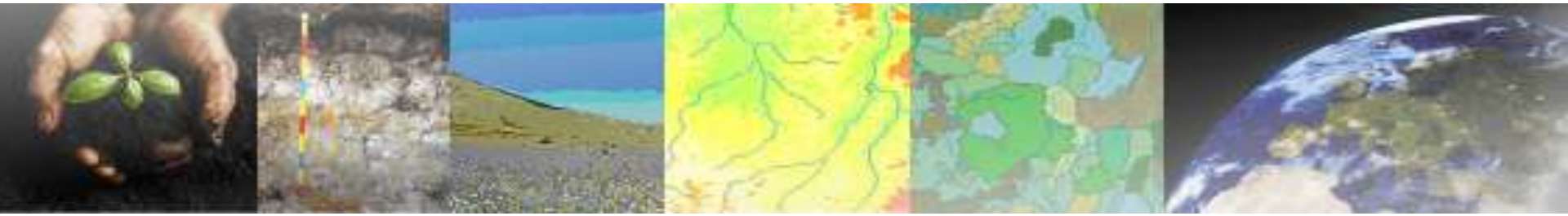
e-SOTER

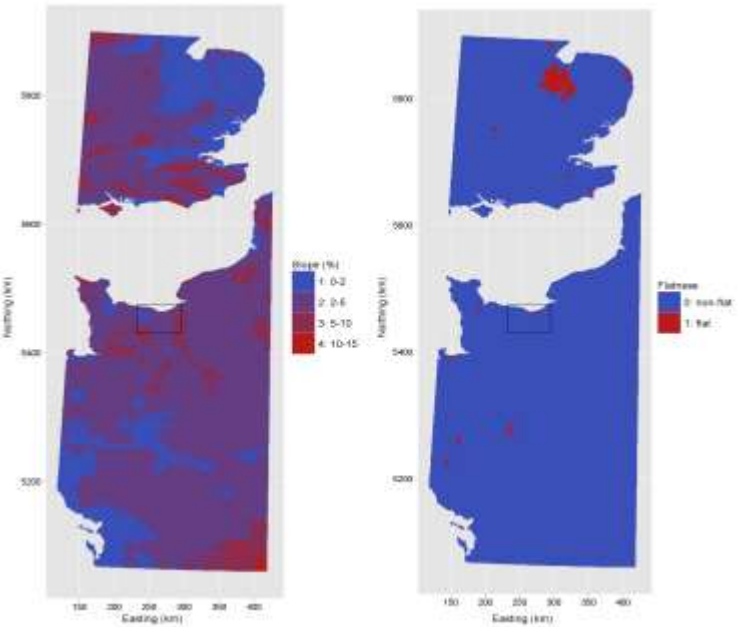
Regional pilot platform as EU contribution to a Global Soil Observing System

Validation and Uncertainty Analysis of e-SOTER products

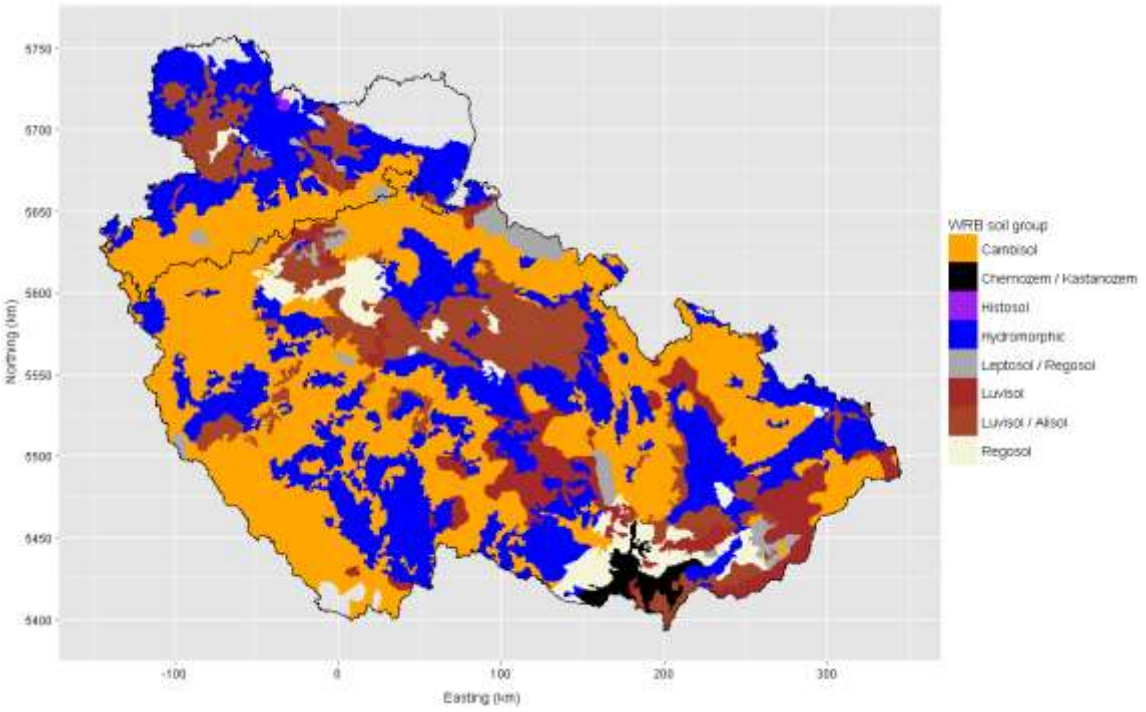
Bas Kempen, Gerard Heuvelink, Tereza Zádorová, Vít Penížek, Jacqueline Hannam, Rainer Baritz and Ulrich Schuler



e-SOTER produced nice maps



BUT HOW GOOD ARE THESE?



