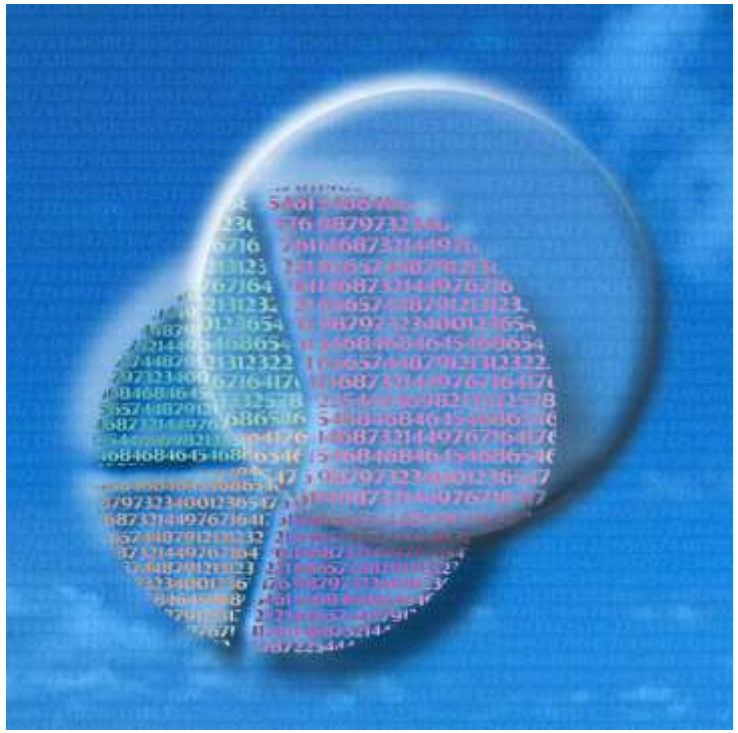


How not to do a sensitivity analysis

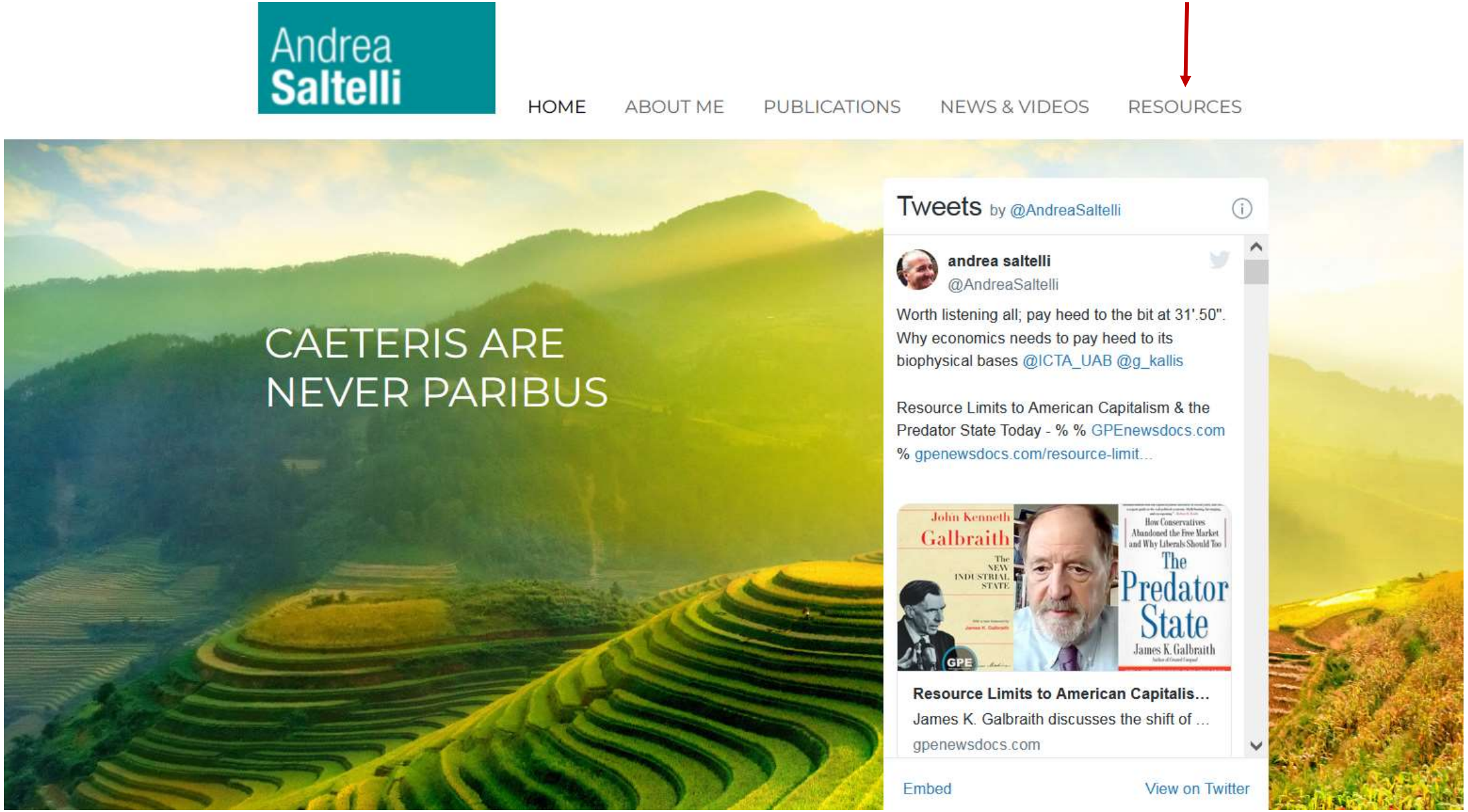
Andrea Saltelli
University College London,
Institute for Innovation and Public Purpose,
March 2nd 2022

upf. BARCELONA SCHOOL OF MANAGEMENT

CNR ISTC ISTITUTO DI SCIENZE E TECNOLOGIE DELLA COGNIZIONE

UNIVERSITAS BERGENSIS

Where to find this talk: www.andreasaltelli.eu

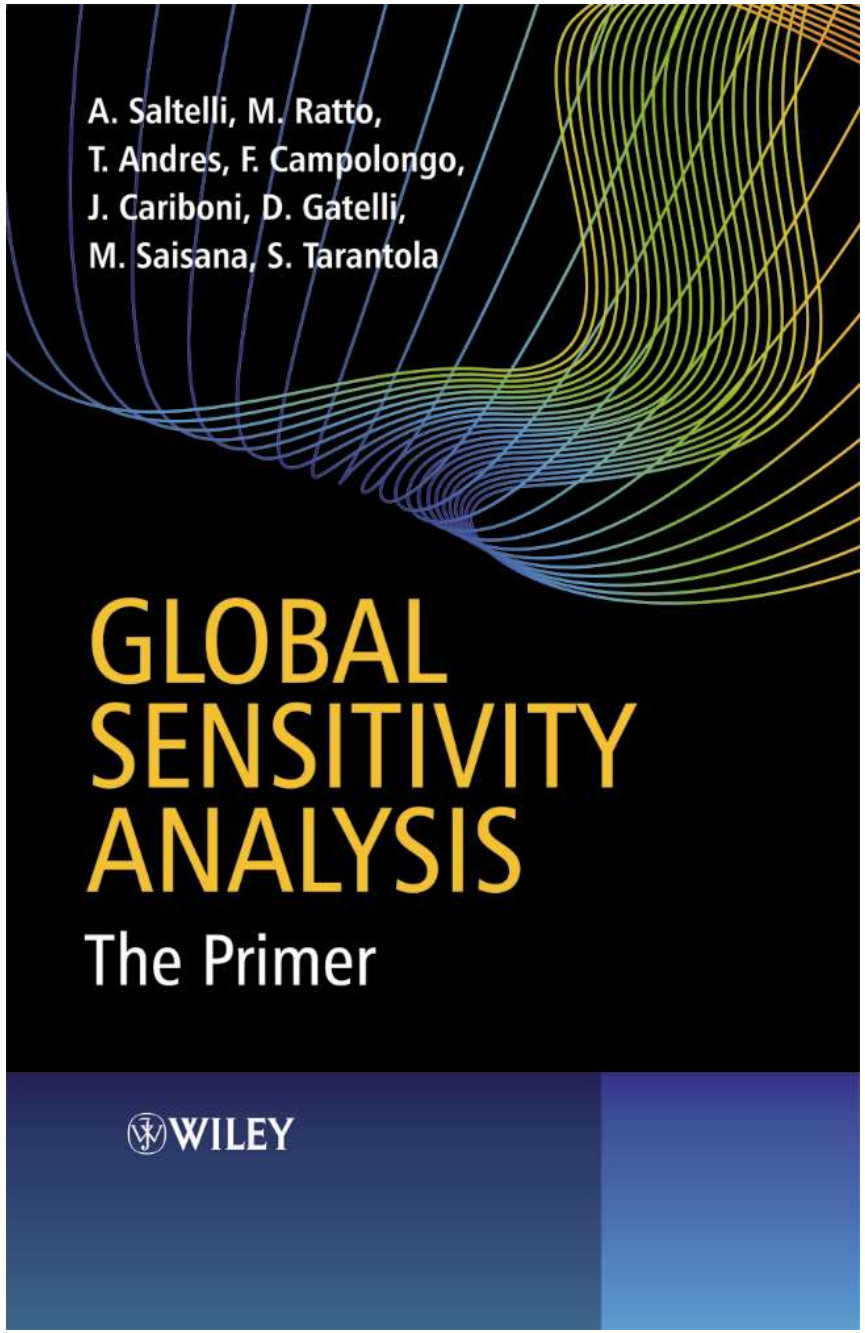


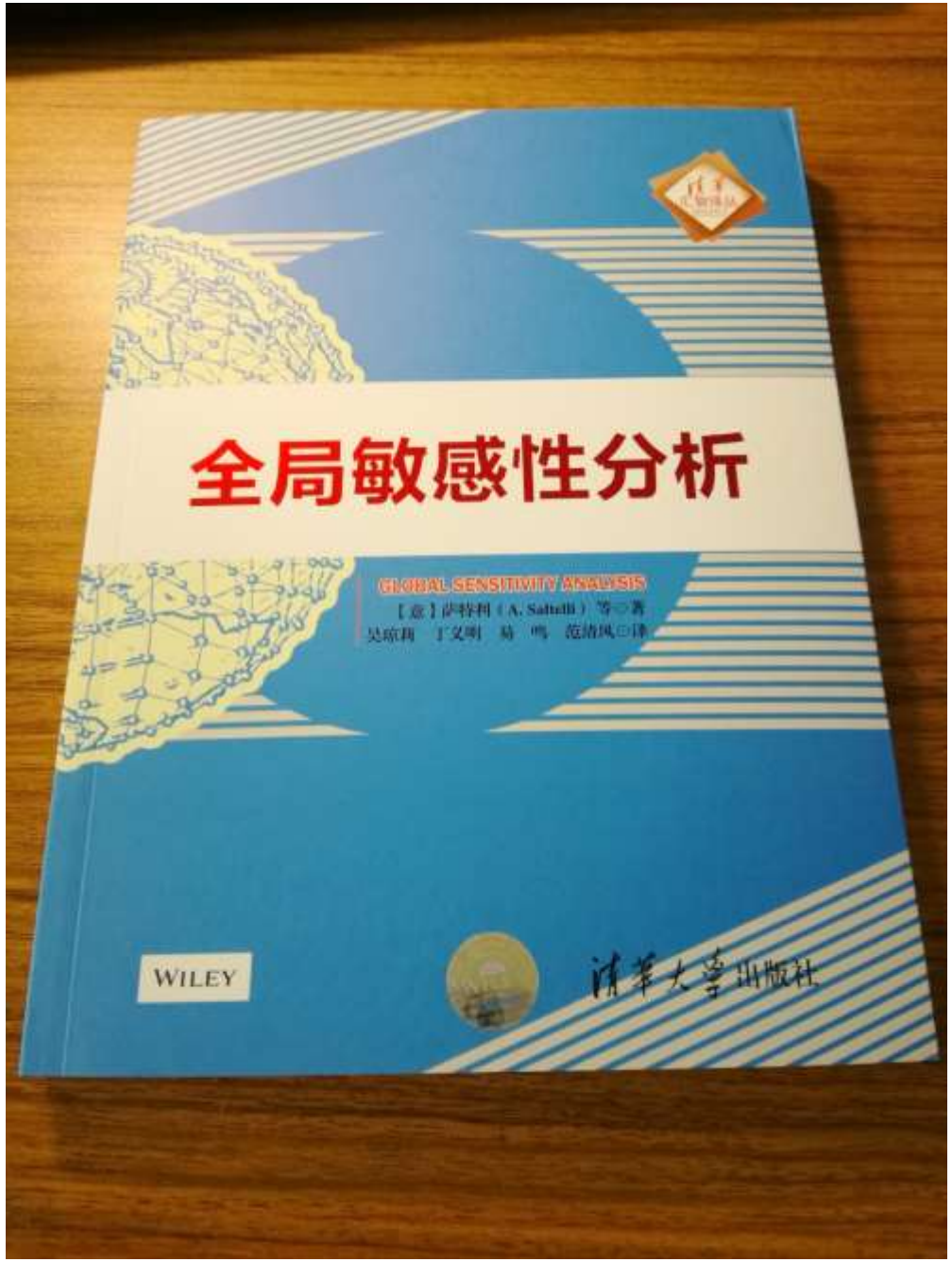
- Don’t use just any method
- Don’t use One factor At a Time (OAT)
- Don’t use method that are not model-independent
- Don’t use either LHS or optimized LHS
- Don’t run the model just once
- Don’t use Morris’ method
- Don’t confuse the map with the territory
- Beware the dimension of your model
- Don’t sample just parameters and boundary conditions
- Don’t go public with your results without having seen your SA
- NEVER vary all factors of the same amount (5%, 10%, 20%)

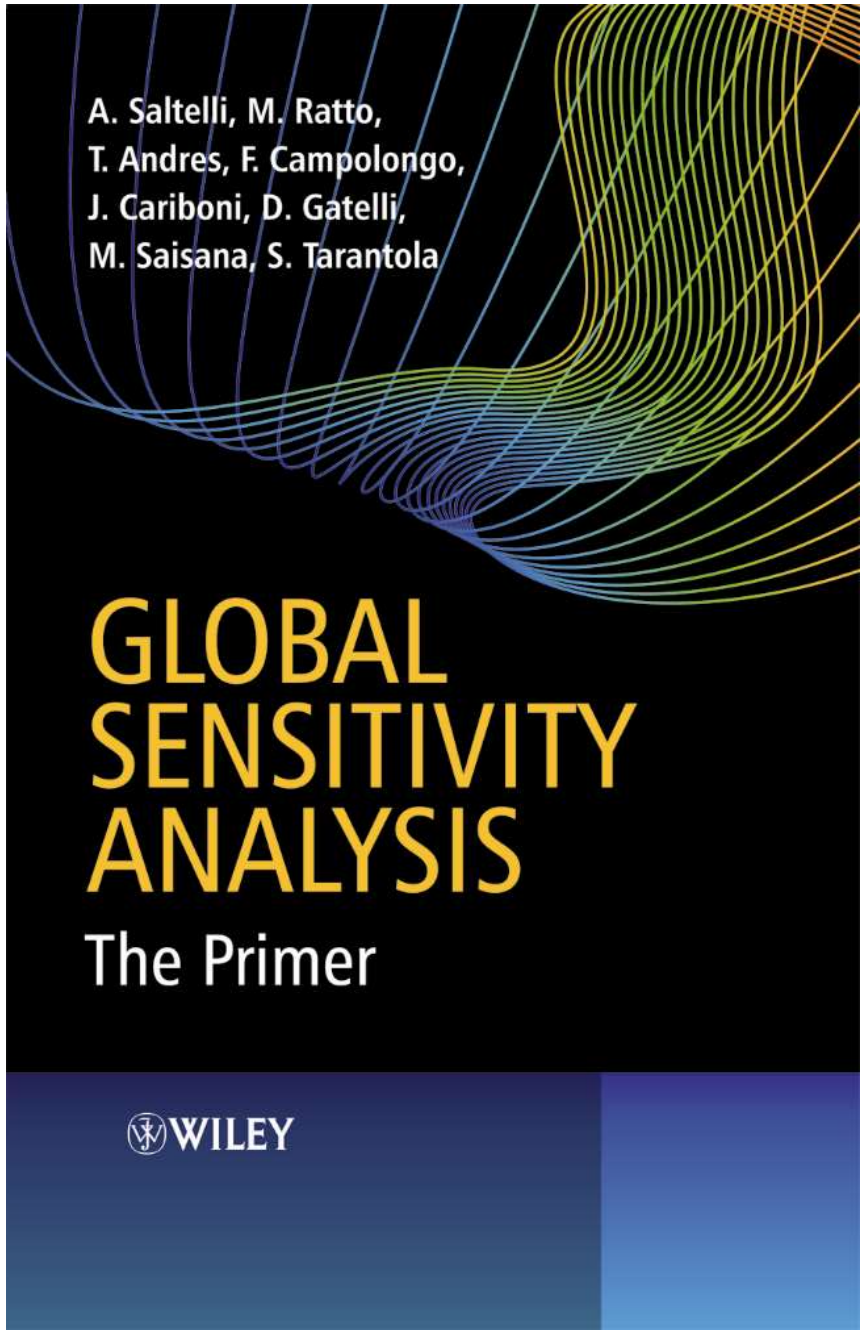
Don’t use just any method

Use the method appropriate to context and purpose

An introduction to variance based methods

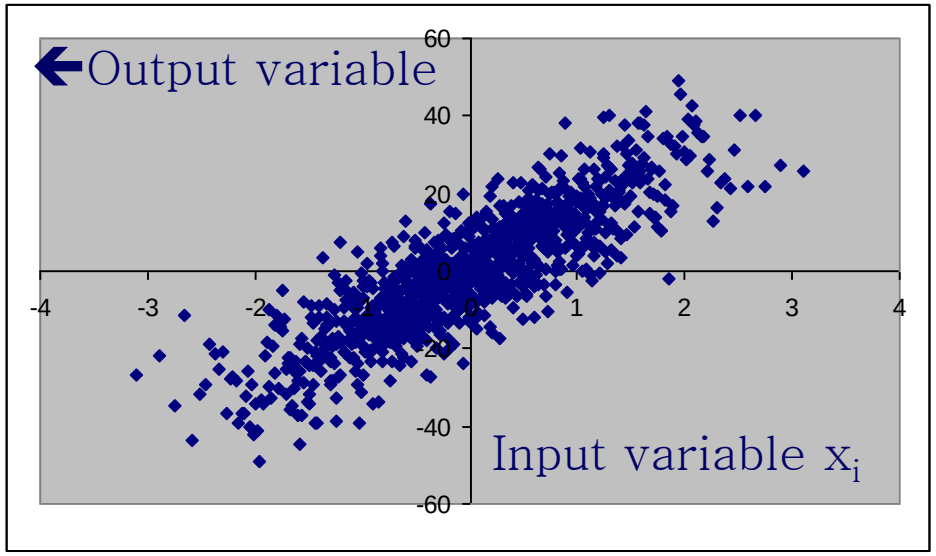

A. Saltelli, M. Ratto,
T. Andres, F. Campolongo,
J. Cariboni, D. Gatelli,
M. Saisana, S. Tarantola
GLOBAL
SENSITIVITY
ANALYSIS
The Primer
WILEY

