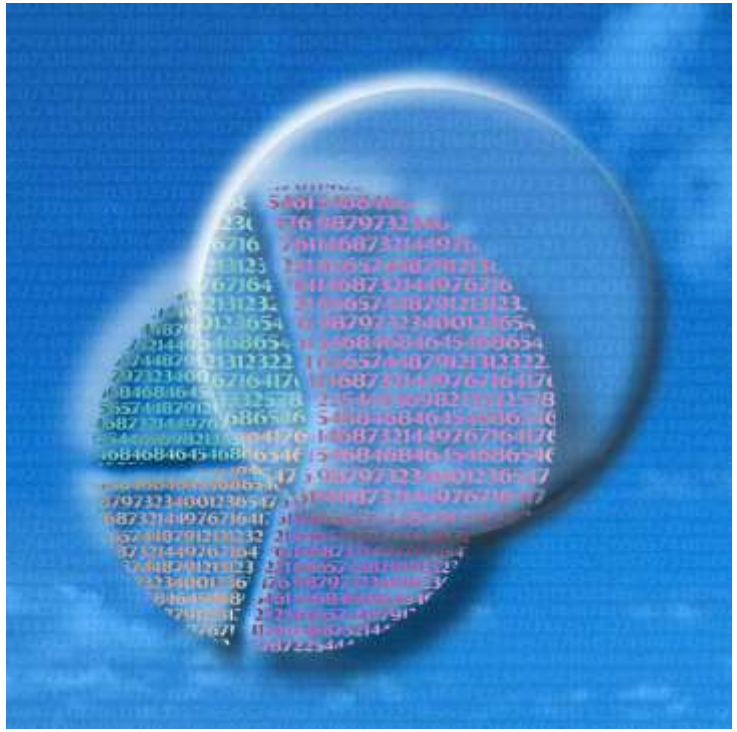


How not to do a sensitivity analysis

Andrea Saltelli

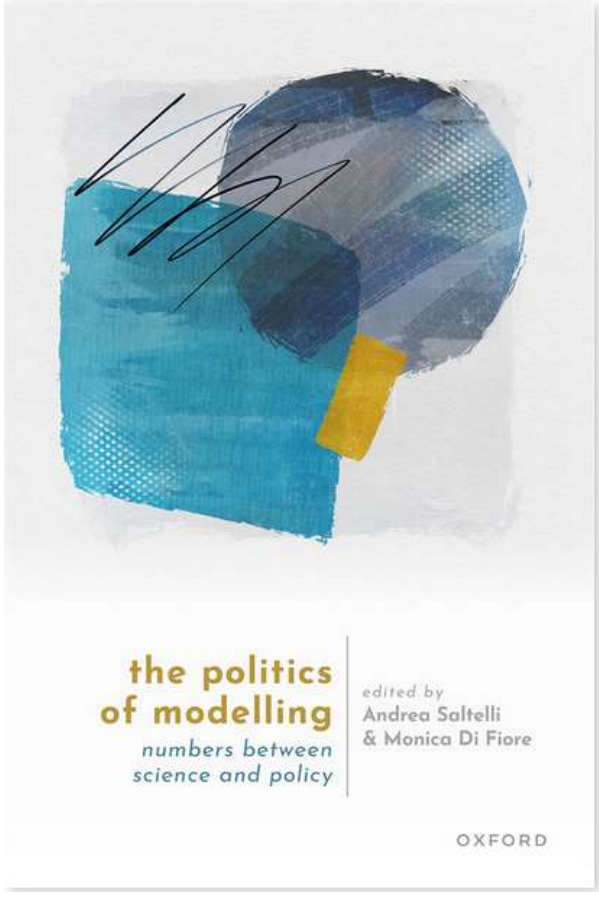
Course MNF990, University of Bergen, September 22, 2023

Where to find this talk: www.andreasaltelli.eu

Andrea Saltelli

HOME ABOUT ME PUBLICATIONS NEWS & VIDEOS RESOURCES

August 25 2023: The politics of modelling is out!



Praise for the volume

"A long awaited examination of the role —and obligation —of modeling."
Nassim Nicholas Taleb , Distinguished Professor of Risk Engineering, NYU Tandon School of Engineering. Author, of the 5 -volume series Incerto.

***

"A breath of fresh air and a much needed cautionary view of the ever-widening dependence on mathematical modeling."
Orrin H. Pilkey, Professor at Duke University's Nicholas School of the Environment, co-author with Linda Pilkey-Jarvis of Useless Arithmetic: Why Environmental Scientists Can't Predict the Future, Columbia University Press 2009.

***

"The methods by which power insinuates itself

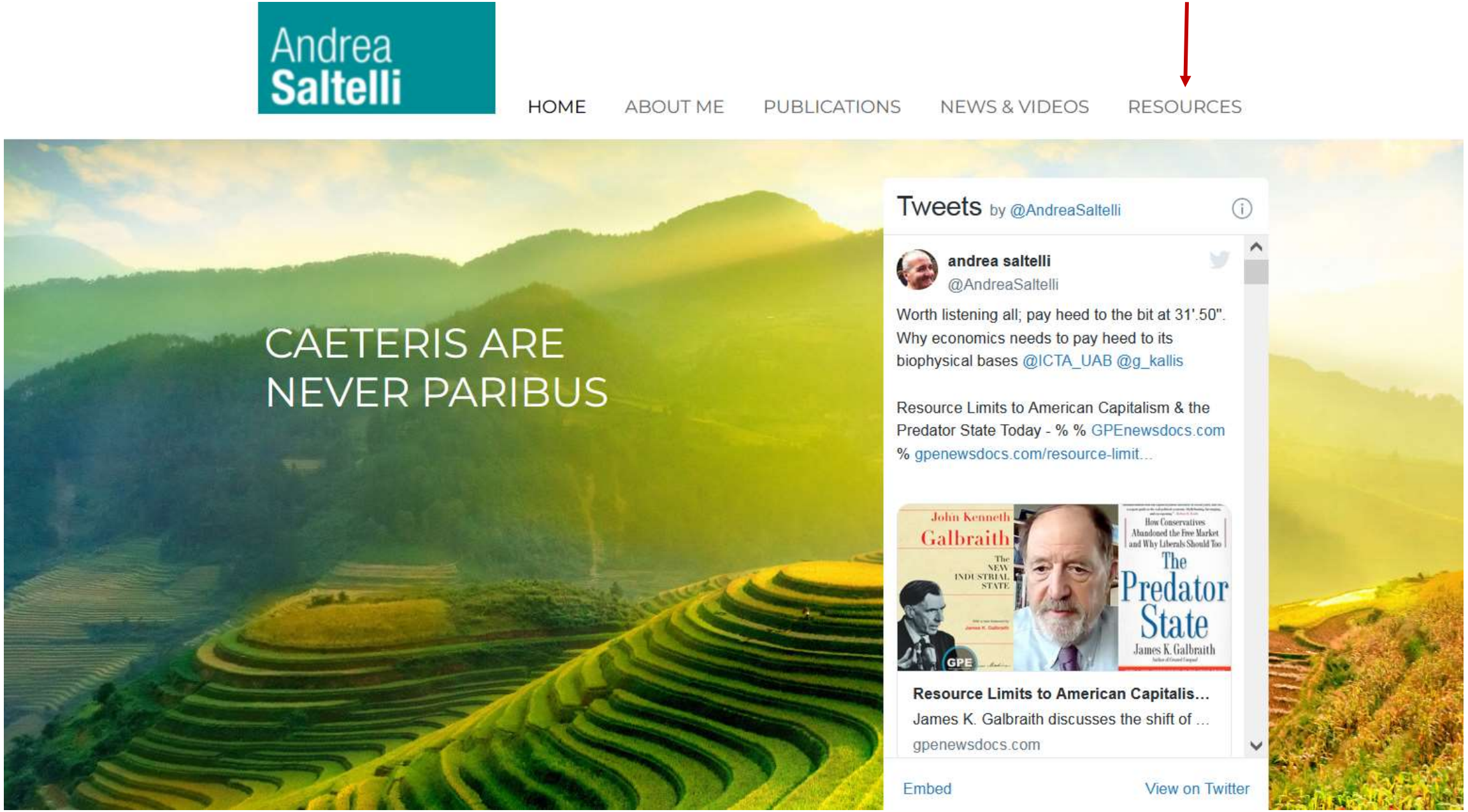
Mastodon Toots by
@AndreaSaltelli

AndreaSaltelli
2023/8/28 11:24
August 26 Podcast (16m) - interview for ABC NET RADIO, AUS: Assumptions and consequences: the politics of modelling, Guests: Ehsan Nabavi and Andrea Saltelli, Producer - Chris Bullock.

abc.net.au/listen/programs/sun

View on mstdn.social

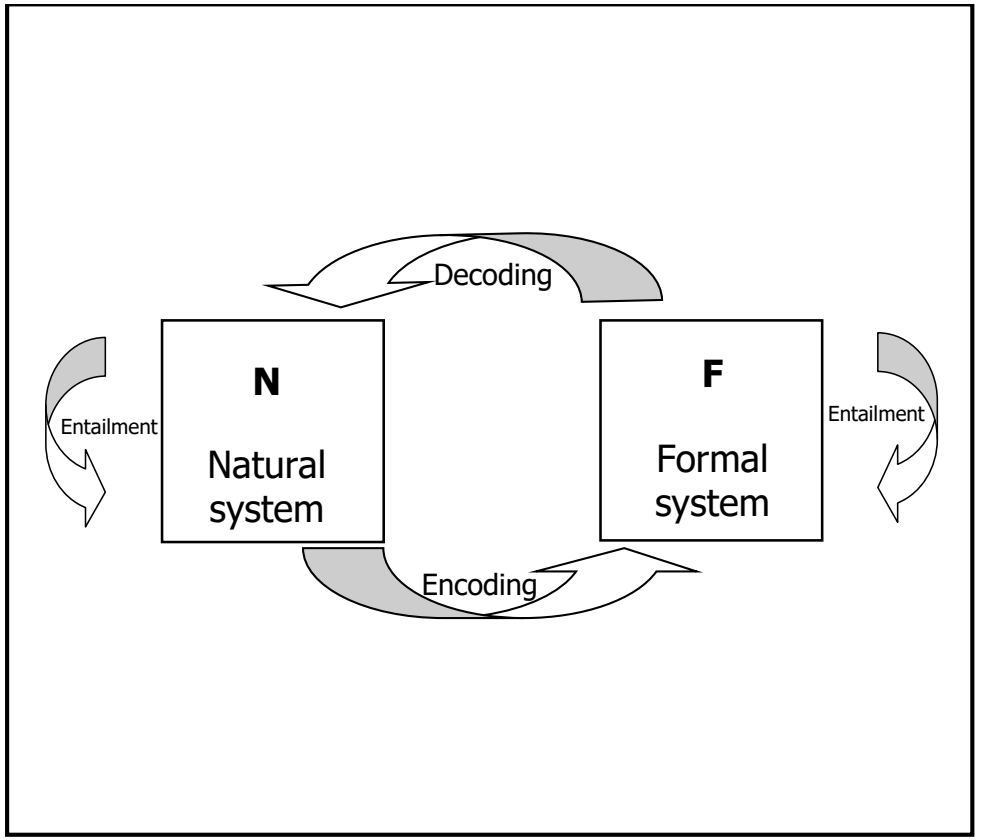
Where to find this talk: www.andreasaltelli.eu

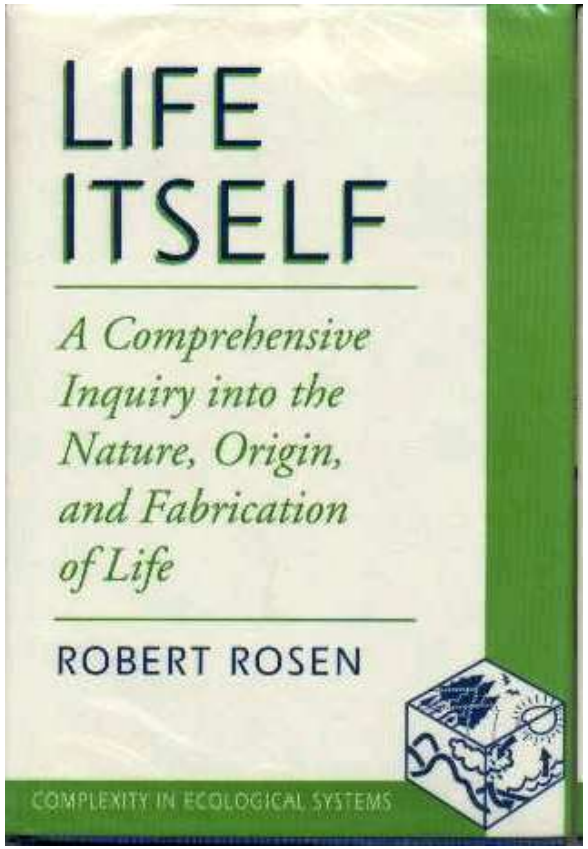


Something general about mathematical modelling

Modelling is a craft more than a science

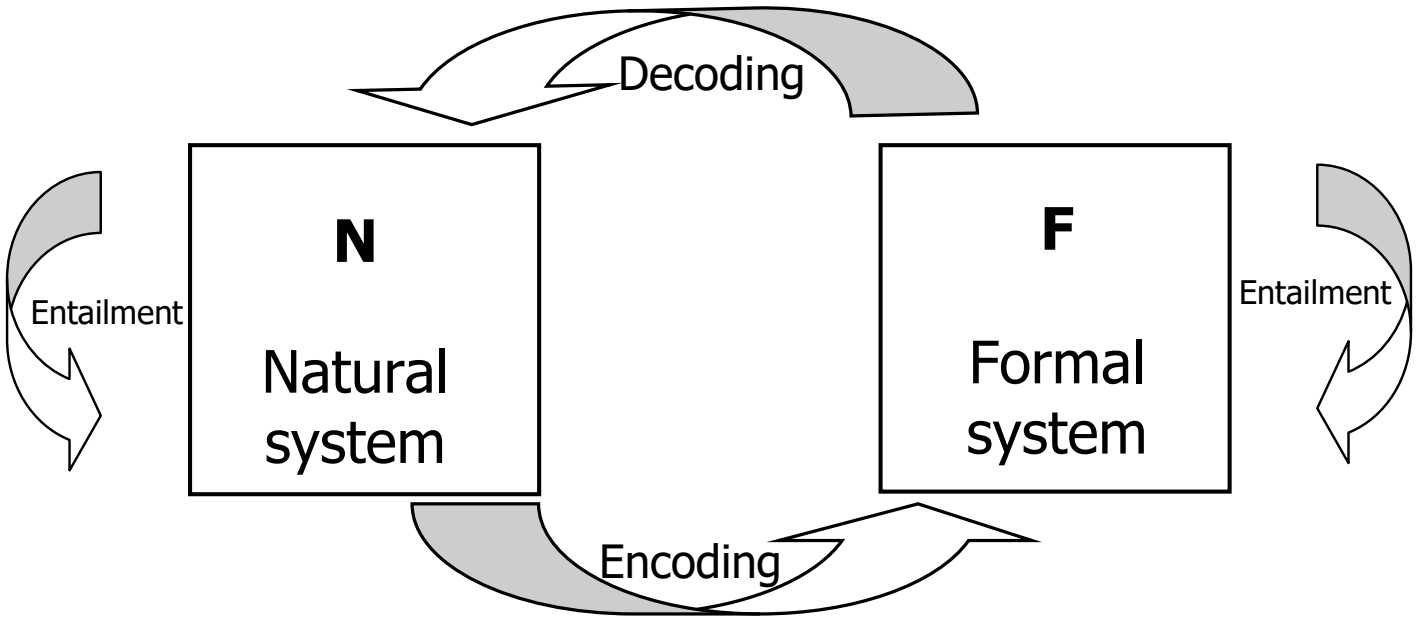
Modelling as a craft rather than as a science for Robert Rosen





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 ?

