

Sensitivity Analysis

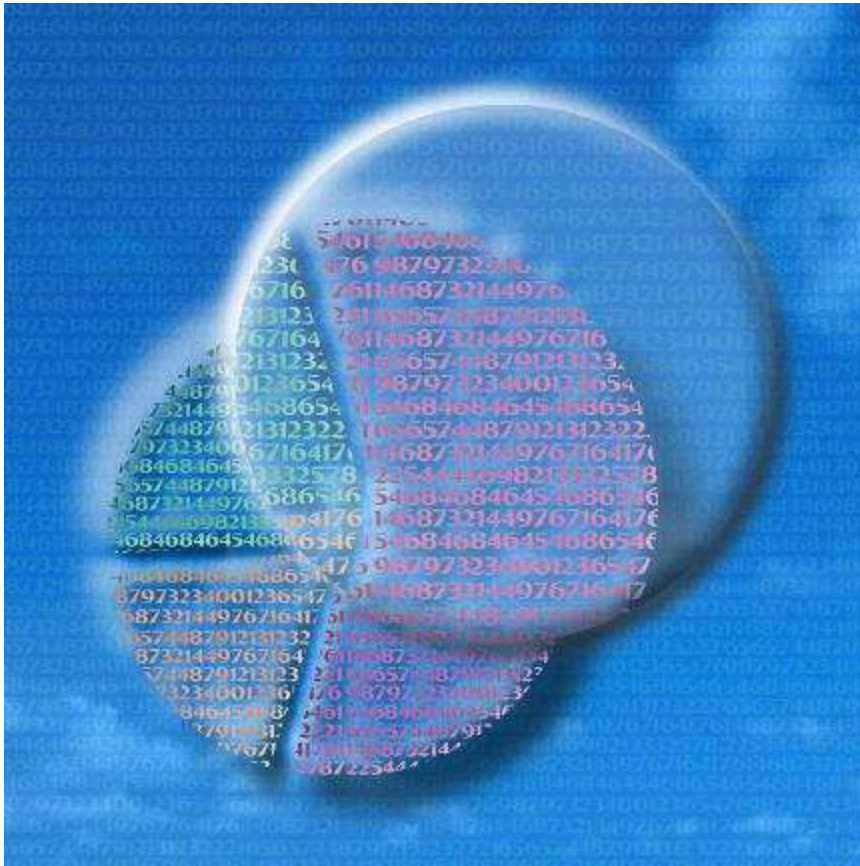
Andrea Saltelli

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

PhD-course

Numbers for policy: Practical
problems in quantification

Bergen, March 13–17, 2017



CAETERIS ARE
NEVER PARIBUS

Tweets by @AndreaSaltelli



andrea saltelli

@AndreaSaltelli

Sign and donate. What these people are doing is unique. twitter.com/Jeroen_vdSluijs



24/11



andrea saltelli

@AndreaSaltelli

Lovely (also in the sense of 'of love') piece by an Italian scholar [@robertocalasso](#):

[nybooks.com/articles/2016/...](http://nybooks.com/articles/2016/)



Embed

[View on Twitter](#)

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

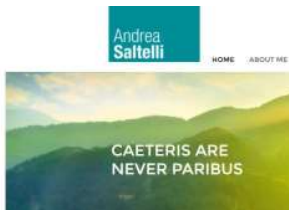
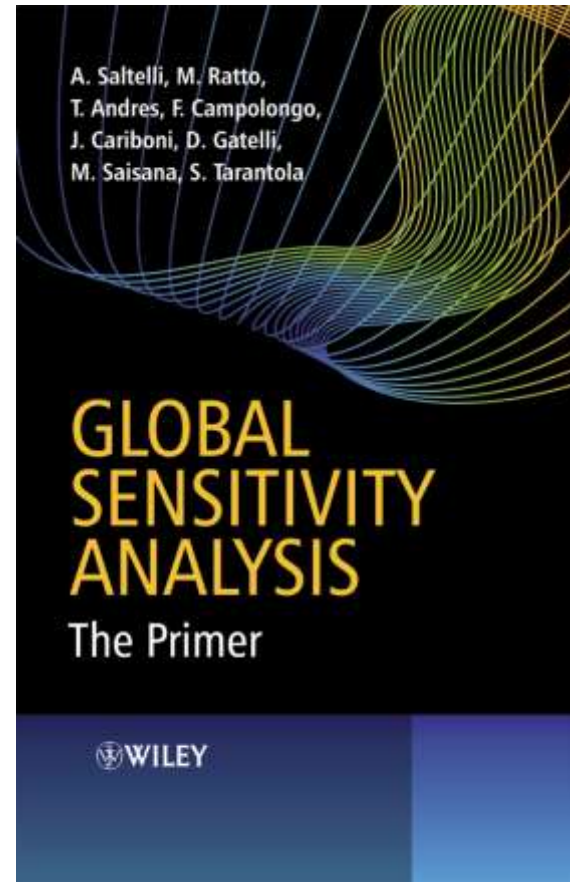
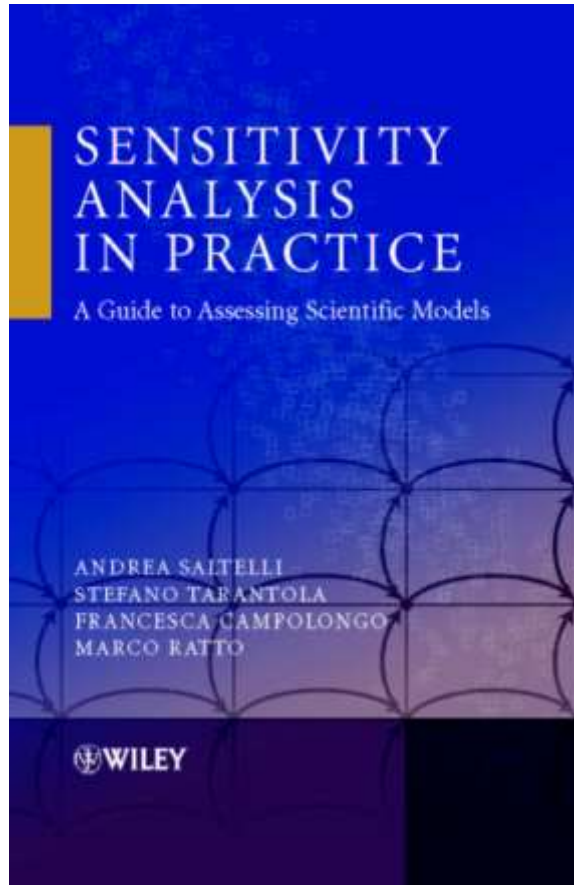


= more material on my web site



= discussion time

Sensitivity analysis books available on LibGen



What is sensitivity
analysis?

Definitions

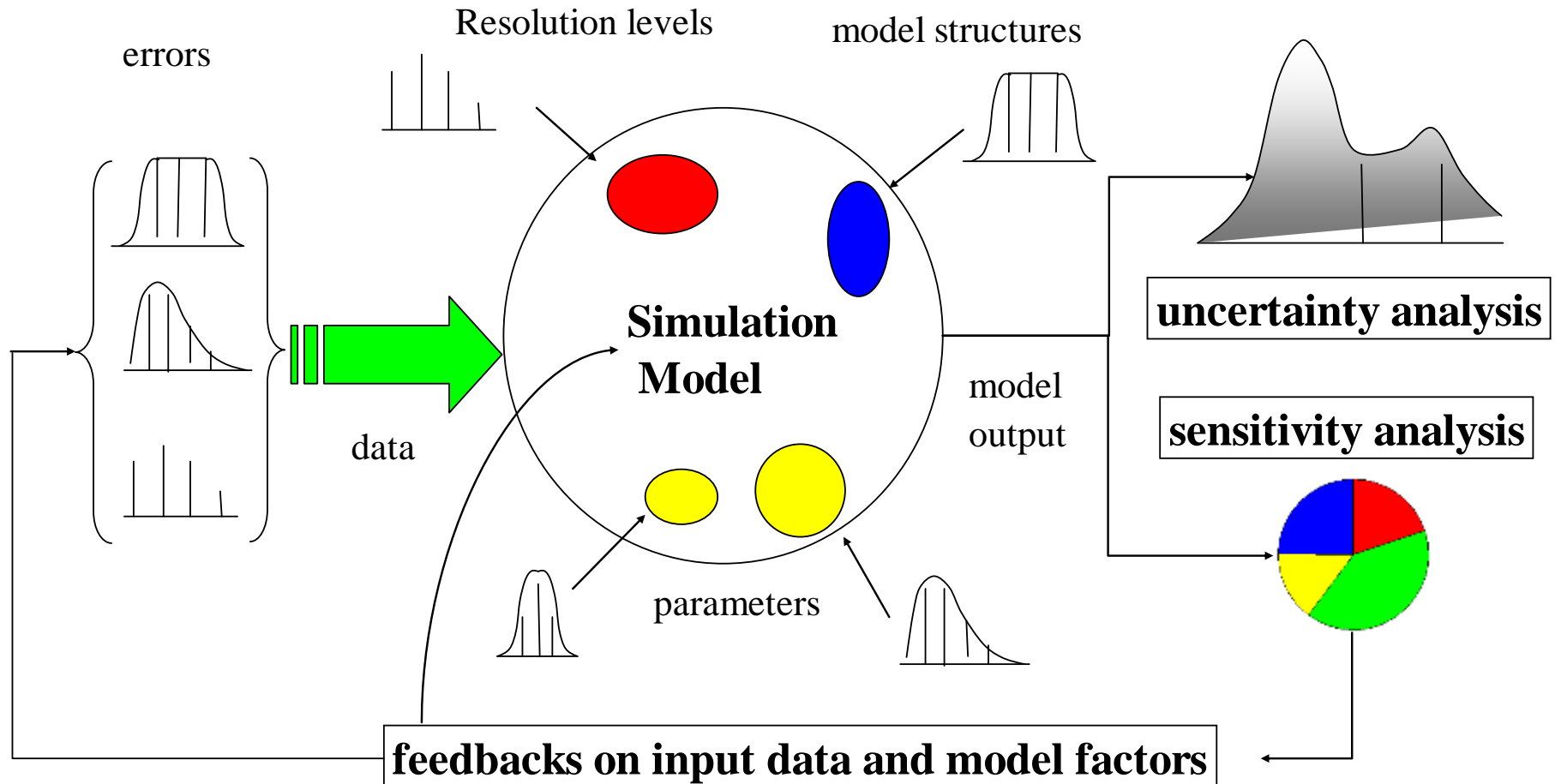
Uncertainty analysis: Focuses on just quantifying the uncertainty in model output

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”

Saltelli A., 2002, Sensitivity Analysis for Importance Assessment, Risk Analysis, 22 (3), 1–12.

An engineer's vision of UA, SA



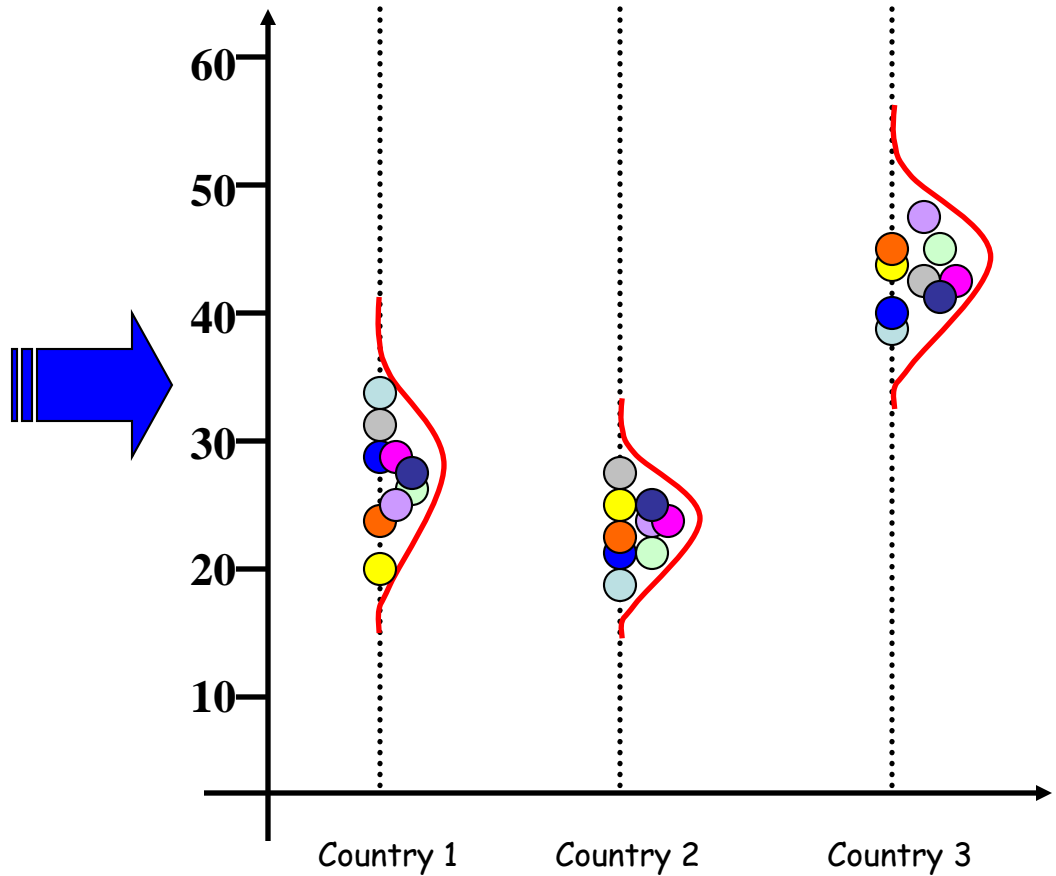
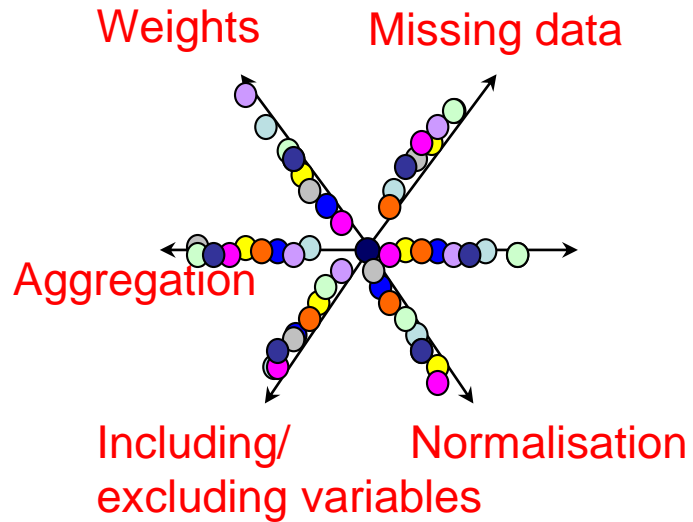
One can sample more than just factors

