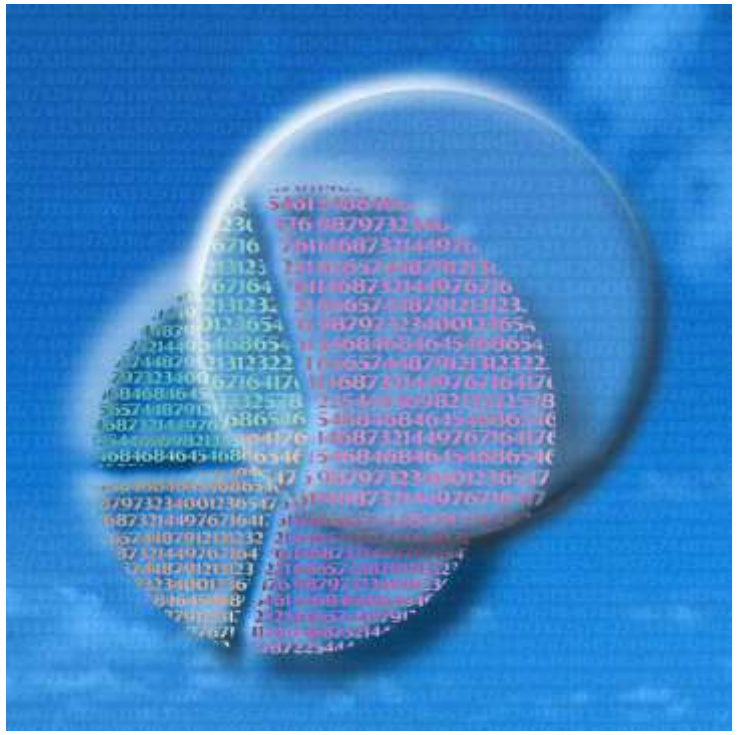


upf. BARCELONA SCHOOL OF MANAGEMENT

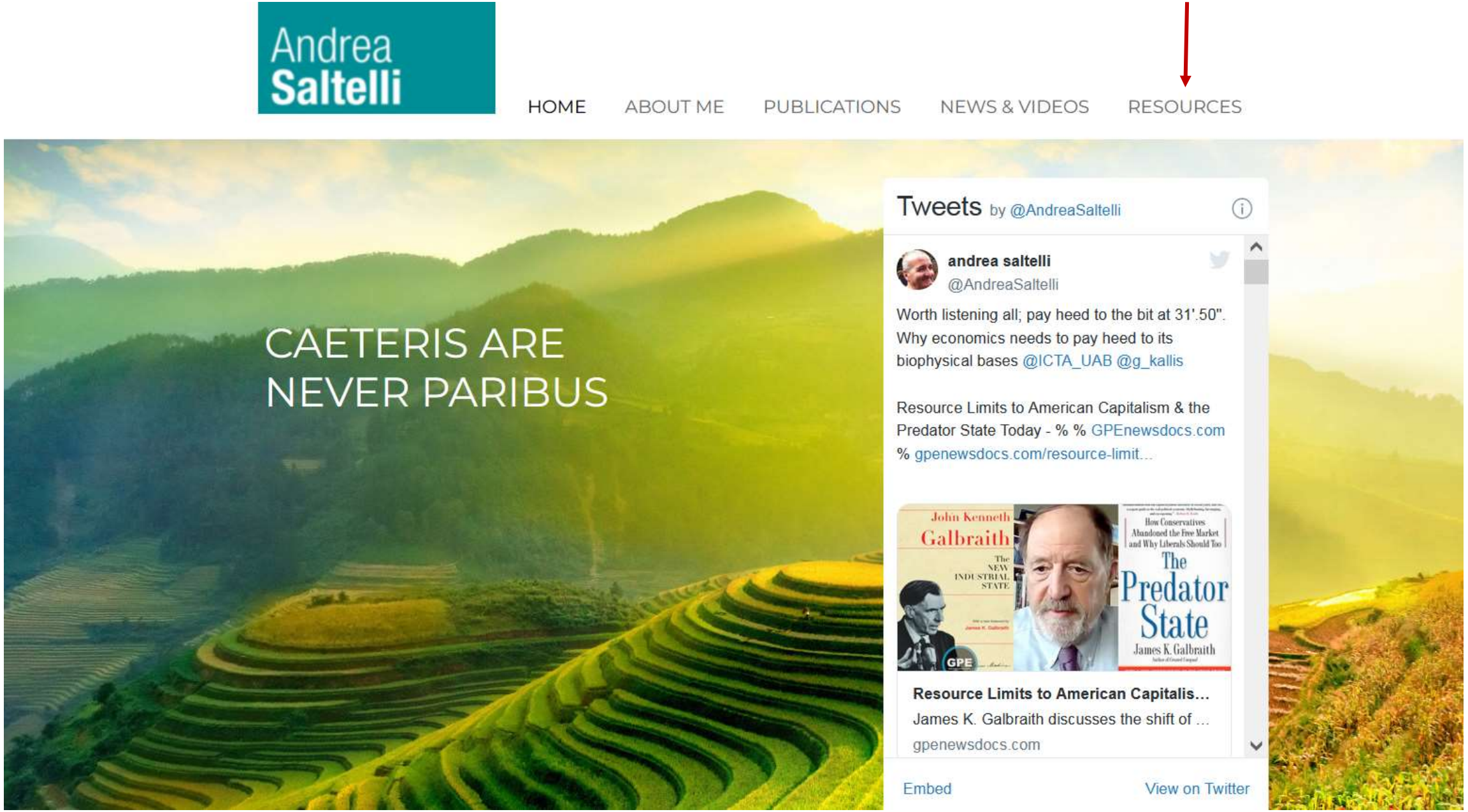
CNR ISTC ISTITUTO DI SCIENZE E TECNOLOGIE DELLA COGNIZIONE

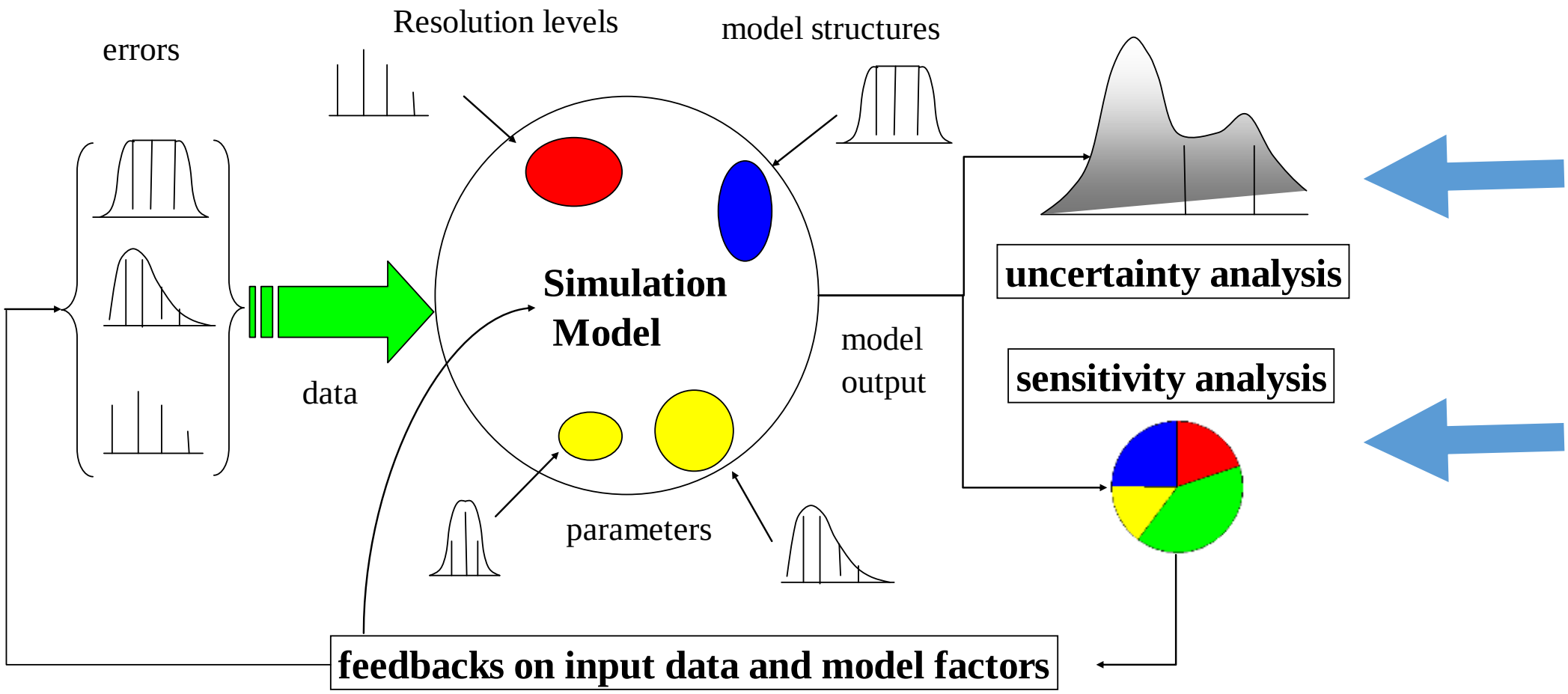
UNIVERSITAS BERGENSIS

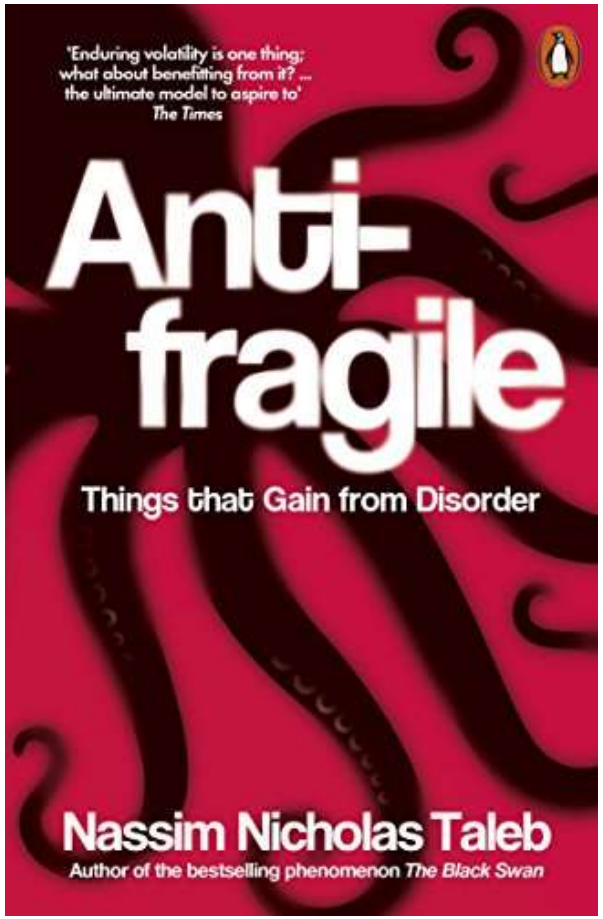
How not to do a sensitivity analysis

Andrea Saltelli, UPF-BSM, Barcelona,
June, 28 2022

Where to find this talk: www.andreasaltelli.eu




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



A short trip through sensitivity analysis borrowing Taleb's *via negativa*

ELSEVIER

Environmental Modelling & Software

Volume 114, April 2019, Pages 29-39

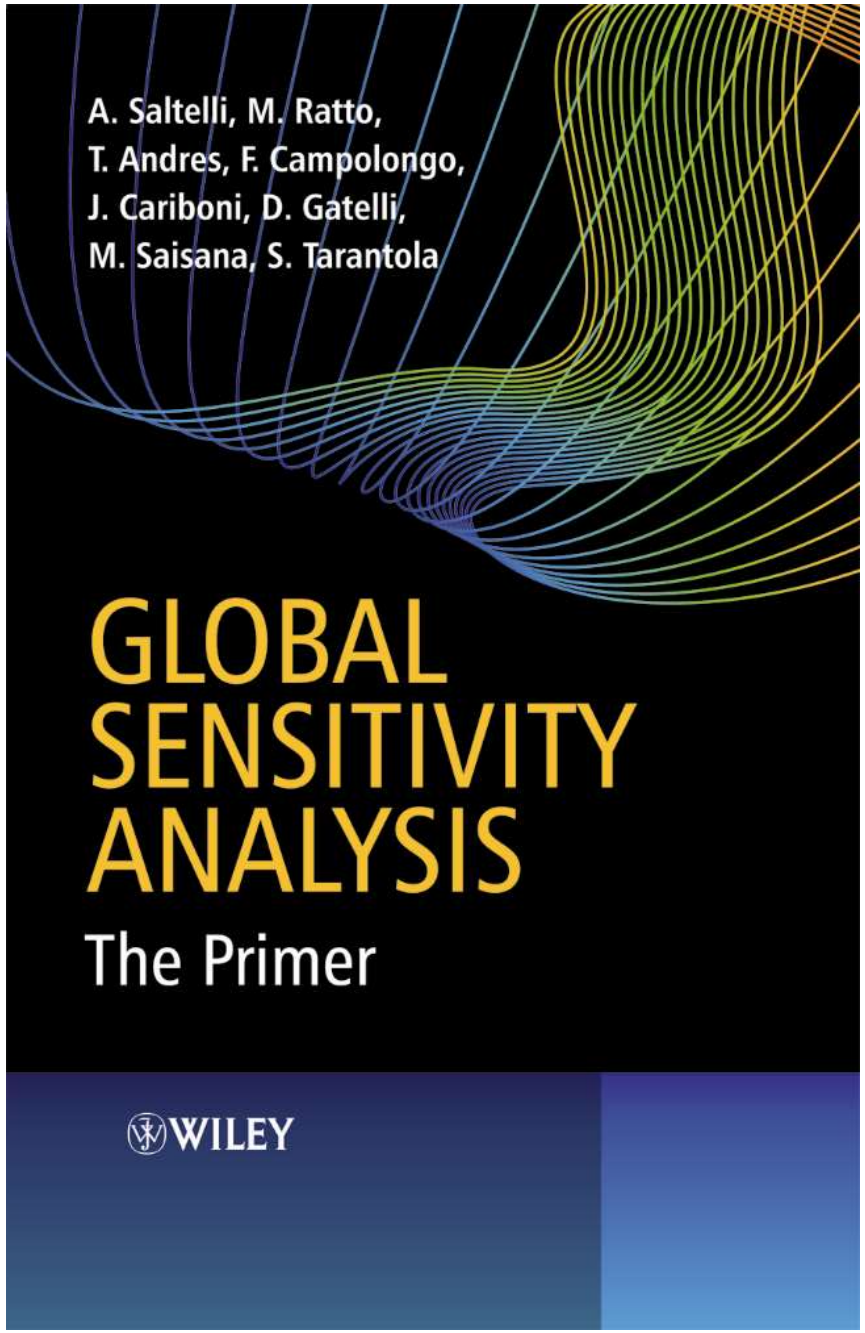
Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices

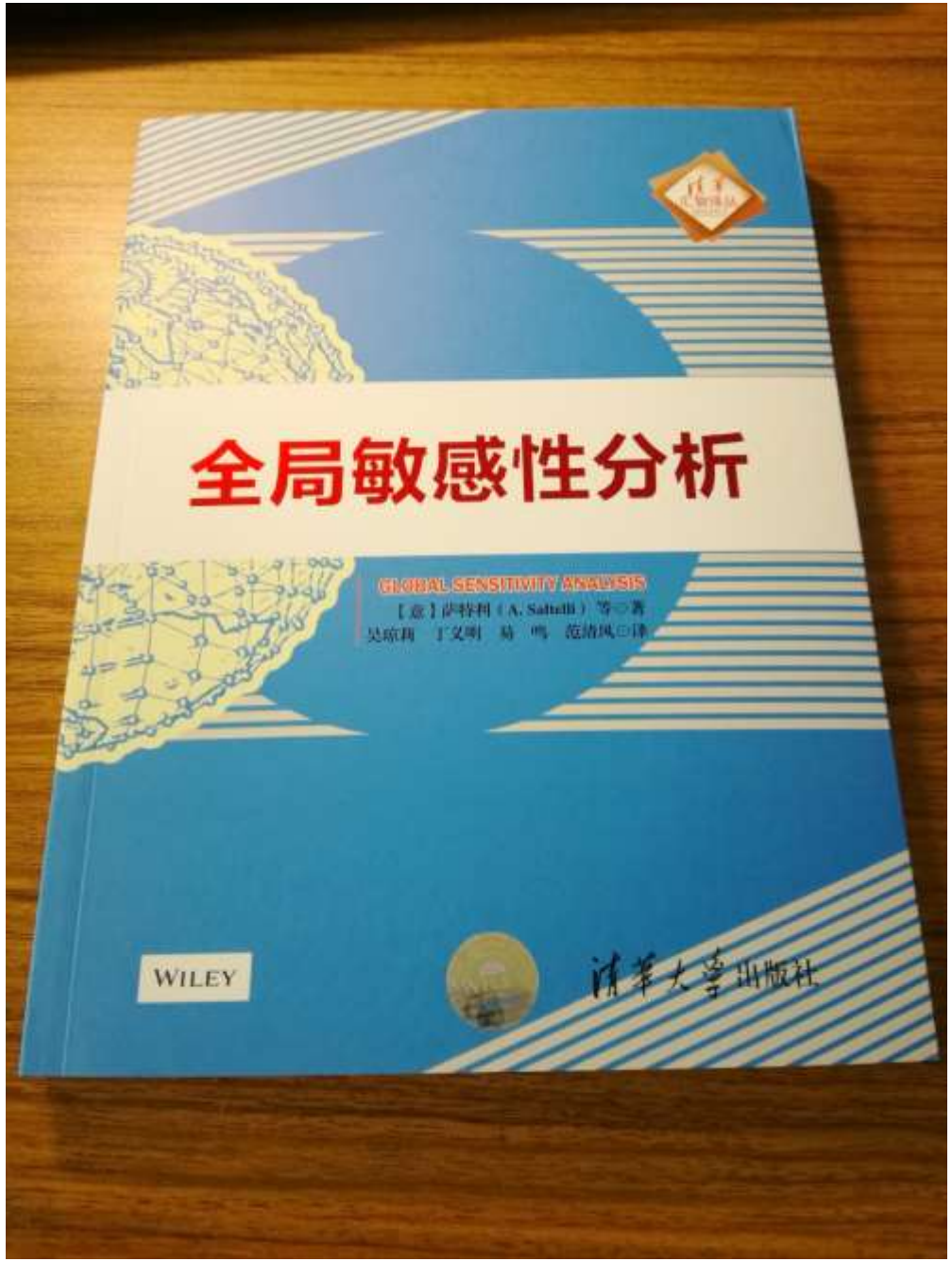
Andrea Saltelli [a, b], Ksenia Aleksankina [c], William Becker [d], Pamela Fennell [e], Federico Ferretti [d], Niels Holst [f], Sushan Li [g], Qiongli Wu [h]

