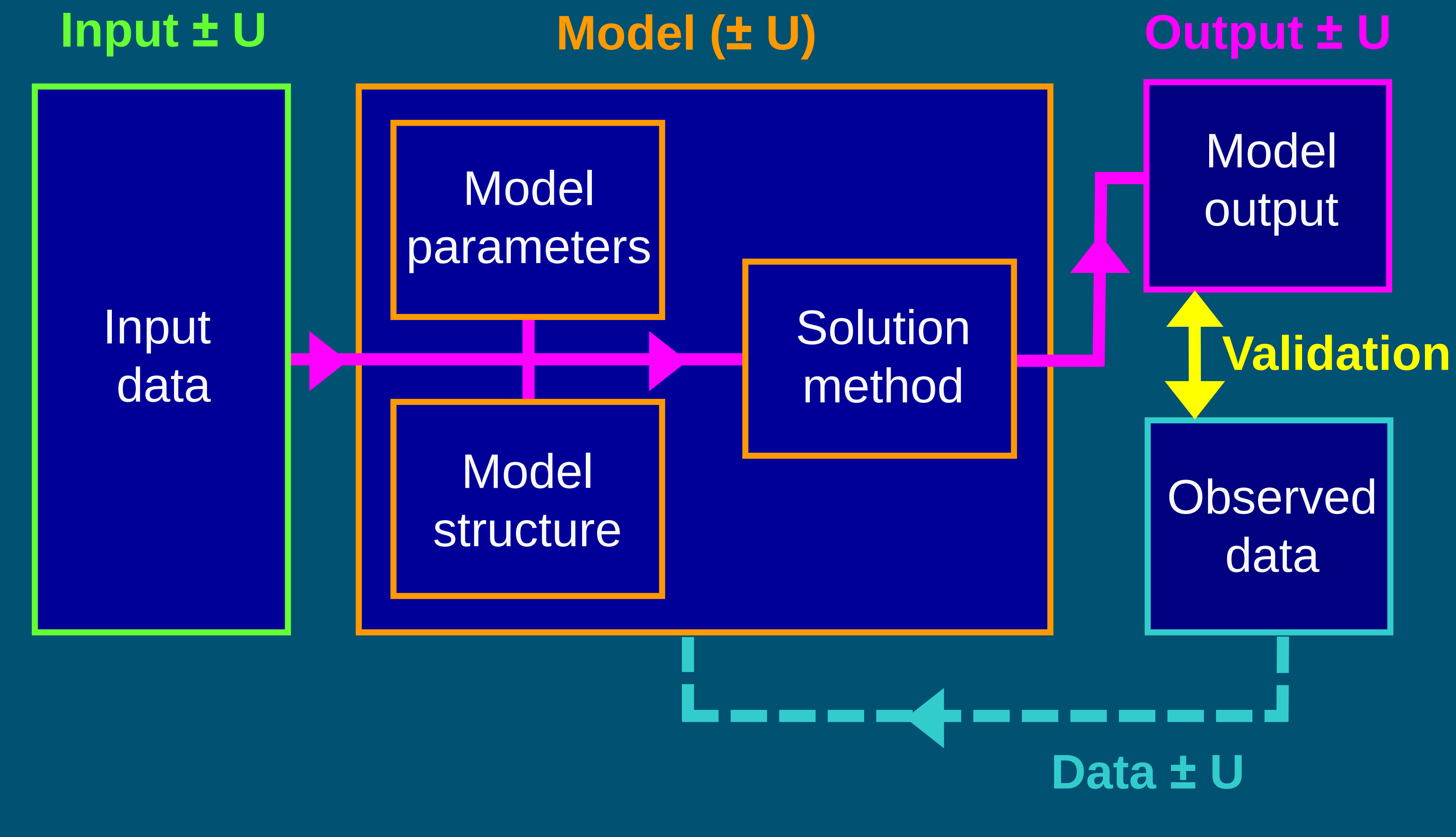
Uncertainty Analysis and Statistical Validation of Spatial Environmental Models

PE&RC Course 9-13 December 2024

Gerard Heuvelink and Sytze de Bruin

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Uncertainty propagation and model validation overview



3. Actual UNCERTAINTY PROPAGATION ANALYSIS

We discuss two methods:

1. Taylor series approximation (yesterday)
2. Monte Carlo simulation (today)

But today we will also address some other topics:

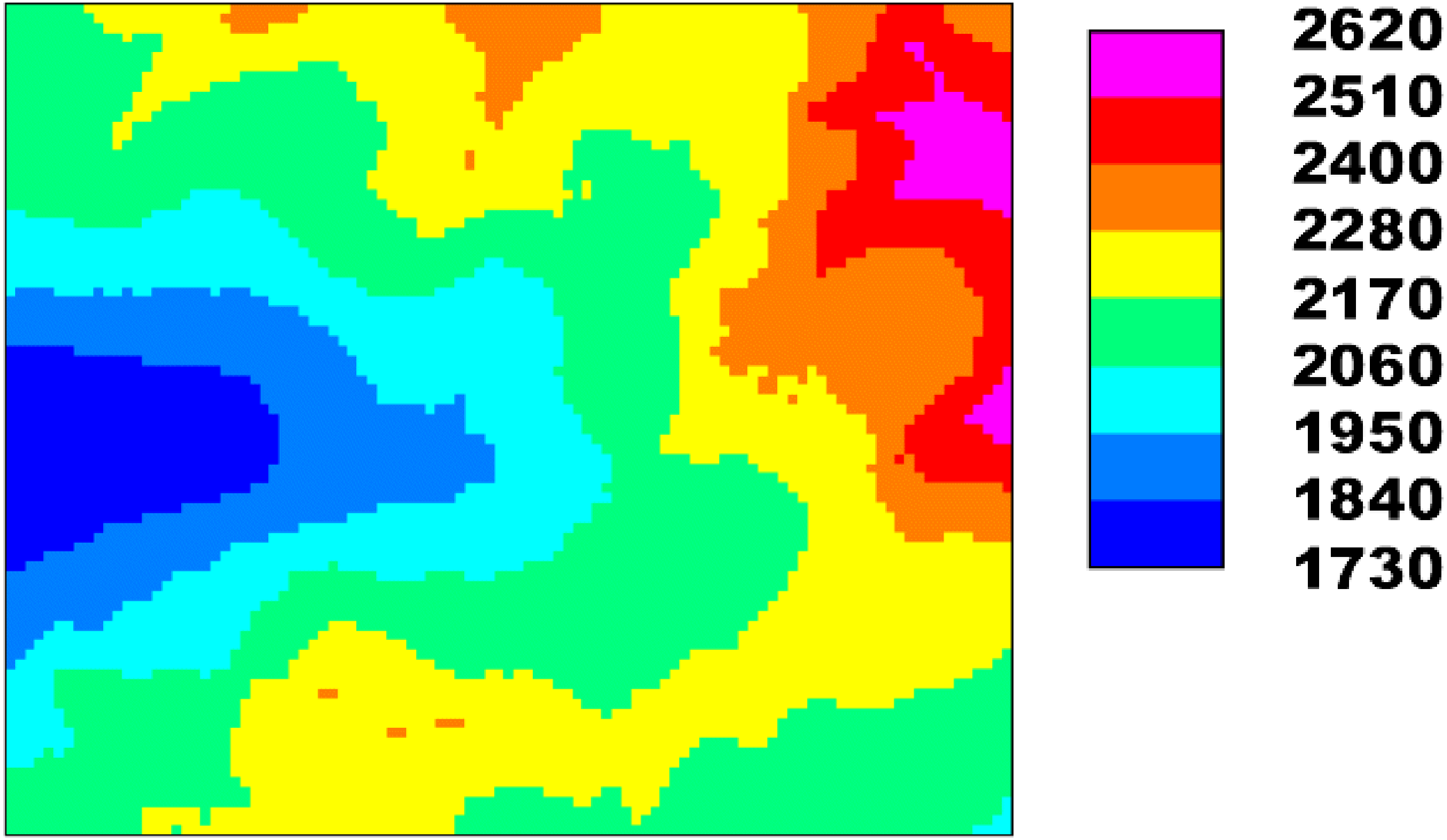
1. Uncertainty source contributions
2. Added value of uncertainty quantification
3. Uncertainty about the uncertainty
4. Real-world application of spatial environmental uncertainty propagation

Monte Carlo method

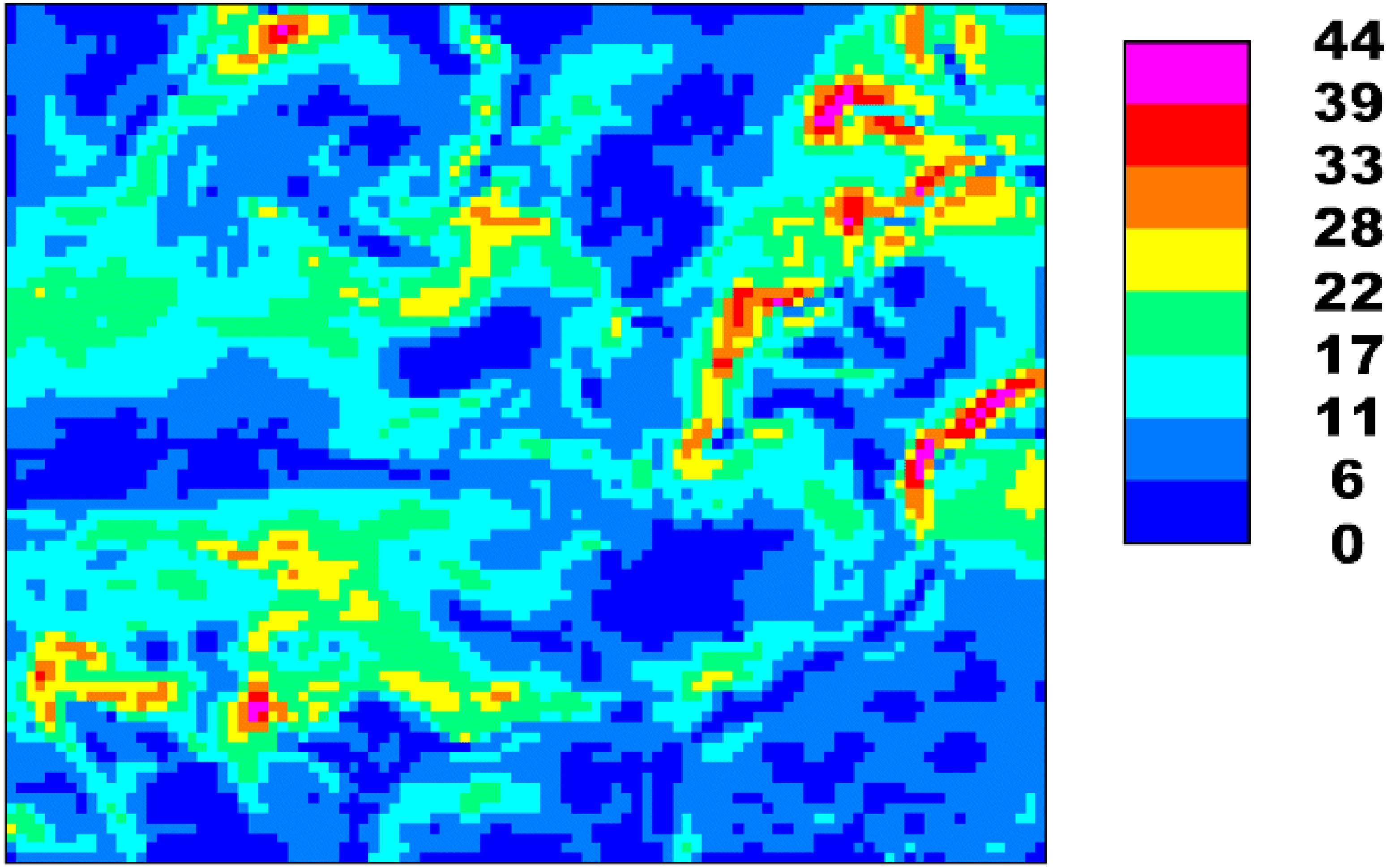
Introduce by means of an example

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Example: computing slope from DEM for a 2 km by 2.5 km area in the Austrian Alps

