Course at ICTA: ‘Sensitivity analysis, sensitivity auditing and beyond’
Lesson 1: Sensitivity Analysis

Andrea Saltelli
Centre for the Study of the Sciences and the Humanities (SVT) – University of Bergen (UIB)
Institut de Ciència i Tecnologia Ambientals (ICTA) – Universitat Autonoma de Barcelona (UAB)

Barcelona, Bellaterra Campus, February 6–8 2017
CAETERIS ARE NEVER PARIBUS
= more material on my web site

= discussion time
Sensitivity analysis books available on LibGen
What is 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
[Global*] sensitivity analysis: “The study of how the uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input”

An engineer’s vision of UA, SA

Simulation
Model
parameters
Resolution levels
data
errors
model structures
uncertainty analysis
sensitivity analysis
model output
feedbacks on input data and model factors

errors
data
Resolution levels
model structures
model output
uncertainty analysis
sensitivity analysis
feedbacks on input data and model factors
One can sample more than just factors
One can sample modelling assumptions
Example: The output is a composite indicator
<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 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

Country 1

Country 2

Country 3

Sensitivity analysis
Is this an uncertainty analysis or a sensitivity analysis?
If I did a sensitivity analysis what information would I obtain?
Sample matrix for uncertainty and sensitivity analysis

\[
\begin{array}{cccc}
  x_{11} & x_{12} & \cdots & x_{1k} \\
  x_{21} & x_{22} & \cdots & x_{2k} \\
  \vdots & \vdots & \ddots & \vdots \\
  x_{N1} & x_{N2} & \cdots & x_{Nk} \\
\end{array}
\]

Each row is a sample trial for one model run. Each column is a sample of size N from the distribution of the factor.
Each *column* is a sample of size N from the distribution of factor.

\[ x_{11} \quad x_{12} \quad \ldots \quad x_{1k} \\
 x_{21} \quad x_{22} \quad \ldots \quad x_{2k} \\
\ldots \quad \ldots \quad \ldots \quad \ldots \\
 x_{N1} \quad x_{N2} \quad \ldots \quad x_{Nk} \]
Model results:
Each entry is the error-free result of the model run.

\[ y_1 \]

\[ y_2 \]

\[ \ldots \]

\[ y_N \]
In the simplest case $y$ could be a function of – a simple mathematical expression of – the $x_1, x_2, \ldots x_k$

e.g. $y = x_1 \sin(x_2)/x_3$

Or it could be a more complicate mathematical model in a computer code to generate $y$ given $x_1, x_2, \ldots x_k$

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{11}$</td>
<td>$x_{12}$</td>
<td>...</td>
<td>$x_{1k}$</td>
<td></td>
<td>$y_1$</td>
</tr>
<tr>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
<td>...</td>
<td>$x_{2k}$</td>
<td></td>
<td>$y_2$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>$x_{N1}$</td>
<td>$x_{N2}$</td>
<td>...</td>
<td>$x_{Nk}$</td>
<td></td>
<td>$y_N$</td>
</tr>
</tbody>
</table>

Input matrix

Output vector:
Why Sensitivity analysis?
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
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.
Limits of sensitivity analysis
Useless Arithmetic: Why Environmental Scientists Can't Predict the Future
by Orrin H. Pilkey and Linda Pilkey-Jarvis

Orrin H. Pilkey
Duke University, NC
<<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 $^{36}\text{Cl}$ 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.

\[\rightarrow\ldots\text{SA useless if it is instead } \sim 3,000\text{ millimetres per year.}\]
“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.
Can I lie with sensitivity analysis?
Will any sensitivity analysis do the job? Can I lie with sensitivity analysis as I can lie with statistics?

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
“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 = ?

\[ \sim \frac{3}{4} \]
OAT in 3 dimensions

Volume sphere / volume cube = ?

