

Sensitivity analysis, an introduction

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

Seminar at Auditorio de Instituto de Hidráulica Ambiental de la Universidad de Cantabria, March 11, 2019







Where to find this talk: www.andreasaltelli.eu



On modelling

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

Jakeman et al. call for a 10 points participatory checklist $% \left(\frac{1}{2}\right) =0$

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.

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.







Modelling is 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

Saltelli, A., 2018, Discussion Paper: Should statistics rescue mathematical modelling? https://arxiv.org/abs/1712.06457

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							
	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

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

Why Sensitivity analysis?

IDENCE. RGUMENT.& SI IASION IF POLICY

"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?"

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_bulleti n_010906.pdf, pp. 16–17, accessed December 2015.

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



Source: IA Toolbox, p. 391



EUROPEAN COMMISSION



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.

Futoment	Better Regulation			
Commission Sumpson Communica > Better Angulate	m y Guddelinna			
Home	Better Regulation Guidelines	SHR.	Share COA	
Stakeholder consultations	These guidelines explain what Better Regulation is and how it st	Search		
Roadmaps / Inception Impact	to day practices when preparing new initiatives and proposals or managing existing policies and legistation		Stay connected	
Assessments	They cover the whole policy cycle, from policy preparation and ad	Tacatoos 💟 Texter 🎘 EL Taba		
Impact Assessment	and application, to evaluation and revision of EU law. For each o number of Better Regulation principles, objectives, tools and pro			
Evaluation	the EU has the best regulation possible. These relate to plannin states holder consultation, implementation and evaluation	Entrated designments		
Regulatory Scrutiny Board	The Dater Devision Guidelines are shurbined into charters of	to charden which ever such of the	Latest documents	
Guidelines Better Regulation Guidelines Better Regulation "Toolbor"	instruments of the law-making process. The corresponding logicar gives more detailed and technical information		 <u>1905/2015 - Better Regulation</u> Package 	
	Better Regulation Guidelines are based on the outcomes of publ carried out in 2013 and 2014.	Help us improve		
Key documents	 Public consultation on the redision of the Commission's Innext Assessment Consumers 		Find what you wanted?	
	 Stakeholder Consultation Goldelines 	Vis C No C		
	Consultation on the draft Commission Evaluation Police Gui	What were you looking for?		
			Any suggestions?	
			Send	

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-Jancis

Useless arithmetic Wy Twine world Sciences Carl Productive Focus Orme 11, Pillery & Lable Pillery-Javes <>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 ³⁶Cl 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.





Robert K. Merton

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

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

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.

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

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



Was this an uncertainty or a sensitivity analysis?







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

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



TOT SA/TOT MOD (%)

TOT GSA/TOT MOD (%)

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.201 6.02.133

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





How would you test the scaffolding?

How coupled ladders are shaken in most of available literature How to shake coupled ladders



WHY WE UNDERESTIMATE RISK IN

THE FACE OF UNCERTAINTY

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 ~
$$E_{\mathbf{X}_{i}}(Y|X_{i})$$



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

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





Which factor has the highest $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)$?

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





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

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

Why?

Because:

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

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…

must be the expected reductionin variance that would be achievedif factor Xi could be fixed

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: second order terms for a model with k factors are as many as $k(k-1)/2 \cdots$ 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)

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



Is this conditional variance smaller or larger than V(Y)?

What is the shortcoming of S_{Ti} ?


Coding Si 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'





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



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? 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. Sixth secret:

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

<text>

Doing the right thing

or

Avoiding something wrong?

Nassim Nicholas Taleb

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





Why?

An example: Sensitivity analysis: the case of the Stern review 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



William Nordhaus, University of Yale



Nicholas Stern, London School of Economics

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).

<u>The Stern – Nordhaus exchange on SCIENCE</u>

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'



... 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)



END



EC impact assessment guidelines: sensitivity analysis & auditing

Commission	in) Gudelmaa	
Home	Better Regulation Guidelines	Share 🛛 🖸
Stakeholder consultations Roadmaps / Inception Impact	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.	Search
		Stay connected
Assessments	They cover the whole policy cycle, from policy preparation and adoption to implementation	Tacatoos 💟 Tactler 🚜 C
Impact Assessment	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	a true
* Evaluation	the EU has the best regulation possible. These relate to planning, impact assessment, stakeholder consultation, implementation and evaluation.	Latest documents
Regulatory Scrutiny Board	The Better Regulation Guidelines are structured into chapters which cover each of the	
Guidennes Defen Gemutation Cuidelines	instruments of the law-mailing process. The corresponding tooloor gives more detailed and technical information	Package
Better Regulation 'Toolbox' Better Regulation 'Toolbox' Key documents	Better Regulation Guidelines are based on the outcomes of public consultation exercises carried out in 2013 and 2014.	Help us improve
	 Public consultation on the revision of the Commission's Impact Assessment Guidelines 	Find what you wanted?
	 Stakeholder Consultation Goldelines 	YHS C No C
	 Genauftation on the draft Commission Evaluation Policy Guidelines 	What were you looking for?
		Any suggestions?

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

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 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"

Paul N. Edwards, 1999, Global climate science, uncertainty and politics: Data-laden models, model-filtered data. More than a technical uncertainty and sensitivity analysis? 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 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



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

"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"

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.

The rules of sensitivity auditing

- 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.