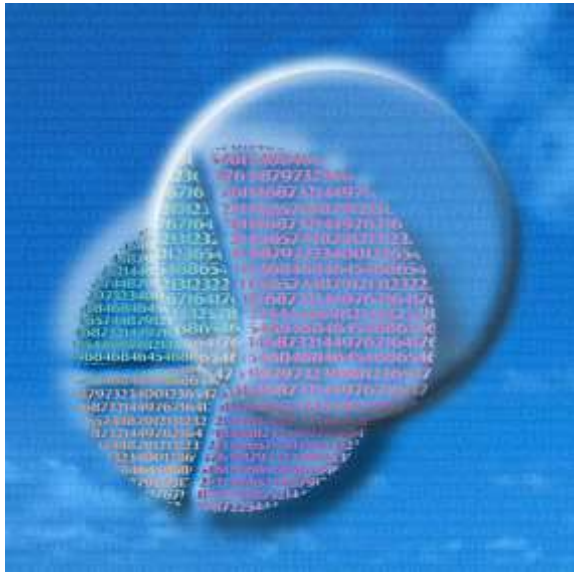


Sensitivity Analysis

Andrea Saltelli
Open Evidence Research, Open University of Catalonia

Lecture to the Department of Statistics,
Harvard University, September 20, 2021

Where to find this talk: www.andreasaltelli.eu

Definitions

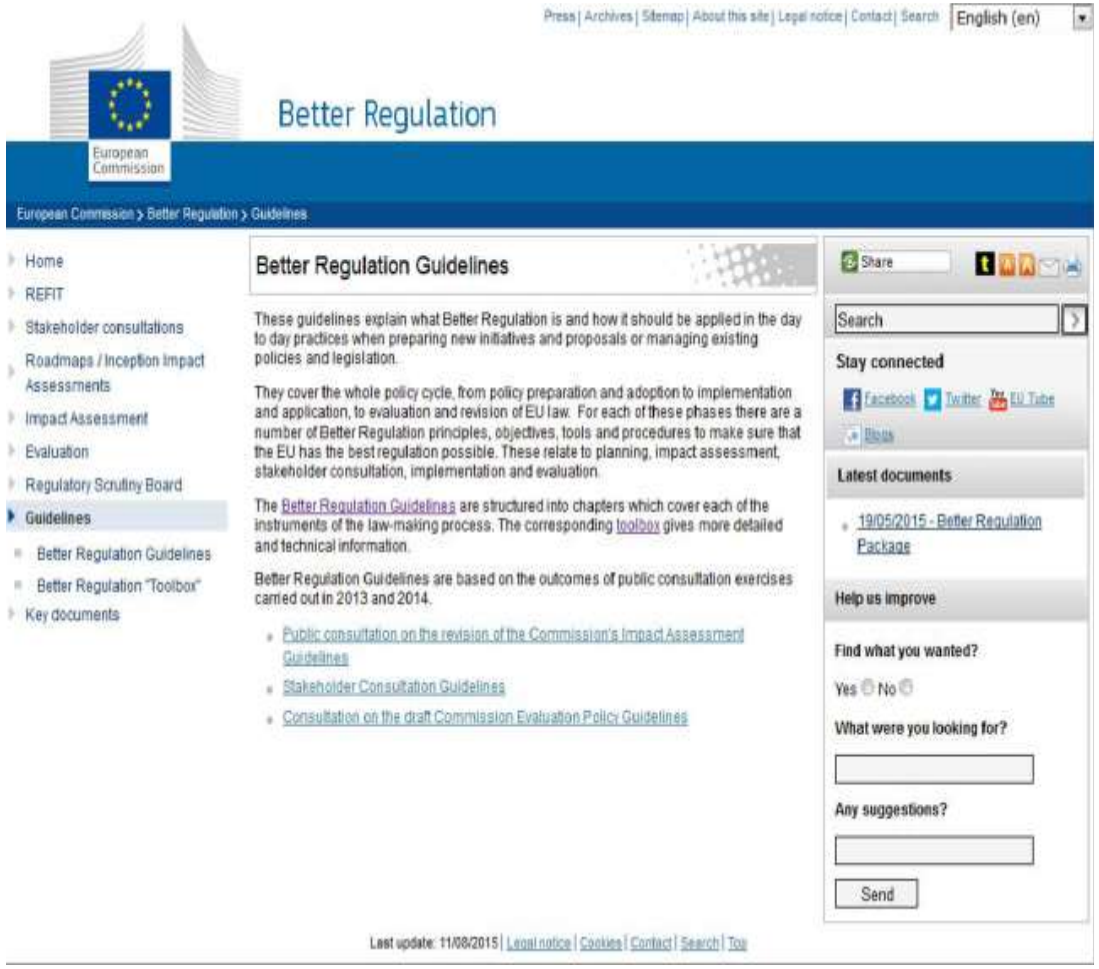
Uncertainty analysis: Focuses on just quantifying the uncertainty in model output

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

Why Sensitivity analysis?

It is in the guidelines!

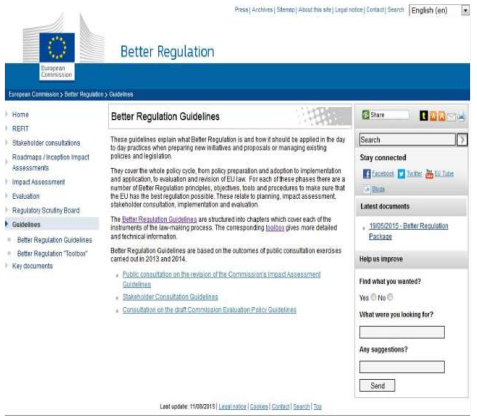
EC impact assessment guidelines: sensitivity analysis & auditing



http://ec.europa.eu/smart-regulation/guidelines/docs/br_toolbox_en.pdf

4. SENSITIVITY AND UNCERTAINTY ANALYSES

Page 391



Six steps for a global SA:

1. Select <u>one</u> output of interest;
2. Participatory step: discuss which input may matter;
3. Participatory step (extended peer review): define distributions;
4. Sample from the distributions;
5. Run (=evaluate) the model for the sampled values;
6. Obtain in this way bot the uncertainty of the prediction and the relative importance of variables.

Is something wrong with this statement (p. 384 of EC guidelines)

The influence of the key variables

should be investigated by a 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.

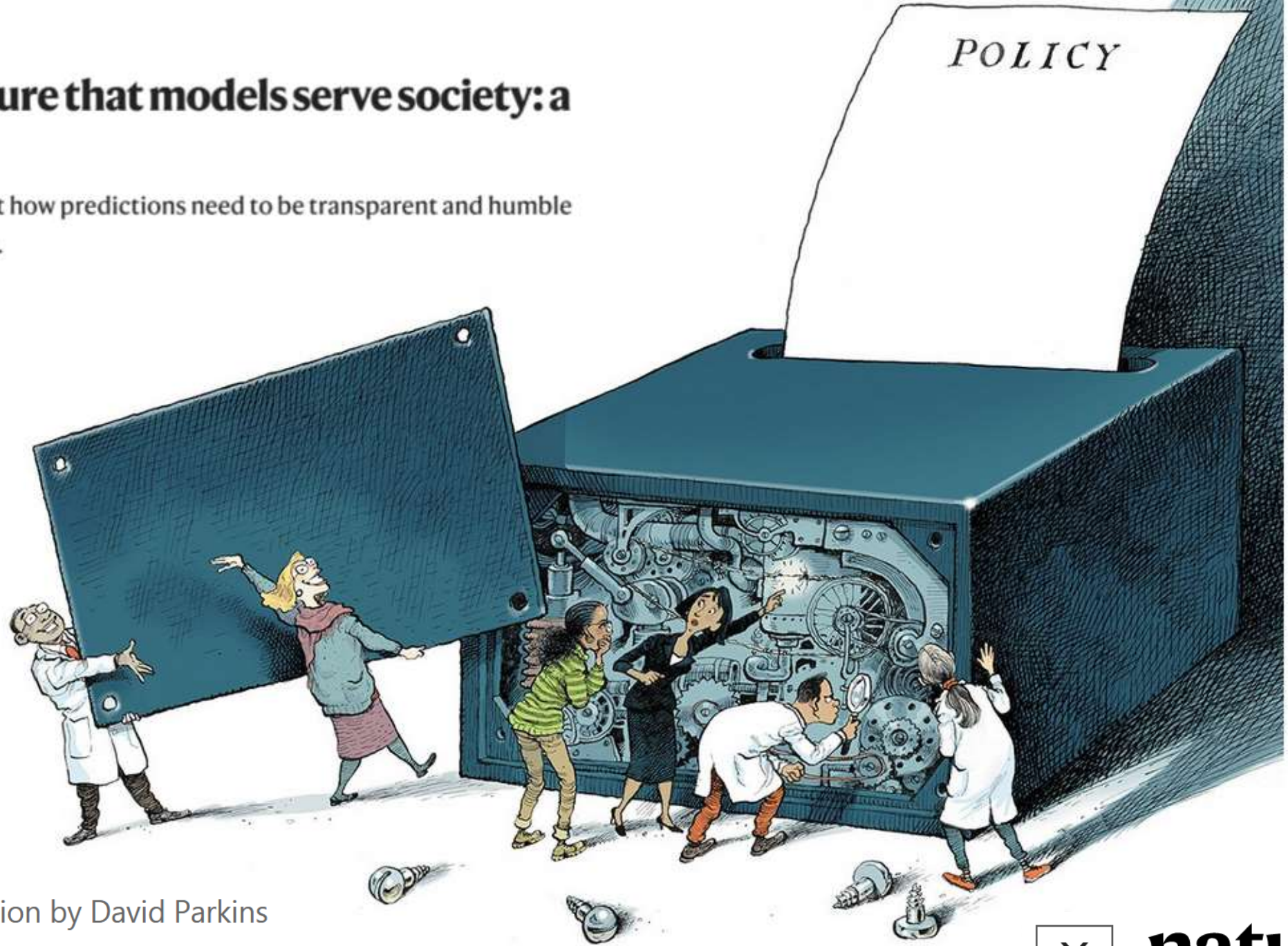


Illustration by David Parkins

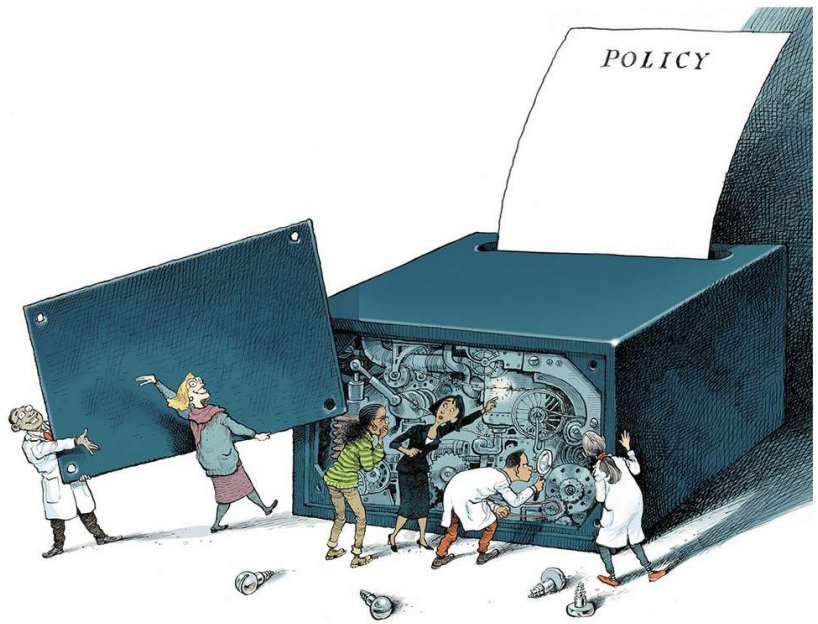
COMMENT · 24 JUNE 2020

nature

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.

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



Mind the assumptions

Assess uncertainty and sensitivity

Mind the hubris

Complexity can be the enemy of relevance

Mind the framing

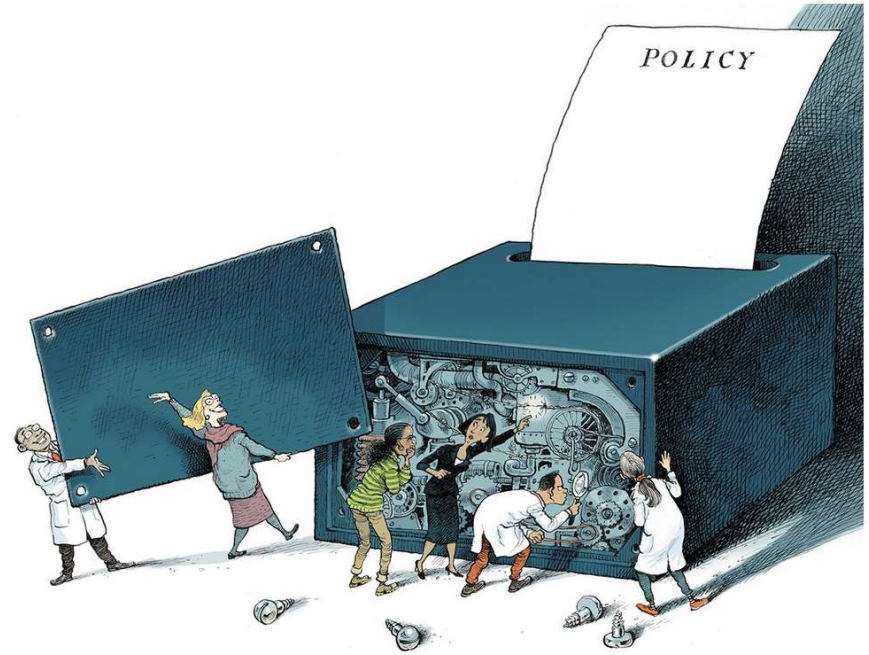
Match purpose and context

Mind the consequences

Quantification can backfire

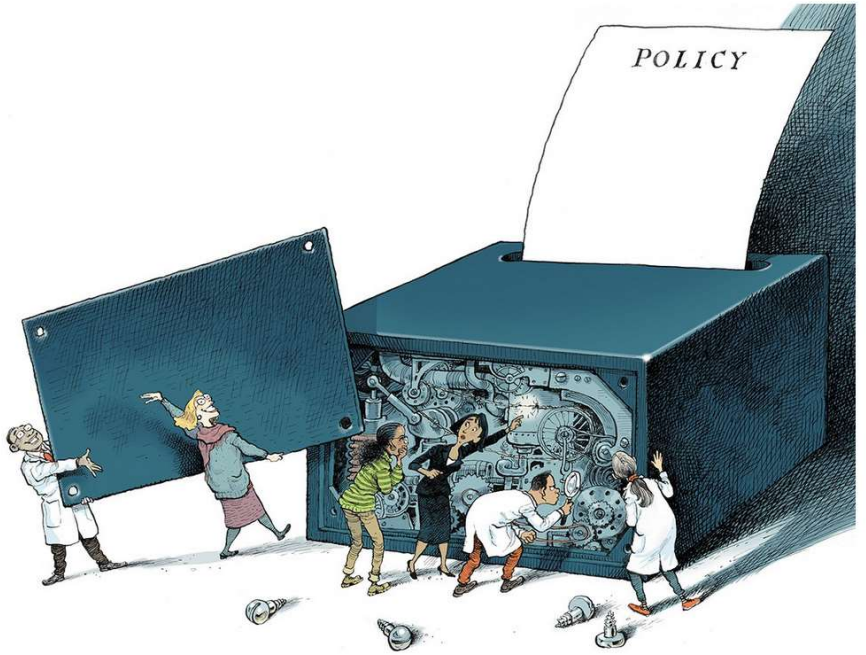
Mind the unknowns

Acknowledge ignorance



Mind the assumptions

Assess uncertainty and sensitivity

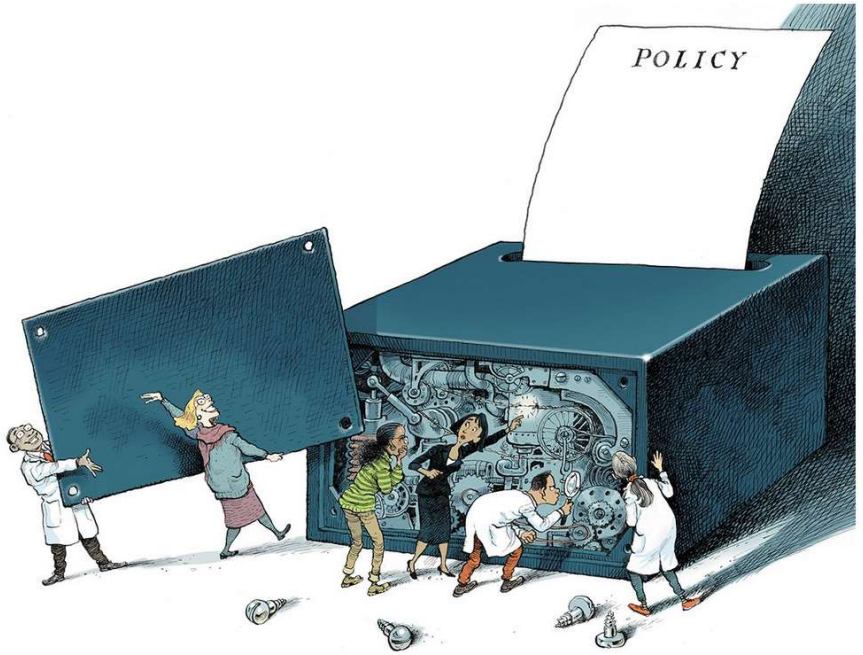


… assumptions that are reasonable in one situation can become nonsensical in another…

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

Mind the assumptions

Assess uncertainty and sensitivity



… to mitigate these issues: perform global uncertainty and sensitivity analyses … to make interesting discoveries