Validation

Validation = comparison of map predictions with independent, (nearly) error-free observations of reality

Landform validation

- e-SOTER – WP1 defines landform as a composite of four landform attributes: elevation class, slope class, relief intensity class and flatness index class
- Therefore reality can be derived exhaustively from an accurate DEM
- However, the e-SOTER landform map deviates from reality because the e-SOTER algorithm involves several aggregation and generalization steps

Soil validation

- e-SOTER – WP2 creates a soil type map with a WRB legend, often as associations of multiple soil components per e-SOTER mapping unit
- Reality can be observed at validation locations by digging pits and classifying the soil
- It is cheaper to use independent existing soil data, although this may require a 'correlation' between the local soil legend and WRB

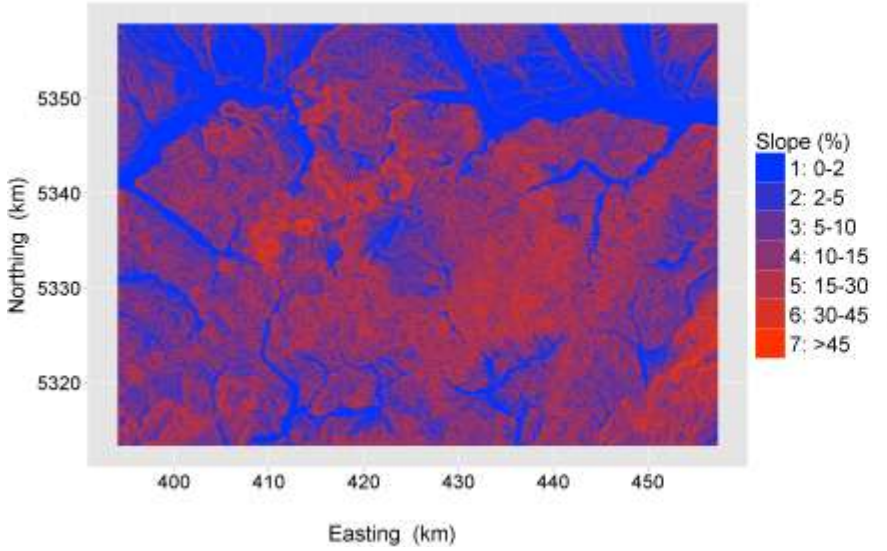
Validation measures are derived from the error matrix

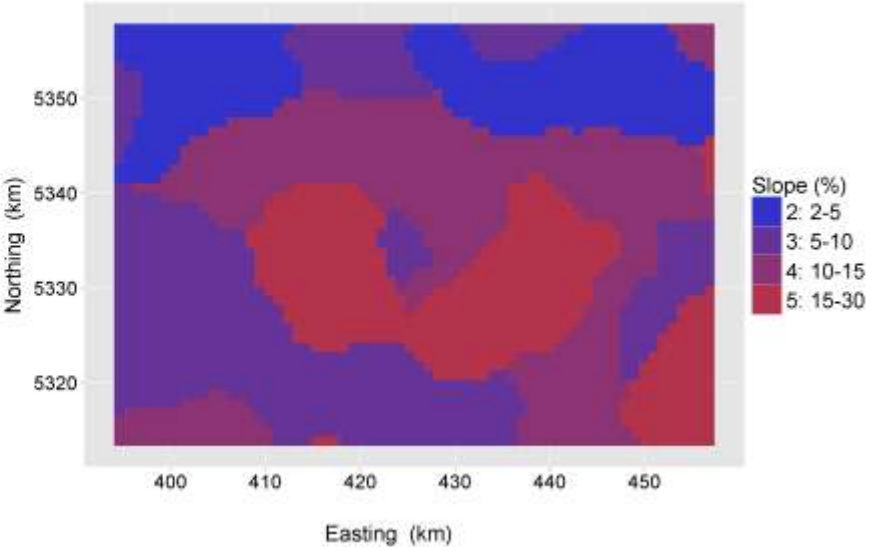
Mapped class	True class						
	1	2	…	…	…	U	Σ
1	n_{11}	n_{12}	…	…	…	n_{1U}	n_{1+}
2	n_{21}	n_{22}	…	…	…	n_{2U}	n_{2+}
.	…	…	…	…	…	…	…
.	…	…	…	…	…	…	…
.	…	…	…	…	…	…	…
U	n_{U1}	n_{U2}	…	…	…	n_{UU}	n_{U+}
Σ	n_{+1}	n_{+2}	…	…	…	n_{+U}	N

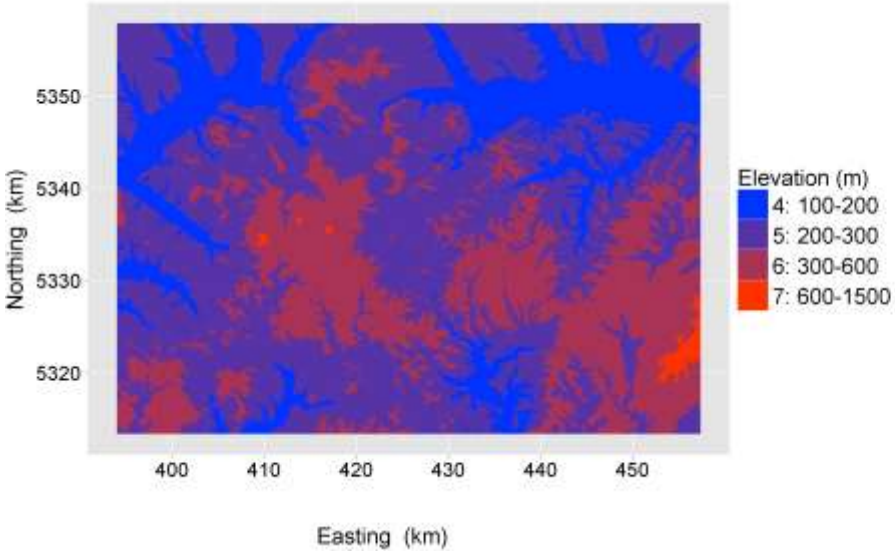
strict purity = percentage correctly classified = $\frac{100}{N}\sum_{u=1}^{U} n_{uu}$

