Uncertainty and sensitivity analysis: who, how and especially why!

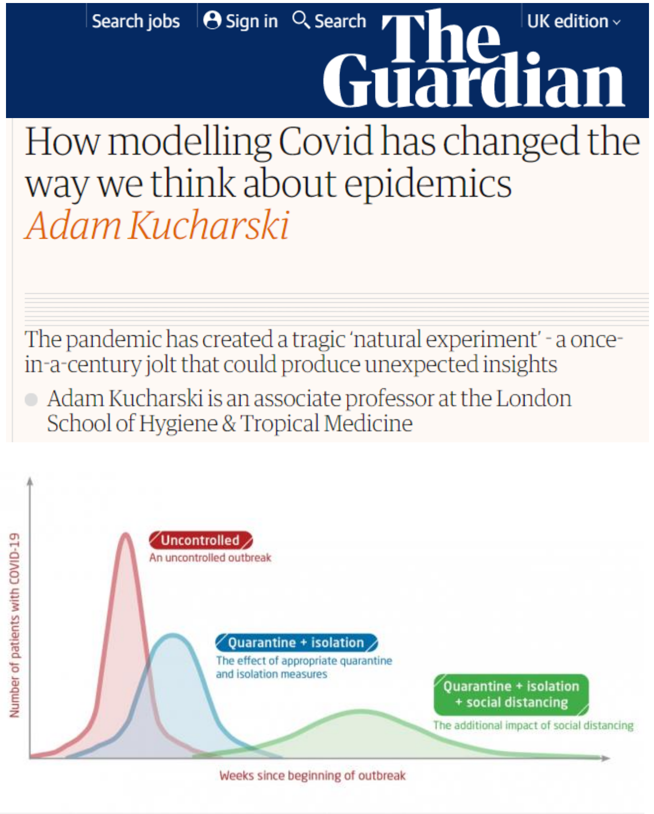
Francesca Pianosi
University of Bristol

Access to software, papers, example: https://safetoolbox.github.io/

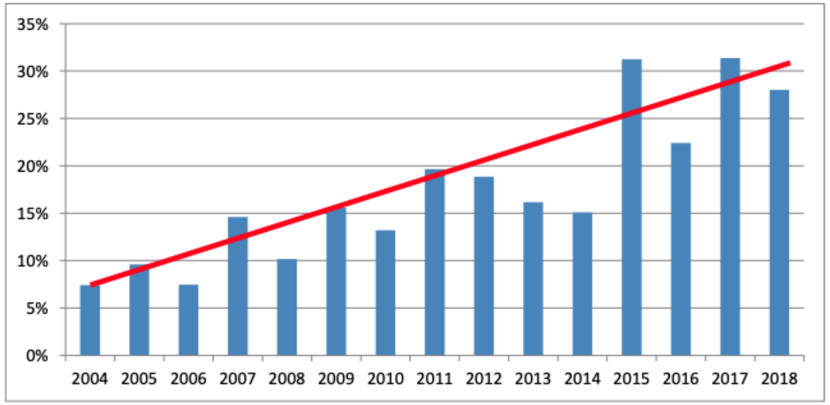
Pianosi is supported through grant ST/Y003713/1

francesca.pianosi@bristol.ac.uk

Mathematical models are increasingly used to inform decisions in a variety of sectors



Share of EC Impact Assessments supported by modelling



[JRC 2019 *Modelling for EU Policy support*]

What do these models have in common?

"*All models are wrong, but some are useful*" George Box (1976)

francesca.pianosi@bristol.ac.uk

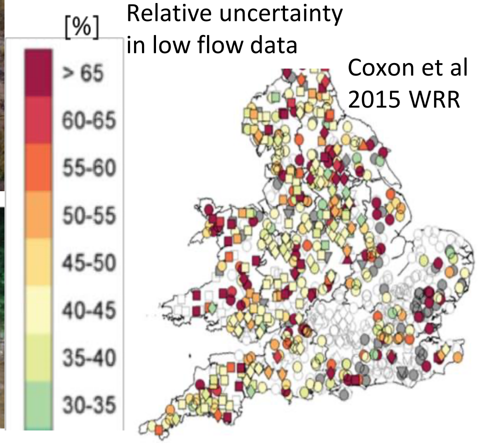
Model outputs are conditional on many uncertain assumptions about the system properties and drivers

- The data we use to build and test models have errors and gaps

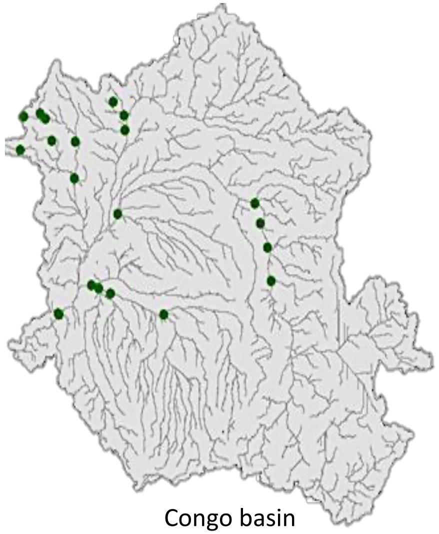
Example: river flow data

Relative uncertainty in low flow data

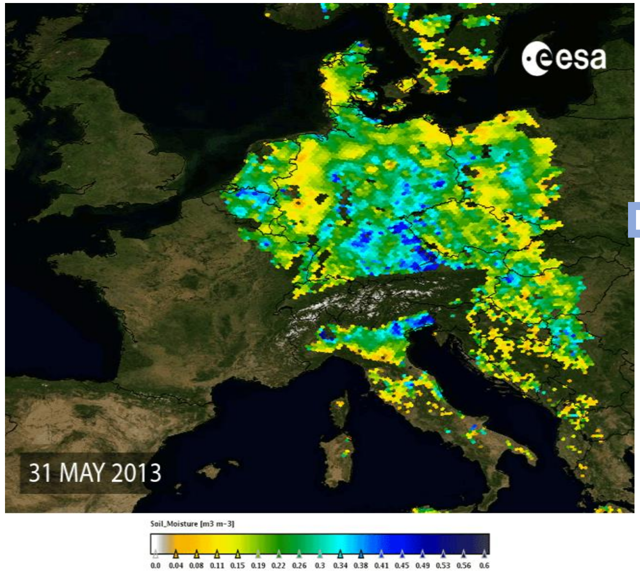
Coxon et al 2015 WRR



England & Wales (242,500 km^2)



Congo basin (3,700,000 km^2)

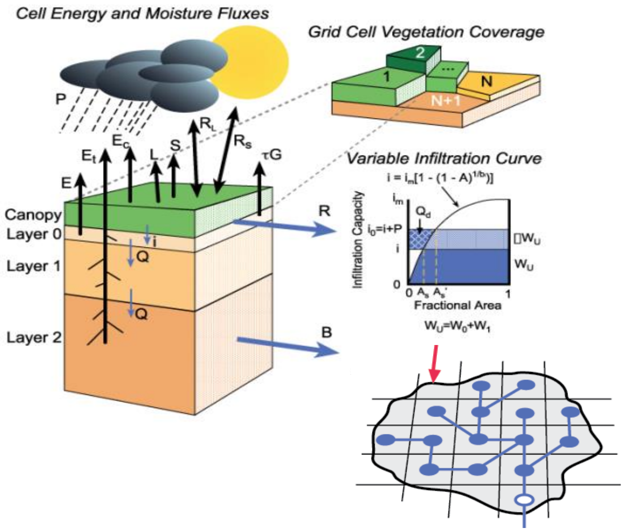


francesca.pianosi@bristol.ac.uk

Model outputs are conditional on many uncertain assumptions about the system properties and drivers

- The data we use to build and test models have errors and gaps
- Models use simplifying assumptions whose adequacy is uncertain

Water flowing over and through hillslopes



Rainfall-runoff processes represented in a mathematical model (VIC)

These uncertainties can be "aleatory" (=due to intrinsic randomness, e.g. from noise in measurement devices) or "epistemic" (=due to lack of knowledge, e.g. on the mechanics of a particular process)

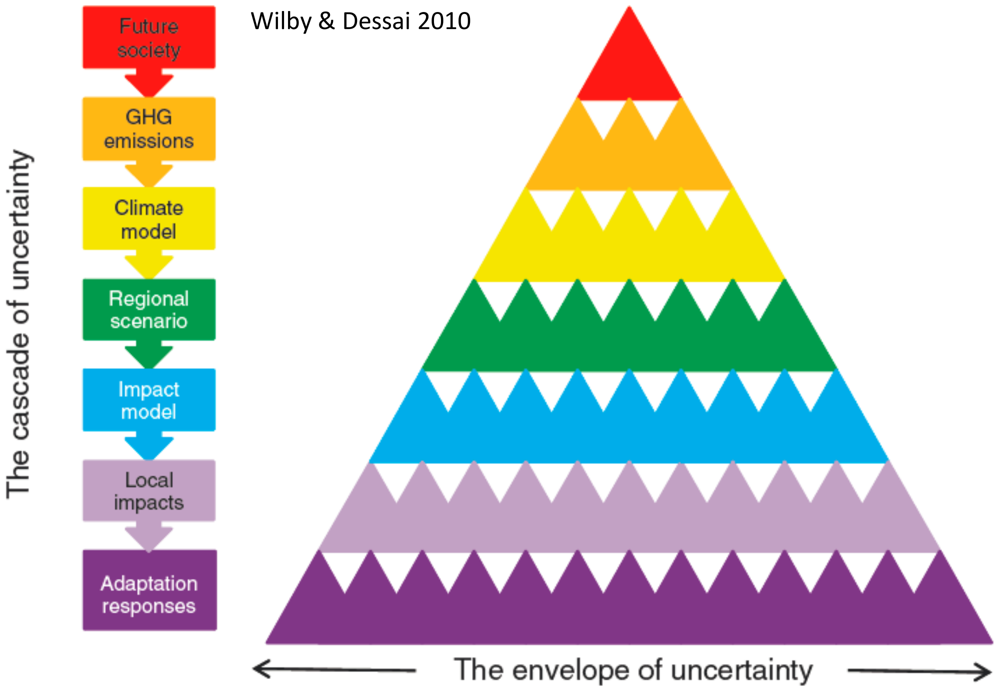
Aleatory and epistemic uncertainties are - to some extent - quantifiable and reducible

francesca.pianosi@bristol.ac.uk

Model outputs are conditional on many uncertain assumptions about the system properties and drivers *now*, and about how they will evolve *in the future*

