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 CERFACS, December 3rd, 2019, 10.30, 42 Avenue Gaspard Coriolis, Toulouse
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.


Define model purpose

Specify modelling context

Conceptualise system, specify data and other prior knowledge

Select model features: nature, family, form of uncertainty specification

Determine how model structure and parameter values are to be found

Choose estimation/performance criteria and algorithm

Identify model structure and parameter values

Respecify objectives if necessary

Reassess if necessary

Don’t start here!

May need to revisit previous steps
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

Model-based knowing is conditional
For John Kay modelling may need as input information which we don’t have

## WebTAG: Annual Percentage Change in Car Occupancy (% pa) up to 2036

<table>
<thead>
<tr>
<th>Journey Purpose</th>
<th>7am-10am</th>
<th>10am-4pm</th>
<th>4pm-7pm</th>
<th>7pm-7am</th>
<th>Weekday Average</th>
<th>Weekend</th>
<th>All Week</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work</td>
<td>-0.48</td>
<td>-0.4</td>
<td>-0.62</td>
<td>-0.5</td>
<td>-0.44</td>
<td>-0.48</td>
<td>-0.45</td>
</tr>
<tr>
<td>Non-Work (commuting and other)</td>
<td>-0.67</td>
<td>-0.65</td>
<td>-0.53</td>
<td>-0.47</td>
<td>-0.59</td>
<td>-0.52</td>
<td>-0.56</td>
</tr>
</tbody>
</table>
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?
"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


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

Source: IA Toolbox, p. 391

EUROPEAN COMMISSION

Better Regulation "Toolbox"
4. SENSITIVITY AND UNCERTAINTY ANALYSES

Page 391

Six steps for a global SA:
1. Select one 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 both 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.
Limits of sensitivity analysis
Useless Arithmetic: Why Environmental Scientists Can't Predict the Future
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 → one is the low permeability of the geological formation → 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 $^{36}$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 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.
Hacked! available for free at

http://www.andreasaltelli.eu
An engineer’s vision of UA, SA

Simulation Model

- Resolution levels
- Model structures
- Parameters
- Data
- Errors

Uncertainty analysis
Sensitivity analysis
Feedbacks on input data and model factors
One can sample more than just factors
One can sample modelling assumptions, alternative data sets, resolution levels, scenarios …
<table>
<thead>
<tr>
<th>Assumption</th>
<th>Alternatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of indicators</td>
<td>▪ all six indicators included or one-at-time excluded (6 options)</td>
</tr>
<tr>
<td>Weighting method</td>
<td>▪ original set of weights,</td>
</tr>
<tr>
<td></td>
<td>▪ factor analysis,</td>
</tr>
<tr>
<td></td>
<td>▪ equal weighting,</td>
</tr>
<tr>
<td></td>
<td>▪ data envelopment analysis</td>
</tr>
<tr>
<td>Aggregation rule</td>
<td>▪ additive,</td>
</tr>
<tr>
<td></td>
<td>▪ multiplicative,</td>
</tr>
<tr>
<td></td>
<td>▪ Borda multi-criterion</td>
</tr>
</tbody>
</table>
Space of alternatives

- Weights
- Missing data
- Aggregation
- Including/excluding variables
- Normalisation

Sensitivity analysis

Country 1
Country 2
Country 3
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 $y$.

<table>
<thead>
<tr>
<th>$x_{11}$</th>
<th>$x_{12}$</th>
<th>...</th>
<th>$x_{1k}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
<td>...</td>
<td>$x_{2k}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$x_{N1}$</td>
<td>$x_{N2}$</td>
<td>...</td>
<td>$x_{Nk}$</td>
</tr>
</tbody>
</table>

Examples of distributions of input factors.
Can one lie with sensitivity analysis as one can lie with statistics?

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.

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
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 \( \sim E_{X_{-i}} \left( Y | X_i \right) \)
Take the variance of the pink points and you have a sensitivity measure

\[ V_{X_i} \left( E_{X \sim i} (Y|X_i) \right) \]
Which factor has the highest

\[ V_{X_i}(E_{X_{\sim i}}(Y|X_i)) \]?
\[ S_i \equiv \frac{V\left( E\left( Y|X_{i} \right) \right)}{V_Y} \]
First order sensitivity index:

\[
\frac{\operatorname{V}_{x_i} \left( \mathbb{E}_{x \sim i} (y \mid x_i) \right)}{\operatorname{V} (y)}
\]

Smoothed curve:

\[
\mathbb{E}_{x \sim i} (y \mid x_i)
\]
First order sensitivity index

Pearson’s correlation ratio

Smoothed curve

First order sensitivity index

Unconditional variance

$$S_i \equiv \eta_i^2 : = \frac{V_{x_i} (E_{x \sim i} (y \mid x_i))}{V(y)}$$
\[ V_{X_i} \left( E_{X \sim i} (Y|X_i) \right) \]

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

Why?
Because:

\[ V_{X_i} \left( E_{X_{-i}} \left( Y \mid X_i \right) \right) + \]
\[ + E_{X_i} \left( V_{X_{-i}} \left( Y \mid X_i \right) \right) = V(Y) \]