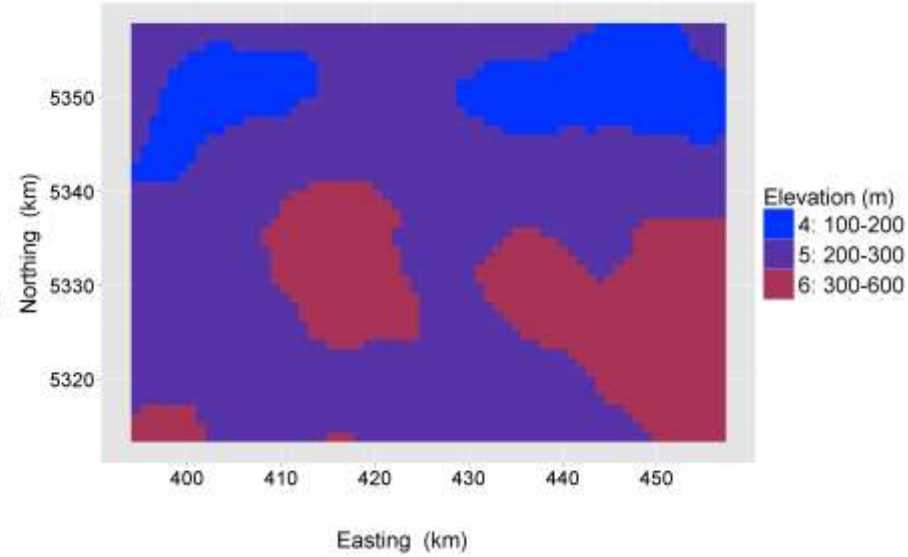
'one-off' purity → allow differences by one class

Results landform validation (Central European window)