- The data we use to build and test models have errors and gaps
- Models use simplifying assumptions whose adequacy is uncertain
- If we use models to simulate long-term system behaviour, then we need to build scenarios for how the system drivers will evolve in the future

"it's difficult to make predictions, particularly when they concern the future" (Danish proverb)



"Deep uncertainty" refers to a situation where there is no consensus regarding the appropriate models or probability distributions that should be used to conceptualise or represent a particular process or problem

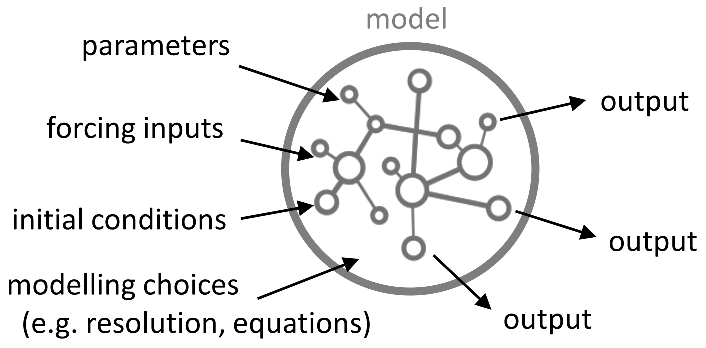
Deep uncertainty is unquantifiable and irreducible

francesca.pianosi@bristol.ac.uk

Uncertainty in >>> Uncertainty out

uncertainty
/ʌnˈsəːtnti/
noun
something that is uncertain or that causes one to feel uncertain
plural noun: **uncertainties**

“**input uncertainties**” when the ‘right’ value or choice for some of the model input factors is uncertain



“**output uncertainty**” as a consequence of input uncertainties

francesca.pianosi@bristol.ac.uk

Uncertainty in >>> Uncertainty out
So, what do we do with it?

- Ignore it

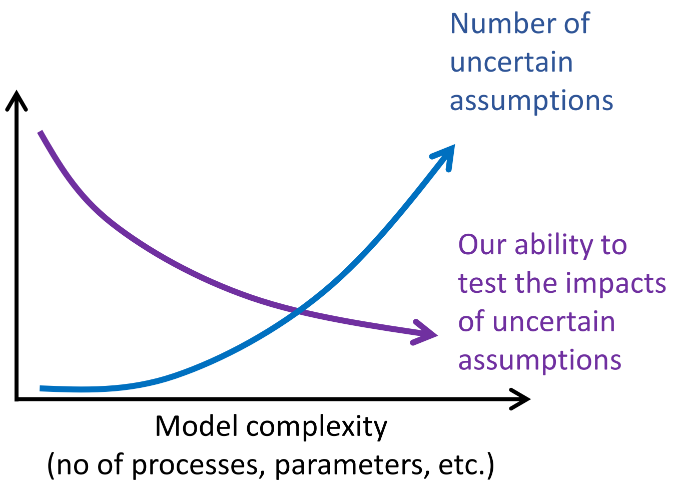
Uncertainty in >>> Uncertainty out
So, what do we do with it?

- Ignore it
- Fight it

  - more monitoring, fieldwork, experiments, thinking, etc.: yes 🤓
  - more detailed and "realistic" models: meh 🤔

The complexity paradox: *"A complex model may be more realistic, yet it is ironic that as we add more factors to a model, the certainty of its predictions may decrease even as our intuitive faith in the model increases (…) The more we strive for realism by incorporating as many as possible of the different processes and parameters that we believe to be operating in the system, the more difficult it is for us to know if our tests of the model are meaningful"*

Oreskes 2003 *The role of quantitative models in science*



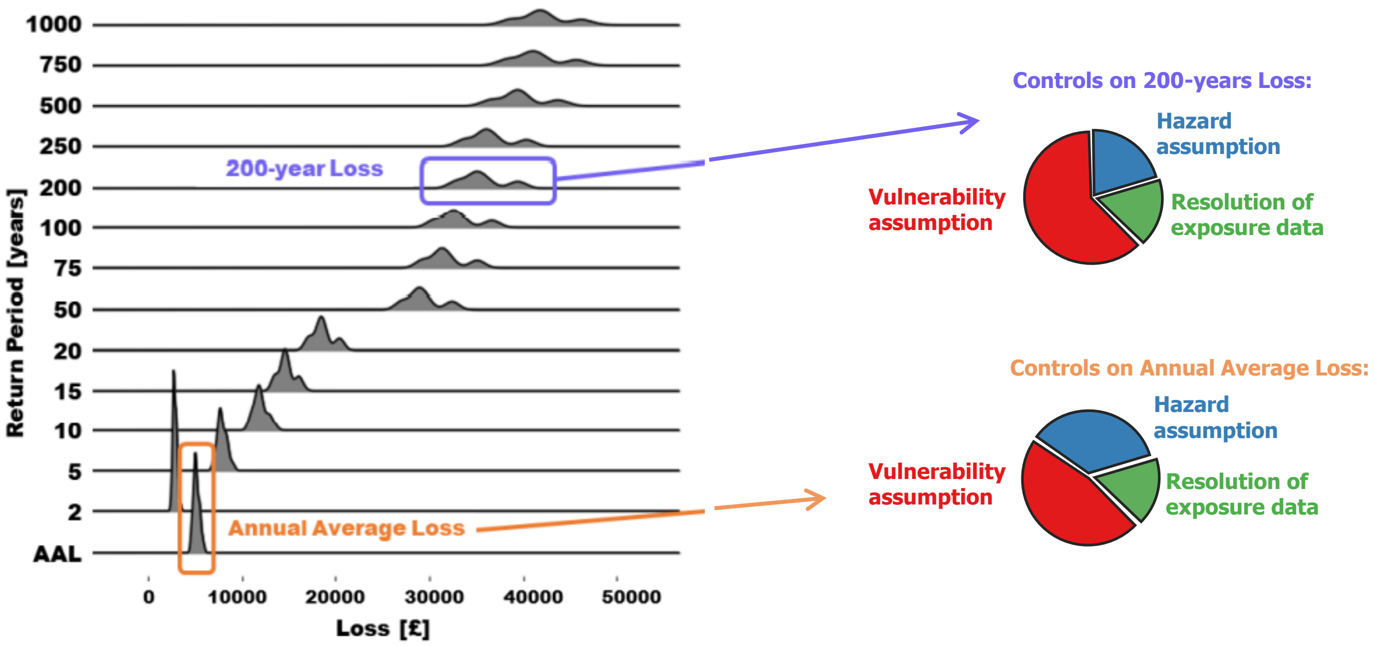
francesca.pianosi@bristol.ac.uk

Uncertainty in >>> Uncertainty out
So, what do we do with it?

- Ignore it
- Fight it
- Acknowledge it

Uncertainty analysis:
quantify uncertainty in the model output(s) as a consequence of known (assumed) input uncertainties

Sensitivity analysis:
assess the relative contribution of different input uncertainties to the output(s) uncertainty



francesca.pianosi@bristol.ac.uk

What kind of questions can we answer through uncertainty and sensitivity analysis and who are these interesting for?

- Questions to prioritise efforts for model improvement

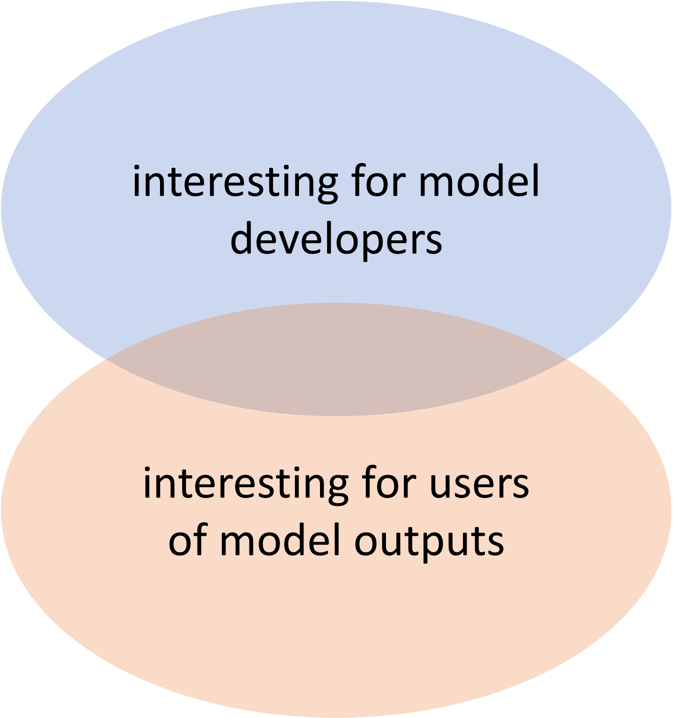
What are the input data, parameters or model components that control the model outputs uncertainty most and where reduction of uncertainty would be most beneficial?

- Questions to evaluate models and establish they are fit for purpose

Does the “right” inputs/components control the “right” outputs? Are model outputs sufficiently controlled by the “decision levers” relative to other input uncertainties?

