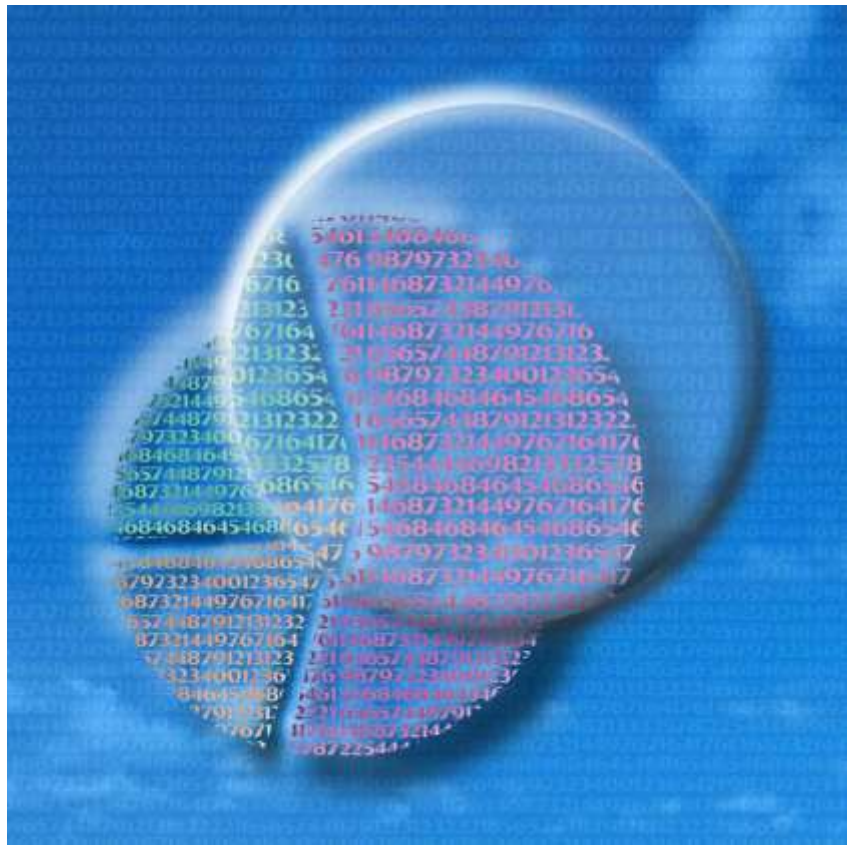


Sensitivity analysis for mathematical modelling

Andrea Saltelli
Centre for the Study of the Sciences and the Humanities, University of Bergen, and Open Evidence Research, Open University of Catalonia

MNF990 19H / Theory of Science and Ethics, Bergen, Thormøhlens gate 51 (VilVite), Konferanserom CD, September 16, 2019

UNIVERSITAS BERGENSIS

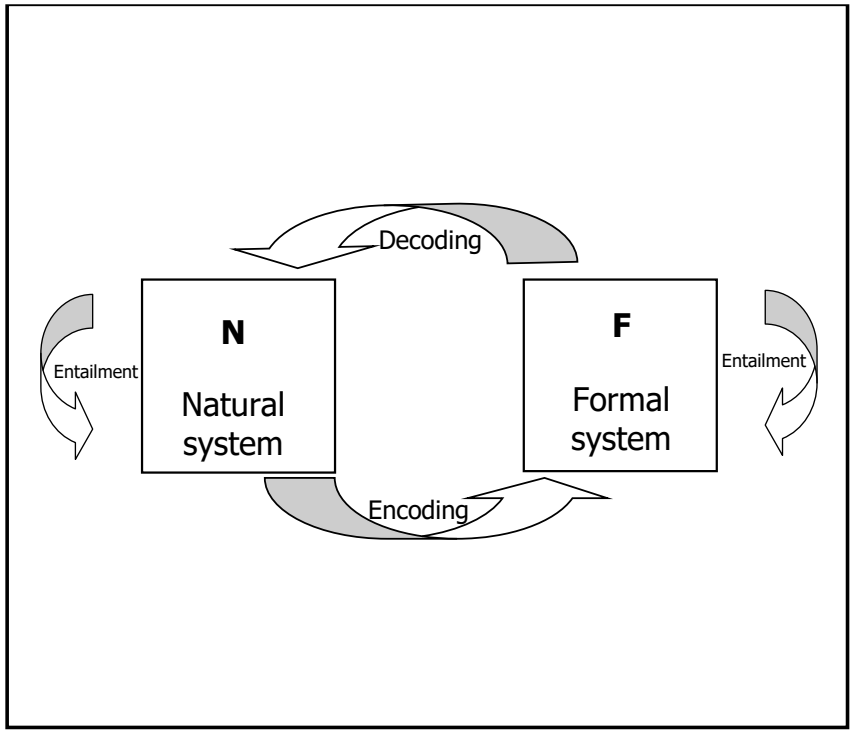
OPEN EVIDENCE

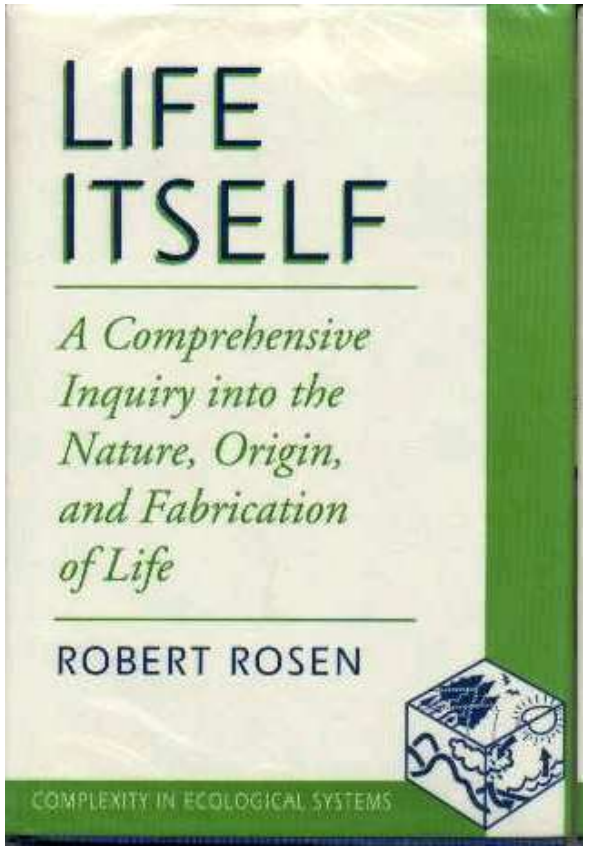
UOC Universitat Oberta de Catalunya

Where to find this talk: www.andreasaltelli.eu

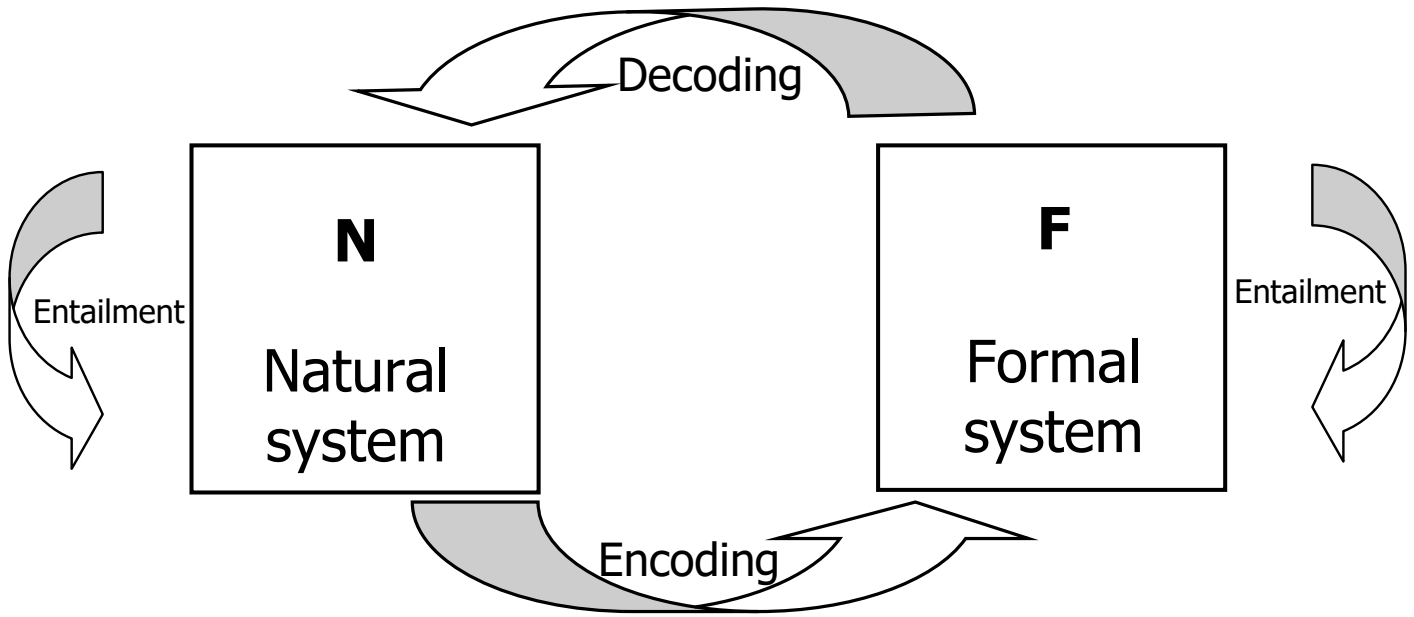
What is modelling?

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.



What is a model ?

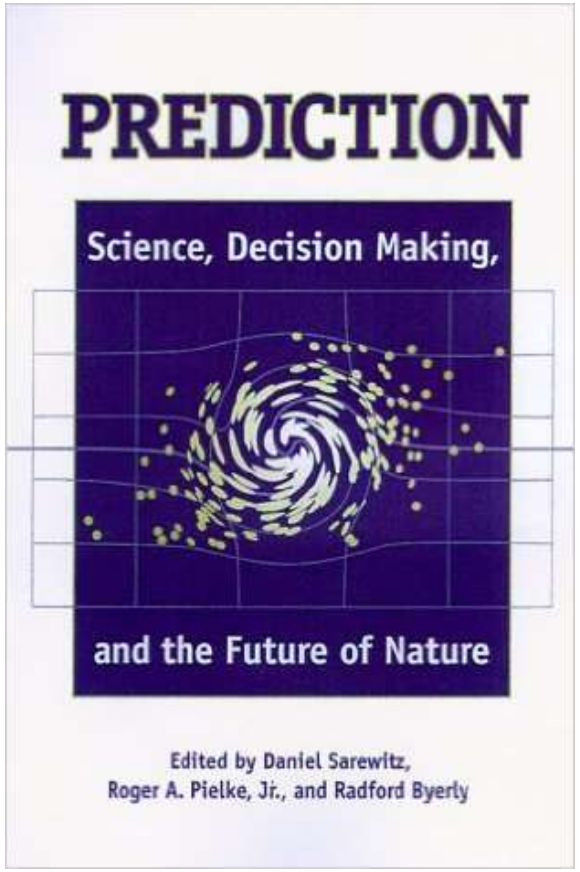
Robert Rosen

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

Naomi Oreskes

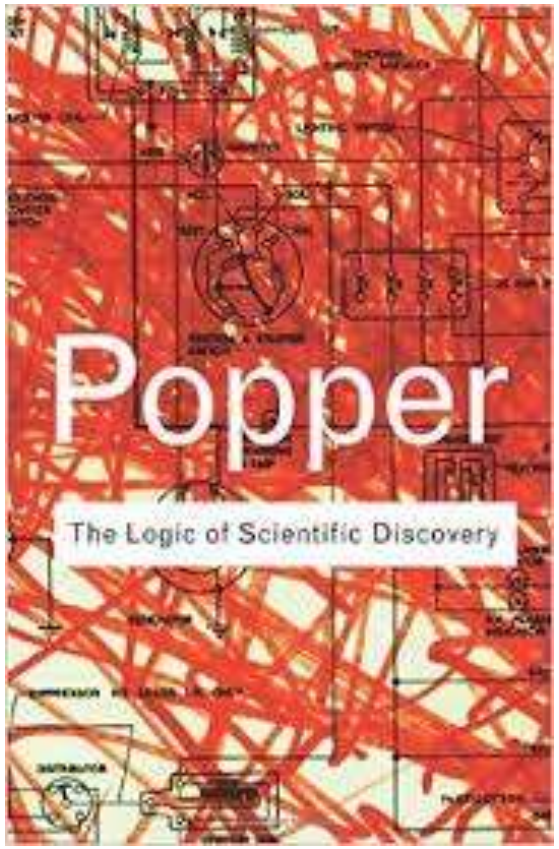
N. Oreskes, K. Shrader-Frechette, and K. Belitz, “Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences,” Science, 263, no. 5147, 1994.

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)

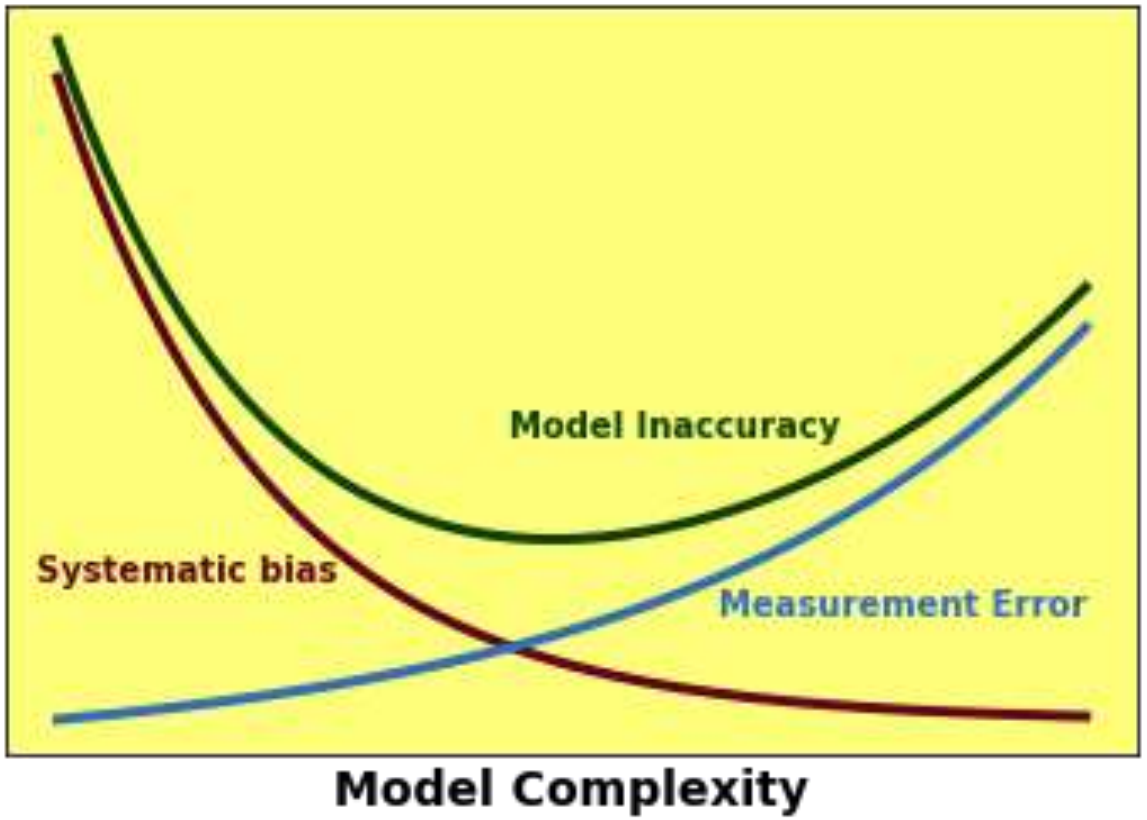


“In many cases, these temporal predictions **are treated with the same respect** that the hypothetic-deductive model of science accords to logical predictions. But this respect is largely misplaced”

“[…] models are complex amalgam of theoretical and phenomenological laws (and the governing equations and algorithms that represent them), empirical input parameters, and a model conceptualization […] 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 are hence vulnerable to misuse, more than e.g. a physical law or a standardized statistical test ➔

- Mathematical hubris
- Lack of standards for quality control
- Forgetting the conditionality of model knowing



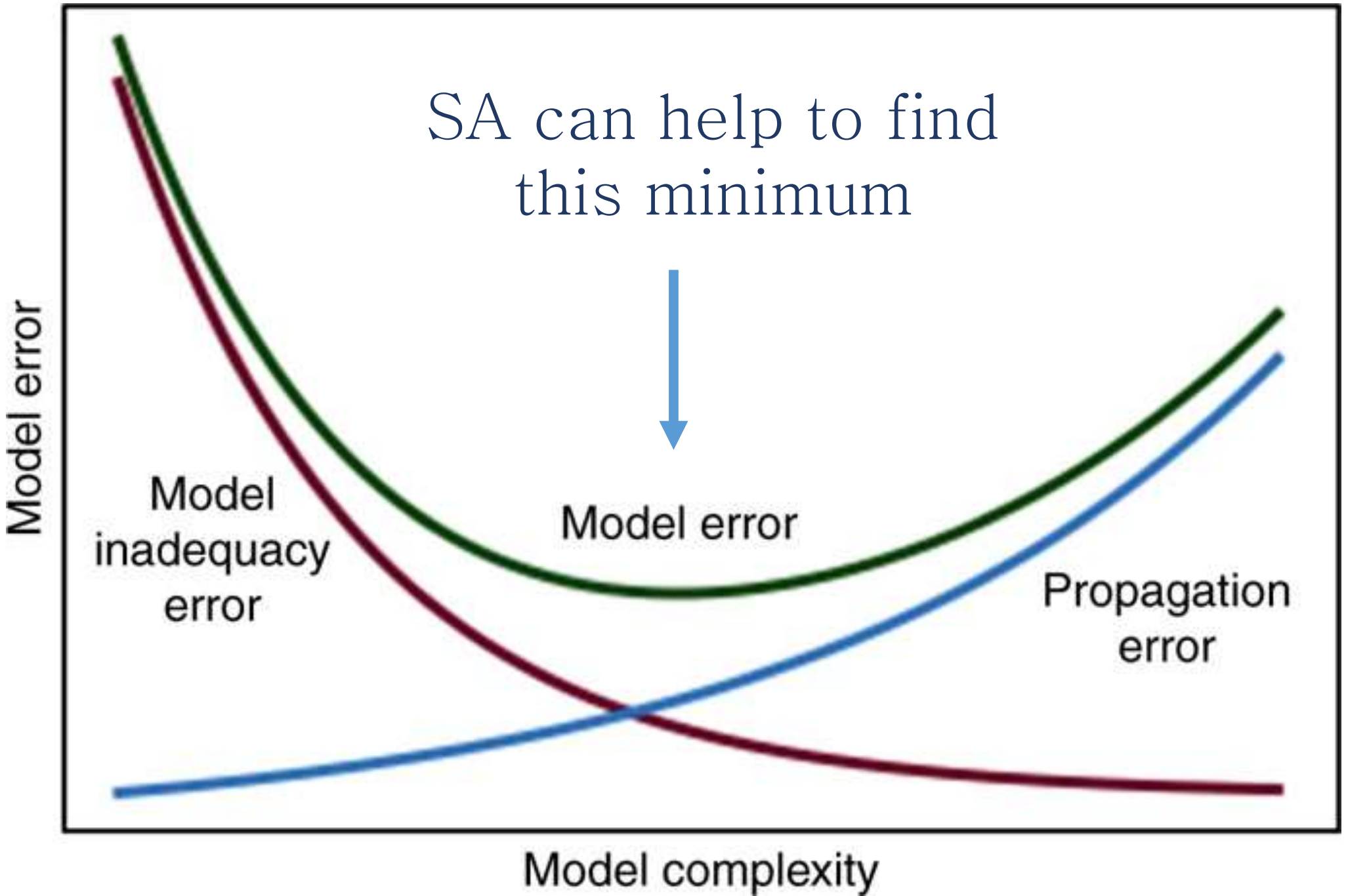
Presented as ‘Conjecture by O’Neill’

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.