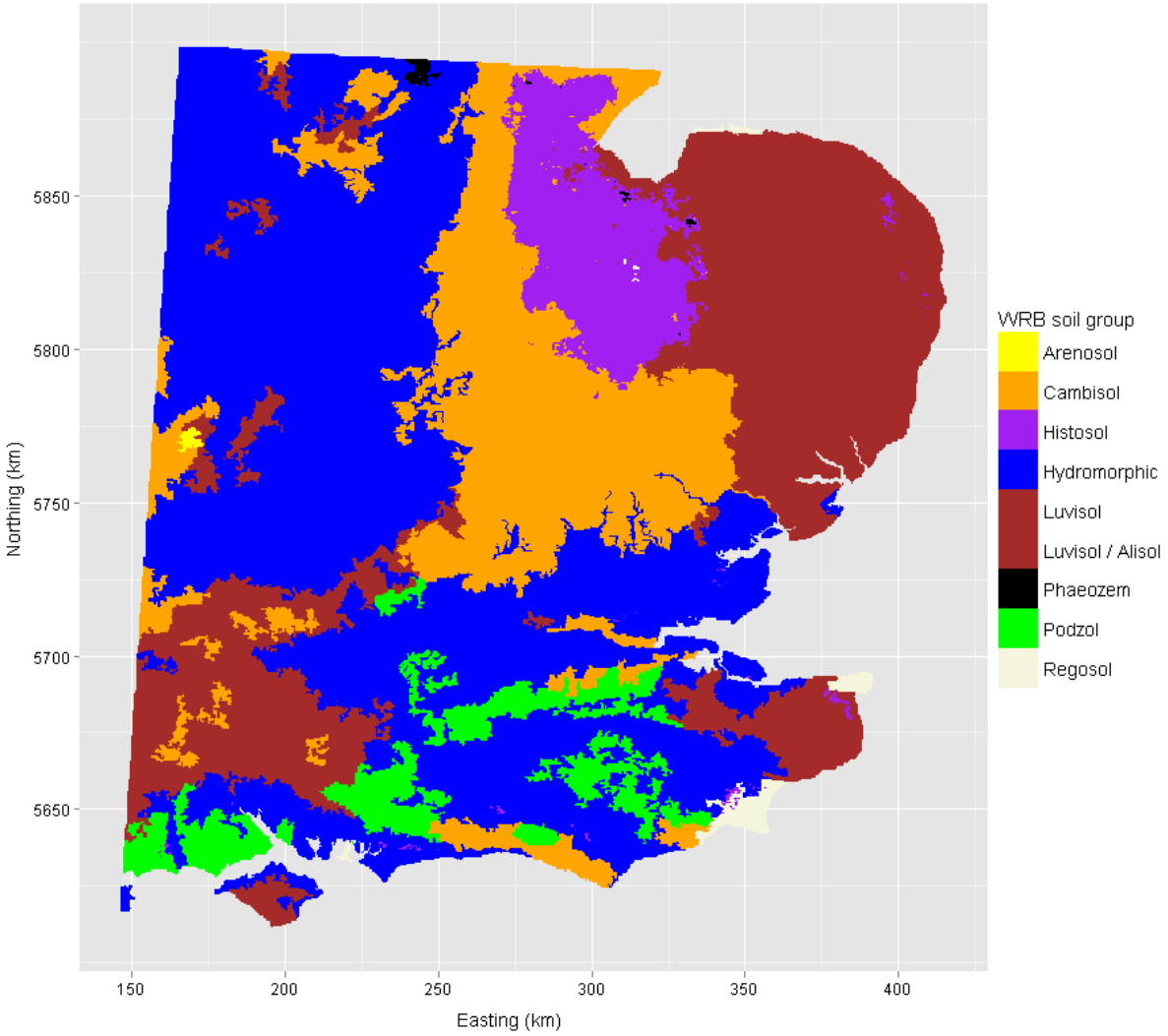


Results landform validation

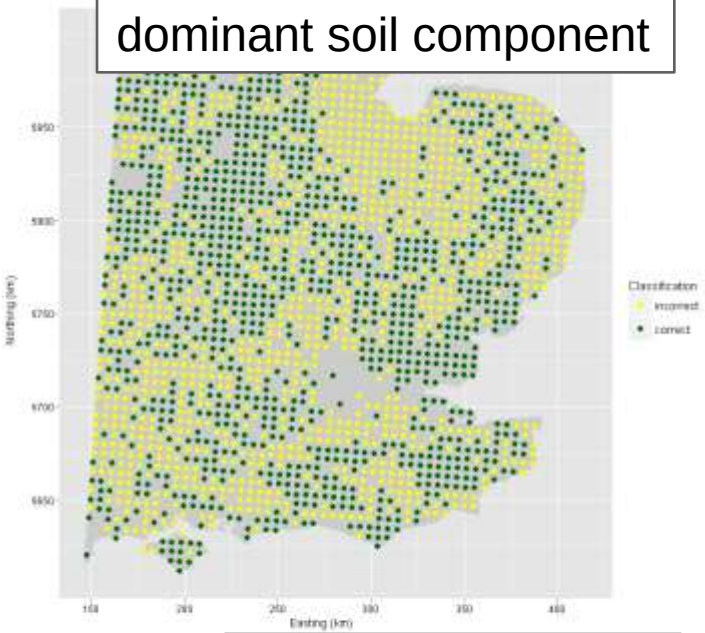
Landform attribute	Western European window		Central European window	
	strict purity	'one-off' purity	strict purity	'one-off' purity
Elevation	81.1	99.4	87.8	99.8
Relief intensity	92.0	99.7	81.1	98.5
Slope	44.6	94.8	50.6	86.6
Flatness	98.3		98.1	

➔ smallest purity for slope because original map highly fragmented
➔ WE and CE windows have comparable purities

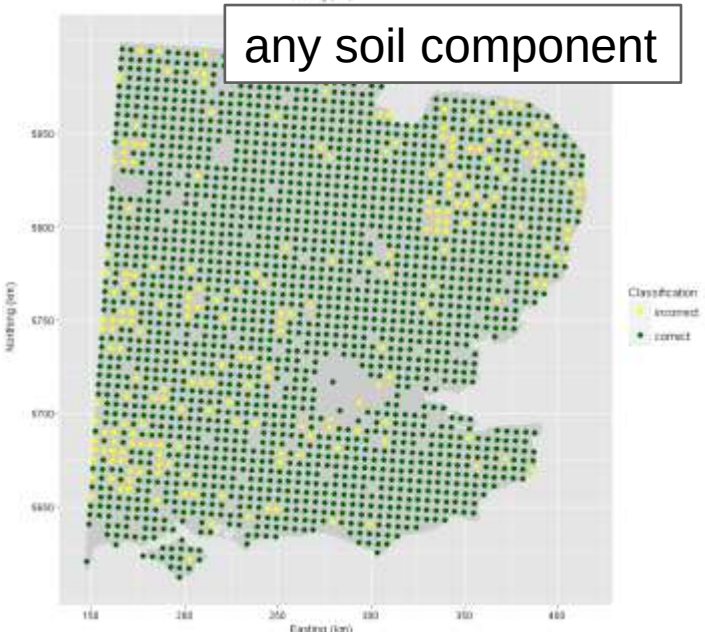
Results soil validation (UK part Western European window)



dominant soil component



any soil component



Results soil validation

Nr of soil components in mapping unit	UK part Western European window	G/CZ part Central European window
1	51.0	31.2
2	65.4	
3	76.8	
4	83.7	
5	87.6	
Any soil component (in the association)	91.6	86.0

- ➔ purities are fairly small, particularly in stringent case when match required with dominant soil component
- ➔ UK area has a larger purity than the G/CZ area
- ➔ Small purities may also be caused by errors in validation data (e.g. 'correlation' errors)

Uncertainty analysis

Uncertainty (propagation) analysis = analysis how errors/uncertainties in the input of the e-SOTER algorithm propagate to the output

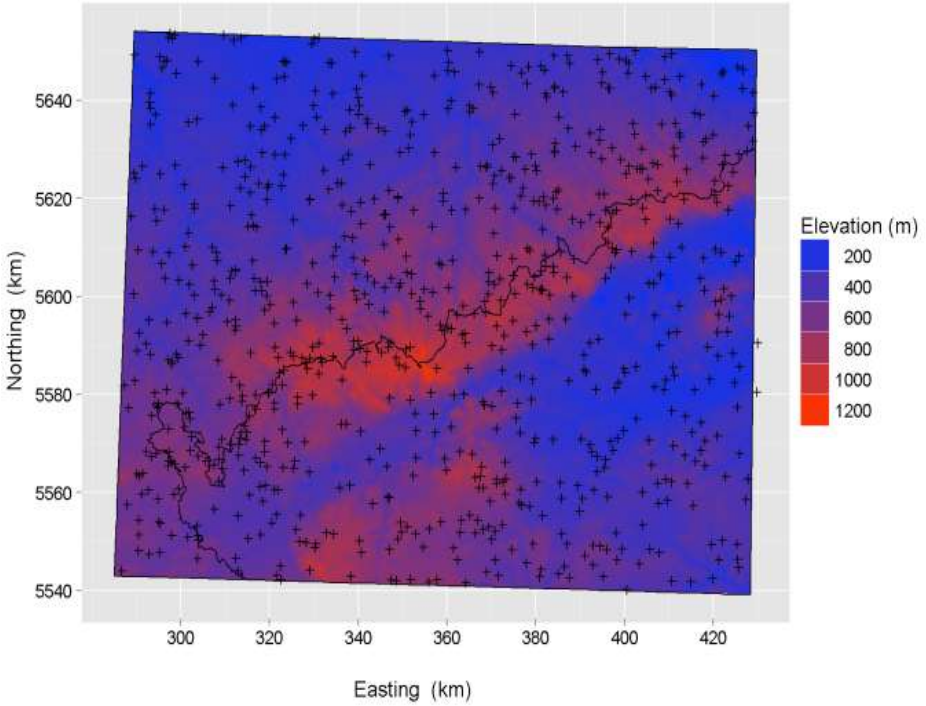
Uncertainty analysis, WHY?