- Questions about the systems behaviour (under the assumption: system=model)

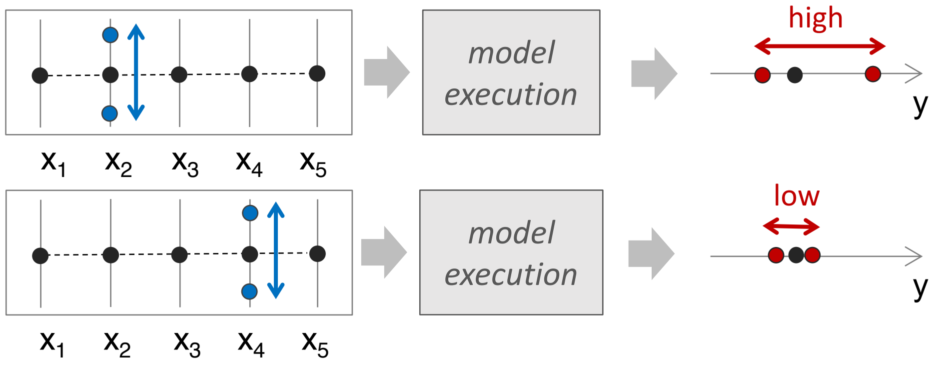
What are the key drivers of the system in the face of deep uncertainty? Are there “robust” decisions that work sufficiently well across a range of uncertainties?



francesca.pianosi@bristol.ac.uk

Two ways of approaching Uncertainty and Sensitivity Analysis

Local approaches: investigate the sensitivity of the output (y) to varying the uncertain inputs (x_1,x_2,….) **one at the time** around a **baseline**



Limitation:

- ignore interactions
- results only valid for the chosen baseline: output uncertainty and sensitivity may be very different for a different baseline

Global approaches: investigate the sensitivity of the output (y) to varying all uncertain inputs **simultaneously** across their entire variability space

francesca.pianosi@bristol.ac.uk

Global approaches are based on repeated executions of the model against different inputs' combinations and a statistical analysis of the resulting input-output dataset

[1] characterize uncertainty in input factors

[2] propagate input uncertainties through the model

[3] quantify output uncertainty

[4] attribute output uncertainty

RUNNING MONTE-CARLO SIMULATIONS

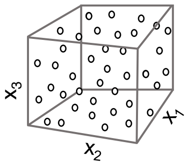
ANALYSING INPUT-OUTPUT DATASET

characterize uncertainty in the inputs

(e.g. via a list of possible values, variability range, probability distribution.)

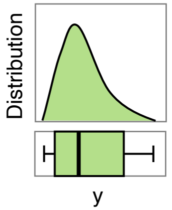
sample N combinations of inputs

(e.g. via random sampling, Latin Hypercube sampling)



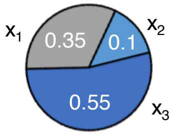
execute the model against each inputs' combination

derive the output(s)' distribution or CIs



calculate sensitivity indices

contribution to output uncertainty



- Elementary Effects method
- Variance-based (Sobol') method
- etc.

francesca.pianosi@bristol.ac.uk

WHY

doing UA/SA?

francesca.pianosi@bristol.ac.uk

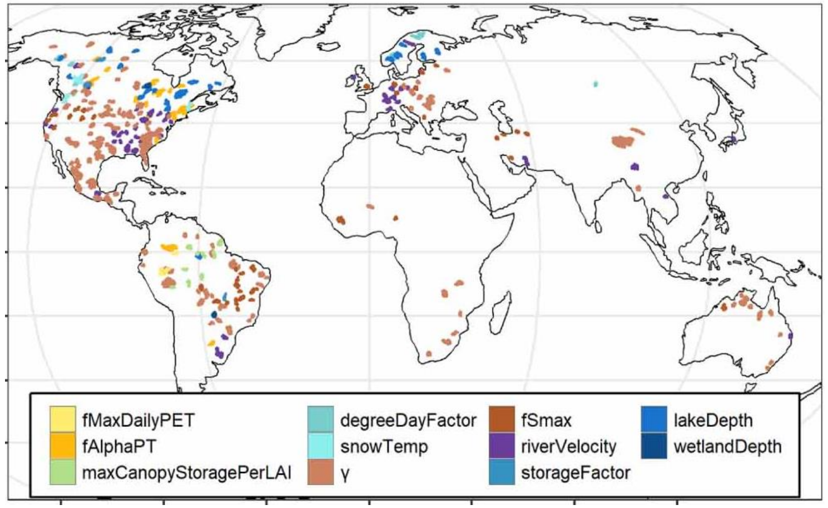
Guiding model calibration

Application to global hydrological model WaterGAP3

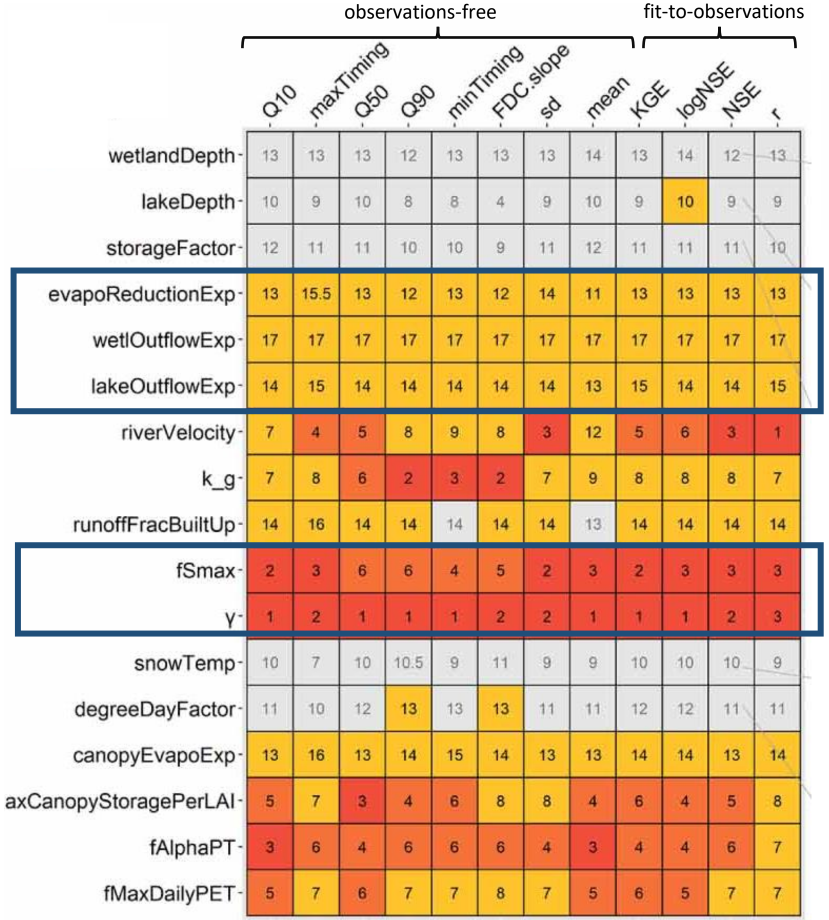
Kupzig et al 2023
Towards parameter estimation in global hydrological models *Environ. Res. Lett.*
doi:10.1088/1748-9326/acdae8

Which parameters are consistently important across output metrics and/or spatial domain and thus should be the focus of computationally-expensive calibration?

Identifying most influential parameter on NSE of daily streamflow across 347 gauged catchments



Identifying most influential parameters on different output metrics



Ranks
≥7 [less important]
4 - 6
≤ 3 [more important]

Some parameters are consistently less important across metrics and basins → exclude from calibration

Some parameters are consistently very important across metrics and basins → prioritise for calibration

francesca.pianosi@bristol.ac.uk

Prioritizing efforts for uncertainty reduction

Application to natural risk assessment model

Dawkins et al 2023
Quantifying uncertainty and sensitivity in climate risk assessments: Varying hazard, exposure and vulnerability modelling choices *Clim. Risk Man.* **doi:** 10.1016/j.crm.2023.100511

Which source of uncertainty influence the precision of model output the most?
Where efforts for model/data improvement will have most significant effects towards improving precision?
(... and where they won't?)

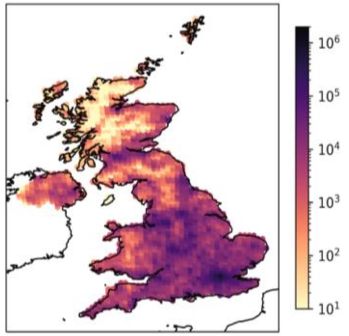
hazard maps | vulnerability curve | exposure map