Why sensitivity analysis

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
Model error
Model
inadequacy
error
Model error
Propagation
error
Model complexity

Model-based knowing is conditional

For John Kay modelling may need as input information which we don't have

John Kay

J. A. Kay, "Knowing when we don't know," 2012, https://www.ifs.org.uk/docs/john_kay_feb2012.pdf

WebTAG: Annual Percentage Change in Car Occupancy (% pa) up to 2036

Journey Purpose	Weekday					Weekend	All Week
	7am-10am	10am-4pm	4pm-7pm	7pm-7am	Weekday Average		
Work	-0.48	-0.4	-0.62	-0.5	-0.44	-0.48	-0.45
Non - Work (commuting and other)	-0.67	-0.65	-0.53	-0.47	-0.59	-0.52	-0.56

Economics

Paul Romer's Mathiness = use of mathematics to veil normative stances

P. M. Romer, "Mathiness in the Theory of Economic Growth," Am. Econ. Rev., vol. 105, no. 5, pp. 89–93, May 2015.

Paul Romer

Erik Reinert: scholastic tendencies in the mathematization of economics

Erik Reinert

E. S. Reinert, “Full circle: economics from scholasticism through innovation and back into mathematical scholasticism,” J. Econ. Stud., vol. 27, no. 4/5, pp. 364–376, Aug. 2000.

The Equality Assumption: Uneven development and the history of economic thought

Seminar på engelsk med professor Erik Reinert, professor of Technology Governance and Development Strategies, Tallinn University of Technology, om alternativ økonomi.

Foto/ill.: Kroglund

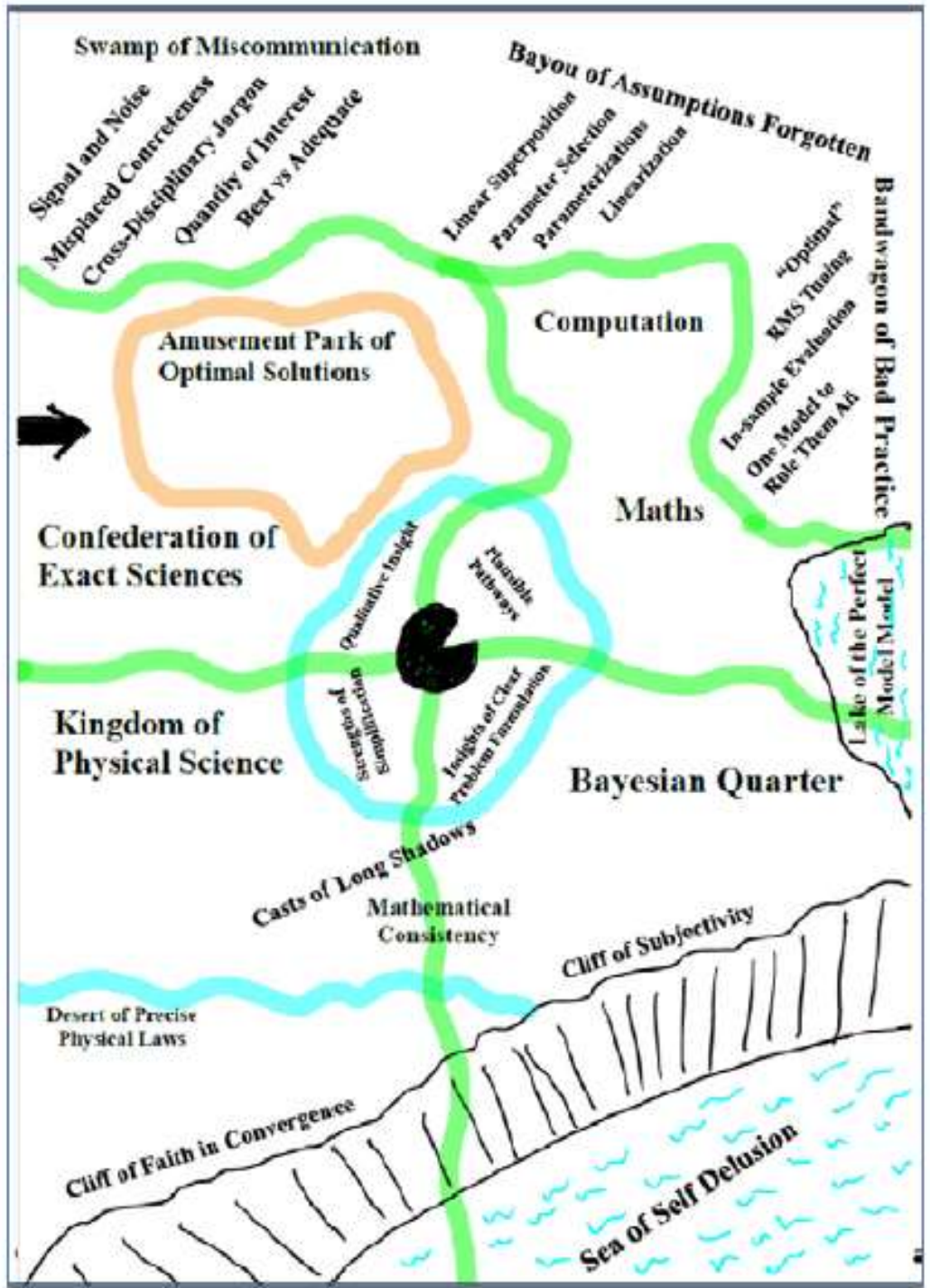
SEMINAR

22.10.2019 - 15.00–20.00

Sjøfartsmuseet, Haakon
Sheteligsplass 15, Bergen

Escape from model-land!

E. L. . Smith and T. L. A., “Escape from model-land,” Econ. Open access, Open Assess. J., vol. 23, no. March 08, 2019.



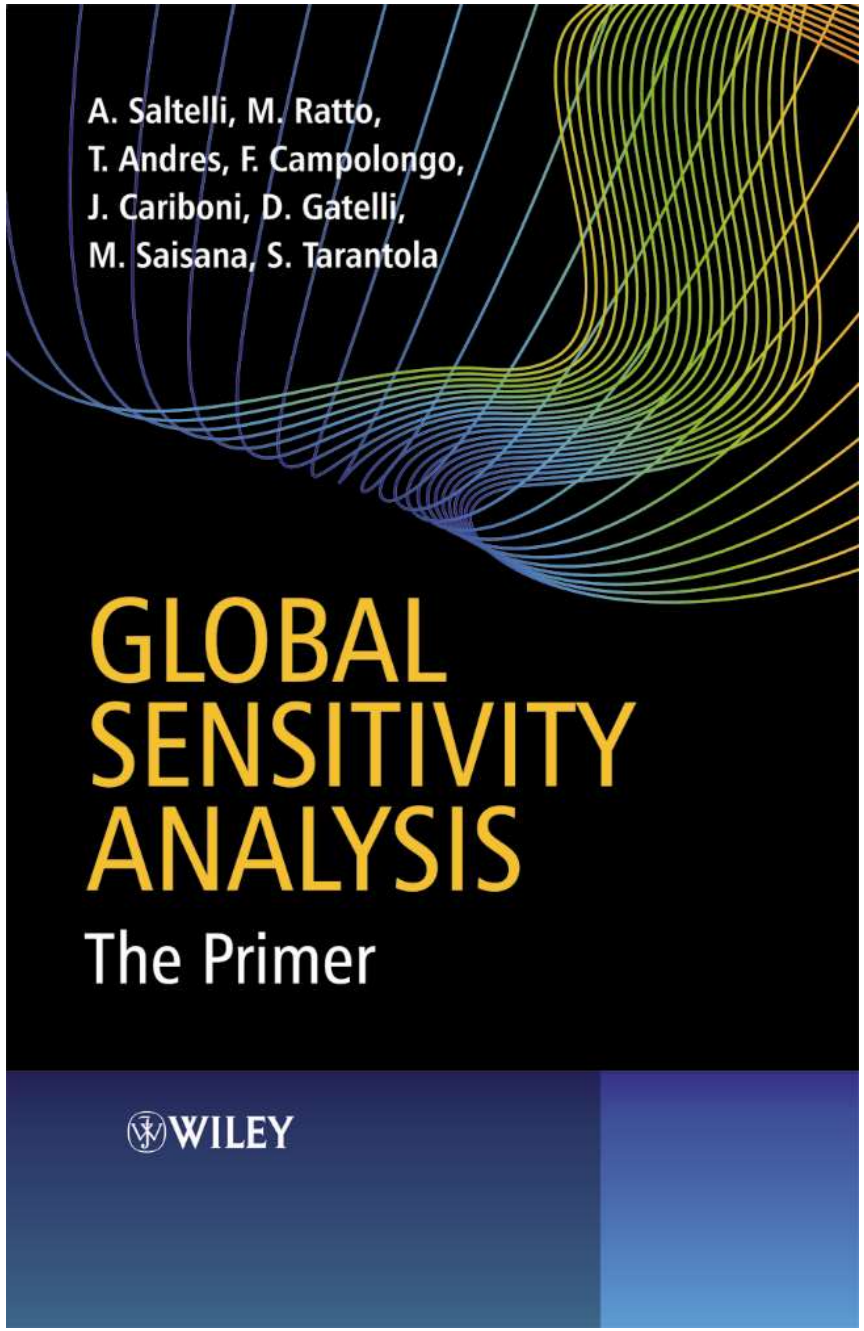
E. L. . Smith and T. L. A., “Escape from model-land,” Econ. Open access, Open Assess. J., vol. 23, no. March 08, 2019.

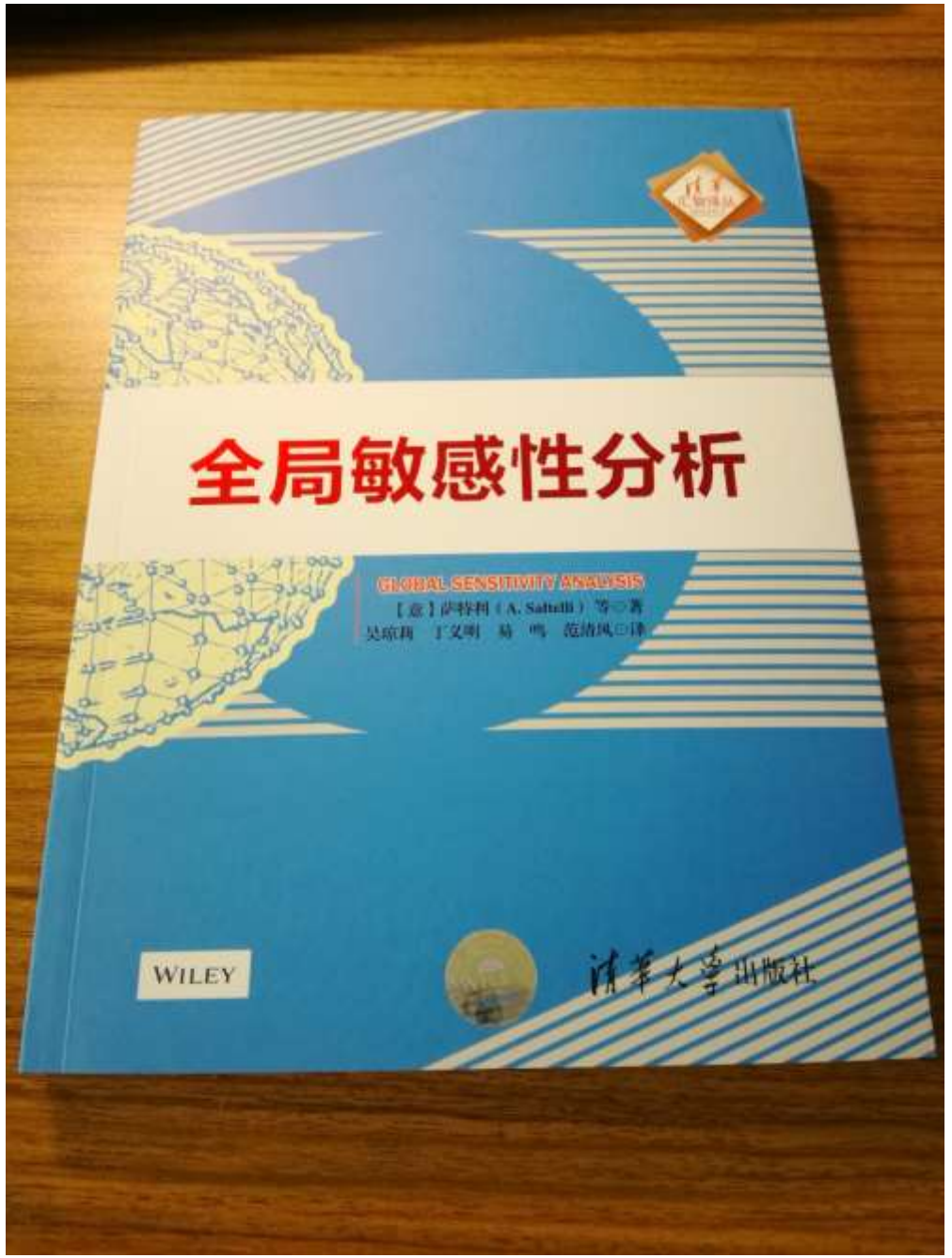
On sensitivity analysis and its take up

Definitions

Uncertainty analysis: Focuses on quantifying the uncertainty in model output

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


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

European Commission, 2015

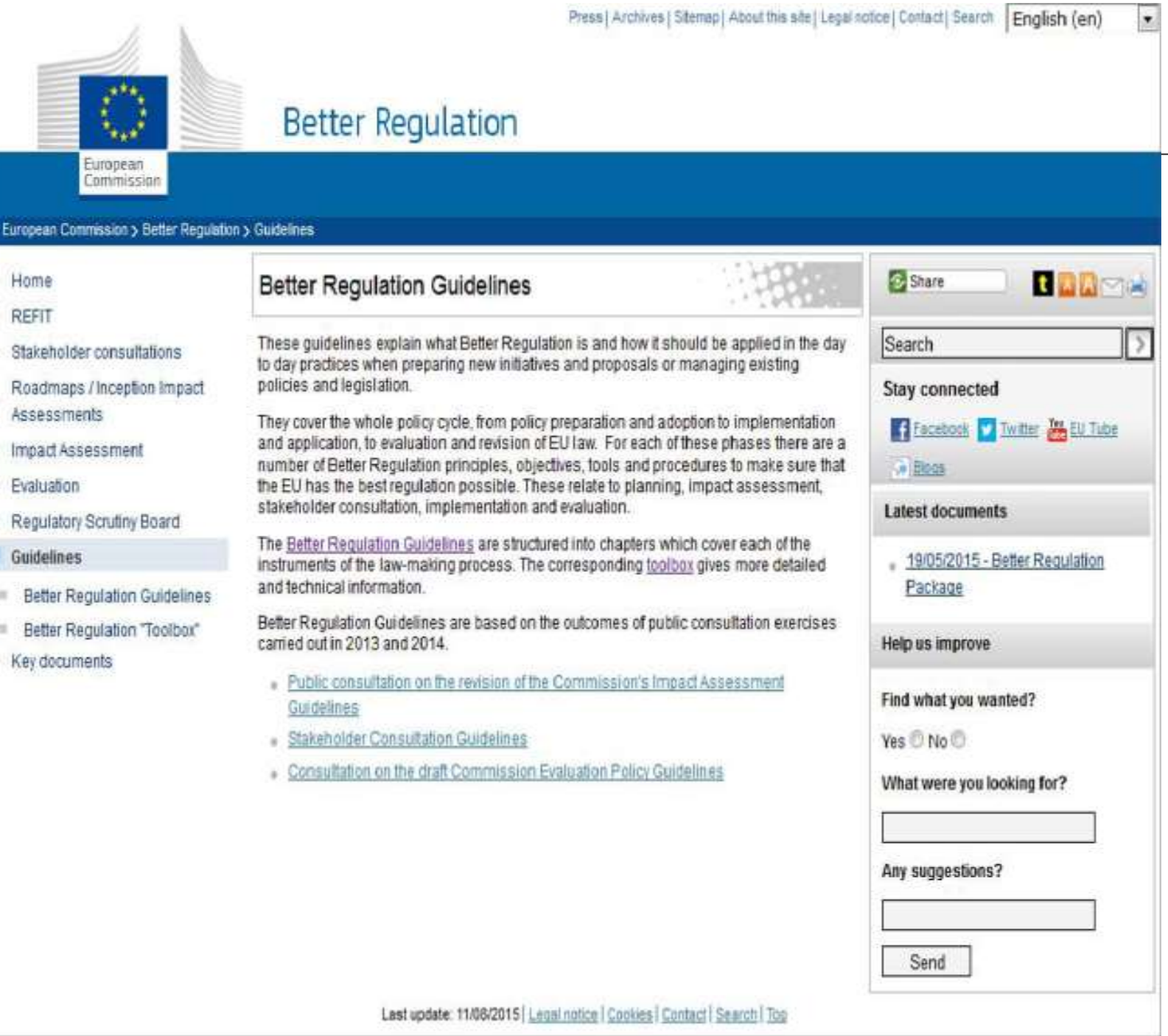
Office for the Management and Budget, 2006

Environmental Protection Agency, 2009

EPA, 2009, March. Guidance on the Development, Evaluation, and Application of Environmental Models. Technical Report EPA/100/K-09/003. Office of the Science Advisor, Council for Regulatory Environmental Modeling, http://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P1003E4R.PDF, Last accessed December 2015.