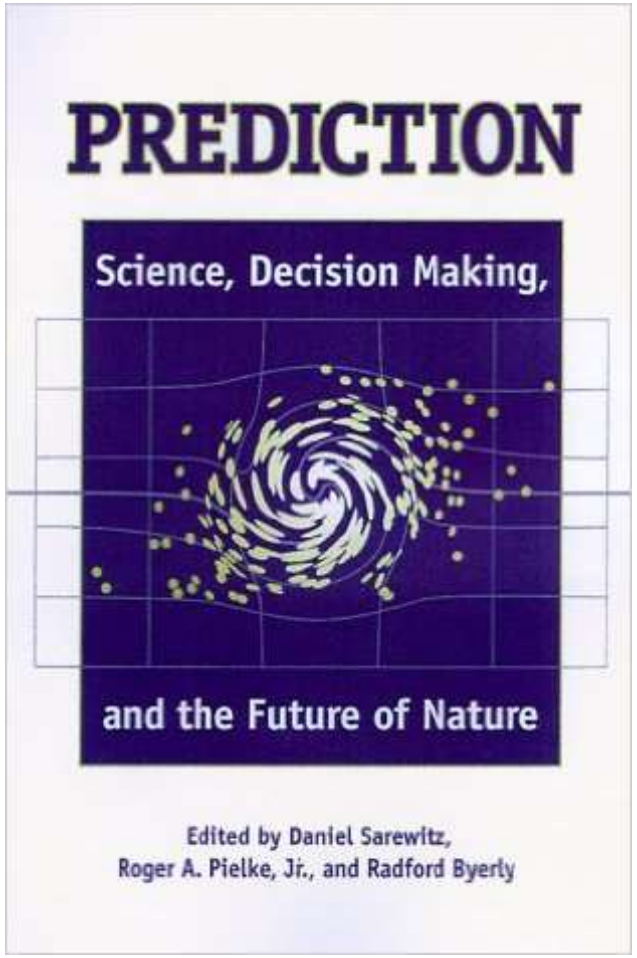
Robert Rosen

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

Naomi Oreskes

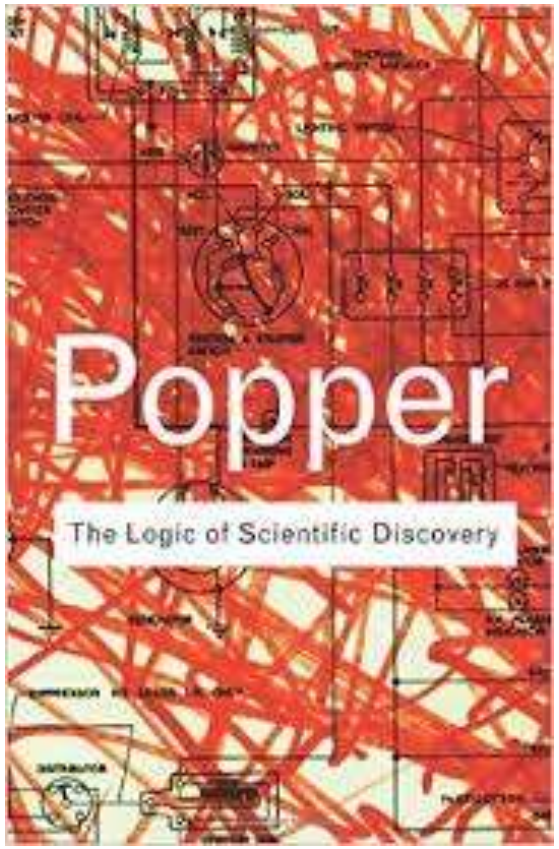
N. Oreskes, K. Shrader-Frechette, and K. Belitz, “Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences,” Science, 263, no. 5147, 1994.

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 is a Latin phrase meaning "all other things being equal" or "other things held constant" or "all else unchanged“ (Wikipedia)

The case of DSGE, dynamic stochastic general equilibrium models

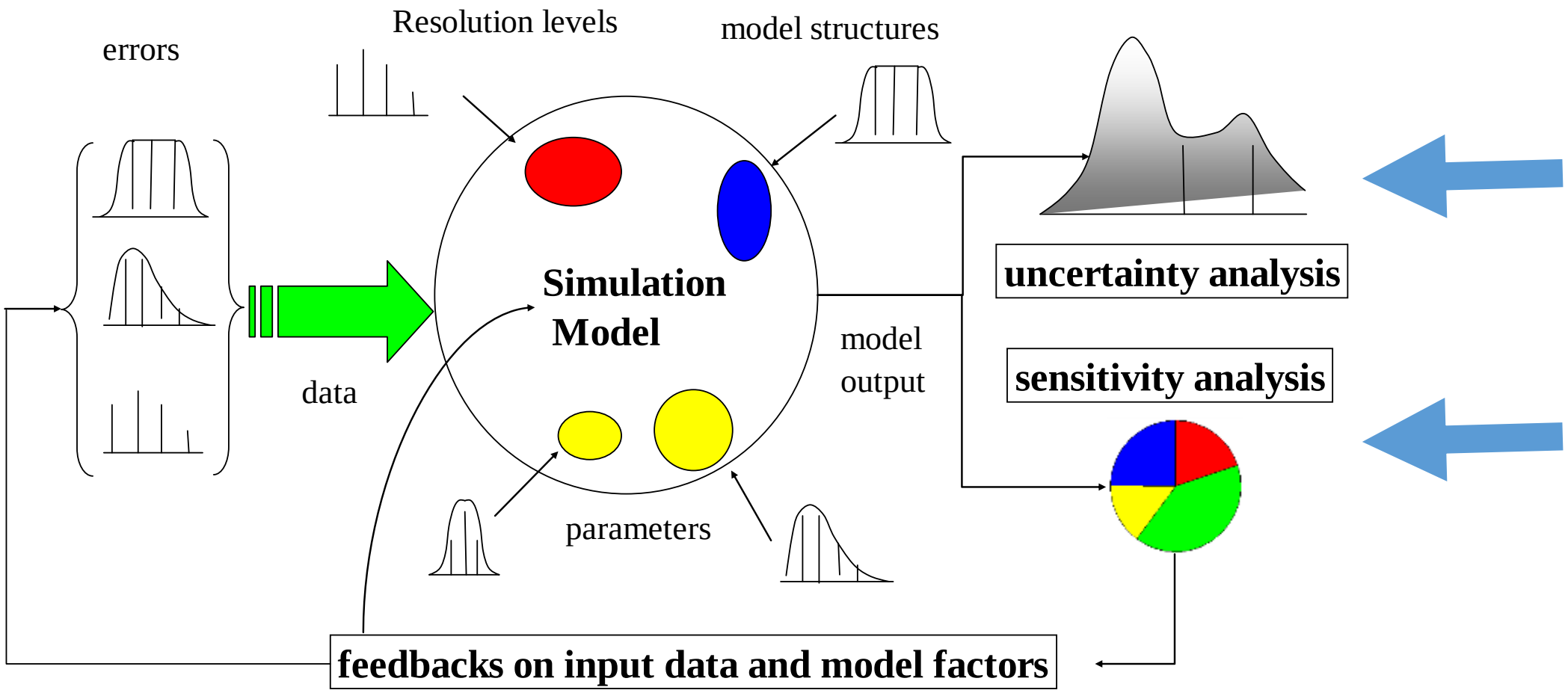
Rational expectations of agents
Efficient market hypothesis

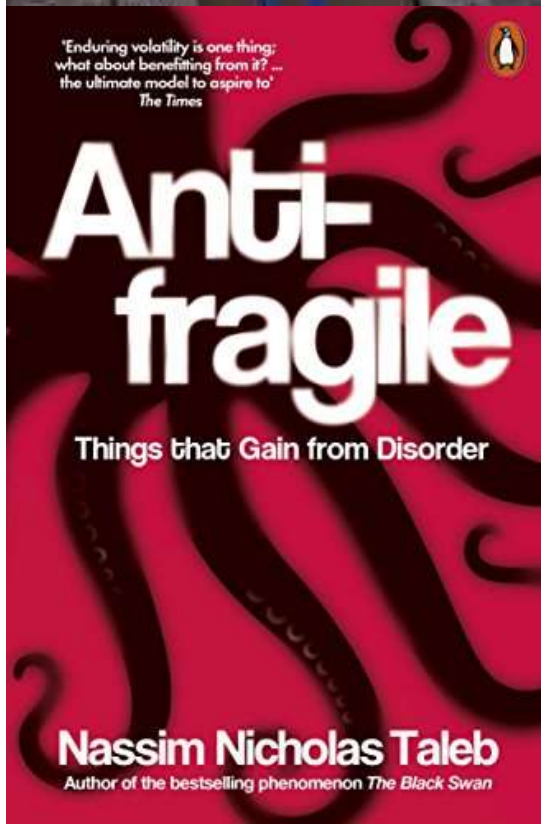
Philip Mirowski

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.


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 N. N. Taleb's *via negativa*

ELSEVIER

Environmental Modelling & Software

Volume 114, April 2019, Pages 29-39

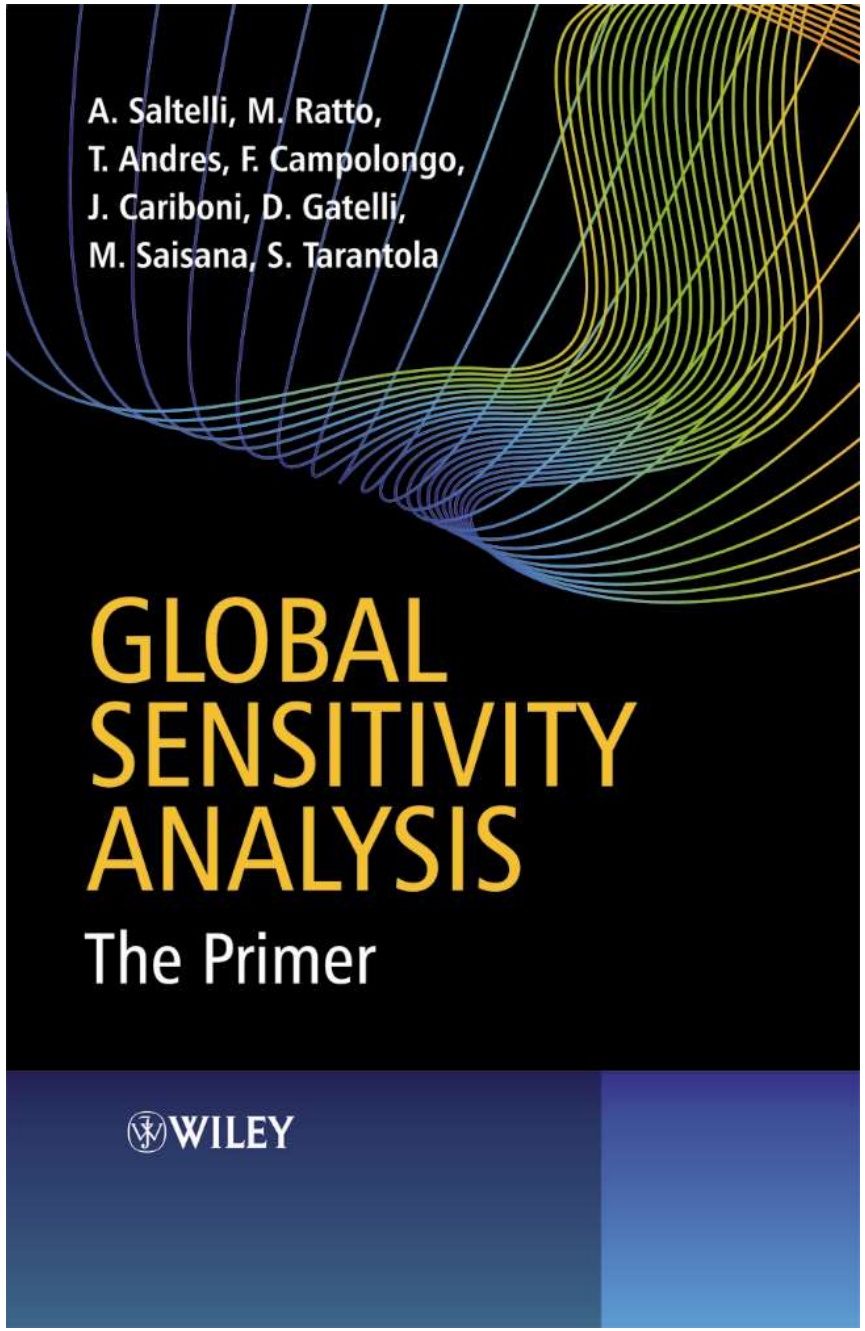
Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices

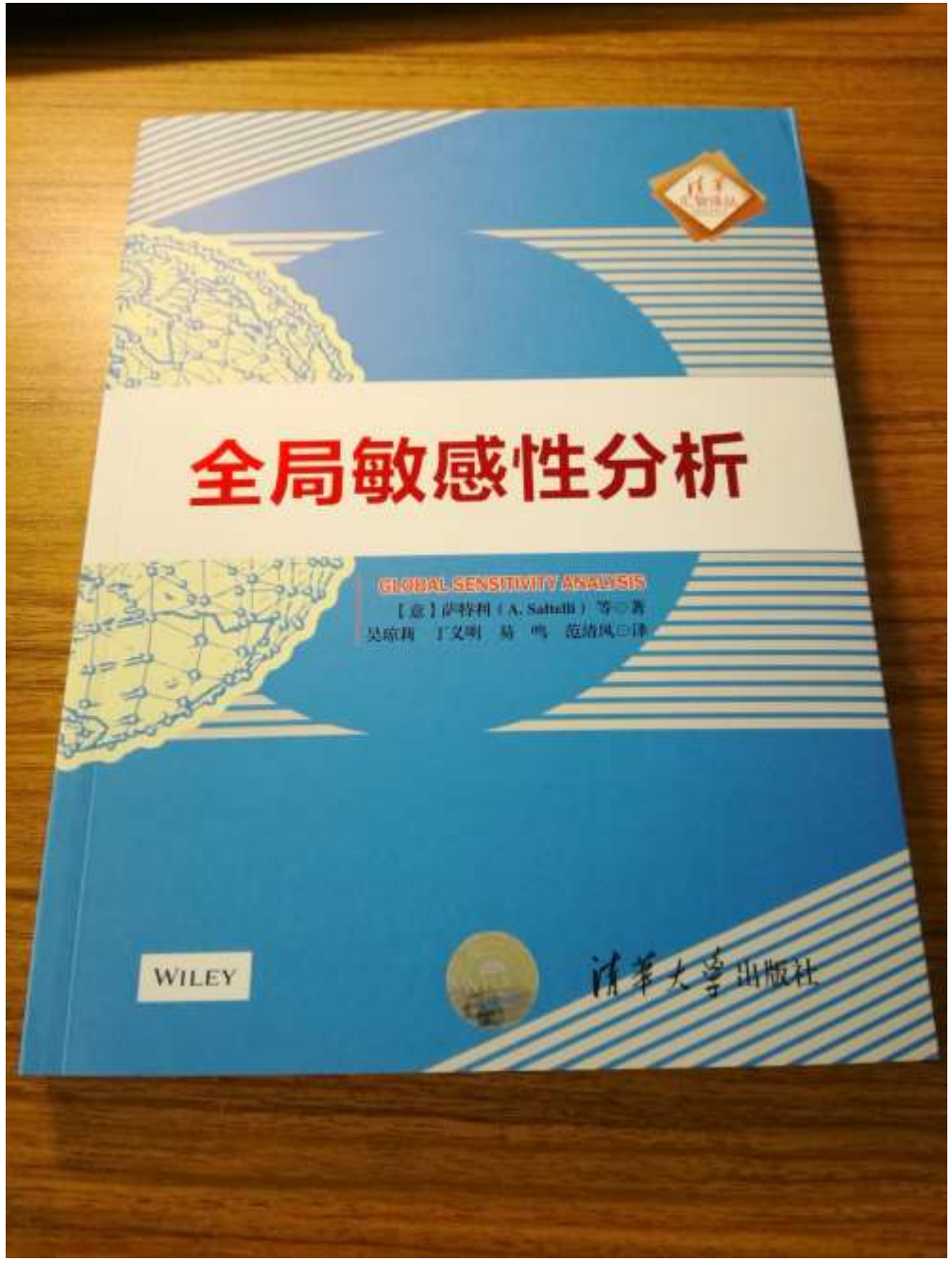
Andrea Saltelli [a, b], Ksenia Aleksankina [c], William Becker [d], Pamela Fennell [e], Federico Ferretti [d], Niels Holst [f], Sushan Li [g], Qiongli Wu [h]