↓

risk calculation
vulnerability(hazard_intensity) x exposure

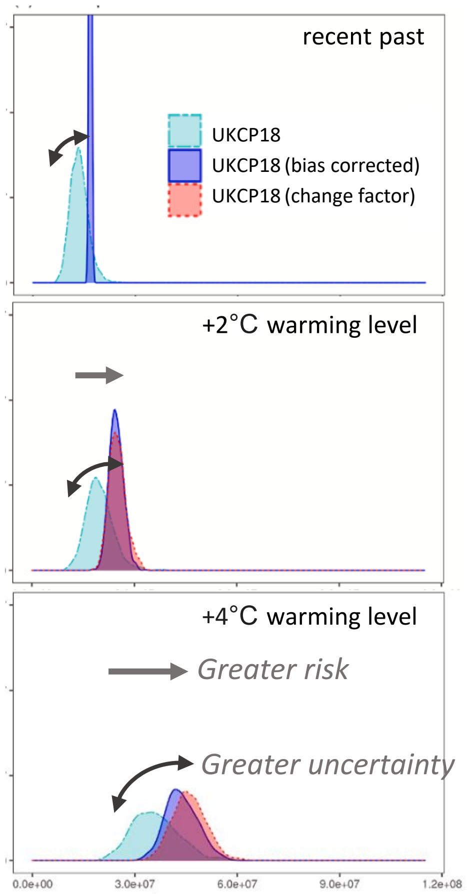
↓

aggregation over time (events) and/or space



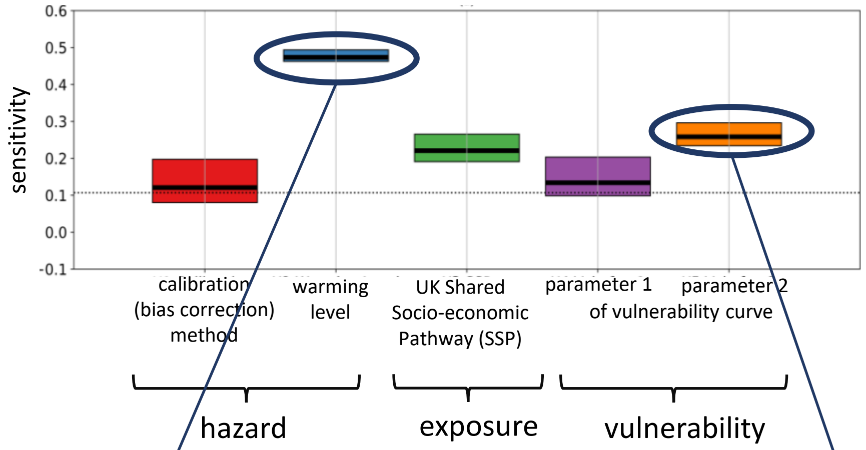
Risk: Expected Annual Impact (EAI): number of days of work lost due to heat stress

Quantifying future risk
(using different hazard maps)



Risk: spatially aggregated EAI

Quantifying sensitivity of risk (mean of spatially aggregated EAI) to input uncertainties



warming level most influential factor

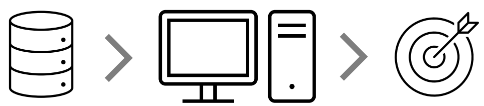
vulnerability curve parameter 2nd most influential (more than bias correction!)

francesca.pianosi@bristol.ac.uk

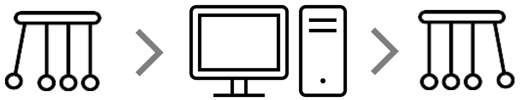
Model evaluation

Wagener et al 2022
On the evaluation of climate change impact models *WIREs-CC*
doi:10.1002/wcc.772

"Data-based" evaluation: → "Response-based" evaluation:



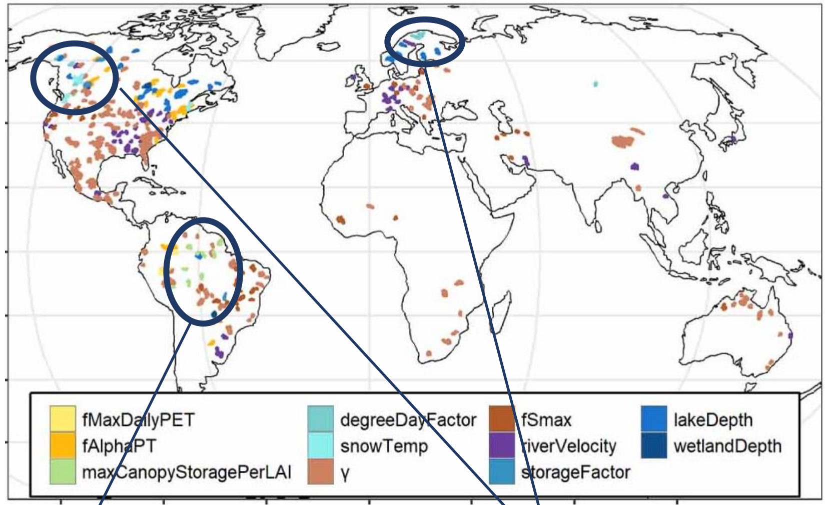
Fit-to-data: are model outputs consistent with observations?



Is the model input-output response consistent with our understanding of the systems functioning?

Is the most influential parameter the "right one" at the "right place"?

Is parameter importance explained by catchment attributes and/or climate zones?



CanopyStorage key in forested catchments

☺

SnowTemp key in snowmelt dominated catchments

Is the "right" parameter most important on the "right" output metric?

	Q10											
wetlandDepth	13	13	13	12	13	13	13	14	13	14	12	13
lakeDepth	10	9	10	8	8	4	9	10	9	10	9	9
storageFactor	12	11	11	10	10	9	11	12	11	11	11	10
evapoReductionExp	13	15.5	13	12	13	12	14	11	13	13	13	13
wetlOutflowExp	17	17	17	17	17	17	17	17	17	17	17	17
lakeOutflowExp	14	15	14	14	14	14	14	13	15	14	14	15
riverVelocity	7	4	5	8	9	8	3	12	5	6	3	1
k_g	7	8	6	2	3	2	7	9	8	8	8	7
runoffFracBuiltUp	14	16	14	14	14	14	14	13	14	14	14	14
fSmax	2	3	6	6	4	5	2	3	2	3	3	3
γ	1	2	1	1	1	2	2	1	1	1	2	3
snowTemp	10	7	10	10.5	9	11	9	9	10	10	10	9
degreeDayFactor	11	10	12	13	13	13	11	11	12	12	11	11
canopyEvapoExp	13	16	13	14	15	14	13	13	14	14	13	14
axCanopyStoragePerLAI	5	7	3	4	6	8	8	4	6	4	5	8
fAlphaPT	3	6	4	6	6	6	4	3	4	4	6	7
fMaxDailyPET	5	7	6	7	7	8	7	5	6	5	7	7

Ranks
≥7 [less important]
4 - 6
≤ 3 [more important]

Evapotranspiration parameter (fAlphaPT) more important on high flow (Q10) than routing parameter (riverVelocity)

☹

Model structural error?

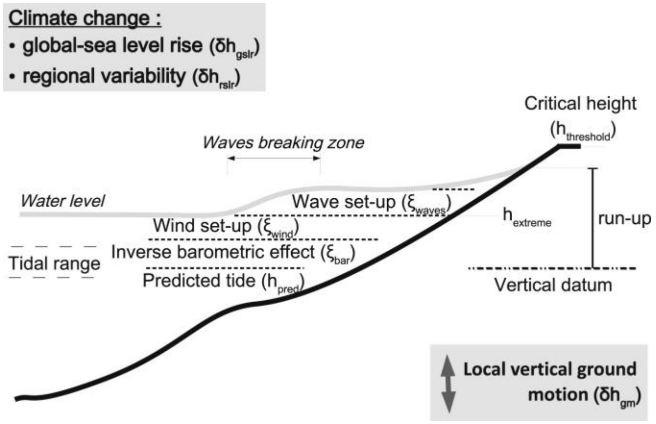
francesca.pianosi@bristol.ac.uk

Model evaluation

Wagener et al 2022
On the evaluation of climate change impact models *WIREs-CC*
doi:10.1002/wcc.772

Application to flood defence assessment model

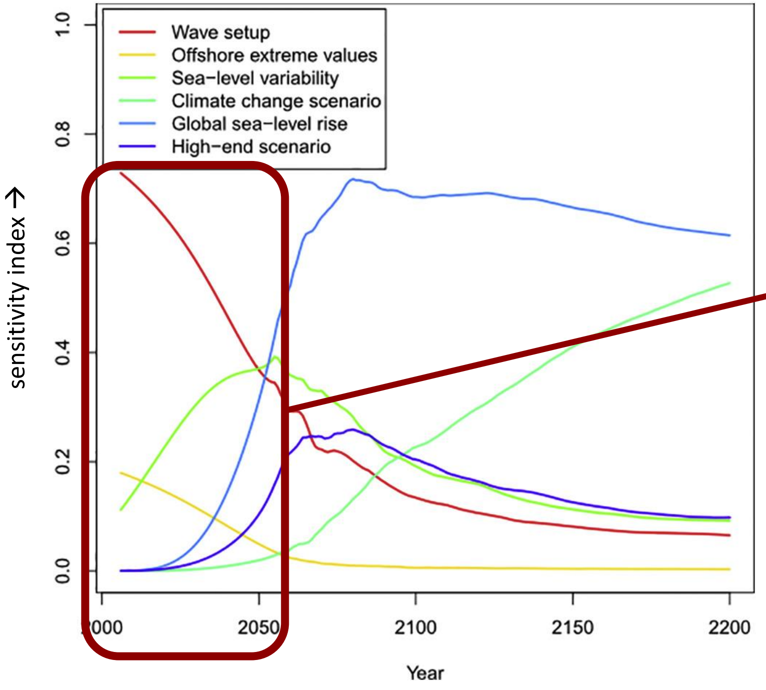
Le Cozannet et al 2015
doi: 10.1016/j.envsoft.2015.07.021



What are the dominant control of the model output?
Are model outputs sensitive to decision-relevant inputs?

