



Project acronym

e-SOTER

Project full title

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

Project No

211578

**Deliverable  
D10**

**e-SOTER validation and  
accuracy assessment**

March 2012

---

**SEVENTH FRAMEWORK PROGRAMME  
Environment**

ENV.2007.4.1.3.3 Development of a Global Soil Observing System



## Document Information

Deliverable number	D10
Deliverable title	e-SOTER validation and accuracy assessment
Period covered	n.a.
Due date of deliverable	29/02/2012
Actual date of deliverable	
Author(s)	Bas Kempen, Gerard Heuvelink (Alterra), Tereza Zádorová, Vít Penížek (CULS), Jacqueline Hannam (CU), Rainer Baritz, Ulrich Schuler (BGR)
Participants	All WP4 partners
Work package	4
Work package title	Accuracy assessment of terrain and soil platforms
Dissemination level	PU
Version	1.0

## History of Versions

Version	Date	Status	Author	Approval level
1.0	06/03/2012	Draft	Bas Kempen	
2.0	14/03/2012	Final	Bas Kempen	

## Table of Contents

Table of Contents.....	3
1. Executive Summary.....	4
2. Introduction .....	6
3. Landform Validation .....	8
3.1 Methodology .....	8
3.2 Application to Western European window .....	11
3.3 Application to CE window.....	16
4. Soil Validation .....	21
4.1 Methodology .....	21
4.2 Application to UK part of the Western European window.....	24
4.3 Application to German-Czech part of the Central European window.....	34
5. Uncertainty Propagation Analysis.....	46
5.1 Methodology .....	46
5.2 Application to the Western European pilot area .....	49
5.3 Application to Central European pilot area.....	55
6. Validation of the WP3 landform map .....	62
6.1 Methodology .....	62
6.2 Results .....	63
7. Conclusions and recommendations.....	70
8. References .....	73
Appendix 1. Site, soil profile and topsoil data of the NSI dataset (UK) .....	74
Appendix 2. Landform classification levels of the WP3 maps .....	77

## 1. Executive Summary

This deliverable reports on the validation and uncertainty propagation analysis of e-SOTER products by comparing the e-SOTER soil and landform maps with independent validation data and by analysing how errors in a Digital Elevation Model (DEM) propagate to the e-SOTER landform map.

Landform validation was done by comparing the true landform in the Western European (WE) and Central European (CE) windows with the landform as depicted on the e-SOTER maps. The various simplification and generalization steps of the e-SOTER landform classification methodology cause discrepancies between the true and predicted landform. Validation showed that the accuracy of the e-SOTER landform maps is high for landform classes 'elevation', 'relief intensity' and 'flatness', with agreement between true and predicted class in 81-98% of cases. For landform class 'slope' the agreement is only around 50%, which means that in 1 out of 2 cases the e-SOTER map does not agree with reality. This is caused by the highly fragmented spatial pattern of the true slope map, a feature that cannot be reproduced in the e-SOTER map because it must generalize the map to the 1:1000,000 million scale. Comparison of landform validation results between the WE and CE windows did not show large or meaningful differences.

Soil validation was based on a comparison of the dominant soil classes on the e-SOTER maps with independent soil observations derived from existing legacy data. This first required a conversion of soil classes from the local classification system to the World Reference Base (WRB). For this purpose correlation tables were prepared. The validation results show that the e-SOTER soil map of the UK pilot area reproduced the large scale patterns and had an agreement with the 'true' soil class of 51%. The e-SOTER map overrepresented Histosols and Podzols and lacked Leptosols as a dominant soil group. Validation results in the German / Czech pilot area revealed an overall agreement of only 32%. The rather low purity in this case can be assigned to high variability of the soil cover and often low dominancy of the dominant soil unit in the SOTER units. Also, in this case the more detailed validation data allow the use of more strict validation criteria than in the UK case. Overall, the fairly low agreements between e-SOTER soil maps and validation data in both pilot areas show that the soil maps have large uncertainties, even at the coarse soil group level that was considered here. Part of the disagreement may be caused by errors in the validation data, but this is unlikely to be the major cause of the discrepancy.

The uncertainty propagation analysis showed that DEM uncertainty mainly affects the e-SOTER slope class. The real elevation is more noisy than the smoothed DEM and this causes the DEM-derived slope to be too small. Elevation was hardly affected while flatness is only sensitive to DEM errors in relatively flat areas and relief intensity in areas with a more intense relief (i.e. the CE pilot area). Overall DEM uncertainty does not seriously impair landform accuracy and is mainly restricted to zones along class boundaries. It is a smaller source of uncertainty than the simplification and generalization steps that were analysed in the landform validation procedure.

Validation of the WP3 landform maps for the UK area of the Western European window in terms of predictability of the WRB reference soil groups indicated that the hillshed analysis gives the best overall results. Both the hillshed and the object-oriented approach give better results than the WP1 map at subclass level, although differences in predictability and purity are modest. The entropy is equal for all three landform maps, indicating that the WP3 landform units are not internally more homogeneous with respect to soil distribution than the WP1 units.

## 2. Introduction

The e-SOTER project used much of its resources to automate and improve the SOTER methodology such that it yields reproducible results that utilise new sources of (remote sensing) information that have recently become available. The new methodology and results for pilot areas and windows have been described in great detail in project deliverables D3, D5 and D8. However, the new methodology and resulting products only have merit when these are sufficiently accurate for the intended use. It is therefore important to validate the resulting maps and analyse how errors in the inputs to the e-SOTER algorithms propagate to the output. This deliverable tackles these issues by confronting the e-SOTER maps of landform and soil with independent validation data and by analysing how errors in the Digital Elevation Model (DEM) propagate through e-SOTER landform classification algorithms.

The term 'validation' is defined here as "the process of determining the degree to which a product is an accurate representation of the real world". Thus landform validation requires a comparison of 'true' landform with landform produced with the e-SOTER methodology, and likewise soil validation makes a comparison of 'true' soil classes with soil classes predicted by the e-SOTER methodology. In order to derive the true landform it is first necessary to define it. In the e-SOTER project the landform at some location is completely determined by the altitude at the location and its surroundings. This implies that landform validation data can be obtained from a detailed and accurate DEM of the area of interest. Discrepancies between true landform and 'e-SOTER landform' are then caused by the various simplification and generalization steps of the e-SOTER methodology. For soil it is different, because in this case the validation data have to be obtained from independent observations in the field. This can be done by collecting new soil data but a more efficient way is to use existing legacy soil data that have not been used by the e-SOTER methods. Since landform and soil class are variables that are measured on a categorical scale, validation measures must be based on the entries of contingency tables that tabulate the various combinations of true and predicted classes observed at validation sites. In short, the e-SOTER product will be an accurate representation of the real world when the predicted class at validation locations often agrees with the observed class, but there are also additional, more specific measures of accuracy that can be derived from contingency tables.

When validation shows that the e-SOTER product fails to accurately describe the real world it may be interesting to analyse what are the main causes of the discrepancy. This is particularly useful when one needs to take rational decisions on how to improve the product. Attention should be focused on the weakest link in the chain of actions, i.e. improving those inputs or processing steps that are the main source of uncertainty in the end product. This is where uncertainty propagation analysis becomes useful. Using a stochastic simulation approach, this method allows to trace the propagation of individual sources of uncertainty and to calculate the relative contribution of each uncertainty source. In this deliverable the uncertainty propagation analysis is limited to an analysis of DEM uncertainty propagation to the e-SOTER landform map. The analysis provides valuable information on whether improvement of the quality of the DEM may lead to a substantial increase of the accuracy of the e-SOTER landform map.

This deliverable presents the theory of validation and uncertainty propagation analyses and applies these to validation and uncertainty propagation analysis of the e-SOTER landform and soil maps of the Western European and Central European windows. The deliverable first addresses landform validation (Chapter 3), next soil validation (Chapter 4), DEM uncertainty propagation (Chapter 5) and finally validation of the work package 3 landform maps (Chapter 6).

## 3. Landform Validation

### 3.1 Methodology

#### 3.1.1 Introduction

Landform validation focuses on analysing how the various aggregation and generalization steps in the e-SOTER procedure affect the output: the Physiographic Unit (PU) map. The PUs are defined by four landform attributes (LFAs): elevation, slope, relief intensity and flatness index.

It is useful to provide a short summary of the e-SOTER procedure to clarify the chosen validation approach. More details are given in deliverables D3 and D5. First, the LFAs are derived from the 90-m SRTM DEM and classified into  $n$  classes (slope 7 classes, elevation 9, relief intensity 4 and flatness 2). Class definitions are given in Dobos et al. (2005). Next the most frequently occurring attribute class within a  $990 \times 990$  m<sup>2</sup> block is calculated for each pixel in the 90-m raster map. The LFA maps are resampled to 990-m resolution and again the most frequently occurring attribute class within a six cell radius is calculated for each pixel in the 990-m map. Finally the elevation, slope and relief intensity maps are combined into one map. Each unique combination of the three LFAs defines a PU. One last generalization step eliminates the PU units smaller than a given threshold. The fourth LFA map, the flatness index, is 'burned' into the PU map.

The above shows that construction of the final PU map involves several steps of generalizing and aggregating the LFA maps and the PU map, losing detail and introducing generalization errors in each step. These steps are needed for cartographic principles (e.g. minimum size of mapping unit) but do impair the quality of the final product. Landform validation assesses how well the mapped LFA classes characterize the PUs after all aggregation and generalization steps have been done, by comparing the 90-m classified LFA maps (which are considered to be the ground truth) with the content of the PUs.

Landform validation was done for the Western and Central European windows. The validated PU map was obtained with version 6 of the e-SOTER procedure (dated 28 April 2011).

#### 3.1.2 Estimation of map quality measures

The elevation, slope and relief intensity attributes are measured on an ordinal scale. This means that these are categorical data (classes) with a logical ordering to the classes. The flatness index is a binary (0/1) variable where 1 indicates a flat area and 0 a non-flat area. This difference between attribute scales has to be taken into account for validation. For example, mapping a location as slope class 2 when in reality it is slope class 3 can be considered 'less wrong' than when the real slope class would have been 6.

The classified 90-m LFA rasters and the PU polygon maps were sampled by randomly selecting point locations with an overall density of 1 point per km<sup>2</sup> (yielding a very large sample of 184,000 point

locations). At these points the ‘true’ LFAs were compared with the LFAs of the PUs. From this comparison several quality measures were computed. We distinguish *single* quality measures, calculated for individual LFAs, and *composite* quality measures, calculated for a combination of LFAs. The quality measures are based on Stehman (1997) and Brus et al. (2011), who provide a thorough methodology for statistical validation of digital (soil) maps.

