European Commission COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT



Sensitivity analysis and auditing

Andrea Saltelli andrea.saltelli@jrc.ec.europa.eu

12th JRC Annual Training on Composite Indicators & Multicriteria Decision Analysis (COIN 2014) European Commission Joint Research Centre Econometrics and Applied Statistics Unit Composite Indicators Research Group (JRC-COIN)

Andrea Saltelli



Some recap



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Andrea Saltelli

Some conclusions.

- Composite indicators are models. They are built for analysis and advocacy, and are defined by their quality.
- Analysis, advocacy and quality are not independent from one another. Example: most developers adopt for transparency and simplicity linear aggregation procedures to build composite indicators which are fraught with considerable difficulties. In this case quality may suffer at the expenses of advocacy (Saltelli and Saisana, 2013).
- Composite indicators sit between analysis and advocacy, but quality discriminates the plausible from the rhetorical.

Saltelli, A., and Saisana, M., Advocacy, analysis and quality. The Bermuda triangle of Statistics, International Statistical Institute Conference, Hong Kong, August 2013, Statistics and Policy.



Some recap

Andrea Saltelli



Commission

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Some conclusions.

- Building a composite indicator can be seen as a process of social discovery for which a model of extended participation comes natural.
- Frames and indicators are co-produced in the process which must be designed as to have a meaningful 'interpretant', or 'end-in-sight'.





Some recap



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Andrea Saltelli

Issues with trust / quality in the scientific enterprise

• The centrality of ethics for the health of the scientific enterprise is well described by Ravetz (1971).

"Two separate factors are necessary for the achievement of worthwhile scientific results: a community of scholars with shared knowledge of the standards of quality appropriate for their work and a shared commitment to enforce those standards by the informal sanctions the community possesses; and individuals whose personal integrity sets standards at least high as those required by their community."

Ravetz, J., 1971, Scientific Knowledge and its Social Problems, Oxford University Press, p.22.



Jerome R. Ravetz





COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Cargo Cult Science

by RICHARD P. FEYNMAN

Some remarks on science, pseudoscience, and learning how to not fool yourself. Caltech's 1974 commencement address.



CALTECH

Feynman's cargo cult

Andrea Saltelli



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

"[...] cargo cult science. In the South Seas there is a cargo cult of people. During the war they saw airplanes land with lots of good materials, and they want the same thing to happen now. So they've arranged to imitate things like runways, to put fires along the sides of the runways, to make a wooden hut for a man to sit in, with two wooden pieces on his head like headphones and bars of bamboo sticking out like antennas--he's the controller--and they wait for the airplanes to land."



Andrea Saltelli



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

"They're doing everything right. The form is perfect. It looks exactly the ay it looked before. But it doesn't work. No airplanes land. So I call these things cargo cult science, because they follow all the apparent precepts and forms of scientific investigation, but they're missing something essential, because the planes don't land."





"[...] there is one feature I notice that is generally missing in cargo cult science. That is the idea that we all hope you have learned in studying science in school [...]."





Andrea Saltelli



"It's a kind of scientific integrity, a principle of scientific thought that corresponds to a kind of utter honesty--a kind of leaning over backwards. [...] Details that could throw doubt on your interpretation must be given, if you know them. [...] give all of the information to help others to judge the value of your contribution."



The innocence lost

Andrea Saltelli



Peter Høeg, a Danish novelist, in Borderliners (Høeg, 1995)







Andrea Saltelli



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

"That is what we meant by science. That both question and answer are tied up with uncertainty, and that they are painful. But that there is no way around them. And that you hide nothing; instead, everything is brought out into the open."



Mertonian scientific norms

Andrea Saltelli

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT



Robert K. Merton, sociologist of science, 1910-2003, the father of Science and Technology Studies

JRC-COIN ©





COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Communalism - the common ownership of scientific discoveries, according to which scientists give up intellectual property rights in exchange for recognition and esteem ...

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 Scepticism - all ideas must be tested and are subject to rigorous, structured community scrutiny.



Counter-norms

Andrea Saltelli



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

The same R.K. Merton realized later in life that norms have corresponding counter norms ... see Mitroff, I. I. 1974, Am. Soc. Rev. 39, 579-595.

