# Máster Universitario en Administración y Dirección de Empresas

# Full Time MBA

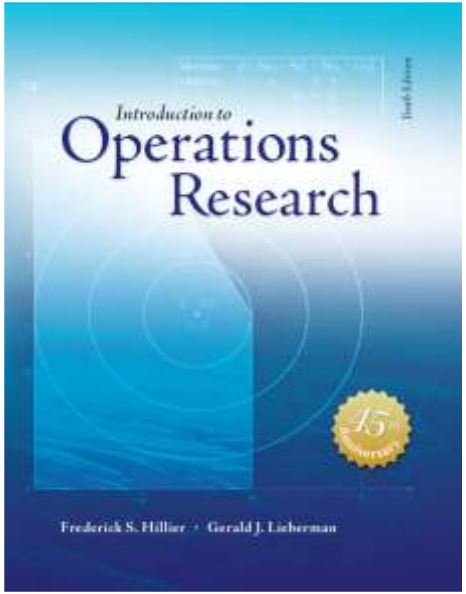
## Quantitative methods for decision making

Professor Andrea Saltelli

upf. BARCELONA SCHOOL OF MANAGEMENT

# Elements of quantification for decision making with emphasis on operation research

upf. BARCELONA SCHOOL OF MANAGEMENT




# Important – please use this newer version

Download this version (10th)
https://www.andreasaltelli.eu/file/repository/Introduction_to_Operations_Research_10th_Frederick_S_Hillier.pdf

# Notes from the homework

All delivered and accepted

Two (Claudia Cuevas and Michel Maynard) gave references for the bio.

Some discussed the bridges problem by reflecting on whether one should or not be in a closed loop or instead cross the river. Good, though this was not in the assignment (see Lesson one on extra assumptions made by the analyst).

Madoka Shimura extended the discussion of the seven bridges to Euler's work on cyclic graphs.

# Notes from the homework

Yay Patel noted the Newton–Leibniz concurrent discovery calculus

Isac Newton
1643-1727

Gottfried Wilhelm
Leibniz 1646-1716

**Historical note:** this is also the history of a bitter controversy where one institution, the Royal Society, led by the Newton, weighted in in favour of the same Newton against Leibniz who was unjustly accused of plagiarism

# Notes from the homework

Yitian Deng and others spoke about **John Graunt** and the 1.08 birth rate boys versus girls.

**Historical note:** in 1713 **William Derham**, an English clergyman and natural philosopher argued that this was a sign of God's providence because men were more likely to die earlier.

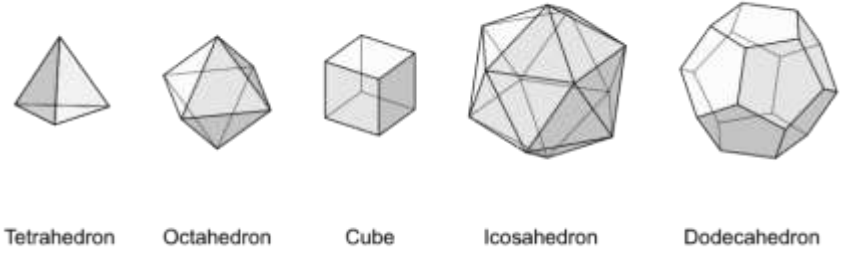
Quin Yun good discussion of **Isac Newton** and the necessity of alchemy to his theory. She also notes Newton's religiosity – he can explain Kepler's heliocentric system and motions but God is the maker of the heavens.

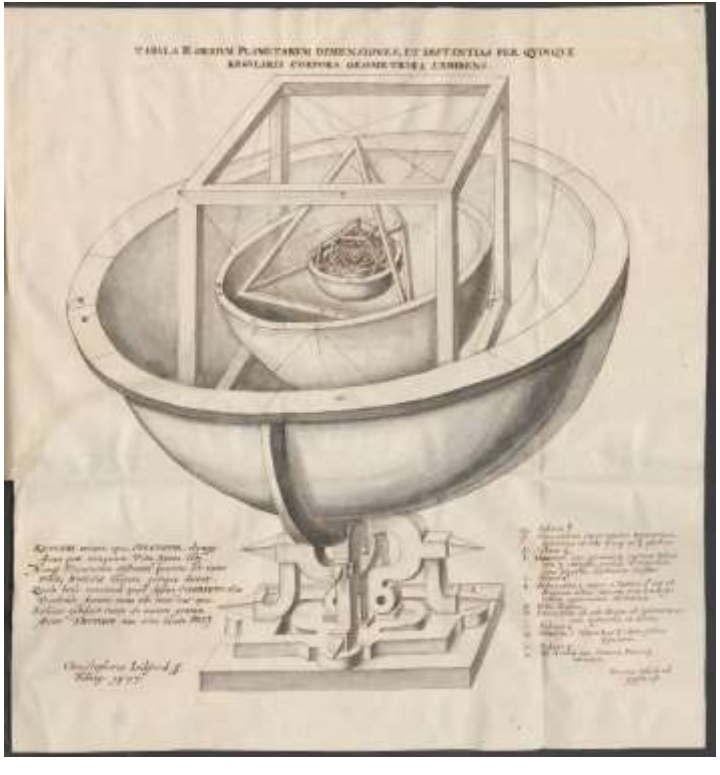

John Graunt, founder of demography
1620-1674

# Notes from the homework

## Kepler though that the ‘heavens’ followed the geometry of the platonic solids

Johannes Kepler
1571 –1630



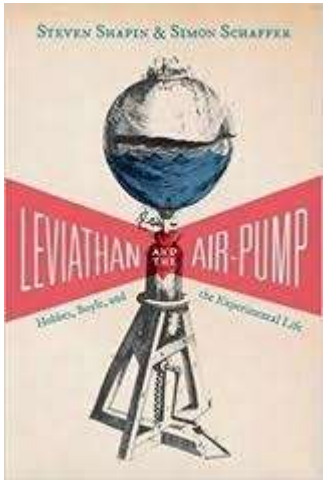
# Digression

Common religious (or superstitious) beliefs among the protagonists of the scientific revolution, see **Francis Bacon**'s 'Prayer or Psalm', or **Robert Boyle** and his work on demons …

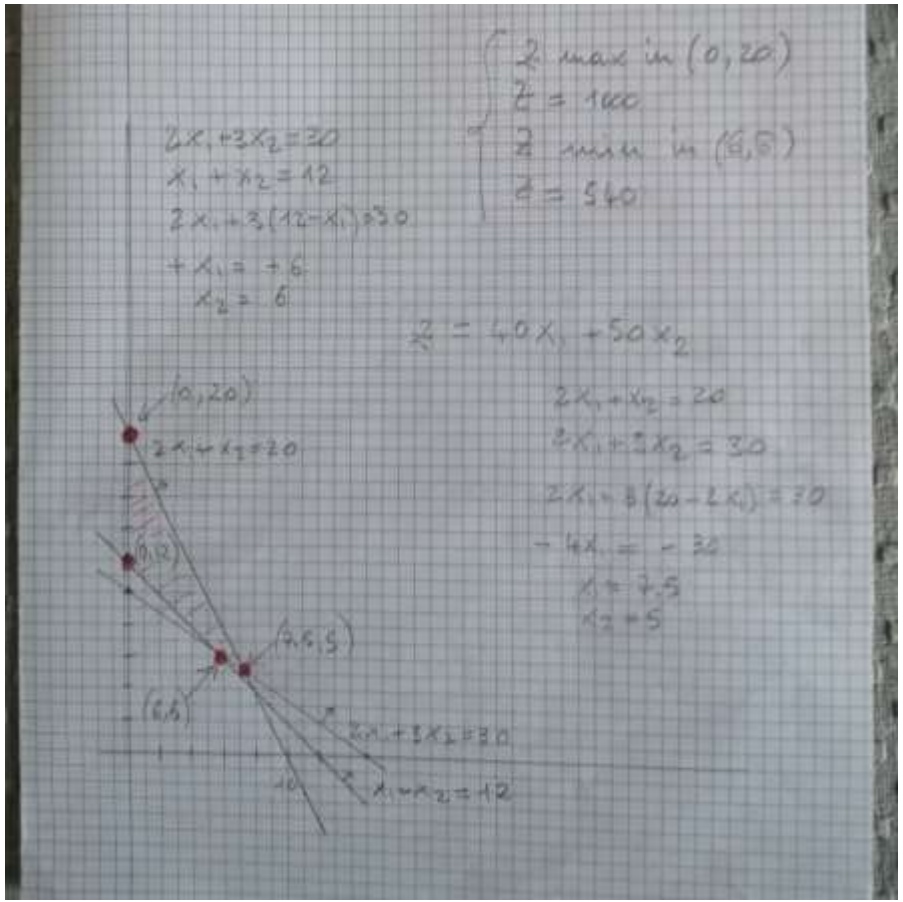
Francis Bacon
1617-1621

Robert Boyle
1627-1691

Protagonist of another famous controversy with Thomas Hobbes on the virtues of experimental method



# Solutions homework lesson 2




Consider the following model: Maximize

$$Z = 40x_1 + 50x_2$$

subject to

$$2x_1 + 3x_2 \geq 30$$ wrong

$$x_1 + x_2 \geq 12$$

$$2x_1 + x_2 \geq 20$$

and

$$x_1 \geq 0$$

$$x_2 \geq 0$$

There are four CPF (red) and the maximum corresponds to (0,20), $Z = 1{,}000$

I should have asked the minimum, that is more laborious to find in (6,6), $Z = 540$

# Solutions homework lesson 2

## The nut-mix problem

*The nut-mix problem of Charnes and Cooper (1953):*

A manufacturer wishes to determine an optimal program for mixing three grades [A, B, D] of nuts consisting of cashews [C], hazels [H], and peanuts [P] according to the specifications and prices given in table 1. Hazels may be introduced into the mixture in any quantity, provided the specifications are met. The amounts of each nut available each day and their costs are given in table 2. Determine the pounds of each mixture that should be manufactured each day to maximize the gross return (contribution margin).

Page 94 Gass, S. I., & Assad, A. A. (2006). *An Annotated Timeline Of Operations Research: An Informal History* (1st Corrected ed. 2005. Corr. 2nd printing 2006 edition). Springer-Verlag New York Inc.

upf. BARCELONA SCHOOL OF MANAGEMENT

# Solutions homework lesson 2

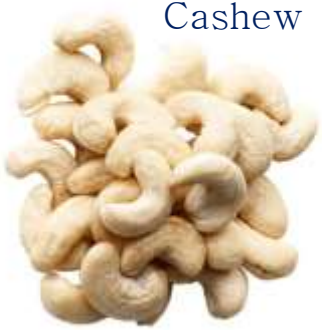
Table 1

| Mixture | Specifications | Selling price: ¢/pound |
|---|---|---|
| A | Not less than 50% cashews<br>Not more than 25% peanuts | 50 |
| B | Not less than 25% cashews<br>Not more than 50% peanuts | 35 |
| D | No specifications | 25 |

Peanuts

https://www.nutsforlife.com.au

Cashew



Source: https://www.cashews.org

# The nut-mix problem

Hazels

Source: https://www.woodlandtrust.org.uk/

Table 2

| Inputs | Capacity: pounds/day | Price: ¢/pound |
|---|---|---|
| C | 100 | 65 |
| H | 60 | 35 |
| P | 100 | 25 |
| Total | 260 | |

upf. BARCELONA SCHOOL OF MANAGEMENT

# Solutions homework lesson 2

Table 1

| Mixture | Specifications | Selling price: ¢/pound |
|---|---|---|
| A | Not less than 50% cashews<br>Not more than 25% peanuts | 50 |
| B | Not less than 25% cashews<br>Not more than 50% peanuts | 35 |
| D | No specifications | 25 |

Table 2

| Inputs | Capacity: pounds/day | Price: ¢/pound |
|---|---|---|
| C | 100 | 65 |
| H | 60 | 35 |
| P | 100 | 25 |
| Total | 260 | |

## Hint 1

Reckon in terms of pounds per day of the three nuts type

## Hint 2

C pounds cashew/day
H pounds hazels/day
P pounds peanuts/day
---
C_A pounds cashew/day in A
C_B pounds cashew/day in B
...
C_P pounds peanuts/day in C
**(nine variables)**

# Solutions homework lesson 2

Table 1

| Mixture | Specifications | Selling price: ¢/pound |
|---|---|---|
| A | Not less than 50% cashews<br>Not more than 25% peanuts | 50 |
| B | Not less than 25% cashews<br>Not more than 50% peanuts | 35 |
| D | No specifications | 25 |

Table 2

| Inputs | Capacity: pounds/day | Price: ¢/pound |
|---|---|---|
| C | 100 | 65 |
| H | 60 | 35 |
| P | 100 | 25 |
| Total | 260 | |

## My solution

Maximize

$50(C\_A+H\_A+P\_A)-(65C\_A+35H\_A+25P\_A)+$
$35(C\_B+H\_B+P\_C)-(65C\_B+35H\_B+25P\_B)+$
$25(C\_D+H\_D+P\_D)-(65C\_D+35H\_D+25P\_D)$

Subject to

$C\_A+C\_B+C\_D \leq 100$
$H\_A+H\_B+H\_D \leq 60$
$P\_A+P\_B+P\_D \leq 100$

… how about these specifications?

# Solutions homework lesson 2

Table 1

| Mixture | Specifications | Selling price: ¢/pound |
|---|---|---|
| A | Not less than 50% cashews<br>Not more than 25% peanuts | 50 |
| B | Not less than 25% cashews<br>Not more than 50% peanuts | 35 |
| D | No specifications | 25 |

Table 2

| Inputs | Capacity: pounds/day | Price: ¢/pound |
|---|---|---|
| C | 100 | 65 |
| H | 60 | 35 |
| P | 100 | 25 |
| Total | 260 | |

## My solution

Maximize

$50(C\_A+H\_A+P\_A)-(65C\_A+35H\_A+25P\_A)+$
$35(C\_B+H\_B+P\_C)-(65C\_B+35H\_B+25P\_B)+$
$25(C\_D+H\_D+P\_D)-(65C\_D+35H\_D+25P\_D)$

Subject to

$C\_A+C\_B+C\_D\leq 100$
$H\_A+H\_B+H\_D\leq 60$
$P\_A+P\_B+P\_D\leq 100$

$$\frac{C\_A}{C\_A+H\_A+P\_A}\geq 0.5 \qquad \frac{C\_B}{C\_B+H\_B+P\_B}\geq 0.25$$

$$\frac{P\_A}{C\_A+H\_A+P\_A}\leq 0.25 \qquad \frac{P\_B}{C\_B+H\_B+P\_B}\leq 0.50$$

+ non-negativity constraints for all C's, H's and P's

**4.48** A statistical experiment has 10 equally likely outcomes that are denoted by 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. Let event $A = \{3, 4, 6, 9\}$ and event $B = \{1, 2, 5\}$.

a. Are events $A$ and $B$ mutually exclusive events?

b. Are events $A$ and $B$ independent events?

c. What are the complements of events $A$ and $B$, respectively, and their probabilities?

Yes, No, $\bar{A} = \{1,2,5,7,8,10\}, \bar{B} = \{1,4,6,7,8,9,10\}, P(\bar{A}) = \frac{6}{10}, P(\bar{B}) = \frac{7}{10}$

**4.70** Find the joint probability of $A$ and $B$ for the following.

a. $P(A) = .40$ and $P(B \mid A) = .25$

b. $P(B) = .65$ and $P(A \mid B) = .36$

$P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$ for Bayes

a) $.40 * .25 = .1$

b) $.65 * .36 = .234$

**4.76** Given that $P(A) = .30$ and $P(A \text{ and } B) = .24$, find $P(B \mid A)$.

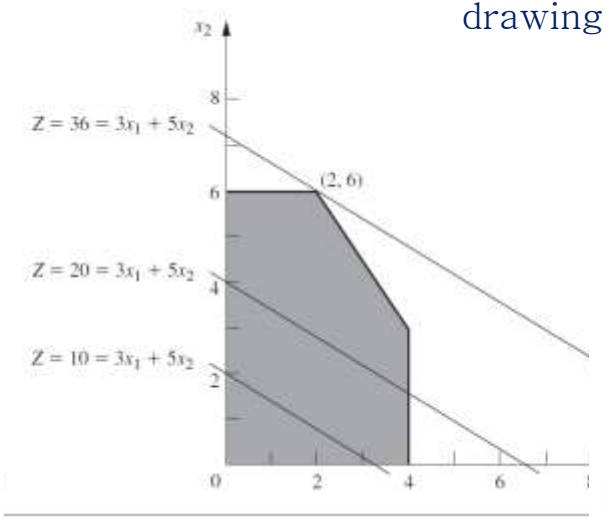
$P(A \cap B) = P(B \mid A)P(A)$ so $P(B \mid A) = .24/.30$

**4.99** Suppose that 20% of all adults in a small town live alone, and 8% of the adults live alone and have at least one pet. What is the probability that a randomly selected adult from this town has at least one pet given that this adult lives alone?

$P(\text{alone} \cap \text{pet}) = P(\text{pet} \mid \text{alone})P(\text{alone})$ so $P(\text{pet} \mid \text{alone}) = 8/20$

# What should we recall from the previous lesson?

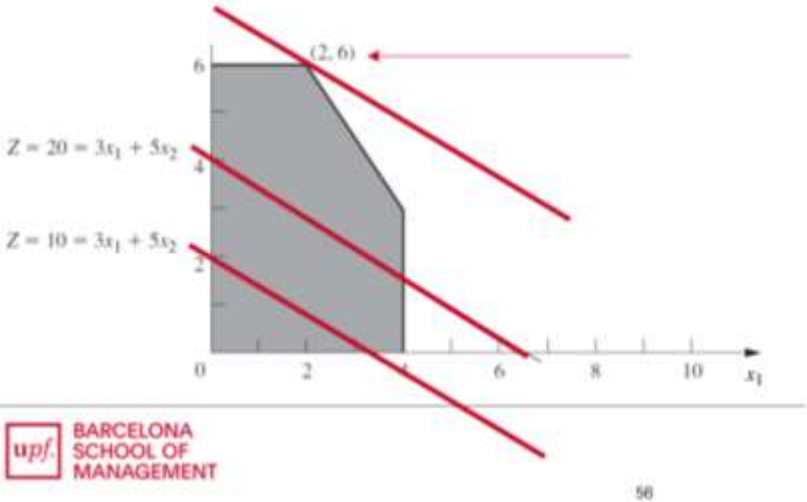
We solved this graphically by identifying the feasible region, then drawing some possible lines for Z



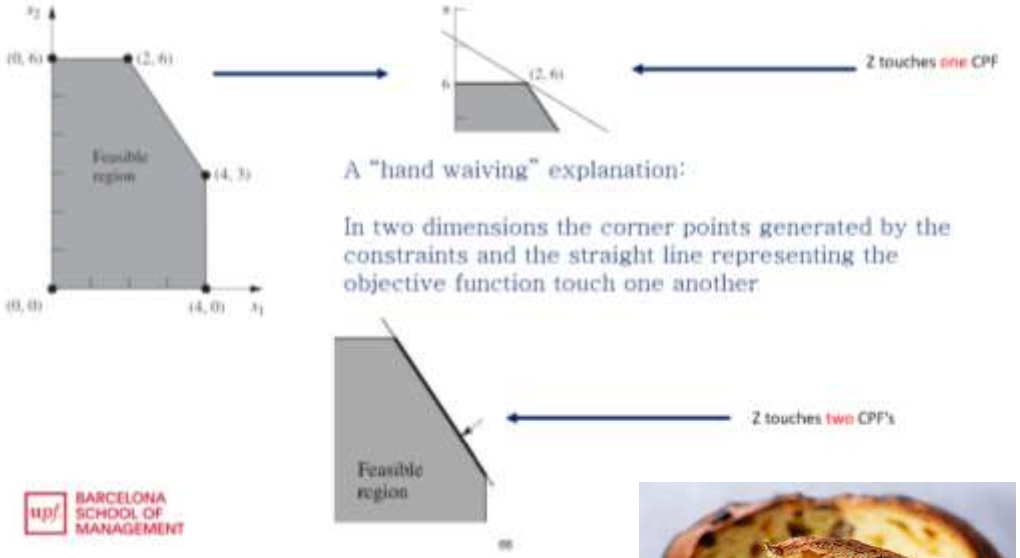

First we need to note that Z grows from bottom up.

Then to note that if we must respect $Z = 3x_1 + 5x_2$ the best point is $(x_1, x_2)=(2,6)$

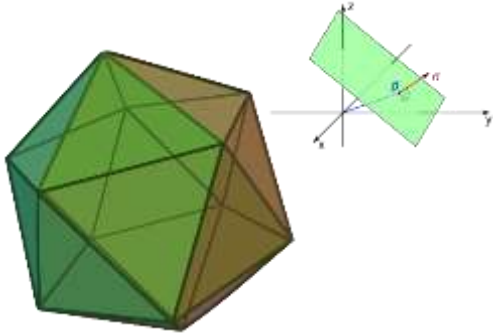
Then we just plug $(x_1, x_2)=(2,6)$ into $Z = 3x_1 + 5x_2$ to get $Z = 3 * 2 + 5 * 6 = 36$



# What should we recall from the previous lesson?




The beauty of **linear programming** is in that the solution is on the frontier of the feasible region, and more specifically it will touch one or more Corner Point Feasible (CPF)



This will not be the case in **integer programming** where the solution points will be 'hidden' inside the feasible region

# What should we recall from the previous lesson?

**Table 4.4** Two-Way Classification of Employee Responses with Totals

| | In Favor (A) | Against (B) | Total |
|---|---|---|---|
| Male (M) | 15 | 45 | 60 |
| Female (F) | 4 | 36 | 40 |
| Total | 19 | 81 | 100 |

In favour or against higher pay for CEO's?

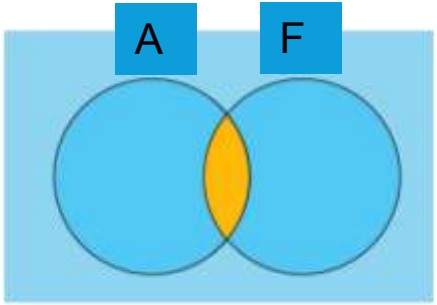
$P(A \cap F) = P(A \text{ and } F)$, Bayes' formula

$P(A \cap F) = P(A \mid F)P(F) = P(F \mid A)P(A)$

$$\frac{4}{100} = \frac{4}{40}\frac{40}{100} = \frac{4}{19}\frac{19}{100}$$

One step more: what about P(A∪F) = P(A **or** F)

Is it P(A) + P(F) =
=(40 + 19)/100 ?




No, it is:

P(A) + P(F) − P(A∩F)

$(40 + 19 - 4)/100 = 55/100$

Why?

