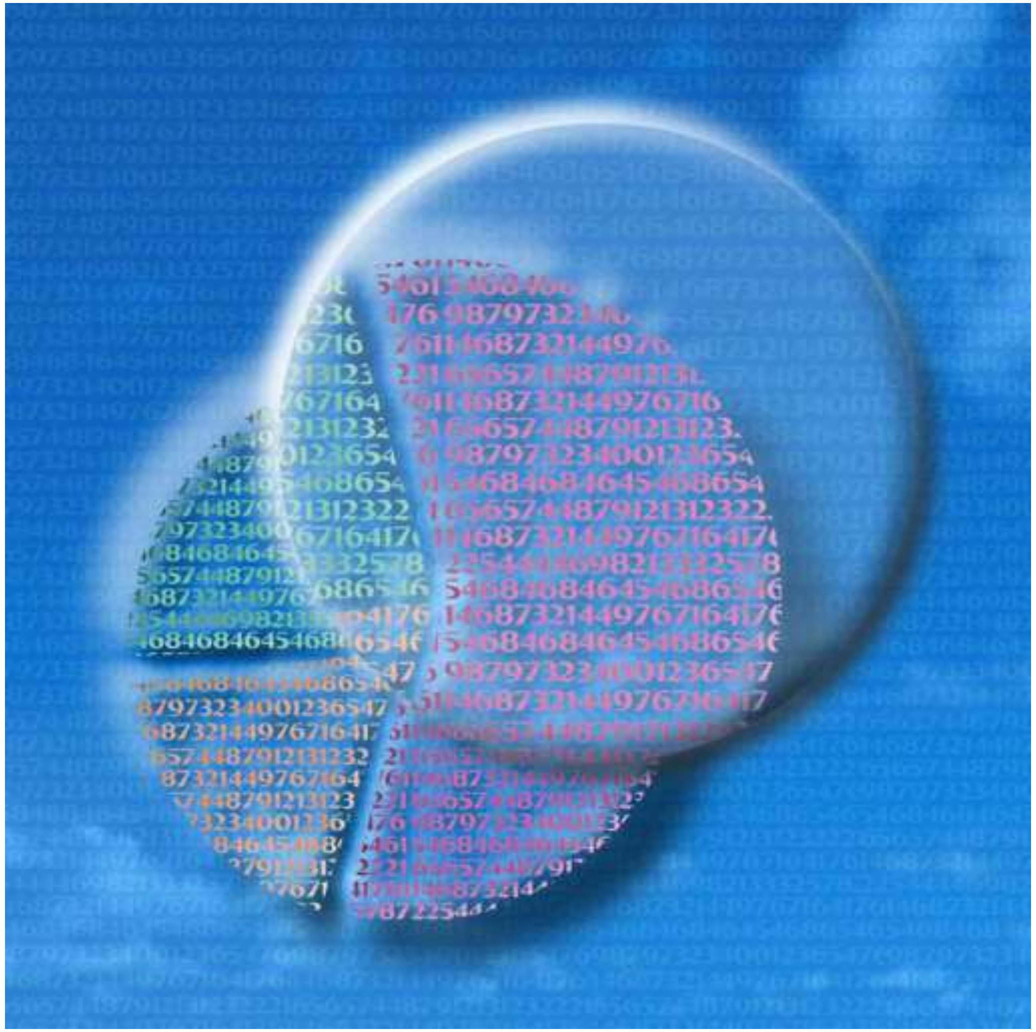


An introduction of sensitivity analysis

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

Summer School on Sensitivity Analysis – SAMO 2018
Ranco (Italy) – June 11-15, 2018
Conca Azzurra Hotel

Where to find this talk: www.andreasaltelli.eu

= more material on www.andreasaltelli.eu

About modelling

Statistics and algorithms in the spotlight; how about models?
What is a model? Models versus data: a blurring boundary

Statistics in the fray

The discipline of statistics has been going through a phase of critique and self-criticism, due to mounting evidence of poor statistical practice of which misuse and abuse of the P-test is the most visible sign

ASA News
AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®
732 North Washington Street, Alexandria, VA 22314 • (703) 684-1221 • Toll Free: (888) 231-3473 • www.amstat.org • www.twitter.com/AmstatNews

AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON STATISTICAL SIGNIFICANCE AND *P*-VALUES

Provides Principles to Improve the Conduct and Interpretation of Quantitative Science

March 7, 2016

+ twenty 'dissenting' commentaries

Wasserstein, R.L. and Lazar, N.A., 2016. 'The ASA's statement on p-values: context, process, and purpose', The American Statistician, DOI:10.1080/00031305.2016.1154108.

See also Christie Aschwanden at http://fivethirtyeight.com/features/not-even-scientists-can-easily-explain-p-values/

P-hacking (fishing for favourable p-values) and **HARKing** (formulating the research **H**ypothesis **A**fter the **R**esults are **K**nown);
Desire to achieve a sought for – or simply publishable – result leads to fiddling with the data points, the modelling assumptions, or the research hypotheses themselves

Leamer, E. E. Tantalus on the Road to Asymptopia. J. Econ. Perspect. 24, 31–46 (2010).

Kerr, N. L. HARKing: Hypothesizing After the Results are Known. Personal. Soc. Psychol. Rev. 2, 196–217 (1998).

A. Gelman and E. Loken, "The garden of forking paths: Why multiple comparisons can be a problem, even when there is no 'fishing expedition' or 'p-hacking' and the research hypothesis was posited ahead of time," 2013.

Big data and algorithms

Algorithms decide upon an ever-increasing list of cases, such as recruiting, carriers – including of researchers, prison sentencing, paroling, custody of minors…

Alexander, L. Is an algorithm any less racist than a human? The Guardian. Available at https//www.theguardian.com/technology/2016/aug/03/algorithm-racist-human-employers-work (2016) (Accessed: 30th August 2017).

Abraham C. Turmoil rocks Canadian biomedical research community. Statnews, Available at https://www.statnews.com/2016/08/01/cihr-canada-research/ (2016) (Accessed: 30th August 2017).

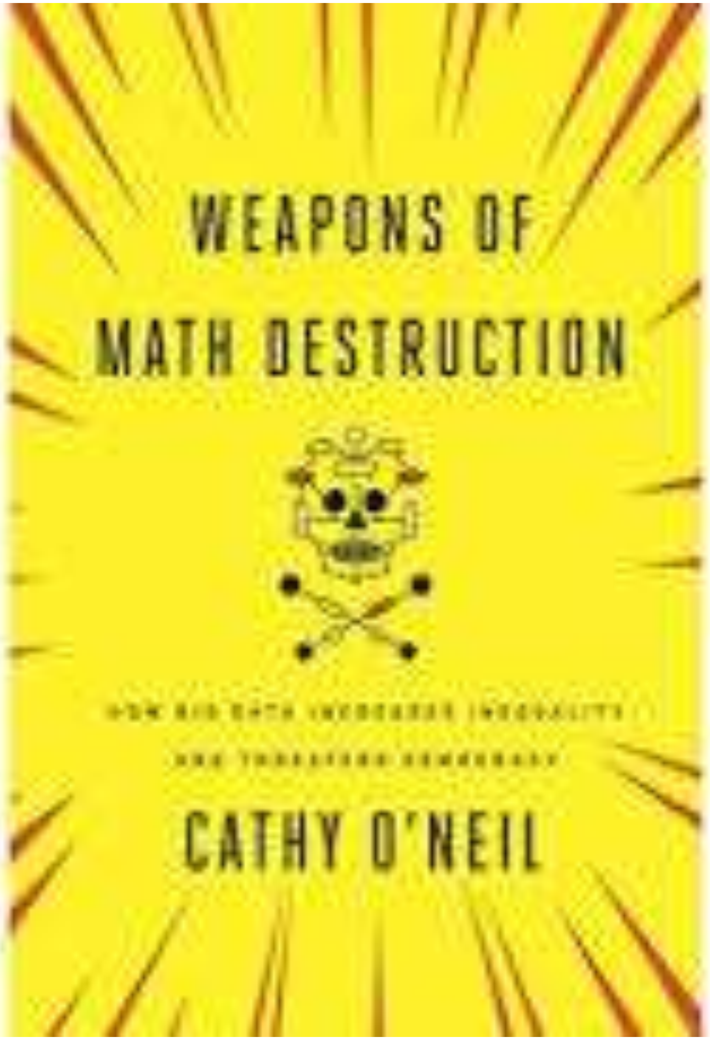
R. Brauneis and E. P. Goodman, "Algorithmic Transparency for the Smart City," Algorithmic Transpar. Smart City, vol. 20, pp. 103–176, 2018.

Weapons of Math Destruction

O’Neil, C. Weapons of math destruction : how big data increases inequality and threatens democracy. (Crown/Archetype, 2016).

Algorithmic audit in New York city

Dwyer J. Showing the Algorithms Behind New Y
City Services - The New York Times. New Yorl
Times Aug. 24, (2014).



Mathematical modelling does not make it to the headlines but …

Statistical modelling

Algorithms

Mathematical modelling

Blurring lines:

"what qualities are specific to rankings, or indicators, or models, or algorithms?"

E. Popp Berman and D. Hirschman, The Sociology of Quantification: Where Are We Now?, Contemp. Sociol., vol. in press, 2017.

“[in climate modelling] it looks very little like our idealized image of science, in which pure theory is tested with pure data

[impossible to] eliminate the model-dependency of data or the data-ladenness of models”

Paul N. Edwards, 1999, Global climate science, uncertainty and politics: Data-laden models, model-filtered data.

“[For] philosophers Frederick Suppe and Stephen Norton the blurry model/data relationship pervades all science”

Paul N. Edwards, 1999, Global climate science, uncertainty and politics:

Data-laden models, model-filtered data.

Two concerned papers:
Padilla et al. & Jakeman et al.

The heterogeneous nature of the modelling and simulation community prevents the emergence of consolidated paradigms ➔

➔verification and verification procedures are a rather trial and error business

This is a survey involving 283 responding modellers in J. J. Padilla, S. Y. Diallo, C. J. Lynch, and R. Gore, "Observations on the practice and profession of modeling and simulation: A survey approach," Simulation, vol. I14, 2017

Most users unaware of limitations, uncertainties, omissions and subjective choices in models ➔ over-reliance in the quality of model-based inference

Modellers oversimplify or overelaborate, obfuscating model use

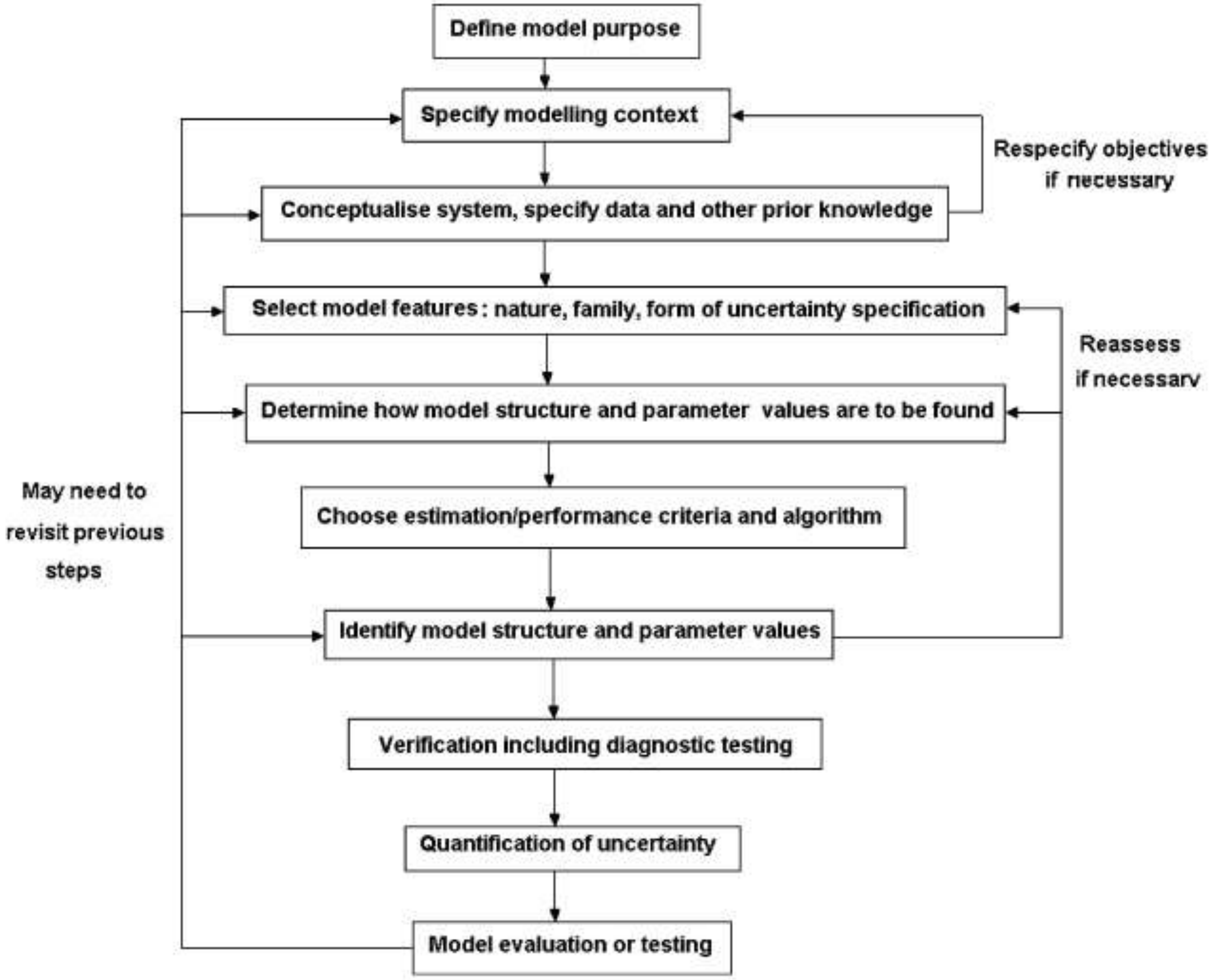
A large review of several existing checklists model quality: A. J. **Jakeman**, R. A. Letcher, and J. P. Norton, “Ten iterative steps in development and evaluation of environmental models,” Environ. Model. Softw., vol. 21, no. 5, pp. 602–614, 2006.

