

### Sensitivity analysis: Scope and limitations

Statistical and participatory tools for Impact Assessment 28-29 April, 2016 Brussels - PLB3 - 4/59 Rue P.Lebon 3 - ROOM 3.36

Andrea Saltelli

#### andrea.saltelli@uib.no

European Centre for Governance in Complexity, Barcelona

#### www.andreasaltelli.eu



Welcome to the home page of Andrea Saltelli

Caeteris are never paribus

#### Where to find this presentation

sensitivity analysis, sensitivity auditing, science for policy, impact assessment, ...

andrea.saltelli@uib.no

#### Topics for this course

#### Sensitivity analysis - first day

- Why sensitivity analysis?
- Limits of sensitivity analysis: Will any sensitivity analysis do?
- Sensitivity analysis in the EC impact assessment toolbox
- How this looks like in practice

International conferences on sensitivity analysis every 3 years since 1995; next (8<sup>th</sup>) in November 2016, La Reunion (FR)

Summer schools every two years; next July 2016, Capri Island (sold out)



# Why sensitivity analysis

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

European Commission	Press   Archives   Stemap   About this ste   Legal r Better Regulation	otice Contact Search English (en)	
European Commission > Better Regulati			
REELT	Better regulation		
Stakeholder consultations	Better regulation is about designing EU policies and laws so that they achieve their	Search	
Roadmaps / Inception Impact Assessments	objectives at minimum cost. It ensures that policy is prepared, implemented and reviewed in an open, transparent manner, informed by the best available evidence and backed up by involving stakeholders.	Stay connected	
<ul> <li>Impact Assessment</li> <li>Evaluation</li> </ul>	To ensure that EU action is effective, the Commission assesses the expected and actual impacts of policies, legislation and other important measures at every stage of the policy cycle - from planning to implementation, to review and subsequent revision.	El Facebook V Twitter 🚜 El Tube	
<ul> <li>Regulatory Scrutiny Board</li> <li>Guidelines</li> </ul>	The Commission decided on 19 May 2015 (172 kB) (2 to create a REFIT Platform to advise the Commission on simplifying and making EU laws more effective and efficient.	Latest documents	
Key documents	The Commission publishes regularly provisional dates of adoption of Commission Initiatives 2.	Guidelines 19/05/2015 - Better Regulation	
	Before the EU takes action	Package	
	<ul> <li>The Commission publishes roadmaps and inception impact assessments describing planned new initiatives and evaluations of existing legislation.</li> </ul>	Help us improve	
	<ul> <li>Commission impact assessments examine the potential economic, social and environmental consequences of proposed options for action.</li> </ul>	Find what you wanted? Yes O No O	

#### http://ec.europa.eu/smartregulation/guidelines/docs/br\_toolbox\_en.pdf

European Commission	Better Regulation			
European Commission > Better Regulation > Guidelines				
Home	Better Regulation Guidelines	Share		
REFIT				
Stakeholder consultations	These guidelines explain what Better Regulation is and how it should be applied in the day to day practices when preparing new initiatives and proposals or managing existing policies and legislation. They cover the whole policy cycle, from policy preparation and adoption to implementation and application, to evaluation and revision of EU law. For each of these phases there are a number of Better Regulation principles, objectives, tools and procedures to make sure that the EU has the best regulation possible. These relate to planning, impact assessment, stakeholder consultation, implementation 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 <u>toolbox</u> gives more detailed and technical information.	Search		
Roadmaps / Inception Impact Assessments		Stay connected		
Impact Assessment		Blogs		
Evaluation Regulatory Scrutiny Board		Latest documents		
Guidelines		19/05/2015 - Better Regulation		
<ul> <li>Better Regulation Guidelines</li> </ul>		Package		
Better Regulation "Toolbox"	<ul> <li>Better Regulation Guidelines are based on the outcomes of public consultation exercises carried out in 2013 and 2014.</li> <li><u>Public consultation on the revision of the Commission's Impact Assessment Guidelines</u></li> <li><u>Stakeholder Consultation Guidelines</u></li> <li><u>Consultation on the draft Commission Evaluation Policy Guidelines</u></li> </ul>	Heln us improve		
Key documents		Find what you wanted? Yes O No O What were you looking for?		