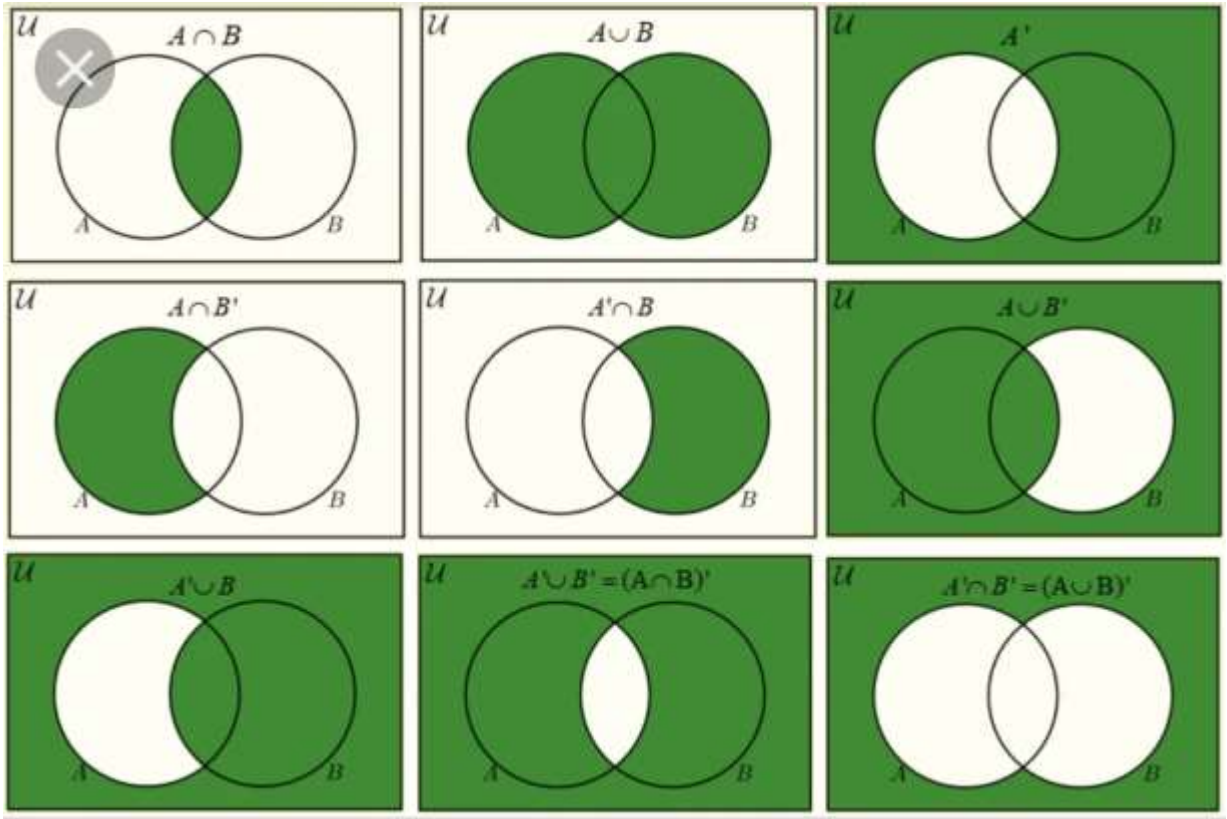
Table 4.4 **Two-Way Classification of Employee Responses with Totals**

| | In Favor (A) | Against (B) | Total |
|---|---|---|---|
| Male (M) | 15 | 45 | 60 |
| Female (F) | 4 | 36 | 40 |
| Total | 19 | 81 | 100 |

Otherwise this would have been counted twice

## A useful summary
## Note the notation $A'$ for the complementary set of $A$

(In our Mann stats textbook the complementary set of $A$ is indicated as $\bar{A}$)




Source: https://www.tuitionwithjason.sg/tag/venn-diagram/

# Using Excel Solver

## How to instal and open EXCEL SOLVER?

### In MAC

https://www.youtube.com/watch?v=ge4FMyZEUF0

### In Windows

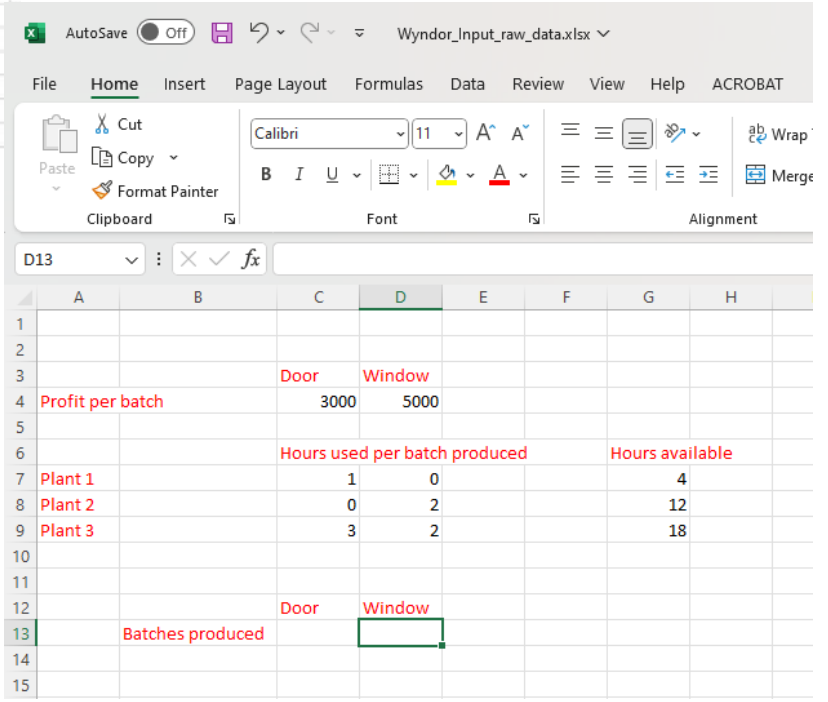
https://www.youtube.com/watch?v=W6tIS4JZ5J0

# 1) Open a white excel sheet; 2) Create a table as this:

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | | Wyndor Glass Co. Product-Mix Problem | | | | | |
| 2 | | | | | | | | |
| 3 | | | Doors | Windows | | | | |
| 4 | | Profit Per Batch | 3000 | 5000 | | | | |
| 5 | | | | | | | | |
| 6 | | | Hours Used Per Batch Produced | | Hours Used | Hint | Hours Available | |
| 7 | | Plant 1 | 1 | 0 | | <= | 4 | |
| 8 | | Plant 2 | 0 | 2 | | <= | 12 | |
| 9 | | Plant 3 | 3 | 2 | | <= | 18 | |
| 10 | | | | | | | | |
| 11 | | | Doors | Windows | | | Total Profit | |
| 12 | | Batches Produced | | | | | | |
| 13 | | | | | | | | |

A compiled version available at

http://www.andreasaltelli.eu/file/repository/Wyndor_Input_raw_data.xlsx




| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | | | | | | | | |
| 3 | | | Door | Window | | | | |
| 4 | Profit per batch | | 3000 | 5000 | | | | |
| 5 | | | | | | | | |
| 6 | | | Hours used per batch produced | | | | Hours available | |
| 7 | Plant 1 | | 1 | 0 | | | 4 | |
| 8 | Plant 2 | | 0 | 2 | | | 12 | |
| 9 | Plant 3 | | 3 | 2 | | | 18 | |
| 10 | | | | | | | | |
| 11 | | | | | | | | |
| 12 | | | Door | Window | | | | |
| 13 | | Batches produced | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |

Suggestion: use the same cell numbers throughout, e.g. Profit per batch in cells C4, D4

We shall first introduce data (now), then equations/relations

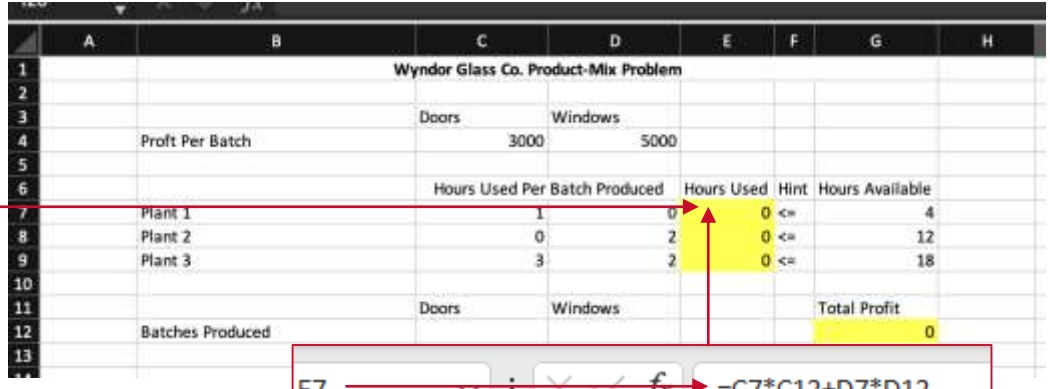
# 3) Insert the following Excel formulae

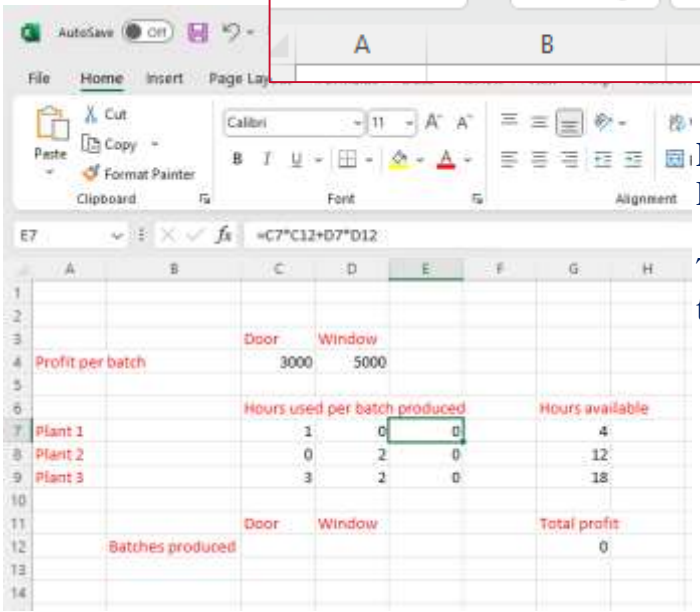
In the cell **E7** write: = C7*C12 + D7*D12

In the cell **E8** write: = C8*C12 + D8*D12

In the cell **E9** write: = C9*C12 + D9*D12

In the cell **G12** write: = C4*C12 + D4*D12




Note that C12 and D12 are still empty

They will be filled by the solver

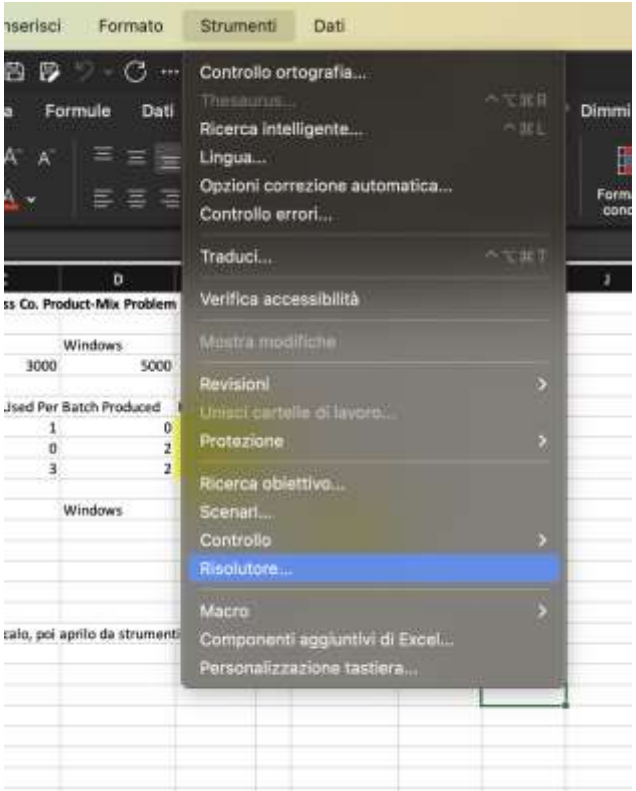
## Introducing relations

A compiled version available at

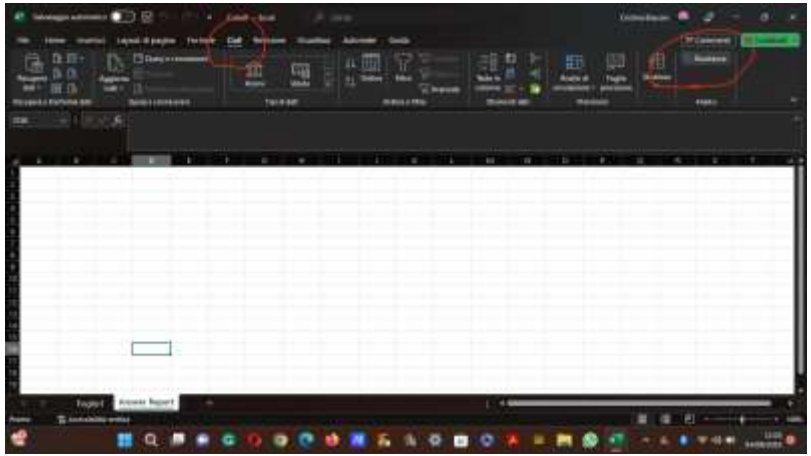
http://www.andreasaltelli.eu/file/repository/Wyndor_Input_relations.xlsx