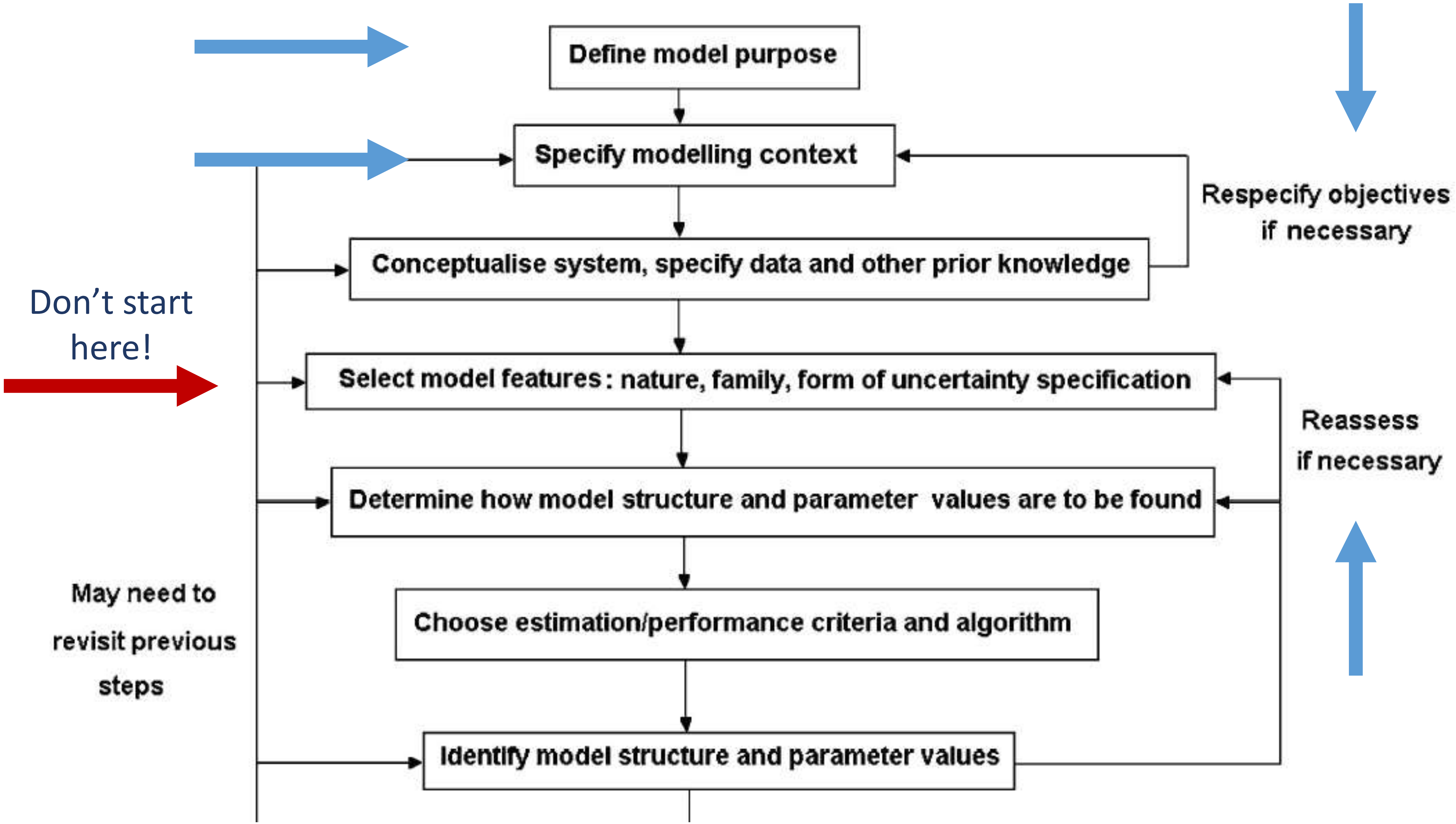
Padilla et al. call for **a more structured, generalized and standardized approach to verification**

Jakeman et al. call for a **10 points participatory checklist** including **NUSAP** and **J. R. Ravetz's process based approach**

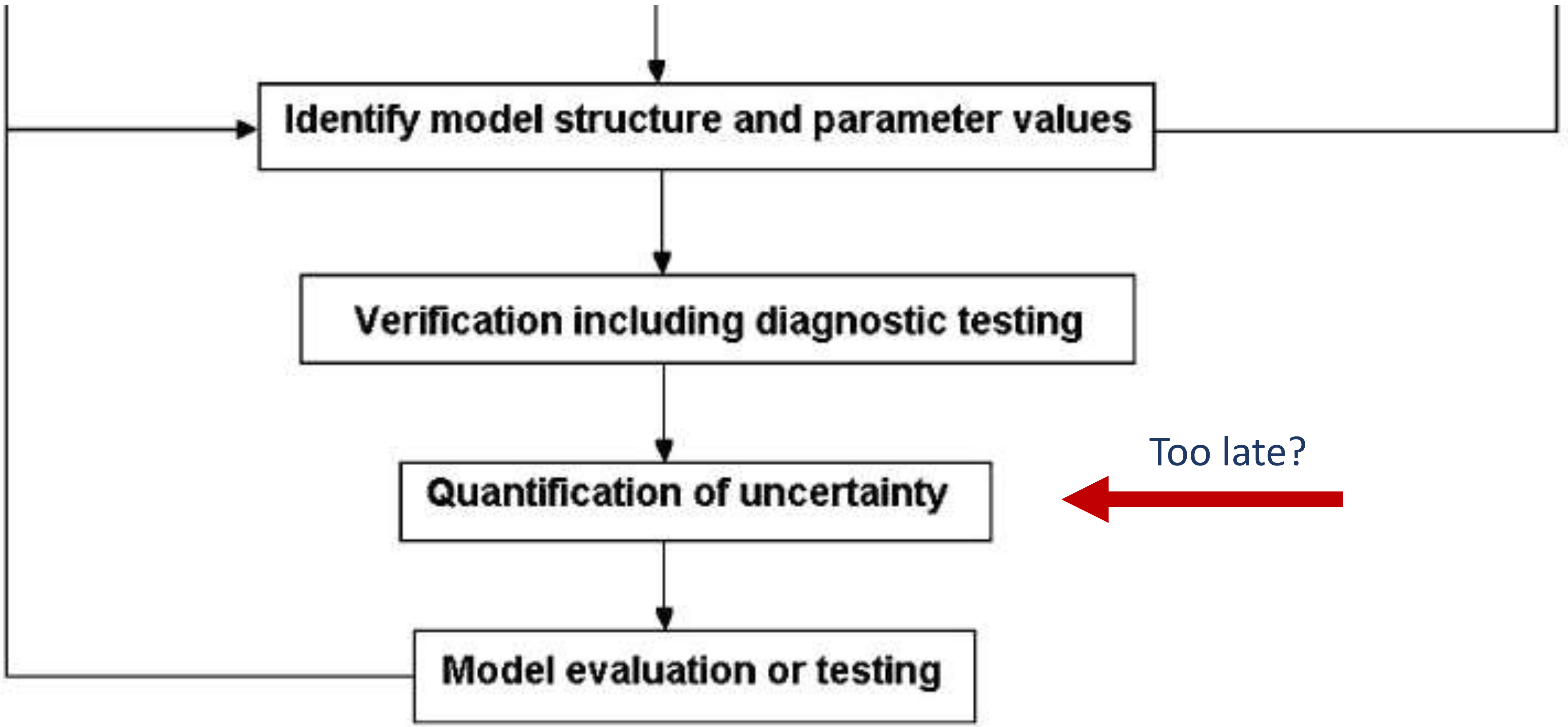
For NUSAP: Funtowicz, S.O., Ravetz, J.R., 1990. Uncertainty and Quality in Science and Policy. Kluwer, Dordrecht

J. R. Ravetz, "Integrated Environmental Assessment Forum, developing guidelines for 'good practice', Project ULYSSES.," 1997.http://www.jvds.nl/ulysses/eWP97-1.pdf


Define model purpose
Specify modelling context
Respecify objectives if necessary
Conceptualise system, specify data and other prior knowledge
Select model features : nature, family, form of uncertainty specification
Reassess if necessarv
Determine how model structure and parameter values are to be found
May need to revisit previous steps
Choose estimation/performance criteria and algorithm
Identify model structure and parameter values
Verification including diagnostic testing
Quantification of uncertainty
Model evaluation or testing



Define model purpose
Specify modelling context
Respecify objectives
if necessary
Conceptualise system, specify data and other prior knowledge
Don't start
here!
Select model features : nature, family, form of uncertainty specification
Reassess
if necessary
Determine how model structure and parameter values are to be found
May need to
revisit previous
steps
Choose estimation/performance criteria and algorithm
Identify model structure and parameter values


Identify model structure and parameter values
Verification including diagnostic testing
Quantification of uncertainty
Model evaluation or testing
Too late?

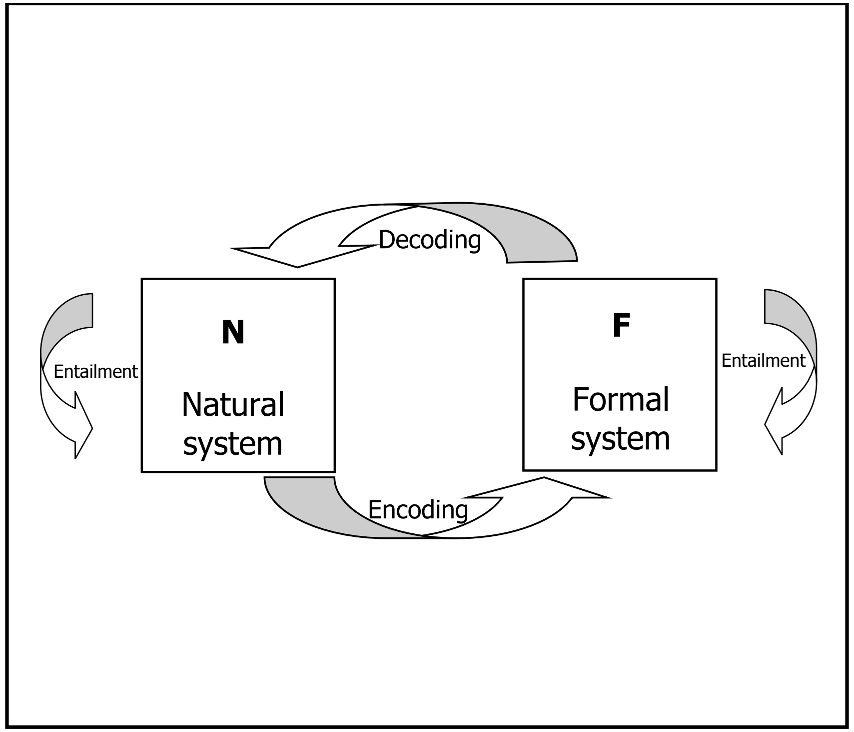
Not a discipline

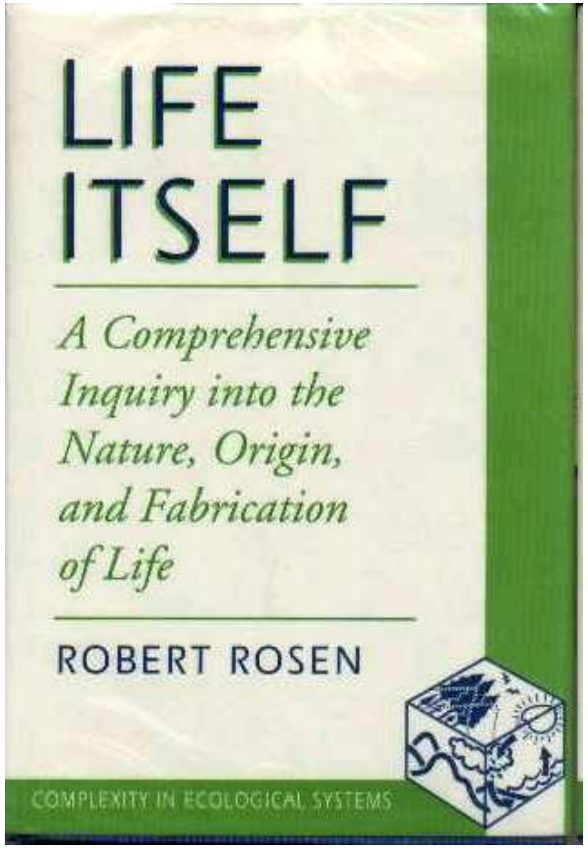
Unlike statistics, mathematical modelling is not a discipline, hence the lack of universally accepted quality standards, disciplinary fora and journals and recognized leaders

Making sensitivity analysis part of the syllabus of statistics?

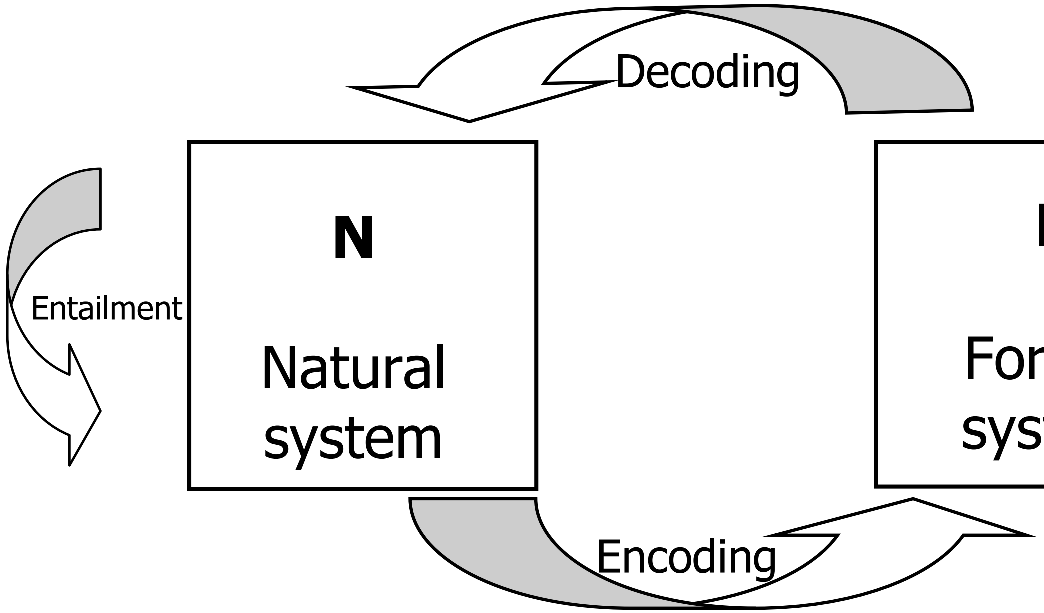
Saltelli, A., Does Modelling need a reformation? Ideas for a new grammar of modelling, available at https://arxiv.org/abs/1712.06457

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 ?