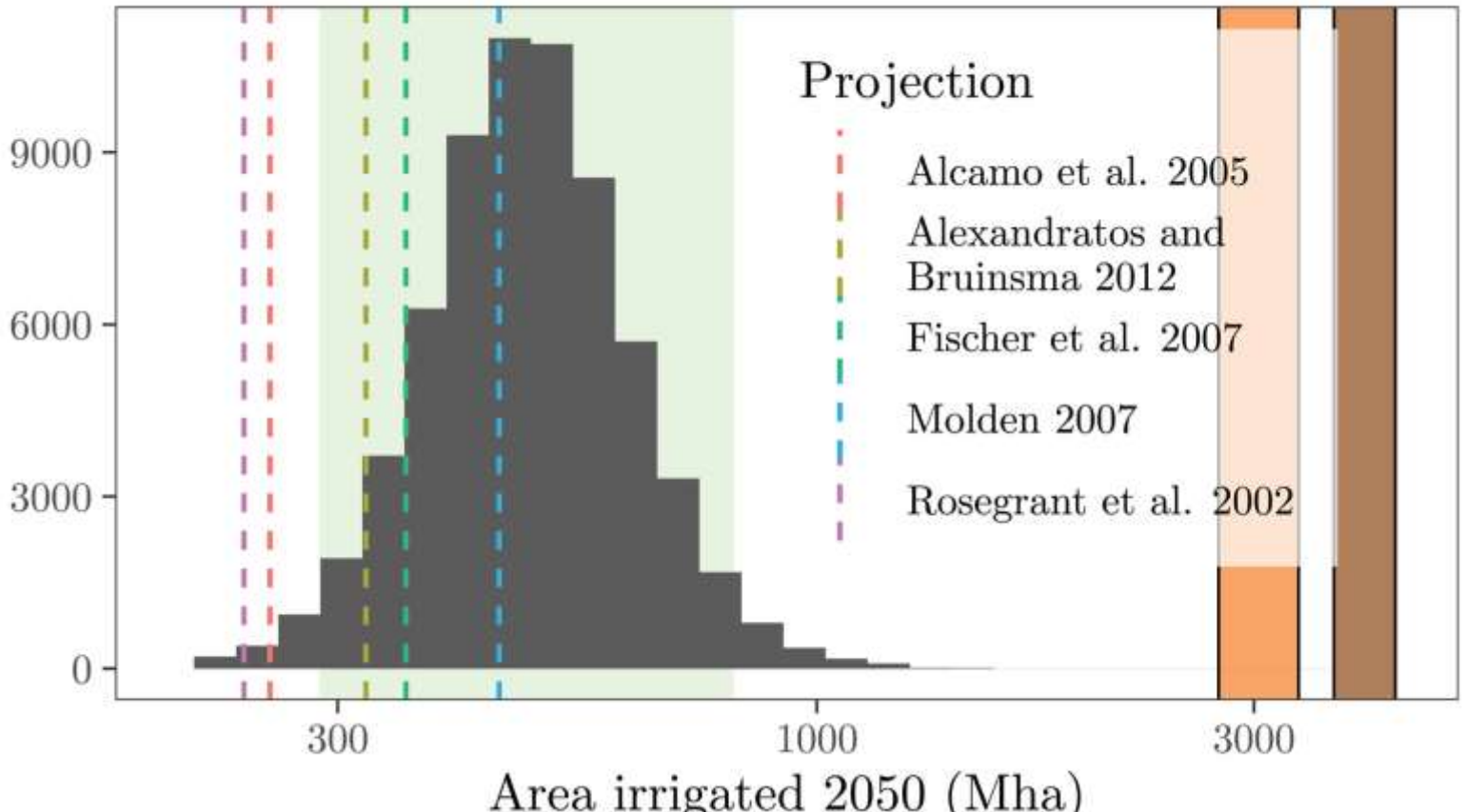
Geophysical Research Letters

Research Letter | Open Access

Current Models Underestimate Future Irrigated Areas

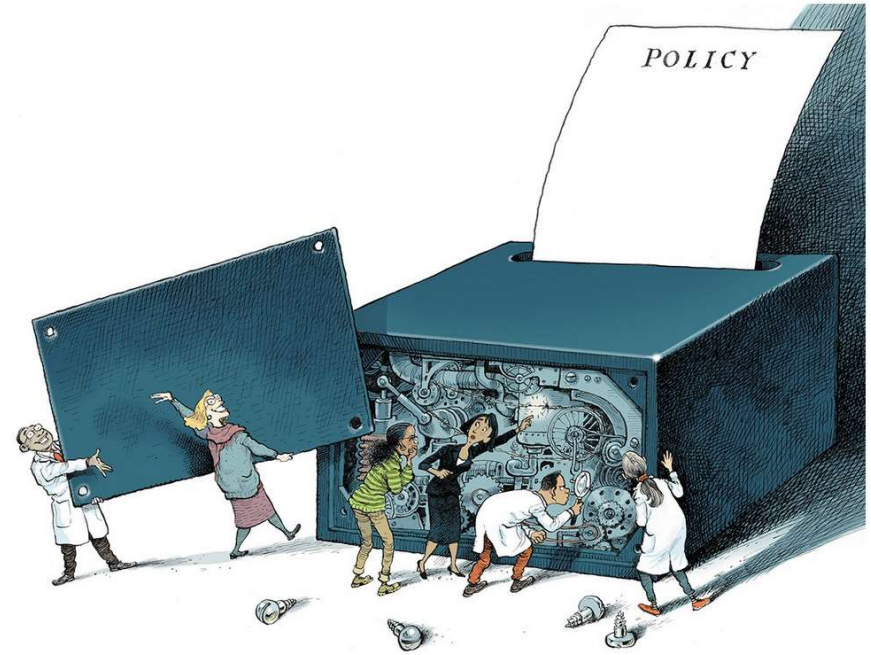
A. Puy ✉, S. Lo Piano, A. Saltelli

First published: 17 April 2020 | https

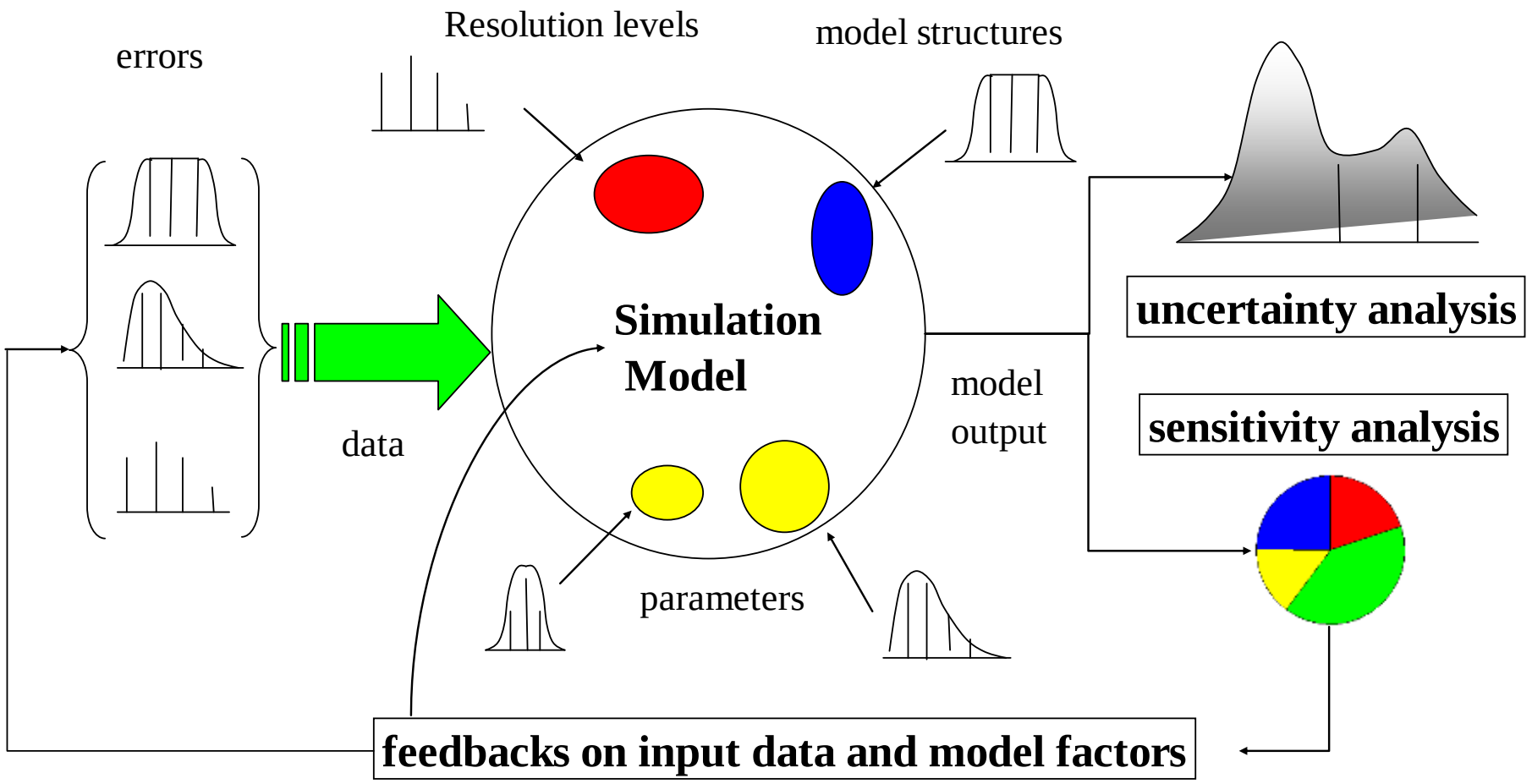


SUPPLEMENTARY INFORMATION

1. Additional information and references



An engineer's vision of UA, SA

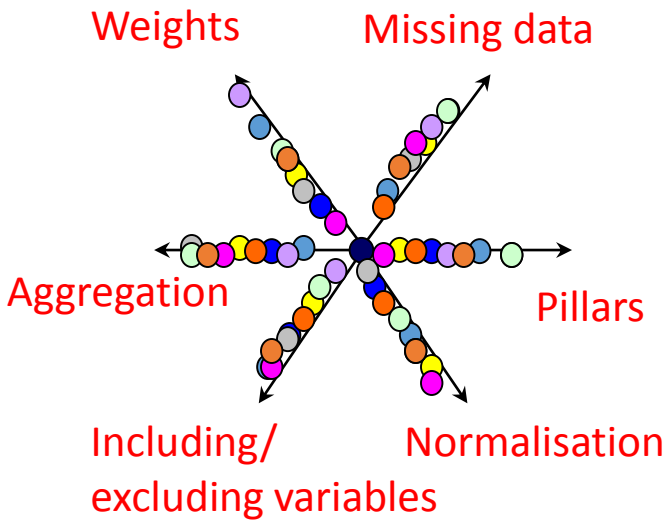


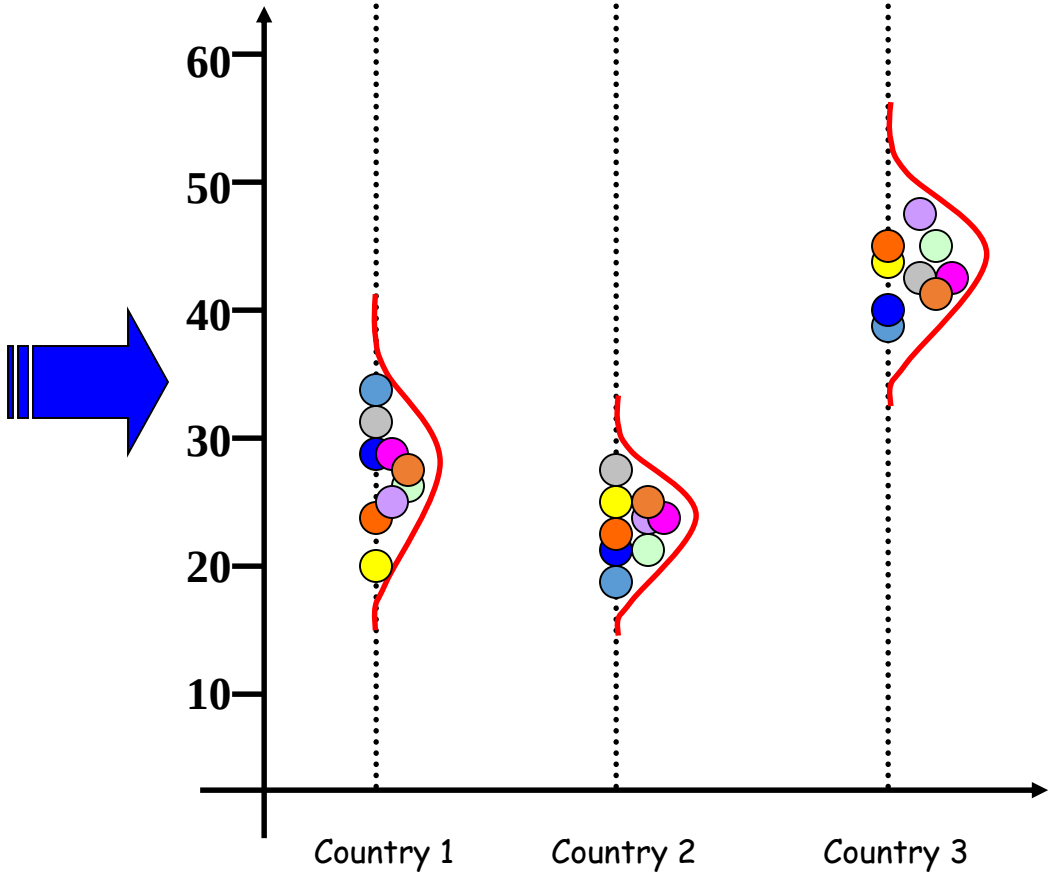
One can sample more than just factors:

- 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

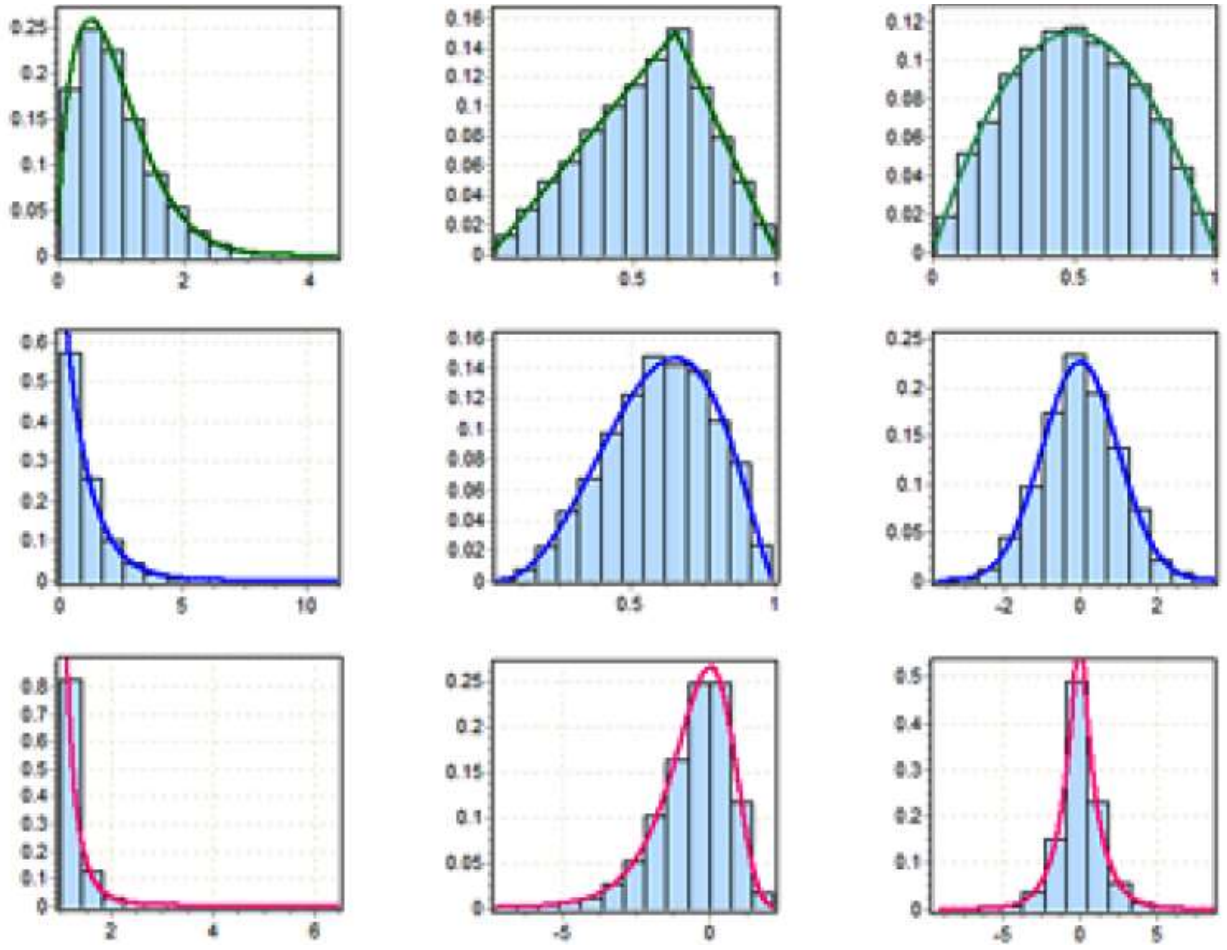
Space of alternatives





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

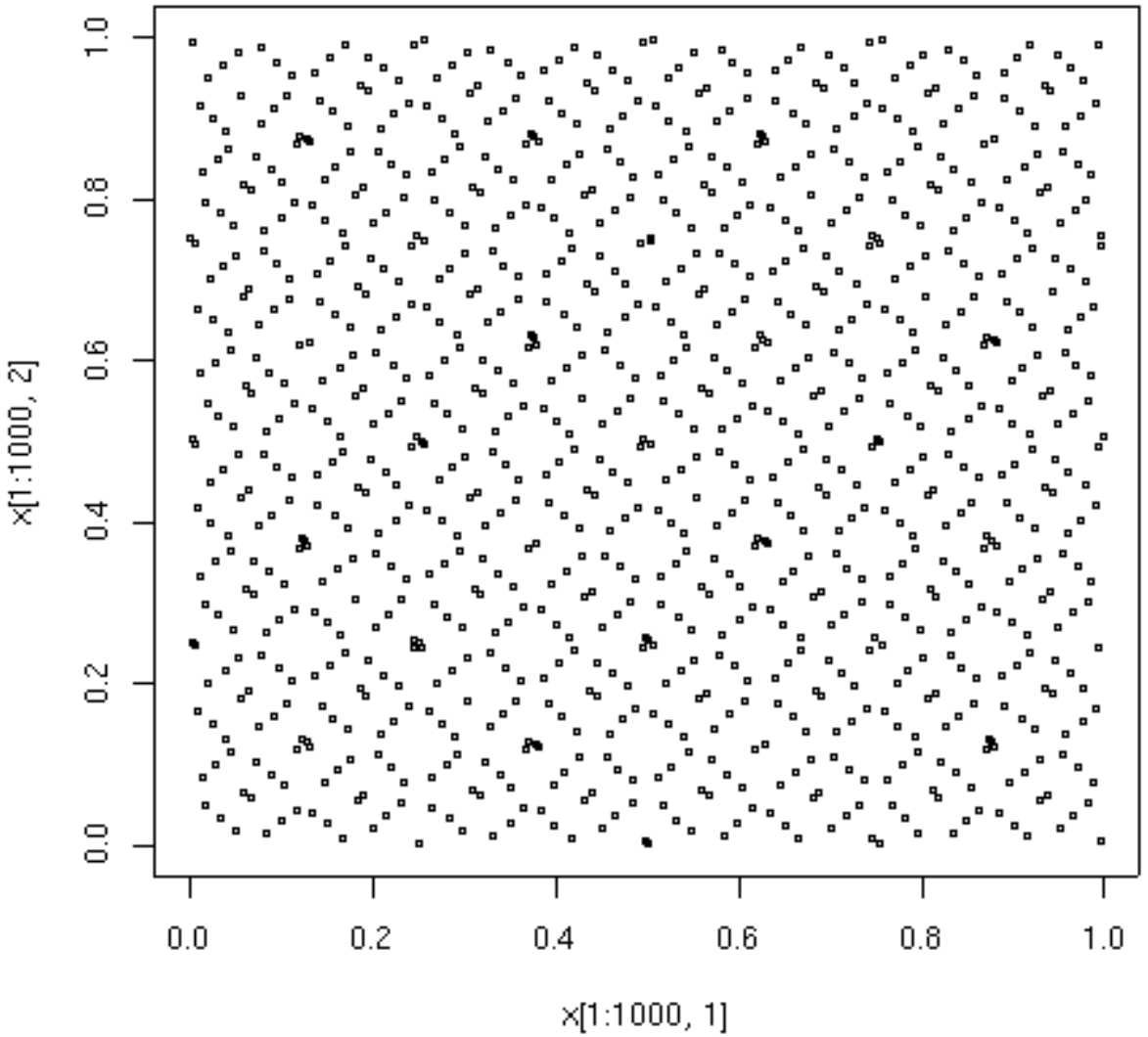
Each column is a sample from the distribution of a factor
Each row is a sample trial to generate a value of y



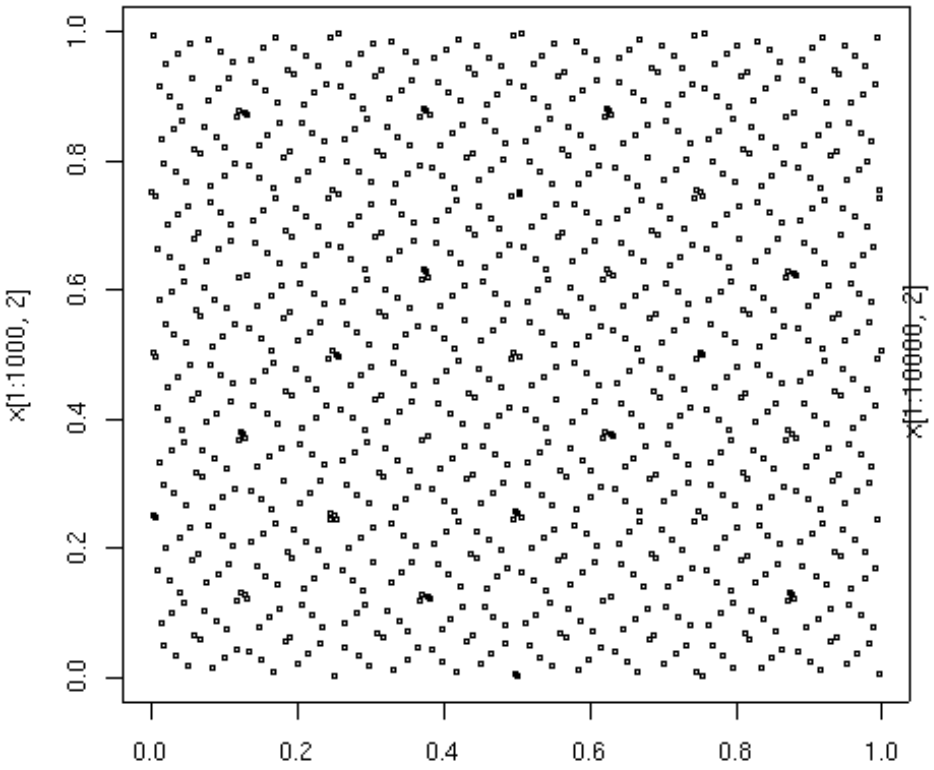
Examples of distributions of input factors

How to generate the random sample?