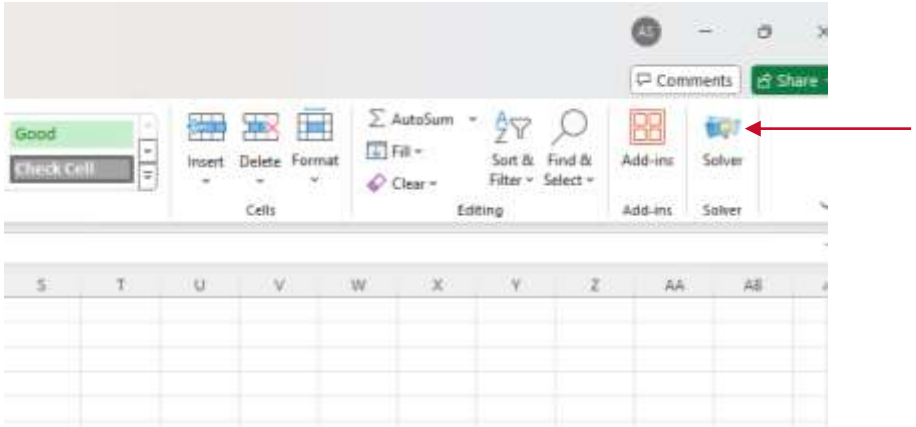
# 4) Open the solver

## In MAC

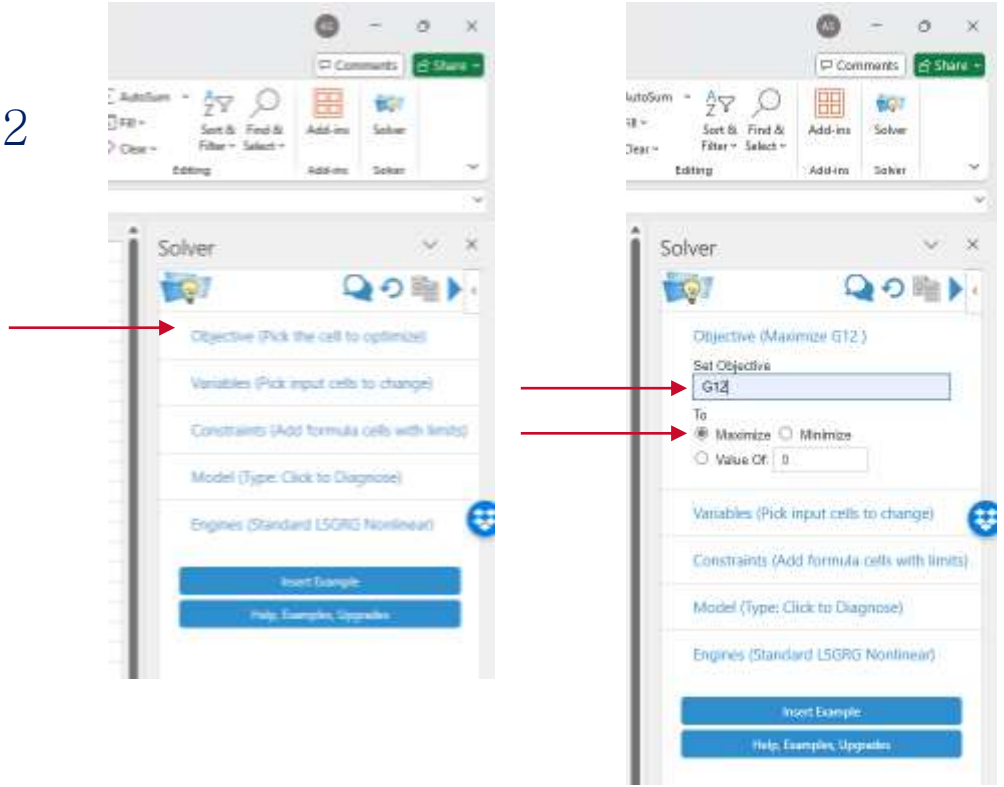



## In Windows (different versions)



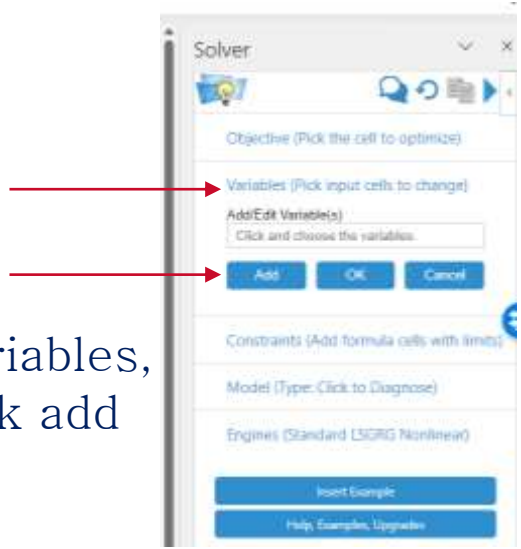


5) In Set objective insert the cell G12




Next, select the option Maximize

6) How? Changing Variable Cells
insert the cells C12:D12

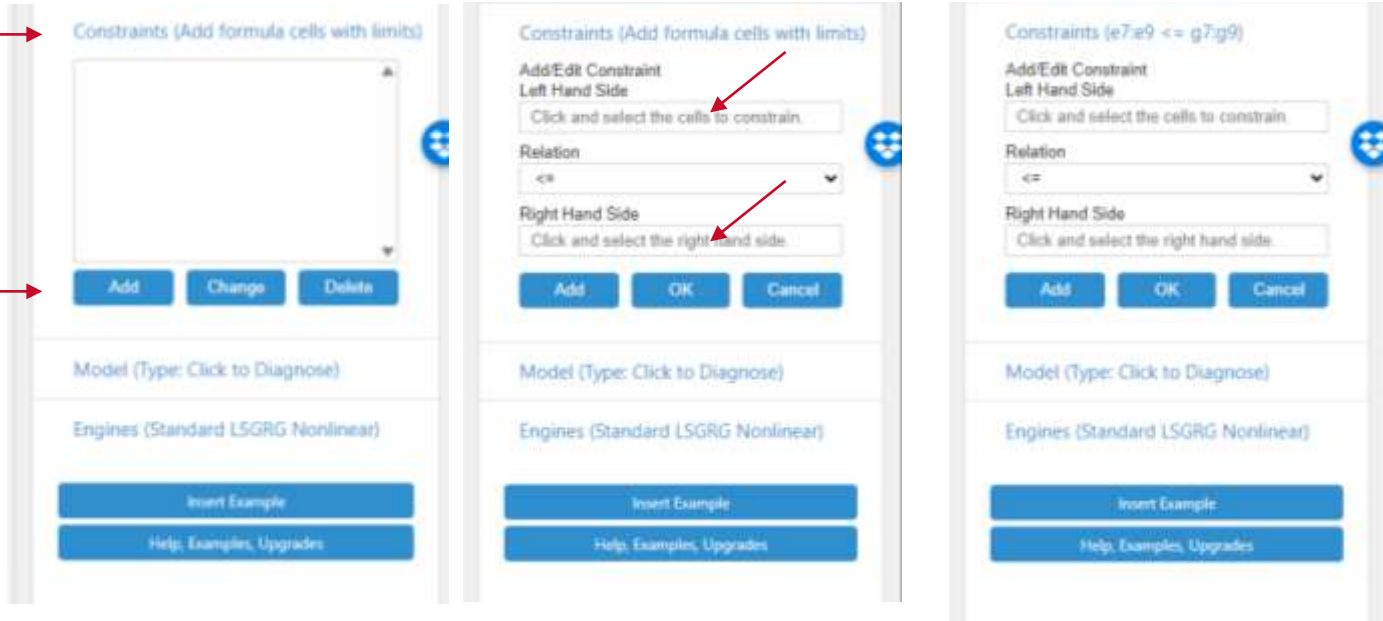



Click variables,
then click add


And this
appears

# 7) In **Constraints** click **Add** and **Insert** that cells **E7:E9 <= G7:G9**

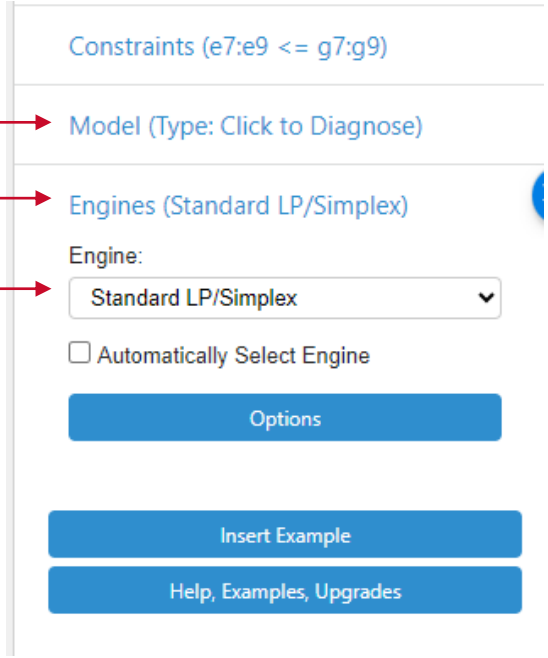
Click constraints,
then click add

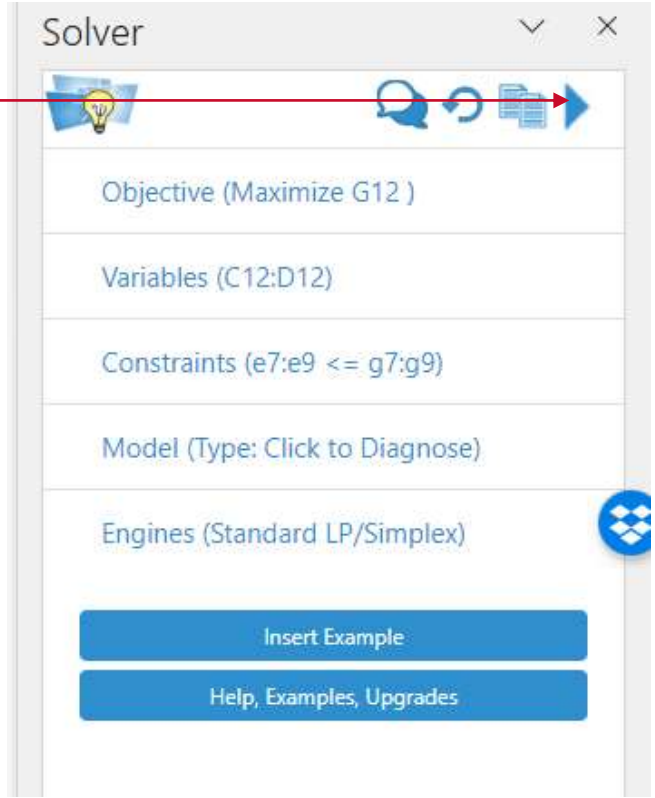


# 8) Select under Engines
## Simplex LP

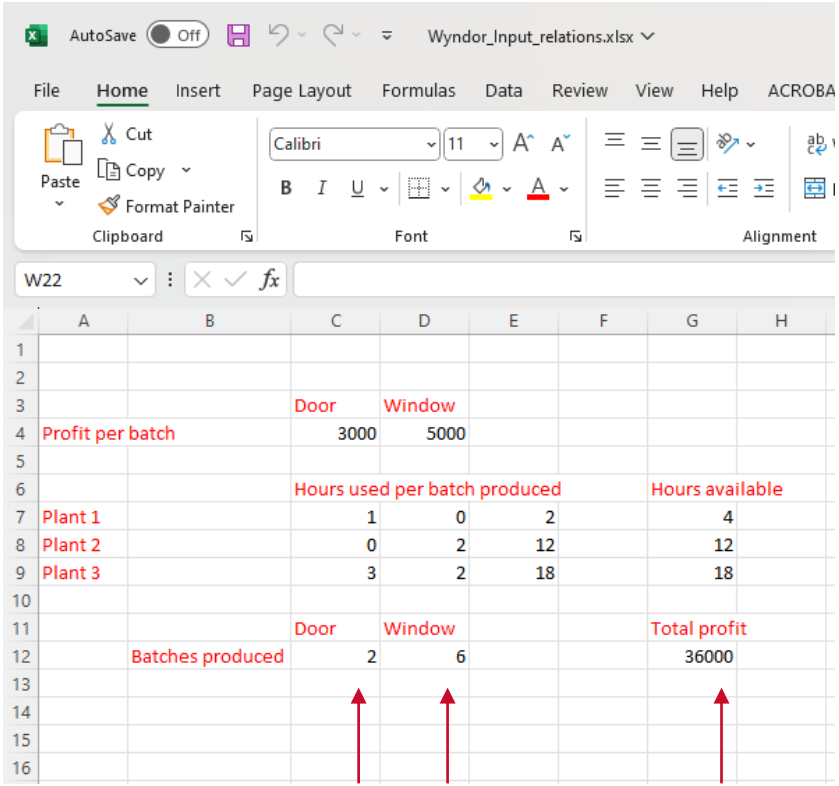
Then to solve
click the arrow

Skip



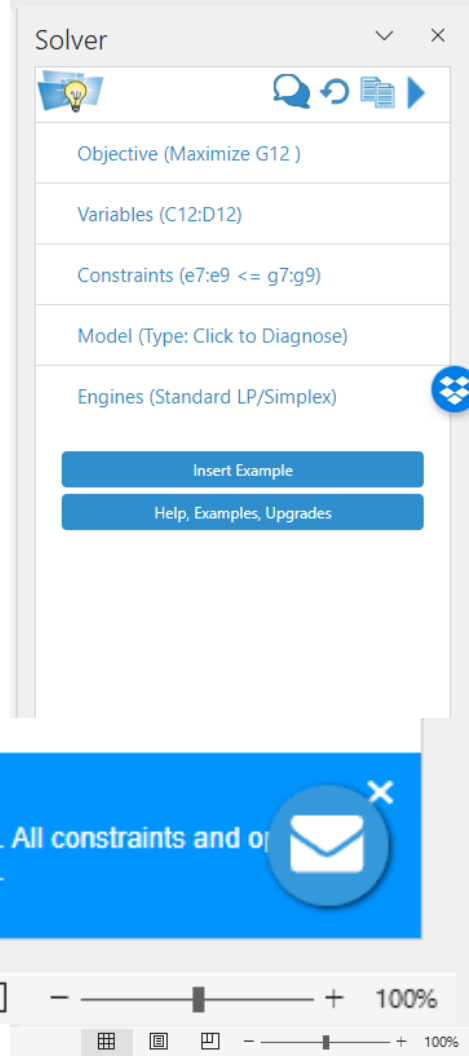



# ...and see the results in the cells C12:D12 and G12



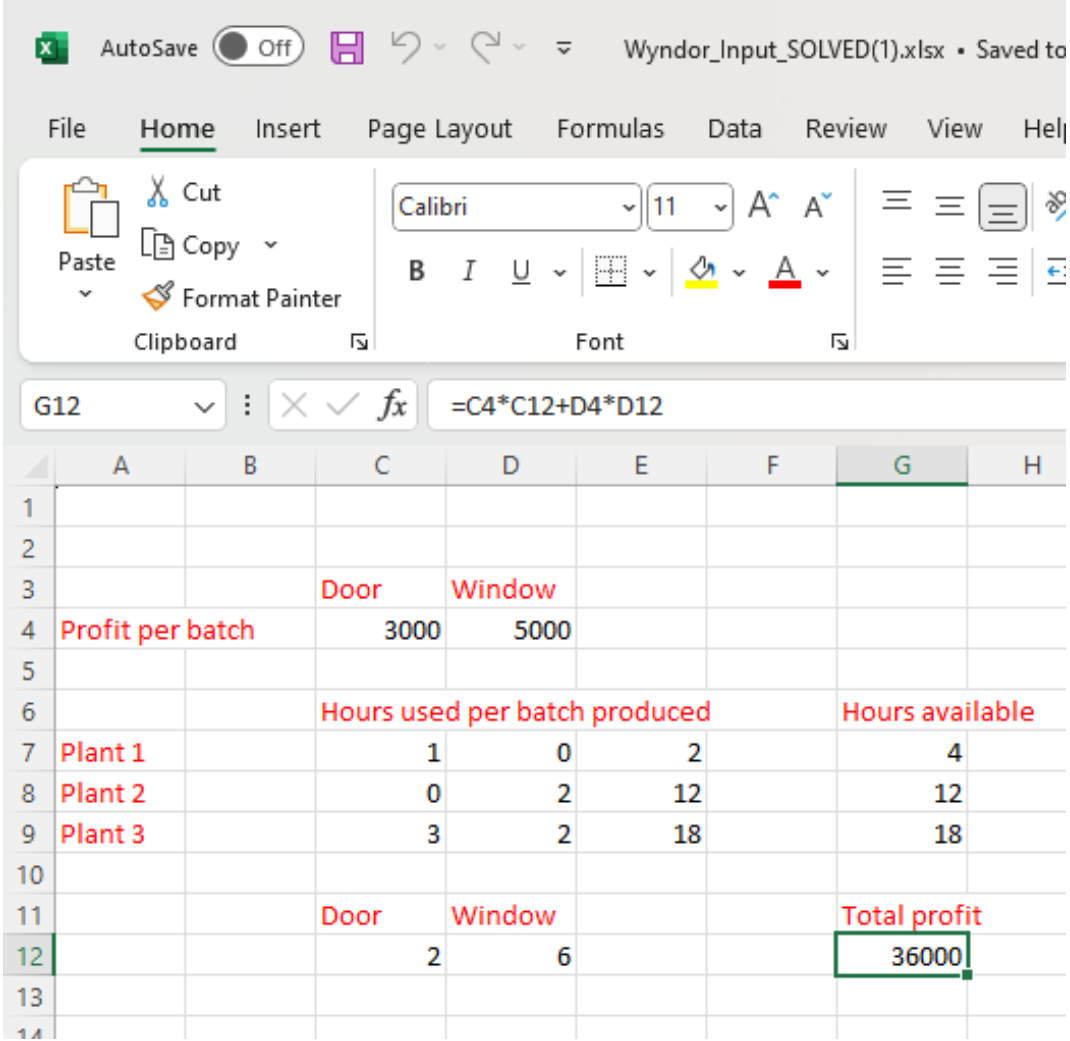

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | | | | | | | | |
| 3 | | | Door | Window | | | | |
| 4 | Profit per batch | | 3000 | 5000 | | | | |
| 5 | | | | | | | | |
| 6 | | | Hours used per batch produced | | | | Hours available | |
| 7 | Plant 1 | | 1 | 0 | 2 | | 4 | |
| 8 | Plant 2 | | 0 | 2 | 12 | | 12 | |
| 9 | Plant 3 | | 3 | 2 | 18 | | 18 | |
| 10 | | | | | | | | |
| 11 | | | Door | Window | | | Total profit | |
| 12 | | Batches produced | 2 | 6 | | | 36000 | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |
| 16 | | | | | | | | |

The vector X=(x1, x2) that maximize the profit is (2 doors, 6 windows) and the total profit is 36000



A compiled version available at

http://www.andreasaltelli.eu/file/repository/Wyndor_Input_SOLVED.xlsx




| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | | | | | | | | |
| 3 | | | Door | Window | | | | |
| 4 | Profit per batch | | 3000 | 5000 | | | | |
| 5 | | | | | | | | |
| 6 | | | Hours used per batch produced | | | | Hours available | |
| 7 | Plant 1 | | 1 | 0 | 2 | | 4 | |
| 8 | Plant 2 | | 0 | 2 | 12 | | 12 | |
| 9 | Plant 3 | | 3 | 2 | 18 | | 18 | |
| 10 | | | | | | | | |
| 11 | | | Door | Window | | | Total profit | |
| 12 | | | 2 | 6 | | | 36000 | |
| 13 | | | | | | | | |

upf. BARCELONA SCHOOL OF MANAGEMENT

# In the next slides:

09 Introduction to uncertainty and sensitivity analysis

10 What is a model?

11 Methods for uncertainty and sensitivity analysis

# 9.

## Introduction to uncertainty and sensitivity analysis

Sensitivity analysis as performed in linear programming relies often on 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.

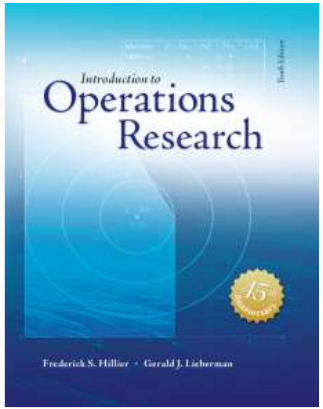
# Linear programming and sensitivity analysis

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"




p. 265

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

Let us try the standard one-factor-at-a time analysis

## 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 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 | 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 | |

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

For example we change this constraint

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

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

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

**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 | |

BARCELONA SCHOOL OF MANAGEMENT

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

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

New constraint line

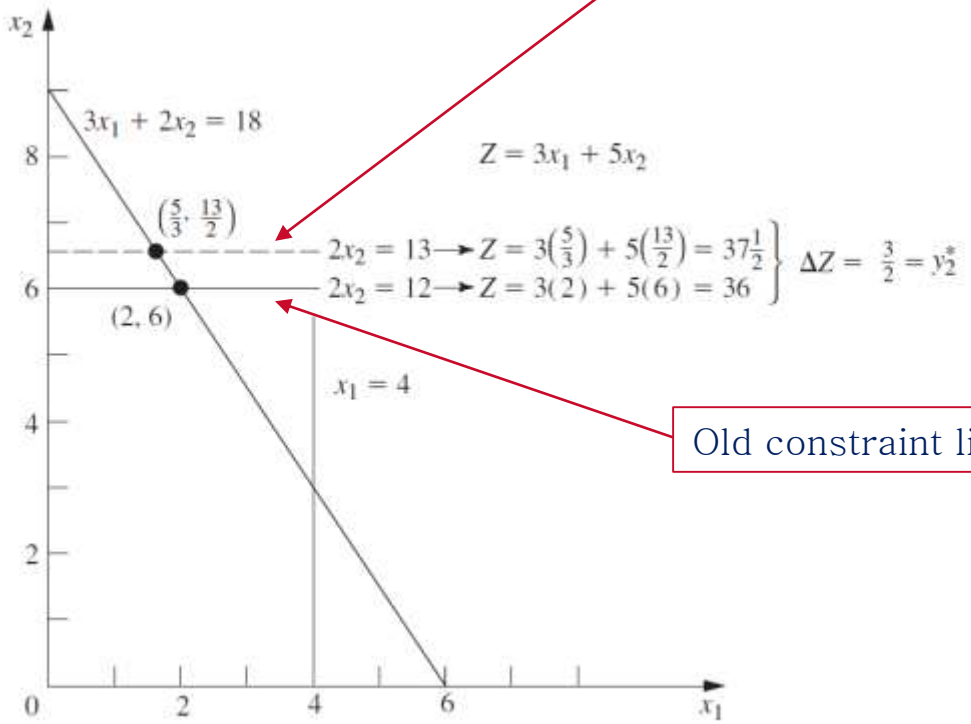
Find the new intercept between the old constraint line $3x_1+2x_2 = 18$ and the new constraint line $2x_2 = 13$

This is $\frac{5}{3}+\frac{13}{2}$

With the new constraint $Z$ becomes 37.5 instead of 36

So the shadow price is

$$y_2^* = 37.5 - 36 = 1.5$$




Old constraint line

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

**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 | |

Six $a_i$'s, three $b_i$'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

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


"Oh, Honey - our first bankruptcy!"

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

upf. BARCELONA SCHOOL OF MANAGEMENT

# McNamara's fallacy: 'if it cannot be measured, it is not important'

Source: www.history.com

"The first step is to measure whatever can easily be measured. This is OK as far as it goes. The second step is to disregard that which can't be easily measured or to give it an arbitrary quantitative value. This is artificial and misleading. The third step is to presume that what can't be measured easily really isn't important. This is blindness. The fourth step is to say that what can't be easily measured really doesn't exist. This is suicide." (after sociologist Daniel Yankelovich, see https://en.wikipedia.org/wiki/McNamara_fallacy)

# 10.

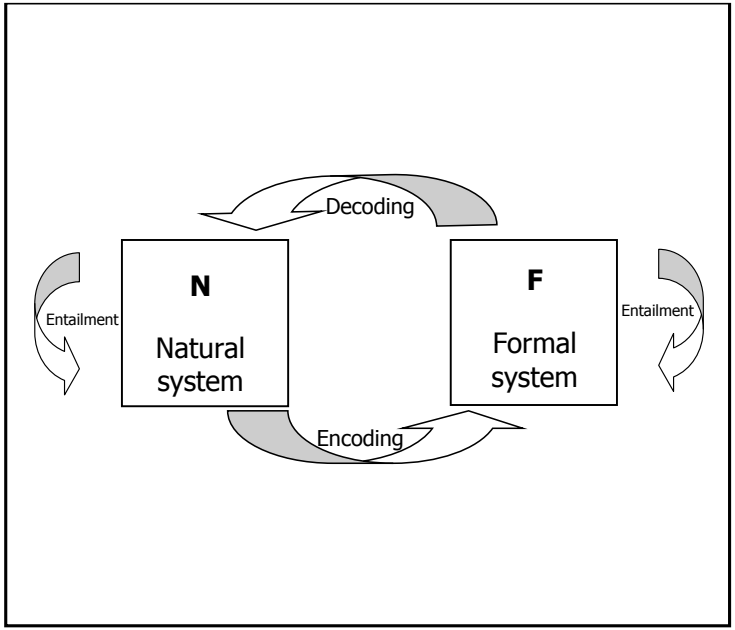
## Models

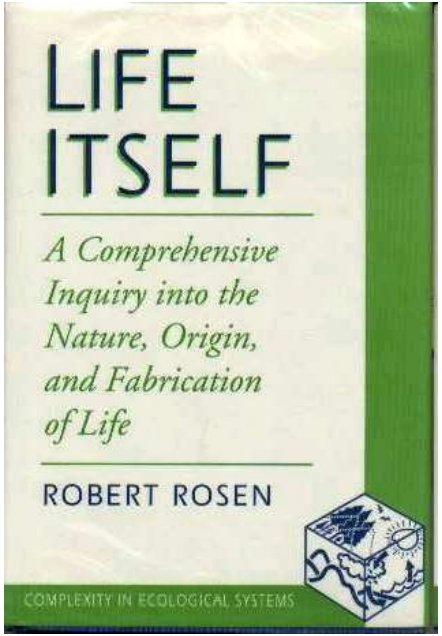
Why this section. Modelling as a craft or an art. Model versus straight physical laws. Models and their memory. Models in economics. Maps and the territory. Underfitting versus overfitting. Uncertainty versus ignorance.

upf. BARCELONA SCHOOL OF MANAGEMENT

# Modelling is a craft more than a science

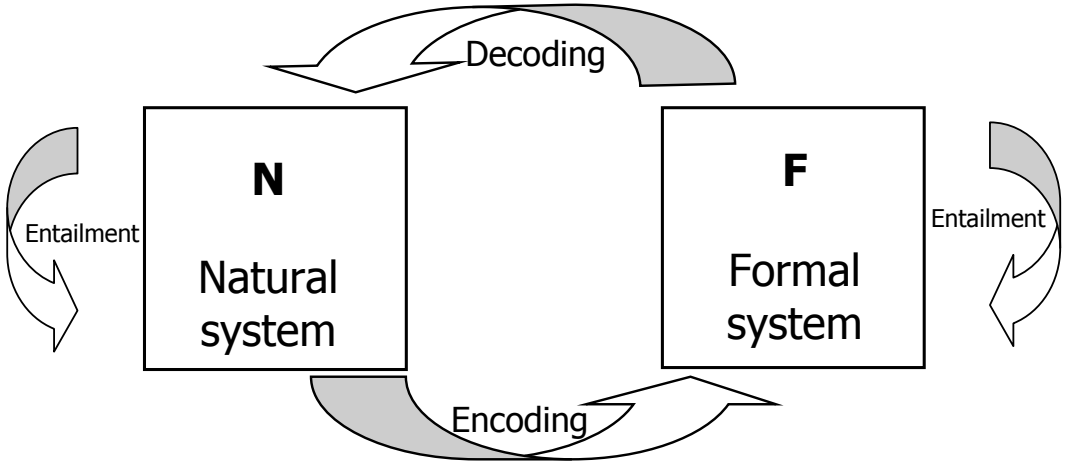
# Modelling as a craft rather than as a science for Robert Rosen




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 ?

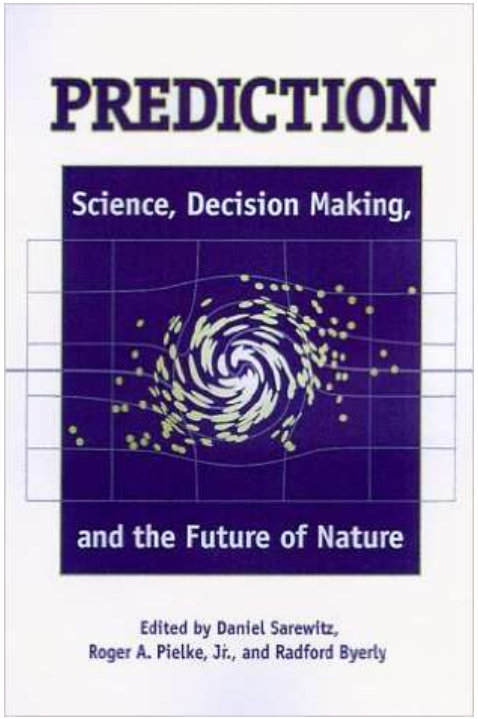
Robert Rosen
(1934-1998)

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

Naomi Oreskes

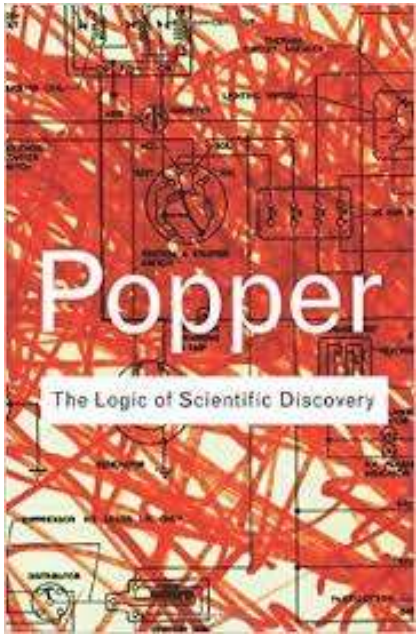
N. Oreskes, K. Shrader-Frechette, and K. Belitz, “Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences,” Science, 263, no. 5147, 1994.

# Models are not physical laws

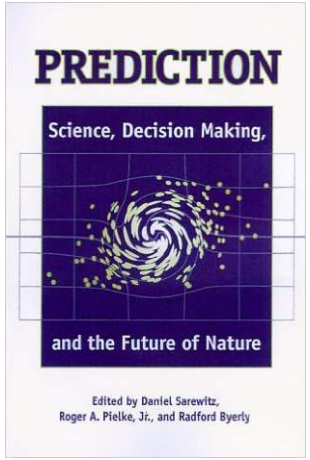



Oreskes, N., 2000, Why predict? Historical perspectives on prediction in Earth Science, in Prediction, Science, Decision Making and the future of Nature, Sarewitz et al., Eds., Island Press, Washington DC

“[…] 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?

