

# Sensitivity Analysis

Andrea Saltelli Open Evidence Research, Open University of Catalonia



MNF990 / Theory of Science and Ethics, Bergen, February 15, 2021



#### Where to find this talk: www.andreasaltelli.eu



HOME ABOUT ME

JT ME PUBLICATIONS

NEWS & VIDEOS RESOURCES

(1)

#### CAETERIS ARE NEVER PARIBUS



🔁 andrea saltelli Retweeted

i-site ULNE @isiteULNE

#statistiques #probabilités #modélisation #prédiction Isabelle Bruno du #CERAPS @univ\_lille @CNRS\_HdF @ScPoLille nous parle des dérives de la #quantophrénie dans un article à lire sur le media @FR\_Conversation https://twitter.com/FR\_Conversation/status /1302651033164881920



Sep 7, 2020

#### @Andrea saltelli

Pour mes amis francophones. Honoured to be coauthor of a statactivist like Isabelle Bruno du #CERAPS @univ\_lille @CNRS\_HdF @ScPoLille @OpenEvidence @UOCNews Statistiques et modèles mathématiques : doit-on

Embed

View on Twitter



# Why something technical in this course



#### Environmental Science & Policy Volume 106, April 2020, Pages 87-98



The technique is never neutral. How methodological choices condition the generation of narratives for sustainability

Andrea Saltelli <sup>a, b</sup> 은 쩓, Lorenzo Benini <sup>c</sup>, Silvio Funtowicz <sup>a</sup>, Mario Giampietro <sup>d, e</sup>, Matthias Kaiser <sup>a</sup>, Erik Reinert <sup>a, f</sup>, Jeroen P. van der Sluijs <sup>a, g,</sup> <sub>h</sub> Something general about mathematical modelling

# Caeteris are never paribus

Ceteris paribus or caeteris paribus is a Latin phrase meaning "all other things being equal" or "other things held constant" or "all else unchanged" (Wikipedia) The case of DSGE, dynamic stochastic general equilibrium models

Rational expectations of agents Efficient market hypothesis





#### Philip Mirowski

Philip Mirowski, 2013, Never let a serious crisis go wasted, Verso Books.

#### The US senate and Queen Elisabeth perplexed…







Philip Mirowski, 2013, Never let a serious crisis go wasted, Verso Books.

An ethical problem in the use of models in economics? 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.

Models have little memory

"[…] The process of constructing and validating [value-at risk] models is time consuming and detail oriented; normally even the people who produced the model will not remember many of the assumptions incorporated into it, short of redoing their work, which means that the client cannot simply ask then what went into it."

E. Millgram The Great Endarkenment, p. 29

Modelling is a craft more than a science

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



R. Rosen, Life Itself: A Comprehensive Inquiry Into the Nature, Origin, and Fabrication of Life. Columbia University Press, 1991.





#### Robert Rosen

Can models be falsified?

"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"

# Model-based knowing is conditional

# When models 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						
	7am- 10am	10am- 4pm	4pm-7pm	7pm-7am	Weekday Average	Weekend	All Week
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

# Models and their data

"[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 modeldependency 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

# 3

## Why Atmospheric Modeling Is Good Science

Stephen D. Norton and Frederick Suppe

Changing the Atmosphere: Expert Knowledge and Environmental Governance, edited by Clark A. Miller, Paul N. Edwards,

Changing the Atmosphere Expert Knowledge and Environmental Governance

## Deep Uncertainty methods (Steinman et al., 2020)

- Exploratory modelling → see also Quantitative storytelling
- Global sensitivity analysis → This talk
- Robust decision making → see also Quantitative storytelling

Steinmann, P., Wang, J. R., Voorn, G. A. K. van & Kwakkel, J. H. Don't try to predict COVID-19. If you must, use Deep Uncertainty methods. Rev. Artif. Soc. Soc. Simul. (2020).

# Robust decision making and Exploratory modeling: iterative processes; compares candidate strategies, possible future states of the world and related vulnerabilities → hedging for the trade-offs among ulnerabilities

Lempert, R. J., Groves, D. G., Popper, S. W. & Bankes, S. C. A general, analytic method for generating robust strategies and narrative scenarios. Manage. Sci. 52, 514–528 (2006).

Bankes, S. Exploratory Modeling for Policy Analysis. Oper. Res. 41, 435–449 (1993).

'Quantitative storytelling': stress-test policy options against criteria of feasibility (e.g., do we have the biophysical resources?), viability (can it be done in the present social arrangement?) and desirability (is it what society wants?)