EUROPEAN COMMISSION, Better regulation toolbox, appendix to the Better Regulation Guidelines, Strasbourg, 19.5.2015, SWD(2015) 111 final, COM(2015) 215 final, http://ec.europa.eu/smart-regulation/guidelines/docs/swd_br_guidelines_en.pdf.

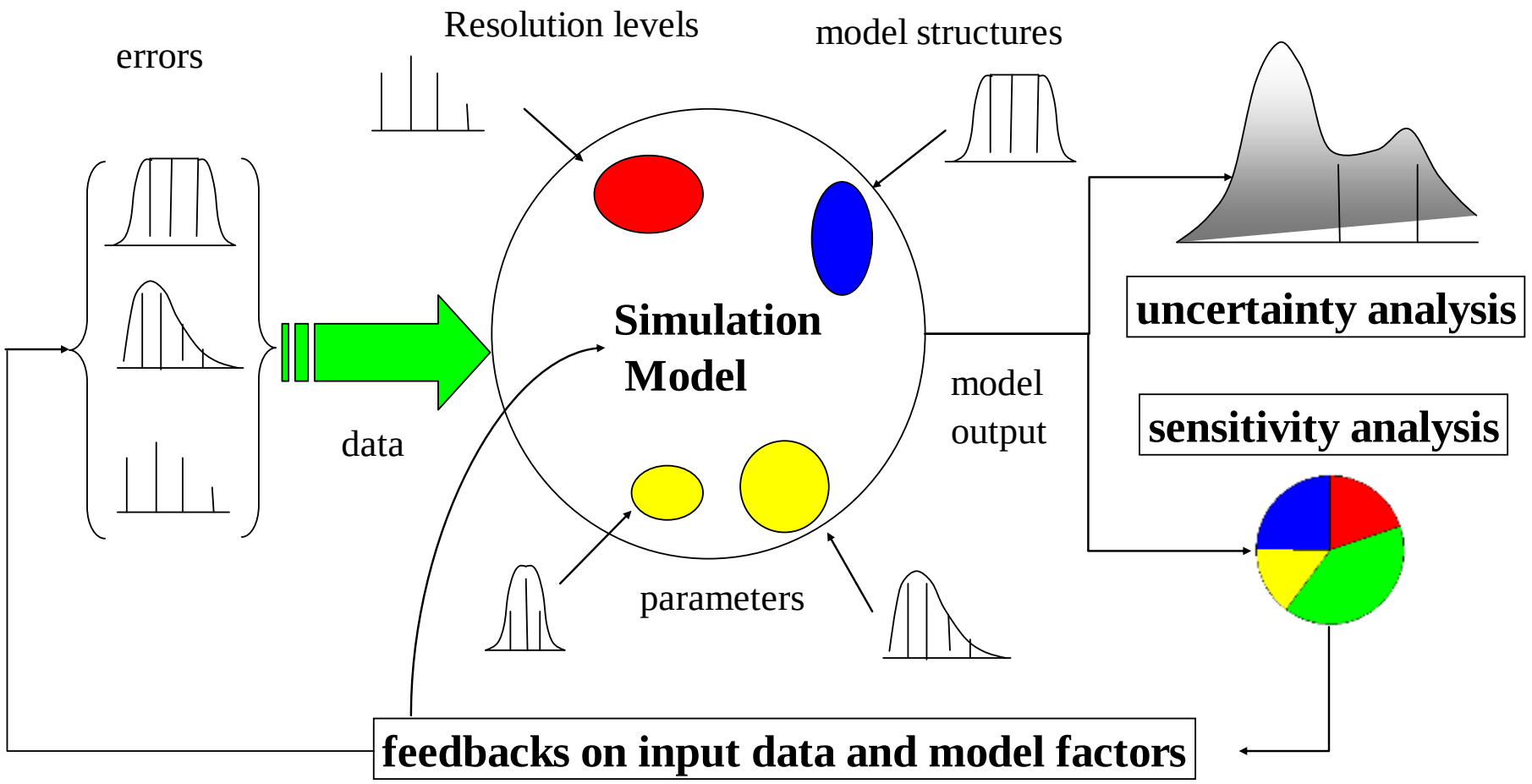
OMB, Proposed risk assessment bulletin, Technical report, The Office of Management and Budget's – Office of Information and Regulatory Affairs (OIRA), January 2006, https://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/proposed_risk_assessment_bulletin_010906.pdf, pp. 16–17, accessed December 2015.



http://ec.europa.eu/smart-regulation/

Source: IA Toolbox, p. 391

An engineer's vision of UA, SA

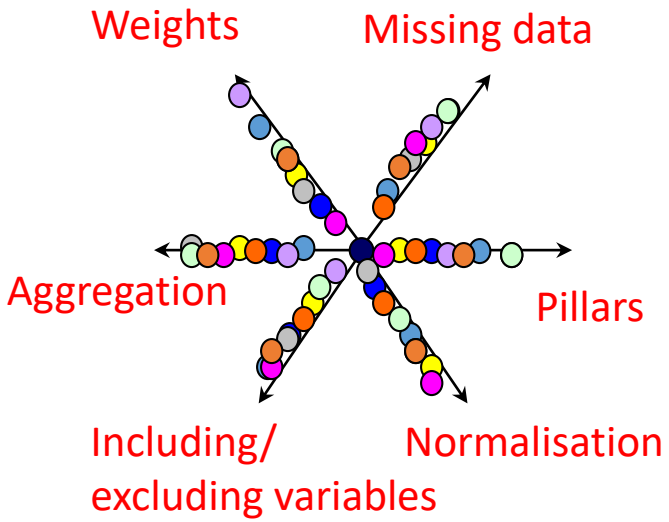


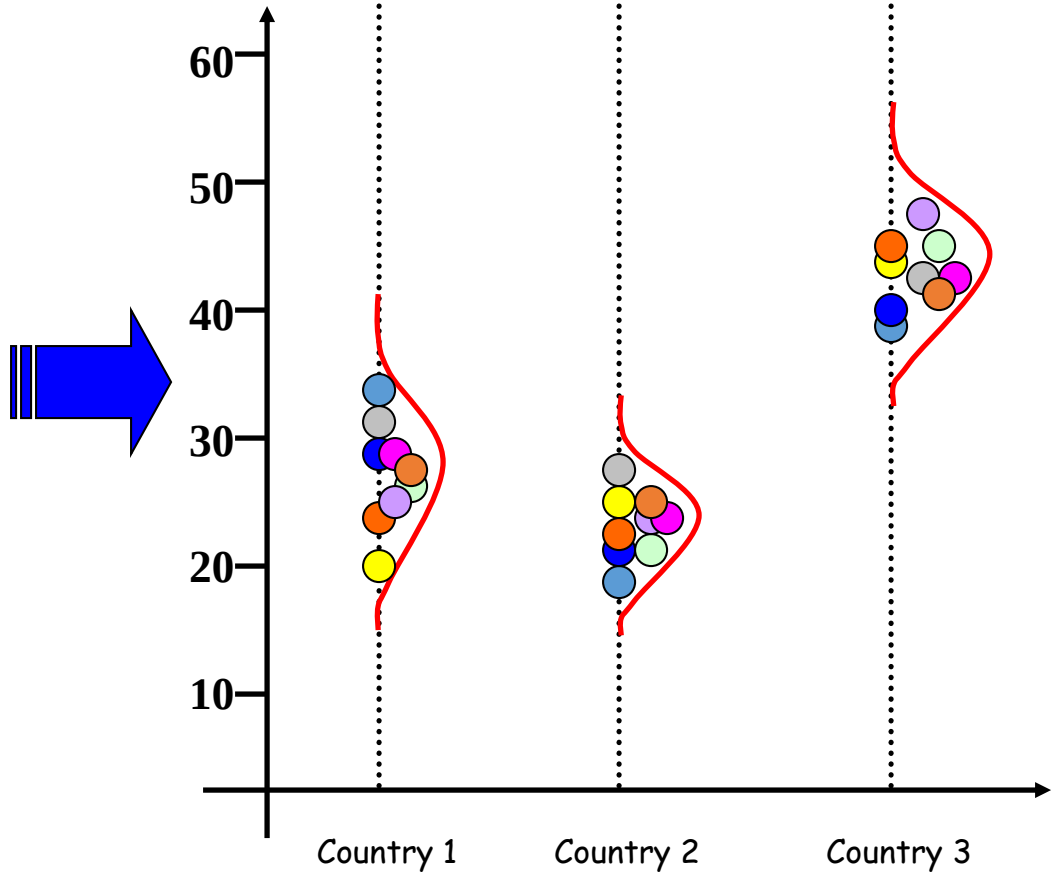
One can sample more than just factors:

- 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

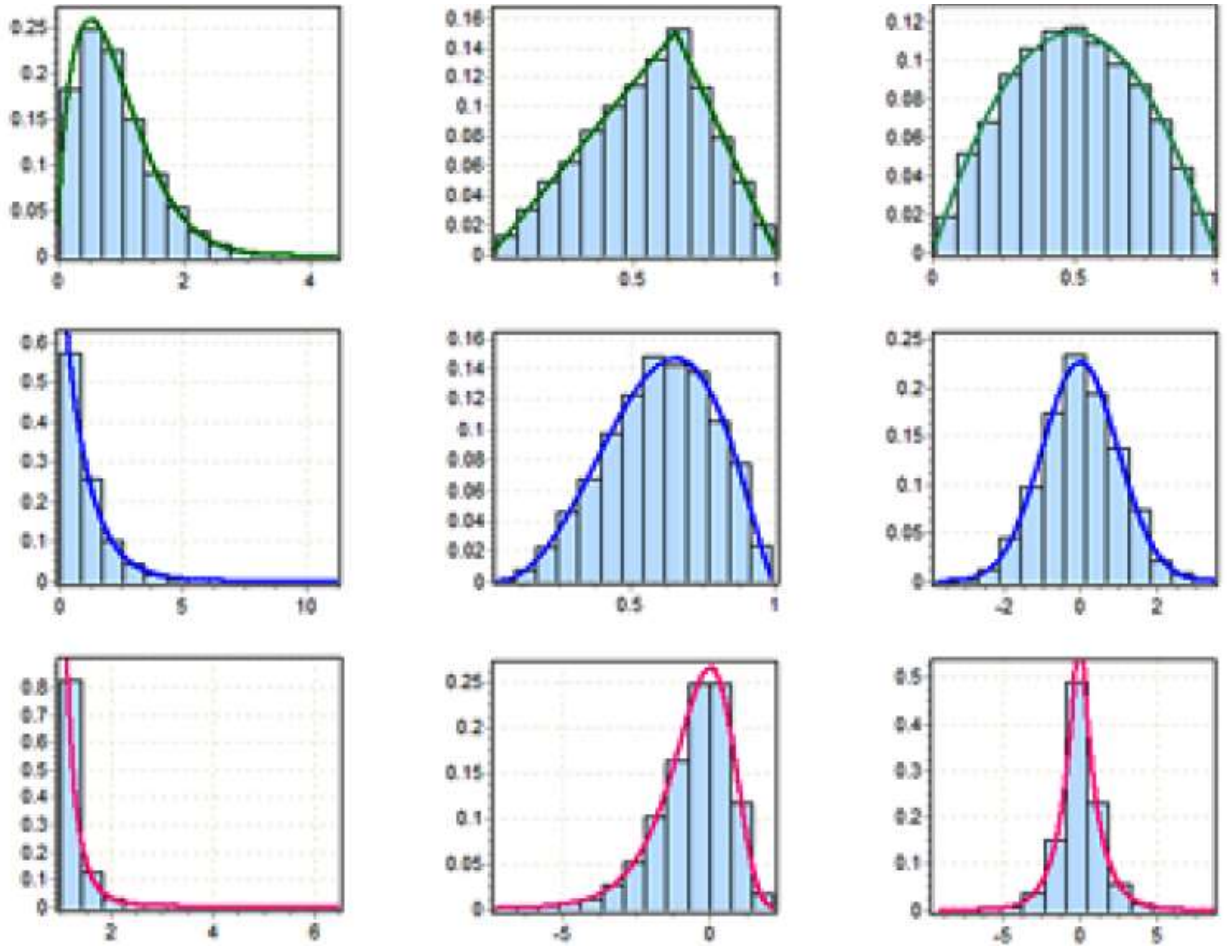
Space of alternatives





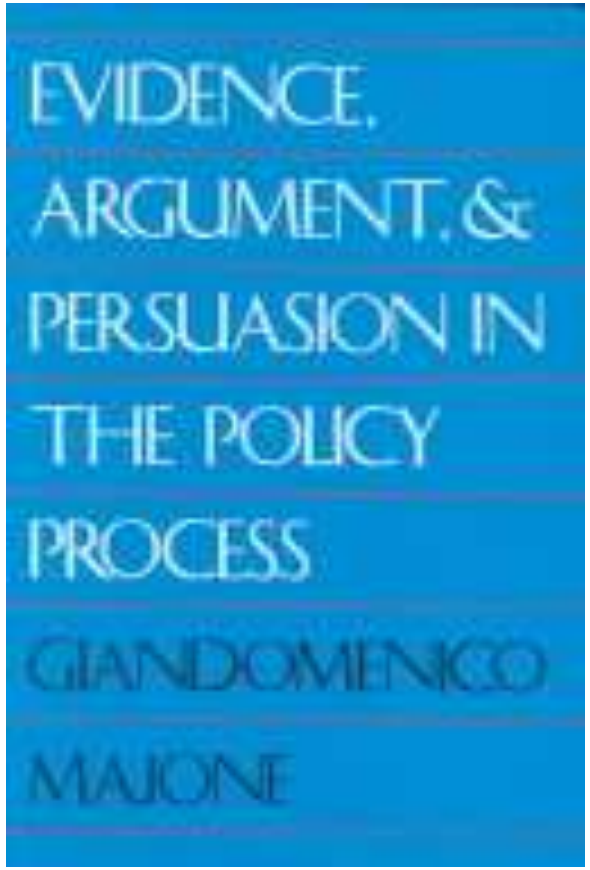
$$\begin{matrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ x_{N1} & x_{N2} & \cdots & x_{Nk} \end{matrix}$$

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



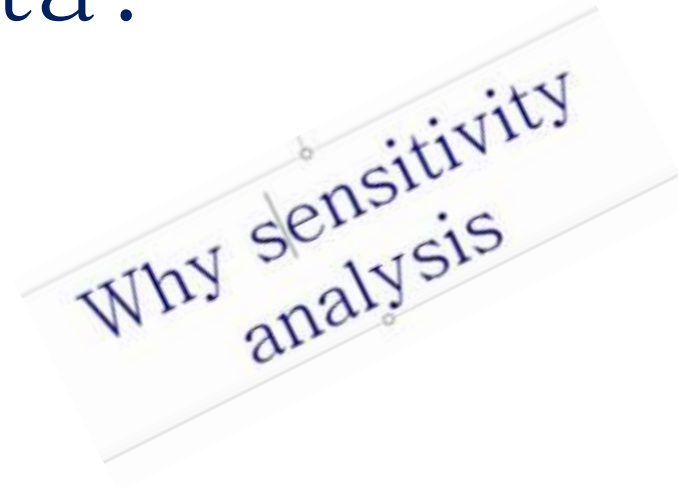
Examples of distributions of input factors

Why sensitivity analysis

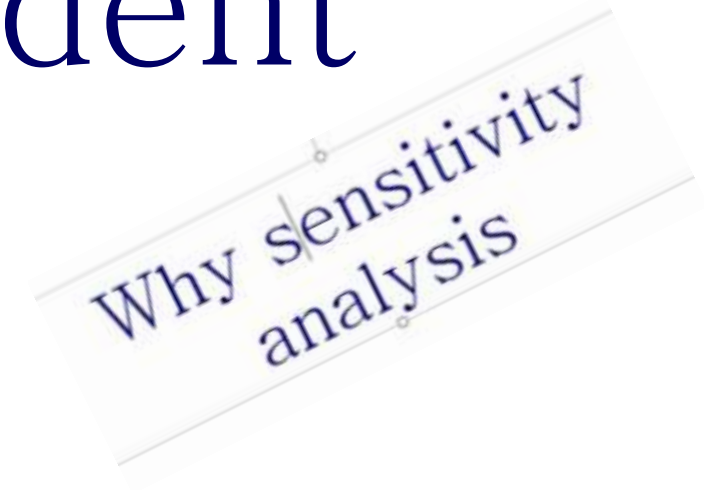


"Are the results from a particular model more sensitive to changes in the model and the methods used to estimate its parameters, or to changes in the data?"

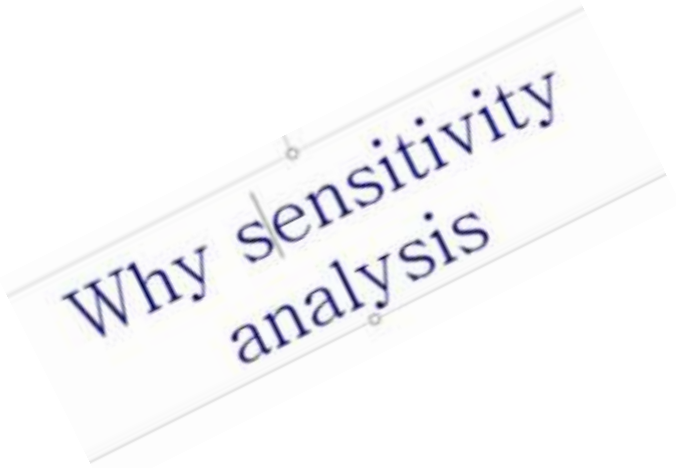
… SA can tell



Model-based knowing is conditional; SA makes the conditionality evident



SA can detect garbage in garbage out (GIGO)



Funtowicz & Ravetz's GIGO (Garbage In, Garbage Out) Science "where uncertainties in inputs must be suppressed least outputs become indeterminate"

Leamer's "Conclusions are judged to be sturdy only if the neighborhood of assumptions is wide enough to be credible and the corresponding interval of inferences is narrow enough to be useful"

S. Funtowicz and J. R. Ravetz, *Uncertainty and Quality in Science for Policy*. Dordrecht: Kluwer, 1990; E. E. Leamer, "Sensitivity Analyses Would Help," *Am. Econ. Rev.*, vol. 75, no. 3, pp. 308–313, 1985.