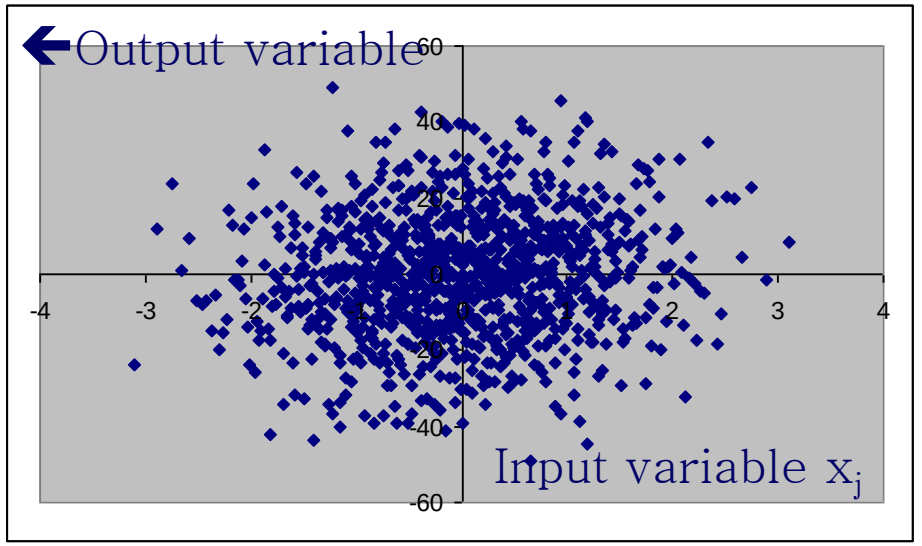

全局敏感性分析
GLOBAL SENSITIVITY ANALYSIS
WILEY



Available for free at

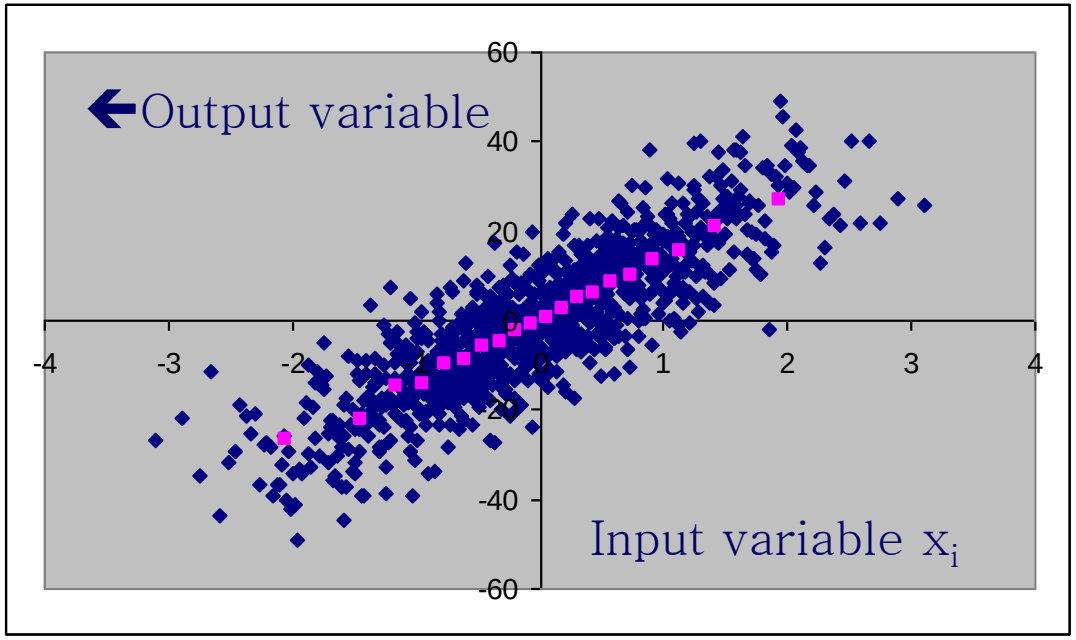
http://www.andreasaltelli.eu

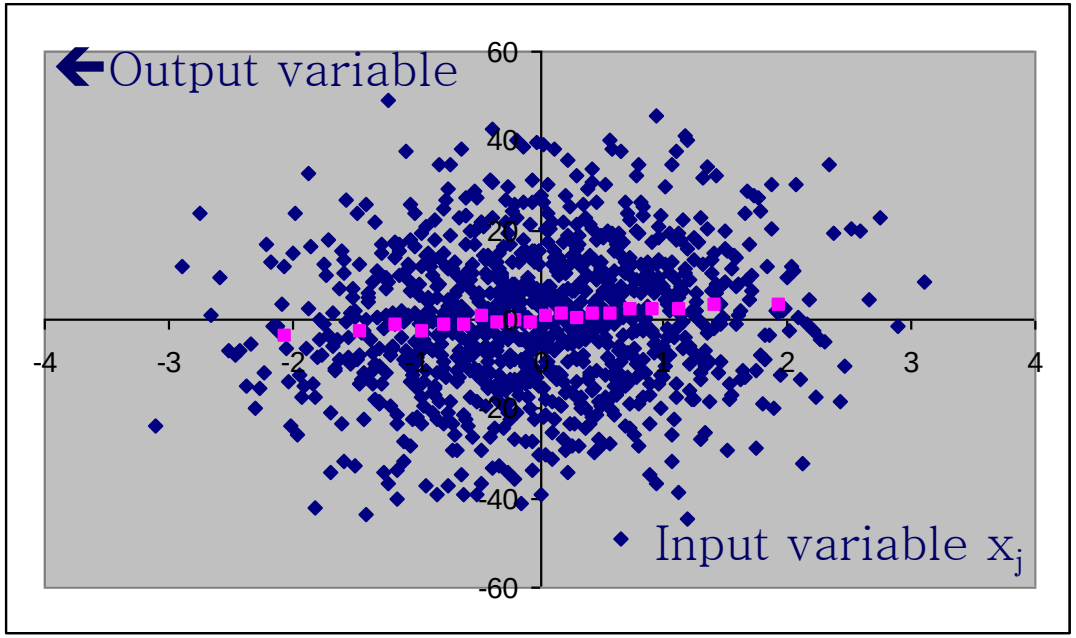




Plotting the output as a function of two different input factors

Which factor is more important?

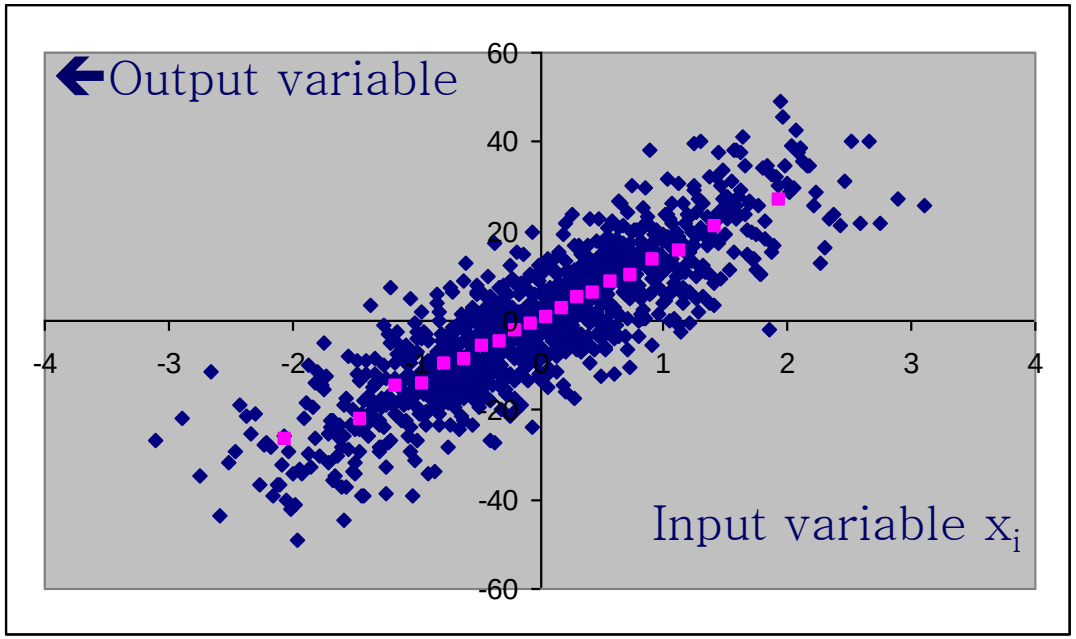




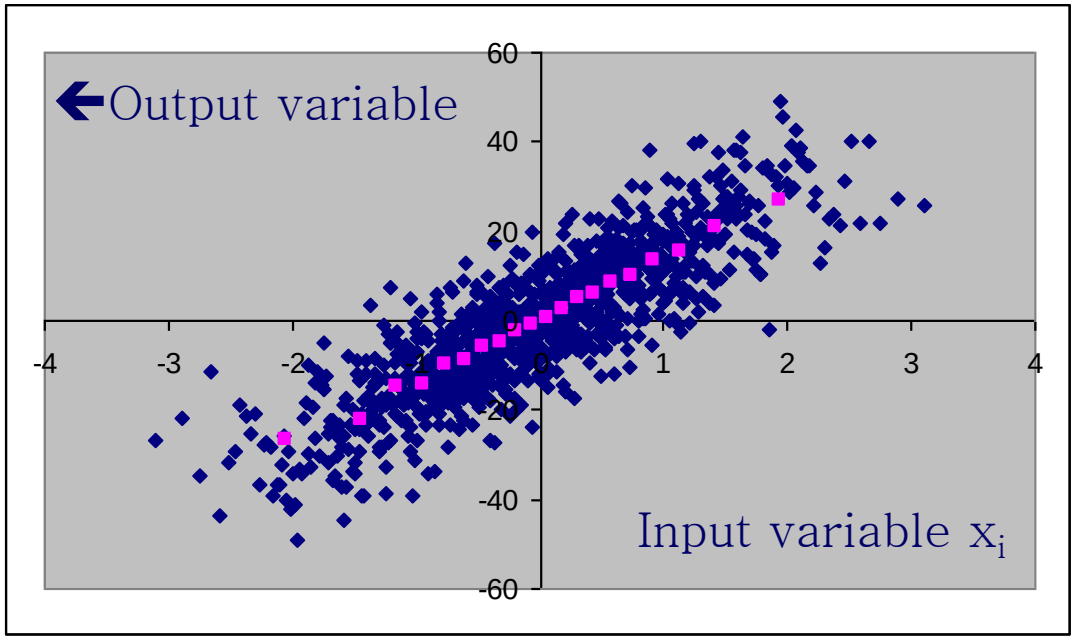
~1,000 blue points

Divide them in 20 bins of ~ 50 points

Compute the bin's average (pink dots)

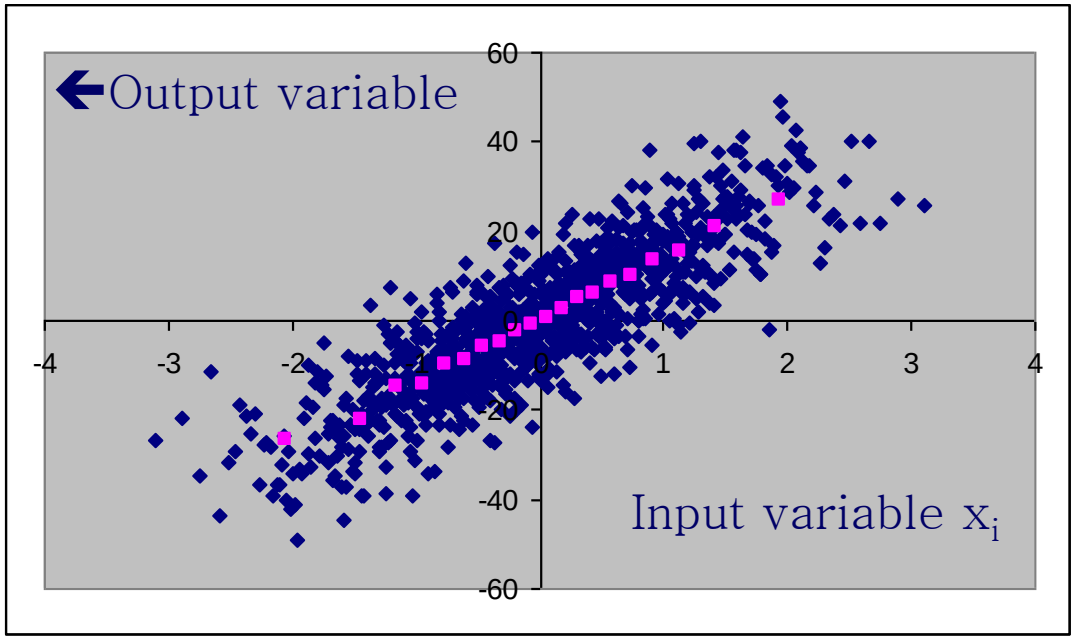


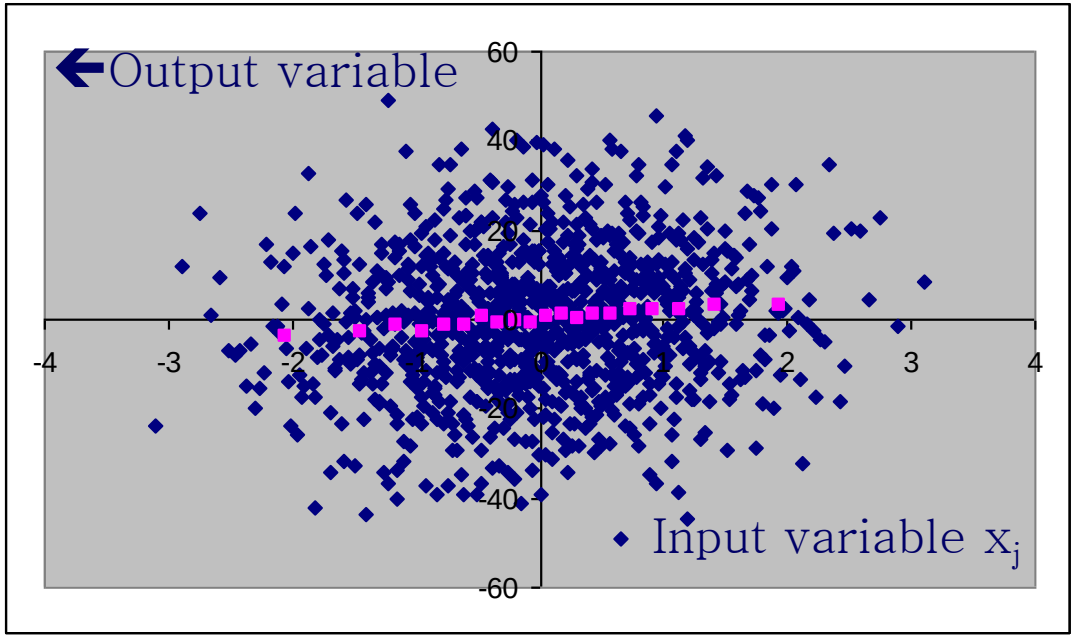
Each pink point is ~ $E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)$