~ 1/2
OAT in 10 dimensions

Volume hypersphere / volume ten dimensional hypercube \approx 0.0025
OAT in k dimensions

The graph shows the volume of an n-ball inscribed in the unit hypercube as a function of the number of dimensions. The graph demonstrates how the volume decreases as the number of dimensions increases. Markers indicate the volume for different values of k:

- K=2
- K=3
- K=10
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
OAT is still the most largely used technique in SA. Out of every 100 papers with modelling & SA only 4 are ‘global’ in the sense discussed here.

In 2014 out of 1000 papers in modelling 12 have a sensitivity analysis and < 1 a global SA
Fig. 4. GSA in the different scientific domains.
Discussion points (1)

• Is the geometric argument necessary? Anyone experience in design of experiment (DOE)?

• Can OAT be justified in some cases?
The influence of the key variables should be investigated by a sensitivity analysis.

- Is something wrong about the statement above (p. 384 of EC guidelines)
Discussion points (3)

- If I keep a parameter 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 is sensitivity analysis done?
\[
\begin{array}{cccc}
  x_{11} & x_{12} & \ldots & x_{1k} \\
  x_{21} & x_{22} & \ldots & x_{2k} \\
  \ldots & \ldots & \ldots & \ldots \\
  x_{N1} & x_{N2} & \ldots & x_{Nk} \\
\end{array}
\]

Input matrix

\[
\begin{array}{c}
  y_1 \\
  y_2 \\
  \ldots \\
  y_N \\
\end{array}
\]

Output vector:
Input matrix:

- 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
Output vector:

- Just one output of interest; but $y$ could also be a vector (function of time) or a map, etc. …

- $Y$ can be plotted against any of the $x_i$
$Y$ plotted against two different factors $x_i$ and $x_j$

The values of the output on the ordinate are the same
Can I do a sensitivity analysis just looking at the plots?

Scatterplots of $y$ versus sorted factors

- Input variable $x_i$
- Input variable $x_j$
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 \( \sim E_{X \sim_i} \left( Y \mid X_i \right) \)
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

\[ V_{X_i} \left( E_{X_{-i}} (Y|X_i) \right) \]?
\[ S_i \equiv \frac{V(E(Y|X_i))}{V_Y} \]
First order sensitivity index $S_i \equiv \eta_i^2 := \frac{\text{Var}_{x_i} (\text{Ex}_{x_i} (y \mid x_i))}{\text{Var}(y)}$
First order sensitivity index:

\[ E_{x \sim i} (y \mid x_i) \]

\[ \frac{V_{x_i} (E_{x \sim i} (y \mid x_i))}{V(y)} \]
\[ V_{X_i} \left( E_{X \sim i} (Y|X_i) \right) \]

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:

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

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

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 what variance would be left (on average) if \( X_i \) could be fixed…
\[ \text{... then this...} \]

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

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

\[ \text{... must be the expected reduction in variance that would be achieved if factor } X_i \text{ could be fixed} \]
For additive models one can decompose the total variance as a sum of first order effects

$$\sum_{i} V_{X_i} \left( E_{X_{\sim i}} \left( Y | X_i \right) \right) \approx V(Y)$$

… which is also how additive models are defined
If an additive model is one where the $V$ of the output is a linear combination of the partial variances of the inputs then:

- can I guess a formula for an additive model?

- and for a non additive?
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)

\[ V_{X_i} \left( E_{X \sim i} \left( Y \big| X_i \right) \right) = V_i \]

\[ V_{X_i X_j} \left( E_{X \sim ij} \left( Y \big| X_i, X_j \right) \right) = \]

\[ = V_i + V_i + V_{ij} \]

\[ \ldots \]
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.
Variance decomposition (ANOVA)

When the factors are not independent the decomposition loses its unicity (and hence its appeal)
If fact interactions terms are awkward to handle: second order terms 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 =

\[ E_{x_{-i}} \left( V_{x_i} \left( Y \mid X_{\sim i} \right) \right) \]

= the expected variance that would be left if all factors but \( X_i \) could be fixed.
\[ S_{Ti} \equiv \frac{E(V(Y|X_{\sim i}))}{V_Y} \]
What is the shortcoming of $S_{Ti}$?
\[
\frac{V_{X_i} \left( E_{X_{\sim i}} \left( Y \mid X_i \right) \right)}{V(Y)} = S_i
\]

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

Scaled to [0,1]; first order and total order sensitivity coefficient.
Why these measures?

