Máster Universitario en Administración y Dirección de Empresas
Full Time MBA

Quantitative methods for decision making

Professor Andrea Saltelli

upf. BARCELONA SCHOOL OF MANAGEMENT

Elements of quantification for decision making with emphasis on operation research

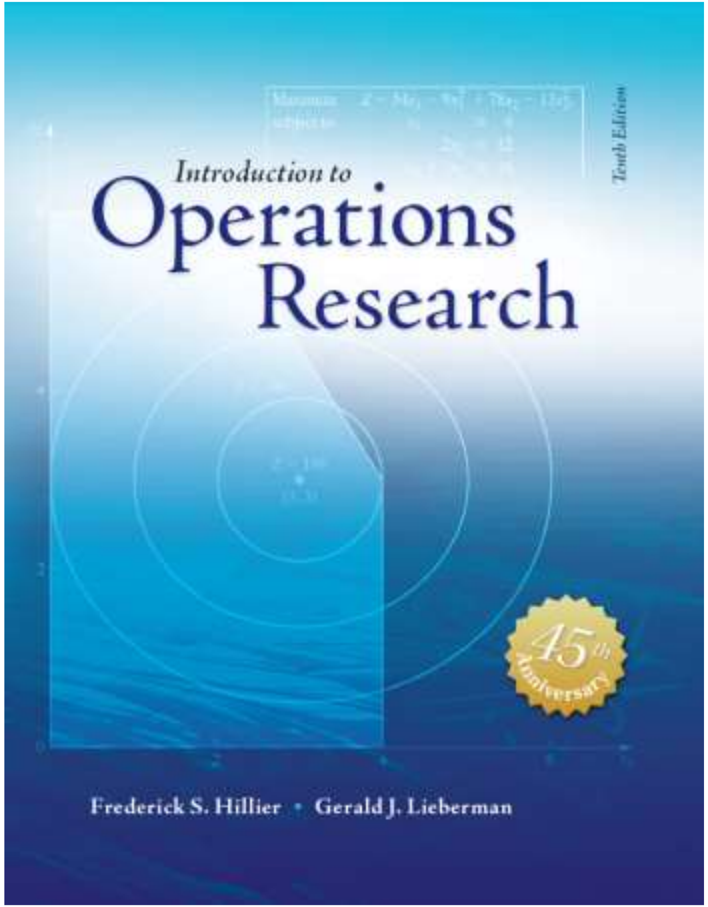
BARCELONA SCHOOL OF MANAGEMENT

In this set of slides:

BARCELONA SCHOOL OF MANAGEMENT

Where to find this book:

https://www.dropbox.com/sh/ddd48a8jguinbcf/AABF0s4eh1lPLVxdx0pes-Ofa?dl=0&preview=Introduction+ to+ Operations+ Research+ -
+ Frederick+ S.+ Hillier.pdf



upf. BARCELONA SCHOOL OF MANAGEMENT

9.

Introduction to uncertainty and sensitivity analysis

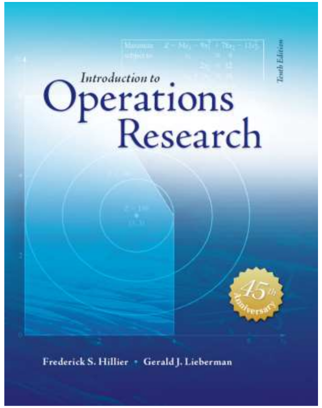
Sensitivity analysis as performed in linear programming relies of local, one– at-a-time (OAT) methods. This vision can be complemented by vision of SA coming by other disciplines. From shadow prices to global sensitivity analysis. Hillier 2014, chapter 7.

Linear programming and sensitivity analysis

Linear programming viewpoint: testing which parameter, when changed in isolation, lead to a change in the optimal solution

Global SA viewpoints: explore the distribution of the optimal solution when all uncertain coefficients are allowed to vary over their plausible range

“However, the situation is quite different when dealing with the larger linear programming problems that are typically encountered in practice. For example, **Selected Reference 1** at the end of the chapter describes what happened when dealing with the problems in a library of 94 large linear programming problems (hundreds or thousands of constraints and variables).



It was assumed that the parameters could be randomly in error by as much as 0.01 percent. Even with such tiny errors throughout the model, the optimal solution was found to be infeasible in 13 of these problems and badly so for 6 of the problems”

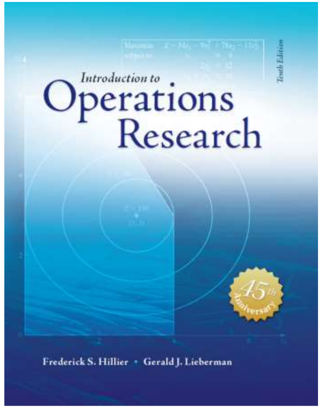
Selected reference 1 of Hillier is: Ben-Tal, A., L. El Ghaoui, and A. Nemirovski: Robust Optimization, Princeton University Press, Princeton, NJ, 2009.

upf. BARCELONA SCHOOL OF MANAGEMENT

We should not be surprised that the sensitivity analysis practiced in linear programming is linear!

Yet so much can be lost by neglecting that part of the uncertainty that escapes linearity that this needs to be mentioned in a course of Quantitative methods for decision making

The advantages of understating global methods for uncertainty and sensitivity analysis are very large, including the possibility to test to flexibility of managerial decision when ‘all the rest’ is varying as well



Linear programming: Compute shadow prices

The shadow price for resource i (denoted by y_i^*) measures the marginal value of this resource, i.e., the rate at which Z could be increased by (slightly) increasing the amount of this resource (b_i) being made available

TABLE 3.1 Data for the Wyndor Glass Co. problem

Plant	Production Time per Batch, Hours — Product 1	Production Time per Batch, Hours — Product 2	Production Time Available per Week, Hours
1	1	0	4
2	0	2	12
3	3	2	18
Profit per batch	\$3,000	\$5,000	

→

$$x_1 \leq 4$$
$$2x_2 \leq 12$$
$$3x_1 + 2x_2 \leq 18$$

BARCELONA SCHOOL OF MANAGEMENT

In our classic example the structural constraint b_2 for decision variable x_2 was $2x_2 \leq 12$; imagine we change 12 into 13 i.e. we are willing to allow one more hour in plant two

$$\begin{aligned} x_1 &\leq 4 \\ 2x_2 &\leq 12 \\ 3x_1 + 2x_2 &\leq 18 \end{aligned} \longrightarrow \begin{aligned} x_1 &\leq 4 \\ 2x_2 &\leq 13 \\ 3x_1 + 2x_2 &\leq 18 \end{aligned}$$

TABLE 3.1 Data for the Wyndor Glass Co. problem

	Production Time per Batch, Hours		
	Product		
Plant	**1**	**2**	**Production Time Available per Week, Hours**
1	1	0	4
2	0	2	12
3	3	2	18
Profit per batch	$3,000	$5,000	

upf. BARCELONA SCHOOL OF MANAGEMENT

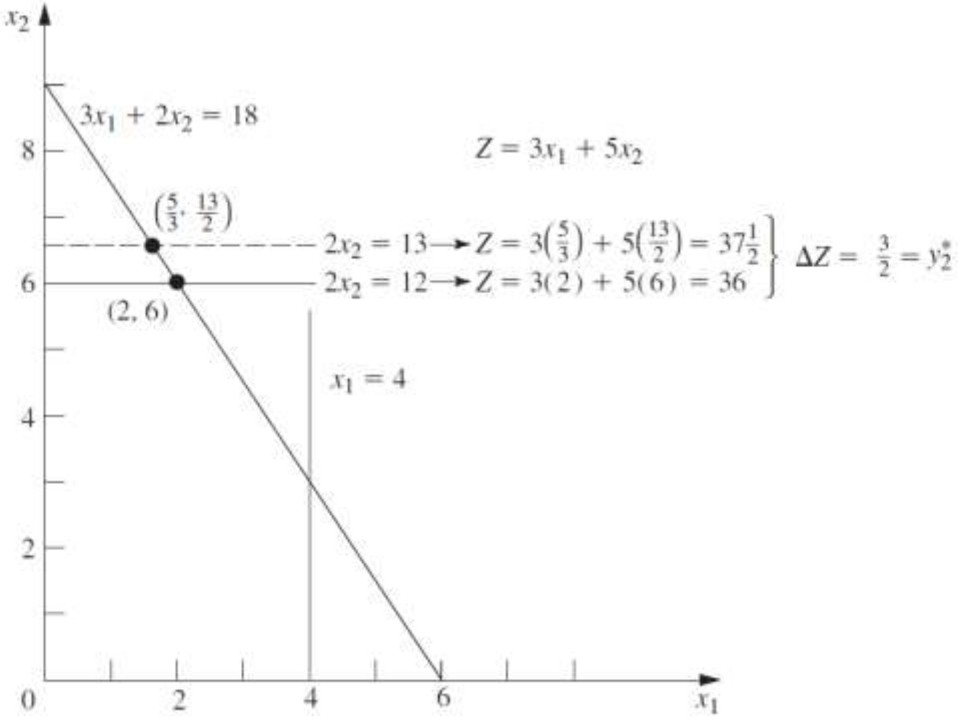
$x_1 \le 4$
$2x_2 \le 12$
$3x_1 + 2x_2 \le 18$

→

$x_1 \le 4$
$2x_2 \le 13$
$3x_1 + 2x_2 \le 18$

Find the new intercept between Z and the new structural constraint $2x_2 \le 13$

With the new constraint Z becomes 37.5 instead of 36
➔ $y_2^* = 1.5$ or $\frac{3}{2}$



BARCELONA SCHOOL OF MANAGEMENT

This is neat, but how many inputs does this problem have?

TABLE 3.1 Data for the Wyndor Glass Co. problem

	Production Time per Batch, Hours		
	Product		
Plant	1	2	Production Time Available per Week, Hours
1	1	0	4
2	0	2	12
3	3	2	18
Profit per batch	$3,000	$5,000	

Six a_i's, three a_i's, two c_i's

=11 inputs

As a manager, I might want to explore more broadly; computing is relatively inexpensive while error in the production planning can be expensive

Source: The Simpson, 20th Television Animation (The Walt Disney Company)

BARCELONA SCHOOL OF MANAGEMENT

As a manager, I might want to explore more broadly; computing is relatively inexpensive while error in the production planning can be expensive

Also interesting to explore what should **not** happen – what conditions might jeopardize the entire enterprise

"Oh, Honey - our first bankruptcy!"

By John Klossner at https://www.cartoonstock.com

upf. BARCELONA SCHOOL OF MANAGEMENT

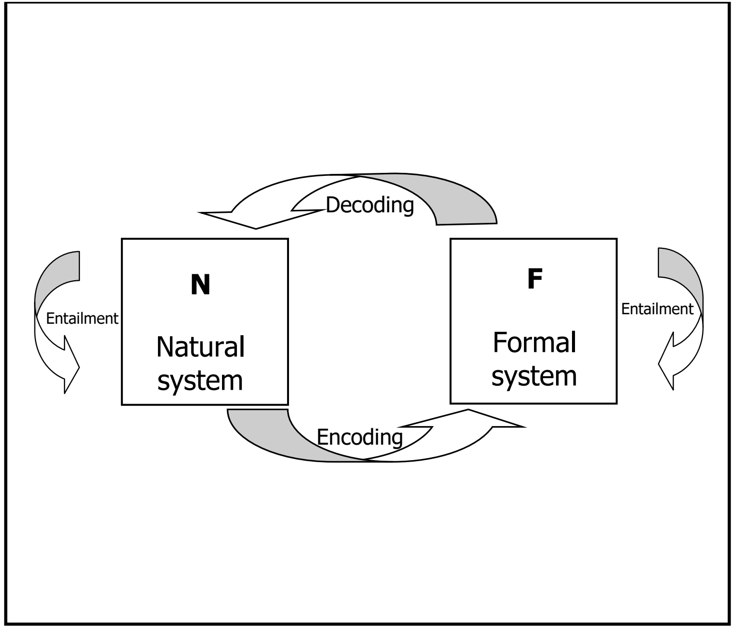
10.

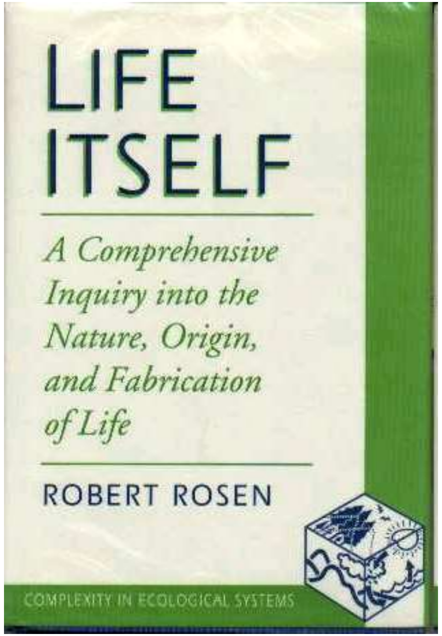
Models

Why this section. Modelling as a craft or an art. Model versus straight physical laws. Models and their memory. Models in economics. Maps and the territory. Underfitting versus overfitting. Uncertainty versus ignorance.

Modelling is a craft more than a science

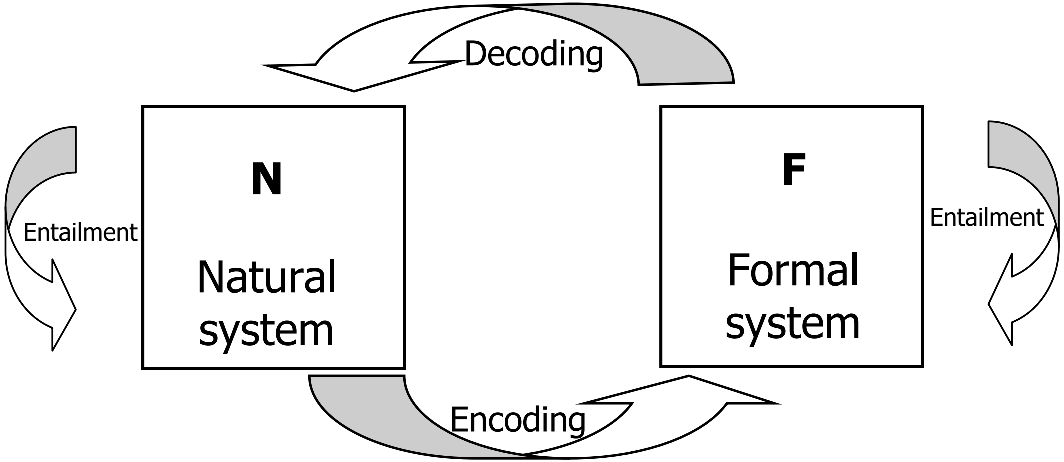
Modelling as a craft rather than as a science for Robert Rosen





R. Rosen, Life Itself: A Comprehensive Inquiry Into the Nature, Origin, and Fabrication of Life. Columbia University Press, 1991.

Louie, A.H. 2010. “Robert Rosen’s Anticipatory Systems.” Edited by Riel Miller. Foresight 12 (3): 18–29. https://doi.org/10.1108/14636681011049848.



What is a model ?

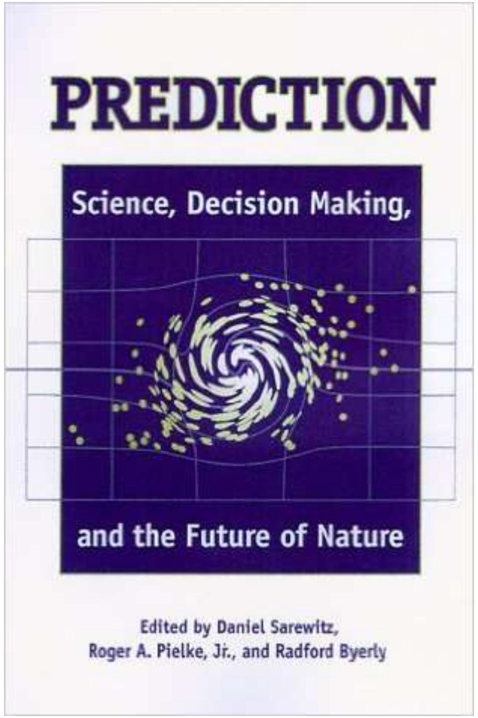
Robert Rosen
(1934-1998)

“models are most useful when they are used to challenge existing formulations, rather than to validate or verify them”

Naomi Oreskes

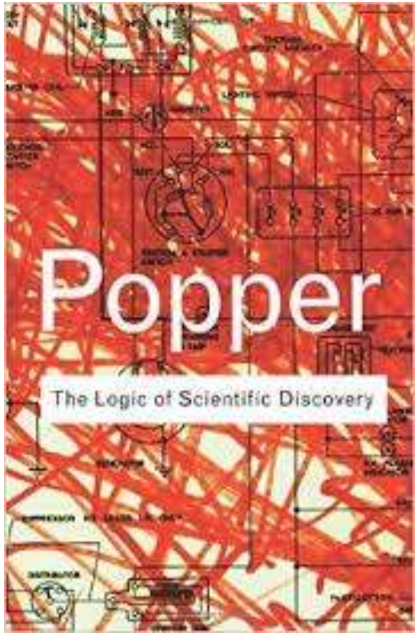
N. Oreskes, K. Shrader-Frechette, and K. Belitz, “Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences,” Science, 263, no. 5147, 1994.

Models are not physical laws

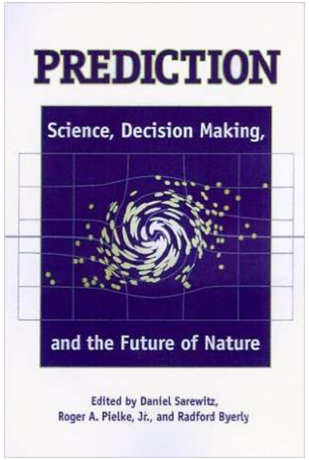


Oreskes, N., 2000, Why predict? Historical perspectives on prediction in Earth Science, in Prediction, Science, Decision Making and the future of Nature, Sarewitz et al., Eds., Island Press, Washington DC

“[⋯] to be of value in theory testing, the predictions involved must be capable of refuting the theory that generated them” (N. Oreskes)



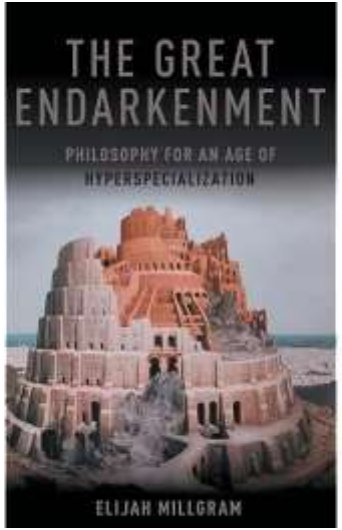
“When a model generates a prediction, of what precisely is the prediction a test? The laws? The input data? The conceptualization?



Any part (or several parts) of the model might be in error, and there is no simple way to determine which one it is”

Models have little memory

“[…] The process of constructing and validating [value-at risk] models is time consuming and detail oriented; normally even the people who produced the model will not remember many of the assumptions incorporated into it, short of redoing their work, which means that the client cannot simply ask then what went into it.”