One can sample modelling assumptions

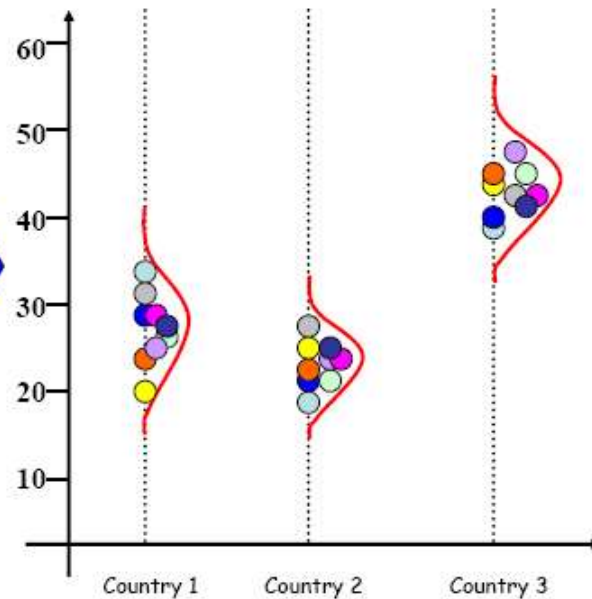
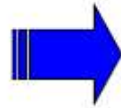
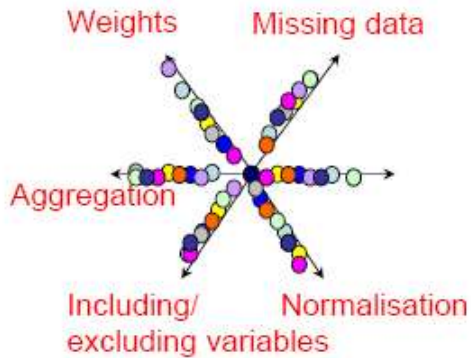
Example: The output is a composite indicator

Assumption	Alternatives
Number of indicators	<ul style="list-style-type: none">▪ all six indicators included or one-at-time excluded (6 options)
Weighting method	<ul style="list-style-type: none">▪ original set of weights,▪ factor analysis,▪ equal weighting,▪ data envelopment analysis
Aggregation rule	<ul style="list-style-type: none">▪ additive,▪ multiplicative,▪ Borda multi-criterion

Space of alternatives

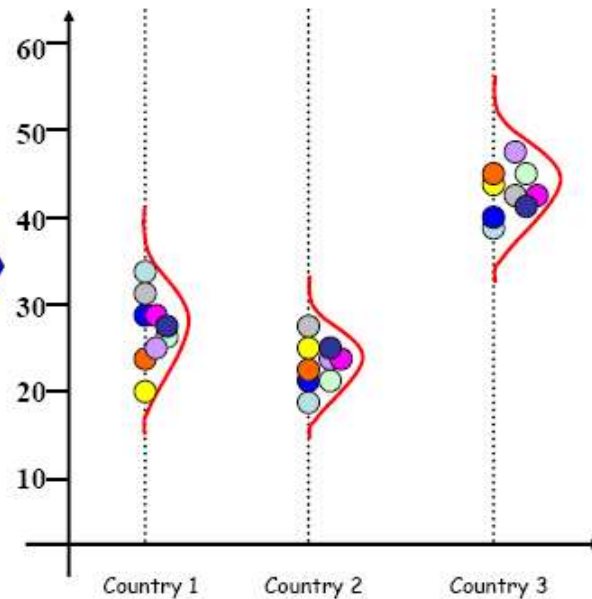
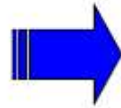
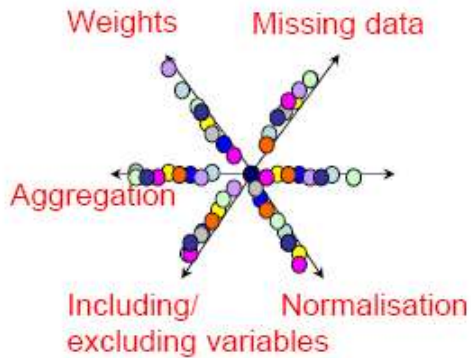


Space of alternatives



Is this an uncertainty analysis or a sensitivity analysis?

Space of alternatives



If I did a sensitivity analysis what information would I obtain?

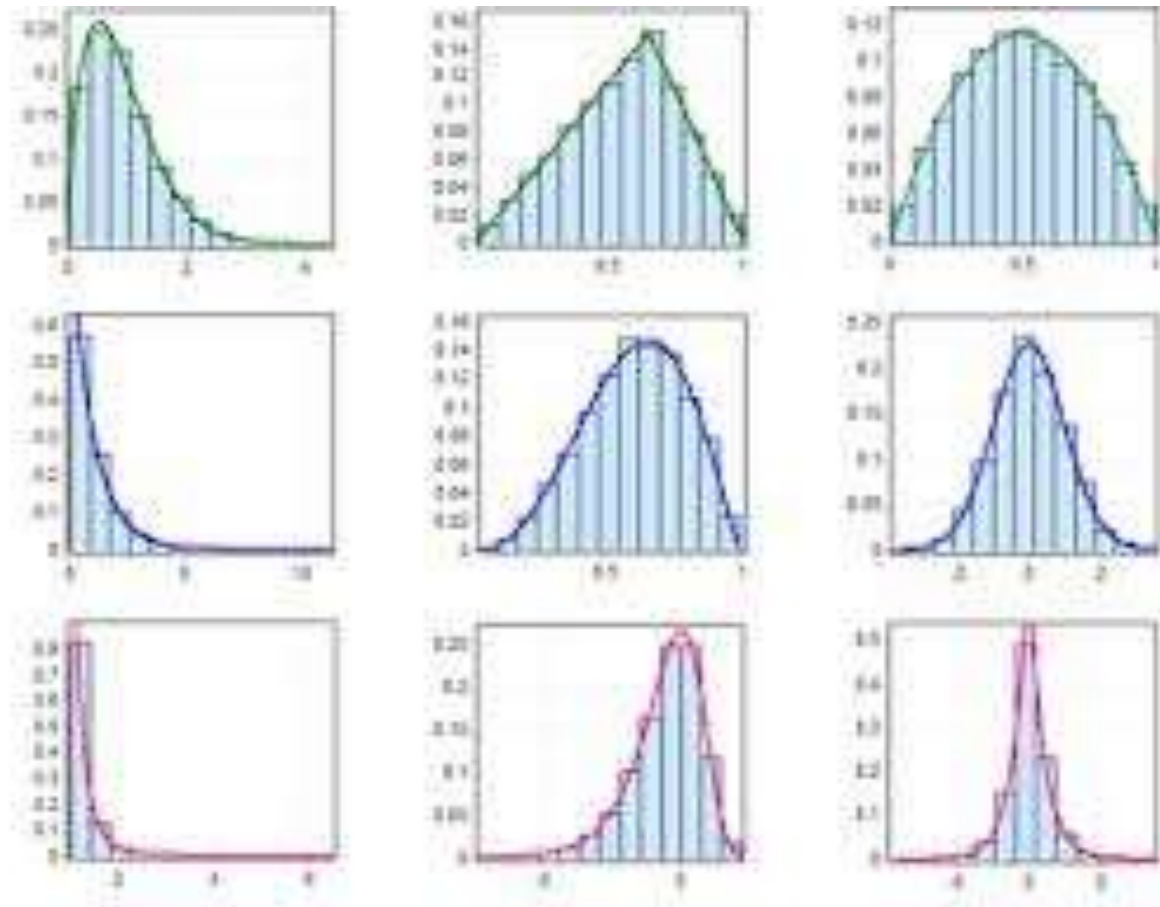
Sample matrix for
uncertainty and
sensitivity analysis

$$\begin{array}{cccc} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots \\ x_{N1} & x_{N2} & \dots & x_{Nk} \end{array}$$

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.

$$\begin{array}{cccc}
 x_{11} & x_{12} & \dots & x_{1k} \\
 x_{21} & x_{22} & \dots & x_{2k} \\
 \dots & \dots & \dots & \dots \\
 x_{N1} & x_{N2} & \dots & x_{Nk}
 \end{array}$$



Model results:

Each entry is the
error-free result of
the model run.

 y_1 y_2 \dots y_N

$$\begin{array}{ccccccc}
 x_{11} & x_{12} & \dots & x_{1k} & & y_1 \\
 x_{21} & x_{22} & \dots & x_{2k} & & y_2 \\
 \dots & \dots & \dots & \dots & & \dots \\
 x_{N1} & x_{N2} & \dots & x_{Nk} & & y_N
 \end{array}$$

Input matrix

Output vector:

In the simplest case y could be a function of – a simple mathematical expression of – the x_1, x_2, \dots, 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, \dots, x_k

Why Sensitivity analysis?

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_bulletin_010906.pdf, pp. 16–17, accessed December 2015.

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



The screenshot shows the 'Better Regulation' section of the European Commission's website. The header features the European Commission logo and the title 'Better Regulation'. A navigation menu on the left includes links to Home, REFIT, Stakeholder consultations, Roadmaps / Inception Impact Assessments, Impact Assessment, Evaluation, Regulatory Scrutiny Board, and Guidelines. The 'Guidelines' section is expanded, showing links to 'Better Regulation Guidelines', 'Better Regulation "Toolbox"', and 'Key documents'. The main content area is titled 'Better Regulation Guidelines' and contains text explaining the guidelines' purpose and structure. It mentions that the guidelines cover the whole policy cycle and are based on public consultation exercises. A list of related documents is provided, including 'Public consultation on the revision of the Commission's Impact Assessment Guidelines', 'Stakeholder Consultation Guidelines', and 'Consultation on the draft Commission Evaluation Policy Guidelines'. On the right side, there are sections for 'Share', 'Search', 'Stay connected' (with social media links), 'Latest documents' (listing '10/05/2015 - Better Regulation Package'), 'Help us improve', and a 'Find what you wanted?' section with a search bar and a 'Send' button.

European Commission

Better Regulation

European Commission > Better Regulation > Guidelines

Home

REFIT

Stakeholder consultations

Roadmaps / Inception Impact Assessments

Impact Assessment

Evaluation

Regulatory Scrutiny Board

Guidelines

Better Regulation Guidelines

Better Regulation "Toolbox"

Key documents

Better Regulation Guidelines

These guidelines explain what Better Regulation is and how it should be applied in the day to day practices when preparing new initiatives and proposals or managing existing policies and legislation.

They cover the whole policy cycle, from policy preparation and adoption to implementation and application, to evaluation and revision of EU law. For each of these phases there are a number of Better Regulation principles, objectives, tools and procedures to make sure that the EU has the best regulation possible. These relate to planning, impact assessment, stakeholder consultation, implementation and evaluation.

The [Better Regulation Guidelines](#) are structured into chapters which cover each of the instruments of the law-making process. The corresponding [toolbox](#) gives more detailed and technical information.

Better Regulation Guidelines are based on the outcomes of public consultation exercises carried out in 2013 and 2014.

- Public consultation on the revision of the Commission's Impact Assessment Guidelines
- Stakeholder Consultation Guidelines
- Consultation on the draft Commission Evaluation Policy Guidelines

Share

Search

Stay connected

Facebook Twitter YouTube

Blog

Latest documents

- 10/05/2015 - Better Regulation Package

Help us improve

Find what you wanted?

Yes No

What were you looking for?

Any suggestions?

Send

Last update: 11/09/2015 | [Legal notice](#) | [Cookies](#) | [Contact](#) | [Search](#) | [Top](#)

Source: IA Toolbox, p. 391



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.

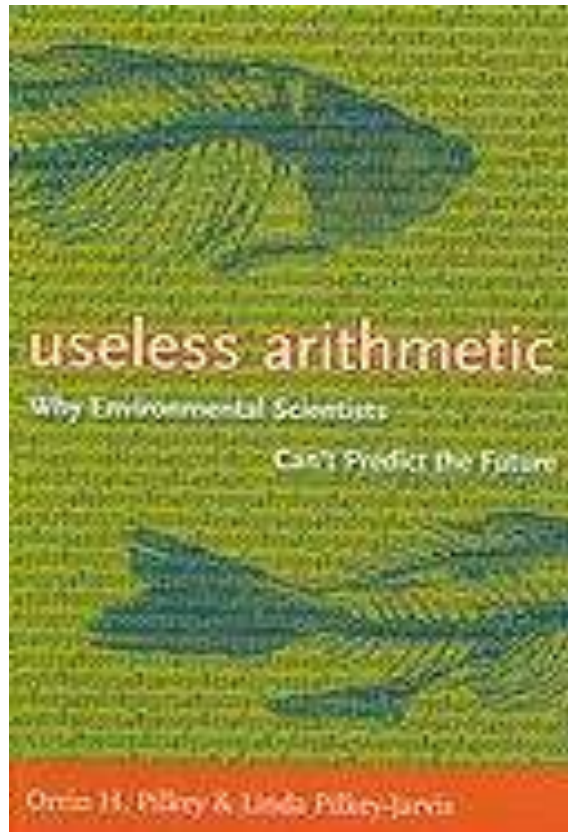
Limits of sensitivity analysis

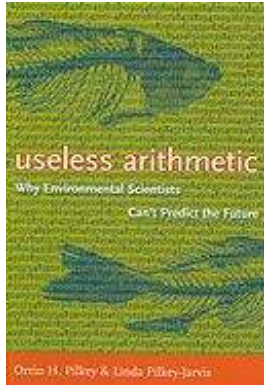


Orrin H. Pilkey
Duke University,
NC

Useless Arithmetic: Why Environmental Scientists Can't Predict the Future

by Orrin H. Pilkey and Linda
Pilkey-Jarvis

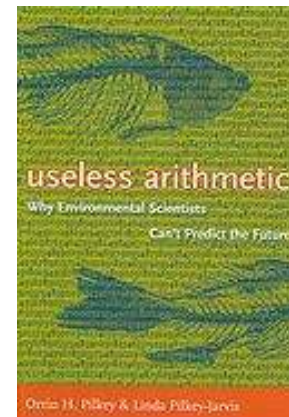


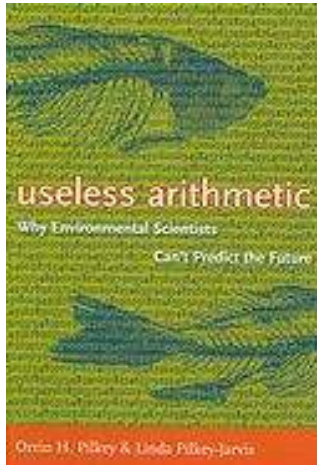


<<It is important, however, to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature.

