Tobias K Andersen, Xiangzhen Kong

Why should all modelers do sensitivity analysis?
SA for complex models

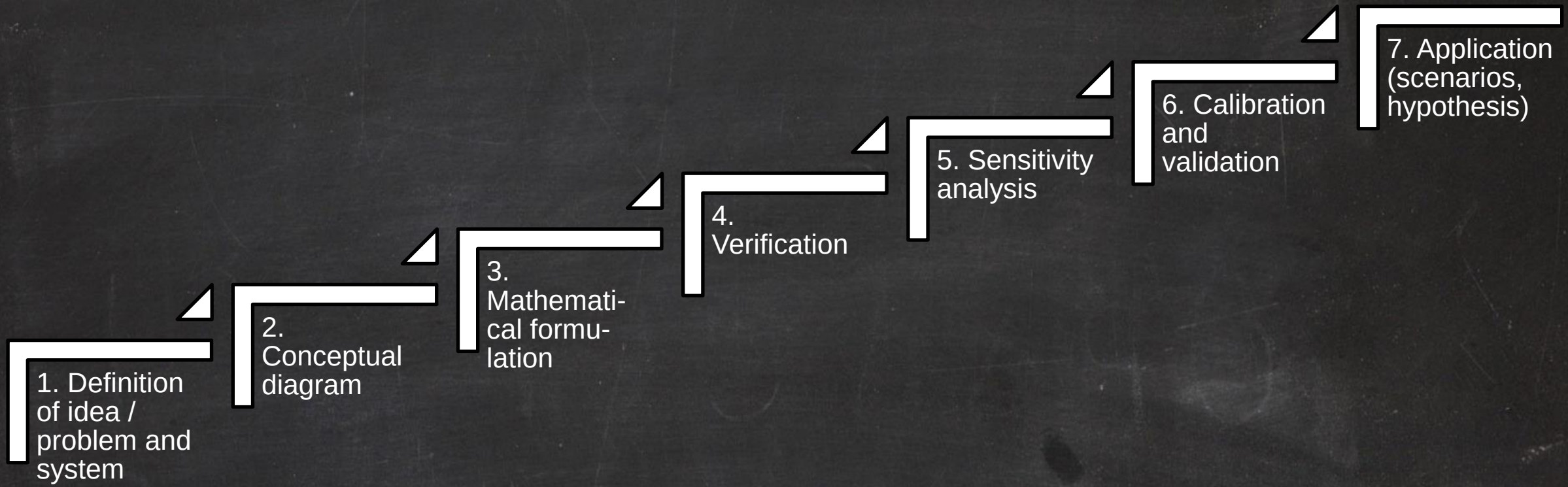
Aim of presentation

Learning outcomes:

Participants will get an introduction to

- The importance of sensitivity analysis
- SA connection to uncertainty analysis
- Local and global SA methods
- Potential SA errors
- SA methods with parsac
- SA examples

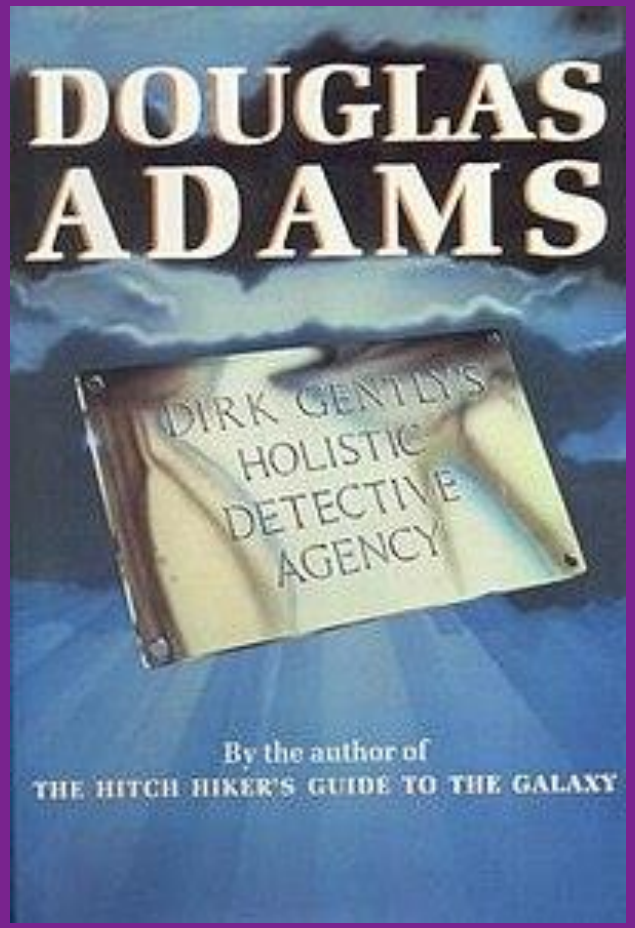
The modelling process



What modelling is not

Well, Gordon's great insight was to design a program which allowed you to specify in advance what decision you wished it to reach, and only then to give it all the facts. The program's task, […], was to construct a plausible series of logical-sounding steps to connect the premises with the conclusion.

Douglas Adams, 1987, p. 69, *Dirk Gently's Holistic Detective Agency*



There are many opinions (and feelings) about models

Because models influence decisions, which in turn influence people's lives.

LandbrugsAvisen

Søg her

FORSIDE KVÆG MARK SVIN TEKNIK ØKONOMI NOTERINGER DEBAT EJENDOMSMARKED JOBLANDET

Mathematical models misregulate agriculture

ESTE NYHEDER

rnes tillid til økonomien dykker (10:22)

n billigere end F1-lån ved indfrielse (10:19)

Fejrer 40 år med ny plov (10:14)

Vindmøllemodstandere sure over "svinestreg" i

Torsdag 26. februar 2015 | 21:04 Skrevet af Frederik Thalbitzer Udskriv

" The entire environmental regulation is dependent on a model of emissions that does not match reality …"

POLITIKEN ≡ SEKTIONER Q SØG ♥ DIT POLITIKEN MERE

PÅ UDVALGTE STEL

*GÆLDER VED KØB AF KOMPLET BRILLE 19. – 31. JANUAR 2016 OG KAN IKKE KOMBINERES MED ANDRE TILBUD OG RABATTER.

2050

Analyse Feature Grafik 2050 forklarer Klima Energi Grøn omstilling

Farmer demands sampling – not modelling

Landmand får lov til at gøde mere af kommunen. Konkrete målinger viser, at han forurener langt mindre, end staten siger.

SKÅL! Byrådsmedlem Asger Møller Madsen (V) (t.v.) og landmand Carsten Søborg (t.h.) skåler i drænvandet fra Carsten Søborgs marker. Målinger fra Aarhus Universitet viser, at det er tre gange renere end kravet til drikkevand. - Foto: SØREN SCHNOOR

Gem + Følg

THOMAS FÆRGEMAN
Journalist

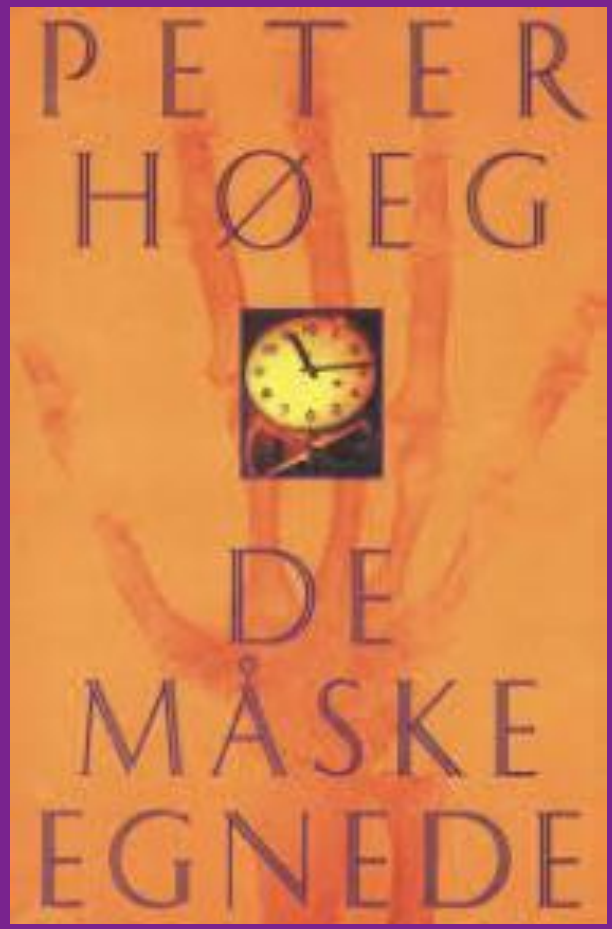
»Skål!« Vi lægger halsen tilbage og

HVAD ER FORSKELLEN?

Why understand uncertainty?