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

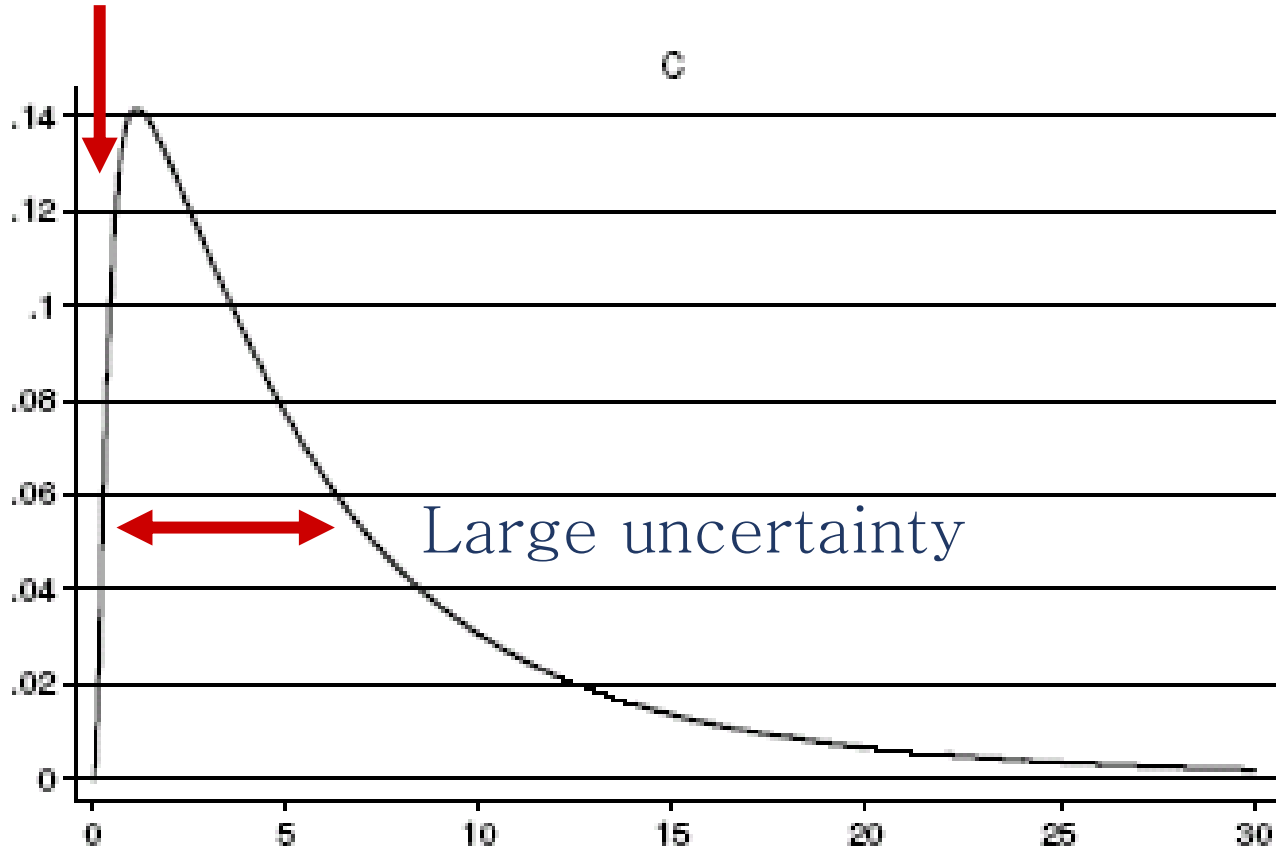
William Nordhaus, University of Yale Nobel 'Economics' 2018

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).

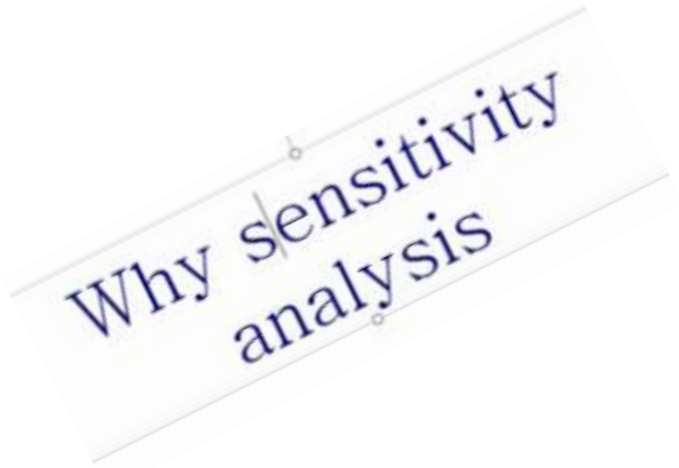
How was it done? A reverse engineering of the analysis

Missing points



% loss in GDP per capita

Finding all sorts of surprises



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?

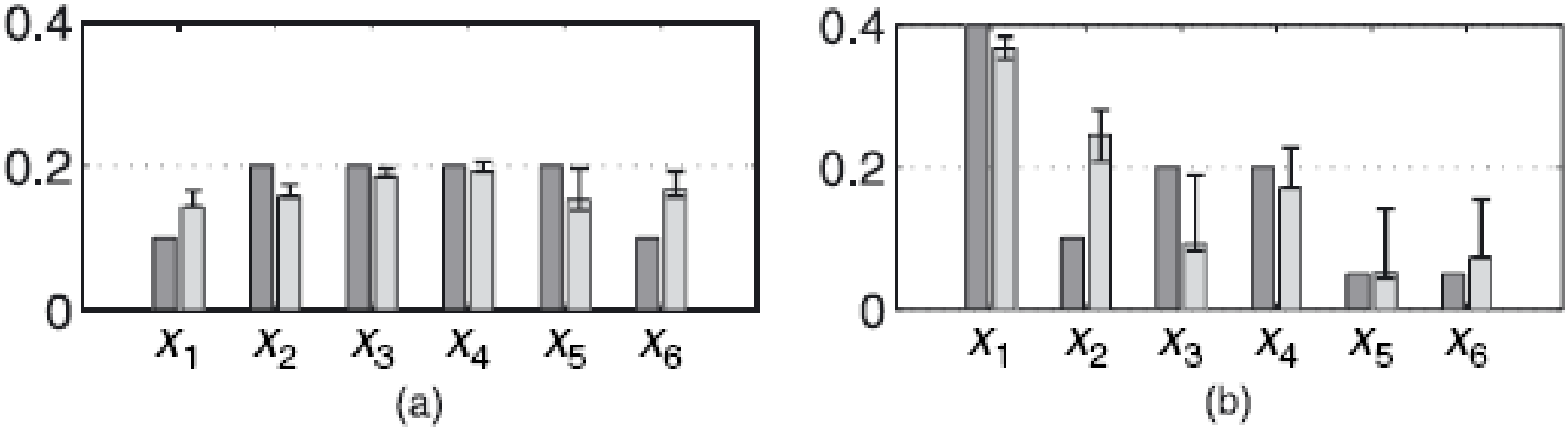
Paolo Paruolo

University of Insubria, Varese, Italy

and Michaela Saisana and Andrea Saltelli

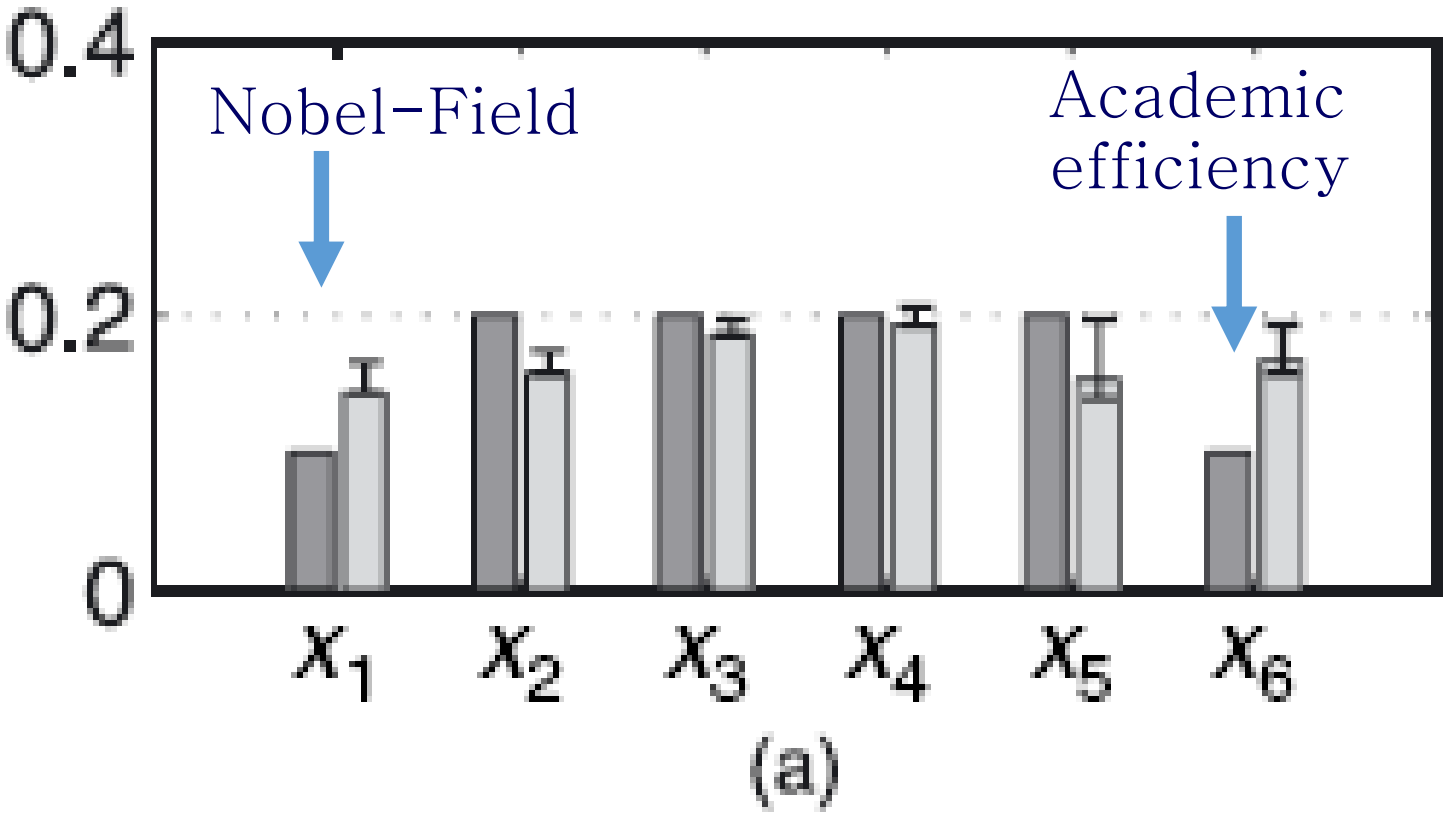
European Commission, Ispra, Italy

Why sensitivity analysis



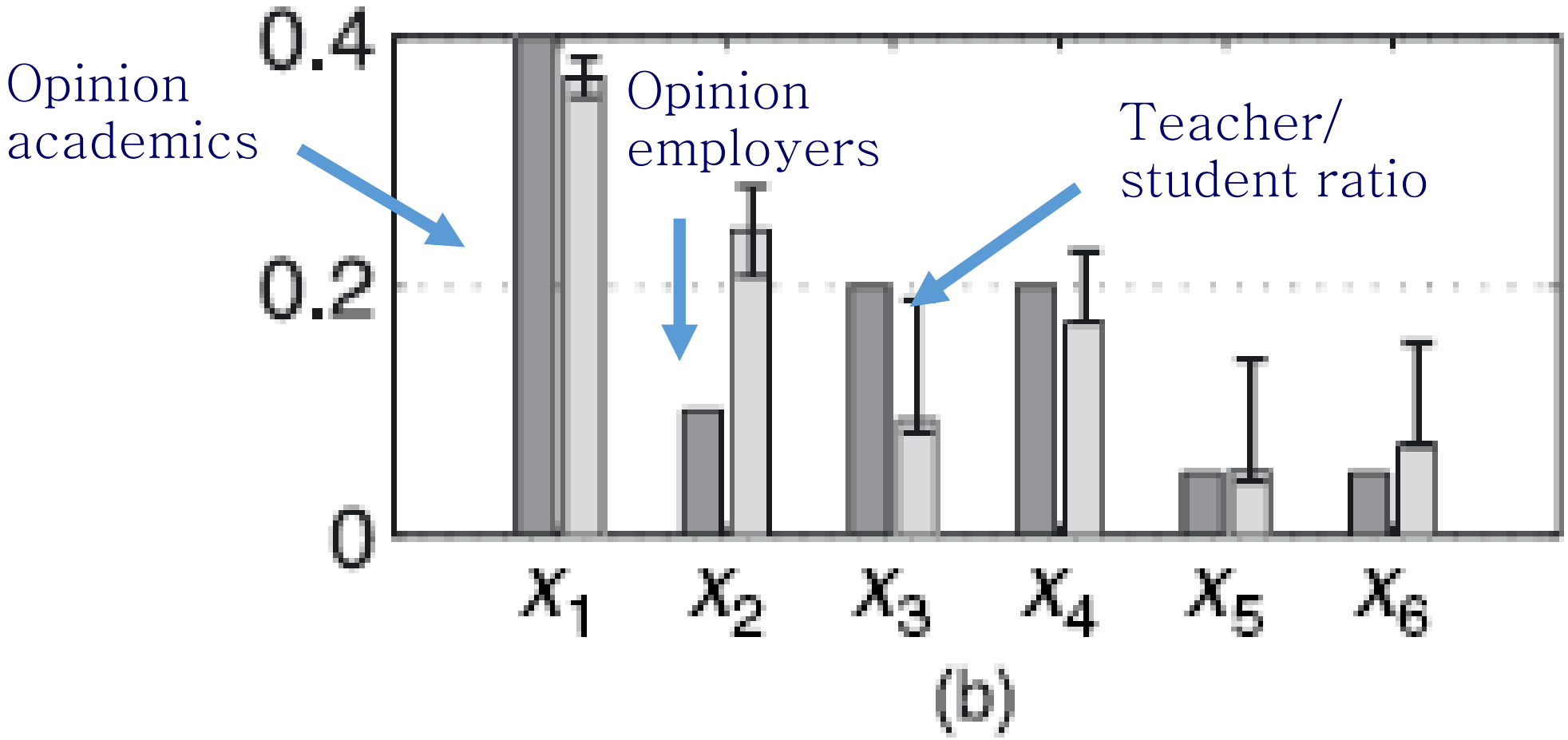
Importance of the factors according to the proponents (dark gray) and according to theory for ARWU (a) and THES (b)

Shanghai versus THES university ranking 2008



Proponents (dark grey) versus theory (light) for ARWU

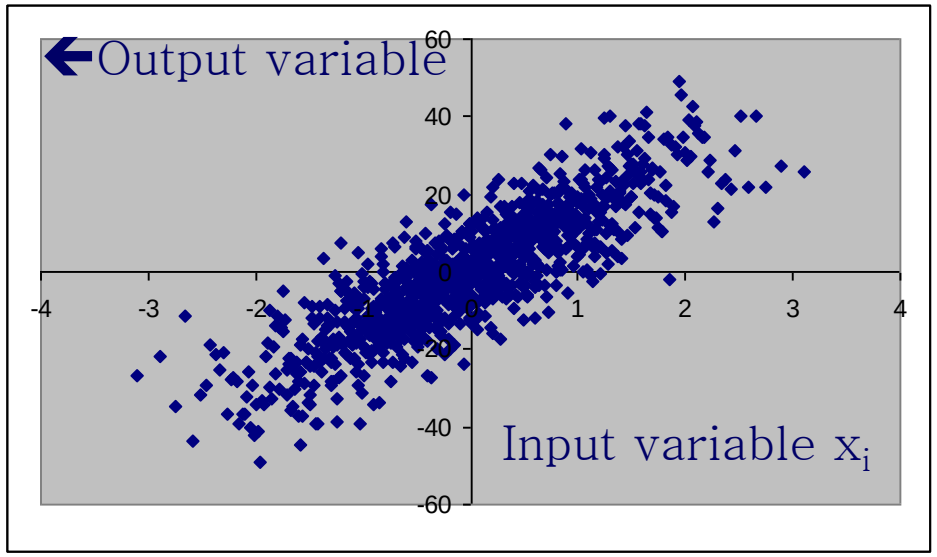
Shanghai university ranking 2008

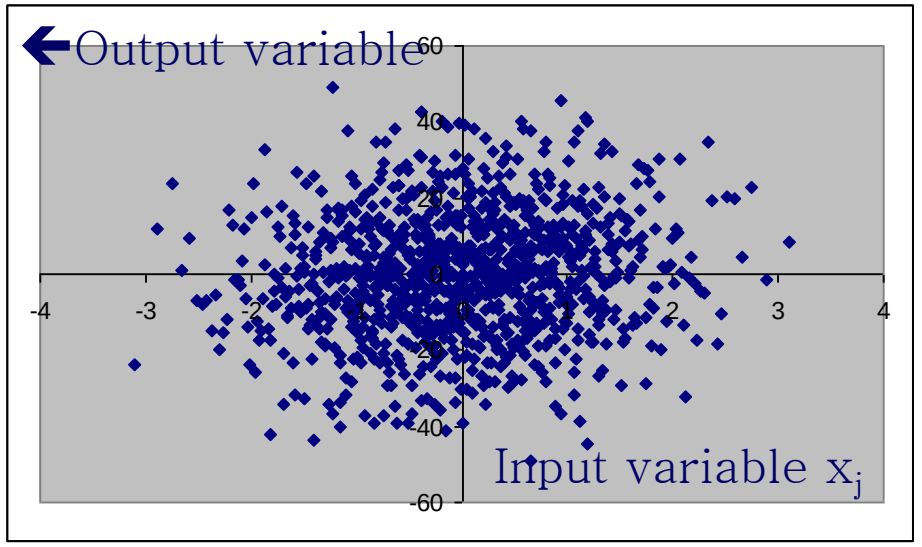


Proponents (dark grey) versus theory (light) for THES

THES university ranking 2008

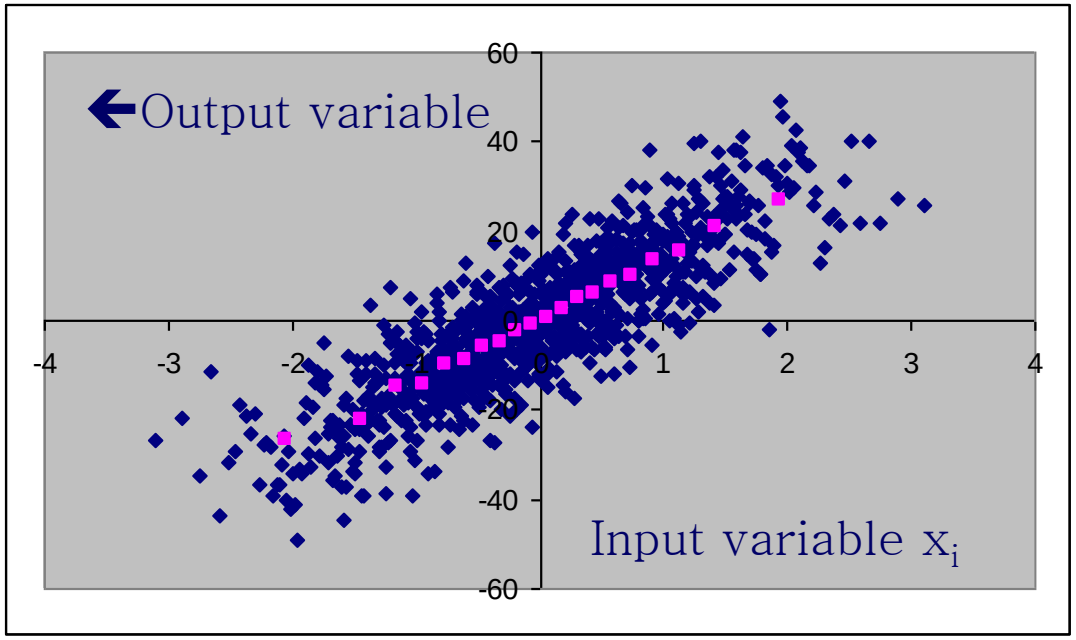
Why using variance-based sensitivity analysis methods