[...] If the model is wrong or if it is a poor representation of reality, determining the sensitivity of an individual parameter in the model is a meaningless pursuit.>>

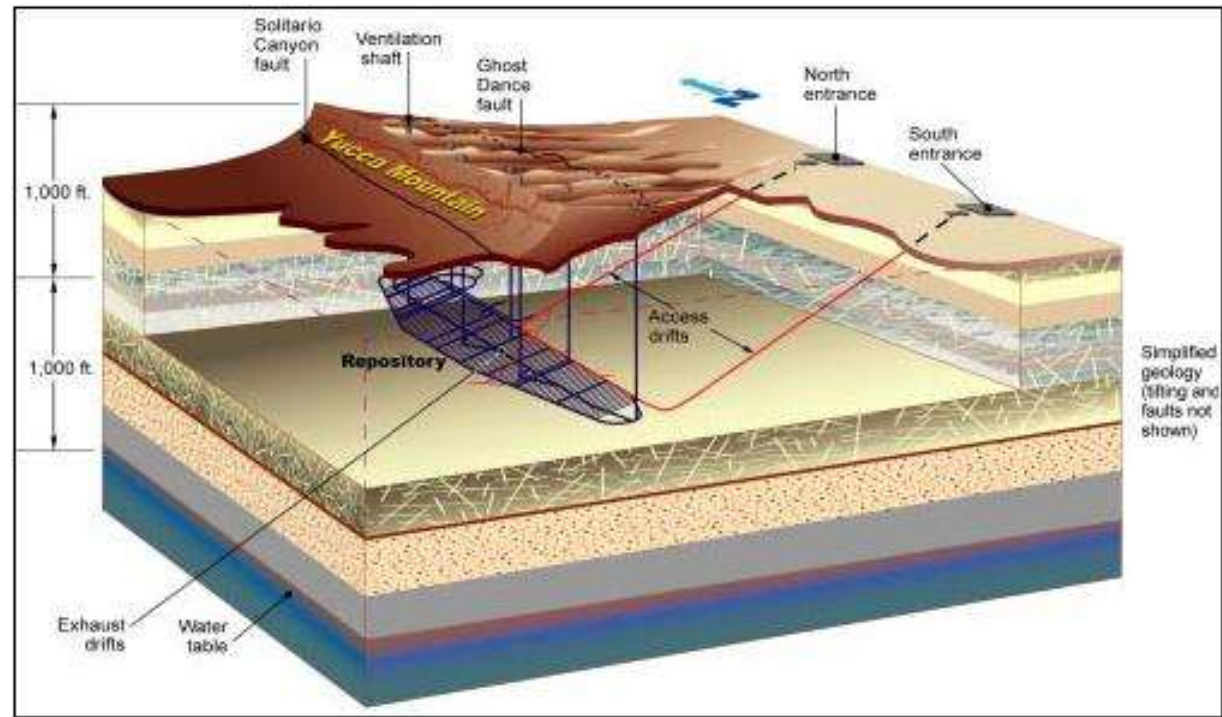
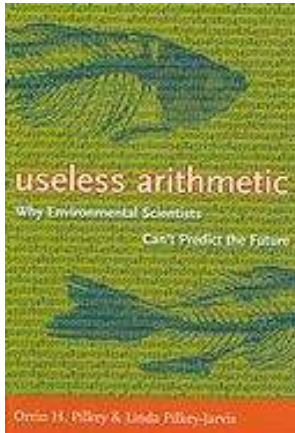
One of the examples discussed concerns the **Yucca Mountain** repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis. TSPA is Composed of 286 sub-models.





TSPA (like any other model)
relies on assumptions → one is
the low permeability of the
geological formation → long
time for the water to percolate
from surface to disposal.



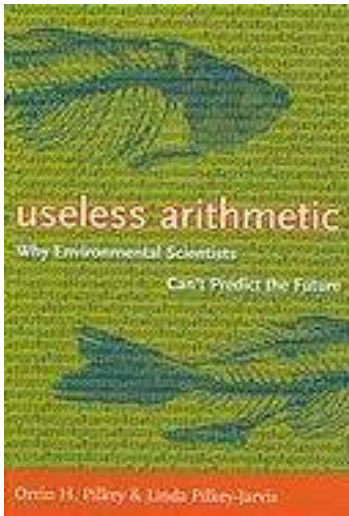


The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter (the ^{36}Cl story)

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'



Robert K. Merton

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);

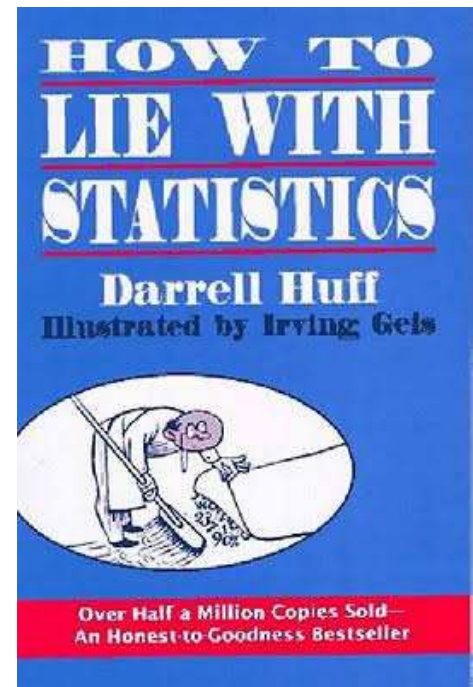
Universalism – according to which claims to truth are evaluated in terms of universal or impersonal criteria, and not on the basis of race, class, gender, religion, or nationality;

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?



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

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

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

Source: IA Toolbox, p. 391



Better Regulation "Toolbox"

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

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

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

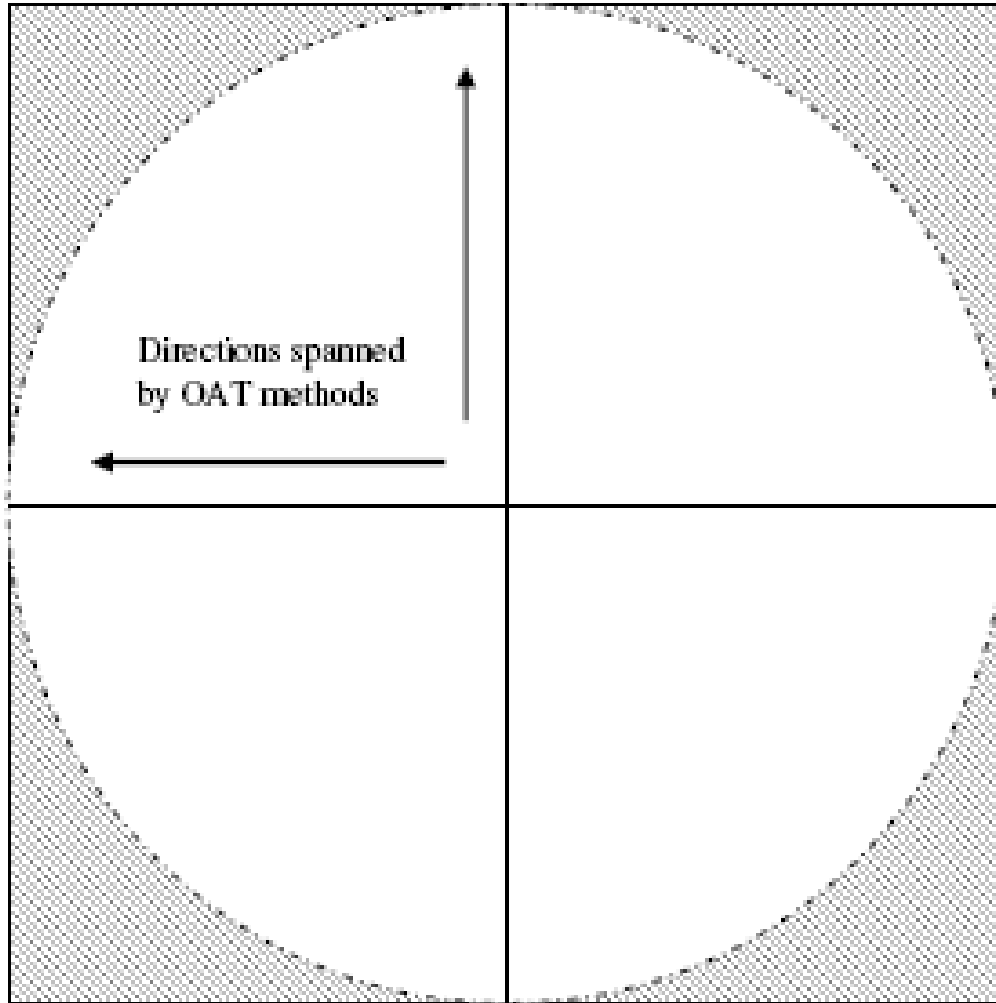
https://www.whitehouse.gov/omb/circulars_a004_a-4/



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

Because it is a bad idea!

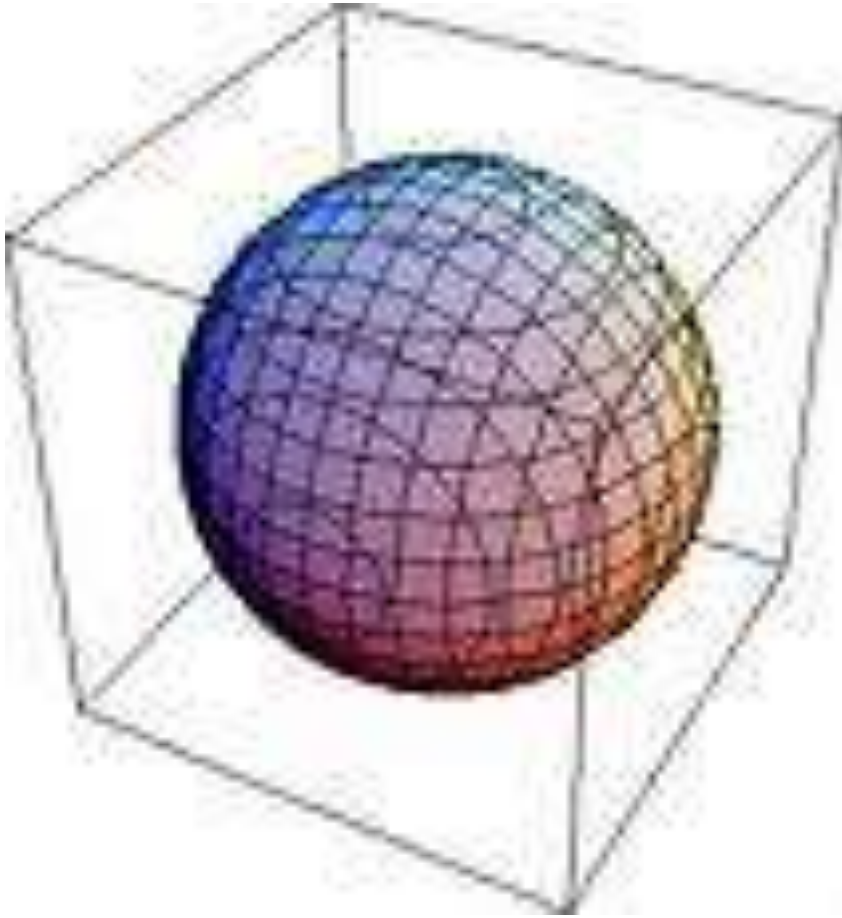
OAT in 2 dimensions



Area circle / area
square = ?

~ 3/4

OAT in 3 dimensions



Volume sphere /
volume cube =?