Single quality measures include the *overall purity*, *LFA class purity* and *LFA class representation*. These three quality measures can be easily estimated from a sample error matrix (or confusion table, Table 3.1) (Foody, 2002; Stehman, 1997). These measures, however, do not take into account the ordinality of the LFA classes. The entries of the error matrix are the number of observations for each combination of mapped and true LFA classes. The row marginals,  $n_{u+}$ , are the number of locations mapped as LFA class  $u$ ,  $u=1,2,\dots,U$ . The column marginals,  $n_{+u}$ , are the number of locations with true LFA class  $u$ . Dividing a matrix entry by a row or column total or the total number of locations gives a proportion.

Table 3.1. Error matrix of mapped against true LFA class.

Mapped LFA class (90 m)	True LFA class (90 m)				$\Sigma$
	1	2	...	U	
1	$n_{11}^a$	$n_{12}$	...	$n_{1U}$	$n_{1+}$
2	$n_{21}$	$n_{22}$	...	$n_{2U}$	$n_{2+}$
.	.	.	...	.	.
.	.	.	...	.	.
.	.	.	...	.	.
U	$n_{U1}$	$n_{U2}$	...	$n_{UU}$	$n_{U+}$
$\Sigma$	$n_{+1}$	$n_{+2}$	...	$n_{+U}$	$N$

<sup>a</sup>  $n_{ij}$  = the number of validation locations mapped as LFA class  $c_i$  with true LFA class  $c_j$ .

Overall purity is defined as the proportion of the total observations (i.e., the mapped area) in which the mapped LFA class equals the true LFA class. In other words, it is the areal proportion of the LFA map correctly classified:

$$p = \sum_{u=1}^U \frac{n_{uu}}{N}, \quad (3.1)$$

where  $U$  denotes the number of LFA classes,  $n_{uu}$  denotes the number of correctly classified validation locations in class  $u$  and  $N$  denotes the total number of validation locations.

The purity can also be defined at the level of the LFA classes (the map unit purity), leading to the proportion of the area of an LFA class correctly classified:

$$p_u = \frac{n_{uu}}{n_{u+}}, \quad (3.2)$$

where  $n_{u+}$  denotes the number of locations mapped as LFA class  $u$ . The complement of  $p_u$ ,  $1-p_u$  is referred to as the error of commission for class  $u$ . It is the proportion of the area incorrectly mapped as class  $u$ .

The LFA class representation for class  $u$  is the proportion of the area where in reality LFA class  $u$  occurs that is also mapped as class  $u$ :

$$r_u = \frac{n_{uu}}{n_{+u}}, \quad (3.3)$$

where  $n_{+u}$  denotes the number of locations mapped as class  $u$ ;  $r_u$  is also referred to as the sensitivity, and its complement,  $1-r_u$ , is referred to as the error of omission, i.e. the proportion of the area with true LFA class  $u$  not mapped as class  $u$ .

When taking the ordinality of the slope, elevation and relief intensity attributes into account, less strict purity measures can be calculated for these attributes. For these measures we not only considered the diagonal entries of the error matrix as correctly classified locations but also the 'one off'-diagonal entries. The overall purity of an LFA is then defined as the areal proportion of the mapped area in which the mapped LFA class equals the true LFA or one class higher or lower than the true class. Less strict purity measures were not estimated for the flatness index (because this has only two classes).

Another quality measure that can be used for ordinal data is the Spearman's rank correlation coefficient or Spearman's rho (Liu et al., 2007). This statistic measures the statistical dependence between pairs of observations (the mapped and true LFAs) after converting the data to ranks. Spearman's rho is computed as:

$$\rho = \frac{1 - \frac{6 \sum_i^N D_i^2}{N(N^2 - 1)} - \frac{1}{2}(T_{map} + T_{true})}{\sqrt{(1 - T_{map})(1 - T_{true})}}, \quad (3.4)$$

in which  $T_*$  is defined as:

$$T_* = \frac{\sum_k t_k(t_k^2 - 1)}{N(N^2 - 1)}, \quad (3.5)$$

where  $D_i$  is the difference between the ranks of the mapped and true LFA class for the  $i$ -th validation observation.  $N$  is the number of validation observations, as before, and  $t_k$  is the number of validation observations with tied rank. The ranks are calculated as follows. First, the validation locations are arranged by mapped LFA class in ascending order. The rank assigned to each class equals the average of the positions of the locations in the ascending order of the classes. For example, suppose that mapped slope class is 1 at 10 validation locations, then the rank of each of these locations is 5.5  $([1+2+\dots+10]/10)$ . If mapped slope class is 2 at 20 locations, then the rank of each of these locations is 20.5  $([11+12+\dots+30]/20)$  because these locations take up positions 11 to 30, when all locations are arranged in ascending order according to slope class. In this way the rank is computed for each mapped LFA class. This is repeated for the true LFA class so that in the end each validation location is assigned two ranks: one based on mapped LFA and one on true LFA. The difference between these two ranks at validation location  $i$  is  $D_i$  in Eq. 4 above. Spearman's rho was computed for the slope, elevation and relief intensity attributes.

Finally we defined the composite purity to indicate how well the combination of four mapped LFA classes characterize the PU. The composite purity is the proportion of the mapped area where  $k$  LFAs

are correctly mapped ( $k = 1, 2, 3$  or  $4$ ). To calculate the composite purity an indicator variable is created for each LFA that equals 1 if the mapped LFA class equals the true LFA class and 0 otherwise. Next the locations are counted for which the sum of the four indicators equals  $k$  and this sum is divided by the total number of validation locations  $N$ .

### 3.2 Application to Western European window

Figure 3.1 shows maps of the four LFAs derived from the PU map at 990-m resolution. In addition, four LFA maps with 90-m resolution were obtained with the same version of the e-SOTER procedure. Both maps were sampled at 184,000 randomly selected locations (one validation location per  $\text{km}^2$ ). Tables 3.2 to 3.5 show the error matrices of the four LFAs. The overall purity, the less-strict overall purity based on the diagonal and one off-diagonal elements of the error matrices and the Spearman's rank correlation coefficient are listed in Table 3.6, while the map unit purities and class representations are shown in Table 3.7.

The LFA most affected by the generalization steps in the e-SOTER procedure is slope. Overall purity of the slope map is 45% and Spearman's rank correlation is 38%. The error matrix shows that slope class 1 as derived from the 90-m SRTM DEM covers 48% of the window. After generalization this is 21% (resulting in a class representation of 32%), whereas the areal proportion of slope class 2 increases from 33% to almost 70% after generalization. This is because large areas with slope class 1 and 3 are incorporated in class 2. This results in a map unit purity of class 2 of only 37%. Slope classes 4 to 7 covered less than 5% of the 90-m slope map. After generalization slope classes 5 to 7 have disappeared. The large effect of the generalization procedure on slope is also evident in Figure 3.2 which shows details (45x63 km) of the four LFA maps before and after generalization. The fact that slope is most affected by generalization can be explained by the fragmented appearance of the slope classes on the 90-m map. Slope classes mainly occur in small 'islands'. When such map is generalized it induces more errors in the PU map — in which map delineations have a minimum size of  $25 \text{ km}^2$  (Dobos et al., 2005) — than a 90-m LFA map with spatially more contiguous classes.

Elevation is the least affected LFA. Areal proportions on the PU map are very close to those on the 90-m map (Table 3.2). Elevation regions are relatively large, contiguous areas that are therefore little affected by generalization (see for example the detail in Figure 3.2). Map unit purities as well as class representations are roughly between 70 and 90%. Spearman's rank correlation is 0.91.

Overall purity of LFA Relief intensity (RI) is 92% (Table 3.6). This large purity is mainly attributed to the 94% purity of RI class 1 (which covers 91% of the area on the 90-m map and 95% on the PU map). The purity of class 2 is 58% and that of class 3 is 45%. Like the slope classes, RI classes 2 and 3 are occur in many small islands at 90-m resolution (Figure 3.1) that are typically not large enough to be retained in the PU map. This explains the small class representations of these two classes (28% and 4%, respectively) and the modest correlation coefficient of 0.43.

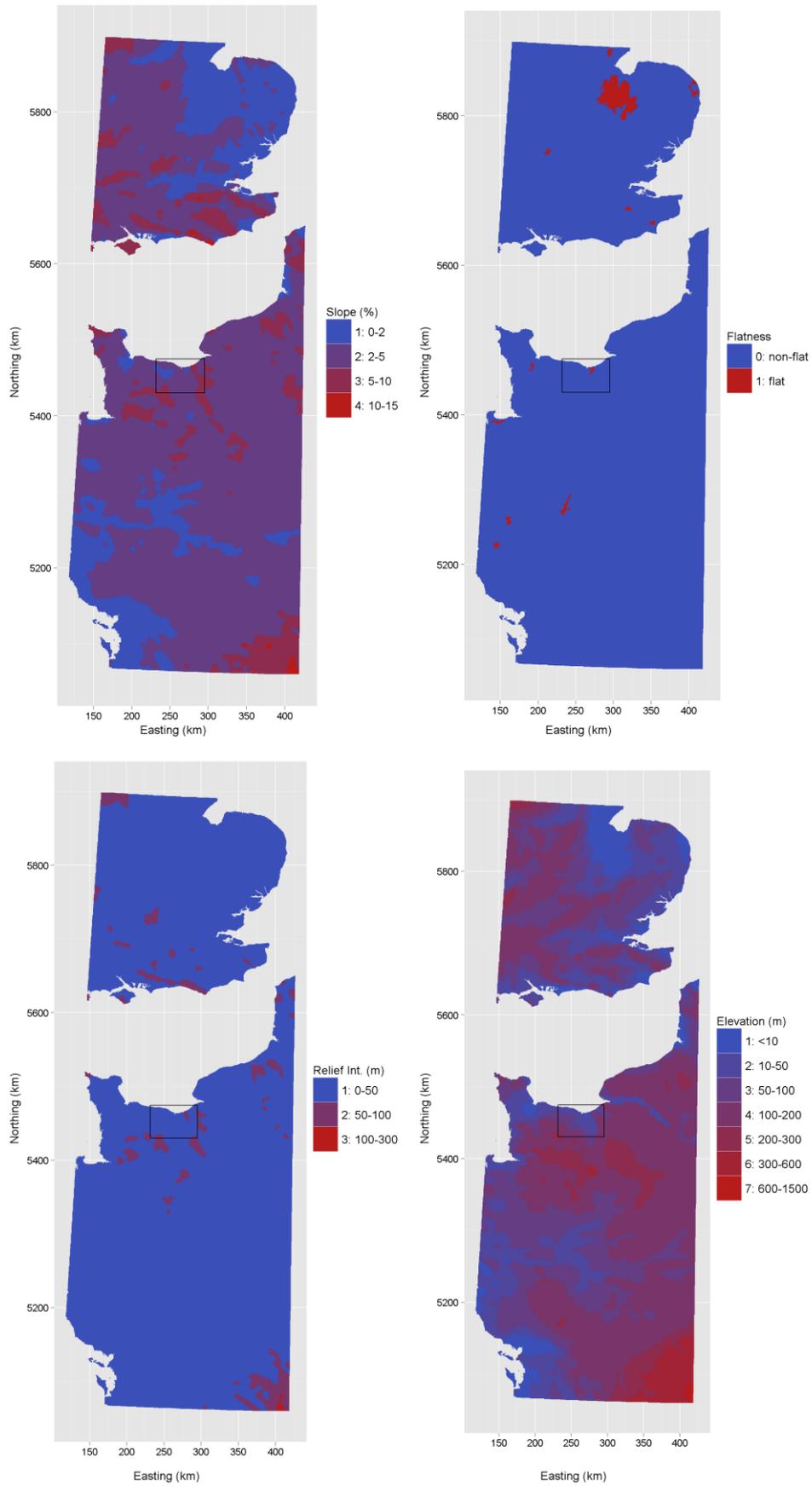


Figure 3.1. Landform attribute maps derived from the e-SOTER Physiographic Units. The rectangle indicates the area from which the map details maps were obtained.