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

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

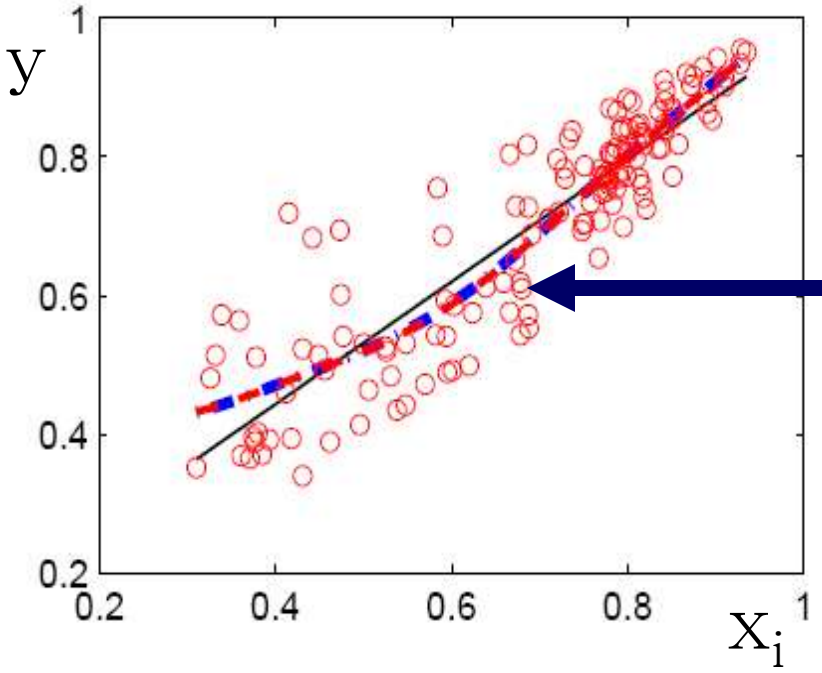
For additive models one can decompose the total variance as a sum of those partial variances

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

… 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

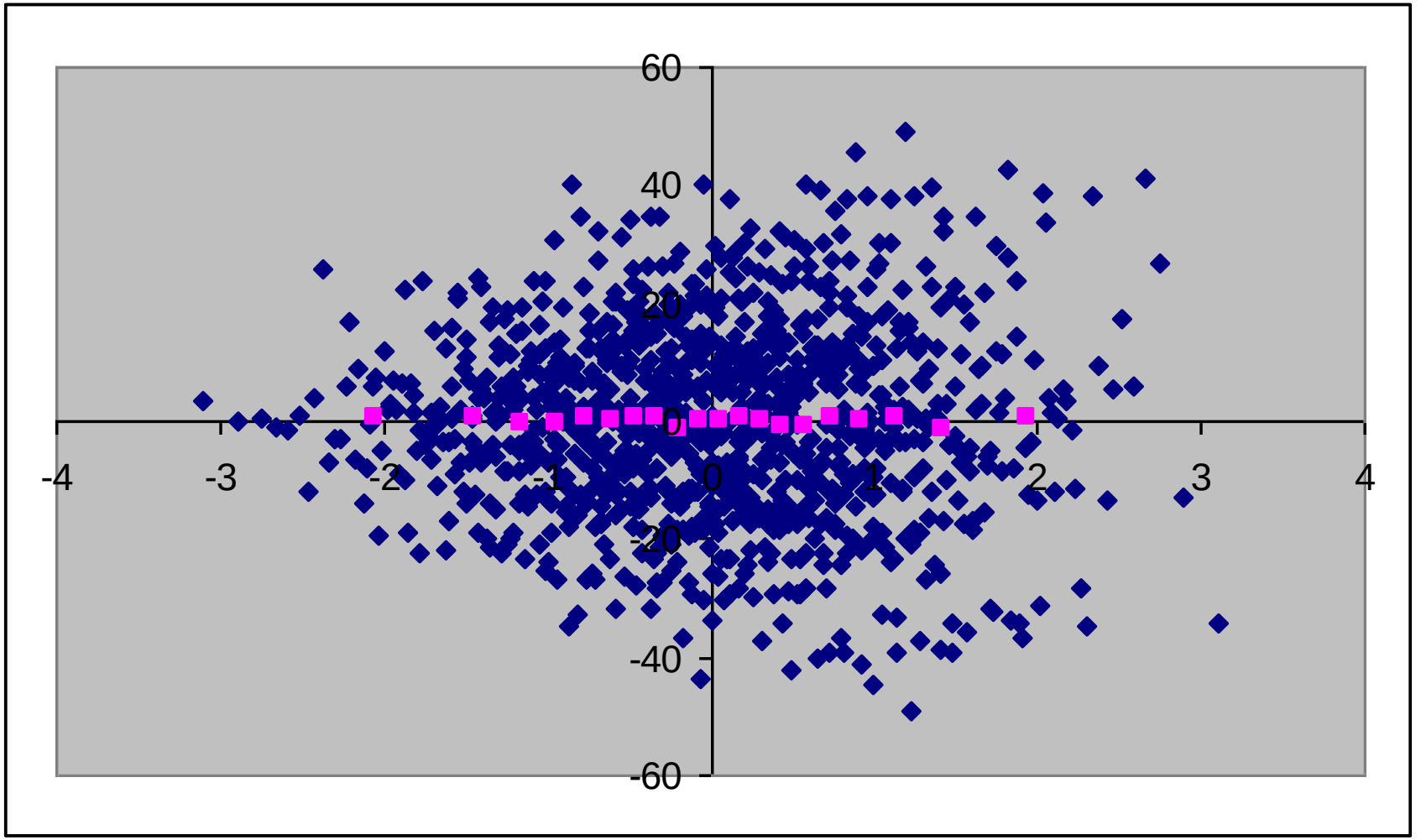
$$S_i \equiv \eta_i^2 := \frac{\mathrm{V}_{x_i}\left(\mathrm{E}_{\mathbf{x}_{\sim i}}\left(y \mid x_i\right)\right)}{\mathrm{V}\left(y\right)}$$

First order sensitivity index

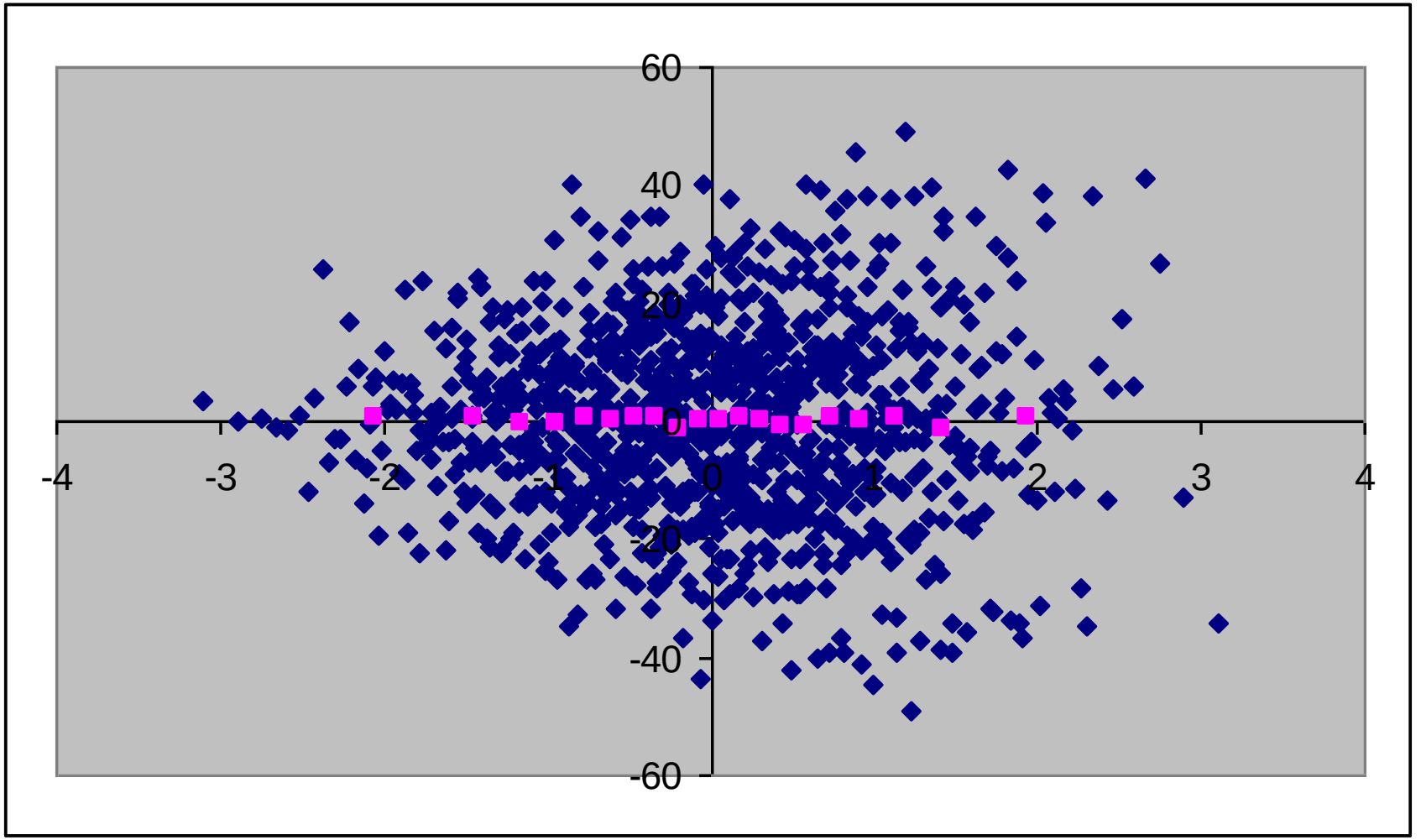
Unconditional variance

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)

Wouldn’t it be handy to have just a single ‘importance’ terms for all effects, inclusive of first order and interactions?

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

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

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** ▾

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

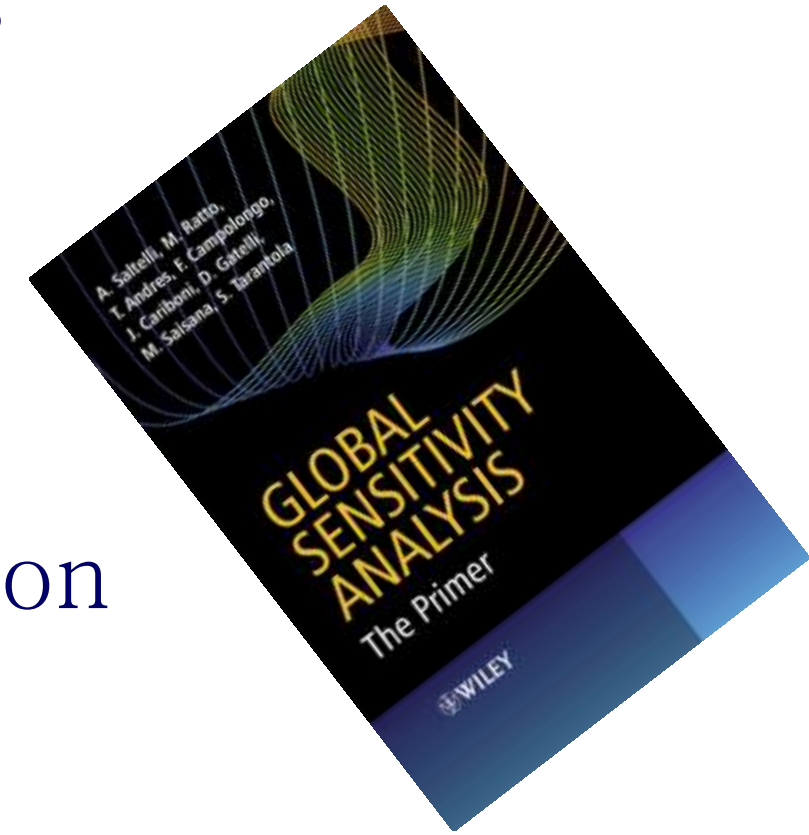
The difficulty of a function/model is not in its number of dimensions but in the number of effective dimensions, either in the **truncation** or **superposition** sense

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
- working with sets
- 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

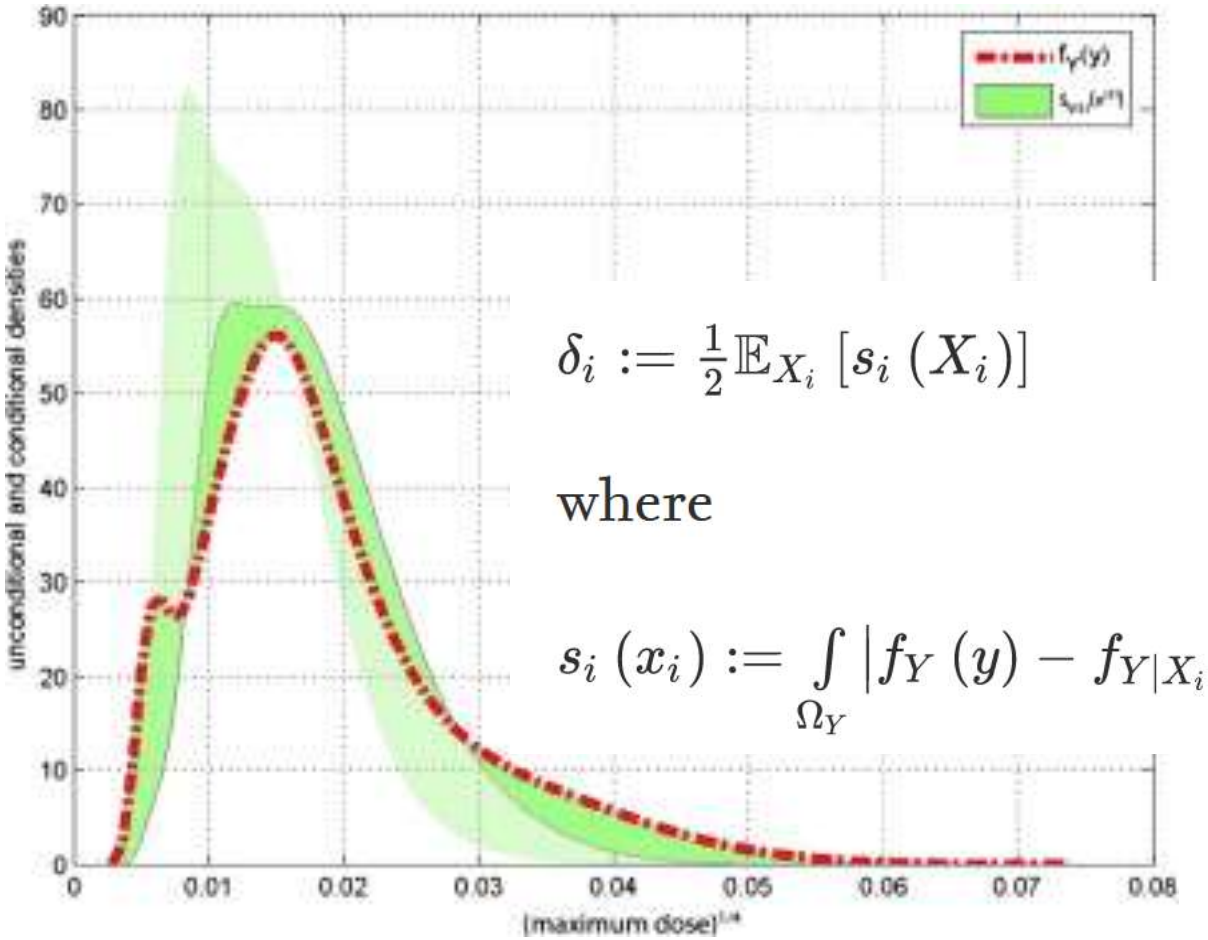
Environmental Modelling & Software

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

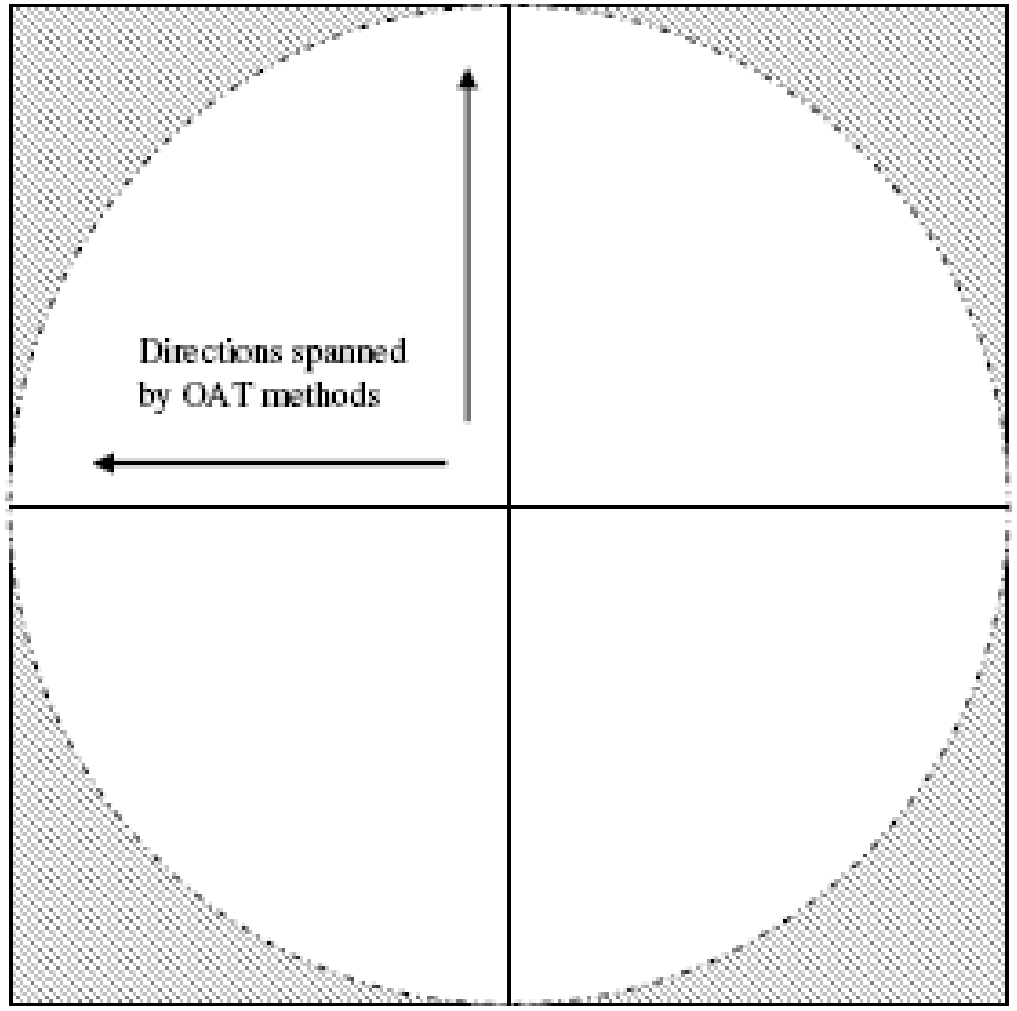
journal homepage: www.elsevier.com/locate/envsoft