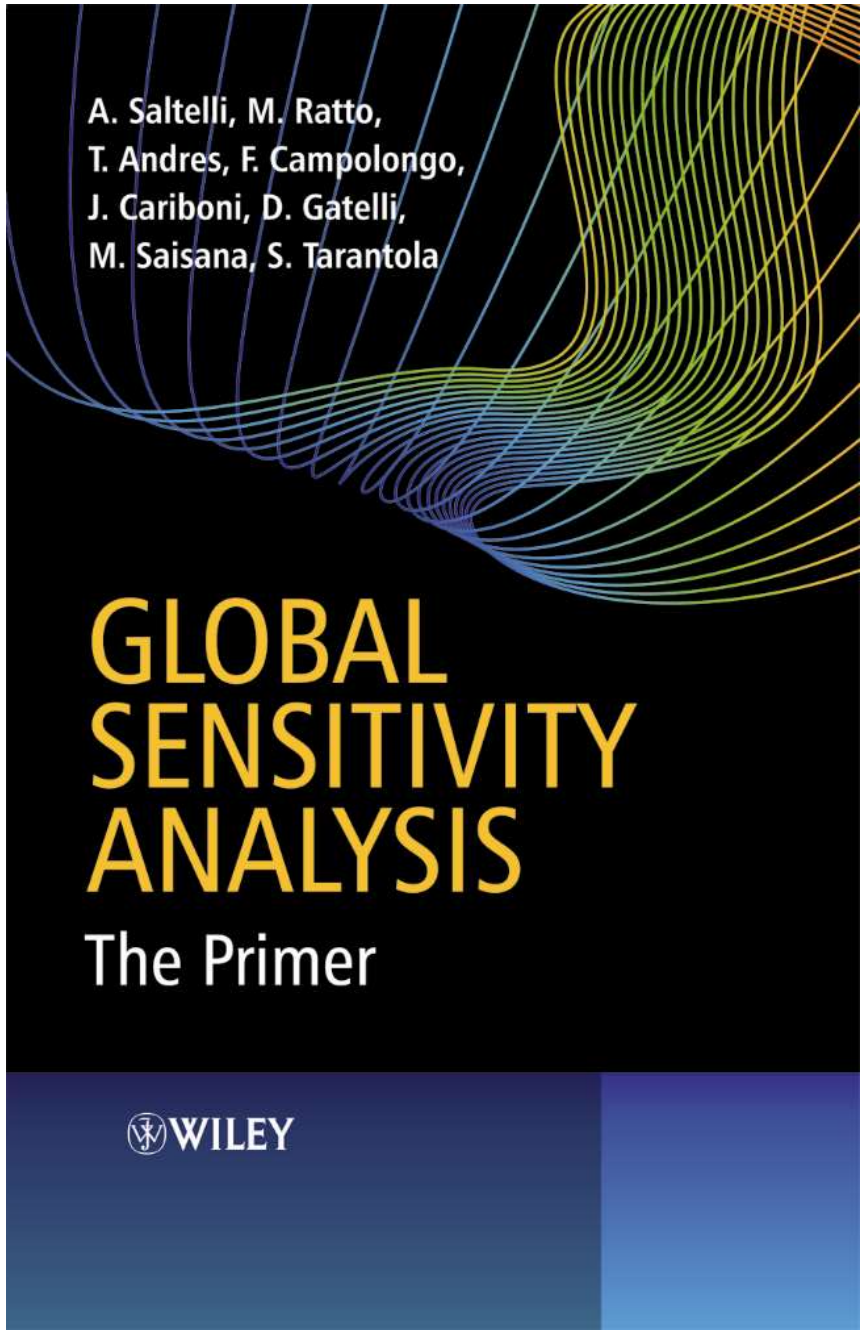
Don’t use just any method

Use the method appropriate to context and purpose

An introduction to variance based methods

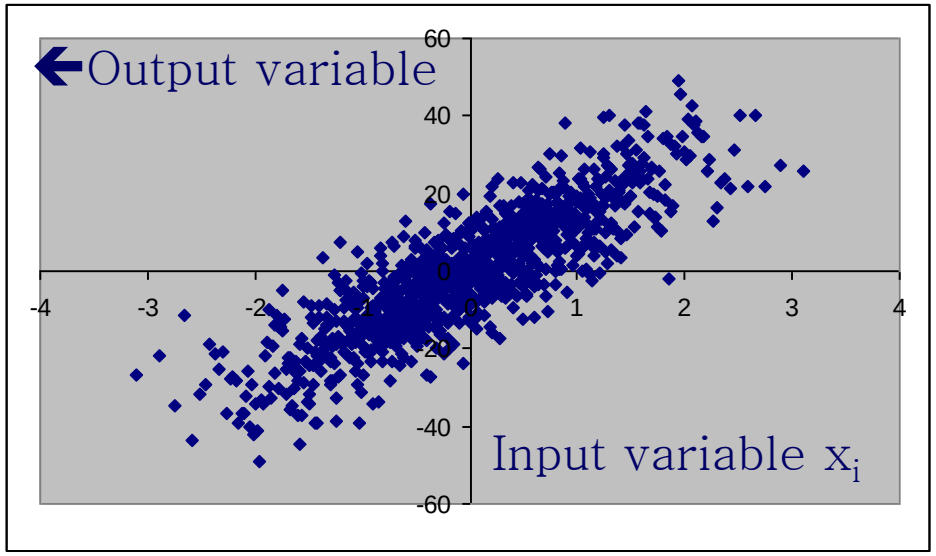

A. Saltelli, M. Ratto,
T. Andres, F. Campolongo,
J. Cariboni, D. Gatelli,
M. Saisana, S. Tarantola
GLOBAL
SENSITIVITY
ANALYSIS
The Primer
WILEY

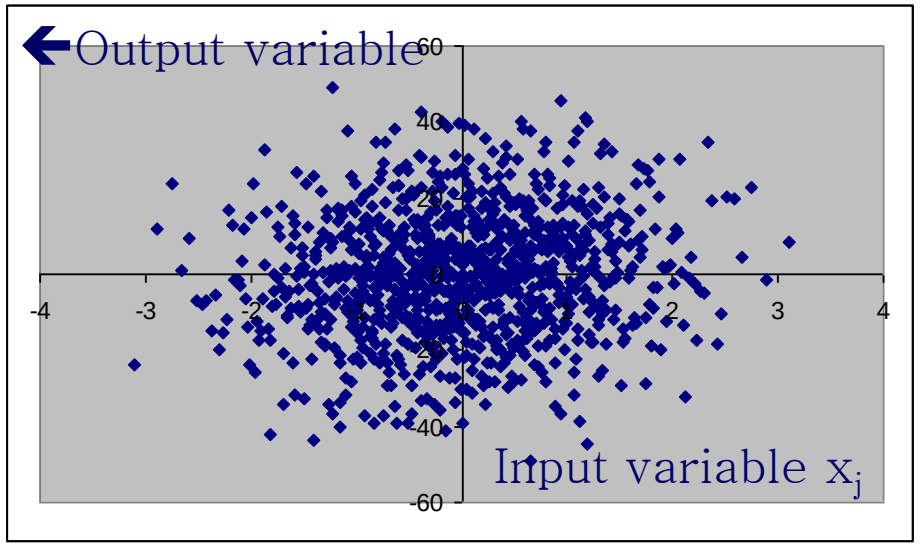

全局敏感性分析
GLOBAL SENSITIVITY ANALYSIS
WILEY



Available for free at

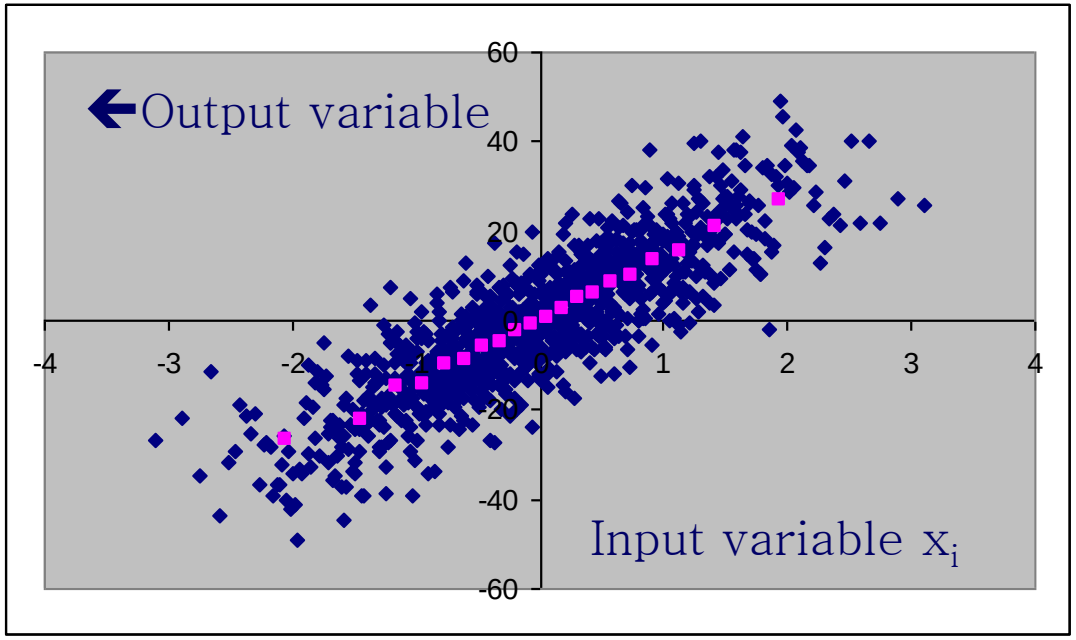
http://www.andreasaltelli.eu

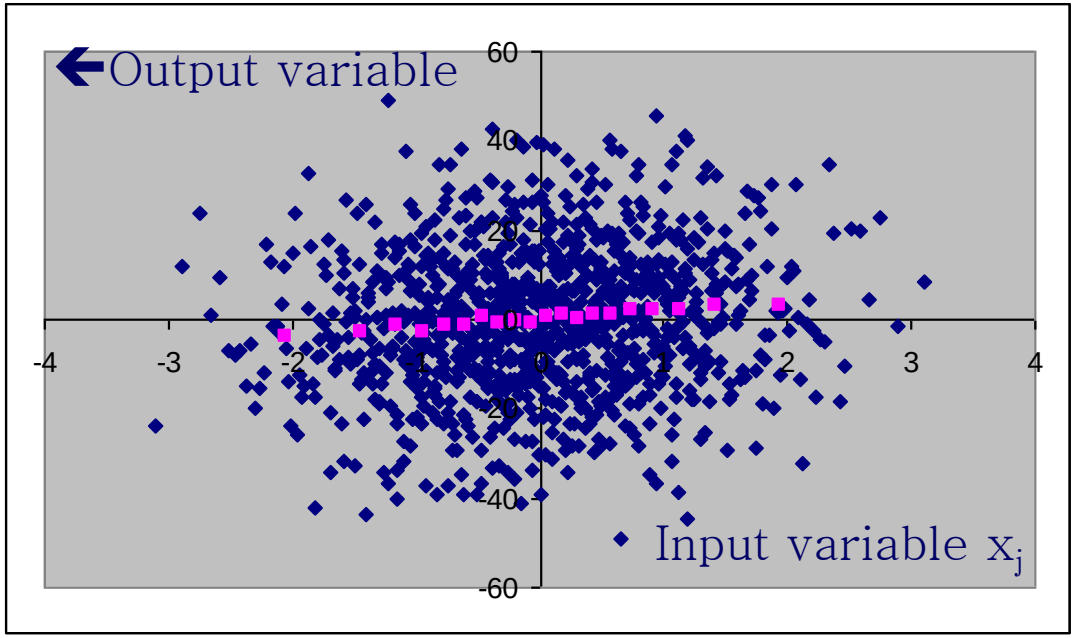




Plotting the output as a function of two different input factors

Which factor is more important?

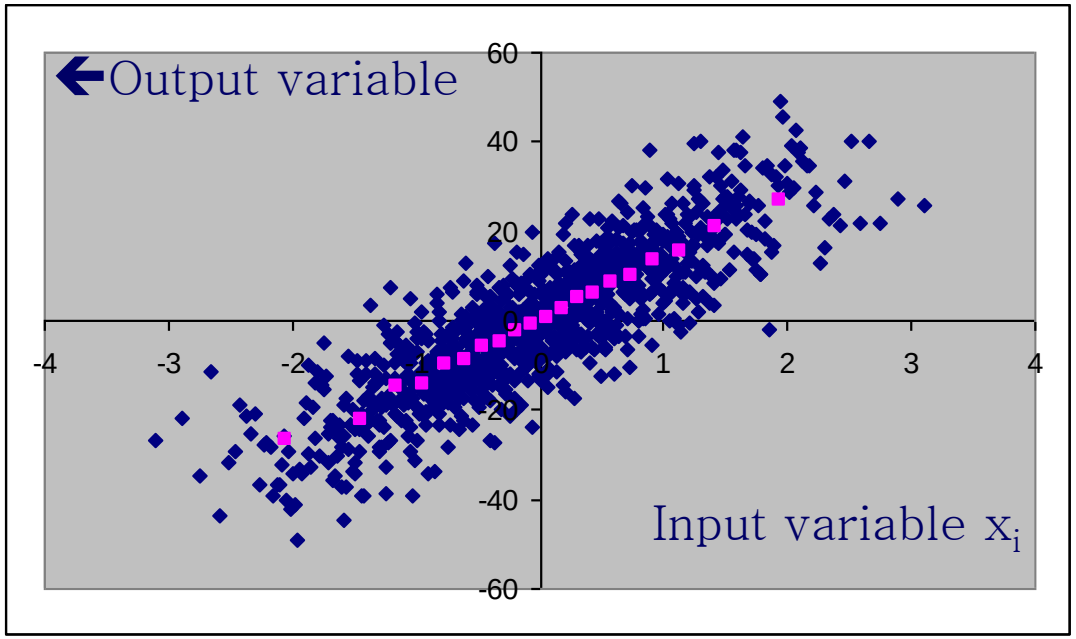




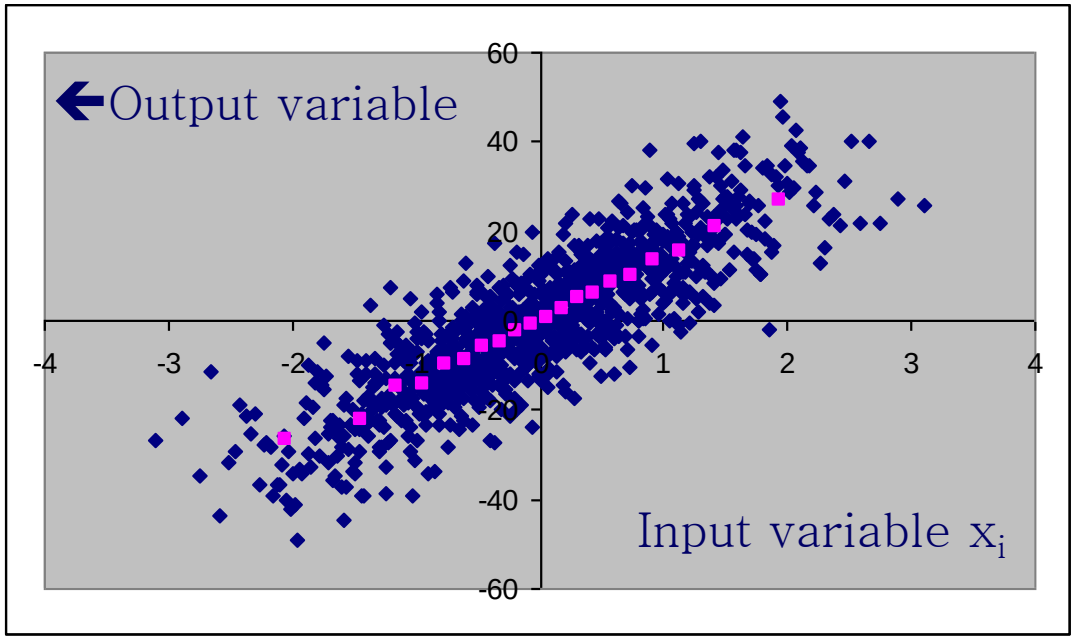
~1,000 blue points

Divide them in 20 bins of ~ 50 points

Compute the bin's average (pink dots)

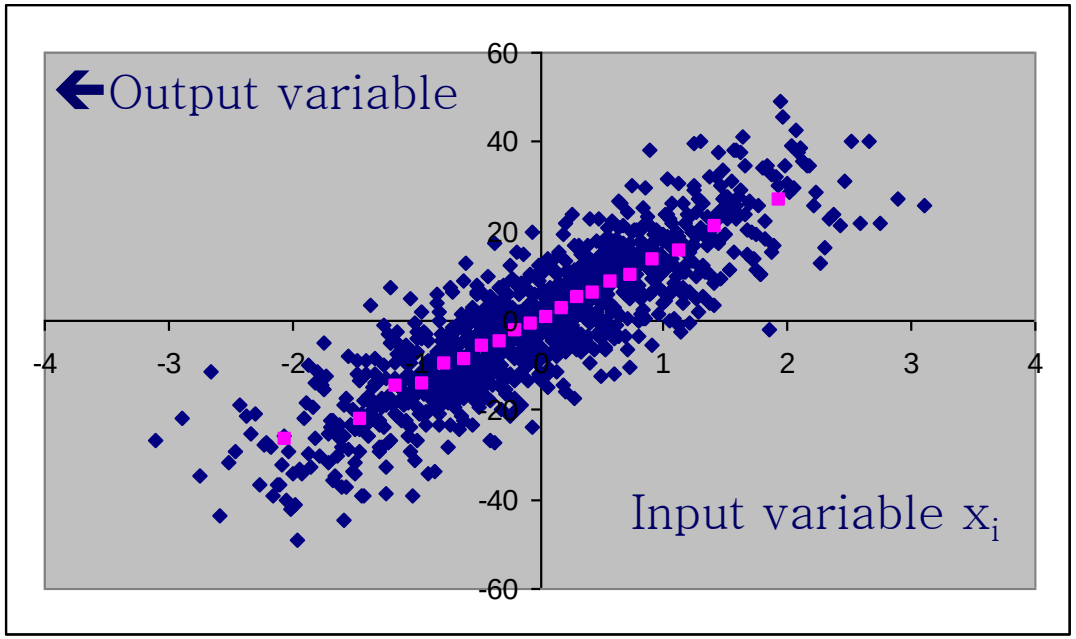


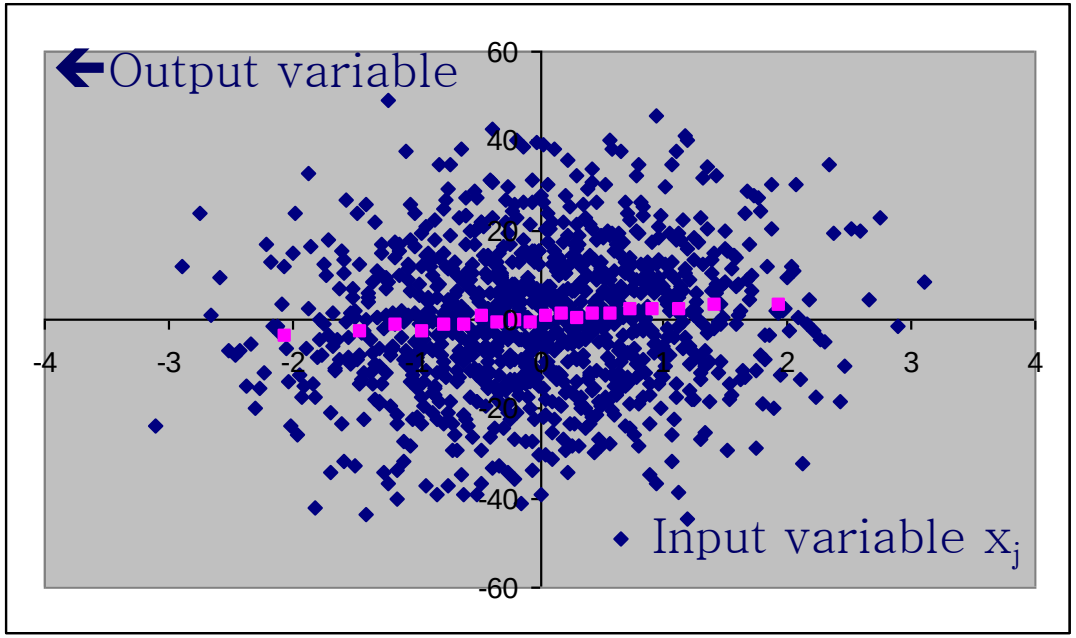
Each pink point is ~ $E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)$