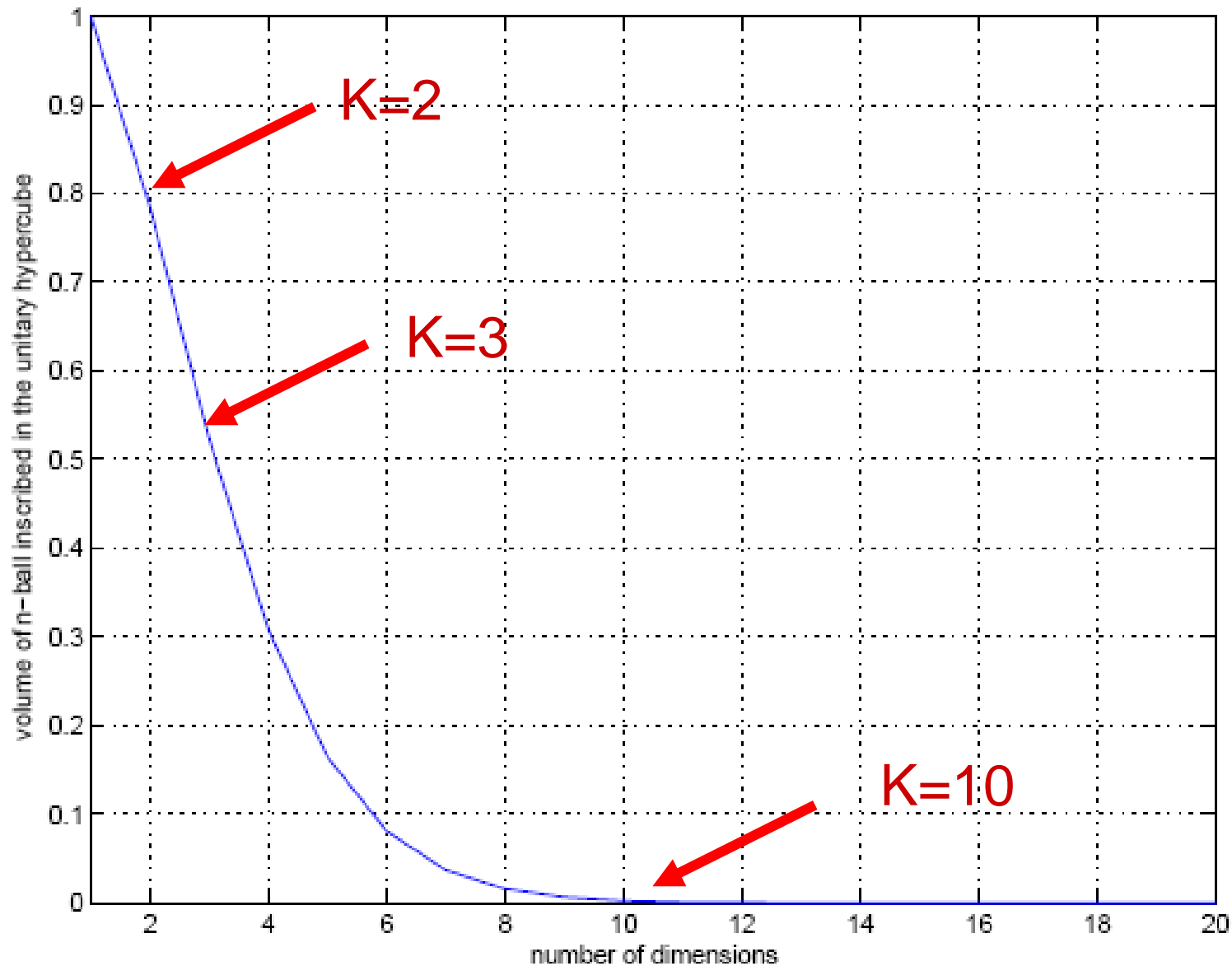
~ 1/2

OAT in 10 dimensions

Volume hypersphere / volume
ten dimensional hypercube ~ 0.0025



OAT in k dimensions



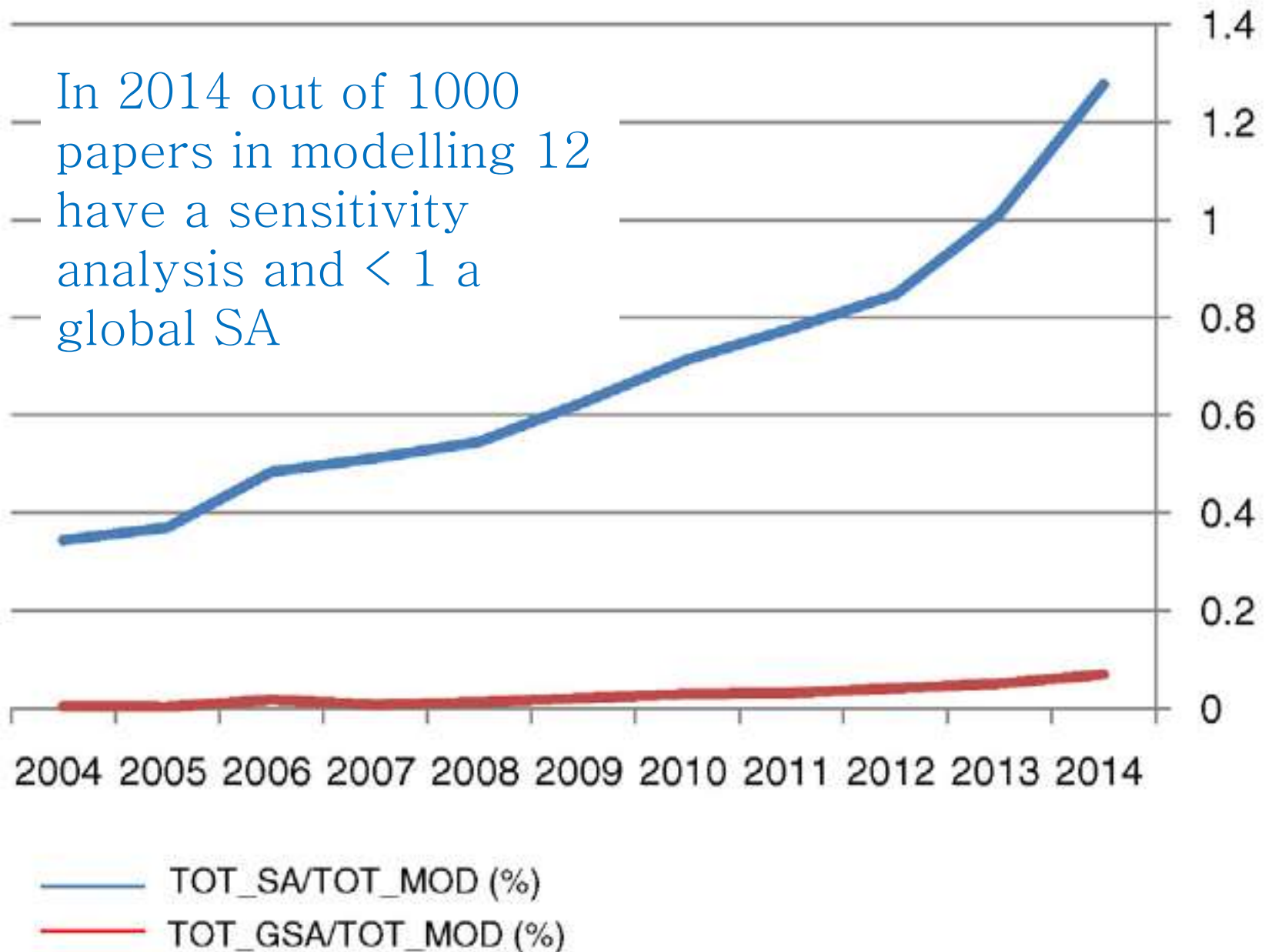
Bottom-line: once a sensitivity analysis is done via OAT there is no guarantee that either uncertainty analysis (UA) or sensitivity analysis (SA) is any good:

- ➔ 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.

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.2016.02.133>

In 2014 out of 1000
papers in modelling 12
have a sensitivity
analysis and < 1 a
global SA



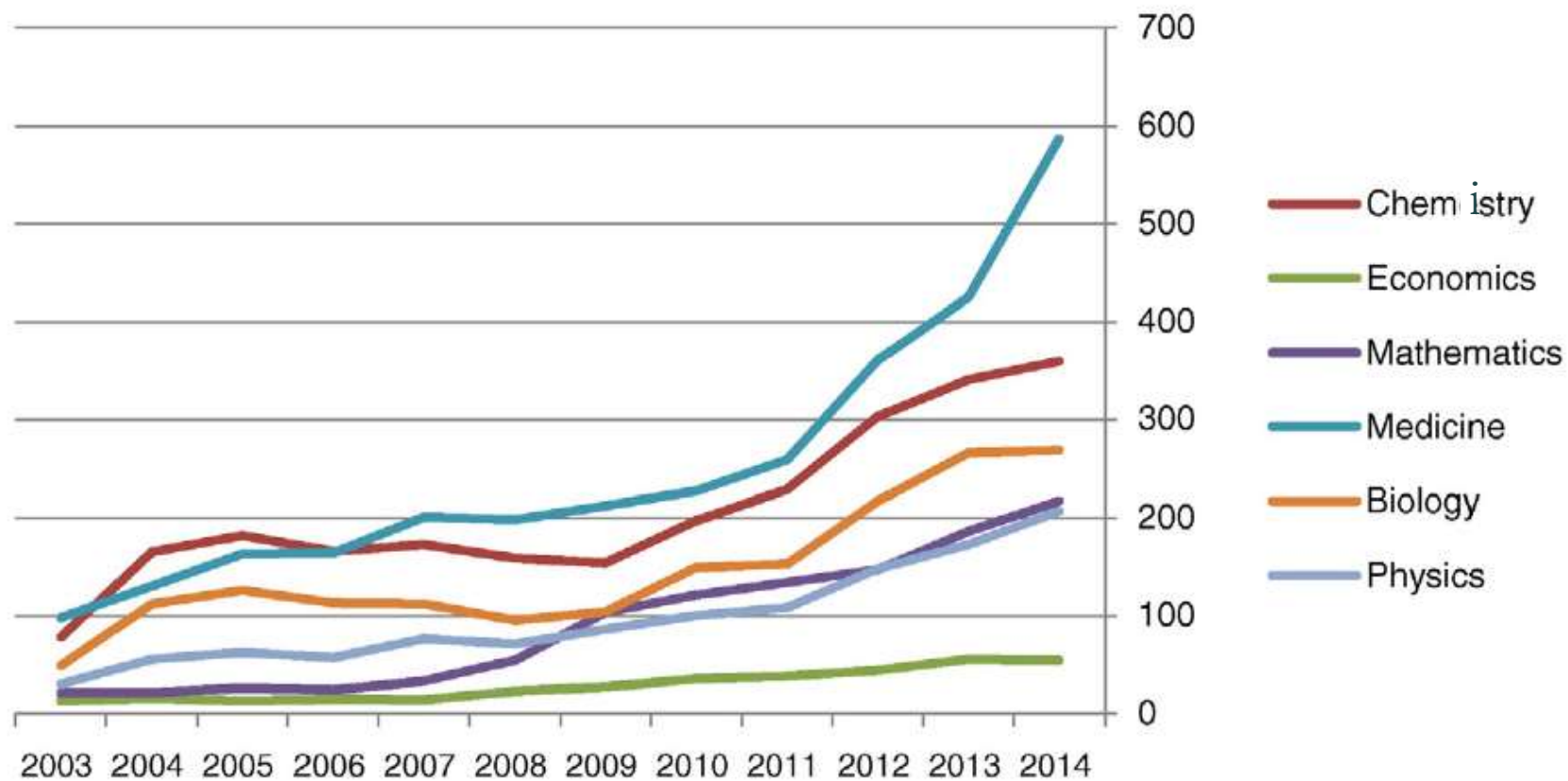


Fig. 4. GSA in the different scientific domains.



Discussion points (1)



- Is the geometric argument necessary? Anyone experience in design of experiment (DOE)?
- Can OAT be justified in some cases?

Discussion points (2)



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?

$$\begin{array}{cccc}
 x_{11} & x_{12} & \dots & x_{1k} \\
 x_{21} & x_{22} & \dots & x_{2k} \\
 \dots & \dots & \dots & \dots \\
 x_{N1} & x_{N2} & \dots & x_{Nk}
 \end{array}$$

Input matrix

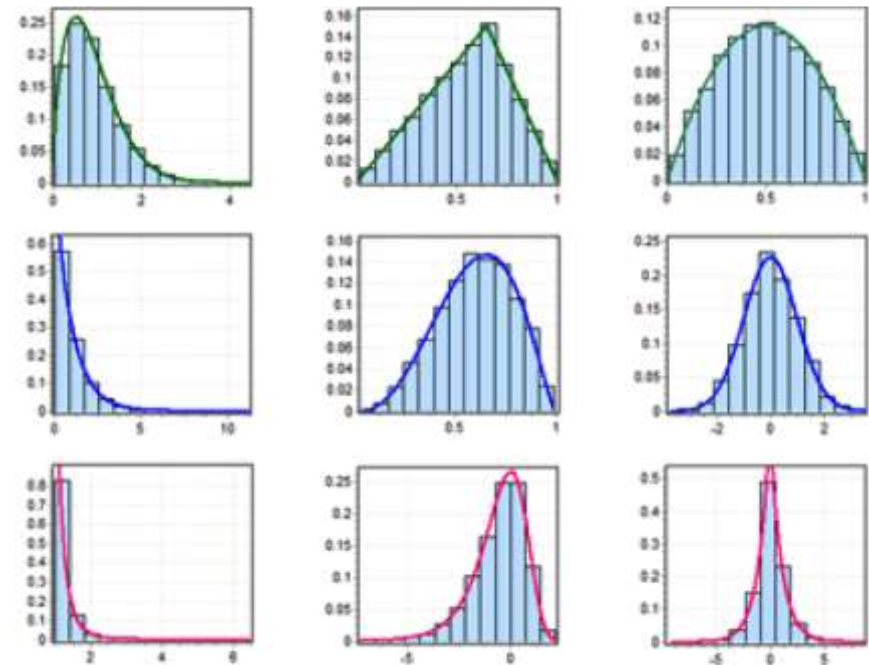
$$\begin{array}{c}
 y_1 \\
 y_2 \\
 \dots \\
 y_N
 \end{array}$$

Output vector:

$$\begin{array}{cccc}
 x_{11} & x_{12} & \dots & x_{1k} \\
 x_{21} & x_{22} & \dots & x_{2k} \\
 \dots & \dots & \dots & \dots \\
 x_{N1} & x_{N2} & \dots & x_{Nk}
 \end{array}$$

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

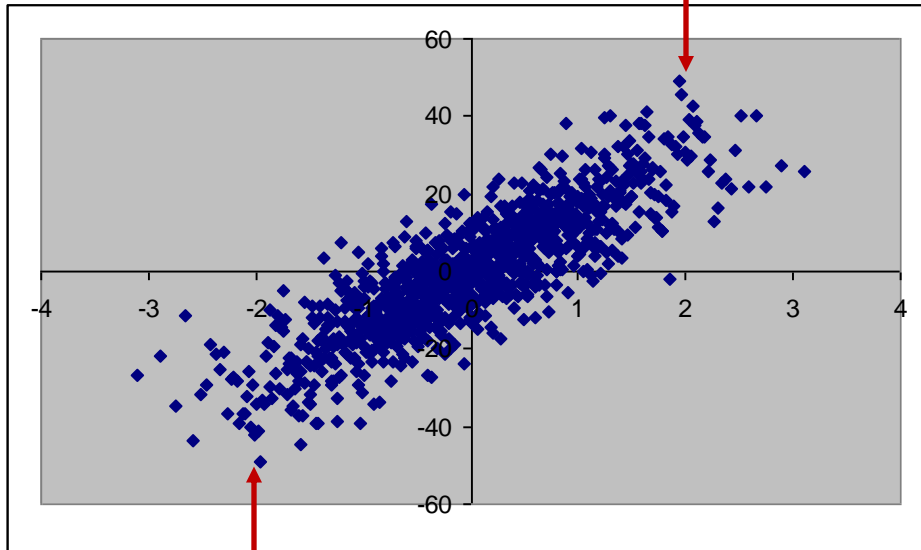
$$\begin{array}{ccccccc}
 x_{11} & x_{12} & \dots & x_{1k} & \longrightarrow & y_1 \\
 x_{21} & x_{22} & \dots & x_{2k} & \longrightarrow & y_2 \\
 \dots & \dots & \dots & \dots & & \dots \\
 x_{N1} & x_{N2} & \dots & x_{Nk} & \longrightarrow & y_N
 \end{array}$$

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

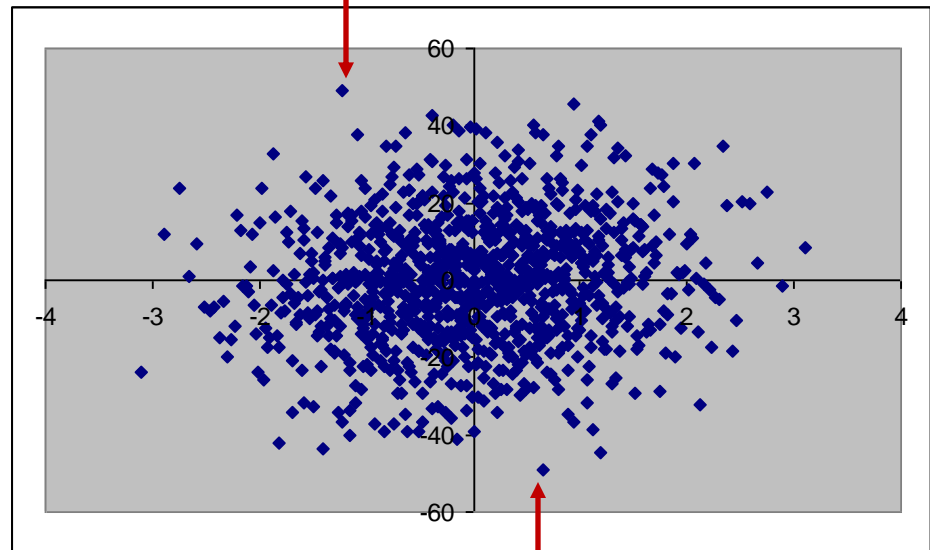
Y plotted against two different factors x_i and x_j