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

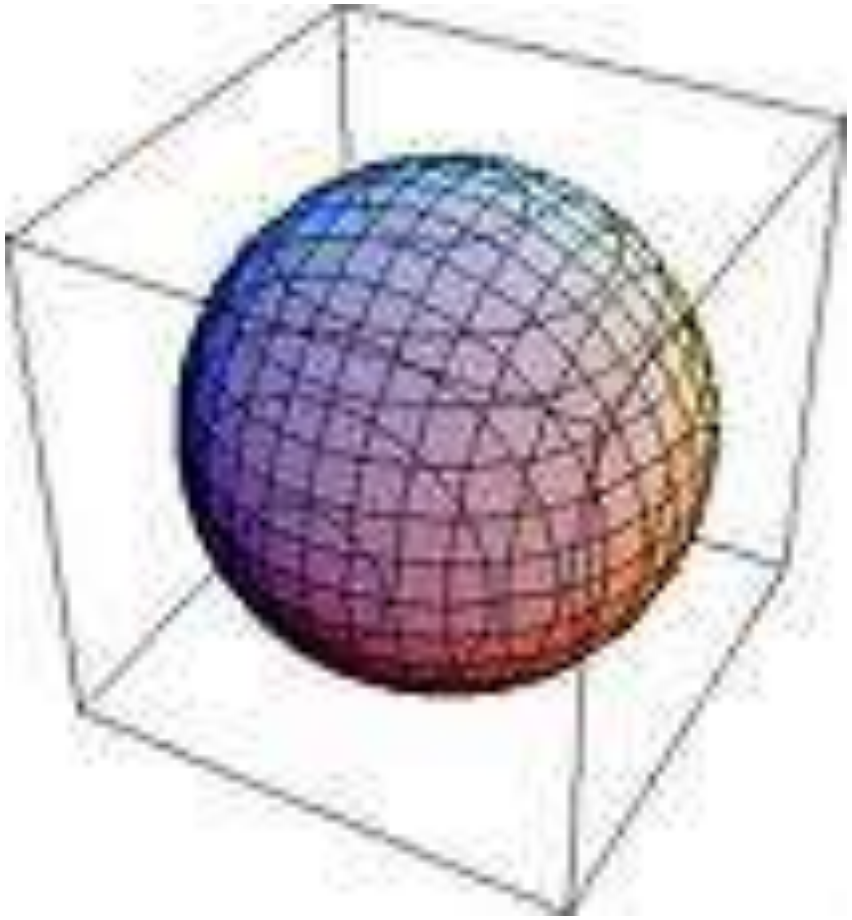
OAT in 2 dimensions



Area circle / area square =?

~ 3/4

OAT in 3 dimensions

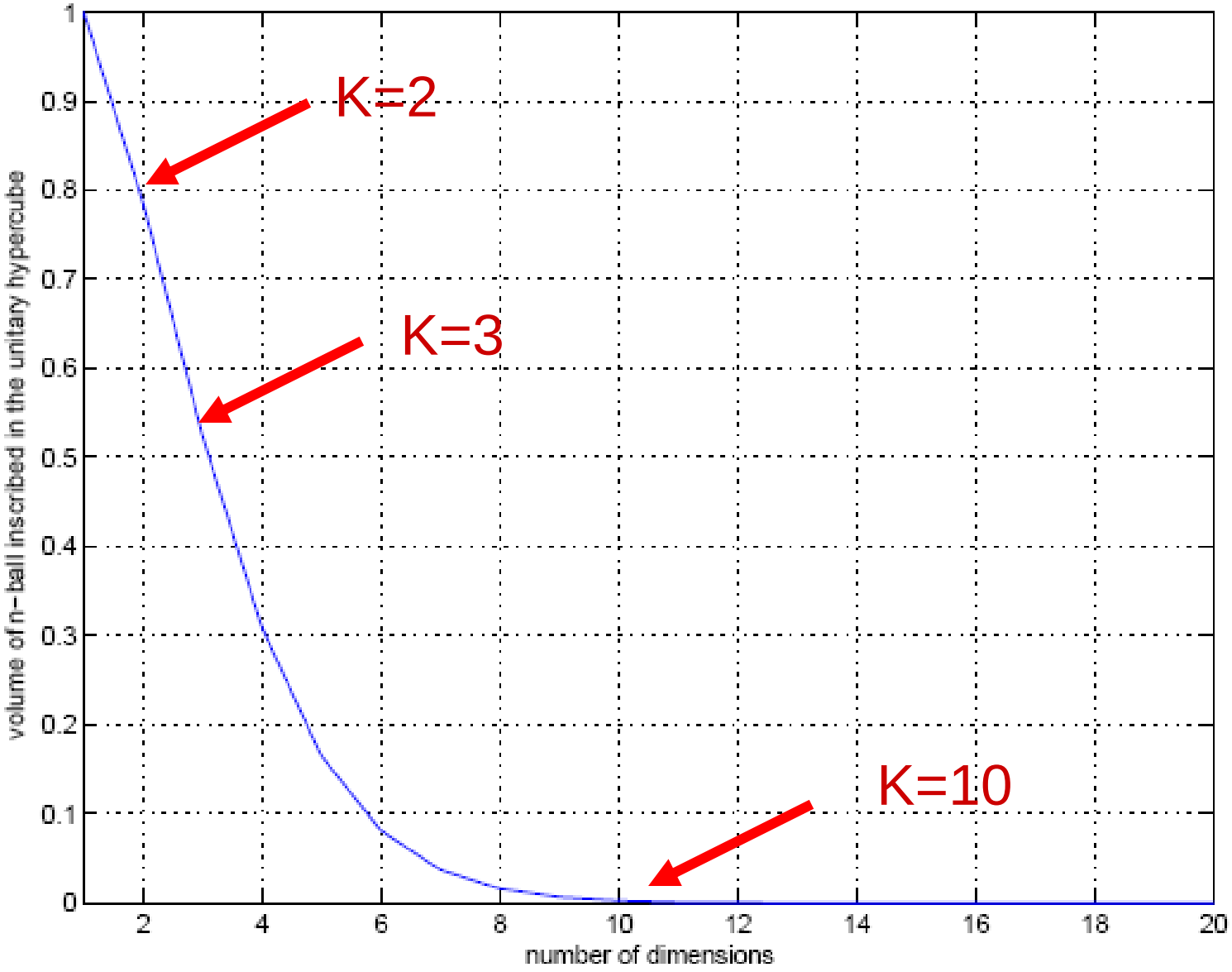


Volume sphere / volume cube =?

~ 1/2

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

OAT in k dimensions

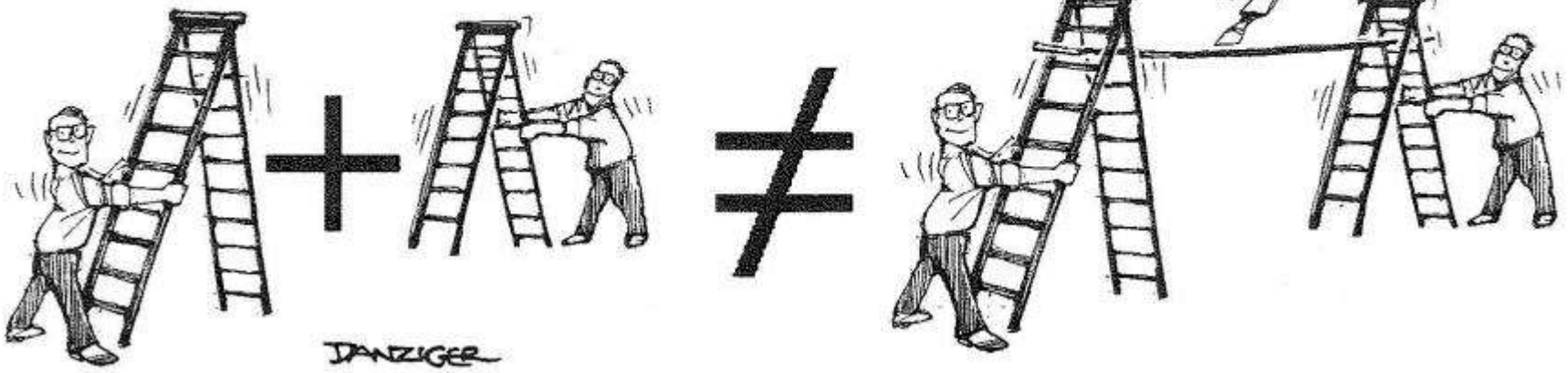


OAT does not capture interactions

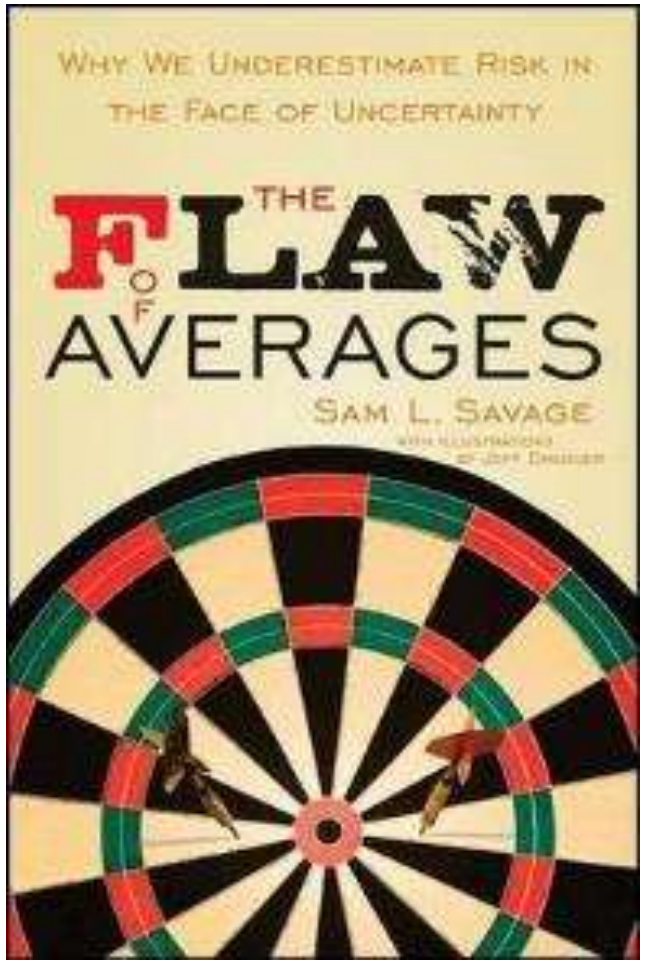
➔ The resulting analysis is non conservative

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

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

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

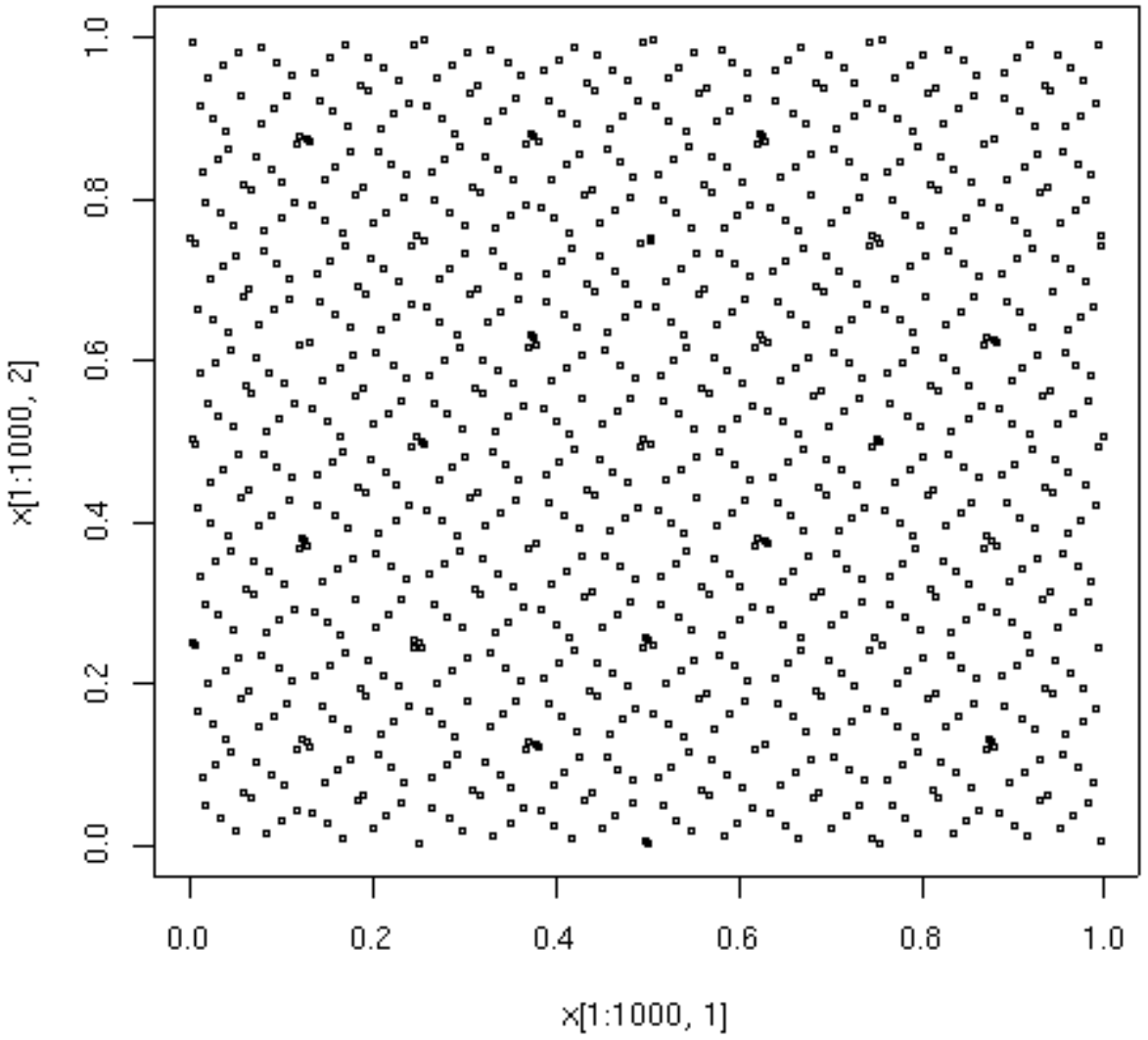
➔They assume linearity (PCC) or monotonicity (PRCC), which is difficult to know *ex-ante*

Don’t use either LHS or optimized LHS

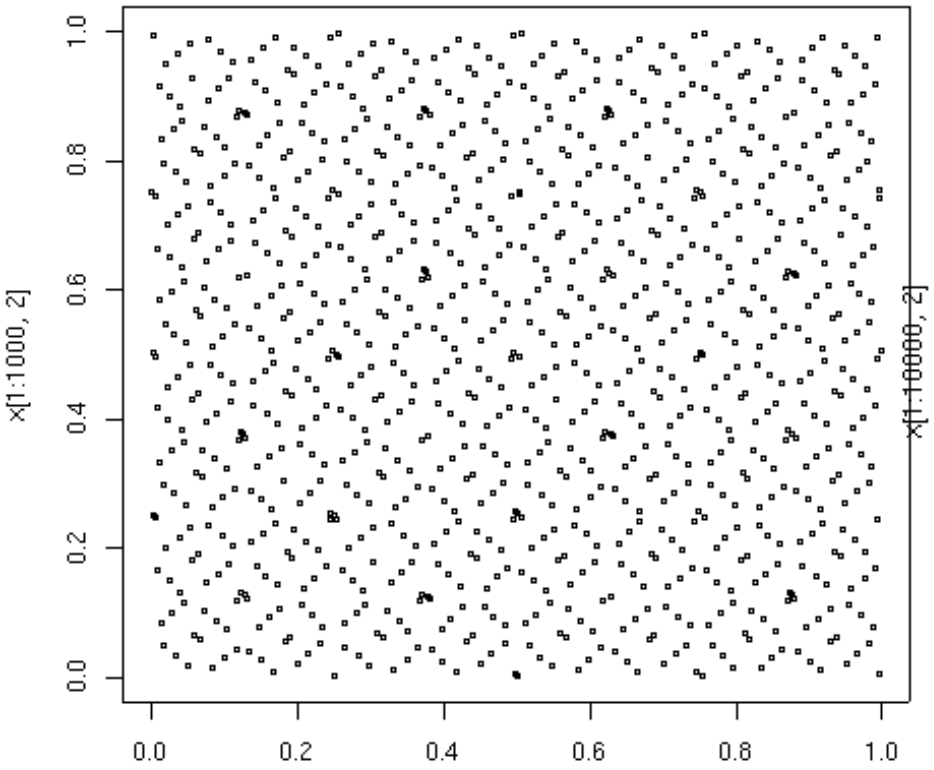
Quasi-random sequences are better

Ilya M. Sobol’