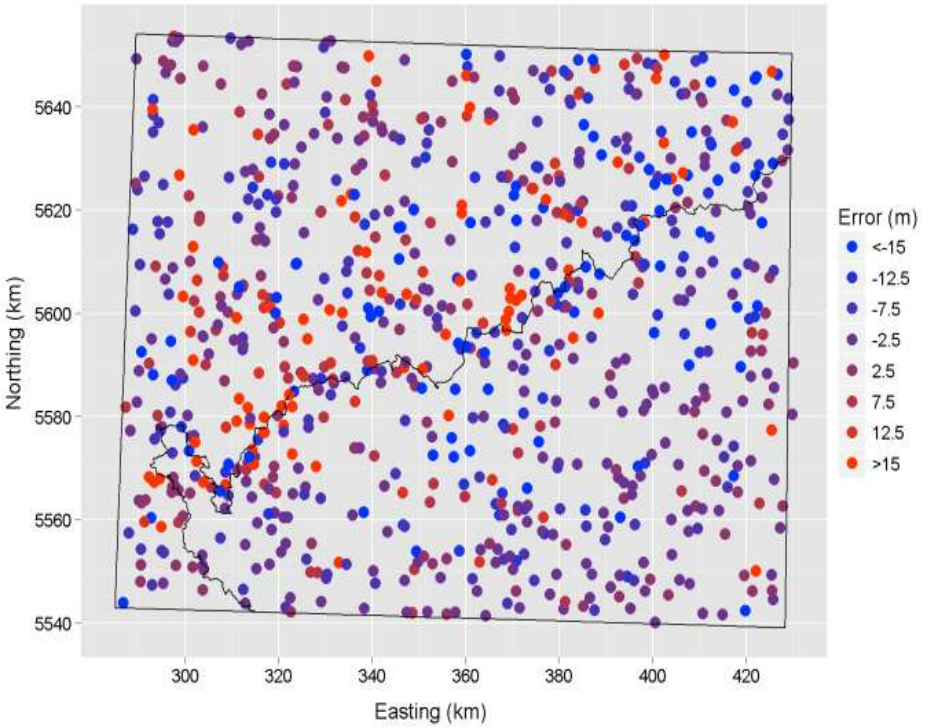
- Unlike validation, uncertainty analysis can quantify the contribution of sources of uncertainty, i.e. it can identify the weakest links
- In e-SOTER we only looked at propagation of DEM error through the landform classification algorithm

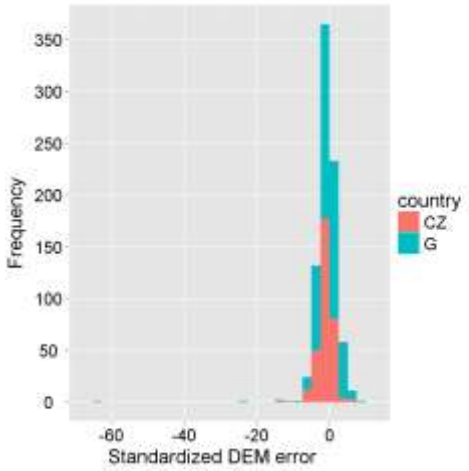
Uncertainty analysis, HOW?

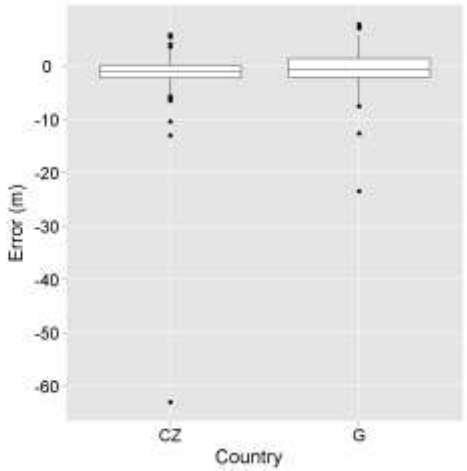
- Monte Carlo simulation:
  - Sample from the probability distribution of uncertain inputs using a random number generator
  - Run model and store result
  - Repeat the above many times (we used 1,000 Monte Carlo runs)
  - Calculate summary statistics of results (e.g. probability for each class, dominant class, entropy)

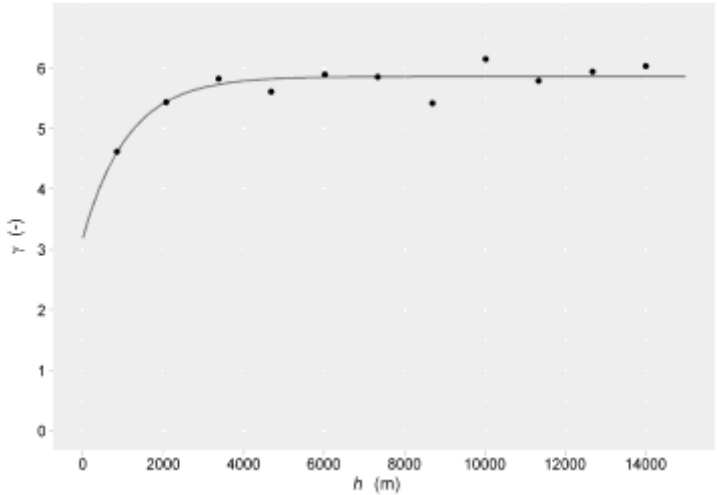
Results for CE pilot area (G/CZ border area)



