Transferring expertise on Global Sensitivity Analysis to insurance practitioners using workflows

Valentina Noacco[1], Francesca Pianosi[1,2], Thorsten Wagener[1,2], Tom Philp[3], John Wardman[3], Mike Maran[3]

[1]Department of Civil Engineering, University of Bristol, UK; [2]Cabot Institute, University of Bristol, UK; [3]AXA XL, London, UK

EGU2019 - 10450

Background on losses from natural disasters and insurance

- Insurance companies provide insurance against a wide range of threats, such as natural hazards.
- Losses from natural disasters are increasing globally (in 2017 they were estimated to be USD 340 billion) [1].
- To monitor risk and support investment decisions, insurance companies often use mathematical models to help calculate fair and robust risk premiums, so as to ensure appropriate capital allocation, so that there is little risk of the company finding itself in financial trouble should deleterious events occur.

Challenges for (re)insurers

- Cost of model development;
- Increased regulation (European Solvency II) [2];
- Lack of standardised regulatory guidelines for model validation.

What is Global Sensitivity Analysis?

GSA is a set of statistical analysis techniques to investigate the complex behaviour of mathematical models in a structured, transparent and comprehensive way [3].

Sensitivity analysis answers the question: how much does varying each input factor contribute to the variability of the model output?

Project's goal

To improve decision-making under uncertainty by transferring state-of-the-art methods for GSA, tools (SAFE software toolbox) [4,5] and workflows to the (re)insurance sector and catastrophe community.

By better capturing uncertainty, we aim to make decision-making more robust, as the chance of unexpected surprises is reduced, and increase the transparency of an insurer's risk profile

[1] Munich Re (2018) https://www.munichre.com/topics-online/en/2018/01/2017-year-in-figures

[2] EC (2009) Directive 2009/138/EC *The taking-up and pursuit of the business of Insurance and Reinsurance* (Solvency II).

[3] Saltelli *et al.* (2008) *Global Sensitivity Analysis. The Primer.* Wiley

[4] Pianosi *et al.* (2015) *An open-source Matlab Toolbox for GSA.* Env. Mod. & Soft.