Saltelli, Andrea; Giampietro, Mario (2017). "What is wrong with evidence based policy, and how can it be improved?". Futures. 91:

The philosophy of these approaches:

Exhaustive description of a given hypothesis on the state of the world

or

Exhaustive exploration of possible/plausible worlds

# Economics

The Open-Access, Open-Assessment E-Journal

#### **Discussion Paper**

No. 2019-23 | March 08, 2019 | http://www.economics-ejournal.org/economics/discussionpaper:

#### **Escape from model-land**

Erica L. Thompson and Leonard A. Smith



Beware <<"optimal" model-land quantities obtained from imperfect simulations>>



PATHWAYS TO SUSTAINABILITY

#### THE POLITICS OF UNCERTAINTY

**Challenges of Transformation** 



# 3

# SHARING RISKS OR PROLIFERATING UNCERTAINTIES?

Insurance, disaster and development

Leigh Johnson

Model-based parametric insurance … one more example of the tragedy of reductionism?

**Open access:** https://www.taylorfrancis.com/books/politicsuncertainty-ian-scoones-andystirling/e/10.4324/9781003023845

## 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

European Commission			
Home	Better Regulation Guldelines		🖸 Share 🛛 🖸 🔯 🕅
Stakeholder consultations	These guidelines explain what Better Regulation is and how it sho to day practices when preparing new initiatives and proposals or n policies and laxialation	Search	
Roadmaps / Inception Impact Assessments Impact Assessment Evaluation	policies and legislation. They cover the whole policy cycle, from policy preparation and adog and application, to evaluation and revision of EU law. For each of th number of Better Regulation principles, objectives, tools and proce the EU has the best regulation possible. These relate to planning,	Stay connected	
Regulatory Scrutiny Board Guidelines	staxenoider consultation, implementation and evaluation. The <u>Better Reoutlation Guidelines</u> are structured into chapters while instruments of the law-making process. The corresponding <u>loolog</u> and technical information	Latest documents <u>19/05/2015 - Better Regulation</u> Package	
Better Regulation "Toolbox" Key documents	Better Regulation Guidelines are based on the outcomes of public carried out in 2013 and 2014.	Help us improve	
	Public consultation on the revision of the Commission's Impa <u>Guidelines</u> Stakeholder Consultation Guidelines	Find what you wanted?	
	Consultation on the draft Commission Evaluation Policy Guide Consultation on the draft Commission Evaluation Policy Guide	What were you looking for?	
			Any supporting 2

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.

	Press Archives   Stemep   About this site   Legal	rotice   Contact   Search   English (en) 💌
European Commission European Commission > Better Regulatio	m y Guldelfres	
Home	Better Regulation Guidelines	Share 🚺 🖬 🗋 🖂 📥
REFIT Stakeholder consultations	These guidelines explain what Beller Regulation is and how it should be applied in the day to day practices when preparing new initiatives and proposals or managing existing evidence and indicated by the statement of the stat	Search
Assessments	policies and registration.	Stay connected
Impact Assessment	and application, to evaluation and revision of EU law. For each of these phases there are a	facebook 🔛 Twitter 🊟 EU Tube
Evaluation	number of Better Regulation principles, objectives, tools and procedures to make sure that the EU has the bestregulation opsile. These elide to planning, impact assessment, stateholder consultation and evaluation. The <u>Better Regulation Guidelines</u> are structured into chapters which cover each of the instruments of the law-making process. The corresponding toolbox gives more detailed and technical information. Better Regulation Guidelines are based on the outcomes of public consultation exercises carried out in 2013 and 2014.	1 223
Regulatory Scrutiny Board		Latest documents
Guidelines		19/05/2015 - Better Regulation Package
<ul> <li>Better Regulation "Goldbar"</li> <li>Better Regulation "Toolbar"</li> <li>Key documents</li> </ul>		Help us improve
	<ul> <li>Public consultation on the revision of the Commission's Impact Assessment Guidelines</li> </ul>	Find what you wanted?
	<ul> <li><u>Stakeholder Consultation Guidelines</u></li> </ul>	Yes O No O
	Consultation on the draft Commission Evaluation Policy Guidelines	What were you looking for?
		Any suggestings?
		with and Accounts :
		Send
	Lest update: 11/08/2015   Leost notice   Contect   Search   Top	

## A 10 points participatory checklist (Jakeman et al. 2006)

Jakeman, A. J., Letcher, R. A., & Norton, J. P. (2006). Ten iterative steps in development and evaluation of environmental models,. Environmental Modelling & Software, 21(5), 602–614.







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.



**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.

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**



#### **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	<ul> <li>original set of weights,</li> </ul>
	<ul> <li>factor analysis,</li> </ul>
	<ul> <li>equal weighting,</li> </ul>
	<ul> <li>data envelopment analysis</li> </ul>
Aggregation rule	<ul> <li>additive,</li> </ul>
	<ul> <li>multiplicative,</li> </ul>
	<ul> <li>Borda multi-criterion</li> </ul>





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'





Sobol' sequences of quasirandom 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=numericversus-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

the

Search...