If outputs are more strongly controlled by uncertain assumptions/parameters than by the policy/scenario inputs, then the model will tell us more about the consequences of the assumptions embedded in it than it will tell us about the different policy options/scenarios

Sensitivity of coastal defence vulnerability (= annual probability of exceeding the threshold height of coastal defences)



In the 'mid-term', the model predictions are controlled by the modeller's choice of the (very uncertain) 'wave setup' parameter way more than they are by climate change and sea level rise scenarios

→ the model should not be used for impact assessment on such temporal scales

francesca.pianosi@bristol.ac.uk

How

to do UA/SA?

francesca.pianosi@bristol.ac.uk

Characterizing uncertainty in input factors

What is the appropriate distribution / range for the uncertain inputs?

characterize uncertainty in the inputs	sample N combinations of inputs	execute the model against each inputs' combination	calculate output(s) ranges or distribution	calculate sensitivity indices

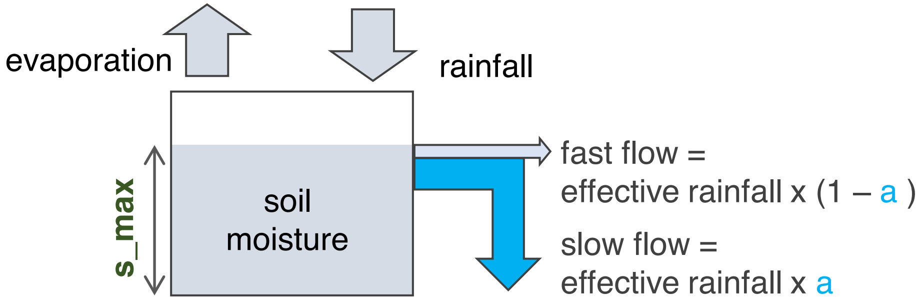
francesca.pianosi@bristol.ac.uk

Characterizing uncertainty in input factors

Depending on the type of input (scalar parameter, time series, discrete modelling choice, etc.) and on our level of uncertainty about it, we can use:
- a list of possible values
- a uniform distribution within an uncertainty range
- a probability distribution

… and define them using literature sources, historical observations, experts' judgment, etc.

Sometimes the range (distribution) is univocally defined by the physical meaning of the input, but most often different definitions are possible



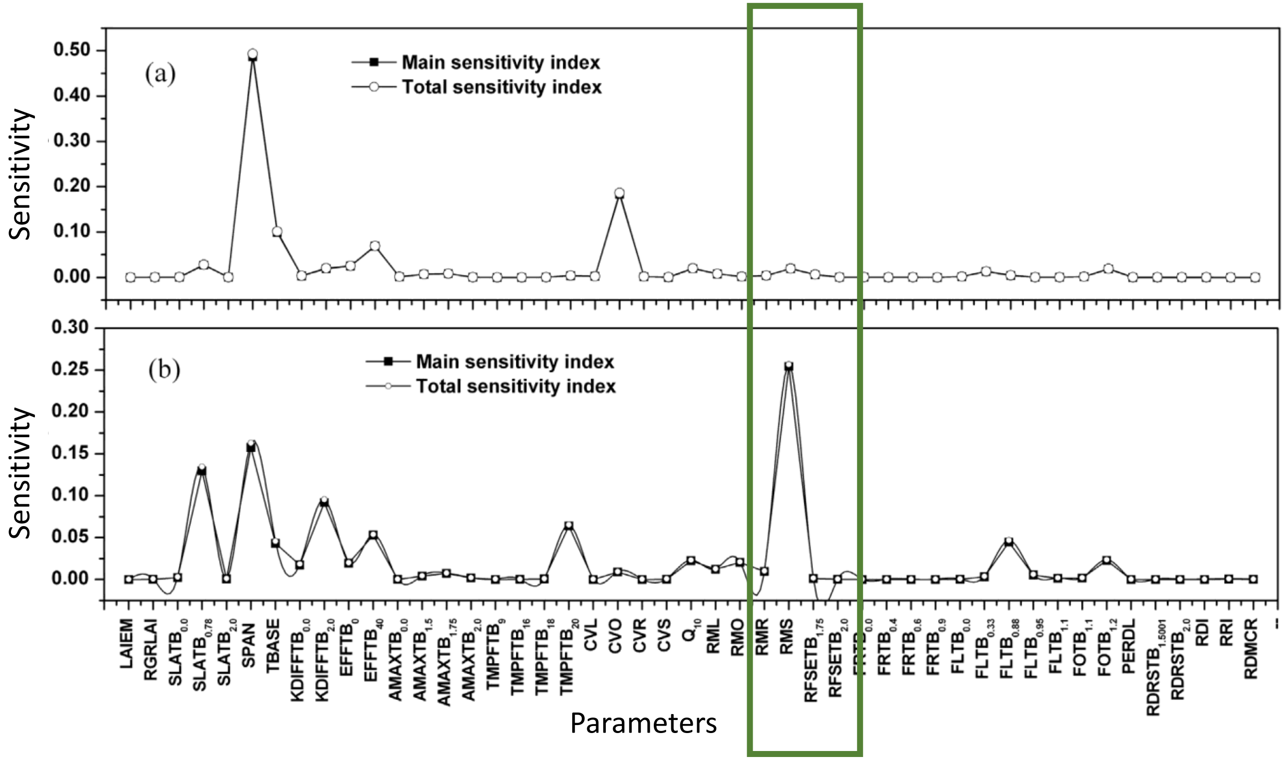
The maximum soil capacity (s_max) varies between 0 and an upper bound that may be difficult to define

The repartition coefficient (a) varies between 0 and 1 by definition

francesca.pianosi@bristol.ac.uk

When different definitions of the inputs ranges (distributions) are possible, the choice can significantly condition UA/SA results

Example from SA of a crop growth model (Wang et al EMS 2013)



Sensitivity of simulated crop yield based on 'literature' parameter ranges

Sensitivity of simulated crop yield based on 'expanded' parameter ranges

francesca.pianosi@bristol.ac.uk

Choosing the sampling strategy and size

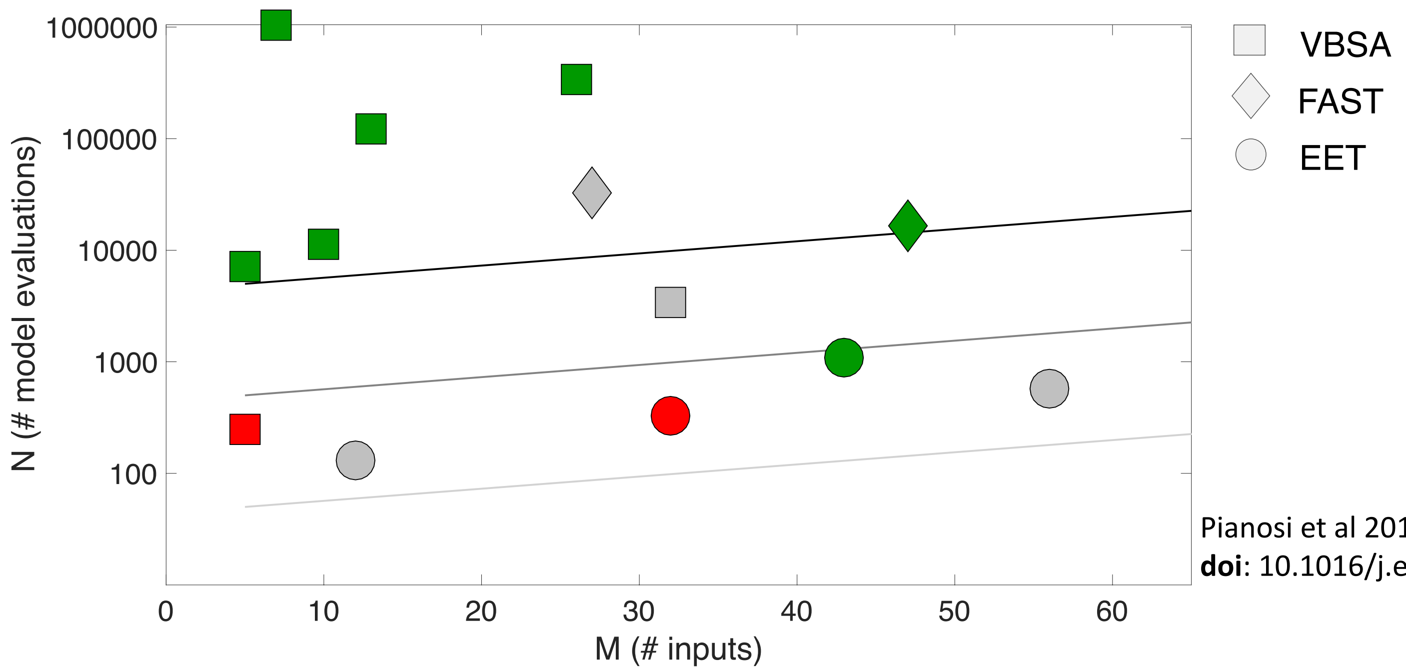
Which sampling technique to use (e.g. random sampling, Latin Hypercube, quasi-random sequences)?