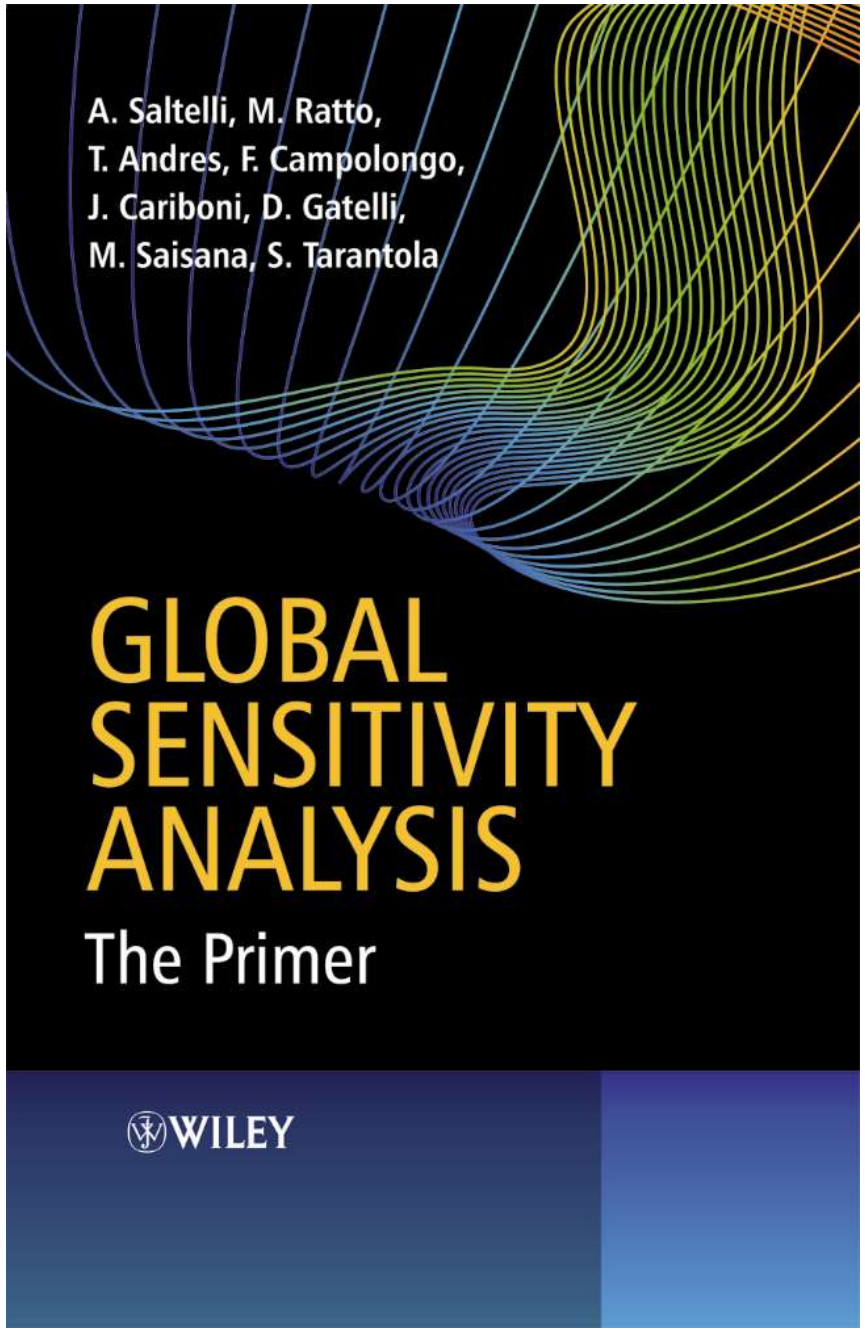
Don’t use just any method

Use the method appropriate to context and purpose

An introduction to variance based methods

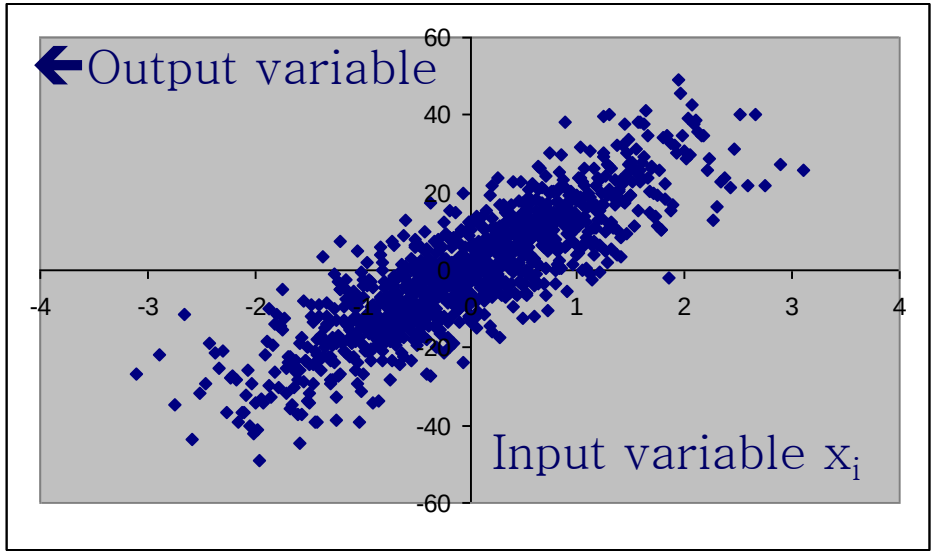

A. Saltelli, M. Ratto,
T. Andres, F. Campolongo,
J. Cariboni, D. Gatelli,
M. Saisana, S. Tarantola
GLOBAL
SENSITIVITY
ANALYSIS
The Primer
WILEY

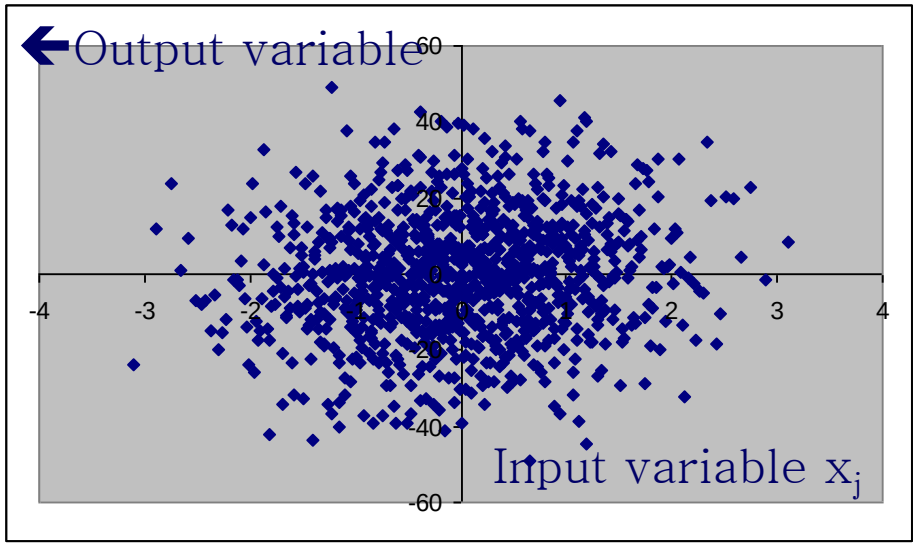

全局敏感性分析
GLOBAL SENSITIVITY ANALYSIS
WILEY



Available for free at

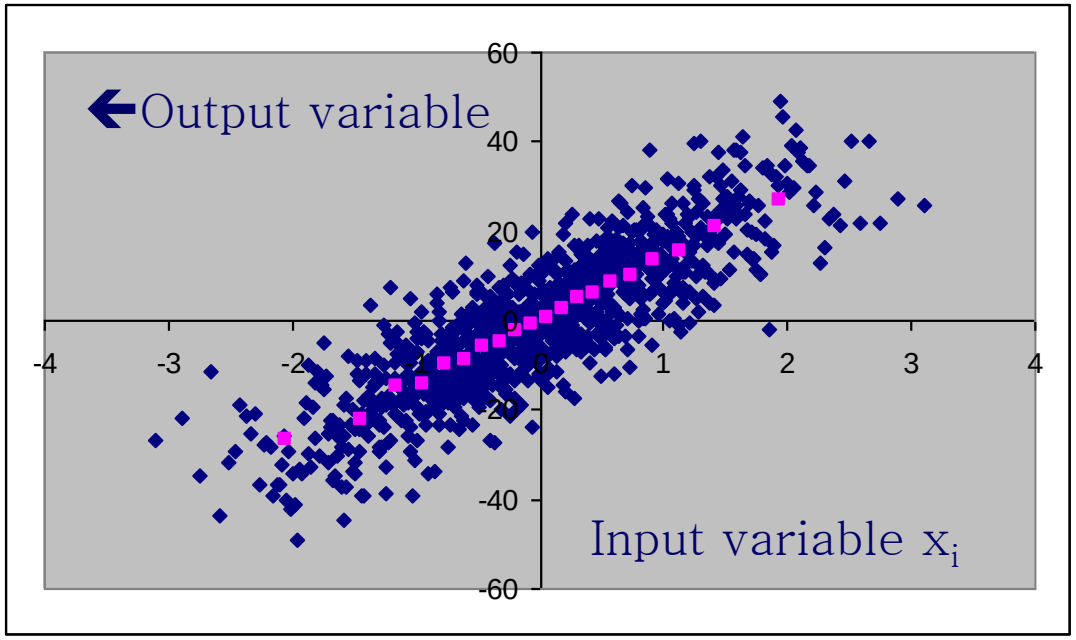
http://www.andreasaltelli.eu

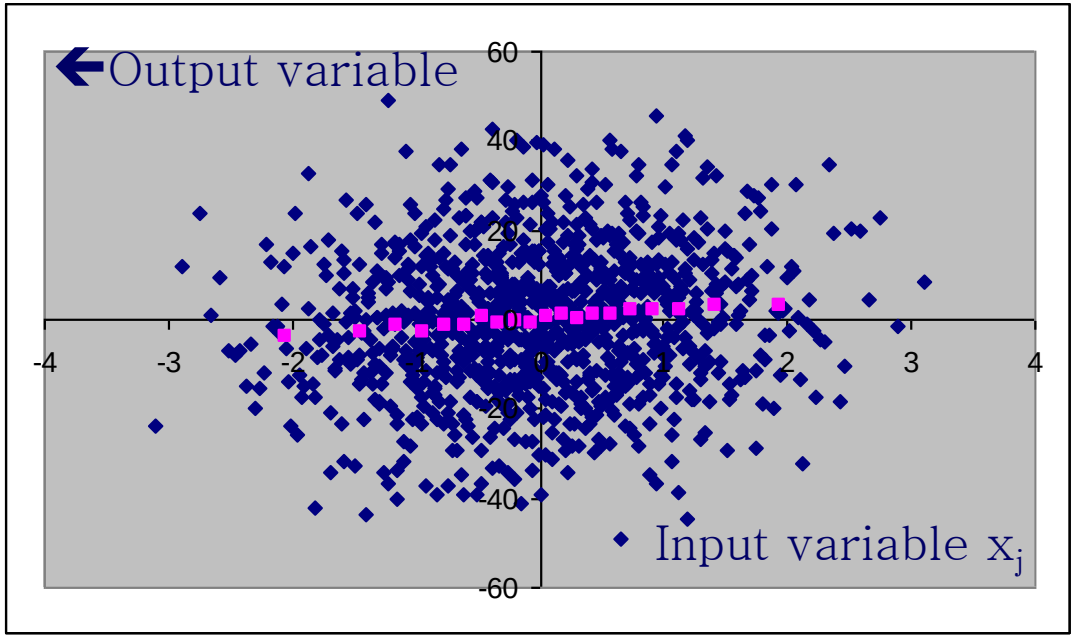




Plotting the output as a function of two different input factors

Which factor is more important?

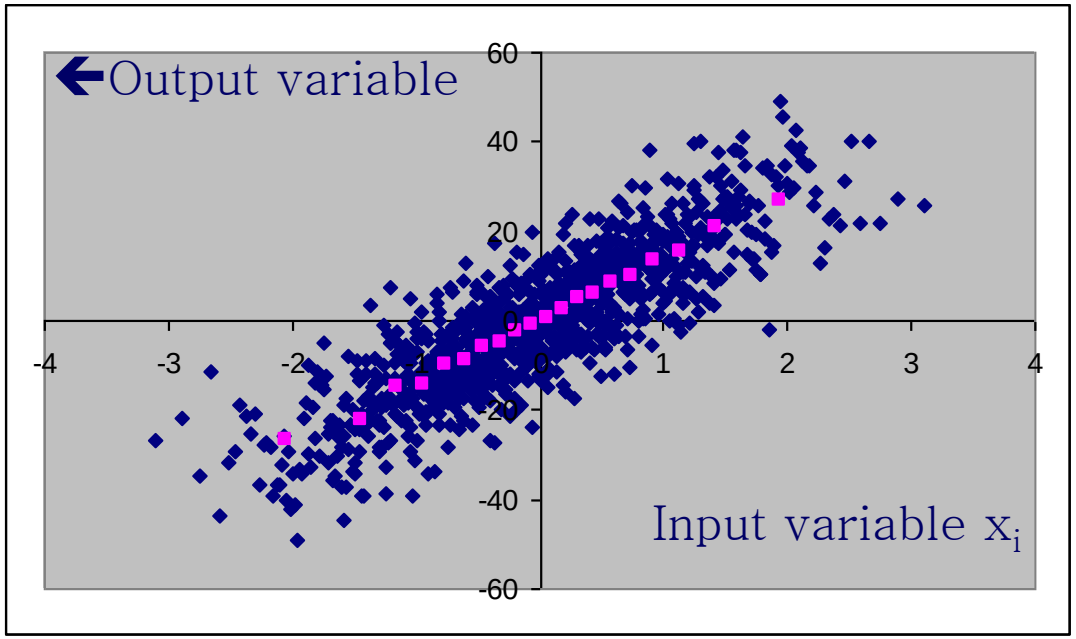




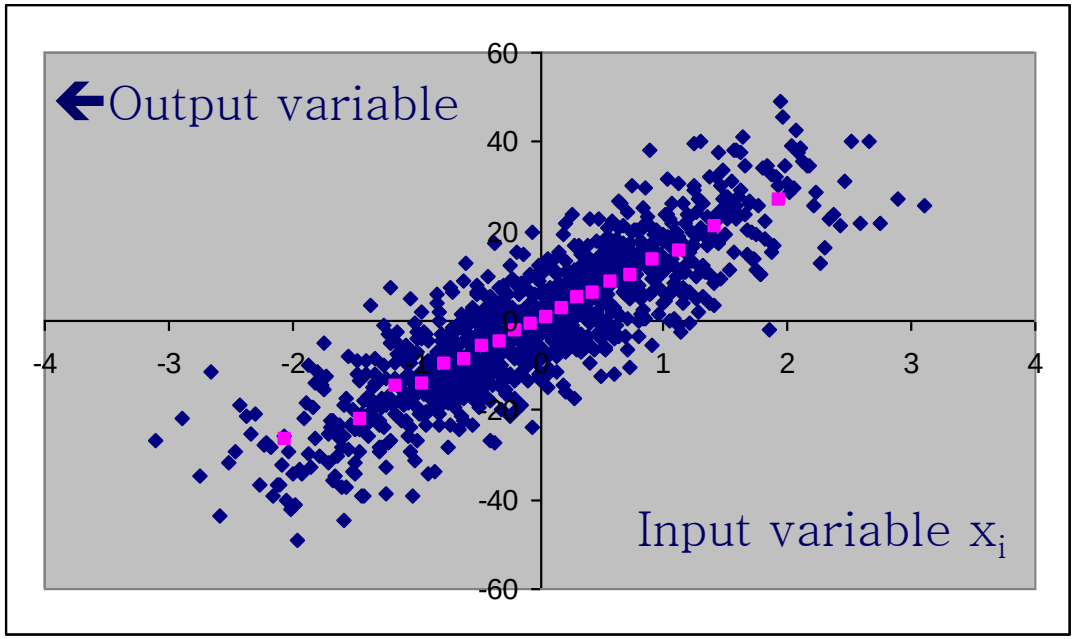
~1,000 blue points

Divide them in 20 bins of ~ 50 points

Compute the bin's average (pink dots)