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.***

Peter Høeg, 1993, *Borderlines*



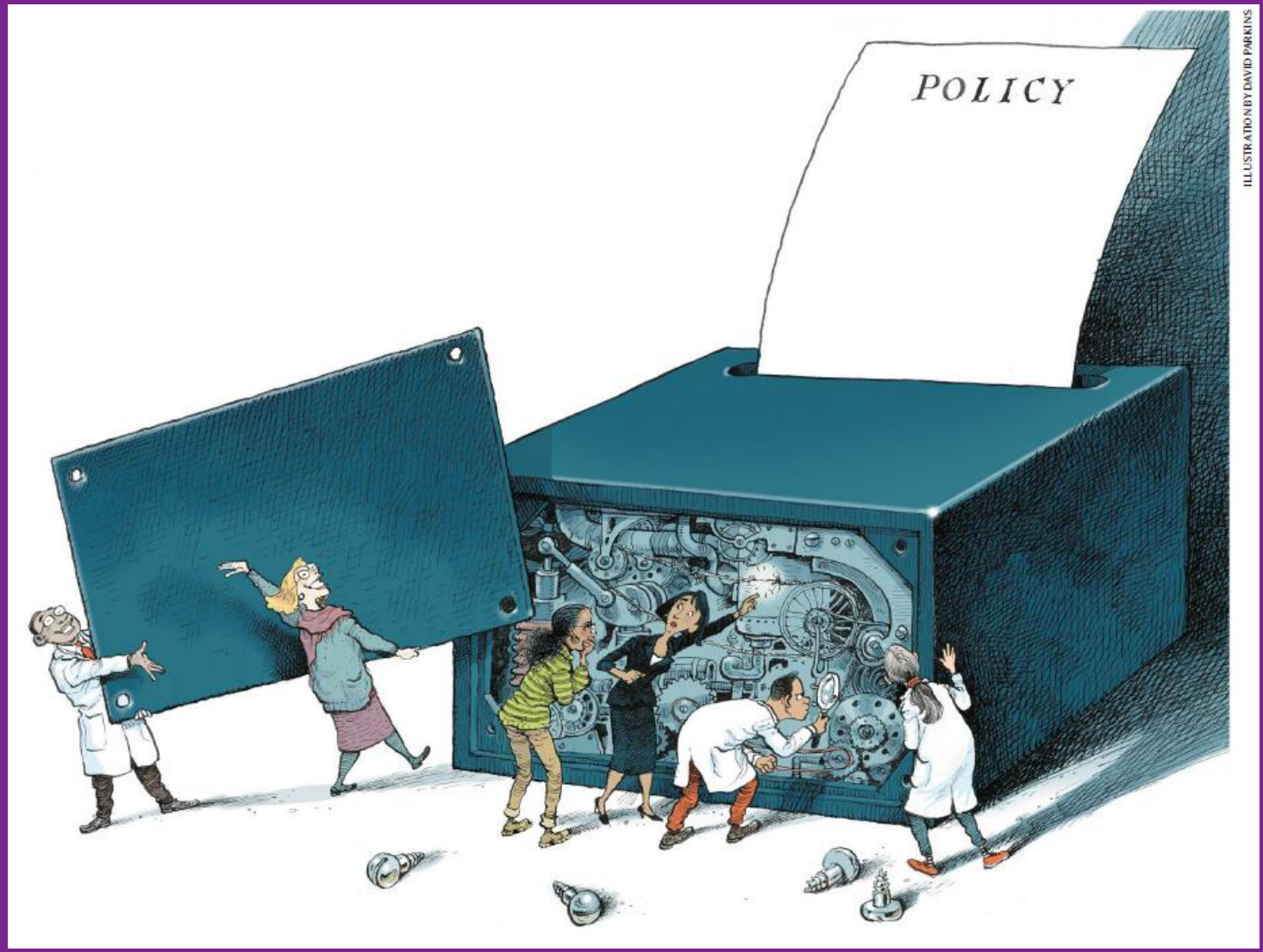
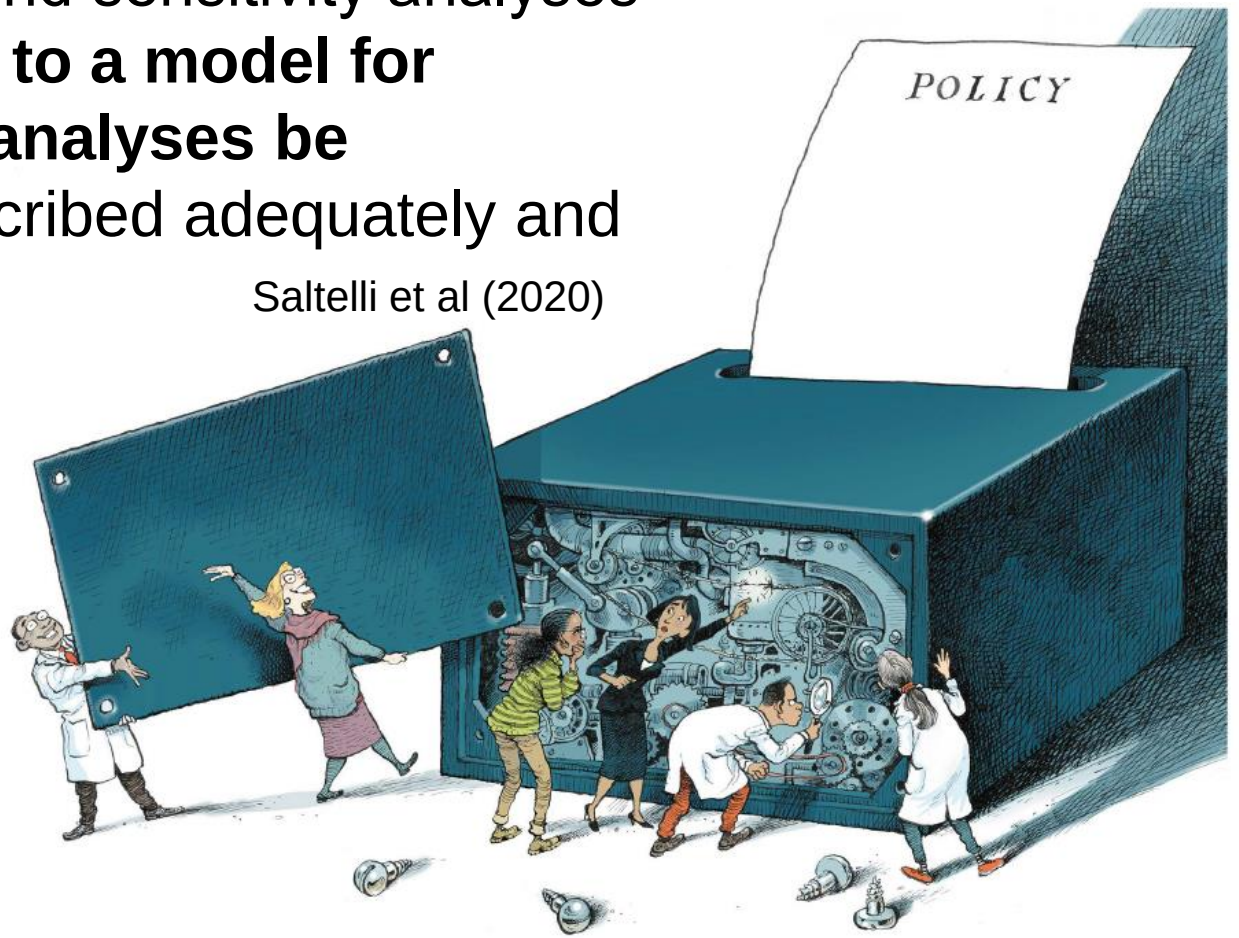


Illustration from Nature article by Saltelli et al (2020)

SA should be common practice

"However, these global uncertainty and sensitivity analyses are often not done. **Anyone turning to a model for insight should demand that such analyses be conducted**, and their results be described adequately and made accessible."

Saltelli et al (2020)



SA is also recommended by both the EU Commision and the US Environmental Proctection Agency.

Best practice for SA

Pre-analysis

- 1) Establish goals for analysis
- 2) Decide on input factors
- 3) Decide on distribution function for each input factor
- 4) Choose SA strategy and method(s)

Step 1

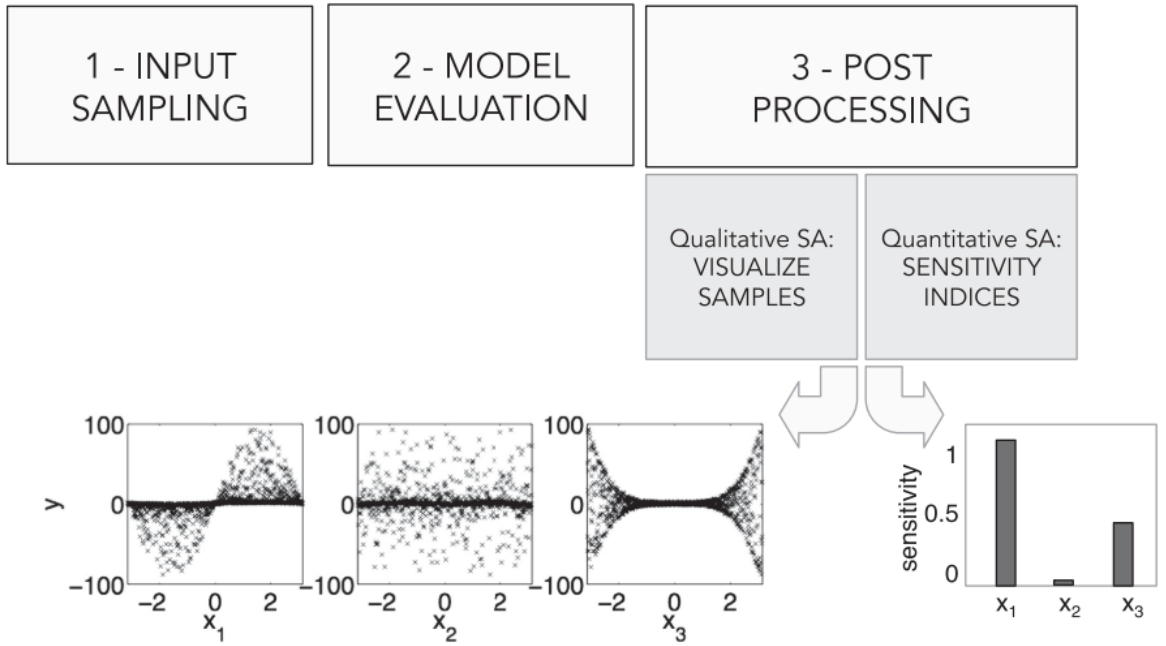
- 5) Generate input sample

Step 2

- 6) Execute model on generated input sample and produce output

Step 3

- 7) Analyze model outputs with SA methods, evaluate and draw conclusion of SA



Pianosi et al (2016)

Fig. 2. The three basic steps in sampling-based Sensitivity Analysis, with an example of qualitative or quantitative results produced by the post-processing step.

SA settings

- **Factor prioritization**

  objective: to identify most the most important factor

- **Factor fixing**

  objective: fixing of non-influential factors (screening)

- **Variance cutting**

  objective: reduce variance of output from its unconditional value to a lower pre-established threshold value.

- **Factors mapping**

  objective: which factor is most responsible for producing realisations of output in the region of interest?

Caution: Here be errors!

- **Type I error**

Classify an un-important parameter/factor as influential.

- **Type II error**

Classify an important parameter/factor as noninfluential.

- **Type III error**

Wrong framing of the sensitivity analysis.

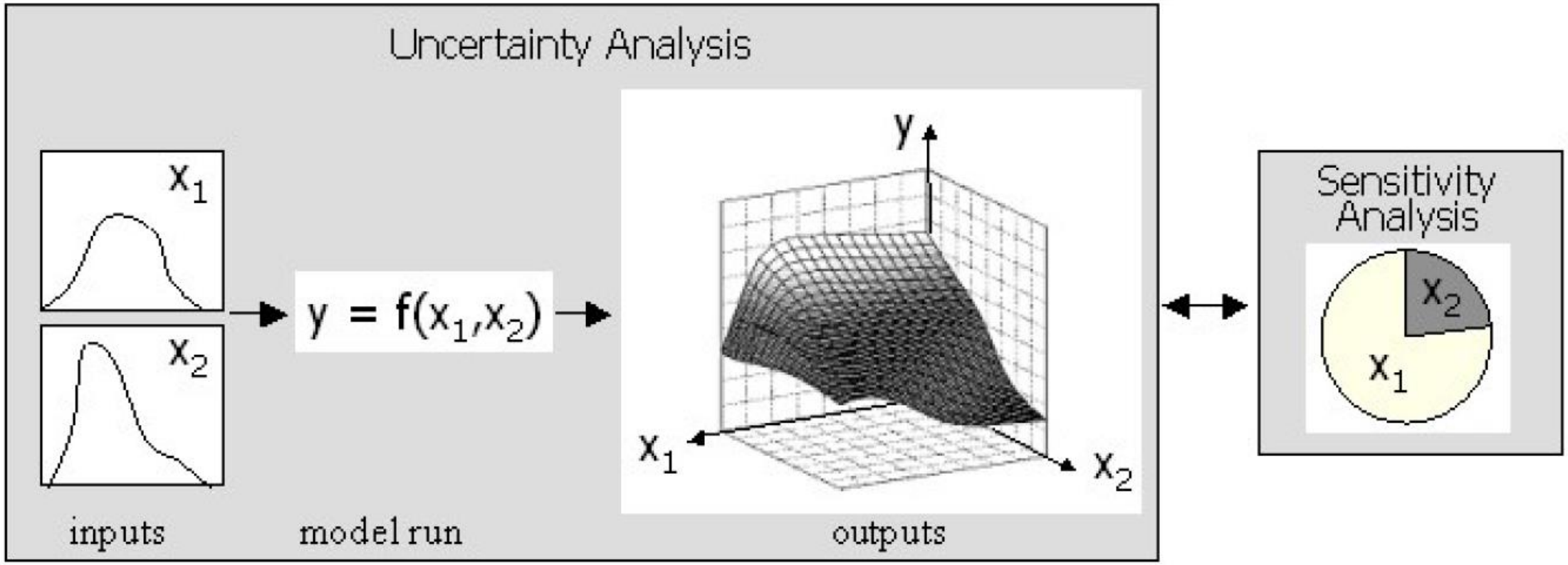


Figure D.5.1. Uncertainty and sensitivity analyses. Uncertainty analysis investigates the effects of lack of knowledge or potential errors of model inputs. Sensitivity analysis evaluates the respective contributions of inputs x_1 and x_2 to output y.

Sensitivity analysis: local vs global

Local sensitivity analysis

- Evaluates the effect of changing one parameter (from baseline) on the model output. This is repeated for all parameters one-at-a-time (OAT). All parameters considered in these OAT analyses are taken to be independent from one another.