How many samples are needed?

characterize uncertainty in the inputs	sample N combinations of inputs	execute the model against each inputs' combination	calculate output(s) ranges or distribution	calculate sensitivity indices

francesca.pianosi@bristol.ac.uk

In general, the required sample size (N) increases with the number of uncertain inputs (M). However, the proportionality rate varies significantly from one method to another, and from one application of the same method to another



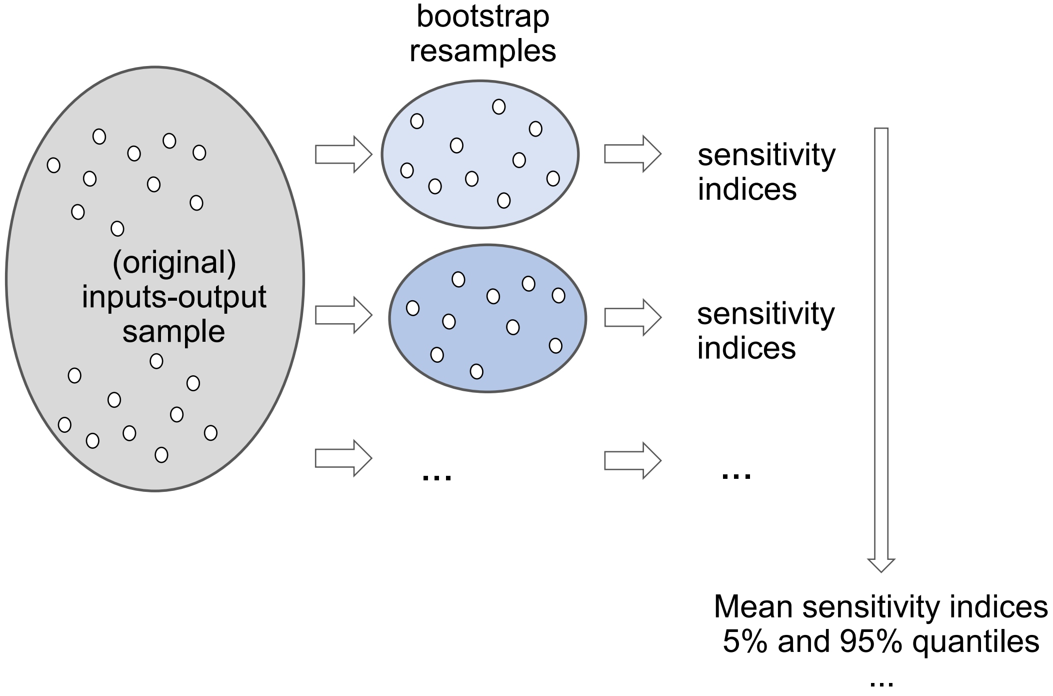
Pianosi et al 2016

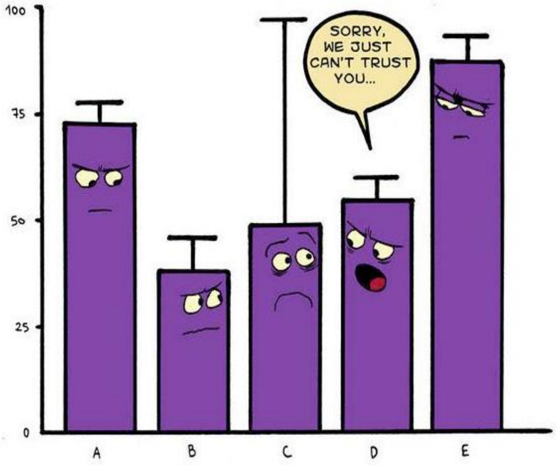
doi: 10.1016/j.envsoft.2016.02.008

francesca.pianosi@bristol.ac.uk

Sensitivity indices are calculated from a sample, so if the sample size is small, their values may be poorly approximated

In order to assess the robustness of our sensitivity estimates to the chosen sample, *without re-running the model*, we can use bootstrapping

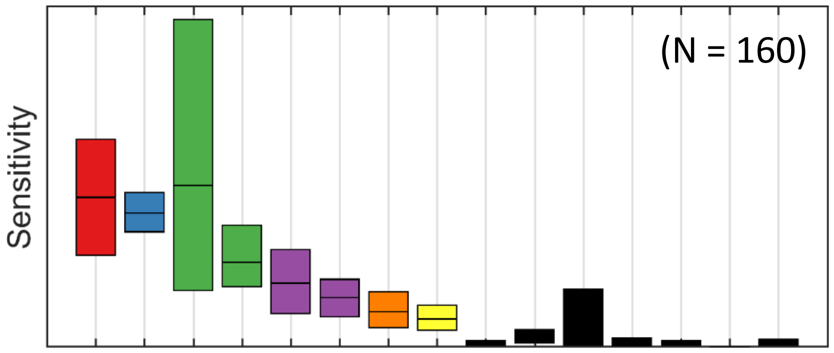




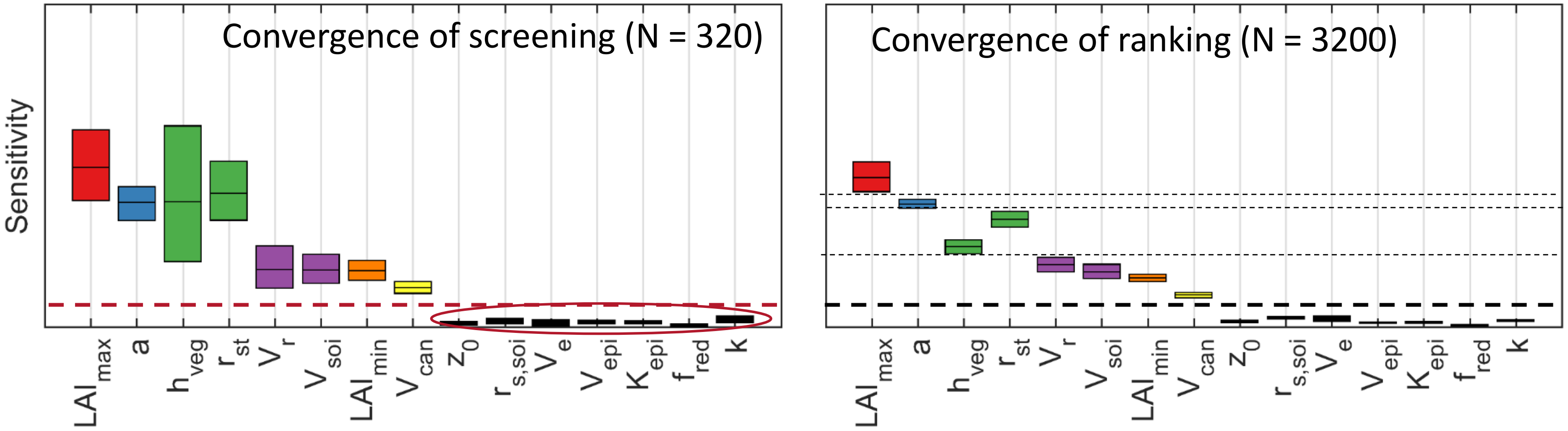
SOURCE: Reading Local Group of the Royal Statistical Society

francesca.pianosi@bristol.ac.uk

If the confidence intervals of our sensitivity indices are not “small enough” we must increase the sample size (what is “small enough” depends on the goal of our GSA)



black line: mean sensitivity index
bar: 90% confidence interval



francesca.pianosi@bristol.ac.uk

Defining scalar output metric(s)

Which output metrics
should we look at?

characterize uncertainty in the inputs

sample N combinations of inputs

execute the model against each inputs’ combination

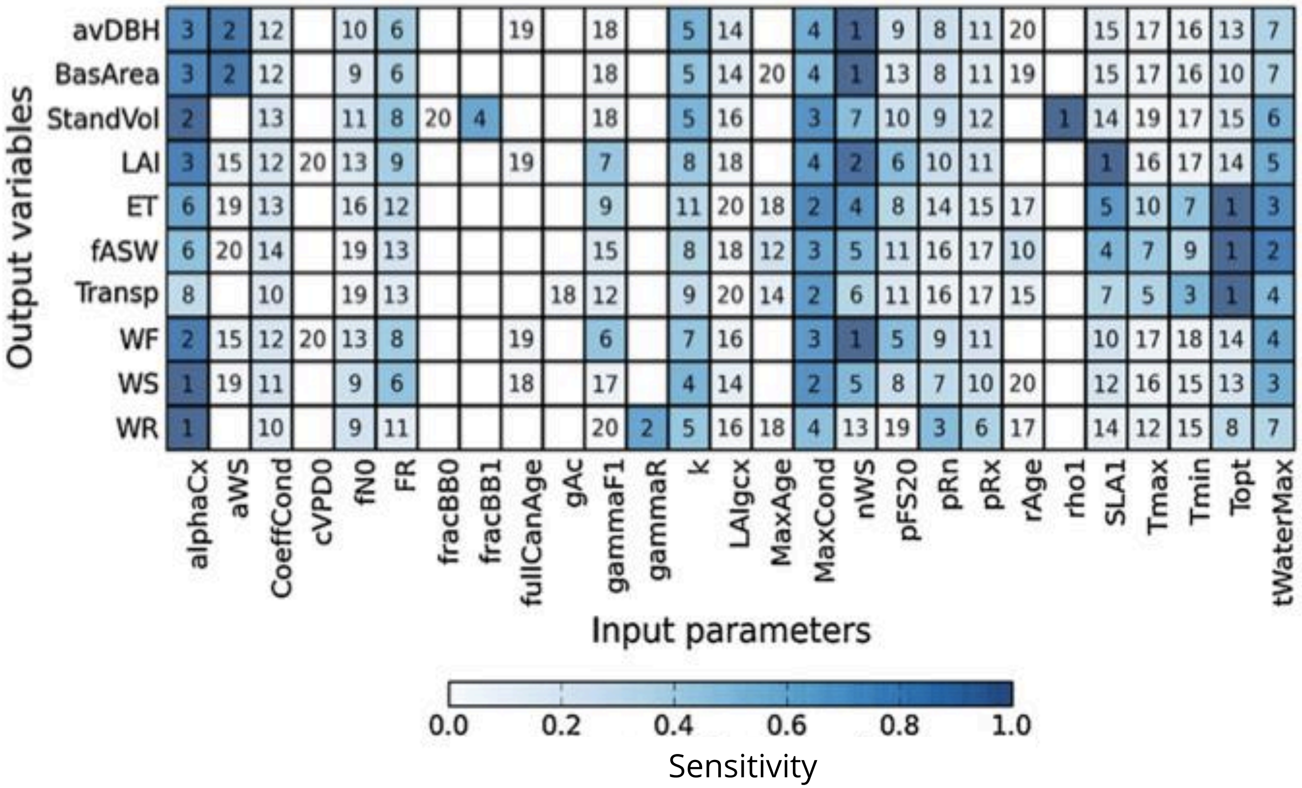
calculate output(s) ranges or distribution