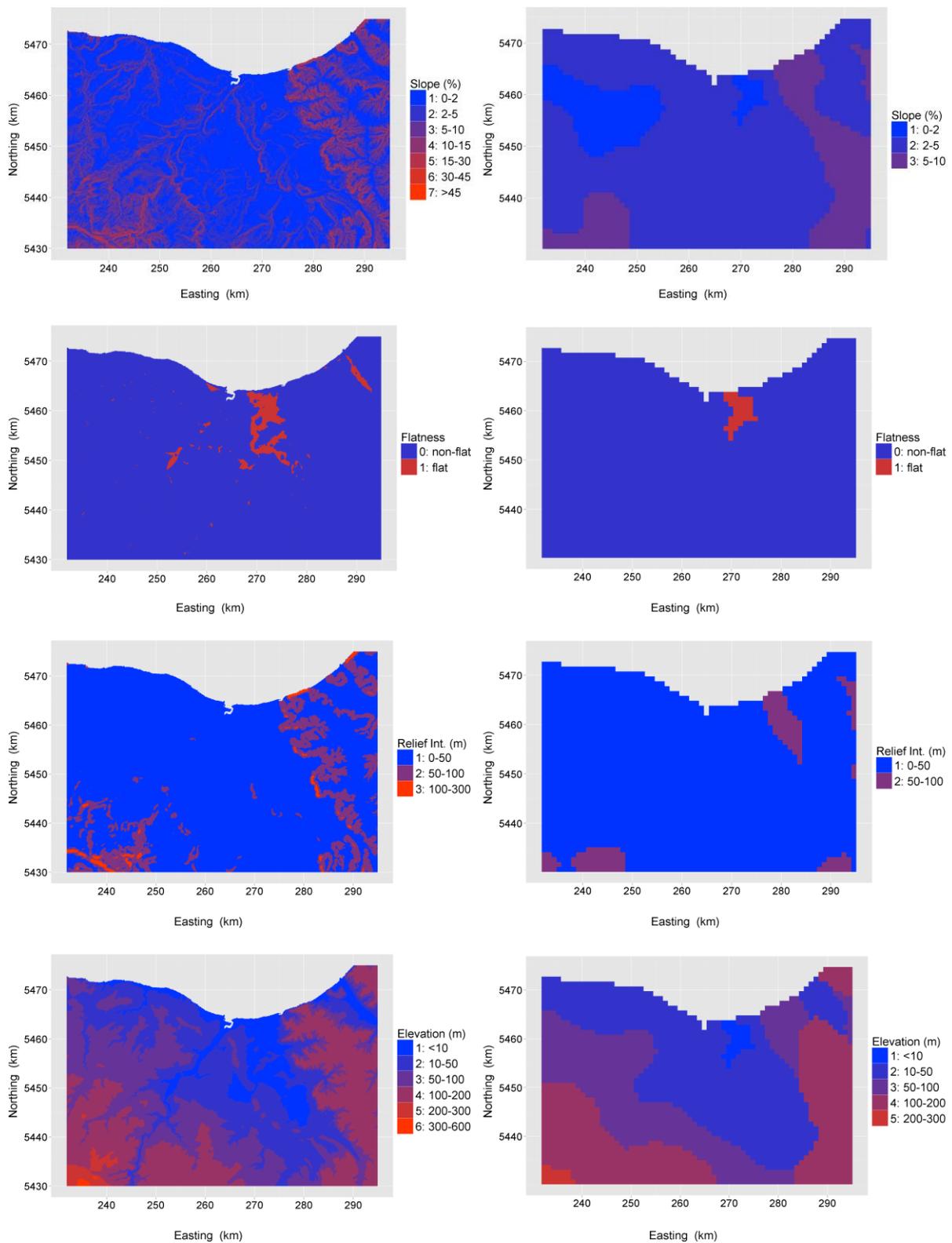


Figure 3.2. Details of the four LFA maps with 90-m resolution (left) as derived from the SRTM DEM and with 990-m resolution (right) after the generalization steps in the e-SOTER procedure.

Table 3.2. Error matrix of LFA 'Elevation'.

Mapped class (990 m)	'Observed' class (90 m)							Total	%
	1	2	3	4	5	6	7		
1	<b>9,320</b>	1,259	43	1	0	0	0	10,623	5.8
2	3,679	<b>25,832</b>	4,505	204	2	0	0	34,222	18.6
3	192	5,373	<b>38,854</b>	6,143	37	0	0	50,599	27.5
4	98	458	7,444	<b>62,487</b>	2,837	9	0	73,333	39.8
5	0	0	4	1,224	<b>8,398</b>	603	0	10,229	5.6
6	0	0	0	0	387	<b>4,250</b>	119	4,756	2.6
7	0	0	0	0	0	28	<b>210</b>	238	0.1
Total	13,289	32,922	50,850	70,059	11,661	4,890	329	184,000	
%	7.2	17.9	27.6	38.1	6.3	2.7	0.2		100

Table 3.3. Error matrix of LFA 'Relief intensity'.

Mapped class (990 m)	'Observed' class (90 m)			Total	%
	1	2	3		
1	<b>164,954</b>	10,960	464	176,378	94.8
2	2,413	<b>4,346</b>	742	7,501	4.1
3	6	61	<b>54</b>	121	0.1
Total	167,373	15,367	1,260	184,000	
%	90.9	8.4	0.7		100

Table 3.4. Error matrix of LFA 'Slope'.

Mapped class (990 m)	'Observed' class (90 m)							Total	%
	1	2	3	4	5	6	7		
1	<b>28,252</b>	8,236	1,451	124	39	0	0	38,102	20.7
2	58,643	<b>47,256</b>	17,775	3,028	1,187	44	5	127,938	69.5
3	2,106	5,236	<b>6,395</b>	2,336	1,326	75	2	17,476	9.5
4	16	68	149	<b>117</b>	122	12	0	484	0.3
5	0	0	0	0	<b>0</b>	0	0	0	0
6	0	0	0	0	0	<b>0</b>	0	0	0
7	0	0	0	0	0	0	<b>0</b>	0	0
Total	89,017	60,796	25,770	5,605	2,674	131	7	184,000	
%	48.4	33.0	14.0	3.0	1.5	0.07	0.004		100

Table 3.5. Error matrix of LFA 'Flatness'.

Mapped class (990 m)	'Observed' class (90 m)		Total	%
	0	1		
0	<b>178,966</b>	2,854	181,820	98.8
1	250	<b>1,930</b>	2,180	1.2
Total	179,216	4,784	184,000	
%	97.4	2.6		100

Table 3.6. Overall purities (%) and Spearman's rank correlation coefficient of the LFAs.

Landform attribute	Western European window		Central European window	
	Overall purity	Spearman's rho	Overall purity	Spearman's rho
Elevation	81.1 / 99.4 <sup>a</sup>	0.91	87.8 / 99.8	0.96
Relief intensity	92.0 / 99.7	0.43	81.1 / 98.5	0.86
Slope	44.6 / 94.8	0.38	50.6 / 86.6	0.81
Flatness	98.3 / -	-	98.1 / -	-

<sup>a</sup> Overall purity / Overall purity based on one off-diagonal entries of the error matrix.

Table 3.7. Map unit purities (%) and class representations (%) of the LFA classes.

LFA class	Western European window		Central European window	
	Map unit purity	Class representation	Map unit purity	Class representation
<i>Elevation</i>				
1	87.7	70.1	-	-
2	75.5	78.5	80.8	32.2
3	76.8	76.4	95.3	94.3
4	85.2	89.2	91.8	93.0
5	82.1	72.0	83.2	80.2
6	89.4	86.9	87.6	89.3
7	88.2	63.8	84.4	84.8
8	-	-	83.4	81.6
9	-	-	-	0
<i>Relief intensity</i>				
1	93.5	98.6	91.3	91.5
2	57.9	28.3	57.7	53.4
3	44.6	4.3	71.4	71.5
4	-	-	73.1	83.9
<i>Slope</i>				
1	74.1	31.7	81.0	68.9
2	36.9	77.7	35.6	36.1
3	36.6	24.8	34.7	48.3
4	24.2	2.1	25.5	14.3
5	-	0	41.0	52.4
6	-	0	30.1	28.7
7	-	0	56.4	56.5
<i>Flatness</i>				
0	98.4	99.9	98.2	99.9
1	88.5	40.3	87.0	26.5

Table 3.8. The composite purity (%) based on the number of LFAs correctly mapped at each validation location.

LFAs correctly mapped	Western European window		Central European window	
	Purity	Cumulative	Purity	Cumulative
4	34.7	34.7	41.2	41.2
3	49.3	84.0	38.8	80.0
2	13.5	97.5	16.7	96.7
1	2.5	100	3.3	100
0	0	100	0	100

Overall purity of the flatness attribute is 98%. However, one should take care with the interpretation of this (large) value. If the entire window would have been mapped as flatness class 0 (i.e. non-flat), then the purity would have been 97.4%. It is therefore better to focus on how well the PU map represents the flat areas. Class representation of the flat areas is a moderate 40%, meaning that 40% of the area of the WE window that was classified as ‘flat’ on basis of the analysis of the 90-m SRTM DEM is classified as ‘flat’ on the PU map. Only the relatively large flat areas, such as for example in the north-central part of the UK area, are depicted on the PU map (Figures 3.1 and 3.2). The majority of the flat areas form small ‘islands’ in the non-flat area and are lost in the generalization procedure.

Validation based on the diagonal plus one off-diagonal entries of the error matrices shows that at 95 to 100% of the validation locations the mapped LFA class is not more than 1 class off the ‘true’ LFA class, that is the LFA class according to the 90-m raster. Finally, Table 3.8 presents the composite purity. From this table it can be observed that for 35% of the window the four LFA classes mapped at 990-m resolution equal the LFA classes according to the 90-m maps and that for 84% of the window at least three of the mapped LFA classes correspond to the classes according to the 90-m maps.