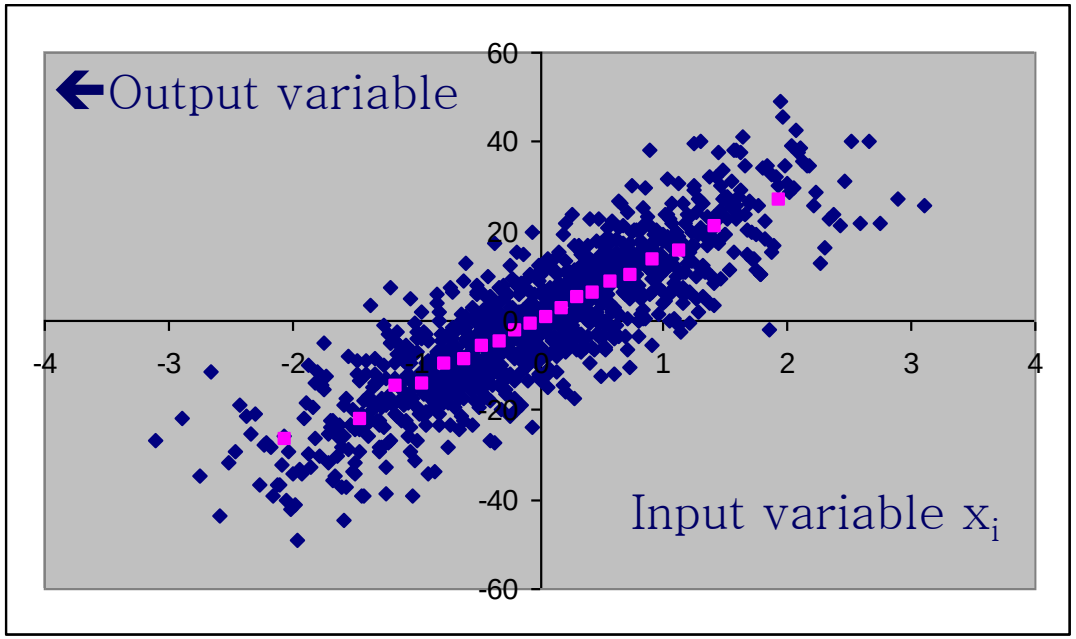


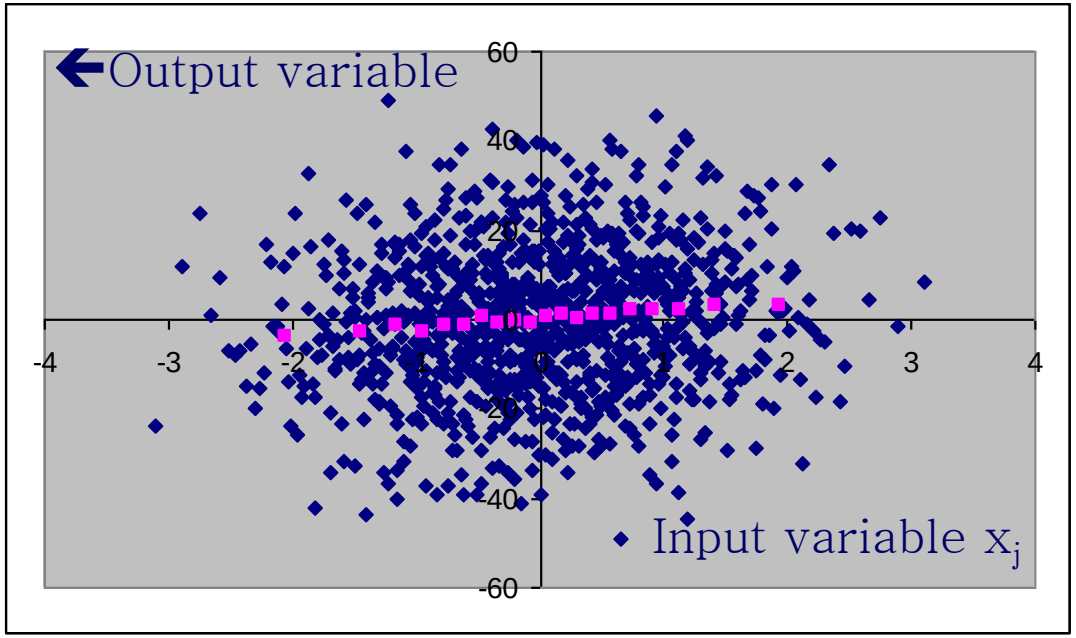
Each pink point is ~ $E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)$



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

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

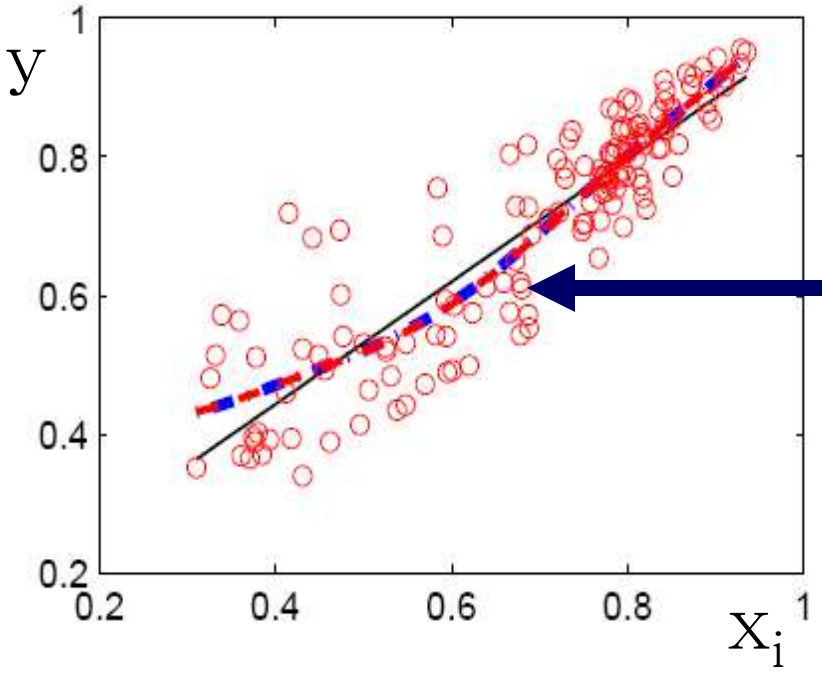
For additive models one can decompose the total variance as a sum of those partial variances

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

… 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

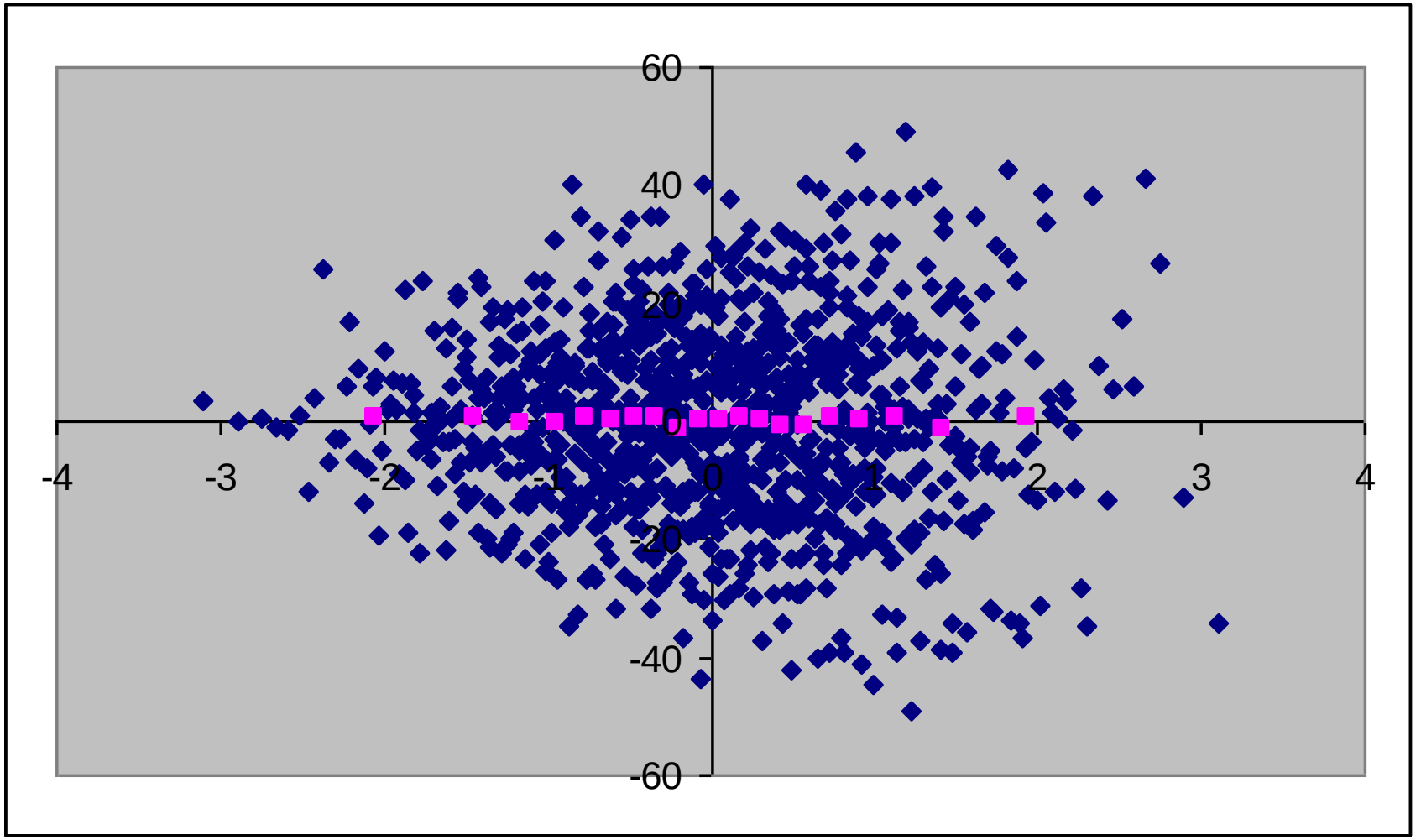
$$S_i \equiv \eta_i^2 := \frac{\mathrm{V}_{x_i}\left(\mathrm{E}_{\mathbf{x}_{\sim i}}\left(y \mid x_i\right)\right)}{\mathrm{V}\left(y\right)}$$

First order sensitivity index

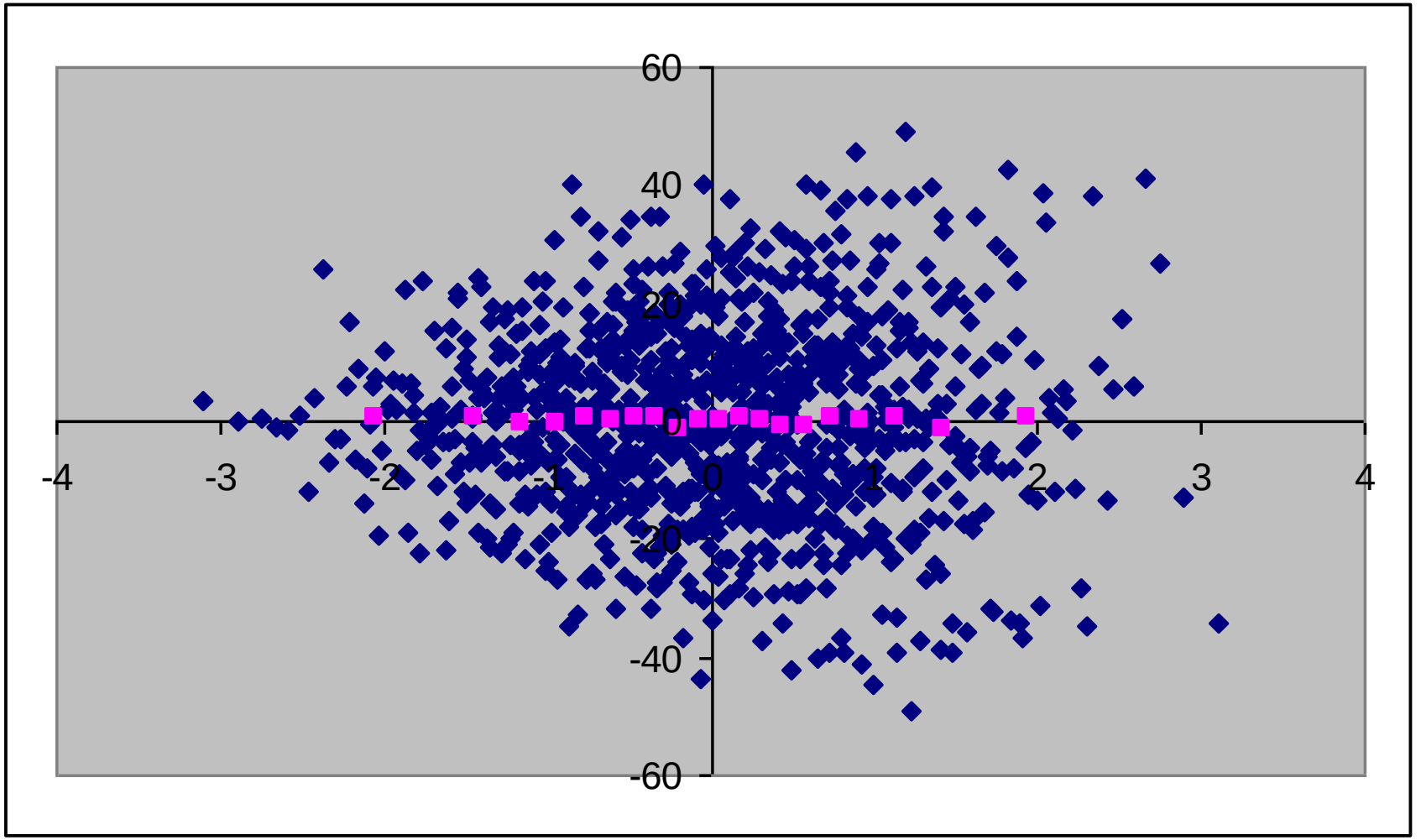
Unconditional variance

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)

$$V(Y) =$$

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

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 (only for 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)

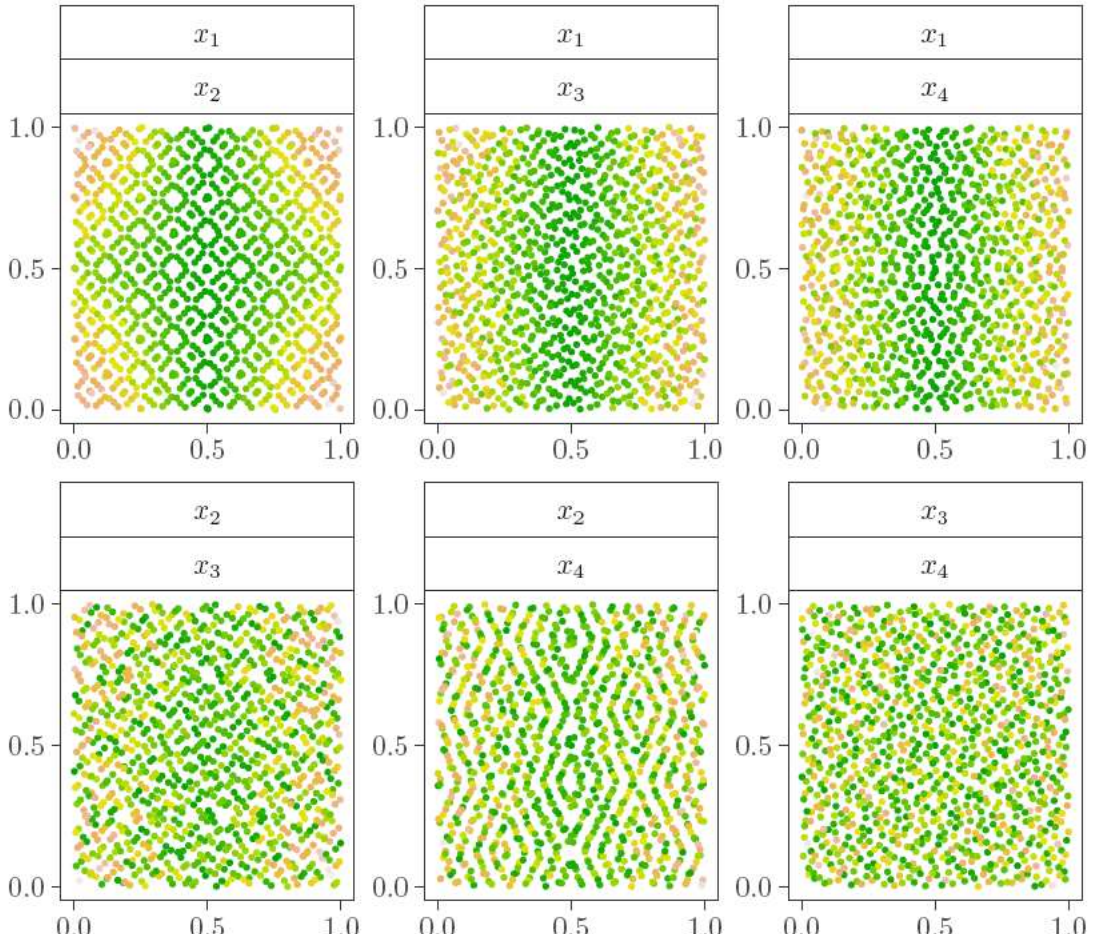
SALib https://salib.readthedocs.io/en/latest/

