Why pay attention to uncertainty?

- Any self-respecting researcher should want to check the quality of his/her results before these are made public.

- Quantified uncertainty allows to compare the performance of models – scientists, clients and end users must know the quality of model outputs to judge their usability for specific purposes.

- Uncertainty source contributions helps to decide rationally and economically how best to improve models.

- Quantified uncertainty can be included in decision making, e.g. risk analysis: Expected costs = \( \sum P(\text{outcome}) \cdot \text{Cost}(\text{outcome}) \)
Uncertainty propagation overview

Model parameters
Model structure
Solution method
Observ. data

Input ± U

Data ± U

Output ± U
Environmental data vary in space

E.g. Rainfall

Categories like soil type

Environmental quality indicators like soil pH
Deriving maps is encumbered with error

E.g. interpolation error or error associated with maps derived from other maps
Fact of life: maps stored in the GIS database are rarely if ever error-free

Causes: generalisation, digitisation, measurement, classification and interpolation errors

Consequence: errors will propagate through GIS operations and spatial models

Key research question: given the errors in the inputs to the GIS operation, how large are the errors in the output?
Uncertainty propagation analysis involves three steps:

1. **DEFINITION** of a (statistical) uncertainty model for variables of interest

2. **IDENTIFICATION** of the uncertainty model (estimate its parameters)

3. Perform the actual **UNCERTAINTY PROPAGATION ANALYSIS**
DEFINE and IDENTIFY a (statistical) uncertainty model

**DEFINITION**

\[ Z(x) = \mu(x) + \sigma(x) \cdot \varepsilon(x) \]

- \( \mu \) is the (deterministic) mean of the variable of interest \( Z \)
- \( \sigma \) is a spatially variable standard deviation associated with the prediction \( \mu \)
- \( \varepsilon(x) \) is the (stochastic) error about it (typically zero mean, but non-zero variance and spatially correlated)

**IDENTIFICATION**

- Marginal pdf at each location (both its shape and parameters)
- Spatial correlation (correlogram or semivariogram)
- Cross-correlation with other uncertain inputs
Performing the actual UNCERTAINTY PROPAGATION ANALYSIS

1. Taylor series approximation method
2. Monte Carlo method
Monte Carlo method

Introduce by means of an example
Example: computing slope from DEM for a 2 by 2.5 area in the Austrian Alps
Slope map computed from the DEM (percent):
Now let the uncertainty about the elevation be ± 10 meter
Realisations of uncertain DEM:
Corresponding slope maps:
Histograms capture uncertainty in slope:
Monte Carlo algorithm:

1. Repeat N times (N ≥ 100):
   1. Simulate a realisation from the probability distribution of the uncertain inputs using a pseudo-random number generator
   2. Run the model with these inputs and store the result

2. Analyse the N model outputs by computing summary statistics such as the mean and standard deviation (the latter is a measure of the output uncertainty)
Performing the actual UNCERTAINTY PROPAGATION ANALYSIS

1. Taylor series approximation method
2. Monte Carlo method

Advantages Monte Carlo method:

- Yields the full output pdf, not only the mean and variance
- Works with any model
- Easy to implement
Monte Carlo method with spatial inputs

- Requires that we simulate from the probability distribution of a spatially distributed variable

- Can be done with spatial stochastic simulation: instead of kriging, which produces the most likely value (conditional expectation), we generate a possible reality, by simulating from the probability distribution of the spatial variable, using a pseudo-random number generator

- Spatial correlations are taken into account
Why is it important to include spatial correlation?

- The spatial pattern of the generated realizations depends on the degree of spatial autocorrelation.

- Important for neighbourhood or global operations (opposite to point operations).
R package **spup** - motivation

R packages:
- propagate
- FME
- mcmcse
- ArArRedux
- betaper
- UncerIn2
- usdm
- sensitivity
- and others...

@RISK
COSSAN
Crystall Ball
SAFE

UQLab
UNCSIM
UNCSAM
UCODE
TIME

OpenTURNS
OSTRICH
PEST
PSUADE

SimLAB
DAKOTA
DUE
FRAMES

WAGENINGEN UNIVERSITY
WAGENINGEN UR

QUICS
R package **spup** – underlying methodology

Monte Carlo approach principle

Variable 1

Sampling

Model run (e.g. Output = Var1 x Var2)

Deterministic (D) → Random (R) → Parameter (P) → Result (RV)

Variable 2

Sampling

c(10, 14, 3, 16, ....)
R package **spup** - functionality

- **D**: Import data and objects and define UM
- **R**: Monte Carlo sampling techniques
- **P**: Model call and simulations
- **RV**: Visualization of results
Computer exercise – calculating slope from DEM

- Analyse how uncertainty in DEM map of Zlatibor region in Serbia propagates to slope calculations
- Check how uncertain the DEM and slope map is, and where in Zlatibor it is most uncertain
- Identify areas in Zlatibor for which you are 90% certain that the soil is in risk of serious erosion
Computer exercise – calculating slope from DEM

- Analyse how uncertainty in DEM map of Zlatibor region in Serbia propagates to slope calculations
- Check how uncertain the DEM and slope map is, and where in Zlatibor it is most uncertain
- Identify areas in Zlatibor for which you are 90% certain that the soil is in risk of serious erosion

**Folder ‘install’:**

**Required:**
- R >= v.3.3.3

**Packages:**
- spup
- sp
- gstat
- GGally
- gridExtra
- purrr
- magrittr

**Folder ‘practical’:**

**Data:**
- mean DEM: dem30m.RData
- sd DEM: dem30m.RData

**R script:**
- DEM_exercise.R

**PDF tutorial:**
- DEM_excercise.pdf