### 3.3 Application to CE window

Figure 3.3 shows maps of the four LFAs for the Central European (CE) window as derived from the PU map. With the same version the four LFA maps with 90-m resolution were obtained. Both maps were sampled at 685,000 randomly selected locations (one validation location per km<sup>2</sup>). Tables 3.9 to 3.12 show the error matrices of the four LFAs. The overall purity, the less-strict overall purity based on the diagonal and one off-diagonal elements of the error matrices and the Spearman’s rank correlation coefficient are listed in Table 3.6, while the map unit purities and class representations are shown in Table 3.7.

Validation results for the CE window are comparable to those of the WE window. The overall purities for LFAs elevation and slope are somewhat larger than those of the WE window, whereas the overall purity of RI is somewhat smaller. Also in the CE window, slope has the smallest purity of the four LFAs for reasons explained in previous sections. A detail of the 90-m slope map (Figure 3.4) clearly shows the fragmented occurrence of the slope classes.

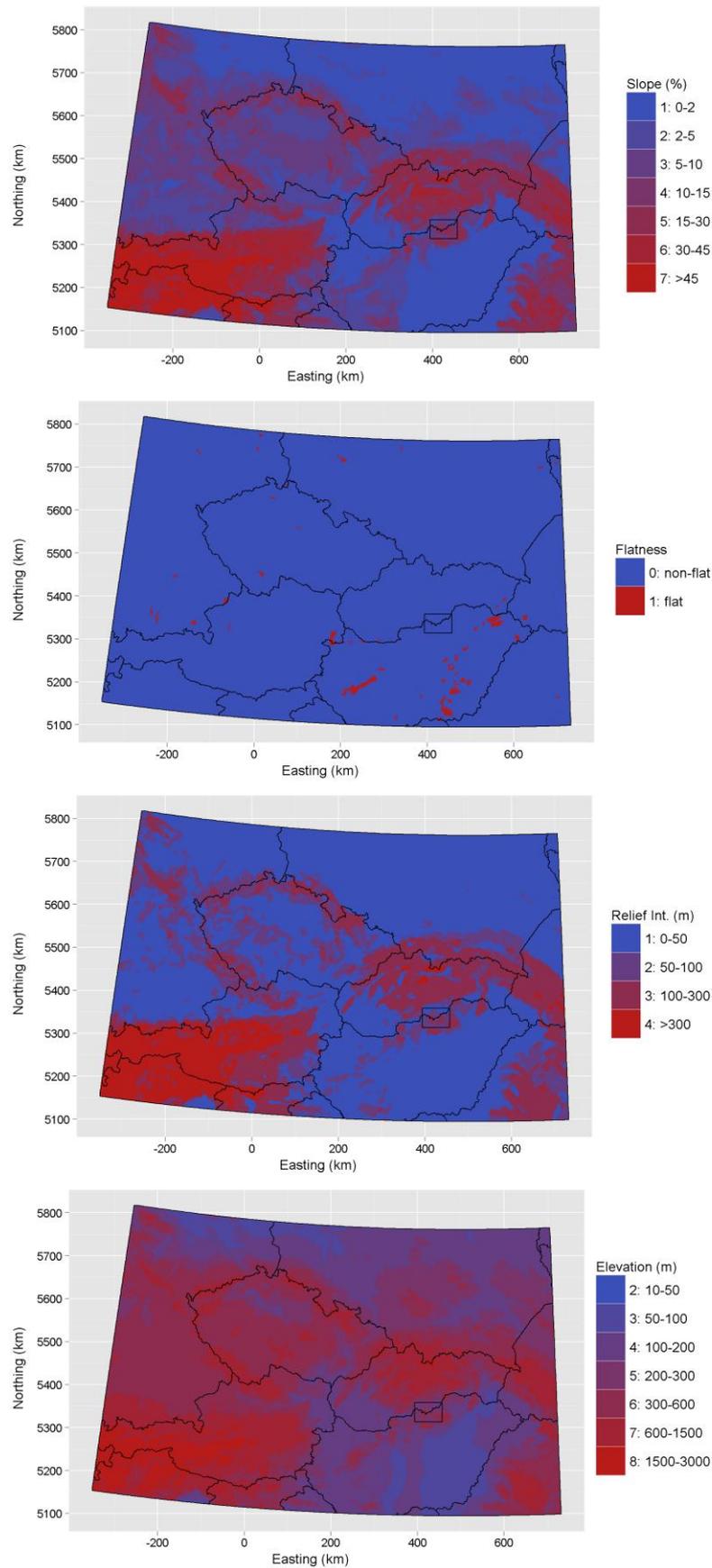


Figure 3.3. Landform attribute maps derived from the e-SOTER Physiographic Units. The rectangle indicates the area from which the map details were obtained.

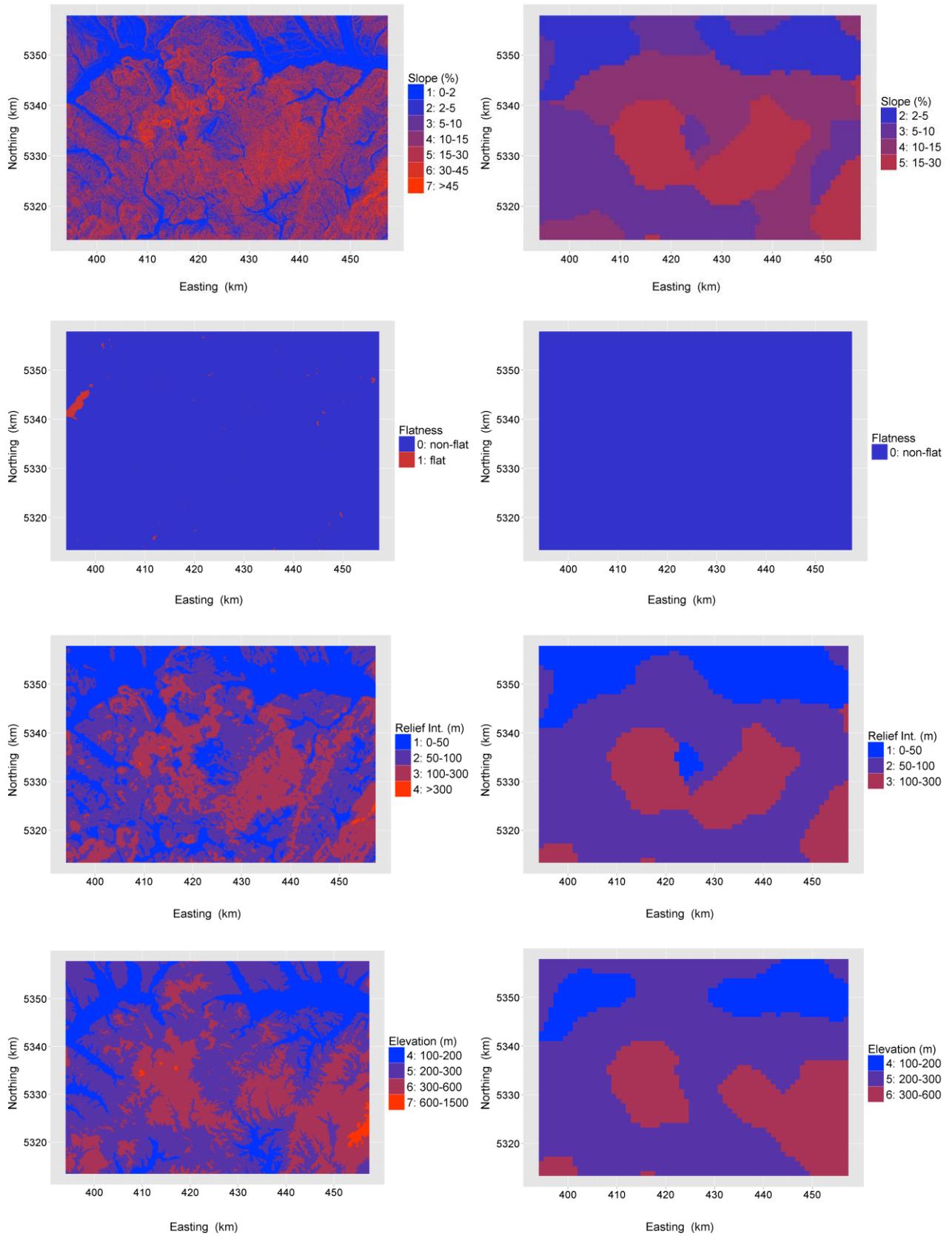


Figure 3.4. Details of the four LFA maps with 90-m resolution (left) as derived from the SRTM DEM and with 990-m resolution (right) after the generalization steps in the e-SOTER procedure.

Table 3.9. Error matrix of LFA 'Elevation'.

Mapped class (990 m)	'Observed' class (90 m)								Total	%
	2	3	4	5	6	7	8	9		
2	<b>101</b>	24	0	0	0	0	0	0	125	0.02
3	196	<b>48,573</b>	2,201	10	5	0	0	0	50,985	7.4
4	13	2,890	<b>157,300</b>	10,708	344	5	0	0	171,260	25.0
5	0	3	9,111	<b>96,274</b>	10,265	12	0	0	115,665	16.9
6	0	3	532	12,817	<b>180,756</b>	12,339	0	0	206,447	30.1
7	4	2	66	267	11,145	<b>93,173</b>	5,685	0	110,342	16.1
8	0	0	0	0	1	4,400	<b>25,171</b>	604	30,176	4.4
9	0	0	0	0	0	0	0	<b>0</b>	0	0.0
Total	314	51,495	169,210	120,076	202,516	109,929	30,856	604	685,000	
%	0.05	7.5	24.7	17.5	29.6	16.0	4.5	0.1		100

Table 3.10. Error matrix of LFA 'Relief intensity'.

Mapped class (990 m)	'Observed' class (90 m)				Total	%
	1	2	3	4		
1	<b>367,863</b>	31,102	3,884	41	402,890	58.8
2	28,204	<b>60,667</b>	16,146	74	105,091	15.3
3	5,031	21,090	<b>85,265</b>	7,994	119,380	17.4
4	726	816	13,962	<b>42,135</b>	57,639	8.4
Total	401,824	113,675	119,257	50,244	685,000	
%	58.7	16.6	17.4	7.3		100

Table 3.11. Error matrix of LFA 'Slope'.

Mapped class (990 m)	'Observed' class (90 m)							Total	%
	1	2	3	4	5	6	7		
1	<b>171,631</b>	31,715	6,421	1,342	760	74	10	211,953	30.9
2	54,450	<b>46,042</b>	21,944	4,593	2,090	151	31	129,301	18.9
3	17,480	39,270	<b>49,142</b>	22,017	12,573	876	98	141,456	20.7
4	1,834	3,705	8,508	<b>7,869</b>	8,096	761	111	30,884	4.5
5	2,692	5,616	13,624	16,550	<b>40,327</b>	15,341	4,233	98,383	14.4
6	778	811	1,309	1,646	7,677	<b>10,458</b>	12,118	34,797	5.1
7	283	372	757	1,115	5,402	8,740	<b>21,557</b>	38,226	5.6
Total	249,148	127,531	101,705	55,132	76,925	36,401	38,158	685,000	
%	36.4	18.6	14.8	8.0	11.2	5.3	5.6		100

Table 3.12. Error matrix of LFA 'Flatness'.