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

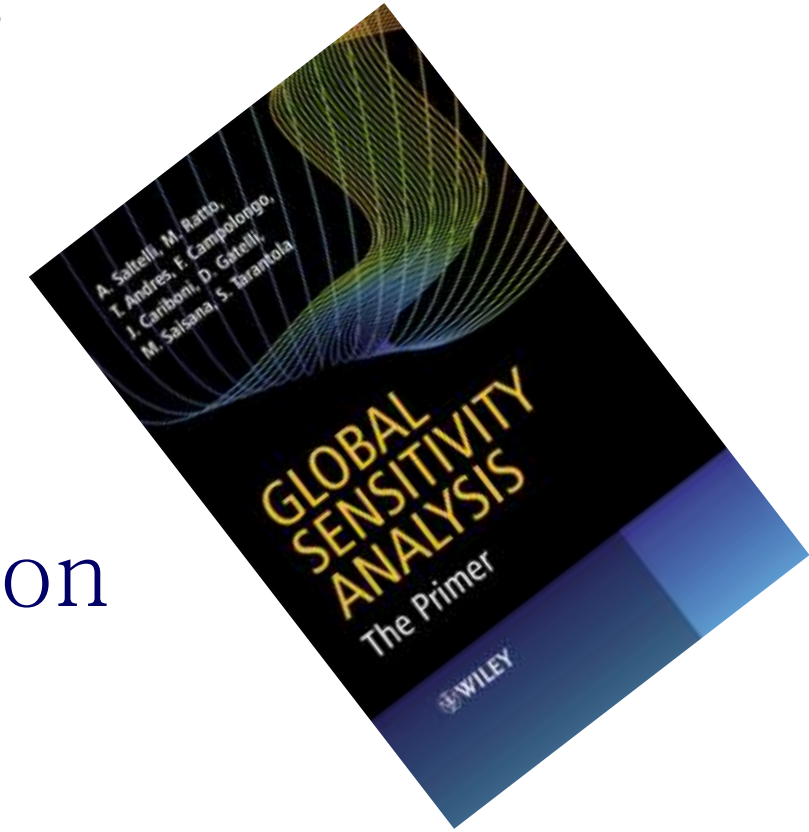
The difficulty of a function/model is not in its number of dimensions but in the number of effective dimensions, either in the **truncation** or **superposition** sense

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

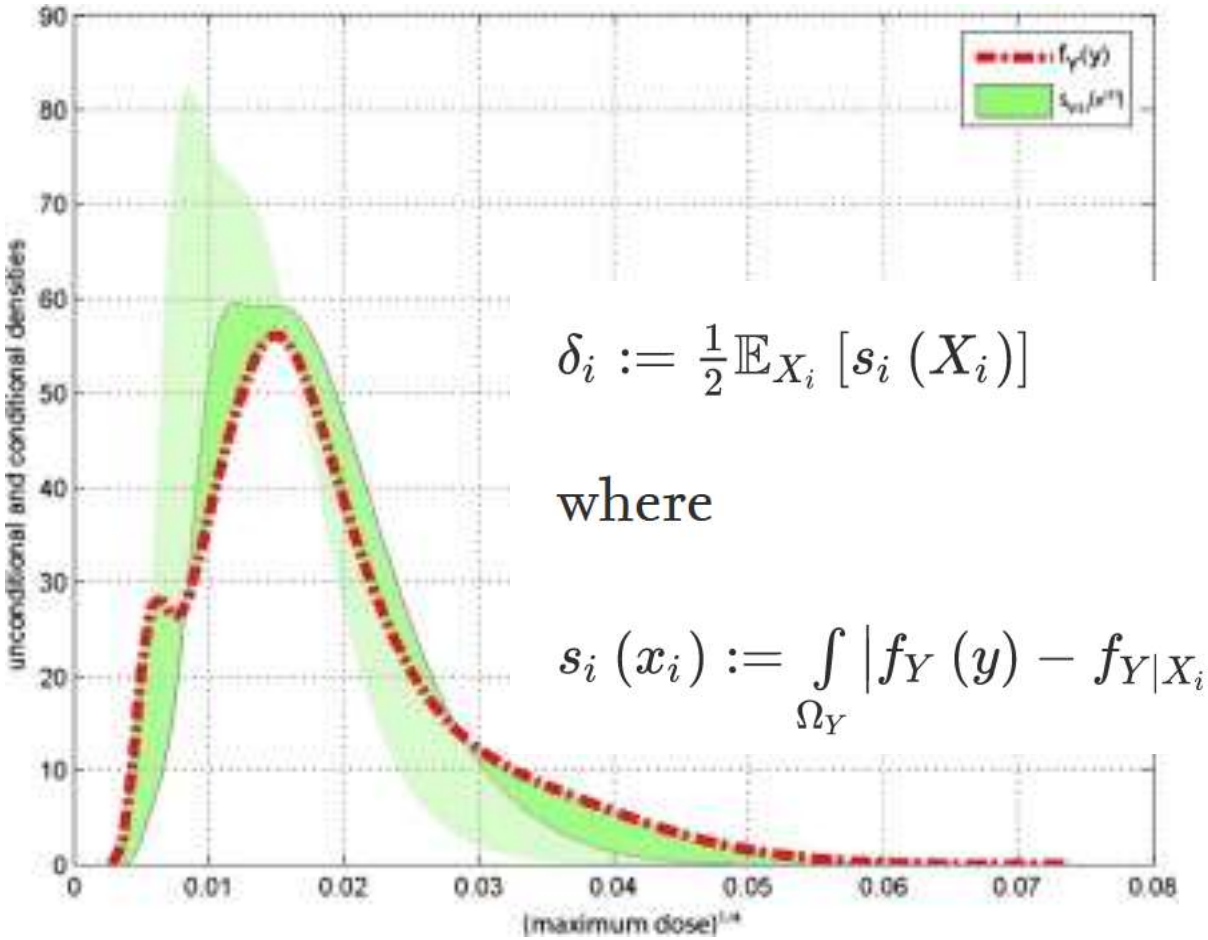
Environmental Modelling & Software

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

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

ELSEVIER

Environmental Modelling & Software

Volume 34, June 2012, Pages 105-115

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

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

Don’t use One factor At a Time (OAT)

A geometric proof

Environmental Modelling & Software 25 (2010) 1508–1517

Contents lists available at ScienceDirect

Environmental Modelling & Software

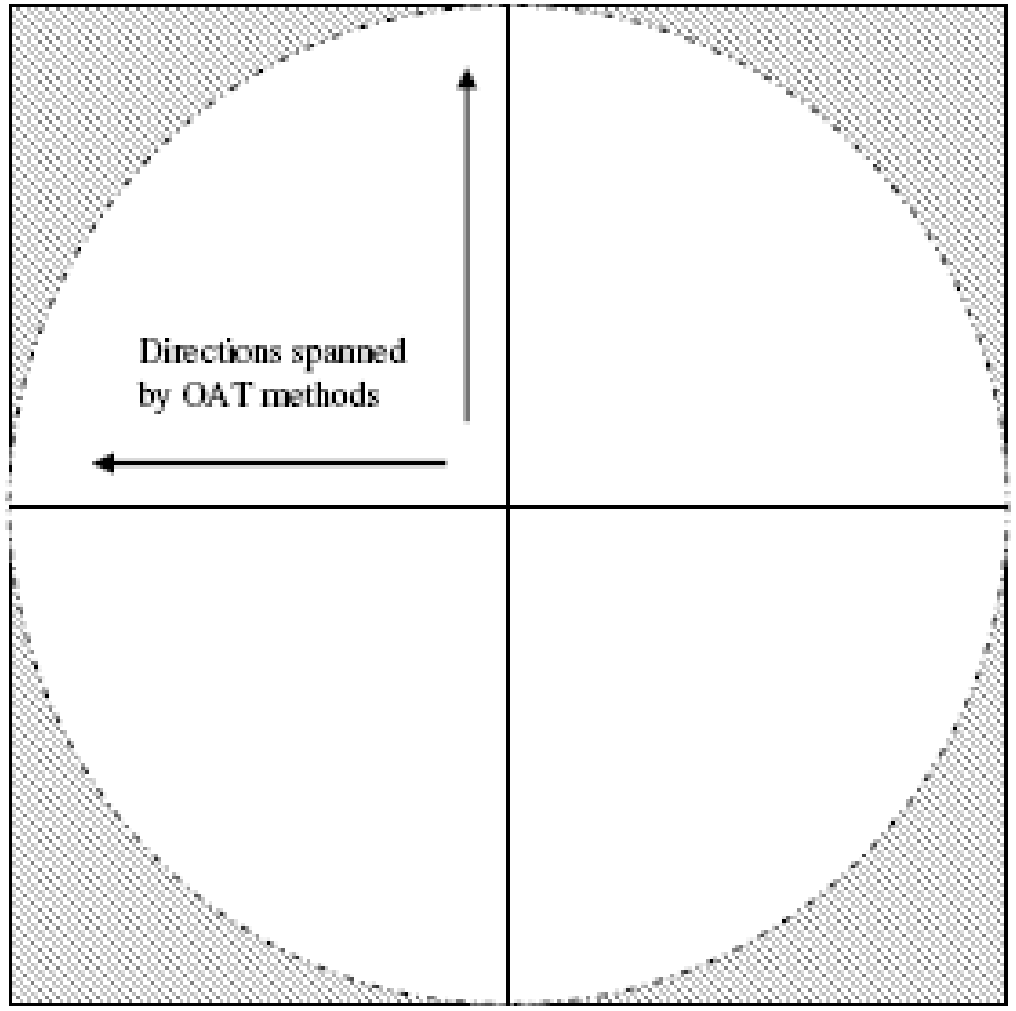
journal homepage: www.elsevier.com/locate/envsoft

How to avoid a perfunctory sensitivity analysis

Andrea Saltelli*, Paola Annoni

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

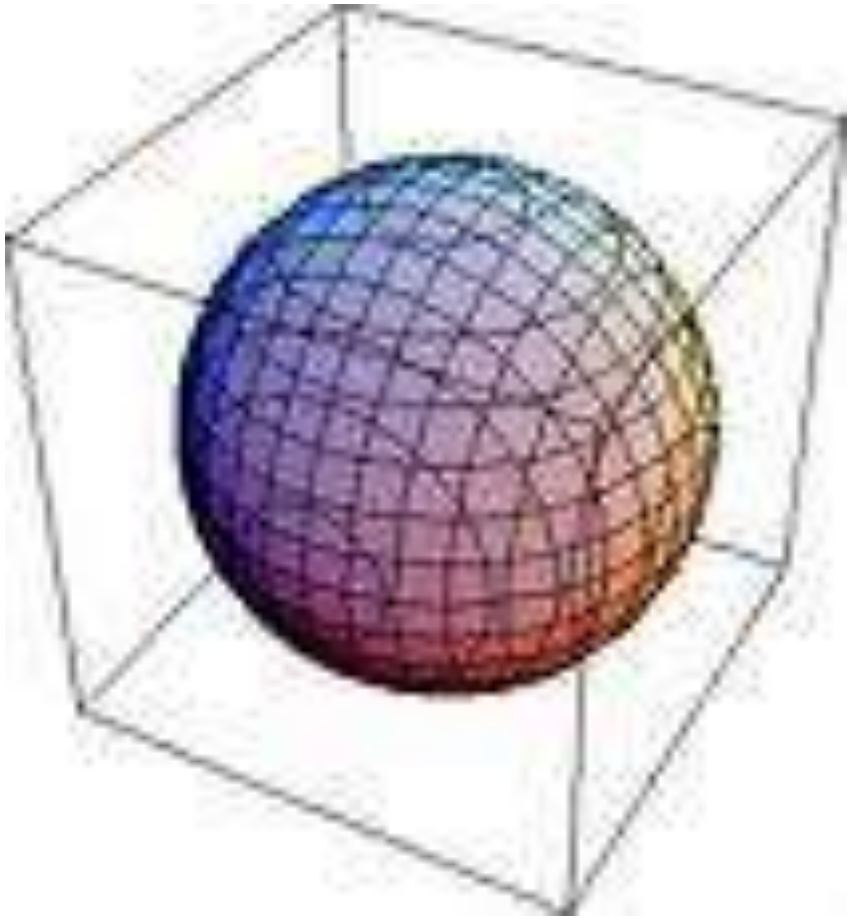
OAT in 2 dimensions



Area circle / area square =?

~ 3/4

OAT in 3 dimensions

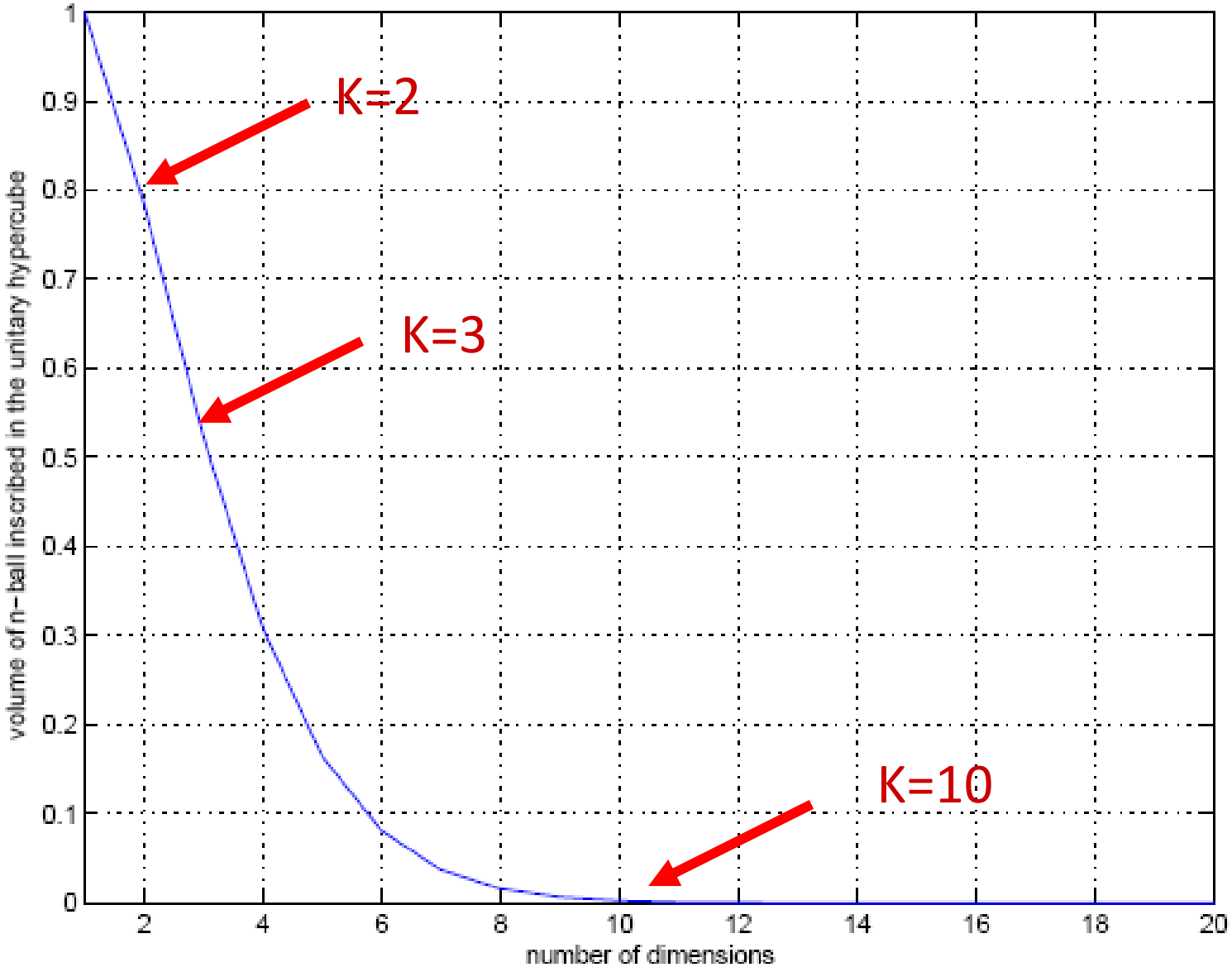


Volume sphere / volume cube =?