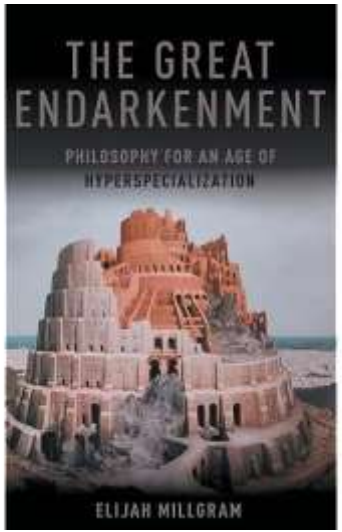



Any part (or several parts) of the model might be in error, and there is no simple way to determine which one it is"

# 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

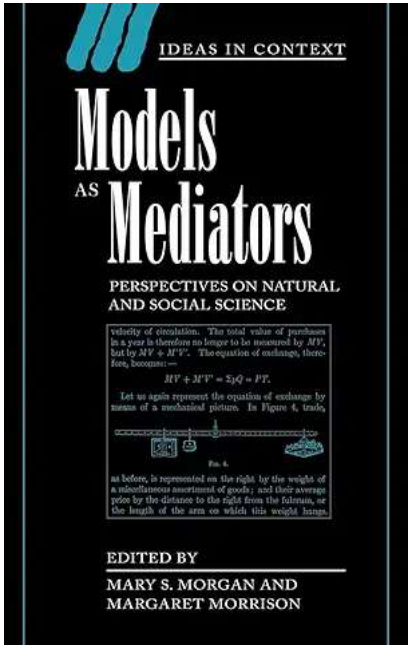

Philip Mirowski, 2013, Never let a serious crisis go wasted, Verso Books.

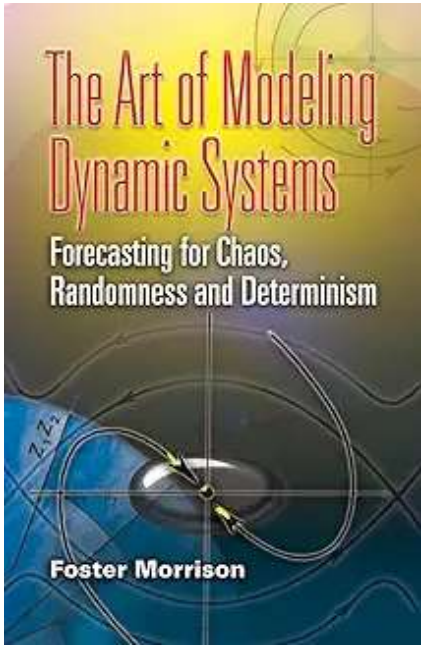
# The US senate and Queen Elisabeth perplexed…

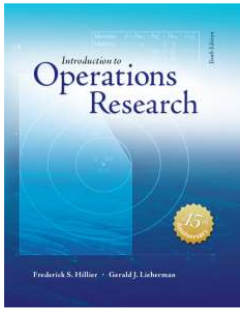


Philip Mirowski, 2013, Never let a serious crisis go wasted, Verso Books.

# What is a model? What is modelling?




"A mathematical model forms a bridge to the use of high-powered mathematical techniques and computers to analyze the problem...

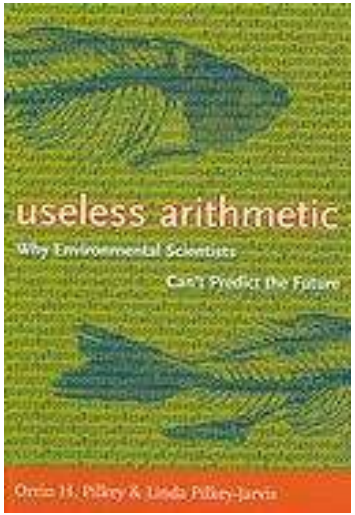
The proper criterion for judging the validity of a model is whether the model predicts the relative effects of the alternative courses of action with sufficient accuracy to permit a sound decision." p. 14

# Don’t confuse the map with the territory

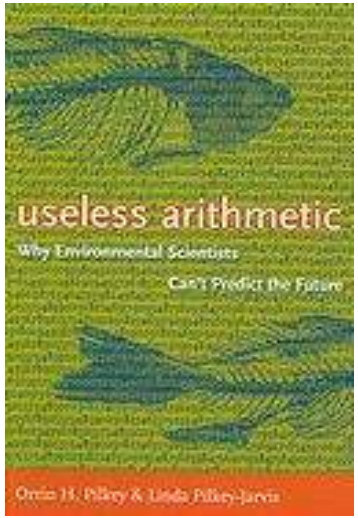
If you do, sensitivity analysis will not save you

<<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



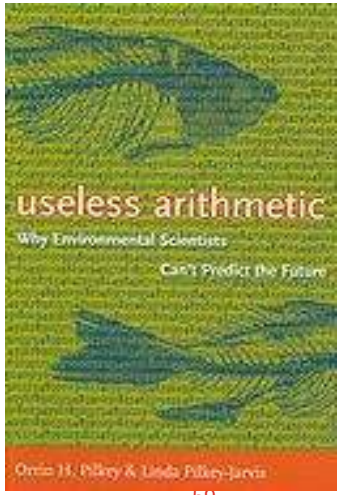

Useless Arithmetic: Why Environmental Scientists Can't Predict the Future
by Orrin H. Pilkey and Linda Pilkey-Jarvis, Columbia University Press, 2009.

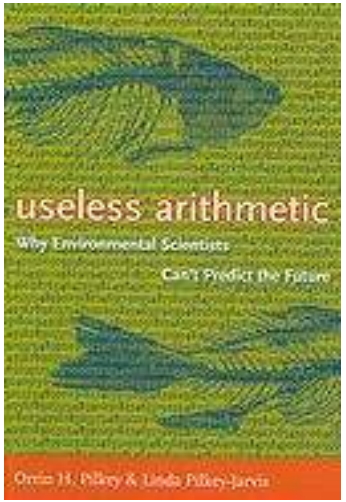



<<…If the model is wrong or if it is a poor representation of reality, determining the sensitivity of an individual parameter in the model is a meaningless pursuit>>

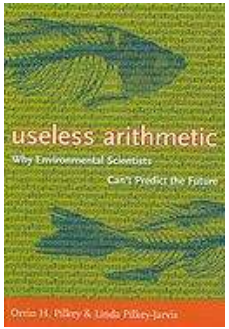
One of the examples discussed concerns the Yucca Mountain repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis.

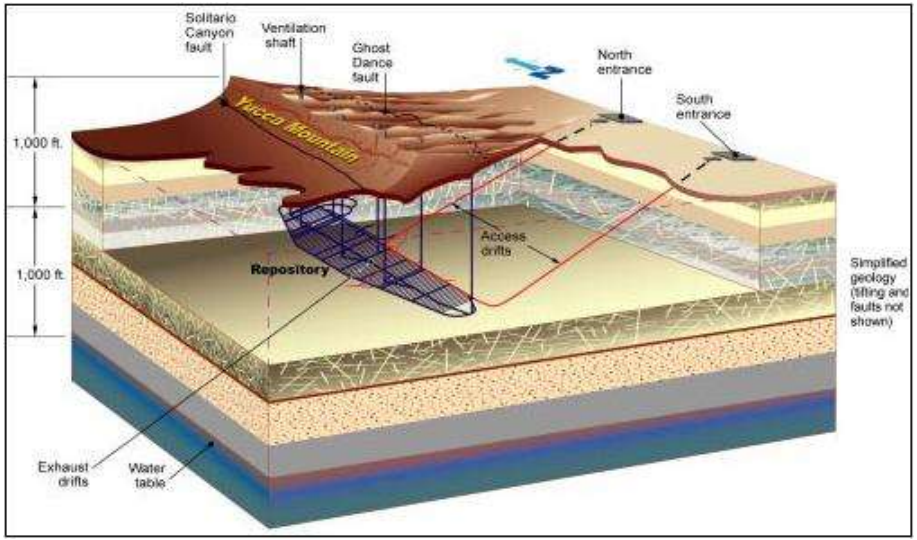
TSPA is Composed of 286 sub-models.






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

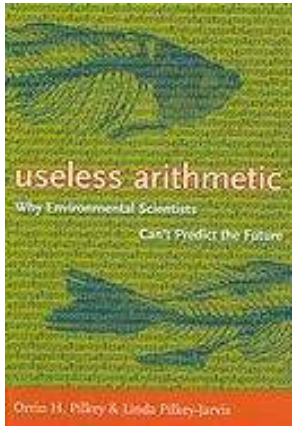



The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter
(the $^{36}Cl$ story)

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)

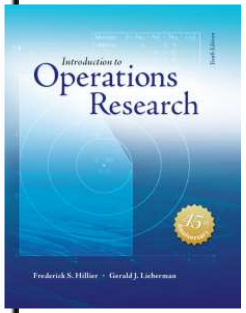
Communalism, Universalism, Disinterestedness, Organized Skepticism, from sociology of science, Robert K. Merton.

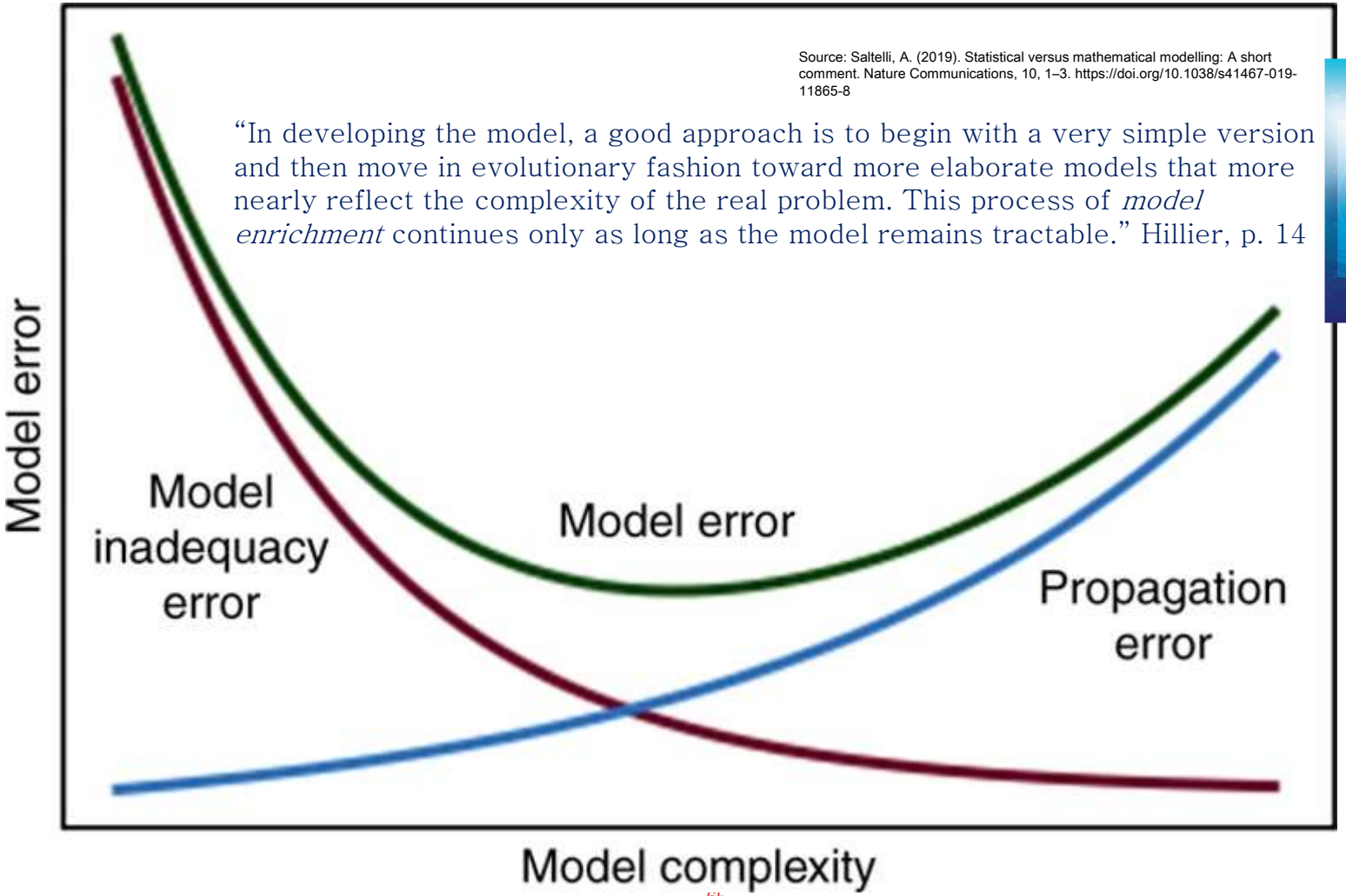
# Beware the size of your model

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

"In developing the model, a good approach is to begin with a very simple version and then move in evolutionary fashion toward more elaborate models that more nearly reflect the complexity of the real problem. This process of *model enrichment* continues only as long as the model remains tractable." Hillier, p. 14



# Simple principles of responsible modelling

## Mind the assumptions

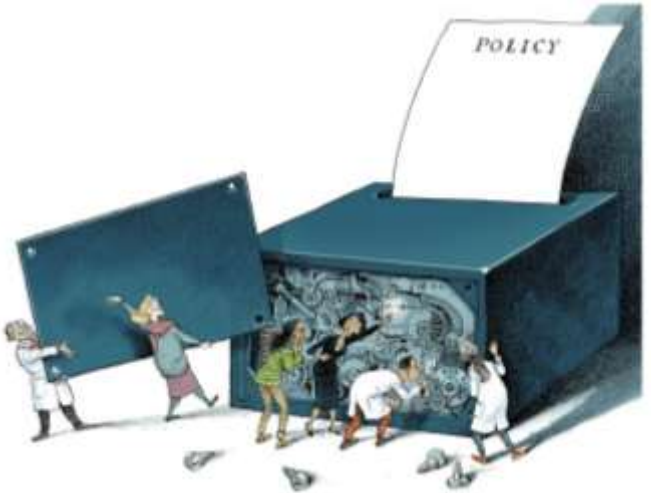
Assess uncertainty and sensitivity

## 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

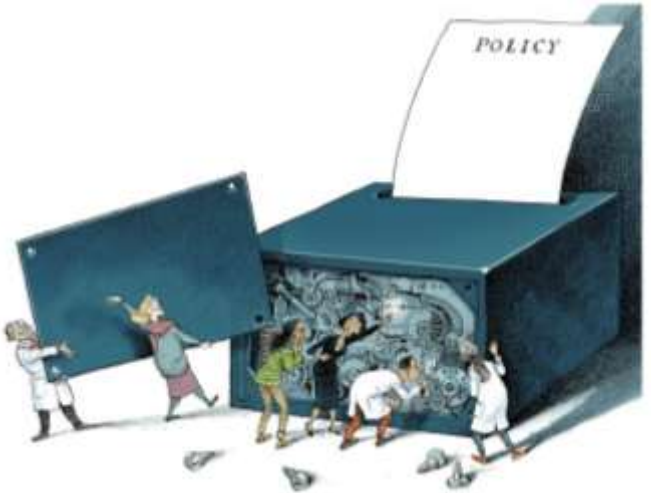
## Mind the unknowns

Acknowledge ignorance

# Mind the assumptions

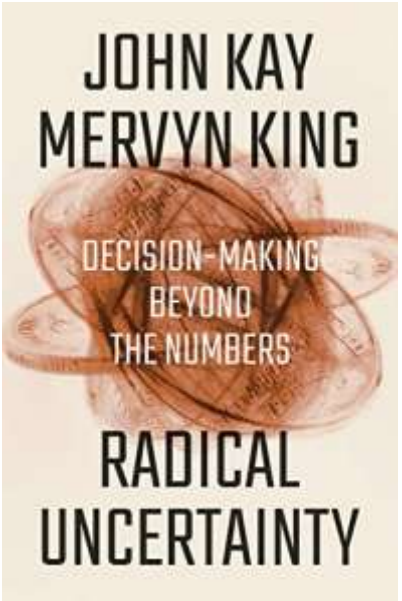
## Assess uncertainty and sensitivity

… 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

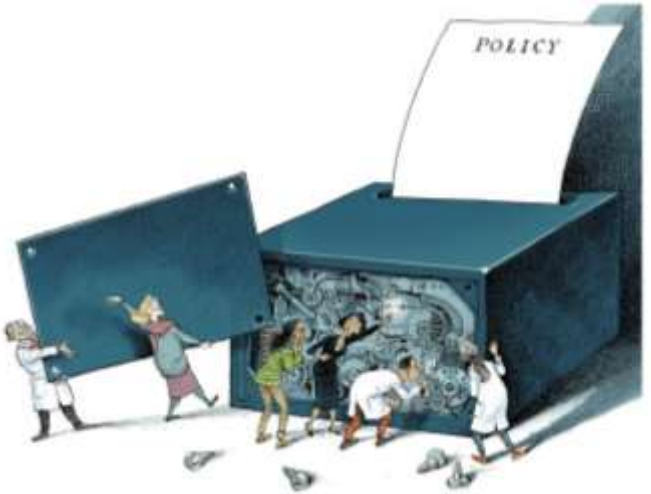
Assess uncertainty and sensitivity

## 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

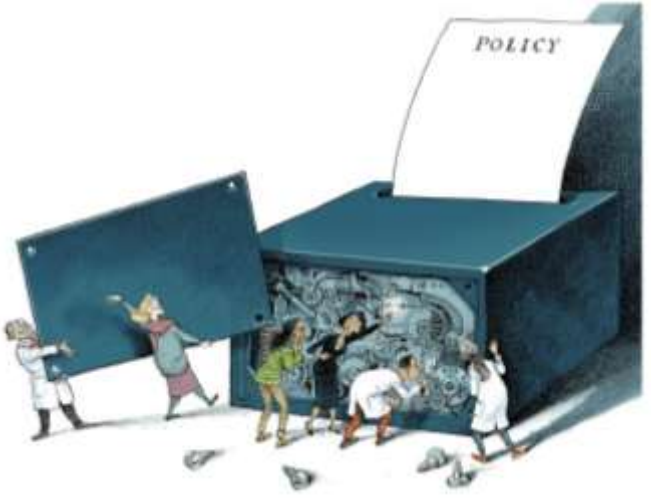
Quantification can backfire

## 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


**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

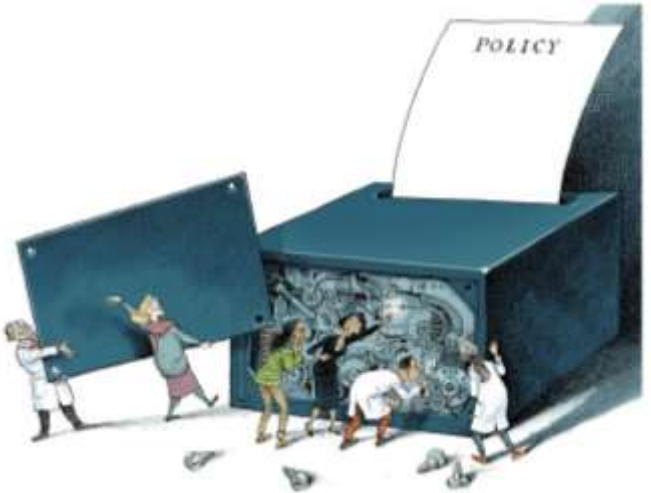
Assess uncertainty and sensitivity

## 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

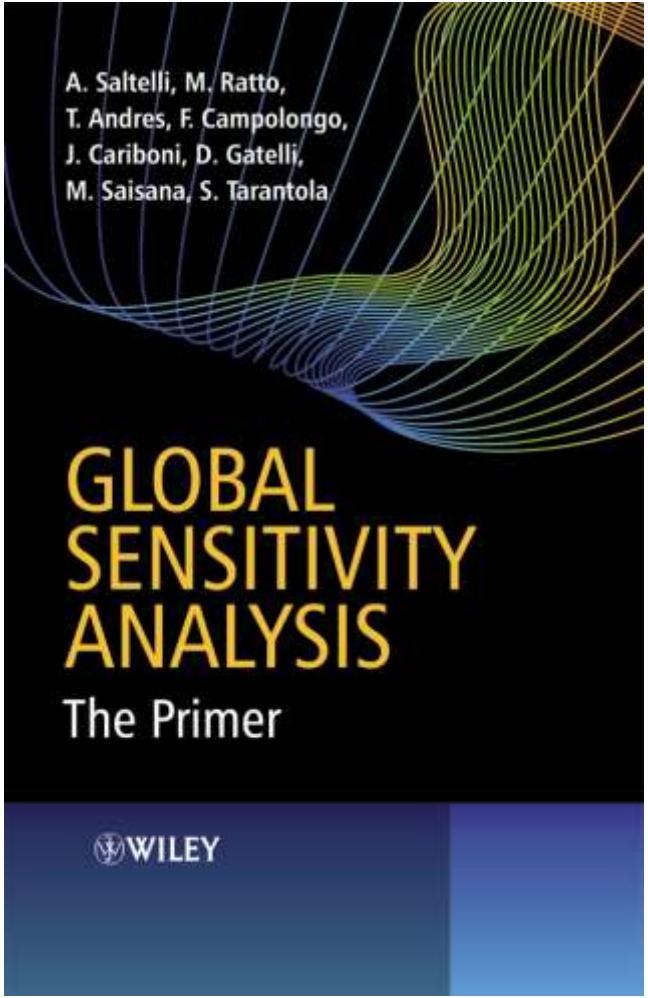

Source: A. Saltelli, G. Bammer, I. Bruno, et al., Five ways to ensure that models serve society: a manifesto, Nature 582 (2020) 482–484.