Careford and a

When testing the evidence behind inference some reasonable people (and guidelines) suggest that 'sensitivity analysis would help'



Edward E. Leamer (UCLA, CA); Peter Kennedy; Orrin H. Pilkey (Duke, NC). Edward E. Leamer, 1990, Let's Take the Con Out of Econometrics, American Economics Review, 73 (March 1983), 31-43.



<<I have proposed a form of organised sensitivity analysis that I call "global sensitivity analysis" in which a neighborhood of alternative assumptions is selected and the corresponding interval of inferences is identified.>> Edward E. Leamer, 1990, Let's Take the Con Out of Econometrics, American Economics Review, 73 (March 1983), 31-43.



<<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.>> With the ashes of the mathematical models used to rate mortgagebacked securities still smoldering on Wall Street, now is an ideal time to revisit the sensitivity issues.

Tantalus on the Road to Asymptopia Edward E. Leamer, 2010 *Journal of Economic Perspectives*, **24**, (2), 31–46.



"... my observation of economists at work who routinely pass their data through the filters of many models and then choose a few results for reporting purposes."

"One reason these methods [global sensitivity analysis are rarely used is their honesty seems destructive;" Ihidem

"or, to put it another way, a fanatical commitment to fanciful formal models is often needed to create the appearance of progress." *Ibidem* 



# Peter Kennedy, A Guide to Econometrics.

Anticipating criticism by applying sensitivity analysis. This is one of the ten commandments of applied econometrics according to Peter Kennedy:



<<Thou shall confess in the presence of sensitivity. Corollary: Thou shall anticipate criticism >>



<<When reporting a sensitivity analysis, researchers should explain fully their specification search so that the readers can judge for themselves how the results may have been affected. This is basically an 'honesty is the best policy' approach, [...]'.>>



A GUIDE TO Econometrics FITTER GATION FRITER KENNEDT



Sixth edition 2008

Is this debate over? The reproducibility crisis in cancer research, organic chemistry, psychology, behavioural studies,... The p-values saga and its climax; the ASA statement and the 20 commentaries.



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, DOI:10.1080/00031305.2016.1154108.

Misuse of the P value — a common test for judging the strength of scientific evidence — is contributing to the number of research findings that cannot be reproduced, the American Statistical Association (ASA) warned on 8 March.

nature

NEWS

IN FOCUS

# **Statisticians issue warning on P values**

Statement aims to halt missteps in the quest for certainty.

Baker, M., 2016, Statisticians issue warning on P values, Nature, 531, 151.



Shanks et al. (2015) JEP:General

J Exp Psychol Gen. 2015 Oct 26. "Romance, Risk, and Replication: Can Consumer Choices and Risk-Taking Be Primed by Mating Motives?", Shanks DR, Vadillo MA, Riedel B, Clymo A, Govind S, Hickin N, Tamman AJ, Puhlmann LM.: http://www.ncbi.nlm.nih.gov/pubmed/26501730

#### FEATURE 13 April 2016

#### Statistical and mathematical modelling are at the hearth of

- science for policy
- storm about malpractices.

#### New Scientists talks of "statistical sausage factory"

## Why so much science research is flawed – and what to do about it

Dodgy results are fuelling flawed policy decisions and undermining medical advances. They could even make us lose faith in science. **New Scientist** investigates



LEADER 13 April 2016

#### Science isn't as solid as it should be – but science can fix it

An alarming amount of research is flawed Brett Ryder

Unconscious biases and data-torturing are weakening our knowledge base – but unlike politicians and bankers, scientists aren't covering up their failings







#### Discussion points



- Why doing a sensitivity analysis if it can undermine an laborious quantification exercise?
- What do I do if this happens to be the case?
- What do I do if my consultant returns me an assessment/quantification without sensitivity analysis?