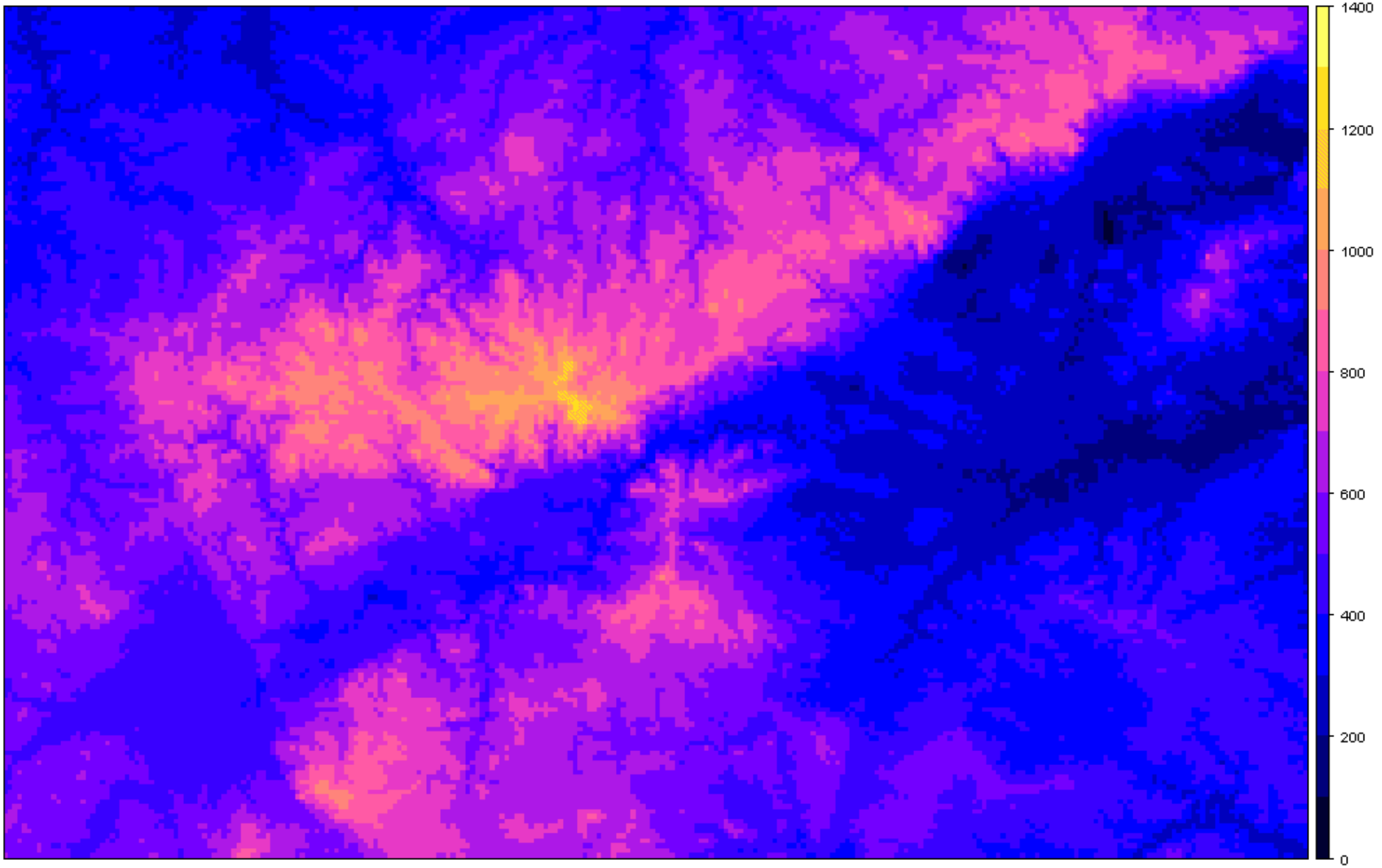




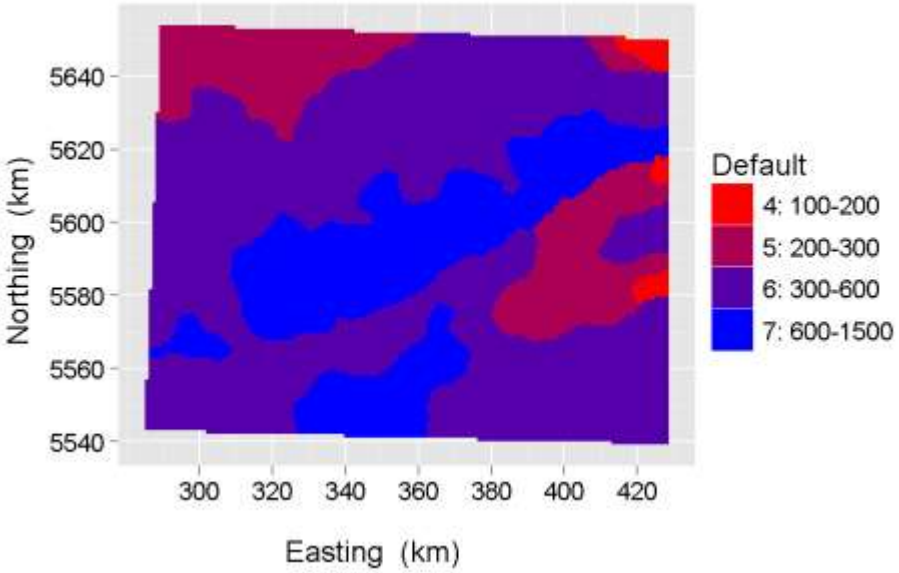


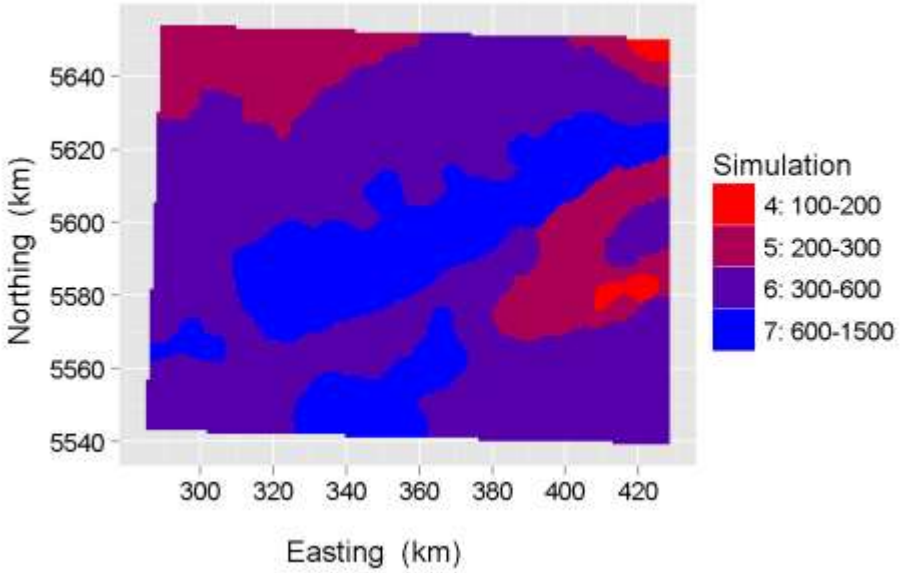


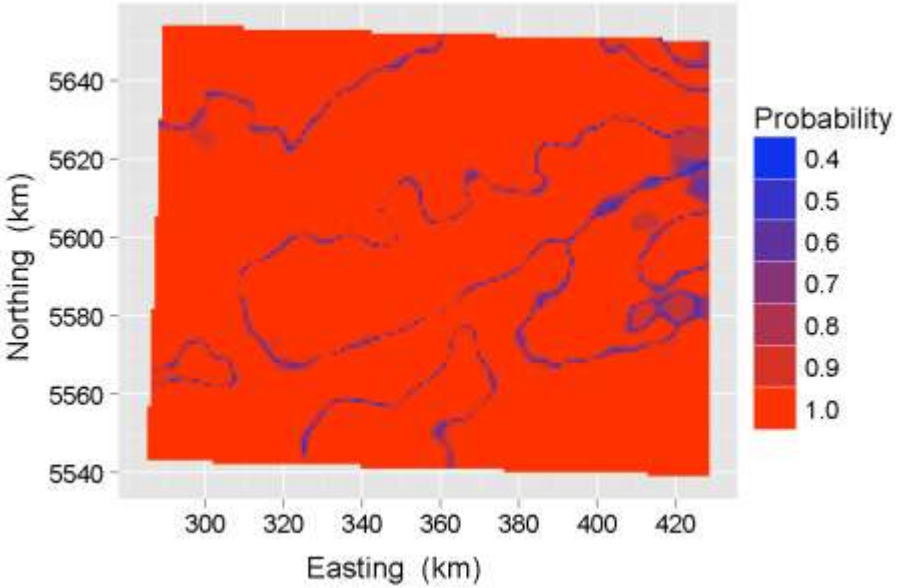
Possible realities of the ‘true’ DEM

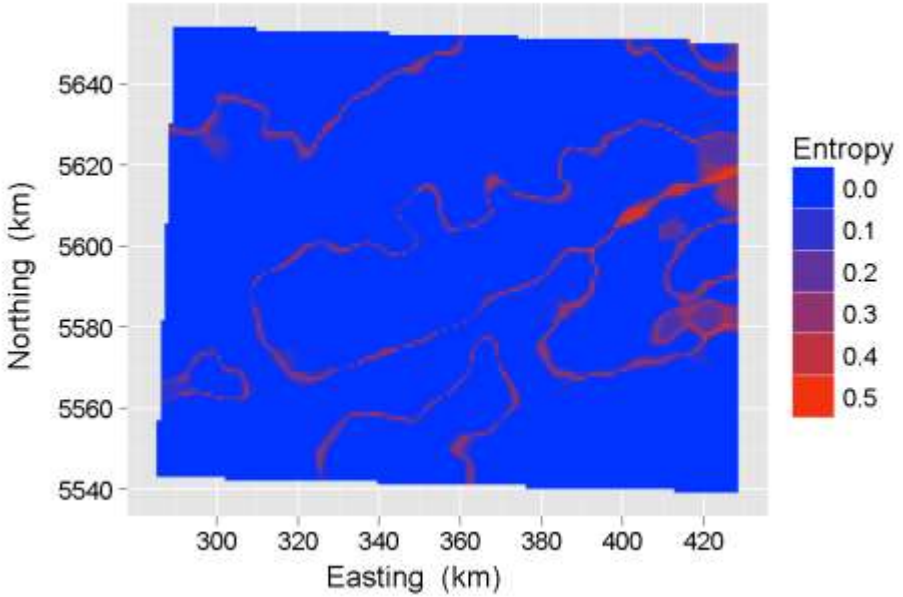


Uncertainty propagation to elevation class

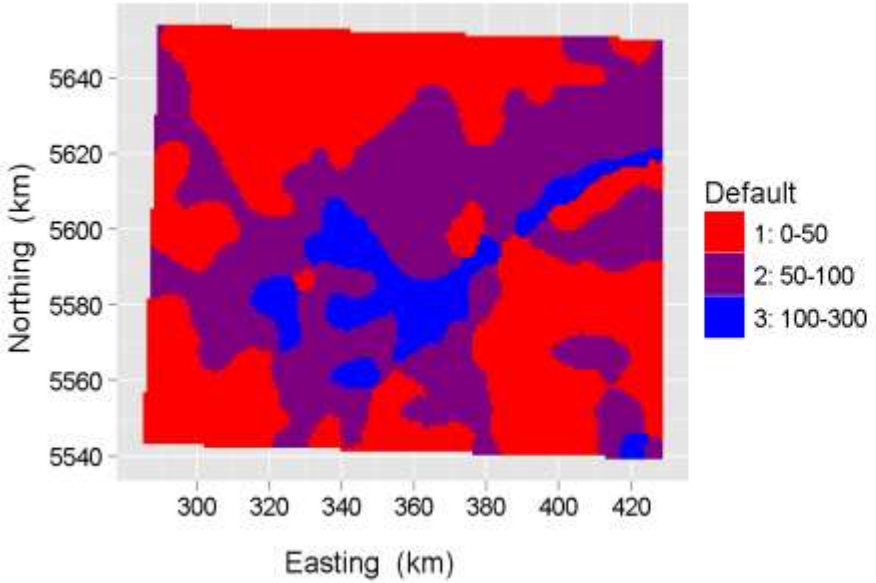


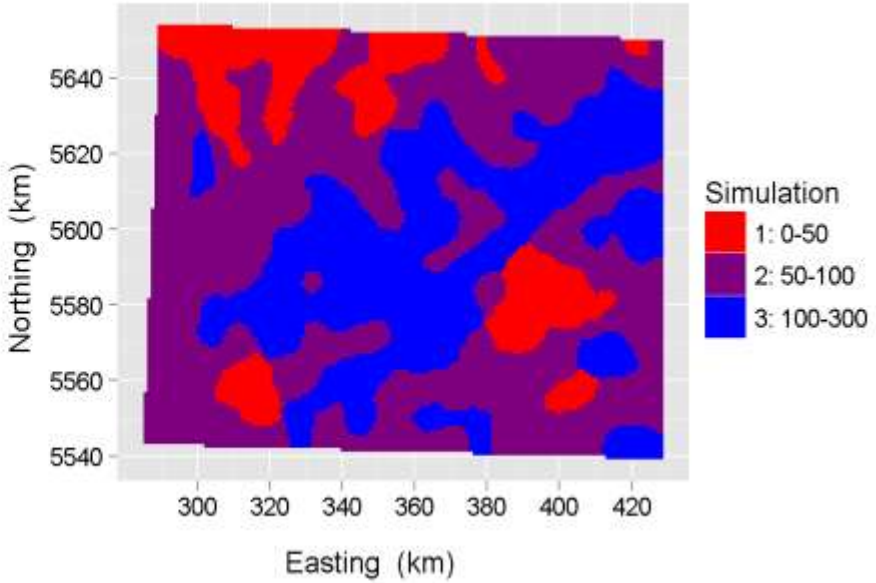


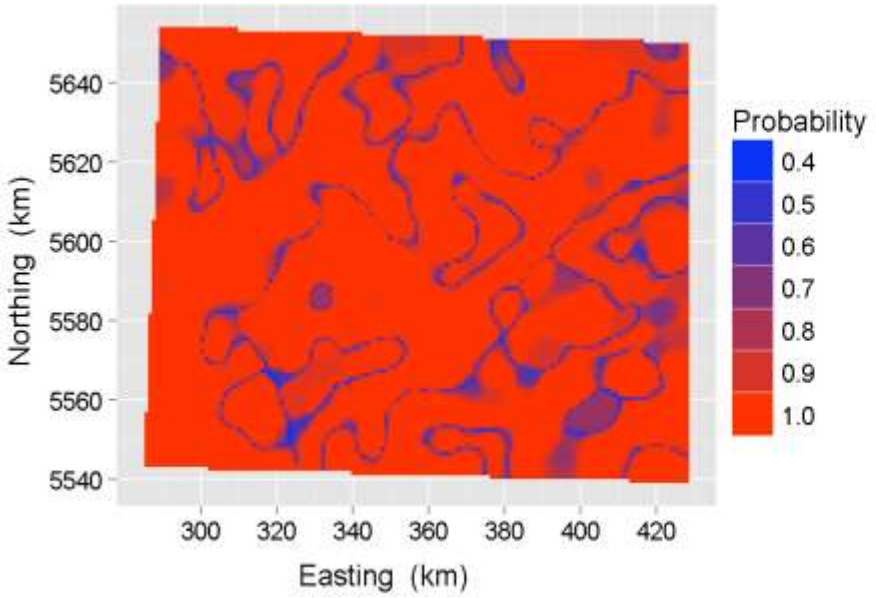


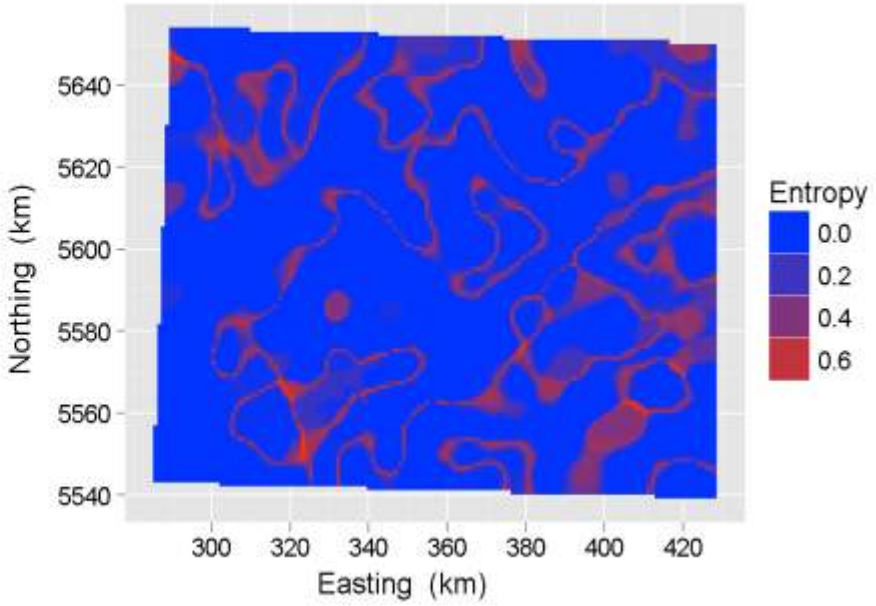


Uncertainty propagation to slope class









Conclusions (1/2)

Landform validation

- Generally large purities: negative effect of generalization steps is limited
- Slope most affected: can be solved by using fewer than seven slope classes, which will reduce the generalization effect

Soil validation