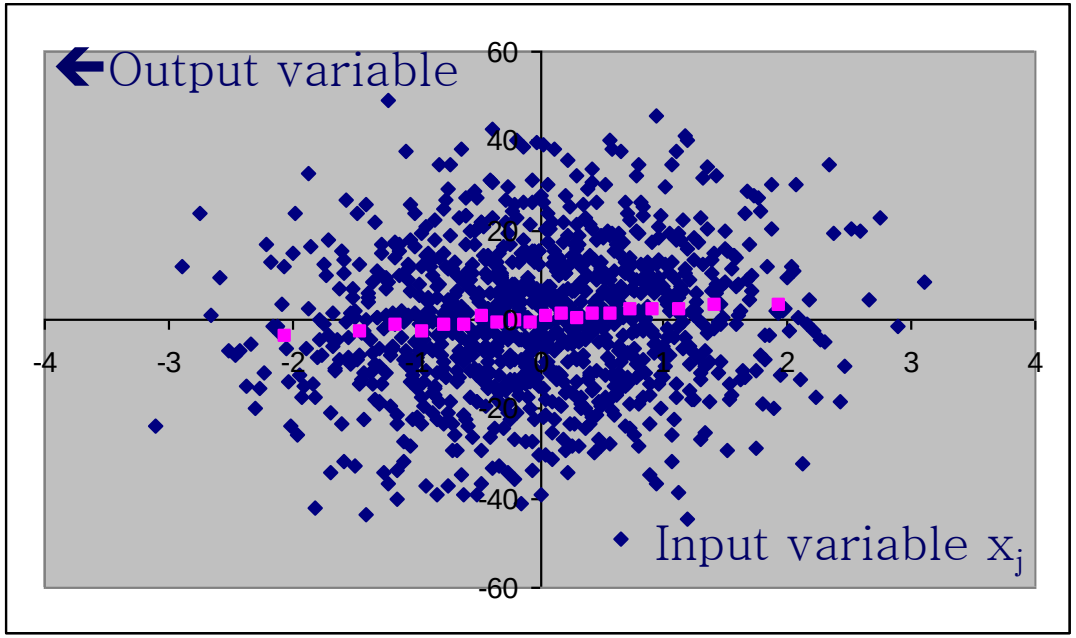




Plotting the output as a function of two different input factors

Which factor is more important?

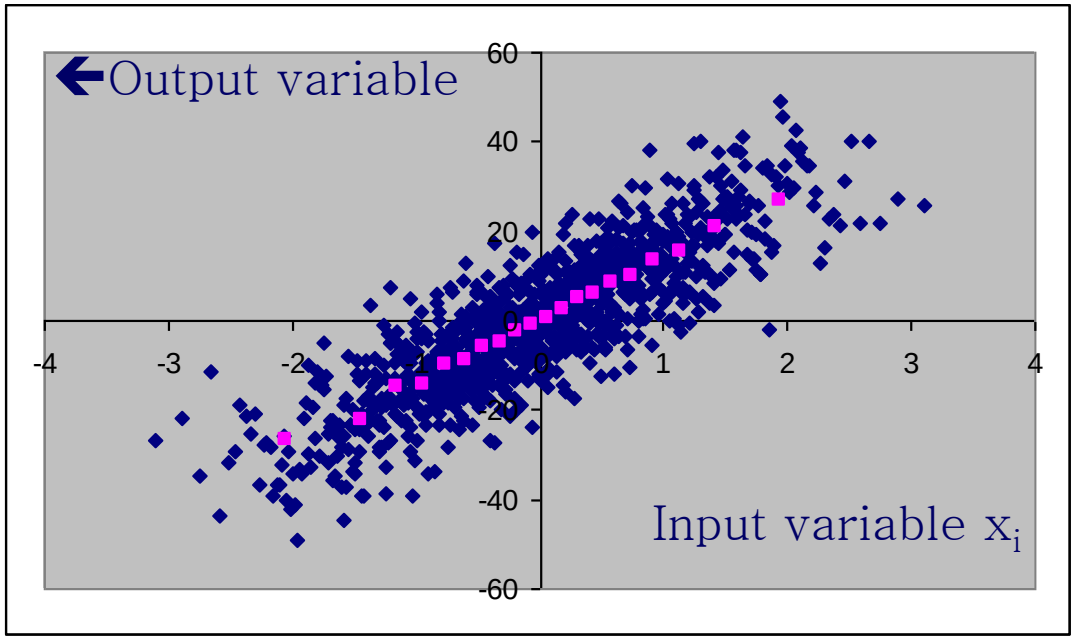




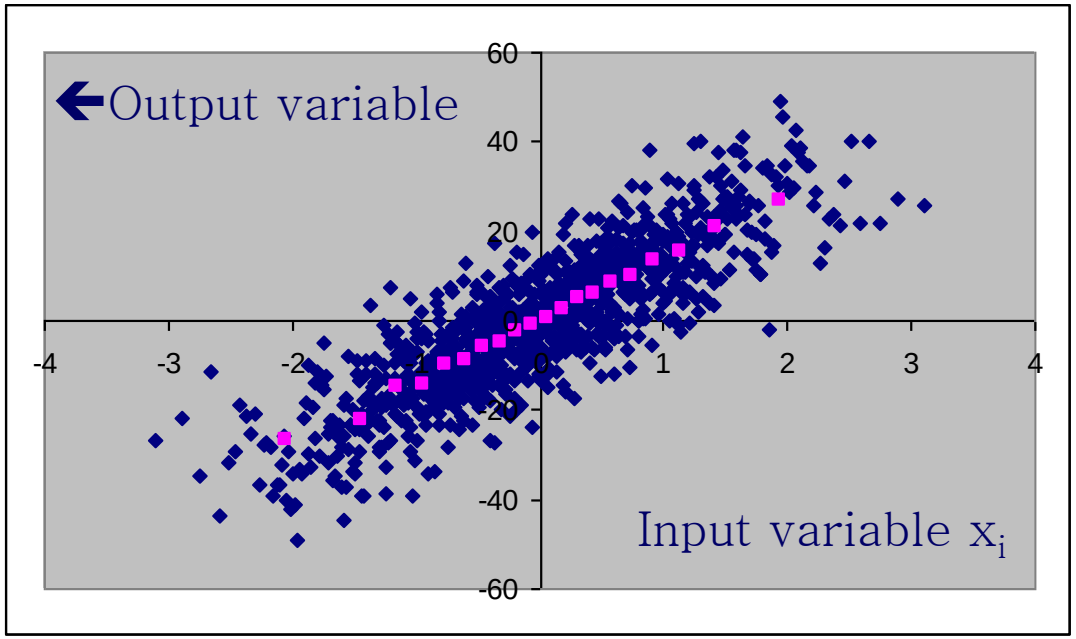
~1,000 blue points

Divide them in 20 bins of ~ 50 points

Compute the bin's average (pink dots)

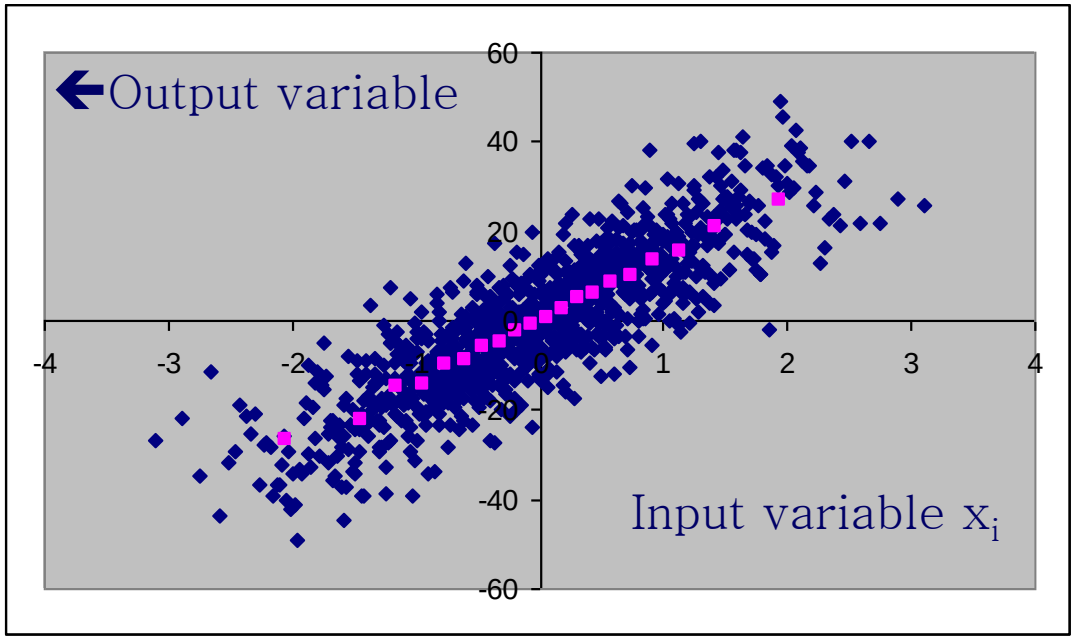


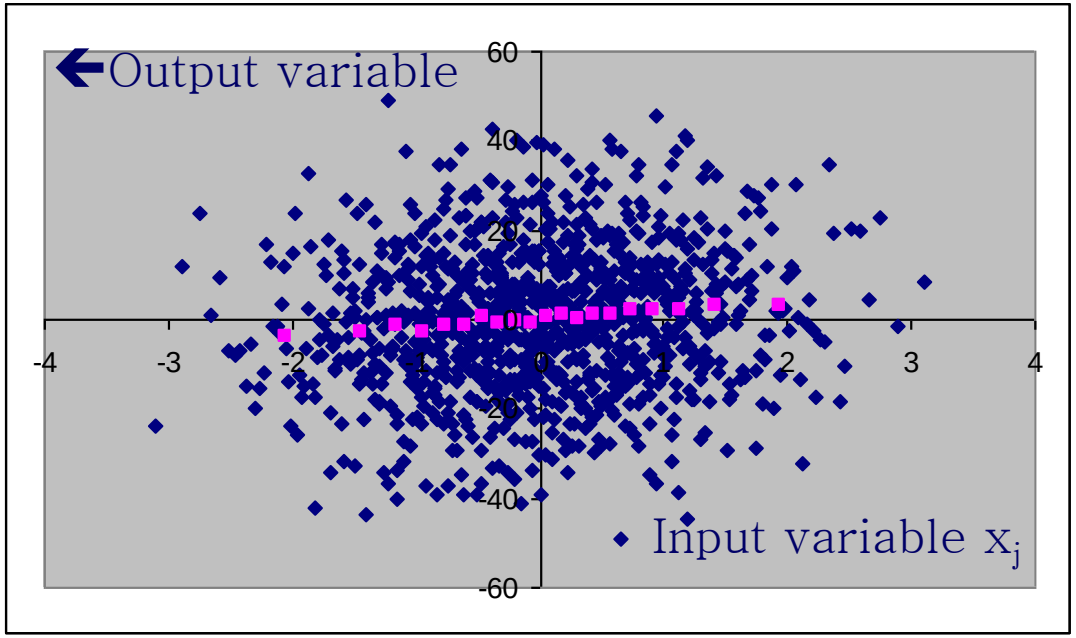
Each pink point is ~ $E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)$



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

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

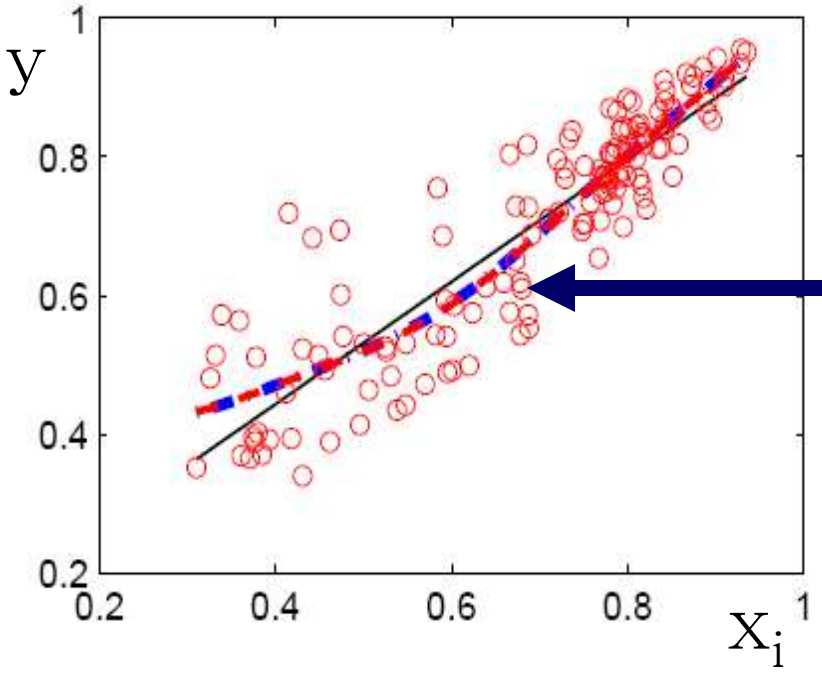




Which factor has the highest

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

$$S_i \equiv \frac{V\left(E\left(Y|X_i\right)\right)}{V_Y}$$



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

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

First order effect, or top marginal variance = the expected reduction in variance that would be achieved if factor Xi could be fixed.

Why?

Because:

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

Easy to prove using $V(Y)=E(Y^2)-E^2(Y)$

Because:

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

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

This is the variance when a factor Xi is fixed …

Because:

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

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

This is what variance would be left (on average) if Xi could be fixed…

… then this …

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

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

… must be the expected reduction in variance that would be achieved if factor Xi could be fixed

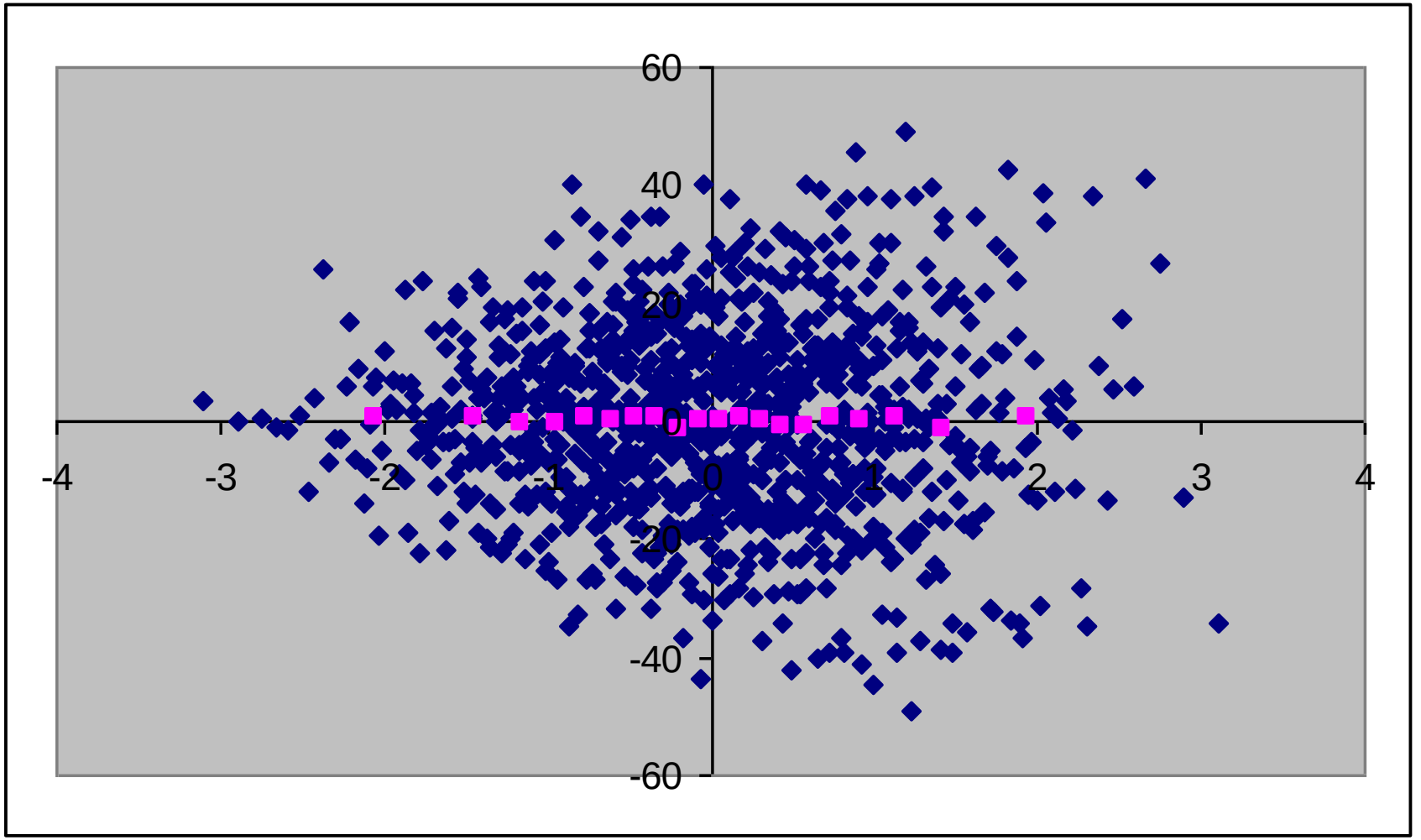
For additive models one can decompose the total variance as a sum of first order effects