Output variable



Input variable x_i

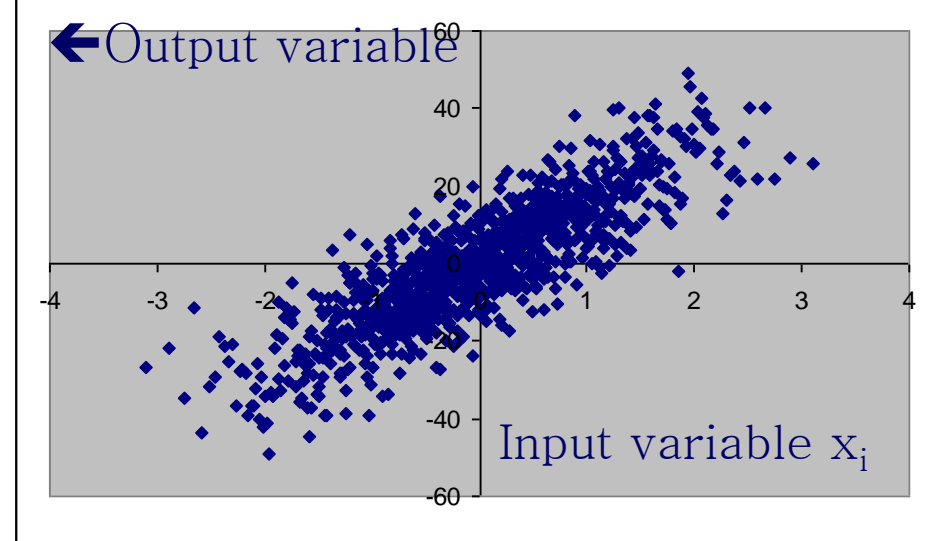
Output variable



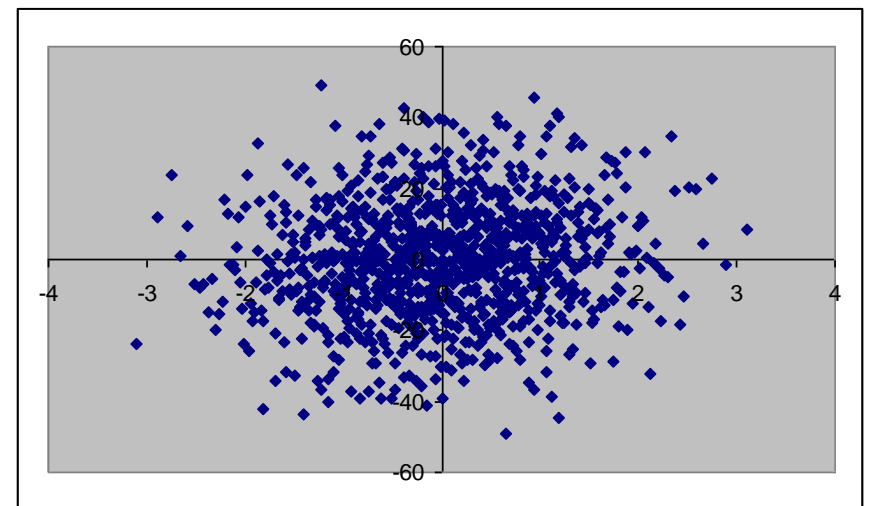
Input variable x_j

The values of the output on the ordinate are the same

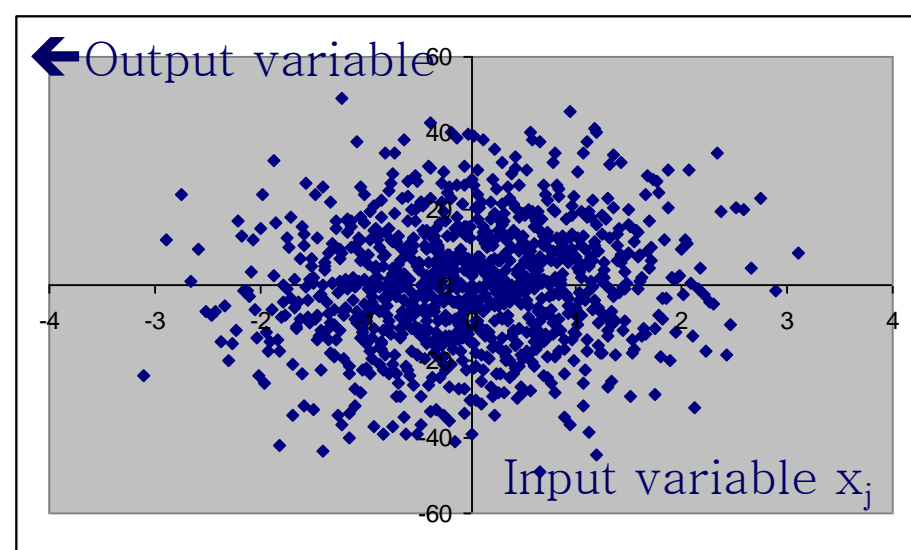
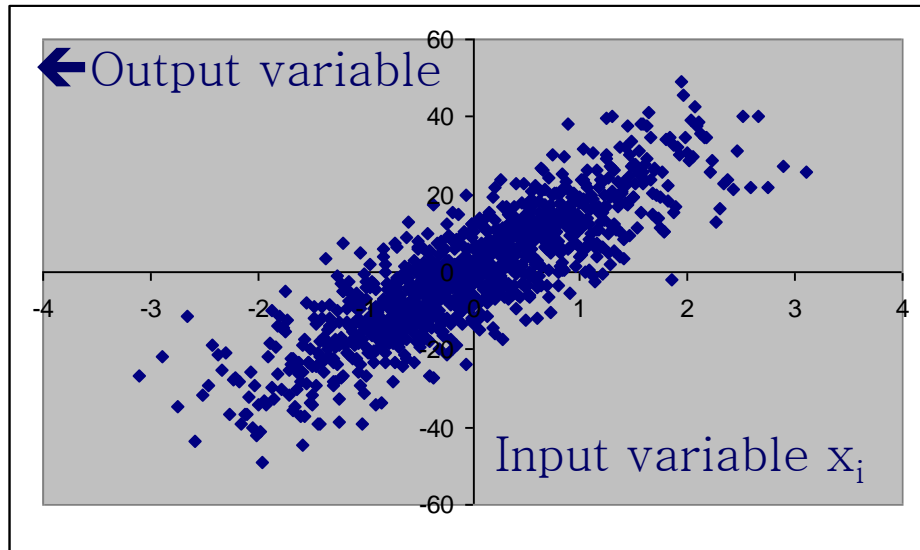
Can I do a
sensitivity analysis
just looking at the
plots?



Output variable →

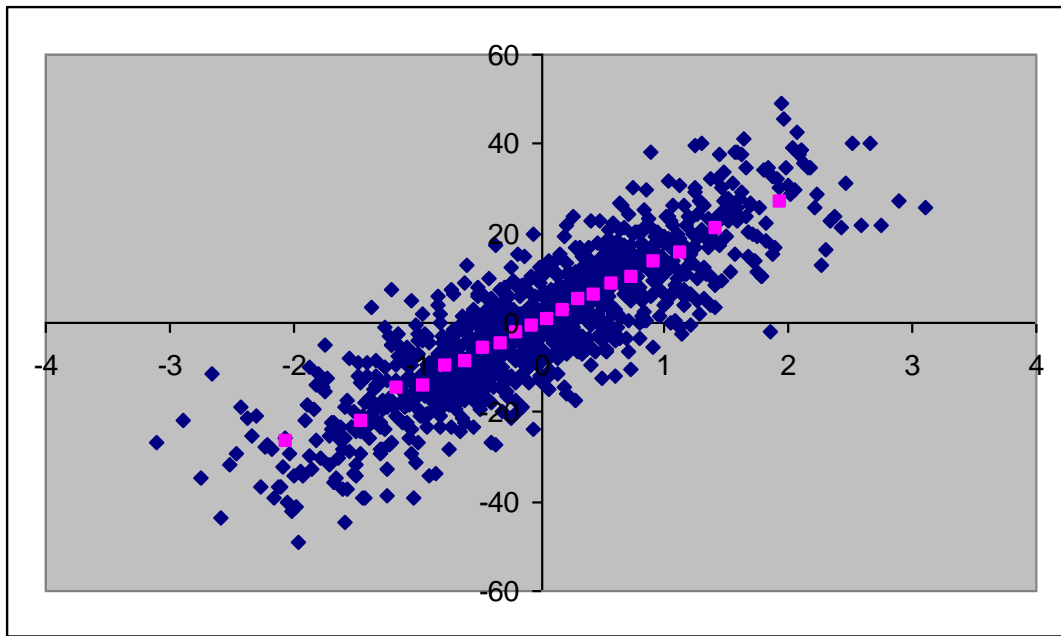


Scatterplots of y versus
sorted factors



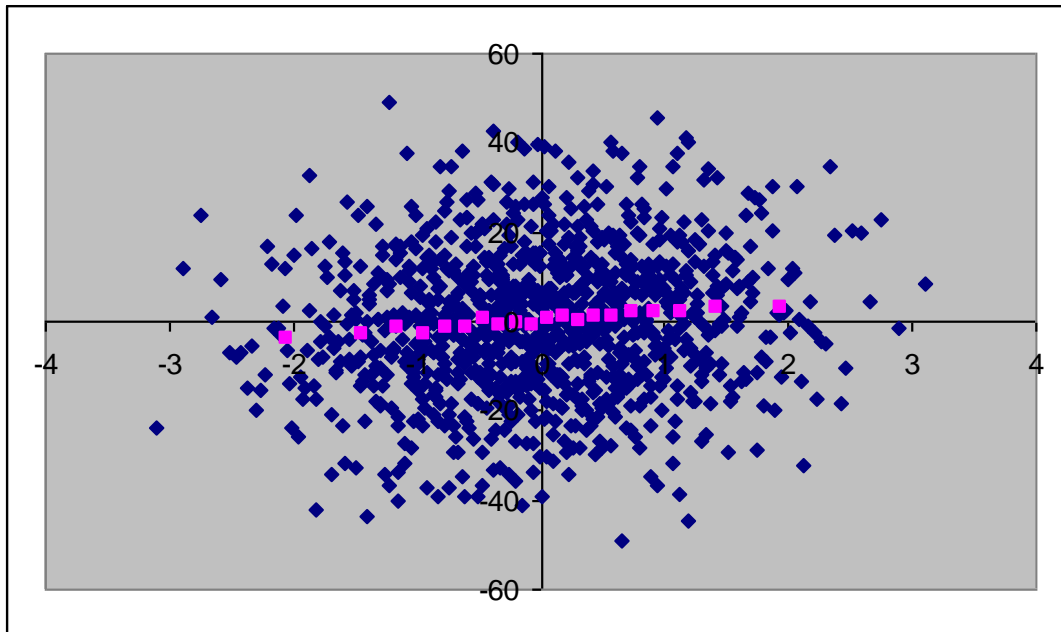
Which factor is more important?

Why?

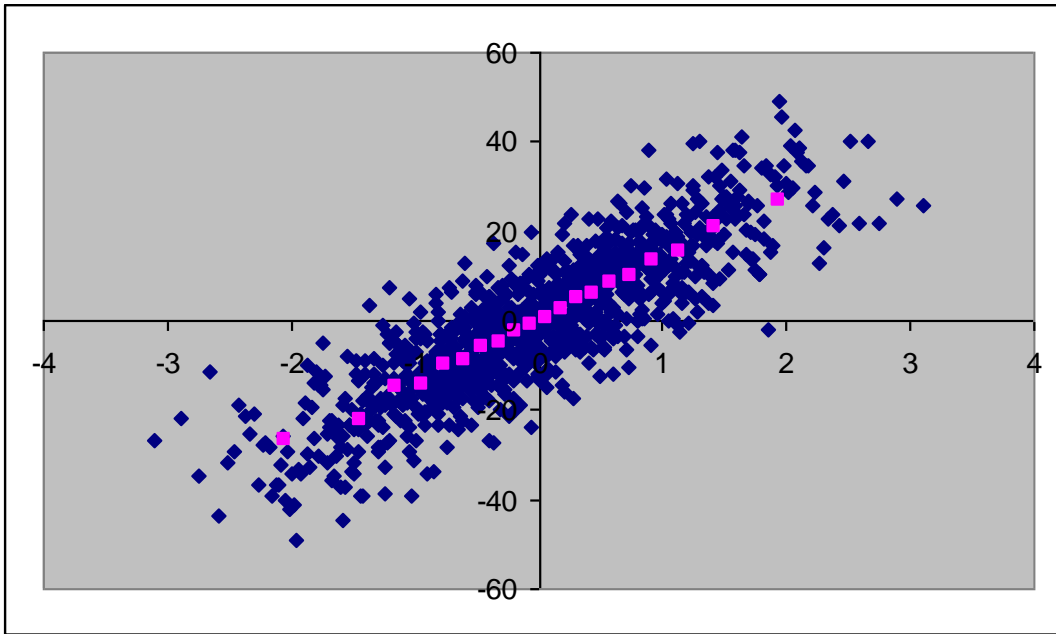


~1,000 blue
points

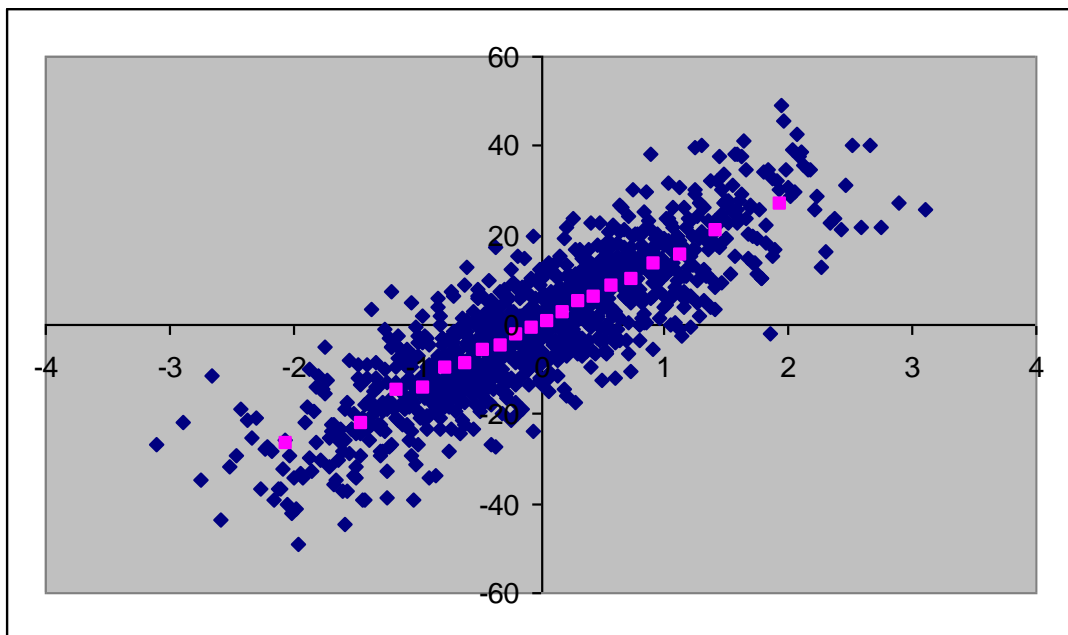
Divide them
in 20 bins of
~ 50 points



Compute the
bin's average
(pink dots)

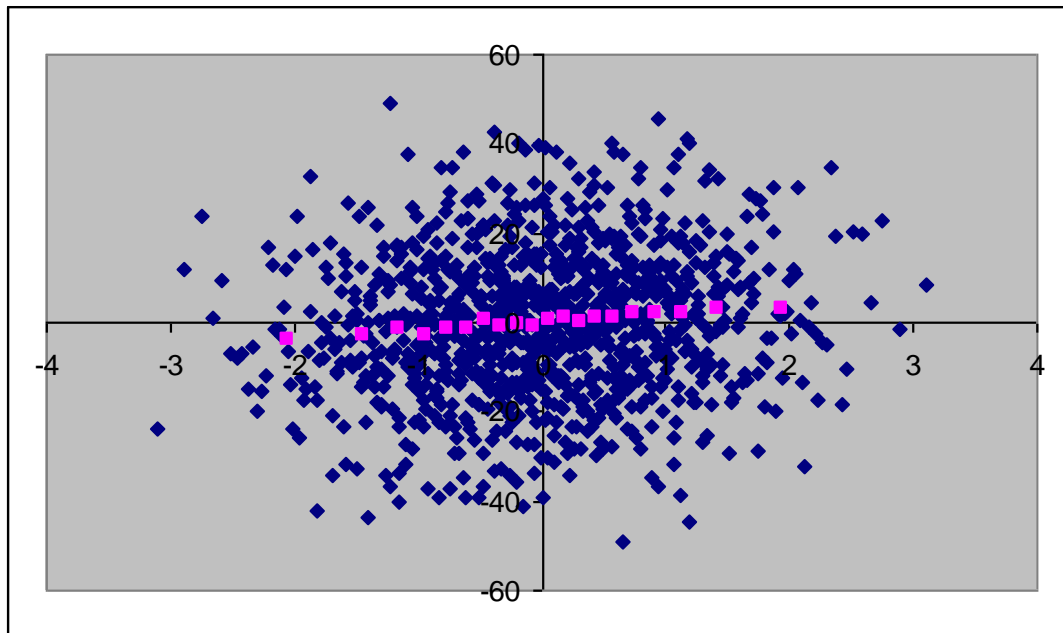
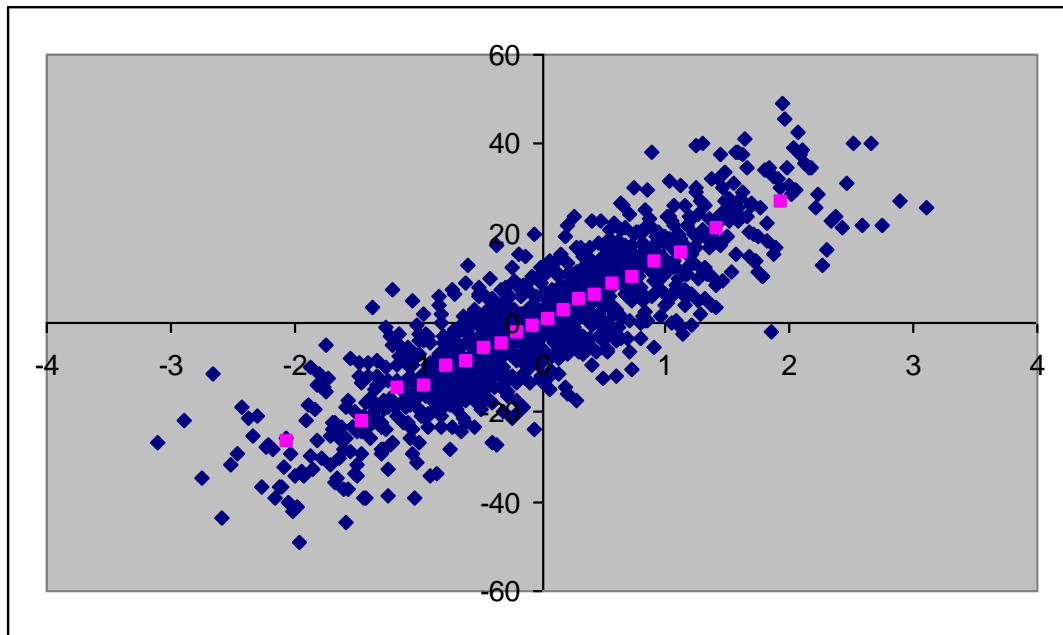