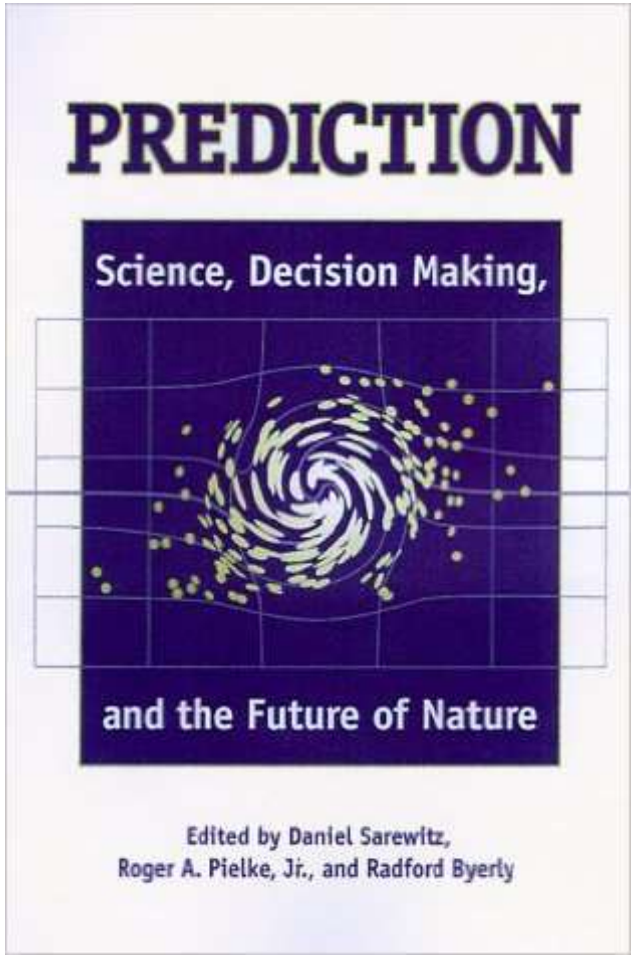
Robert Rosen

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

Naomi Oreskes

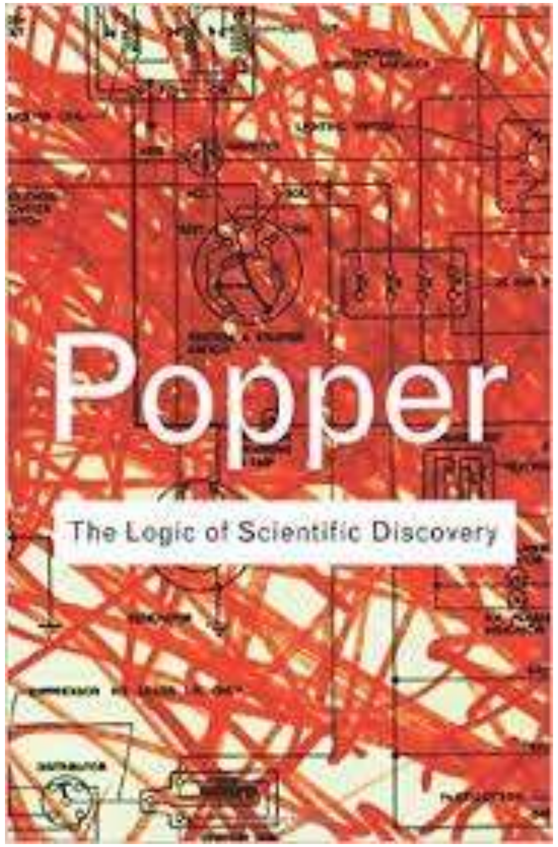
N. Oreskes, K. Shrader-Frechette, and K. Belitz, “Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences,” Science, 263, no. 5147, 1994.

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)

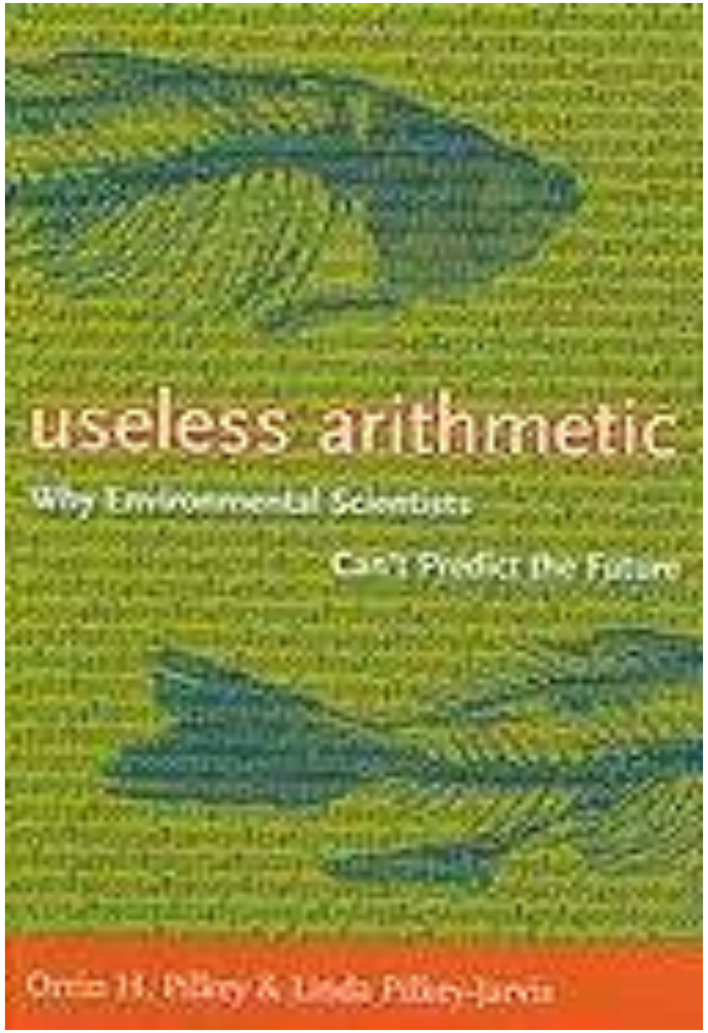


“In many cases, these temporal predictions **are treated with the same respect** that the hypothetic-deductive model of science accords to logical predictions. But this respect is largely misplaced”

“[⋯] 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”

Egregious modelling failure from Pilkey and Pilkey-Jarvis

(from AIDS to coastal erosion to nuclear waste disposal …)



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

For John Kay modelling may need as input information which we don't have (The case of WEBTAG; knowing car passengers number decades into futures)

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

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

Erik Reinert: scholastic tendencies in the mathematization of economics

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

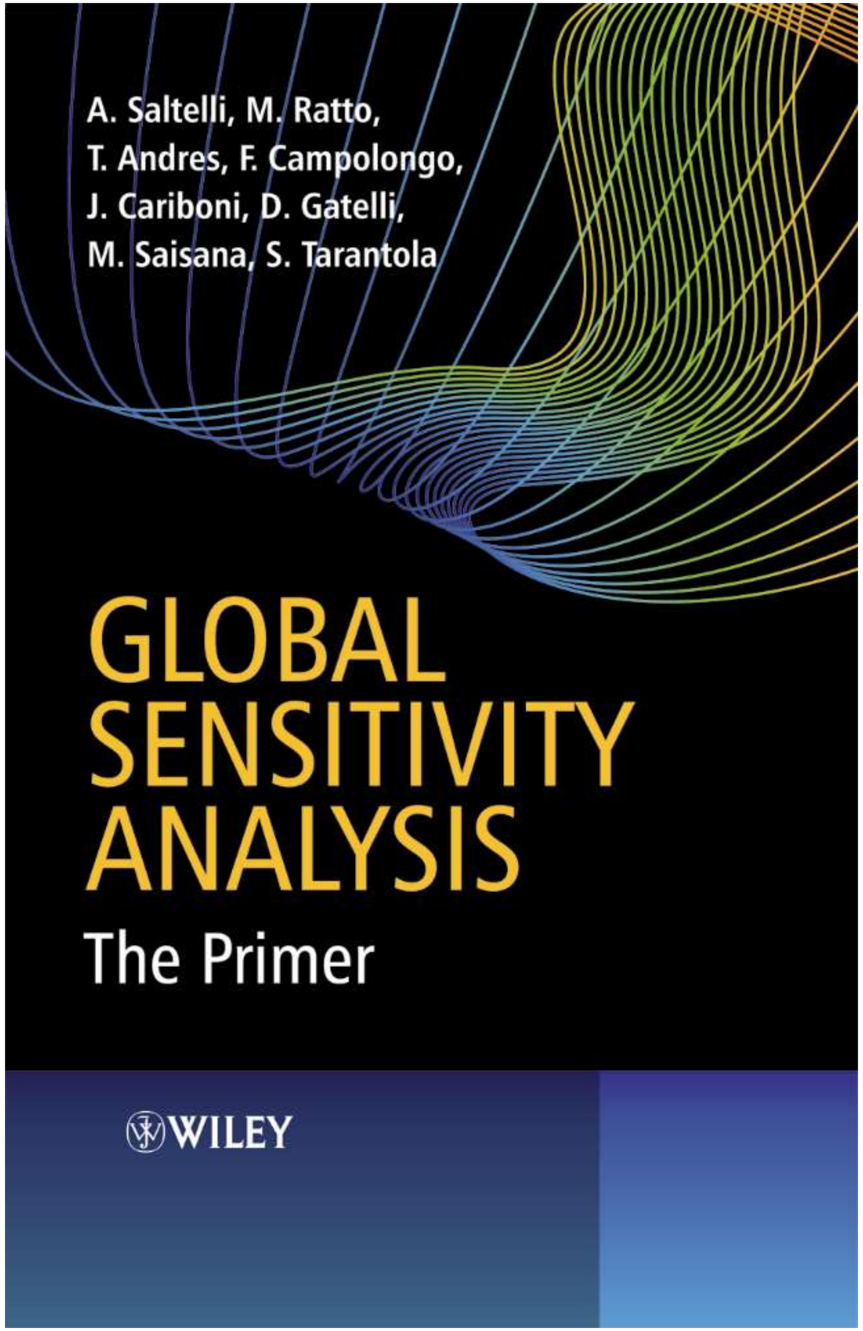
E. S. Reinert, "Full circle: economics from scholasticism through innovation and back into mathematical scholasticism," J. Econ. Stud., vol. 27, no. 4/5, pp. 364–376, Aug. 2000.