E. Millgram The Great Endarkenment, p. 29

Caeteris are never paribus

Ceteris paribus or caeteris paribus (Latin) = "all other things being equal" or "other things held constant" or "all else unchanged”

The case of DSGE, dynamic stochastic general equilibrium models

Rational expectations of agents
Efficient market hypothesis

Philip Mirowski

Philip Mirowski, 2013, Never let a serious crisis go wasted, Verso Books.

The US senate and Queen Elisabeth perplexed…

Philip Mirowski, 2013, Never let a serious crisis go wasted, Verso Books.

Dangers of mathematization of economics

Wolfgang Drechsler

Erik S. Reinert

Paul Romer

Philip Mirowski

W. Drechsler, “On the possibility of quantitative-mathematical social science, chiefly economics,” *J. Econ. Stud.*, vol. 27, no. 4/5, pp. 246–259, 2000.

E. S. Reinert, “Full circle: economics from scholasticism through innovation and back into mathematical scholasticism,” *J. Econ. Stud.*, vol. 27, no. 4/5, pp. 364–376, Aug. 2000.

P. Romer, “Mathiness in the Theory of Economic Growth,” *Am. Econ. Rev.*, vol. 105, no. 5, pp. 89–93, May 2015.

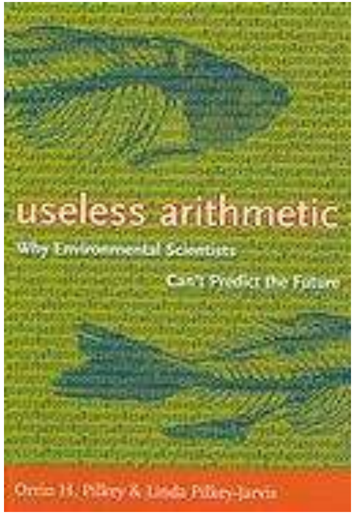
Mirowski, Philip. 2013. Never Let a Serious Crisis Go to Waste: How Neoliberalism Survived the Financial Meltdown. Verso.

Don’t confuse the map with the territory

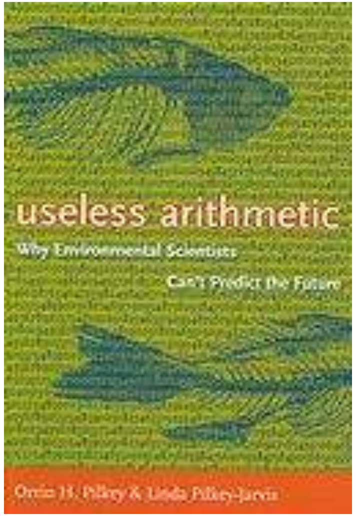
If you do, sensitivity analysis will not save you

<<It is important, however, to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature>>

Orrin H. Pilkey



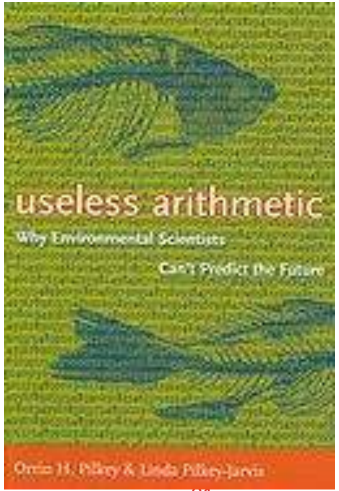
Useless Arithmetic: Why Environmental Scientists Can't Predict the Future
by Orrin H. Pilkey and Linda Pilkey-Jarvis, Columbia University Press, 2009.

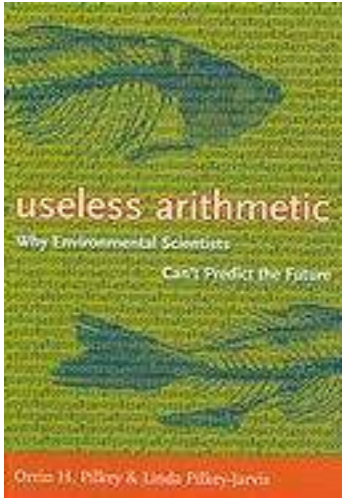


<<…If the model is wrong or if it is a poor representation of reality, determining the sensitivity of an individual parameter in the model is a meaningless pursuit>>

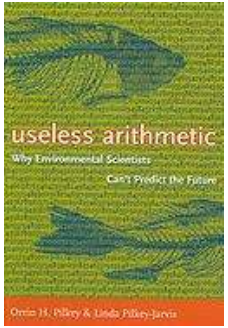
One of the examples discussed concerns the Yucca Mountain repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis.

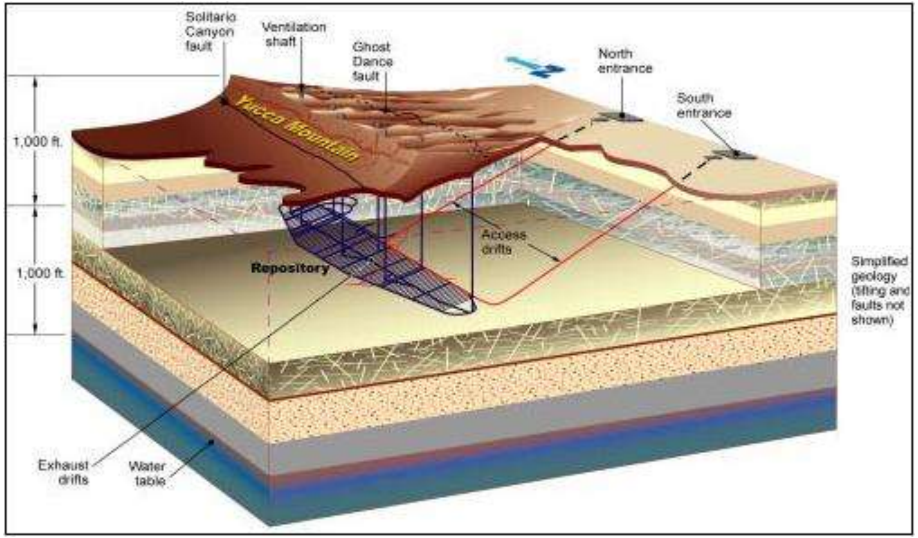
TSPA is Composed of 286 sub-models.





TSPA (like any other model) relies on assumptions → one is the low permeability of the geological formation → long time for the water to percolate from surface to disposal



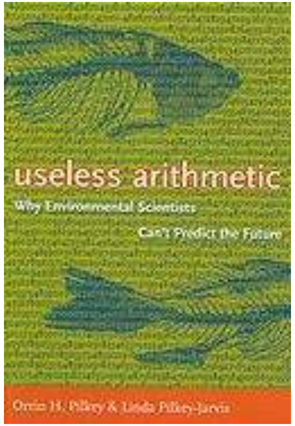


The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter

(the ^{36}Cl story)

In the case of TSPA (Yucca mountain) a range of 0.02 to 1 millimetre per year was used for percolation of flux rate.

→… SA useless if it is instead ~ 3,000 millimetres per year.



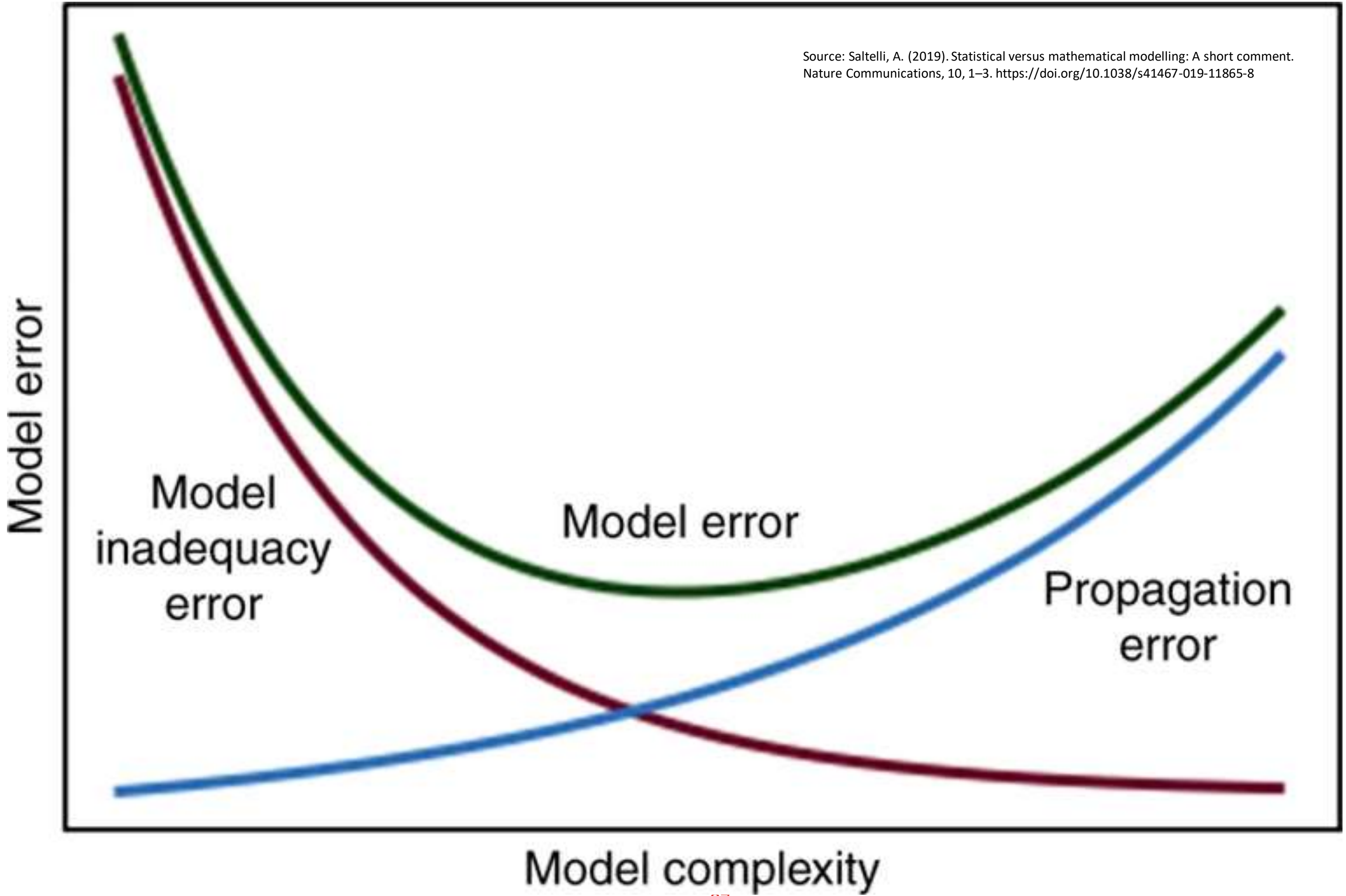
“Scientific mathematical modelling should involve constant efforts to falsify the model”

➔ Organized skepticism (as per CUDOS)

Communalism, **U**niversalism, **D**isinterestedness, **O**rganized **S**kepticism, from sociology of science, Robert K. Merton.

Beware the size of your model

Mind the conjecture of O'Neil



Source: Saltelli, A. (2019). Statistical versus mathematical modelling: A short comment. Nature Communications, 10, 1–3. https://doi.org/10.1038/s41467-019-11865-8

Conjecture by O'Neill, also known as Zadeh's principle of incompatibility, whereby as complexity increases "precision and significance (or relevance) become almost mutually exclusive characteristics"

In M. G. Turner and R. H. Gardner, "Introduction to Models" in Landscape Ecology in Theory and Practice, New York, NY: Springer New York, 2015, pp. 63–95.

L. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," IEEE Trans. Syst. Man. Cybern., vol. 3, no. 1, pp. 28–44, 1973.

Puy, Arnald, Pierfrancesco Beneventano, Simon A. Levin, Samuele Lo Piano, Tommaso Portaluri, and Andrea Saltelli. 2022. "Models with Higher Effective Dimensions Tend to Produce More Uncertain Estimates." Science Advances 8 (eabn9450).

Simple principles of responsible modelling

Mind the assumptions

Assess uncertainty and sensitivity

Mind the hubris

Complexity can be the enemy of relevance

Mind the framing

Match purpose and context

Source: A. Saltelli, G. Bammer, I. Bruno, et al., Five ways to ensure that models serve society: a manifesto, Nature 582 (2020) 482–484.

Mind the consequences

Quantification can backfire.

Mind the unknowns

Acknowledge ignorance

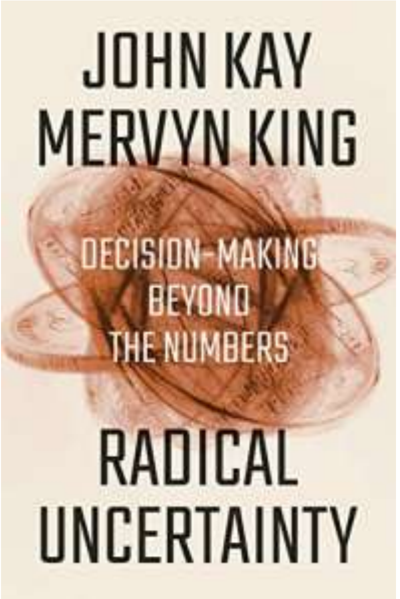
Mind the assumptions

Assess uncertainty and sensitivity

… models require input values for which there is no reliable information…

Source: A. Saltelli, G. Bammer, I. Bruno, et al., Five ways to ensure that models serve society: a manifesto, Nature 582 (2020) 482–484.

Models ask as input information which we don't have – The case of WEBTAG



John Kay

WebTAG: Annual Percentage Change in Car Occupancy (% pa) up to 2036

Journey Purpose	Weekday					Weekend	All Week
	7am-10am	10am-4pm	4pm-7pm	7pm-7am	Weekday Average		
Work	-0.48	-0.4	-0.62	-0.5	-0.44	-0.48	-0.45
Non - Work (commuting and other)	-0.67	-0.65	-0.53	-0.47	-0.59	-0.52	-0.56

Source: J. A. Kay, "Knowing when we don't know," 2012, https://www.ifs.org.uk/docs/john_kay_feb2012.pdf

Mind the assumptions

Assess uncertainty and sensitivity

Mind the hubris

Complexity can be the enemy of relevance

Mind the framing

Match purpose and context

Source: A. Saltelli, G. Bammer, I. Bruno, et al., Five ways to ensure that models serve society: a manifesto, Nature 582 (2020) 482–484.

Mind the consequences

Quantification can backfire.

Mind the unknowns

Acknowledge ignorance

Mind the consequences

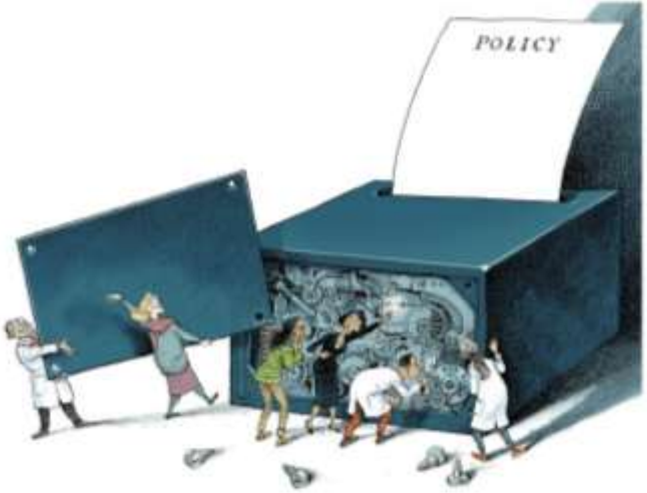
Quantification can backfire.

<< Back to Article

WIRED MAGAZINE: 17.03

Recipe for Disaster: The Formula That Killed Wall Street

By Felix Salmon 02.23.09



$$\Pr[T_A<1,T_B<1]=\phi_2(\phi^{-1}(F_A(1)),\phi^{-1}(F_B(1)),\gamma)$$

Here's what killed your 401(k) *David X. Li's Gaussian copula function as first published in 2000. Investors exploited it as a quick—and fatally flawed—way to assess risk. A shorter version appears on this month's cover of Wired.*

Here is what killed your 401(k)…

Li's Gaussian copula function …

Nassim Nicholas Taleb, hedge fund manager and author of *The Black Swan*, is particularly harsh when it comes to the copula. "People got very excited about the Gaussian copula because of its mathematical elegance, but the thing never worked," he says. "Co-association between securities is not measurable using correlation," because past history can never prepare you for that one day when everything goes south. "Anything that relies on correlation is charlatanism."

Felix Salmon, Wired, February 2009

WIRED

Source: https://www.wired.com/2009/02/wp-quant/

Mind the assumptions

Assess uncertainty and sensitivity

Mind the hubris

Complexity can be the enemy of relevance

Mind the framing

Match purpose and context

Source: A. Saltelli, G. Bammer, I. Bruno, et al., Five ways to ensure that models serve society: a manifesto, Nature 582 (2020) 482–484.

Mind the consequences

Quantification can backfire.

→ **Mind the unknowns**

Acknowledge ignorance

Mind the unknowns

Acknowledge ignorance

Source: A. Saltelli, G. Bammer, I. Bruno, et al., Five ways to ensure that models serve society: a manifesto, Nature 582 (2020) 482–484.