[5] www.safetoolbox.info

NERC SCIENCE OF THE ENVIRONMENT

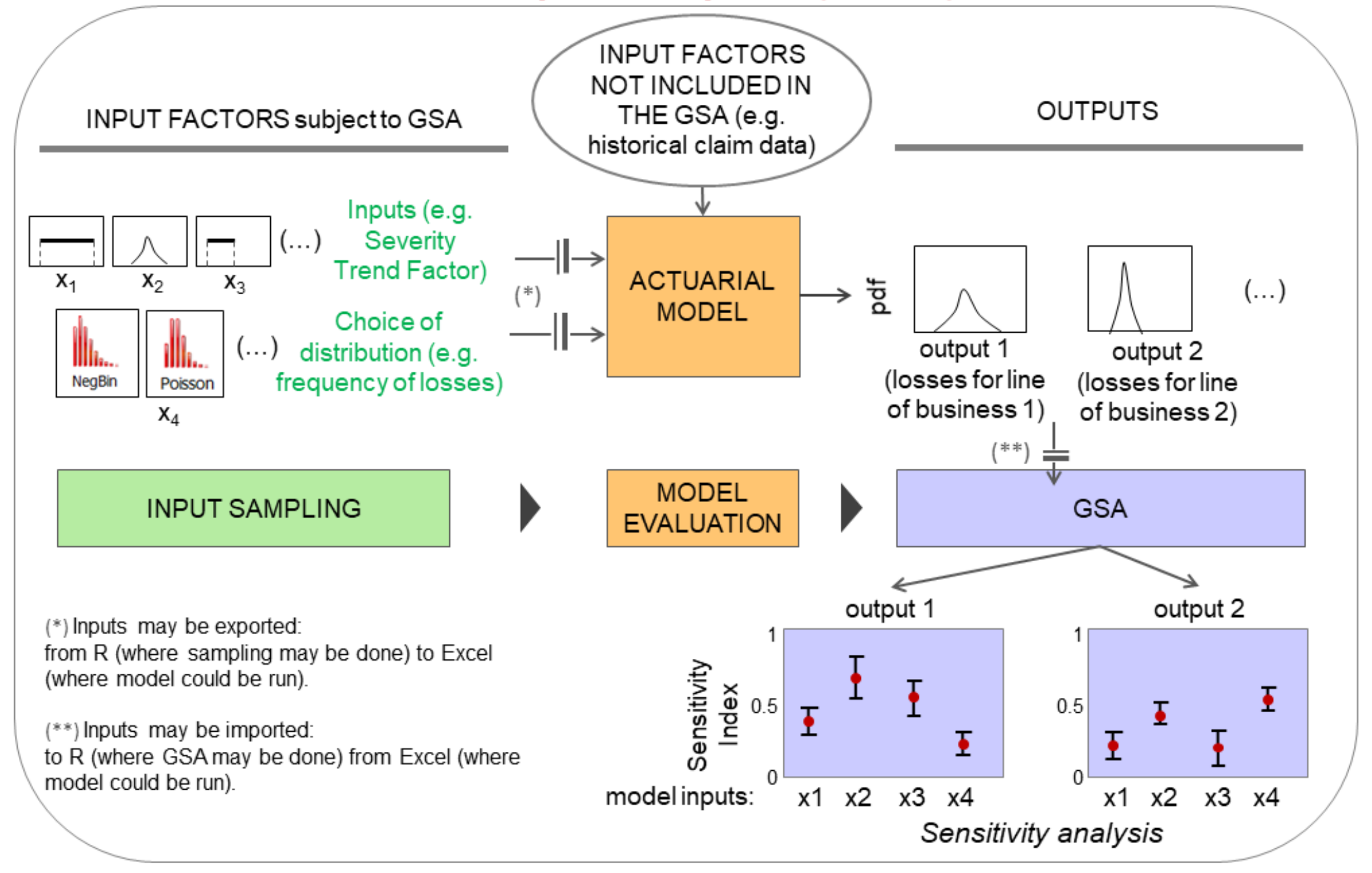
AXA

University of BRISTOL

KE Fellowship: NE/R003734/1

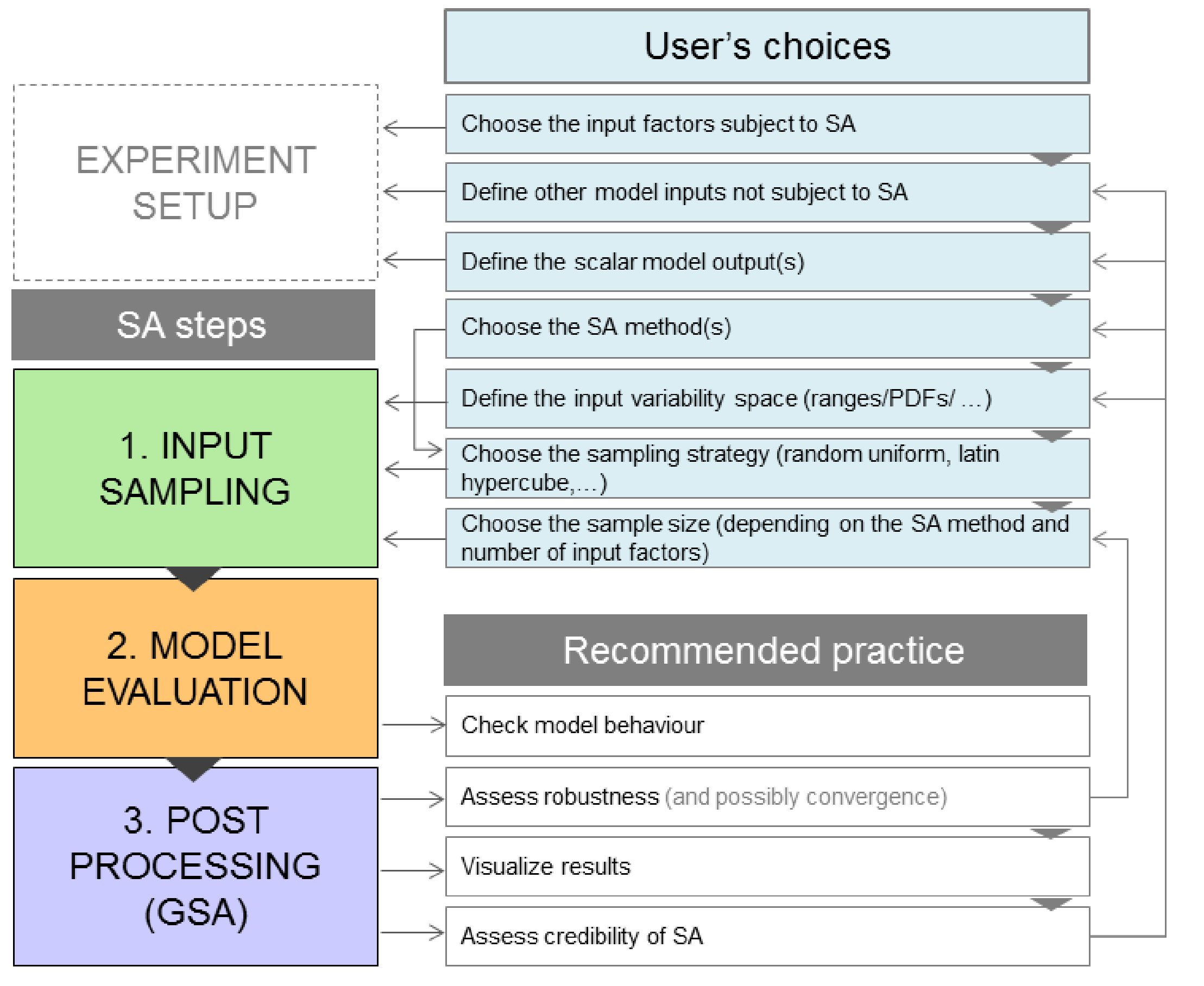
Workflows can accelerate practitioners' **uptake** of methods and software to better attribute uncertainty

How Global Sensitivity Analysis (GSA) works


INPUT FACTORS subject to GSA
INPUT FACTORS NOT INCLUDED IN THE GSA (e.g. historical claim data)
OUTPUTS
x1
x2
x3
(…)
Inputs (e.g. Severity Trend Factor)
NegBin
Poisson
x4
(…)
Choice of distribution (e.g. frequency of losses)
(*)
ACTUARIAL MODEL
pdf
output 1 (losses for line of business 1)
output 2 (losses for line of business 2)
(…)
(**)
INPUT SAMPLING
MODEL EVALUATION
GSA
output 1
output 2
Sensitivity Index
1
0.5
0
model inputs:
x1 x2 x3 x4
x1 x2 x3 x4
Sensitivity analysis

(*) Inputs may be exported: from R (where sampling may be done) to Excel (where model could be run).

(**) Inputs may be imported: to R (where GSA may be done) from Excel (where model could be run).


User's choices
EXPERIMENT SETUP
SA steps
1. INPUT SAMPLING
2. MODEL EVALUATION
3. POST PROCESSING (GSA)
Choose the input factors subject to SA
Define other model inputs not subject to SA
Define the scalar model output(s)
Choose the SA method(s)
Define the input variability space (ranges/PDFs/ …)
Choose the sampling strategy (random uniform, latin hypercube,…)
Choose the sample size (depending on the SA method and number of input factors)
Recommended practice
Check model behaviour
Assess robustness (and possibly convergence)
Visualize results
Assess credibility of SA

safe-insurance.uk

@ValentinaNoacco

valentina.noacco@bristol.ac.uk

Photography encouraged

Load the packages

Install and load the packages below.

```
library(caTools)
library(calibrater) # Install from tar file, also available at: https://people.maths.bris.ac.uk/~mazjcr/calibrater_0.51.tar.
gz
library(SAFER) # Install from zip
library(ggplot2)
```