~ 1/2

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

OAT in k dimensions

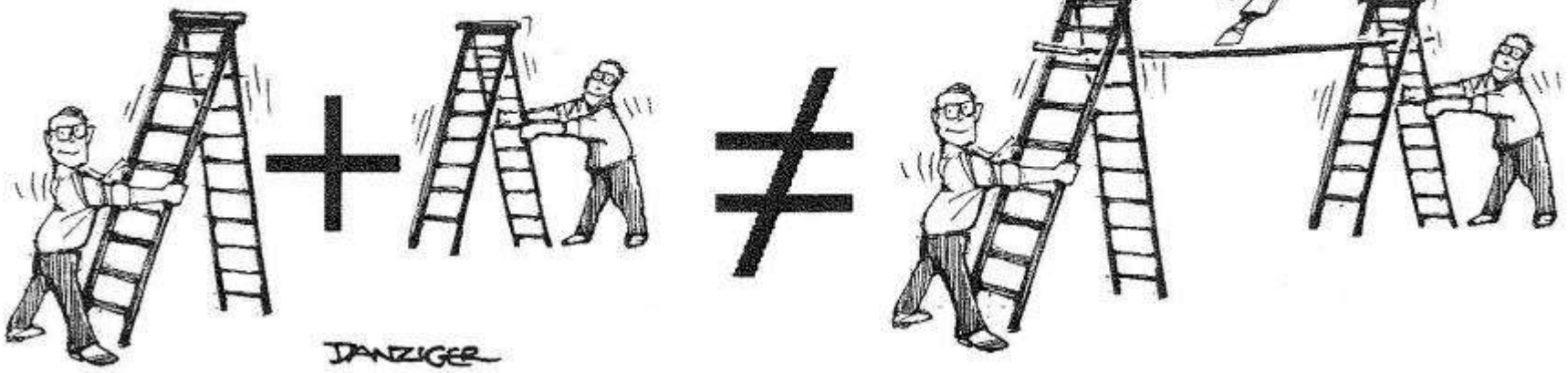


OAT does not capture interactions

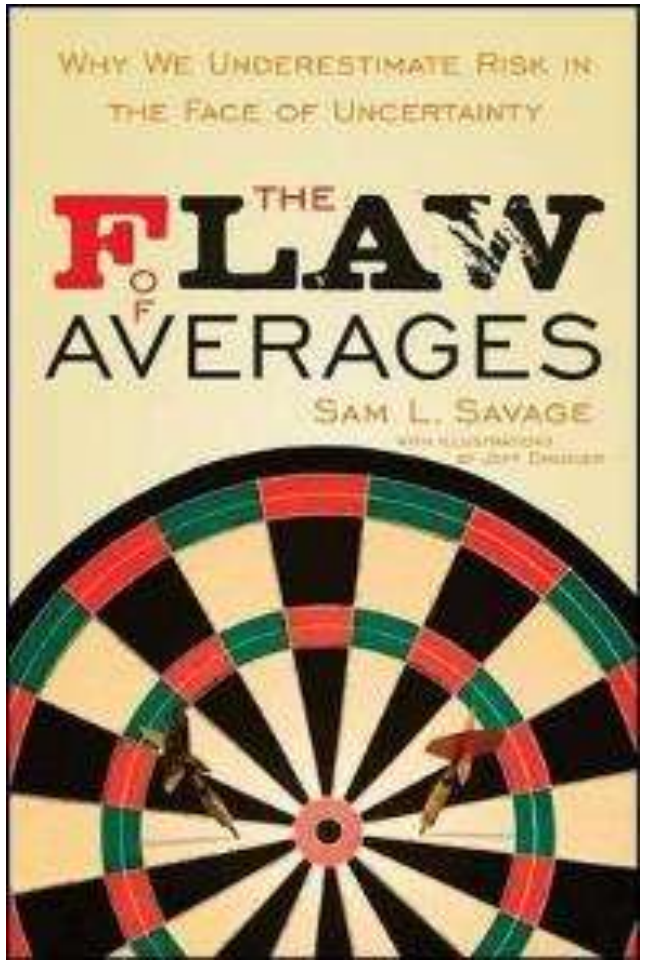
➔ The resulting analysis is non conservative

How would you test the scaffolding?

How coupled ladders are shaken in most of available literature



How to shake coupled ladders



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

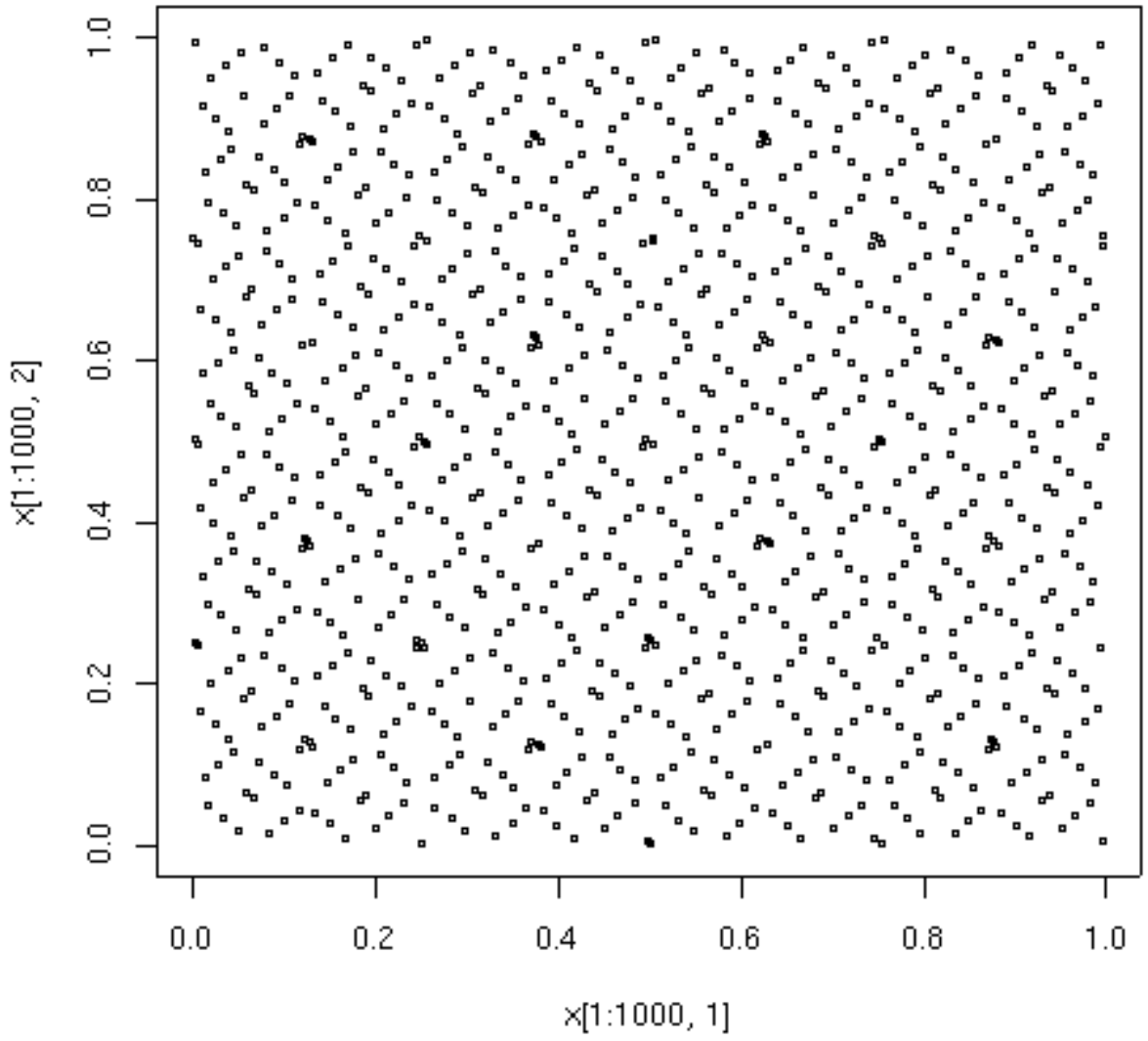
➔They assume linearity (PCC) or monotonicity (PRCC), which is difficult to know *ex-ante*

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

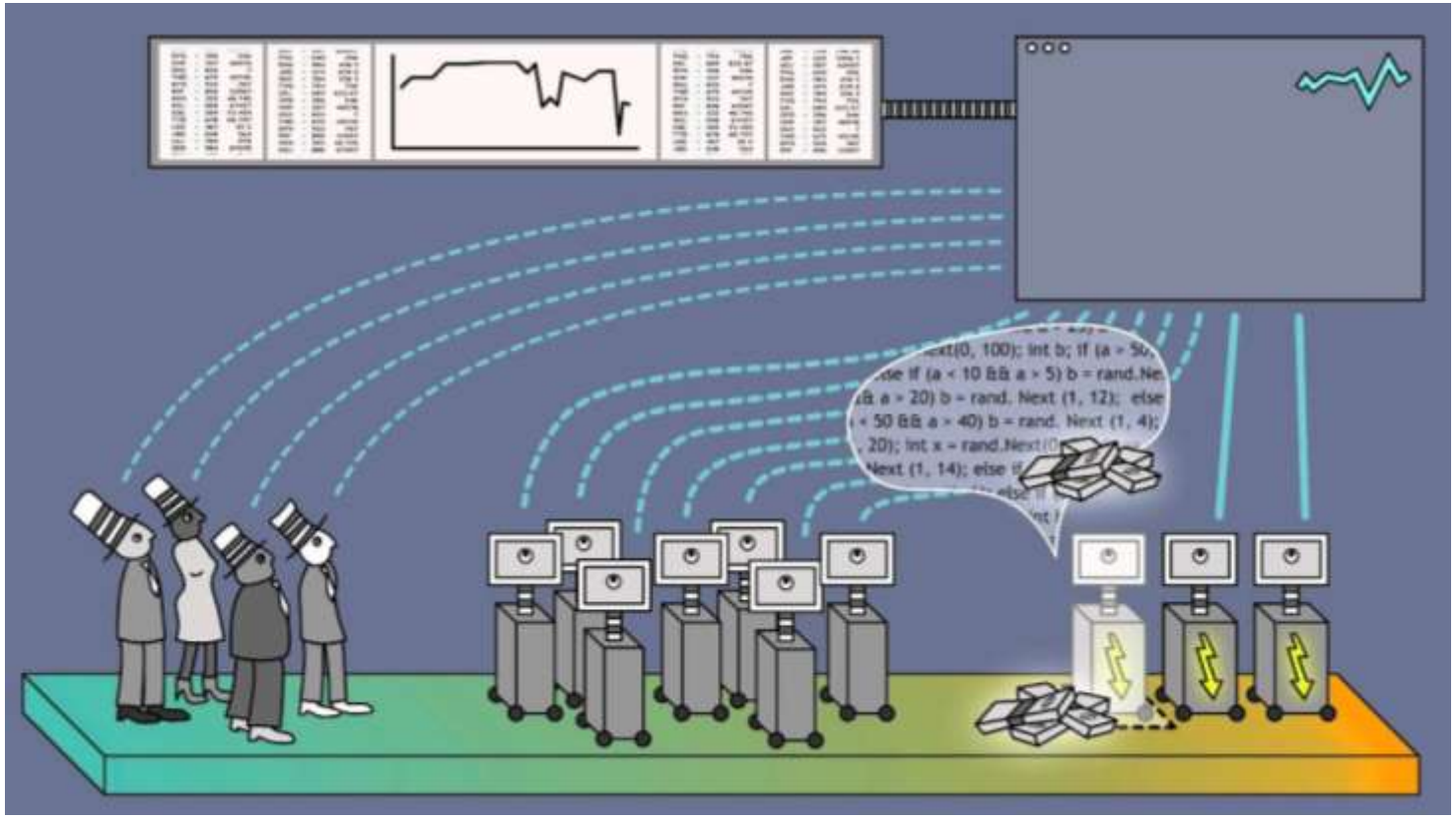
Statistics > Applications

[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!

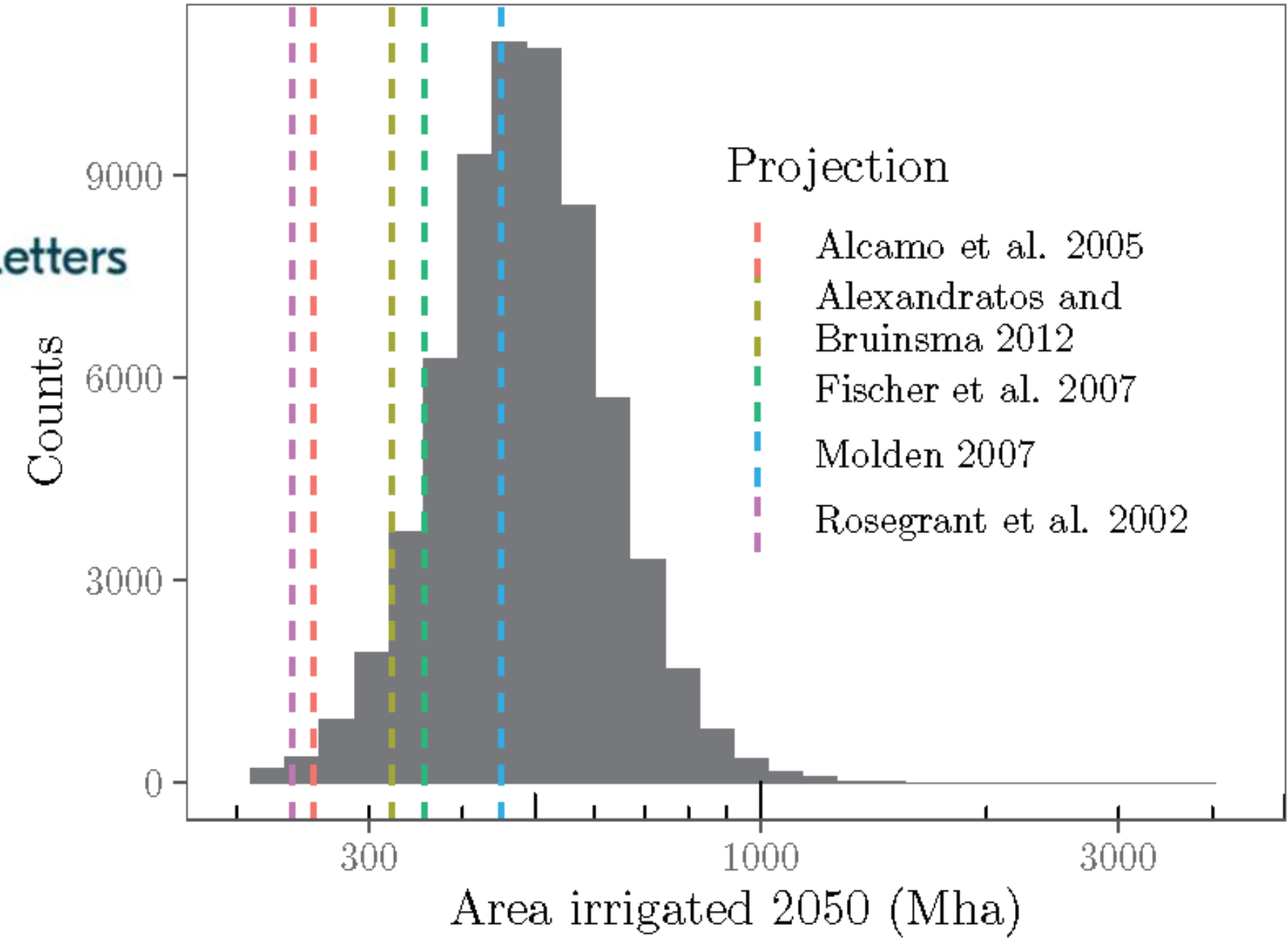
Model routinely used to produce point estimates may becomes non conservative when the uncertainty is plugged in

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

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

Solution? Modelling of the modelling process

Jorge Luis Borges
(1899–1986)

Ts'ui Pên
(character of the novel)