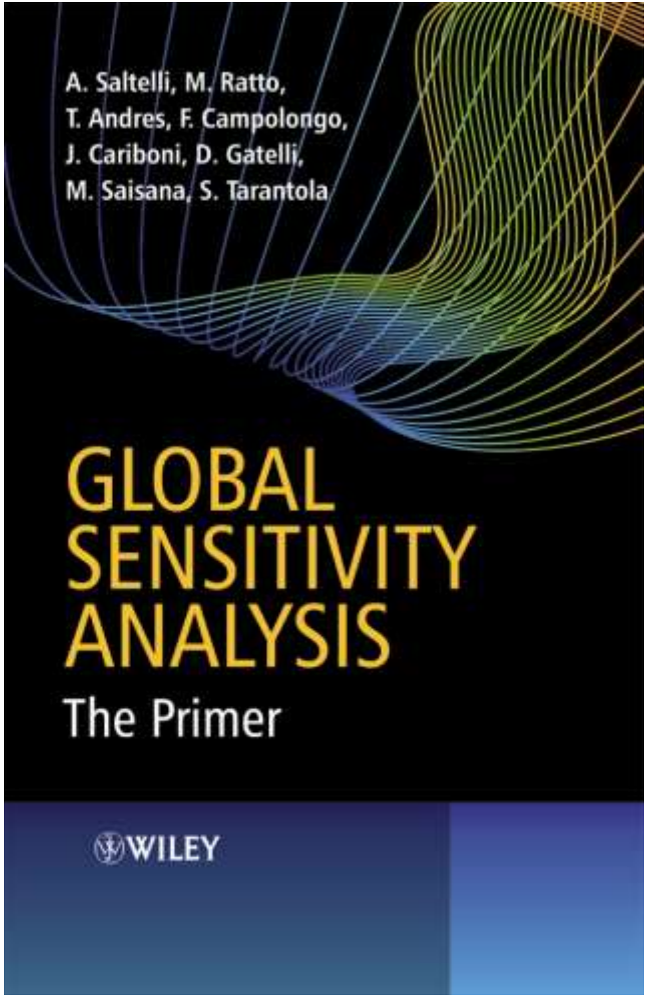
From Socrates's "knowing of not knowing" to Nicolaus Cusanus' Docta Ignorantia, ignorance was a virtue until Descartes

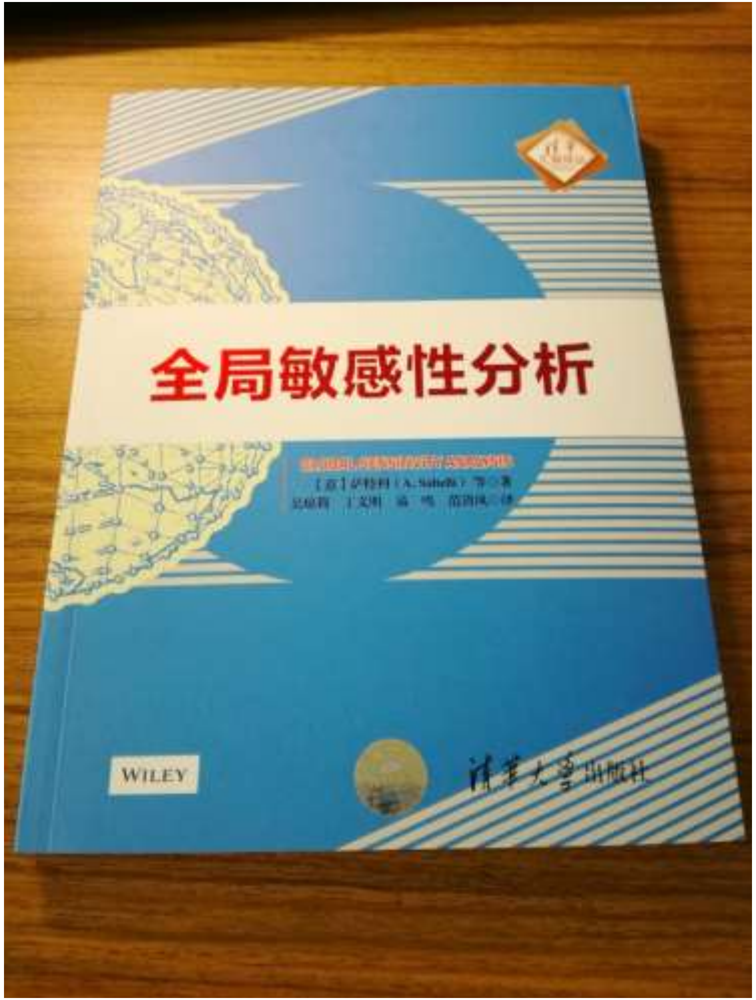
11.

Methods for Uncertainty and sensitivity analysis

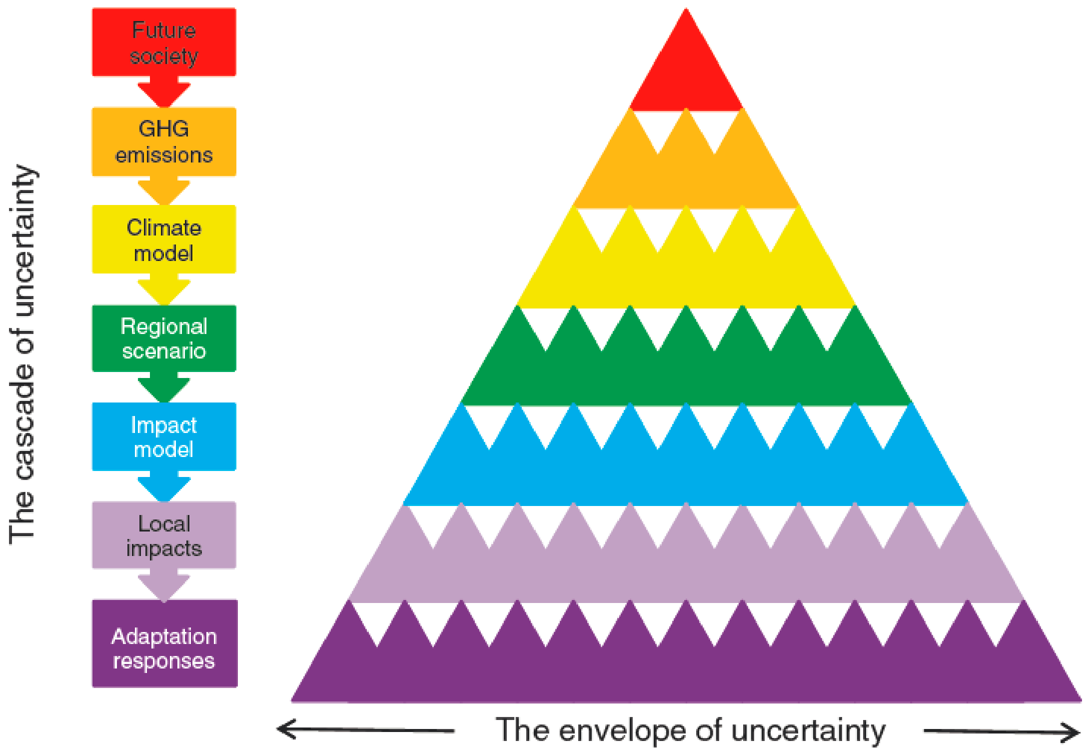
Uncertainty analysis versus sensitivity analysis; sensitivity analysis made simple; Type 1 and type 2 errors. About ‘fishing expeditions’. A trivial Monte Carlo with Excel. Mostly based on Saltelli, A. et al. (2008) Global sensitivity analysis : the primer. John Wiley.

BARCELONA SCHOOL OF MANAGEMENT


A. Saltelli, M. Ratto,
T. Andres, F. Campolongo,
J. Cariboni, D. Gatelli,
M. Saisana, S. Tarantola
GLOBAL
SENSITIVITY
ANALYSIS
The Primer
WILEY


全局敏感性分析
WILEY

Uncertainty analysis: the study of the uncertainty in model output—see also uncertainty cascade



Source: https://www.climate-lab-book.ac.uk/2014/cascade-of-uncertainty/

Sensitivity analysis: the study of the relative importance of different input factors on the model output

Sensitivity analysis can:

- surprise the analyst,
- uncover technical errors in the model,
- identify critical regions in the space of the inputs,
- avoid type II errors ···

Source: The Simpson, 20th Television Animation
(The Walt Disney Company)

BARCELONA SCHOOL OF MANAGEMENT

Sensitivity analysis can:

- avoid type II errors …

Tests are normally set in the negative, the so called null hypothesis (here: **you are not pregnant**).

Erroneously rejecting a true null hypothesis is a Type one error (the man is indeed not pregnant)

Erroneously accepting a null hypothesis that is instead false is a Type two error

BARCELONA SCHOOL OF MANAGEMENT

Type one and two have different implications: e.g. null hypothesis= **chemical X is not carcinogenic**

If X is not carcinogenic, but I reject the true null hypotheses (Type one error), this is bad for the firm producing the chemical and for innovation

If X is carcinogenic, but I accept the false null hypotheses (Type two error), this is bad for people

Type one = erroneously rejecting a true null hypothesis

Type two = erroneously accepting a false null hypothesis

Philosopher Richard Rudner used this example to make the point that scientists do need to make value judgments

Philosophy of Science

VOL. 20 January, 1953 NO. 1

THE SCIENTIST *QUA* SCIENTIST MAKES VALUE JUDGMENTS*

RICHARD RUDNER

R. Rudner, "The Scientist Qua Scientist Makes Value Judgments," Philosophy of Science, vol. 20. The University of Chicago Press Philosophy of Science Association, pp. 1–6, 1953. http://www.andreasaltelli.eu/file/repository/00_Rudnerphs53.pdf

“How sure we need to be before we accept a hypothesis will depend on how serious a mistake would be”

Philosophy of Science

VOL. 20 January, 1953 NO. 1

THE SCIENTIST *QUA* SCIENTIST MAKES VALUE JUDGMENTS*

RICHARD RUDNER

R. Rudner, “The Scientist Qua Scientist Makes Value Judgments,” Philosophy of Science, vol. 20. The University of Chicago Press Philosophy of Science Association, pp. 1–6, 1953. http://www.andreasaltelli.eu/file/repository/00_Rudnerphs53.pdf

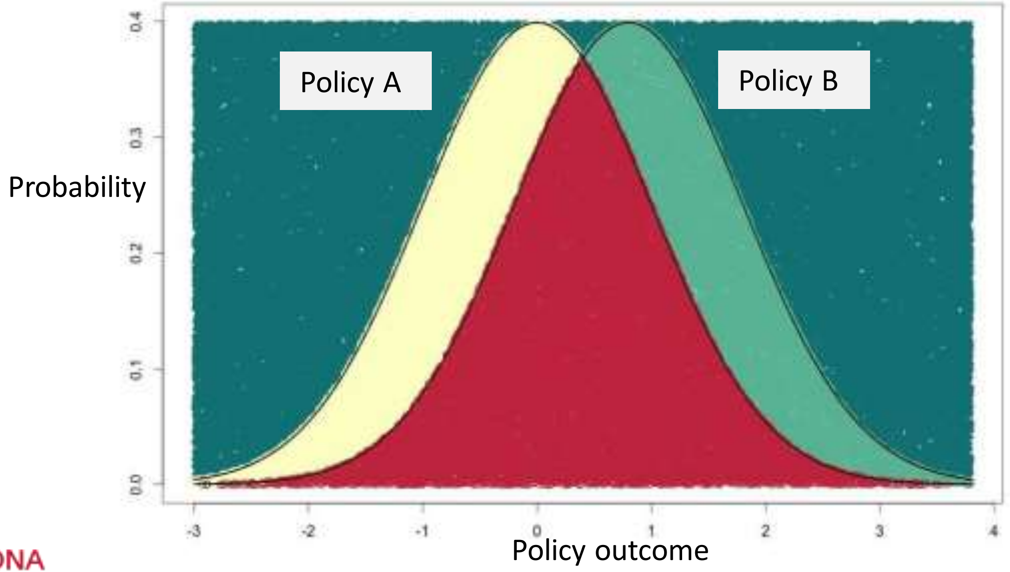
Sensitivity analysis can :

surprise the analyst,
uncover technical errors in the model,
identify critical regions in the space of the inputs,
avoid type II errors
establish priorities for research,
simplify models
falsify models (show that a model is false or irrelevant)
defend against your own model being falsified

BARCELONA SCHOOL OF MANAGEMENT

Sensitivity analysis can:

verify whether policy options (or marketing strategies) can be distinguished from one another given the uncertainties in the system, …



What method would one choose to perform sensitivity analysis?

Source: iStock by Getty images

What method would one choose to perform sensitivity analysis?

Most of the sensitivity analysis found in the literature are local or otherwise OAT (One factor At a Time)

$$y = f(x_1, x_2, \dots x_k)$$

$$\left.\frac{\partial y}{\partial x_i}\right|_{x_i = x_i^0} \longleftarrow \text{Local}$$

BARCELONA SCHOOL OF MANAGEMENT

What method would one choose to perform sensitivity analysis?

Most of the sensitivity analysis found in the literature are local or otherwise OAT (One factor At a Time)

$$y = f(x_1, x_2, \dots x_k)$$

$$\left.\frac{x_i^0}{y^0}\frac{\partial y}{\partial x_i}\right|_{x_i = x_i^0} \longleftarrow \text{Local}$$

What method would one choose to perform sensitivity analysis?

Most of the sensitivity analysis found in the literature are local or otherwise OAT (One factor At a Time)

$$y = f(x_1, x_2, \dots x_k)$$

$$\left.\frac{std(x_i)}{std(y)} \frac{\partial y}{\partial x_i}\right|_{x_i = x_i^0} \longleftarrow \text{Hybrid}$$

BARCELONA SCHOOL OF MANAGEMENT

$$\left.\frac{\partial y}{\partial x_i}\right|_{x_i=x_i^0}$$

← Relative effect on y of perturbing x_i around its nominal value

$$\left.\frac{x_i^0}{y^0}\frac{\partial y}{\partial x_i}\right|_{x_i=x_i^0}$$

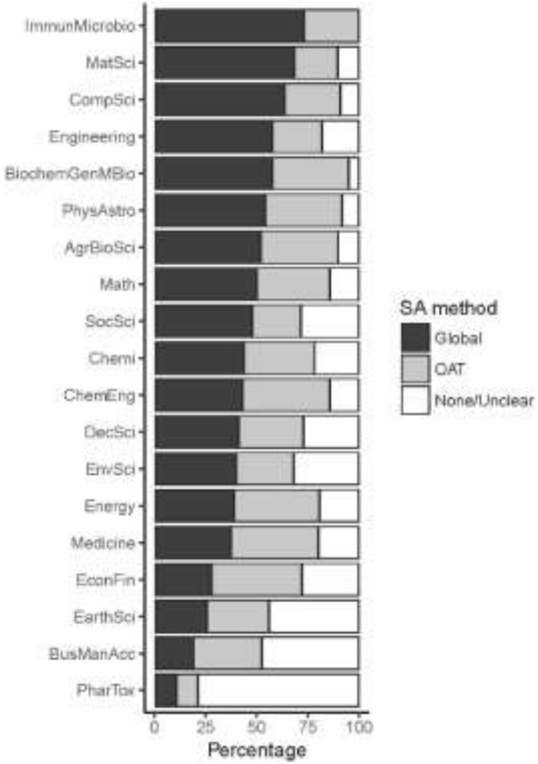
← Relative effect on y of perturbing x_i by a fixed fraction of its nominal value

$$\left.\frac{std(x_i)}{std(y)}\frac{\partial y}{\partial x_i}\right|_{x_i=x_i^0}$$

← Relative effect on y of perturbing x_i by a fixed fraction of its standard deviation

Jump to 83

Instead of derivatives (local), incremental ratios can be taken by moving factors one at a time away from their baseline value by some (e.g. 5%) fraction. These methods are not local, but still OAT.



Source: Saltelli, Andrea, Ksenia Aleksankina, William Becker, Pamela Fennell, Federico Ferretti, Niels Holst, Sushan Li, and Qiongli Wu. 2019. "**Why so Many Published Sensitivity Analyses Are False**: A Systematic Review of Sensitivity Analysis Practices." Environmental Modelling & Software 114 (April): 29–39. https://doi.org/10.1016/J.ENVSOFT.2019.01.012.

upf. BARCELONA SCHOOL OF MANAGEMENT

A self-evident problem, to understand the methods applied to it.

A simple linear form:

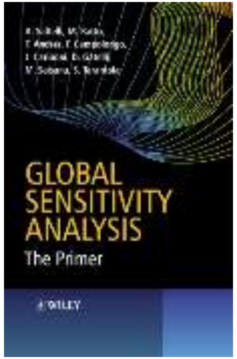
Not a vector

$y = \sum_{i=1}^{k} \Omega_i Z_i$

Where y (a scalar) is the output of interest, the Ω_i's are fixed coefficients and Z_i's are uncertain input factors following a Normal distribution

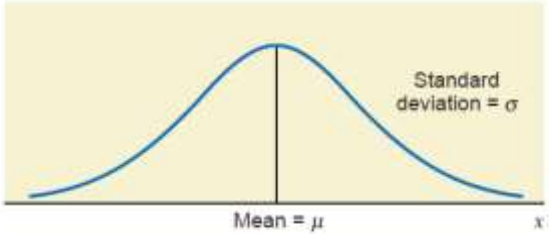
$Z_i \sim N(\bar{z}_i, \sigma_{Z_i})$ ← A distribution

Where $\bar{z}_i = 0, i = 1,2,\ldots k$ are the means of the factors Z_i's andσ_{Z_i}their standard deviations



upf. BARCELONA SCHOOL OF MANAGEMENT

$$Z_i \sim N\left(\overline{z_i}, \sigma_{Z_i}\right)$$

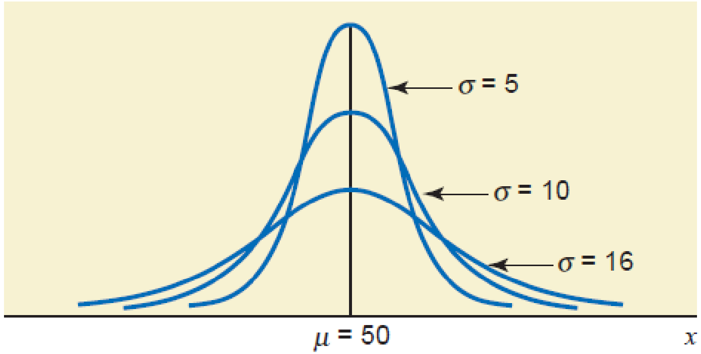


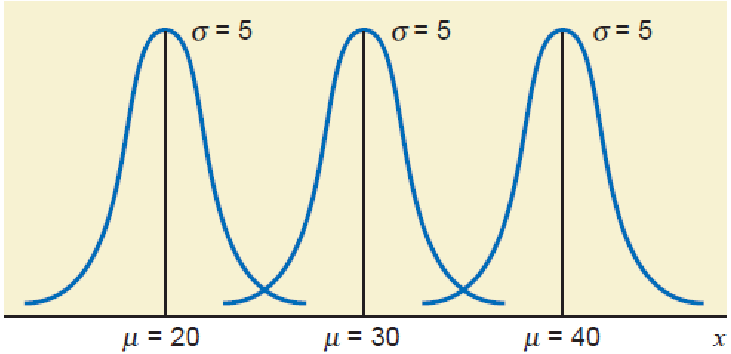
Normal Probability Distribution A *normal probability distribution*, when plotted, gives a bell-shaped curve such that:

1. The total area under the curve is 1.0.
2. The curve is symmetric about the mean.
3. The two tails of the curve extend indefinitely.

$$Z_i \sim N\left(\overline{z_i}, \sigma_{Z_i}\right)$$

The mean and the standard deviation are the distribution's parameter that determine its shape