Define input factors subject to SA

The distribution of the inputs `DistrFun` and their range `DistrIn` can be chosen by expert judgement, available data (e.g. portfolio or market data) or literature.

```
DistrFun  <- "unif" # Inputs distribution
DistrIn  <- list( c(0, 1), c(0, 1), c(0, 1), c(0, 1)) # Range of each input
x_labels <- c("Frequency trend","Severity trend","Exposure trend", "Development pattern") # Name of inputs
```

Sample inputs space

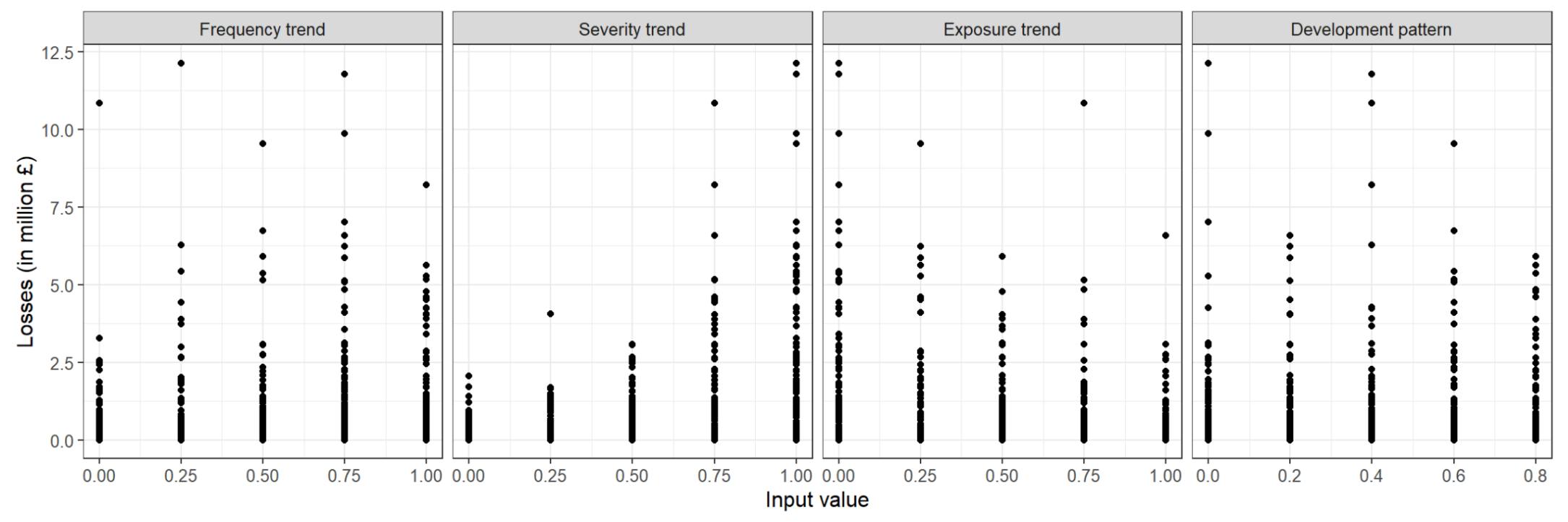
```
SampStrategy <- "lhs" # Here the sampling strategy for All At the Time (AAT) sampling is
# Latin hypercube (another option is random uniform)
N <- 500 # Sample size
M <- length(DistrIn) # Number of inputs
X <- AAT_sampling(SampStrategy, M, DistrFun, DistrIn, N) # Sample inputs space
colnames(X) <- x_labels # Set columns names
```

Run the model

```
Y <- actuarial_model(X) # Where 'actuarial_model' is your chosen model
```

Check model behaviour by visualising input/output samples

```
scatter_plots(X, Y, prnam = x_labels) + ylab("Losses (in million £)")  +
  xlab("Input value") + theme(text = element_text(size=sz_tx))
```



Question: *From these scatter plots, which input factor would you say is most influential? Why? Are there input factors that are not influential at all?*