- The 51% purity for the UK area is not that bad. The map shows the general soil spatial patterns, which is the purpose of a 1:1M soil map
- Important error sources in the UK e-SOTER soil map are the over-representation of Histosols and Podzols and the absence of Leptosols as a dominant soil group
- Overall purity for the G/CZ validation area is 32%. The difference with the UK area can be partly explained by stricter validation criteria
- Important error sources in the G/CZ e-SOTER soil map are the under-representation of Chernozems and Podzols and the confusion between Hydromorphic soils, Cambisols and Luvisols

Conclusions (2/2)

Uncertainty analysis

- DEM uncertainty has the largest effect on slope class. The dominant slope on the basis of 1,000 simulations is typically one class above the default class. This is because uncertainty adds ‘noise’
- In the more rugged CE pilot area slope and relief intensity are most affected by DEM error
- Uncertainty about the prevailing landform attribute, quantified by the entropy, is generally small. The largest uncertainties are found in zones along the class boundaries

WP3 landform validation

- We also did a validation of the WP3 landform maps for the UK part of the WE window, by verifying how homogeneous these are with respect to the soil component (results not presented)
- Both the hillshed and the object-oriented approach give better results than the WP1 landform map at subclass level, although improvements in predictability and purity are modest

Many more details:

Project acronym	e-SOTER
Project full title	Regional pilot platform as EU contribution to a Global Soil Observing System
Project No	211578

e-SOTER validation and accuracy assessment

Deliverable D10

March 2012

Thank you