Each pink point is $\sim E_{\mathbf{X}_{\sim i}}(Y|X_i)$



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

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



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

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

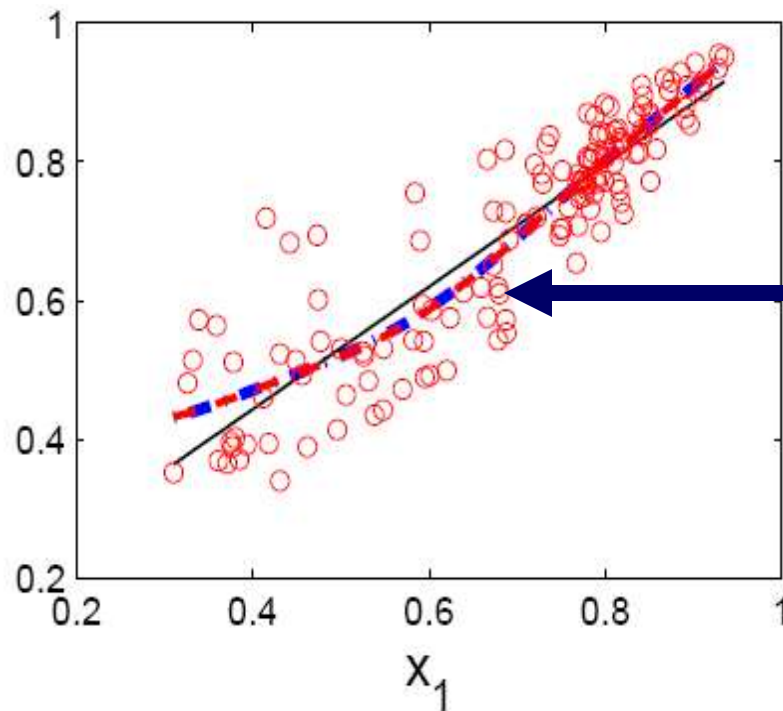
Pearson's correlation
ratio

Smoothed curve

$$S_i \equiv \eta_i^2 := \frac{V_{x_i} (\mathbf{E}_{\mathbf{x}_{\sim i}} (y \mid x_i))}{V(y)}$$

First order sensitivity index

Unconditional
variance



Smoothed curve

$$\mathbf{E}_{\mathbf{x} \sim i} (y \mid x_i)$$

First order sensitivity
index:

$$\frac{V_{x_i} (\mathbf{E}_{\mathbf{x} \sim i} (y \mid x_i))}{V(y)}$$

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

First order effect, or top marginal
variance=

= the expected reduction in variance that
would be achieved if factor X_i could be
fixed.

Why?

We need first to prove that

$$V(Y) = E(Y^2) - E^2(Y)$$

Because:

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

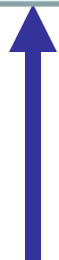
Easy to prove using $V(Y) = E(Y^2) - E^2(Y)$

Because:

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

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

$$= V(Y)$$



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

... then this ...



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

... 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

$$\sum_i V_{X_i} \left(E_{\mathbf{X}_{\sim i}} (Y | X_i) \right) \approx V(Y)$$

... which is also how additive models are defined

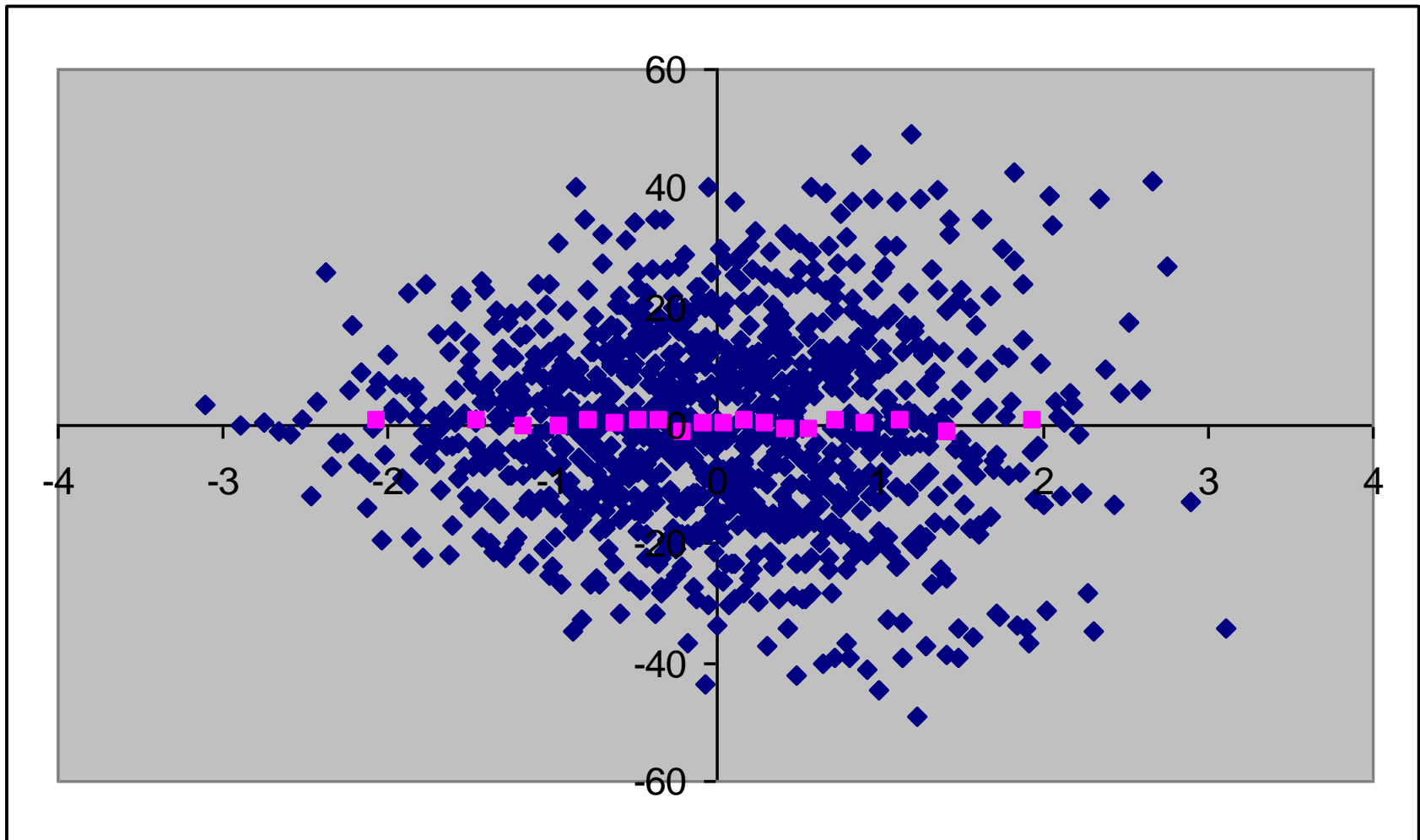
If an additive model is one where the variance V of the output is a linear combination of the partial variances of the inputs then:

- can I guess a formula for an additive model $f(X_1, X_2, X_3, \dots)$?
- and for a non additive one?

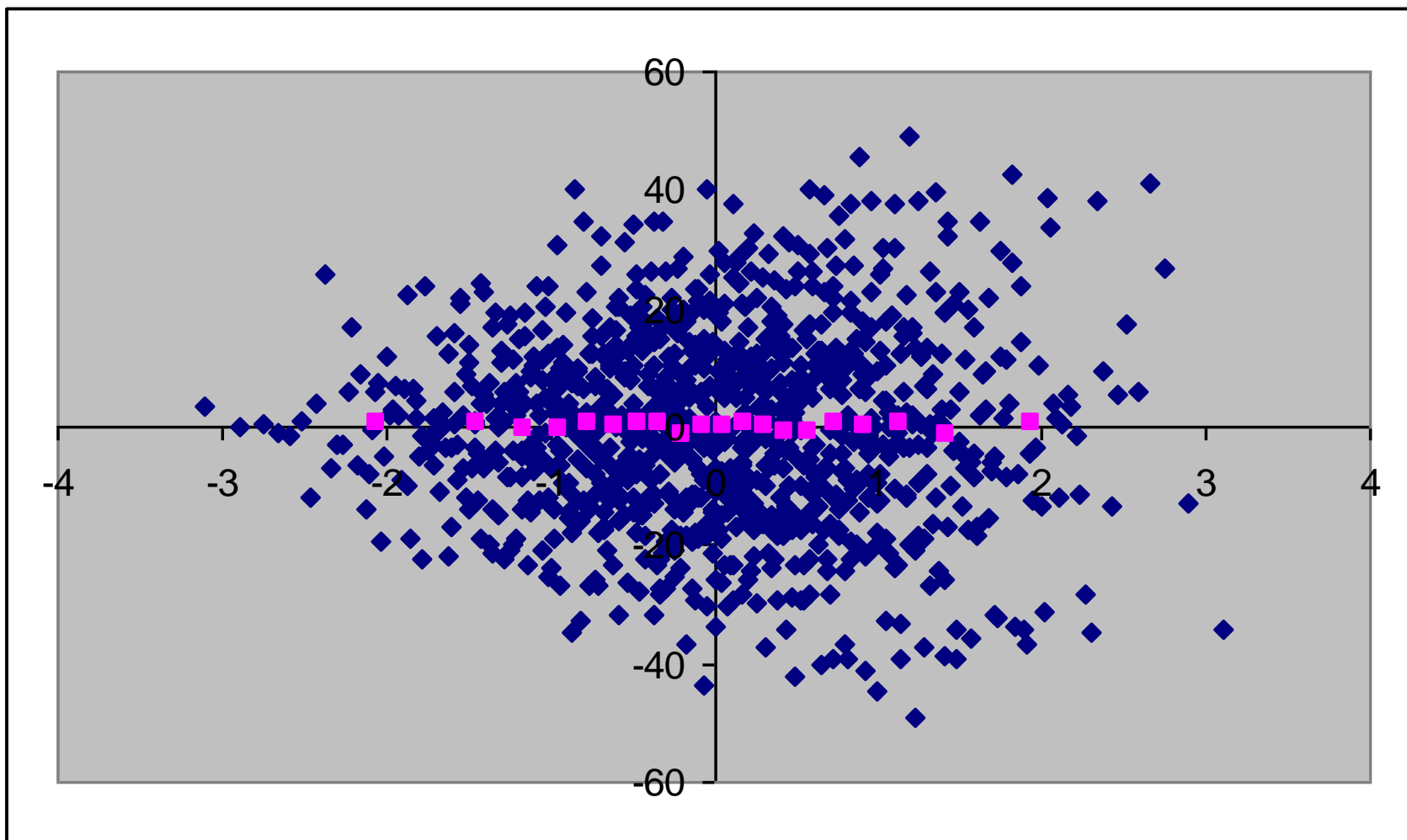


Non additive models

Is $S_i = 0$?



Is this factor non-important?



There are terms which capture two-way, three way, \cdots 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)

$$\begin{aligned} V_{X_i} \left(E_{\mathbf{X}_{\sim i}} \left(Y \mid X_i \right) \right) &= V_i \\ V_{X_i X_j} \left(E_{\mathbf{X}_{\sim ij}} \left(Y \mid X_i X_j \right) \right) &= \\ &= V_i + V_j + V_{ij} \end{aligned}$$

...

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 $k(k-1)/2 \dots$

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

$$\begin{aligned} V = & V_1 + V_2 + V_3 + \\ & + V_{12} + V_{13} + V_{23} + \\ & + V_{123} \end{aligned}$$

which divided by V becomes \dots

$$\begin{aligned}
1 = & S_1 + S_2 + S_3 + \\
& + S_{12} + S_{13} + S_{23} + \\
& + S_{123}
\end{aligned}$$

We can compute an index S_{Ti} so that e.g. for factor x_1 :

$$S_{T1} = S_1 + S_{12} + S_{13} + S_{123}$$

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 X_i could be fixed.

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

What is the shortcoming
of S_{Ti} ?



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

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

Scaled to $[0,1]$; first order and total order
sensitivity coefficient

Why these measures?

$V_{X_i} \left(E_{\mathbf{X}_{\sim i}} \left(Y | X_i \right) \right)$ Factors
prioritization

$E_{\mathbf{X}_{\sim i}} \left(V_{X_i} \left(Y | \mathbf{X}_{\sim i} \right) \right)$ Fixing (dropping)
non important
factors

Saltelli A. Tarantola S., 2002, On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal, *Journal of American Statistical Association*, **97** (459), 02–709.

More about the settings:

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

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

•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?

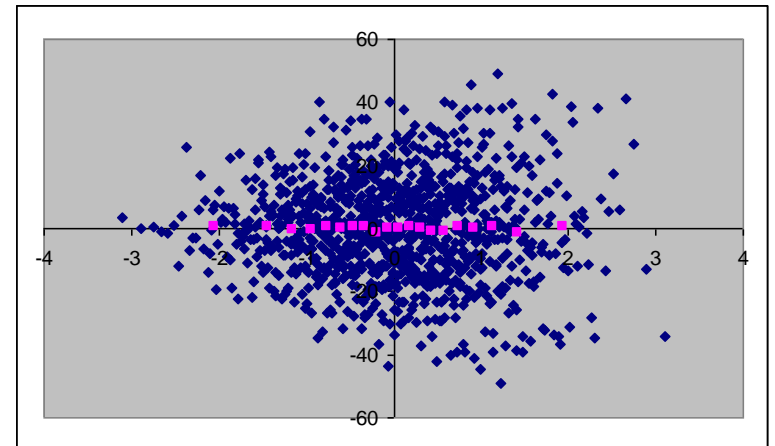
$$S_{Ti} \equiv \frac{E(V(Y|\mathbf{X}_{\sim i}))}{V_Y}$$

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 influent at the second
order

Can we use S_{T_i} to fix a factor?