Mapped class (990 m)	'Observed' class (90 m)		Total	%
	0	1		
0	<b>667,950</b>	12,046	679,996	99.3
1	651	<b>4,353</b>	5004	0.7
Total	668,601	16,399	685,000	
%	97.6	2.4		100

Elevation has the largest overall purity (88%). Like for WE, the areal proportions of the PU map are very close to those of the 90-m map (Table 3.9). Map unit purities and class representations vary between 80 and 95% (Table 3.7), except for the representation of class 2 (32%). However this class only covers 0.05% (or 340 km<sup>2</sup>) of the 90-m map. Like for WE, the elevation classes form relatively large, contiguous regions at 90-m resolution (Figure 3.3) which reduces the effect of the generalization steps on the outcome classes.

The overall purity of the RI attribute is 81% with map unit purities ranging from 58 to 91%. Class representation varies between 53 and 92%. The areal proportion of the RI classes on the PU map are close to those on the 90-m map (Table 3.8). The Spearman's rank correlation coefficient is 0.86. These results are in sharp contrast with the results for the WE window. An explanation for the difference in the effect of the generalization steps on the final PU map for the RI attribute between the two windows is again the nature of spatial variation of the RI at 90-m resolution. In the WE window the RI classes occur fragmented in relatively small islands, whereas in the CE window the RI classes form large, contiguous areas (Figure 3.4). Generalization of the RI classes therefore induces fewer errors in the PU map in the CE window than in the WE window.

The flatness attribute has an overall purity of 91%. However, more relevant is the representation of the 'flat' class, which is only 26%. A large part of the area classified as 'flat' at 90-m resolution is lost during the generalization steps. Representation of the 'flat' class in the CE window is much smaller than in the WE window (Table 3.7). The CE window has more pronounced relief than the WE window (compare Figure 3.1 with Figure 3.3). In areas with pronounced relief, flat areas are likely to be smaller than in areas with less pronounced relief. (Note, however, that the fraction of flat areas is almost the same for the two windows; 2.4% for CE and 2.6% for WE). Smaller flat areas are more affected by the generalization procedure than larger areas and are thus more likely to be lost on the final PU map (see for example Figure 3.4).

Validation based on the diagonal plus one off-diagonal entries of the error matrices shows similar results as for the WE window. Between 87 and 100% of the validation locations the mapped LFA class is not more than one class off the 'true' LFA class. Finally, Table 3.8 presents the composite purity which is also comparable to that of the WE window. From this table it can be observed that for 41% of the window the four LFA classes mapped at 990-m resolution equal the LFA classes according to the 90-m maps and that for 80% of the window at least three of the mapped LFA classes correspond to the classes according to the 90-m maps.

## 4. Soil Validation

### 4.1 Methodology

#### 4.1.1 Introduction

Validation means putting a product to the test. In this case the product is the soil component of the e-SOTER terrain units. The mapped soil component is compared with field data, which are assumed to have negligible error compared to errors contained in the map. From this comparison various map quality measures can be calculated. Preferably, validation is done with field data obtained by probability sampling (Brus et al., 2011). The advantage of using such data for validation is that unbiased estimates of the map quality measures and their standard errors can be obtained. The standard errors quantify the uncertainty associated to the quality measures. This uncertainty arises from the fact that the quality measures are estimated from a limited dataset.

Validation of the soil component of the terrain units is done for the UK part of the western European window and the German/Czech Republic part of the central European window. Unfortunately validation data collected by probability sampling were not available for these areas and there were insufficient resources to collect such data. Independent legacy soil data were therefore used for validation. Independent means that the data were not used by WP2 to define the soil components. It should be noted, however, that the validation data in both windows were obtained by purposive sampling instead of probability sampling. This has consequences when one wants to estimate map quality measures as we will explain below. The advantage of using legacy data is that these were abundantly available and hence a large sample of validation observations was available for both areas.

The validation data consisted of described and classified soil profiles at geo-referenced point locations. The soils were classified according to local (British, German and Czech) classification systems. Before these data could be used for validation, the local soil classification systems had to be correlated to the WRB soil classification system so that the observed soil types could be reclassified to Reference Soil Groups (RSGs), including the key qualifiers. Once soil correlation and reclassification were completed, the observed RSGs were compared to the mapped soil groups of the soil components of the terrain units.

Because the validation data were collected by purposive sampling, spatial dependence between the data points should be taken into account when estimating map quality measures from these data. This can be done by assuming a model for the spatial variation of the error (the difference between mapped soil attribute and observed soil attribute) and using this model when estimating the quality measures. While this is relatively easy when the variable of interest is continuous (e.g. a soil property such as organic matter content or the available water capacity), it is much more difficult for categorical variables such as soil group. For this reason we treated the validation observations as spatially independent (i.e. as if collected by probability sampling), which allowed us to use

straightforward (design-based) estimation methods for map quality measures (Brus et al., 2011). We refrained however from estimating the standard errors of the quality measures because these cannot be properly estimated from data obtained by purposive sampling methods.

#### 4.1.2 Soil data to be validated

The e-SOTER soil database dated 20 June 2011 was validated. This database distinguishes seventeen WRB soil groups or combinations of groups. These are:

- Arenosol (AR)
- Calcisol (CA)
- Cambisol (CM)
- Chernozem/Kastanozem (CK)
- Histosols (HS)
- Hydromorphic soils (Gleysol/Fluvisol/Stagnosol) (HY)
- Leptosol/Regosol (LR)
- Lixisols (LX)
- Luvisol (LV)
- Luvisols/Alisols (LA)
- Nudilithic soils (ND)
- Phaeozem (PH)
- Podzol (PZ)
- Regosol (RG)
- Salt-affected soils (Solonetz/Solonchak) (SA)
- Umbrisol (UB)
- Vertisol (VT)

The soil component of an e-SOTER terrain unit is made up of one or more (up to twelve) soil groups. For each soil group in a soil component the proportion of the terrain unit that this soil group covers is given. There is no information on where the individual groups are found within the terrain unit.

The validation soil groups should follow the same legend as soil groups that comprise the soil components. This means that after correlation and reclassification of the validation soil data, some of the soil groups must be grouped into one composite group (e.g. Chernozems and Kastanozems).

Validation of the soil components is done in two modes:

1. *The stringent mode.* Correct prediction of the soil component only when the validation soil group equals the dominant soil group in the soil component. The dominant group is the group with the largest terrain unit proportion and is typically used for interpretation of the e-SOTER map for environmental applications (e.g. for the assessment of soil threats by WP5).
2. *The flexible mode.* Correct prediction of the soil component when the validation soil group equals any of the soil groups in the component.

WRB qualifiers of the soil components were not validated because at the time of validation (June 2011) the qualifier database was not yet finalized and spatially referenced, i.e. the qualifiers were not linked to the terrain unit map.

The overall purity is a very strict accuracy measure in which all errors are treated equally. This means that confusion between, for example, Gleysol and Chernozem is 'as wrong' as confusion between Chernozem and Phaeozem although Gleysols and Chernozems are taxonomically farther apart than Chernozems and Phaeozems. For the soil map user confusion between Chernozem and Phaeozem might be less severe than confusion between Chernozem and Gleysol (of course this depends on the application of the soil map). Validation was therefore also done for grouped RSGs (i.e. a generalized legend). The grouping is based on presence of key diagnostic horizons or properties. The generalized legend has twelve entries:

- Arenosol (AR)
- Argic soils (Luvisol/Alisol/Lixisol) (AG)
- Calcisol (CA)
- Cambisol (CM)
- Dark soils (Chernozem, Kastanozem, Phaeozem) (DK)
- Histosols (HS)
- Hydromorphic soils (Gleysol/Fluvisol/Stagnosol) (HY)
- Shallow soils (Leptosol/Regosol/Nudilithic) (SH)
- Umbrisols (UB)
- Podzol (PZ)
- Salt-affected soils (Solonetz/Solonchak) (SA)
- Vertisol (VT)

Note that for each soil component in the e-SOTER soil database the percentage coverage of the associated SOTER unit is recalculated after generalizing the legend. The dominant soil group in the SOTER unit is then derived from these.

#### **4.1.3 Estimation of map quality measures**

For the stringent mode three map quality measures for the soil components were estimated from the validation data: overall purity, map unit purity (user's accuracy) and group representation (producer's accuracy) (Stehman, 1997; Brus et al., 2011). Like for landform validation, these measures are estimated from a sample error matrix (Table 3.1) by Equations. 3.1, 3.2 and 3.3. The entries of the error matrix are the number of observations for each combination of mapped and true (observed) soil group. For the flexible mode only the overall purity was estimated from the validation data.

It should be noted that the three quality measures are global measures. This means that these give quality estimates for the map as a whole. They do not provide local information on accuracy, i.e. where predictions are correct and where they are wrong.

## 4.2 Application to UK part of the Western European window

### 4.2.1 Soil data

The e-SOTER soil database of the UK part of the Western European window was validated with the National Soil Inventory (NSI) data, which is a set of point data collected around 1980 on a regular grid with a 5-km spacing between the grid nodes. A total of 2,354 NSI sampling locations are located in the UK area. Locations in the sea (38) or in the major urban centres (131) were discarded as these lacked a profile description and classification. In addition, nine sampling locations fell outside the extent of the e-SOTER map and were discarded as well. This left 2,176 data points that make up the validation dataset. A site description is available for each sampling location and includes 18 site properties, among them are soil subgroup and soil series name according to the Soil Survey of England and Wales (SSEW) classification scheme. Profile descriptions include 24 soil properties (Appendix 1).

In addition, the NSI topsoil dataset was available that contains information on 35 soil properties (Appendix 1) — including pH and soil organic carbon (SOC) — for 2,138 data points. Fifty profile lacked SOC data and 73 profiles pH data. For each of these profiles the average SOC and pH values of the soil subgroup that the profile belongs to were used as estimates.

To construct the e-SOTER soil map the polygon map with the e-SOTER units obtained with Version 6 (28 April 2011) of the e-SOTER procedure was used. The e-SOTER units are linked to the soil database by a common identifier.

### 4.2.2 Soil correlation

The validation datasets did not contain full profile descriptions and horizon analytical data (such as those used to populate the WP2 Representative profile database), thus a comprehensive classification to WRB reference soil groups and qualifiers was not possible. However, detailed correlation was not required as the mapped product contains information only on the RSG, which is sufficient detail for the scale of the e-SOTER product at 1:1M. WRB diagnostic horizons, features and properties used to determine the RSG can be complex and require detailed soil information. Thus a pragmatic approach is necessary to correlate the SSEW classification to the WRB to ensure that the primary WRB criteria are met but absence of some data does not limit the ability to correlate. The simplified criteria for WRB diagnostics derived in WP2 (Annex II, Annex III Deliverable D5) were used to aid this process. The resulting conceptual approach for correlation used a combination of direct correlation using criteria inherent in the SSEW classification scheme, classifier rules that require numerical data thresholds and expert judgment. The statistical software package R (R Development Core Team, 2008) was used to write scripts to assign hard-wired correlations in the case of direct correlations from the definitions and classifier rules where numerical or class requirements were needed.