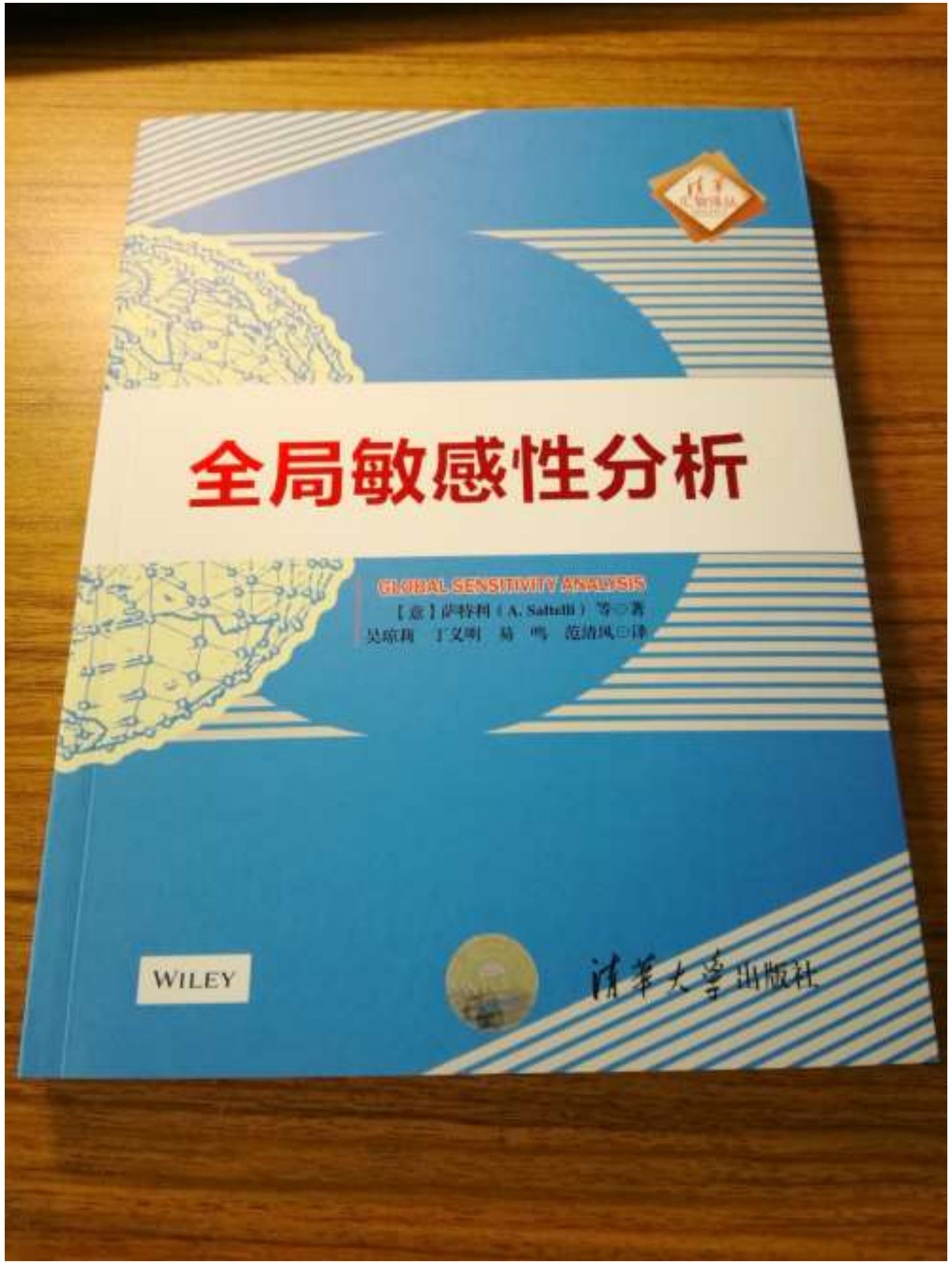
Uncertainty and sensitivity analysis

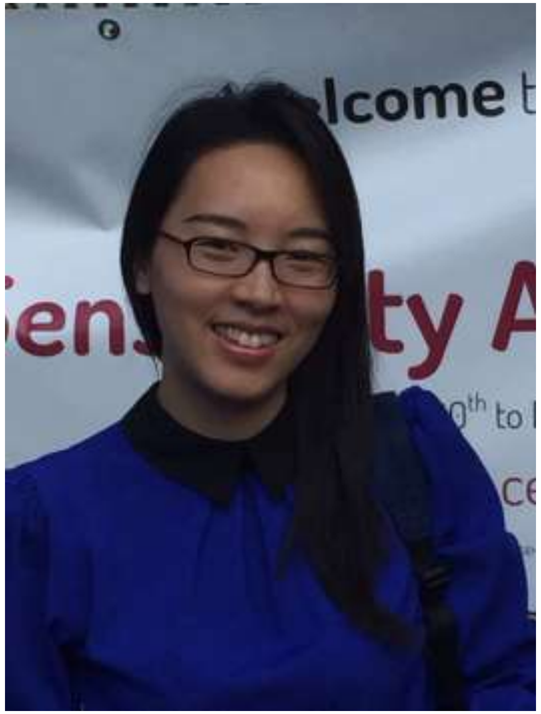
Definitions

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

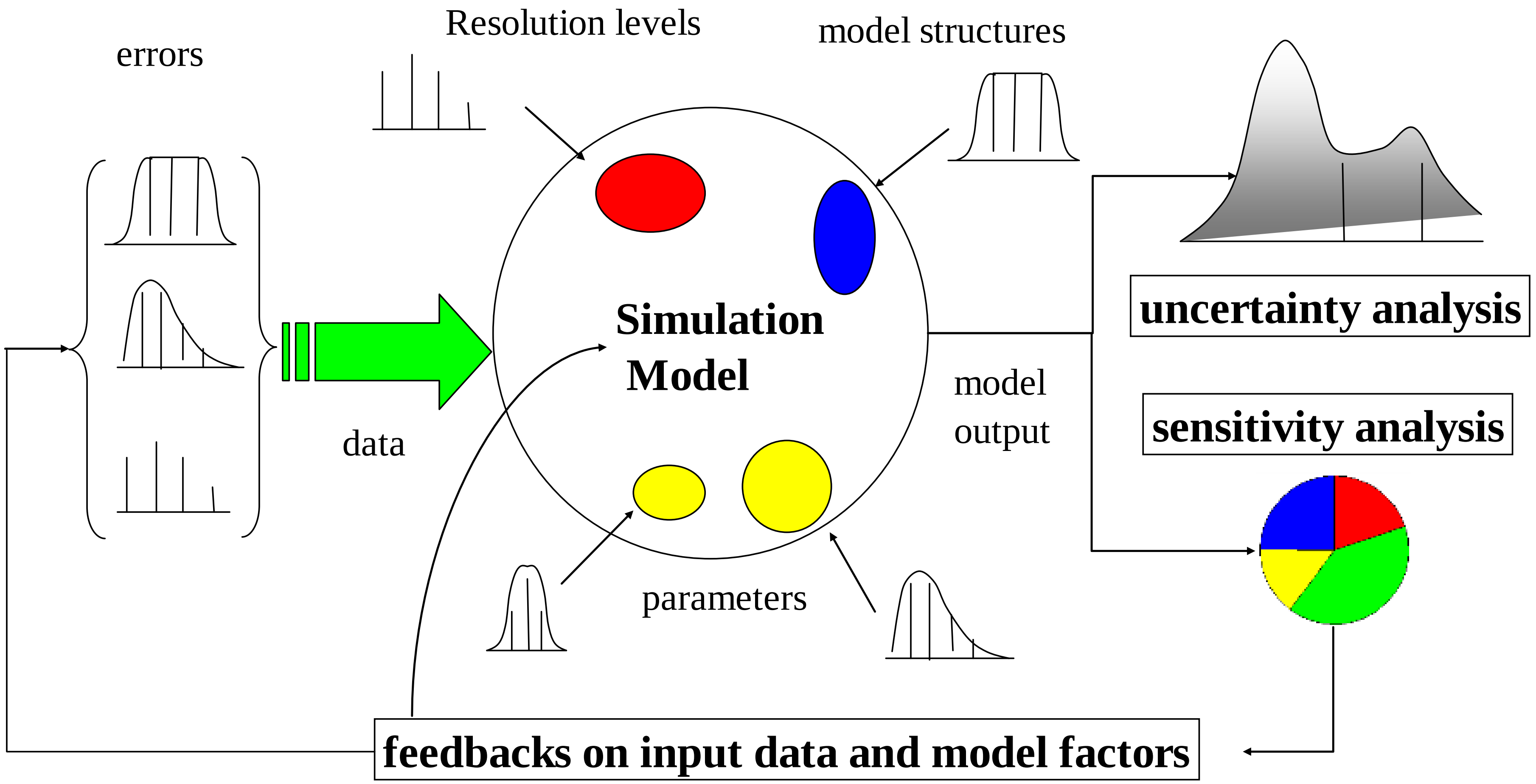

A. Saltelli, M. Ratto,
T. Andres, F. Campolongo,
J. Cariboni, D. Gatelli,
M. Saisana, S. Tarantola
GLOBAL
SENSITIVITY
ANALYSIS
The Primer
WILEY


全局敏感性分析
GLOBAL SENSITIVITY ANALYSIS
WILEY
清华大学出版社



Wu
Qiongli

An engineer's vision of UA, SA

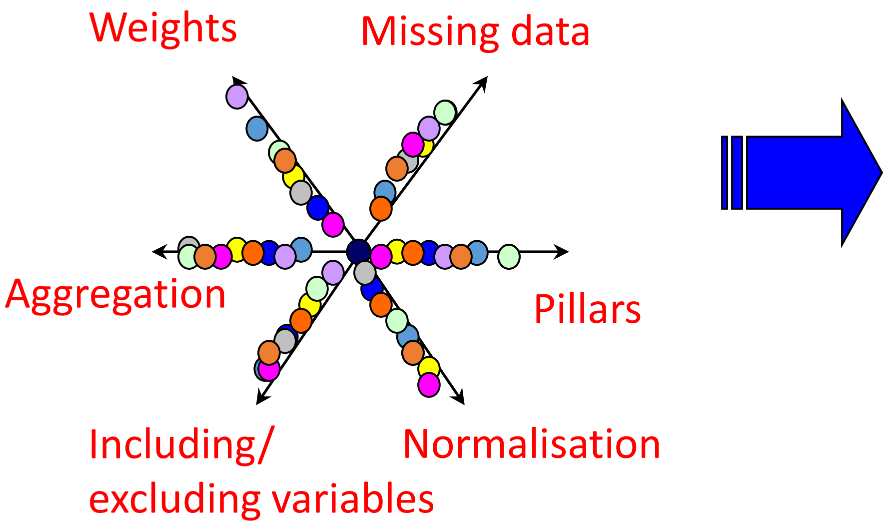


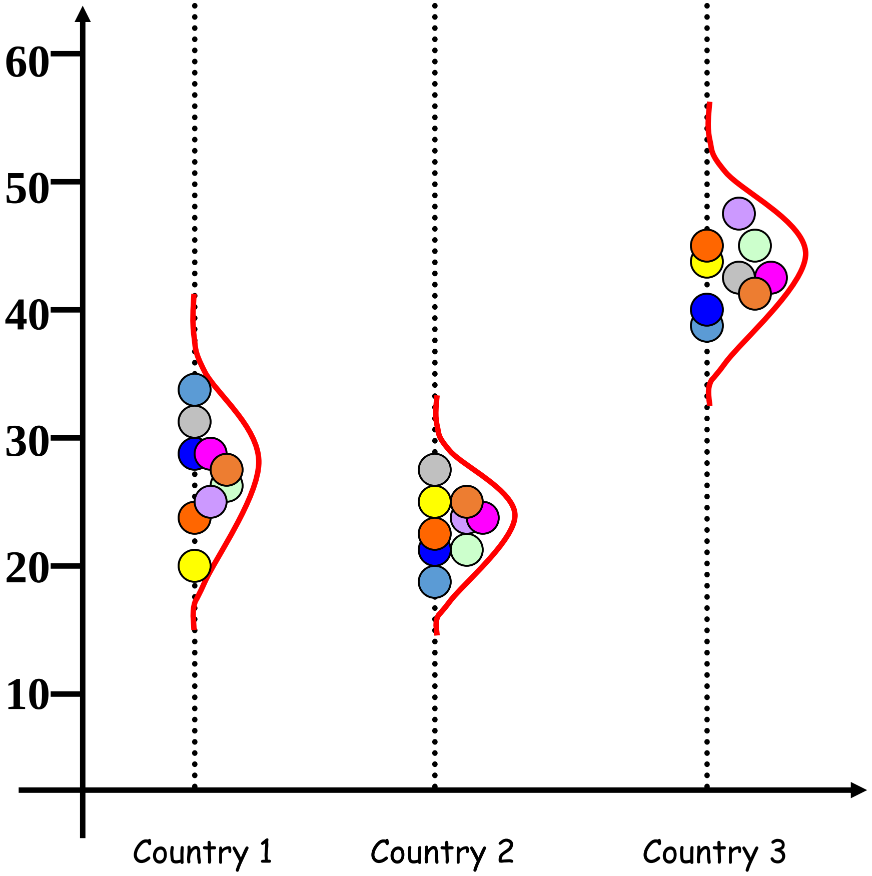
One can sample more than just factors

One can sample modelling assumptions, alternative data sets, resolution levels, scenarios …

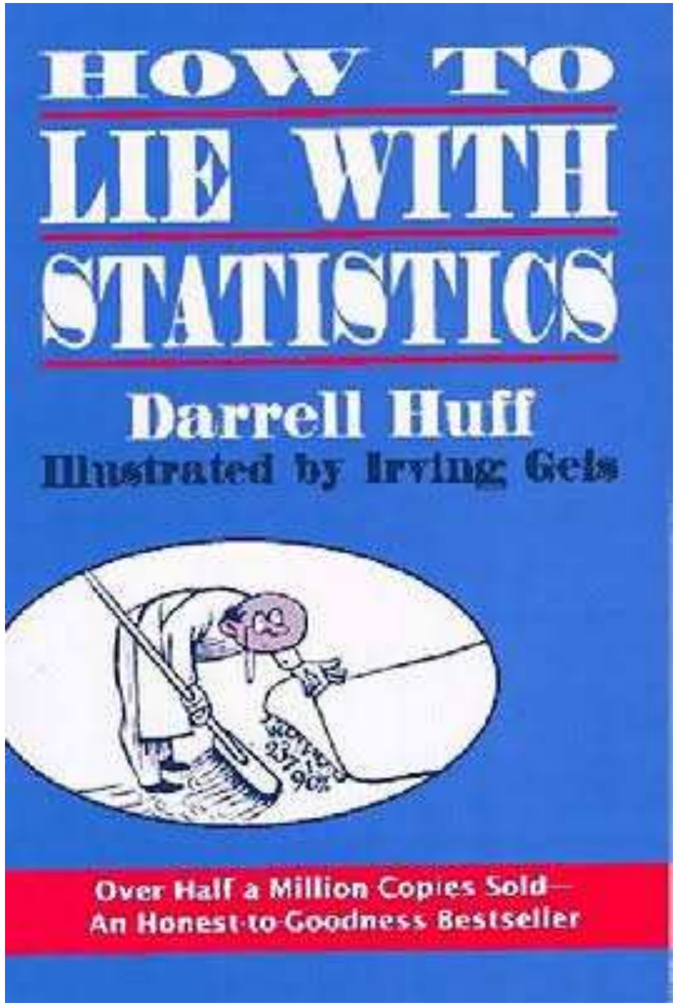
Assumption	Alternatives
Number of indicators	▪ all six indicators included or one-at-time excluded (6 options)
Weighting method	▪ original set of weights, ▪ factor analysis, ▪ equal weighting, ▪ data envelopment analysis
Aggregation rule	▪ additive, ▪ multiplicative, ▪ Borda multi-criterion

Space of alternatives



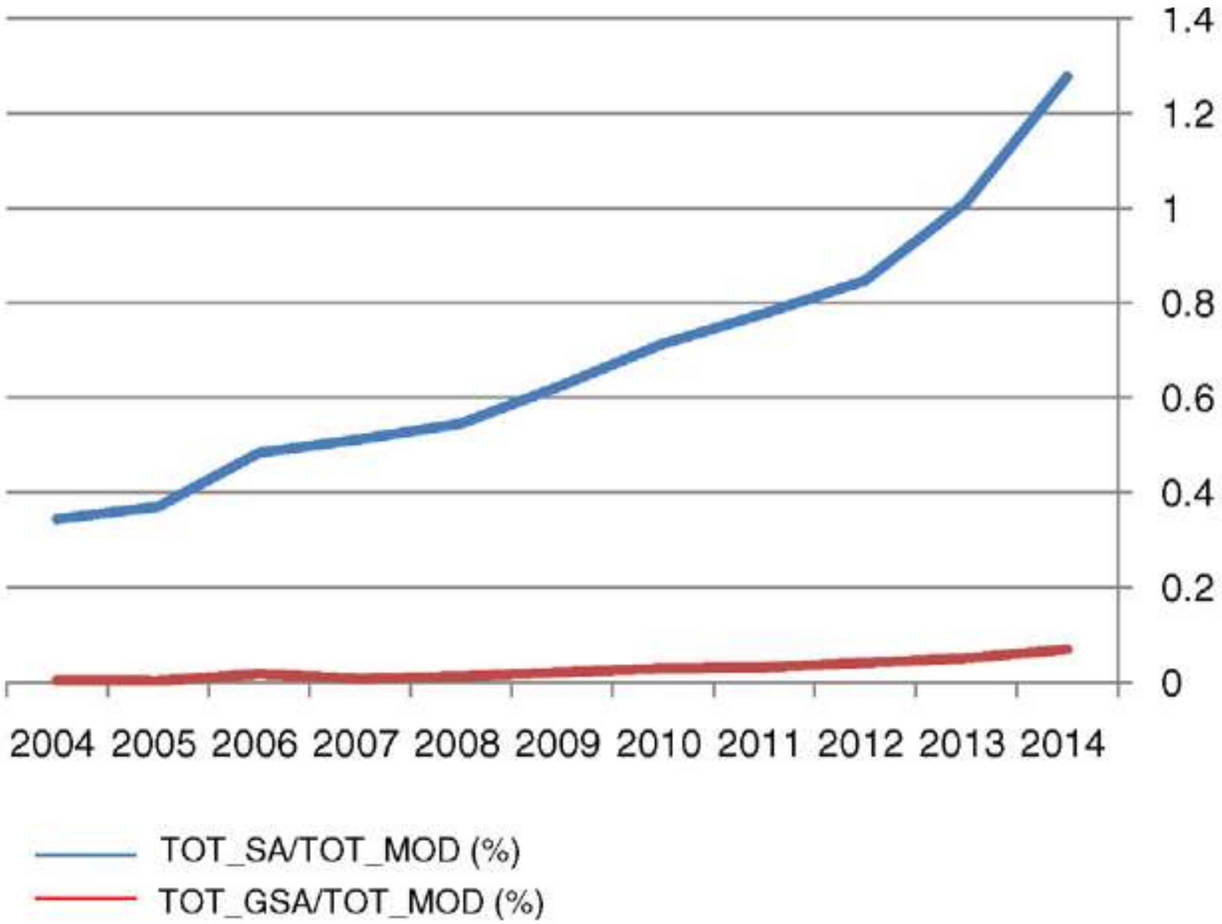


Can one lie with sensitivity analysis as one can lie with statistics?