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

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





Which factor has the highest

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

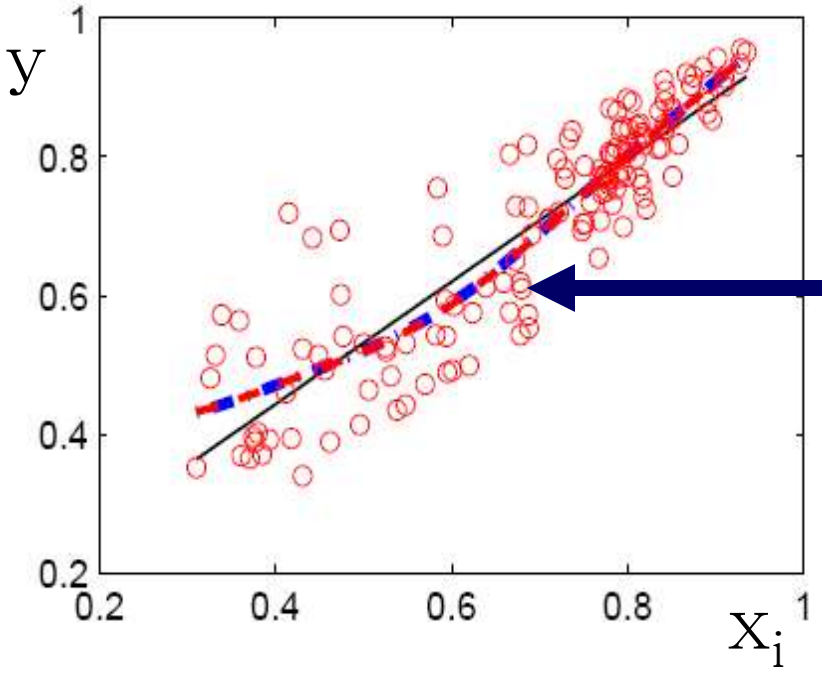
For additive models one can decompose the total variance as a sum of those partial variances

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

… which is also how additive models are defined

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

The partial variance divided by the total variance is the so-called sensitivity index of the first order, identical in formulation to Pearson's correlation ratio



Smoothed curve:

$$\mathrm{E}_{\mathbf{x}_{\sim i}}\left(y \mid x_i\right)$$

First order sensitivity index:

$$\frac{\mathrm{V}_{x_i}\left(\mathrm{E}_{\mathbf{x}_{\sim i}}\left(y \mid x_i\right)\right)}{\mathrm{V}\left(y\right)}$$

Pearson's correlation ratio

Smoothed curve

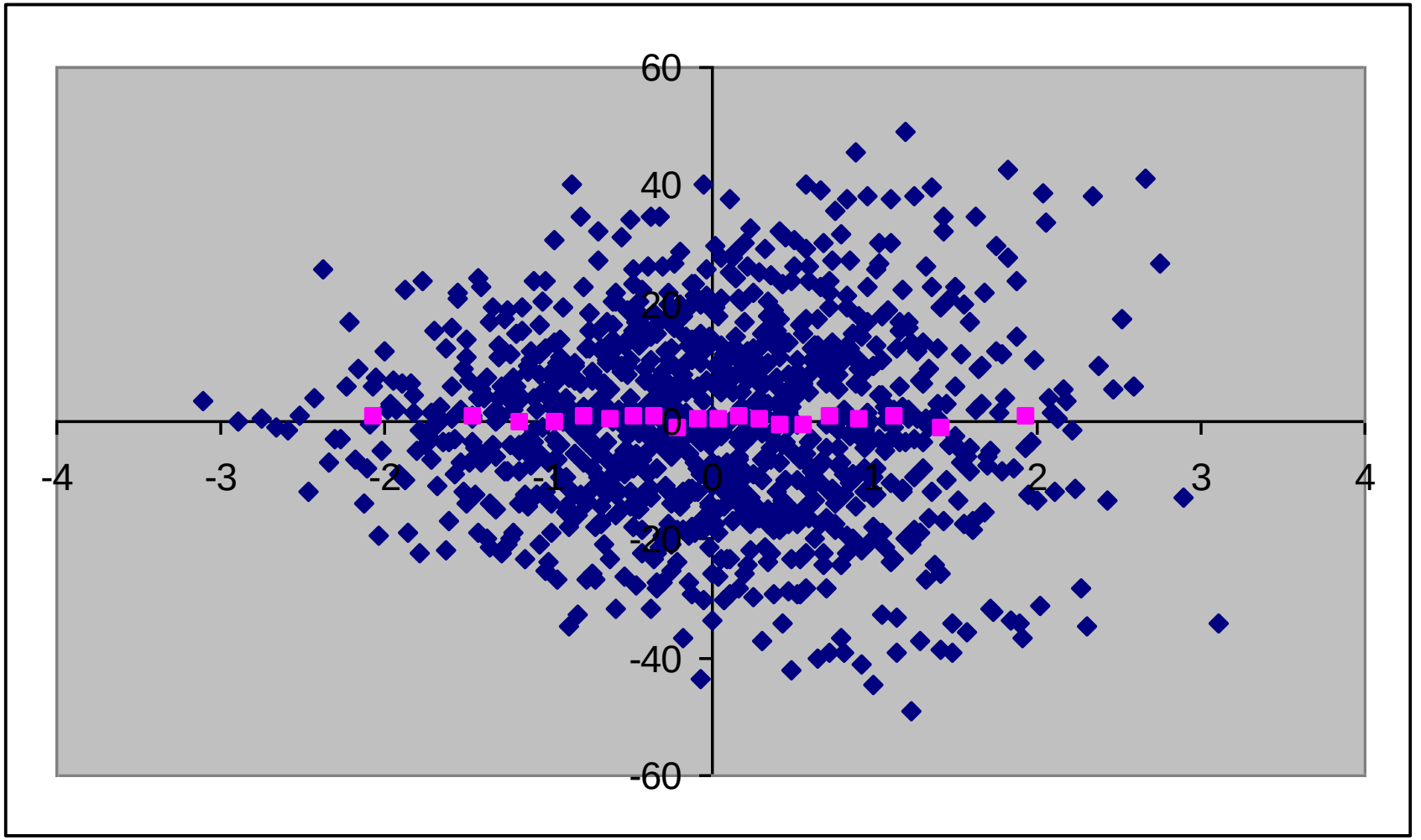
$$S_i \equiv \eta_i^2 := \frac{\mathrm{V}_{x_i}\left(\mathrm{E}_{\mathbf{x}_{\sim i}}\left(y \mid x_i\right)\right)}{\mathrm{V}\left(y\right)}$$

First order sensitivity index

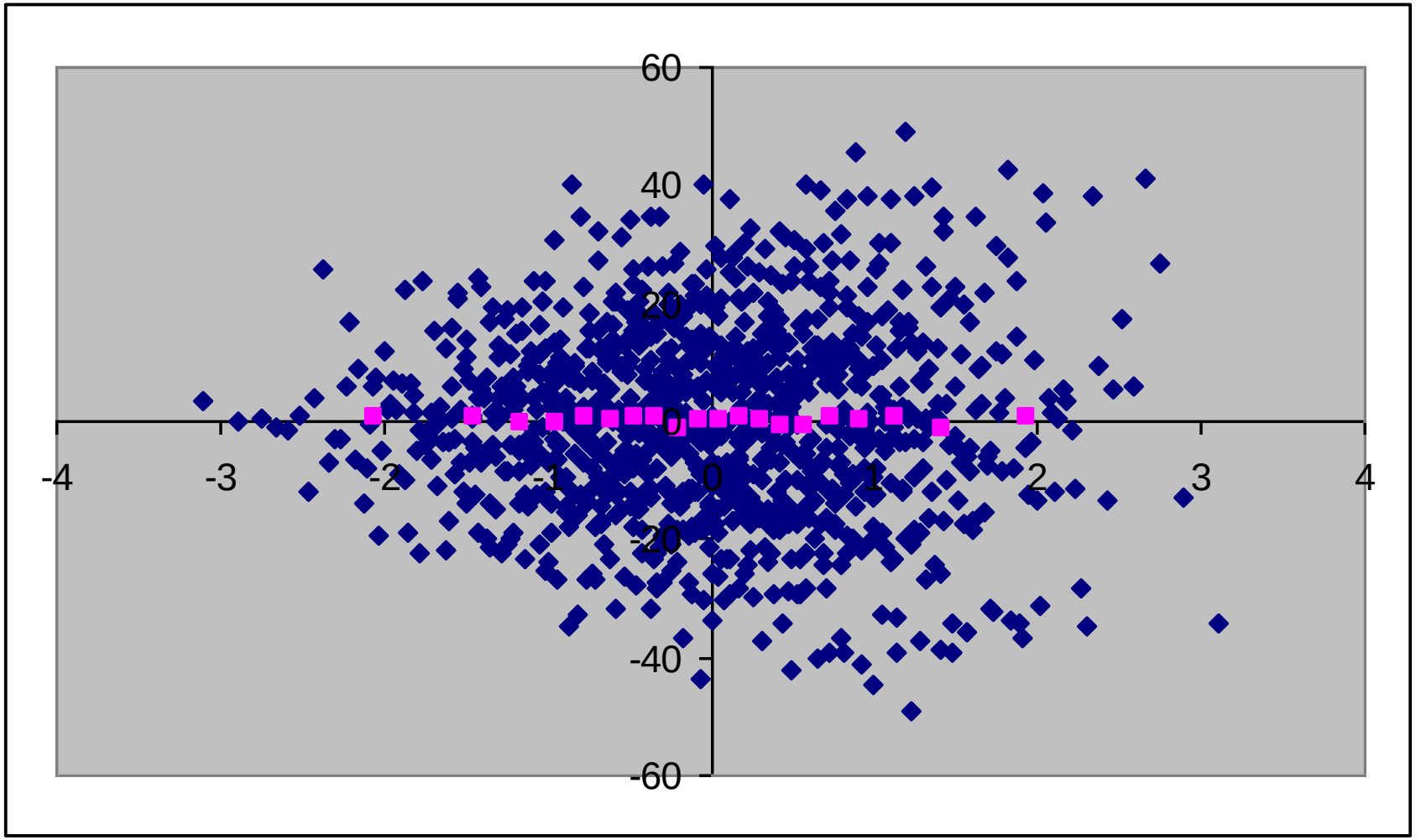
Unconditional variance

Non additive models

Is $S_i = 0$?



Is this factor non-important?



There are terms which capture two-way, three way, … interactions among variables

All these terms are linked by a formula

Variance decomposition (ANOVA)

$$V(Y) =$$

$$\sum_{i} V_i + \sum_{i,\, j>i} V_{ij} + \ldots + V_{123..k}$$

Variance decomposition (ANOVA)

The total variance can be decomposed into main effects and interaction effects up to the order k, the dimensionality of the problem (independent factors)

If fact interactions terms are awkward to handle: **just** the **second** order terms for a model with k factors are as many as k(k-1)/2 …

(10 factors=45 second order terms)

How about a single ‘importance’ terms for all effects?

In fact such terms exist and can be computed easily, without knowledge of the individual interaction terms

Thus given a model $f(X_1, X_2, \ldots, X_3)$

Where the variance decomposition would read $1 = S_1 + S_2 + S_3 + S_{12} + S_{13} + S_{23} + S_{123}$

We compute

$$T_1 = S_1 + S_{12} + S_{13} + S_{123}$$

$$T_2 = S_2 + S_{12} + S_{23} + S_{123}$$

$$T_3 = S_3 + S_{13} + S_{23} + S_{123}$$

The measures and their ‘settings’ = when to use them

Journal of the American Statistical Association ›
Volume 97, 2002 - Issue 459

Enter keywords, authors, DOI, C

Submit an article | Journal homepage

979
Views
286
CrossRef citations to date
6
Altmetric

Primary Article

On the Relative Importance of Input Factors in Mathematical Models

Safety Assessment for Nuclear Waste Disposal

Andrea Saltelli & Stefano Tarantola

Pages 702-709 | Published online: 31 Dec 2011

Download citation https://doi.org/10.1198/016214502388618447

The measures and their ‘settings’ = when to use them

First order effect	Factor prioritization (orienting research)
Total effect	Factor fixing (model simplification)

ELSEVIER

Computer Physics Communications

Volume 145, Issue 2, 15 May 2002, Pages 280-297

Making best use of model evaluations to compute sensitivity indices

Andrea Saltelli ✉ ⊕

Computing the indices efficiently

Higher order Sobol' indices

Get access >

Art B. Owen ✉, Josef Dick, Su Chen

Information and Inference: A Journal of the IMA, Volume 3, Issue 1, March 2014, Pages 59–81, https://doi.org/10.1093/imaiai/iau001

Published: 01 March 2014 **Article history** ▾

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)

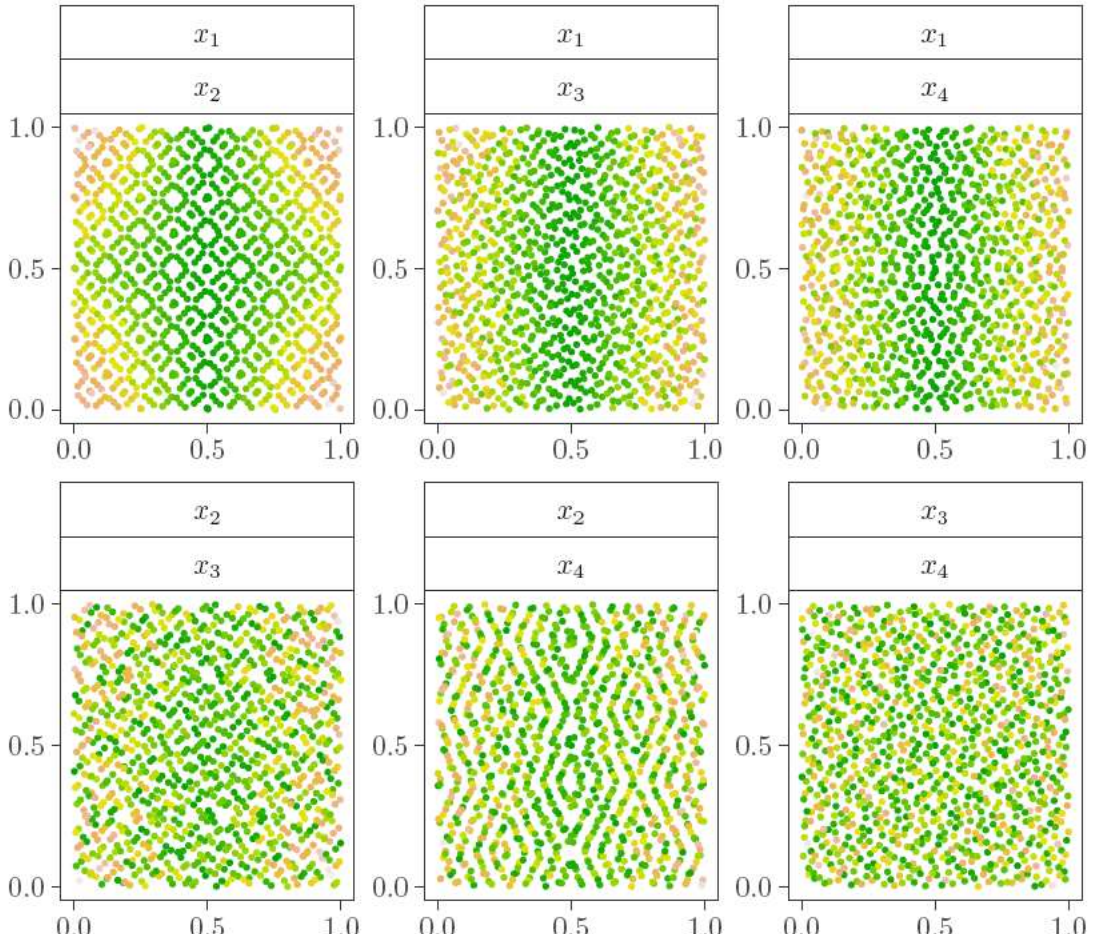
SALib https://salib.readthedocs.io/en/latest/

Journal of Statistical Software

Home / Archives / Vol. 102 (2022) / Issue 5

sensobol: An R Package to Compute Variance-Based Sensitivity Indices

Arnald Puy, Samuele Lo Piano, Andrea Saltelli, Simon A. Levin



Model's effective dimension

ACADEMIC PRESS

Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT®

Journal of Complexity 19 (2003) 101–124

Journal of COMPLEXITY

http://www.elsevier.com/locate/jco

The effective dimension and quasi-Monte Carlo integration ☆

Xiaoqun Wang[a,b,*] and Kai-Tai Fang[c]

[a] *Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China*
[b] *School of Mathematics, University of New South Wales, Sydney 2052, Australia*
[c] *Department of Mathematics, Hong Kong Baptist University, Hong Kong, China*

Received 12 February 2002; accepted 6 November 2002

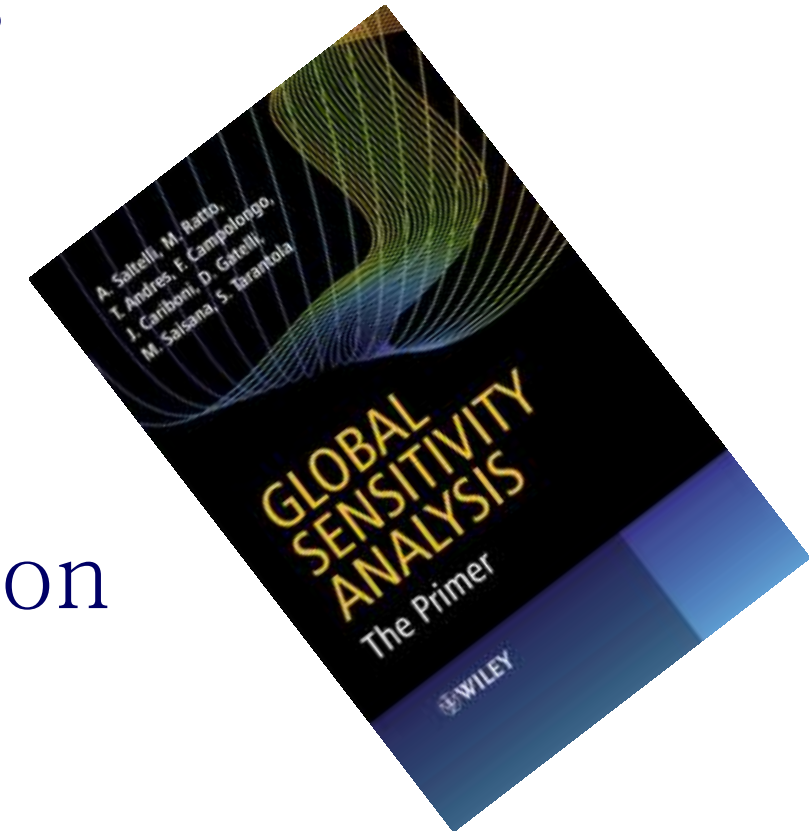
The difficulty of a function/model is not in its number of dimensions but in the number of effective dimensions, either in the **truncation** or **superposition** sense

truncation sense = how many factors are important?
superposition sense=how high is the highest interaction?