An example of classifier rules based on numerical data (with some pragmatic decisions and expert judgement to deal with missing data) are those used to determine the WRB qualifiers 'Mollic' and 'Umbric'. These qualifiers were determined based on colour, pH (as a proxy for base saturation (BS), which was not included in the horizon analytical data) and organic carbon (OC), and topsoil thickness data. First value and chroma were extracted from the colour coding and the average value and chroma for the 0-20-cm soil layer was computed. Next for each profile the colour (value and chroma <3.5) and thickness (topsoil >10 cm) requirements were evaluated. The OC content (0.6-20 %) and pH (where it was assumed that a pH >6.5 correlates to a BS >50% and pH <6.5 to a BS <50%) requirements were evaluated. Soil layers that met the colour, thickness and OC requirement and with pH <6.5 were classified as 'Umbric'. Those with pH >6.5 as 'Mollic'. Soil layers that lacked colour information but with an OC content between 3 and 20% were also classified as 'Mollic' or 'Umbric' (depending on pH) as it was assumed that these horizons were dark enough to fulfil the colour requirement.

The SSEW classification is a hierarchical system comprising the following classes from the least to the most detail in terms of soil description (Avery, 1980): Major Soil Group (e.g. 5. Brown Soils), Soil Group (e.g. 5.4 Brown Earths), Soil Sub-Group (e.g. 5.42 Stagnogleyic Brown Earth), Soil Series (e.g. Papworth). The classes are differentiated primarily on the composition of soil material within specified depths and the presence or absence of diagnostic horizons or features. Where these features have similar classification rules to WRB a direct correlation is achieved from the soil class (soil group or sub-group) to a WRB RSG. For example the classification of the sub-group '5.4 Brown Earth' requires a weathered B-horizon in the SSEW system and this definition correlates with the Cambic horizon in WRB, thus correlating to a Cambisol RSG. In some cases the definitions did not correlate completely or there was some ambiguity in the SSEW definitions, thus additional criteria were necessary to assign the RSG. For example '3.4 Rendzinas' have a depth definition of calcareous bedrock within 30 cm and this encompasses the requirements for Leptosols and Cambisols in WRB, the former having shallower depth criteria. Therefore additional information on soil depth was used to differentiate between Cambisol and Leptosol.

The correlation resulted following RSGs identified in the validation dataset: Anthrosols, Arenosols, Cambisols, Leptosols, Regosols, Histosols, Luvisols, Podzols, Fluvisols, Gleysols and Stagnosols. The e-SOTER legend comprises some simplification of the RSGs as a result of the methodologies in WP2 (Deliverable D5) used to create the spatial dataset for the soil component of the e-SOTER units. This results in Fluvisols, Gleysols and Stagnosols grouped as Hydromorphic soils. Anthrosols are not distinguished in the legend used to classify the soils of the soil component. In the generalized legend Anthrosols are included in the Dark Soil group. Table 4.1 summarizes the soil correlation from SSEW classification to WRB, including the method used to determine the WRB qualifiers.

Because soil information contained in the profile descriptions was often not detailed enough to determine the RSG with negligible error, we relaxed the validation criteria for several RSGs. Validation locations where a Gleyic/Stagnic Luvisol or Gleyic/Stagnic Cambisol was observed but where a Hydromorphic soil was mapped were considered correctly mapped. The same holds for Luvic Hydromorphic soils mapped as Luvisols or as the Luvisol/Alisol association.

Table 4.1. Soil correlation from SSEW classification to WRB.

SSEW Soil Group <sup>a</sup>	SSEW Soil Group descriptor	RSG direct correlation <sup>b</sup>	RSG other criteria	Qualifiers by sub-group or soil series definition <sup>c</sup>	Qualifiers by numerical thresholds
1.1	Raw sands	Arenosol			Histic, Mollic
1.2	Raw alluvial soils	Fluvisol			Umbric, Mollic, Arenic, Clayic
1.3	Raw Skeletal soils	Regosol			Arenic
1.5	Man-made raw soils	Regosol			Arenic
2.2	Unripened gley soils	Fluvisol		Gleyic	Umbric, Mollic, Arenic, Clayic
3.1	Rankers	Cambisol or Leptosol	Depth and Stoniness	Leptic, Lithic, Skeletic	Umbric, Mollic, Vertic, Clayic
3.3	Ranker-like alluvial soils	Fluvisols		Skeletic, Calcaric, Gleyic	Umbric, Mollic, Arenic, Clayic
3.4	Rendzinas	Cambisol or Leptosol	Depth and Stoniness	Rendzic, Calcaric, Skeletic, Gleyic, Stagnic, Leptic, Lithic	Umbric, Mollic, Vertic, Clayic
3.5	Pararendzinas	Cambisol or Leptosol	Depth and Stoniness	Rendzic, Calcaric, Skeletic, Gleyic, Stagnic	Umbric, Mollic, Vertic, Clayic
3.7	Rendzina-like alluvial soils	Fluvisols		Skeletic, Calcaric, Gleyic	Umbric, Arenic
4.1	Calcareous Pelosols	Cambisol		Calcaric, stagnic	Histic, Mollic, Vertic, Clayic
4.2	Non-calcareous Pelosols	Cambisol		Stagnic	Histic, Mollic, Vertic, Clayic
4.3	Argillic Pelosols	Luvisol		Stagnic	Mollic, Vertic
5.1	Brown Calcareous Earths	Cambisol		Calcaric, Gleyic, Stagnic, Leptic, Skeletic	Histic, Mollic, Vertic, Clayic
5.2	Brown Calcareous Sands	Cambisol or Arenosol	Particle size	Calcaric	Histic, Umbric, Mollic, Vertic, Clayic
5.3	Brown Calcareous Alluvial soils	Cambisol or Fluvisol	Expert judgement	Calcaric, Gleyic	Histic, Umbric, Mollic, Vertic, Arenic, Clayic
5.4	Brown Earths	Cambisol		Leptic, Stagnic	Histic, Mollic, Vertic, Clayic
5.5	Brown Sands	Cambisol or Arenosol	Particle size	Gleyic	Histic, Mollic, Vertic, Clayic
5.6	Brown Alluvial soils	Cambisol		Stagnic, Gleyic	Histic, Mollic, Vertic, Clayic
5.7	Argillic Brown Earths	Luvisol		Stagnic, Gleyic	Mollic, Vertic, , Arenic
5.8	Paleo-argillic Brown Earths	Luvisol		Stagnic, Gleyic	Mollic, Vertic, , Arenic
6.1	Brown Podzolic	Cambisol			Histic, Mollic, Vertic,

	soils			Clayic
6.3	Podzols	Podzol	Skeletal, Albic, Gleyic, Stagnic	Histic, Umbric
7.1	Stagnogley soils	Stagnosol	Luvic	Umbric, Mollic, Vertic, Arenic, Clayic
7.2	Stagnohumic gley soils	Stagnosol		Umbric, Mollic, Vertic, Arenic, Clayic
8.1	Alluvial gley soils	Fluvisol		Histic, Umbric, Mollic, Arenic, Clayic
8.2	Sandy gley soils	Gleysol		Histic, Umbric, Mollic, Arenic, Clayic
8.3	Cambic gley soils	Gleysol		Histic, Umbric, Mollic, Arenic, Clayic
8.4	Argillic gley soils	Gleysol	Luvic	Histic, Umbric, Mollic, Arenic, Clayic
8.5	Humic-alluvial gley soils	Fluvisol		Histic, Umbric, Mollic, Arenic, Clayic
8.6	Humic-sandy gley soils	Gleysol		Histic, Umbric, Mollic, Arenic, Clayic
8.7	Humic gley soils	Gleysol	Luvic	Histic, Umbric, Mollic, Arenic, Clayic
9.1	Man-made humus soils	Anthrosol		Arenic, Clayic
9.2	Disturbed soils	Regosol		Arenic
10.1	Raw Peat soils	Histosol		
10.2	Earthy peat soils	Histosol		

<sup>a</sup> For Soil Groups in validation dataset only.

<sup>b</sup> Soil group definition meets criteria for WRB definition for RSG.

<sup>c</sup> Soil sub-group or soil series definition meets criteria for WRB qualifier.

#### 4.2.3 Validation of the soil component of the e-SOTER units

Figure 4.1 (bottom) shows the e-SOTER soil map. Here the dominant soil group in the soil component associated to an individual e-SOTER unit is used for mapping.

##### *Stringent mode, full legend*

Table 4.2 presents the sample error matrix of mapped soil group, which corresponds to the dominant soil group of the e-SOTER units that contain the validation locations (Figure 4.2, bottom), against the observed soil group (Figure 4.2, top). The overall purity is 51.0%, which means that for 51% of the mapped area the dominant soil group correctly predicts the actual soil group. By contrast, the overall purity of the 1:250,000 soil map with WRB legend (Figure 4.1, top) is 68.0%, as estimated from the validation sample. Nevertheless, a purity between 50-60% can be considered an adequate result for a map constructed with digital soil mapping (Kempen et al., 2009; Kempen et al.,

2012), especially given the 1:1M scale of the e-SOTER soil map (given a legend, the map units of a soil map generally become less pure when map scale decreases) and the fact that for each SOTER unit the dominant soil group is used to predict the soil at any location within that unit. The expected (or theoretical) purity (Brus et al., 2008; Kempen et al., 2009) of the e-SOTER soil map is computed by averaging the areal proportions of the dominant soil groups of the e-SOTER units and equals 70.1%. The 19% difference between theoretical and overall purity might indicate that soil spatial variation within the terrain units is somewhat larger than indicated by the soil database.

Figure 4.3 (top) shows the validation locations with correct and incorrect classification. A major source of mapping error is the area mapped as 'Histosol' in the north-central part of the validation area. The map purity of this map unit is only 8% (Table 4.3). At 116 validation locations in this map unit (77%) a hydromorphic soil is observed, indicating wet conditions but absence of peat. The 'Podzol' map unit, found in the south-central part of the validation area, is another important error source. Podzols seem over-represented on the e-SOTER soil map. On this map the Podzol map unit covers 2949 km<sup>2</sup>, whereas on the England-Wales map (Figure 4.1, top) these soils cover 1269 km<sup>2</sup>. Furthermore, at only 14 out of 118 validation locations in the Podzol map unit a Podzol is observed, resulting in a map unit purity of only 11.9%. Luvisols (map unit purity around 40%) are mainly confused with Cambisols and Cambisols (map unit purity 50%) with Luvisols. The 'Hydromorphic' map unit is the most pure (71%). This soil also has the largest class representation (68%), followed by Luvisols, Histosols and Cambisols.