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

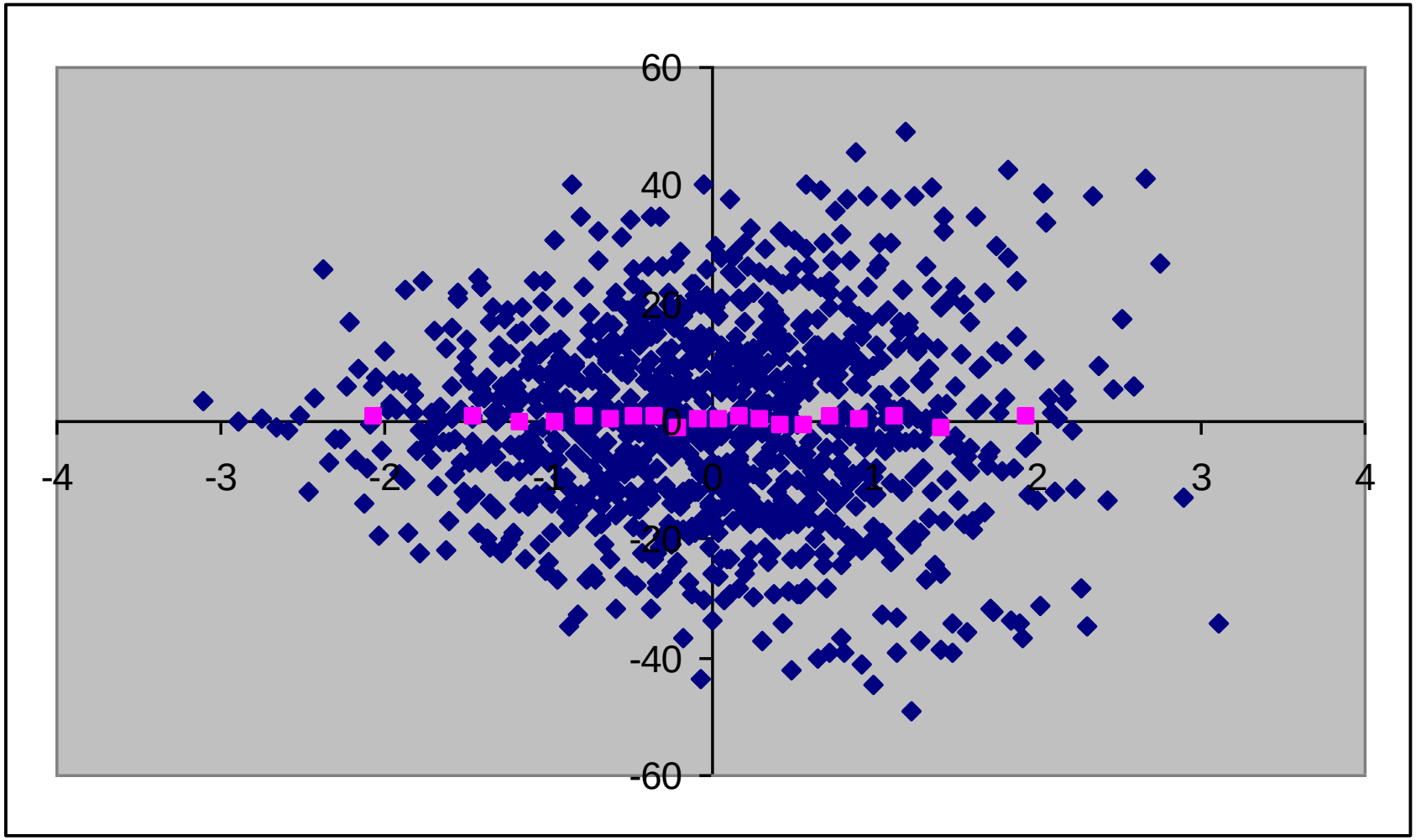
… which is also how additive models are defined

Non additive models can also be effectively dealt with, in the context of variance decomposition theory

Is $S_i = 0$?



Is this factor non-important?



Variance decomposition (ANOVA)

$$V(Y) =$$

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

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.

Thus given a model $Y=f(X_1,X_2,X_3)$

Decomposition:

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

Thus given a model $Y=f(X_1,X_2,X_3)$

Same decomposition divided by V

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

One may be interested in the total effects:

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

(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)

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

Why? Because it is the complement (resp. unconditional variance) of the first order effect of non Xi

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

What is the shortcoming of S_{Ti}?

Summary: advantages with variance based methods:

- graphic interpretation via scatterplots
- statistical interpretation
- expressed in plain English
- (working with sets)
- (relation to settings)

Coding S_i and S_{Ti} yourself?

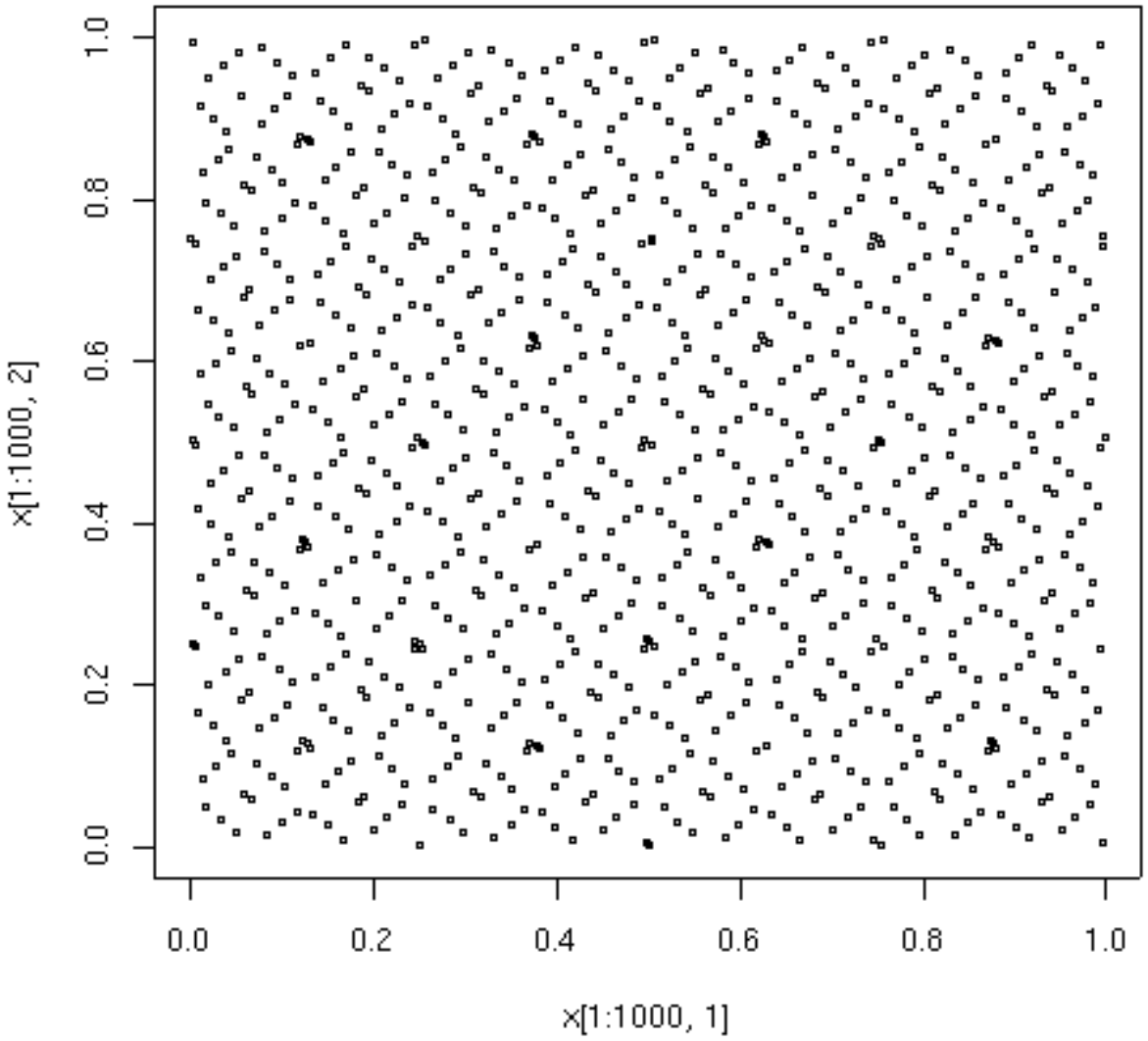
Use this work:

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010, Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Computer Physics Communications, 181, 259–270.

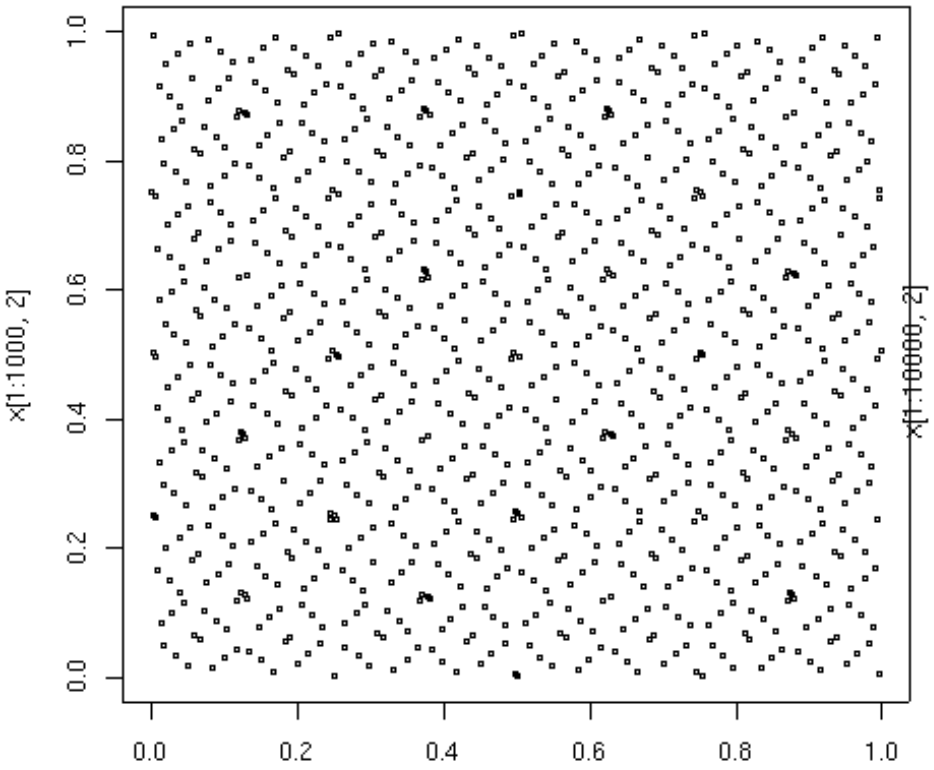
http://www.andreasaltelli.eu/file/repository/PUBLISHED_PAPER.pdf

How to generate the random sample?

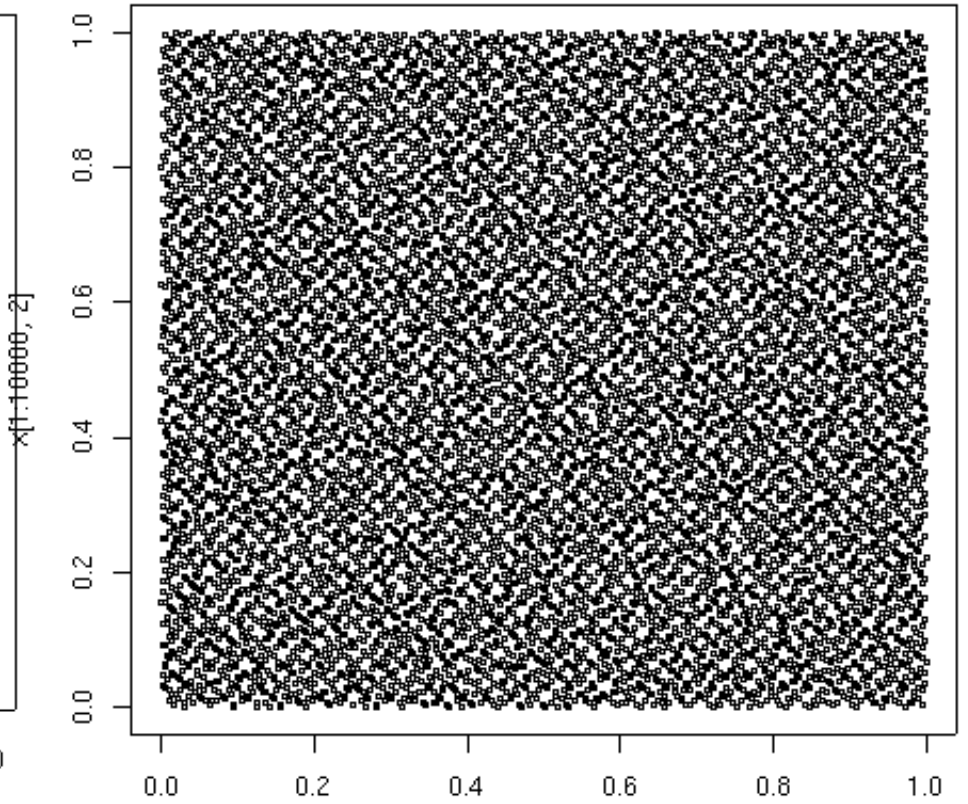
Quasi random sequences developed by I.M. Sobol’