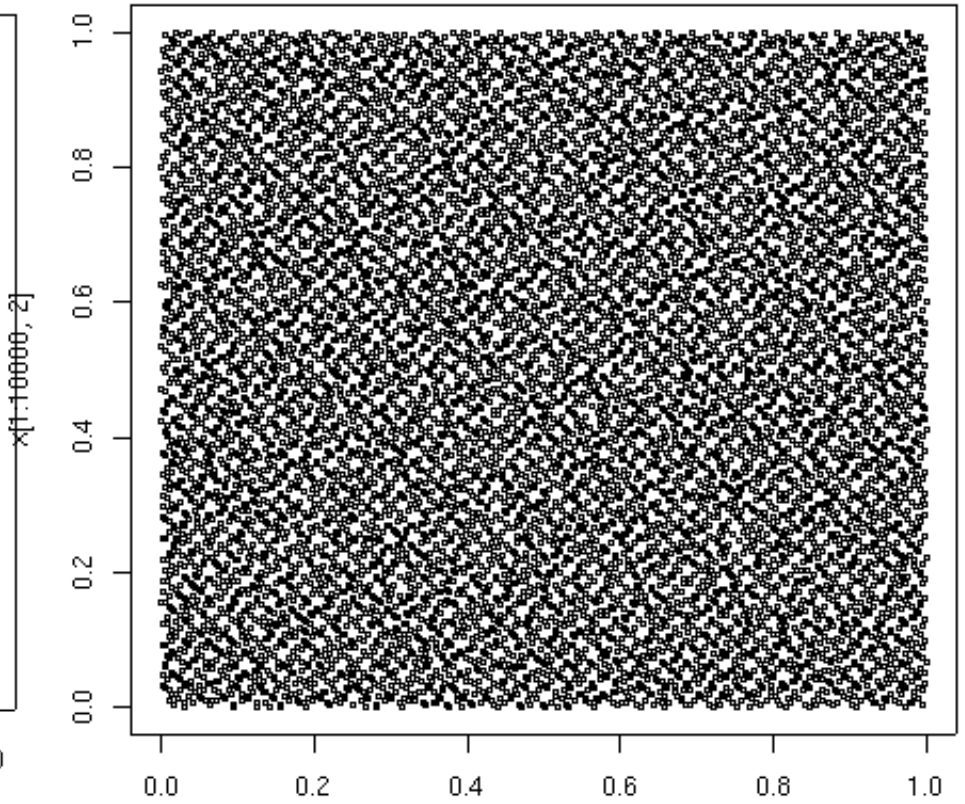
Quasi random sequences



An LP_τ sequence

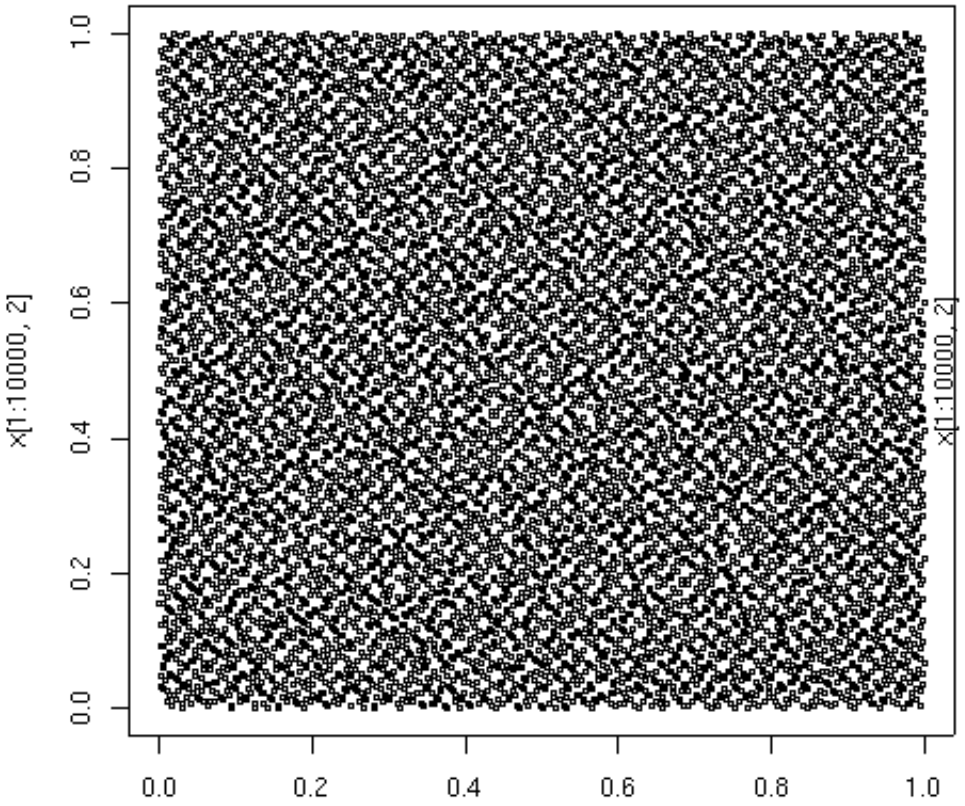


X1,X2 plane, 1000 Sobol' points

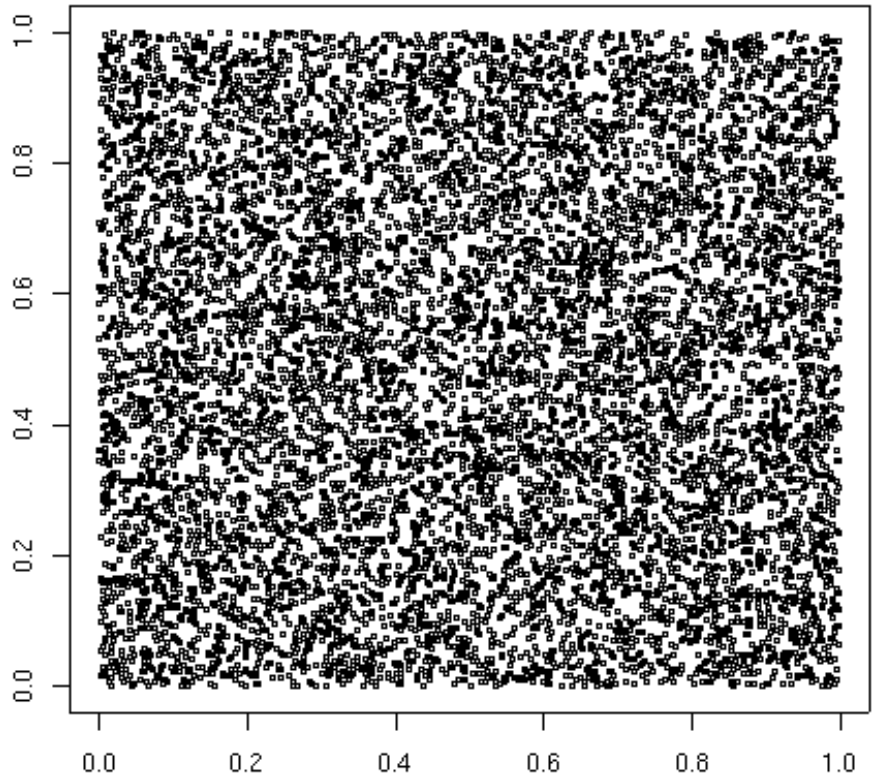


X1,X2 plane, 10000 Sobol' points

Sobol' sequences of quasi-random points



X1,X2 plane, 10000 Sobol' points

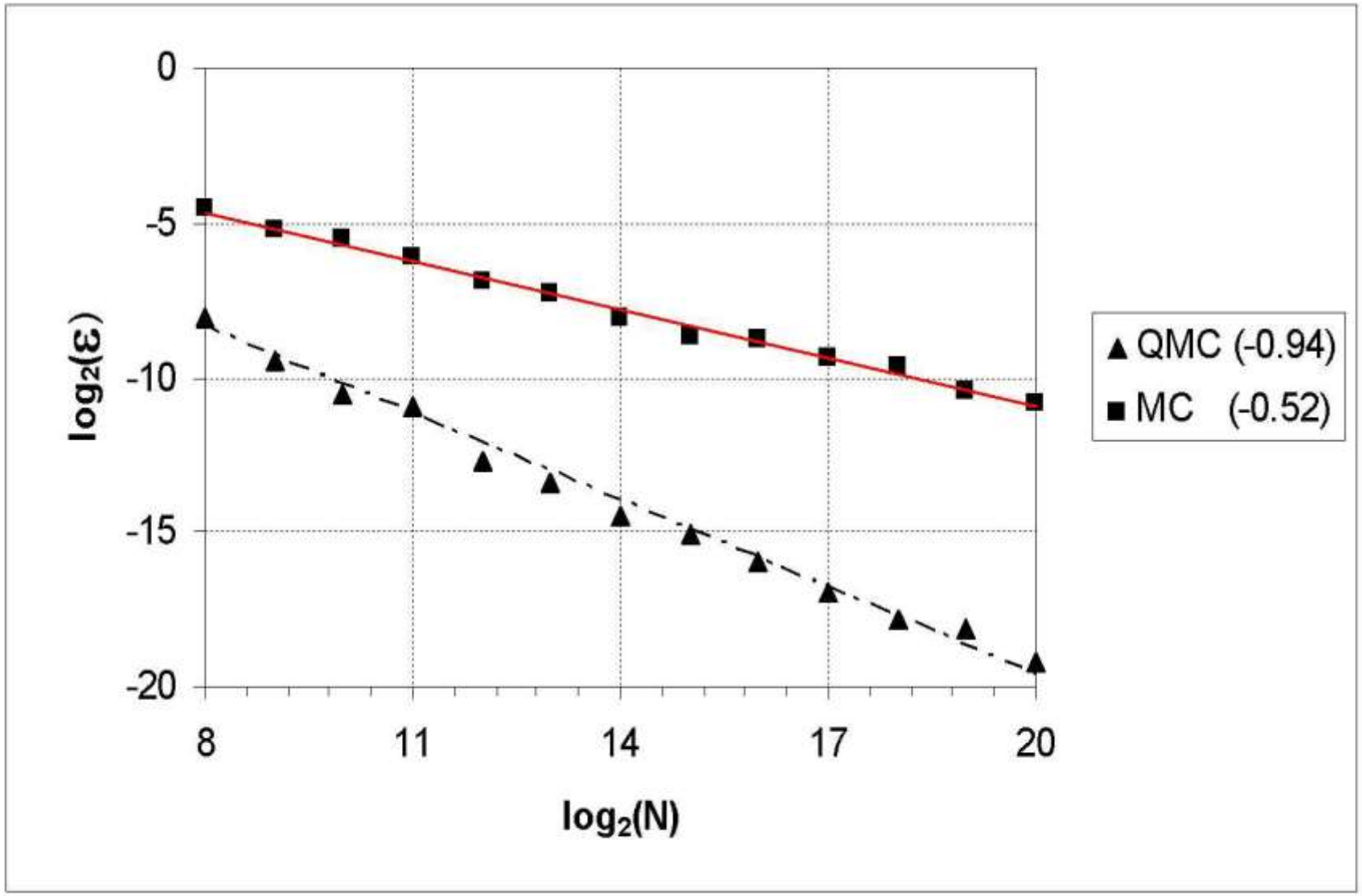


X1,X2 plane, 10000 random points

Sobol' sequences of quasi-random points against random points

Why quasi-random: they have faster convergence

Kucherenko S., Feil B., Shah N., Mauntz W. The identification of model effective dimensions using global sensitivity analysis Reliability Engineering and System Safety 96 (2011) 440–449.



$$\varepsilon = \left(\frac{1}{K} \sum_{k=1}^{K} (I[f] - I_k[f])^2 \right)^{1/2}$$

$$\sum_{i=1}^{n} (-1)^i \prod_{j=1}^{i} x_j$$

Error=numeric-versus-analytic value the integral of the function (for n=360) over its dominion.

Root mean square error over K=50 different trials.

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

Comparing three different sampling methods over an array of functions of different dimensionality and difficulty

The concept of effective dimension

Don’t use plain LHS and think twice about optimized LHS

If in doubt try it for yourself with a set of test functions of varying dimensionality

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!

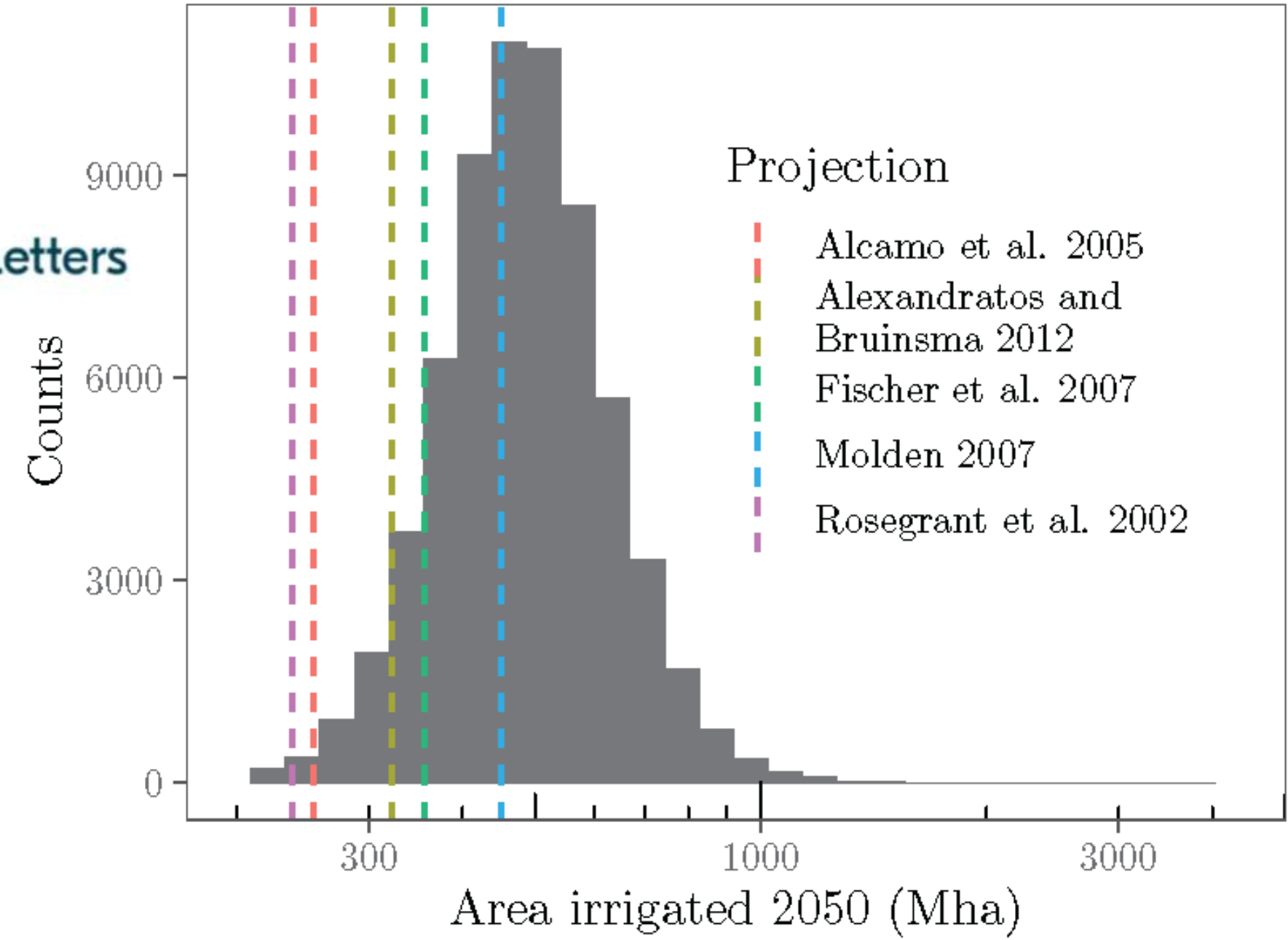
Model routinely used to produce point estimates may becomes non conservative when the uncertainty is plugged in

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



Don’t use Morris’ method

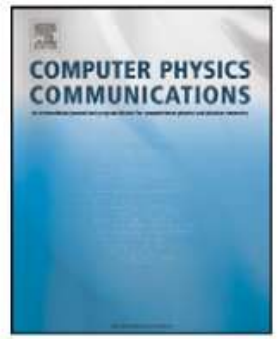
More cumbersome and fragile than the total sensitivity index that is its close equivalent