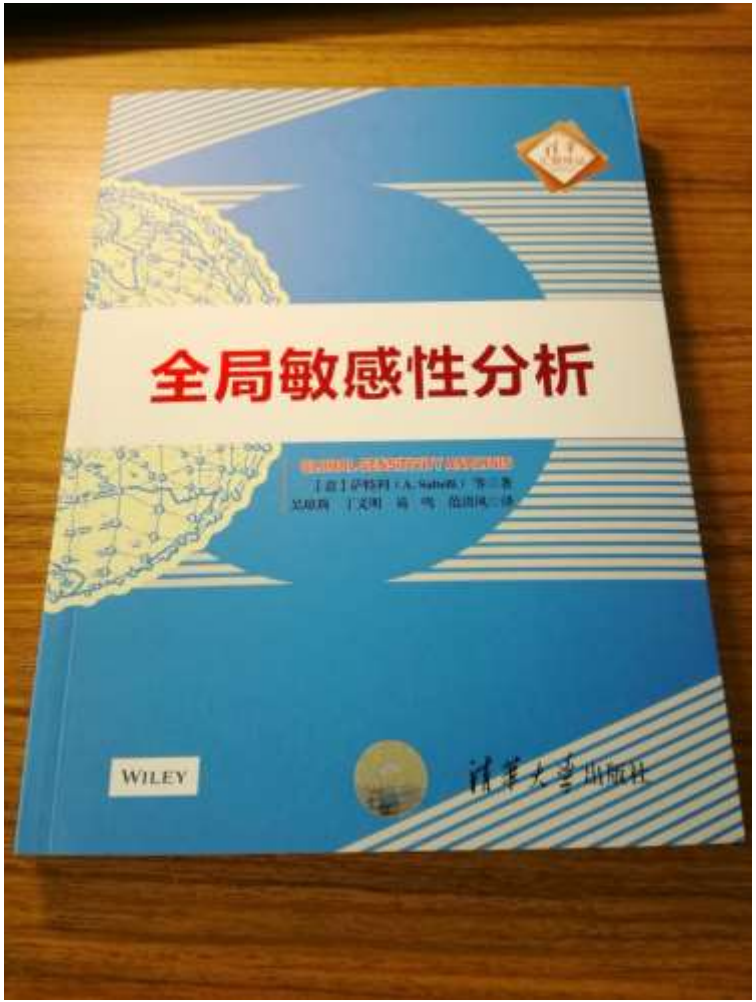
From Socrates's "knowing of not knowing" to Nicolaus Cusanus' Docta Ignorantia, ignorance was a virtue until Descartes

# 11.

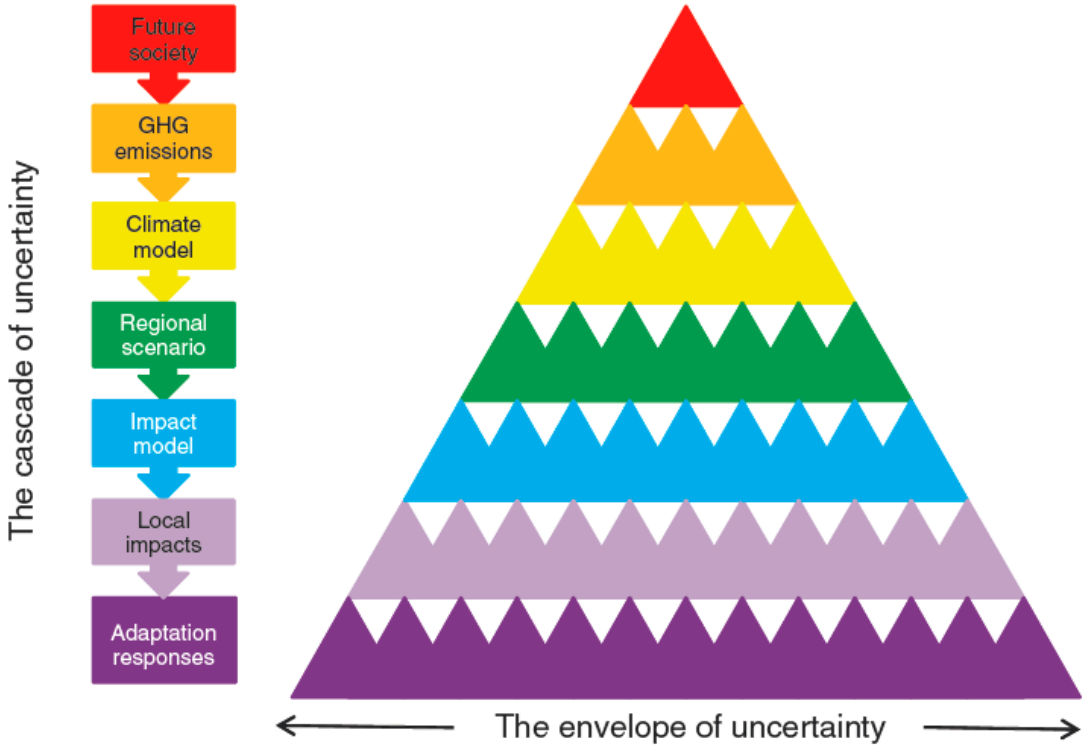
## Methods for Uncertainty and sensitivity analysis

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.


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)

# 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

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

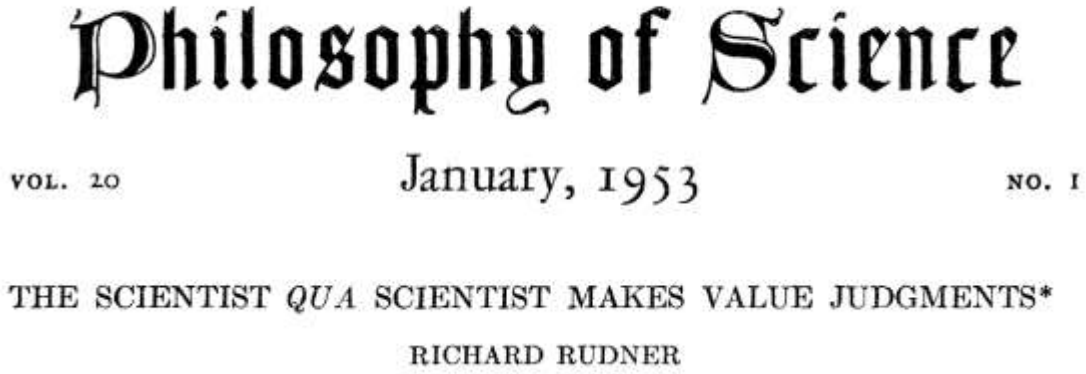
If X is not carcinogenic, but I reject the true null hypotheses (Type one error), this is bad for the firm producing the chemical and for innovation

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

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

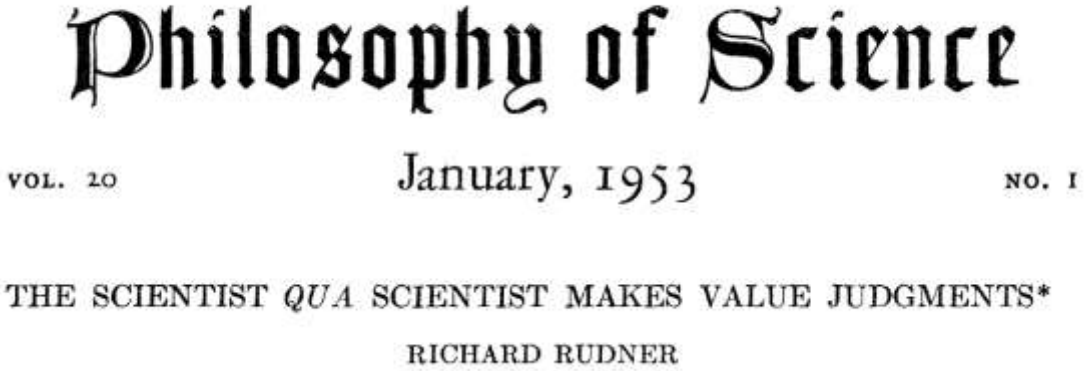
THE SCIENTIST *QUA* SCIENTIST MAKES VALUE JUDGMENTS*

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

upf. BARCELONA SCHOOL OF MANAGEMENT

"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

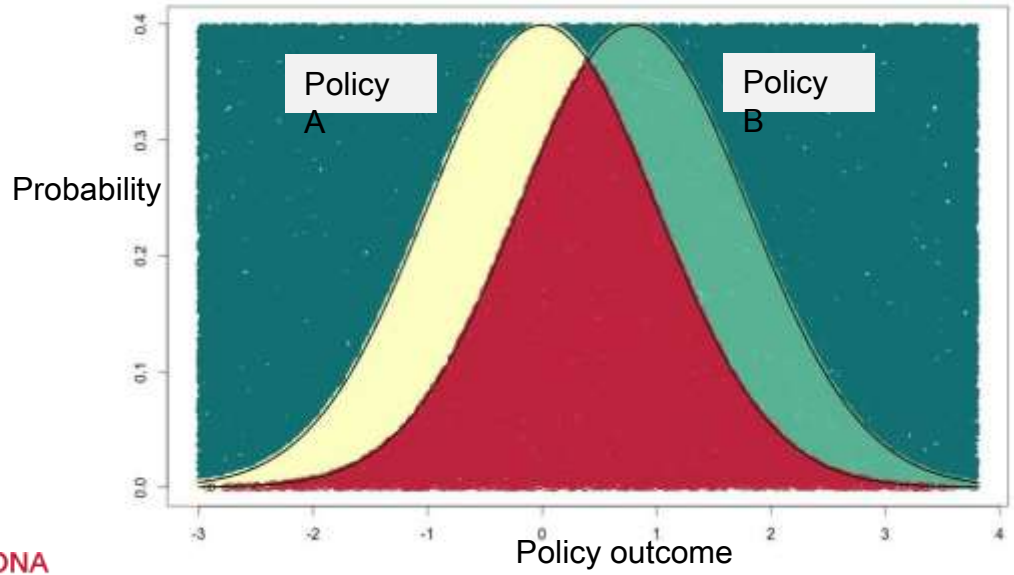
upf. BARCELONA SCHOOL OF MANAGEMENT

# 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

# Sensitivity analysis can:

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




BARCELONA SCHOOL OF MANAGEMENT

# 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)

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

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

upf. BARCELONA SCHOOL OF MANAGEMENT

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

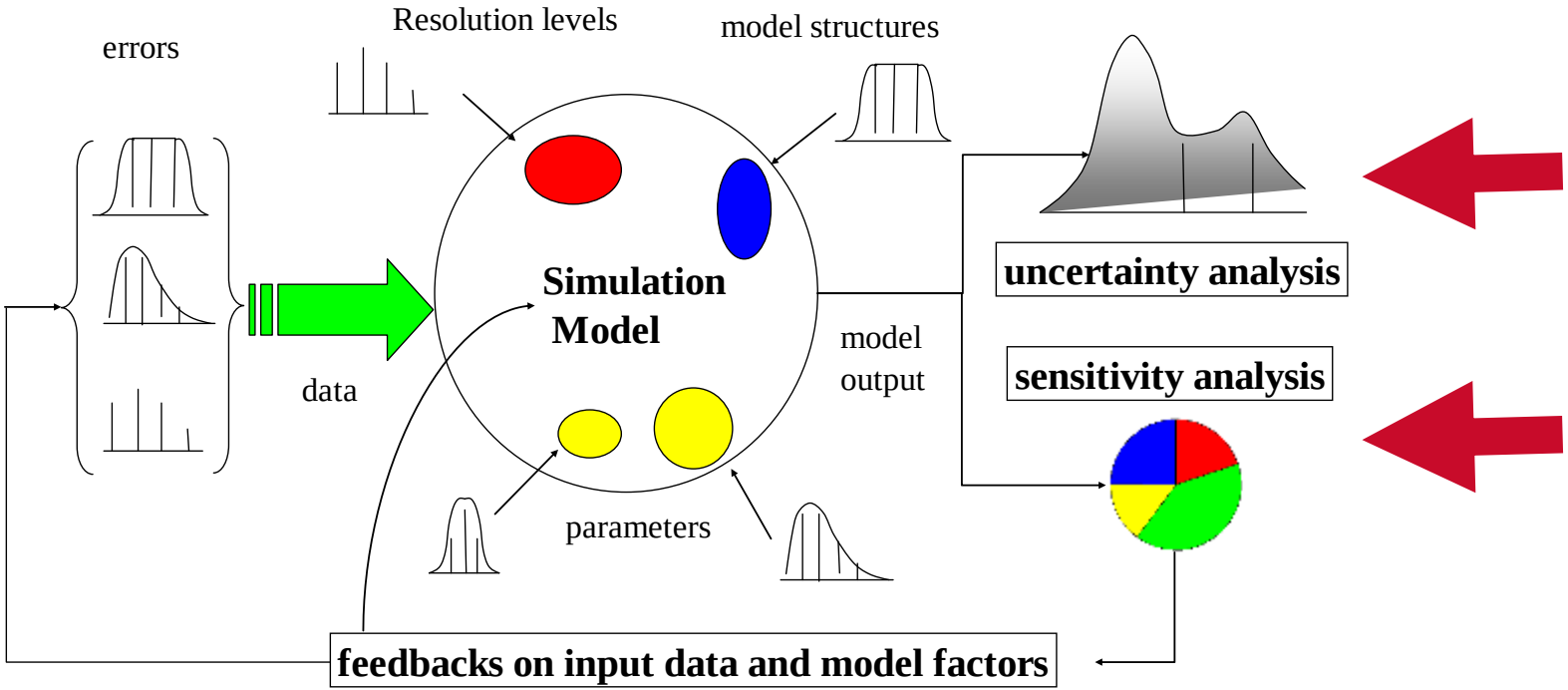
← Effect on $y$ of perturbing $x_i$ around its nominal value

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

← Effect on $y$ of perturbing $x_i$ both 'normalized' (divided by their nominal value)

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

← Effect on $y$ of perturbing $x_i$ both 'standardized' (divided by their standard deviation)

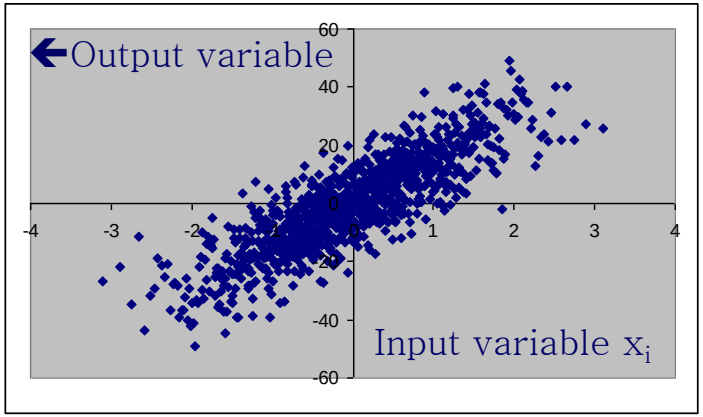

errors
Resolution levels
model structures
Simulation Model
data
model output
parameters
uncertainty analysis
sensitivity analysis
feedbacks on input data and model factors

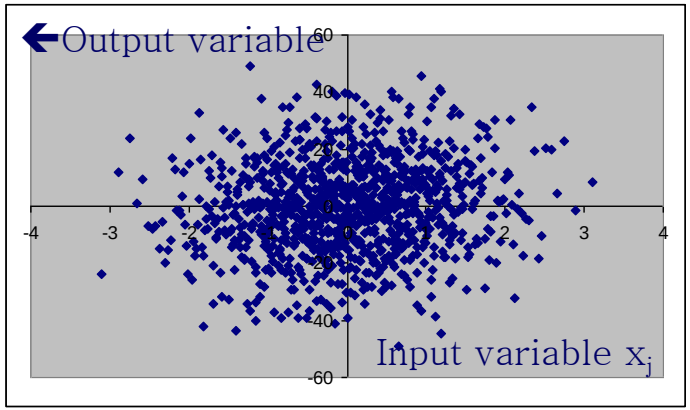


93

# Beyond ‘one factor at a time’

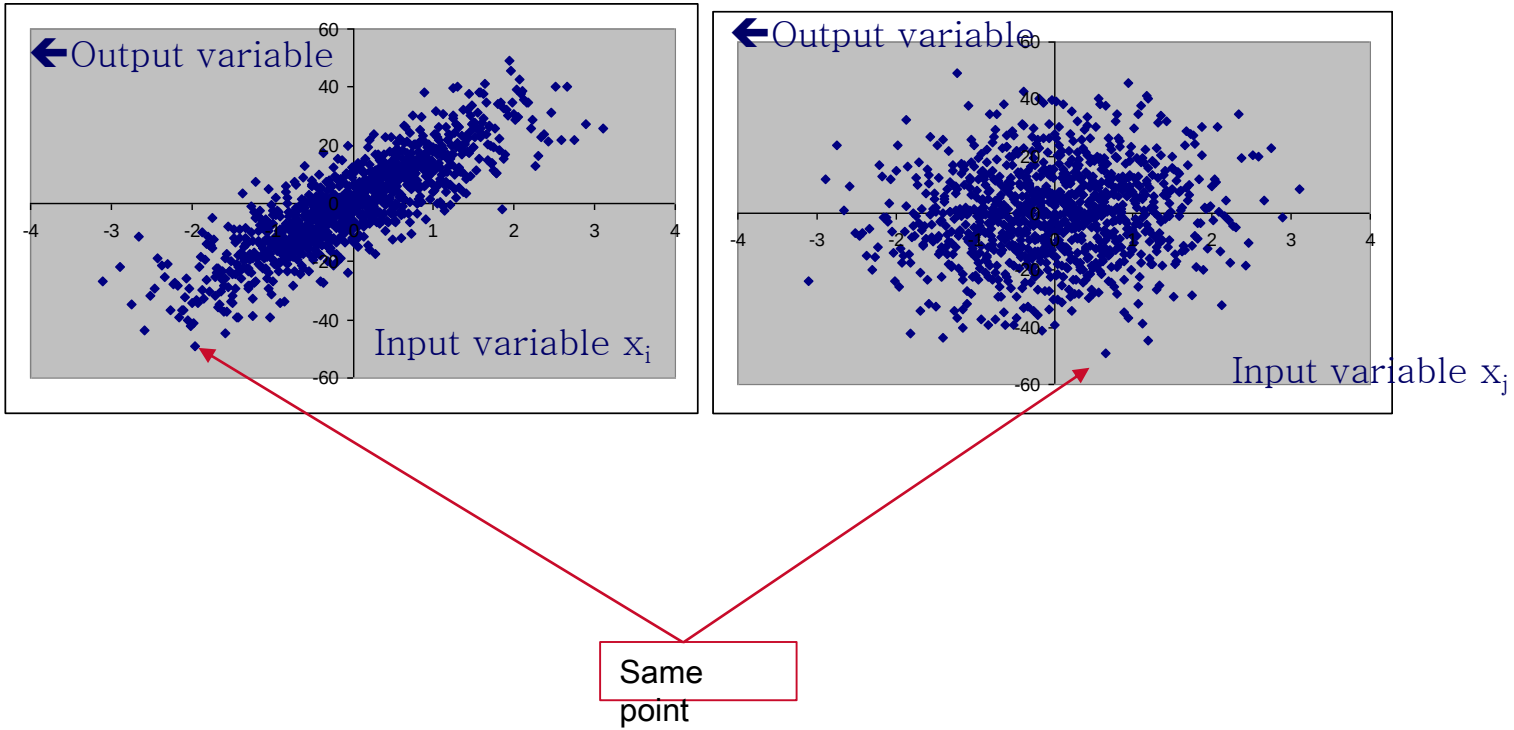
# An introduction to variance based methods



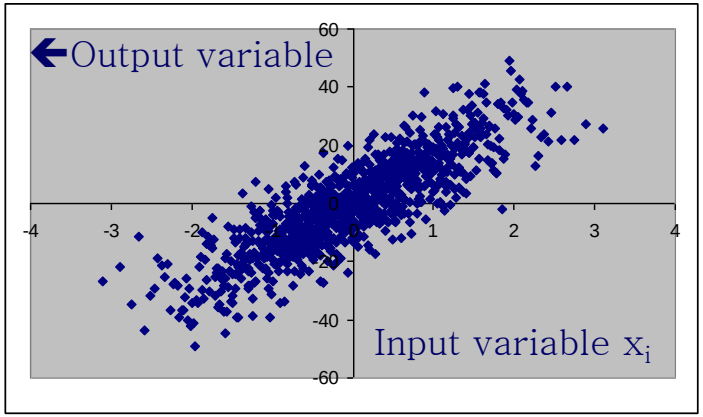



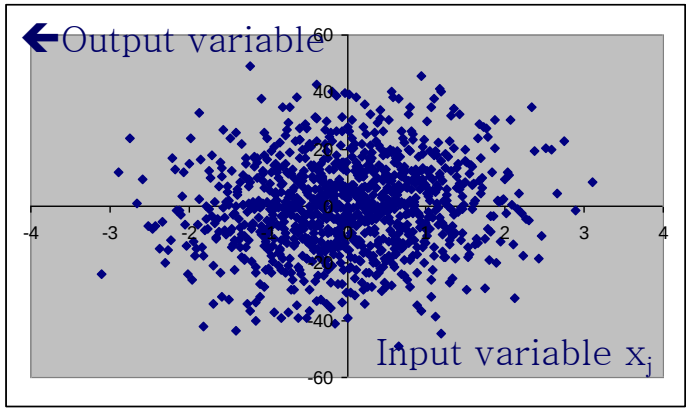

# Plotting the output as a function of two different input factors:

The model has been run 1,000 times sampling all uncertain factors; two of them are shown here, $x_i$ (left) and $x_j$ (right). The output variable has first been sorted by increasing $x_i$ then by increasing $x_j$.



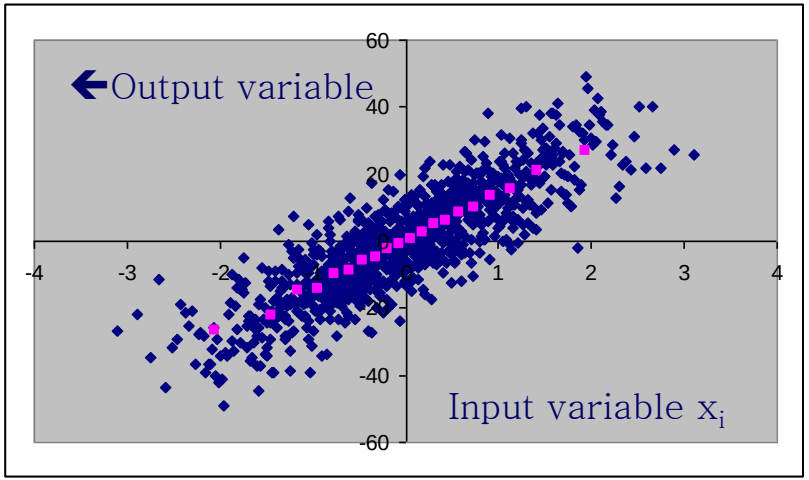

The points are always the same, only sorted differently

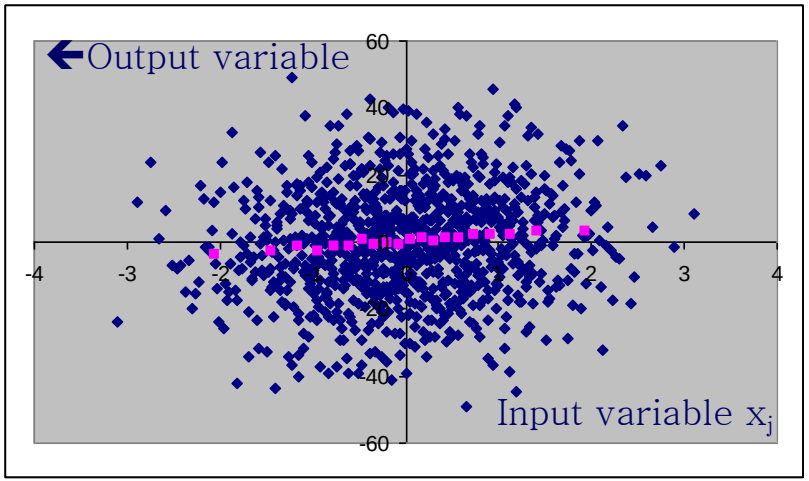



Plotting the output as a function of two different input factors

Which factor is more important?