Why using variance-based sensitivity analysis methods

Advantages with variance based methods:

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



Chapter 1 its exercises

… anyone developing a
new method tests it against S_i, T_i

AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2015WR017558

Companion to
Razavi and Gupta [2016],
doi:10.1002/2015WR017559.

Key Points:
- The VARS framework enables

A new framework for comprehensive, robust, and efficient global sensitivity analysis: 1. Theory

Saman Razavi[1,2] and Hoshin V. Gupta[3]

[1]Global Institute for Water Security & School of Environment and Sustainability, University of Saskatchewan, Saskatoon, Saskatchewan, Canada, [2]Department of Civil and Geological Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada, [3]Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA

S_i, T_i can be used to do a sensitivity analysis of a sensitivity analysis…

ELSEVIER

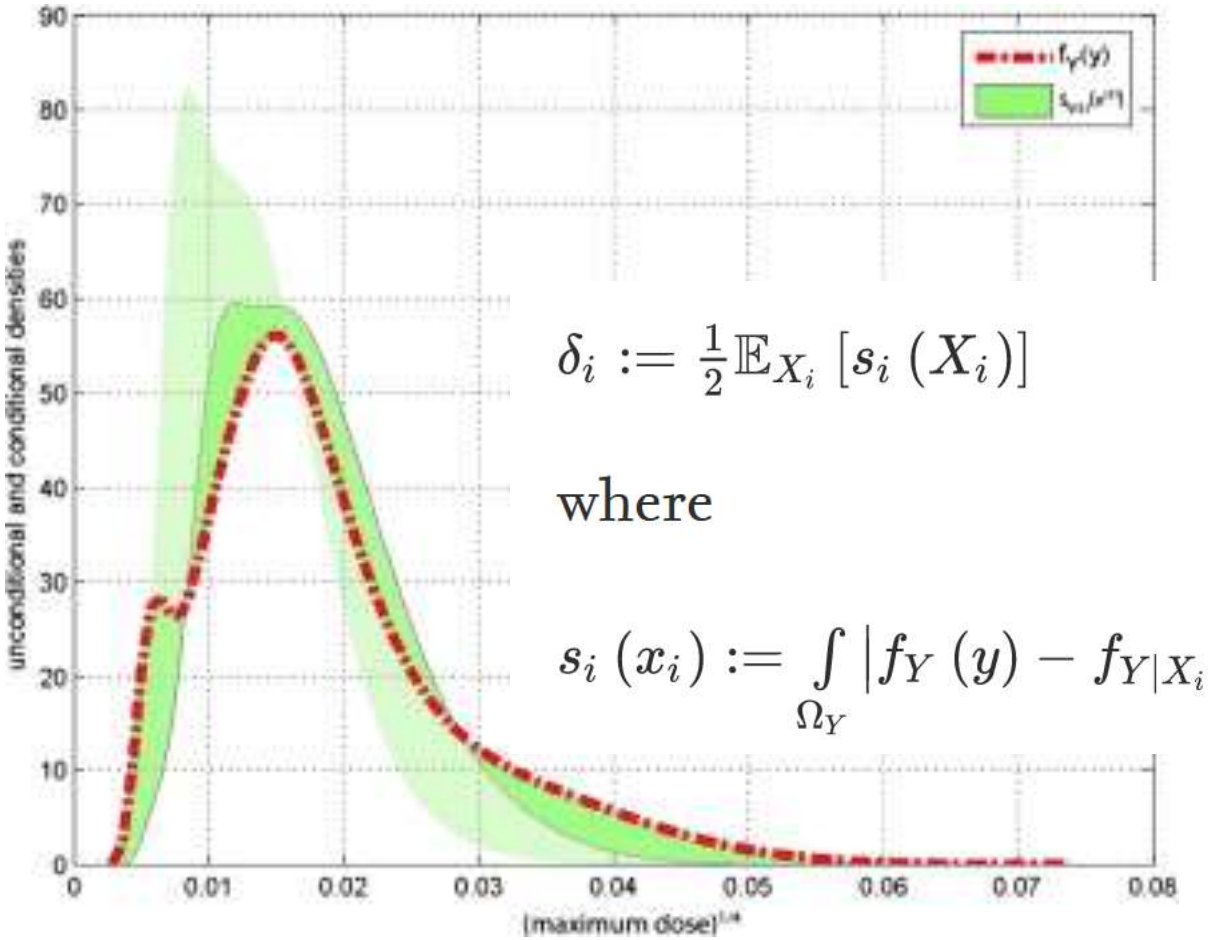
Environmental Modelling & Software

Volume 137, March 2021, 104960

Is VARS more intuitive and efficient than Sobol' indices?

Arnald Puy [a, b], Samuele Lo Piano [c], Andrea Saltelli [d]

… 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

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

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

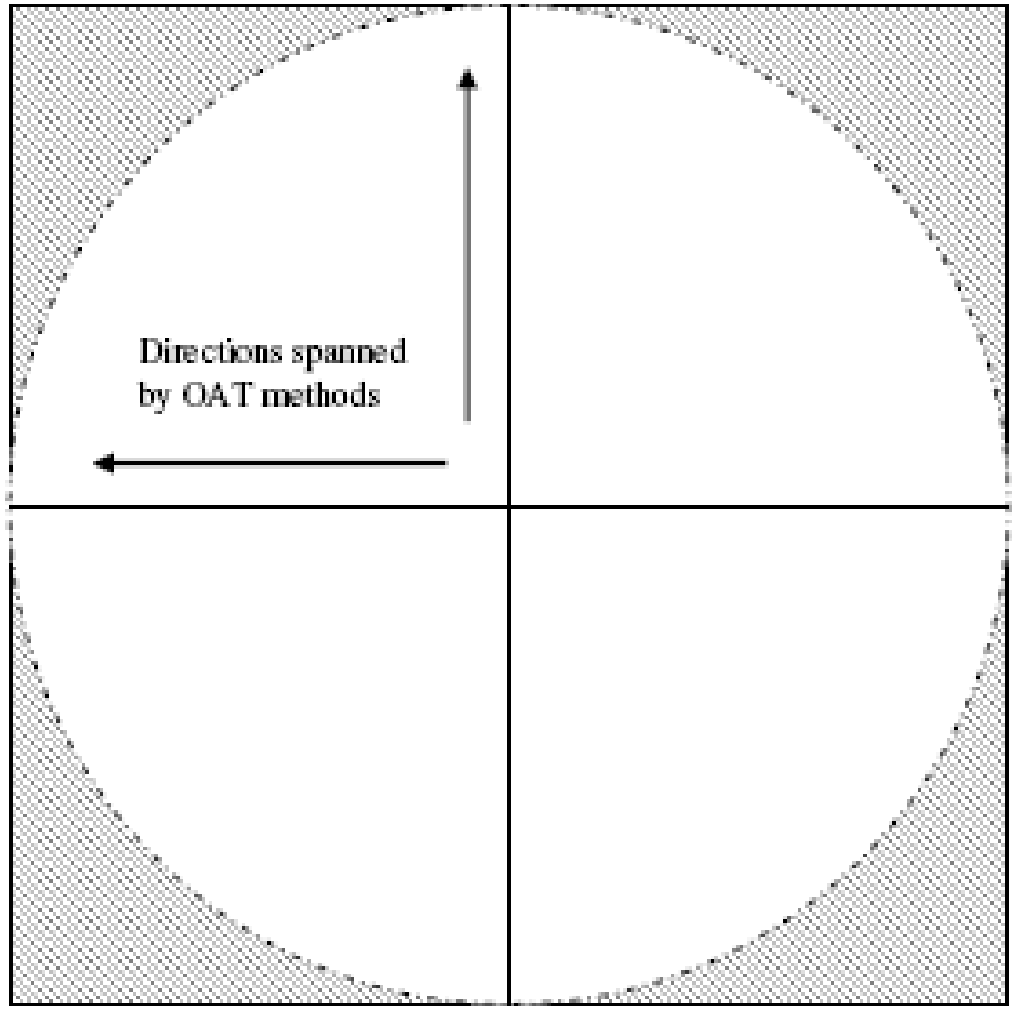
journal homepage: www.elsevier.com/locate/envsoft

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

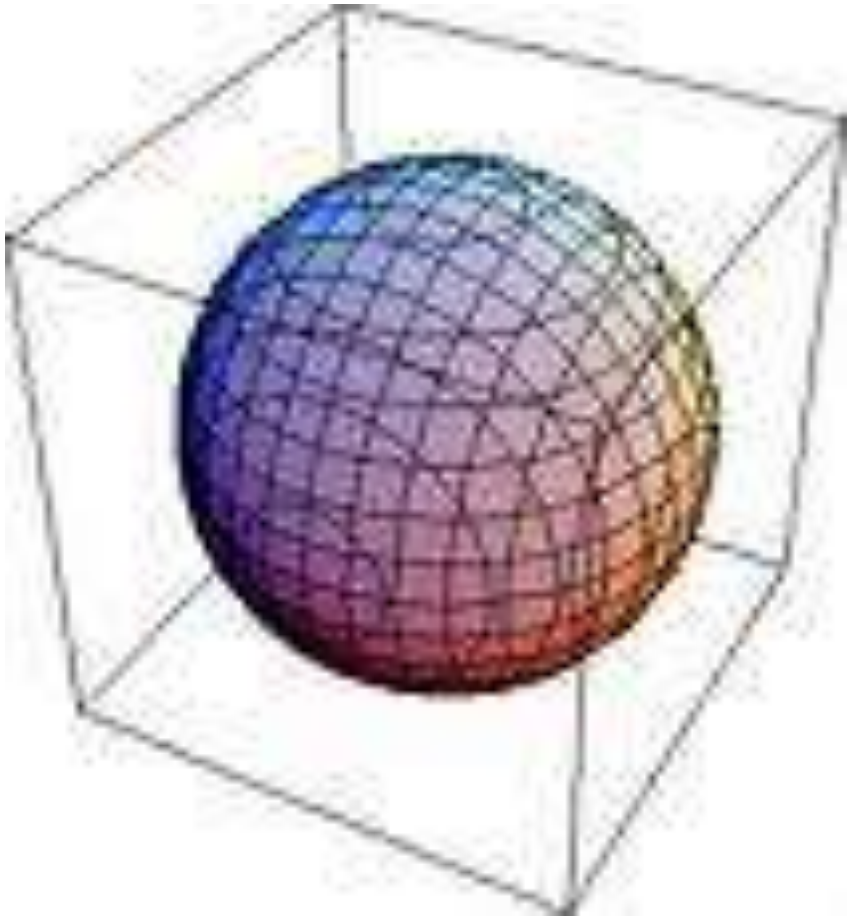
OAT in 2 dimensions



Area circle / area square =?

~ 3/4

OAT in 3 dimensions

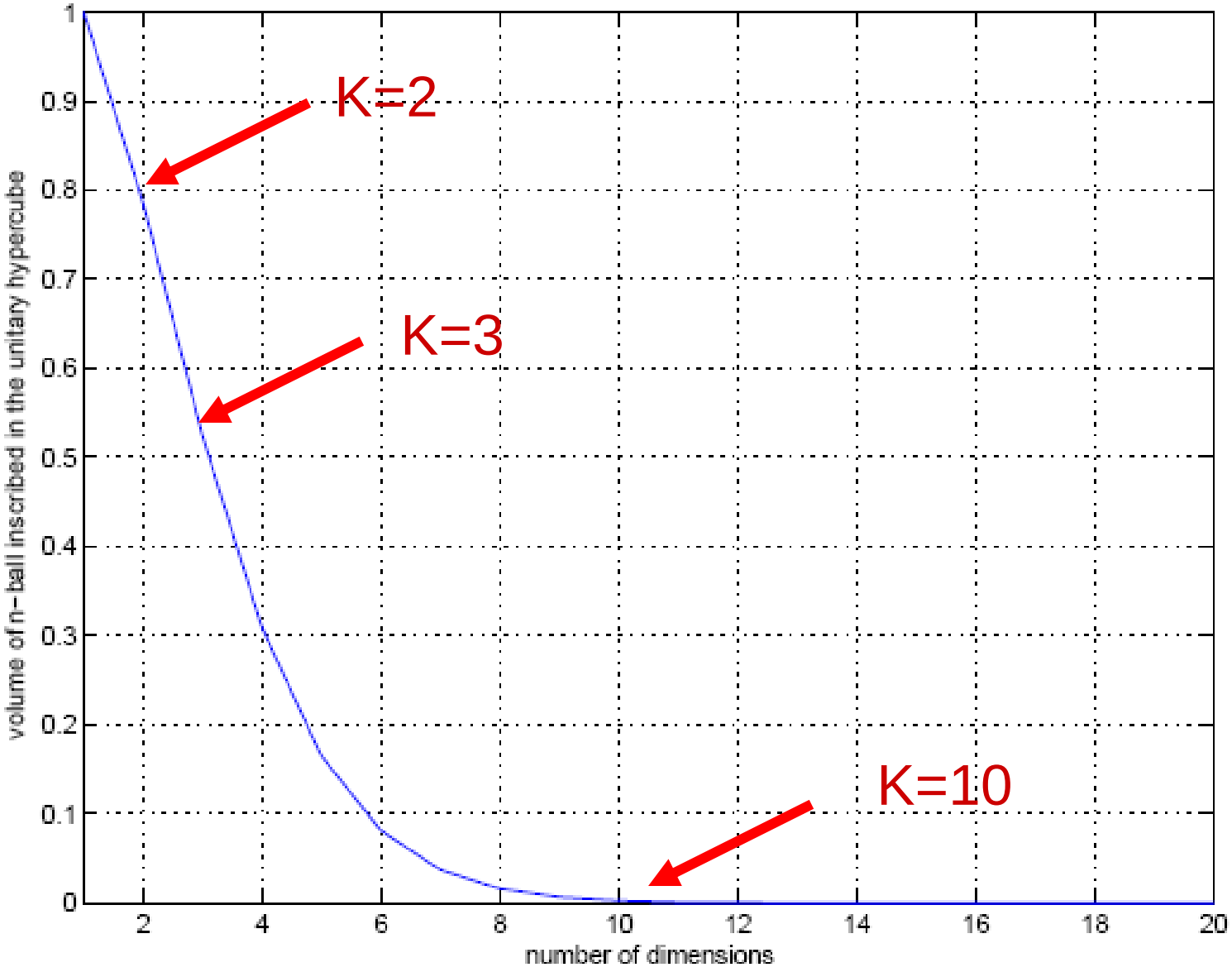


Volume sphere / volume cube =?

~ 1/2

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

OAT in k dimensions

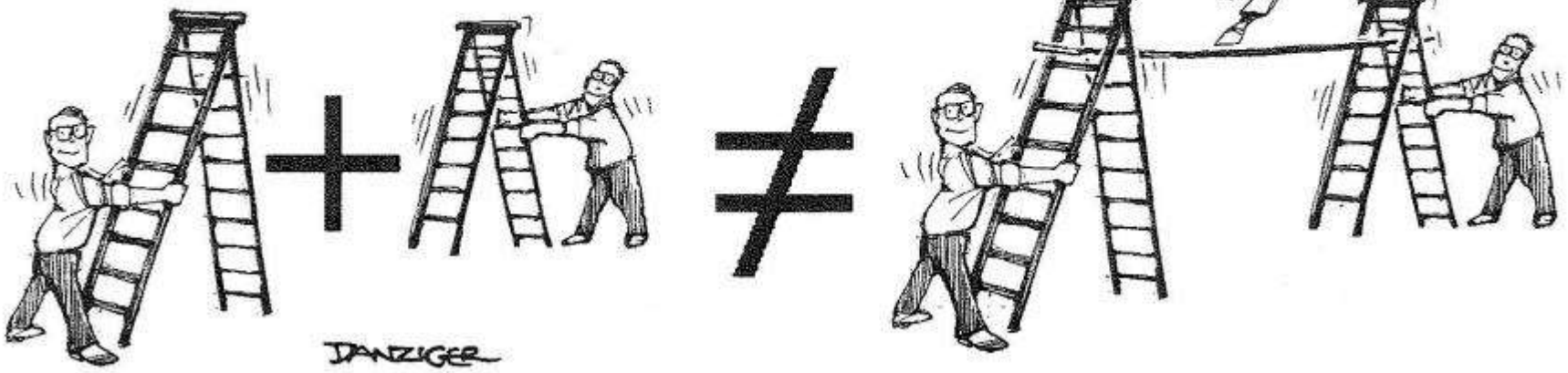


OAT does not capture interactions

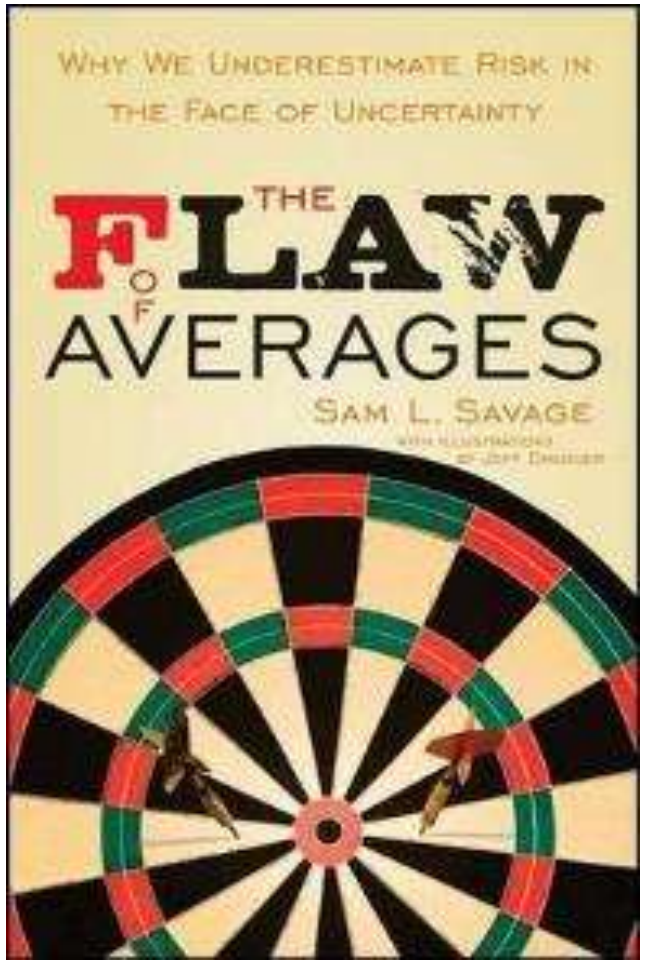
➔ The resulting analysis is non conservative

How would you test the scaffolding?