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‡

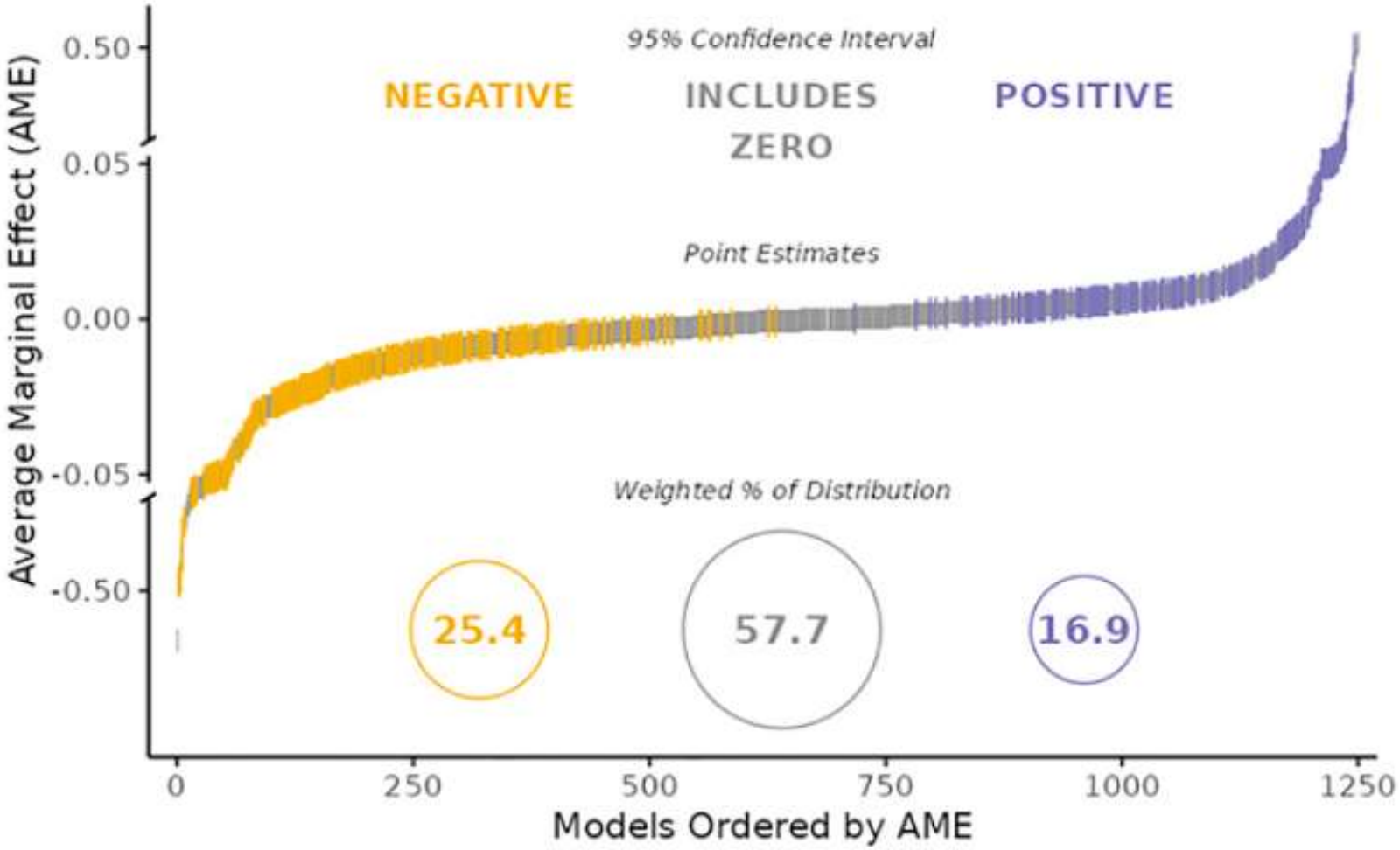
14 Nov 2013

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?

...

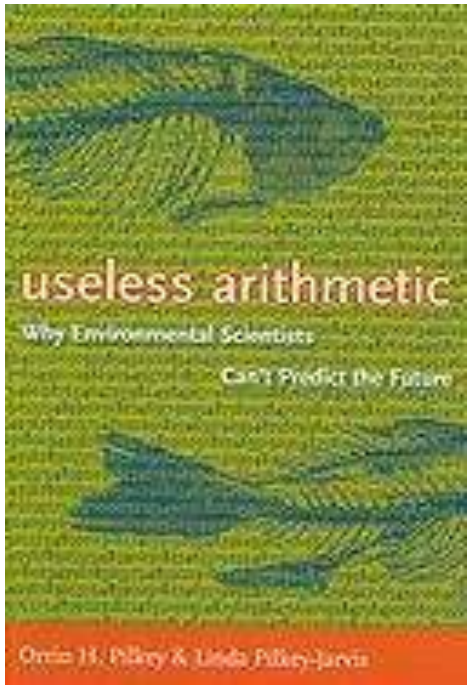
...teams' results varied greatly, ranging from large negative to large positive effects"

(Massey et al. 2022)

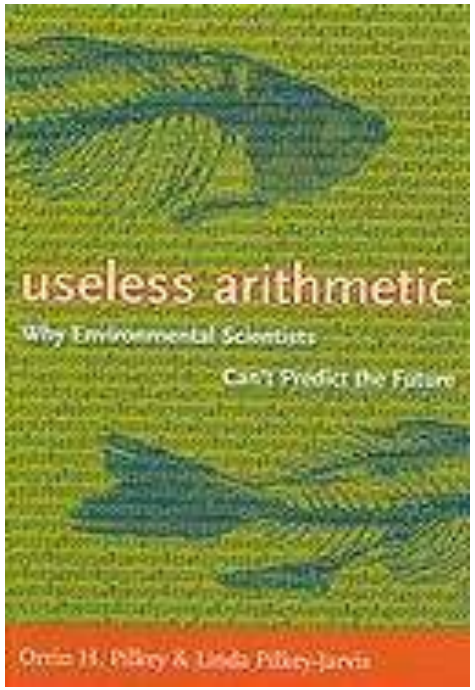
Don’t confuse the map with the territory

If you do, sensitivity analysis will not save you

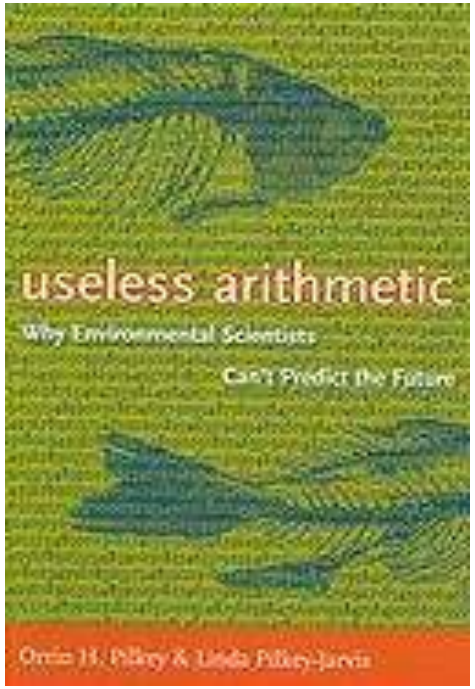
Orrin H.
Pilkey



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



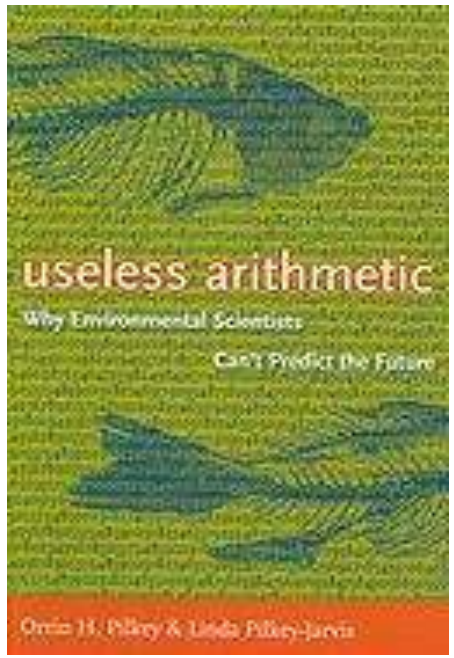
<<It is important, however, to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature>>

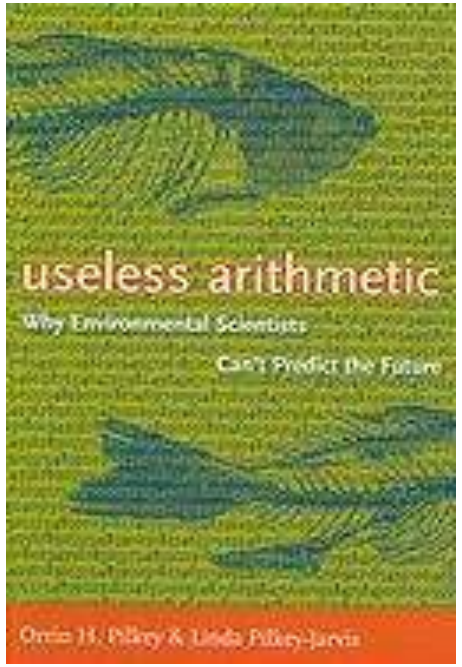


<<…If the model is wrong or if it is a poor representation of reality, determining the sensitivity of an individual parameter in the model is a meaningless pursuit>>

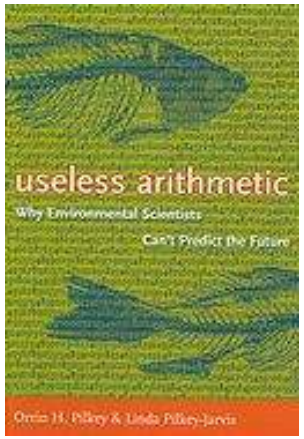
One of the examples discussed concerns the Yucca Mountain repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis.

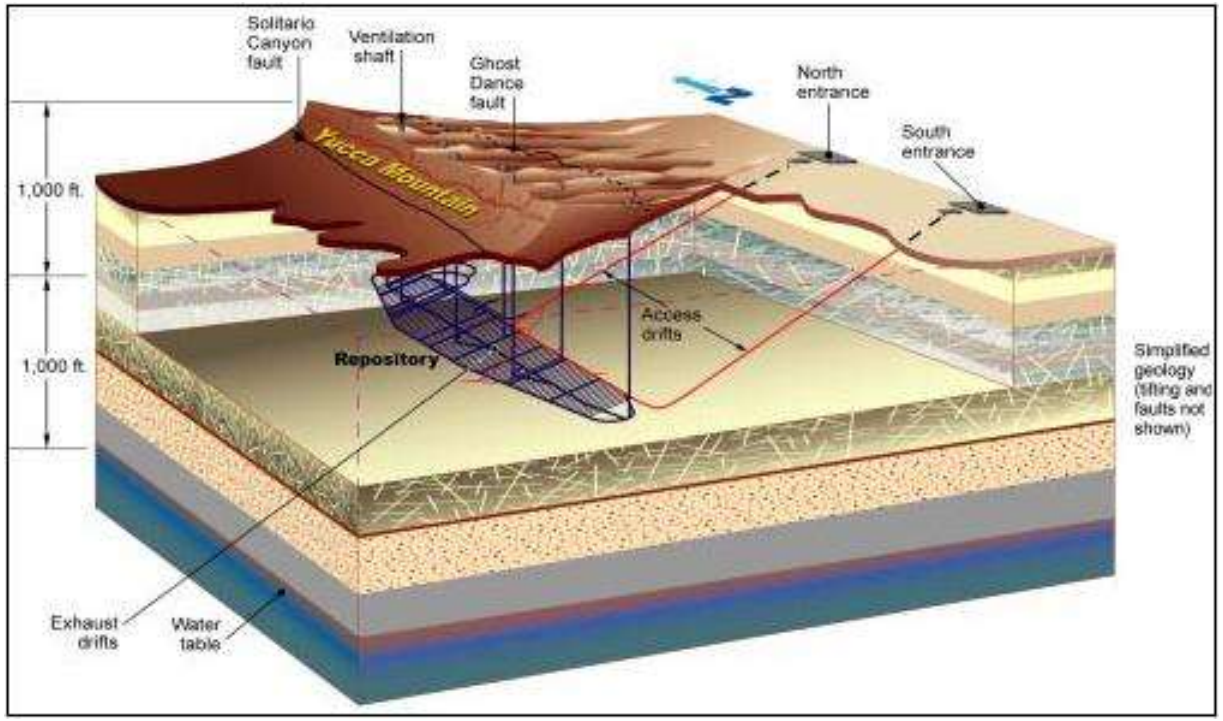
TSPA is Composed of 286 sub-models.





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.





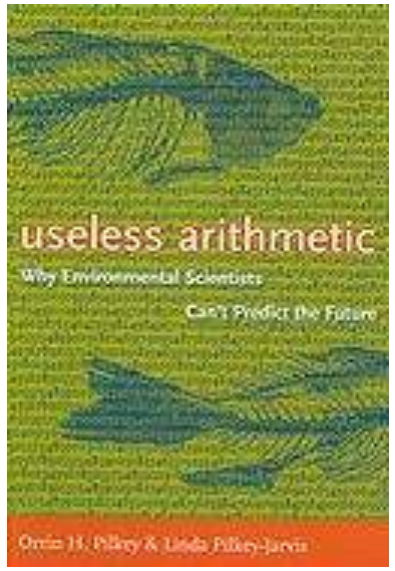
The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter

(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.

Displacement: “The model we have developed tells us that real progress is being achieved” (The focus in now the model not the problem).

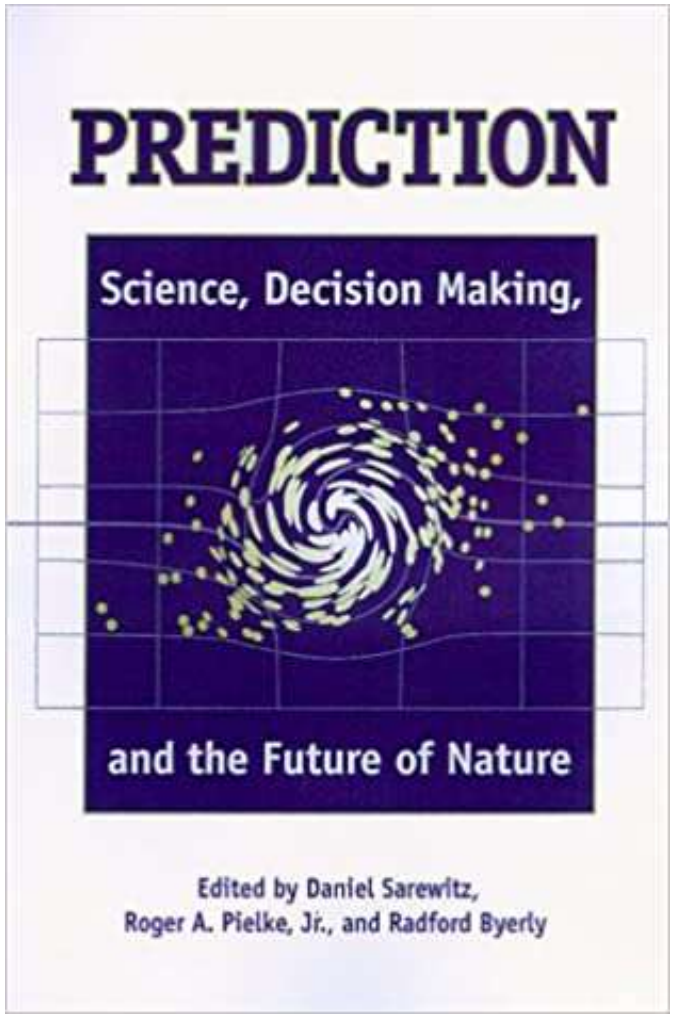
Rayner, S., 2012, Uncomfortable knowledge: the social construction of ignorance in science and environmental policy discourses, Economy and Society, 41:1, 107-125.

Example of displacement: Chesapeake Bay Program (CBP) modelling work

"Bay models are used to track nutrient loads to ensure the cap is not exceeded"

➔The model results – rather than the actual measurements, become the substance of use

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



Model GENESIS for beach erosion

US Army Corps of Engineers ®

Manipulated to support coastal-engineering projects

It neglected the role of extreme event

Sarewitz, D., Pielke, R. A. & Byerly, R. *Prediction: Science, Decision Making, and the Future of Nature* (Island Press, 2000).