[We must] consider, first, how potentially contradictory norms develop in every social institution; next, how in the institution of science conflicting norms generate marked ambivalence in the lives of scientists; and finally, how this ambivalence affects the actual, as distinct from the supposed, relations between men of science (Merton, 1963a:80).



Counter-norms



- Solitariness (secrecy, miserism) is often used to keep findings secret in order to be able to claim patent rights, ...
- Particularism [...] a real issue, particularly when you consider the ratio of researchers in rich countries compared with those in poor countries [...]
- Interestedness arises because scientists have genuine interests at stake in the reception of their research. [...]
- Dogmatism because careers are built upon a particular premise (theory) being true ...

Andrea Saltelli



European Commission COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Andrea Saltelli



Invasive and non invasive sensitivity analysis



Step 7. Sensitivity analysis

Testing (composite) indicators: two approaches



Michaela Saisana, Andrea. Saltelli, and Stefano Tarantola (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *J. R. Statist. Soc. A* **168**(2), 307–323.

Paolo Paruolo, Michaela Saisana, Andrea SaltelliRatings and rankings: Voodoo or Science?, *J. R. Statist. Soc. A*, **176** (2), 1-26

Step 7. Sensitivity analysis

First: The invasive approach



Michaela Saisana, Béatrice d'Hombres, Andrea Saltelli, Rickety numbers: Volatility of university rankings and policy implications *Research Policy* (2011), **40**, 165-177



Commission

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Andrea Saltelli

Classement de Shangai : les universités françaises à la traîne

Seules 23 universités françaises figurent dans le classement de Shanghai des 500 meilleurs établissements mondiaux.

C'est un électrochoc qui secoue chaque année les universités tricolores en pleine torpeur estivale. Le classement de Shanghai, qui évalue les performances des meilleurs établissements d'enseignement supérieur mondiaux, vient d'être dévoilé. Ni chute ni progression spectaculaire : avec 23 universités dans le top 500 (22 l'an passé), et 3 dans le top 100 (4 l'an passé), la France se classe au septième rang des 37 pays, rétrocédant une place à la Suède. Des résultats décevants : en légère baisse par rapport à l'an passé, la France ne parvient toujours pas à rattraper son retard sur ses homologues britanniques et allemandes, dont une quarantaine d'universités sont classées.

Palmarès mo	ndial
des universite	ás

Rang	Institution	Pays
_1	Harvard	Etats-Unis
2	Stanford	Etats-Unis
3	Berkeley	Etats-Unis
4	Cambridge	RoyUni
5	Massachusetts Inst. Tech. (MIT)	Etats-Unis
6	California Inst. Tech.	Etats-Unis
7	Columbia	Etats-Unis
8	Princeton	Etats-Unis
9	Chicago	Etats-Unis
10	Oxford	RoyUni
12	Univ. Paris-VI	France
49	Univ. Paris-XI	France
73	Ecole normale sup. de Paris	France
Source	: université de Jiac	o Tong.

Strasbourg-Ine peut qu'espérer que

Comment sauver l'université française

Des pistes pour remédier à la rétrogradation de nos établissements d'enseignement supérieur dans les classements internationaux

e classement de Shanghaï, qui fait désormais référence quels que soient ses défauts, vient de rétrograder encore une fois les universités françaises par ,rapport à leurs sœurs étrangères. Dans les cent premières, il n'en accepte plus que trois : Paris-VI (42°), Paris-XI (49°) et l'Ecole normale supéricure (73°).

Jacques Blamont

Membre de l'Académie des sciences

0 ; revenus : 0. Ses ressources totales seraient donc sept fois inférieures à celles du Michigan. Les chiffres relatifs à d'autres universités américaines donne-

les entreprises abonderaient les bourses. La gestion des droits d'inscription et des bourses ressortirait à l'autonomie universitaire.

Une deuxième leçon à tirer de la pratique internationale est que l'université ne doit pas être asphyxiée par des gens qui n'y ont pas leur place. Une sélection à l'entrée s'impose, soit par une meilleure utilisation du baccalauréat (seuls seraient admis les titulaires d'une mention), soit

University rankings are used to judge about the performance of university systems



ARWU and THES







COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

- These rankings are relevant to today's discourse on Higher Education reform in the EU
- Also academics use ARWU

P. Aghion, M. Dewatripont, C. Hoxby, A. Sapir, A., "Higher aspirations: An agenda for reforming European universities" (Bruegel Blueprint Series N.5, 2008).