How coupled ladders are shaken in most of available literature



How to shake coupled ladders



Environmental Modelling & Software

Volume 114, April 2019, Pages 29-39

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

Andrea Saltelli [a, b], Ksenia Aleksankina [c], William Becker [d], Pamela Fennell [e], Federico Ferretti [d], Niels Holst [f], Sushan Li [g], Qiongli Wu [h]

Don’t use method that are not model-independent (such as PCC, PRCC)

Use model-free methods

Why not using correlation-regression based techniques? PCC, PRCC, SRC, SRRC

ELSEVIER

Reliability Engineering & System Safety

Volume 28, Issue 2, 1990, Pages 229-253

Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques

A. Saltelli, J. Marivoet

ELSEVIER

Computational Statistics & Data Analysis

Volume 13, Issue 1, January 1992, Pages 73-94

Sensitivity analysis for model output: Performance of black box techniques on three international benchmark exercises

A. Saltelli, T. Homma

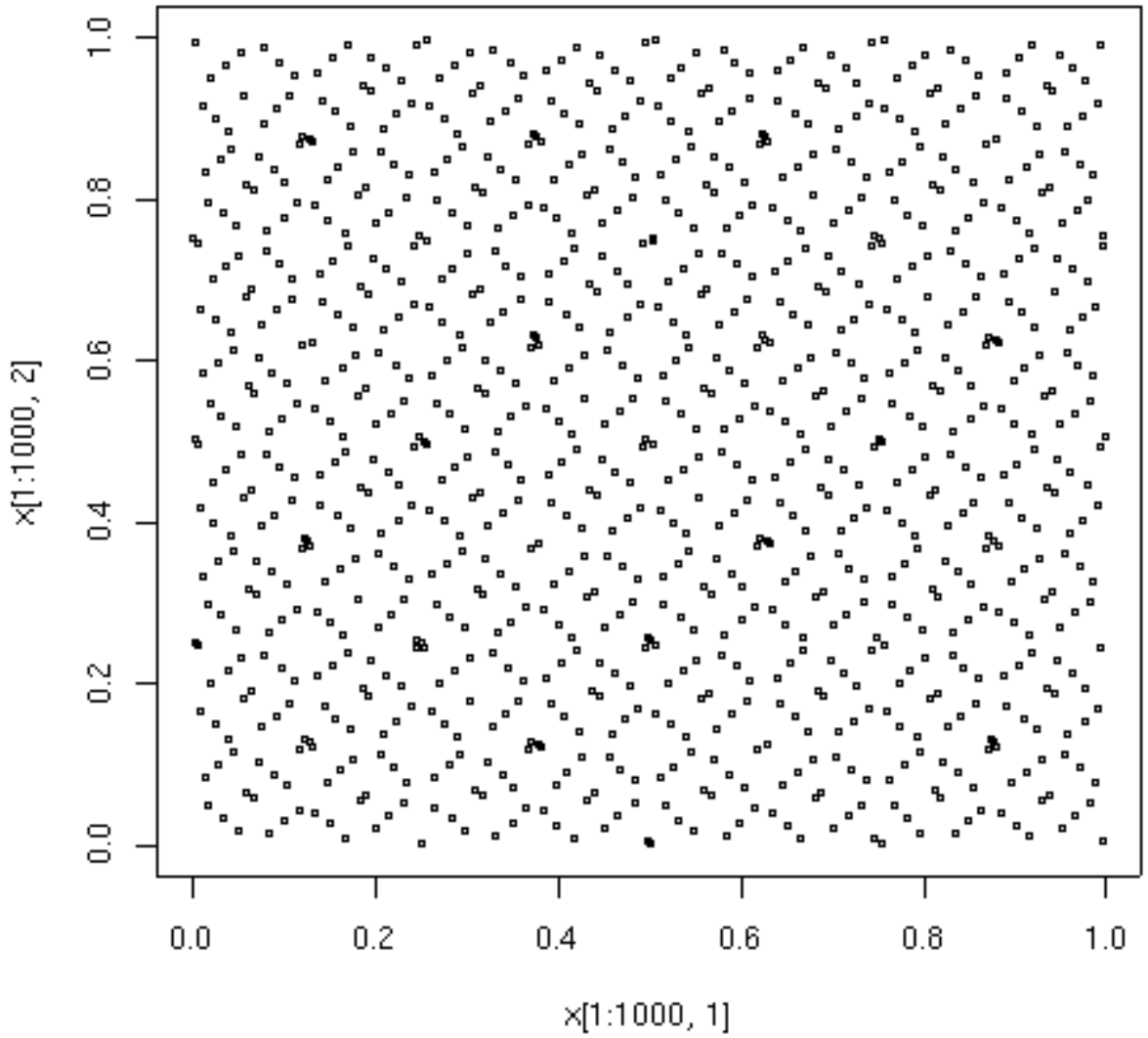
➔They assume linearity (PCC) or monotonicity (PRCC), which is difficult to know *ex-ante*

Don’t use either LHS or optimized LHS

Quasi-random sequences perform better

Quasi random sequences

Ilya M. Sobol’



Cornell University

the

arXiv.org > stat > arXiv:1505.02350

Search...

Help | Adva

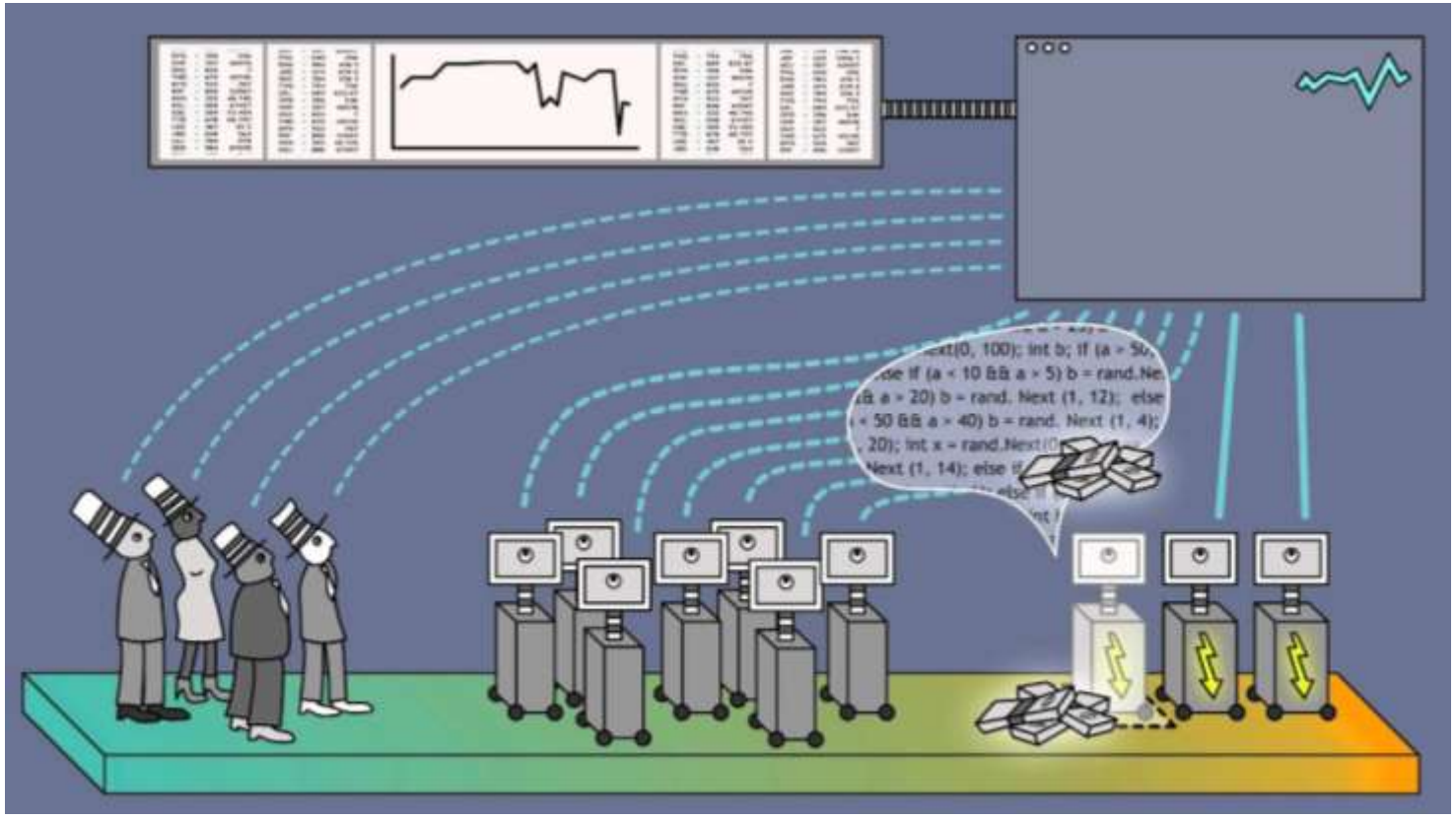
Statistics > Applications

[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=z4nCTdQlH8w

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!

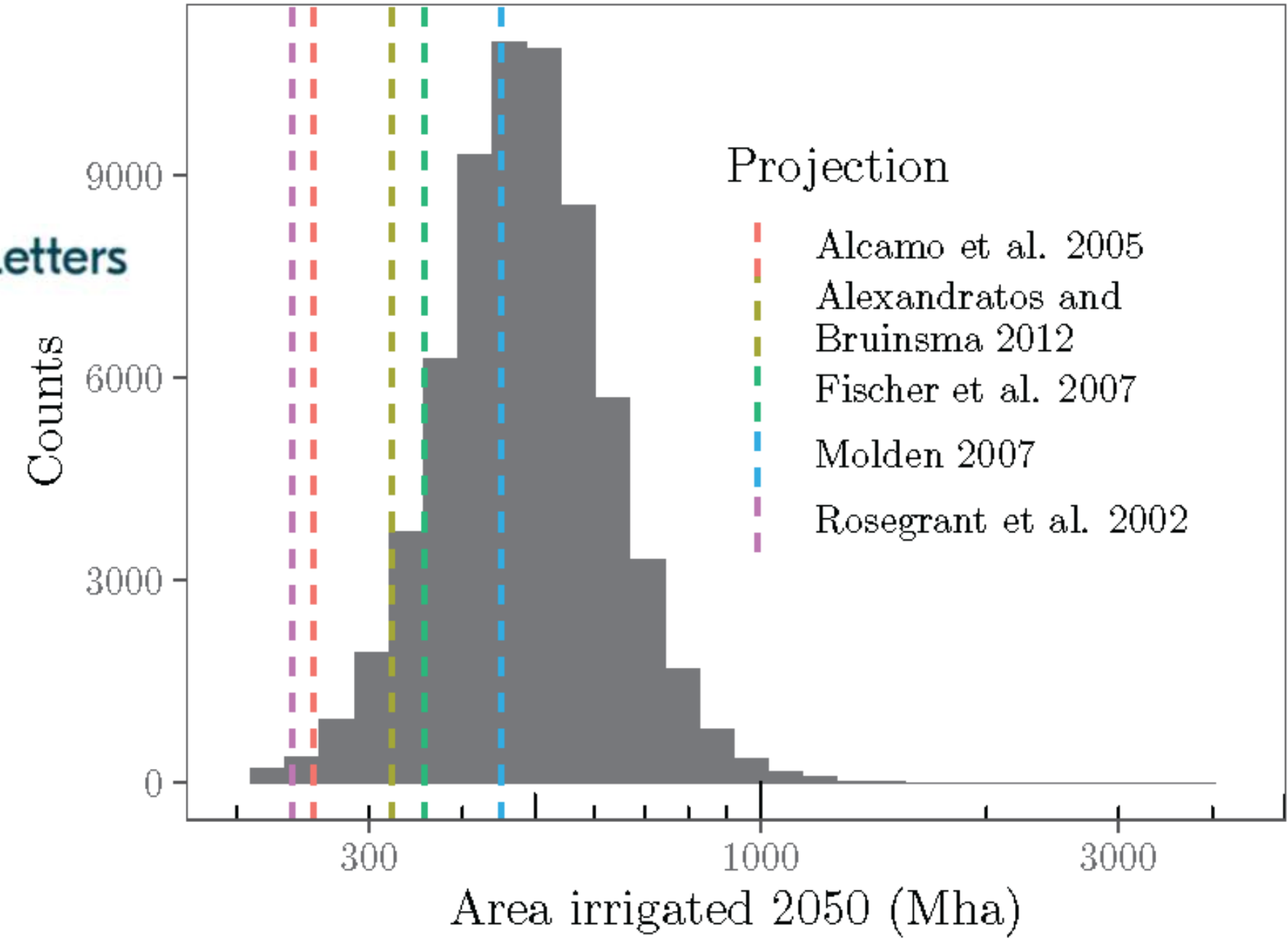
Model routinely used to produce point estimates may becomes non conservative when the uncertainty is plugged in

Current Models Underestimate Future Irrigated Areas

Geophysical Research Letters

A. Puy ✉, S. Lo Piano, A. Saltelli

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



nature communications

Comment | Open Access | Published: 08 June 2022

The delusive accuracy of global irrigation water withdrawal estimates

Arnald Puy ✉, Razi Sheikholeslami, Hoshin V. Gupta, Jim W. Hall, Bruce Lankford, Samuele Lo Piano, Jonas Meier, Florian Pappenberger, Amilcare Porporato, Giulia Vico & Andrea Saltelli

Nature Communications **13**, Article number: 3183 (2022) | Cite this article

3022 Accesses | **102** Altmetric | Metrics

Miscalculating the volumes of water withdrawn for irrigation, the largest consumer of freshwater in the world, jeopardizes sustainable water management. Hydrological models quantify water withdrawals, but their estimates are unduly precise. Model imperfections need to be appreciated to avoid policy misjudgements.

Sustainable Development Goals (SDGs), from Zero Hunger (SDG 2) to Water Stress (SDG 6), will be poorly assessed if irrigation water withdrawal convey an illusion of accuracy

nature communications

Explore content | About the journal | Publish with us

nature > nature communications > comment > article

Comment | Open Access | Published: 08 June 2022

The delusive accuracy of global irrigation water withdrawal estimates

Arnald Puy, Razi Sheikholeslami, Hoshin V. Gupta, Jim W. Hall, Bruce Lankford, Samuele Lo Piano, Jonas Meier, Florian Pappenberger, Amilcare Porporato, Giulia Vico & Andrea Saltelli

Don’t use Morris’ method

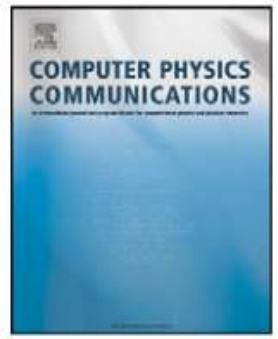
More cumbersome and fragile than the total sensitivity index that is its close equivalent

Computer Physics Communications 182 (2011) 978–988

Contents lists available at ScienceDirect

Computer Physics Communications

www.elsevier.com/locate/cpc



From screening to quantitative sensitivity analysis. A unified approach

Francesca Campolongo *, Andrea Saltelli, Jessica Cariboni

Joint Research Centre, Institute for the Protection and Security of the Citizen, Ispra, Italy

Morris

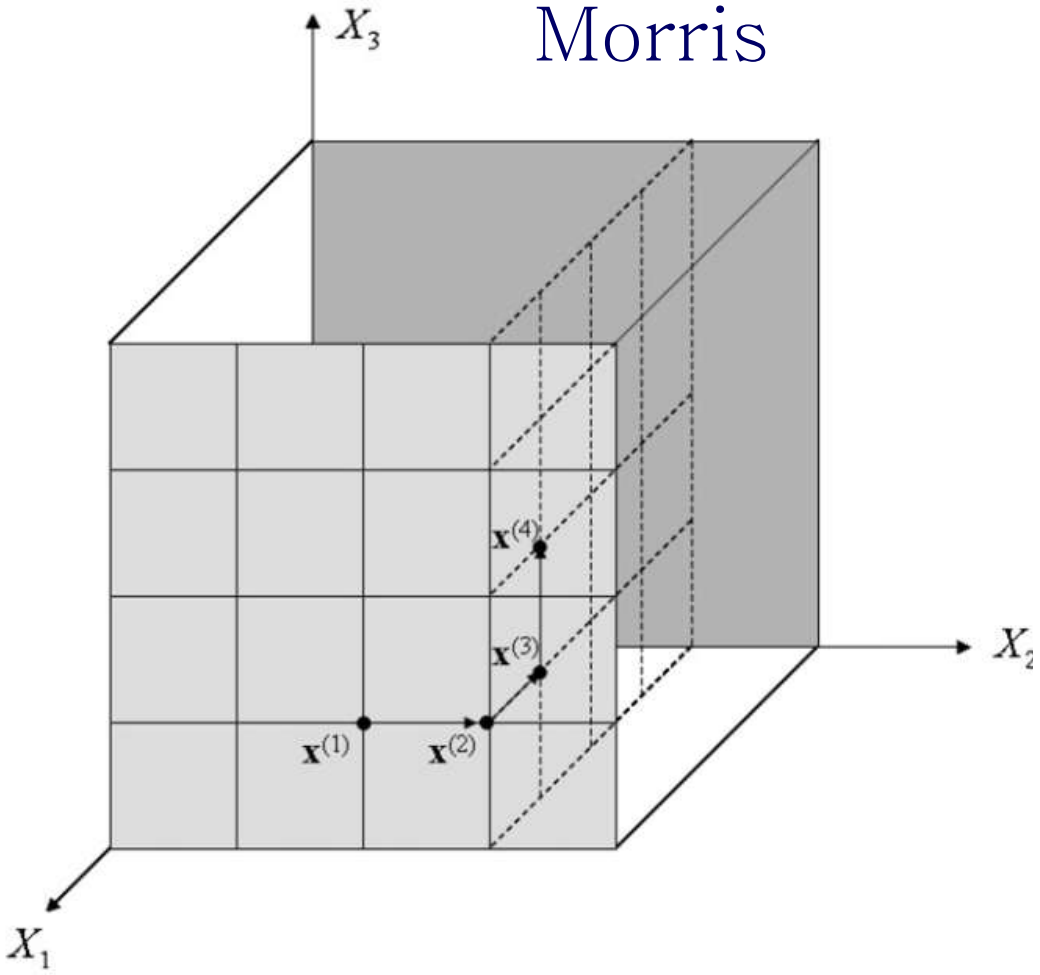


Fig. 1. Example of trajectory in 3 dimensions for the original EE method.