Beware the size of your model

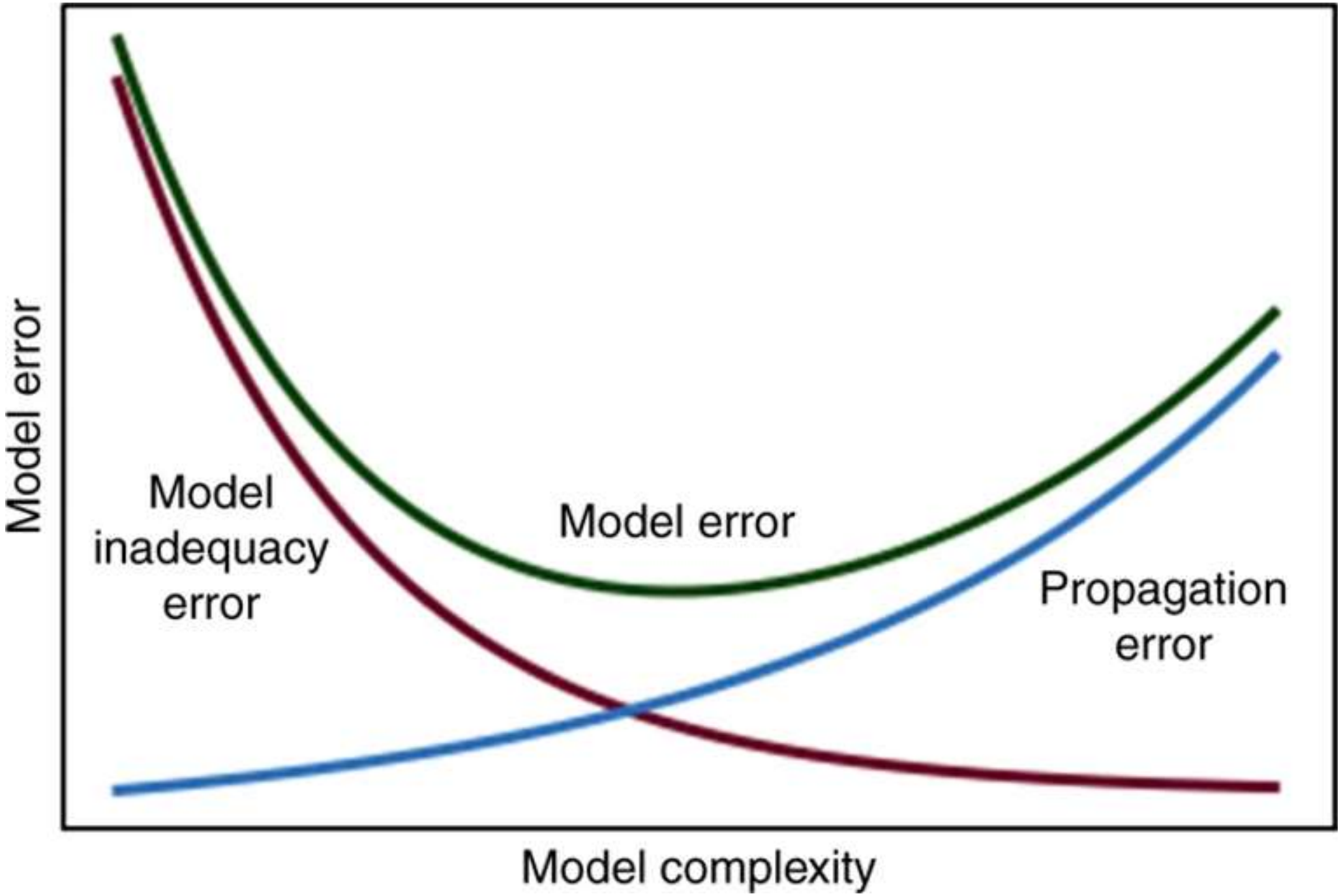
Mind the conjecture of O'Neil

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

ScienceAdvances

Current Issue First release papers Archive About Su

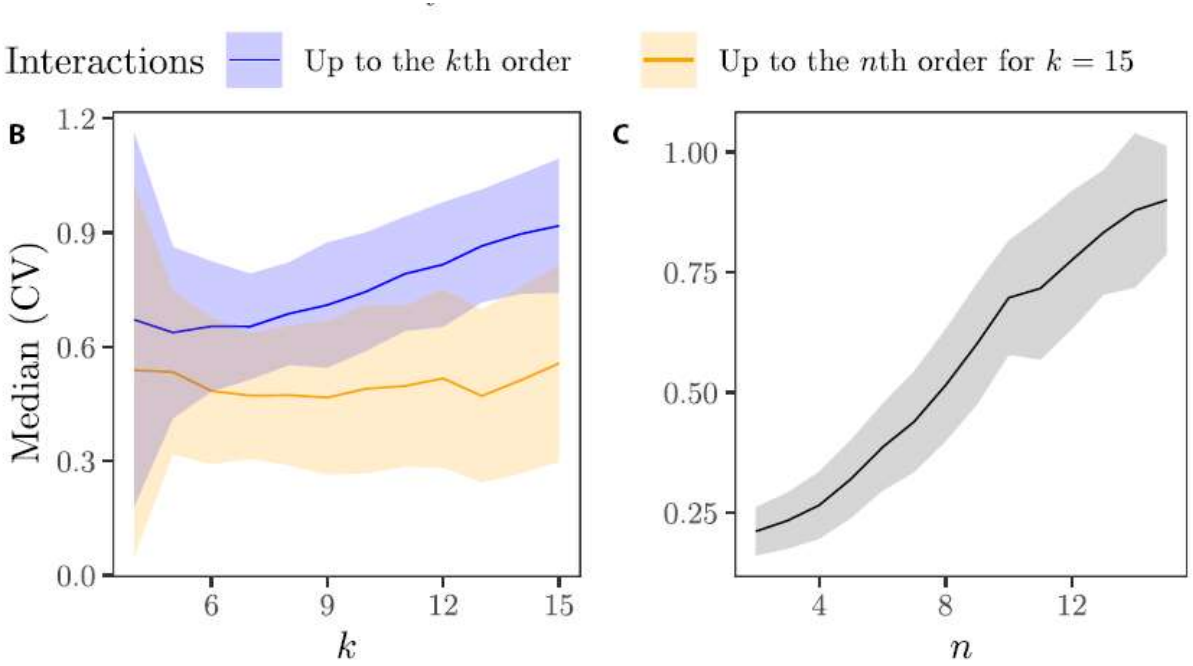
HOME > SCIENCE ADVANCES > VOL. 8, NO. 42 > MODELS WITH HIGHER EFFECTIVE DIMENSIONS TEND TO PRODUCE MORE UNCERTAIN ESTIMATES

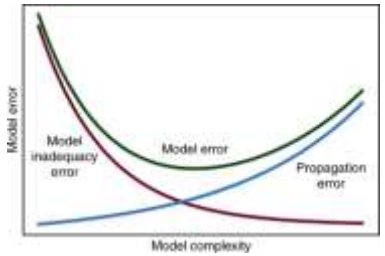
RESEARCH ARTICLE | MATHEMATICS

Models with higher effective dimensions tend to produce more uncertain estimates

ARNALD PUY, PIERFRANCESCO BENEVENTANO, SIMON A. LEVIN, SAMUELE LO PIANO, TOMMASO PORTALURI, AND ANDREA SALTELLI

SCIENCE ADVANCES · 19 Oct 2022 · Vol 8, Issue 42 · DOI: 10.1126/sciadv.abn9450





Empirical test using the SA-based concept of effective dimension

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.

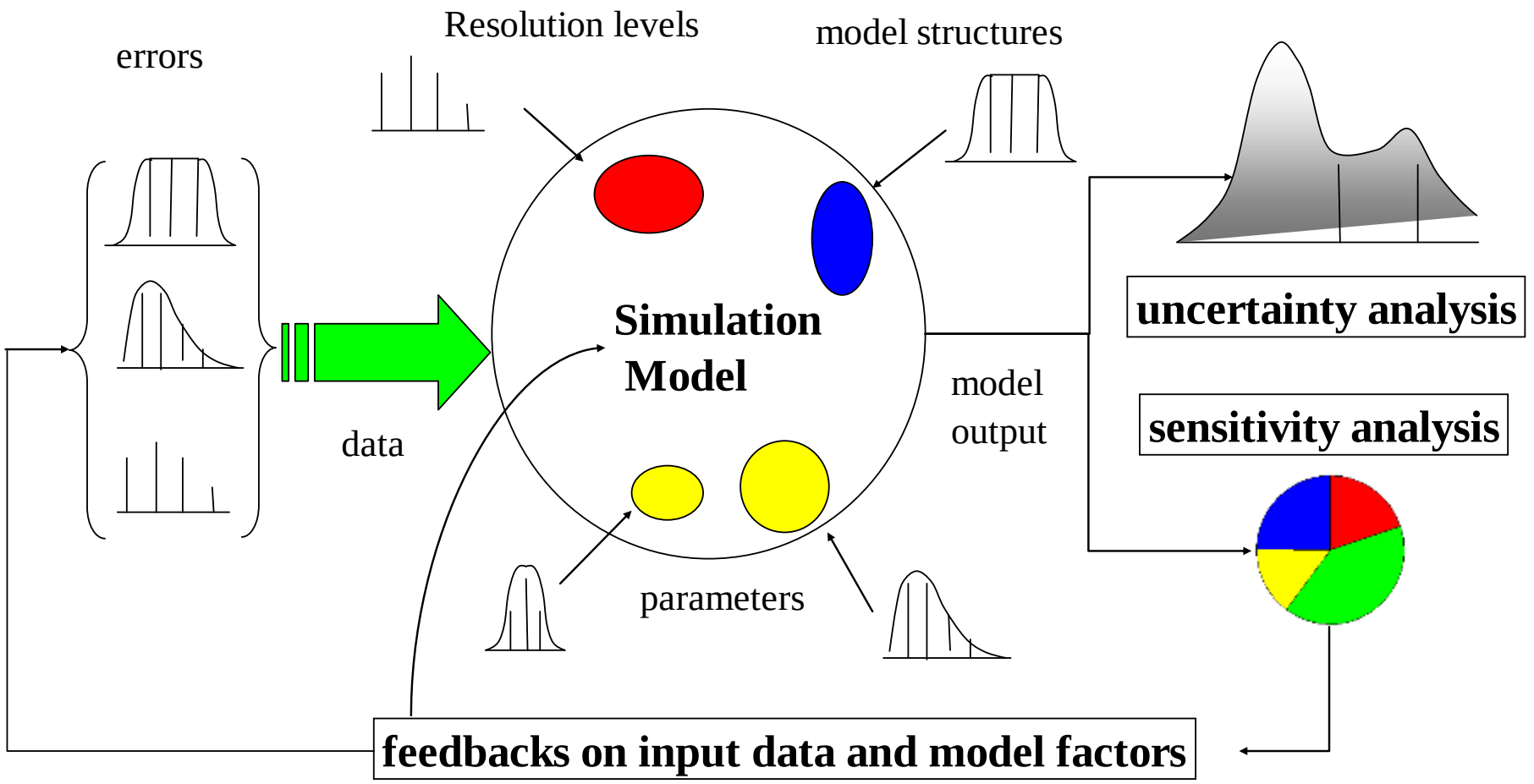
L. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," IEEE Trans. Syst. Man. Cybern., vol. 3, no. 1, pp. 28–44, 1973.

Don’t sample just parameters and boundary conditions

Explore thoroughly the space of the assumptions

An engineer's vision of UA, SA

Remember?

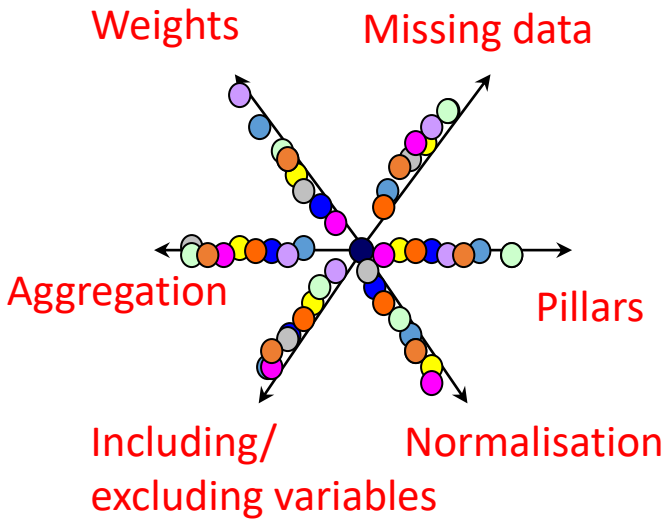


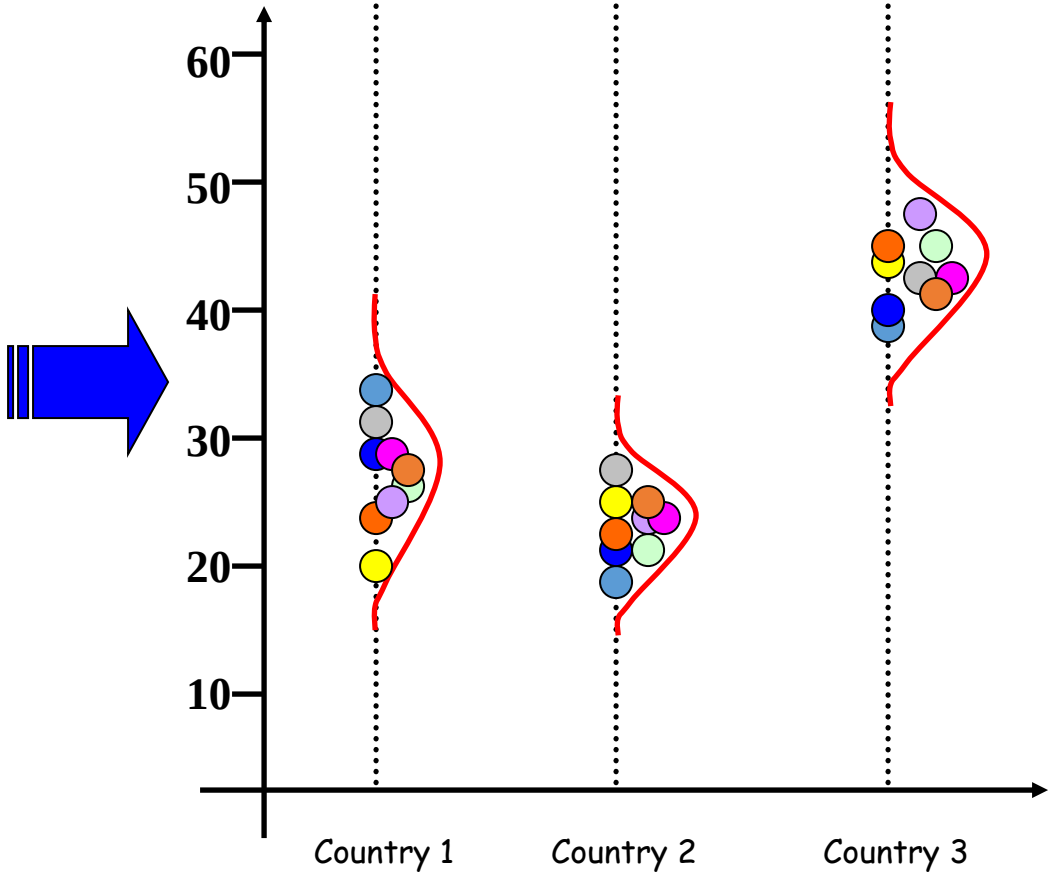
One can sample more than just factors:

- 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



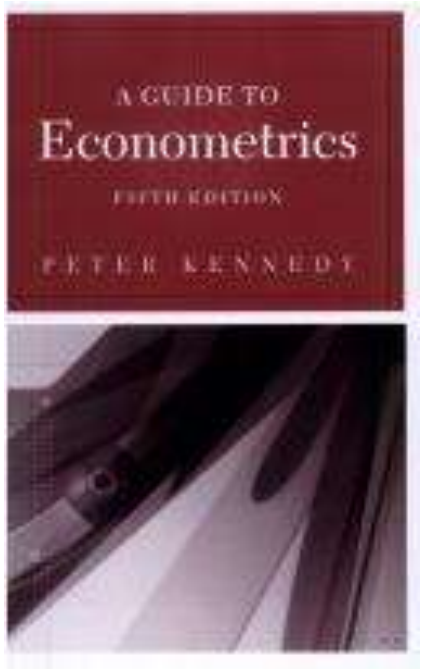


Don’t go public with your results without having seen your SA

Find SA before SA finds you

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%

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

… beside uncertainty can be used instrumentally

The End