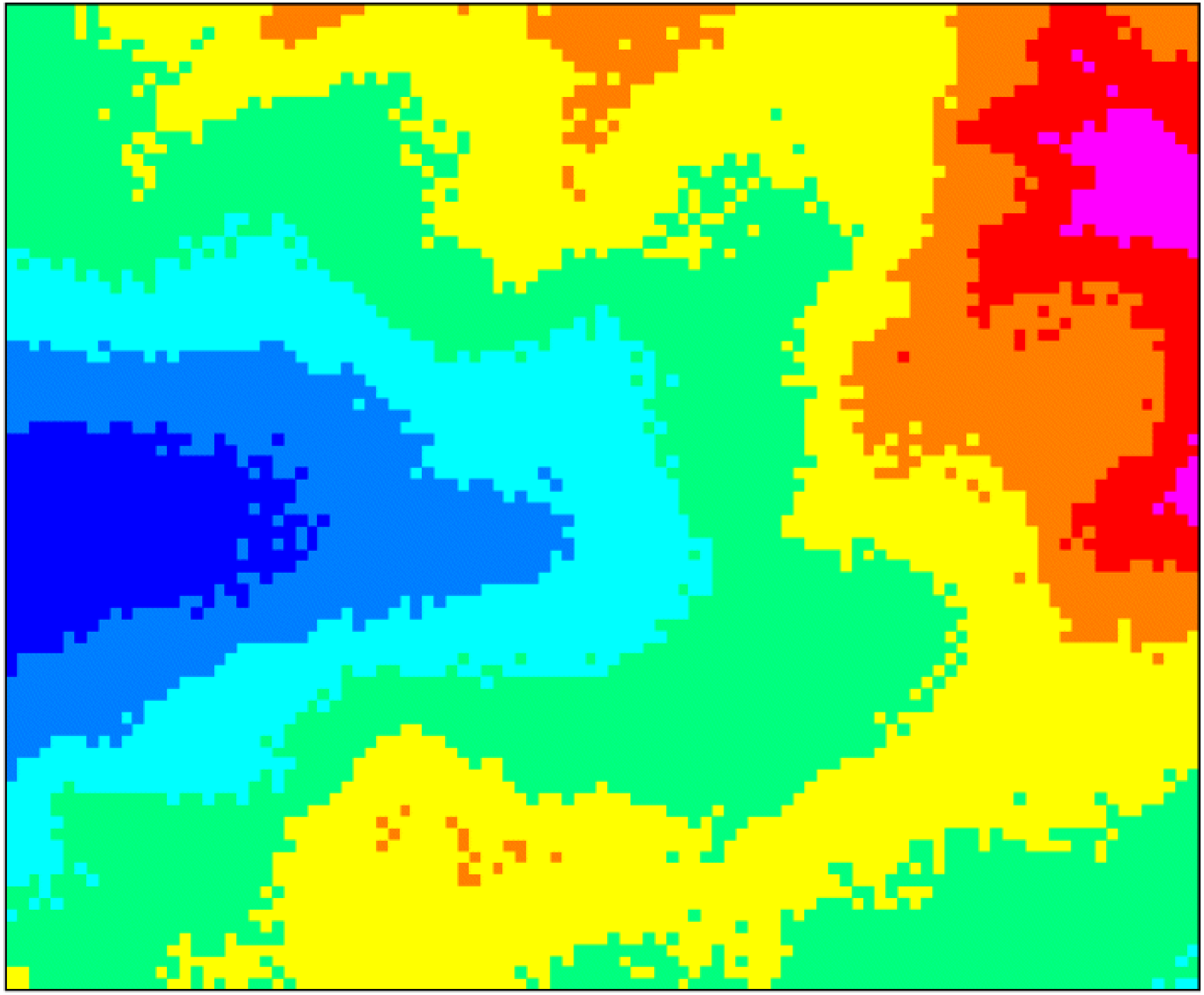


Slope map computed from the DEM (percent):