Limits of sensitivity analysis



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





<<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"

#### Ref. → Robert K. Merton's 'Organized skepticism '

Communalism - the common ownership of scient40

ific 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.

Will any sensitivity analysis do the job?

# Can I lie with sensitivity analysis as I can lie with statistics?



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

# Why not just changing one factor at a time (OAT)?

#### <<"one-at-a-time" (OAT) approach is most commonly used in Commission IAs>>



Source: IA Toolbox, p. 391



#### EUROPEAN COMMISSION

#### **Better Regulation "Toolbox"**

Last update: 11/08/2015 | Legal notice | Cookies | Contact | Search | Top

Why not just changing one factor at a time (OAT)?

"Sensitivity analysis usually proceeds by changing one variable or assumption at a time, but it can also be done by varying a combination of variables simultaneously to learn more about the robustness of your results to widespread changes".

Source: Office for the management and Budget of the White House (OMB), Circular A4, 2003

https://www.whitehouse.gov/omb/circulars\_a004\_a-4/



Why not just changing one factor at a time (OAT)?

Because it is a bad idea!

# 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




Bottom-line: once a sensitivity analysis is done via OAT there is no guarantee that either uncertainty analysis (UA) or sensitivity analysis (SA) is any good:

→ UA will be non conservative

→ SA may miss important factors



### Discussion points



- Does this geometric argument make any sense to me?
- What do I do if my consultant returns with an analysis based on changing a factor at the time (OAT)?
- Can OAT be justified in some cases?

Sensitivity analysis in the EC impact assessment toolbox

### Time to look at the EC own guidelines: what do they about SA?

European Commission	Better Regulation		
Home	Better Regulation Guidelines	Share	
Stakeholder consultations Roadmaps / Inception Impact Assessments Impact Assessment Evaluation Regulatory Scrutiny Board Guidelines Better Regulation Guidelines Better Regulation 'Toolbox'' Key documents	<ul> <li>These guidelines explain what Better Regulation is and how it should be applied in the day to day practices when preparing new initiatives and proposals or managing existing policies and legislation.</li> <li>They cover the whole policy cycle, from policy preparation and adoption to implementation and adoption to display the set of these phases there are a number of Better Regulation principles, objectives, tools and procedures to make sure that the LU has the best regulation principles, objectives, tools and procedures to make sure that the LU has the best regulation principles, objectives, tools and procedures to make sure that the bolter consultation, implementation and evaluation.</li> <li>The <u>Better Regulation Guidelines</u> are structured into chapters which cover each of the instruments of the law-making process. The corresponding <u>toolbox</u> gives more detailed and technical information.</li> <li>Better Regulation Guidelines are based on the outcomes of public consultation exercises carried out in 2013 and 2014.</li> <li><u>Public consultation on the revision of the Commission's Impact Assessment Guidelines</u>.</li> <li><u>Stakeholder Consultation Guidelines</u></li> <li><u>Stakeholder Consultation Guidelines</u></li> <li><u>Consultation on the draft Commission Evaluation Policy Guidelines</u></li> </ul>	Search Stay connected  Stay connected  Stay connected  Stay connected  Stay connected  Search  Search	

### **3.** SENSITIVITY ANALYSIS USING LONG-TERM DISCOUNT RATES

Page 377 SA & discount rates:

- Not OK in general to change discount rates ...
- Unless very long time horizons are considered
- In this latter case OK to decrease with time
- Example: UK Government 3.5% discount rate up to 50 years then down to 1.0% at 300 years.

### 3. SENSITIVITY ANALYSIS USING LONG-TERM DISCOUNT RATES

### Page 377 SA & discount rates:

- SA (using declining discount rates) can affect the results as e.g.
  - long term benefits of new road go up
  - long term cost of biodiversity loss also goes up

# 4. UNCERTAINTY/SENSITIVITY

### Page 384

### - Three ways of doing it:

- Worst/best case scenario analysis
- Partial sensitivity analysis (i.e. changing only some of the assumptions, but not others)
- Monte Carlo but be careful to justify the input distributions

# 4. UNCERTAINTY/SENSITIVITY

Page 384

- A qualitative discussion of the assumptions can help the reader when the numerical analysis is unfeasible

### 4. SENSITIVITY AND UNCERTAINTY ANALYSES

Page 390

- A definition: apportioning the uncertainties of prediction to inputs
- Beware: it needs resources, … do it al least periodically, not for all assessments…
- Can be done one input variable at a time (OAT) or globally in EC IA it is mostly done OAT.