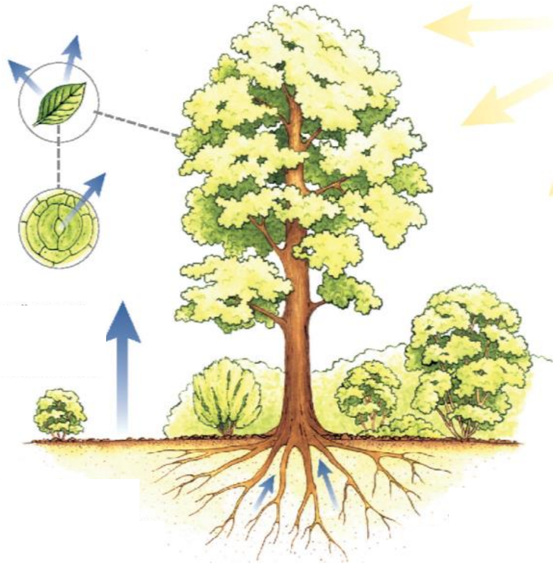
calculate sensitivity indices

francesca.pianosi@bristol.ac.uk

Each output metric is typically sensitive to a small subset of inputs, but which are those inputs will differ from one metric to another

Application to a forest growth model (Song et al EcM 2012)





francesca.pianosi@bristol.ac.uk

Choosing a method for calculating sensitivity indices

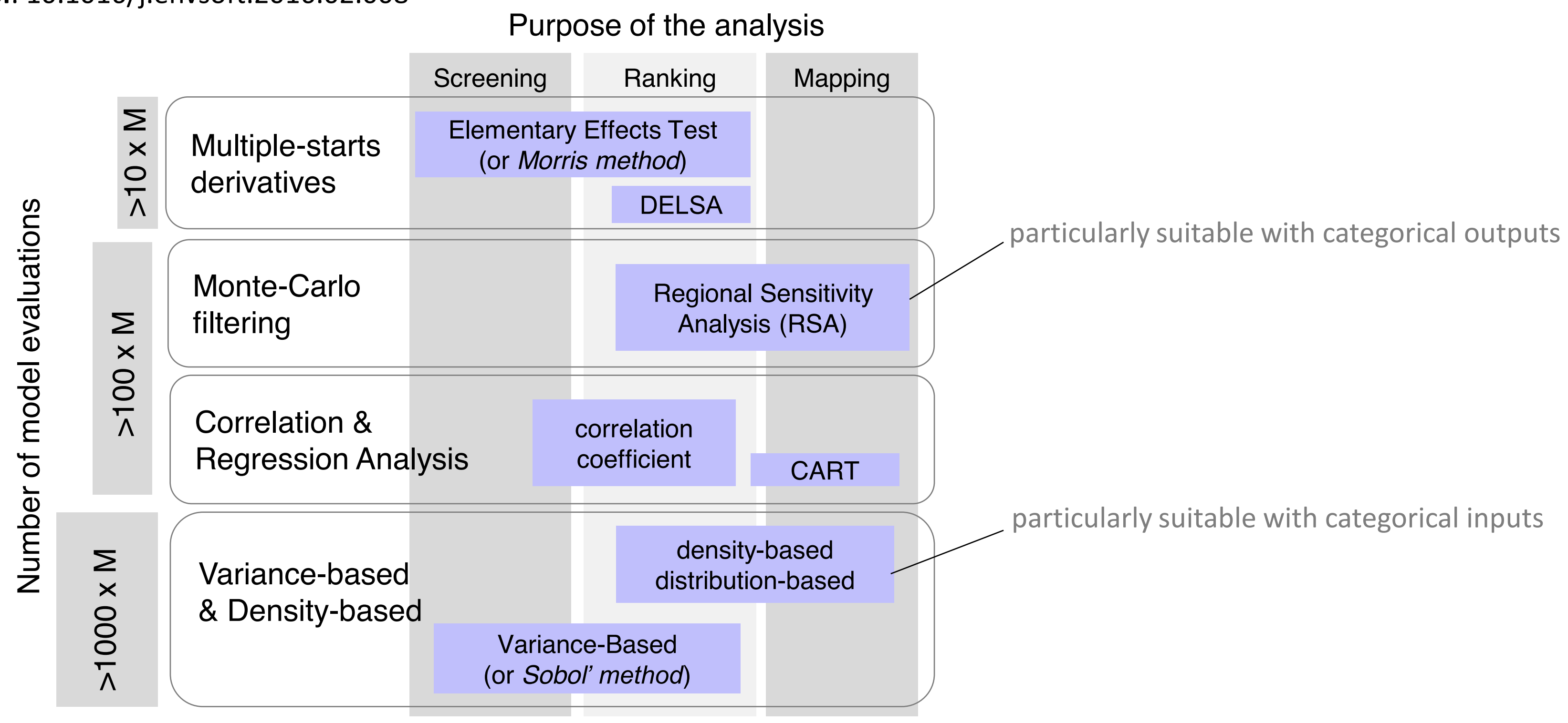
Which global sensitivity analysis method to use (e.g. variance-based, elementary effects test, regional sensitivity analysis, etc.)?

characterize uncertainty in the inputs	sample N combinations of inputs	execute the model against each inputs' combination	calculate output(s) ranges or distribution	calculate sensitivity indices

francesca.pianosi@bristol.ac.uk

Different methods defines “sensitivity” in different ways and are more or less suitable for specific purposes or problems

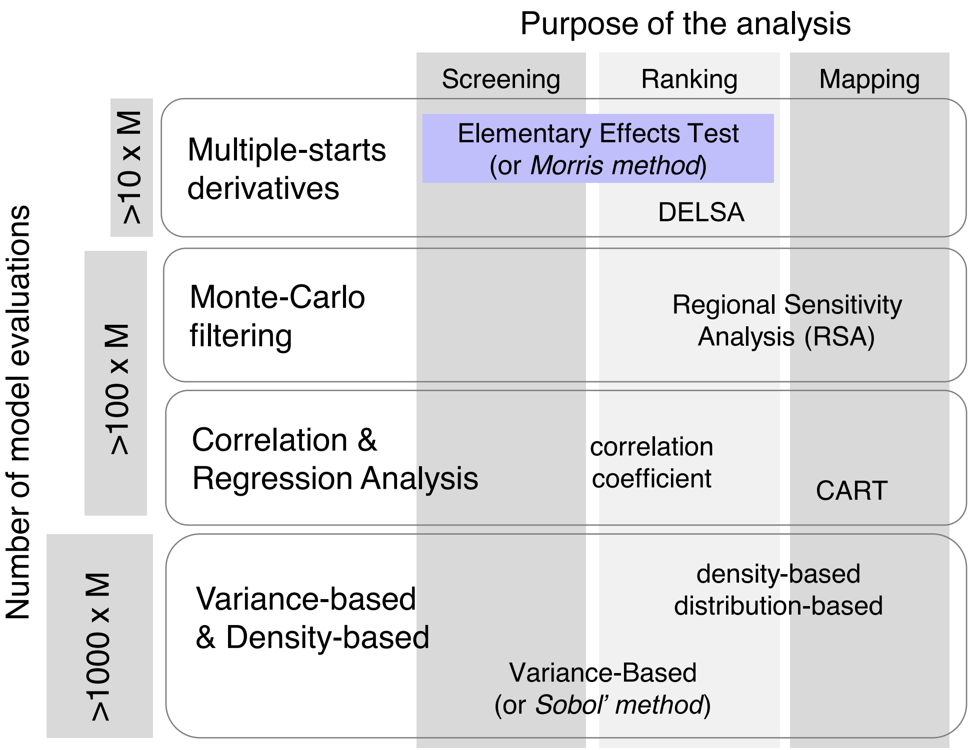
Pianosi et al 2016
doi: 10.1016/j.envsoft.2016.02.008



M = number of input factors

francesca.pianosi@bristol.ac.uk

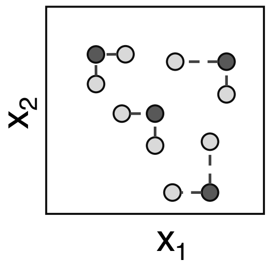
Elementary Effects Test
(Morris, 1991)

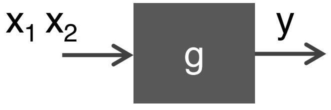


Sensitivity is proportional to...
the mean finite differences of the output across the input space

$$S_i = \frac{1}{r}\sum_{j=1}^{r} EE^j$$

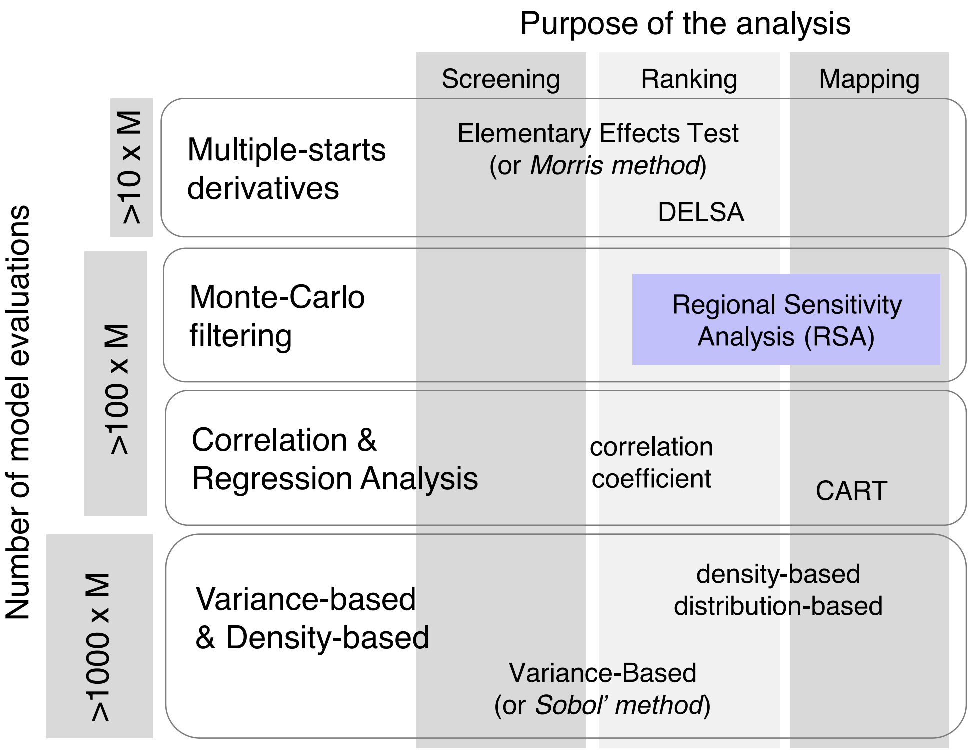
$$EE^j = \frac{g(\bar{x}_1^j, ..., \bar{x}_i^j + \Delta_i^j, ..., \bar{x}_M^j) - g(\bar{x}_1^j, ..., \bar{x}_i^j, ..., \bar{x}_M^j)}{\Delta_i^j} c_i$$