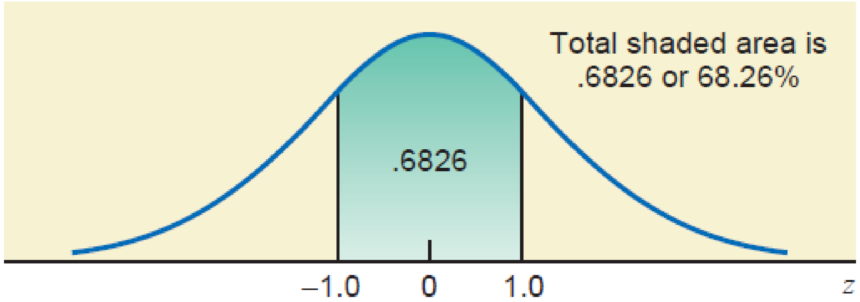




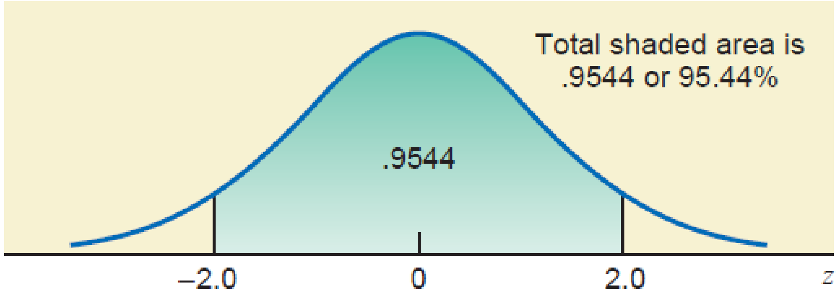
BARCELONA SCHOOL OF MANAGEMENT

$Z_i \sim N(\overline{z_i}, \sigma_{Z_i})$

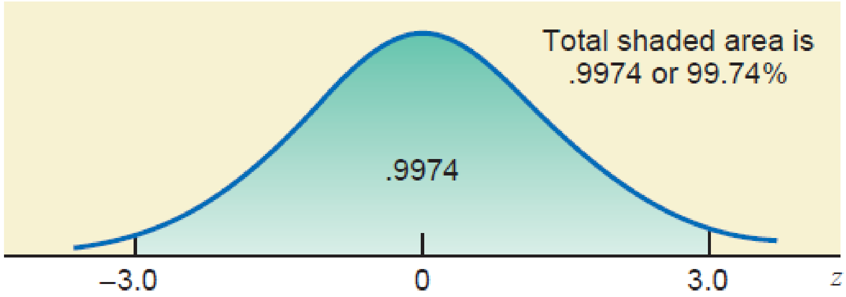
Area under ± σ



Area under ± 2σ



Area under ± 3σ



upf. BARCELONA SCHOOL OF MANAGEMENT

Back to our simple model $y = \sum_{i=1}^{k} \Omega_i Z_i$

Ω_i's are fixed coefficients and Z_i's are uncertain input factors following a Normal distribution $Z_i \sim N\left(\bar{z}_i, \sigma_{Z_i}\right)$ with parameters $\bar{z}_i$ and σ_{Z_i}, for $i = 1,2, \dots k$

It can be shown that y will also be normally distributed with parameters

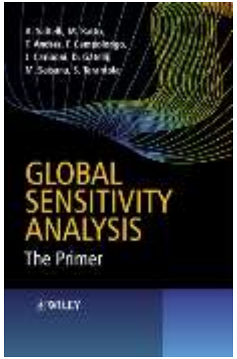
$\bar{y} = \sum_{i=1}^{k} \Omega_i \, \bar{z}_i$

and

$$\sigma_y = \sqrt{\sum_{i=1}^{k} \Omega_i^2 \sigma_{Z_i}^2}$$

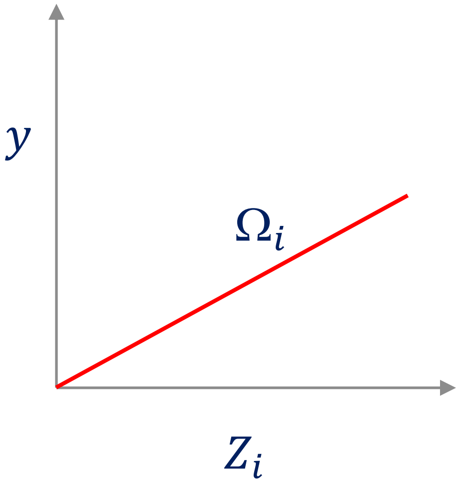
Exercise in the book using the formula for the normal distribution & some calculus

$$N\left(\bar{z}_i, \sigma_{Z_i}\right) = \frac{1}{\sigma_{Z_i}\sqrt{2\pi}} e^{-\frac{z_i^2}{2\sigma_{Z_i}}}$$



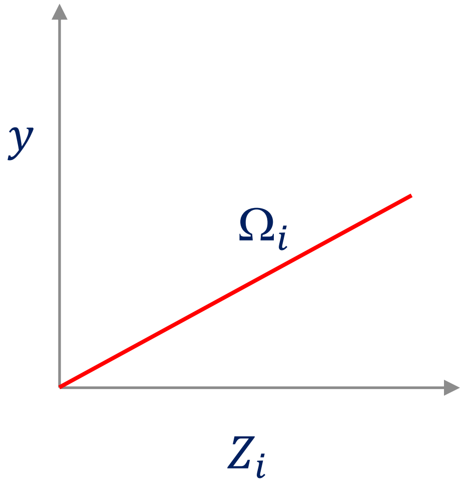
BARCELONA SCHOOL OF MANAGEMENT

Model $y = \sum_{i=1}^{k} \Omega_i Z_i$



Plotted as a function of each of its Z_i (keeping the other Z_i's fixed) gives a straight line of slope Ω_i

Model $y = \sum_{i=1}^{k} \Omega_i Z_i$



Assume

$\sigma_{Z_1} < \sigma_{Z_2} < \cdots < \sigma_{Z_k}$
$\Omega_1 = \Omega_2 = \ldots = \Omega_k$

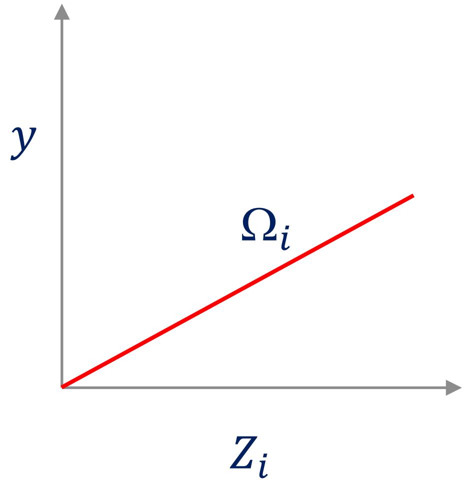
Using local sensitivity analysis the sensitivity of y to each of his input factors Z_i is

$$S_{Z_i}^{d} = \left.\frac{\partial y}{\partial Z_i}\right|_{Z_i = z_i^0}$$

Superscript d in $S_{Z_i}^{d}$ to say that this is a derivative−based sensitivity

i.e. a derivative computed in some point Z_i^0 , for example in $\overline{z_i}$, of of y versus Z_i

Model $y = \sum_{i=1}^{k} \Omega_i Z_i$



Assume

$\sigma_{Z_1} < \sigma_{Z_2} < \cdots < \sigma_{Z_k}$
$\Omega_1 = \Omega_2 = \ldots = \Omega_k$ ← $S_{Z_i}^d = \Omega_i, i = 1,2, \ldots k$

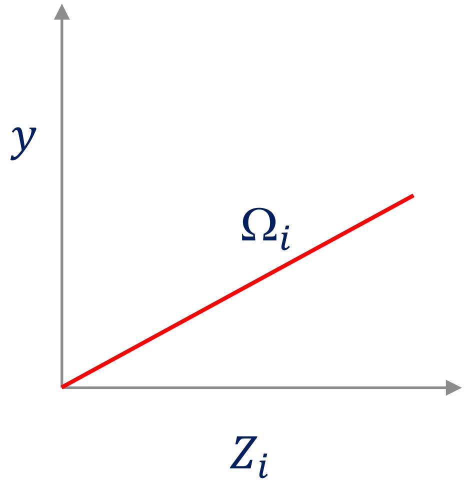
Using local sensitivity analysis the sensitivity of y to each of his input factors Z_i is

$$S_{Z_i}^d = \left. \frac{\partial y}{\partial Z_i} \right|_{Z_i = Z_i^0}$$ ← All derivatives are equal. The model is equally sensitive to all factors

i.e. a derivative computed in some point Z_i^0 , for example in $\overline{z_i}$, of of y versus Z_i

BARCELONA SCHOOL OF MANAGEMENT

Model $y = \sum_{i=1}^{k} \Omega_i Z_i$



Assume

$\sigma_{Z_1} < \sigma_{Z_2} < \cdots < \sigma_{Z_k}$
$\Omega_1 = \Omega_2 = \ldots = \Omega_k$ ← $S_{Z_i}^{d} = \Omega_i, i = 1,2, \ldots k$

All derivatives are equal. The model is equally sensitive to all factors

One would have expected the model to be more sensitive to the factors with the highest standard deviation

Model $y = \sum_{i=1}^{k} \Omega_i Z_i$

Let us generate instead some Monte Carlo points

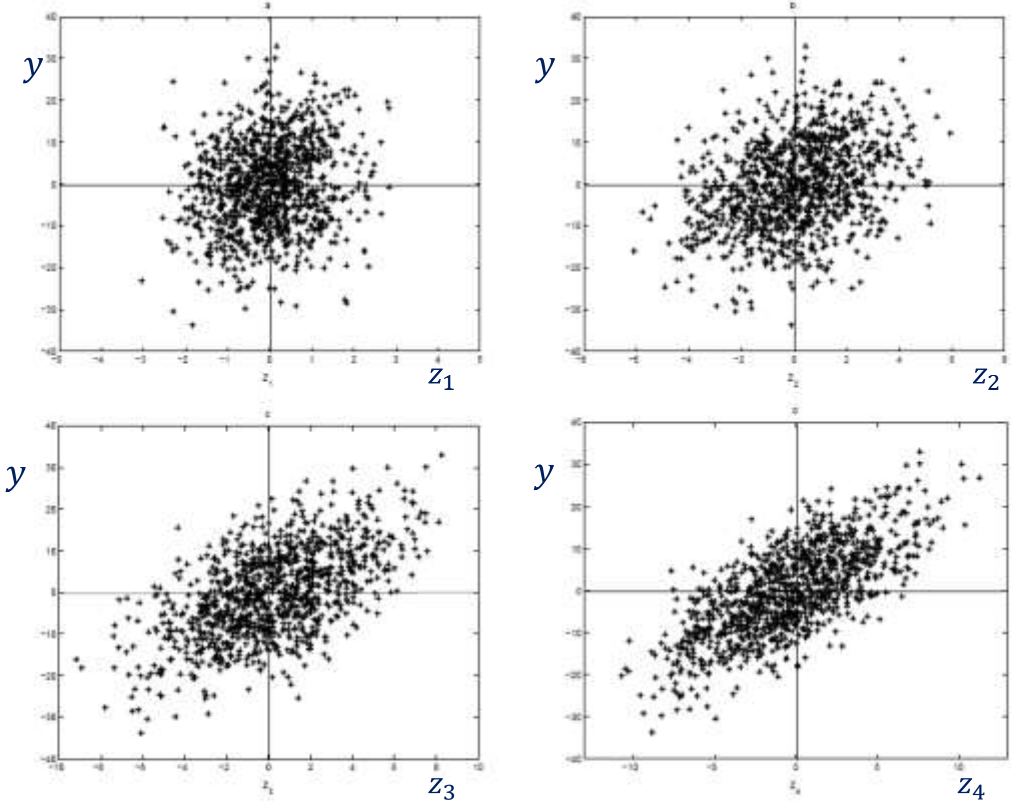
$$
\begin{matrix}
z_1^1 & z_2^1 \dots & z_k^1 \\
z_1^2 & z_2^2 \dots & z_k^2 \\
\dots & \dots & \dots \\
z_1^N & z_2^N \dots & z_k^N
\end{matrix}
$$

Those are random number generated in $(0,1)$ and rescaled for our normal distributions; each row can be used to compute a value of y

$$
\begin{matrix}
z_1^1 & z_2^1 \dots & z_k^1 \\
z_1^2 & z_2^2 \dots & z_k^2 \\
\dots & \dots & \dots \\
z_1^N & z_2^N \dots & z_k^N
\end{matrix}
\longrightarrow
\begin{matrix}
y_1 \\
y_2 \\
\dots \\
y_N
\end{matrix}
$$

We can now plot $\begin{matrix} y_1 \\ y_2 \\ \dots \\ y_N \end{matrix}$ versus the sorted input

$$\begin{matrix} z_1^1 & z_2^1 \dots & z_k^1 \\ z_1^2 & z_2^2 \dots & z_k^2 \\ \dots & \dots & \dots \\ z_1^N & z_2^N \dots & z_k^N \end{matrix}$$

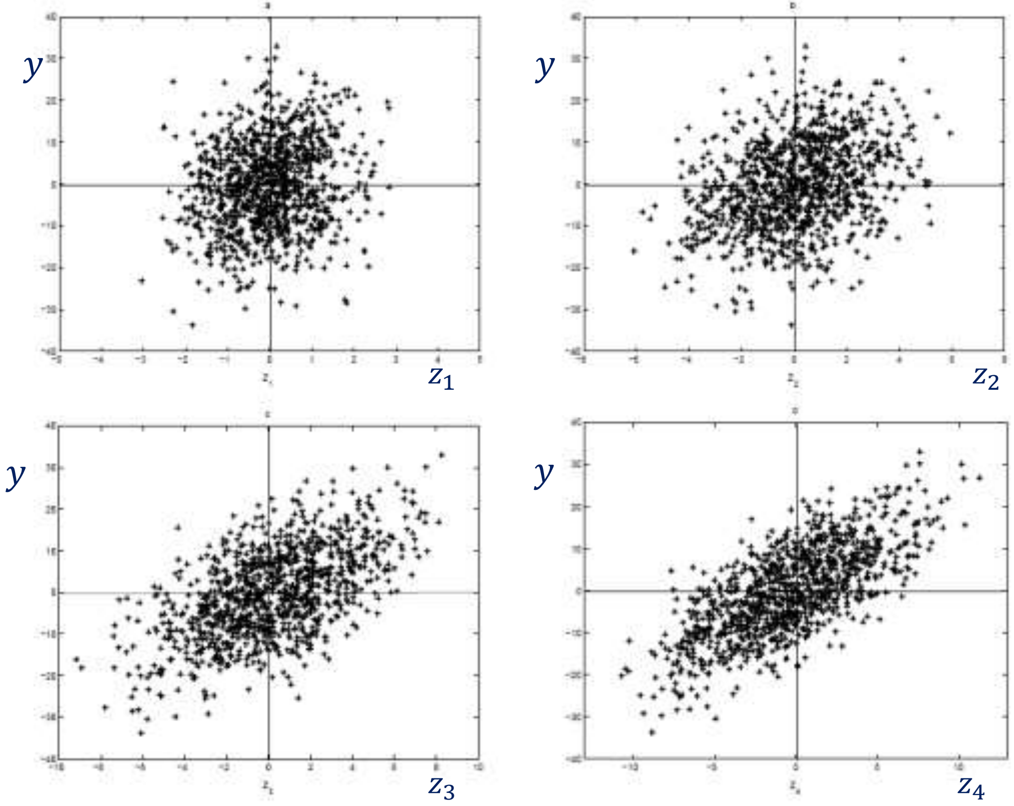


In the plots

$(\sigma_{Z_1}, \sigma_{Z_2}, \sigma_{Z_3}, \sigma_{Z_4}) = (1,2,3,4)$

$\Omega_1 = \Omega_2 = \Omega_3 = \Omega_4 = 2$

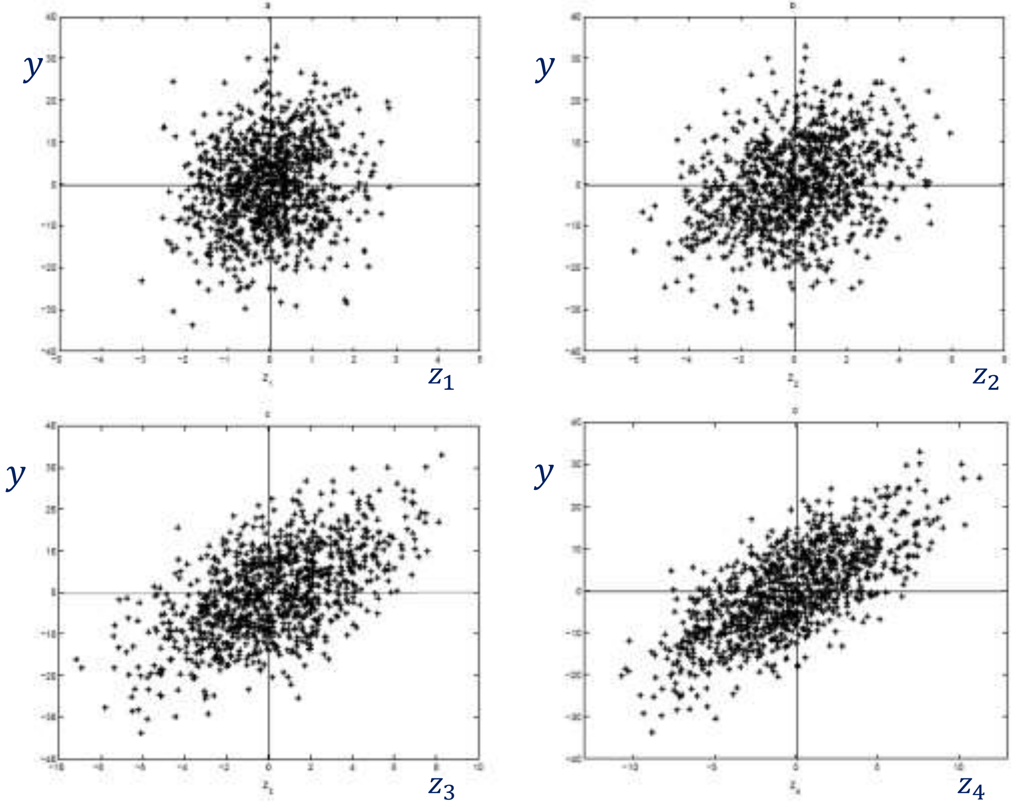
Conclusion: the scatterplot tell us that the importance of variables follows the value of the respective standard deviations, but a derivative based approach does not lead to this result



For this particular problem you can see this also using the derivatives normalized by the standard deviations, i.e.

$$S_{z_i}^{\sigma} = \frac{\sigma_{z_i}}{\sigma_y} \frac{\partial y}{\partial z_i}\bigg|_{z_i = z_i^0}$$

But this is only because the problem is linear



upf. BARCELONA SCHOOL OF MANAGEMENT

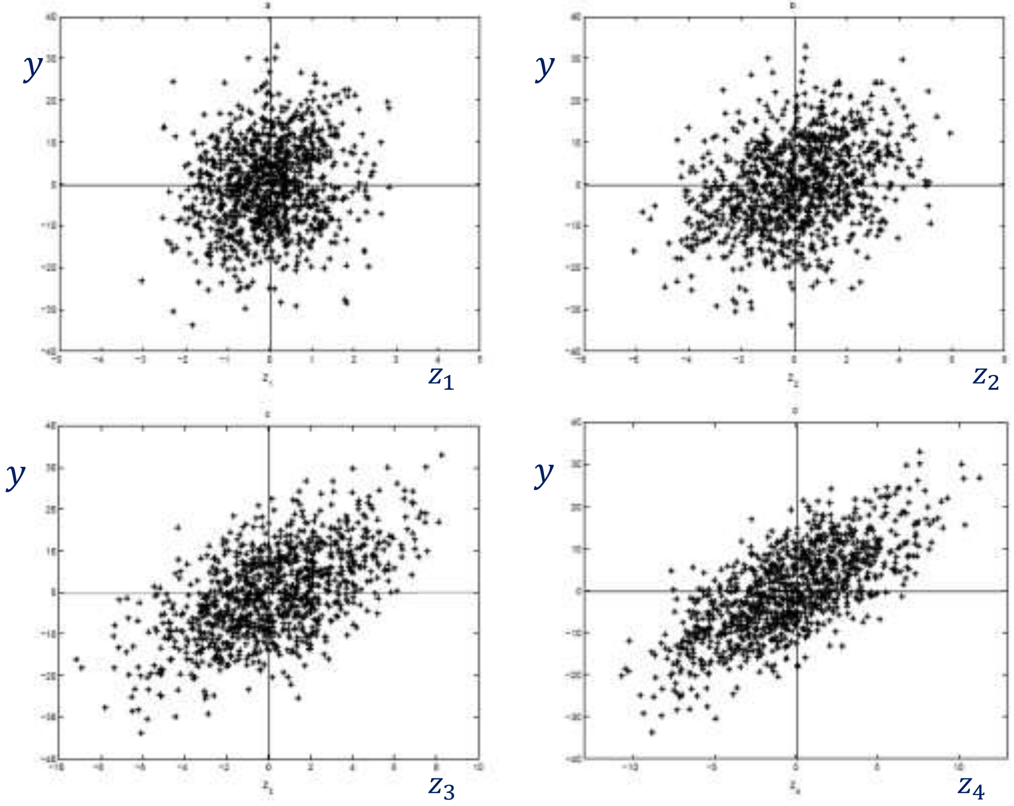
An interesting property of these σ -normalized sensitivity:

$$S_{z_i}^{\sigma} = \frac{\sigma_{z_i}}{\sigma_y} \frac{\partial y}{\partial z_i}\bigg|_{z_i = z_i^0}$$

Computing them for the same data in the scatterplots:

$$(\sigma_{Z_1}, \sigma_{Z_2}, \sigma_{Z_3}, \sigma_{Z_4}) = (1,2,3,4)$$

$$\Omega_1 = \Omega_2 = \Omega_3 = \Omega_4 = 2$$



An interesting property of these σ−normalized sensitivity:

$$S_{z_i}^{\sigma} = \frac{\sigma_{z_i}}{\sigma_y} \frac{\partial y}{\partial z_i}\bigg|_{z_i = z_i^0}$$

Computing them for the same data in the scatterplots:

$$(\sigma_{Z_1}, \sigma_{Z_2}, \sigma_{Z_3}, \sigma_{Z_4}) = (1,2,3,4)$$

$$\Omega_1 = \Omega_2 = \Omega_3 = \Omega_4 = 2$$

	$S_{Z_i}^{d}$	$\left(S_{Z_i}^{\sigma}\right)^2$
Z_1	2	0.036
Z_2	2	0.14
Z_3	2	0.31
Z_4	2	0.56
	Sum	~1

Not only the $S_{Z_i}^{\sigma}$'s reflect the importance of the different standard deviations but once squared they add to one

Where does this lead?