Andrea Saltelli



University rankings- yearly published

- + Very appealing for capturing a university's multiple missions in a single number
- Allow one to situate a given university in the worldwide context
- Can lead to misleading and/or simplistic policy conclusions

Question:

Can we say something about the quality of the university rankings and the reliability of the results?





COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Andrea Saltelli

		European
Criteria	Indicator	Weight
Quality of Education	Alumni of an institution winning Nobel Prizes and Fields Medals	10%
Quality of	Staff of an institution winning Nobel Prizes and Fields Medals	20%
Faculty	Highly cited researchers in 21 broad subject categories	20%
Research	Articles published in Nature and Science	20%
Output	Articles in Science Citation Index- expanded, Social Science Citation Index	20%
Academic performance	Academic performance with respect to the size of an institution	10%

PROS and CONS

- ✓ 6 « objective » indicators
- Focus on research performance, overloo ran
- Biased towards hard sciences intensive in
- Favours large institutions

METHODOLOGY

✓ 6 indicators

- ✓ Best performing institution
 =100; score of other
 institutions calculated as a
 percentage
- Weighting scheme chosen by rankers

Linear aggregation of the 6 indicators



Andrea Saltelli		(\mathcal{A})
Criteria	Indicator Euro	
	Academic Opinion: Peer review, 6,354 academics	
Research Quality	Citations per Faculty: Total citation/ Full Time Equivalent faculty	
Graduate Employability	Recruiter Review: Employers' opinion, 2,339 recruiters	
International Outlook	International Faculty: Percentage of international staff	
	International Students: Percentage of international students	
Teaching Quality	Student Faculty: Full Time Equivalent faculty/student ratio	20%

METHODOLOGY

- ✓ 6 indicators
- \checkmark z-score calculated for each indicator; best performing institution =100; other institutions are calculated as a

percentage PROS and CO

Overview THES ranking

- Attempt to take into account teaching q ✓ Two expert-based indicators: 50% of to of transparency)
- yearly changes in methodology
- Measures research quantity

Weighting scheme: chosen by rankers

✓ Linear aggregation of the 6 indicators

COIN 2014 -

12th JRC Annual Training on **Composite Indicators and MCDA**

22-26/09/2014, Ispra IT

Overview- Comparison (2007)



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT



...1



1 – Same top10: Harvard, Cambridge, Princeton, Caltech, MIT and Columbia

2 - Greater variations in the middle to lower end of the rankings

3 - Europe is lagging behind: both ARWU (else SJTU) and THES rankings

4 – THES favours UK universities: all UK universities below the line (in red) ...1 Either this graph or Table 3 from our paper.

(Invasive) Sensitivity Analysis

/SIS

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Andrea Saltelli





Robustness analysis, of ARWU and THES

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







Question:

Can we say something about the quality of the university rankings and the <u>reliability of the results</u>?

Source: Saisana, D'Hombres, Saltelli, 2011, Research Policy 40, 165–177



A DW/L is simulated maples Tap 20		
ARVVU. SIIIulateu Ia	11KS - 10020	Legend:
	1	Frequency lower 15%
		Frequency between 15 and 30%
		Frequency between 30 and 50%
		Note: Frequencies lower than 4% are not shown
	Simulated rank range - SJTU 2008	note. Hequencies lower than 176 are not shown
		Original
		rank
Harvard Univ	100	1 USA
Stanford Univ	89 11	2 USA
Univ California - Berkeley	97	3 usa
Univ Cambridge	90 10	4 UK
Massachusetts Inst Tech (MIT)	74 26	5 usa
California Inst Tech	27 53 19	6 USA
Columbia Univ	23 77	7 USA
Princeton Univ	71 9 11 7	8 USA
Univ Chicago	51 <mark>34</mark> 13	9 USA
Univ Oxford	99	10 UK
Yale Univ	47 53	11 USA
Cornell Univ	27 73	12 USA
Univ California - Los Angeles	9 <mark>84</mark> 7	13 USA
Univ California - San Diego	41 46 9	14 USA
Univ Pennsylvania	6 <mark>71</mark> 23	15 USA
Univ Washington - Seattle	7 71 21	16 USA
Univ Wisconsin - Madison	27 70	17 USA
Univ California - San Francisco	14 9 14 11 7 10 6 6	18 USA
Tokyo Univ	16 16 49 20	19 Japan
Johns Hopkins Univ	7 54 21 17	20 USA

 Harvard, Stanford, Berkley, Cambridge, MIT: top 5 in more than 75% of our simulations.