- Therefore, local SA should only be used when the model has proved to be linear or at least additive. As local SA does not allow to take parameter interactions into account

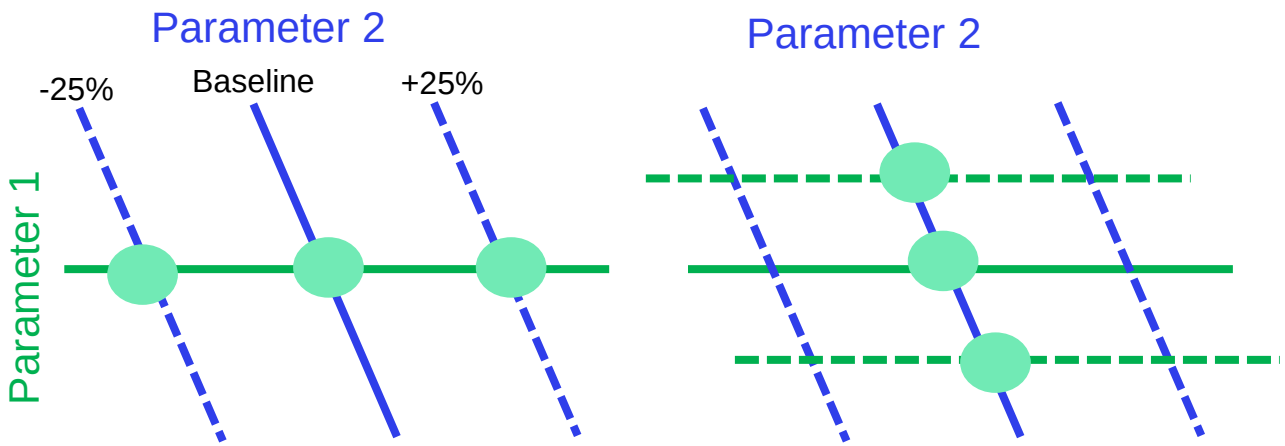
Global sensitivity analysis

- Simultaneously evaluates both the effect of each individual parameter and its interaction with other parameters on model output. This is because all parameters are varied simultaneously over the entire parameter space.

For an example of why local SA is usually perfunctory for complex environmental models see:

Saltelli and Annoni (2010)
How to avoid a perfunctory sensitivity analysis

Illustration of local vs global SA



Step 1

Value of parameter 1 is changed with +/- 25% from baseline. Model is evaluated 3 times.

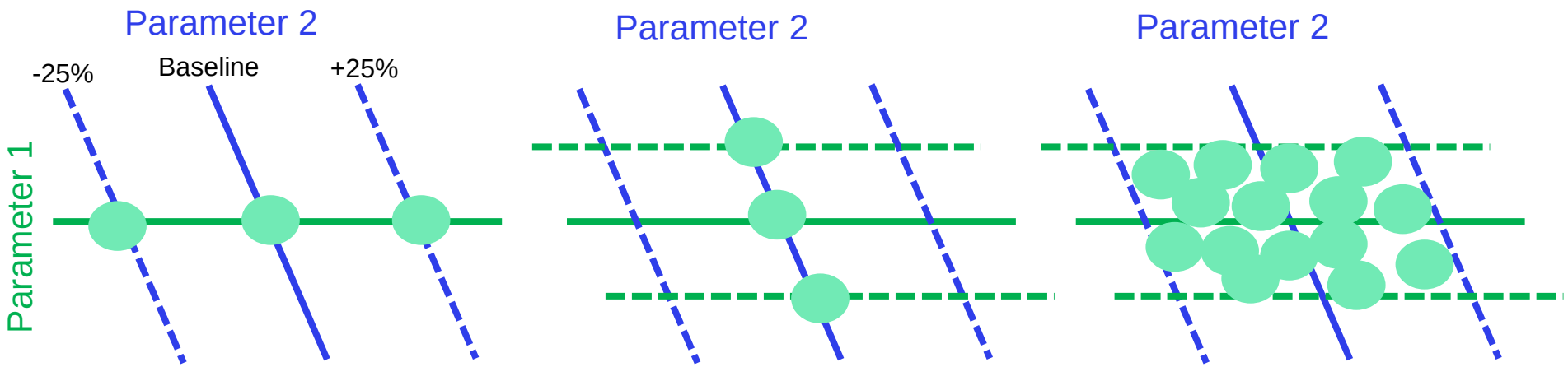
Step 2

Value of parameter 2 is changed with +/- 25% from baseline. Model is evaluated 3 times.

Local SA

(steps repeated as many times as number of parameters)

Illustration of local vs global SA



Step 1

Value of parameter 1 is changed with +/- 25% from baseline. Model is evaluated 3 times.

Step 2

Value of parameter 2 is changed with +/- 25% from baseline. Model is evaluated 3 times.

Entire parameter range of par1 and par2 is sampled by a sampling scheme and model is evaluated as many times as required by global SA method.

Local SA

(steps repeated as many times as number of parameters)

Global SA

For another example of why local SA is usually perfunctory for complex environmental models see:
Saltelli and Annoni (2010) How to avoid a perfunctory sensitivity analysis

Overview of global SA approaches

1. **Multiple-start perturbation approaches**, where global sensitivity is obtained by aggregation of 'OAT' sensitivities obtained at different baseline points (e.g. the Elementary Effects Test or method of Morris);
2. **Correlation and regression approaches**, where sensitivity is measured by the correlation between input and output samples;
3. **Regional sensitivity analysis** (or Monte Carlo filtering) methods, where sensitivity is related to variations in the distributions of input factors induced by conditioning the outputs; and
4. **Variance-based and density-based approaches**, where sensitivity is linked to variations in the output distribution induced by conditioning the inputs.

Wagener & Pianosi (2019): *What has Global Sensitivity Analysis ever done for us? A systematic review to support scientific advancement and to inform policy-making in earth system modelling*

Global SA approaches

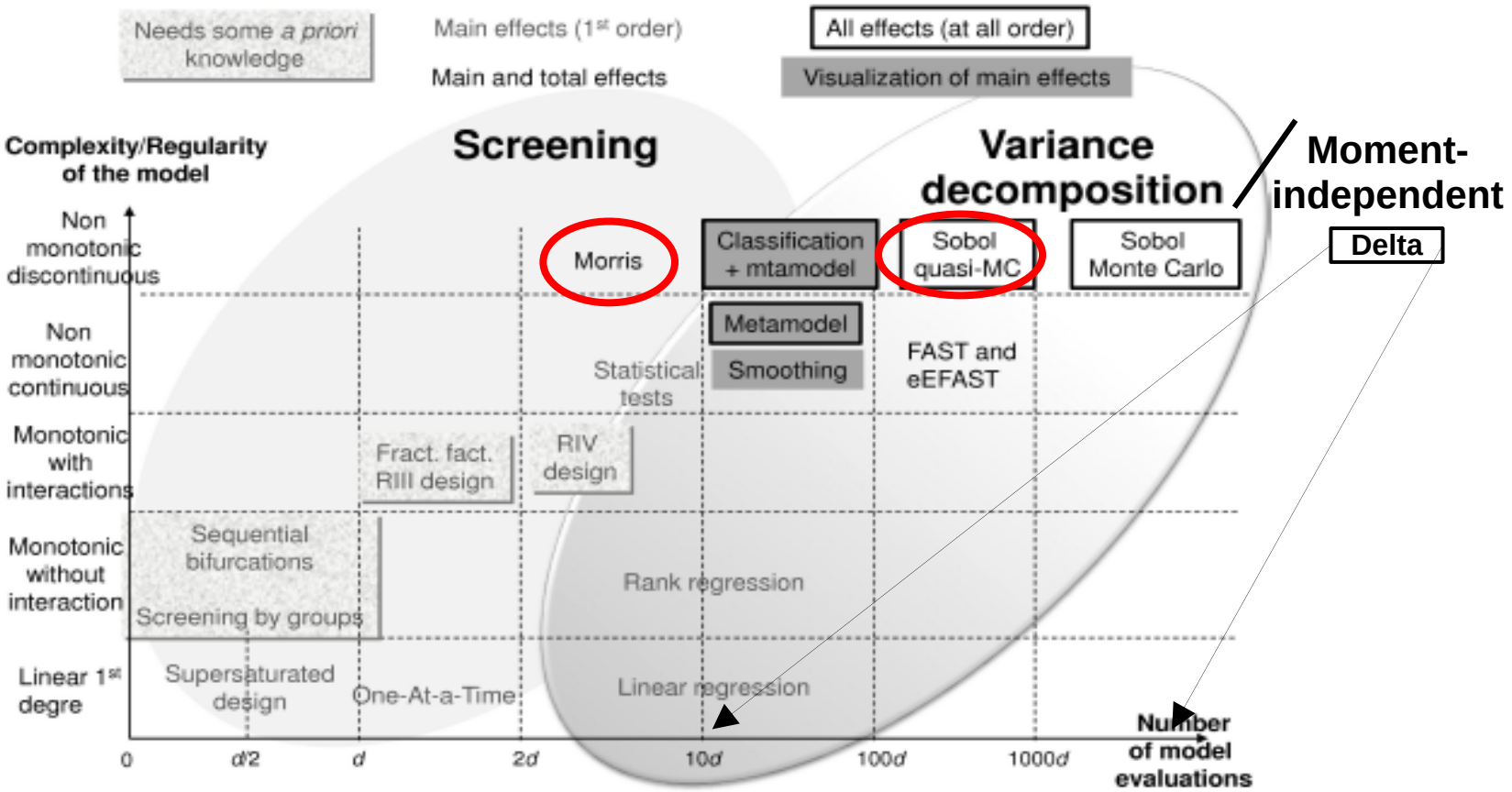
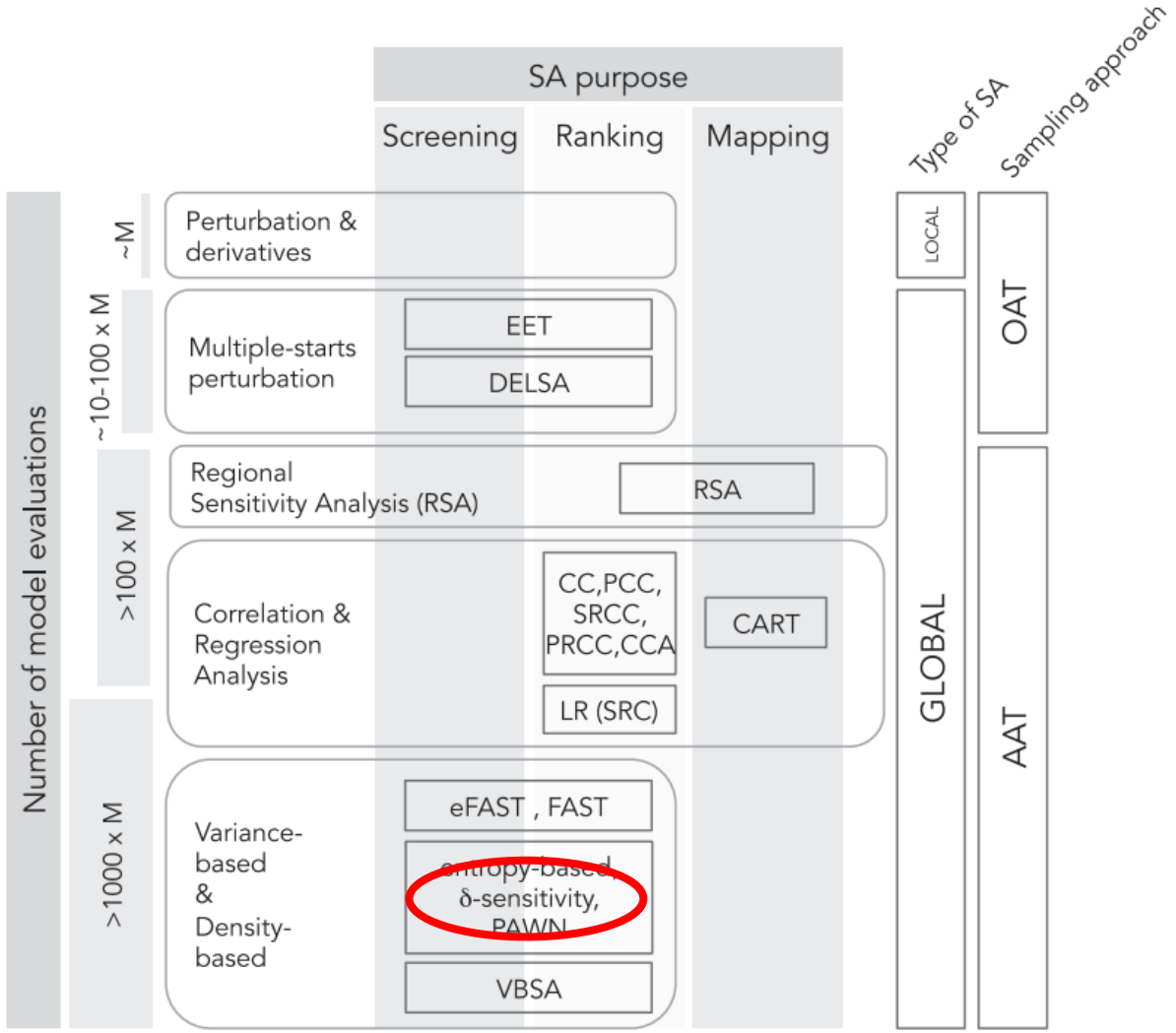


Fig. 6. SA methods graphical synthesis.