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

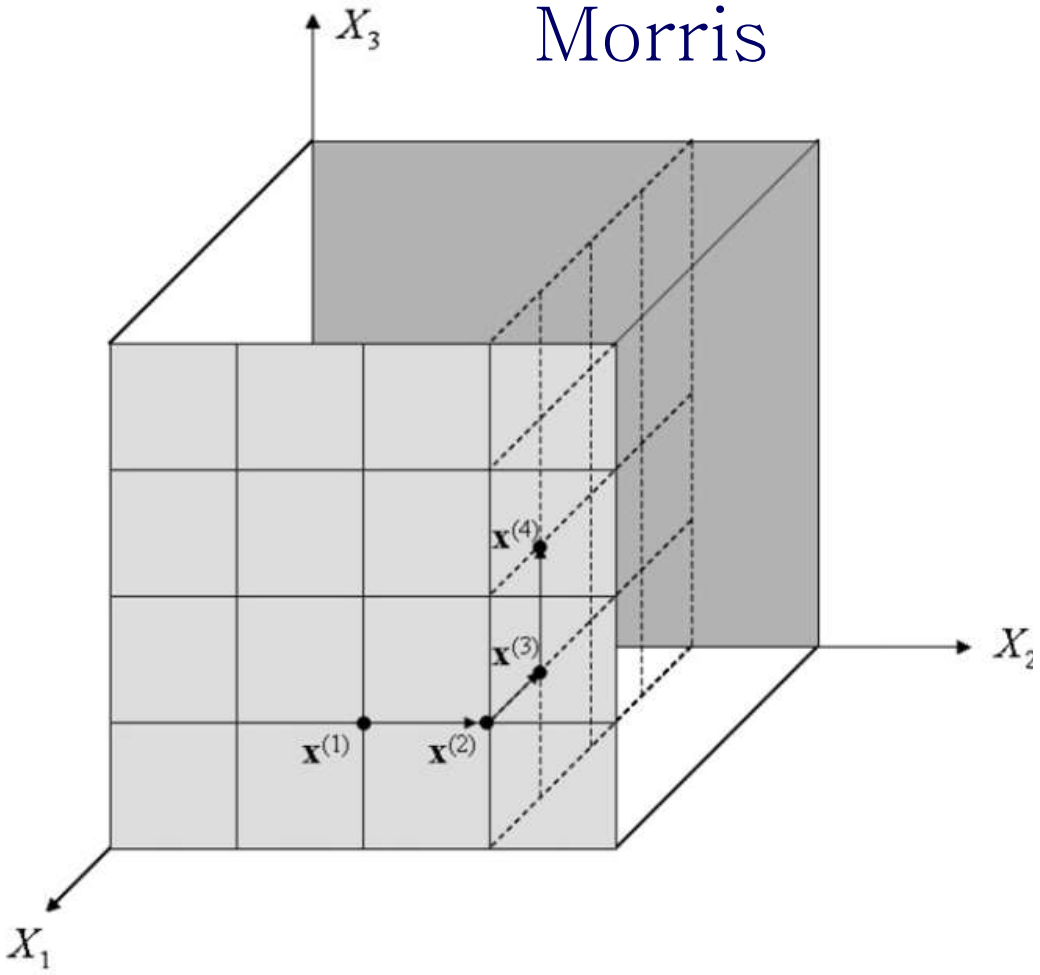


Fig. 1. Example of trajectory in 3 dimensions for the original EE method.

Total sensitivity index T_i

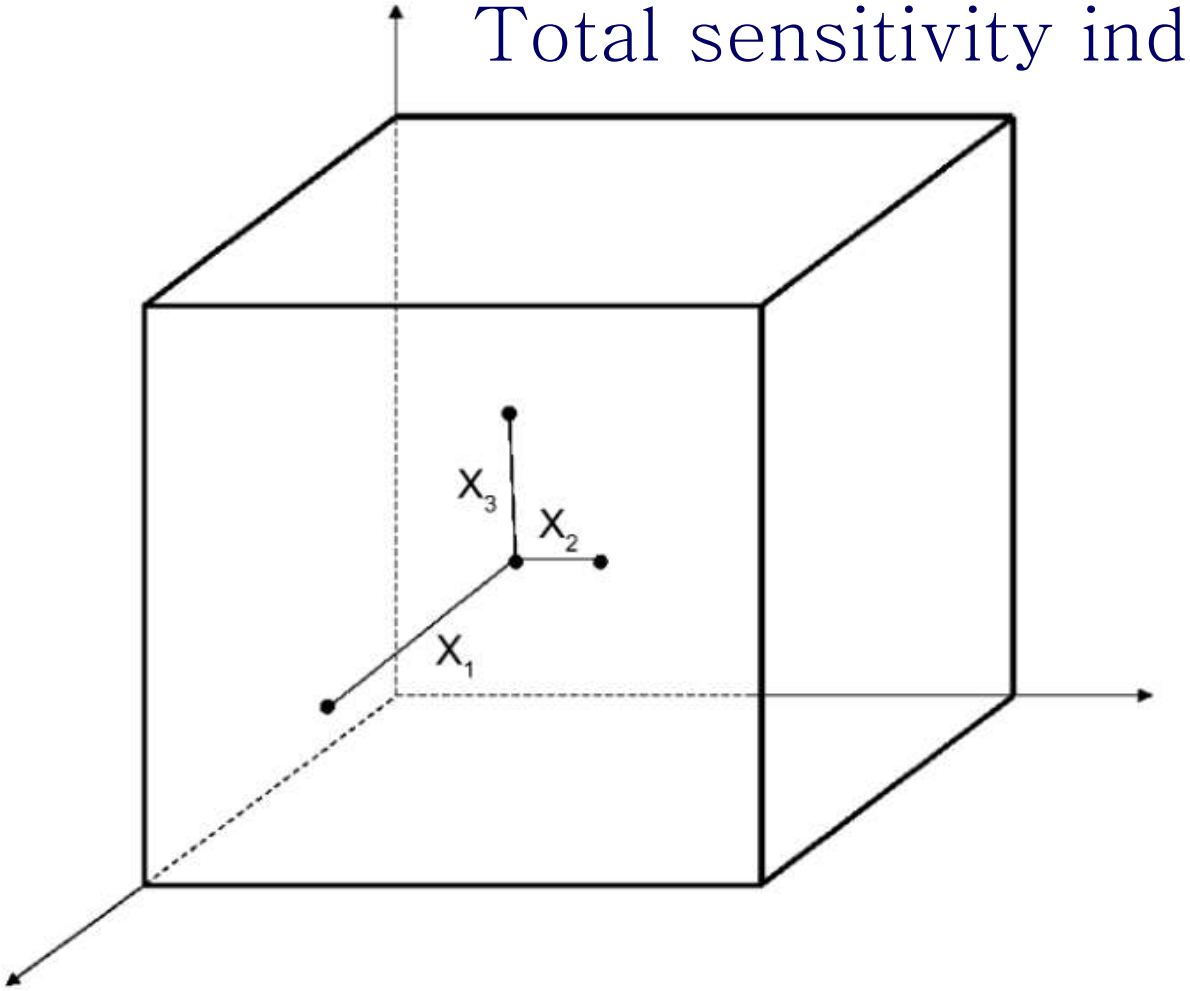
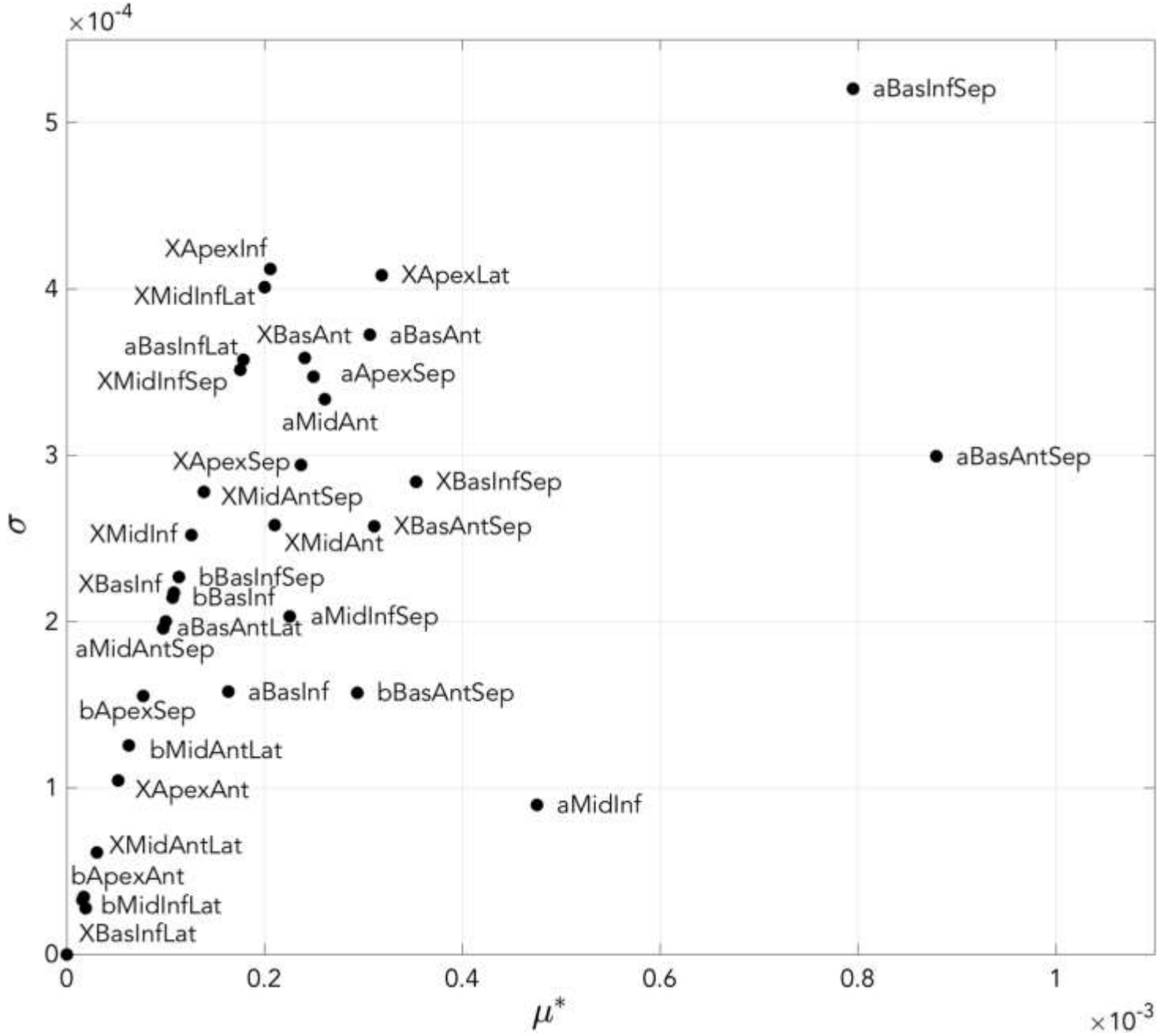


Fig. 3. Example of a radial sample in three dimensions.

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

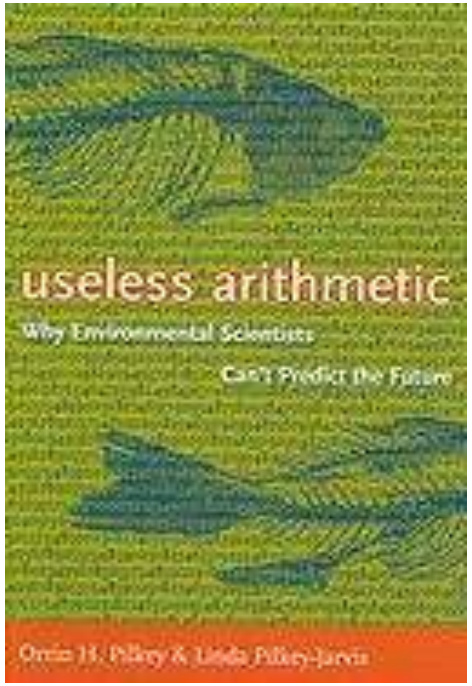
Morris is more cumbersome to interpret (two outputs: mu and sigma)



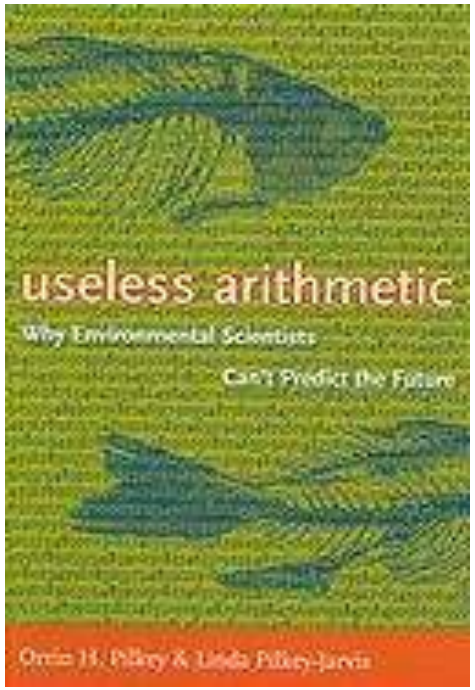
Don’t confuse the map with the territory

If you do, sensitivity analysis will not save you

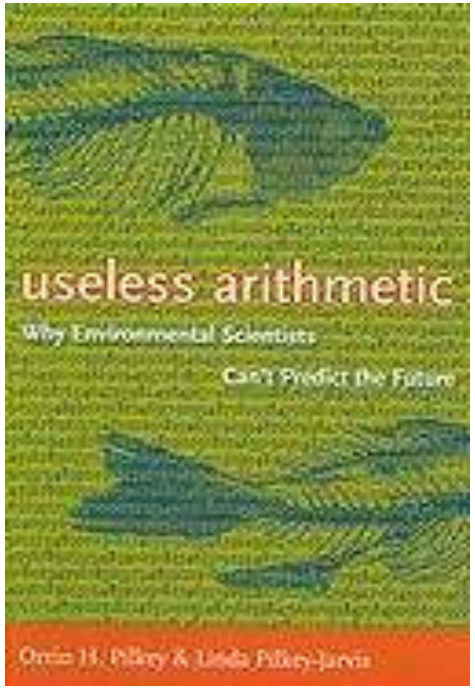
Orrin H.
Pilkey



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



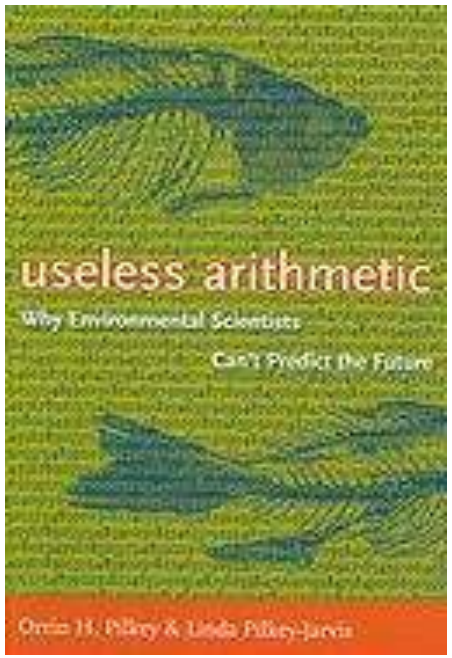
<<It is important, however, to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature>>

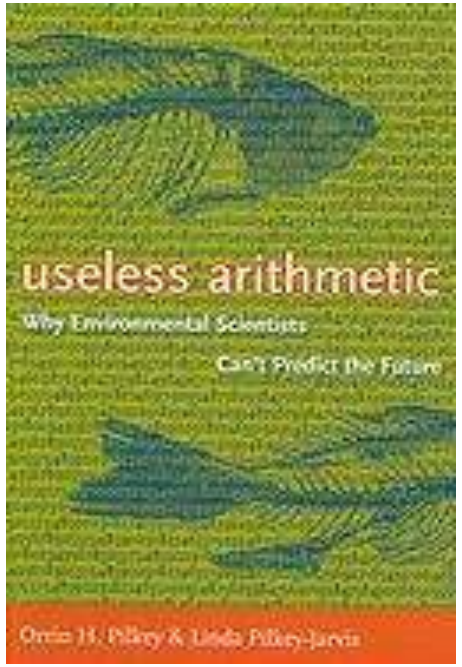


<<…If the model is wrong or if it is a poor representation of reality, determining the sensitivity of an individual parameter in the model is a meaningless pursuit>>

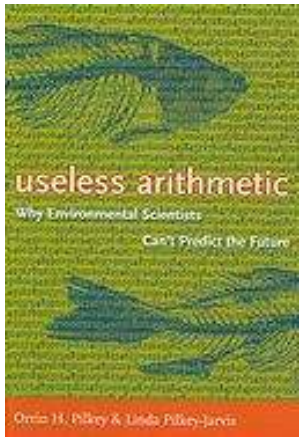
One of the examples discussed concerns the Yucca Mountain repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis.