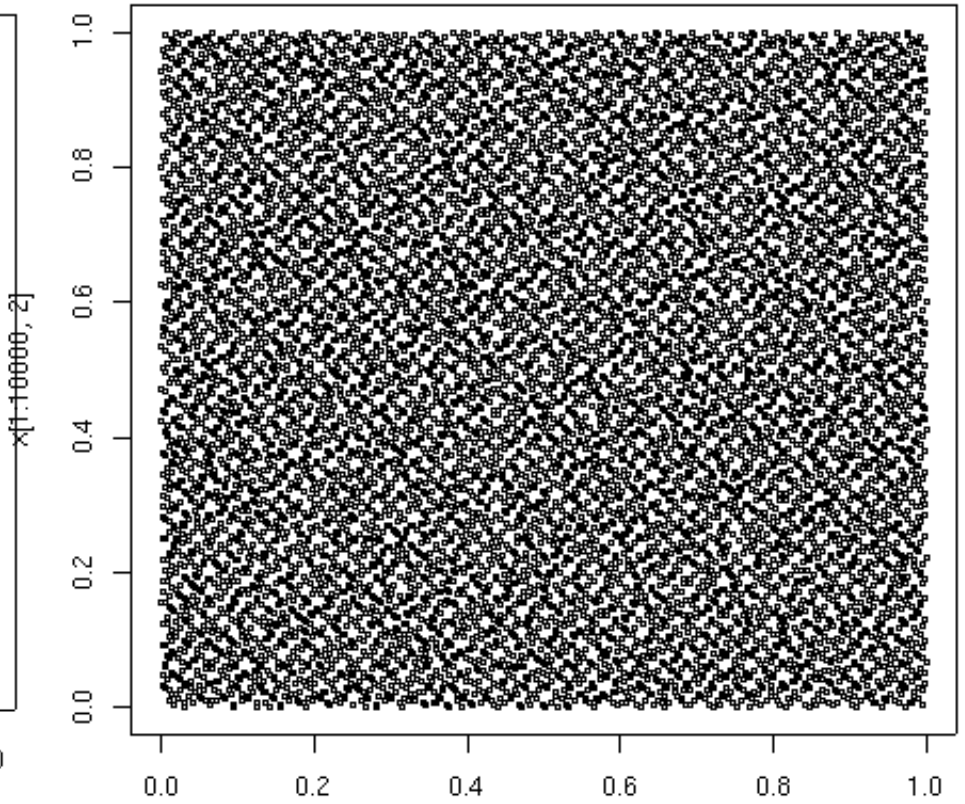
Quasi random sequences developed by I.M. Sobol’



An LP_τ sequence

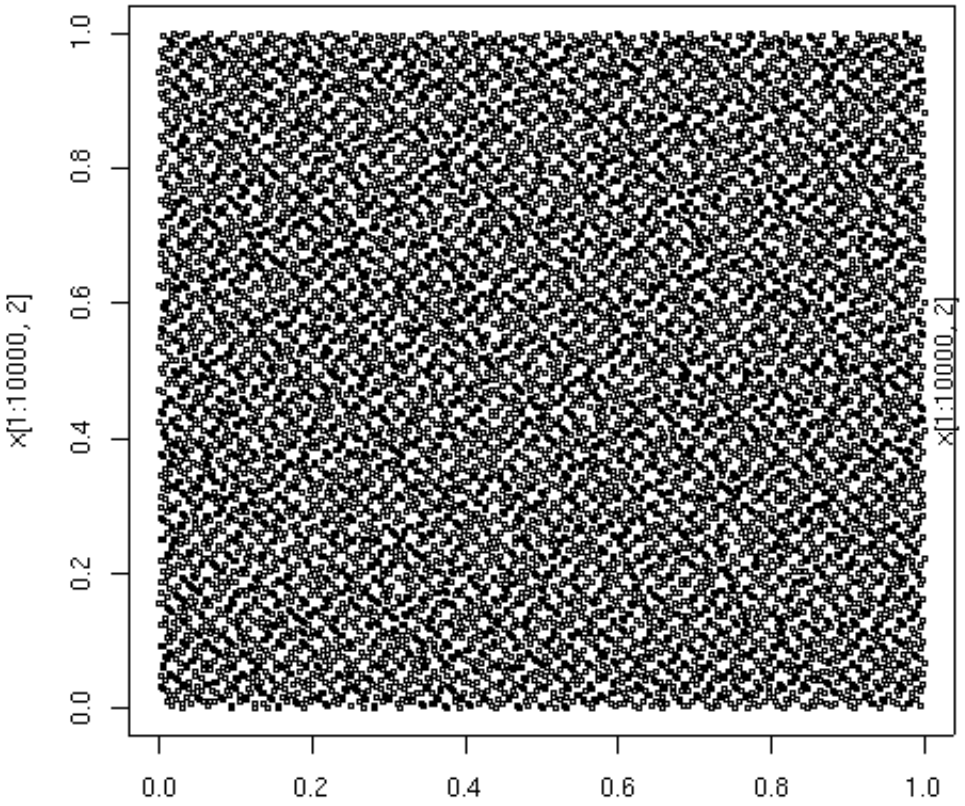


X1,X2 plane, 1000 Sobol' points

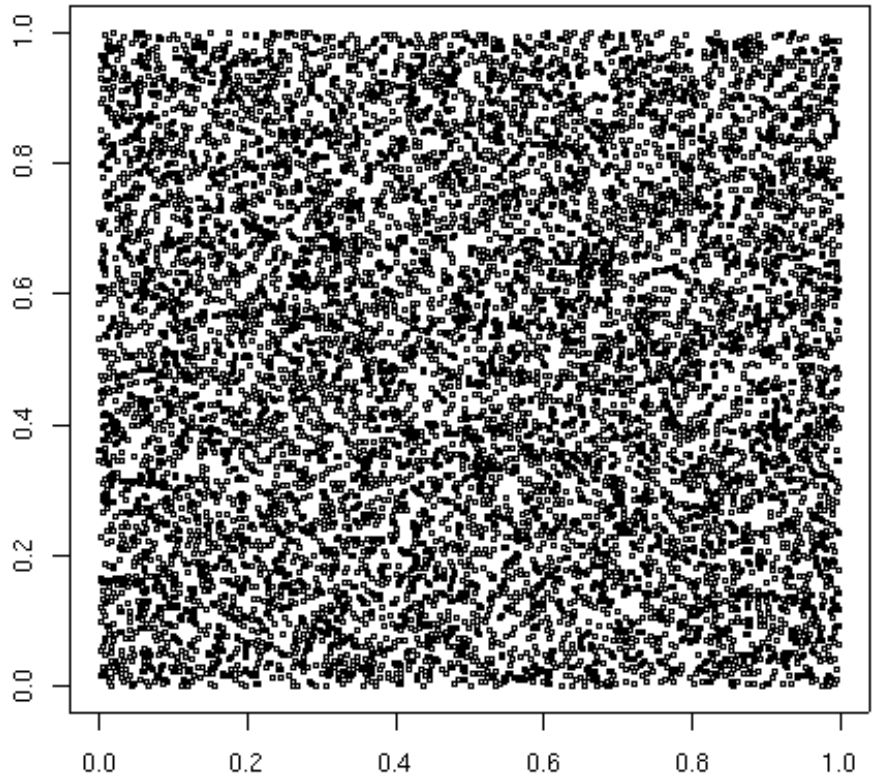


X1,X2 plane, 10000 Sobol' points

Sobol' sequences of quasi-random points



X1,X2 plane, 10000 Sobol' points



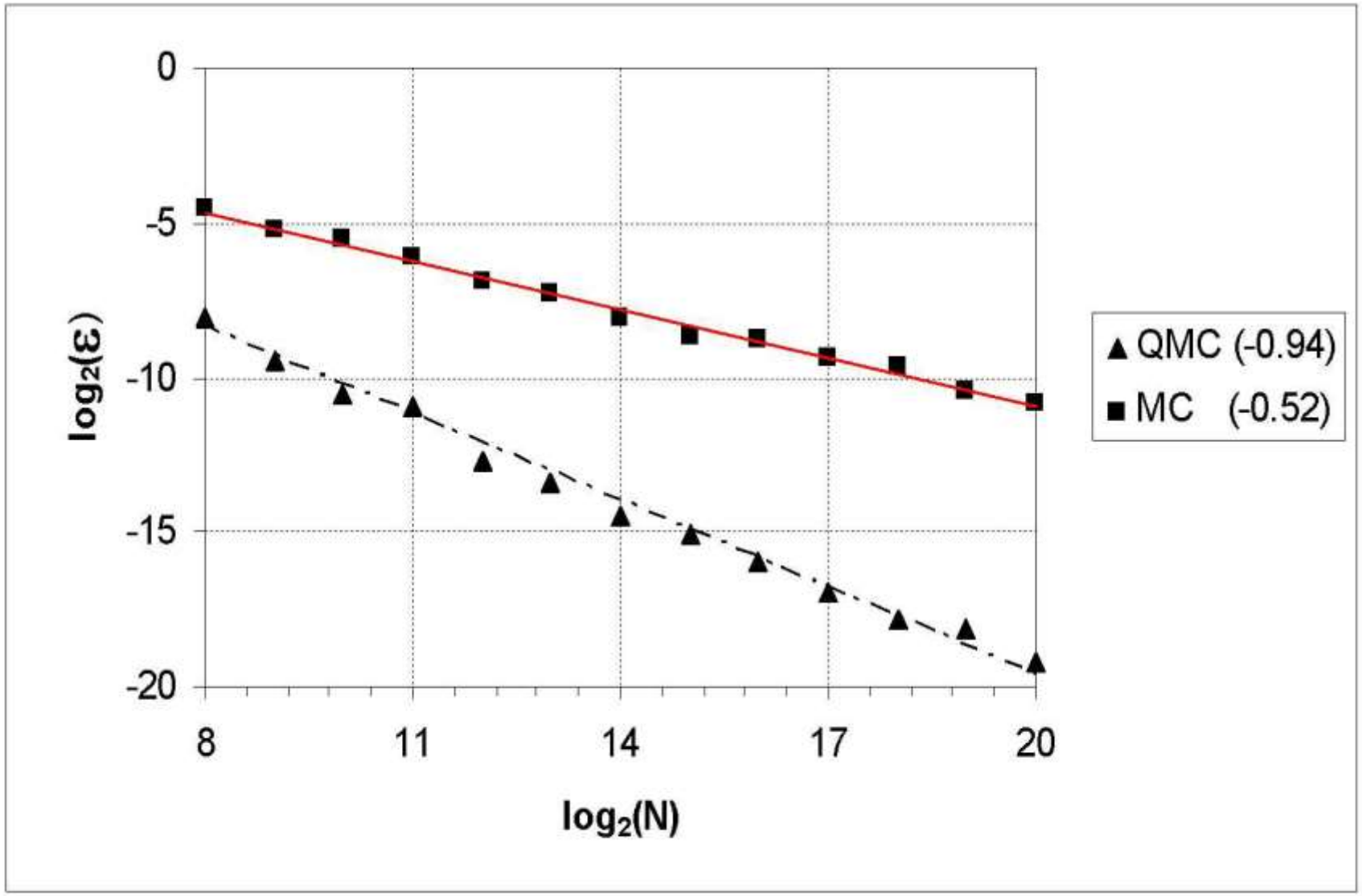
X1,X2 plane, 10000 random points

Sobol' sequences of quasi-random points against random points

Why quasi-random: they have faster convergence

Sergei Kucherenko,
Imperial College London

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.

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

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 Sensitivity analysis?

It can answer interesting questions

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

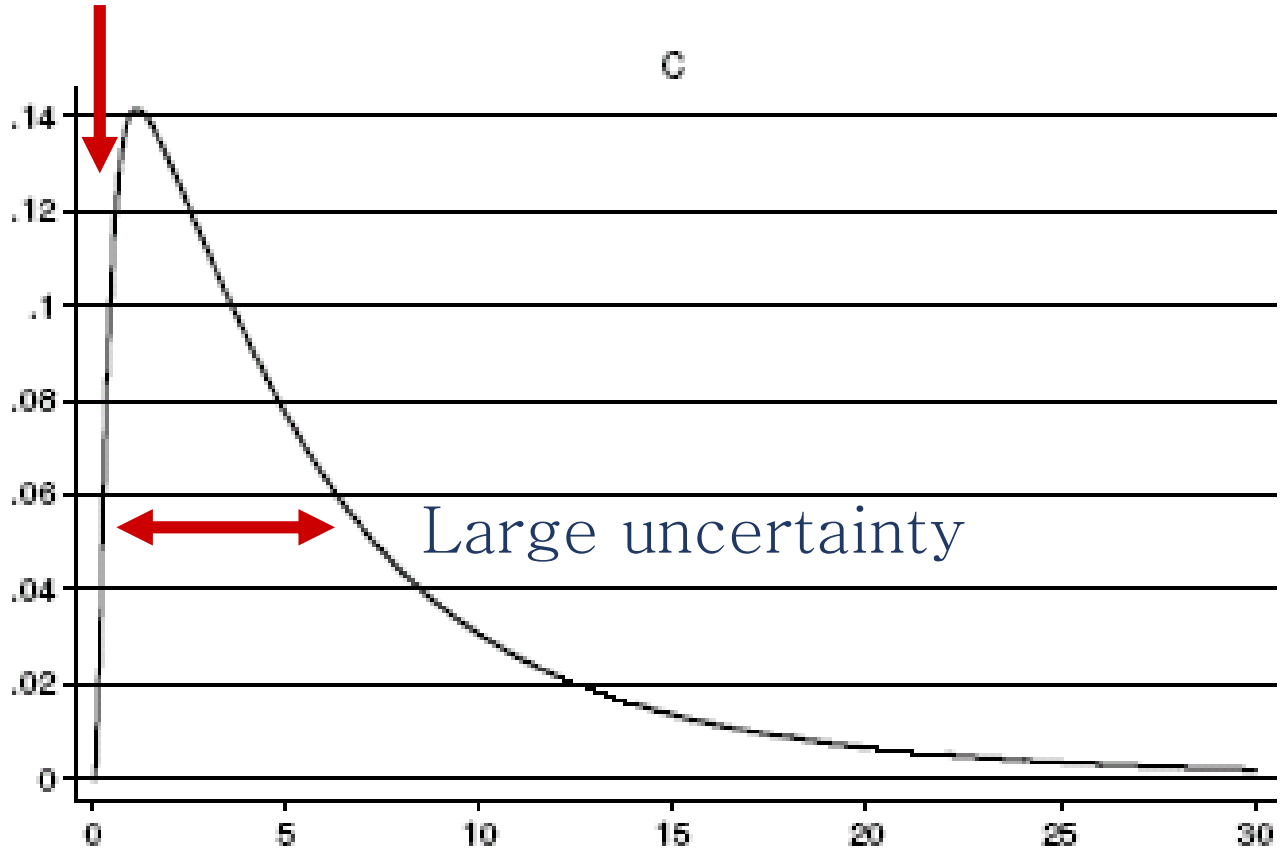
William Nordhaus, University of Yale Nobel 'Economics' 2018

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).

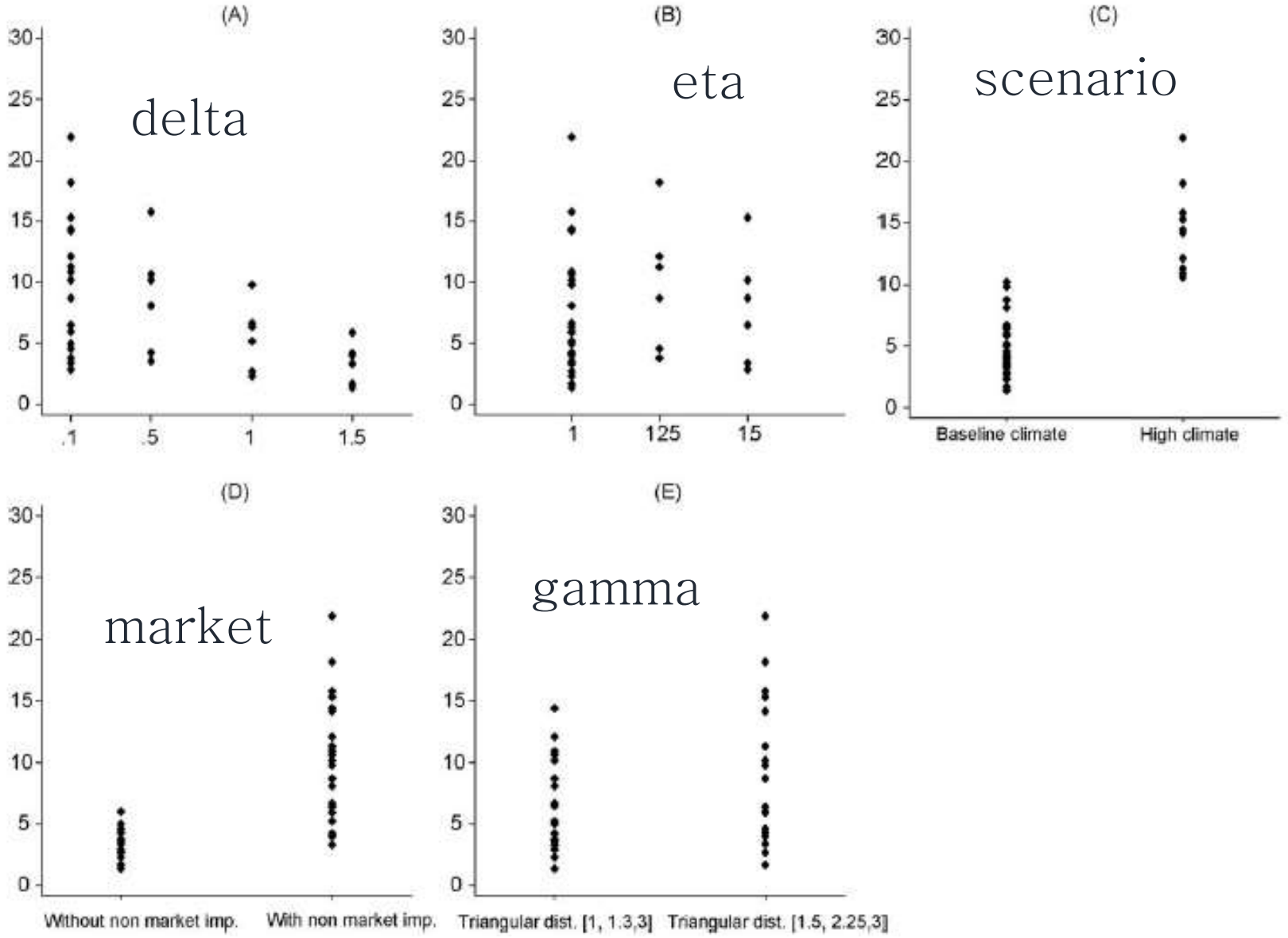
How was it done? A reverse engineering of the analysis

Missing points



% loss in GDP per capita

Sensitivity analysis here (by reverse engineering)



OPINION

PETER COY

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

Sept. 17, 2021

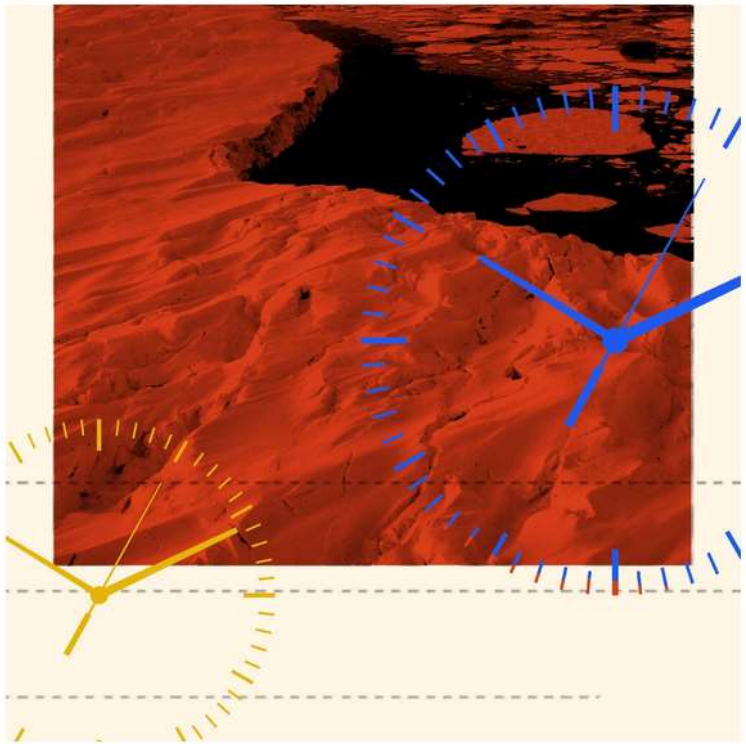


Illustration by Arsh Raziuddin, The New York Times

“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

Why sensitivity analysis?