Fig from Ioos & Lemaitre (2014): A review on global sensitivity analysis methods with own addition of moment-independent method

Global SA approaches



Pianosi et al. (2016)

Fig. 3. Classification system of Sensitivity Analysis methods based on computational complexity (vertical axis; M is the number of input factors subject to SA) and purposes of the analysis. Some of the most widely used methods are reported (acronyms are defined in corresponding paragraphs of Sec. 3). Types of SA and sampling approaches are defined in Sec. 2.2. Figures about computational complexity are indicative, for a further discussion see Sec. 4.5.

Variance-based method: Sobol

First-order effect

XX

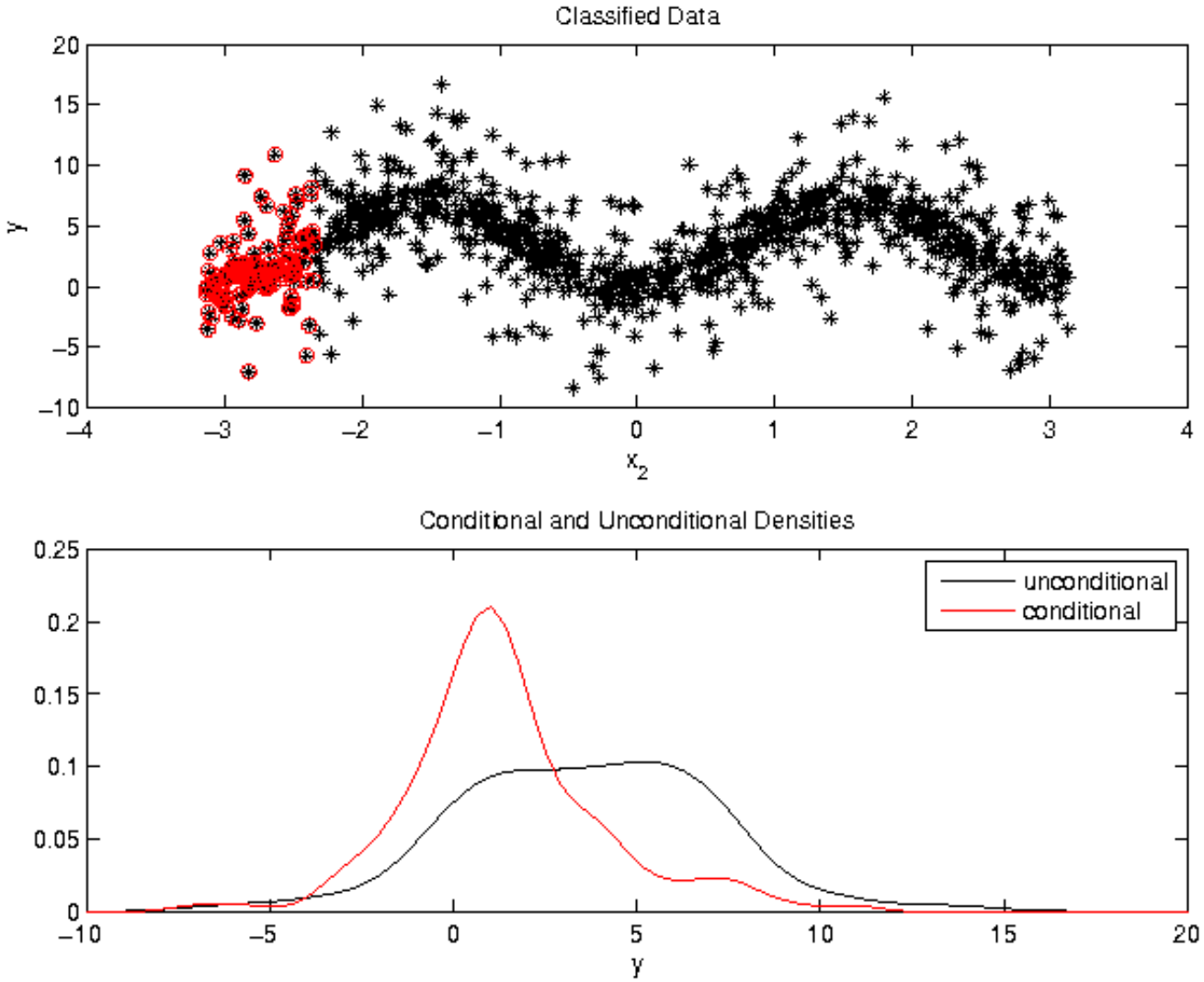
Second-order effect

XX.

Total order effect

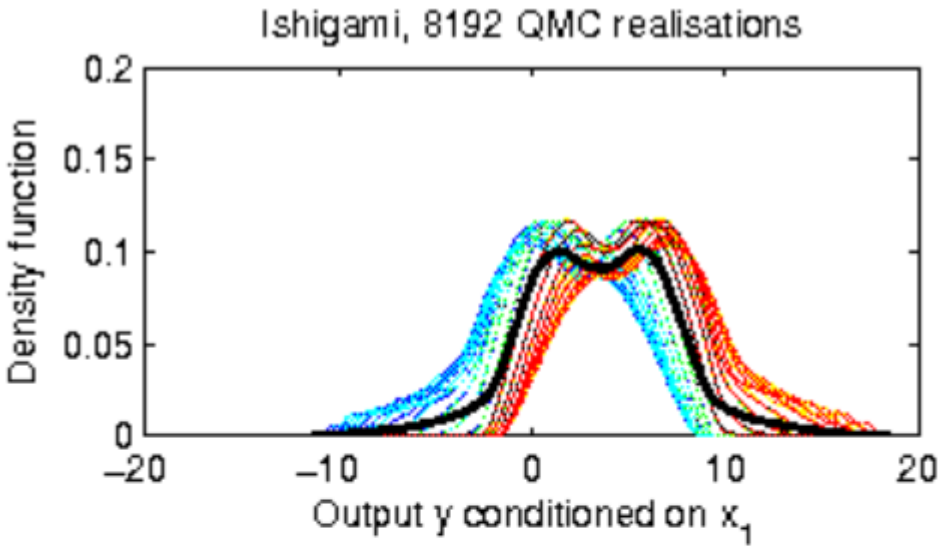
XX.

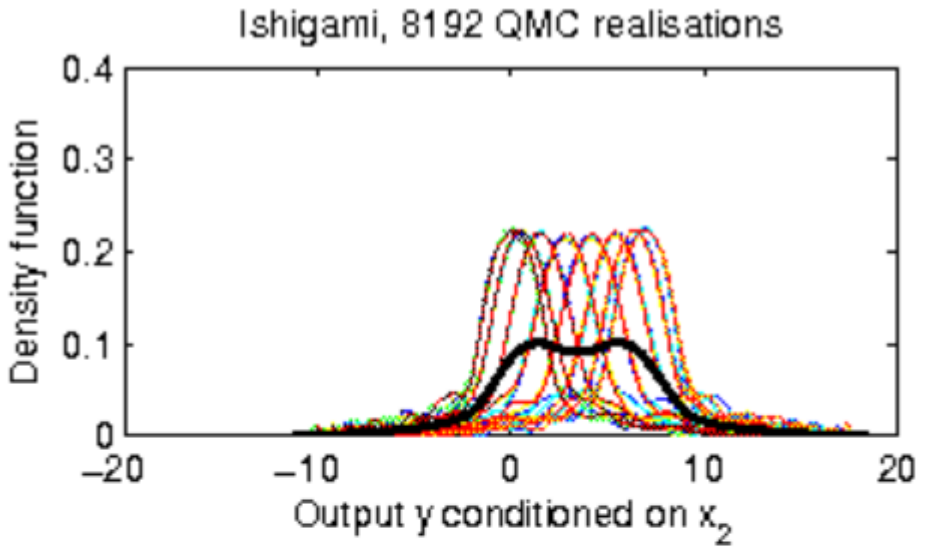
Moment-Independent method: Delta

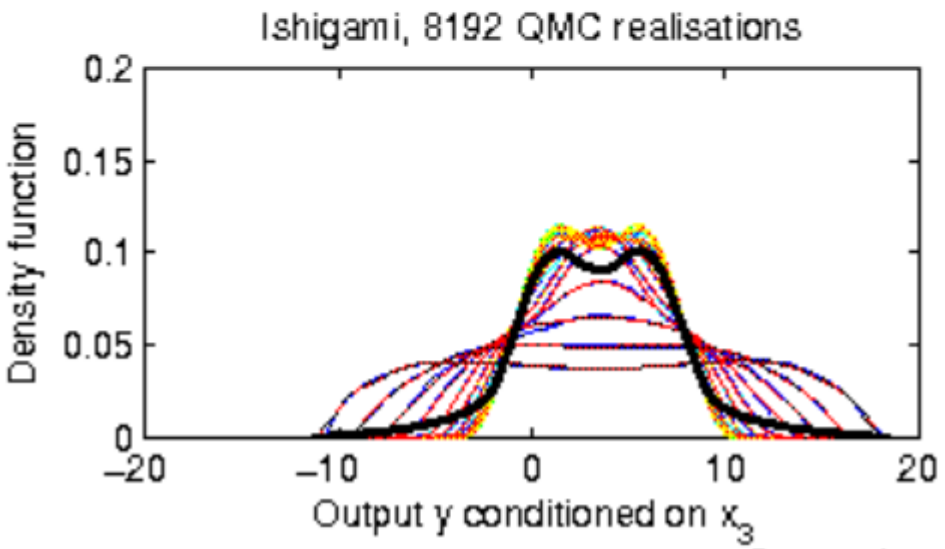


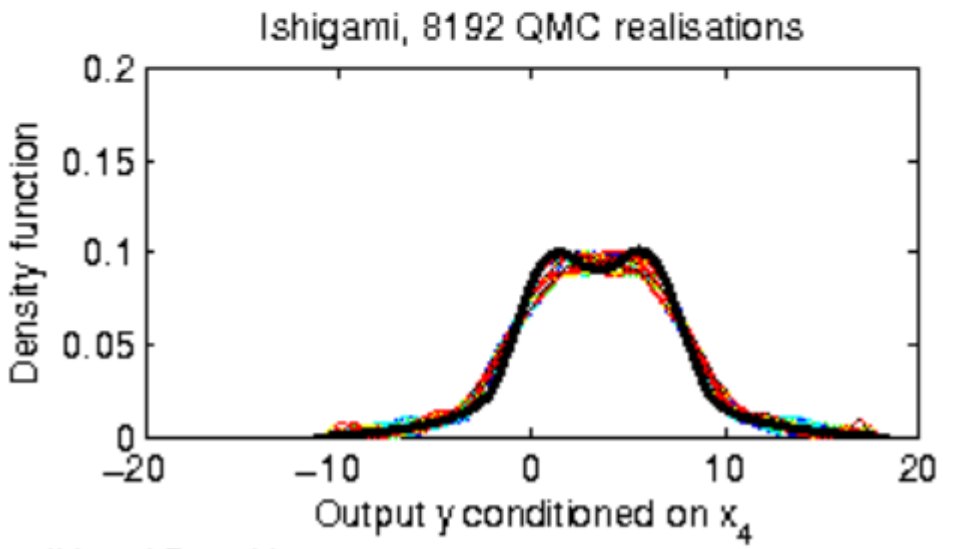
From Dr Plischke homepage: https://artefakte.rz-housing.tuclausthal.de/epl/work/GlobalSA/