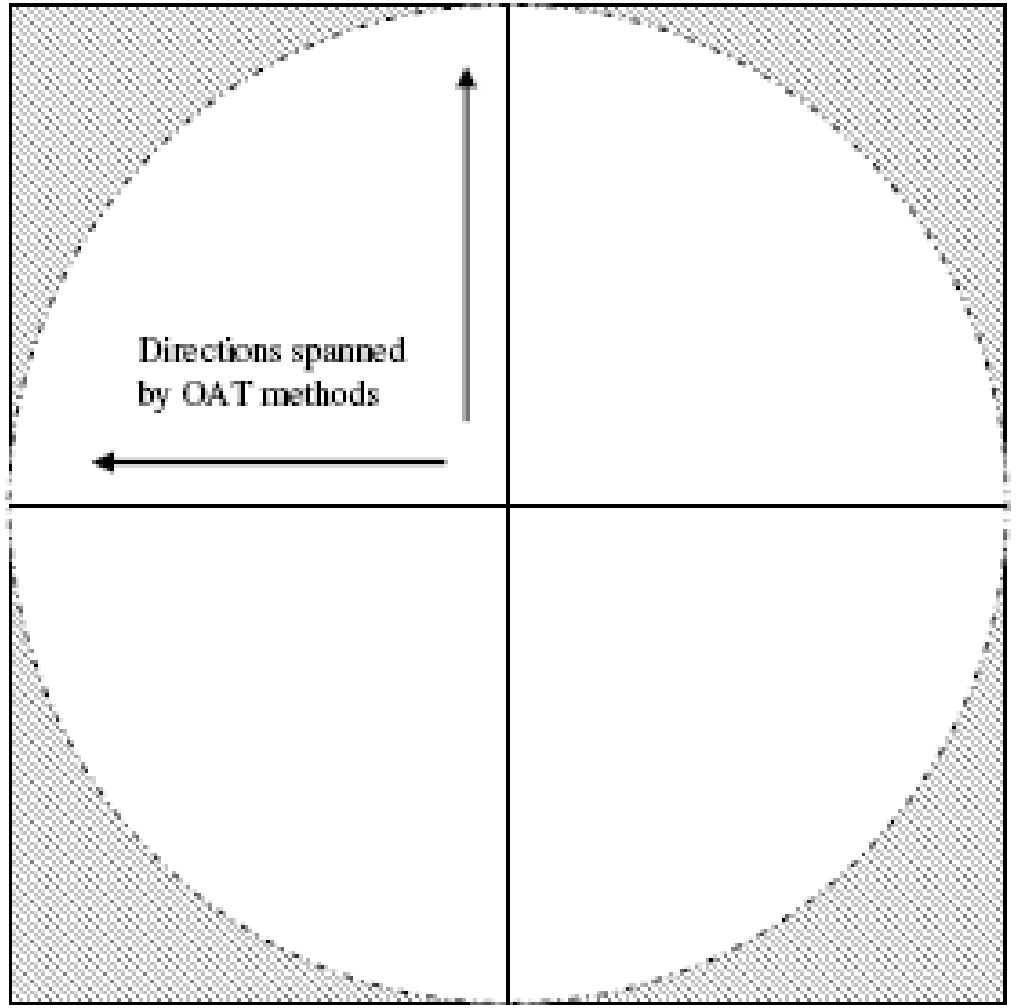
Saltelli, A., Annoni P., 2010, How to avoid a perfunctory sensitivity analysis, Environmental Modeling and Software, 25, 1508–1517.

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



Ferretti, F., Saltelli A., Tarantola, S., 2016, Trends in Sensitivity Analysis practice in the last decade, Science of the Total Environment, http://dx.doi.org/10.1016/j.scitotenv.2016.02.133

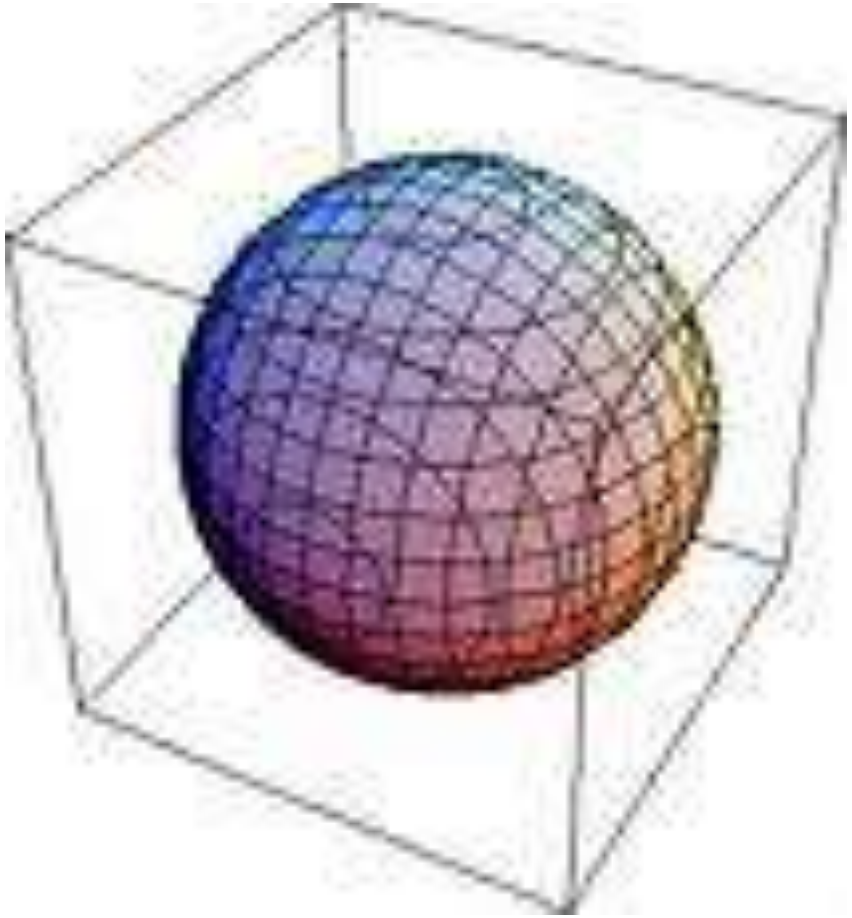
OAT in 2 dimensions



Area circle / area square =?

~ 3/4

OAT in 3 dimensions

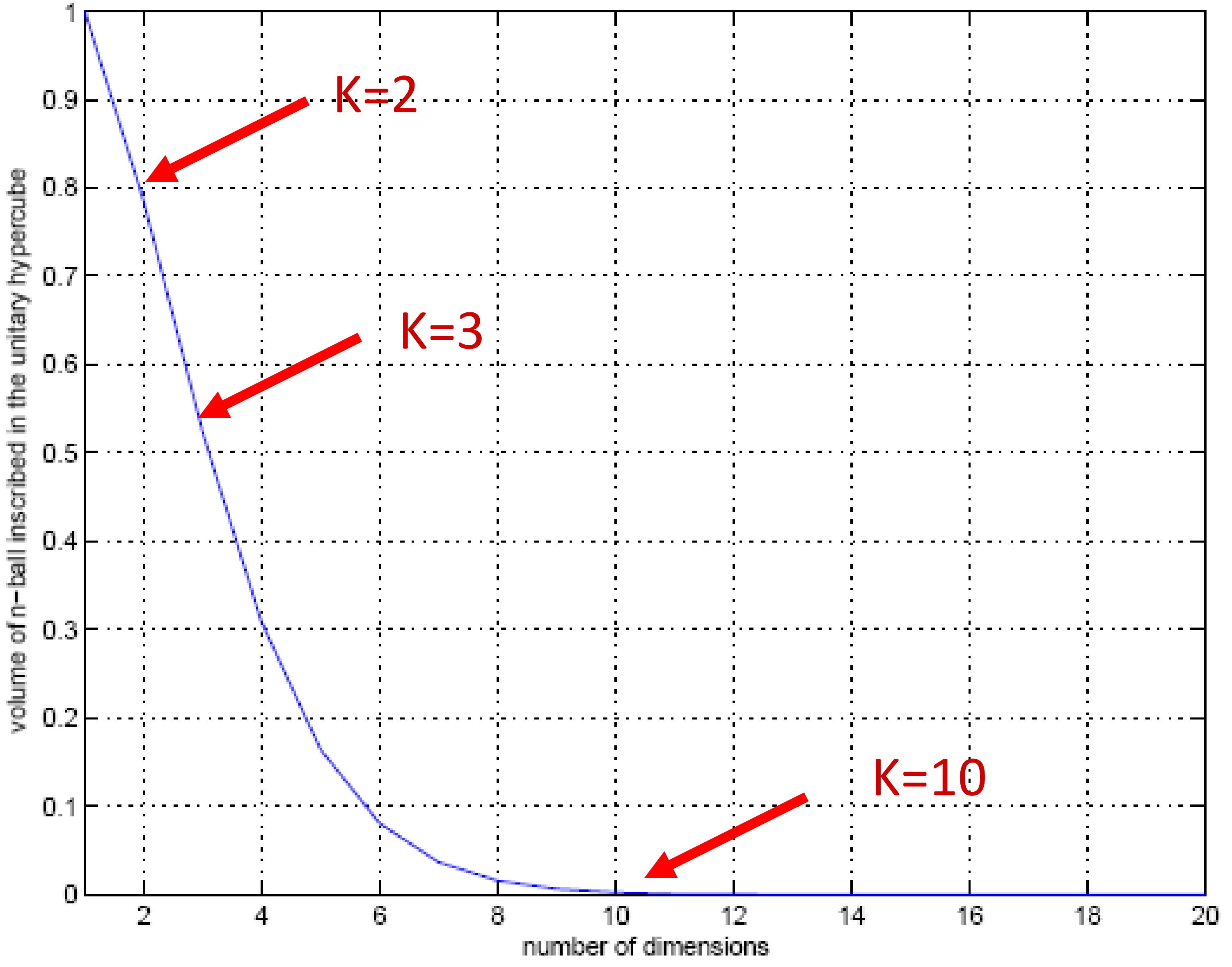


Volume sphere / volume cube =?

~ 1/2

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

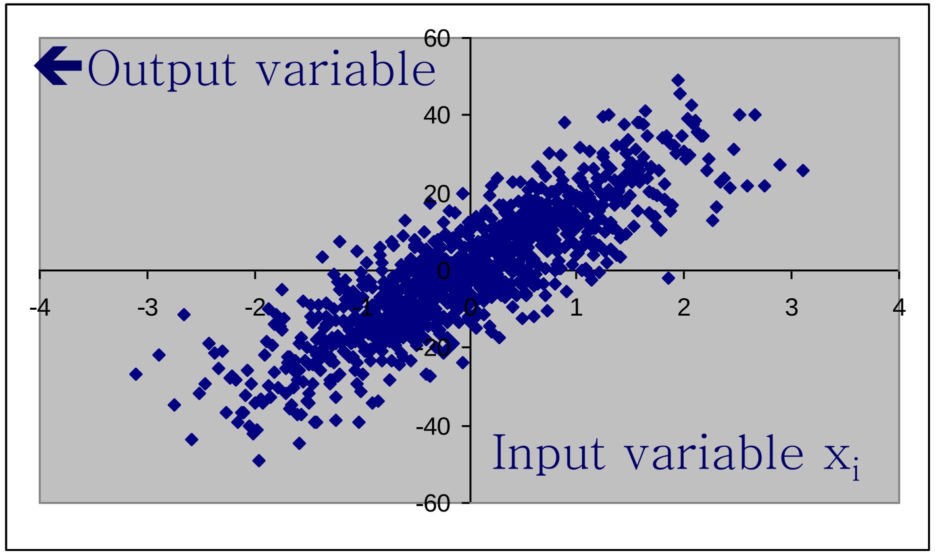
OAT in k dimensions

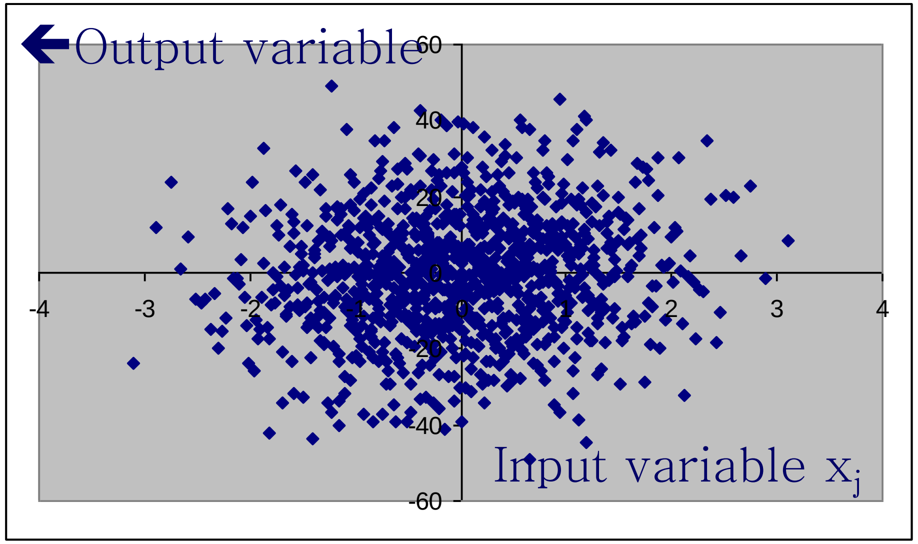


Once a sensitivity analysis is done via OAT there is no guarantee that either uncertainty analysis (UA) or sensitivity analysis (SA) will be any good:

➔ UA will be non conservative

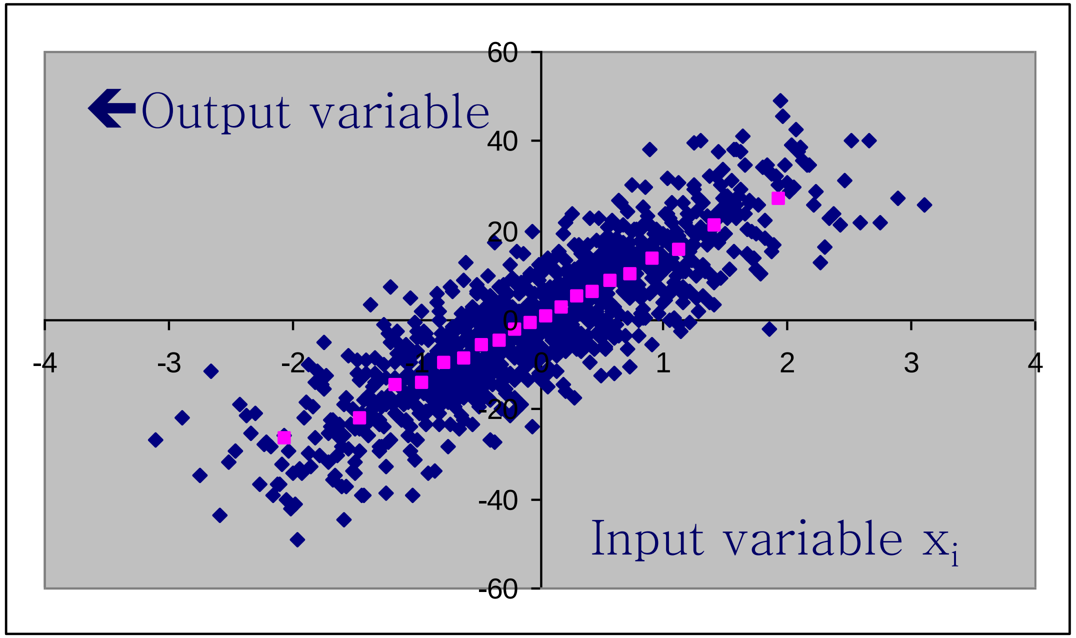
➔ SA may miss important factors

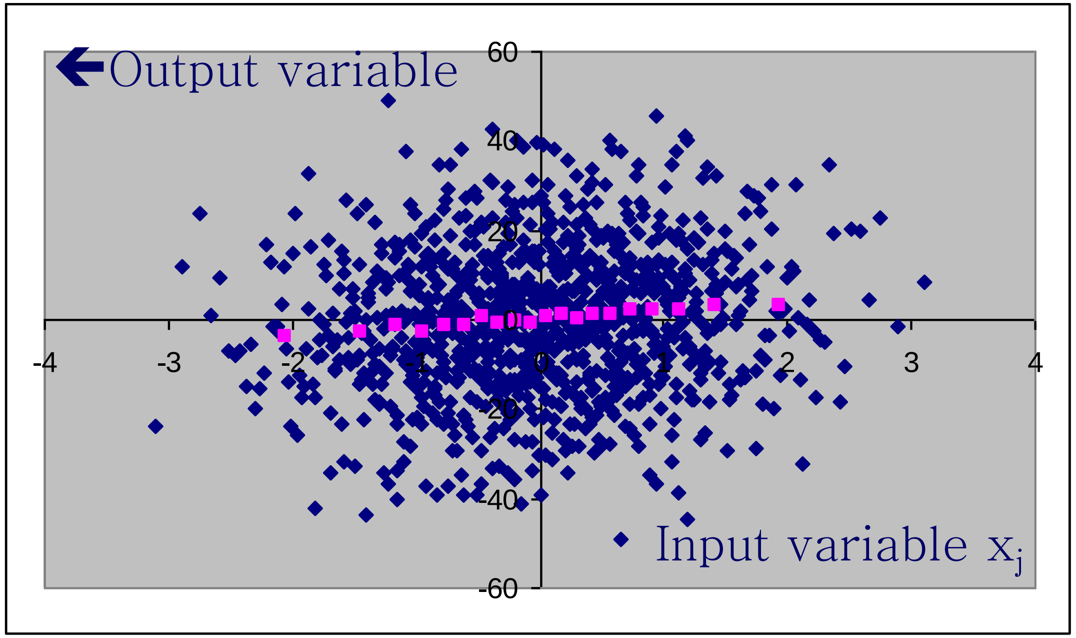




Which factor is more important?

Why?

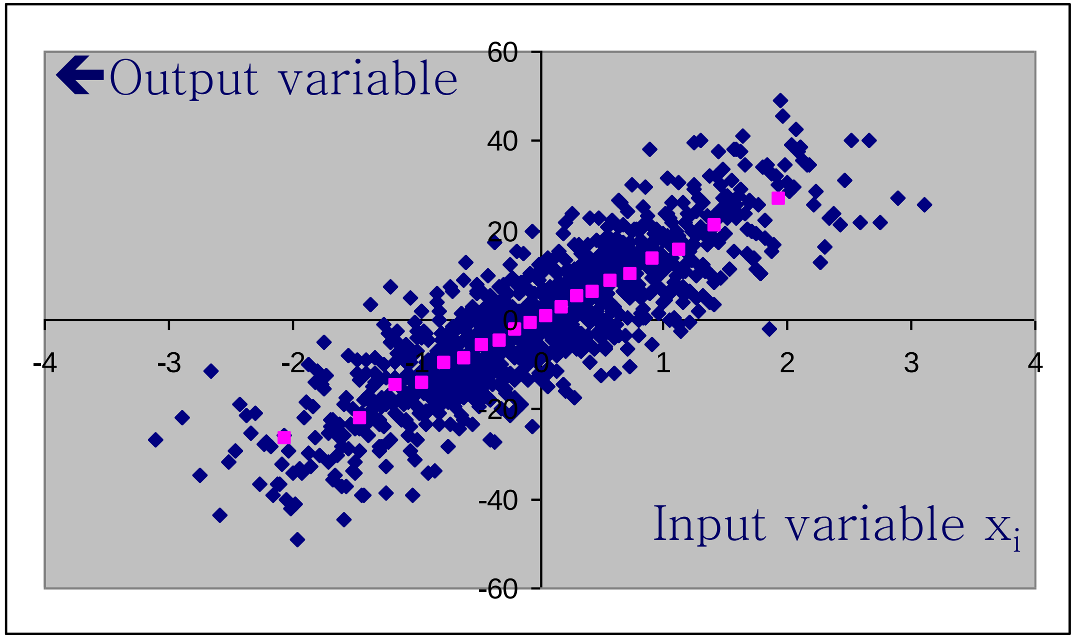




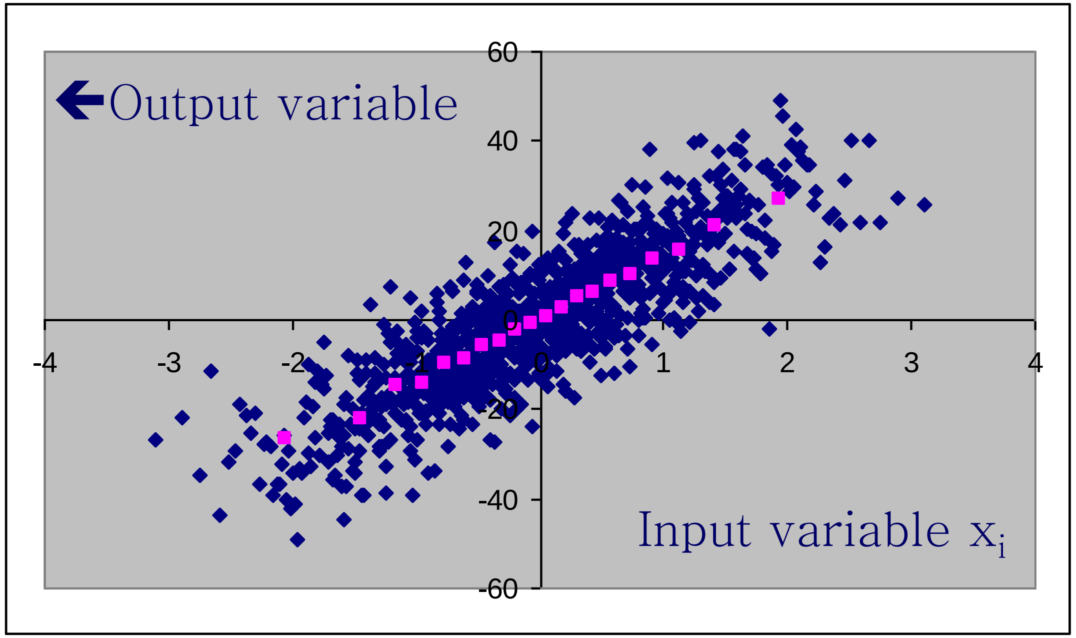
~1,000 blue points

Divide them in 20 bins of ~ 50 points

Compute the bin's average (pink dots)

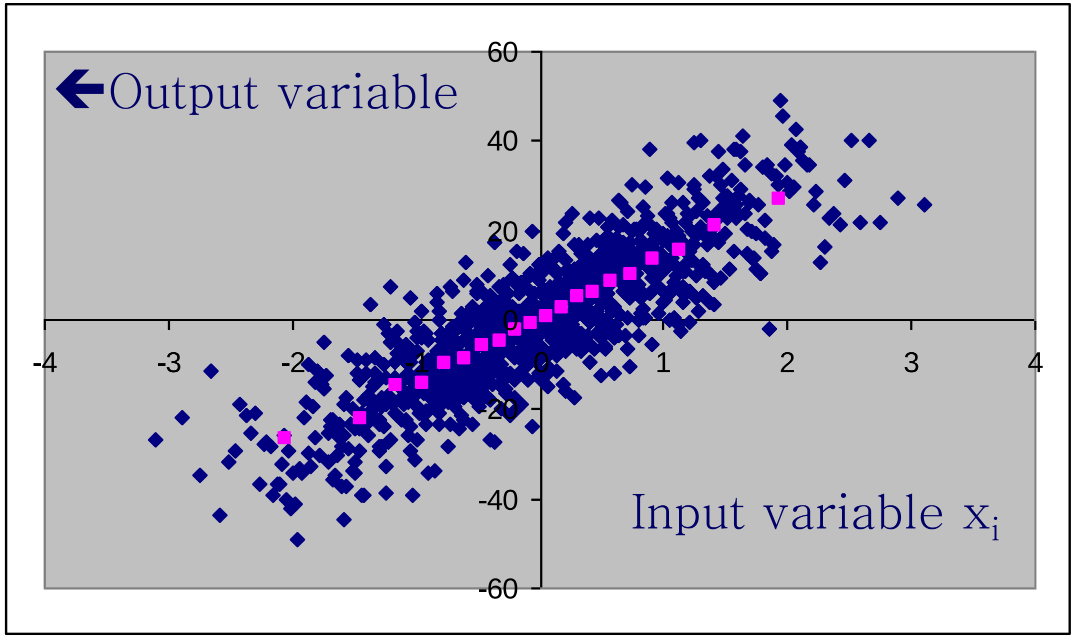


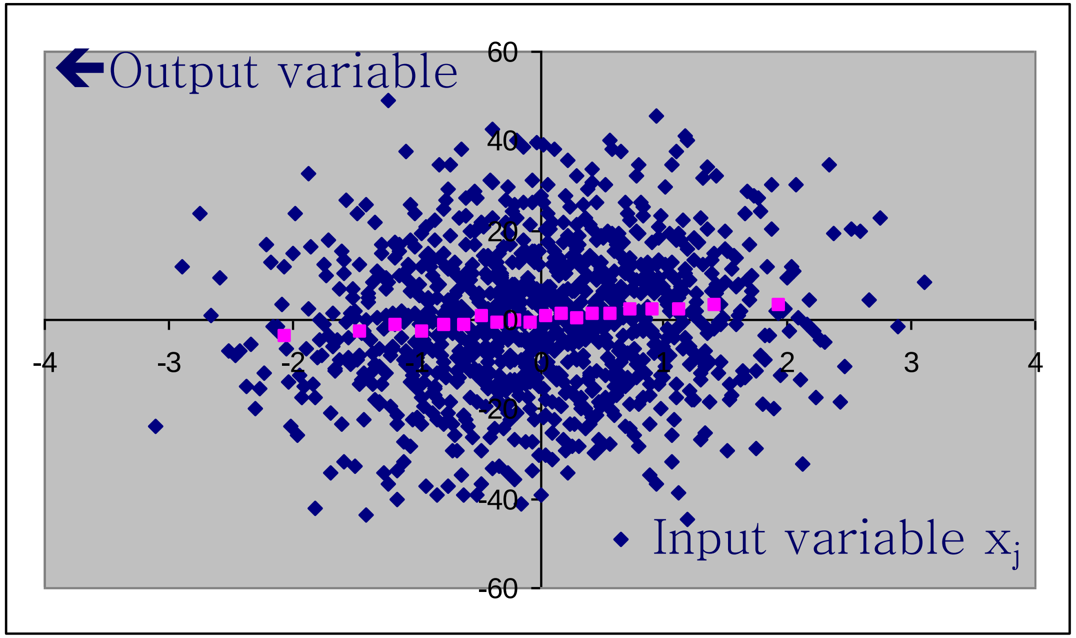
Each pink point is ~ $E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)$



Take the variance of the pink points and you have a sensitivity measure

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





Which factor has the highest

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

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

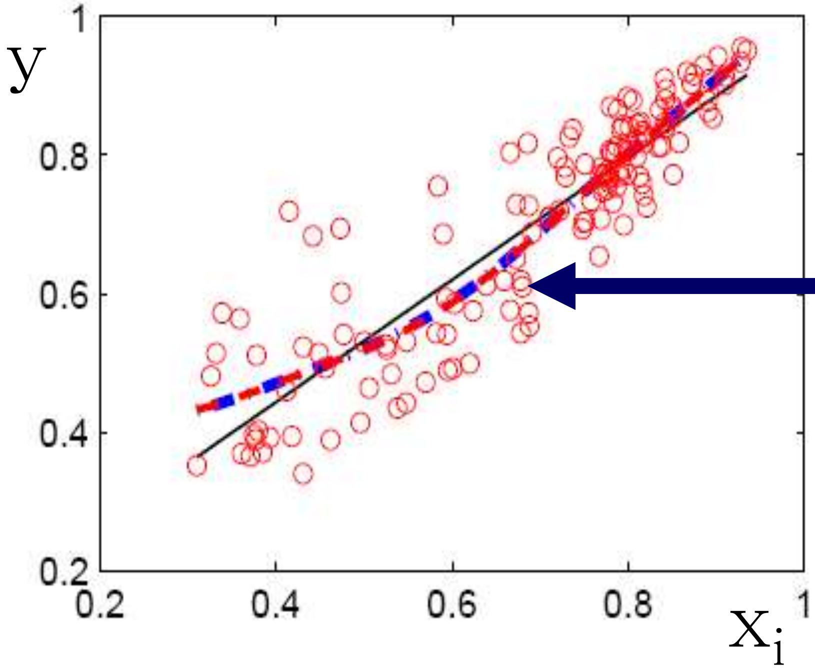
Pearson's correlation ratio

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



Smoothed curve:

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

First order sensitivity index:

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

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

First order effect, or top marginal variance=

= the expected reduction in variance that would be achieved if factor Xi could be fixed.