 $\hfill Univ California SF: original rank 18th but could be ranked anywhere between the 6th and 100th <math display="inline">\hfill position$

•Impact of assumptions: much stronger for the middle ranked universities

THES: simulated ranks – Top 20			
		Legena:	
		Frequency lower 15%	
		Frequency between 30 and 50%	
	Simulated rank range - THES 2008	Frequency greater than 50%	
	0	Note: Frequencies lower than 4% are not shown	
	1-5 6-10 11-15 11-15 16-20 21-25 21-25 56-60 51-65 61-65 66-70 81-85 81-85 81-95 91-95		
HARVARD University	44 56	1 USA	
YALE University	40 49 11	2 USA	
University of CAMBRIDGE	99	3 ик	
University of OXFORD	93 7	4 UK	
CALIFORNIA Institute of Technology	46 50	5 USA	
IMPERIAL College London	74 24	6 UK	
UCL (University College London)	73 23	7 ик	
		8 USA	
MASSACHUSETTS Institute of Technology		9 USA	
	57 50 0		
DUKE University	0 37 27 7 1 0 7 10 6 9 6	13 USA	
IOHNS HOPKINS University	20 10 9 9 7 10 6 6 7 6	13 USA	
CORNELL University	6 24 11 7 6 7 9 9 7	15 USA	
AUSTRALIAN National University	10 30 29 31	16 Australia	
STANFORD University	10 14 7 10 9 10 6 6 7	17 USA	
University of MICHIGAN	6 27 17 9 10 7 14 6	18 USA	
University of TOKYO	16 7 13 7 6 6	19 Japan	
MCGILL University	7 <mark>19 41</mark> 13 9 7	20 Canada	

• Impact of uncertainties on the university ranks is even more apparent.

• M.I.T.: ranked 9th, but confirmed only in 13% of simulations (plausible range [4, 35])

 Very high volatility also for universities ranked 10th-20th position, e.g., Duke Univ, John Hopkins Univ, Cornell Univ.

Non invasive Sensitivity analysis

Second: The <u>non</u>-invasive approach

Comparing the weights as assigned by developers with 'effective weights' derived from sensitivity analysis.



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Andrea Saltelli



This measure Si shall be our ruler for 'importance'; example: Si =0.6 \rightarrow I could reduce the variation of the ARWU score by 60% by fixing 'Papers in Nature & Science'.

Using these points we can compute a statistics (called *S*) that tells us: How much (on average) would the variance of the ARWU scores be reduced if I could fix the variable 'Papers in Nature & Science'? If S=0.6 then I expect to reduce the variance of the index by 60% by fixing this variable.



University Rankings



Comparing the internal coherence of ARWU versus THES by testing the weights declared by developers with 'effective' importance measures.

THES

X1_Academic opinion: 6354 academics	40%
X2_Recruiters' opinion: 2339 recruiters	10%
X3_Full-time equivalent faculty/student ratio	20%
X4_Total citation/full time equivalent faculty	20%
X5_Percentage of full-time international staff	5%
X6_Percentage of full-time international students	5%



Issues with THES: a) 'Opinion' variables' weight overall: >60% instead of 50

b) Faculty/student ratio: 10% instead of 20%
A final word on the issue of robustness in an adversarial context Andrea Saltelli

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Introducing sensitivity auditing

European Commission





Andrea Saltelli

European

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

"If is difficult to get a man to understand something when his salary depends upon his not understanding it"







COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Saltelli, A., van der Sluijs, J., Guimarães Pereira, Â., 2013, Funtowiz, S.O., What do I make of your Latinorum? Sensitivity auditing of mathematical modelling, Submitted to *Foresight and Innovation Policy, Special Issue on Plausibility*, arXiv:1211.2668 [physics.soc-ph]

Andrea Saltelli



Do we need better models?