Moment-Independent method: Delta



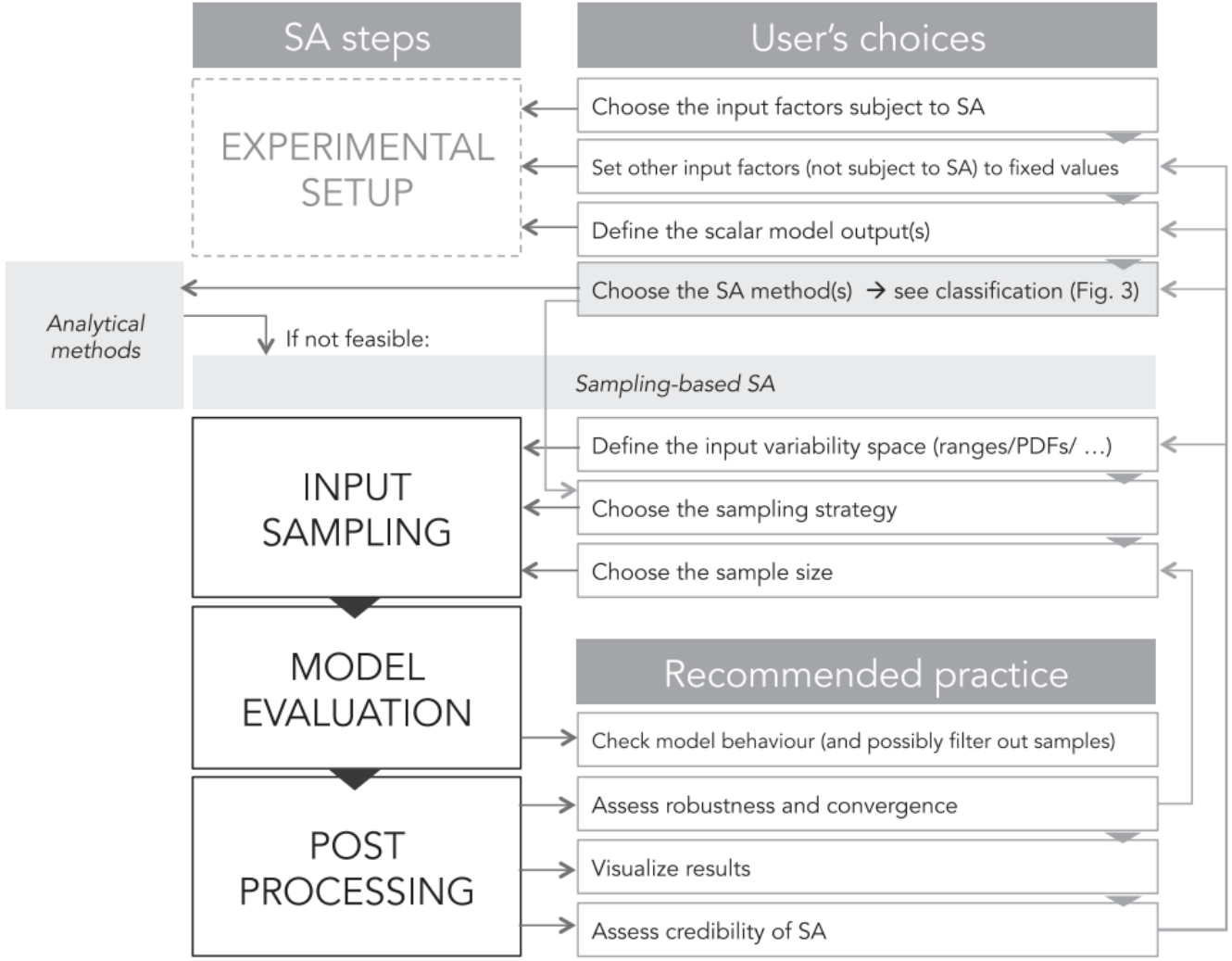






From Dr Plischke homepage: https://artefakte.rz-housing.tuclausthal.de/epl/work/GlobalSA/

SA is an iterative process



Pianosi et al (2016)

Fig. 4. Workflow for the application of Sensitivity Analysis, choices to be made and recommended practice for their revision.

SA with parsac

Andersen et al (2020) *Environ. Model. Softw.*

SCREENING: BORGONOVO'S δ ANALYSIS

Generate samples *with LHS*

Execute model simulations

Compute sensitivity indices

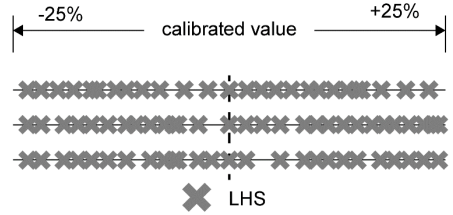
Apply cut-off threshold

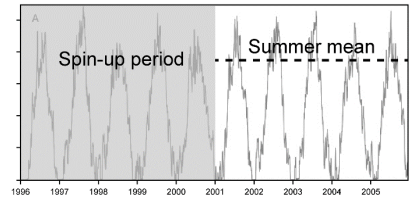
IN-DEPTH: BORGONOVO'S δ AND SOBOL ANALYSIS

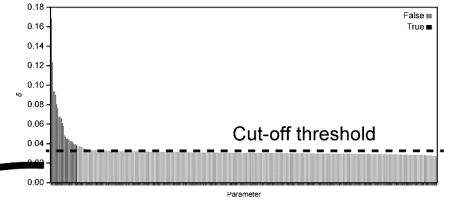
Generate samples *with LHS and Saltelli*