Why?

Because:

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

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

Because:

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

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

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

… then this …

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

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

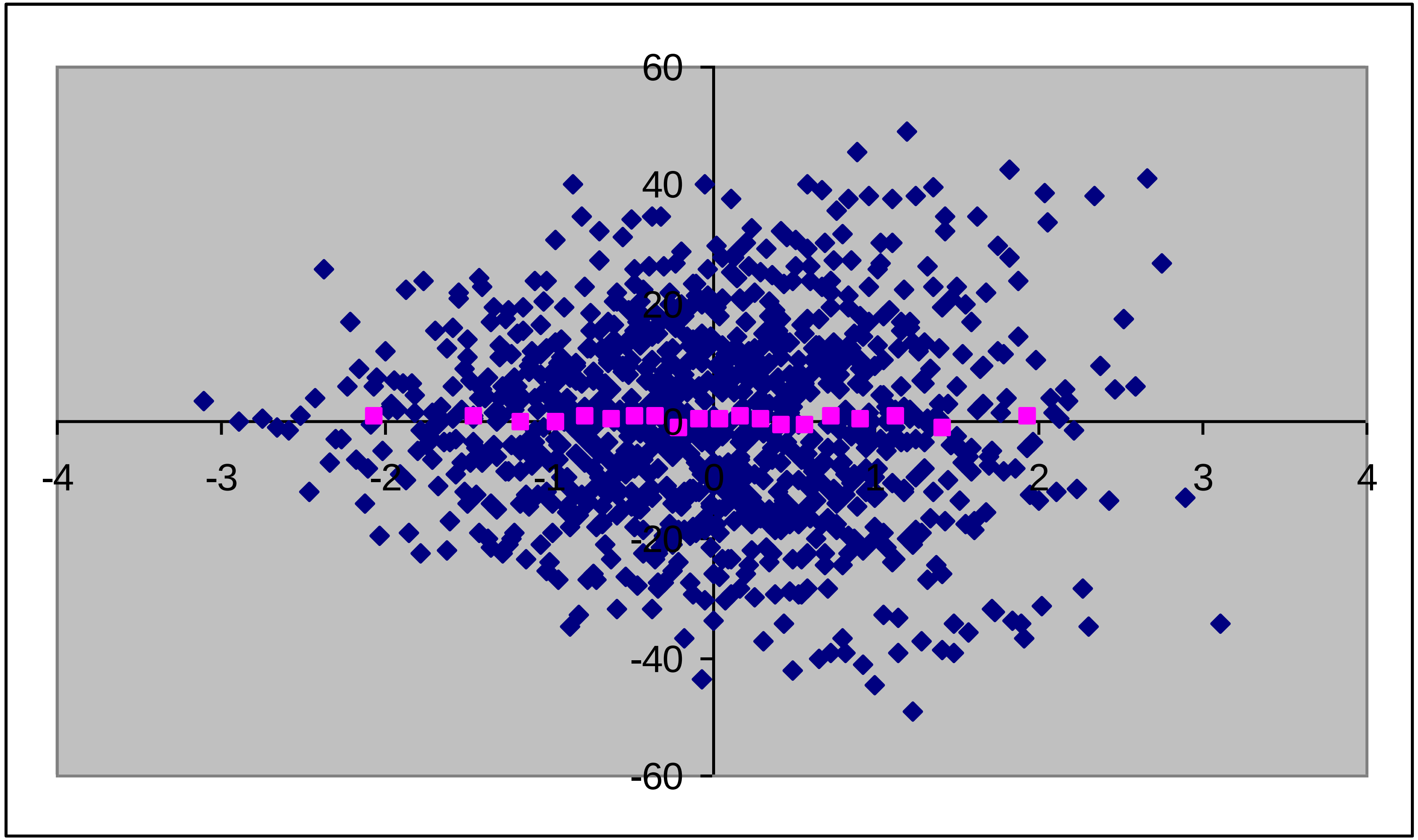
For additive models one can decompose the total variance as a sum of first order effects

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

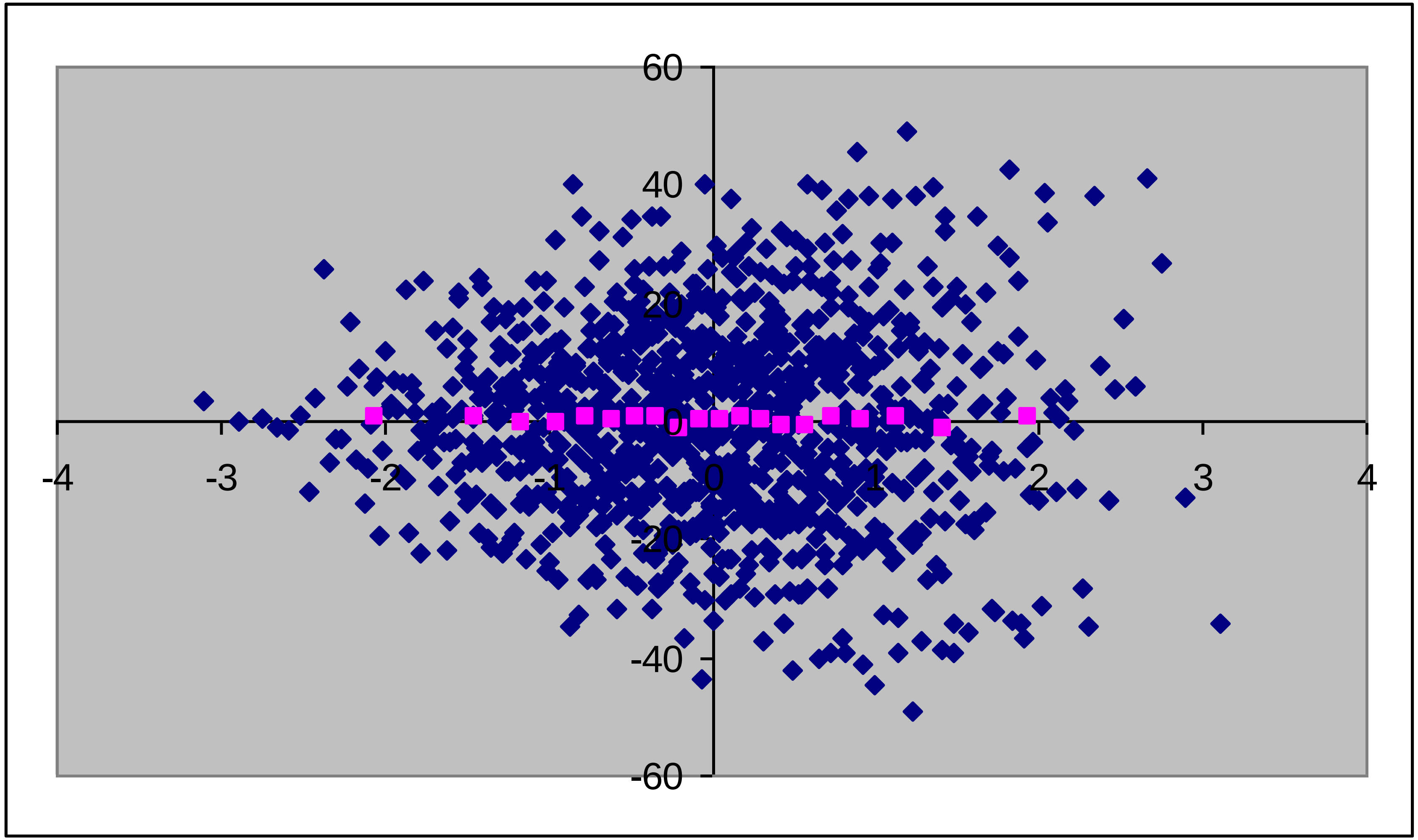
… which is also how additive models are defined

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

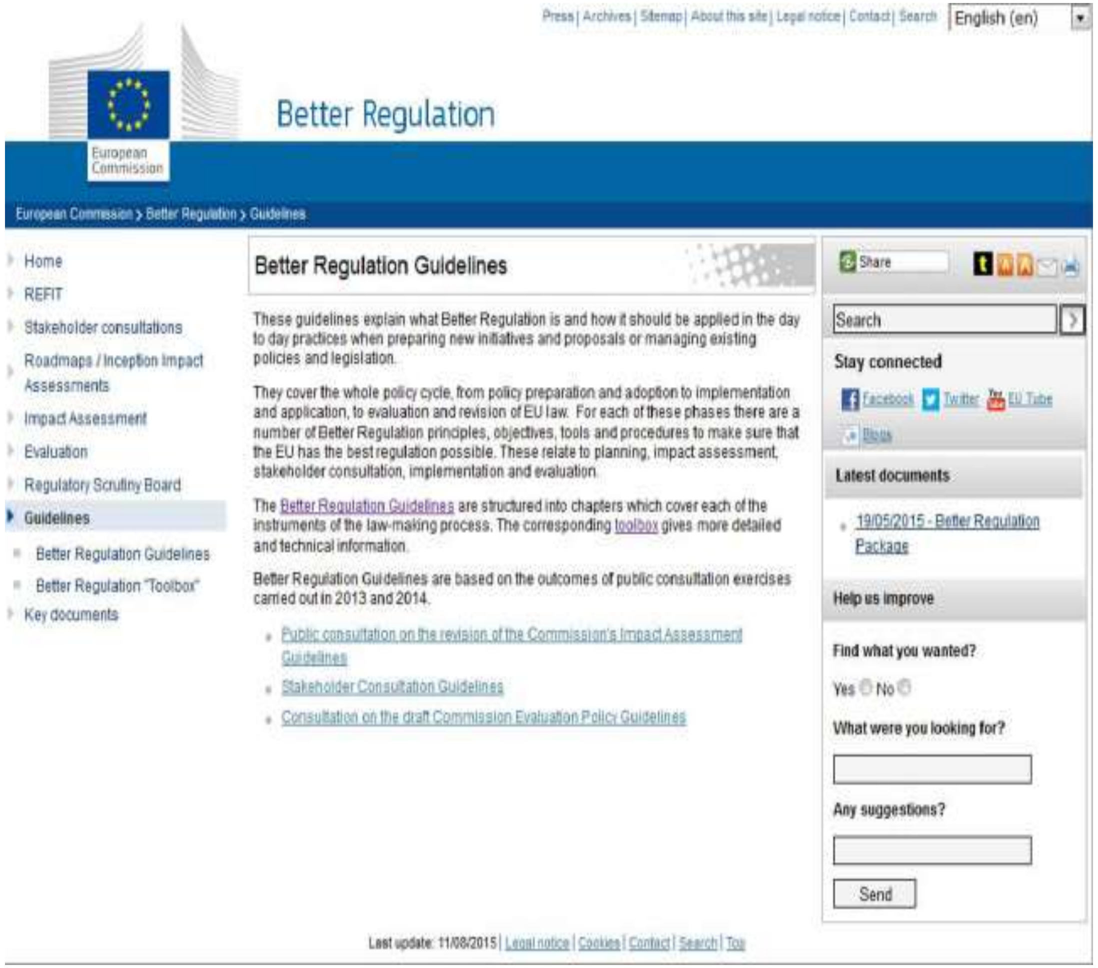
Variance decomposition (ANOVA)

$$V(Y) =$$

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

→ Lesson Stefano Tarantola

EC impact assessment guidelines: sensitivity analysis & auditing



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

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]

We're going to need a bigger rug !

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]

Badly kept secret:

There is always one more bug!

(Lubarsky's Law of Cybernetic Entomology)

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

More than a technical uncertainty and sensitivity analysis?

A new grammar for mathematical modelling?

1. Uncertainty and sensitivity analysis (never execute the model once)
2. Sensitivity auditing and quantitative storytelling (investigate frames and motivations)

Saltelli, A., Guimarães Pereira, Â., Van der Sluijs, J.P. and Funtowicz, S., 2013, 'What do I make of your latinorum? Sensitivity auditing of mathematical modelling', Int. J. Foresight and Innovation Policy, (9), 2/3/4, 213–234.

Saltelli, A., Does Modelling need a reformation? Ideas for a new grammar of modelling, available at https://arxiv.org/abs/1712.06457

3. Replace ‘model to predict and control the future’ with ‘model to help mapping ignorance about the future’ …

… in the process exploiting and making explicit the metaphors embedded in the model

J. R. Ravetz, “Models as metaphors,” in Public participation in sustainability science : a handbook, and W. A. B. Kasemir, J. Jäger, C. Jaeger, Gardner Matthew T., Clark William C., Ed. Cambridge University Press, 2003, available at http://www.nusap.net/download.php?op=getit&lid=11

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; frames; ➔ quantitative storytelling

7. Perform a proper global sensitivity analysis.

An example:

Sensitivity analysis: the case of the Stern review

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

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).