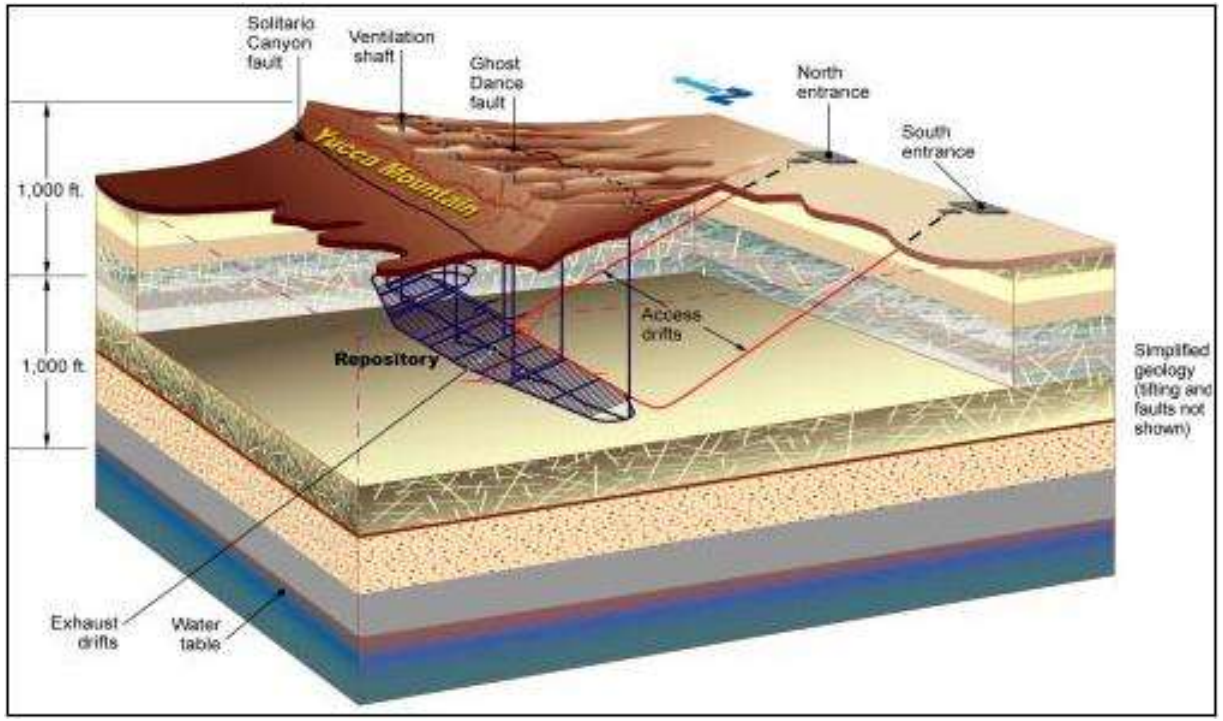
TSPA is Composed of 286 sub-models.





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.





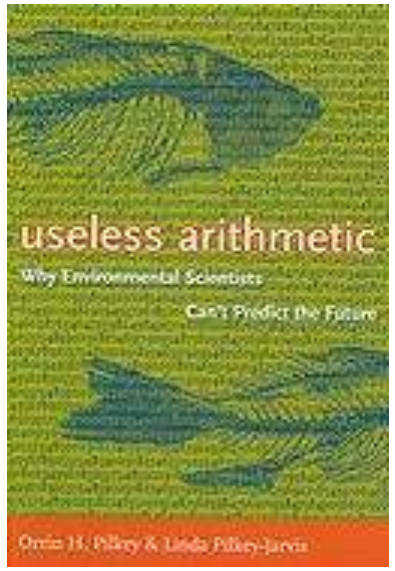
The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter

(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

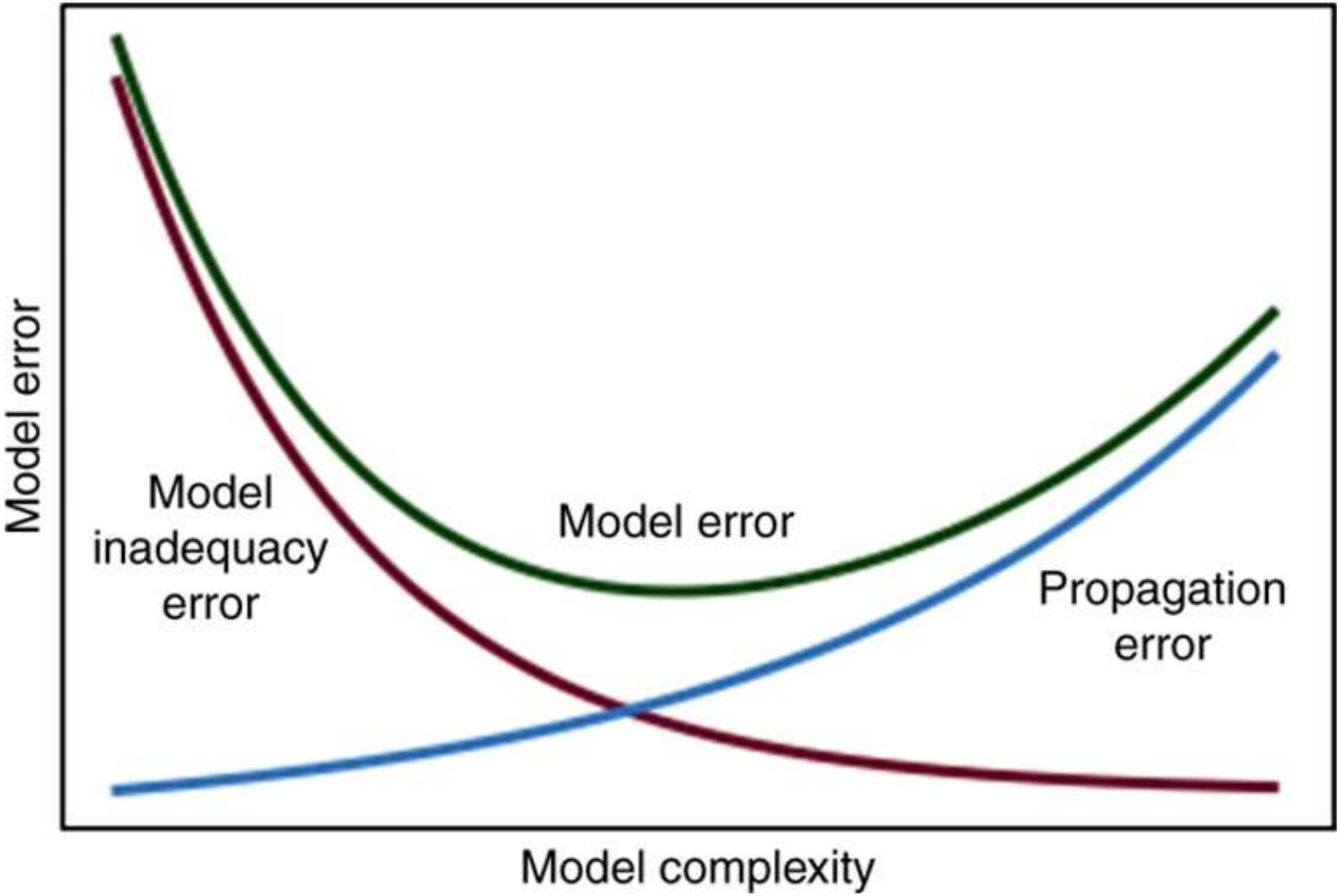
Mind the conjecture of O'Neil

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"

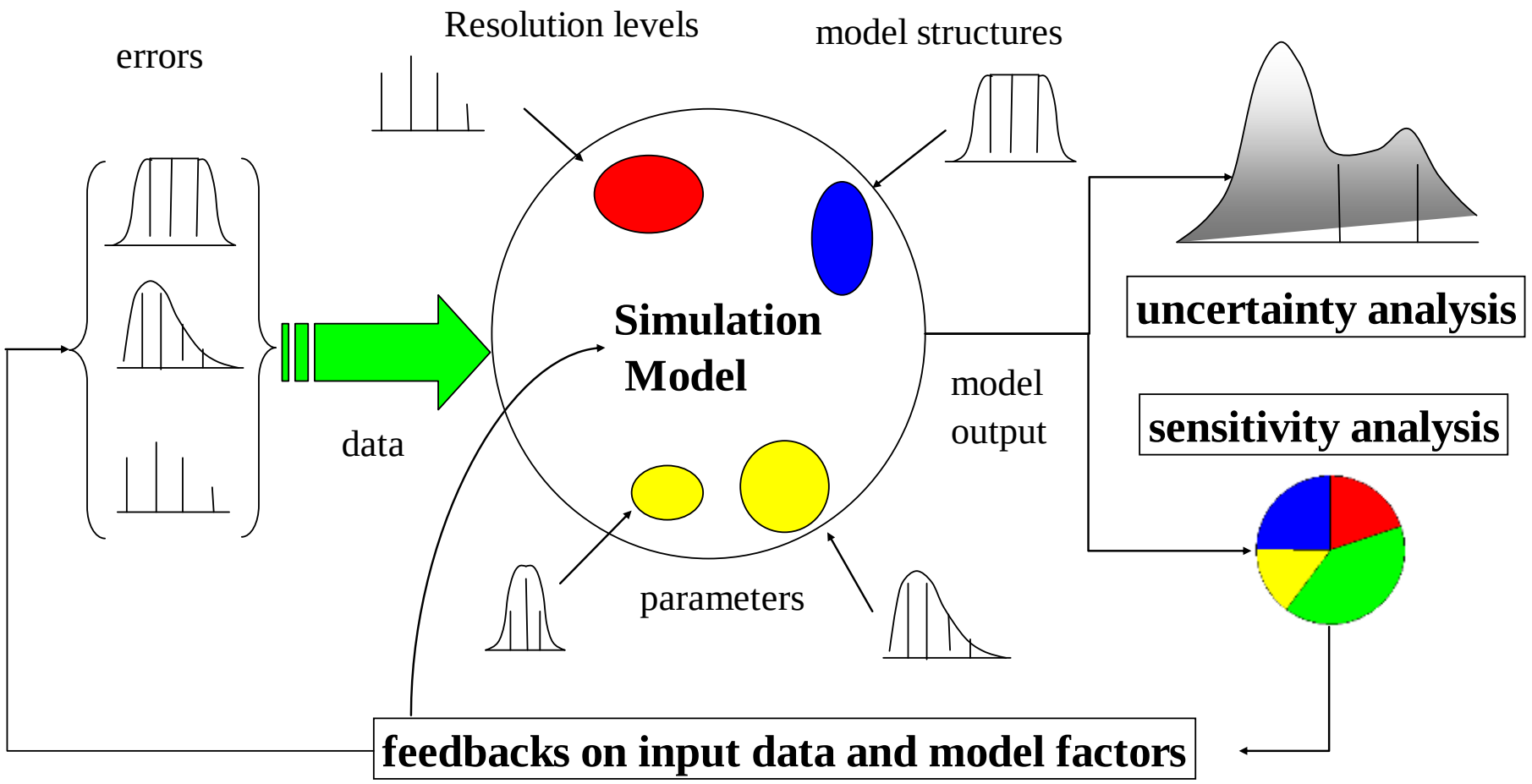
In M. G. Turner and R. H. Gardner, "Introduction to Models" in Landscape Ecology in Theory and Practice, New York, NY: Springer New York, 2015, pp. 63–95.

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

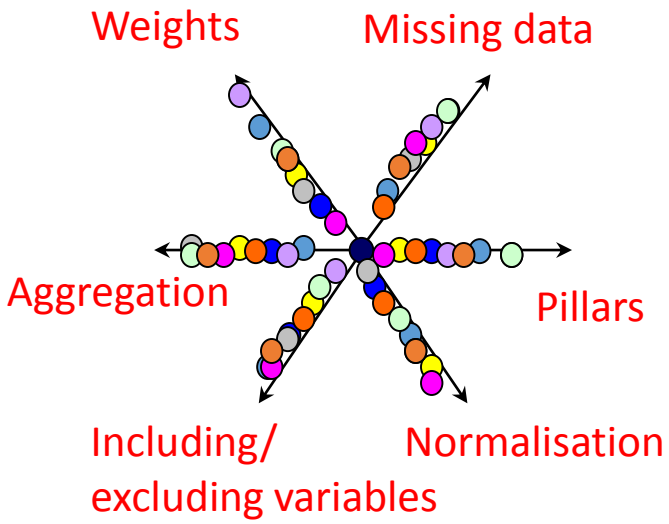


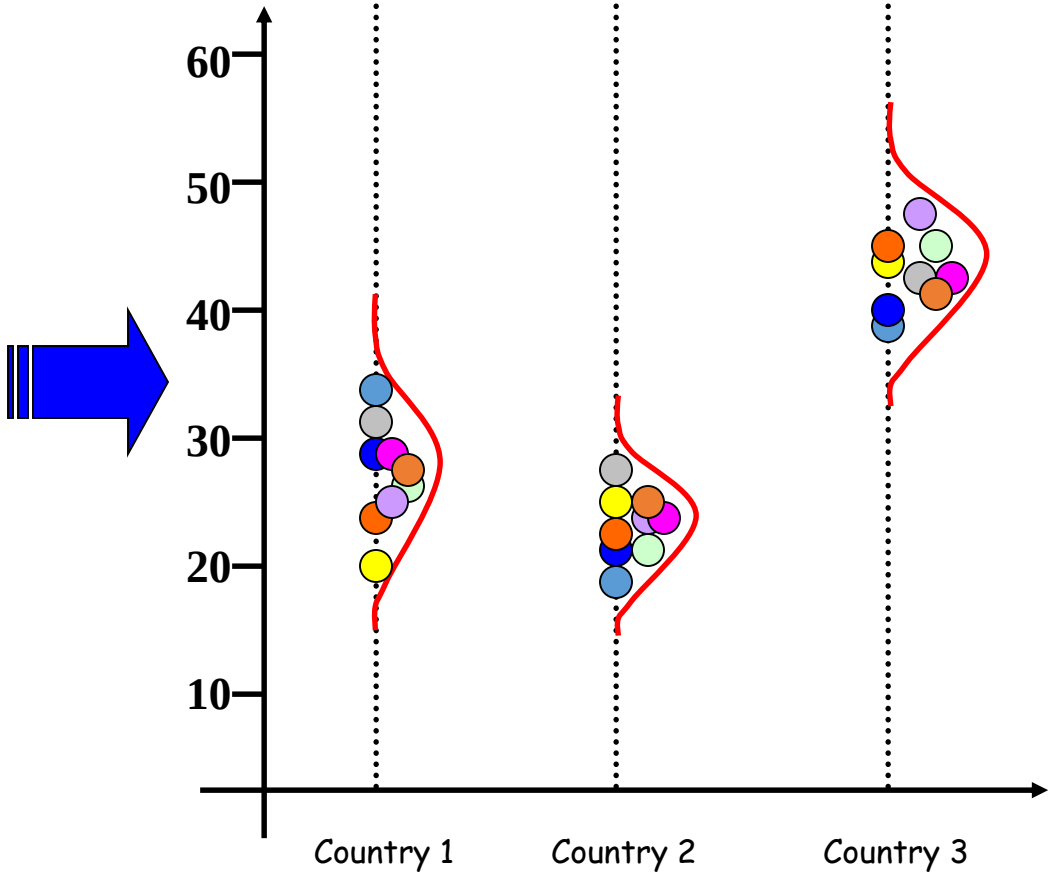
One can sample more than just factors:

- 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.

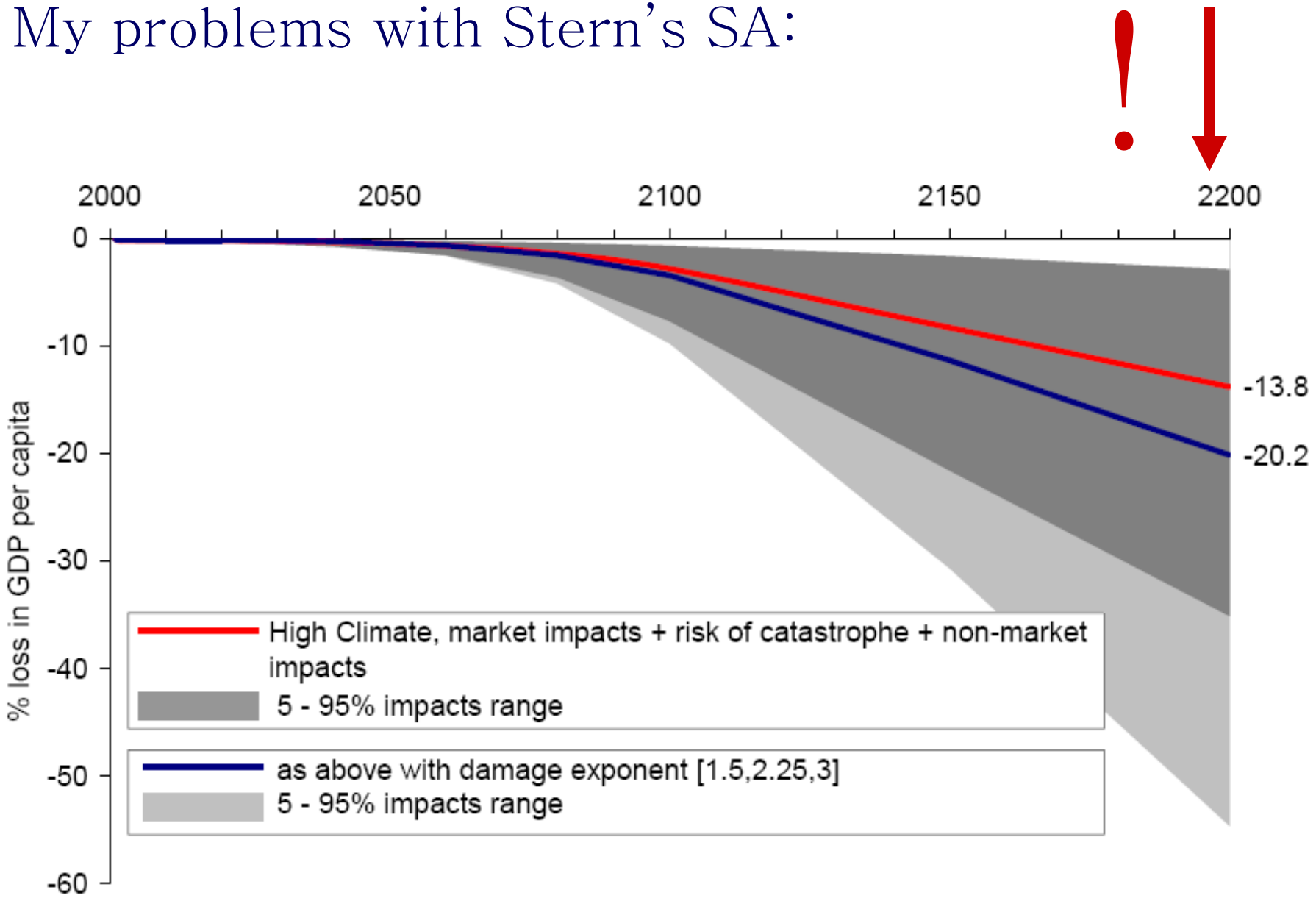
Nordhaus W., Critical Assumptions in the Stern Review on Climate Change, SCIENCE, 317, 201-202, (2007).