$y = \sum_{i=1}^{k} \Omega_i Z_i$ and $S_{z_i}^{\sigma} = \frac{\sigma_{z_i}}{\sigma_y} \frac{\partial y}{\partial z_i}\bigg|_{z_i = z_i^0}$

give

$$S_{z_i}^{\sigma} = \frac{\sigma_{z_i}}{\sigma_y} \Omega_i$$

$$\sum_{i=1}^{4} \left(S_{z_i}^{\sigma}\right)^2 = 1 = \sum_{i=1}^{4} \left(\frac{\sigma_{z_i}}{\sigma_y} \Omega_i\right)^2$$

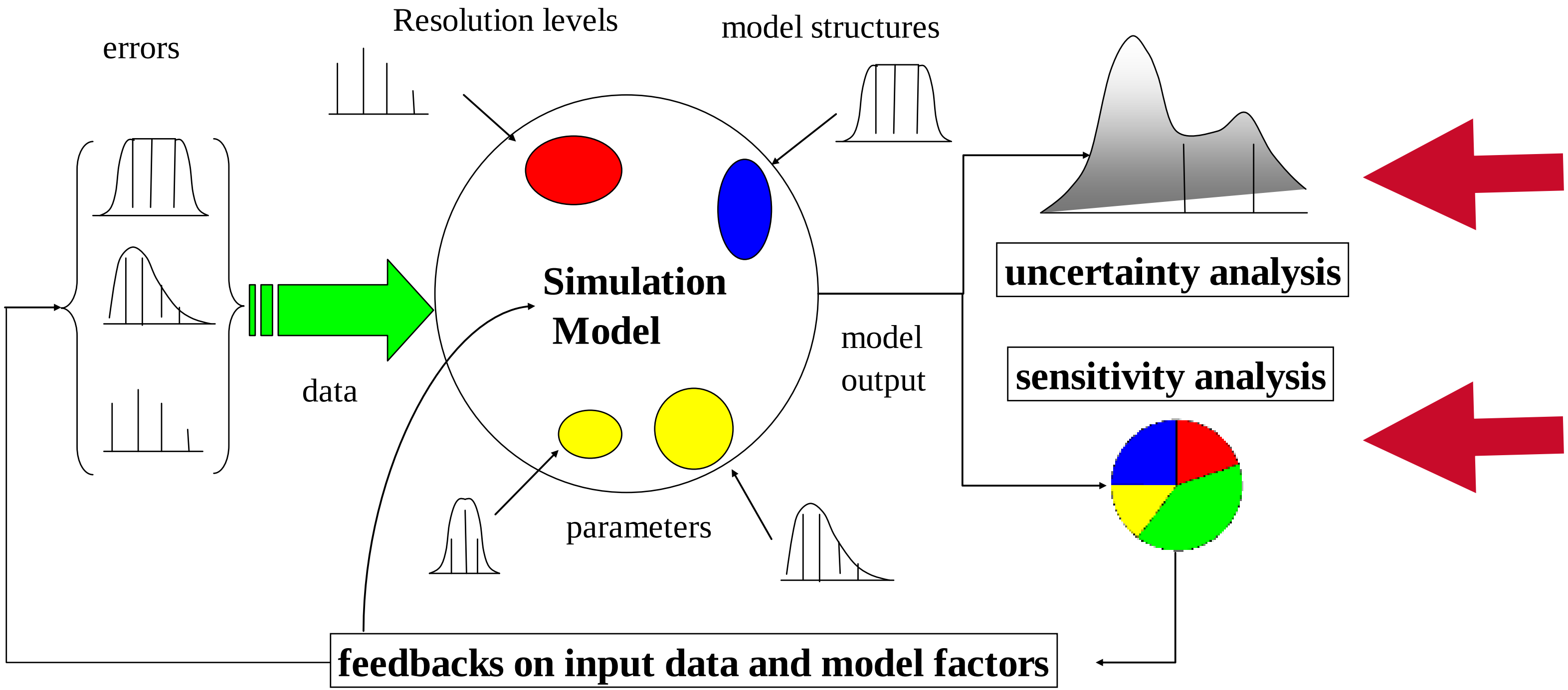
$$\left(\sigma_y\right)^2 = \sum_{i=1}^{4} \sigma_{z_i}^2 \Omega_i^2$$

	$S_{z_i}^{d}$	$\left(S_{z_i}^{\sigma}\right)^2$
Z_1	2	0.036
Z_2	2	0.14
Z_3	2	0.31
Z_4	2	0.56
	Sum	~1

$\left(\sigma_y\right)^2$, the total variance of the output is decomposed

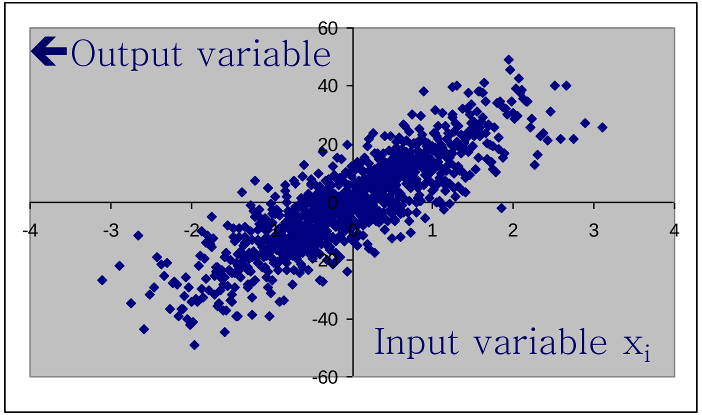
In sensitivity analysis a measure that decomposes the variance of the output is a best practice

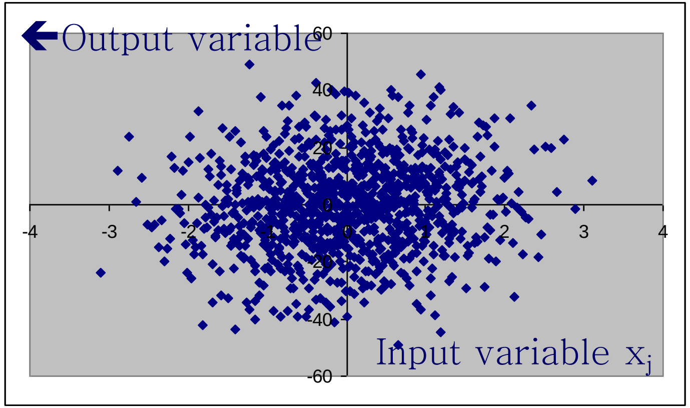
BARCELONA SCHOOL OF MANAGEMENT


Resolution levels
model structures
errors
Simulation Model
data
model output
uncertainty analysis
sensitivity analysis
parameters
feedbacks on input data and model factors

83

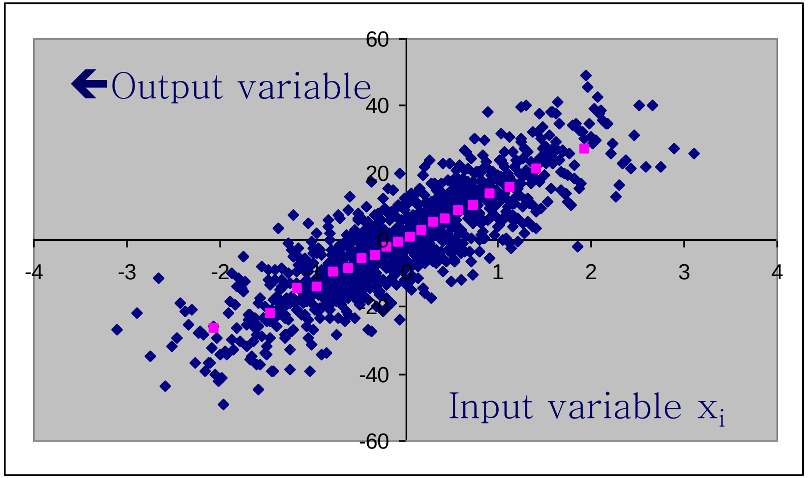
An introduction to variance based methods

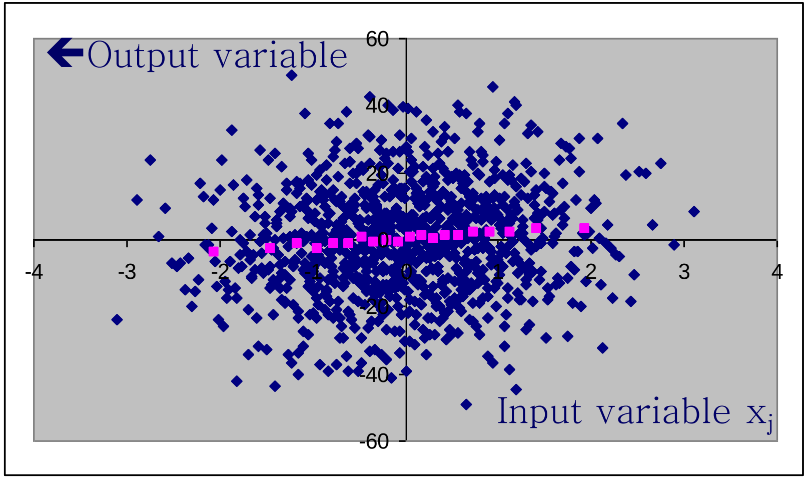




Plotting the output as a function of two different input factors

Which factor is more important?

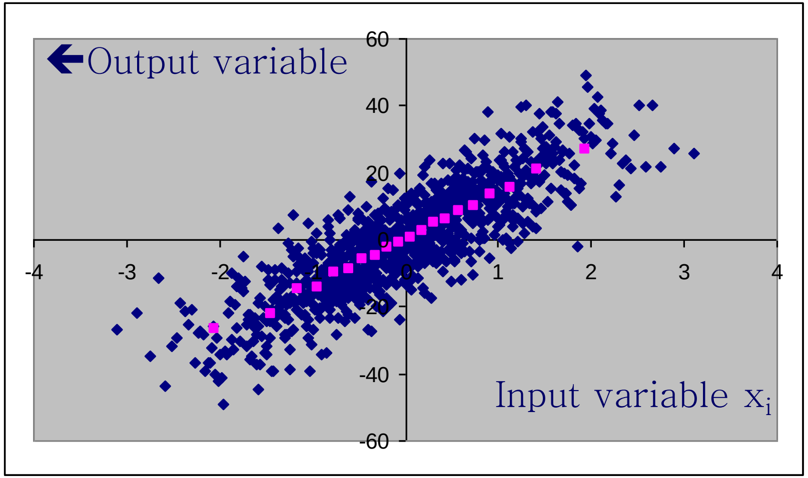




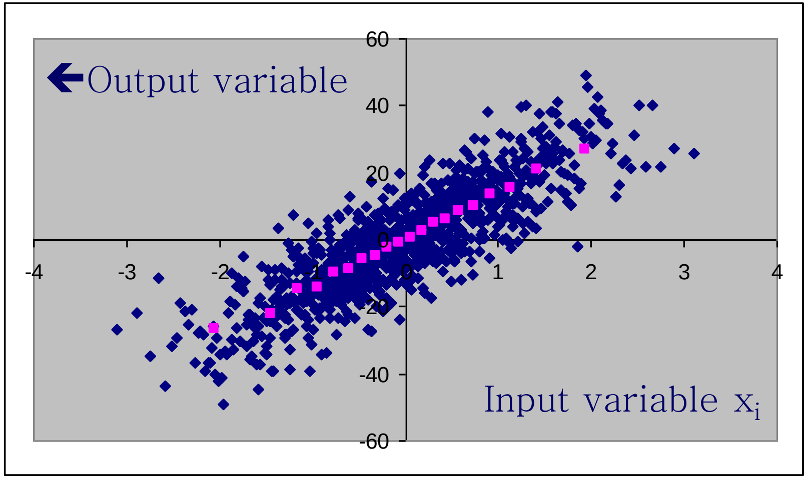
~1,000 blue points

Divide them in 20 bins of ~ 50 points

Compute the bin's average (pink dots)

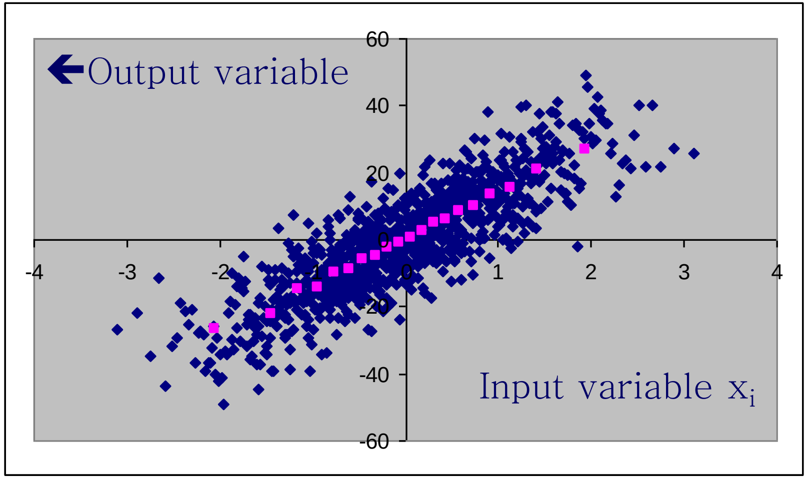


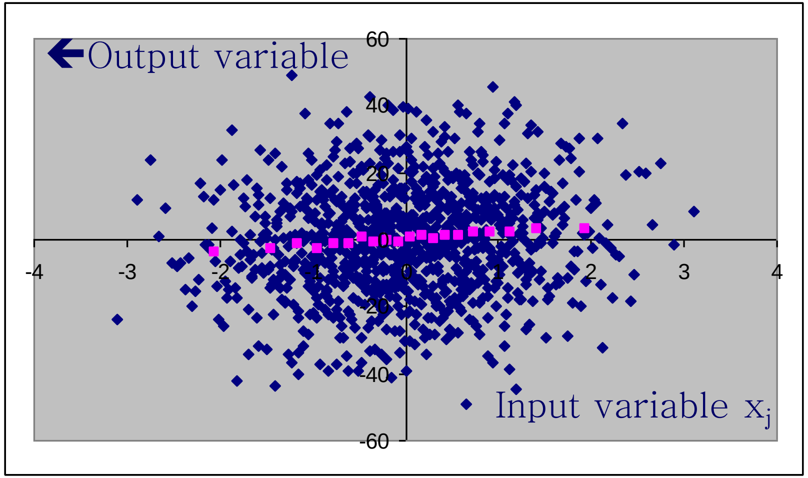
Each pink point is ~ $E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)$



Taking the variance of the pink points one obtains a sensitivity measure

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right)$$





Which factor has the highest

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right) ?$$

$$S_i = \frac{V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right)}{V(Y)}$$

The partial variance divided by the total variance is the so-called sensitivity index of the first order

Plenty of code available in R, MATLAB, and Phyton

https://cran.r-project.org/web/packages/sensitivity/sensitivity.pdf

https://cran.rstudio.com/web/packages/sensobol/index.html

https://www.uqlab.com/ (in MatLab, by Bruno Sudret and his team)

SALib https://salib.readthedocs.io/en/latest/

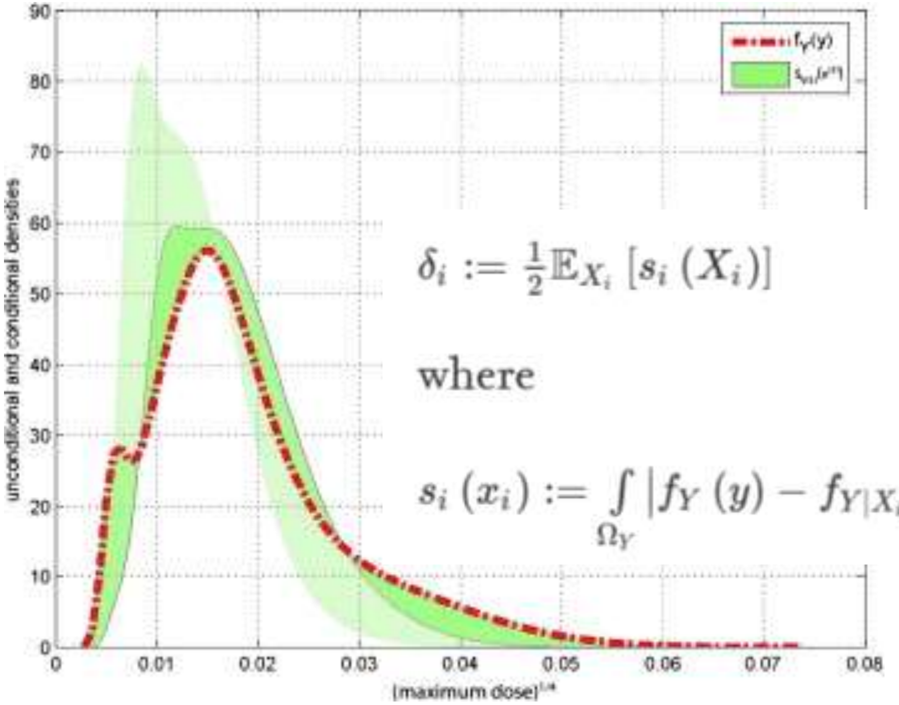
…but there is more, in R, Phython, Giulia …