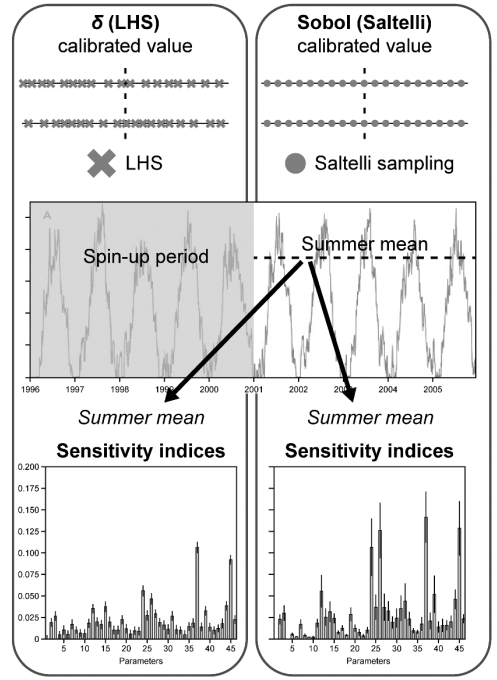
Execute model simulations

Compute sensitivity indices

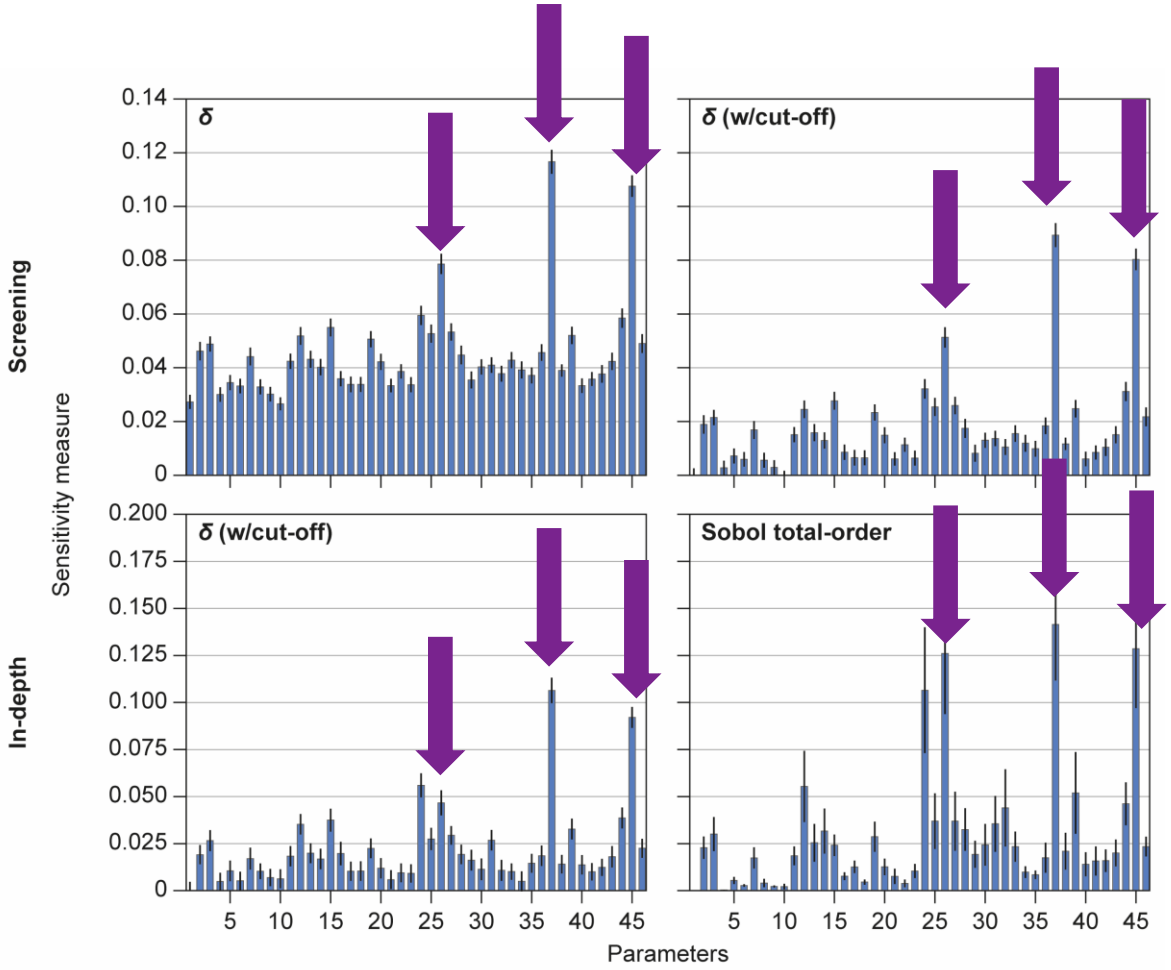








SA with parsac



Andersen et al (2020) *Environ. Model. Softw.*

SCREENING: BORGONOVO'S δ ANALYSIS

Generate samples *with LHS*

Execute model simulations

Compute sensitivity indices

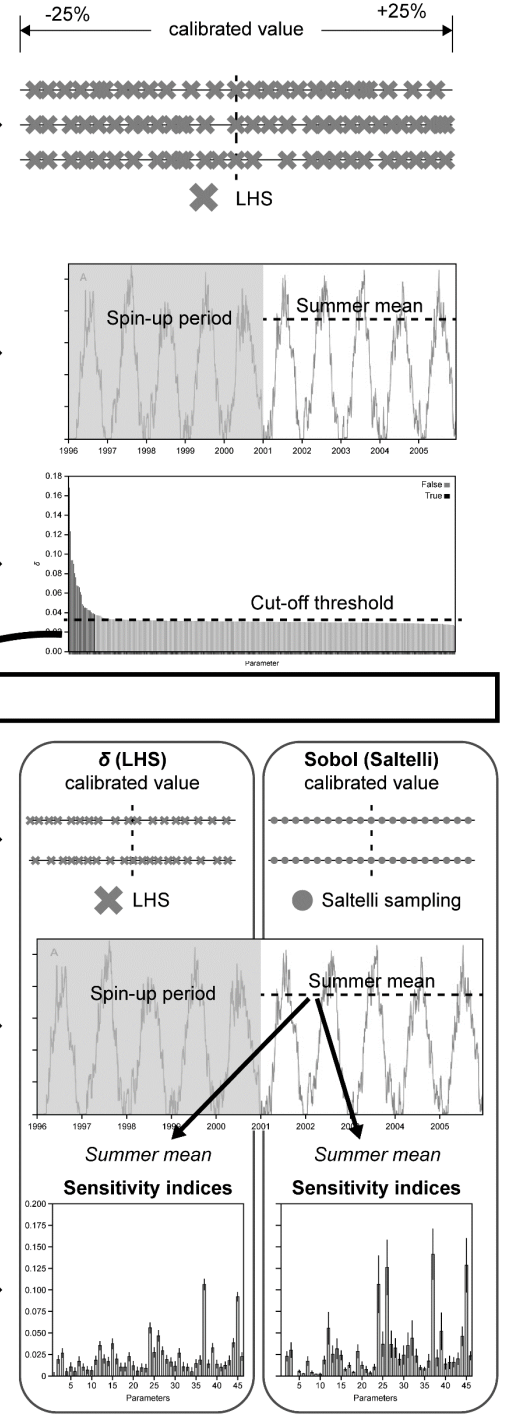
Apply cut-off threshold

IN-DEPTH: BORGONOVO'S δ AND SOBOL ANALYSIS

Generate samples *with LHS and Saltelli*

Execute model simulations

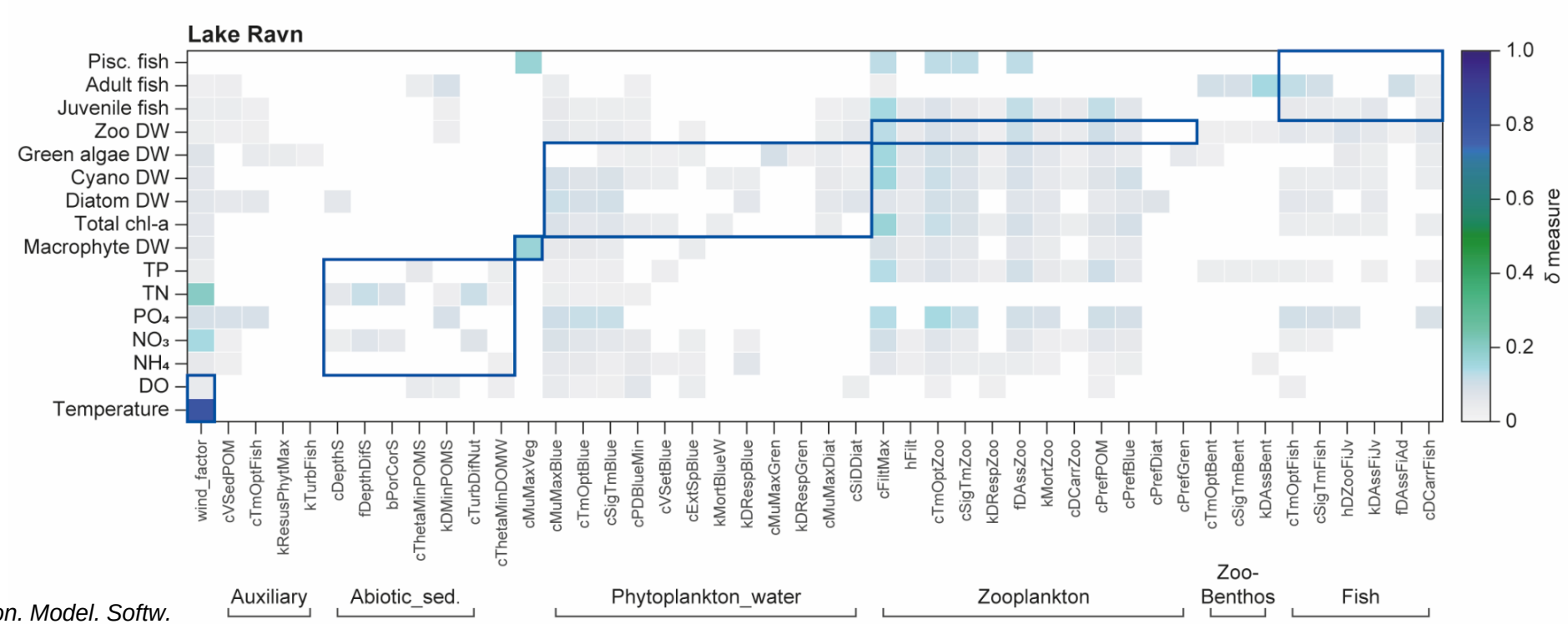
Compute sensitivity indices



SA with parsac

Andersen et al (2020) *Environ. Model. Softw.*

SA with parsac



Andersen et al (2020) *Environ. Model. Softw.*

SA challenges to be aware of

Here be dragons!

Type III errors can be FOOLISH

Type III errors can be DANGEROUS

US government decided to build a repository for radioactive waste in the Yucca Mountains.

A model named Total System Performance Assessment (TSPA) was used for the safety analysis computations. TSPA is composed of 286 sub-models and should predict thousand of years into the future.

Key assumption: Low permeability of the geological formation.
But when construction began, new evidence led to an upward revision of four orders of magnitude of the permeability parameter!

”According to the authors the error was due to the modelling of the granite formation as a homogeneous medium, while a fissures and faults model of the same formation would have been more realistic.”

“It is important […] 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.”

Pilkey and Pilkey-Jarvis, (2007), p.25

Type III errors can be DANGEROUS

US government decided to build a repository for radioactive waste in the Yucca Mountains.

A model named Total System Performance Assessment (TSPA) was used for the safety analysis computations. TSPA is composed of 286 sub-models and should predict thousand of years into the future.

Key assumption: Low permeability of the geological formation.

But when construction began, new evidence led to an upward revision of four orders of magnitude of the permeability parameter.

”According to the authors the error was due to the modelling of the granite formation as a homogeneous medium, while a fissured and faulty model of the same formation would have been more realistic.”

Sensitivity analysis is <u>not</u> immune to Type III errors

“It is important […] 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.”

Pilkey and Pilkey-Jarvis, (2007), p.25

Unquantifiable uncertainties…

Main dimensions of uncertainty in complex environmental models are

- Technical (inexactness)
- Methodological (unrealiability)
- Epistemological (ignorance)
- Societal (social robustness)

"Quantitative methods only address the technical dimension only. They can, however, be complemented with new qualitative approaches addressing aspects of uncertainty that are hard to quantifiy and were largely unadressed in the past." **van der Sluijs et al, 2005**

Sensitivity auditing

An entire research field dedicated to understanding uncertainty in models. This is (maybe even more) important when models are used beyond research interest e.g. in consultancy and to inform policy. Sensitivity auditing belongs to the tradition of **Post-Normal Science**.

1. Is the model redundant?
2. Are there important implicit assumptions?
3. Is uncertainty instrumentally amplified or compressed?
4. Was a sensitivity analysis performed prior to publication of the inference?
5. Is the model transparent?
6. Does the model address the right question?
7. Was sensitivity analysis performed holistically?

A brief introduction to sensitivity audting can be find in Saltelli & Funtowicz (2015): *Evidence-based policy at the end of the Cartesian dream: The case of mathematical modelling*

There are many opinions (and feelings) about models

Because models influence decisions, which in turn influence people's lives.

LandbrugsAvisen

Søg her

FORSIDE KVÆG MARK SVIN TEKNIK ØKONOMI NOTERINGER DEBAT EJENDOMSMARKED JOBLANDET

Mathematical models misregulate agriculture

ESTE NYHEDER

rnes tillid til økonomien dykker (10:22)

n billigere end F1-lån ved indfrielse (10:19)

Fejrer 40 år med ny plov (10:14)

Vindmøllemodstandere sure over "svinestreg" i

Torsdag 26. februar 2015 | 21:04 Skrevet af Frederik Thalbitzer Udskriv

" The entire environmental regulation is dependent on a model of emissions that does not match reality …"

POLITIKEN ≡ SEKTIONER Q SØG ♥ DIT POLITIKEN MERE

PÅ UDVALGTE STEL

*GÆLDER VED KØB AF KOMPLET BRILLE 19. – 31. JANUAR 2016 OG KAN IKKE KOMBINERES MED ANDRE TILBUD OG RABATTER.

2050

Analyse Feature Grafik 2050 forklarer Klima Energi Grøn omstilling

Farmer demands sampling – not modelling

Landmand får lov til at gøde mere af kommunen. Konkrete målinger viser, at han forurener langt mindre, end staten siger.

SKÅL! Byrådsmedlem Asger Møller Madsen (V) (t.v.) og landmand Carsten Søborg (t.h.) skåler i drænvandet fra Carsten Søborgs marker. Målinger fra Aarhus Universitet viser, at det er tre gange renere end kravet til drikkevand. - Foto: SØREN SCHNOOR

Gem + Følg

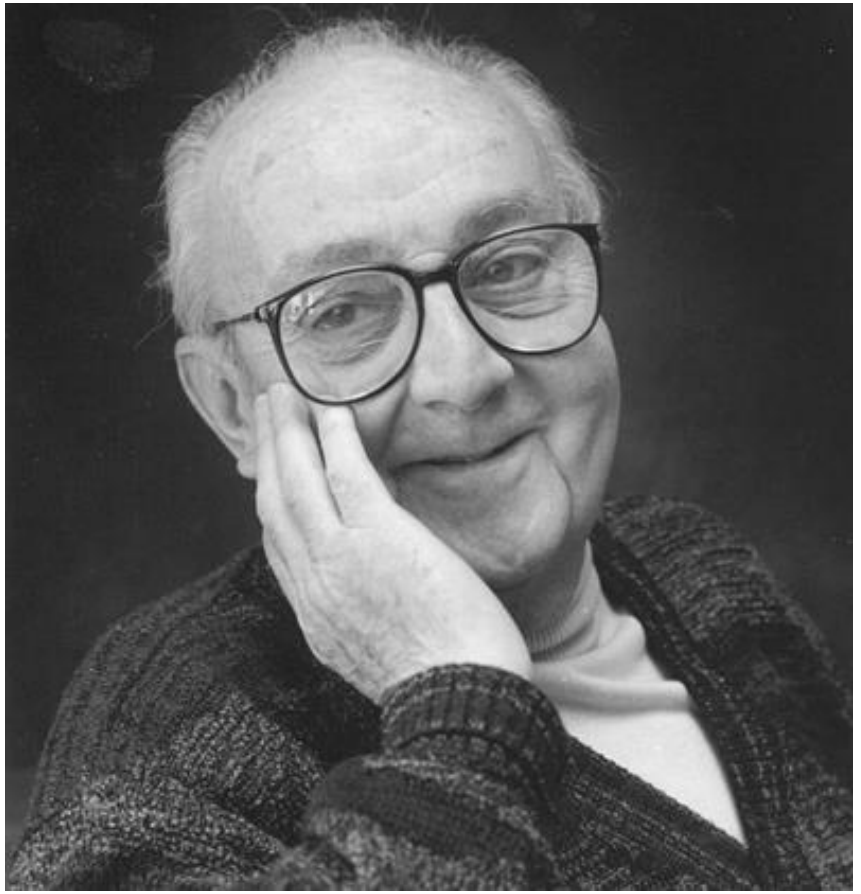
THOMAS FÆRGEMAN
Journalist

»Skål!« Vi lægger halsen tilbage og

HVAD ER FORSKELLEN?

“All models are wrong”

“So the questions you need to ask is not “Is the model true?” (it never is) but “Is the model good enough for this particular application?”



George Box

Welcome to Model Land

*Which **assumptions** did you take when entering ?*

Reality

New insights
Decision-support

Political process

Decisions

Model Land

All assumptions are true.
In here, mathematics rule.

Welcome to Model Land

*Which **assumptions** did you take when entering ?*

Reality

New insights
Decision-support

Political process

Decisions

Model Land

All assumptions are true.
In here, mathematics rule.