It allows interesting discoveries

Journal of the Royal Statistical Society
SERIES A Statistics in Society

J. R. Statist. Soc. A (2013)
176, *Part* 3, *pp.* 609–634

Ratings and rankings: voodoo or science?

Paolo Paruolo

University of Insubria, Varese, Italy

and Michaela Saisana and Andrea Saltelli

European Commission, Ispra, Italy

Why sensitivity analysis

University rankings such as ARWU and THES are technically unsound

Journal of the Royal Statistical Society
SERIES A Statistics in Society

J. R. Statist. Soc. A (2013)
176, *Part* 3, *pp.* 609–634

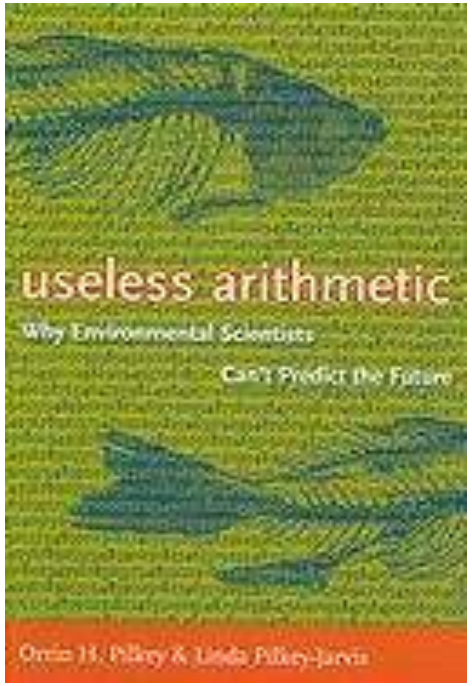
Ratings and rankings: voodoo or science?

Paolo Paruolo
University of Insubria, Varese, Italy
and Michaela Saisana and Andrea Saltelli
European Commission, Ispra, Italy

Why sensitivity analysis

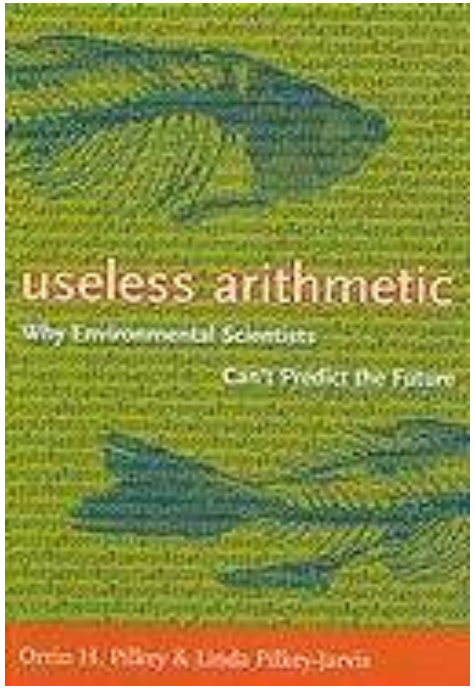
Limits of sensitivity analysis

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.

The map is not the territory

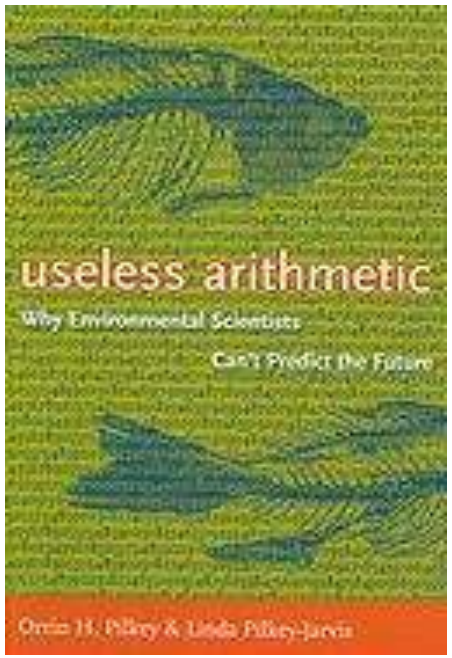


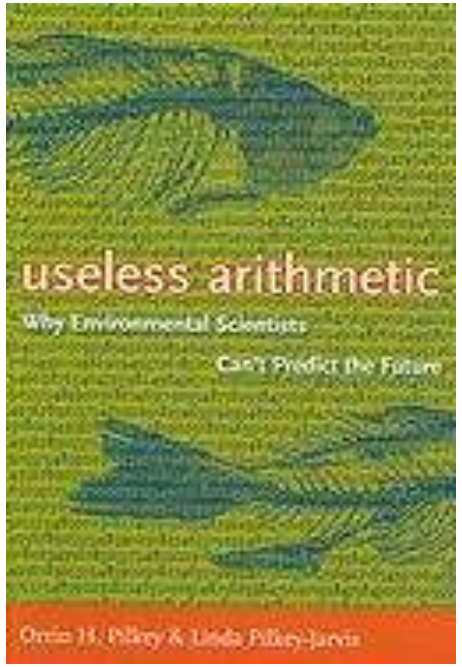
<<It is important, however, to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature.

[…] 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.>>

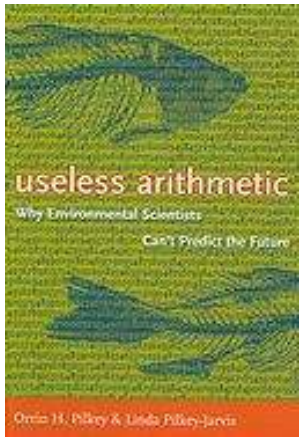
One of the examples discussed concerns the Yucca Mountain repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis.

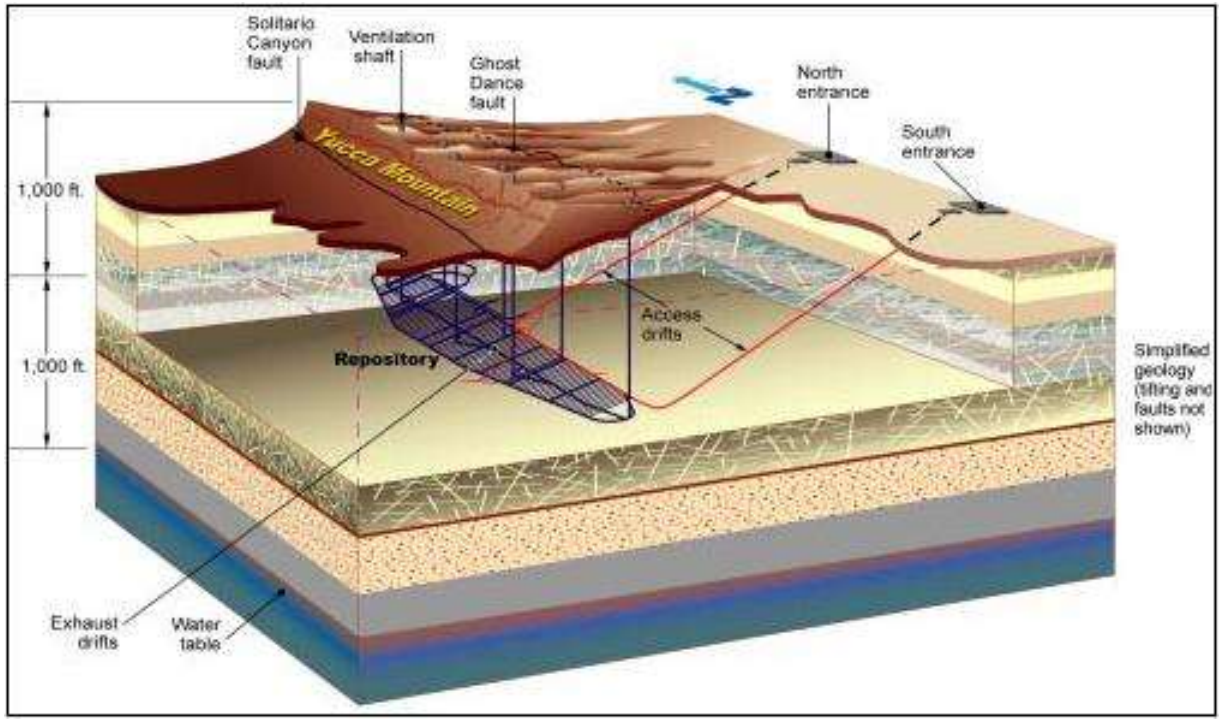
TSPA is Composed of 286 sub-models.





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.





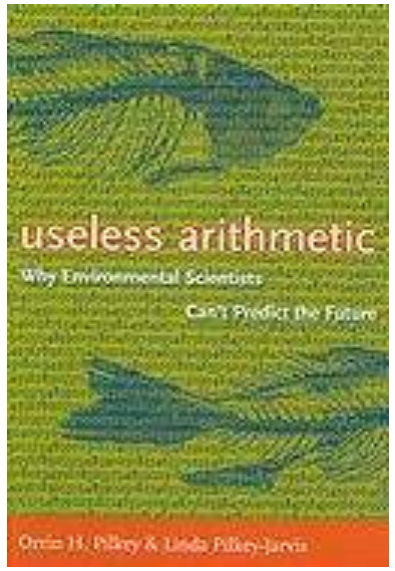
The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter

(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.

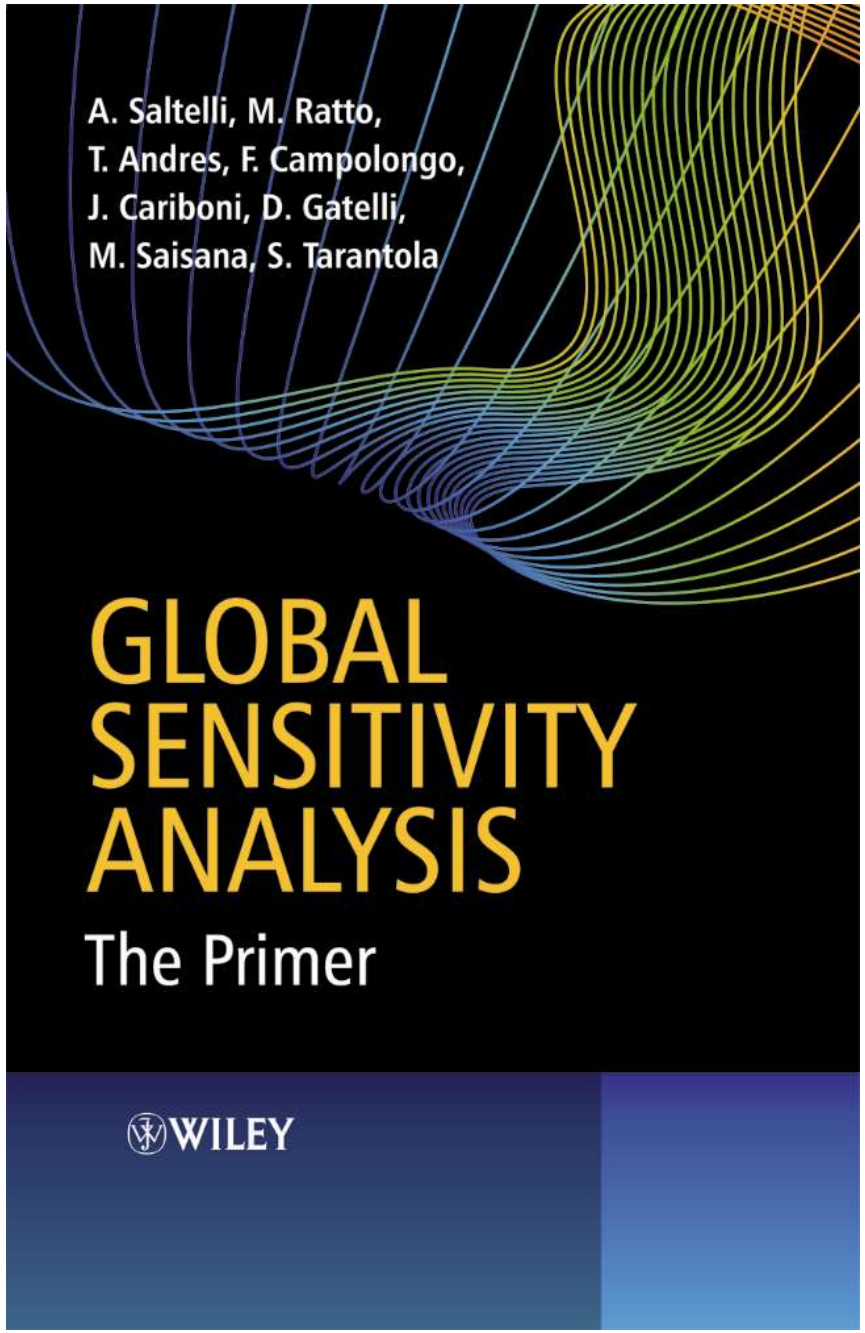
→… SA useless if it is instead ~ 3,000 millimetres per year.

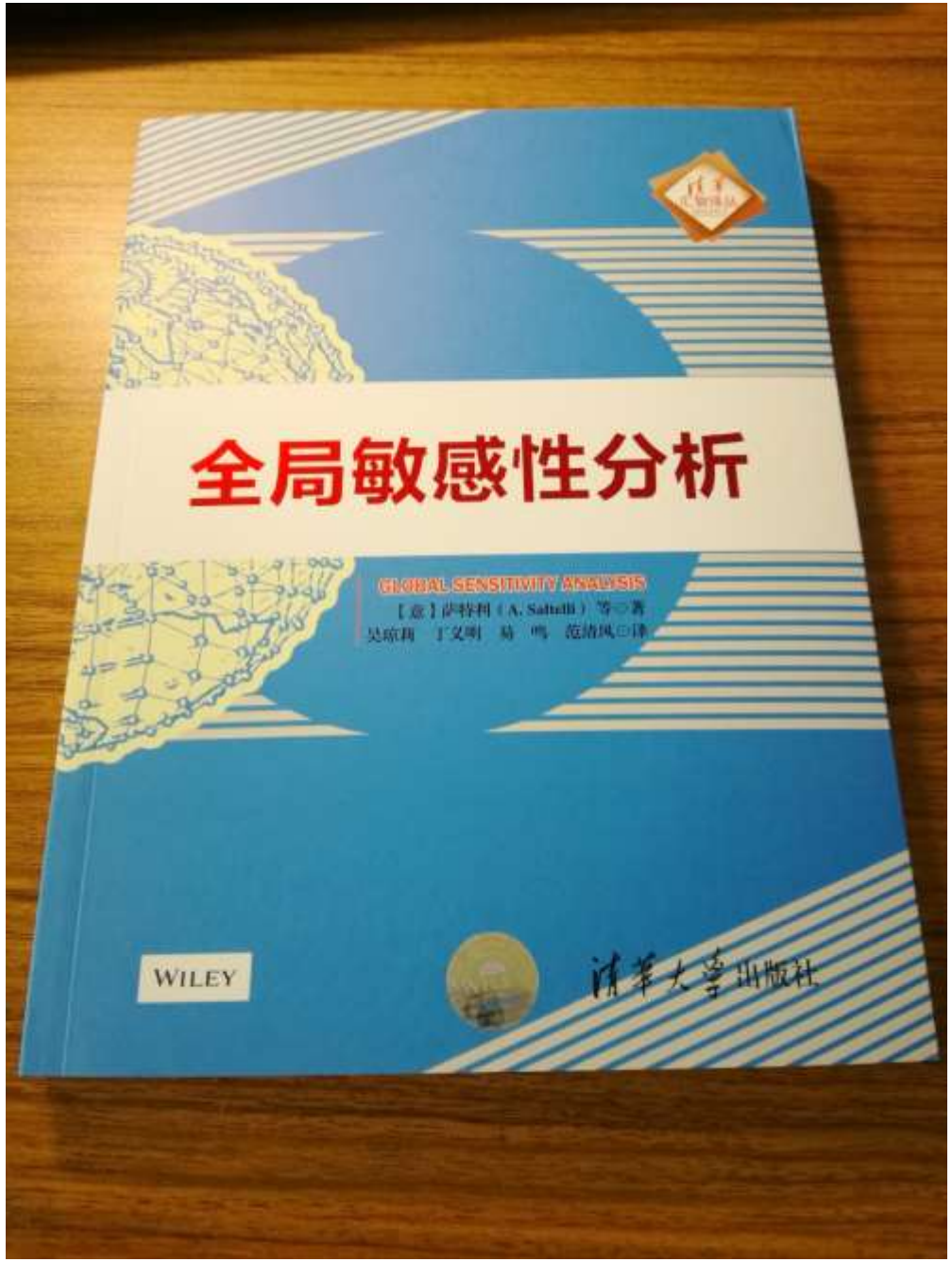


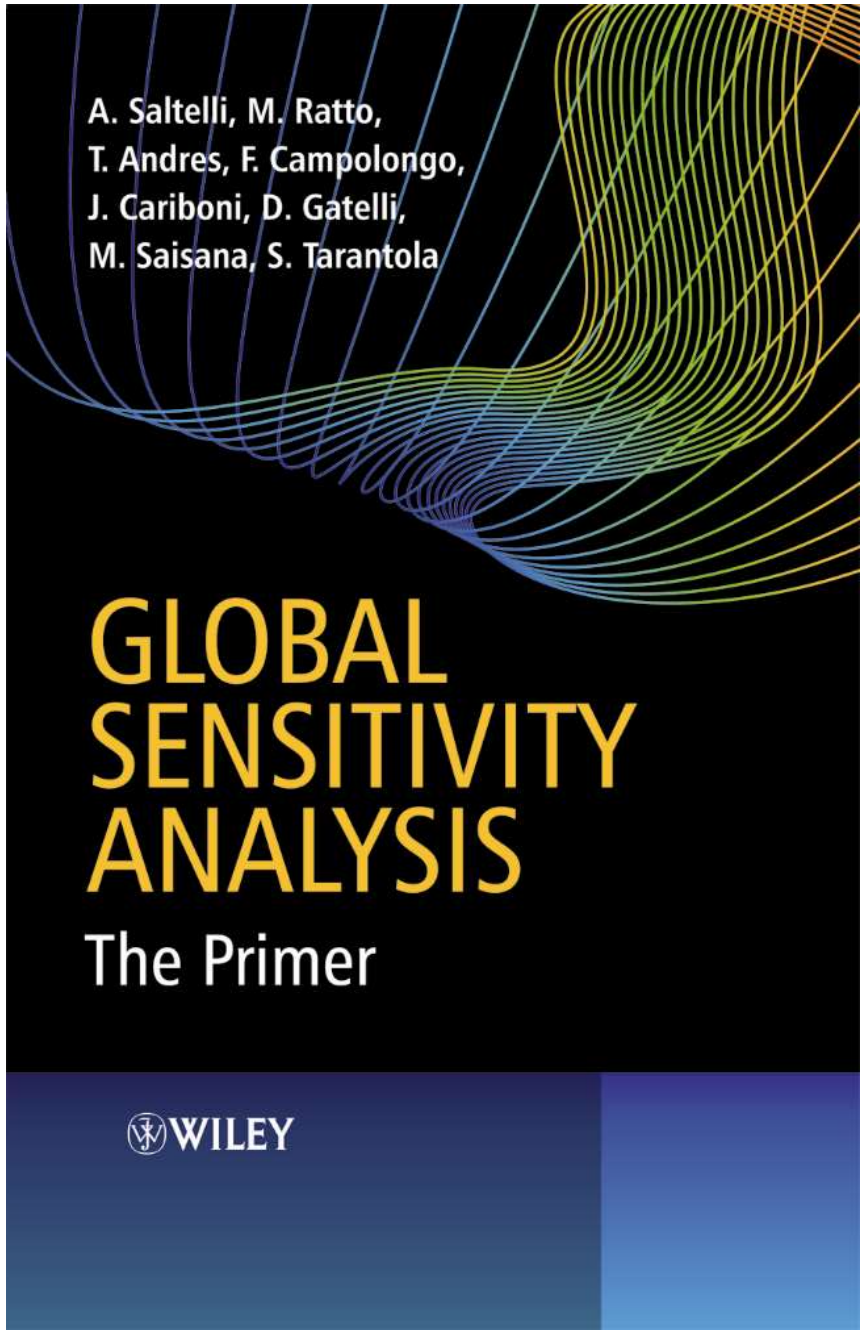
"Scientific mathematical modelling should involve constant efforts to falsify the model"

➔ Organized skepticism (as per CUDOS)

Where to study sensitivity analysis?