francesca.pianosi@bristol.ac.uk

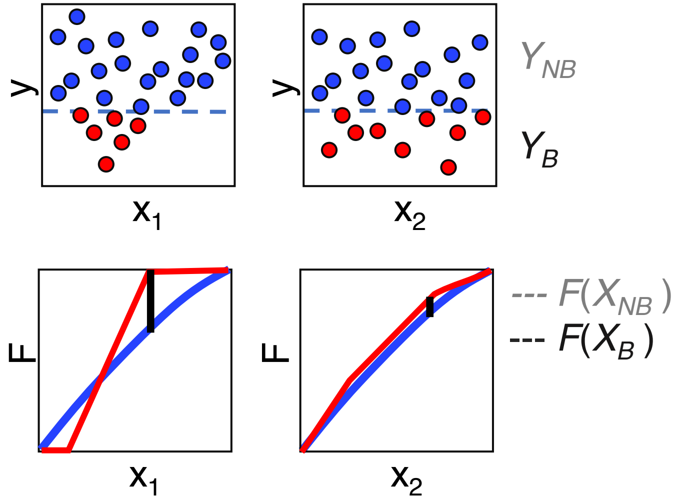
Regional Sensitivity Analysis
(Hornberger & Spear, 1980)



Sensitivity is proportional to...

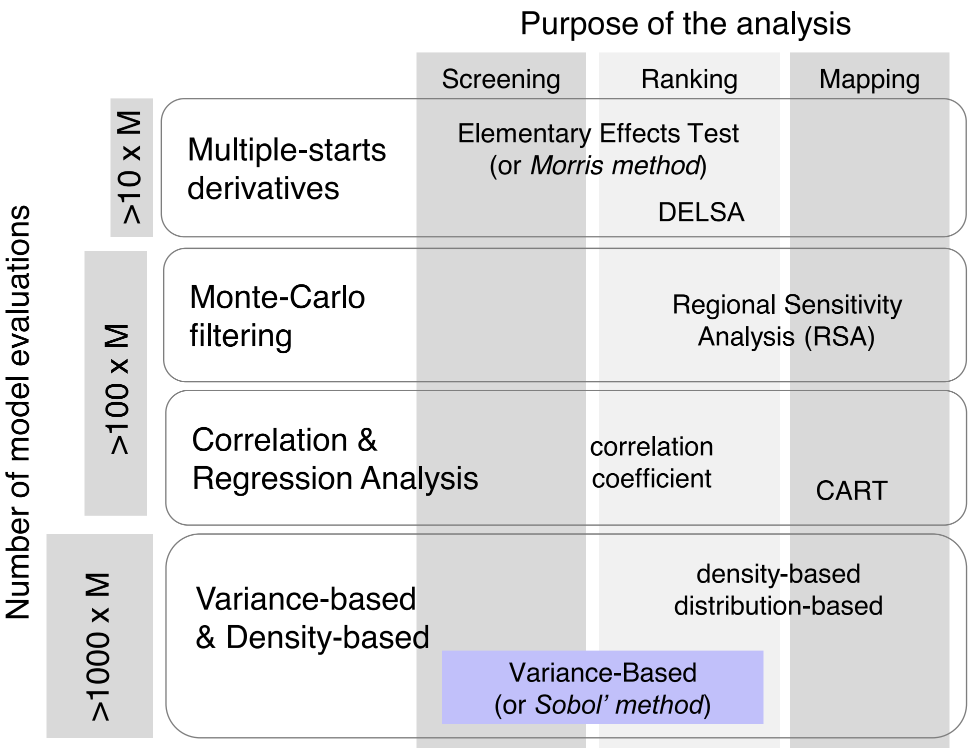
the variation induced in the distribution of an input by conditioning the output

$$S_i = \max_q |F_{x_i}(x|x \in X_B) - F_{x_i}(x|x \in X_{NB})|$$



francesca.pianosi@bristol.ac.uk

Variance-based Sensitivity Analysis (Homma & Saltelli, 1996)



Sensitivity is proportional to...
variation induced in the variance of the output by conditioning an input

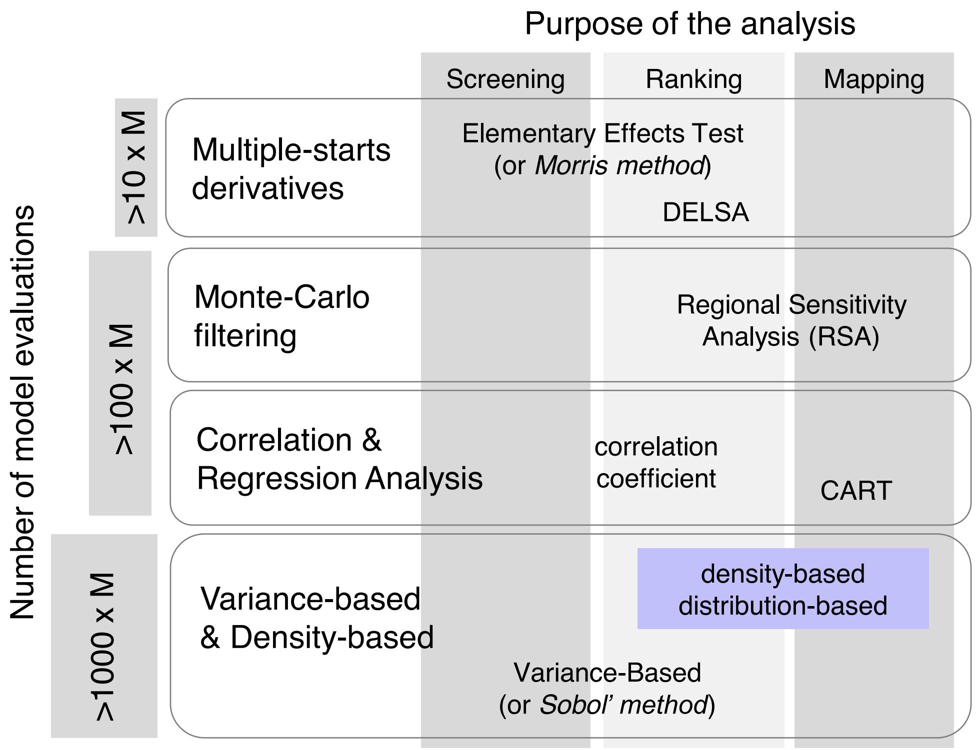
$$S_i = \frac{V_i}{V} = \frac{V_{x_i}[E(y|x_i)]}{V(y)} = \frac{V(y) - E_{x_i}[V(y|x_i)]}{V(y)}$$

$$S_i^T = 1 - \frac{V_{x_{\sim i}}}{V} = 1 - \frac{V_{x_{\sim i}}[E(y|x_{\sim i})]}{V(y)}$$

$$V = \sum_i V_i + \sum_{i<j} V_{ij} + \sum_{i<j<k} V_{ijk} + ... + V_{12...M}$$

francesca.pianosi@bristol.ac.uk

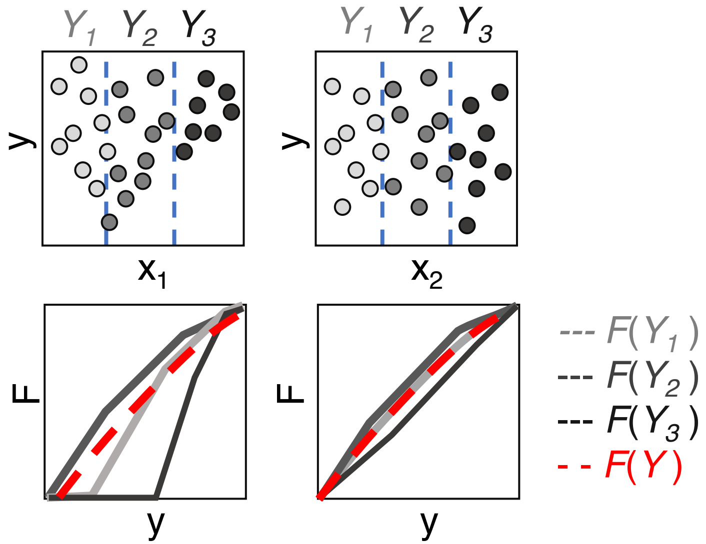
Distribution-based Sensitivity Analysis (PAWN) (Pianosi & Wagener 2015)



Sensitivity is proportional to...

variation induced in the distribution of the output by conditioning an input

$$S_i = \underset{k}{\mathrm{stat}} \ \max_q |F_y(q) - F_{y|x_i}(q|x_i \in I_k)|$$



francesca.pianosi@bristol.ac.uk