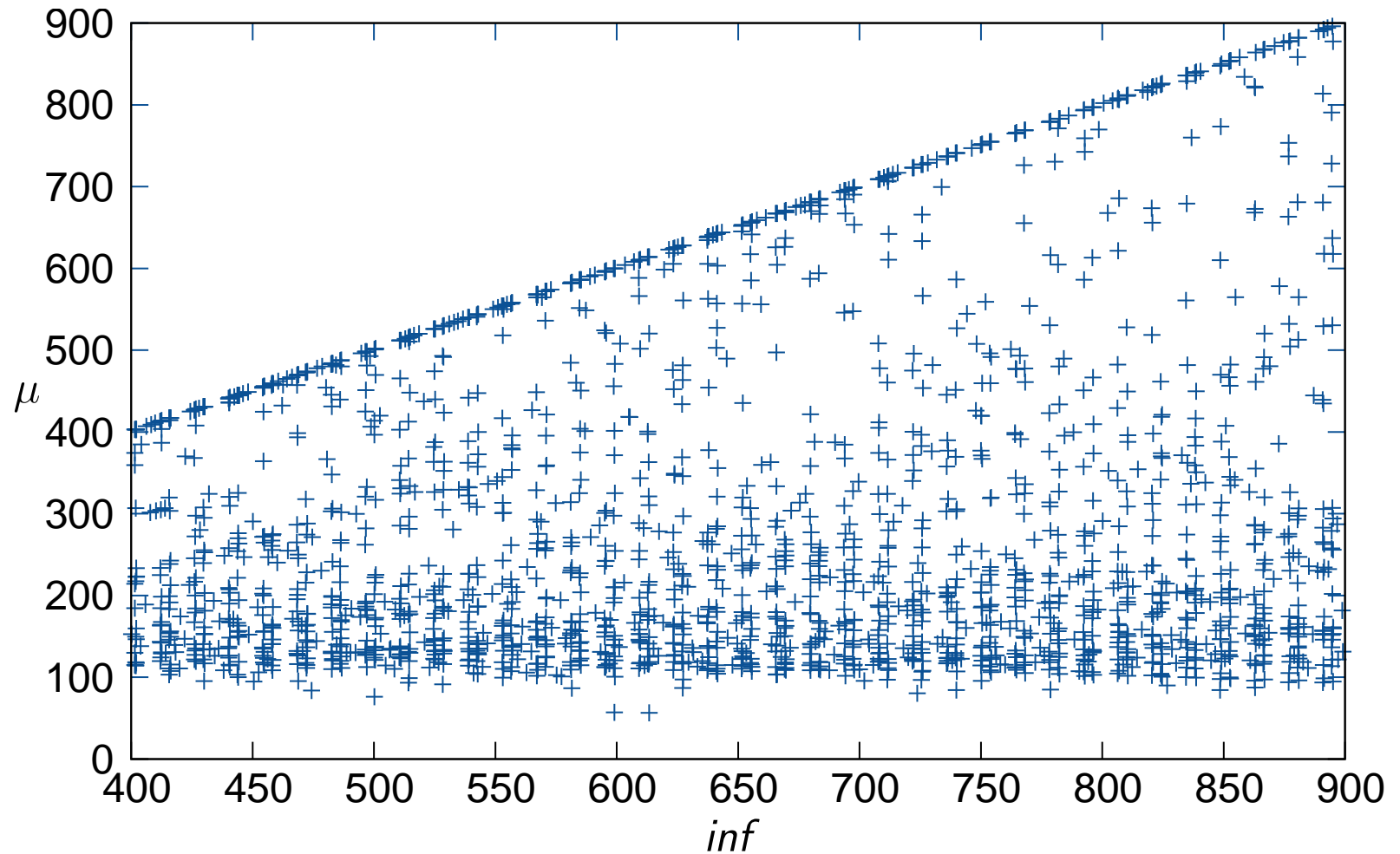
$t$  and  $\mu$

**inf and  $\mu$**

Applying Global SA

Indices

Summary



Introduction

Global Sensitivity Analysis

Approach for SE

Overview of the Approach

Determining Distributions  
Example:

Distributions  
Visualising Using Scatterplots

$p_1$  and  $\mu$

$d$  and  $\mu$

$t$  and  $\mu$

$inf$  and  $\mu$

**Applying Global SA**

Indices

Summary

- Various methods for calculating indices
- FAST (Fourier Amplitude Sensitivity Test) gives quantitative results
- Uses sampled data
- Tool-support: Simlab

Introduction

Global Sensitivity Analysis

Approach for SE

Overview of the Approach

Determining Distributions  
Example:

Distributions  
Visualising Using Scatterplots

$p_1$  and  $\mu$

$d$  and  $\mu$

$t$  and  $\mu$

$inf$  and  $\mu$

Applying Global SA

**Indizes**

Summary

## Main Effect

$t$	6.88e-5
$p_1$	0.0090
$d$	0.9026
$inf$	0.0158

[Introduction](#)

[Global Sensitivity Analysis](#)

[Approach for SE](#)

[Overview of the Approach](#)

[Determining Distributions Example:](#)

[Distributions Visualising Using Scatterplots](#)

[p<sub>1</sub> and μ](#)

[d and μ](#)

[t and μ](#)

[inf and μ](#)

[Applying Global SA](#)

**Indizes**

[Summary](#)

## Main Effect

<i>t</i>	6.88e-5
<i>p</i> <sub>1</sub>	0.0090
<i>d</i>	0.9026
<i>inf</i>	0.0158

## Total Effect

<i>t</i>	0.005788
<i>p</i> <sub>1</sub>	0.022365
<i>d</i>	0.975155
<i>inf</i>	0.086735

[Introduction](#)

[Global Sensitivity  
Analysis](#)

[Approach for SE](#)

**Summary**

[Conclusions](#)

# Summary

Introduction

Global Sensitivity  
Analysis

Approach for SE

Summary

**Conclusions**

- Sensitivity analysis useful tool for predictor models
- SA can help to
  - ◆ find errors in the models
  - ◆ simplify the models
  - ◆ identify interactions between input parameters
  - ◆ identify parameters that should be investigated more
  - ◆ get more robust predictions
- Experience with
  - ◆ reliability model
  - ◆ QA economics model
  - ◆ process model
  - ◆ expert system for IT tools
- Good tool-support available (Simlab:  
<http://simlab.jrc.cec.eu.int/>)