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 9/2014, Ispra IT

••• or <u>better ways of us</u>



JRC-COIN ©

Andrea Saltelli

Introduction 39

Sensitivity Auditing

Andrea Saltelli

Rule 1

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Check against rhetorical use of mathematical modelling;



The instrumental use of mathematical modelling to advance one's agenda can be termed rhetorical, or strategic, like the use of Latin by the elites and the clergy in the classic age.



Sensitivity Auditing

Andrea Saltelli

Check against 'rhetorical', or 'strategic', or non proportional use of mathematical modeling: is the model being used to elucidate or to obfuscate?

→ Can a simplified model representation be produced for the sake of the IA process?

Orrin H. Pilkey's book: Useless Arithmetics





COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT





Why is it so easy to use models rhetorically?

'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.' '[...] to be of value in theory testing, the predictions involved must be capable of refuting the theory that generated them.' What when the 'theory' is not a law but a mathematical model? 'This is where predictions [...] become particularly sticky.'

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



Andrea Saltelli



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

'[...] 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'.





Adopt an 'assumption hunting' attitude:

- What was 'assumed out'?
- Which are the tacit, or pre-analytic, or normative assumptions;
- Which were the caeteris assumed to be paribus?





The UK transport WebTAG model needs as input 'Annual Percentage Change in Car Occupancy up to 2036.'



COIN 2014 –

12th JRC Annual Training on

John Kay



Andrea Saltelli



Sensitivity Auditing

Andrea Saltelli



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Find sensitive assumptions before these find you.



JRC-COIN ©

troduction 46



From: Saltelli, A., D'Hombres, 2010, Sensitivity analysis didn't help. A practitioner's critique of the Stern review, *GLOBAL ENVIRONMENTAL CHANGE*, 20, 298-302.



The case of Stern's Review – Technical Annex to postscript 2014 -

Andrea Saltelli



European Commission 12th JRC Annual Training on Composite Indicators and MCDA / 2014, Ispra IT



William Nordhaus, University of Yale

JRC-COIN ©



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 onClimate Change, SCIENCE, 317, 201-202, (2007).



RULE Three : find sensitivities before sensitivities find converte -12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Commission

Stern's Review – Technical Annex to postscript (a sensitivity analysis of a cost benefit analysis)

The Stern - Nordhaus exchange on *SCIENCE* Nordhaus \rightarrow falsifies Stern based on 'wrong' range of discount rate (~ you GIGOing) Stern \rightarrow 'My analysis shows robustness'







... but foremost Stern says: changing assumptions → important effect when instead he should admit that: changing assumptions → all changes a lot



Introduction 51

How was it done? A reverse engineering of the analysis COIN 2014-12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Commission

Missing points



% loss in GDP per capita



RULE three : find sensitivities before sensitivities find your 2014 -Annual Training on **Composite Indicators and MCDA Andrea Saltelli** 22-26/09/2014, Ispra IT

Commission

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



RULE three : find sensitivities before sensitivities find you 2014-

Commission

Andrea Saltelli



12th JRC Annual Training on Composite Indicators and MCDA

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 >>





RULE three : find sensitivities before sensitivities find your 2014 -

Andrea Saltelli



Commission

<<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, [...]'.>>

<u>Composite Indicators and M</u>

A GUIDE TO Econometrics FRIEKOTION EXTER KENNEDS





Rule (4): detect garbage in garbage out GIGO;

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Andrea Saltelli



European Commission

GIGO

Andrea Saltelli



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

GIGO (Garbage In, Garbage Out) Science – or pseudo-science – "where uncertainties in inputs must be suppressed lest outputs become indeterminate", From: Uncertainty and Quality in Science for Policy by Silvio Funtowicz and Jerry Ravetz, Springer 1990.











COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Detect artificial inflation or deflation of uncertainty.