# **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.



### Discussion points



# The influence of the key variables should be investigated by a sensitivity analysis.

- Is something wrong about the statement above (p. 384)
- If I give an assumption for granted/fixed I am in error, if I give it a distribution there are problems to justify it … is this a law of constant misery?

# How this looks like in practice

# 4. SENSITIVITY AND UNCERTAINTY ANALYSES

### Page 392: an example



Input variables

# An engineer's vision of UA, SA



One can sample more than just factors

One can sample modelling assumptions

Example: The output is a composite indicator

Assumption	Alternatives
Number of indicators	<ul> <li>all six indicators included or</li> </ul>
	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>



Can I make sense of a sensitivity analysis just looking at the plots? <sub>Values of →</sub>

Values of output variable



Scatterplots of y versus sorted factors

Values of input variable



Values of input variable



Values of input variable

# The ordinate axis is always Y

The abscissa are the various factors  $X_i$  in turn.

The points are always the same



Values of input variable



Values of input variable

# Which factor is more important?



Twitter: @andreasaltelli



# These are ~1,000 points Divide them in 20 bins of ~ 50 points





~1,000 blue points Divide them in 20 bins of ~ 50 points

Compute the bin's average (pink dots)



A good sensitivity measure is the variance of values of the pink points



Both indices can be computed via Monte Carlo

We use quasi random sequences developed by I.M. Sobol'



An  $LP_{\tau}$  sequence



X1,X2 plane, 100 Sobol' points

X1,X2 plane, 1000 Sobol' points

Sobol' sequences of quasirandom points



Sobol' sequences of quasirandom points



Sobol' sequences of quasi-random points against random points

## Why quasi-random



Root mean square error over 50 different trials. The error refers to the numeric-versus-analytic value the integral of the function 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.

### Quasi-random versus Latin Hypercube



Discrepancy at different sample sizes. Monte Carlo versus Quasi Monte Carlo and Latin Hypercube Sampling.

**Source:** Kucherenko, S., Albrecht, D., Saltelli, A., 2015, Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling Techniques, Submitted to SIAM/ASA Journal on Uncertainty Quantification.

More about the settings:

•Factor prioritisation 
$$\rightarrow S_i \equiv \frac{V(E(Y|X_i))}{V_Y}$$

If the cost of 'discovering' factors were the same for all factors which factor should I try to discover first?

•<u>Factor fixing</u>: Can I fix a factor [or a subset of input factors] at any given value over their range of uncertainty without reducing significantly the output?

$$S_{Ti} \equiv \frac{E(V(Y|\mathbf{X}_{\sim i}))}{V_Y}$$
Factor fixing is useful to achieve model simplification and 'relevance'.

We cannot use  $S_i$  to fix a factor;  $S_i = 0$  is a necessary condition for  $X_i$  to be non-influential but not a sufficient one

 $X_{\rm i}$  could be influent at the second order.

Remarks on factor fixing: 1

Model simplification supported by factor fixing is useful.

"As the complexity of a system increases ... precision and significance (or relevance) become almost mutually exclusive characteristics"

Zadeh's incompatibility principle (1965).



## Lofti Zadeh

Remarks on factor fixing: 2 The model 'relevance' problem

 $R = \frac{\text{number of factors that truly induce variations in the output of interest}}{\text{total number of factors in the model}}$ 

Low R could flag a model meant to intimidate.



Bruce Beck

Summary:

- Easy-to-code, Monte Carlo better on quasi-random points. Estimate of the error available.
- 2. The main effect can be made cheap; its computational cost does not depend upon k.