Compute sensitivity indices with RSA

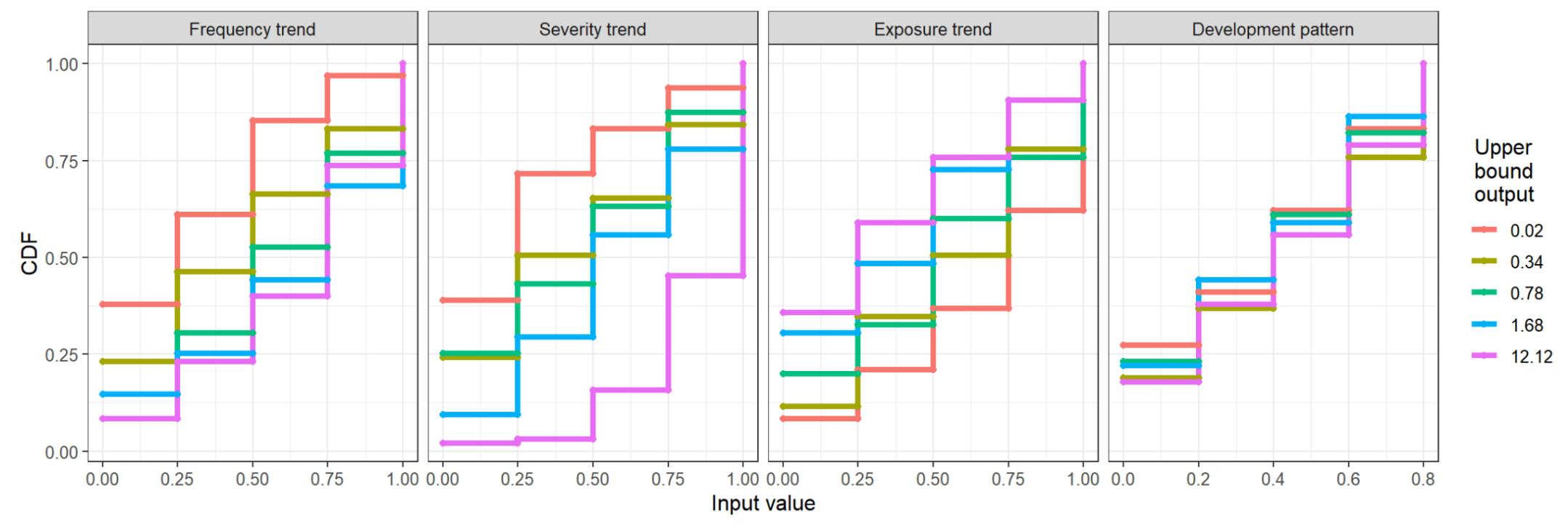
```
n_groups <- 5; # Number of groups into which the output is splitted, default = 10
flag <- 2; # where flag: statistic for the definition of the RSA index (maximum)

rsa_gr <- RSA_indices_groups(X, Y, n_groups,flag)

# Outputs
mvd <- rsa_gr$stat # mvd: maximum vertical distance between CDFs (sensitivity index)
idx <- rsa_gr$idx # idx: index which divides the output into different groups
Yk <- rsa_gr$Yk # Yk: output limit of each group
```

Plot inputs CDFs

```
RSA_plot_groups(X, idx, Yk, prnam = x_labels)  + xlab("Input value") +
  theme(text = element_text(size=sz_tx))
```