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

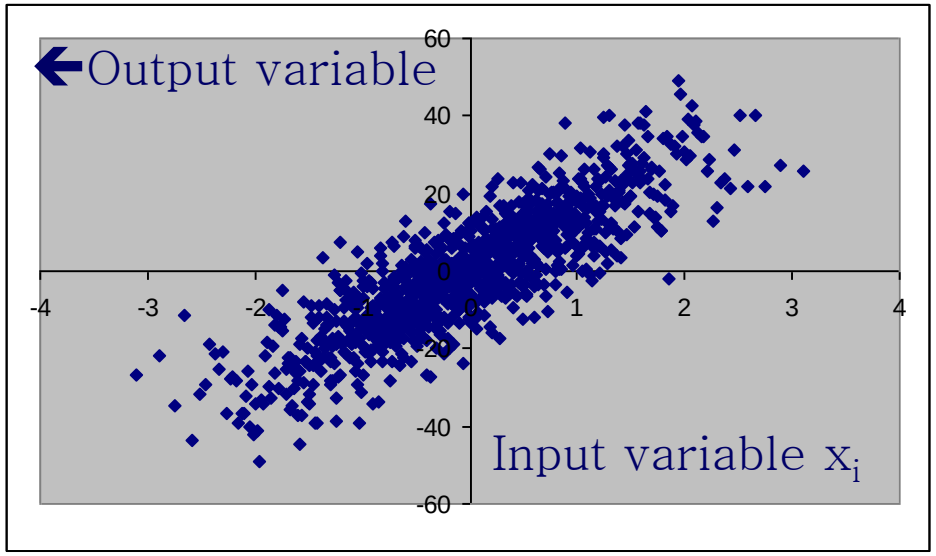

全局敏感性分析
GLOBAL SENSITIVITY ANALYSIS
WILEY

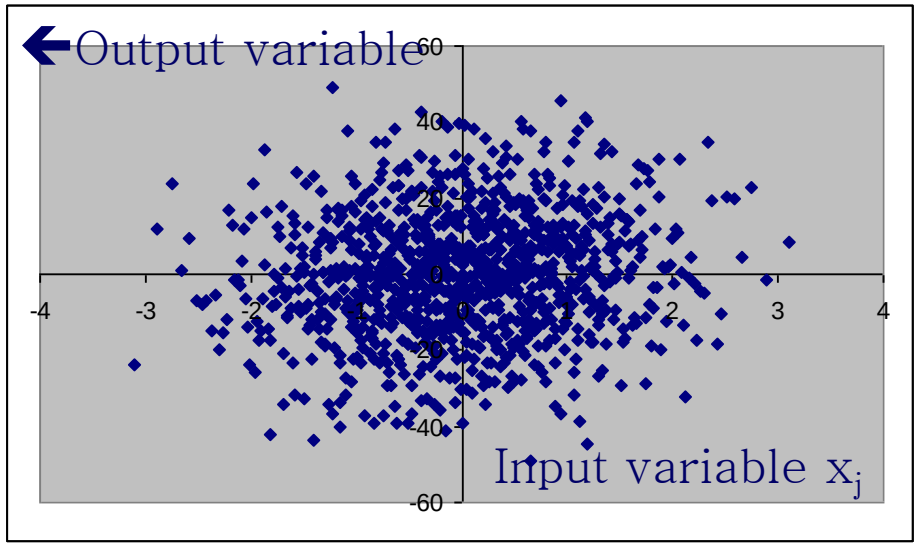


Available for free at

http://www.andreasaltelli.eu

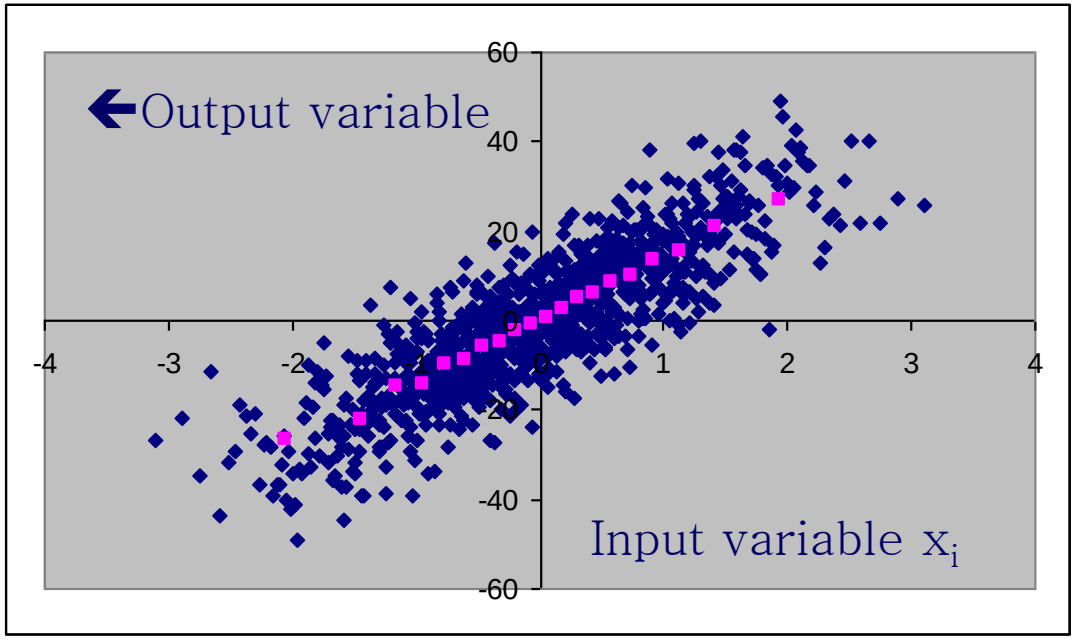
How is it done in practice?

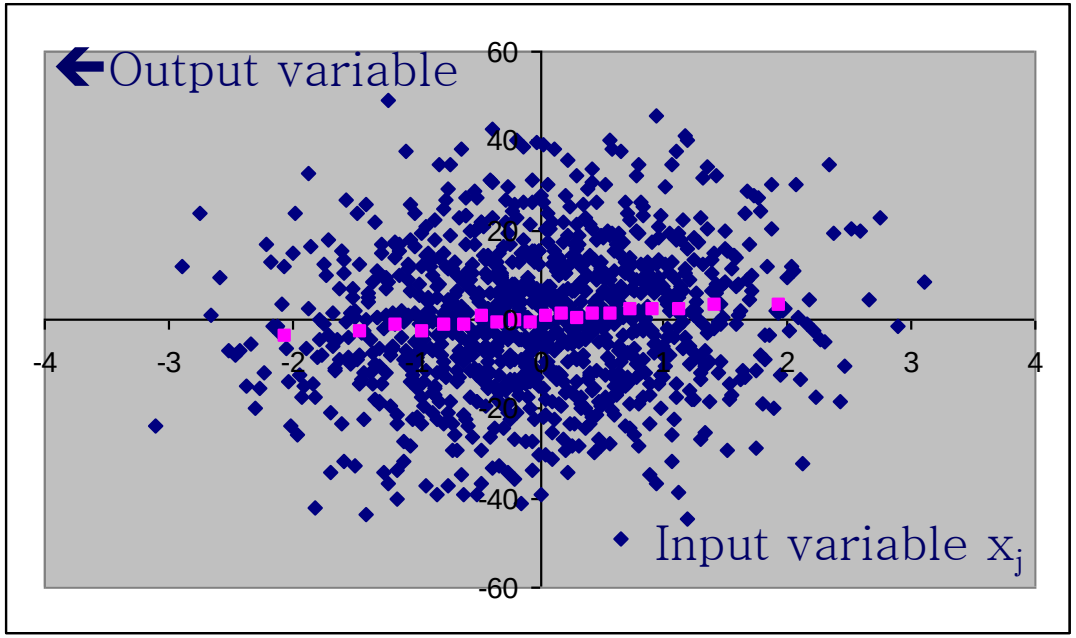




Plotting the output as a function of two different input factors

Which factor is more important?

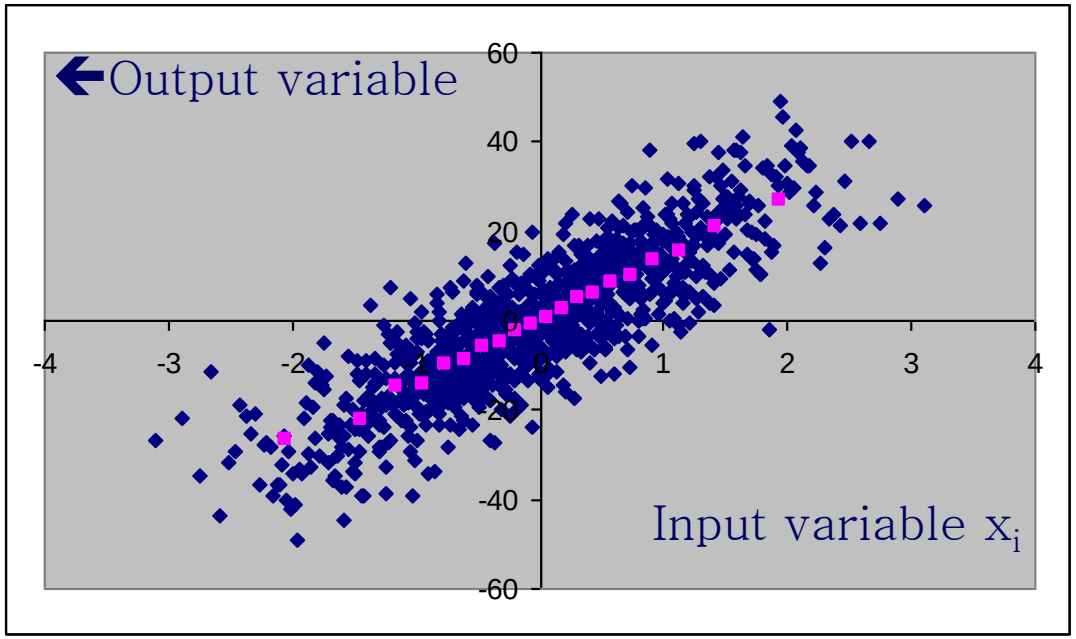




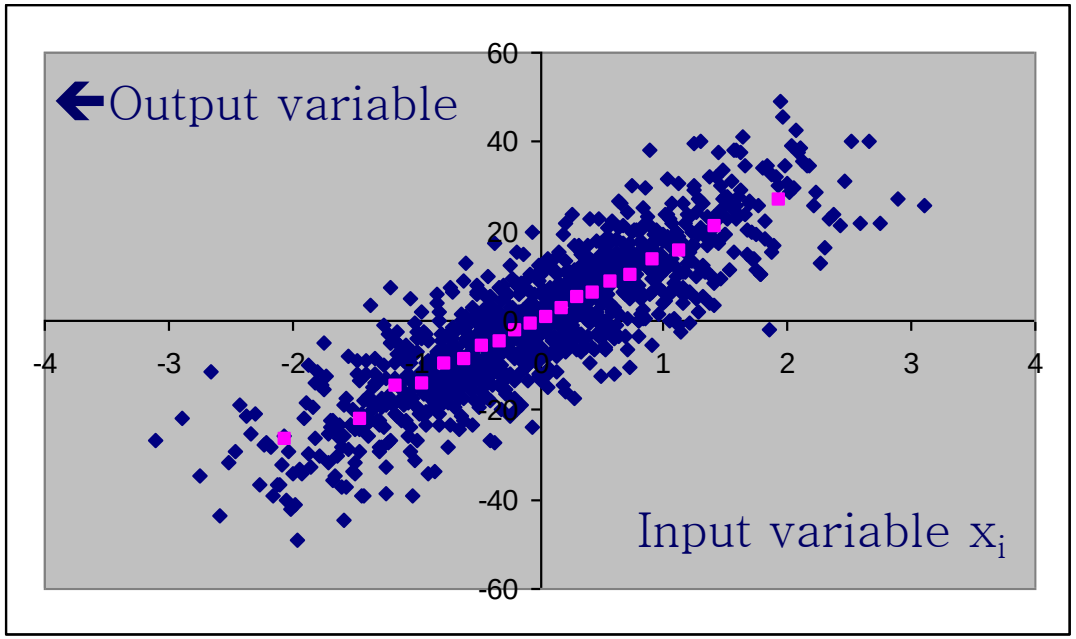
~1,000 blue points

Divide them in 20 bins of ~ 50 points

Compute the bin's average (pink dots)

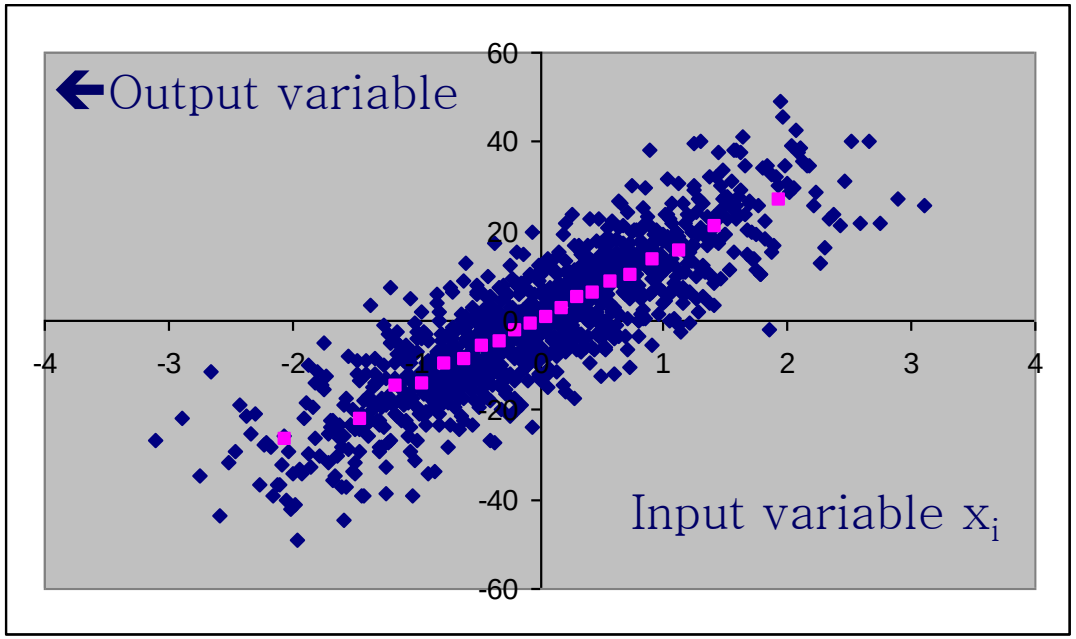


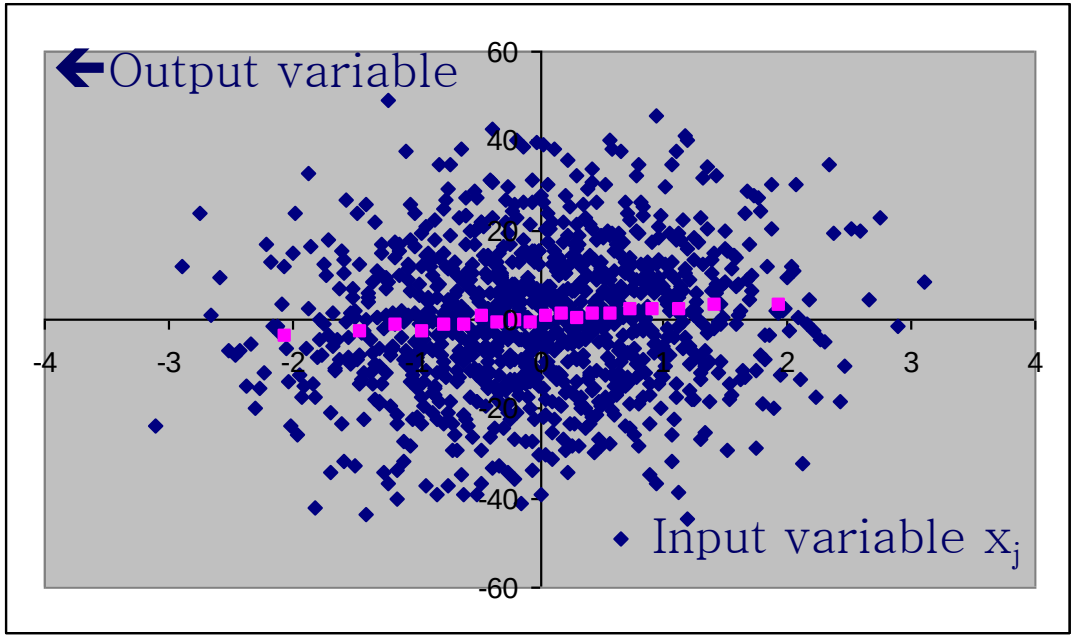
Each pink point is ~ $E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)$