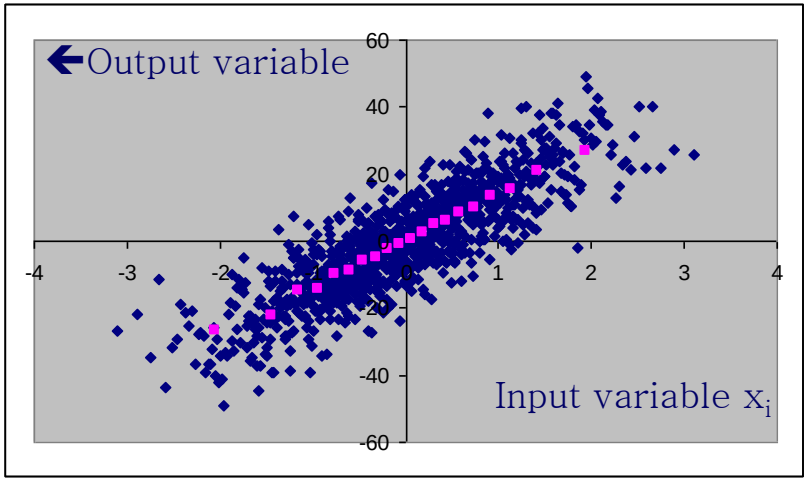


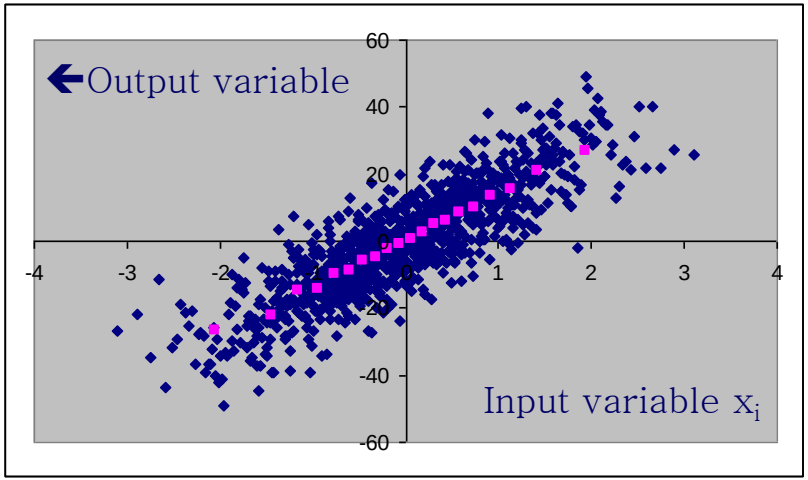

~1,000 blue points

Divide them in 20 bins of ~ 50 points

Compute the bin's average (pink dots)

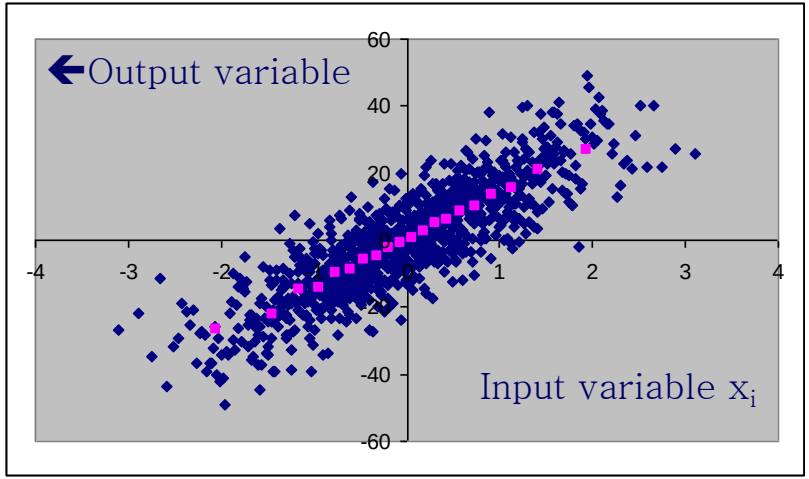


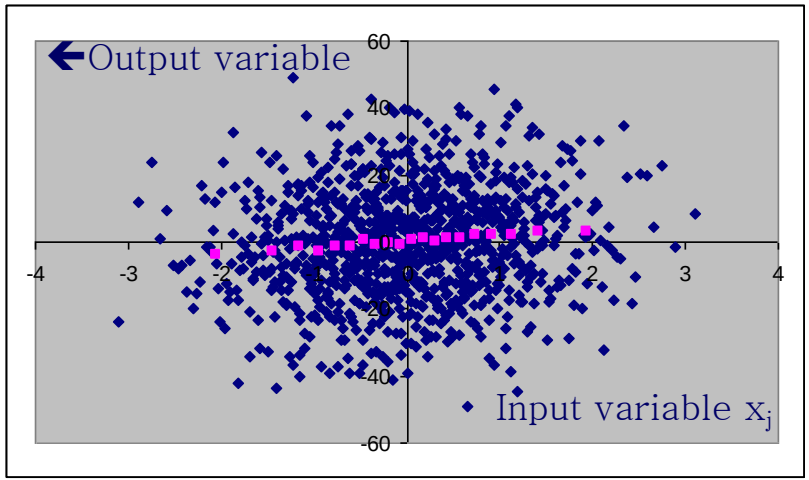

Each pink point is ~ $E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)$




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

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

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

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/

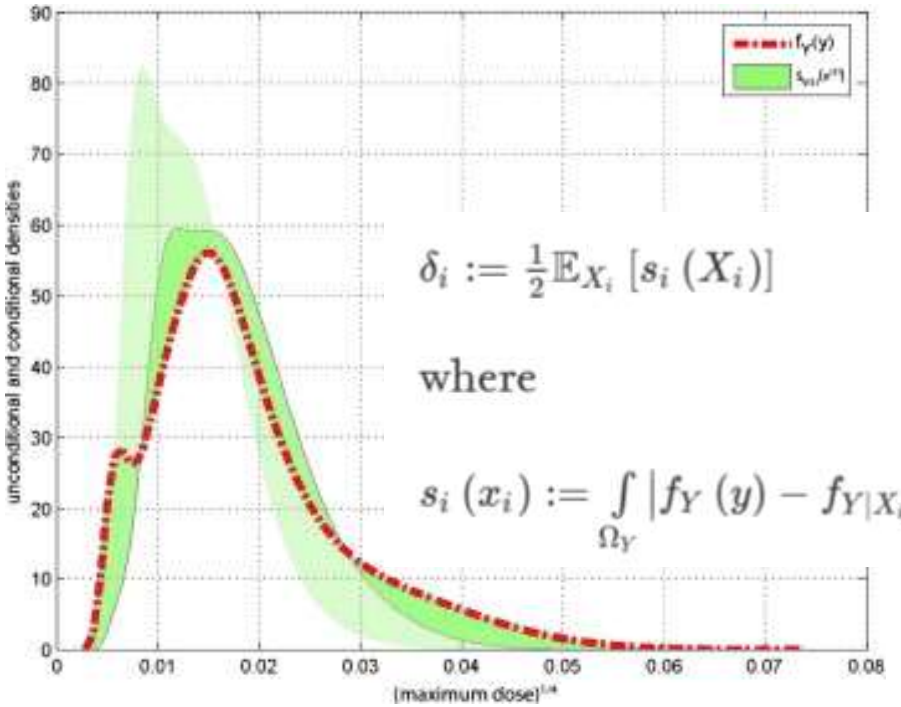
…but there is more, in R, Phython, Giulia …

# 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

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


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

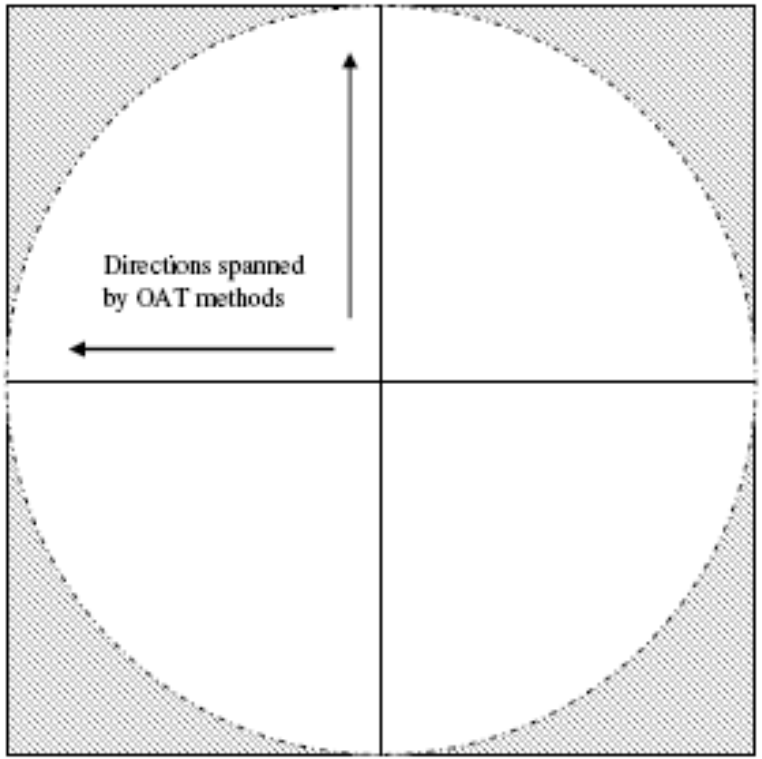
journal homepage: www.elsevier.com/locate/envsoft

# 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*

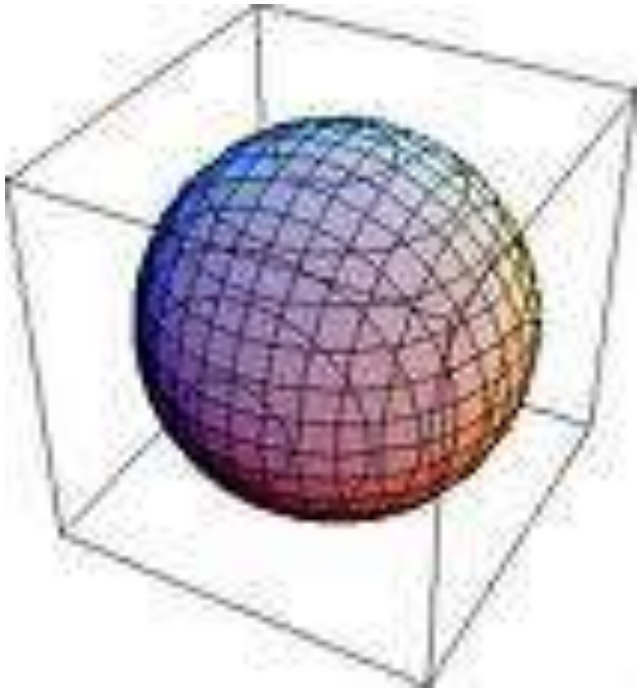
# OAT in 2 dimensions




Area circle / area square =?

~ 3/4

# OAT in 3 dimensions

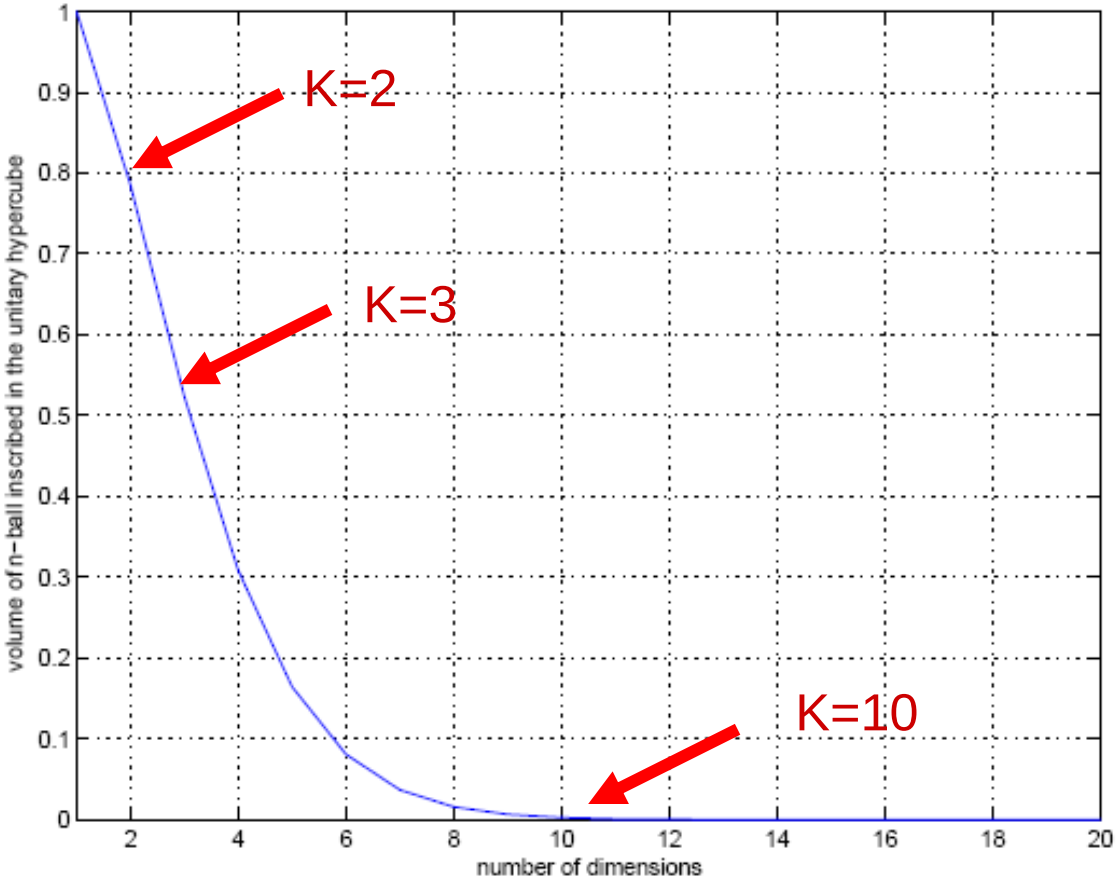


Volume sphere / volume cube =?

~ 1/2

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

# OAT in k dimensions



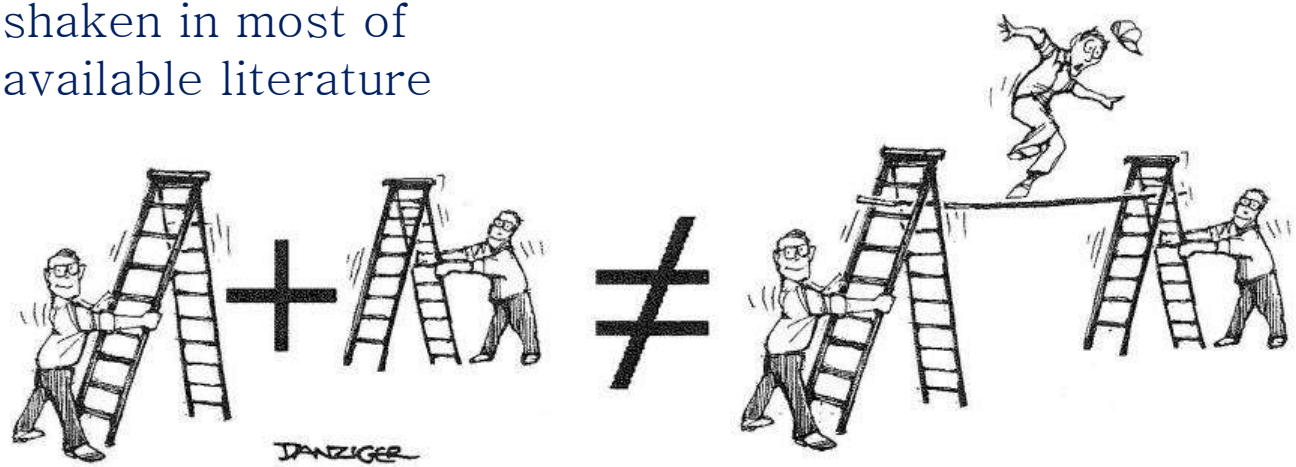
# OAT does not capture interactions

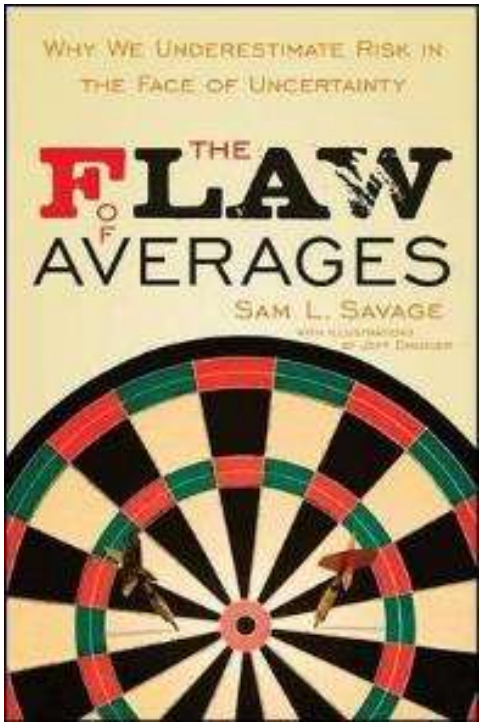
➔ The resulting analysis is non conservative

# How would you test the scaffolding?

How coupled ladders are shaken in most of available literature

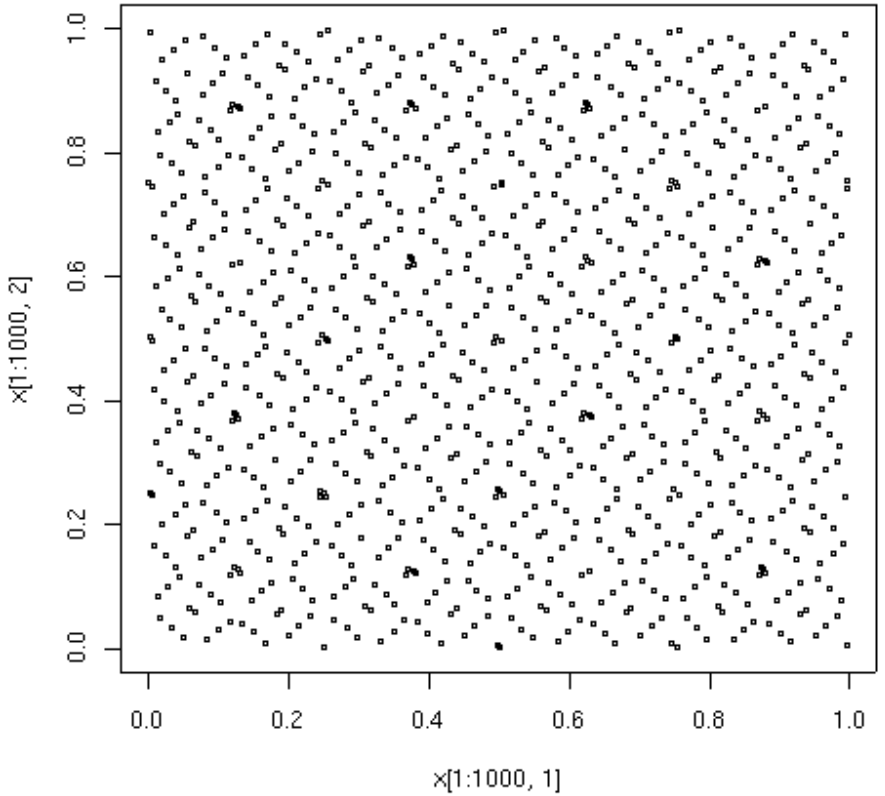
How to shake coupled ladders



# Quasi random sequences

Ilya M. Sobol'



Cornell University

arXiv.org > stat > arXiv:1505.02350

Search...