Leptosols are never the dominant soil group in the e-SOTER soil map. This soil group therefore has a class representation of 0%. However, these soils cover 6,150 km<sup>2</sup> (or 10% of the area) on the soil map of England-Wales and are observed at 87 validation locations, mainly in the south-eastern part of the area (Figure 4.1, top). Arenosols are observed at 73 validation locations but are mainly mapped as Luvisols / Alisols. Anthrosols are not distinguished in the e-SOTER soil legend and therefore have a class representation of 0%. Phaezems are mapped as the dominant soil in a tiny part (70 km<sup>2</sup>) of the validation area. These soils, however, do not occur in the UK and are therefore not observed in the validation dataset. Occurrence of these soils on the e-SOTER soil map is an artefact of the mapping method. Predictive relations (mapping rules) between soils and remote sensing imagery derived for the Central European window (where Phaeozems do occur) are extrapolated to the UK part of the Western European window causing occurrence of Phaeozems. The observed Regosols correlate to soil group 9.2 (Disturbed soils) in the SSEW classification scheme and are typical for areas where for example land has been restored after open-cast mining. The e-SOTER method does not aim to map such soils.

#### *Stringent mode, generalized legend*

Table 4.4 shows the error matrix of mapped against observed soil group for the generalized legend. The overall purity is 50.8%, nearly equal to the overall purity of the map with the full legend. Also the map unit purities and class representations (Table 4.5) are roughly similar to those of the full legend. Grouping the WRB soil groups into twelve classes apparently did not affect the accuracy for the UK validation area.

### Flexible mode

For the flexible mode the overall purity is 89.8% for validation of the full legend. This means that for almost 90% of the validation area the actual soil group is included in the soil component of the SOTER units. Figure 4.3 (bottom) shows the validation locations with correct and incorrect classification. Locations where the actual soil group is not part of the SOTER unit soil component are mainly located in the south-west parts of the UK validation area. In these parts Leptosols are observed but are not part of the soil components associated to the SOTER units in that area, while in the northeast parts this is true for the observed Arenosols and Podzols.

From the stringent mode we saw that for 51% of the validation locations the observed soil group equals the dominant soil group in the SOTER unit. Table 4.6 lists the soil component number of the SOTER unit of which the associated soil group equals the observed soil group. From this table it can be observed that for 77% of the validation area the actual soil type equals the first, second or third dominant soil group of a e-SOTER unit. For the generalized legend the overall purity for the flexible mode is 91.6%.

Table 4.2. Error matrix of mapped (dominant) soil group against observed soil group for the full legend. Bold type indicate the number of validation locations for which the mapped soil group corresponds to the observed soil group.

Mapped RSG	Observed RSG									Total
	AT	AR	CM	HS	HY	LP	LV	PZ	RG	
AR	0	<b>0</b>	1	0	0	1	0	0	0	2
CM	1	8	<b>207</b>	0	69	16	102	0	12	415
HS	0	0	17	<b>12</b>	116	0	2	2	1	150
HY	5	13	104	1	<b>640</b>	17	62	17	37	896
LV	0	1	70	3	13	45	<b>79</b>	1	1	213
LA	9	45	80	7	58	2	<b>157</b>	7	3	368
PH	0	0	1	0	1	0	0	0	0	2
PZ	0	6	19	0	39	6	33	<b>14</b>	1	118
RG	0	0	6	0	5	0	1	0	<b>0</b>	12
Total	15	73	505	23	941	87	436	41	55	2176

Table 4.3. Estimated map unit purities of and class representations based on an e-SOTER soil map with the full legend that depicts the dominant soil group of the soil components.

Soil group	Map unit purity (%)	Class representation (%)
Anthrosol	- <sup>a</sup>	0
Arenosol	0	0
Cambisol	49.9	41.0
Histosol	8.0	52.2
Hydromorphic	71.4	68.0
Leptosol	- <sup>b</sup>	0
Luvisol	37.1	54.1
Luvisol/Alisol	42.7	- <sup>c</sup>
Phaeozem	0	- <sup>c</sup>
Podzol	11.9	34.1
Regosol	0	0

<sup>a</sup> Not distinguished in the e-SOTER soil component database.

<sup>b</sup> Not mapped as dominant soil group in UK validation area.

<sup>c</sup> Not observed in the validation dataset.

Table 4.4. Error matrix of mapped (dominant) soil group against observed soil group for the generalized legend.

Mapped RSG	Observed RSG								Total
	AG	AR	CM	DK	HS	HY	PZ	SH	
AG	<b>239</b>	46	152	9	10	74	8	51	589
AR	0	<b>0</b>	1	0	0	0	0	1	2
CM	101	8	<b>205</b>	1	0	67	0	28	410
DK	1	0	1	<b>0</b>	0	5	0	1	8
HS	2	0	17	0	<b>12</b>	116	2	1	150
HY	62	13	104	5	1	<b>635</b>	17	53	890
PZ	33	6	19	0	0	36	<b>14</b>	7	115
SH	1	0	6	0	0	5	0	<b>0</b>	12
Total	439	73	505	15	23	938	41	142	2176

Table 4.5. Estimated map unit purities of and class representations based on a SOTER soil map with the generalized legend that depicts the dominant soil group of the soil components.

Soil group	Map unit purity (%)	Class representation (%)
AG	40.6	54.4
AR	0	0
CM	50.0	40.6
DK	0	0
HS	8.0	52.2
HY	71.3	67.7
PZ	12.2	34.1
SH	0	0

Table 4.6. Soil component number for which the soil group equals the observed soil group. Here 1 is the dominant (largest areal coverage of the SOTER unit) soil group, 2 is the second dominant soil group, etc. *n* is the number of validation locations, % is the percentage of the total number of validation location and 'Cumulative' is the cumulative percentage.

Soil Component	<i>n</i>	%	Cumulative
1	1109	51.0	51.0
2	315	14.5	65.4
3	248	11.4	76.8
4	150	6.9	83.7
5	85	3.9	87.6
6	32	1.4	89.1
7	8	0.4	89.5
8	6	0.3	89.8

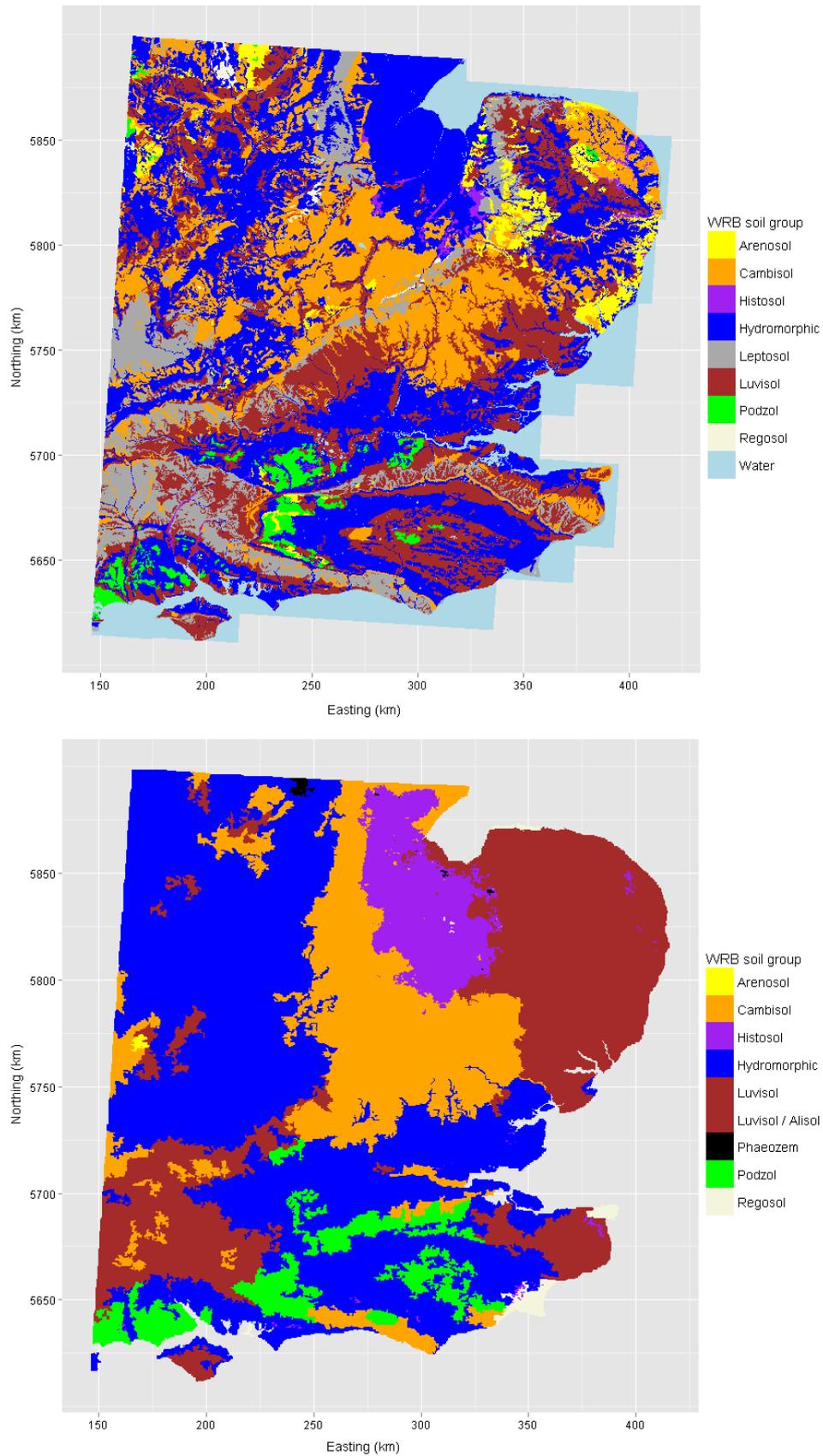


Figure 4.1. The 1:250,000 soil map of England and Wales based on WRB (top) and the 1:1M e-SOTER soil map where the dominant soil group of the e-SOTER units is used for mapping (bottom).