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

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

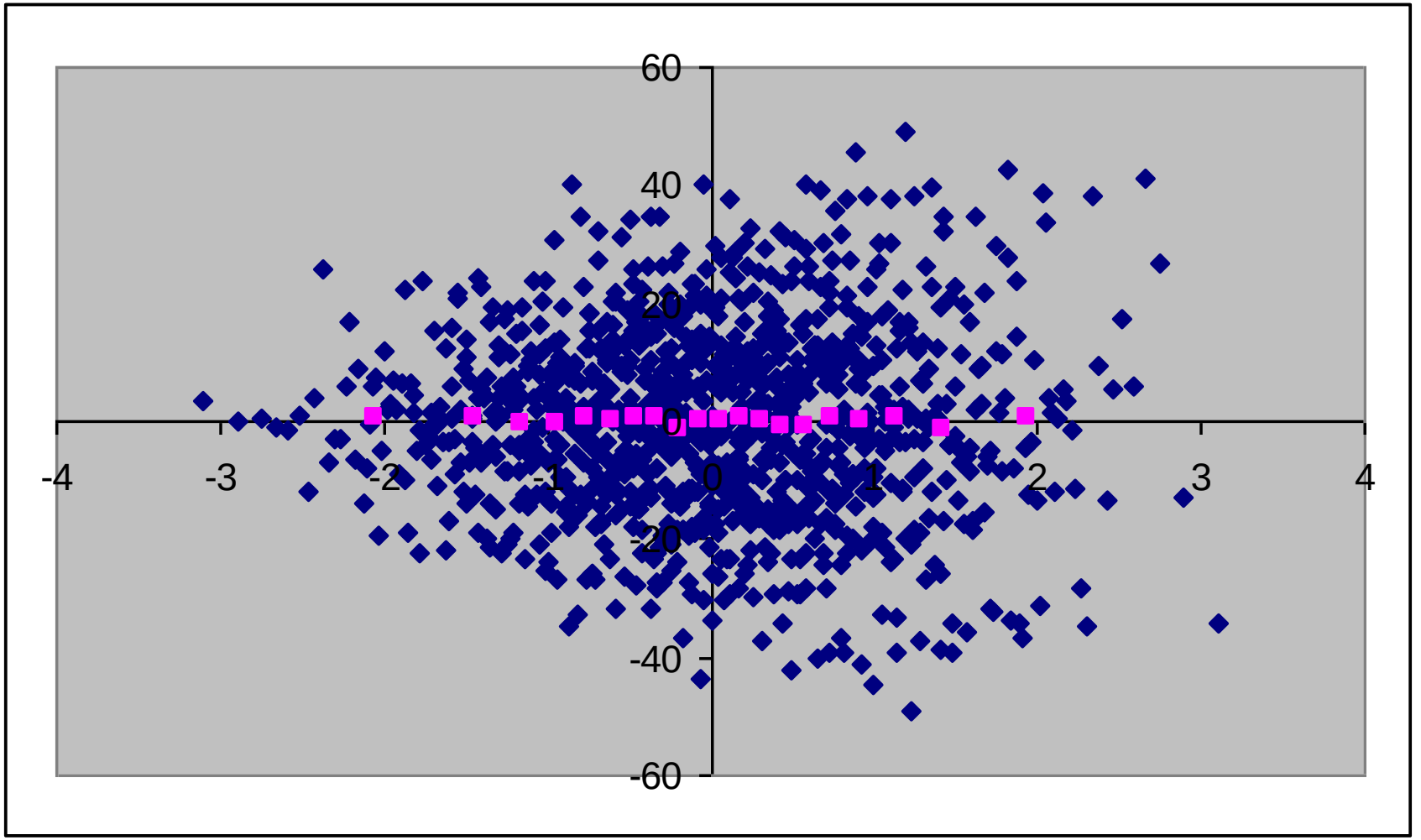
For additive models one can decompose the total variance as a sum of first order effects

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

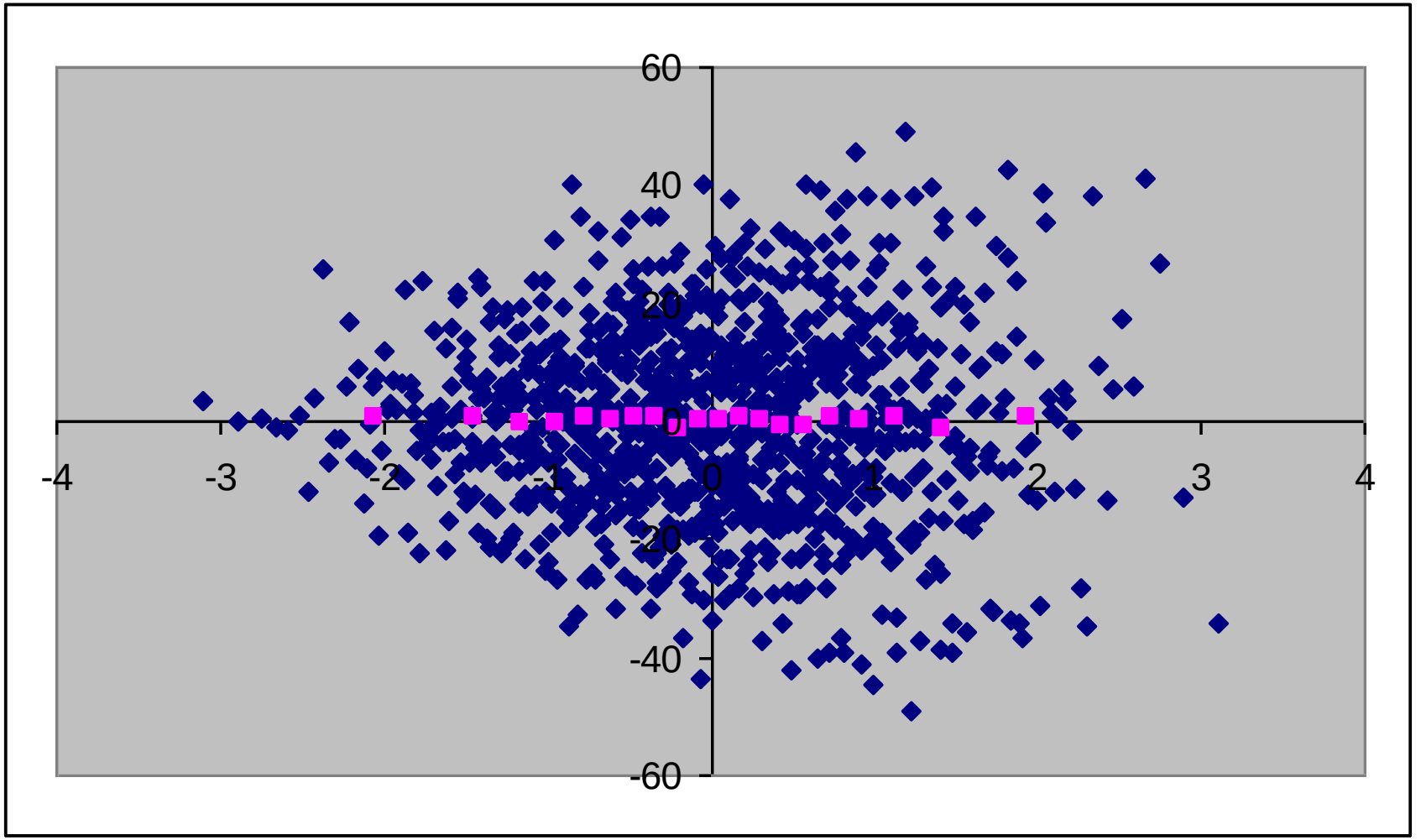
… which is also how additive models are defined

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)

When the factors are independent the total variance can be decomposed into main effects and interaction effects up to the order k, the dimensionality of the problem.

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 $Y=f(X_1,X_2,X_3)$

Instead of

$$V=V_1+V_2+V_3+$$
$$+V_{12}+V_{13}+V_{23}+$$
$$+V_{123}$$

Or – divided by V

$$1=S_1+S_2+S_3+$$
$$+S_{12}+S_{13}+S_{23}+$$
$$+S_{123}$$

We have:

$S_{T1} = S_1 + S_{12} + S_{13} + S_{123}$

(and analogue formulae for S_{T2}, S_{T3}) which can be computed without knowing S_1, S_{12}, S_{13}, S_{123}

S_{T1} is called a total effect sensitivity index

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

Total effect, or bottom marginal variance=

= the expected variance that would be left if all factors but Xi could be fixed (self evident definition)

$$S_{Ti} \equiv \frac{E\left(V\left(Y|\mathbf{X}_{\sim i}\right)\right)}{V_Y}$$

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/

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

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

Reliability Engineering and System Safety **28** (1990) 229–253

Non-parametric Statistics in Sensitivity Analysis for Model Output: A Comparison of Selected Techniques

A. Saltelli

Commission of the European Communities, Joint Research Centre—Ispra Establishment, 21020 Ispra (Varese), Italy

&

J. Marivoet

Belgian Nuclear Research Establishment SCK/CEN, Boeretang 200, B-2400, Belgium

(Received 26 May 1989; accepted 3 August 1989)

Computational Statistics & Data Analysis 13 (1992) 73–94
North-Holland

Sensitivity analysis for model output

Performance of black box techniques on three international benchmark exercises

A. Saltelli

Commission of the European Communities, Joint Research Centre, Ispra, Italy

T. Homma

Japan Atomic Energy Research Institute, Tokai Research Establishment, Department of Environmental Safety Research, Tokai-Mura, Ibaraki, Japan

Received February 1990
Revised October 1990

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

Secrets of sensitivity analysis

Why should one ever run a model just once?

First secret: The most important question is the question.

Or: sensitivity analysis is not “run” on a model but on a model once applied to a question

Second secret: Sensitivity analysis should not be used to hide assumptions [it often is]

Third secret: If sensitivity analysis shows that a question cannot be answered by the model one should find another question or model

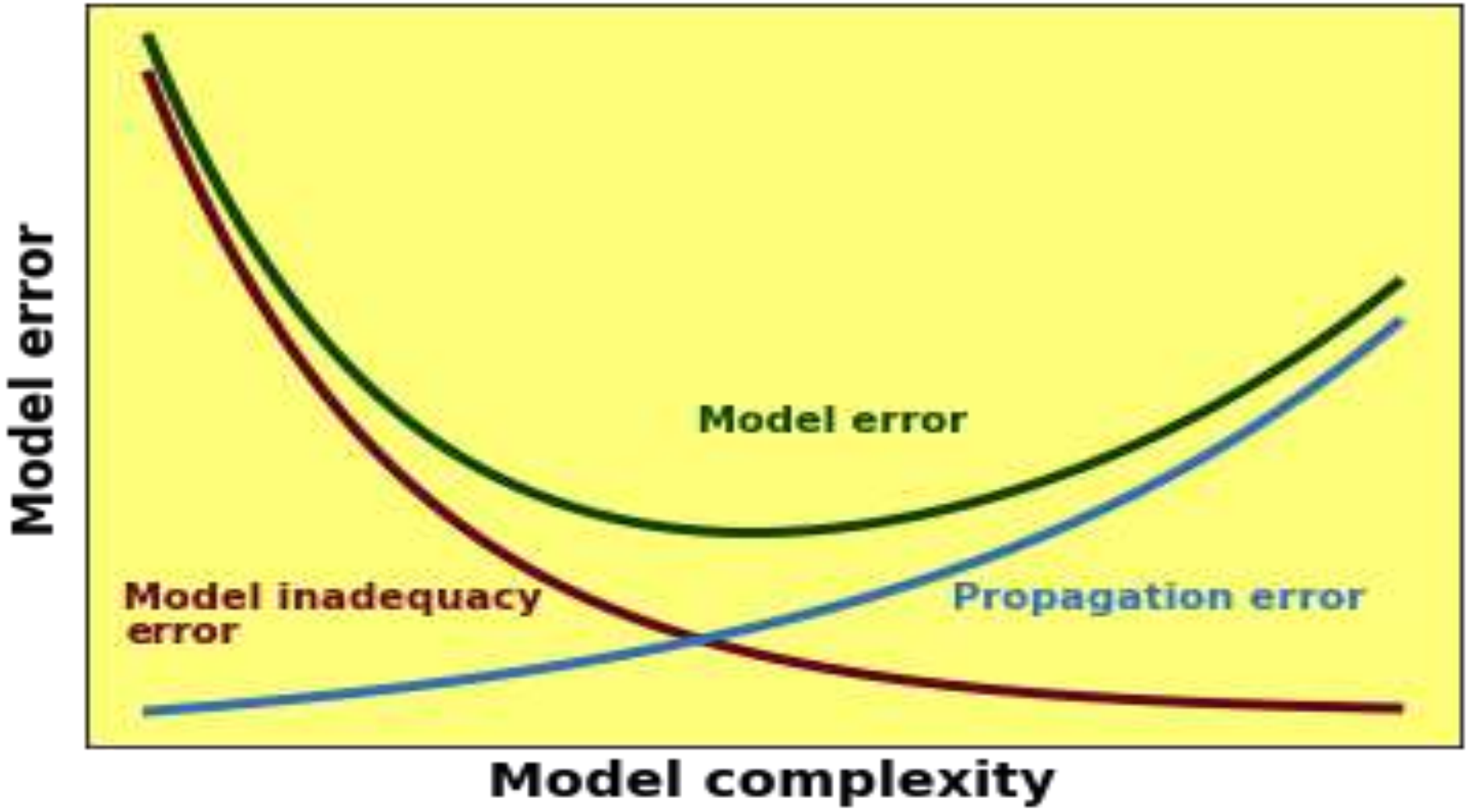
[Often the love for one's own model prevails]

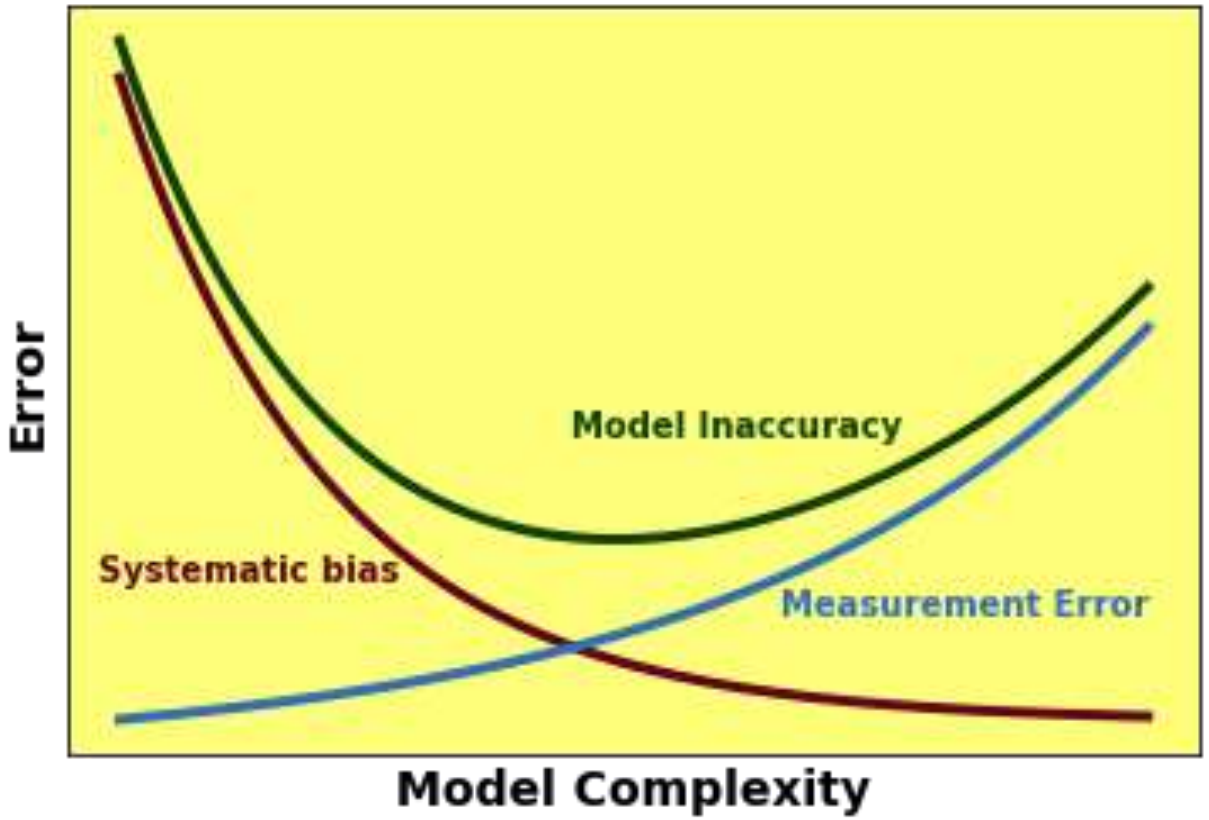
Fourth (badly kept) secret:

There is always one more bug!

=Lubarsky's Law of Cybernetic Entomology

Fifth secret: use SA to calibrate complexity





Presented as ‘Conjecture by O’Neill’

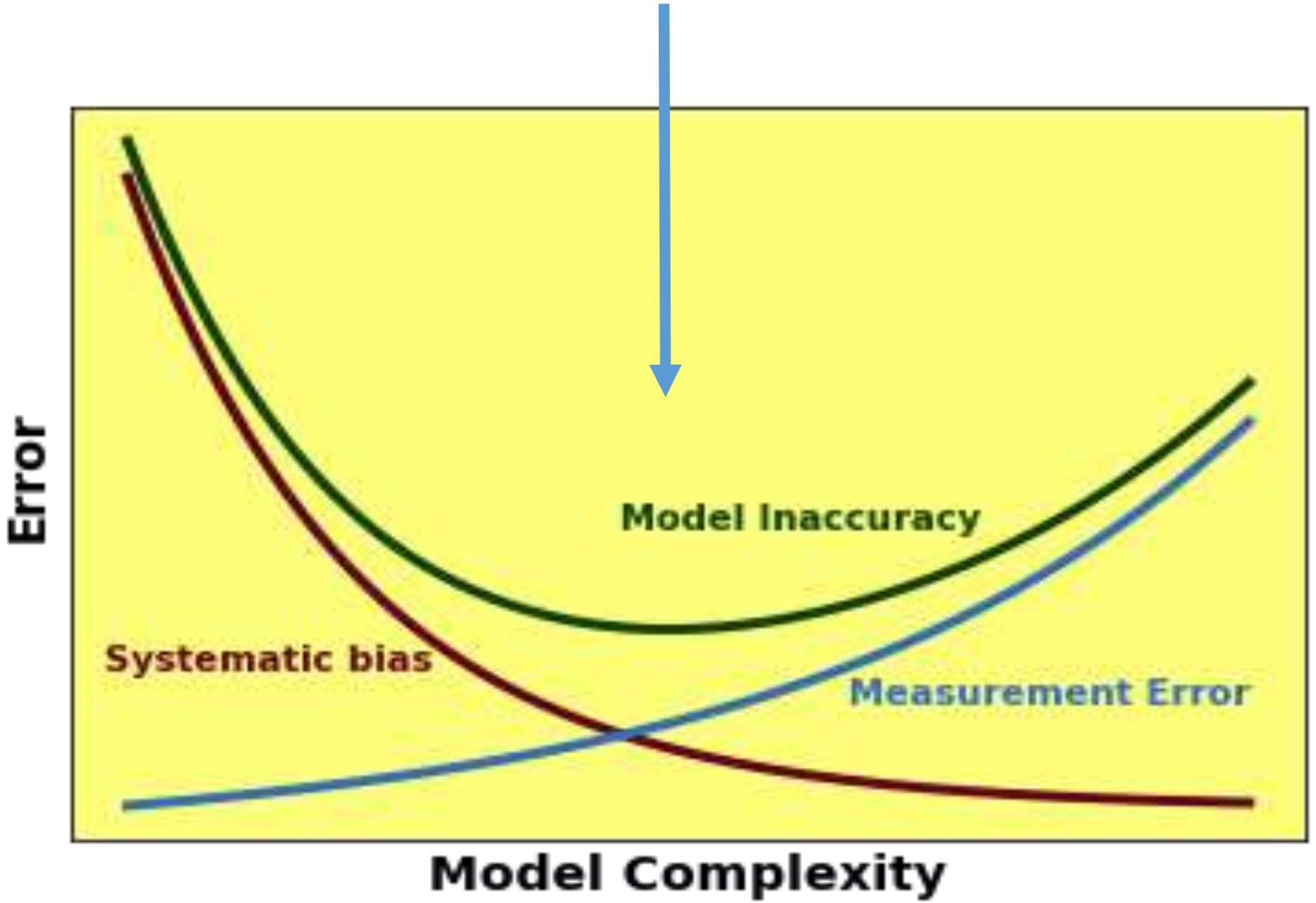
In M. G. Turner and R. H. Gardner, “Introduction to Models” in Landscape Ecology in Theory and Practice, New York, NY: Springer New York, 2015, pp. 63–95.

Lofti Aliasker Zadeh

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

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.