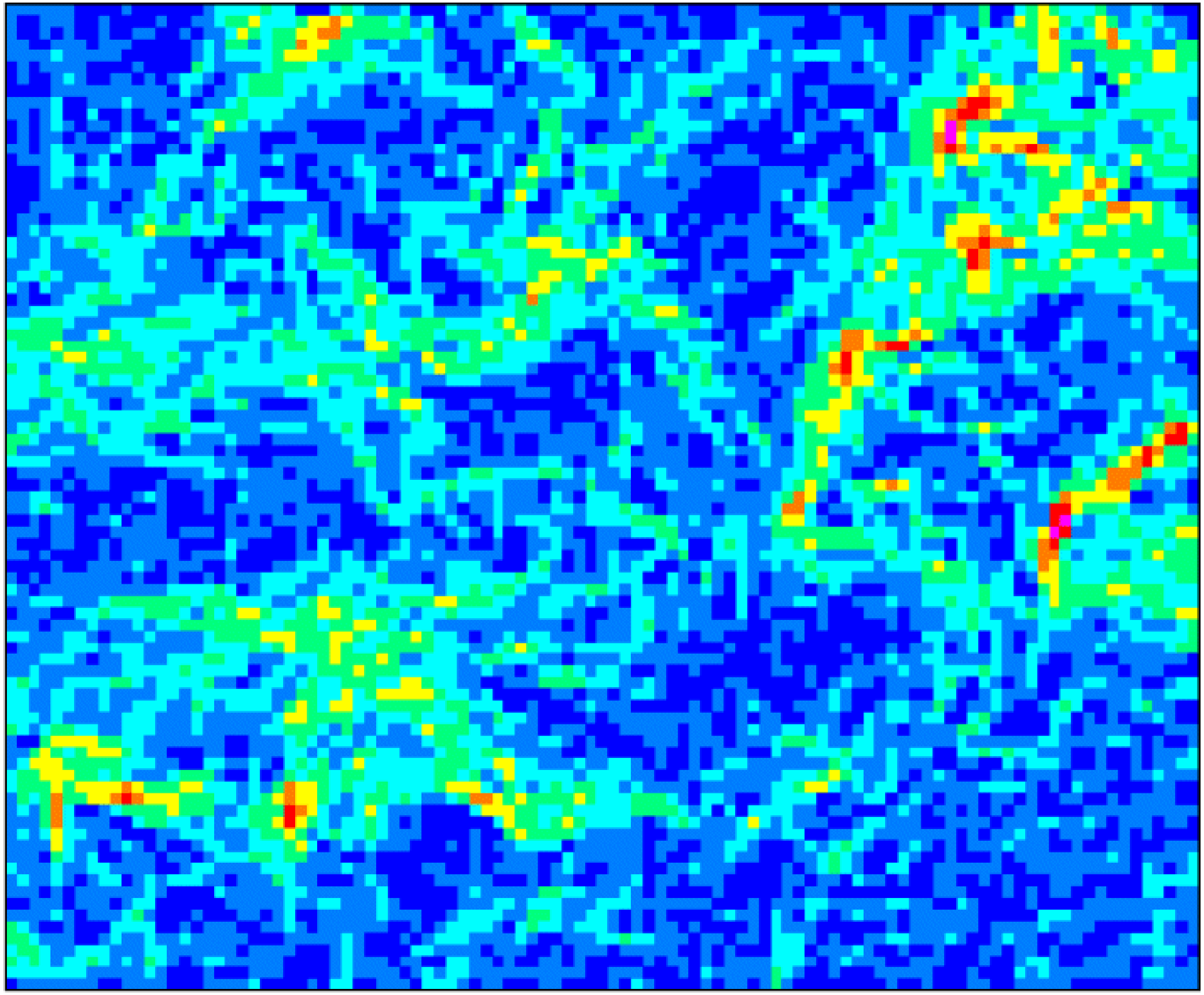


Now let the uncertainty about the elevation be $\pm$ 10 meter

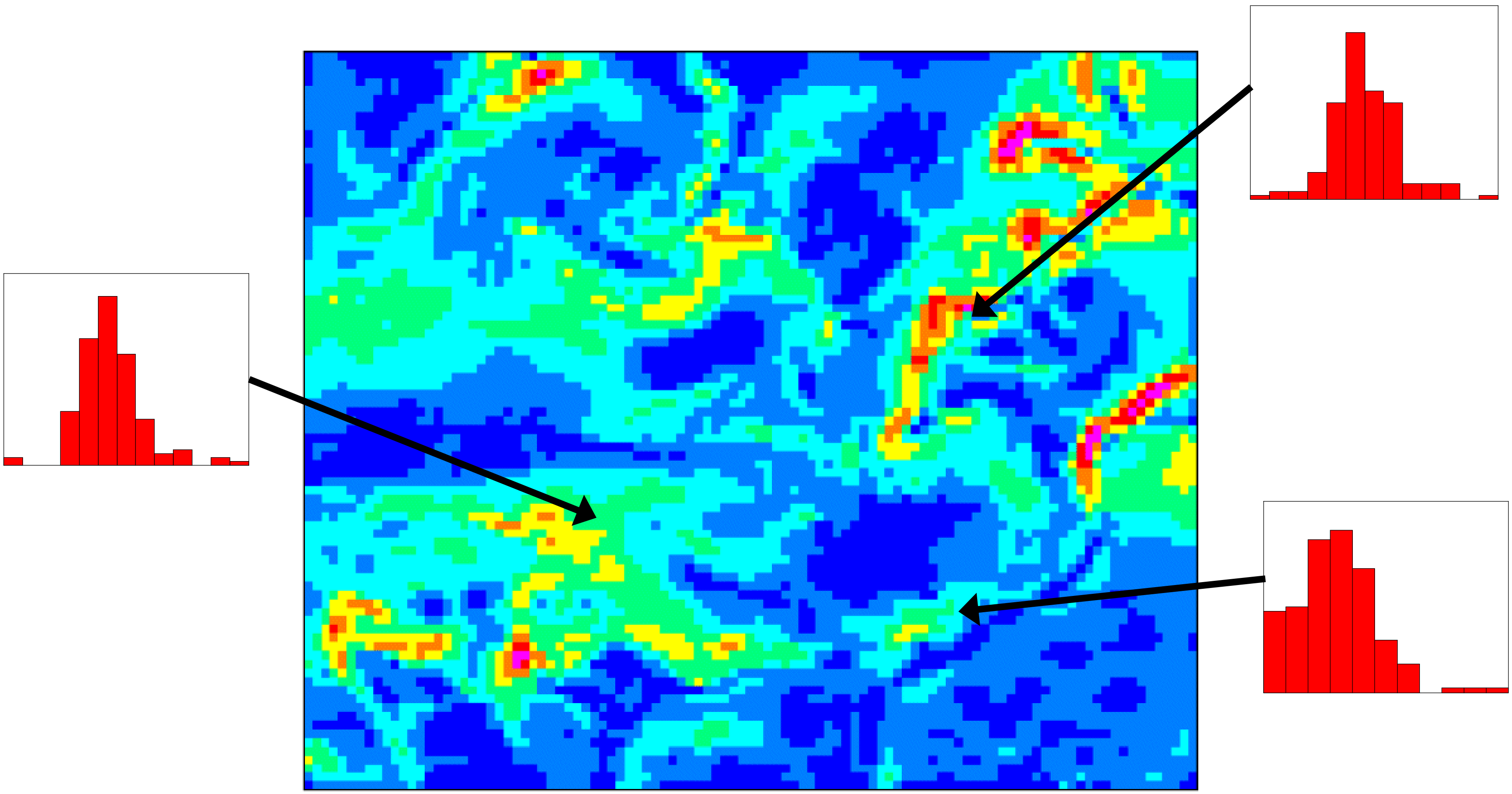
Realisations of uncertain DEM:



Corresponding slope maps:



Histograms capture uncertainty in slope:



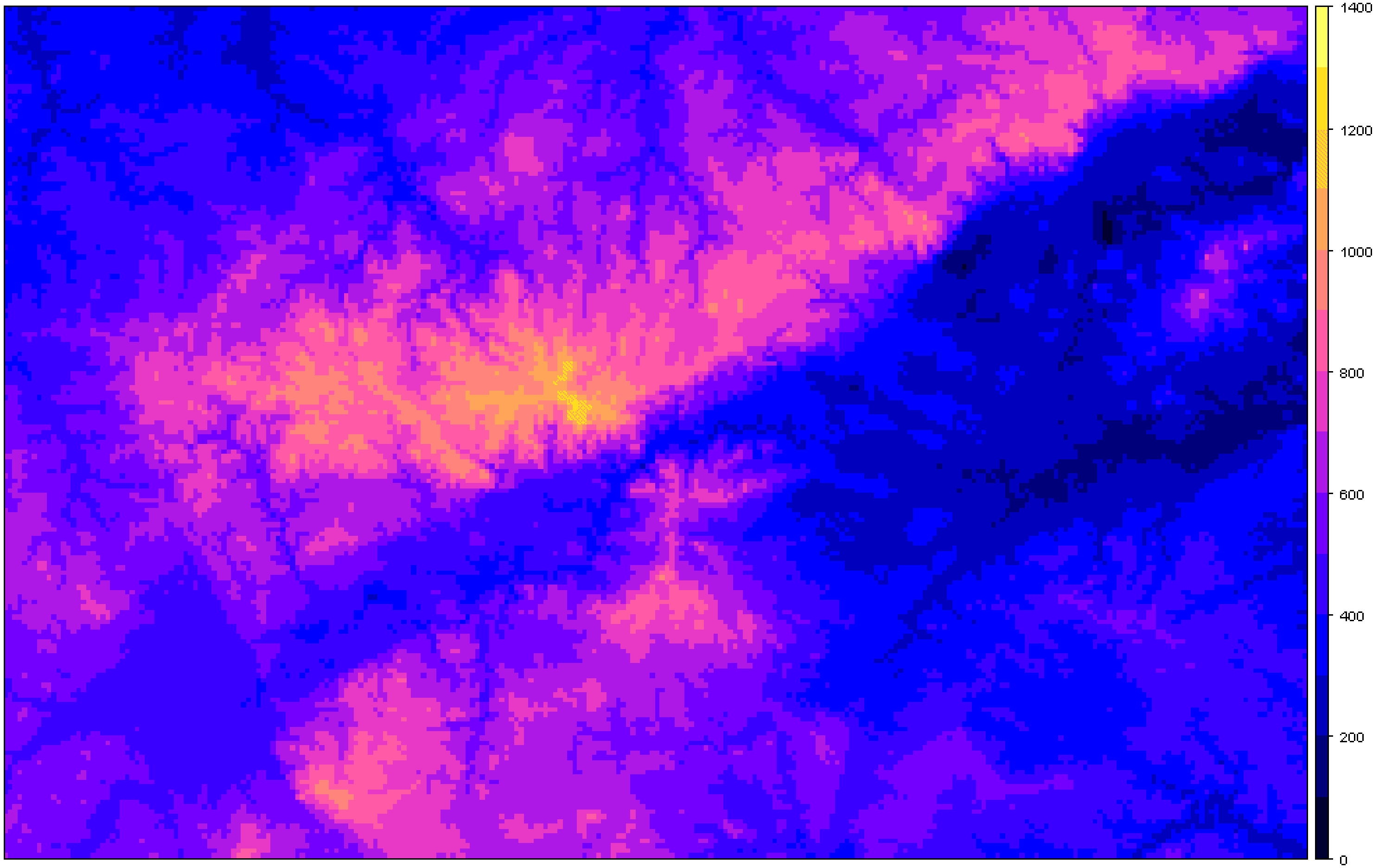
Monte Carlo algorithm:

1. Repeat N times ($N \geq 100$):
    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)

Exercise 1

- Form groups of two or thee persons, make sure to have at least one laptop in your group
- Open file ‘PERC uncertainty Tuesday exercise 1.pdf’ from MS-Teams
- Run the R-script, address al questions, perhaps partly work on your own but also discuss regularly within your group

For spatially distributed variables we might need to take spatial correlation into account. Why is this very important for slope uncertainty propagation analysis? This can be done with geostatistical methods.