Advantages with variance based methods:

- graphic interpretation scatterplots
- statistical interpretation (ANOVA)
- expressed plain English
- working with sets
- relation to settings such as factor fixing and factor prioritization
- give the effective dimension

Chapter 1 and its exercises

… but there are other methods that can be used for different settings, e.g. moment independents methods, Shapley coefficients, reduced spaces, VARS …



$$\delta_i := \frac{1}{2}\mathbb{E}_{X_i}\left[s_i\left(X_i\right)\right]$$

where

$$s_i\left(x_i\right) := \int_{\Omega_Y} \left|f_Y\left(y\right) - f_{Y|X_i=x_i}\left(y\right)\right| dy$$

ELSEVIER
Environmental Modelling & Software
Volume 34, June 2012, Pages 105-115

Model emulation and moment-independent sensitivity analysis: An application to environmental modelling

E. Borgonovo [a], W. Castaings [b, c], S. Tarantola [d]

Don’t use One factor At a Time (OAT)

A geometric proof

Environmental Modelling & Software 25 (2010) 1508–1517

Contents lists available at ScienceDirect

Environmental Modelling & Software

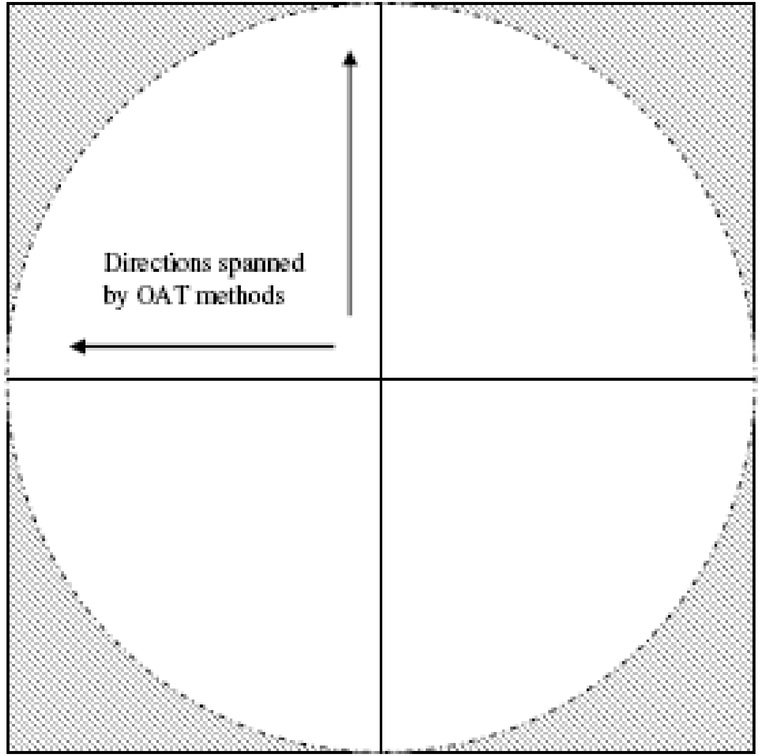
journal homepage: www.elsevier.com/locate/envsoft

How to avoid a perfunctory sensitivity analysis

Andrea Saltelli*, Paola Annoni

Joint Research Center, Institute for the Protection and Security of the Citizen, via E.Fermi, 2749, Ispra VA 21027, Italy

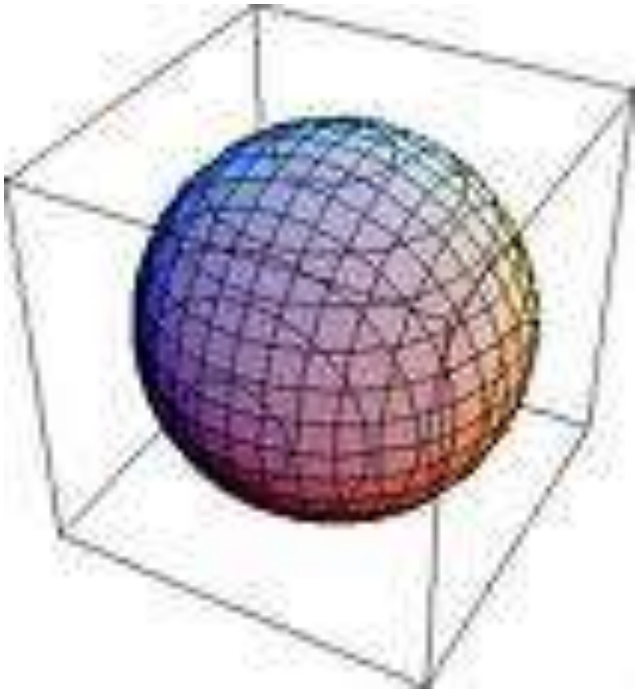
OAT in 2 dimensions



Area circle / area square =?

~ 3/4

OAT in 3 dimensions

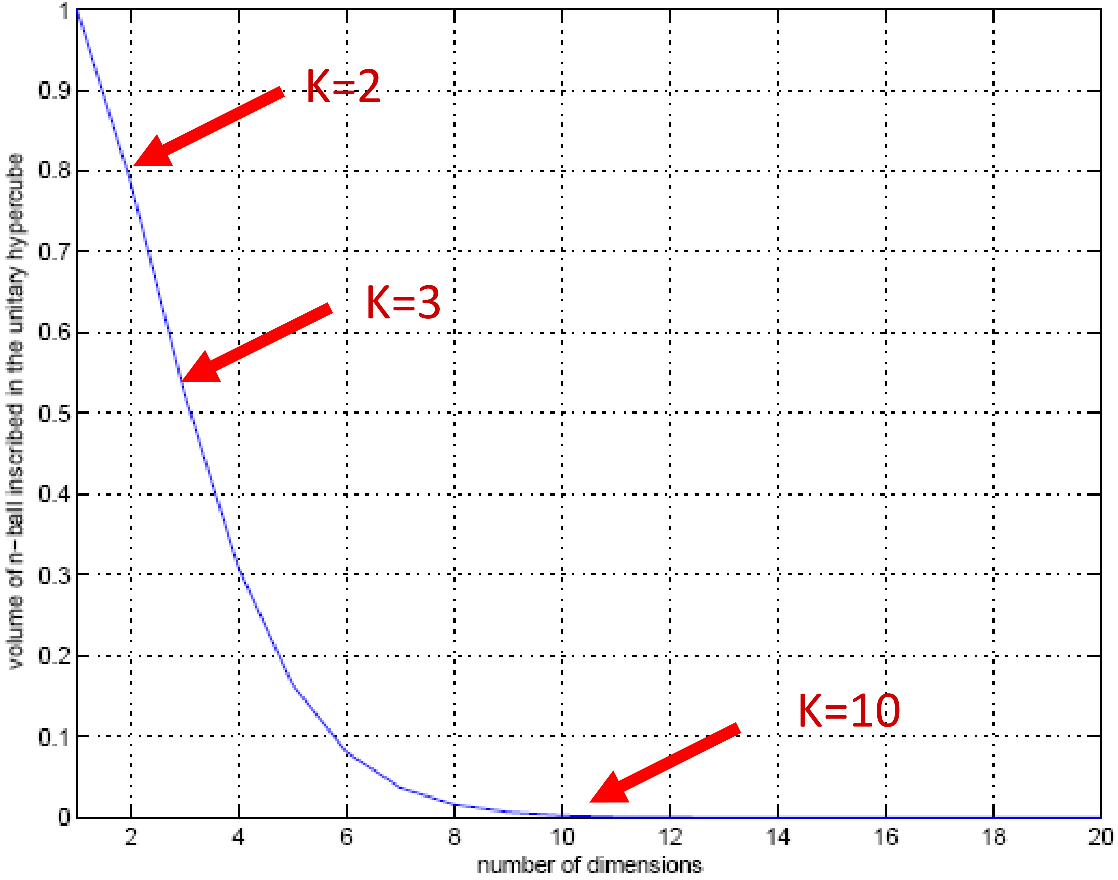


Volume sphere / volume cube =?

~ 1/2

OAT in 10 dimensions; Volume hypersphere / volume ten dimensional hypercube =? ~ 0.0025

OAT in k dimensions



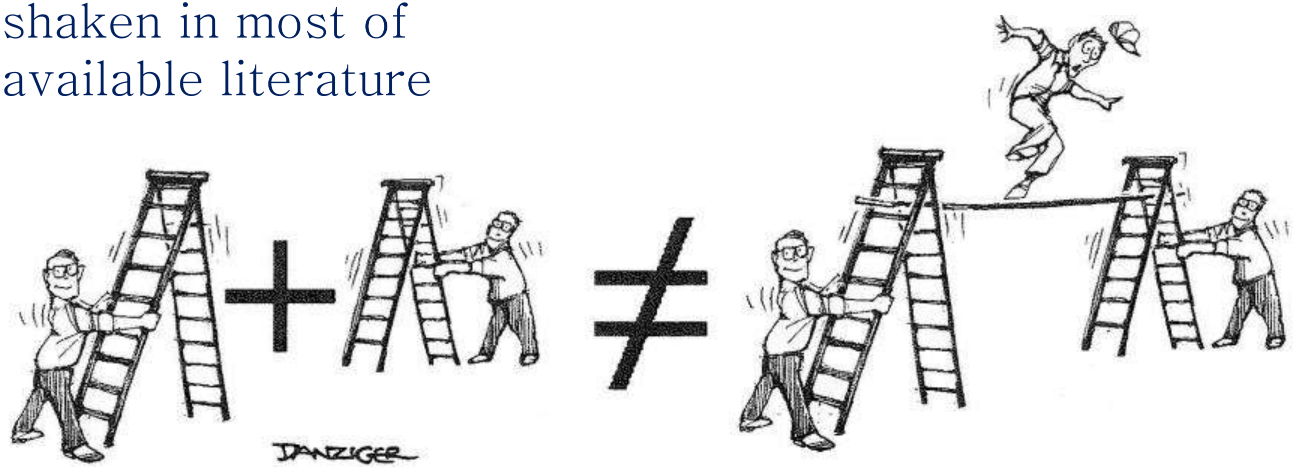
OAT does not capture interactions

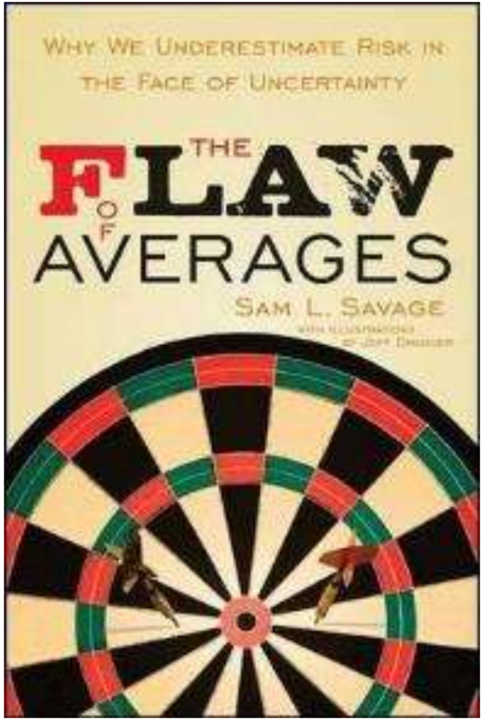
➔ The resulting analysis is non conservative

How would you test the scaffolding?

How coupled ladders are shaken in most of available literature

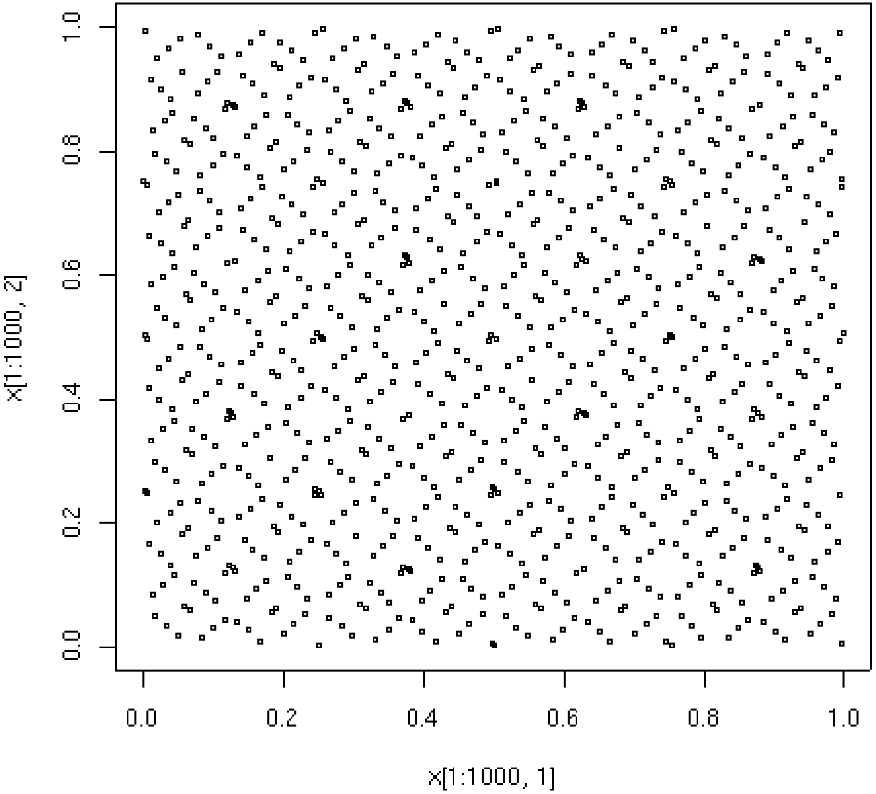
How to shake coupled ladders





Quasi random sequences

Ilya M. Sobol'



Cornell University

arXiv.org > stat > arXiv:1505.02350

Search...

Help | Adva

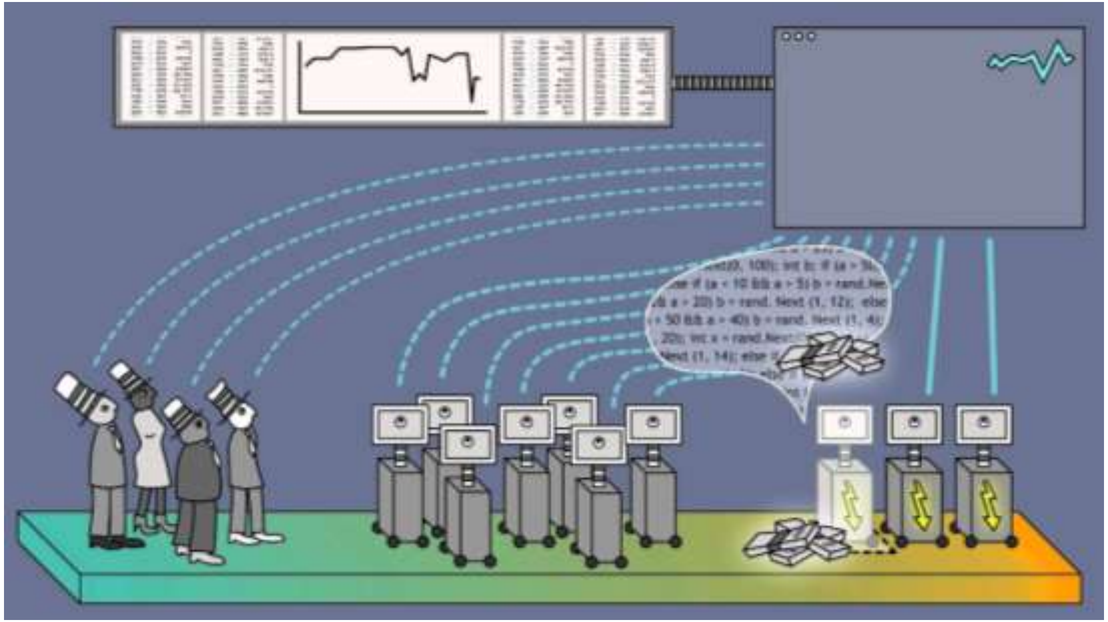
Statistics > Applications

[Submitted on 10 May 2015]

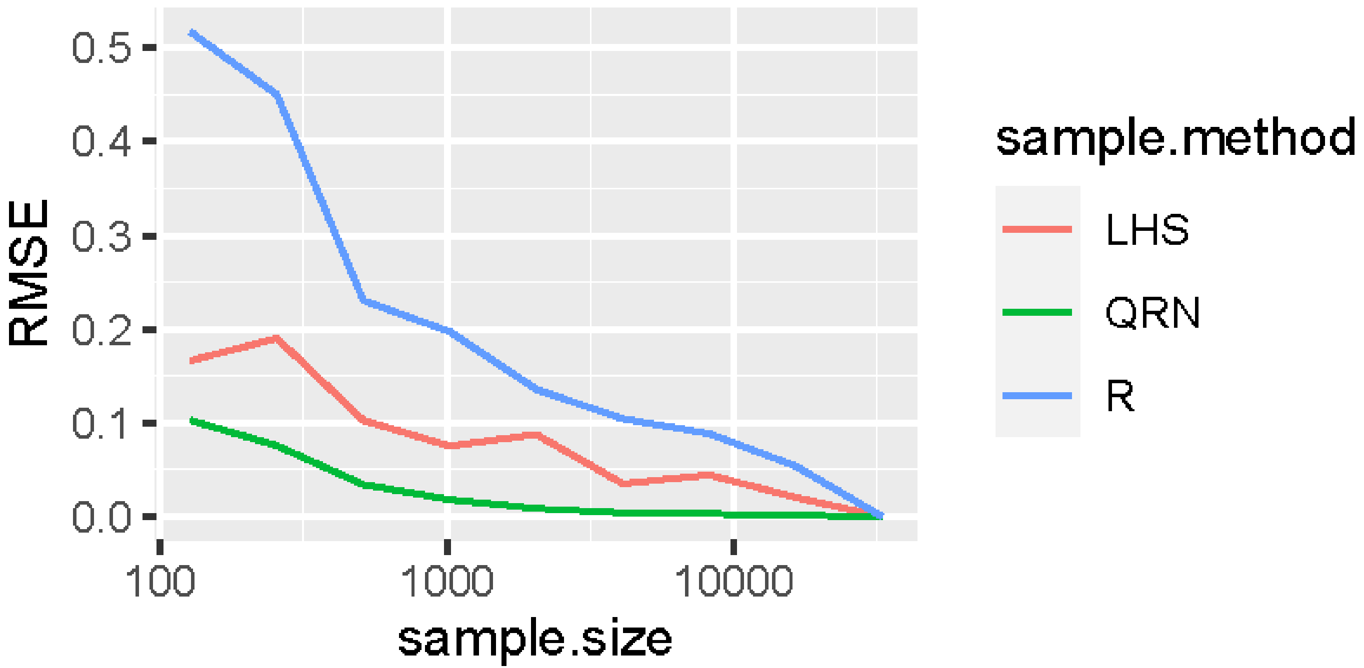
Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling Techniques