If $S_{T_i} = 0$ is X_i a non-influential factor?

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

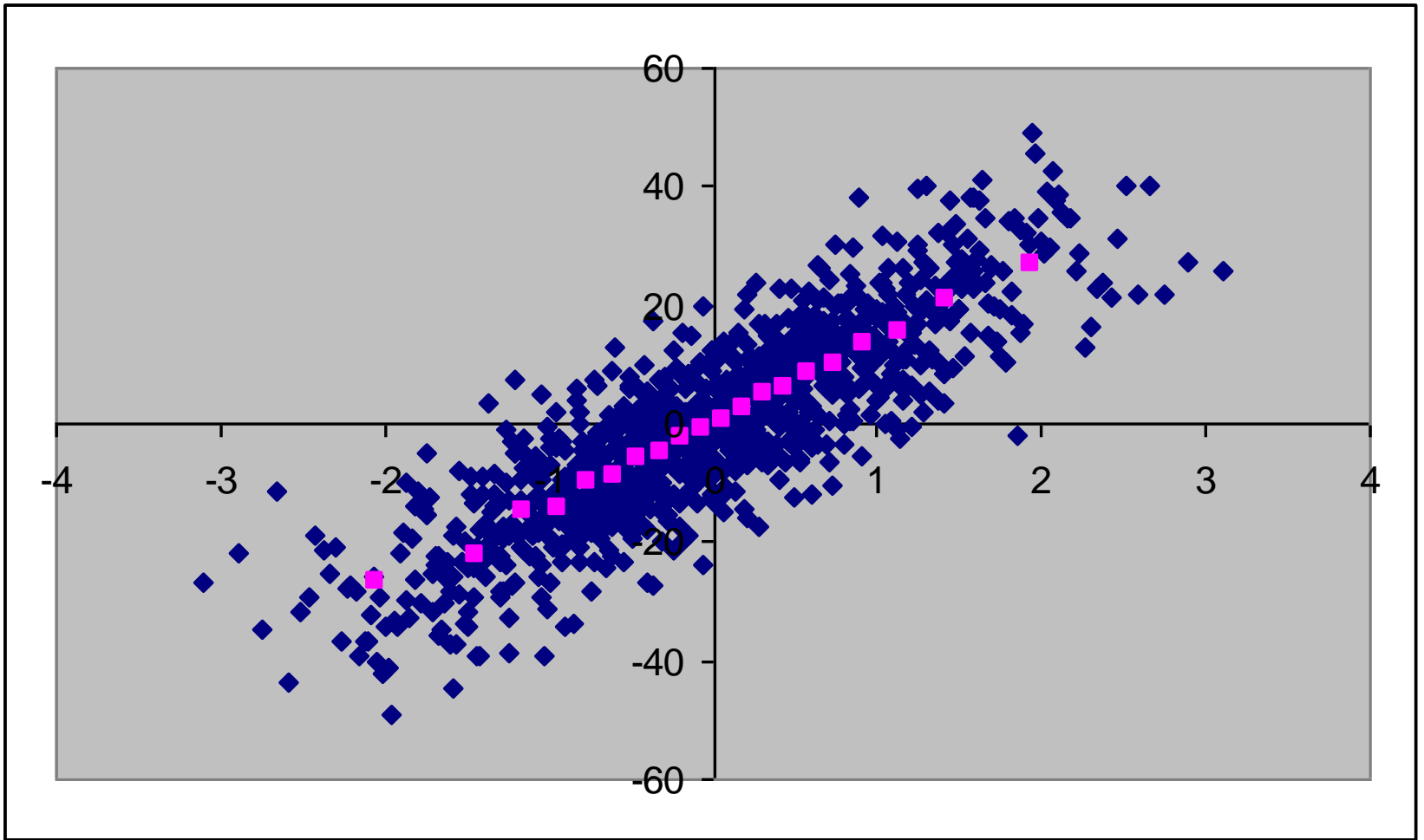
For a mean of non-negative entries to be zero all entries must be zero

Variance is always a positive number

If $S_{T_i} = 0 \rightarrow X_i$ is non influent as there is no point in the hyperspace of the input where x_i has an effect; $S_{T_i} = 0$ 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

y_1 NOT OK

y_2 OK

...

y_N NOT OK

If y can be partitioned into 'ok' and 'not ok' then ...

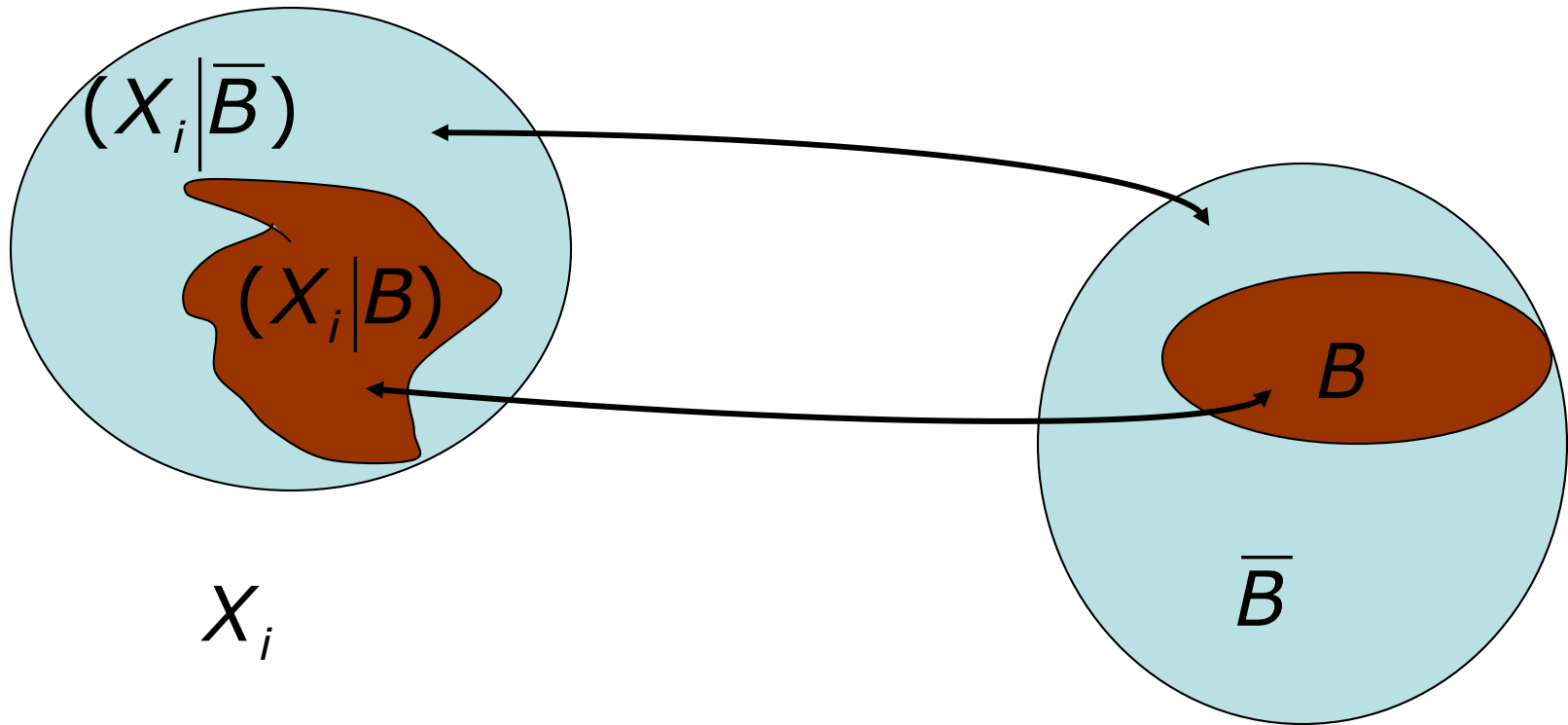
x_{11}	x_{12}	...	x_{1k}	NOT OK	y_1	NOT OK
x_{21}	x_{22}	...	x_{2k}	OK	y_2	OK
...	
x_{N1}	x_{N2}	...	x_{Nk}	NOT OK	y_N	NOT OK

... Then likewise the x_i 's can be partitioned; a 'not ok' x_i is one corresponding to a 'not ok' y_i

NOT OK	x_{11}	x_{12}	\dots	x_{1k}
OK	x_{21}	x_{22}	\dots	x_{2k}
	\dots	\dots	\dots	\dots
NOT OK	x_{N1}	x_{N2}	\dots	x_{Nk}

Taking one column at a time I
can split the sample of each
factor into two subsets

Monte Carlo filtering



B = OK

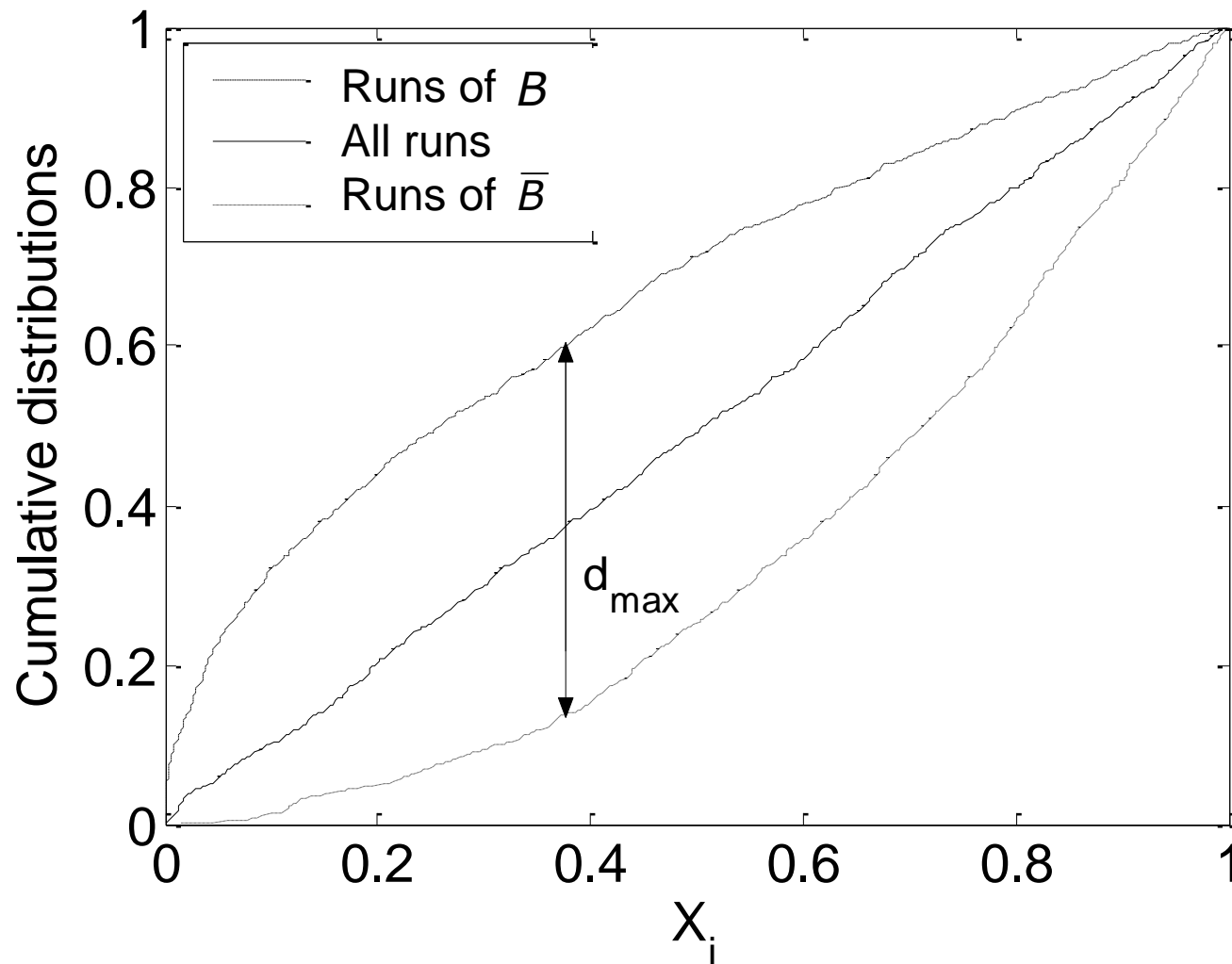
\bar{B} = not OK

Monte Carlo filtering

Step by step:

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

Monte Carlo filtering

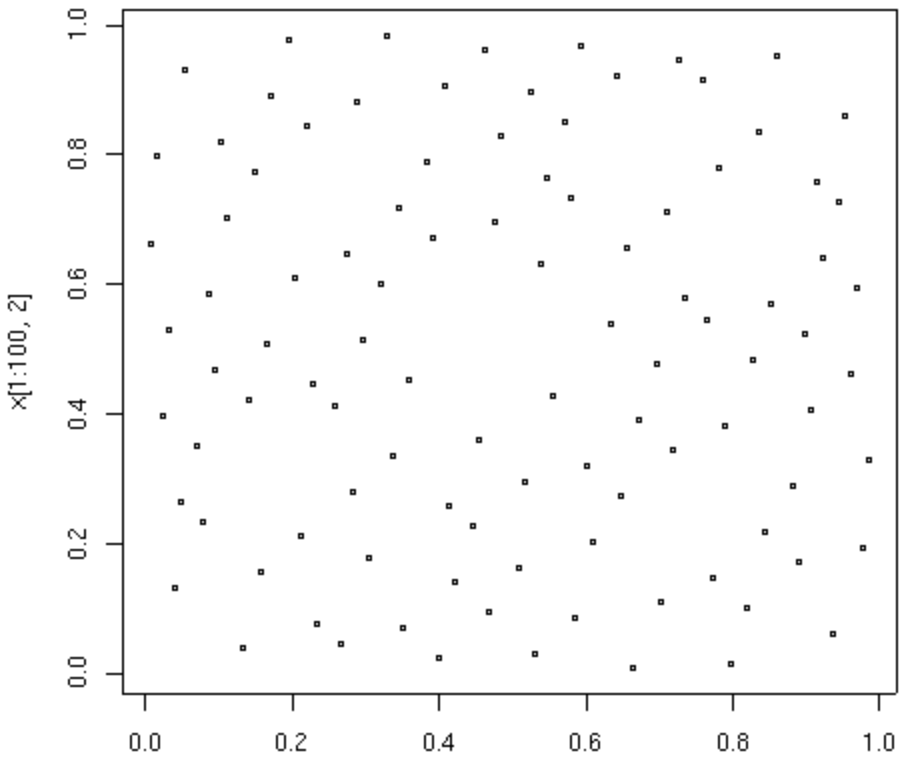




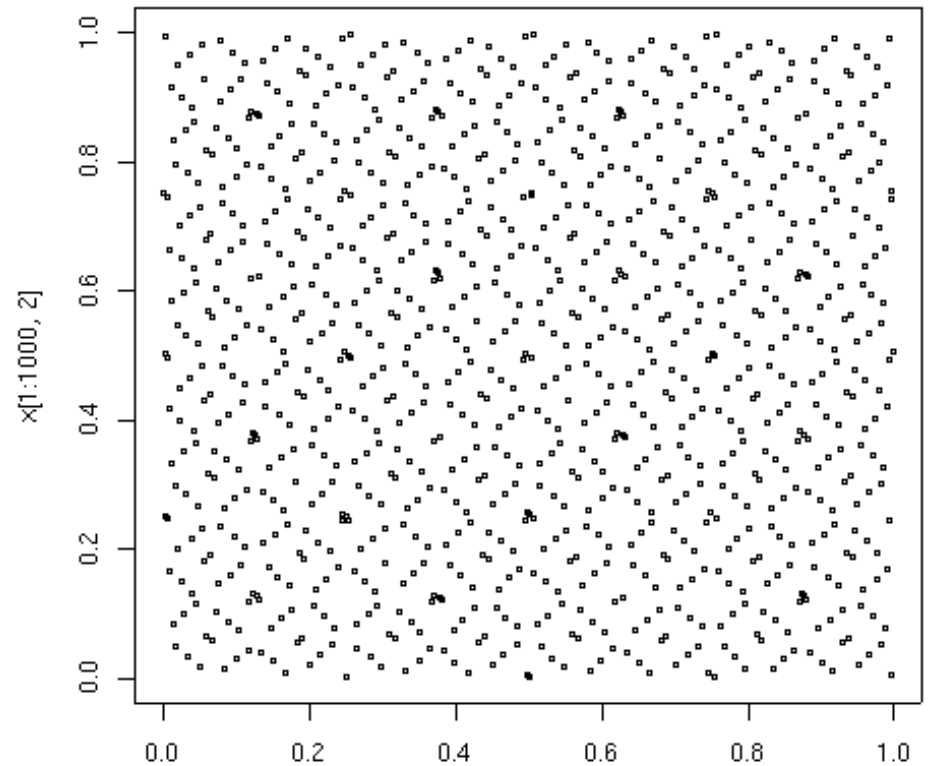
How to generate
the random
sample?