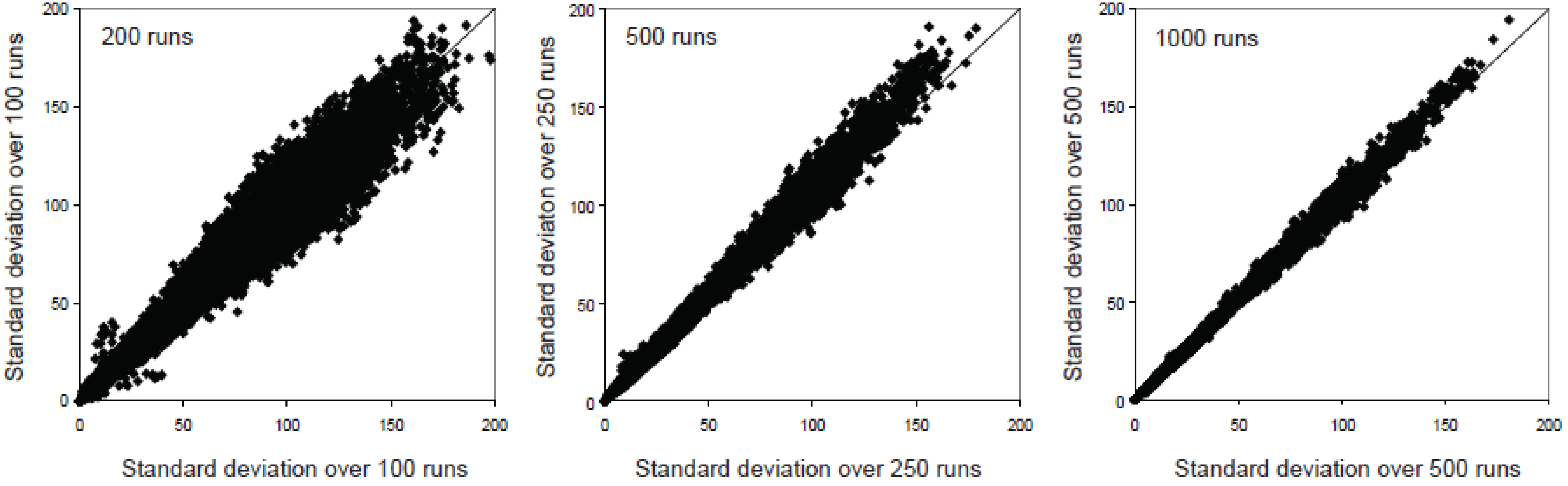
How large must the number of Monte Carlo runs be?

- The larger the better, but this may become a problem when the model is computationally demanding
- Accuracy of estimated output variance is inversely proportional to the square root of the Monte Carlo sample size N (assuming close to normal distribution):

$$sd(\hat{\sigma}_U{}^2) \cong \hat{\sigma}_U{}^2 \cdot \sqrt{\frac{2}{N}}$$

- Efficiency can be improved using stratified sampling from the pdf (e.g. Latin hypercube sampling in case of multiple uncertain inputs)

Required number of Monte Carlo simulations can also be evaluated empirically (keep on increasing until results become stable)



Comparison Taylor and Monte Carlo methods

Advantages Taylor method:

- Fast
- Analytical result

Advantages Monte Carlo method:

- Yields the full output pdf, not only the mean and variance
- Approximation error can be made arbitrarily small
- Works with any model
- Easy to implement

Today we also address some other topics:

1. Uncertainty source contributions
2. Added value of uncertainty quantification
3. Uncertainty about the uncertainty
4. Real-world application of spatial environmental uncertainty propagation

Recall from Taylor method yesterday:

$$Var(U) = \sum_{i=1}^{m} Var(A_i) \cdot \left(\frac{\partial g}{\partial A_i}\right)^2$$

- This shows that the output uncertainty is a sum of various contributions, each caused by one of the uncertain inputs
- This tells you which is the main source of uncertainty
- It also tells you how output uncertainty can best be reduced
- The analysis can also be done in Monte Carlo mode, by fixing individual inputs on their default value (assume that they are certain) and calculating the uncertainty reduction
- Want to know more? Take the PE&RC Statistical Uncertainty Analysis of Dynamic Models course

Exercise 2

- Work in groups as before
- Open file ‘Exercise PERC uncertainty Tuesday 2.pdf’ from MS-Teams
- Address al questions, perhaps partly work on your own but also discuss regularly within your group

Today we also address some other topics:

1. Uncertainty source contributions
2. Added value of uncertainty quantification
3. Uncertainty about the uncertainty
4. Real-world application of spatial environmental uncertainty propagation

Why pay attention to uncertainty?

- Any self-respecting researcher should want to check the quality of his/her results before these are made public
- Clients and end users must know the quality of model outputs to judge their usability for specific purposes
- Quantified uncertainty allows to compare the performance of models, evaluate which performs best, take informed decisions which model to use in which case
- 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: $Expected\ costs = \sum P(outcome) \cdot Cost(outcome)$
- Many users are risk-aversive, you as well, let's play a game:
  - You give me 1 euro, I toss a coin, if heads you loose your euro, if tails I give you 3 euro back
  - Same story but now: 10 euro; 100; euro; 1,000 euro; 10,000 euro,...

Today we also address some other topics:

1. Uncertainty source contributions
2. Added value of uncertainty quantification
3. Uncertainty about the uncertainty
4. Real-world applications of spatial environmental uncertainty propagation

Uncertainty about uncertainty

- We quantified uncertainty with a statistical model (pdf), is it the true uncertainty? Probably not: it is only a model
- This does not make our efforts futile, but it is sensible to critically evaluate the assumptions made:
  - Robustness analysis: would the results of an uncertainty propagation be very different if different assumptions were made?
  - Use independent validation data to check the predicted uncertainty, e.g.

$$SRMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{(O_i - U_i)^2}{Var(U_i)}}$$

  - where O_i is the observed model output. What value should $SRMSE$ have?
- Alternative if pdf cannot be defined: scenarios