Sergei Kucherenko, Daniel Albrecht, Andrea Saltelli

Sobol' LP-TAU are used in high frequency trading



Source: https://www.youtube.com/watch?v=z4nCTdQlH



Root mean square error with different designs

Sensitivity analysis made easy

Cornell University

arXiv > stat > arXiv:2206.13470

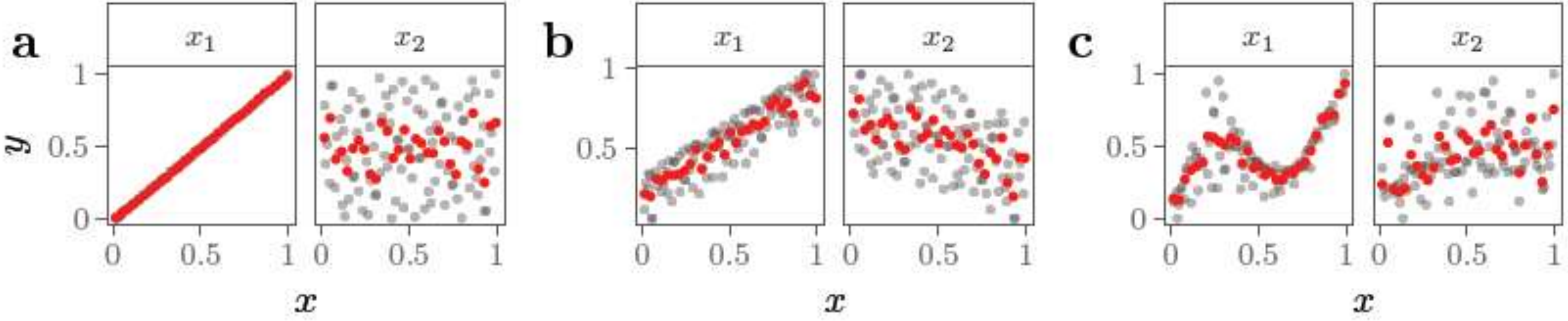
Statistics > Applications

[Submitted on 27 Jun 2022 (v1), last revised 17 Mar 2023 (this version, v2)]

Discrepancy measures for sensitivity analysis

Arnald Puy, Pamphile T. Roy, Andrea Saltelli

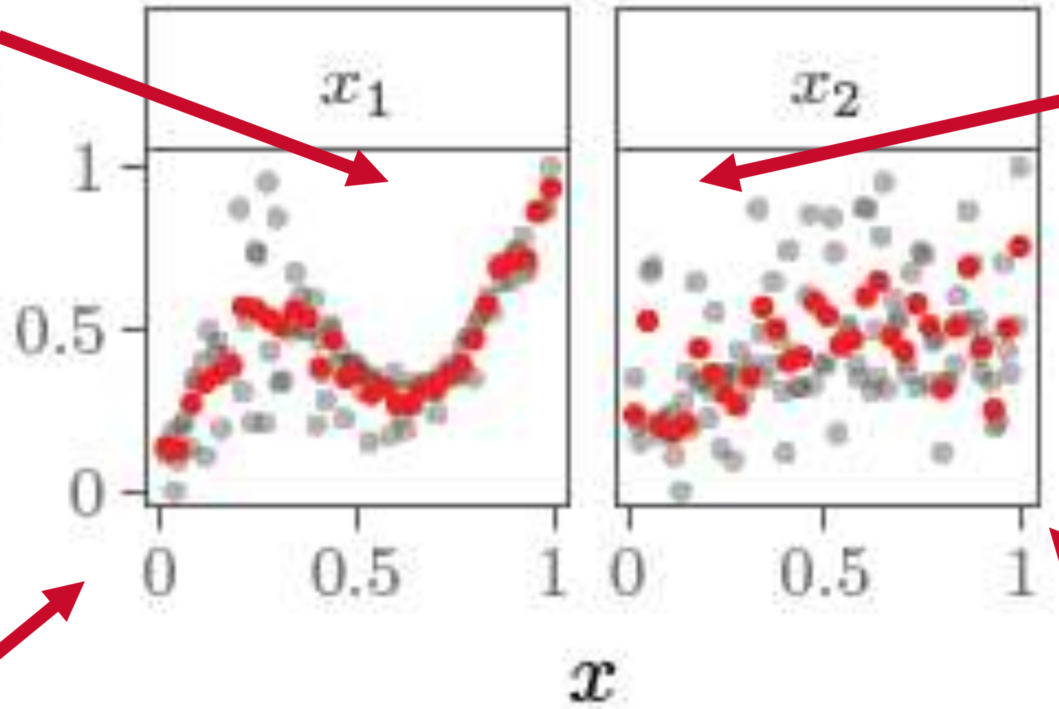
Do we need to compute indices? Can we do without statistics and calculus using the histograms we have met already?



'Stupid' histograms in the
x_i, y plane, both in [0,1], for different $y = f(x_i)$

Bigger 'holes'

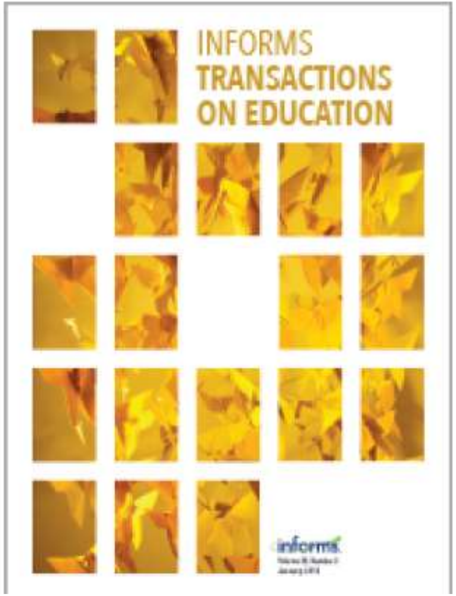
Smaller 'holes'



=more important

=less important

Another way to bypass statistics and calculus



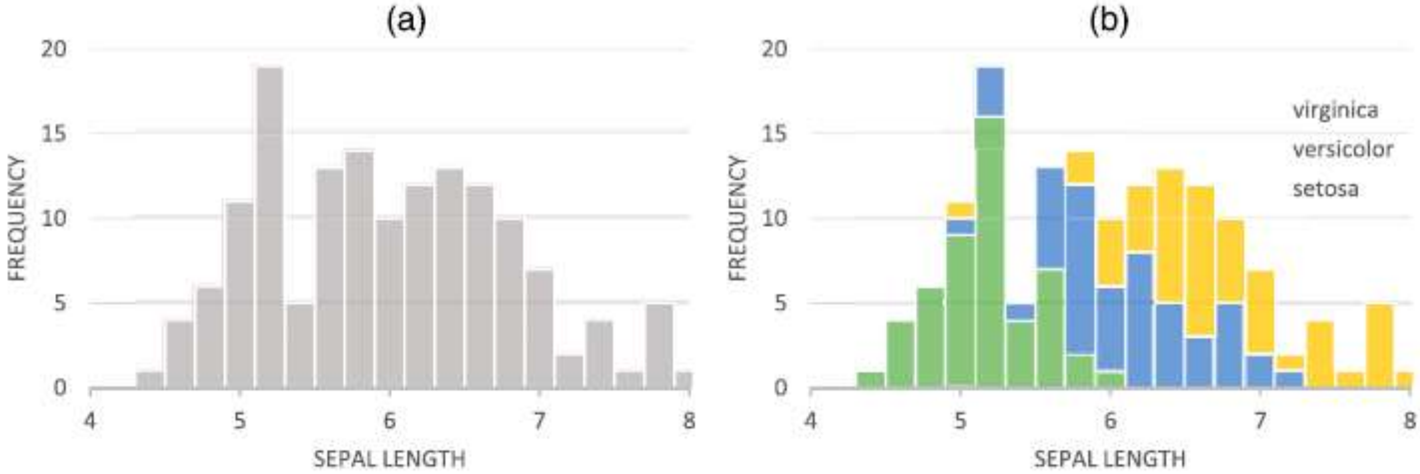
INFORMS Transactions on Education

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

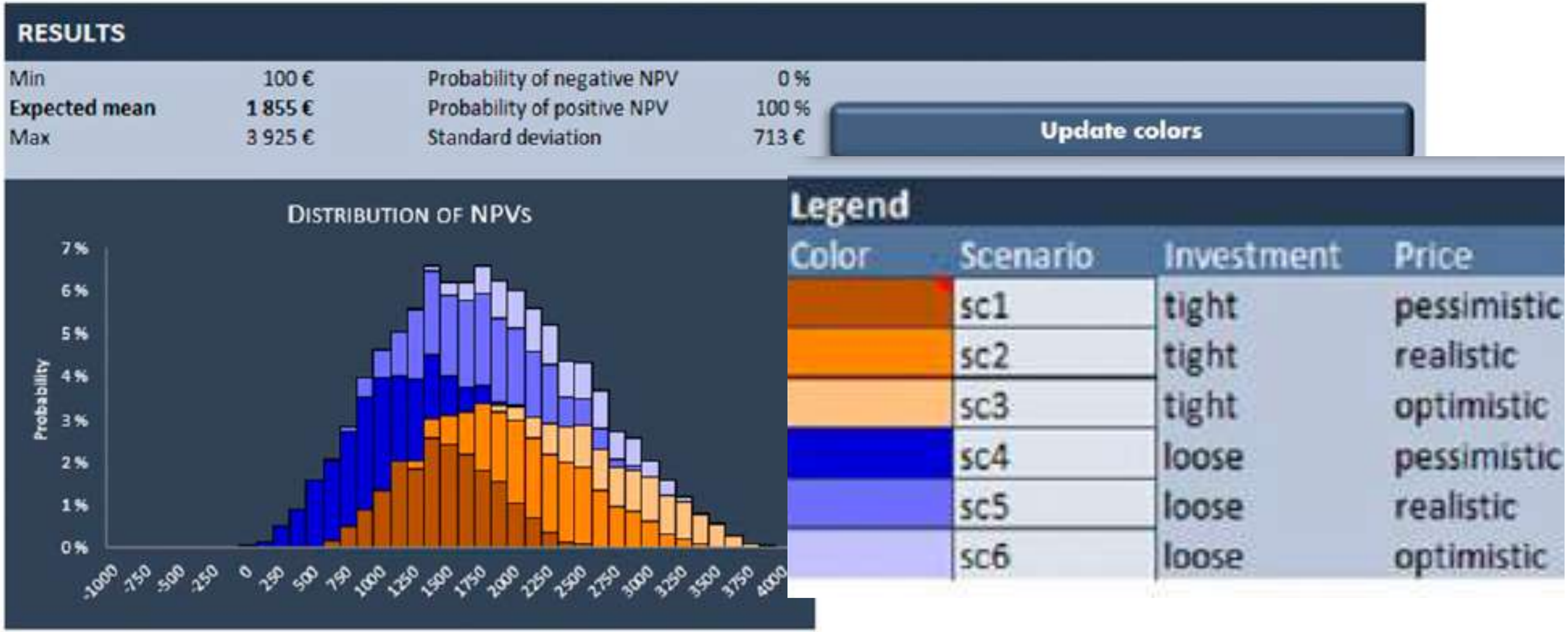
Monte Carlo Enhancement via Simulation Decomposition:
A “Must-Have” Inclusion for Many Disciplines

Mariia Kozlova, Julian Scott Yeomans

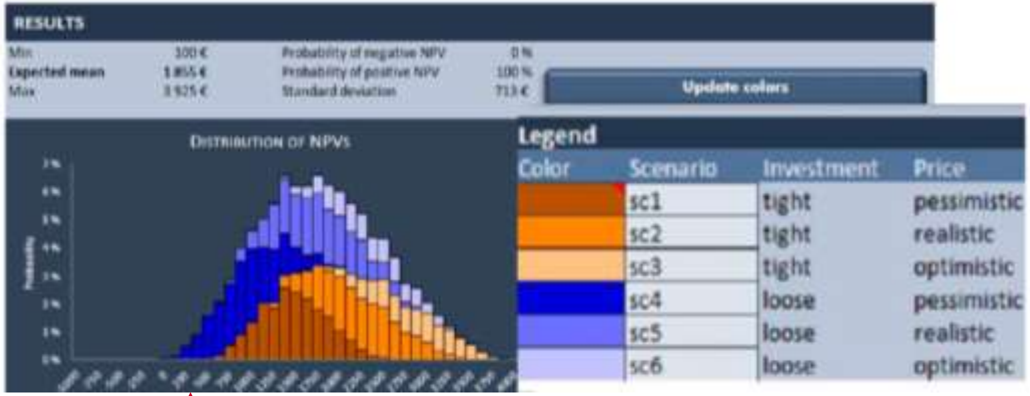
Colouring the output histogram can give sensitivity insights …



… without computing sensitivity indices



… without computing sensitivity indices



What is done here? We have two variables / options:

- Investment= ‘tight’ or ‘loose’
- Price=‘pessimistic’, ‘realistic’ or ‘optimistic’

Combing the 2 levels of investment with the three levels of price gives 2∗3=6 ‘scenarios’

➔ The possibility of very low returns (dark blue) corresponds to loose investment and pessimistic prices

Don't run the model just once

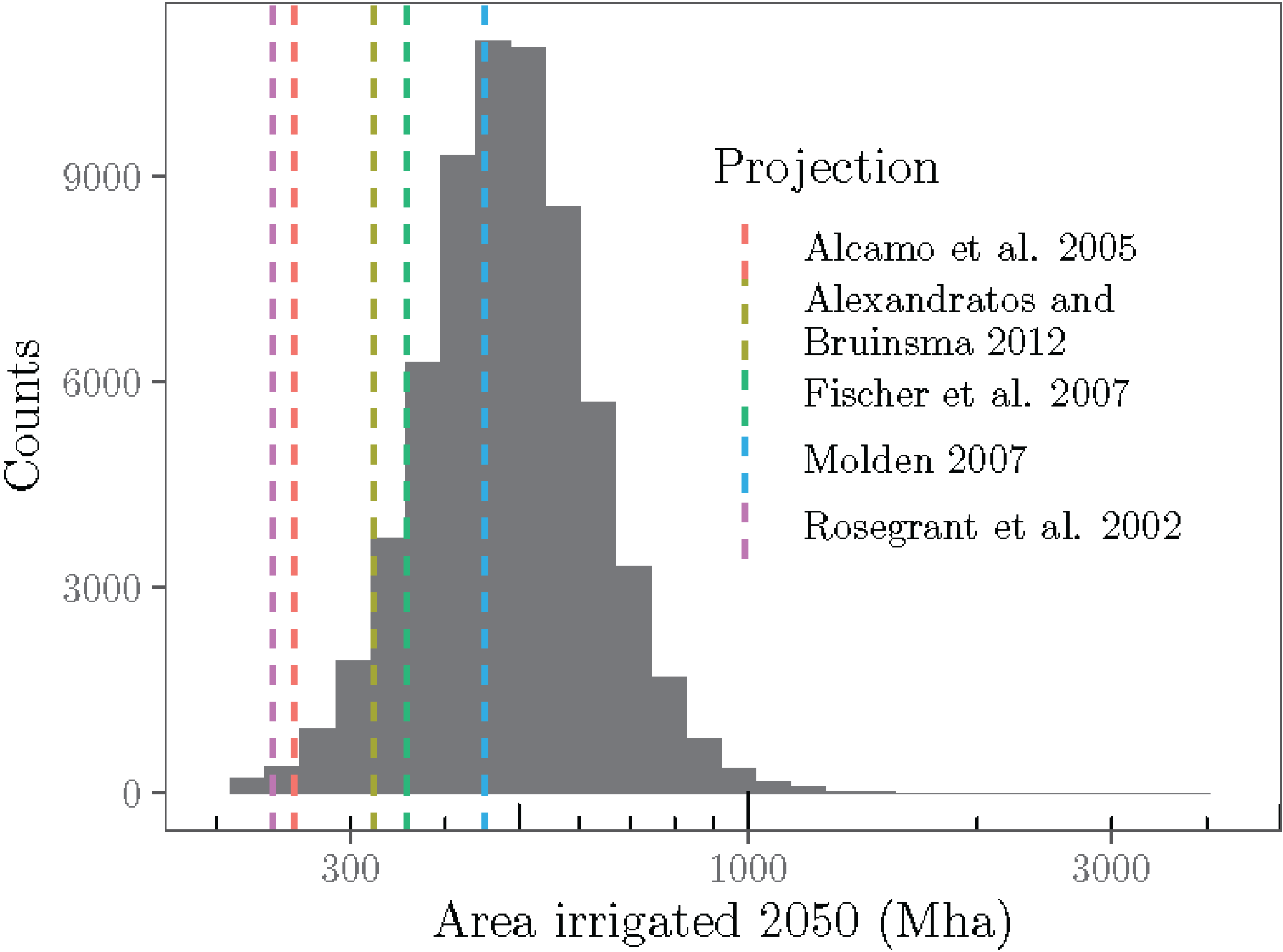
There is much to learn by running the model a few times, especially during model building

Lubarsky's Law of Cybernetic Entomology: there is always one more bug!