An LP_τ sequence

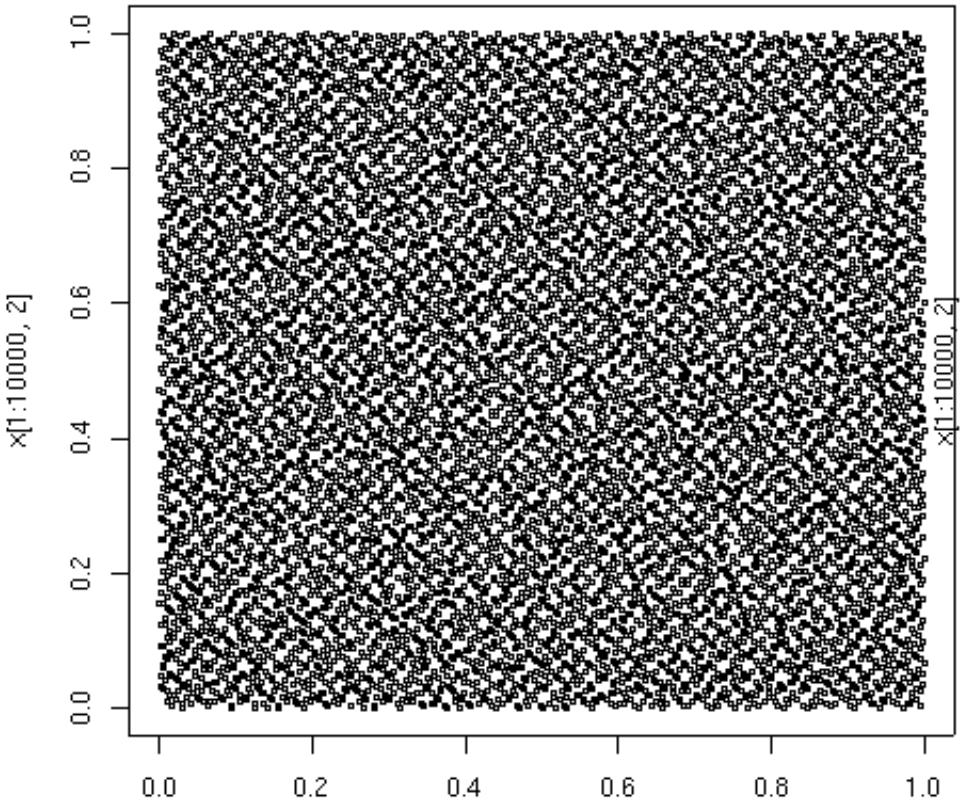


X1,X2 plane, 1000 Sobol' points

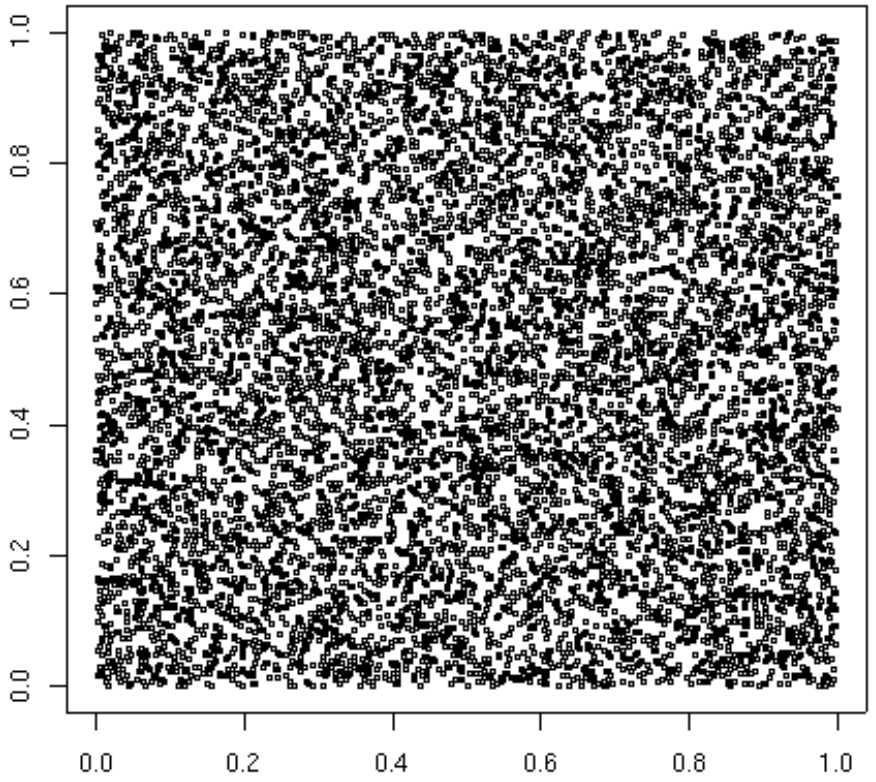


X1,X2 plane, 10000 Sobol' points

Sobol' sequences of quasi-random points



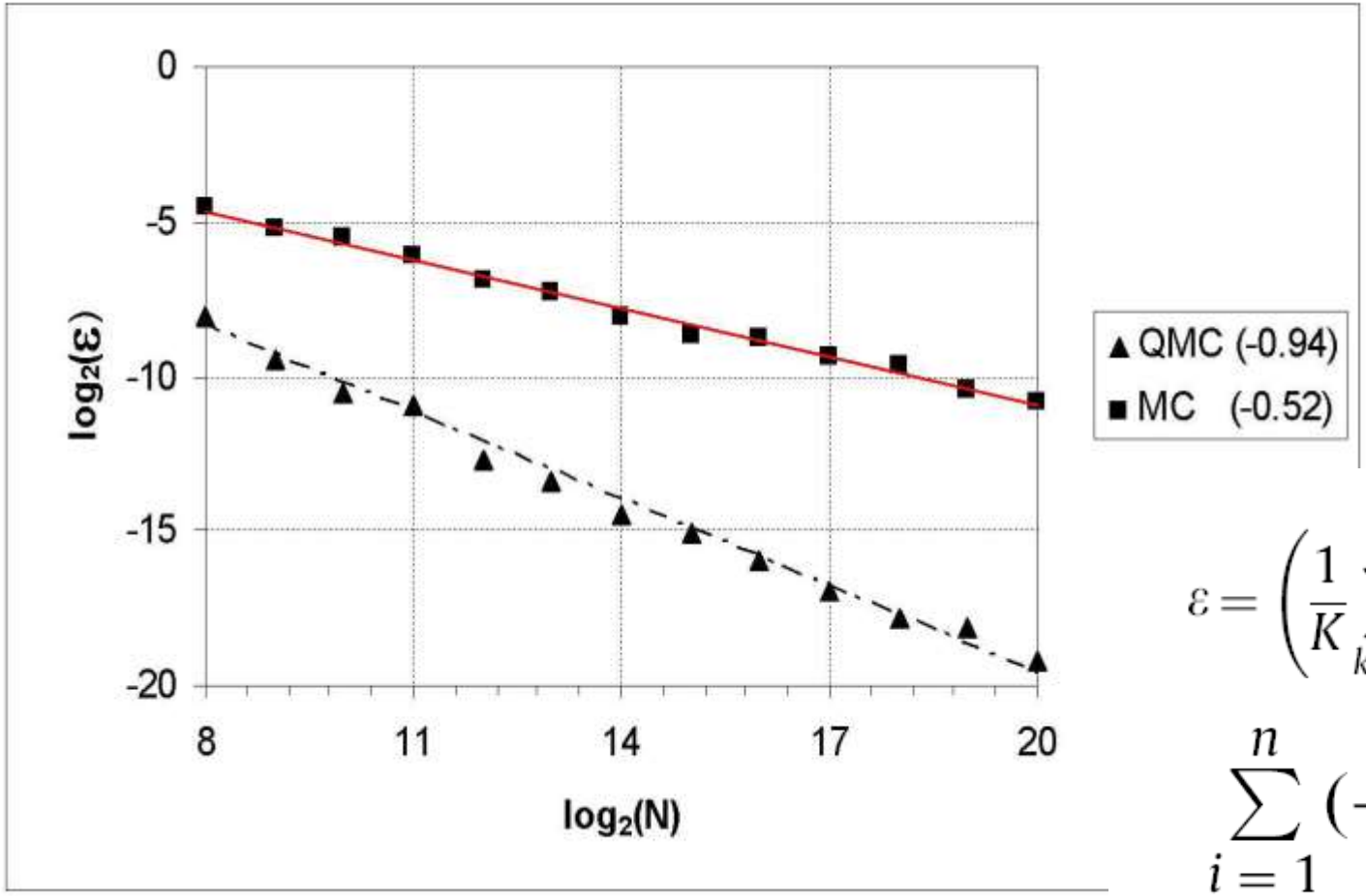
X1,X2 plane, 10000 Sobol' points



X1,X2 plane, 10000 random points

Sobol' sequences of quasi-random points against random points

Why quasi-random



Sergei Kucherenko,
Imperial College London

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

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

Root mean square error over K=50 different trials. The error refers to the numeric-versus-analytic value the integral of the function (for n=360) over its dominion.

Source: 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.

Secrets of sensitivity analysis

Why should one ever run a model just once?

➔ While developing a model run it a few more times at a time (e.g. 16, 32,..)

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

Sensitivity analysis should not be used to hide assumptions

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

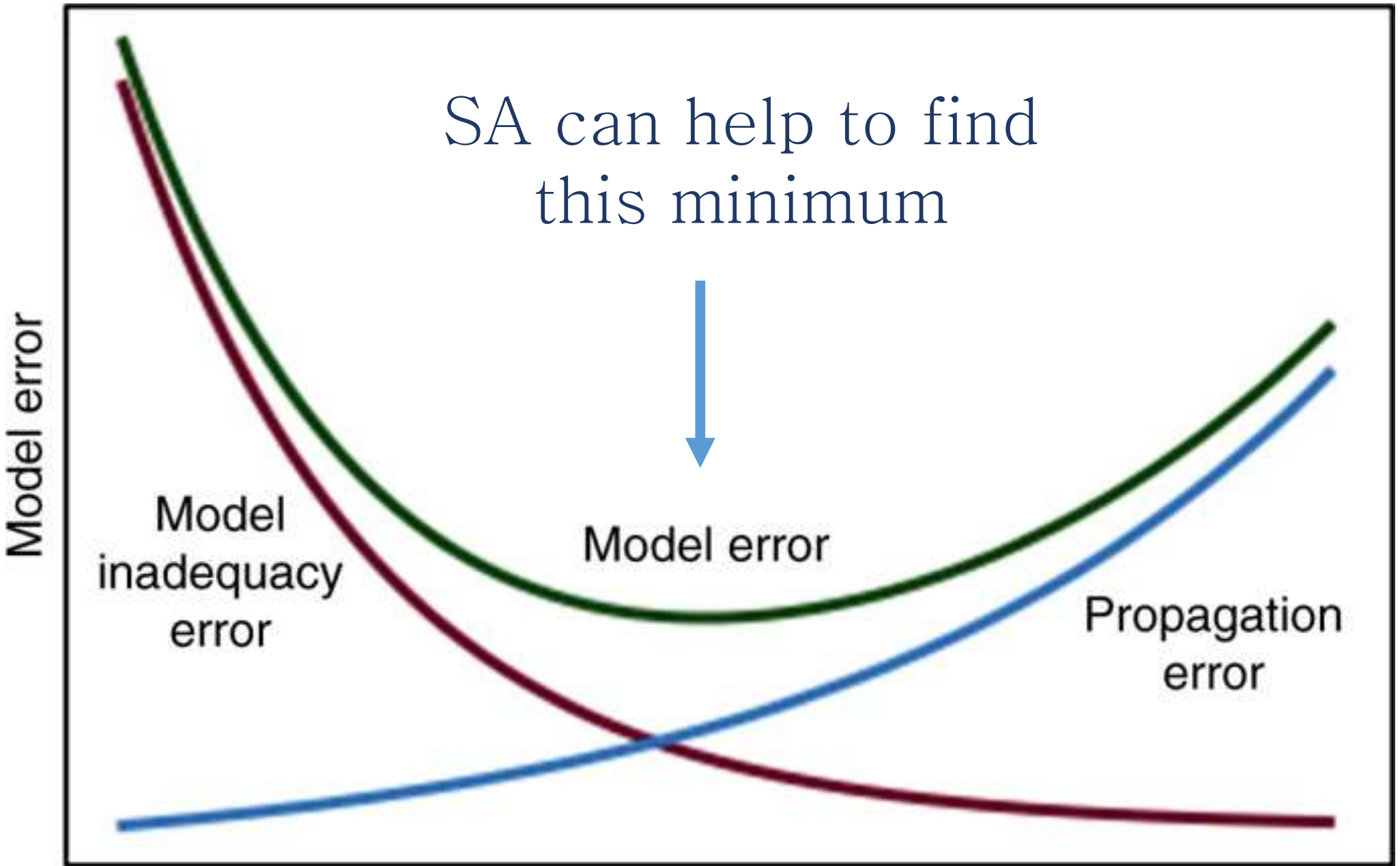
[Often the love for one's own model prevails]

Badly kept secret:

There is always one more bug!

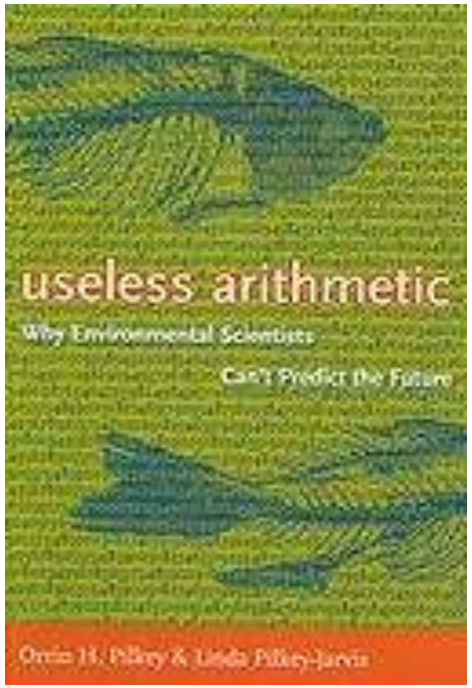
= Lubarsky's Law of Cybernetic Entomology

Use SA to calibrate complexity

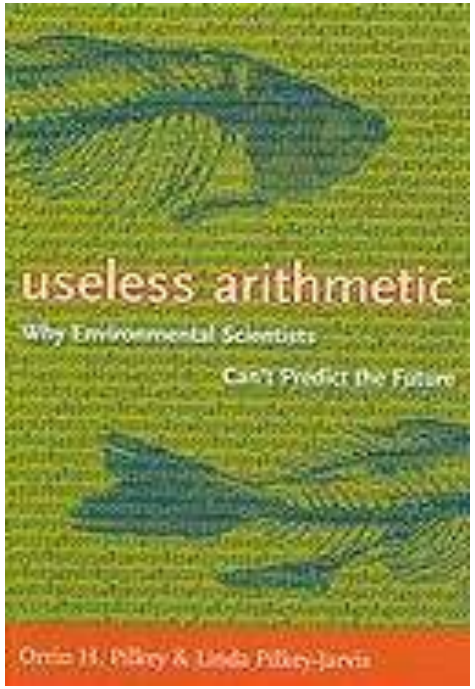


Limit of SA: the map is not the territory

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.

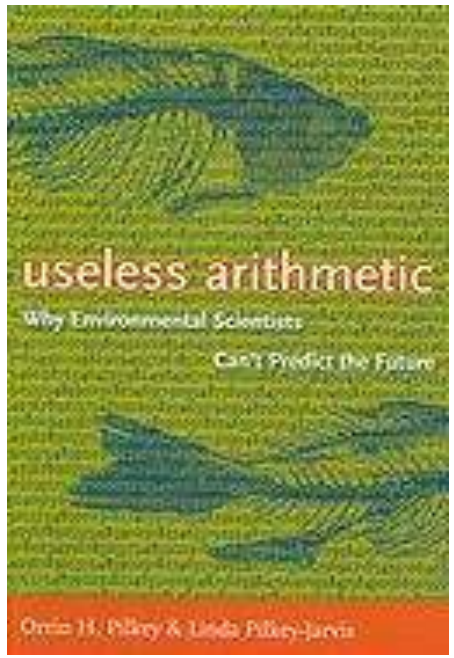


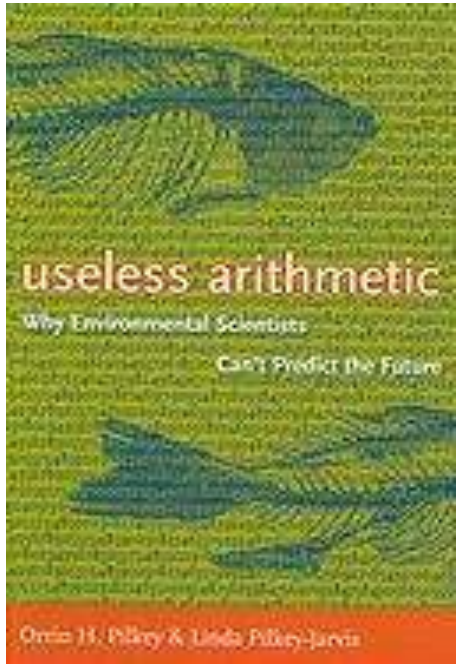
<<It is important to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature.

[…] 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.>>

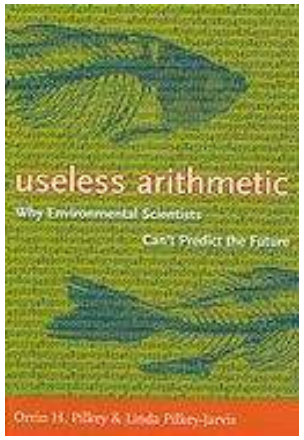
One of the examples discussed concerns the Yucca Mountain repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis.

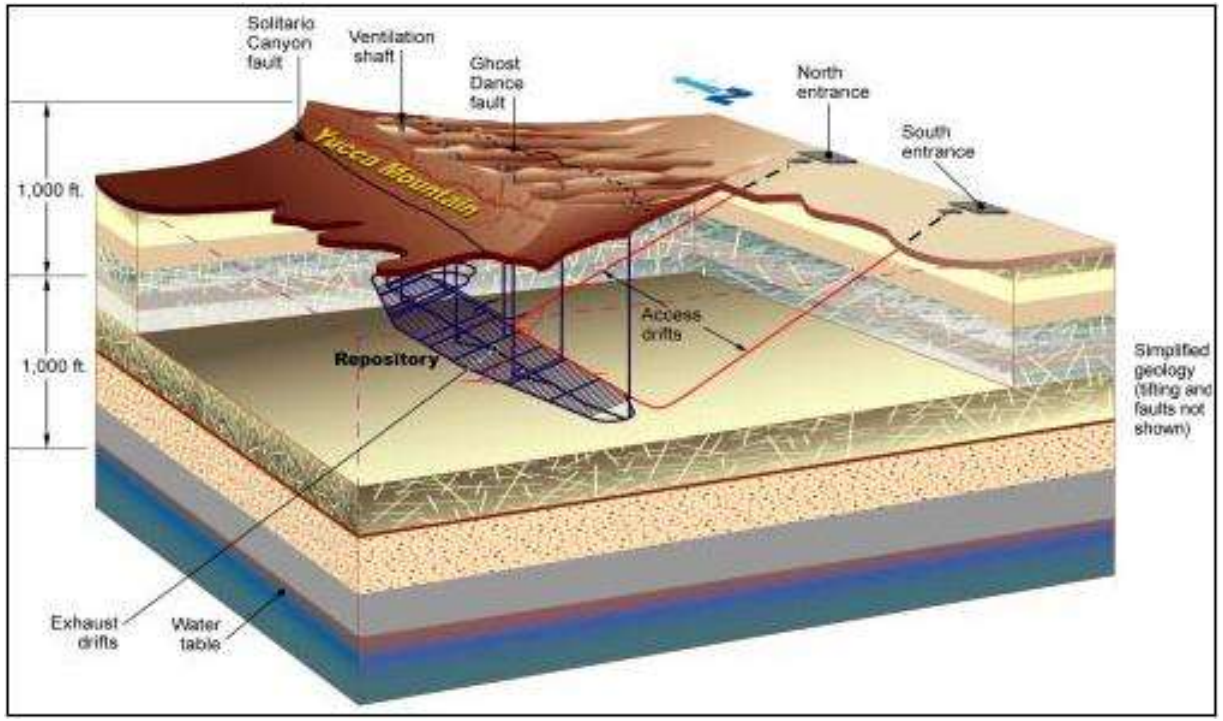
TSPA is Composed of 286 sub-models.





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.

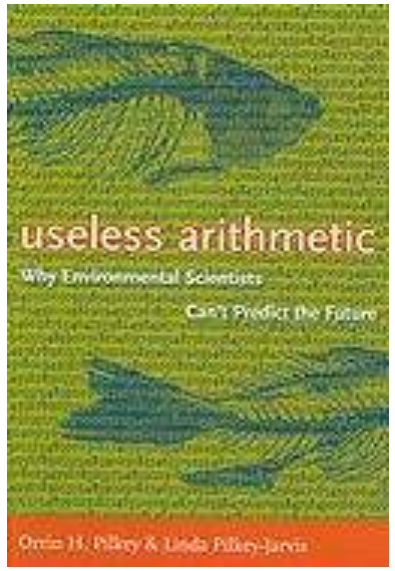




Evidence was produced leading to an upward revision of 4 orders of magnitude of this parameter (the ^{36}Cl story)

In the model 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"

Ref. ➔ Robert K. Merton's 'Organized skepticism '

Robert K. Merton

Communalism – the common ownership of scientific discoveries, according to which scientists give up intellectual property rights in exchange for recognition and esteem (Merton actually used the term Communism, but had this notion of communalism in mind, not Marxism);

Universalism – according to which claims to truth are evaluated in terms of universal or impersonal criteria, and not on the basis of race, class, gender, religion, or nationality;

Disinterestedness – according to which scientists are rewarded for acting in ways that outwardly appear to be selfless;

Organized Skepticism – all ideas must be tested and are subject to rigorous, structured community scrutiny.