Easy to prove using \( V(Y) = E(Y^2) - E^2(Y) \)
Because:

\[ V_{X_i} \left( E_{X_{\sim i}} \left( Y \mid X_i \right) \right) + \]

\[ + E_{X_i} \left( V_{X_{\sim i}} \left( Y \mid X_i \right) \right) = V(Y) \]

This is the variance when a factor \( X_i \) is fixed \( \cdots \)
Because:

\[ V_{X_i} \left( E_{X_{\sim i}} \left( Y \mid X_i \right) \right) + E_{X_i} \left( V_{X_{\sim i}} \left( Y \mid X_i \right) \right) = V(Y) \]

This is what variance would be left (on average) if \( X_i \) could be fixed…
... then this ...

\[ V_{X_i} \left( E_{X \sim i} \left( Y \mid X_i \right) \right) + \\
+ E_{X_i} \left( V_{X \sim i} \left( Y \mid X_i \right) \right) = V(Y) \]

... must be the expected reduction in variance that would be achieved if factor \( X_i \) could be fixed
For **additive** models one can decompose the total variance as a sum of first order effects

\[ \sum_{i} V_{x_i} \left( E_{x_{-i}} \left( Y | X_i \right) \right) \approx V(Y) \]

… which is also how additive models are defined
Non additive models
Is $S_i = 0$?
Is this factor non-important?
There are terms which capture two-way, three way, … interactions among variables.

All these terms are linked by a formula
Variance decomposition (ANOVA)

\[ V(Y) = \sum_i V_i + \sum_{i,j>i} V_{ij} + \ldots + V_{123\ldots k} \]
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 \( k(k-1)/2 \) …
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

$V = V_1 + V_2 + V_3 + V_{12} + V_{13} + V_{23} + V_{123}$

Or divided by $V$

$1 = S_1 + S_2 + S_3 + S_{12} + S_{13} + S_{23} + S_{123}$
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
Total effect, or bottom marginal variance =

\[ E_{X_{\sim i}} \left( V_{X_i} (Y|X_{\sim i}) \right) \]

= the expected variance that would be left if all factors but Xi could be fixed (self evident definition)
\[ S_{Ti} \equiv \frac{E(V(Y|X_{\sim i}))}{V_Y} \]
What is the shortcoming of $S_{Ti}$?
Coding $S_i$ and $S_{Ti}$ yourself?

Use this work:


http://www.andreasaltelli.eu/file/repository/PUBLISHED_PAPER.pdf
How to generate the random sample?

Quasi random sequences developed by I.M. Sobol’
An $L_{P_{\tau}}$ sequence
Sobol’ sequences of quasi-random points

X1,X2 plane, 1000 Sobol’ points

X1,X2 plane, 10000 Sobol’ points

Sobol’ sequences of quasi-random points
Sobol’ sequences of quasi-random points against random points
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.

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]
Badly kept secret:
There is always one more bug!
(Lubarsky's Law of Cybernetic Entomology)
And of course please don’t run a sensitivity analysis where each factor has a 5% uncertainty.
Why?
An example:
Sensitivity analysis: the case of the Stern review
Sensitivity analysis didn't help. A practitioner's critique of the Stern review

Andrea Saltelli *, Beatrice D'Hombres

Joint Research Centre, Institute for the Protection and Security of the Citizen, Ispra, Italy
The case of Stern’s Review – Technical Annex to postscript

Nicholas Stern, London School of Economics


William Nordhaus, University of Yale
The Stern – Nordhaus exchange on *SCIENCE*

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

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

3) Stern thus says: *My analysis shows robustness’*
… but foremost Stern says:

changing assumptions $\rightarrow$ important effect

when instead he should admit that:

changing assumptions $\rightarrow$ all changes a lot
How was it done? A reverse engineering of the analysis

Missing points

Large uncertainty

% loss in GDP per capita
Sensitivity analysis here (by reverse engineering)
EC impact assessment guidelines: sensitivity analysis & auditing

Blurring lines:

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

“[in climate modelling] it looks very little like our idealized image of science, in which pure theory is tested with pure data. [impossible to] eliminate the model-dependency of data or the data-ladenness of models”

Paul N. Edwards, 1999, Global climate science, uncertainty and politics: Data-laden models, model-filtered data.
“[For] philosophers Frederick Suppe and Stephen Norton the blurry model/data relationship pervades all science”
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)


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

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.


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

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

Naomi Oreskes

Models are not physical laws

“[…] 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


The rules of sensitivity auditing

1. Check against rhetorical use of mathematical modelling;
2. Adopt an “assumption hunting” attitude; focus on unearthing possibly implicit assumptions;
3. Check if uncertainty been instrumentally inflated or deflated.
4. Find sensitive assumptions before these find you; do your SA before publishing;

5. Aim for transparency; Show all the data;

6. Do the right sums, not just the sums right; frames; ➔ quantitative storytelling

7. Perform a proper global sensitivity analysis.