Model routinely used to produce point estimates may becomes non conservative when the uncertainty is plugged in

Current Models Underestimate Future Irrigated Areas

- How much land will need to be irrigated by the year 2050?
- Here the dashed lines represent deterministic model predictions from different models and datasets (from FAO & others organizations);
- An uncertainty analysis (grey histogram) reveals that the models are non-conservative: the need might be much larger



Citation:
Puy, A., Lo Piano, S., & Saltelli, A. (2020). Current models underestimate future irrigated areas. *Geophysical Research Letters, 47*, e2020GL087360. https://doi.org/10.1029/2020GL087360

Don’t sample just parameters and boundary conditions

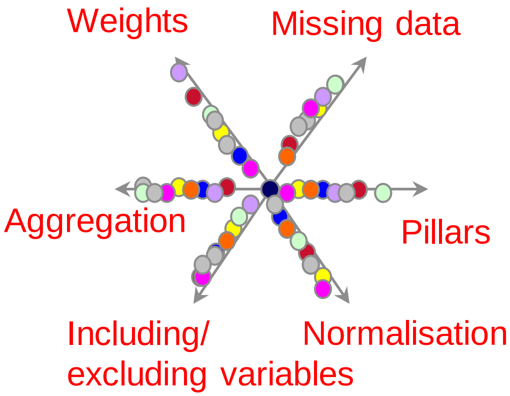
Explore thoroughly the space of the assumptions

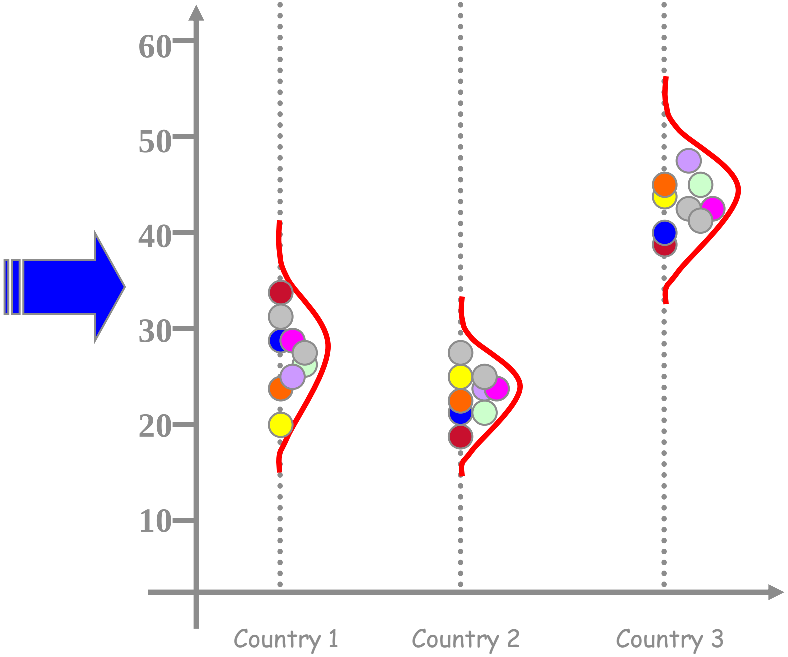
One can sample more than just factors:

- modelling assumptions,
- alternative data sets,
- resolution levels,
- scenarios …

Assumption	Alternatives
Number of indicators	▪ all six indicators included or one-at-time excluded (6 options)
Weighting method	▪ original set of weights, ▪ factor analysis, ▪ equal weighting, ▪ data envelopment analysis
Aggregation rule	▪ additive, ▪ multiplicative, ▪ Borda multi-criterion

Space of alternatives



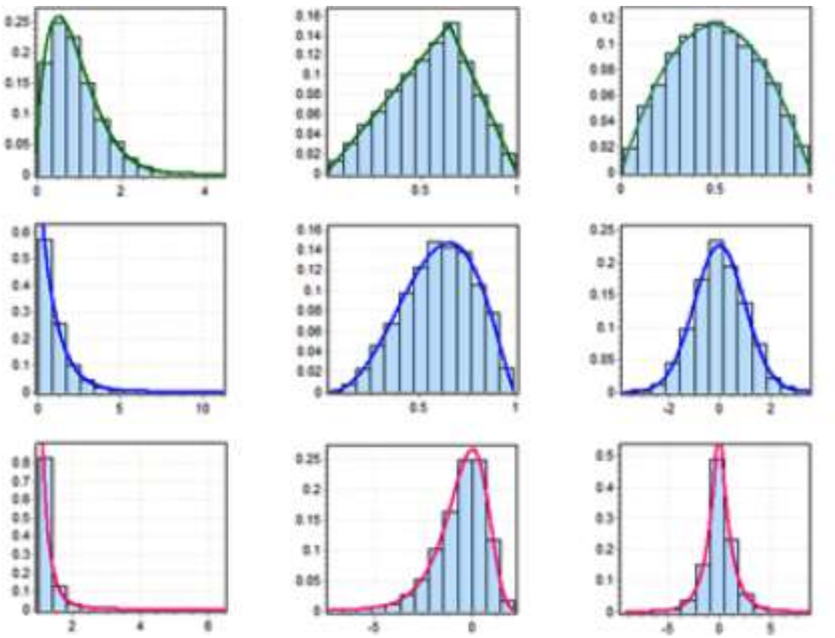


Building a Monte Carlo analysis

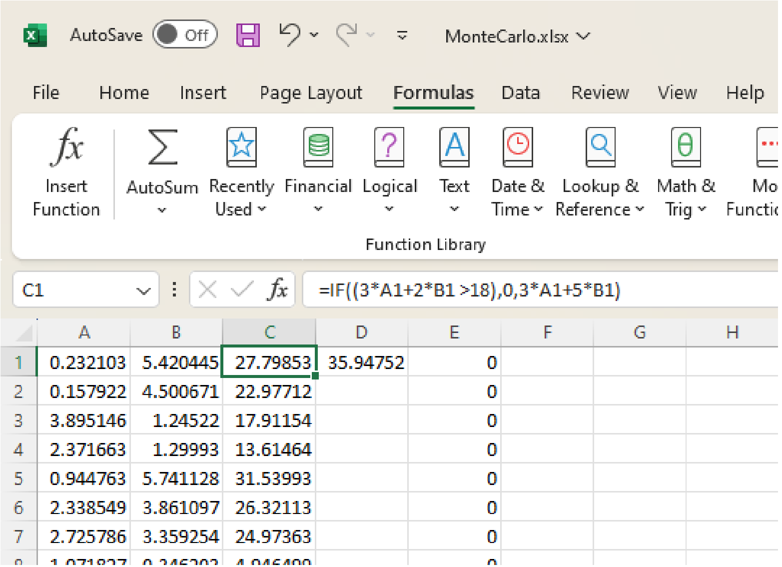
$$\begin{array}{ccc} x_{11} & x_{12\ldots} & x_{1k} \\ x_{21} & x_{22\ldots} & x_{2k} \\ \ldots & \ldots & \ldots \\ x_{N1} & x_{N2} & x_{Nk1} \end{array}$$

Input matrix: each column is a sample of size N from the distribution of a factor

Each row is a sample trial of size k to generate a value of y



Examples of distributions of input factors



$x_1 = 4 * rand()$

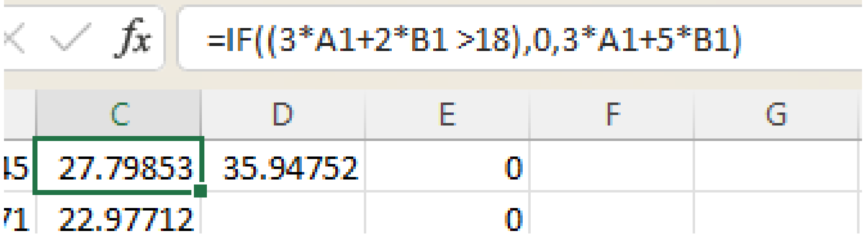
$x_2 = 6 * rand()$

Z in the feasible region

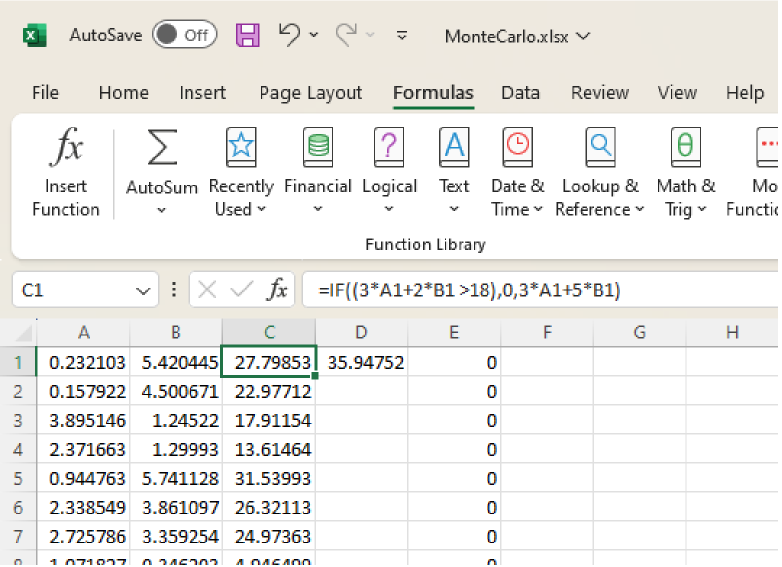
TABLE 3.1 Data for the Wyndor Glass Co. problem

Plant	Production Time per Batch, Hours — Product 1	Production Time per Batch, Hours — Product 2	Production Time Available per Week, Hours
1	1	0	4
2	0	2	12
3	3	2	18
Profit per batch	$3,000	$5,000	

I can search for the solution of our classic example generating x_1 and x_2 in the feasible region, finding an approximate solution for Z



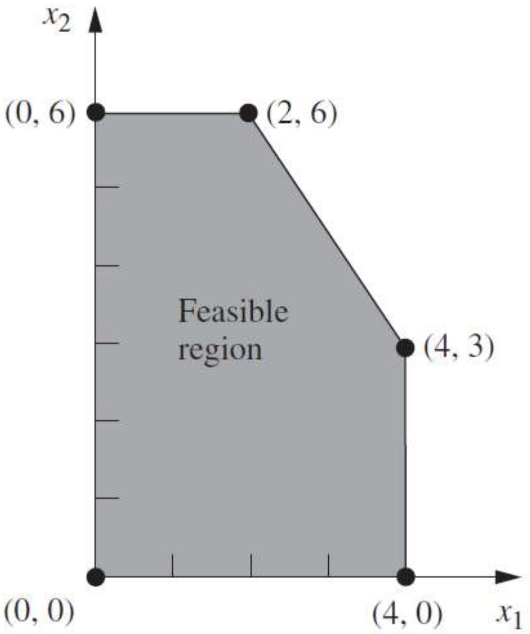
BARCELONA SCHOOL OF MANAGEMENT

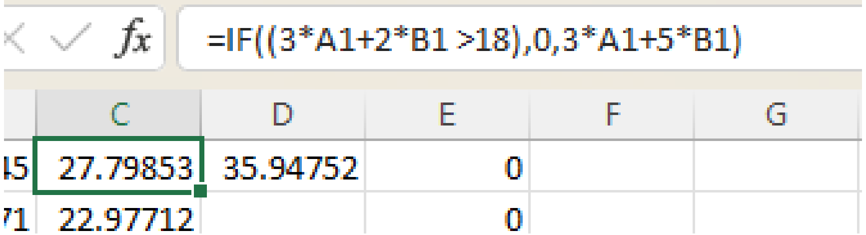


$x_1 = 4 * rand()$

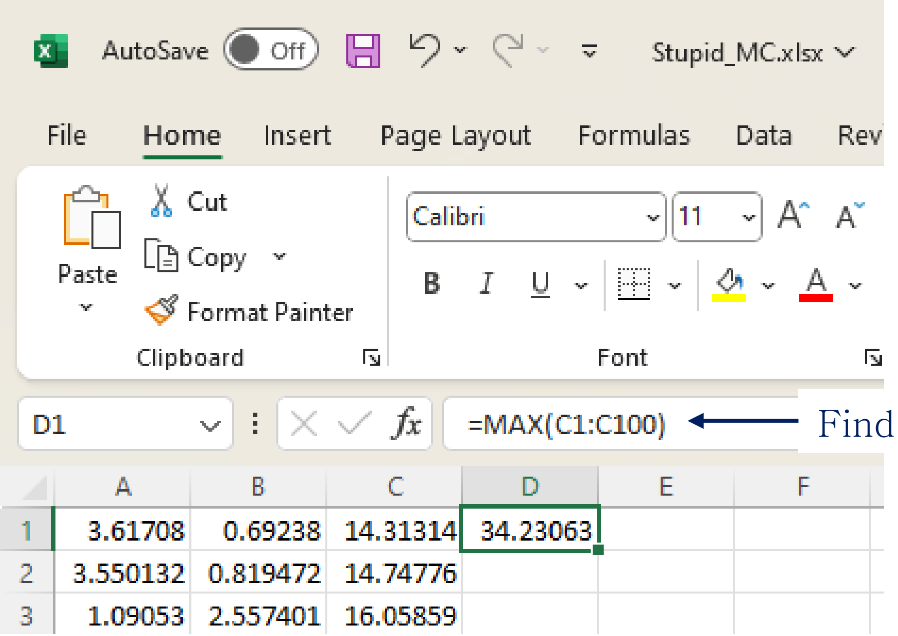
$x_2 = 6 * rand()$

Z in the feasible region





upf. BARCELONA SCHOOL OF MANAGEMENT

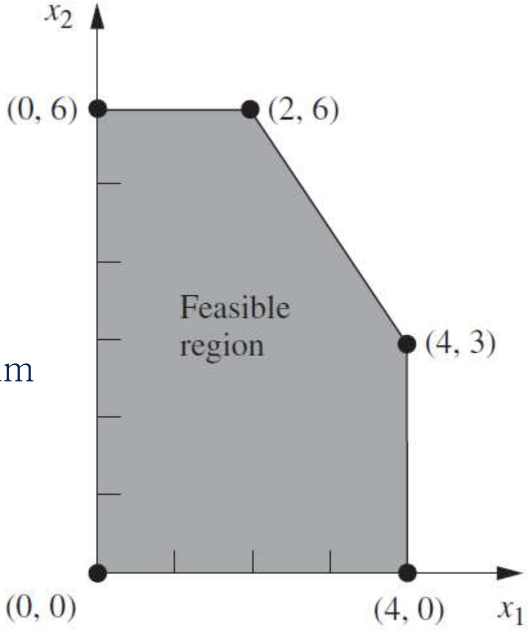


Find the maximum

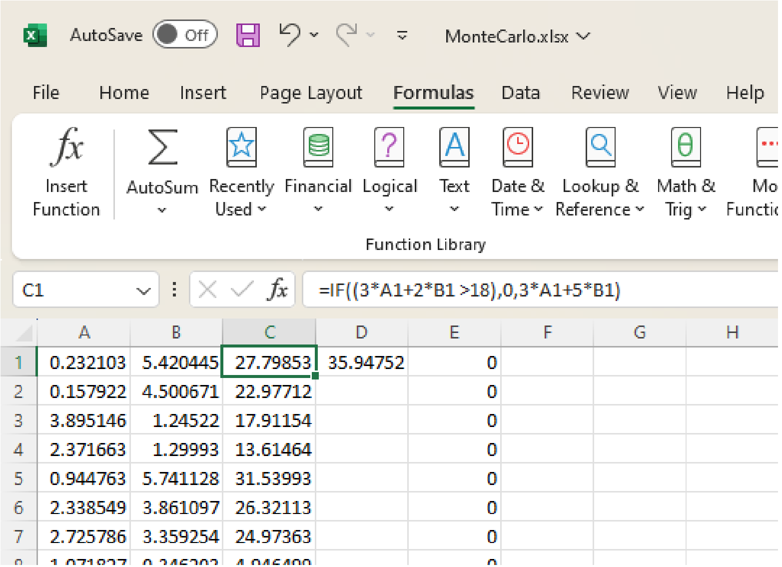
$x_1 = 4 * rand()$

$x_2 = 6 * rand()$

Z in the feasible region



upf. BARCELONA SCHOOL OF MANAGEMENT

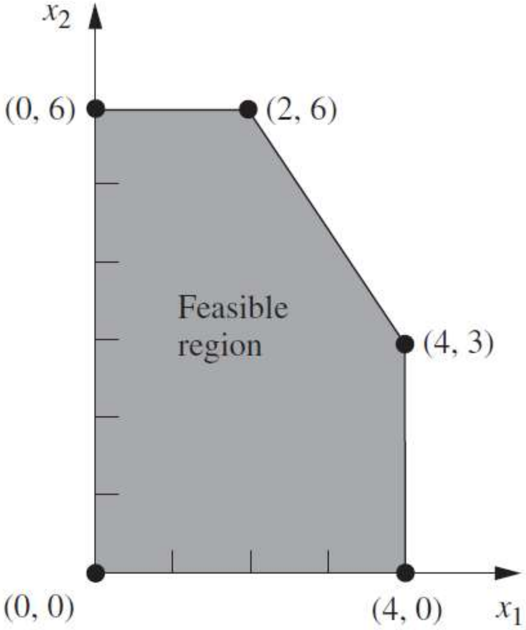


	A	B	C	D	E	F	G	H
1	0.232103	5.420445	27.79853	35.94752	0			
2	0.157922	4.500671	22.97712		0			
3	3.895146	1.24522	17.91154		0			
4	2.371663	1.29993	13.61464		0			
5	0.944763	5.741128	31.53993		0			
6	2.338549	3.861097	26.32113		0			
7	2.725786	3.359254	24.97363		0			

$x_1 = 4 * rand()$

Z in the feasible region

$x_2 = 6 * rand()$



In this case I find Z=35.9 for $x_1 = 1.9$ and $x_2 = 5.9$

Not terribly useful in this case but with this approach one can change the **11** uncertain inputs simultaneously exploring the uncertain in Z

NEVER vary all factors of the same amount

Be it 5%, 10%, or 20%

New WHO estimates: Up to 190 000 people could die of COVID-19 in Africa if not controlled

07 May 2020

Brazzaville – Eighty-three thousand to 190 000 people in Africa could die of COVID-19 and 29 million to 44 million could get infected in the first year of the pandemic if containment measures fail, a new study by the World Health Organization (WHO) Regional Office for Africa finds. The research, which is based on prediction modelling, looks at 47 countries in the

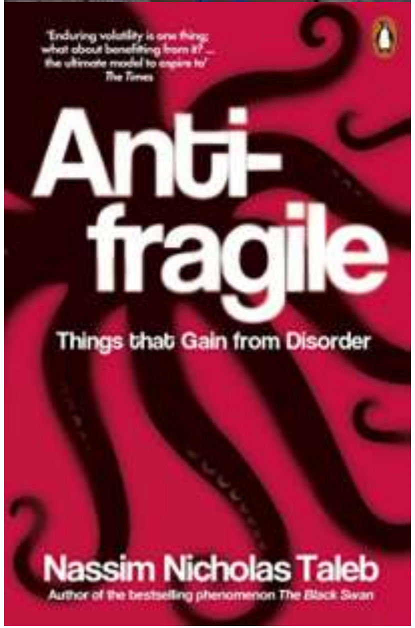
Speculative scenario in which ten uncertain input probabilities are increased by an arbitrary 10% — as if they were truly equally uncertain — with no theoretical or empirical basis for such a choice

World Health Organization
REGIONAL OFFICE FOR Africa

In a numerical experiment relating to a real-life application the range of uncertainty of each input is crucial input to the analysis, and often the most expensive to get

Suggested reading:

- Nassim N. Taleb's books, and his *via negativa,* the science of what is not;
- A paper on why most sensitivity analyses fail




Environmental Modelling & Software

Volume 114, April 2019, Pages 29-39

ELSEVIER

Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices

Andrea Saltelli, Ksenia Aleksankina, William Becker, Pamela Fennell, Federico Ferretti, Niels Holst, Sushan Li, Qiongli Wu

How to play with Python?

Install https://www.anaconda.com/download
Open Anaconda
Run Spyder

Link:
On windows https://www.youtube.com/watch?v=UTqOXwAi1pE
On mac https://www.youtube.com/watch?v=0xYWWFOEBi8
On linux https://www.youtube.com/watch?v=7-naqq9fvZE

my_string = "Hello, World!"
print(my_string)
my_string = "Hello, World!"
print(my_string)

Install packages

https://www.tutorialspoint.com/how-do-i-install-python-packages-in-anaconda

- Import packages:

```
import numpy as np
```

- Try some simple command

```
my_string = "Hello, World!"
print(my_string)
```

BARCELONA SCHOOL OF MANAGEMENT

Homework

- Read slides 66 to 82; what sense can you make (or not make) if it? Write a page (manuscript) on this, possibly in the form "I understood that…" and "The part I did not understand is…"
- Compute the chance of having as many as 5 heads throwing a coin 8 times and write down procedure and result
- Use the slides just given install Anaconda, launch Spider and run some simple Phyton scripts to prepare for work in class next lesson

Thank you

www.andreasaltelli.eu
https://orcid.org/0000-0003-4222-6975
@AndreaSaltelli@mstdn.social
https://www.youtube.com/channel/UCz26ZK04xchekUy4GevA3DA

upf. BARCELONA SCHOOL OF MANAGEMENT