Total sensitivity index T_i

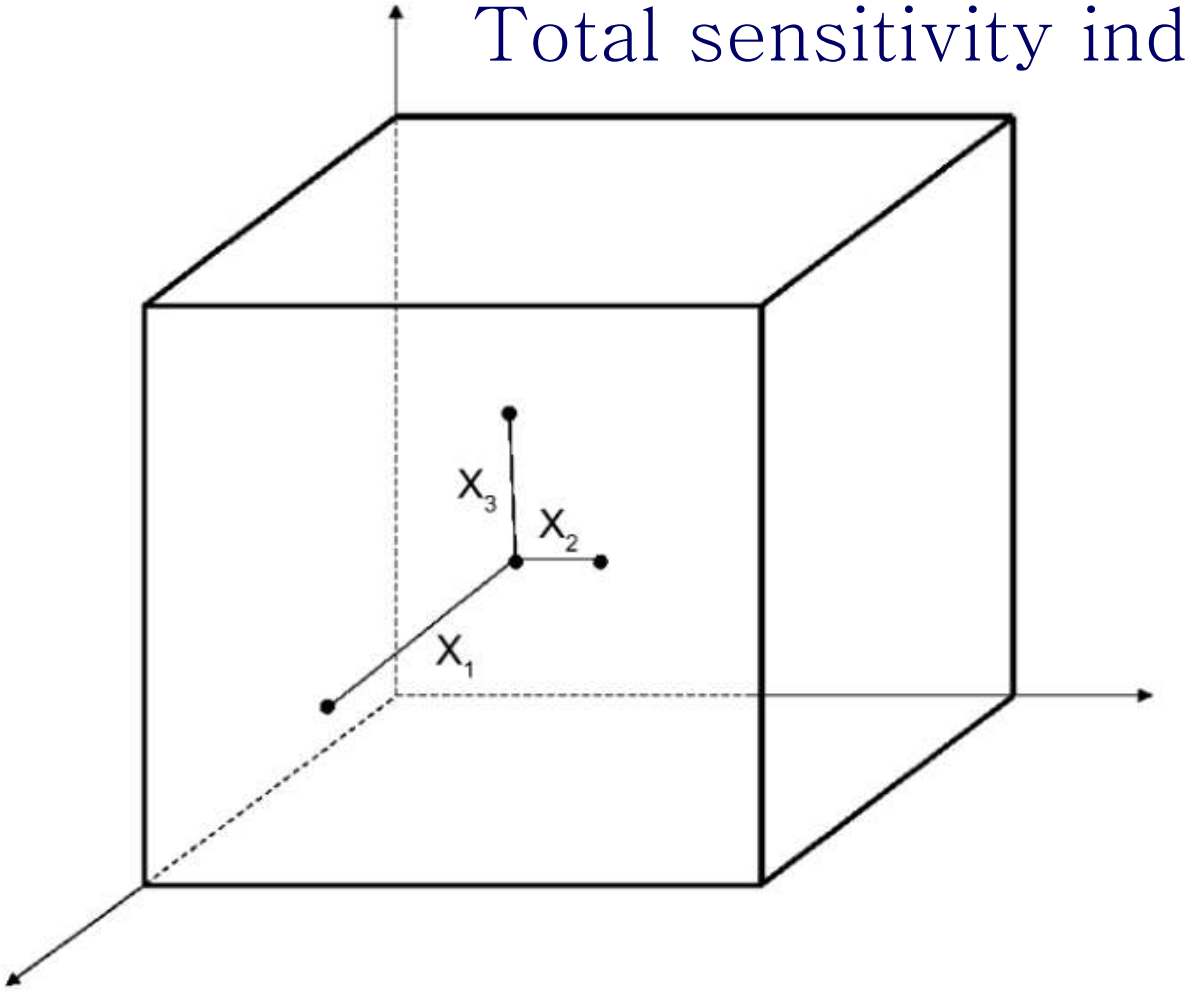
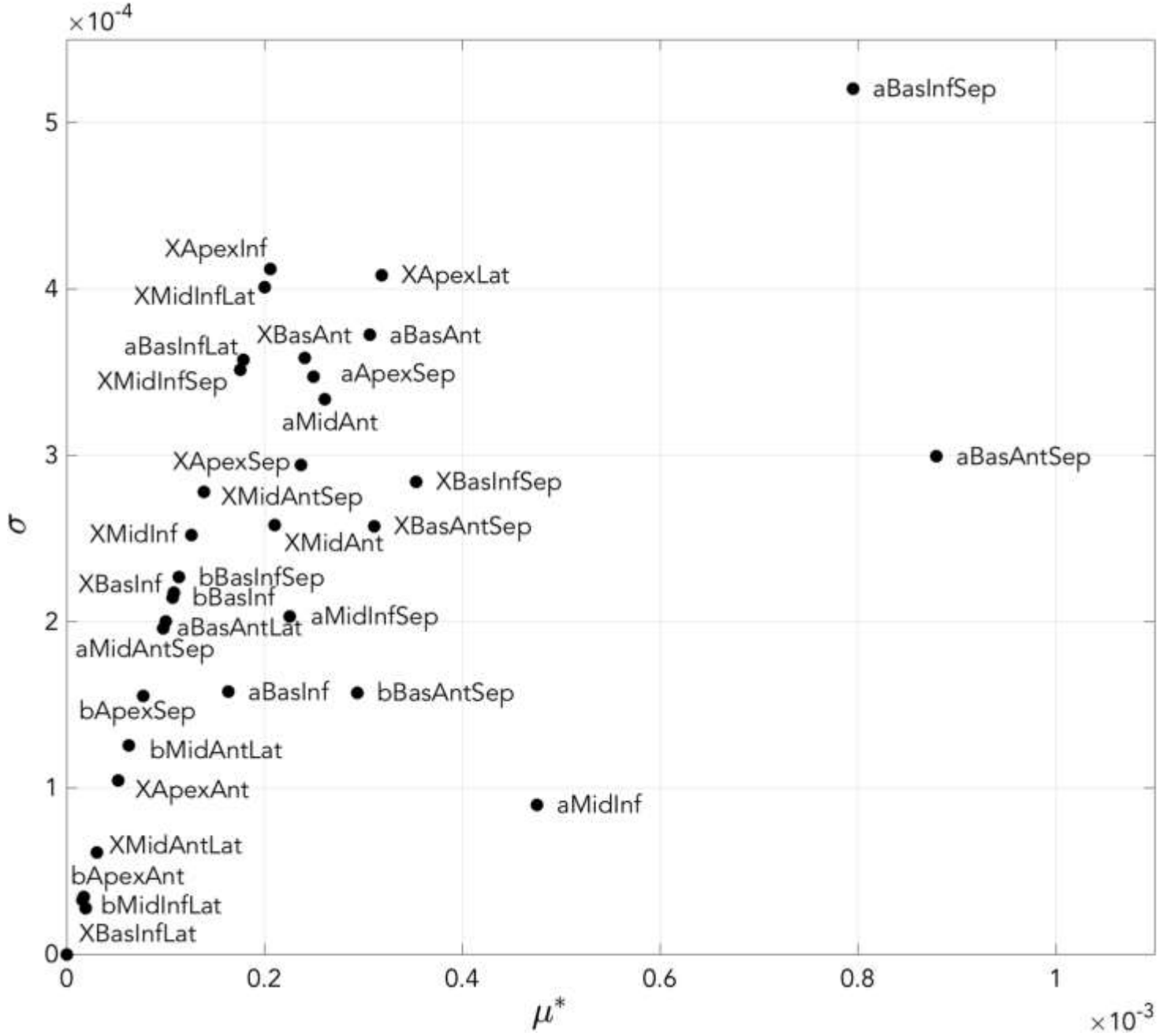


Fig. 3. Example of a radial sample in three dimensions.

Morris needs one more design parameter than T_i:
the space step for the grid

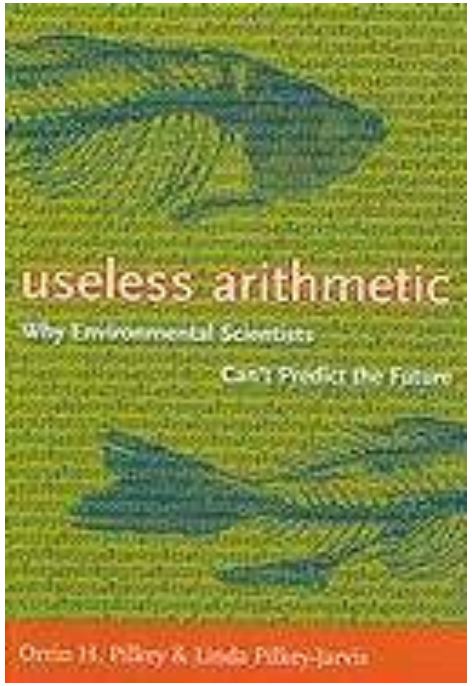
Morris is more cumbersome to interpret (two outputs: mu and sigma)



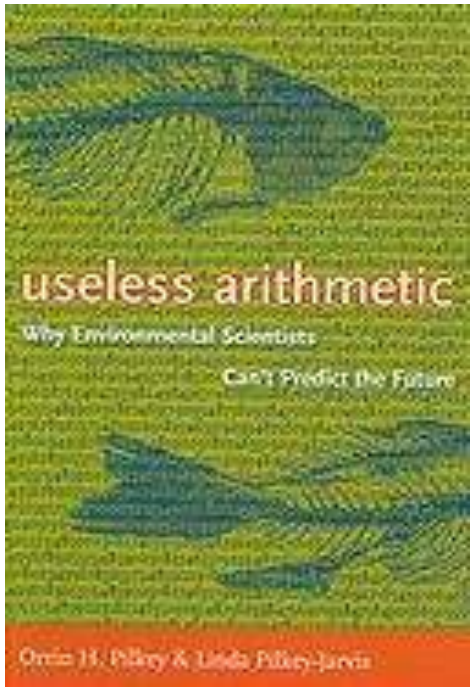
Don’t confuse the map with the territory

If you do, sensitivity analysis will not save you

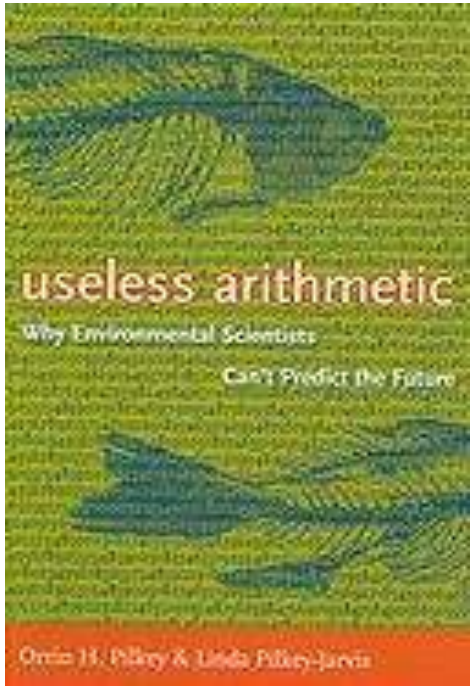
Orrin H.
Pilkey



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



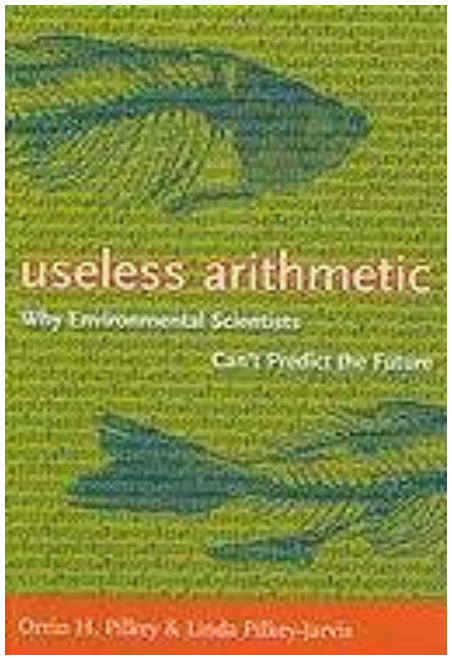
<<It is important, however, to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature>>

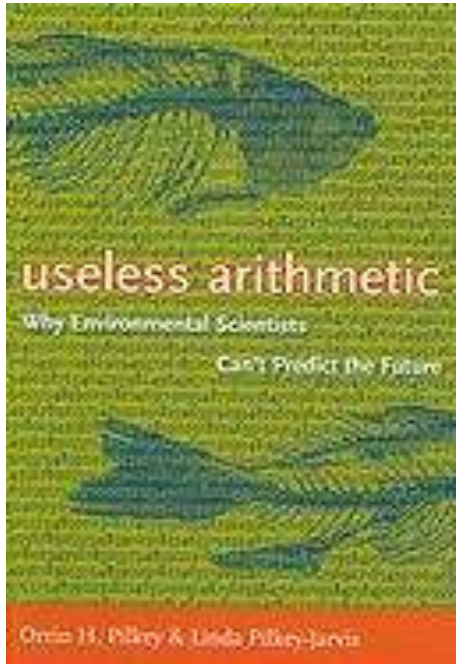


<<…If the model is wrong or if it is a poor representation of reality, determining the sensitivity of an individual parameter in the model is a meaningless pursuit>>

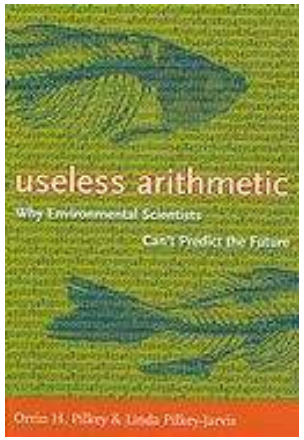
One of the examples discussed concerns the Yucca Mountain repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis.

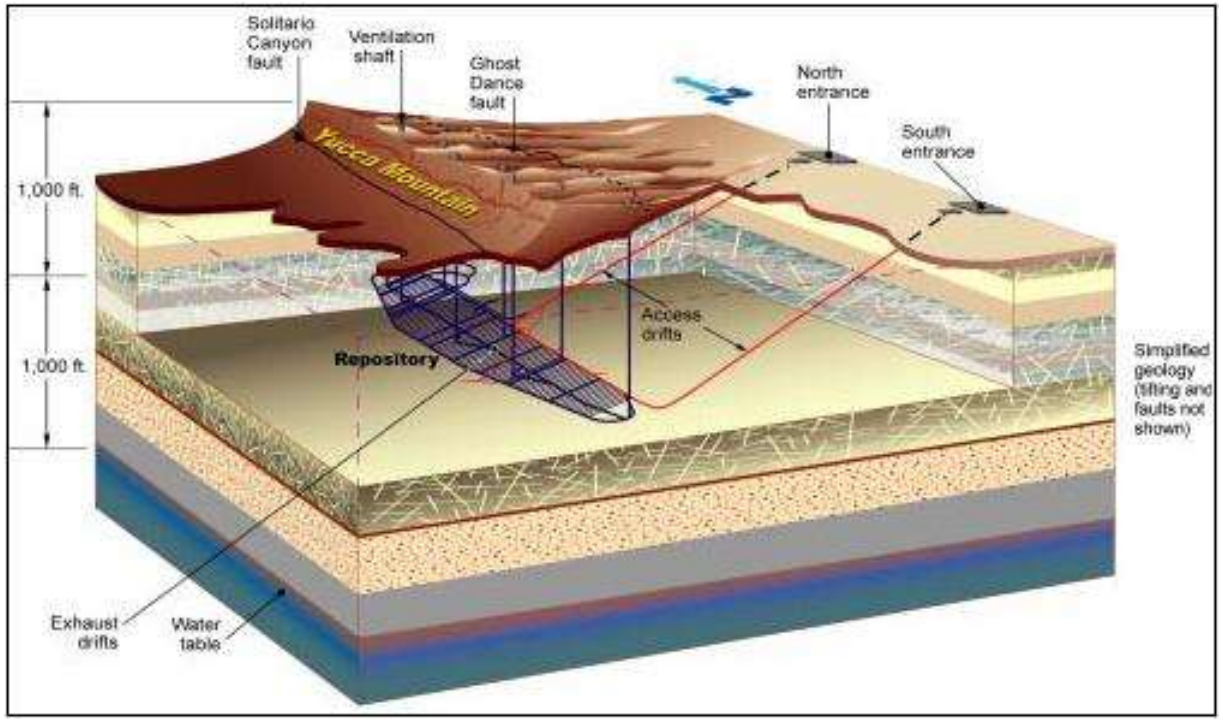
TSPA is Composed of 286 sub-models.





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.





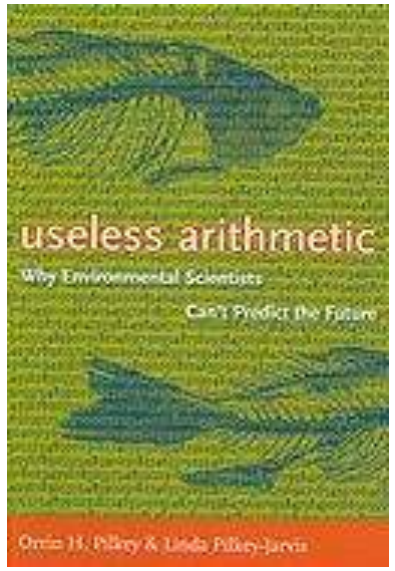
The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter

(the ^{36}Cl story)

Type III error in sensitivity: Examples:

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.

Steve Rayner

Rayner, S., 2012, Uncomfortable knowledge: the social construction of ignorance in science and environmental policy discourses, Economy and Society, 41:1, 107-125.

Rayner's (2012) strategies to deal with "uncomfortable knowledge".

Denial, Dismissal, Diversion, Displacement

Model based

Rayner, S., 2012, Uncomfortable knowledge: the social construction of ignorance in science and environmental policy discourses, Economy and Society, 41:1, 107-125.