Limit of SA: Often no SA or wrong kind of SA applied

Science of the Total Environment 568 (2016) 666–670

Contents lists available at ScienceDirect

Science of the Total Environment

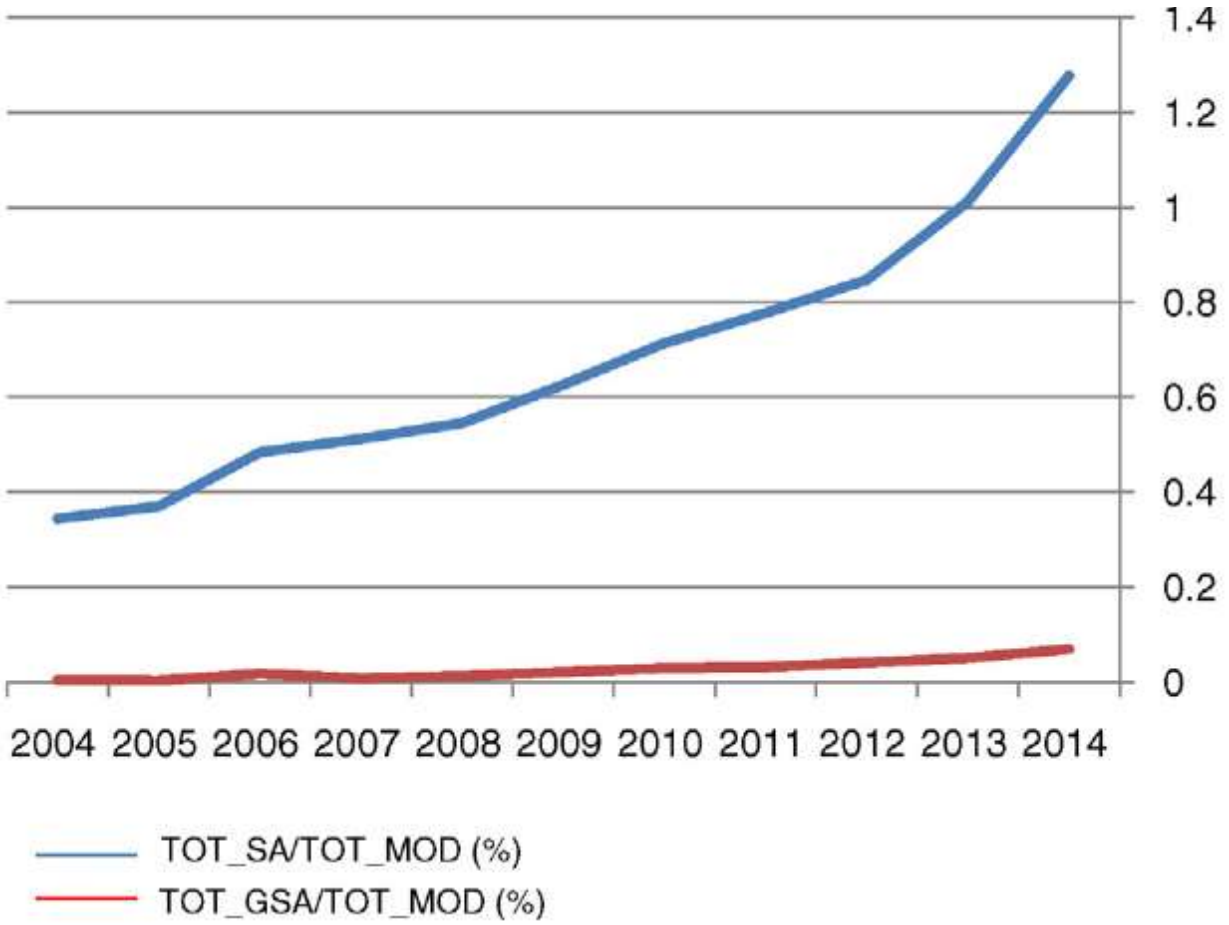
journal homepage: www.elsevier.com/locate/scitotenv

Trends in sensitivity analysis practice in the last decade

Federico Ferretti [a,*], Andrea Saltelli [b,d], Stefano Tarantola [c]

[a] European Commission, Joint Research Centre (JRC), Unit of Econometrics and Applied Statistics, via Enrico Fermi 2749 TP 361, Ispra, 21027 VA, Italy
[b] Centre for the Study of the Sciences and the Humanities (SVT), University of Bergen (UIB), Spain
[c] European Commission, Joint Research Centre (JRC), Institute for Energy and Transport, via Enrico Fermi 2749 TP 690, Ispra, 21027 VA, Italy
[d] Institut de Ciència i Tecnologia Ambientals (ICTA), Universitat Autonoma de Barcelona, Spain

In 2014 out of 1000 papers in modelling 12 have a sensitivity analysis and < 1 a global SA; most SA still move one factor at a time



Environmental Modelling & Software 25 (2010) 1508–1517

Contents lists available at ScienceDirect

Environmental Modelling & Software

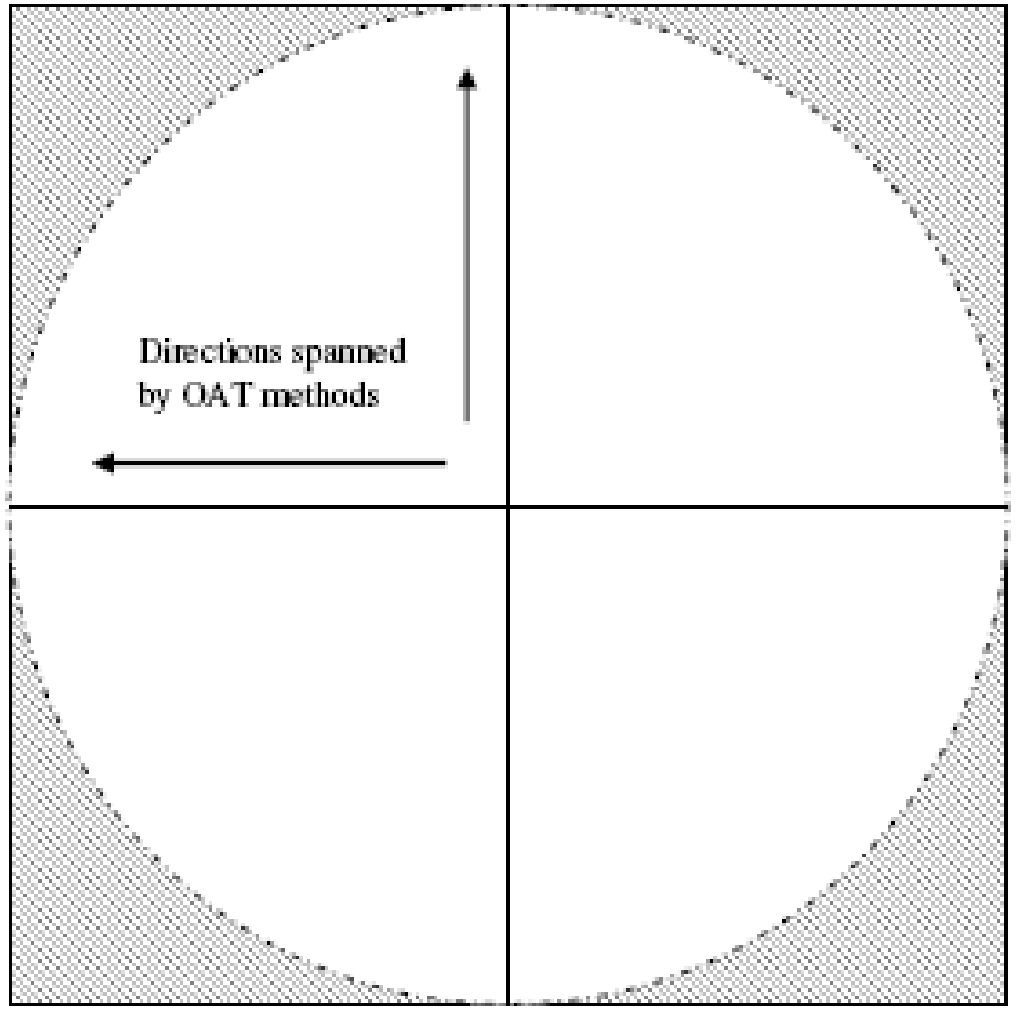
journal homepage: www.elsevier.com/locate/envsoft

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

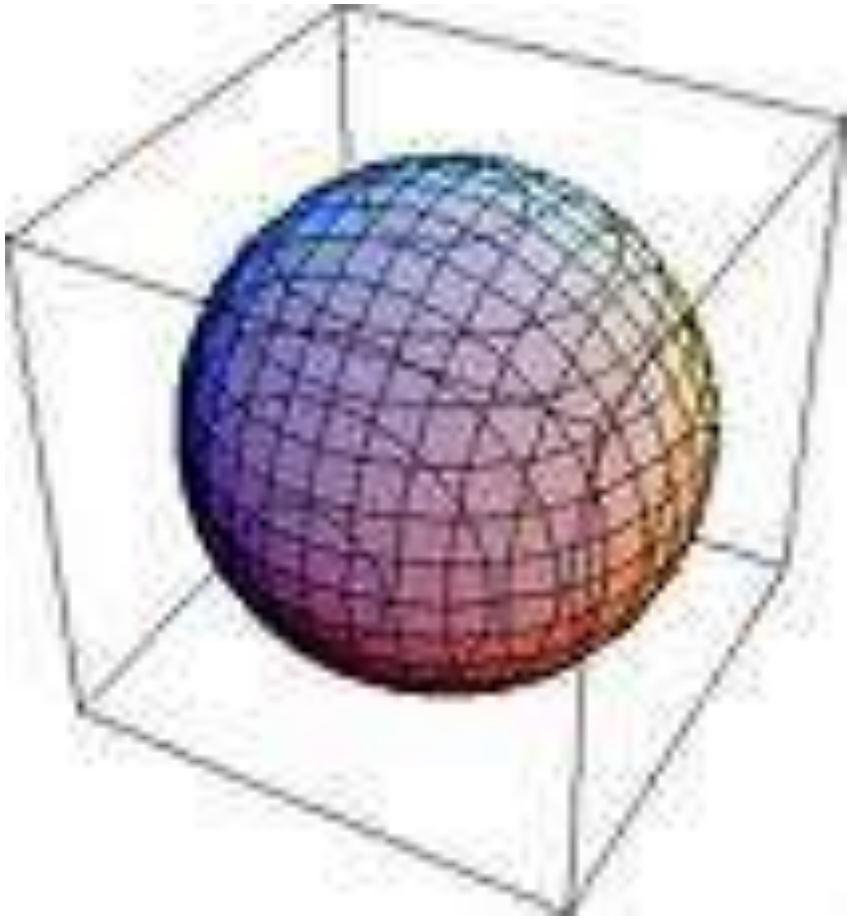
OAT in 2 dimensions



Area circle / area square =?

~ 3/4

OAT in 3 dimensions

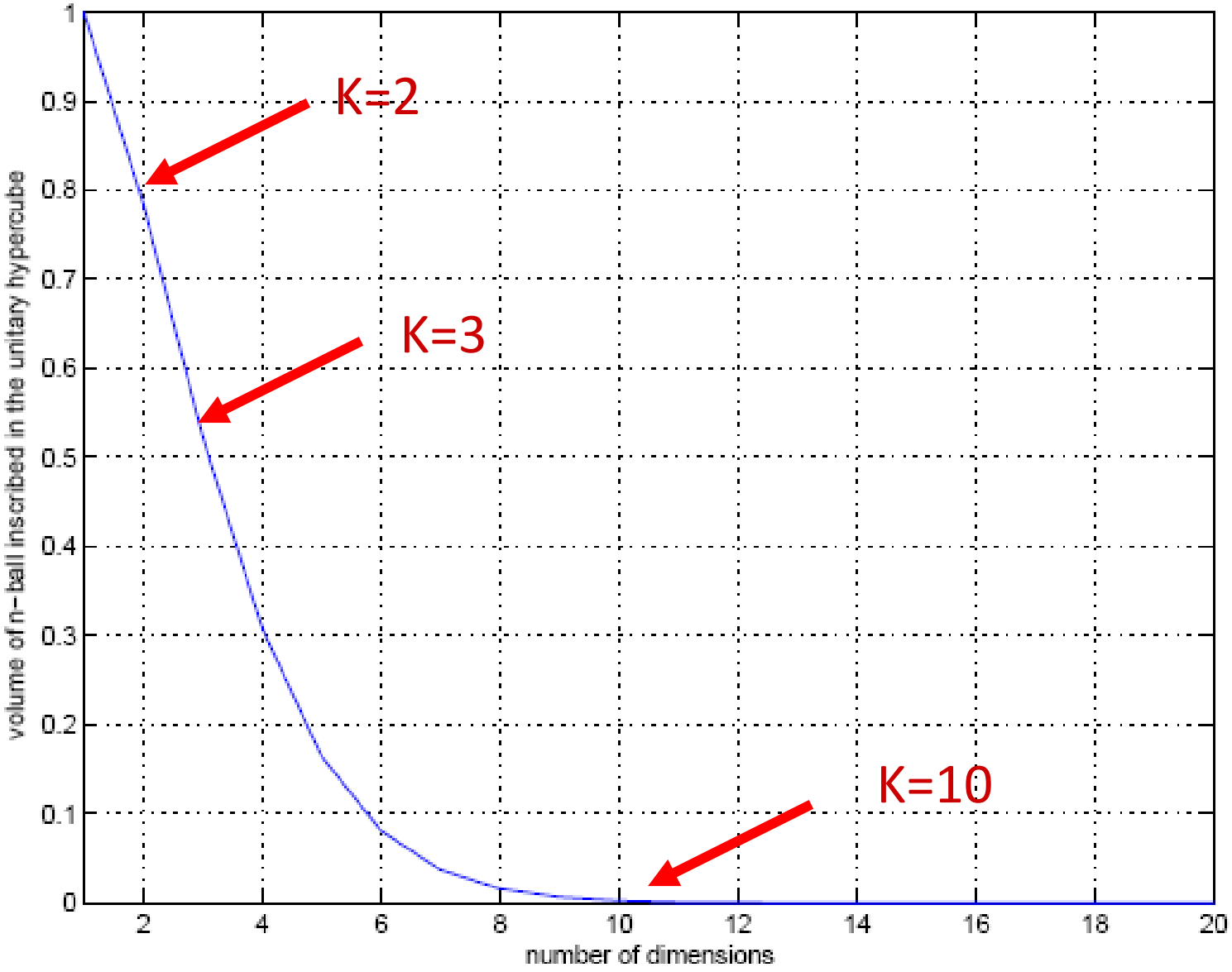


Volume sphere / volume cube =?

~ 1/2

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

OAT in k dimensions



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]

A systematic review of 280 scientific papers mentioning sensitivity analysis, focusing on highly cited works

42% of highly cited papers present a SA of poor quality

Beyond sensitivity analysis: sensitivity auditing

Sensitivity auditing

EC guidelines on impact assessment 2015, and SAPEA report 2019

SAPEA

Science Advice for Policy by European Academies

ISSUES
IN SCIENCE AND TECHNOLOGY

VOL. XXX, NO. 2, WINTER 2014

When All Models Are Wrong

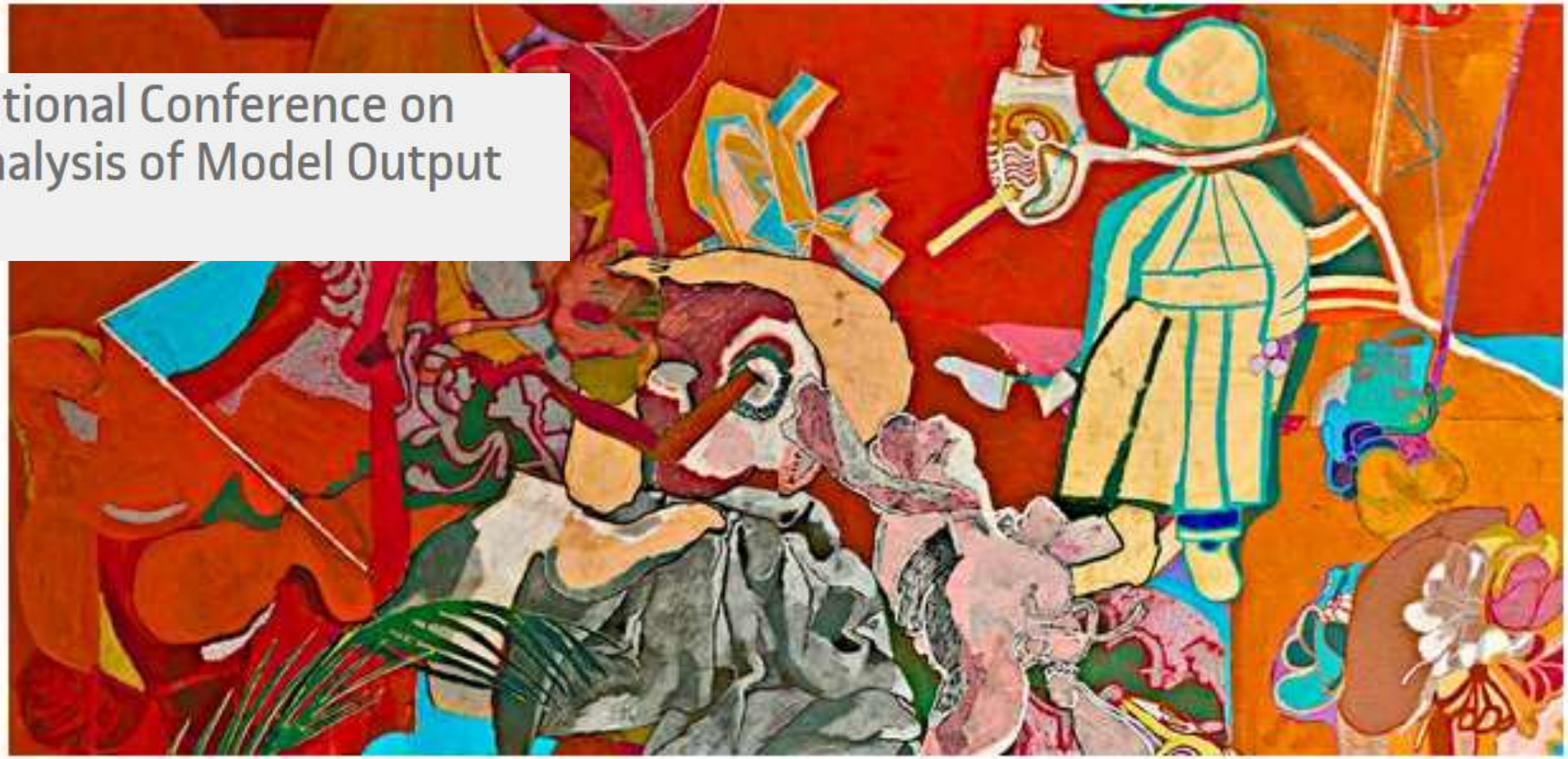
BY ANDREA SALTELLI, SILVIO FUNTOWICZ

The rules of sensitivity auditing

1. Check against rhetorical use of mathematical modelling;
2. Adopt an “assumption hunting” attitude; focus on unearthing possibly implicit assumptions;
3. Check if uncertainty been instrumentally inflated or deflated.

4. Find sensitive assumptions before these find you; do your SA before publishing;
5. Aim for transparency; Show all the data;
6. Do the right sums, not just the sums right;
7. Perform a proper global sensitivity analysis.

Ninth International Conference on Sensitivity Analysis of Model Output



Paula Rego, Self portrait in red (1962), detail, MNAC, Barcelona

CALL FOR PAPERS for the Ninth International Conference on Sensitivity Analysis of Model Output

Barcelona, October Monday 28 - Wednesday 30, 2019

END

@andreasaltelli