SA can help to find this minimum



nature COMMUNICATIONS

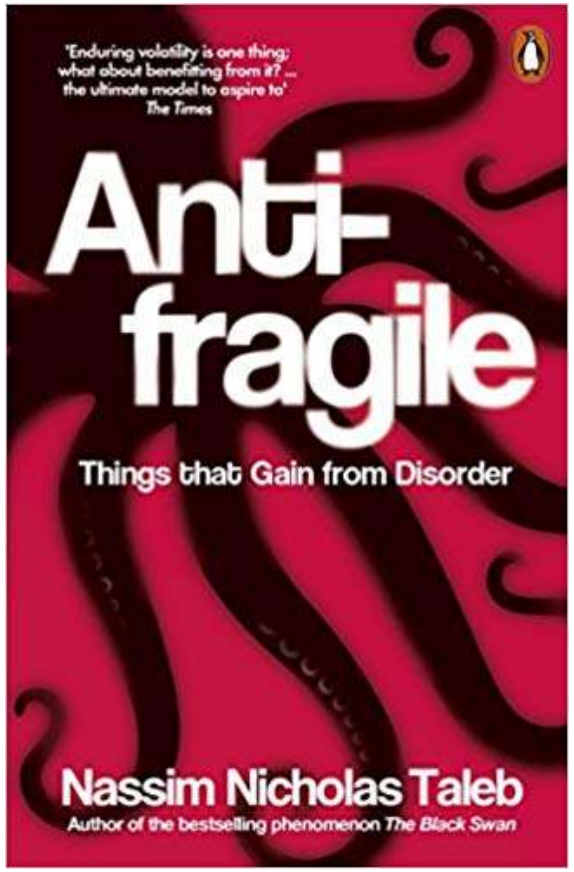
Comment | Open Access | Published: 27 August 2019

A short comment on statistical versus mathematical modelling

Andrea Saltelli

Sixth secret:

With SA it is easier to disprove than to prove; use SA ‘via negativa’:



Doing the right thing

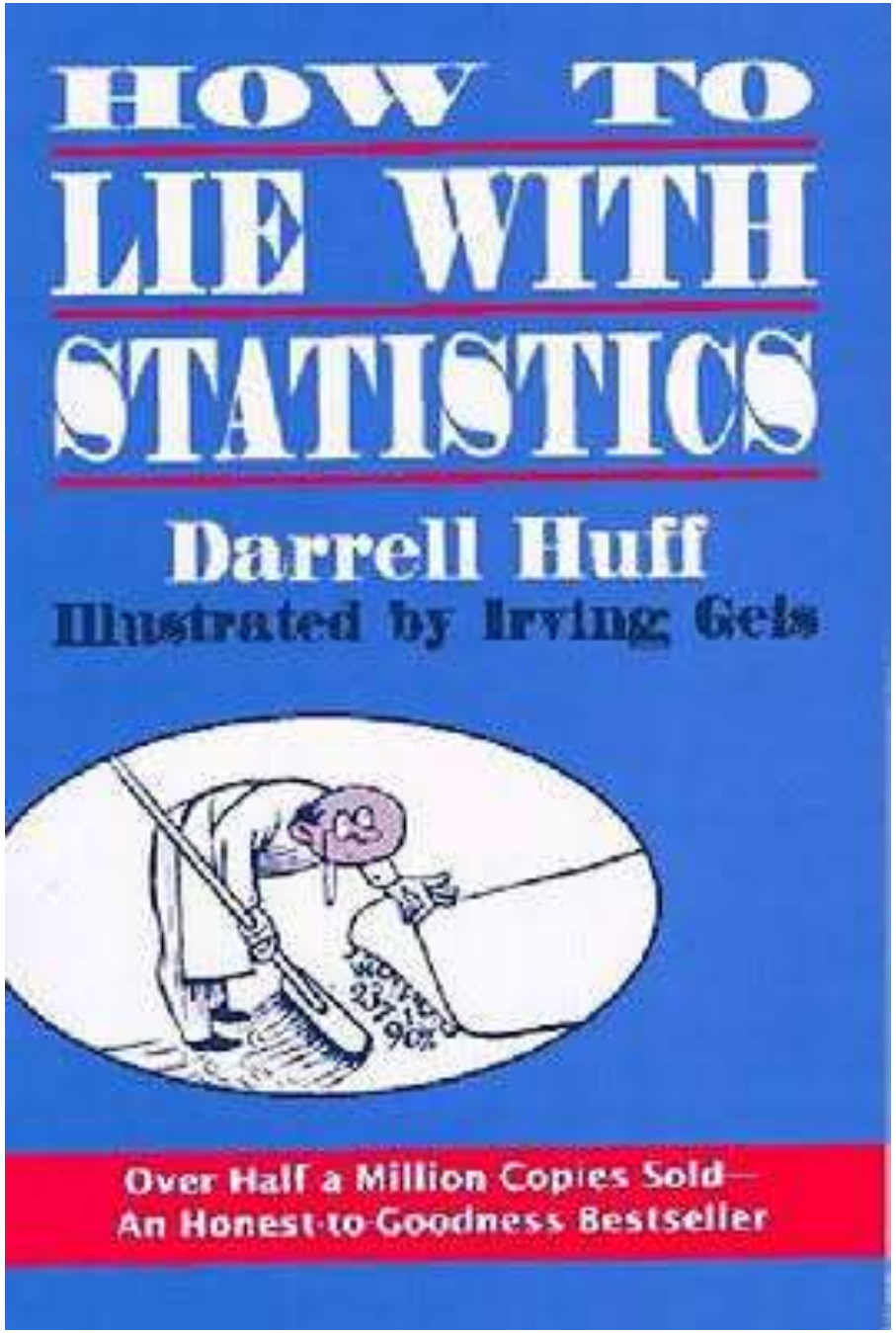
or

Avoiding something wrong?

And of course please don’t run a sensitivity analysis where each factors has a 5% uncertainty

Why?

Can we say that one lies with sensitivity analysis as one can lie with statistics?



Limit of SA: Often no SA (SA conflated with UA e.g. in economics) or one-factor-at-a-time SA

Why is OAT (one-factor-at-a-time) SA so bad?

Environmental Modelling & Software 25 (2010) 1508–1517

Contents lists available at ScienceDirect

Environmental Modelling & Software

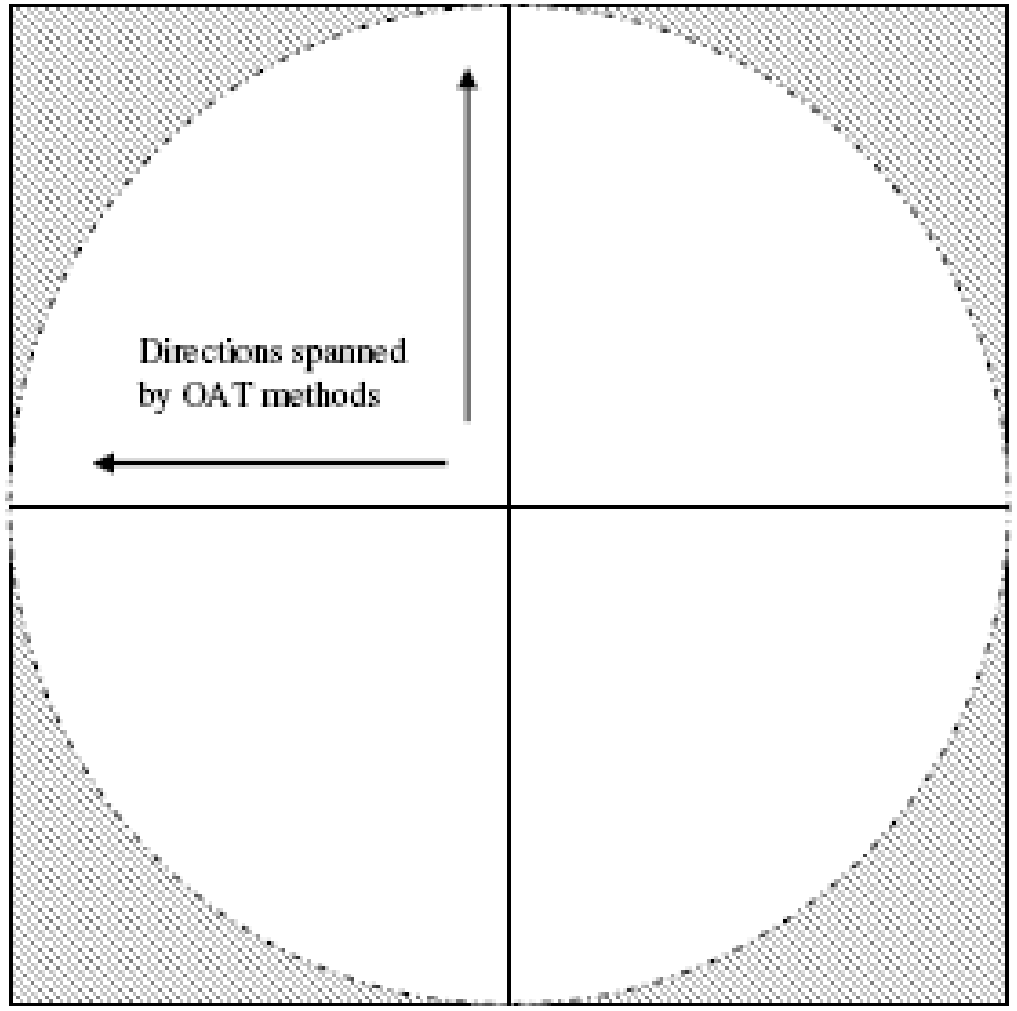
journal homepage: www.elsevier.com/locate/envsoft

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

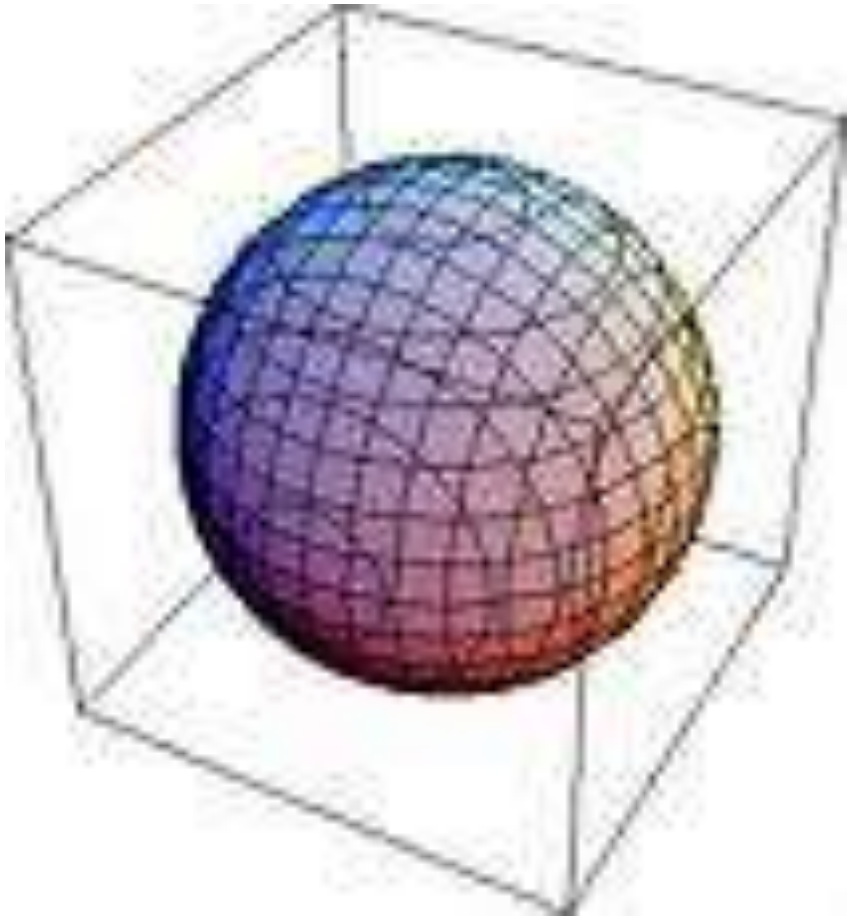
OAT in 2 dimensions



Area circle / area square =?

~ 3/4

OAT in 3 dimensions

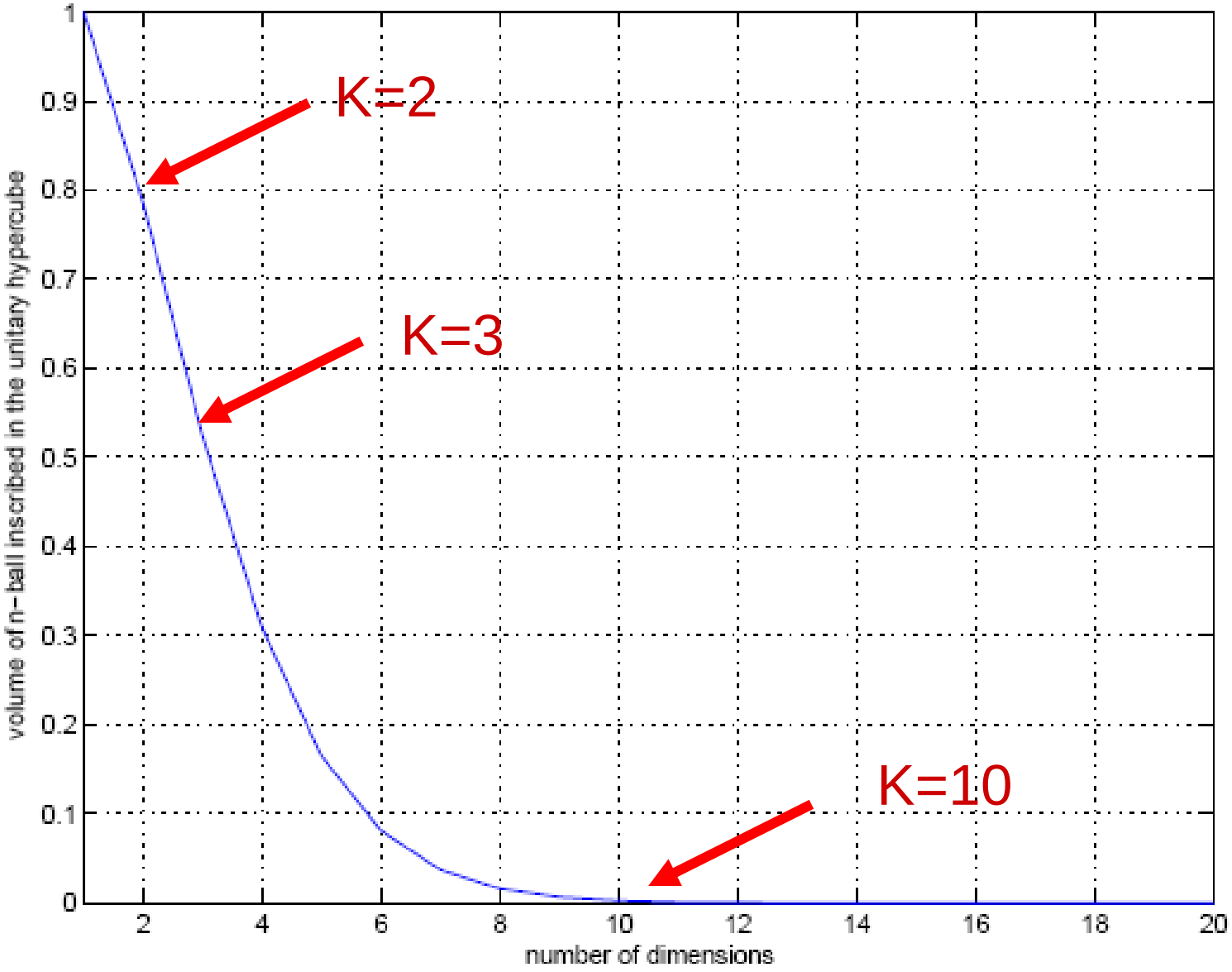


Volume sphere / volume cube =?

~ 1/2

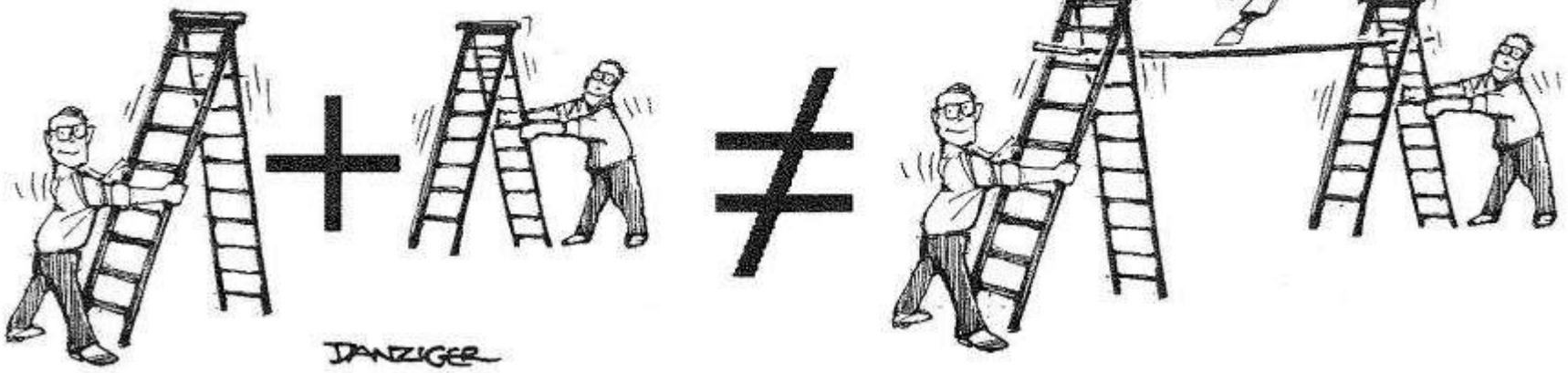
OAT in 10 dimensions; Volume hypersphere / volume ten dimensional hypercube =? ~ 0.0025

OAT in k dimensions

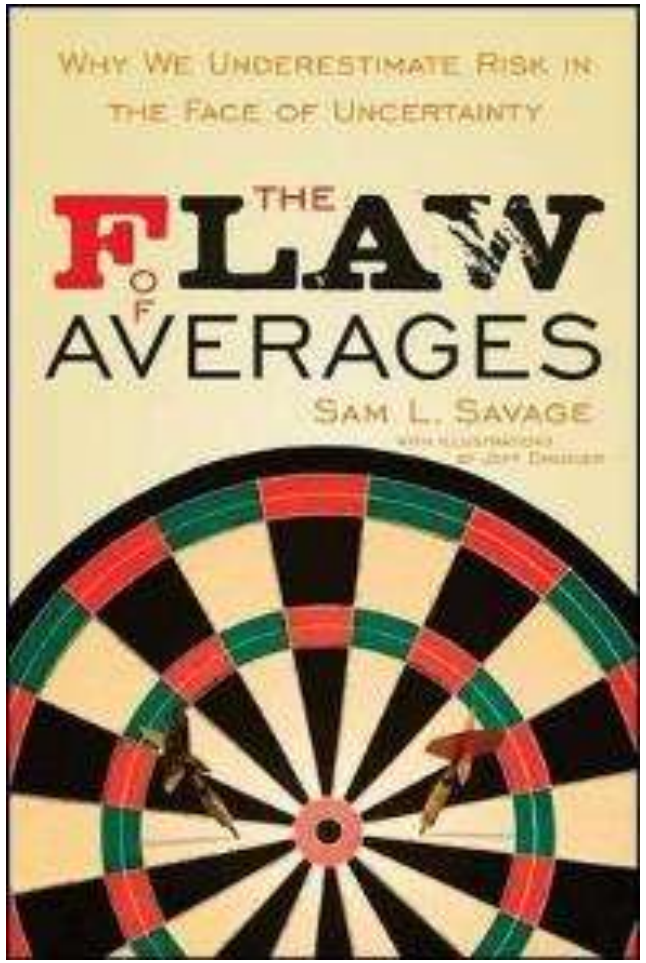


How would you test the scaffolding?