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:

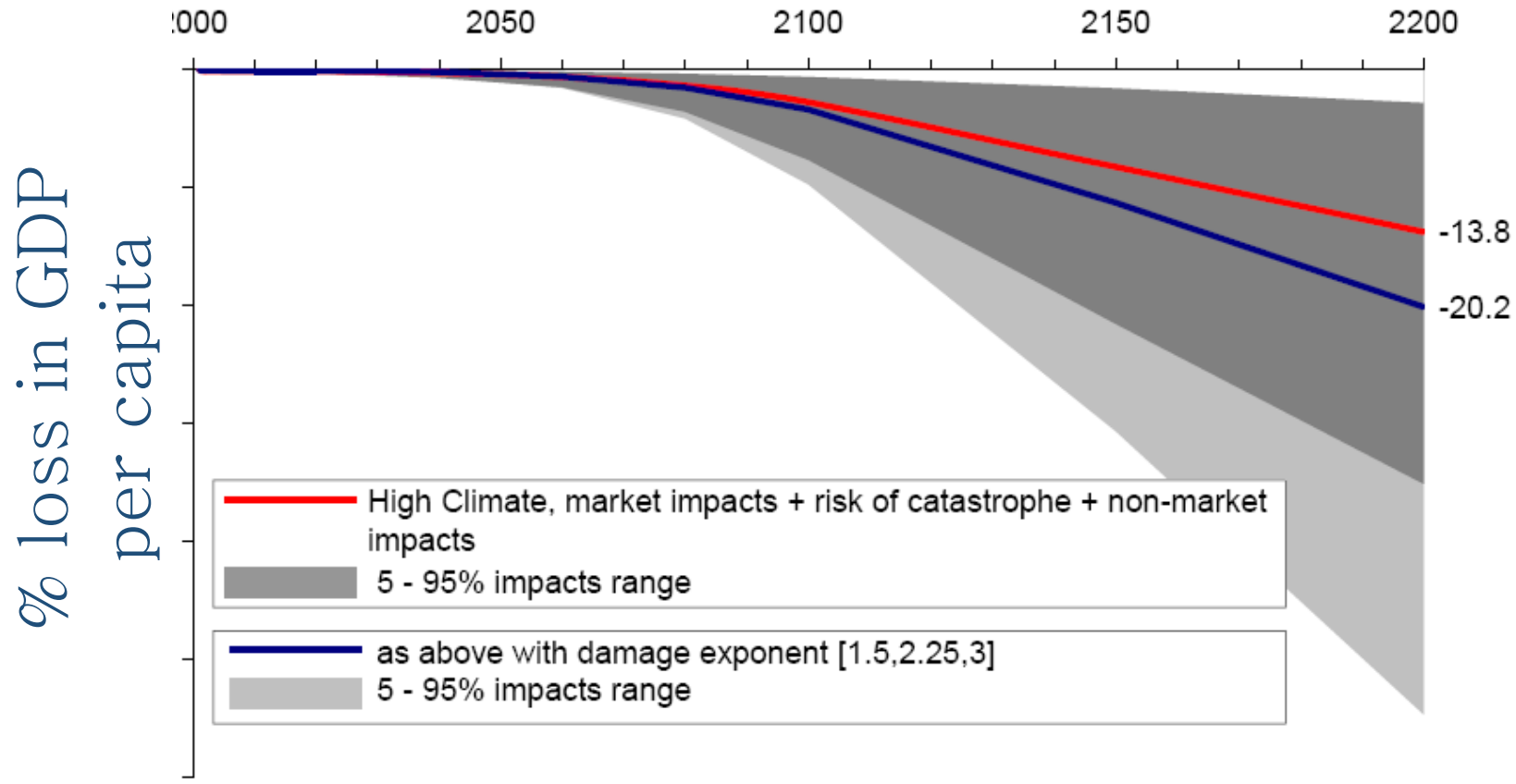


… but foremost Stern says:

changing assumptions → important effect

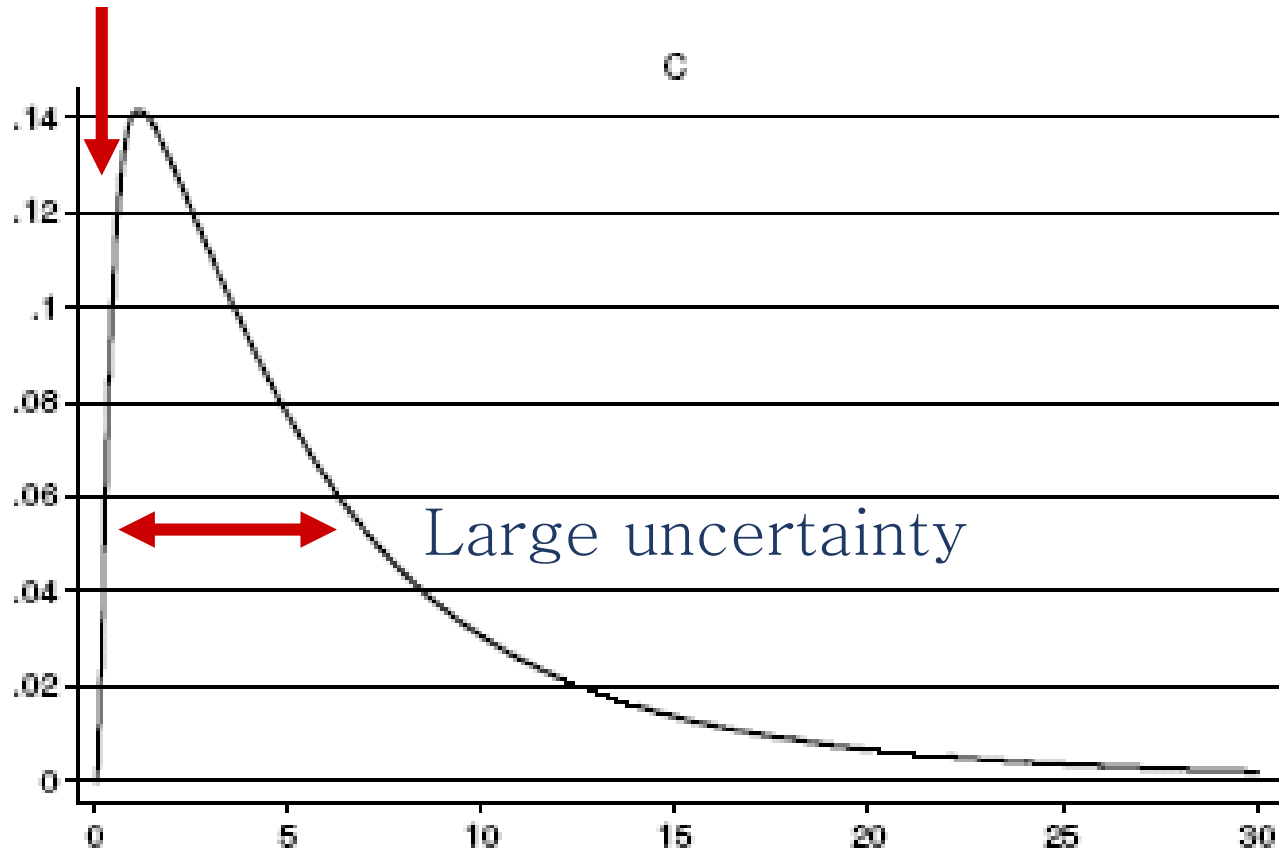
when instead he should admit that:

changing assumptions → all changes a lot



How was it done? A reverse engineering of the analysis

Missing points



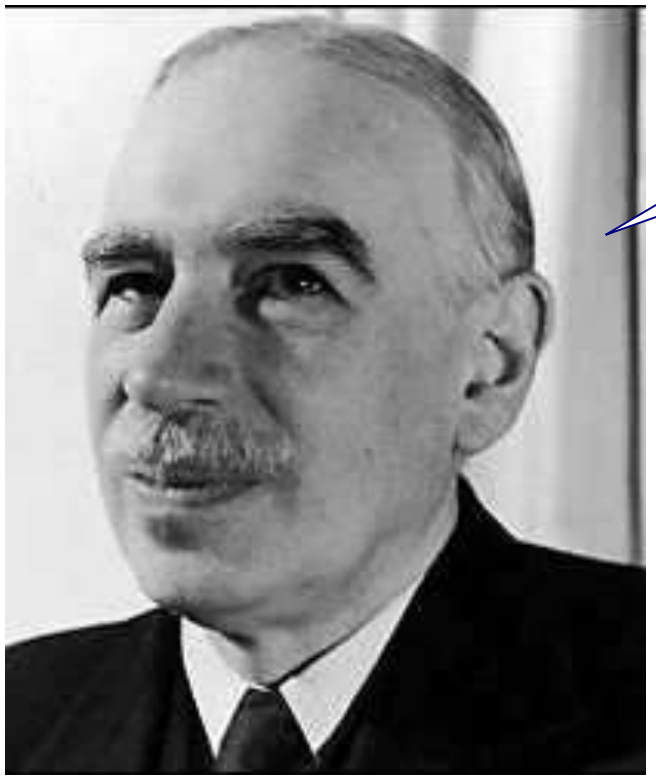
% loss in GDP per capita

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

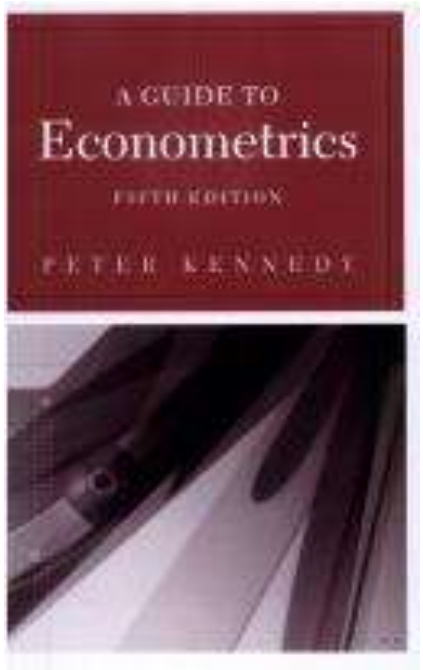
Same criticism applies to Nordhaus – both authors frame the debate around numbers which are …



… 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%

"… a modern pseudo-science where the uncertainty of its inputs must be suppressed, lest they render its outputs totally indeterminate…"

GIGO-Science=Garbage In, Garbage Out

OPINION

PETER COY

‘The Most Important Number You’ve Never Heard Of’

Sept. 17, 2021

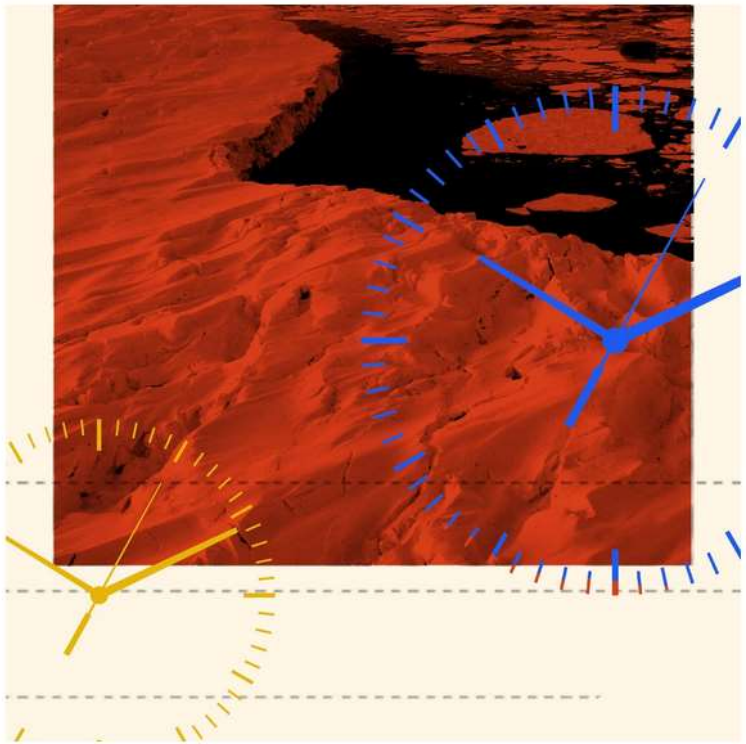


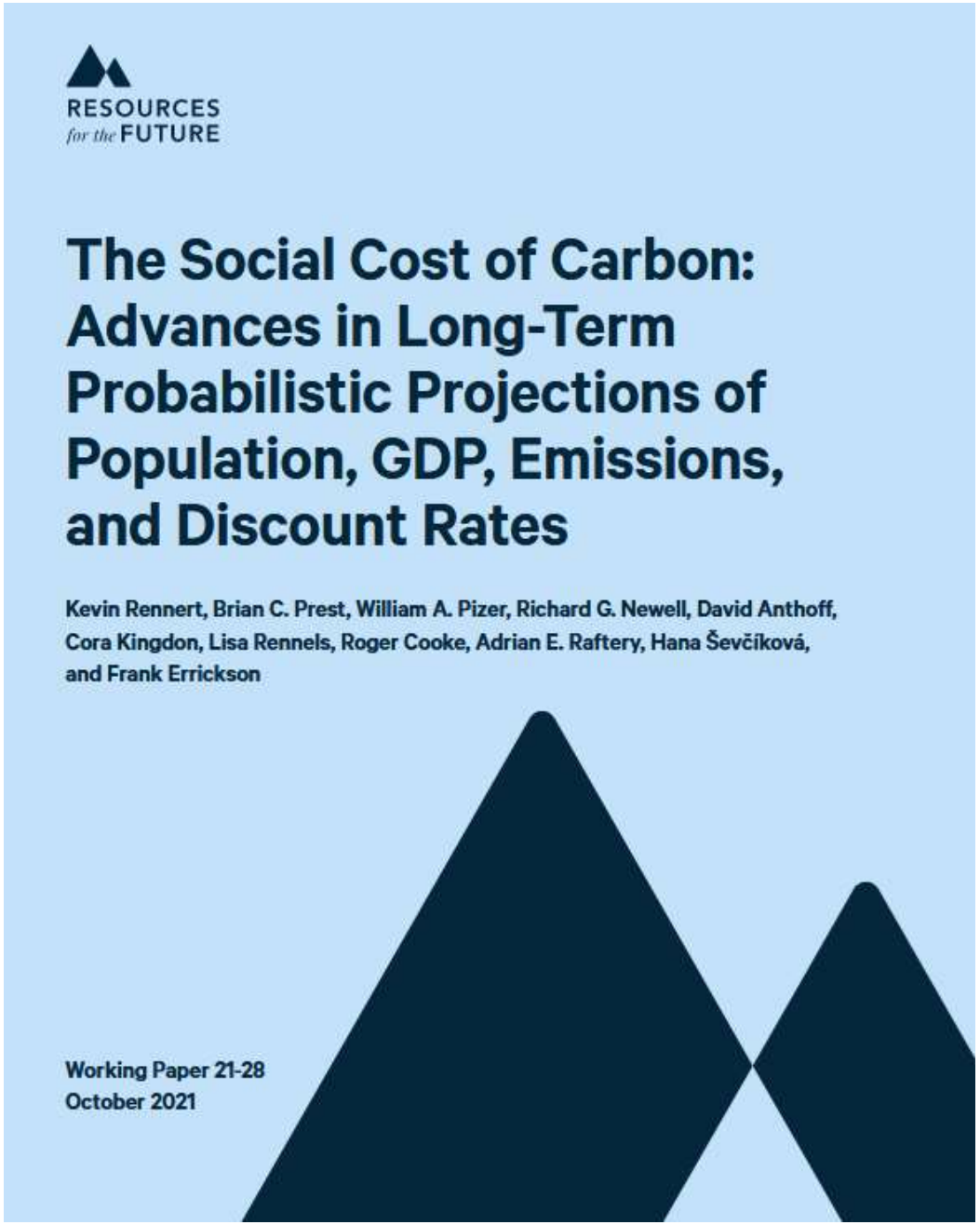
Illustration by Arsh Raziuddin, The New York Times

“social cost of carbon:

=$56 a ton on average at a 3 percent discount rate

=$171 a ton on average at a 2 percent discount rate”

The New York Times



Averaged till year 2300

Feeds into policy design

We have perhaps reached a complex epistemic state, where on the one hand ‘everybody knows’ that some numbers are pseudo-precise and that numbers can be gamed, while the game works only because most people don’t know about it

Jerome R. Ravetz

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 most expensive to get

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

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/

Statistics > Computation

[Submitted on 22 Jan 2021 (v1), last revised 3 Dec 2021 (this version, v3)]

sensobol: an R package to compute variance-based sensitivity indices

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

Journal of Statistical Software

The End

@andreasaltelli