Peter Kennedy

Edward E. Leamer



Edward E. Leamer, 1990, Let's Take the Con Out of Econometrics, American Economics Review, 73 (March 1983), 31-43. Andrea Saltelli



COIN 2014 – ual Training on tors and MCDA 3/2014, Ispra IT

<<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. 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.>>

JRC-COIN ©



Sensitivity Auditing

Andrea Saltelli



COIN 2014 – 12th JRC Annual Training on **Composite Indicators and MCDA** 22-26/09/2014, Ispra IT

Aim for transparency: stakeholders should be able to make sense of, and possibly replicate, the results of the analysis; see the PRIMES history on the FT

Commission

Doubts raised over Europe's green energy plan

'Host of quastions' from advisers

Economic model lacks transparency By PEER Clark in Lamilar

The condicities of a former gracy, Euclasian Euclide, even exact some distabilities approtidad advision who see transported models. plane to only contaction seemsand a best of guestings" ton's simi of a also of the rating, according three Lenhod superi 25

for suscession model trower on Primar, it ternad default about the role of the to the Balanced Terbulen) of find prim summittees Deliversity of Adlance and the white evening a

The Environt Country siers has used in her caupe out to lists golde the Moc's margy pradue, that industry whiles complary East its assessment-forms htpossible to question here as much a priraids could the had

pass savage vertex has has called for the Carpindam. In use other, more The formerhing workers there up to \$256 are hand of the energy read monan sweeperir model which will tweeper the effect writed by a single thread of closeding, een, more wind minerally that ranked he shall bee minibul posses he dependently service and . caset factors's given the The energy attack large gain, has heightened you cent shut for madels low to Burgent Gen transportace, the super-69.9 ampier's ret model mult shows. Our of the group's THAT STREET -"derivated inegaty" he have shown by Texnels the Ostationish was using concastid on the the Primes model to pro-Encars Sundance Dist" door debaral warray and amunica for the read may there was considered a

Bo the Primas multil," sold and to show how the the superior has the ground of different screep of which is theirs by Cheter torgy socaso alloct the finite, on Onderl knivserity explants professors, stad



A setting plant in Learnary: the multiplity of plant to out 10 and alone has been asked into peoplics by apports

incluine hodies such as of the prep pointed on the interactional discourt taking that it does have advecourt

the roads of different and. sight by elementar but not in subodiate" The group's key on

wen "abers die fritzigen franke Gigrun an mone navy of dae Presse weeh, mit Don Die Didiaria and in particular the property rights in the algo-

Inclependent perties cannot replicate the results' Decause the model is private property.

(Elais and detailed televant) workings of the model". the report, which its marked "East dial" "The model remains the property of the Intistud Indusiant Endower stly of Affarms" E mark. "The commission is first independent parties status. Completion, but sombers Parsor Deparation

rememberships for the could Thern were also questions id by of the read man.

The address proop on "Be omneth for the relices sentrytics at public fire made he cashe publich strathetin "he that de seaulte one he reglizated by pass mand painting

> Tailaical Director Addate who hadd the Printers candled, solid the Pleasantial Thusan he opposed that typerviews transienters and model and signif if some of the model's

workings were made public. "But not the ords Hard BOT THE ROTATION? A myslowenitest Est youngs summing GROUPS ONTINGER, MAIL (IN

100

sould not compared on an upperhisted decisions. The ficial method of the advisory press's report would be initiand still the cowers mod map test specify. Post Caprist has here stargy occurations for many years and has held you tions an hodes ranging separate the rend to. This is from Ground course y login commencial marker for the Loker to the estimate 's Petidi

Introduction 60

RULE FIVE: aim for transparency

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT



"Experts have "raised a host of questions" about how the European Commission's use of a non-transparent model could affect the energy review, according to a leaked report by energy specialists chosen by Brussels to advise on the forthcoming "Energy Roadmap to 2050" FT November 6, 2011



European Commission

Andrea Saltelli

RULE FIVE: aim for transparency

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT



"The credibility of a European energy review has been cast into doubt by experts who point out that long-term plans to cut carbon emissions are based on an economic model owned by a single Greek university that cannot be independently scrutinised."



European Commission

Andrea Saltelli



Office of Management and Budget

Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies; Notice; Republication

The OMB about transparency

http://www.whitehouse.gov/omb/inforeg/



COIN 2014 -

ing on ⁄ICDA

pra IT

RULE FIVE: aim for transparency

Andrea Saltelli



COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

[models should be made available to a third party so that it can] use the same data, computer model or statistical methods to replicate the analytic results reported in the original study.