*How do we provide **confidence** in models?*
*How to handle (deep) **uncertainties**?*

Welcome to Model Land

*Which **assumptions** did you take when entering ?*

Reality

Other relevant information

New insights

Decision-support

Political process

Decisions

Model validation

*How do we provide **confidence** in models?*

*How to handle (deep) **uncertainties**?*

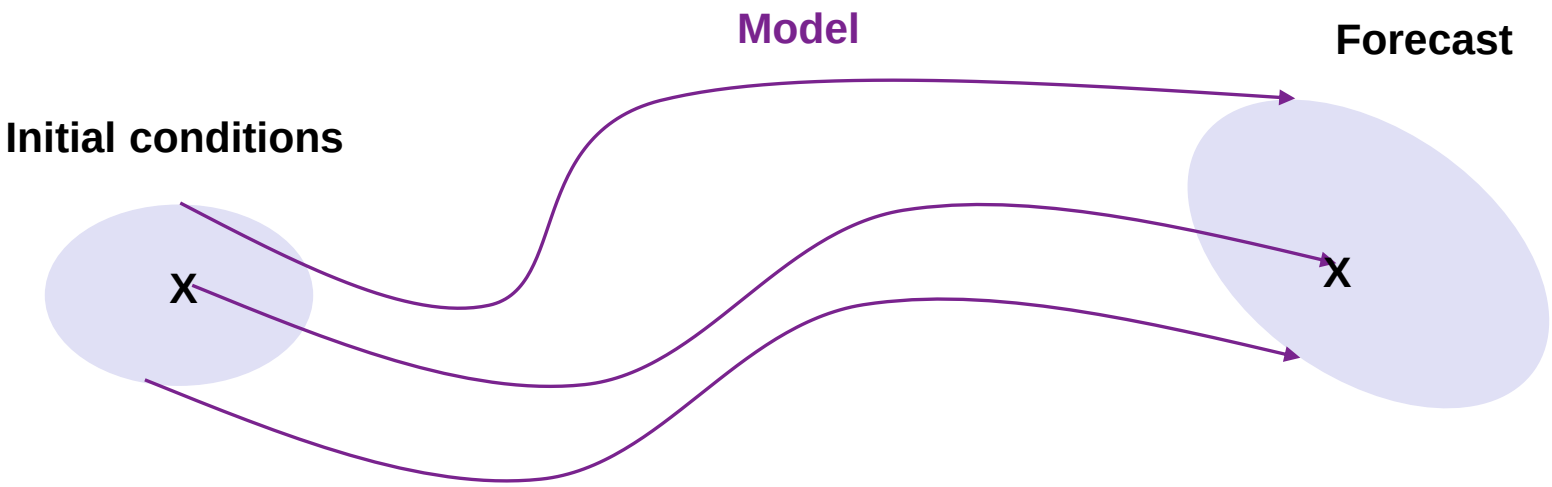
Model Land

All assumptions are true.

In here, mathematics rule.

Sensitivity analysis

Butterflies and Hawkmoths



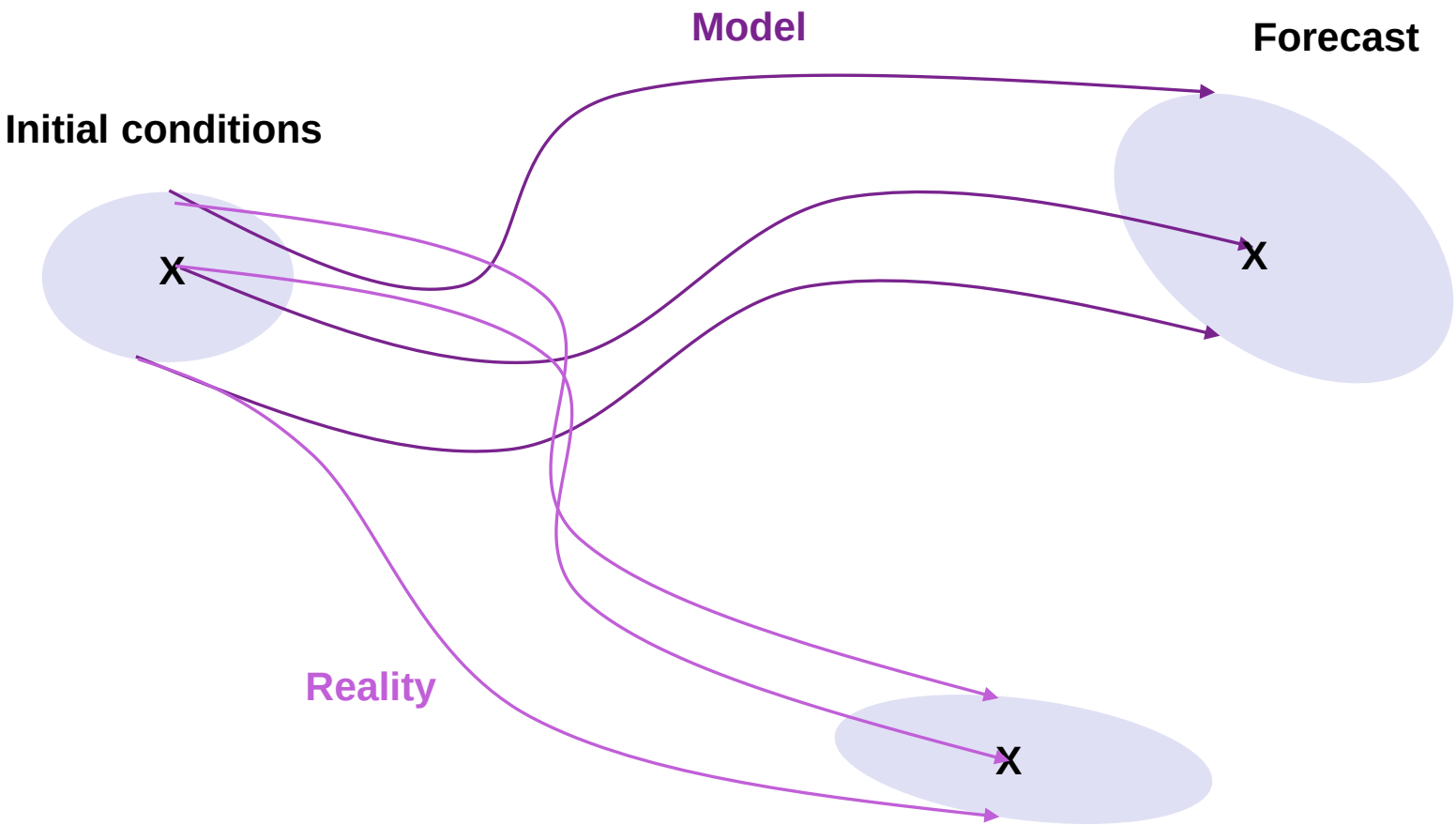
Butterfly Effect

In Model Land

Uncertainties in initial conditions
Limits timescale

Heavily inspired by slides from Erica Thompson and her book *Escape from Model Land*

Butterflies and Hawkmoths



Butterfly Effect

In Model Land

Uncertainties in initial conditions
Limits timescale

Hawkmoth Effect

In the Gap

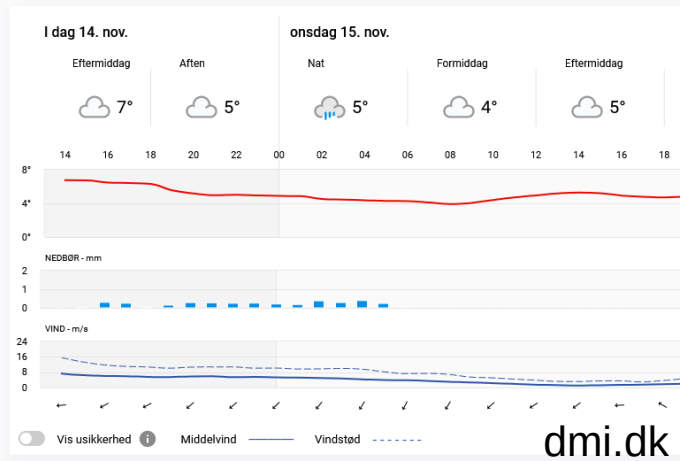
Structural uncertainties in model
Limits timescale **and** accuracy

Heavily inspired by slides from Erica Thompson and her book *Escape from Model Land*

When doing scenarios: Weather- or climate-like predictions?

Weather-like prediction tasks:

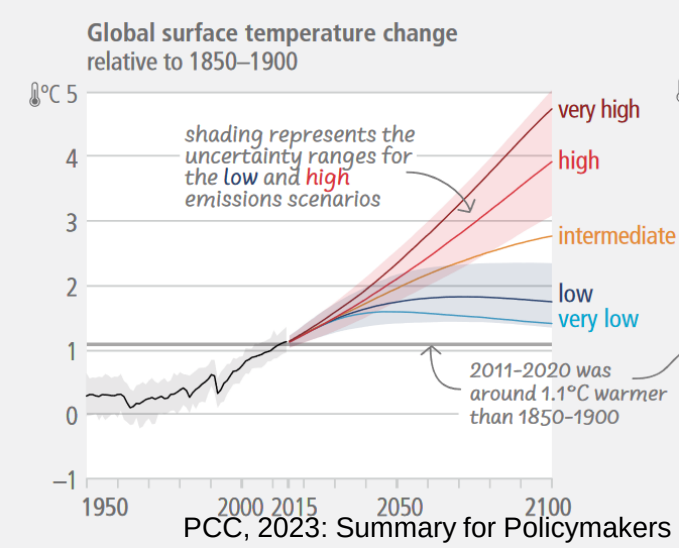
- similar decisions, in similar contexts, are made frequently (e.g. daily, monthly);
- a size-able forecast-outcome data archive available;
- the models involved have a long lifetime relative to the lead time of the forecast.



dmi.dk

Climate-like prediction tasks:

- the decision is effectively one-off, and the decision not to act has potential significant costs;
- the forecast-outcome archive is effectively empty (task is extrapolation)
- the model has a short lifetime compared to the lead time of the forecast.



PCC, 2023: Summary for Policymakers

Welcome to Model Land

*Which **assumptions** did you take when entering ?*

Reality

Other relevant information

New insights

Decision-support

Political process

Decisions

Model validation

*How do we provide **confidence** in models?*

*How to handle (deep) **uncertainties**?*

Model Land

All assumptions are true.

In here, mathematics rule.

Sensitivity analysis

Welcome to Model Land

*Which **assumptions** did you take when entering ?*

Reality

Other relevant information

New insights

Decision-support

Political process

Decisions

Model validation

*How do we provide **confidence** in models?*

*How to handle (deep) **uncertainties**?*

Model Land

All assumptions are true.

In here, mathematics rule.

Sensitivity analysis

Supplementary reading

- Saltelli et al (2020) Five ways to ensure that models serve society: a manifesto, Nature
- Borgonovo et al (2017) Making the most out of a hydrological model data set: Sensitivity analyses to open the model black-box, Water Resources Research
- Razavi, S., Jakeman, A., Saltelli, et al (2021) The Future of Sensitivity Analysis: An essential discipline for systems modeling and policy support, Environmental Modelling and Software

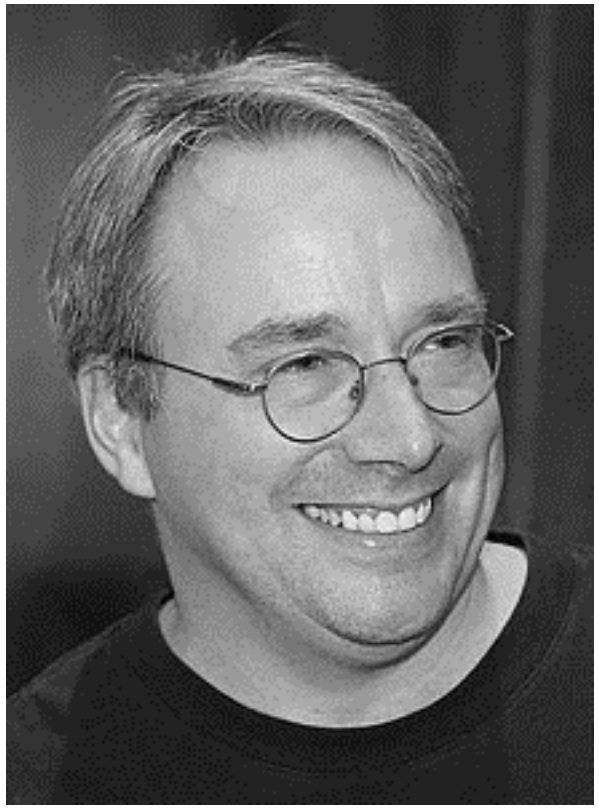
Additional notes

Do open source (when possible)

"*Given enough eyeballs, all bugs are shallow*"

Linus' Law

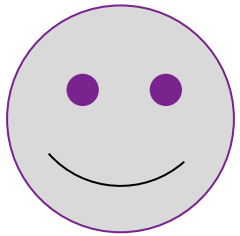
(by Eric S. Raymond)



Linus Thorvalds

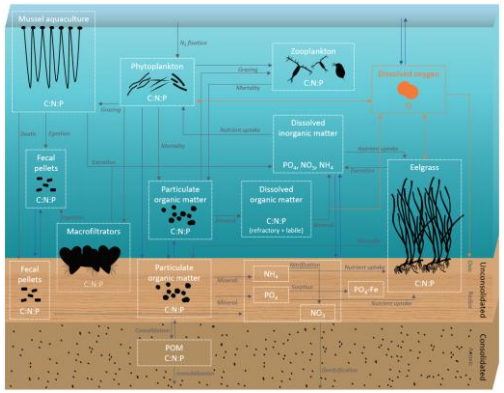
Use of expert knowledge to judge models

What is the relationship between models and experts?