Beware the dimension of your model

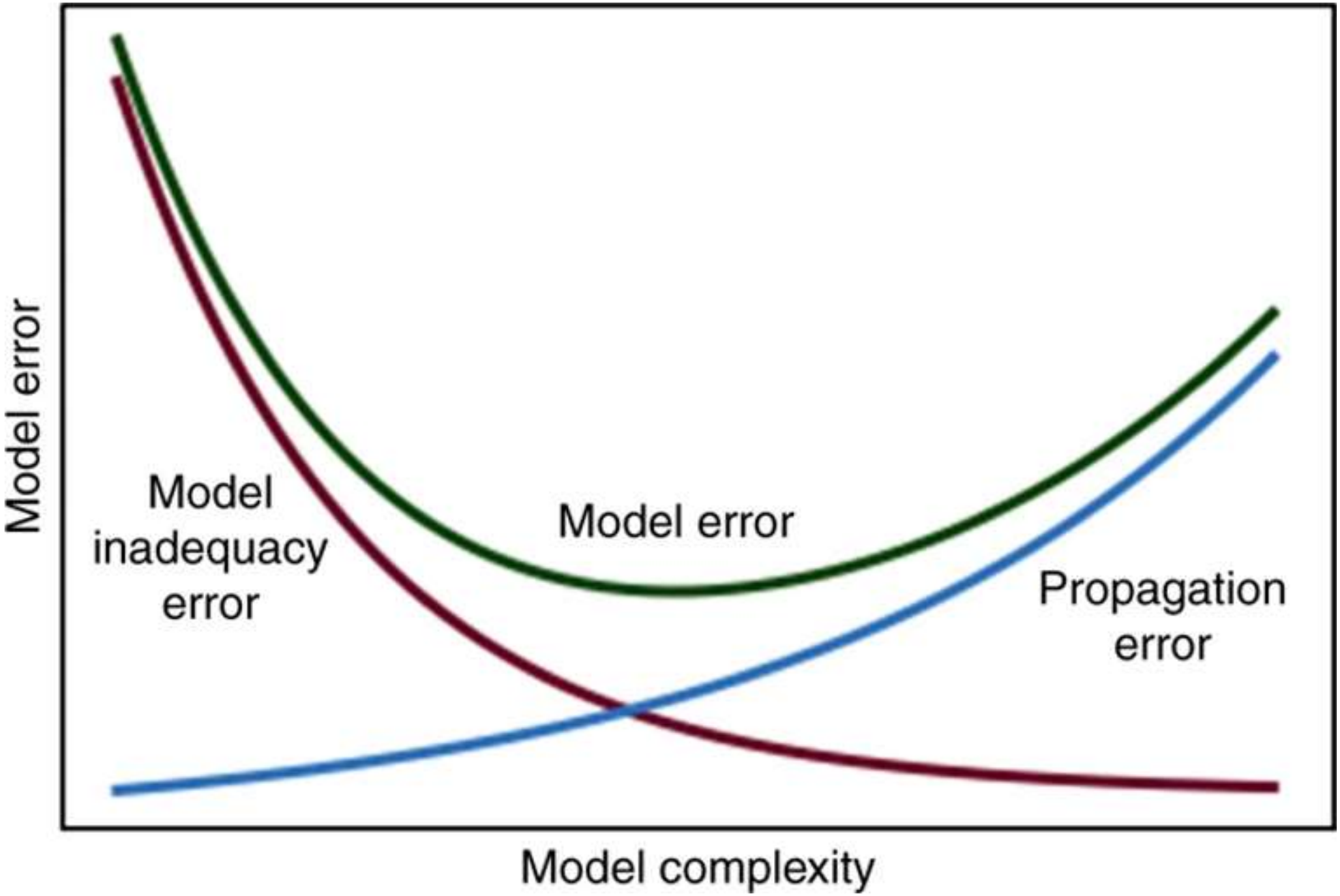
Mind the conjecture of O'Neil

nature COMMUNICATIONS

Comment | Open Access | Published: 27 August 2019

A short comment on statistical versus mathematical modelling

Andrea Saltelli


Model error
Model inadequacy error
Model error
Propagation error
Model complexity

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"

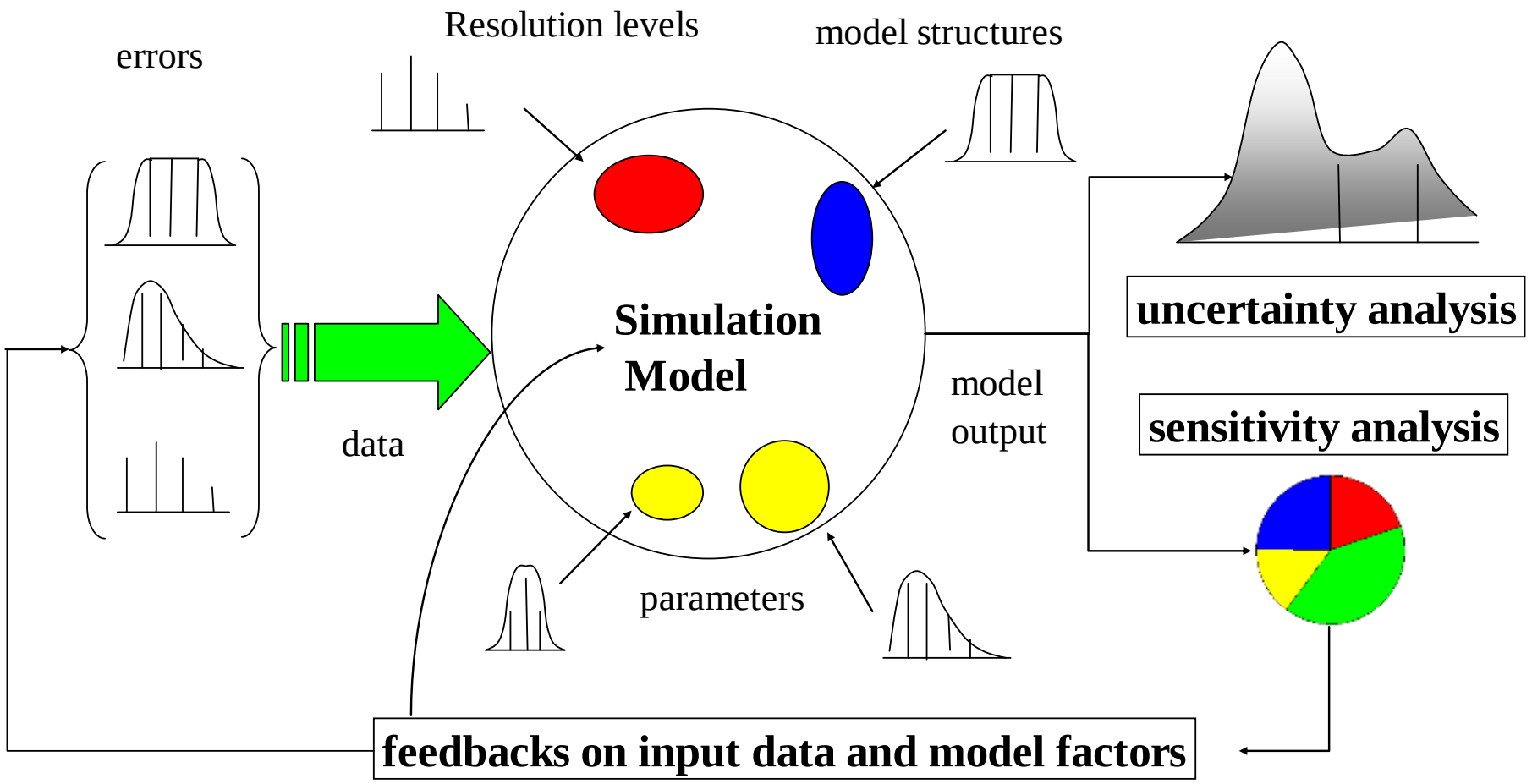
In M. G. Turner and R. H. Gardner, "Introduction to Models" in Landscape Ecology in Theory and Practice, New York, NY: Springer New York, 2015, pp. 63–95.

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.

Don’t sample just parameters and boundary conditions

Explore thoroughly the space of the assumptions

An engineer's vision of UA, SA

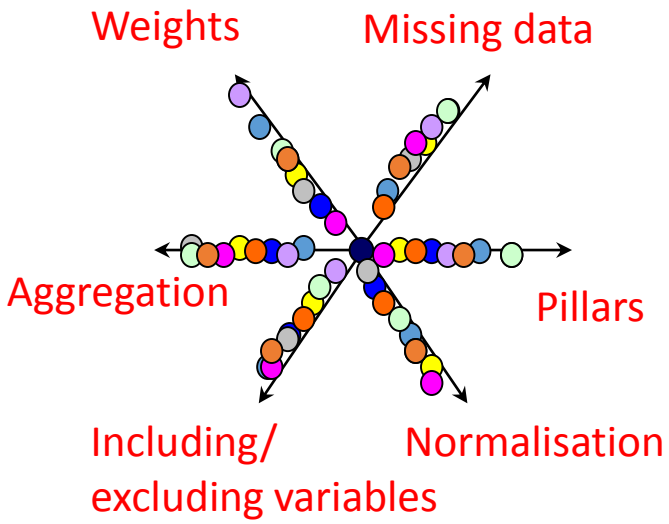


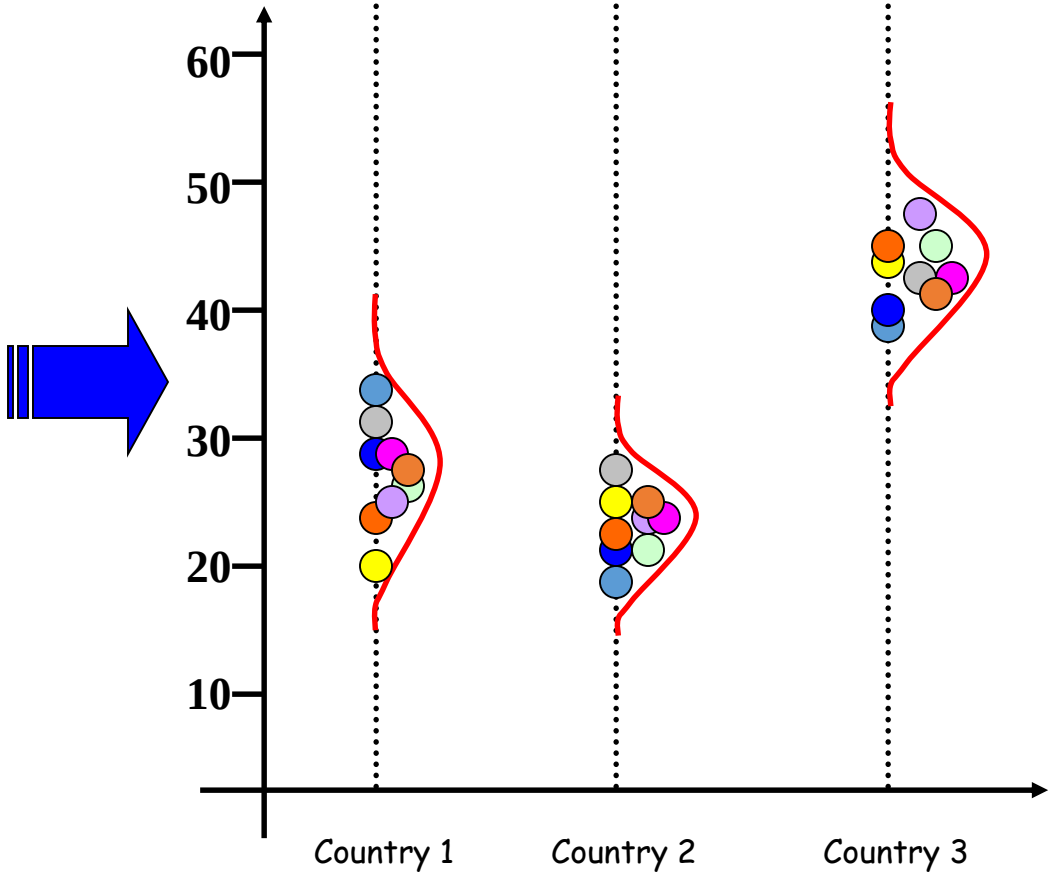
One can sample more than just factors:

- 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





Don’t go public with your results without having seen your SA

Find SA before SA finds you

Global Environmental Change 20 (2010) 298–302

Contents lists available at ScienceDirect

Global Environmental Change

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

Sensitivity analysis didn't help. A practitioner's critique of the Stern review

Andrea Saltelli *, Beatrice D'Hombres

Joint Research Centre, Institute for the Protection and Security of the Citizen, Ispra, Italy

The case of Stern's Review – Technical Annex to postscript

Nicholas Stern, London School of Economics

William Nordhaus, University of Yale Nobel 'Economics' 2018

Stern, N., Stern Review on the Economics of Climate Change. UK Government Economic Service, London, www.sternreview.org.uk.

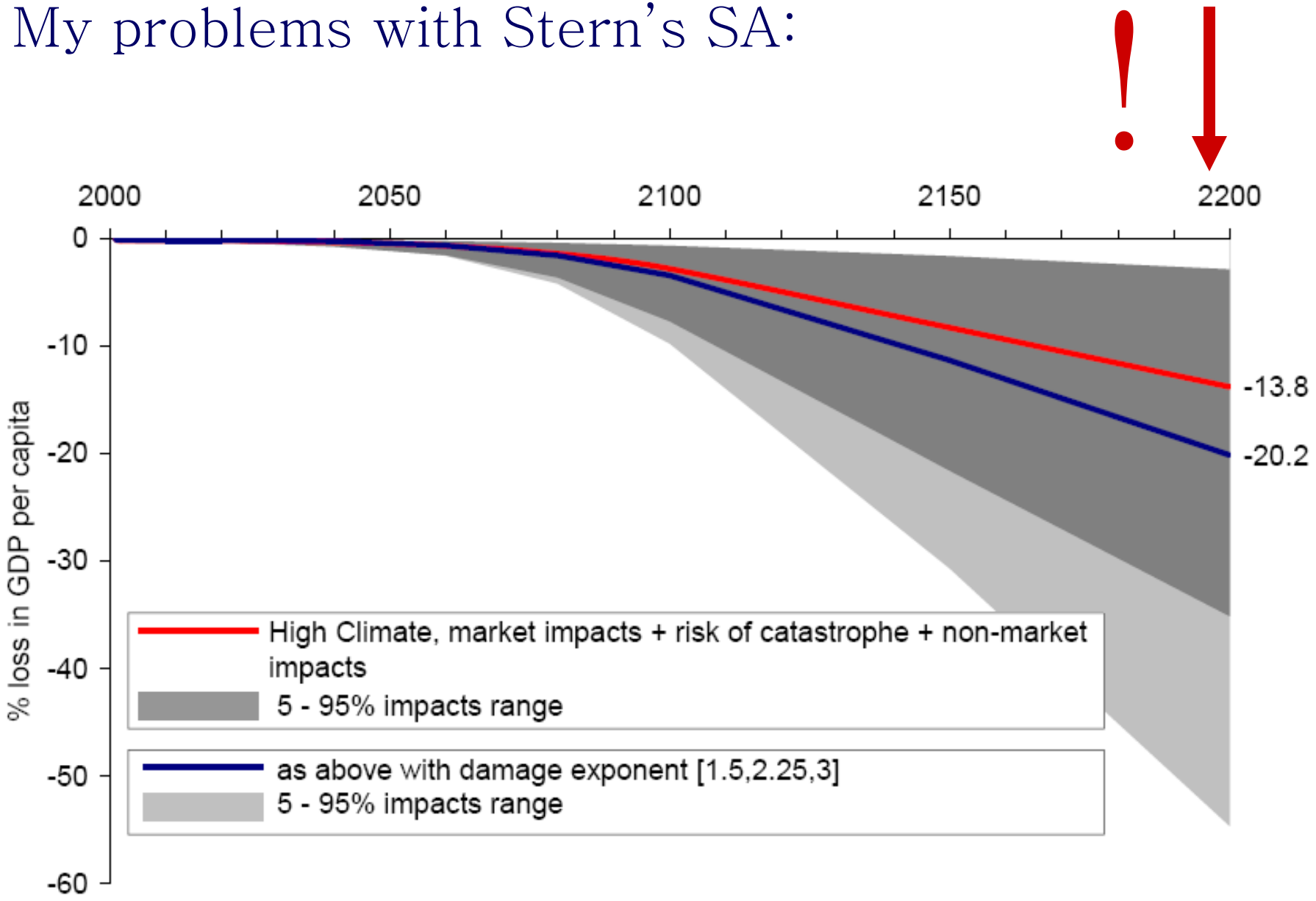
Nordhaus W., Critical Assumptions in the Stern Review on Climate Change, SCIENCE, 317, 201-202, (2007).

The Stern – Nordhaus exchange on *SCIENCE*

Nordhaus → attacks Stern based on 'wrong' range of discount rate (~ you are GIGOing)

Stern → Perform a sensitivity analysis and retorts: 'My analysis shows robustness'

My problems with Stern's SA:

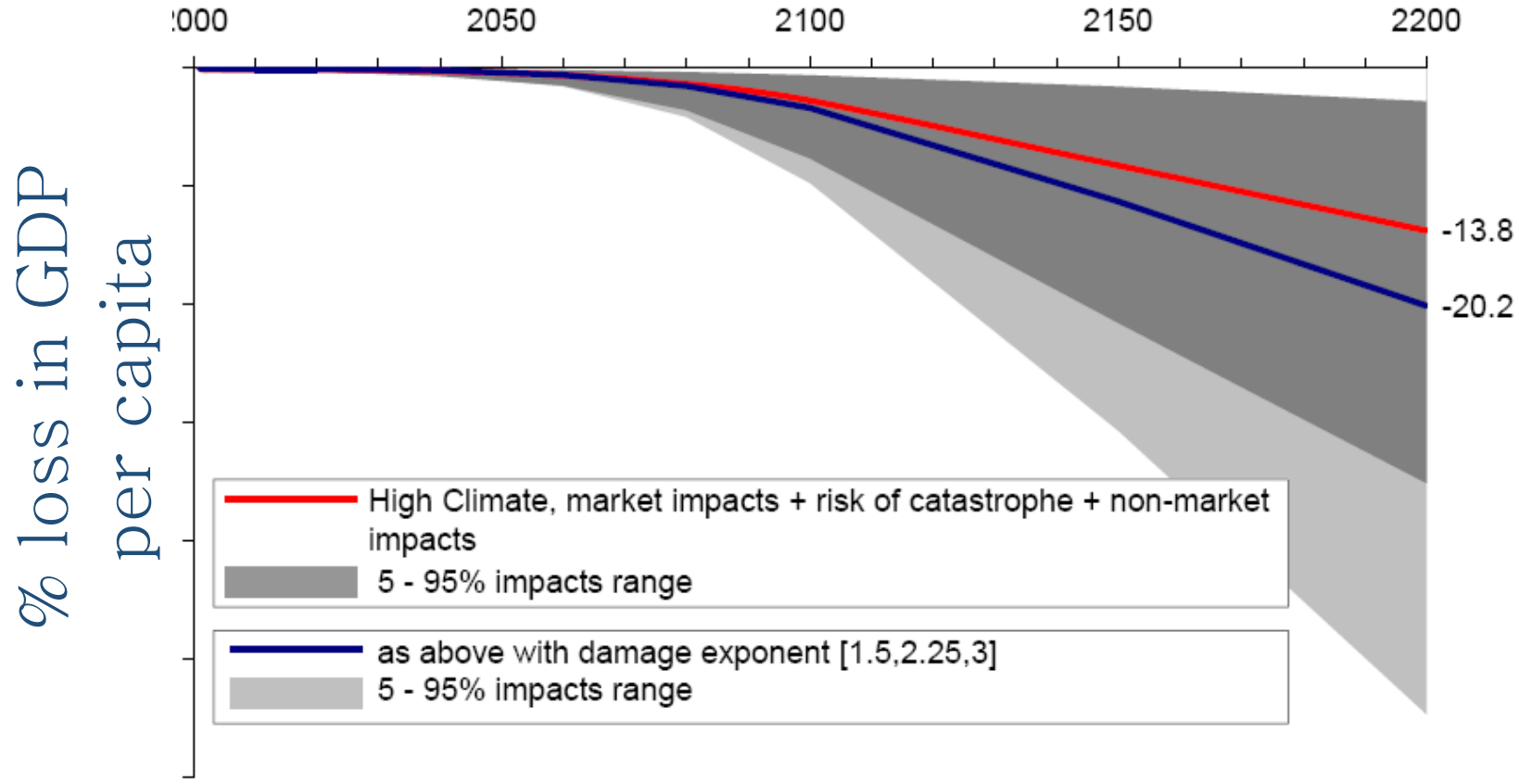


… but foremost Stern says:

changing assumptions → important effect

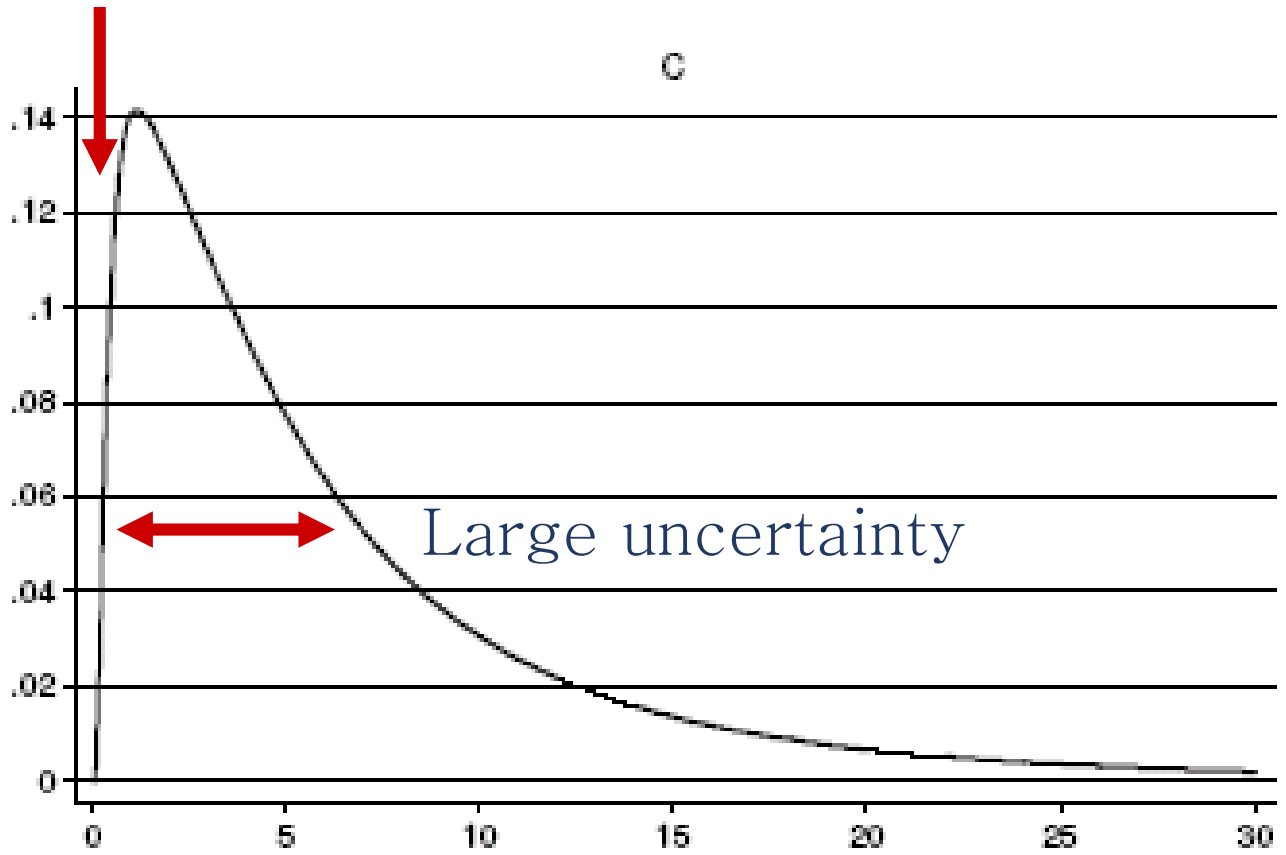
when instead he should admit that:

changing assumptions → all changes a lot



How was it done? A reverse engineering of the analysis

Missing points



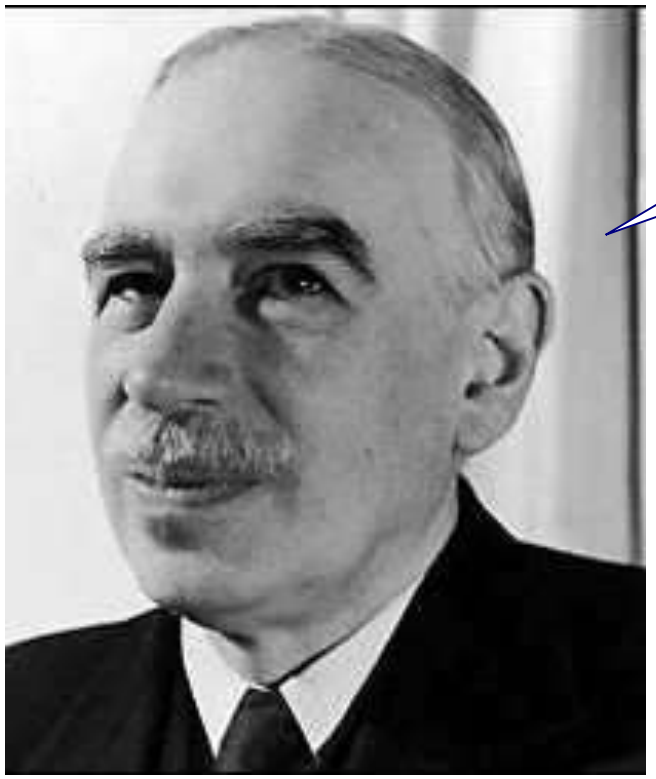
% loss in GDP per capita

Conclusion:

Model prediction are too uncertain to adjudicate the dispute about the urgency of action on climate change;

Both assertion (Stern) and refutation (Nordhaus) are indefensible

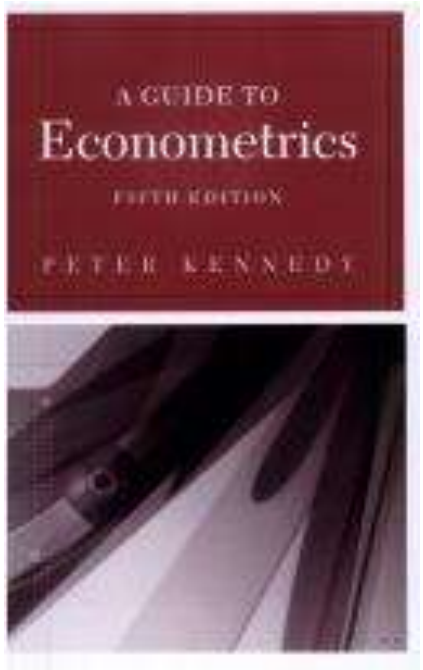
Same criticism applies to Nordhaus – both authors frame the debate around numbers which are …



… precisely wrong

Peter Kennedy, A Guide to Econometrics.

One of the ten commandments of applied econometrics according to Peter Kennedy:



“Thou shall confess in the presence of sensitivity. Corollary: Thou shall anticipate criticism “

NEVER vary all factors of the same amount

Be it 5%, 10%, or 20%

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

… beside uncertainty can be used instrumentally

Don’t stop at sensitivity analysis

COMMENT · 24 JUNE 2020

Five ways to ensure that models serve society: a manifesto

Pandemic politics highlight how predictions need to be transparent and humble to invite insight, not blame.

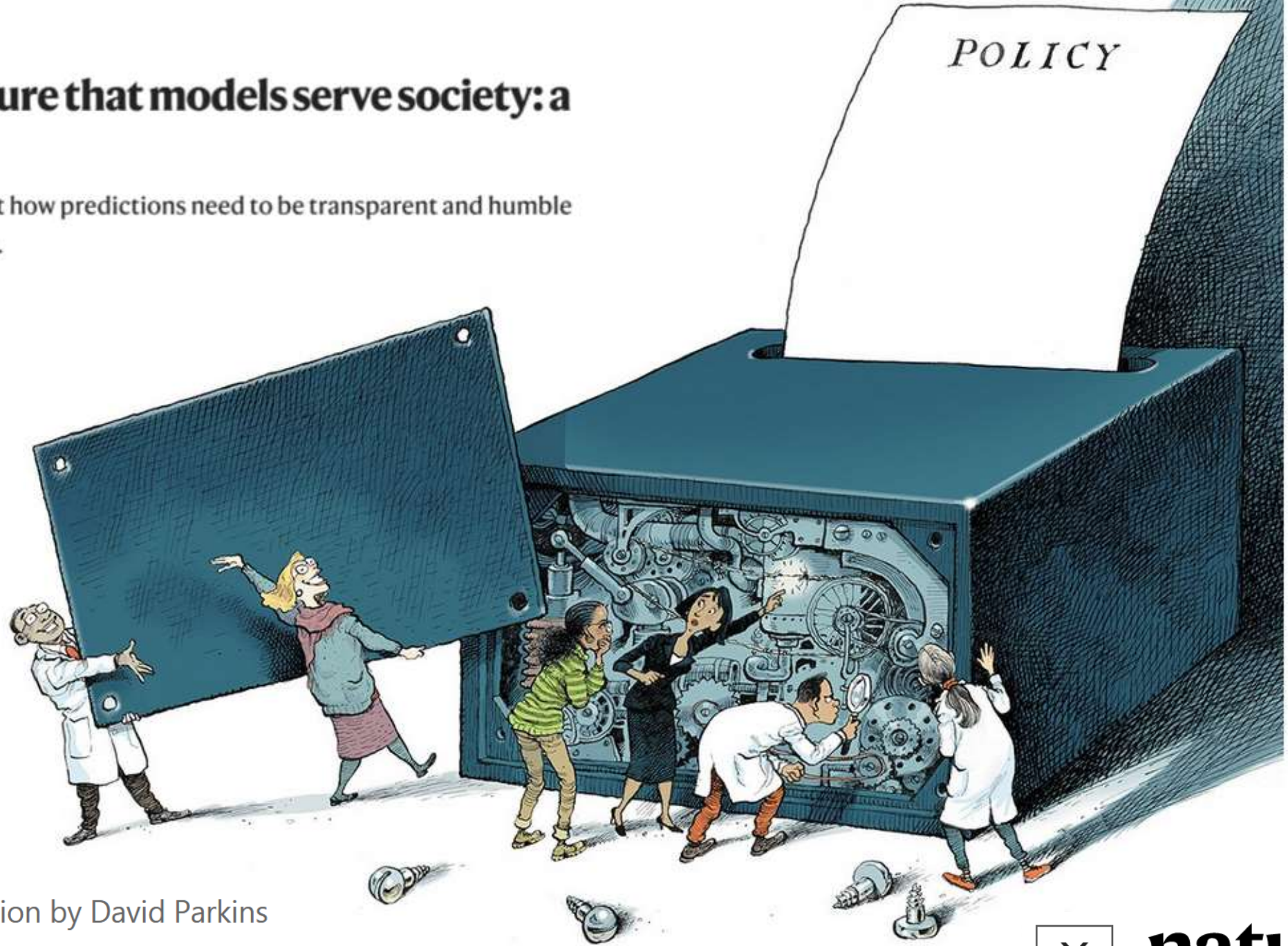


Illustration by David Parkins

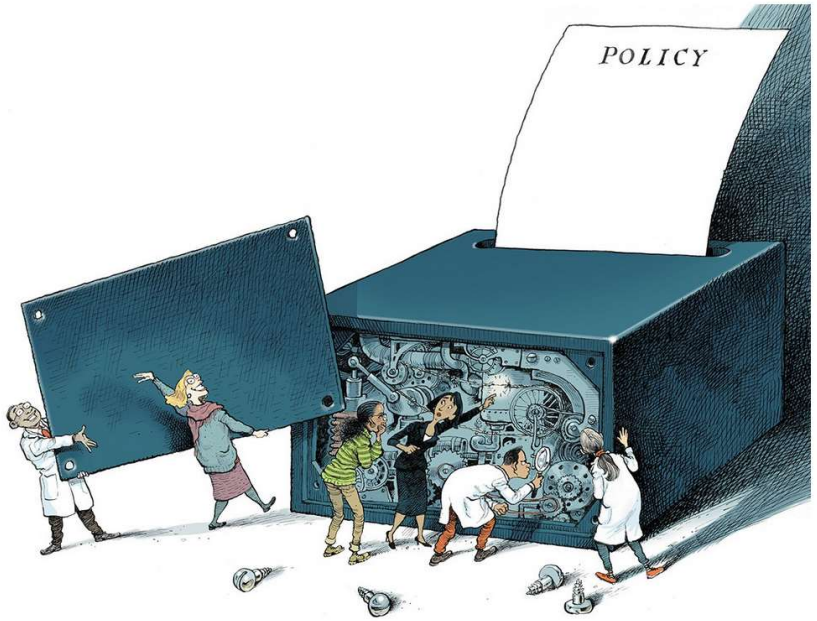
nature

nature

Andrea Saltelli, Gabriele Bammer, Isabelle Bruno, Erica Charters, Monica Di Fiore, Emmanuel Didier, Wendy Nelson Espeland, John Kay, Samuele Lo Piano, Deborah Mayo, Roger Pielke Jr, Tommaso Portaluri, Theodore M. Porter, Arnald Puy, Ismael Rafols, Jerome R. Ravetz, Erik Reinert, Daniel Sarewitz, Philip B. Stark, Andrew Stirling, Jeroen van der Sluijs & Paolo Vineis

3 modellers Lo Piano, Puy, Saltelli
2 experts models and society Pielke, van der Sluijs
3 statisticians Mayo, Stark, Portaluri
2 statactivistes Bruno, Didier
2 economists Kay, Raynert
1 epidemiologist Vineis

2 sociologists of quantification Espeland, Porter
3 STS scholars Bammer, Sarewitz, Stirling
1 philosopher Ravetz
1 historian Charters
1 political scientists Di Fiore
1 expert RRI - Open Science Rafols



Mind the assumptions

Assess uncertainty and sensitivity

Mind the hubris

Complexity can be the enemy of relevance

Mind the framing

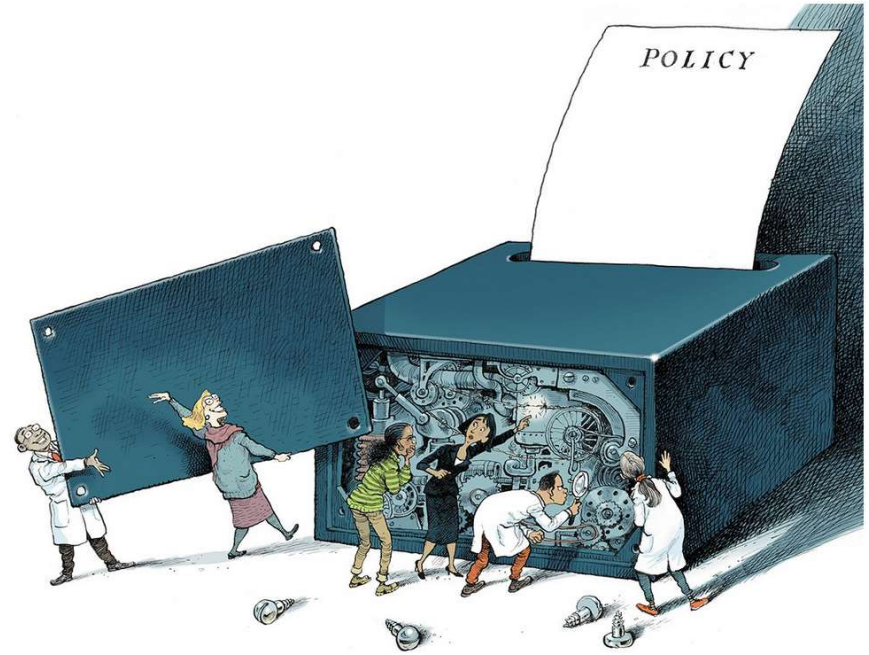
Match purpose and context

Mind the consequences

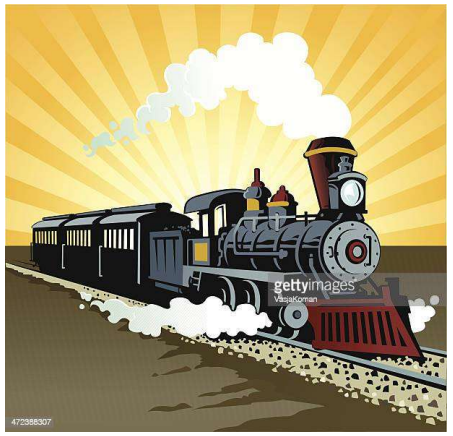
Quantification can backfire

Mind the unknowns

Acknowledge ignorance



Next stop: Ethics of quantification



http://materials.cv.uoc.edu/cdocent/PID_00284929/

Don't use just any method

➔ Use the method appropriate to context and purpose; the example of variance based / moment independent / VARS methods

Don't use One factor At a Time (OAT)

➔ Geometric proof paper plus 'why false SA' paper

Don't use method that are not model-independent (such as PCC, PRCC)

➔ Early SA papers CSDA RESS

Don't use either LHS or optimized LHS

➔ Quasi random numbers and relative papers; mind the constructive dimension (Owen, Kucherenko)

Don't run the model just once

➔ Lubarsky's cybernetic enthomology

Don't use Morris' method

➔ Dependence upon one extra design parameter plus ambiguity in interpretation (mu and sigma); Paper 2011 showing superiority T_j over Morris

Don't confuse the map with the territory

➔ J.L. Borges; Yucca Mountain example; Rayner's displacement

Beware the dimension of your model

➔ The conjecture of O'Neil

Don't sample just parameters and boundary conditions

➔ Use e.g. triggers to explore the effect of other assumptions

Don't go public with your results without having seen your SA

The case of the Stern-Nordhaus controversy

NEVER vary all factors of the same amount (5%, 10%, 20%)

➔ Avoiding GIGO

Don't stop at sensitivity analysis

➔ Think about the sensitivity auditing

The End

@andreasaltelli

Question 1

The influence of the key variables should be investigated by a sensitivity analysis.

- Is something wrong about the statement above (p. 384 of EC guidelines)

Question 2

"Are the results from a particular model more sensitive to changes in the model and the methods used to estimate its parameters, or to changes in the data?" (Majone 1989)

Please comment

Question 3

What uses can you imagine for sensitivity analysis?

Presenter: Andrea Saltelli

When: June 28th, 9.30-10.30

Format: Physical

Room: 511

Summary of the talk: The way SA is mostly done in the literature is wrong. So one can introduce SA 'via negativa' (N.N. Taleb), focusing on what pitfalls should be avoided. Mathematical modelling itself can be effectively used to rule out what cannot be (e.g. in terms of policy options) so that only viable alternatives are left for further investigation. SA also helps to tackle the known tradeoff between model complexity and model error, known in ecology as the conjecture of O'Neil. This becomes particularly relevant where the model does not dispose of a validation data set, e.g. when the model is predicting in either the future or the unknown. Finally, if your friends are sociologists, they might object that even your sensitivity analysis does not go far enough because of embedded or implicit assumptions. In the talk we shall visit these eternal problems with examples.

Reading material:

Andrea Saltelli, Ksenia Aleksankina, William Becker, Pamela Fennell, Federico Ferretti, Niels Holst, Sushan Li, Qiongli Wu, Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices, Environmental Modelling and Software, Volume 114, April 2019, Pages 29-39.

More Reading:
https://www.nature.com/articles/d41586-020-01812-9
https://www.sciencedirect.com/science/article/pii/S1364815221002681?via%3Dihub