Alternative when uncertain inputs cannot be characterised by pdfs: **scenarios**

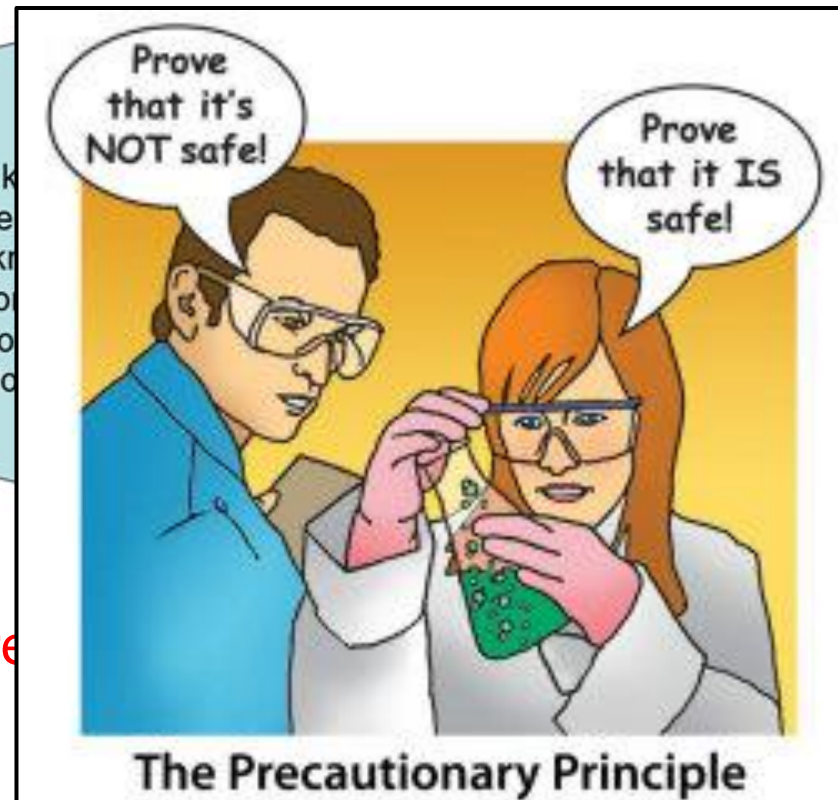
Economic ⟵ ⟶ Environment

Global ↑ ↓ Local

	Economic	Environment
Global	**A1 World Markets** Rapid economic growth, population peaks mid-century then declines, increased globalisation A1FI - Increase use of fossil fuels A1T - non fossil use A1B - balanced growth	**B1 Global Sustainability** A convergent world with rapid change in economic structures, "dematerialization", introduction of clean technologies, global solutions to environmental and social sustainability.
Local	**A2 National Enterprise** Heterogeneous world. strengthening regional cultural identities, family values and local traditions, high population growth, and less concern for rapid economic development	**B2 Local Stewardship** Local solutions to economic, social, and environmental sustainability, move away from globalisation, slower population growth

There are also the unknown unknowns

The Rumsfeld Theorem



This is the rationale behind the pre

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Today we also address some other topics:

1. Uncertainty source contributions
2. Added value of uncertainty quantification
3. Uncertainty about the uncertainty
4. Real-world application of spatial environmental uncertainty propagation

OPEN ACCESS Freely available online

PLOS ONE

Incorporating DEM Uncertainty in Coastal Inundation Mapping

Javier X. Leon[1,2]*, Gerard B. M. Heuvelink[3], Stuart R. Phinn[2]

1 Global Change Institute, The University of Queensland, Brisbane, Queensland, Australia, 2 School of Geography, Planning and Environmental Management, The University of Queensland, St Lucia, Australia, 3 Soil Geography and Landscape group, Wageningen University, Wageningen, The Netherlands

Abstract

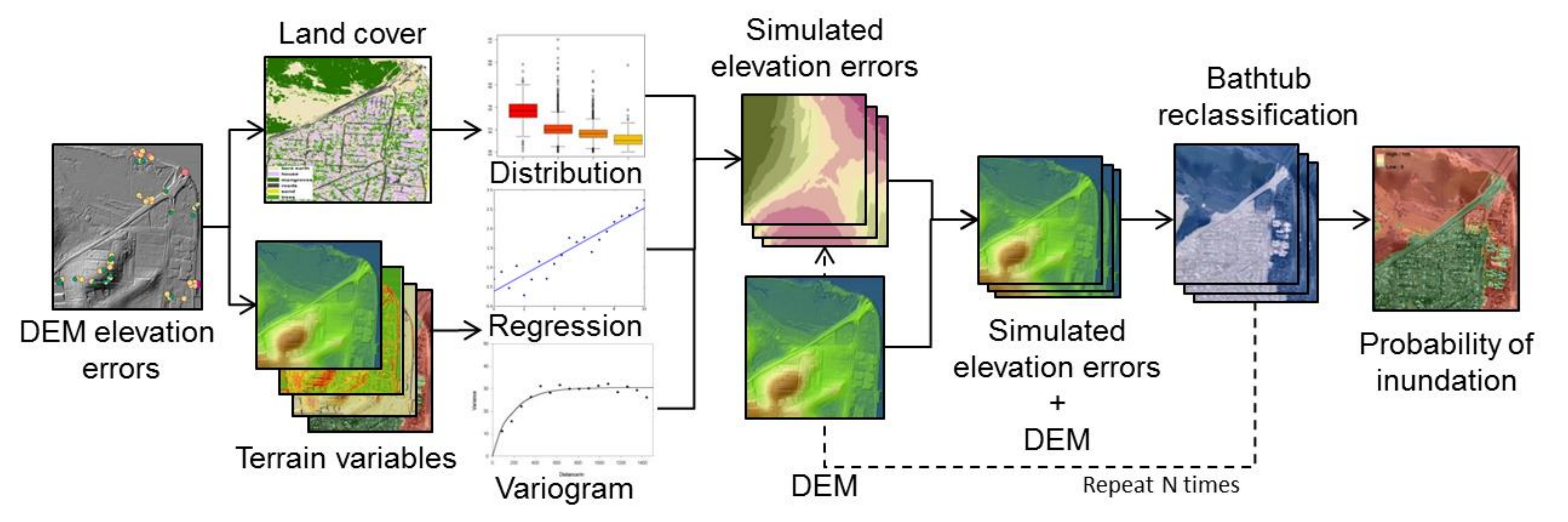
Coastal managers require reliable spatial data on the extent and timing of potential coastal inundation, particularly in a changing climate. Most sea level rise (SLR) vulnerability assessments are undertaken using the easily implemented bathtub approach, where areas adjacent to the sea and below a given elevation are mapped using a deterministic line dividing potentially inundated from dry areas. This method only requires elevation data usually in the form of a digital elevation model (DEM). However, inherent errors in the DEM and spatial analysis of the bathtub model propagate into the inundation mapping. The aim of this study was to assess the impacts of spatially variable and spatially correlated elevation errors in high-spatial resolution DEMs for mapping coastal inundation. Elevation errors were best modelled using regression-kriging. This geostatistical model takes the spatial correlation in elevation errors into account, which has a significant impact on analyses that include spatial interactions, such as inundation modelling. The spatial variability of elevation errors was partially explained by land cover and terrain variables. Elevation errors were simulated using sequential Gaussian simulation, a Monte Carlo probabilistic approach. 1,000 error simulations were added to the original DEM and reclassified using a hydrologically correct bathtub method. The probability of inundation to a scenario combining a 1 in 100 year storm event over a 1 m SLR was calculated by counting the proportion of times from the 1,000 simulations that a location was inundated. This probabilistic approach can be used in a risk-aversive decision making process by planning for scenarios with different probabilities of occurrence. For example, results showed that when considering a 1% probability exceedance, the inundated area was approximately 11% larger than mapped using the deterministic bathtub approach. The probabilistic approach provides visually intuitive maps that convey uncertainties inherent to spatial data and analysis.