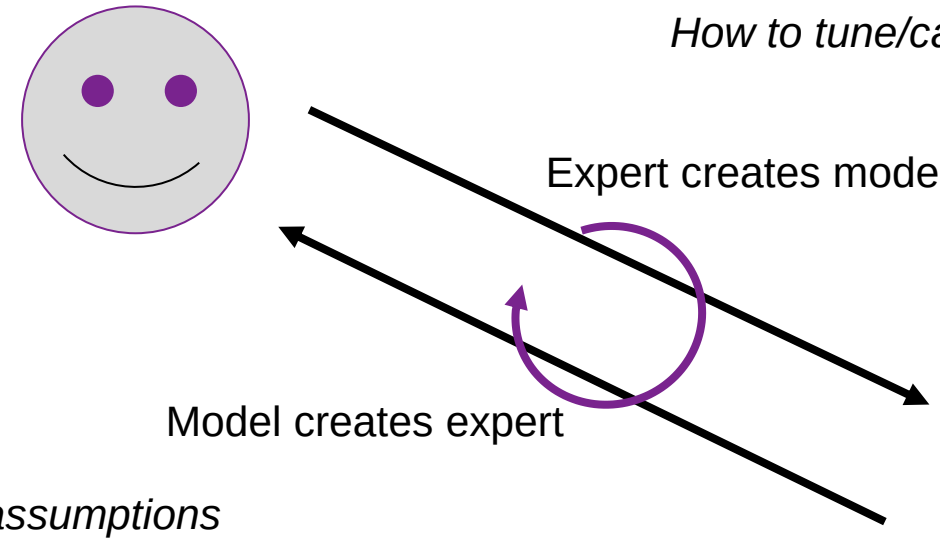
What to include? How to represent?
How to tune/calibrate? How to evaluate?

Expert creates model



Use of expert knowledge to judge models

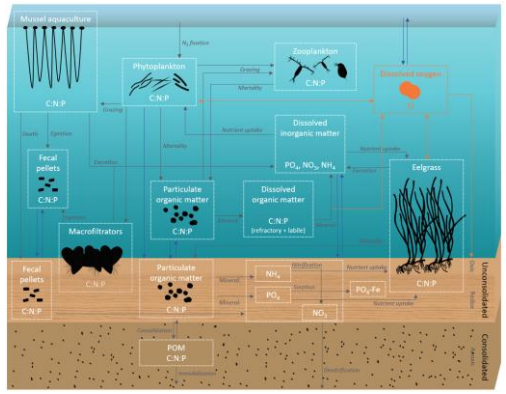
What is the relationship between models and experts?



What to include? How to represent?
How to tune/calibrate? How to evaluate?

Expert creates model

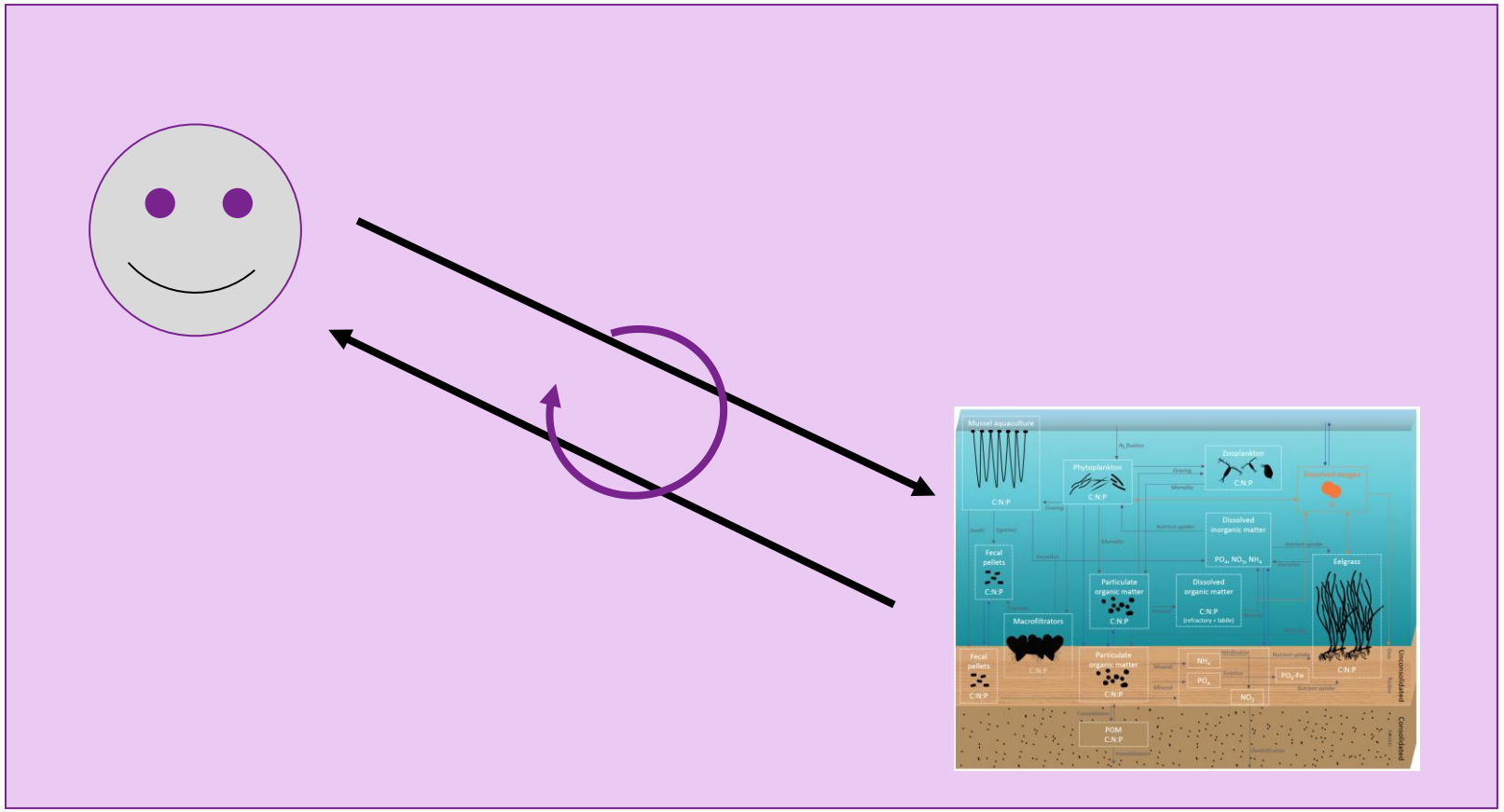
Model creates expert



Play with assumptions
Test hypothesis
Separate system components
Make predictions

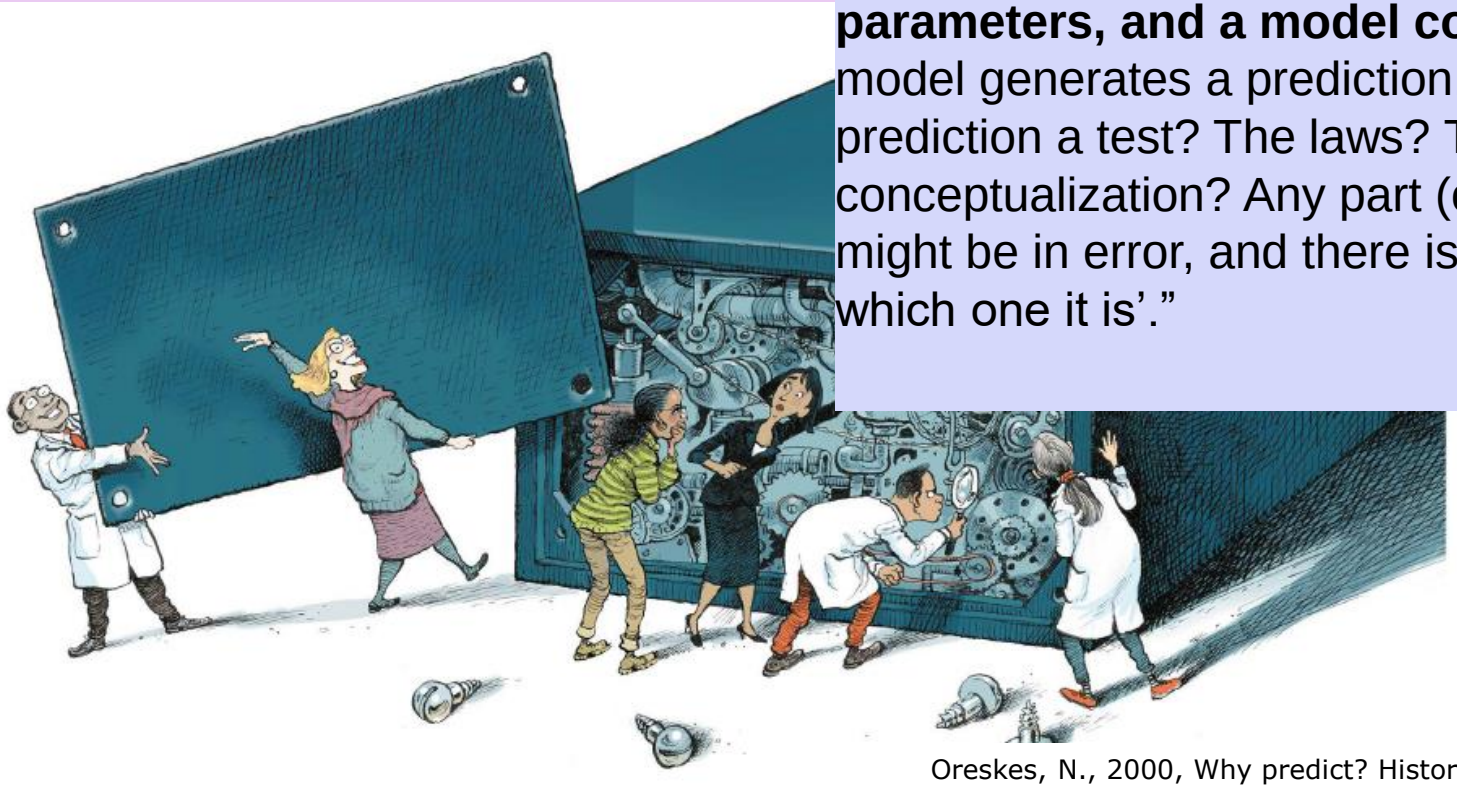
Use of expert knowledge to judge models

Be mindful of expert system



Use of expert knowledge to judge models

Be mindful of expert system



"[…] **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'."

Naomi Oreskes, 2000

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.