Help | Adva

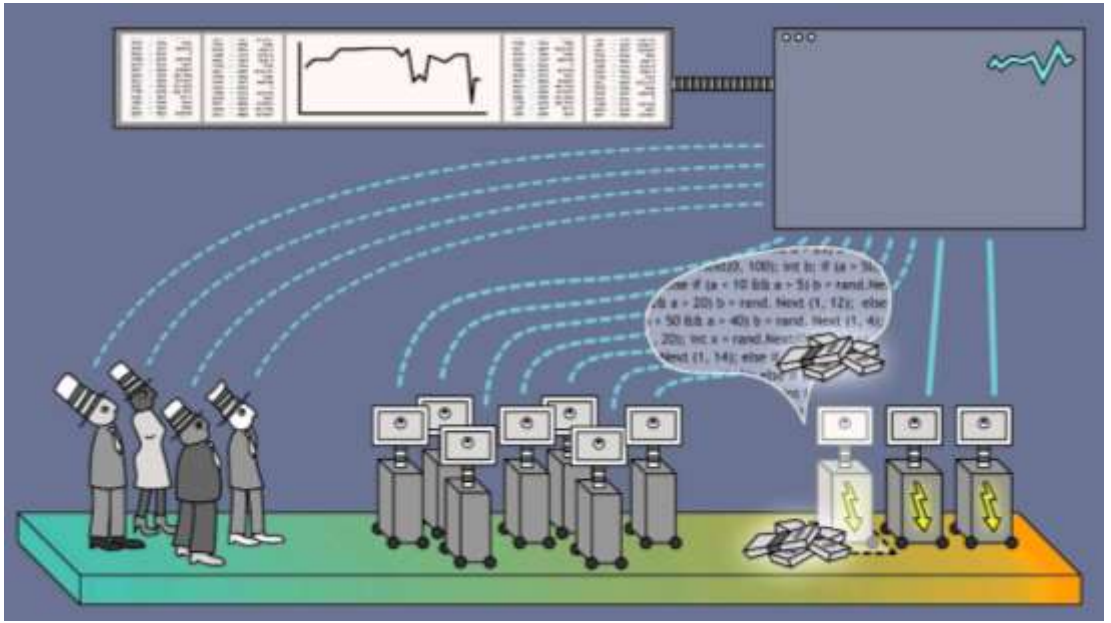
Statistics > Applications

*[Submitted on 10 May 2015]*

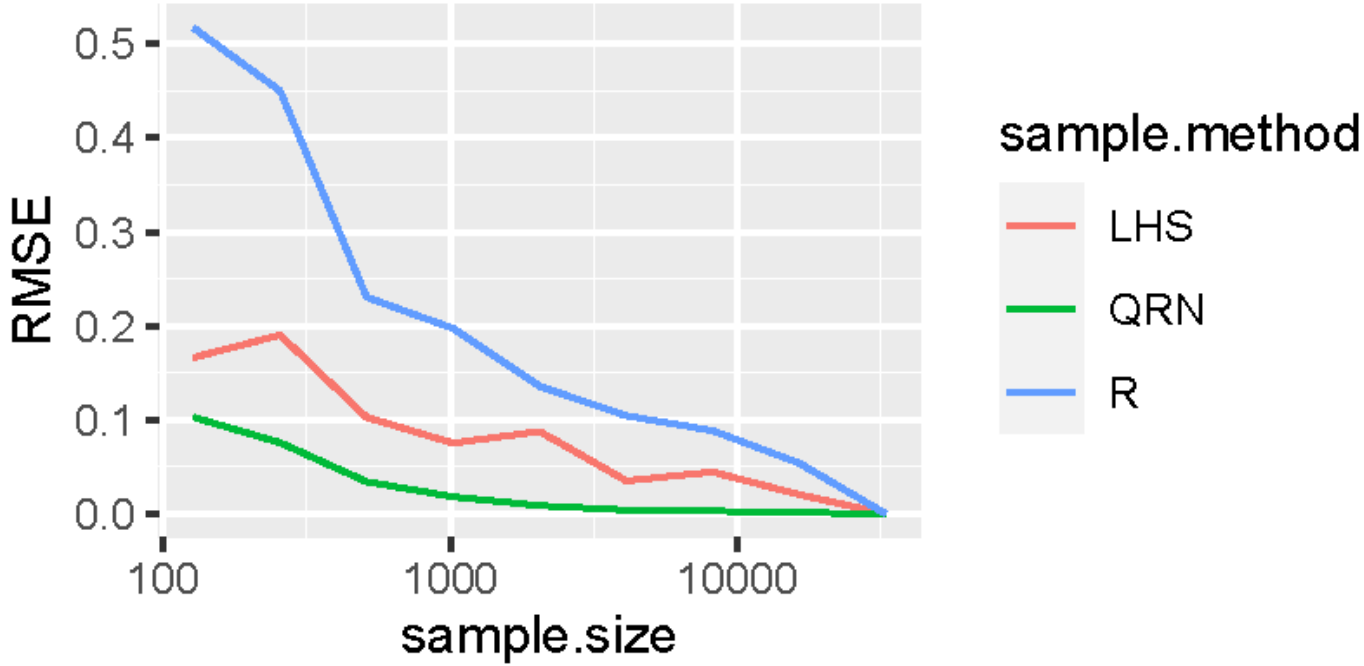
**Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling Techniques**

Sergei Kucherenko, Daniel Albrecht, Andrea Saltelli

# Sobol' LP-TAU are used in high frequency trading



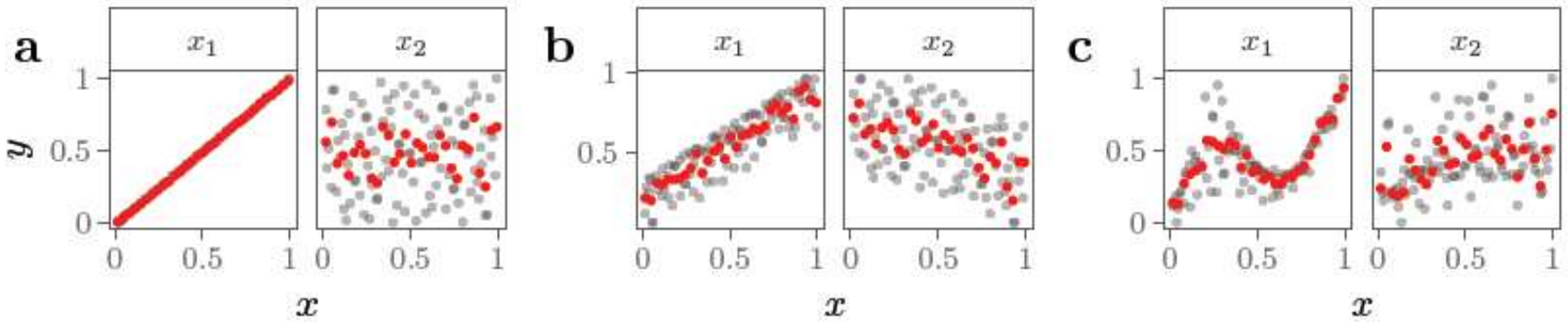
Source: https://www.youtube.com/watch?v=z4nCTdQlH




# Root mean square error with different designs

# Sensitivity analysis made easy

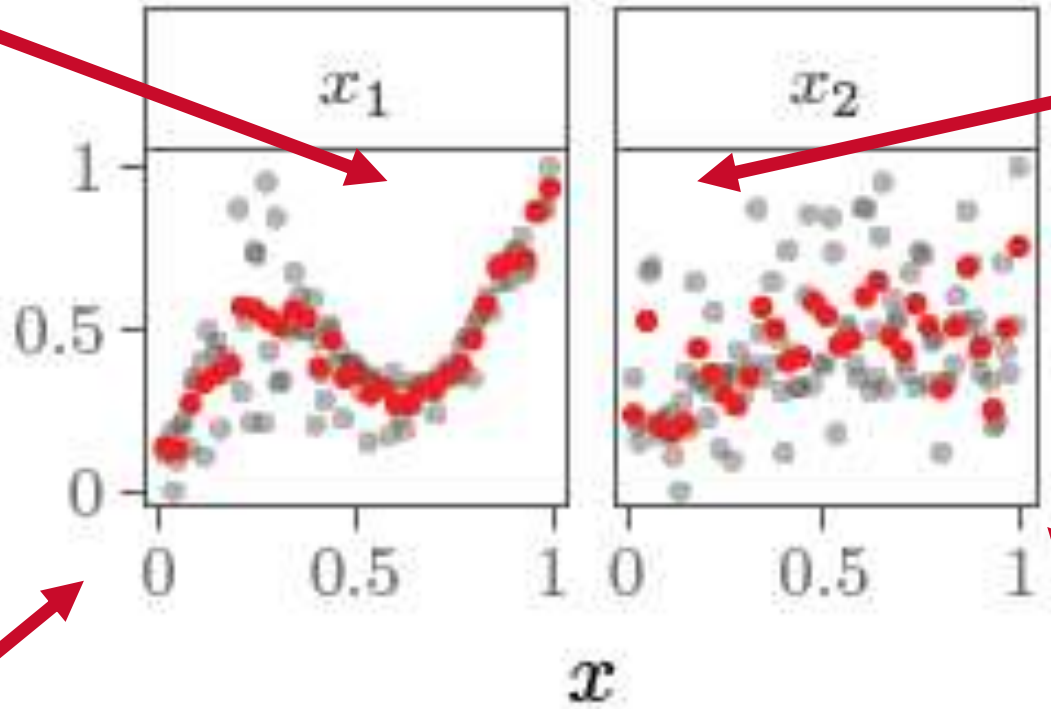
# Do we need to compute indices? Can we do without statistics and calculus using the histograms we have met already?




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

Bigger ‘holes’

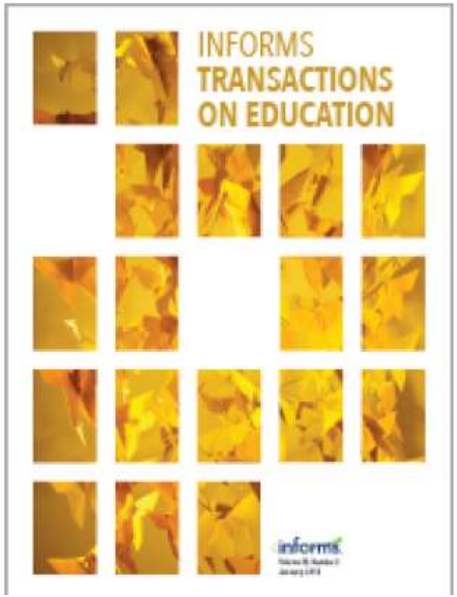
Smaller ‘holes’




=more important

=less important

# Another way to bypass statistics and calculus



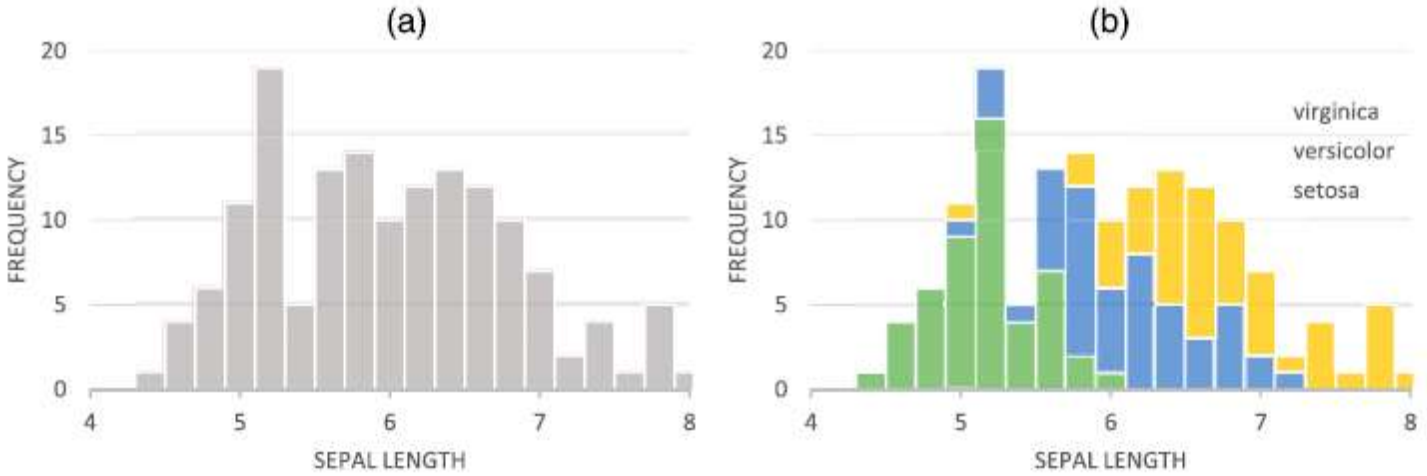

**INFORMS Transactions on Education**

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 …

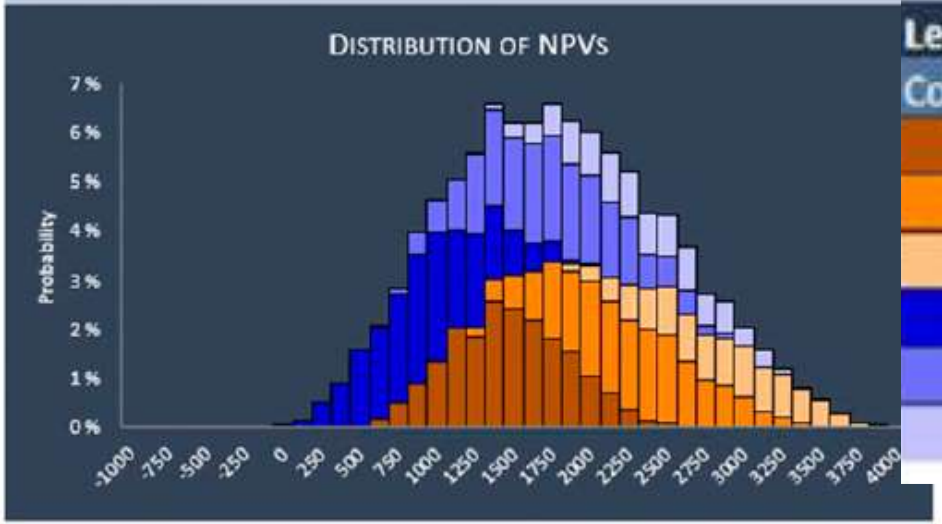


# … without computing sensitivity indices

**RESULTS**

| | | | |
|---|---|---|---|
| Min | 100 € | Probability of negative NPV | 0 % |
| **Expected mean** | **1 855 €** | Probability of positive NPV | 100 % |
| Max | 3 925 € | Standard deviation | 713 € |

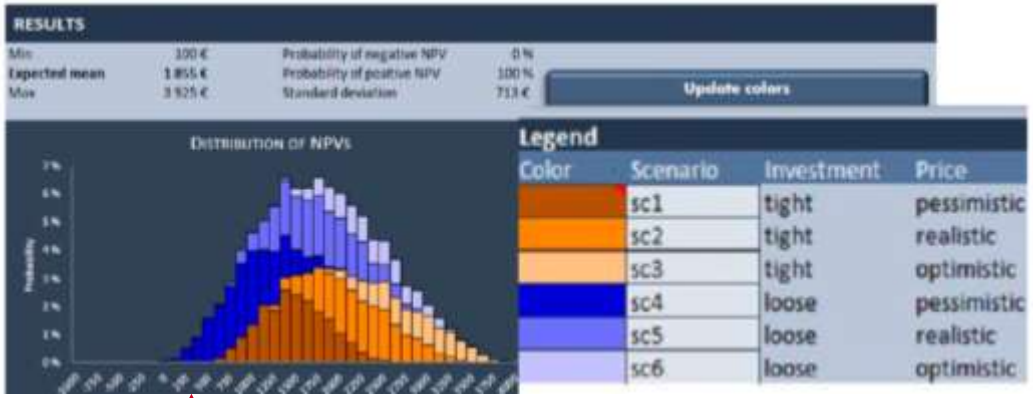
**Update colors**




**Legend**

| Color | Scenario | Investment | Price |
|---|---|---|---|
| | sc1 | tight | pessimistic |
| | sc2 | tight | realistic |
| | sc3 | tight | optimistic |
| | sc4 | loose | pessimistic |
| | sc5 | loose | realistic |
| | sc6 | loose | optimistic |

# … without computing sensitivity indices




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

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’

# Don’t run the model just once

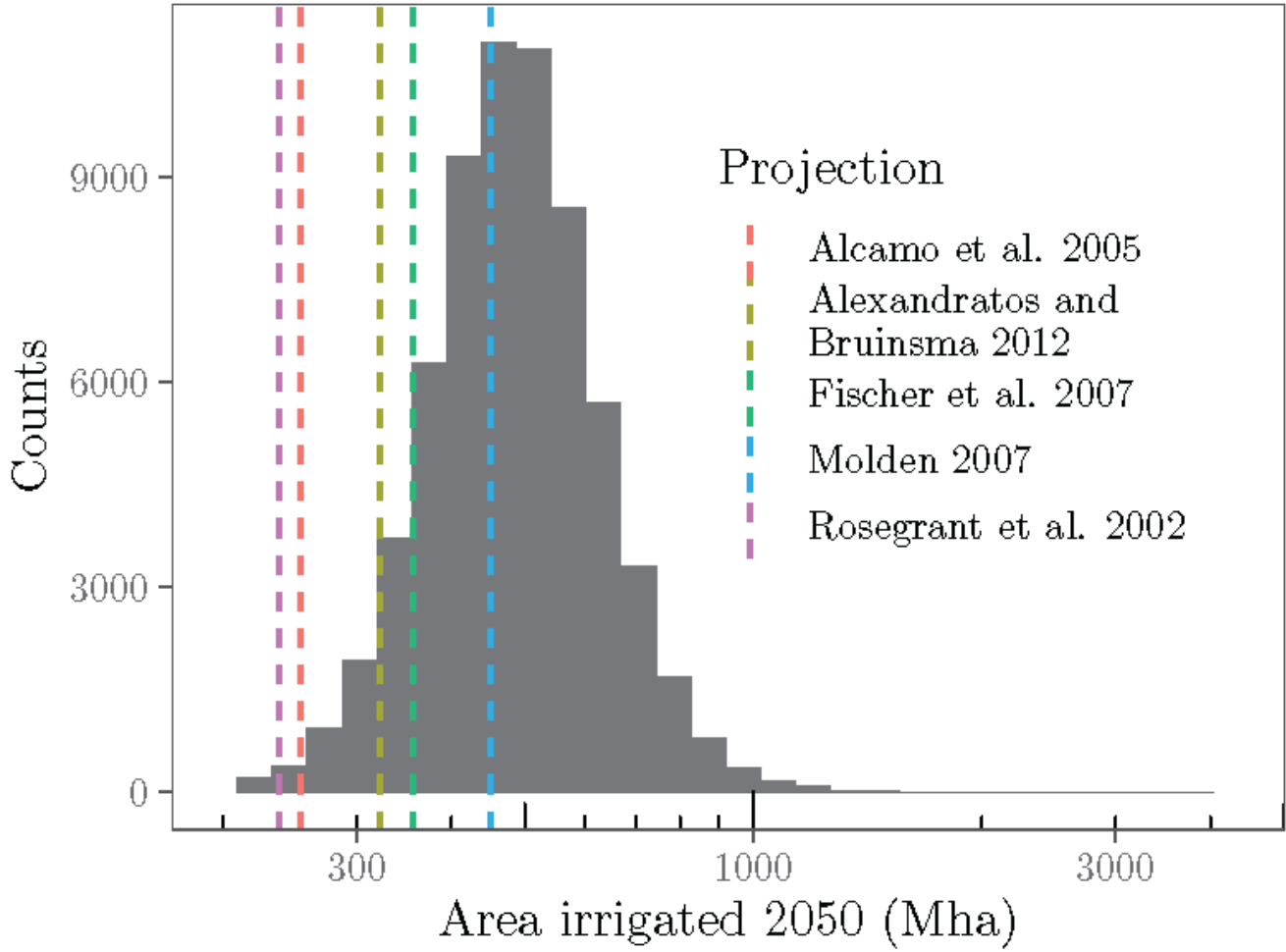
There is much to learn by running the model a few times, especially during model building

# 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

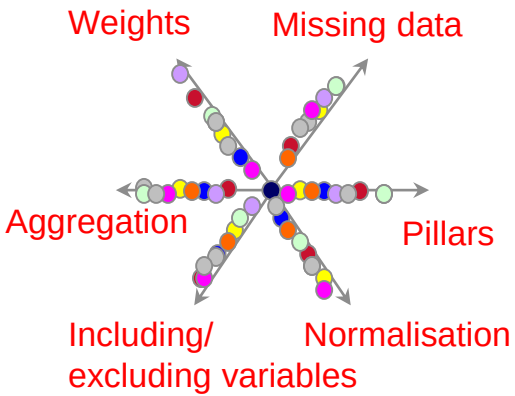
Explore thoroughly the space of the assumptions

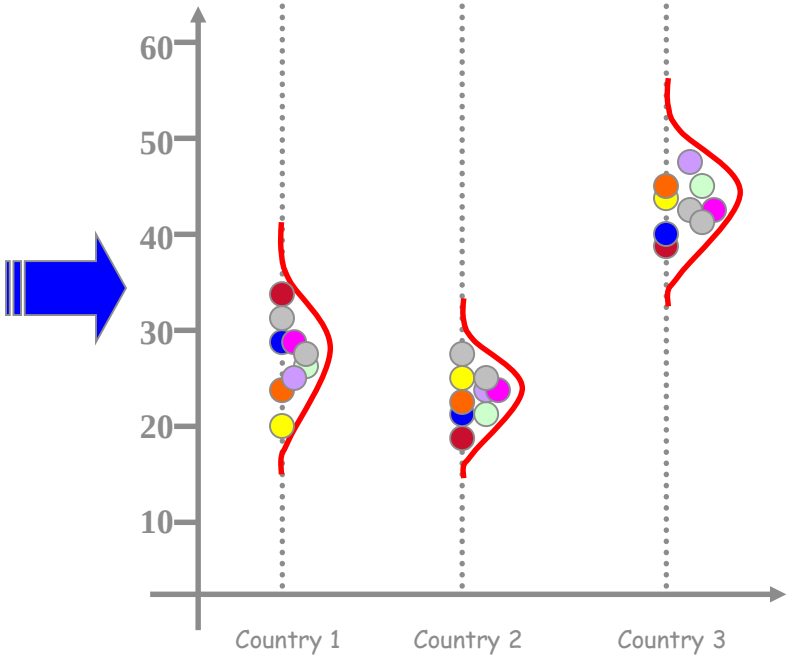
One can sample more than just factors:

- 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,<br>▪ factor analysis,<br>▪ equal weighting,<br>▪ data envelopment analysis |
| Aggregation rule | ▪ additive,<br>▪ multiplicative,<br>▪ Borda multi-criterion |

## Space of alternatives



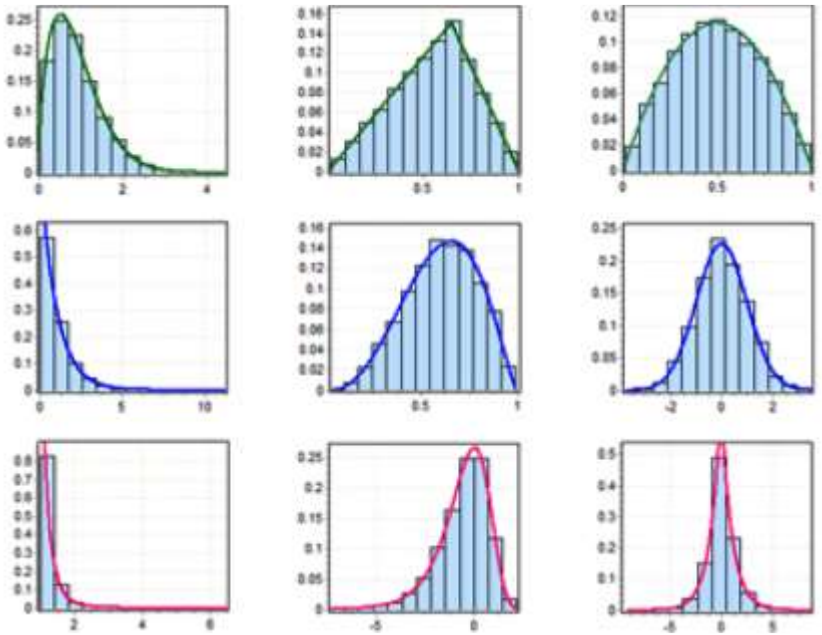



# Building a Monte Carlo analysis

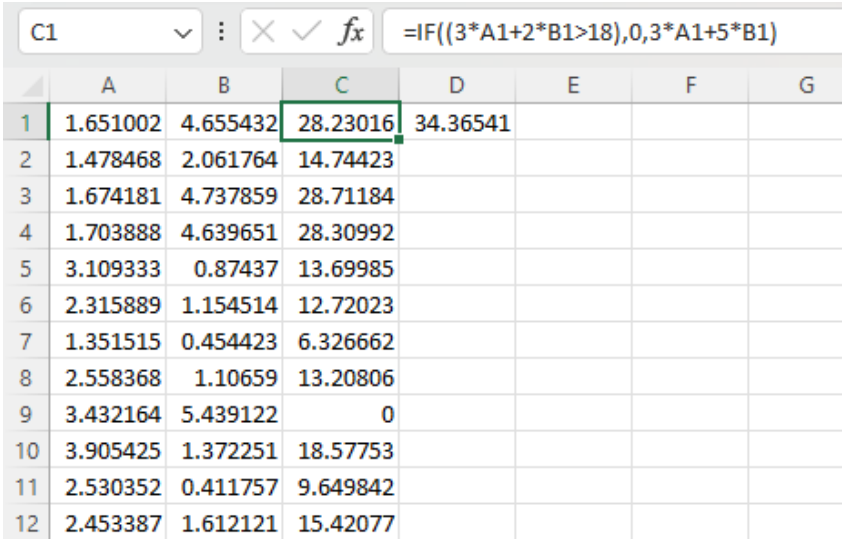
$$\begin{matrix} x_{11} & x_{12\ldots} & x_{1k} \\ x_{21} & x_{22\ldots} & x_{2k} \\ \ldots & \ldots & \ldots \\ x_{N1} & x_{N2} & x_{Nk1} \end{matrix}$$