Citation: Leon JX, Heuvelink GBM, Phinn SR (2014) Incorporating DEM Uncertainty in Coastal Inundation Mapping. PLoS ONE 9(9): e108727. doi:10.1371/journal.pone.0108727

Editor: Guy J.-P. Schumann, University California Los Angeles, United States of America

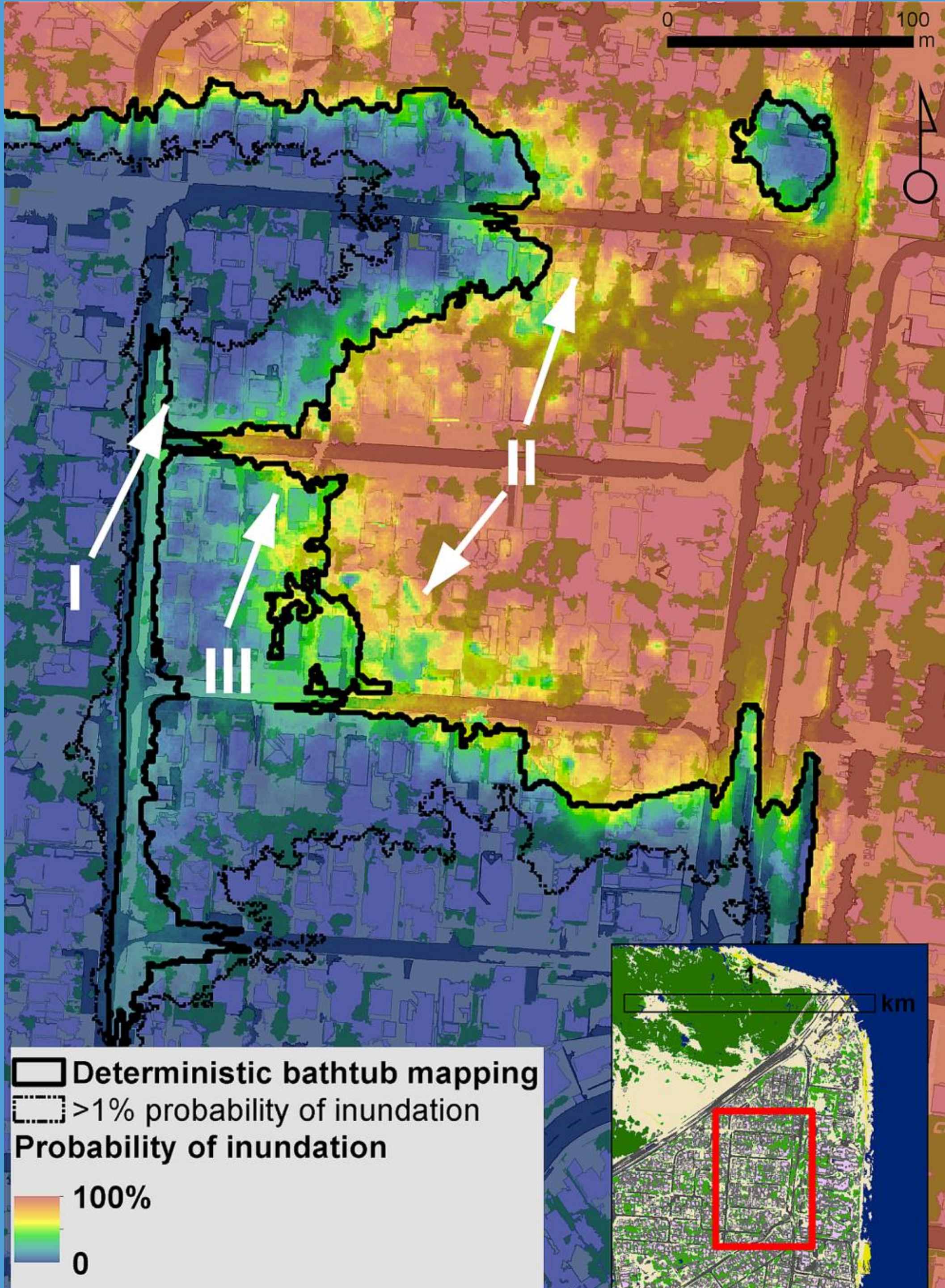
Brisbane (Australia) was flooded in 2010

Analyse propagation of DEM uncertainty using the Monte Carlo method



Results for Brighton, suburb of Brisbane

- The costs of declaring an unsafe area safe are far greater than declaring a safe area unsafe
- One should not go for the deterministic result but put the boundary at a low probability



WAGENINGEN UNIVERSITY
WAGENINGEN UR

Tomorrow's topic: propagation of model parameter and model structure uncertainty