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

Factors prioritization

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

Fixing (dropping) non important factors

More about the settings:

- **Factor prioritisation** → \( S_i \equiv \frac{V(E(Y|X_i))}{V_Y} \)

If the cost of ‘discovering’ factors were the same for all factors which factor should I try to discover first?
**Factor fixing:** 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?

\[
S_{Ti} \equiv \frac{E(V(Y|X_{\sim i})))}{V_Y}
\]
Factor fixing is useful to achieve model simplification and ‘relevance’.
Can we use $S_i$ to fix a factor?

If $S_i = 0$ is $X_i$ a non-influential factor?
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.

$X_i$ could be influential at the second order.
Can we use $S_{Ti}$ to fix a factor?

If $S_{Ti} = 0$ is $X_i$ a non-influential factor?
Variance is always a positive number

$$E_{X \sim_i} \left( V_{X_i} \left( Y \mid X_{\sim_i} \right) \right)$$

For a mean of non-negative entries to be zero, all entries must be zero.

If $S_{Ti} = 0 \quad \Rightarrow \quad X_i$ is non-influential as there is no point in the hyperspace of the input where $x_i$ has an effect; $S_{Ti} = 0$ necessary and sufficient condition for non-influence.
Summary for variance based measures:

1. 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.
Easy to smooth and interpolate!
Summary for variance based measures:

3. The total effect is more expensive; its computational cost is \((k+1)N\) where \(N\) is one of the order of one thousand (unless e.g. using emulators …).
How about other methods?
Monte Carlo filtering
When to use Monte Carlo Filtering?

When we are interested not in the precise value of the output $y$ but on whether or not this value is ‘permitted’ or forbidden.
Partitioning $y$ impose a partitioning on each of the $x_i$’s
Taking one column at a time I can split the sample of each factor into two subsets.
Monte Carlo filtering

\[ (X_i | B) \]

\[ B = \text{OK} \]

\[ \overline{B} = \text{not OK} \]
Monte Carlo filtering

Step by step:

- Classifying simulations as either $B$ or $\bar{B}$. This allows distinguishing two sub-sets for each $X_i$: $(X_i|B)$ and $(X_i|\bar{B})$

- The Smirnov two-sample test (two-sided version) is performed for each factor independently, analyzing the maximum distance between the cumulative distributions of the $B$ and $\bar{B}$ sets.
Monte Carlo filtering

Cumulative distributions

- Runs of $B$
- All runs
- Runs of $\overline{B}$

$d_{\text{max}}$
How to generate the random sample?

We use quasi random sequences developed by I.M. Sobol’
An $LP_\tau$ sequence
Sobol' sequences of quasi-random points
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.

Variance based measures are:
- well scaled,
- concise,
- easy to communicate.

Further
- $S_i$ reduces to squared standard regression coefficients for linear model.
- $S_{Ti}$ detect and describe interactions and
- Becomes a screening test at low sample size

Secrets of sensitivity analysis
First secret: The most important question is the question.

Corollary 1: Sensitivity analysis is not “run” on a model but on a model once applied to a question.
First secret: The most important question is the question.

Corollary 2: The best setting for a sensitivity analysis is one when one wants to prove that a question cannot be answered given the model. It is better to be in a setting of falsification than in one of confirmation (Oreskes et al., 1994).

[Normally the opposite is the case]
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/model which can be treated meaningfully.

[Often the love for the 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
Discussion point

Why should I not run a sensitivity analysis where each factor has a 5% uncertainty
END

Twitter:
@andreasaltelli