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



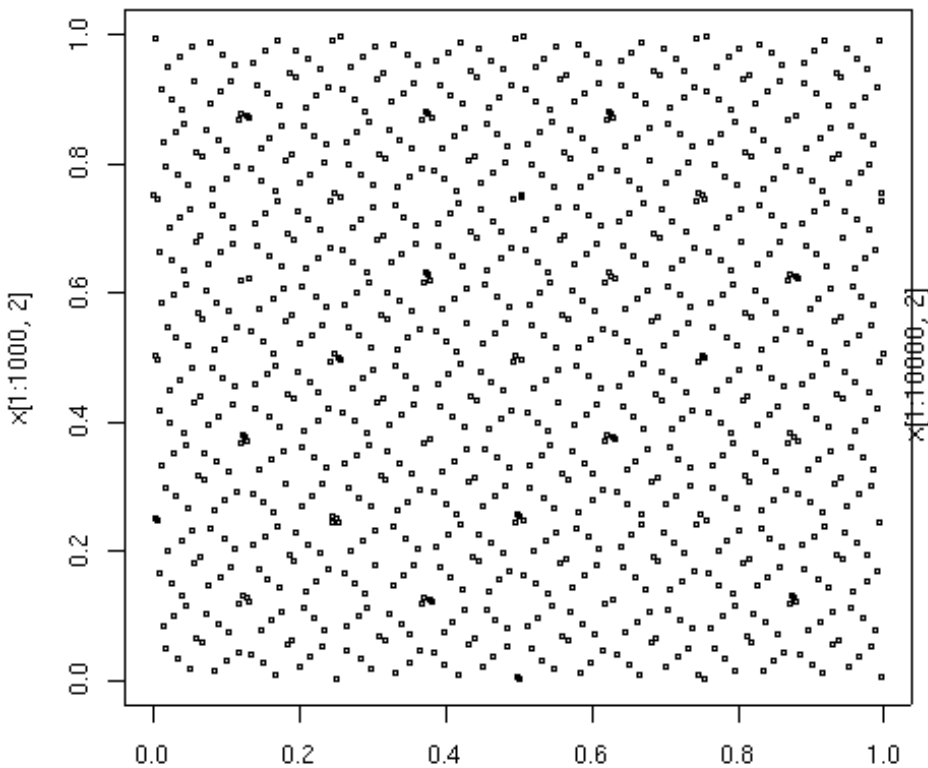


X_1, X_2 plane, 100 Sobol' points

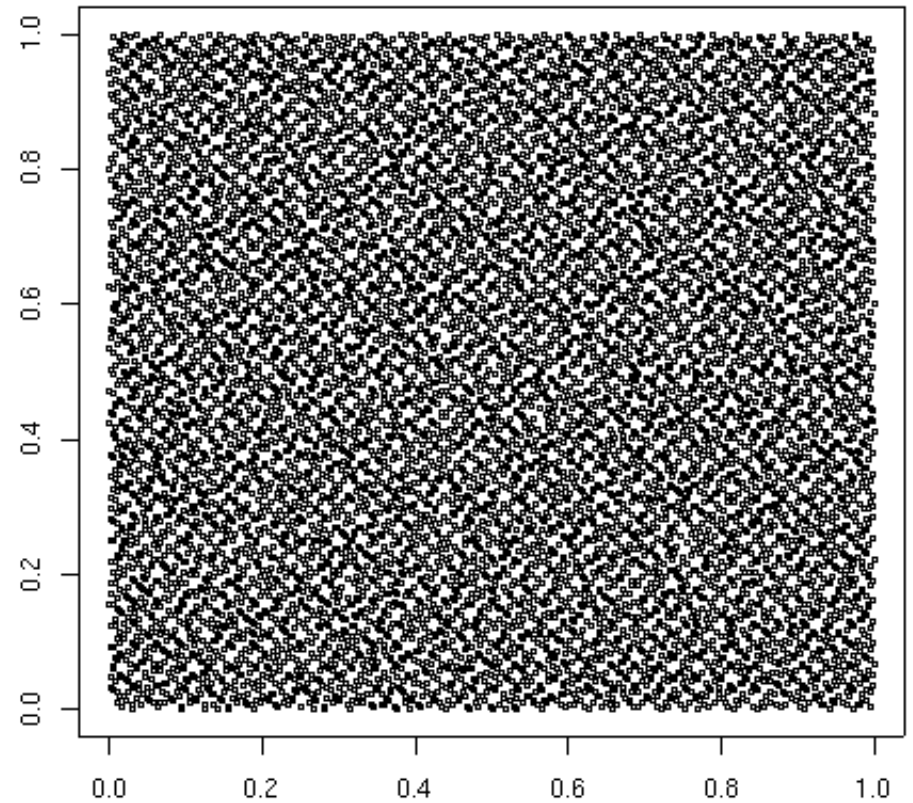


X_1, X_2 plane, 1000 Sobol' points

Sobol' sequences of quasi-random points

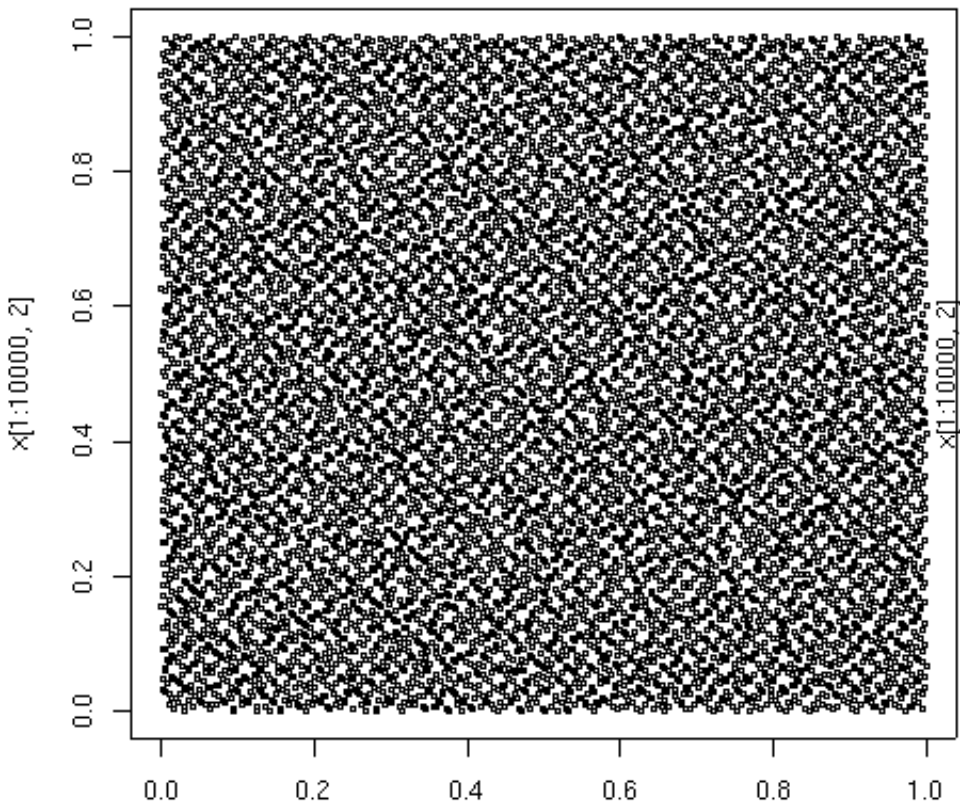


X_1, X_2 plane, 1000 Sobol' points

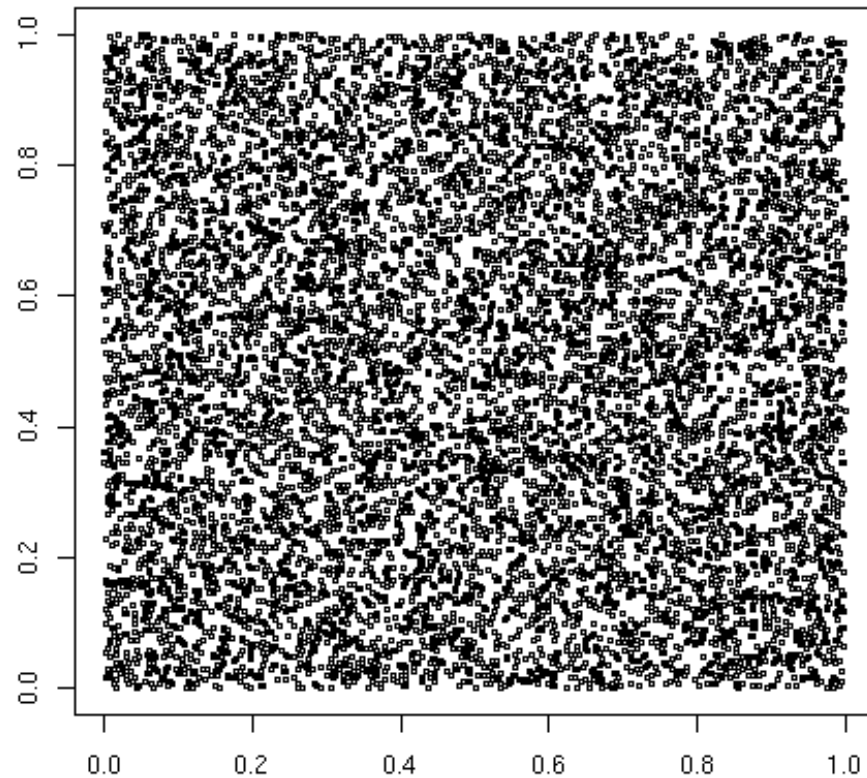


X_1, X_2 plane, 10000 Sobol' points

Sobol' sequences of quasi-random points



X_1, X_2 plane, 10000 Sobol' points



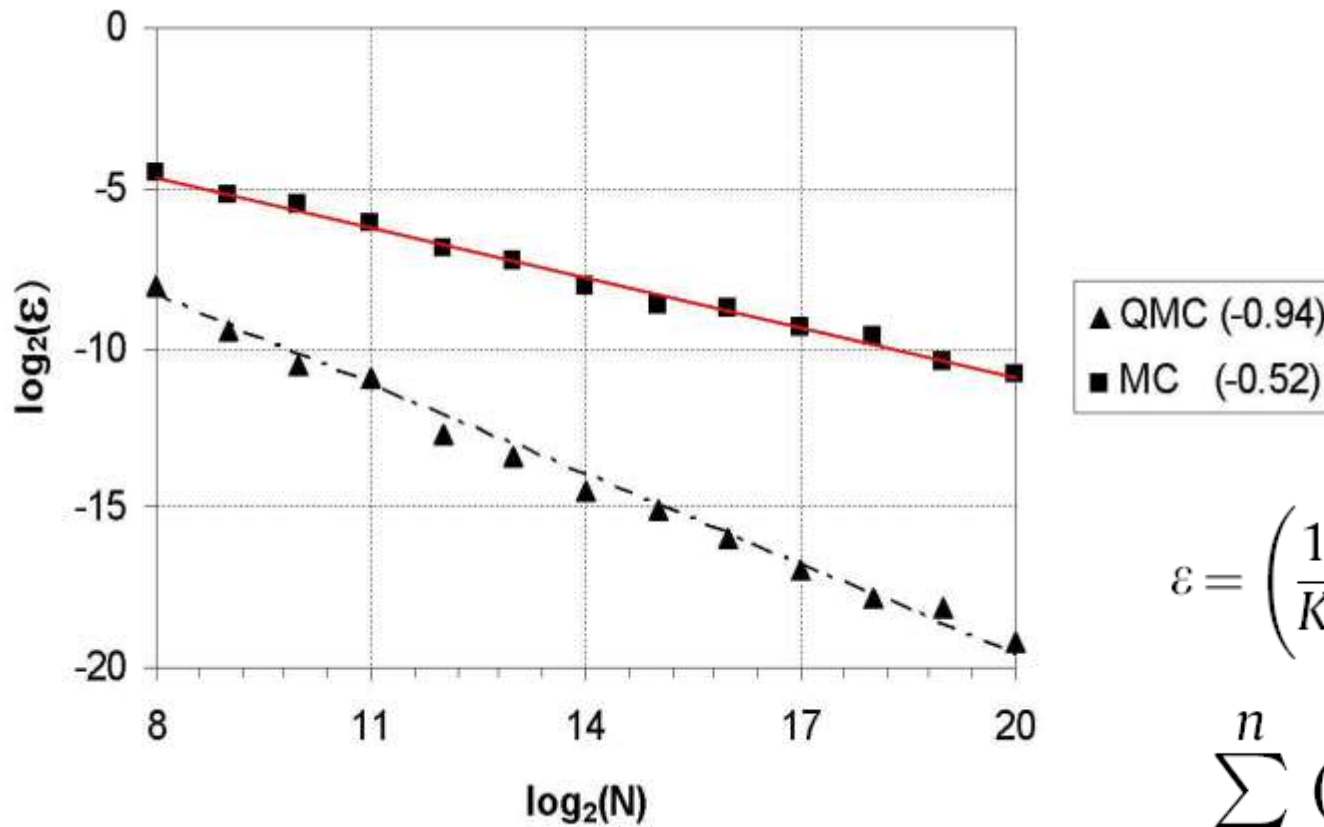
X_1, X_2 plane, 10000 random points

Sobol' sequences of quasi-random points
against random points

Why quasi-random



Sergei Kucherenko,
Imperial College
London



$$\varepsilon = \left(\frac{1}{K} \sum_{k=1}^K (I[f] - I_k[f])^2 \right)^{1/2}$$

$$\sum_{i=1}^n (-1)^i \prod_{j=1}^i x_j$$

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.

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

See Campolongo F, Saltelli A, Cariboni, J, 2011, From screening to quantitative sensitivity analysis. A unified approach, *Computer Physics Communication*, 182 (4), pp. 978–988.

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
factors has a 5% uncertainty



Discussion point



- Why should I not run a sensitivity analysis where each factors has a 5% uncertainty



END

Twitter:
[@andreasaltelli](https://twitter.com/andreasaltelli)