Help | Adva

# Why Sensitivity analysis?

# It can answer interesting questions

Global Environmental Change 20 (2010) 298-302



#### 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

Stern, N., Stern Review on the Economics of Climate Change. UK Government Economic Service, London, <u>www.sternreview.org.uk</u>.

William Nordhaus, University of Yale Nobel 'Economics' 2018

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)



Why sensitivity analysis?

It allows interesting discoveries

Journal of the Royal Statistical 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



University rankings such as ARWU and THES are technically unsound Journal of the Royal Statistical Society



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

#### Ratings and rankings: voodoo or science?

Why sensitivity analysis

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

European Commission, Ispra, Italy

# Limits of sensitivity analysis



Orrin H.

Pilkey

useless arithmetic

Cash's Predict the Futur

By Emironmental Scientists

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

Once H. Pilley & Linda Pilley-Janci

The map is not the territory

Useless arithmetic Wy Indexedul Sciences Carl Productive Focus Orme 11, Pillery & Lable Pillery-Java <>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  $\rightarrow$  one is the low permeability of the geological formation  $\rightarrow$  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 <sup>36</sup>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 SENSITIVIT ANALYSIS The Primer	
<b>⊛</b> ₩ILEY	

全局敏感性分析 【意】萨特利(A. Sahutti)等一著 坚麻斑 丁义明 琦 鸣 液结风口静 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 ~ 
$$E_{\mathbf{X}_{i}}(Y|X_{i})$$



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

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





Which factor has the highest  $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)$ ? For <u>additive</u> models one can decompose the total variance as a sum of first order effects

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

··· which is also how additive models are defined

## Non additive models





### Is 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) =

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

## 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  $k(k-1)/2 \cdots$ 

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

 $+ V_{12} + V_{13} + V_{23} + V_{123}$  $+ V_{123}$ 

 $+ S_{123}$ 

## 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 )

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

## 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

<u>https://www.uqlab.com/</u> (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 correlationregression 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 "tatistics & 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 Environmenial 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



Model complexity



Model 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



#### Model Complexity



#### Comment Open Access Published: 27 August 2019

# A short comment on statistical versus mathematical modelling



Sixth secret:

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

Proving woldstay is one string the viscous de to serve to be to be

Nassim Nicholas Taleb Author of the bestselling phenomenon The Black Swan 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-ata-time) SA so bad?



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





# 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

• Memento

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



Model Complexity



#### 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 <sup>a, b</sup> 은 쩓, Ksenia Aleksankina <sup>c</sup>, William Becker <sup>d</sup>, Pamela Fennell <sup>e</sup>, Federico Ferretti <sup>d</sup>, Niels Holst <sup>f</sup>, Sushan Li <sup>g</sup>, Qiongli Wu <sup>h</sup>

# Literature search in Scopus

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



- 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:

 $\rightarrow$  324 articles, divided among authors

Cleansing manually irrelevant articles:

 $\rightarrow$  280 articles



#### Still many papers apply an OAT SA: 65%

## What if the model is truly linear?

# Linear Nonlinear Unclear

7% 61% 32%

### Linear

# Nonlinear Unclear



## 65% highly cited articles are OAT

Taking all unclear = linear → still over 20% (.32+.07)\*.65 of papers wrong (OAT & non-linear model)



# 5. Discussion

# 5.1. Reasons for bad practice



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:



**EWS** 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

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?  $\rightarrow$  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? 1. Statistics as a discipline takes responsibility for statistical methods for model validation and verification

Example: who can authoritatively suggest to modellers not to overinterpret results from multi-model ensembles?



### Climate Models as Economic Guides: Scientific Challenge or Quixotic Quest?

BY ANDREA SALTELLI, PHILIP B. STARK, WILLIAM BECKER, PAWEL STANO



#### Climate Models as Economic Guides: Scientific Challenge or Quixotic Quest?

BY ANDREA SALTELLI, PHILIP B. STARK, WILLIAM BECKER, PAWEL STANO

A plea against audacious risk or cost-benefit analysis running over centennial time scales; example: crime rate as modified by climate change at US county level in 2100

Solutions? 2. Learn from what happens in statistics where the p-test crisis is being tackled head on



Throw away the concept of statistical significance?





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/

## SIGNIFICANC

**IN PRACTICE** Politics Culture Business **Cargo-cult statistics** and scientific crisis

The mechanical, ritualistic application of statistics is contributing to a crisis in science. Education, software and peer review have encouraged poor practice – and it is time for statisticians to fight back. By **Philip B. Stark** and **Andrea Saltelli** 

## The End

# @andreasaltelli