How coupled ladders are shaken in most of available literature



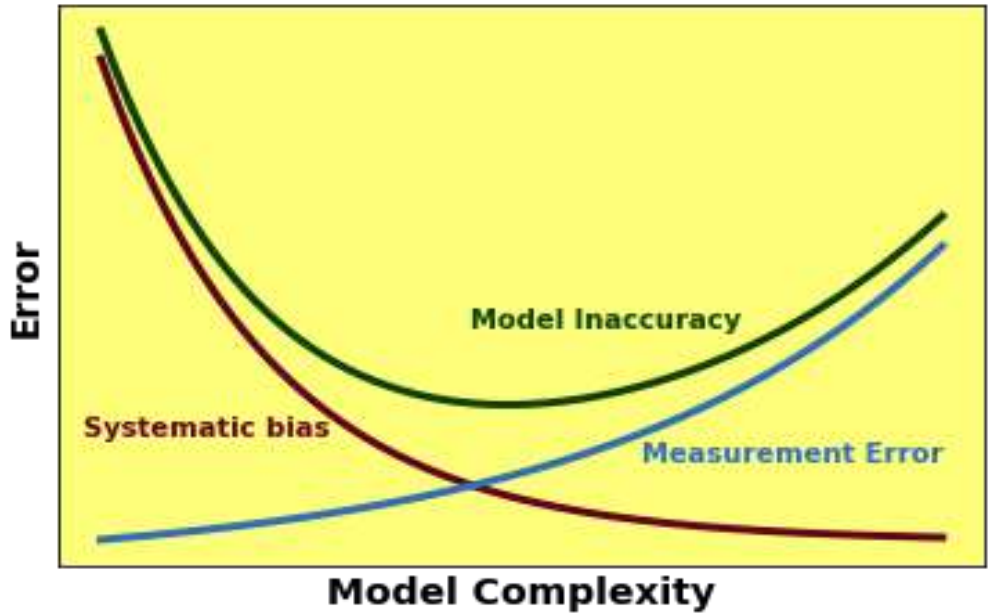
How to shake coupled ladders



Lessons from sensitivity analysis

- Global SA
- UA and SA coupled
- Purpose- & context-specific
- The map is not the territory

- Memento



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]

Literature search in Scopus

Query: “sensitivity analysis” & “model/modelling” & “uncertainty”; years 2012–2017; journal articles; in English

➔ 6000 articles

- AgrBioSci (Agricultural and Biological Sciences)
- BiochemGenMBio (Biochemistry, Genetics and Molecular Biology)
- BusManAcc (Business, Management and Accounting)
- Chemi (Chemistry)
- ChemEng (Chemical Engineering)
- CompSci (Computer Science)
- DecSci (Decisional Science)
- EarthSci (Earth and Planetary Sciences)
- EconFin (Economy and Finance)
- Energy (Energy)
- Engineering (Engineering)
- EnvSci (Environmental Science)
- ImmunMicrobio (Immunology and Microbiology)
- MatSci (Material Science)
- Math (Math)
- Medicine (Medicine)
- PharTox (Pharmacology and Toxicology)
- PhysAstro (Physics and Astronomy)
- SocSci (Social Science)

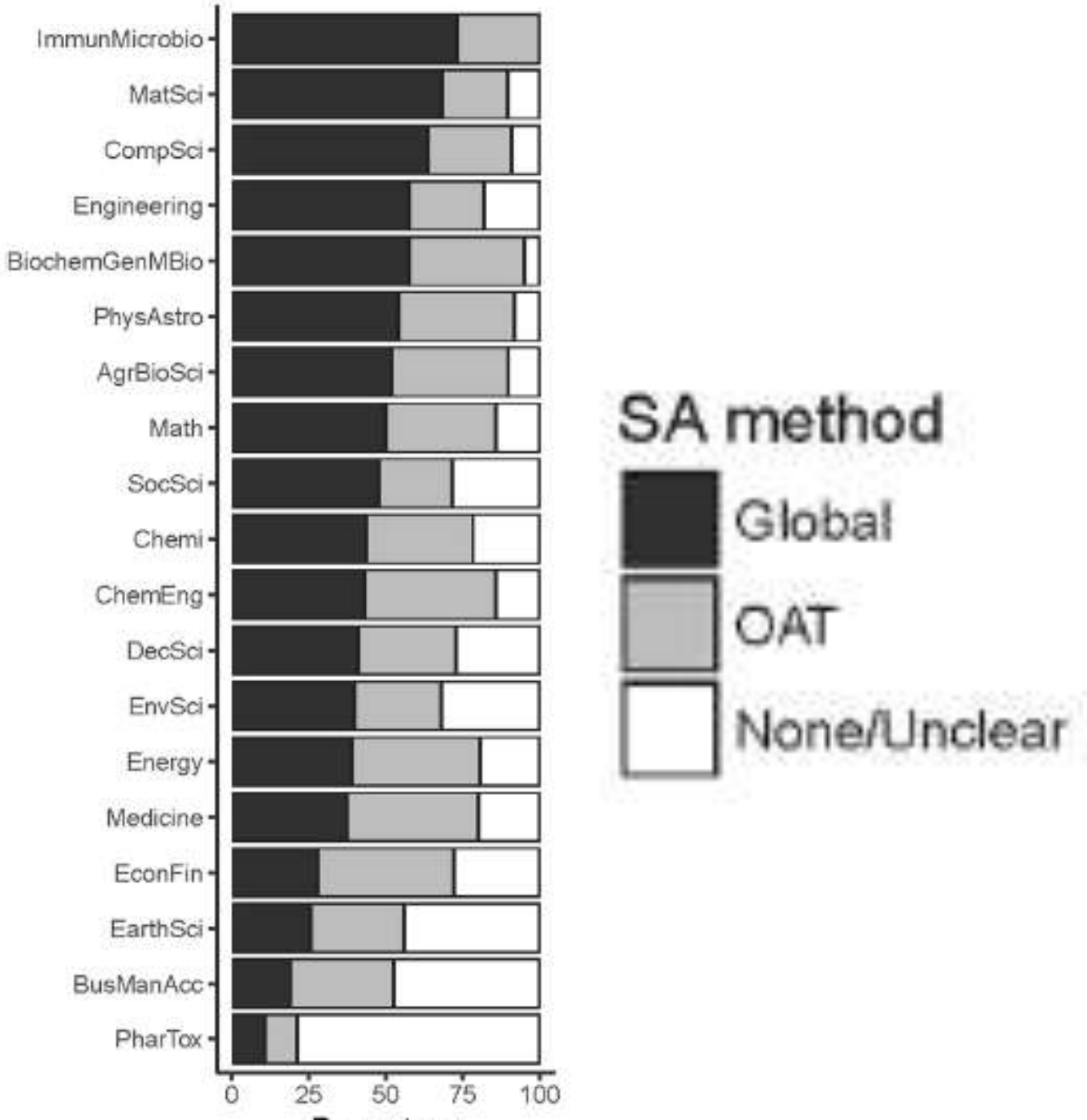
subject areas >100 articles

Taking the top twenty most-cited papers in each subject area:

➔ 324 articles, divided among authors

Cleansing manually irrelevant articles:

➔ 280 articles



Still many papers apply an OAT SA: 65%

What if the model is truly linear?

Linear	7%
Nonlinear	61%
Unclear	32%

Linear	7%
Nonlinear	61%
Unclear	32%

65% highly cited articles are OAT

Taking all unclear = linear ➔ still over 20% (.32+.07)*.65 of papers wrong

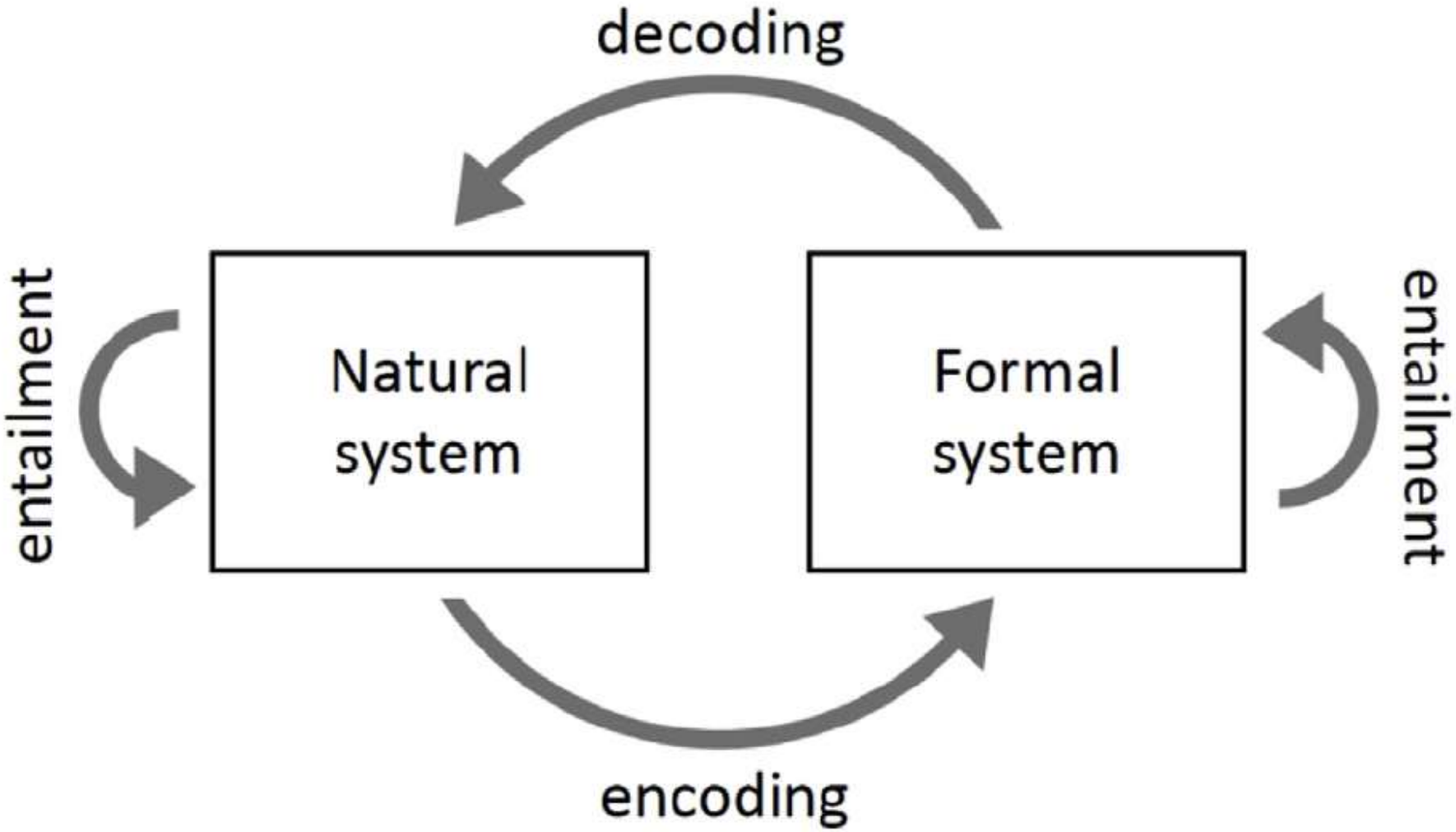
(OAT & non-linear model)

Why?

5. Discussion

5.1. Reasons for bad practice

Why? ➔ 1. Modelling as a craft



Why? ➔ 2. Each discipline going about modelling on its own separate way; pockets of SA practitioners (out of our 280 papers, 35 were methodological, of which 24 suggest global SA)

Why? ➔ 3. Mathematical modelling is not a discipline

Based on a survey of modellers: "there is no dominating paradigm in modelling and simulation… simulation verification is mostly a trial and error activity ➔ challenges model/simulation validity"

Padilla, J. J., Diallo, S. Y., Lynch, C. J., & Gore, R. (2018). Observations on the practice and profession of modeling and simulation: A survey approach. SIMULATION, 94(6), 493–506.

… mathematical modelling cannot do this:

Wasserstein, R.L. and Lazar, N.A., 2016. 'The ASA's statement on p-values: context, process, and purpose', The American Statistician, Volume 70, 2016 - Issue 2, Pages 129-133.

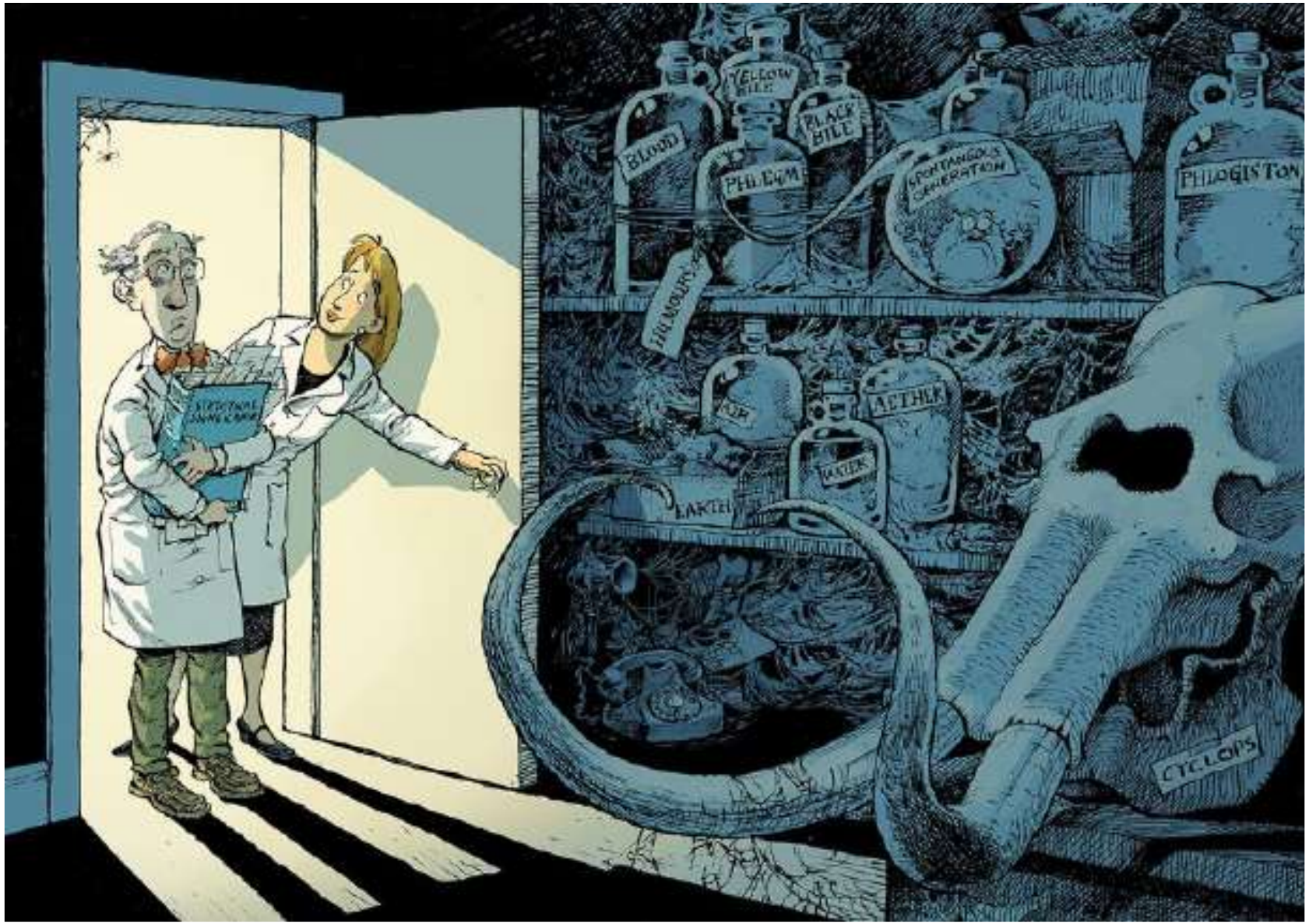
Why? ➔ 4. Good practices require training in statistics

Why? ➔ 5. More time is needed; though mature global sensitivity analysis methods around for more than 25 years researchers tend to emulate methods found in highly cited papers assuming that they are best practice

Why? ➔ 6. Strategical reasons: global SA is bad if one wants to play the uncertainty game, inflating or deflating uncertainties instrumentally

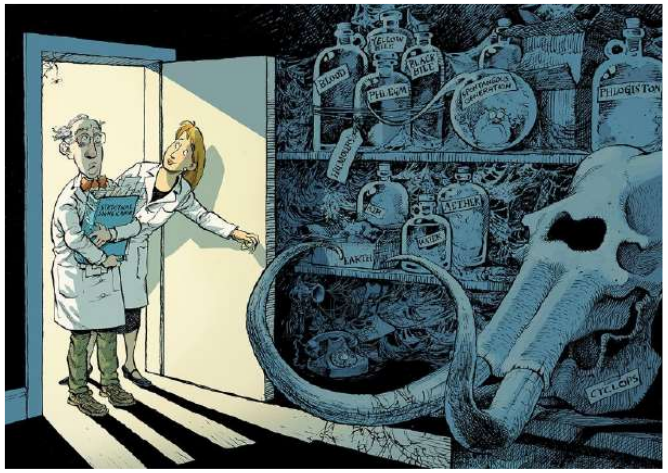
Solutions? Statistics as a discipline takes responsibility for statistical methods for model validation and verification

Learn from what happens in statistics where the p-test crisis is being tackled head on



Throw away the concept of statistical significance?

nature
International journal of science



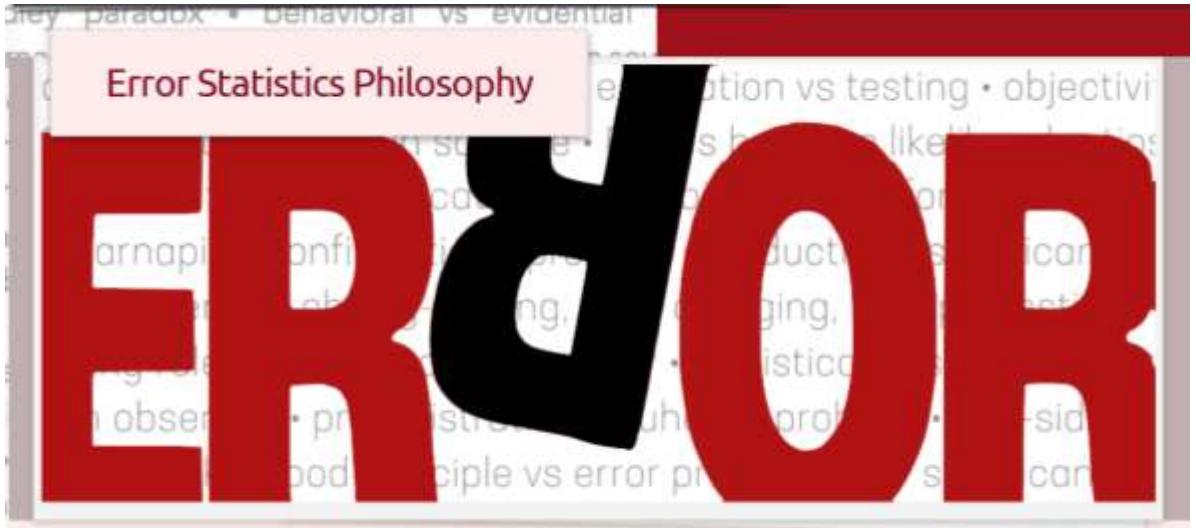
COMMENT · 20 MARCH 2019

Scientists rise up against statistical significance

Valentin Amrhein, Sander Greenland, Blake McShane and more than 800 signatories call for an end to hyped claims and the dismissal of possibly crucial effects.

Valentin Amrhein ✉, Sander Greenland & Blake McShane

See the discussion on the blog of Andrew Gelman https://statmodeling.stat.columbia.edu/



Statistics in the wake of the reproducibility crisis

Statistical wars?

nature communications

Comment | Open Access | Published: 27 August 2019

A short comment on statistical versus mathematical modelling

Andrea Saltelli

Nature Communications **10**, Article number: 3870 (2019) | Cite this article

While the crisis of statistics has made it to the headlines, that of mathematical modelling hasn't. Something can be learned comparing the two, and looking at other instances of production of numbers.Sociology of quantification and post-normal science can help.

The End

@andreasaltelli