Plot inputs CDFs

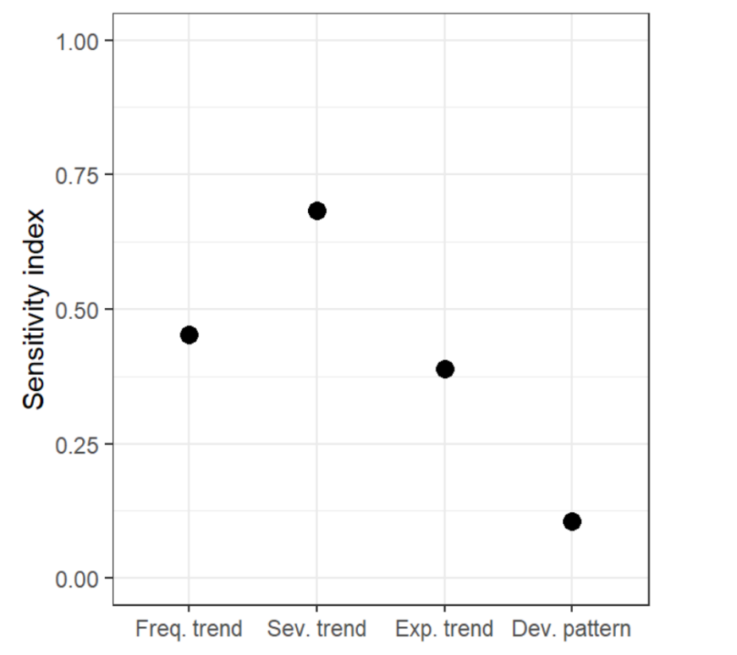
```
RSA_plot_groups(X, idx, Yk, prnam = x_labels)  + xlab("Input value") +
  theme(text = element_text(size=sz_tx))
```



Question: *From the CDF plots, which input factor would you say is most influential? Why? Are these results consistent with the visual analysis of the scatter plots?*

Plot the sensitivity indices

```
boxplot1(mvd, prnam = x_labels) + ylab("Sensitivity index") + ylim(0, 1)
```



Question: *Are the sensitivity indices consistent with the visual inspection of the CDF plots? Are these results credible?*

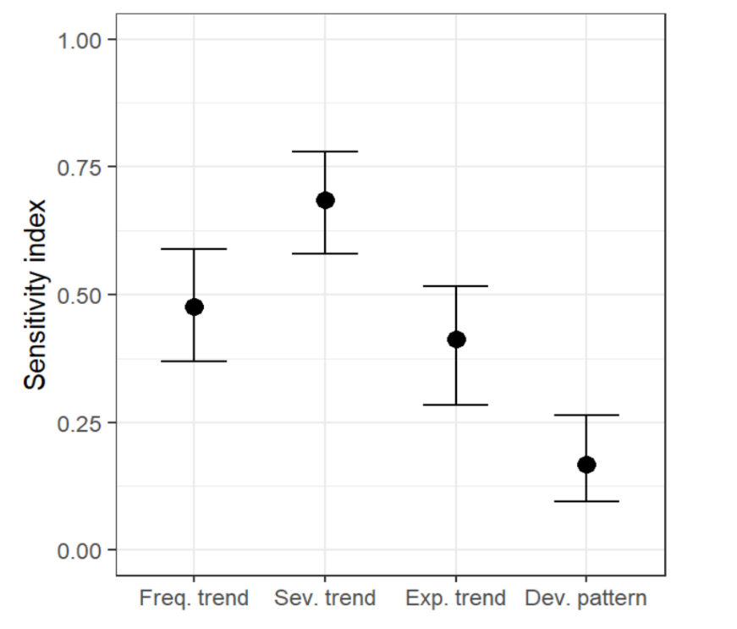
Assess robustness by bootstrapping

In order to assess the robustness of the sensitivity indices bootstrapping in performed (here `Nboot = 1000`).

```
Nboot <- 100 # Number of resamples used for bootstrapping
rsatgr100 <- RSA_indices_groups(X, Y, n_groups, flag, Nboot = Nboot, alfa = 0.05) # By adding the extra argument `Nboot` to
 the function `RSA_indices_groups` bootstrapping is performed,
# 'alfa' is the scalar significance level for the confidence intervals estimated by bootstrapping

mvd_Nb <- rsatgr100$stat
idxb_Nb <- rsatgr100$idxb
mvd_lb <- rsatgr100$stat_lb
mvd_ub <- rsatgr100$stat_ub
```

Here the sensitivity indices with their 95% confidence intervals are plotted.

```
boxplot1(mu = mvd_Nb, lb = mvd_lb, ub = mvd_ub, prnam = x_labels) + ylim(0, 1) + ylab("Sensitivity index")
```



The results show that the Severity Trend factor is the most influential input, followed by the Frequency and Exposure Trend factors.

Question: *Are the sensitivity indices adequately estimated? Is the sample size large enough?*