[...] The more important benefit of transparency is that the public will be able to assess how much an agency's analytic result hinges on the specific analytic choices made by the agency.

Friday, February 22, 2002 Graphic - Federal Register, Part IX Office of Management and Budget Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies; Notice; Republication http://www.whitehouse.gov/omb/inforeg/







JRC-COIN ©





COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

^{*}House Republicans Aim To Limit Power Of Environmental Protection Agency This is 2014

The Huffington Post | by Robin Wilkey (/robin-wilkey)

Posted: 02/07/2014 6:18 pm EST | Updated: 02/08/2014 10:59 am EST





COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

The bill, dubbed the Secret Science Reform Act would force the EPA to publicly release its research on a topic before issuing a policy recommendation, and require that the research be "reproducible." Supporters claim the bill will increase transparency in public policy, while opponents have accused the bill's authors of trying to "keep the EPA from doing its job."





^{113TH CONGRESS} 2D SESSION H.R.4012

To prohibit the Environmental Protection Agency from proposing, finalizing, or disseminating regulations or assessments based upon science that is not transparent or reproducible.

http://beta.congress.gov/bill/113th-congress/house-bill/4012 Accessed May 2014



I





COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Do the right sums. Beware type III errors.

[...] 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"

→ Check that relevant normative stances are not neglected







RULE SIX: Do the right sums

Andrea Saltelli

The spectre of type III errors: Commission **Donald Rumsfeld version: "Reports** that say that something hasn't happened are always interesting to me, because as we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns -- the ones we don't know we don't know."



COIN 2014 -

ig on CDA ra IT




Do the analysis using proper methods

(from Sam L. Savage's The flaw of averages)



RULE SEVEN: Explore diligently the space of the assumptions

The most popular SA practice seen in the literature is that of 'onefactor-at-a-time' (OAT). This consists of analyzing the effect of varying one model input factor at a time while keeping all other fixed.



While the shortcomings of OAT are known from the statistical literature, its widespread use among modelers raises concern on the quality of the associated sensitivity analyses



Sensitivity Analysis

Andrea Saltelli



COIN 2014 – Sensitivity Auditien grc Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

[...] policy-related science calls for an extension of the traditional internal, peer review-based methods of quality assurance to higher levels of supervision, where extended participation and explicit value judgments are necessary [...]

[similarly] \rightarrow

[...] sensitivity analysis must **extend beyond** the technical exploration of the space of uncertain variables and parameters



References and Related Reading

COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Andrea Saltelli

- 1. Paruolo P., Saisana M., Saltelli A., 2013, Ratings and Rankings: voodoo or science?. *J Royal Statistical Society A* **176(2)**.
- 2. Saisana M., Saltelli A., 2012, *JRC audit on the 2012 WJP Rule of Law Index,* In Agrast, M., Botero, J., Martinez, J., Ponce, A., & Pratt, C. WJP Rule of Law Index[®] 2012. Washington, D.C.: The World Justice Project.
- 3. Saisana M., Philippas D., 2012, *Sustainable Society Index (SSI): Taking societies' pulse along social, environmental and economic issues,* EUR 25578, Joint Research Centre, Publications Office of the European Union, Italy.
- 4. Saisana M., D'Hombres B., Saltelli A., 2011, Rickety Numbers: Volatility of university rankings and policy implications. *Research Policy* **40**, 165–177.
- 5. Saisana M., Saltelli A., Tarantola S., 2005, Uncertainty and sensitivity analysis techniques as tools for the analysis and validation of composite indicators. *J Royal Statistical Society A* **168(2)**, 307-323.
- 6. OECD/JRC, 2008, *Handbook on Constructing Composite Indicators. Methodology and user Guide*, OECD Publishing, ISBN 978-92-64-04345-9.





COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Click to edit Master text styles

- Second level
 - Third level
 - Fourth level
 - Fifth level







COIN 2014 – 12th JRC Annual Training on Composite Indicators and MCDA 22-26/09/2014, Ispra IT

Click to edit Master text styles

- Second level
 - Third level
 - Fourth level
 - Fifth level