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




| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | 1.651002 | 4.655432 | 28.23016 | 34.36541 | | | |
| 2 | 1.478468 | 2.061764 | 14.74423 | | | | |
| 3 | 1.674181 | 4.737859 | 28.71184 | | | | |
| 4 | 1.703888 | 4.639651 | 28.30992 | | | | |
| 5 | 3.109333 | 0.87437 | 13.69985 | | | | |
| 6 | 2.315889 | 1.154514 | 12.72023 | | | | |
| 7 | 1.351515 | 0.454423 | 6.326662 | | | | |
| 8 | 2.558368 | 1.10659 | 13.20806 | | | | |
| 9 | 3.432164 | 5.439122 | 0 | | | | |
| 10 | 3.905425 | 1.372251 | 18.57753 | | | | |
| 11 | 2.530352 | 0.411757 | 9.649842 | | | | |
| 12 | 2.453387 | 1.612121 | 15.42077 | | | | |

$x_1 = 4 * rand()$

$Z$ in the feasible region

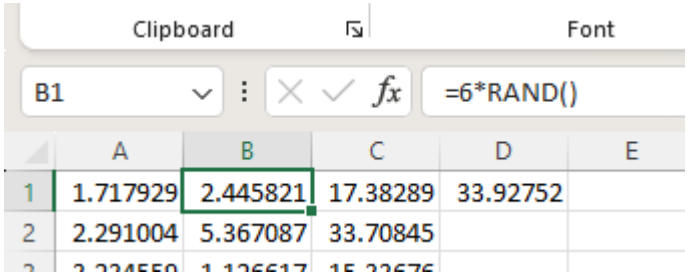
$x_2 = 6 * rand()$

**TABLE 3.1** Data for the Wyndor Glass Co. problem

| Plant | Production Time per Batch, Hours: Product 1 | 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 $Z$

$rand()$ returns a number between zero and one




| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | 1.717929 | 2.445821 | 17.38289 | 33.92752 | |
| 2 | 2.291004 | 5.367087 | 33.70845 | | |

upf. BARCELONA SCHOOL OF MANAGEMENT

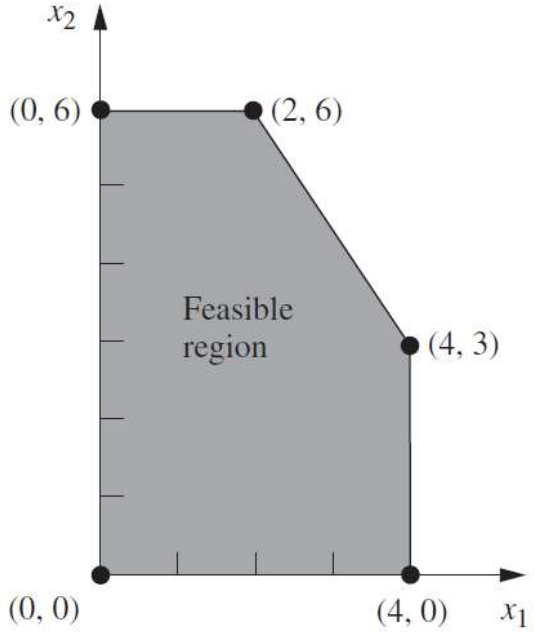
C1 | =IF((3*A1+2*B1>18),0,3*A1+5*B1)

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | 1.651002 | 4.655432 | 28.23016 | 34.36541 | | | |
| 2 | 1.478468 | 2.061764 | 14.74423 | | | | |
| 3 | 1.674181 | 4.737859 | 28.71184 | | | | |
| 4 | 1.703888 | 4.639651 | 28.30992 | | | | |
| 5 | 3.109333 | 0.87437 | 13.69985 | | | | |
| 6 | 2.315889 | 1.154514 | 12.72023 | | | | |
| 7 | 1.351515 | 0.454423 | 6.326662 | | | | |
| 8 | 2.558368 | 1.10659 | 13.20806 | | | | |
| 9 | 3.432164 | 5.439122 | 0 | | | | |
| 10 | 3.905425 | 1.372251 | 18.57753 | | | | |
| 11 | 2.530352 | 0.411757 | 9.649842 | | | | |
| 12 | 2.453387 | 1.612121 | 15.42077 | | | | |

$x_1 = 4 * rand()$

$x_2 = 6 * rand()$

$Z$ in the feasible region




C1 | =IF((3*A1+2*B1>18),0,3*A1+5*B1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | 1.651002 | 4.655432 | 28.23016 | 34.36541 | | |

C1 | =IF((3*A1+2*B1>18),0,3*A1+5*B1)

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | 1.651002 | 4.655432 | 28.23016 | 34.36541 | | | |
| 2 | 1.478468 | 2.061764 | 14.74423 | | | | |
| 3 | 1.674181 | 4.737859 | 28.71184 | | | | |
| 4 | 1.703888 | 4.639651 | 28.30992 | | | | |
| 5 | 3.109333 | 0.87437 | 13.69985 | | | | |
| 6 | 2.315889 | 1.154514 | 12.72023 | | | | |
| 7 | 1.351515 | 0.454423 | 6.326662 | | | | |
| 8 | 2.558368 | 1.10659 | 13.20806 | | | | |
| 9 | 3.432164 | 5.439122 | 0 | | | | |
| 10 | 3.905425 | 1.372251 | 18.57753 | | | | |
| 11 | 2.530352 | 0.411757 | 9.649842 | | | | |
| 12 | 2.453387 | 1.612121 | 15.42077 | | | | |

$x_1 = 4 * rand()$

$x_2 = 6 * rand()$

$Z$ in the feasible region

## Syntax of IF in Excel

Return the value of 0 or 3*A1+ 5*B1 depending upon whether 3*A1+ 2*B1 is larger that 18 or not

C1 | =IF((3*A1+2*B1>18),0,3*A1+5*B1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | 1.651002 | 4.655432 | 28.23016 | 34.36541 | | |

BARCELONA SCHOOL OF MANAGEMENT

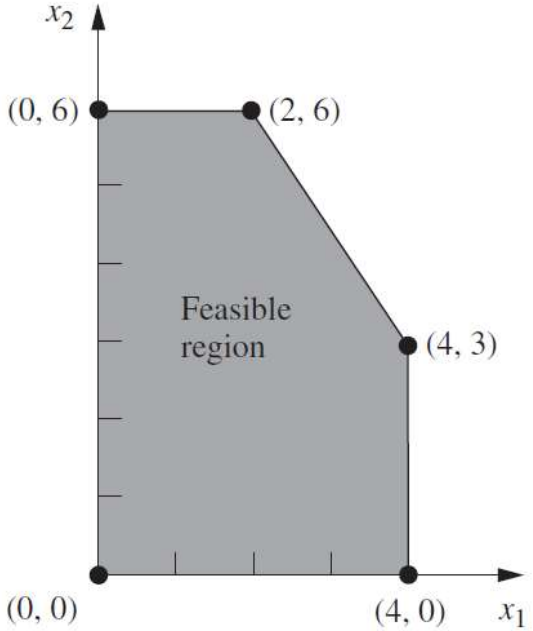
D1 | =MAX(C1:C100)

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | 1.651002 | 4.655432 | 28.23016 | 34.36541 | |
| 2 | 1.478468 | 2.061764 | 14.74423 | | |
| 3 | 1.674181 | 4.737859 | 28.71184 | | |
| 4 | 1.703888 | 4.639651 | 28.30992 | | |
| 5 | 3.109333 | 0.87437 | 13.69985 | | |
| 6 | 2.315889 | 1.154514 | 12.72023 | | |
| 7 | 1.351515 | 0.454423 | 6.326662 | | |
| 8 | 2.558368 | 1.10659 | 13.20806 | | |
| 9 | 3.432164 | 5.439122 | 0 | | |
| 10 | 3.905425 | 1.372251 | 18.57753 | | |
| 11 | 2.530352 | 0.411757 | 9.649842 | | |

$x_1 = 4 * rand()$

$x_2 = 6 * rand()$

$Z$ in the feasible region




Find the maximum over the 100 points tested

BARCELONA SCHOOL OF MANAGEMENT

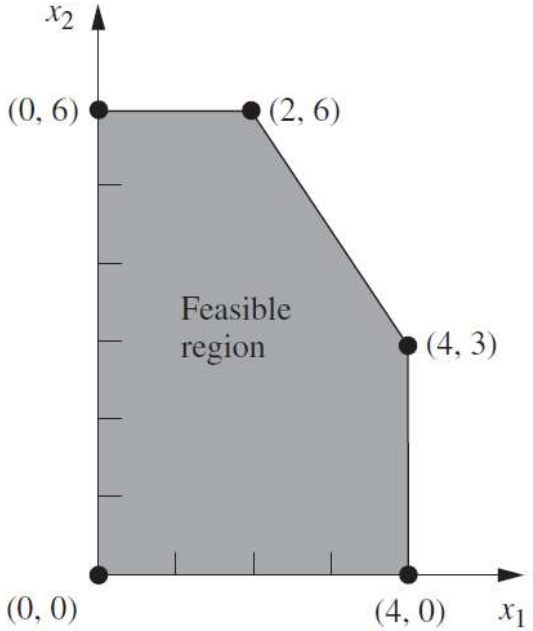
C1 | =IF((3*A1+2*B1>18),0,3*A1+5*B1)

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | 1.651002 | 4.655432 | 28.23016 | 34.36541 | | | |
| 2 | 1.478468 | 2.061764 | 14.74423 | | | | |
| 3 | 1.674181 | 4.737859 | 28.71184 | | | | |
| 4 | 1.703888 | 4.639651 | 28.30992 | | | | |
| 5 | 3.109333 | 0.87437 | 13.69985 | | | | |
| 6 | 2.315889 | 1.154514 | 12.72023 | | | | |
| 7 | 1.351515 | 0.454423 | 6.326662 | | | | |
| 8 | 2.558368 | 1.10659 | 13.20806 | | | | |
| 9 | 3.432164 | 5.439122 | 0 | | | | |
| 10 | 3.905425 | 1.372251 | 18.57753 | | | | |
| 11 | 2.530352 | 0.411757 | 9.649842 | | | | |
| 12 | 2.453387 | 1.612121 | 15.42077 | | | | |

$x_1 = 4 * rand()$

$x_2 = 6 * rand()$

$Z$ in the feasible region




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 $Z$

upf. BARCELONA SCHOOL OF MANAGEMENT

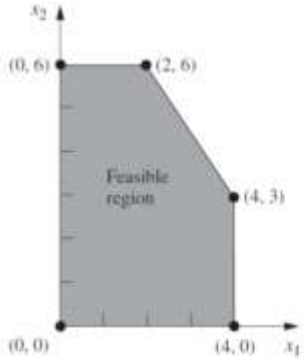
C1 =IF((3*A1+2*B1>18),0,3*A1+5*B1)

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | 1.651002 | 4.655432 | 28.29016 | 34.36541 | | | |
| 2 | 1.478468 | 2.061764 | 14.74423 | | | | |
| 3 | 1.674181 | 4.737859 | 28.71184 | | | | |
| 4 | 1.703888 | 4.639651 | 28.30992 | | | | |
| 5 | 3.109333 | 0.87437 | 13.69985 | | | | |
| 6 | 2.315889 | 1.154514 | 12.72023 | | | | |
| 7 | 1.351515 | 0.454423 | 6.326662 | | | | |
| 8 | 2.558368 | 1.10659 | 13.20806 | | | | |
| 9 | 3.432164 | 5.439122 | 0 | | | | |
| 10 | 3.905425 | 1.372251 | 18.57753 | | | | |
| 11 | 2.530352 | 0.411757 | 9.649842 | | | | |
| 12 | 2.453387 | 1.612121 | 15.42077 | | | | |

$x_1 = 4 * rand()$

$x_2 = 6 * rand()$

$Z$ in the feasible region




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 $Z$

136

This is just for illustration! We treat genetic algorithms later in the course, but the literature on metaheuristic is vast (Chapter 14 in Hillier)

upf. BARCELONA SCHOOL OF MANAGEMENT

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*

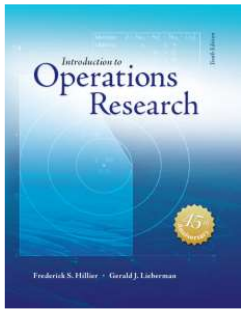
**Brazzaville –** Eighty-three thousand to 190 000 people in Africa could die of COVID-19 and 29 million to 44 million could get infected in the first year of the pandemic if containment measures fail, a new study by the World Health Organization (WHO) Regional Office for Africa finds. The research,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

"OR teams typically spend a surprisingly large amount of time *gathering relevant data* about the problem", Hillier, p. 12



# Why is all this important? Fishing expeditions and forking paths …

Jorge Luis Borges
(1899–1986)


Taking different storylines within the same novel like Ts'ui Pên

**The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time***

Andrew Gelman† and Eric Loken‡

14 Nov 2013

# The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time*

Andrew Gelman† and Eric Loken‡

14 Nov 2013

'Fishing expedition'? An analyst changing the question asked from the data to squeeze out a publishable effect

'P-hacking'? an analyst torturing the data to pass a significance test (remember the tea drinker of Lesson 1)

The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time*

Andrew Gelman† and Eric Loken‡

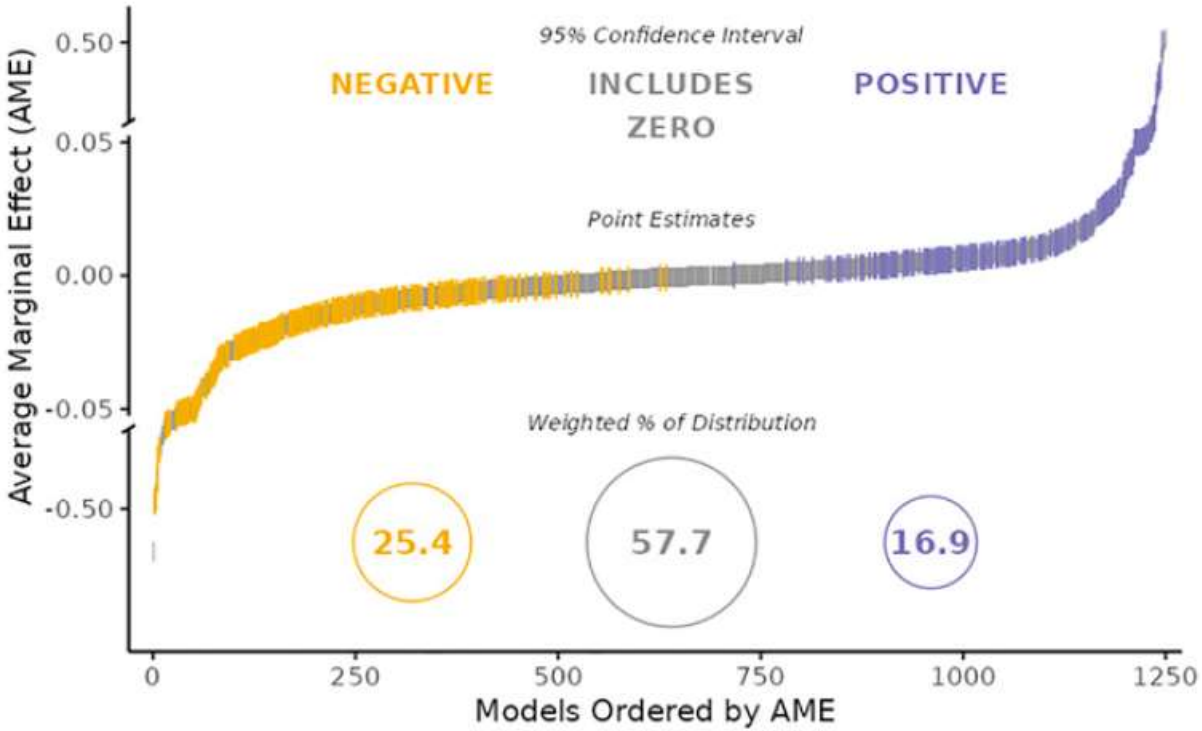
14 Nov 2013

# Why this matters? 🤔

PNAS RESEARCH ARTICLE | SOCIAL SCIENCES OPEN

## Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty

Edited by Douglas Massey, Princeton University, Princeton, NJ; received March 6, 2022; accepted August 22, 2022

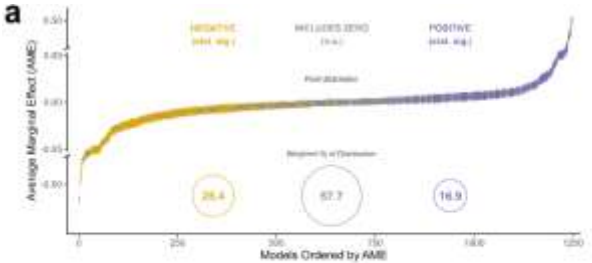



"Will different researchers [73 teams] converge on similar findings when analyzing the same data?

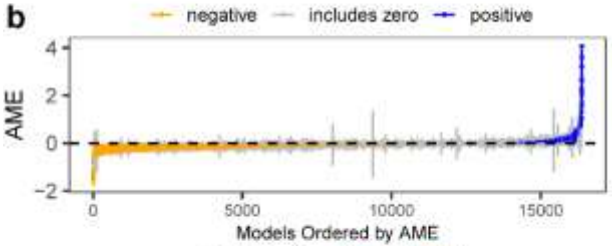
…teams' results varied greatly, ranging from large negative to large positive effects"
(Massey et al. 2022)

# Global sensitivity analysis can chart the garden before you enter …

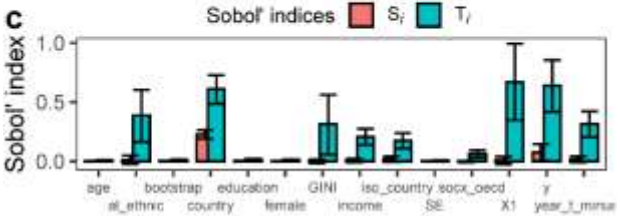
What the analyst did in the garden




A simulation of walking in the garden




What sensitivity analysis says about the choices



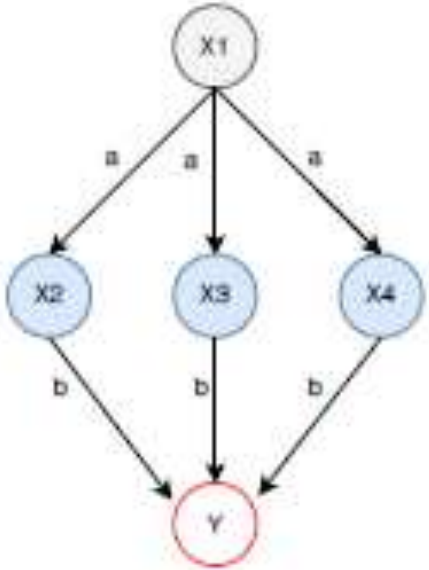

Andrea Saltelli, Arnald Puy, Alessio Lachi, and Nate Breznau, 2024, Global sensitivity analysis unveils the hidden universe of uncertainty in multiverse studies, MetaArXiv Preprints, https://osf.io/preprints/metaarxiv/b67w9.

Another use for uncertainty and sensitivity analysis

In machine learning, check that the algorithm is ‘fair’

PROTECTED ATTRIBUTES:
- Age
- Disability
- National Origin
- Race/color
- Religion
- Sex
- (From the US Equal Opportunity Employment Commission)

Ascertain that an algorithm does not make implicit use of protected attributes (for example in the graph $Y$ must not depend upon $x_1$)




For Machine Learning students: Bénesse, C., Gamboa, F., Loubes, J.-M., & Boissin, T. (2022). Fairness seen as Global Sensitivity Analysis. *Machine Learning*. https://doi.org/10.1007/s10994-022-06202-y

# Homework

1. Compute the chance of having exactly 5 heads throwing a coin 8 times.
2. Re-do manually the exercise of slides 25-27 with these new values for the constraints

| OLD | NEW |
|---|---|
| $x_1 \leq 4$ | $x_1 \leq 5$ |
| $2x_2 \leq 12$ | $2x_2 \leq 13$ |
| $3x_1 + 2x_2 \leq 18$ | $3x_1 + 2x_2 \leq 19$ |

3. The following table gives a two-way classification of all basketball players at a state university who began their college careers between 2001 and 2005, based on gender and whether or not they graduated. If one of these players is selected at random, find the following probabilities:
P(female),
P(male),
P(graduated),
P(non graduated),
P(female and graduated)
=P(female∩)graduated
P(male and did not graduate)
=P(male ∩ did not graduate).

| | Graduated | Did Not Graduate |
|---|---|---|
| Male | 126 | 55 |
| Female | 133 | 32 |

Also find P(graduated and did not graduate) = P(graduated ∩ did not graduate). Is this probability zero? If yes, why?

# Homework

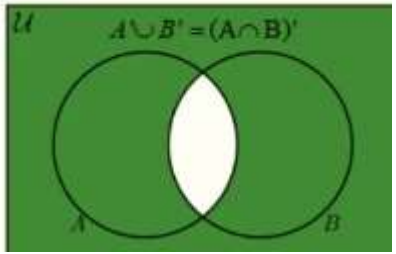
4. Using the results achieved in the preceding exercise compute

P(female or did not graduate) = P(female ∪ did not graduate)

P(graduated or male) = P(graduated ∪ male).

5. A certain state's auto license plates have three letters of the alphabet followed by a three-digit number.
   a. How many different license plates are possible if all three-letter sequences (letter can be repeated) are permitted and any number from 000 to 999 is allowed?
   b. If a witness of a hit-and-run accident says that the first letter on the license plate of the offender's car was a B, that the second letter was an O or a Q, and that the last number was a 5, how many of this state's license plates fit this description?

6. Show graphically that this is true



# 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