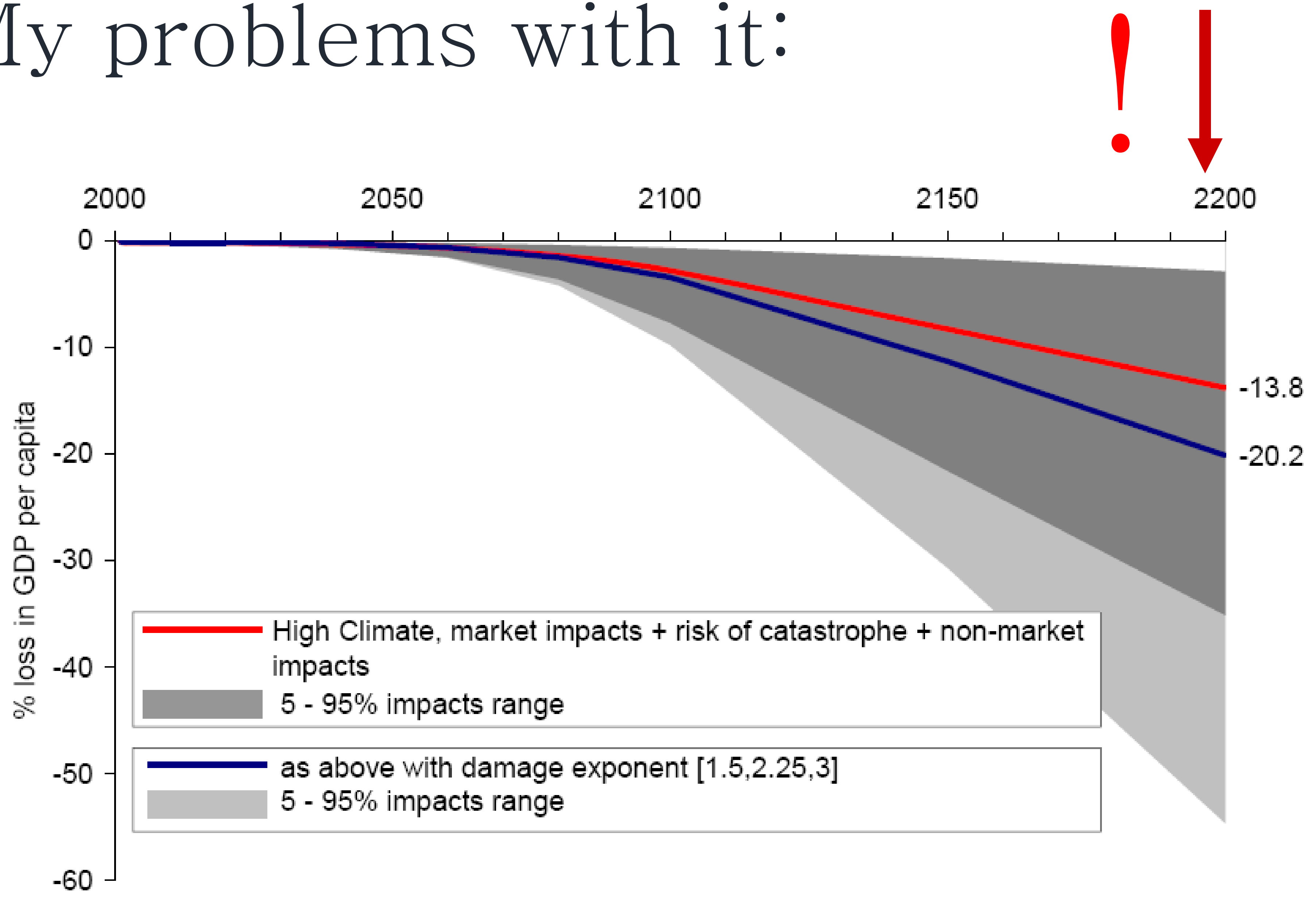
The Stern – Nordhaus exchange on *SCIENCE*

1) Nordhaus falsifies Stern based on 'wrong' range of discount rate

2) Stern's complements its review with a postscript: a sensitivity analysis of the cost benefit analysis

3) Stern thus says: My analysis shows robustness'

My problems with it:

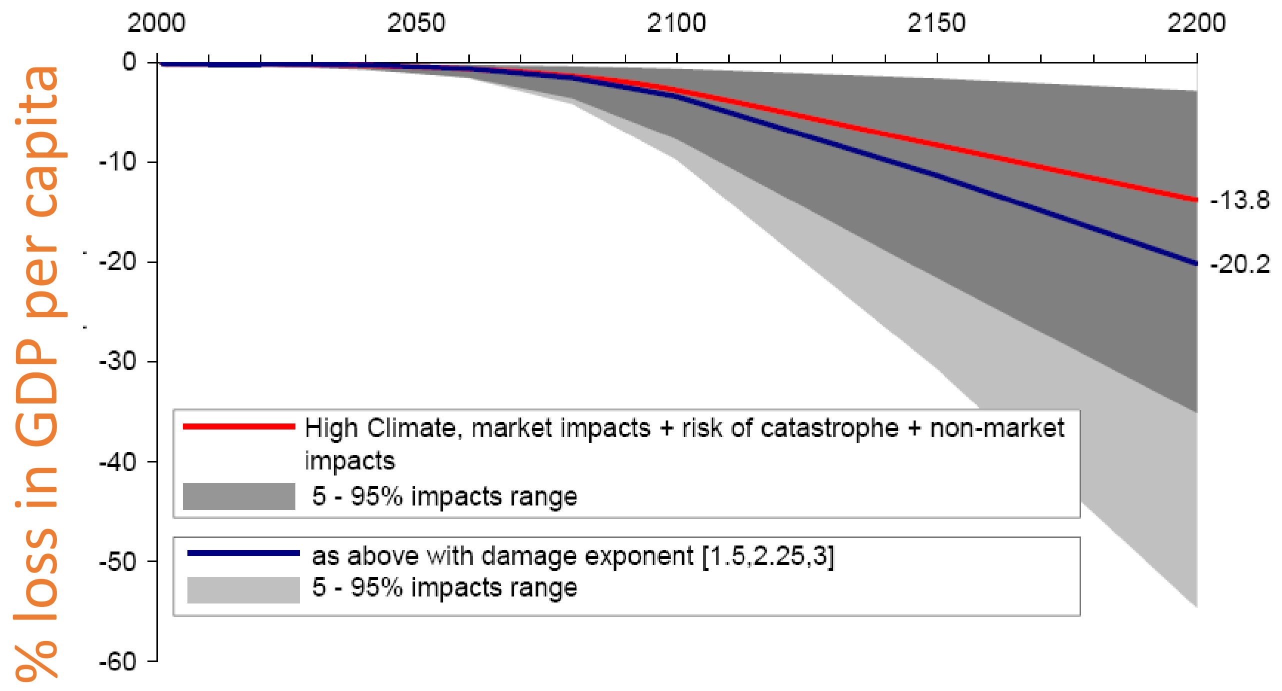


… but foremost Stern says:

changing assumptions → important effect

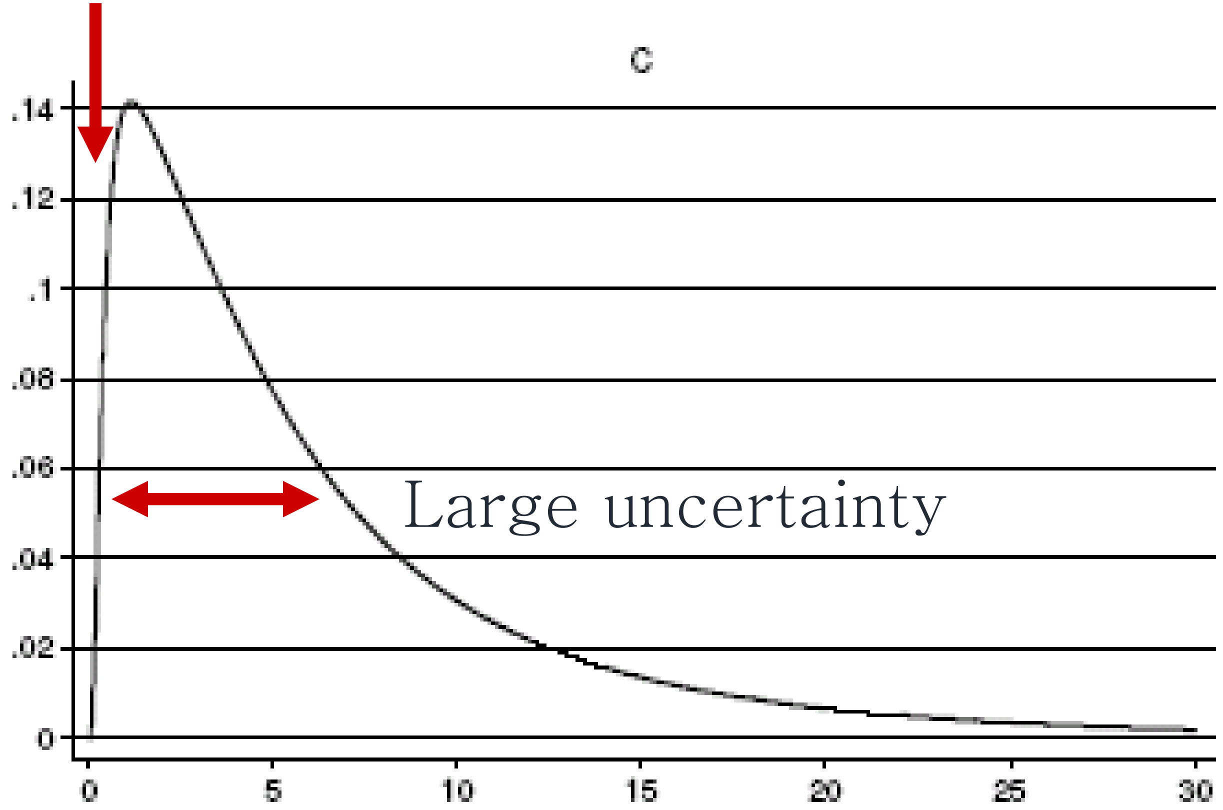
when instead he should admit that:

changing assumptions → all changes a lot



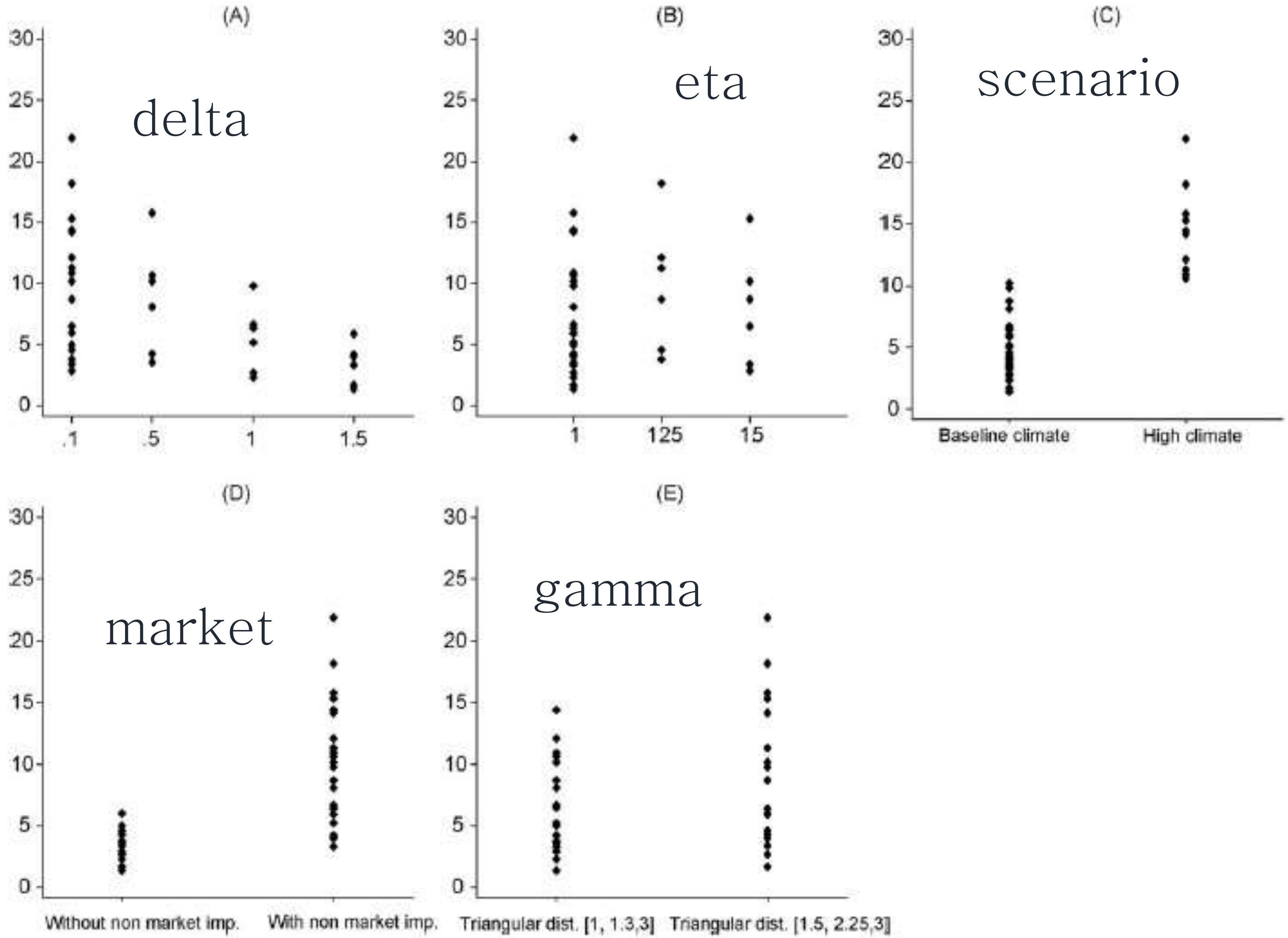
How was it done? A reverse engineering of the analysis

Missing points

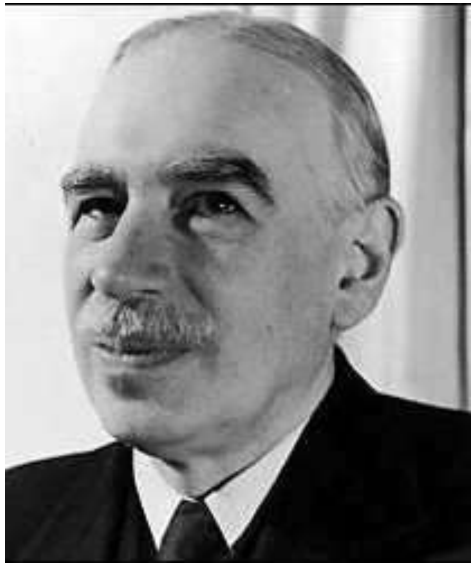


% loss in GDP per capita

Sensitivity analysis here (by reverse engineering)



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



… precisely wrong

Training "Numbers for Policy", Barcelona August 27th – September 1st

http://www.uib.no/en/svt/115575/numbers-policy-practical-problems-quantification

UNIVERSITY OF BERGEN

Universitat Oberta de Catalunya

OPEN EVIDENCE

END

@andreasaltelli

Cooping with uncertainty or quantification hubris

The main issue in existing practices of mathematical modelling is in the management of uncertainty in model-based inference. Modelling studies can be seen which tend to overestimate certainty, pretending to produce crisp numbers precise to the third decimal digits even in situation of pervasive uncertainty or ignorance