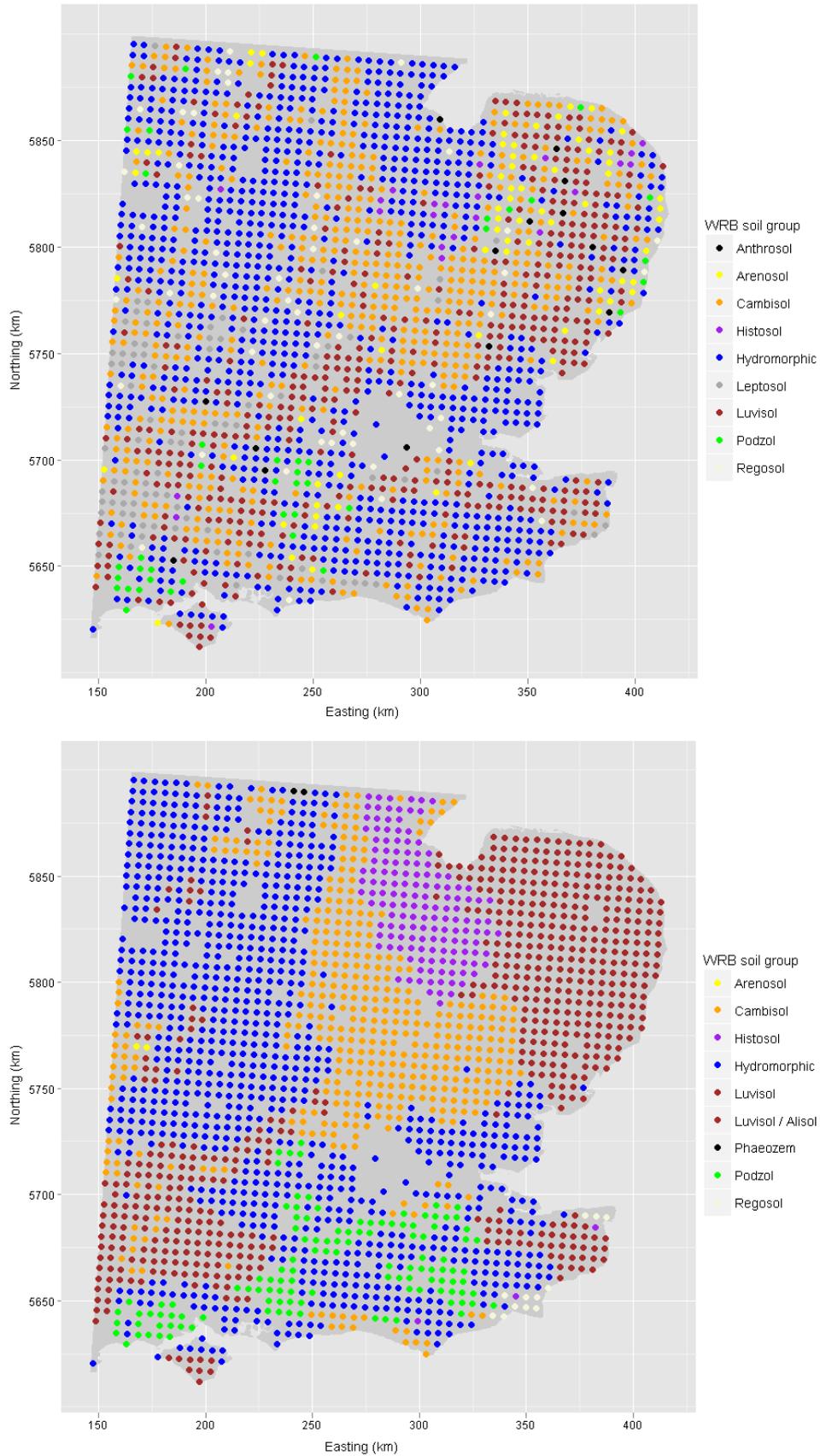


Figure 4.2. Observed WRB soil group (top) and dominant WRB soil group in the e-SOTER soil component (bottom) at the UK validation sites.

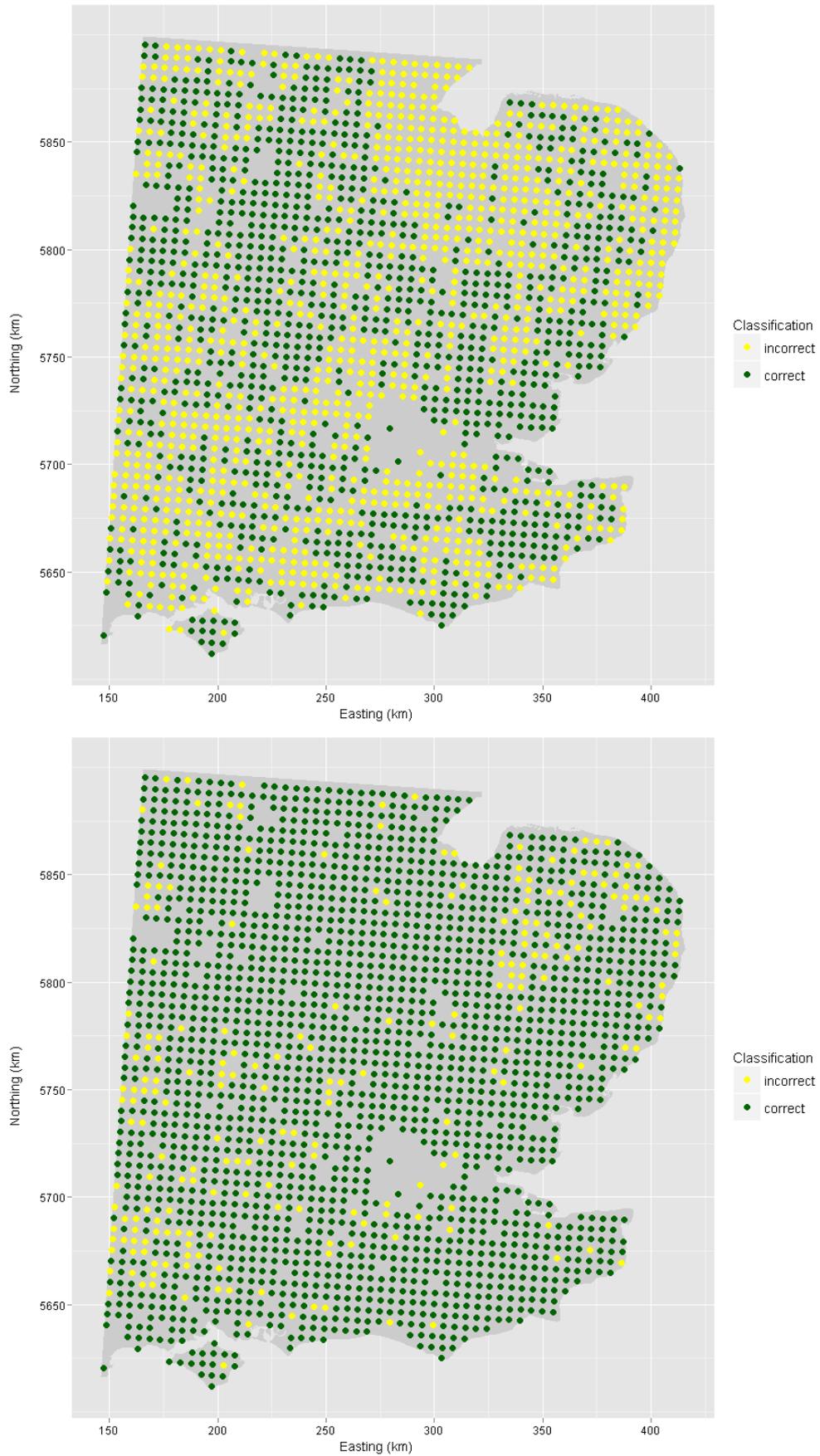


Figure 4.3. Validation results for the stringent mode (top) and flexible mode (bottom) and full legend.

## 4.3 Application to German-Czech part of the Central European window

### 4.3.1 Soil data

The Czech validation dataset contained 437 localized soil profiles distributed evenly in the Czech part of the Central European window. Profiles were selected proportionally according to spatial extent of each soil type (unit) from the Czech soil database. The database includes analysed soil profiles from the detailed maps of the Systematic Soil Survey, a large mapping campaign realized in 1961-1971 that resulted in a set of 1:10,000 soil maps for arable lands. Figure 4.4 gives an overview of the content of each soil profile description. A description includes soil classification according to the Czech Taxonomic Classification System of Soils (CTCSS, Němeček et al. 2011), soil profile structure (horizon thickness), soil horizons definition and naming according to CTCSS up to 150 cm, morphologic soil properties (such as stagnic and gleyic conditions, albeluvisc tonguing etc.), and parent material. Analytic soil properties for each horizon include particle size distribution, soil organic carbon, CaCO<sub>3</sub> content, pH<sub>H2O</sub>, pH<sub>KCl</sub>, CEC and base saturation. Soil profiles were localized and digitized from the maps of the Systematic Soil Survey using their unique identifiers.

Figure 4.4. Example of a validation profile from the Czech soil database with available morphological and analytic data.

Profile ID	Horizon Number	Diagnostic horizon	Diagnostic property	Diagnostic material	Horizon designation	Upper depth	Lower depth	SAND	SILT	CLAY	pH (H2O)	pH KCl	Exchangeable Acidity	CEC
KR09	1				A	0	24	19.9	69.3	10.8	7.7	7.5	0.00	10.30
KR09	2				A/E	24	42	19.3	69.1	11.6	7.6	7.4	0.00	10.30
KR09	3		albeluvisc tonguing		E/Bt	42	52	14.4	68.7	16.9	7.6	7.3	0.00	8.30
KR09	4	argic	stagnic colour pattern		Btg	52	89	20.7	51.9	27.4	7.4	6.8	1.49	15.00
KR09	5				B/C	89	99	5.9	63.1	31.0	5.9	5	4.99	15.40
Profile ID	Lab-ID	Elevation	Land use	Parent material	SOTER classification	Prefix 1	Prefix 2	Prefix 3	WRB soil group	Suffix 1	Suffix 2	Suffix 3	National classification Soil type	National classification specifiers
KR09	UKZUZ	355	arable	loess-like sediments	UTIE		Stagnic	Cutanic	Albeluvisol	Eutric	Siltic		luvizem	modal eubazic

The German validation dataset comprised 252 soil profiles in the German part of the Central European pilot area. The profile data were provided by the Saxon State Agency for Environment, Agriculture and Geology. Generally, soil profile description comprises soil classification, substrate classification, soil horizon specification, parent material, and analytic soil properties for each horizon such as texture, bulk density, field capacity, cation exchange capacity, exchangeable cations and pH.

In total 689 soil profiles were available for validation. Three profiles were discarded because these were located outside the validation area, leaving 683 profiles that were used.

### 4.3.2 Soil correlation

#### *Czech Republic*

The validation dataset provides a solid base for correlation of the Czech and WRB soil classification systems as the majority of analytical data needed for the classification in WRB are available for the whole profile. Quantified data were sufficiently available which allowed us to use an approach based

on classifier rules with numerical data thresholds. Direct semantic (analog soil units provided by expert knowledge) correlation was possible for soil groups where the classification criteria of the WRB and CTCSS were equal. In case of a partial miss of the data, profiles were correlated using expert judgement.

Soil classification in the Czech Republic is guided by the Czech Taxonomic Classification System of Soils (CTSS, Němeček et al., 2011). The CTCSS classification was designed to meet the requirements for effective correlation of the higher taxonomical units with international taxa. Consequently, the principle of morphogenetic feature preference was substituted for the analytical data approach. Assessment of the higher taxonomical units is based on diagnostic horizons and features below 0.25 m of depth (to reduce the influence of different land uses on soil properties). The CTCSS is a multi-categorical system. Reference groups, soil types and subtypes reflect results of long-term soil evolution. Reference groups represent a linkage to the highest taxonomic categories of the world reference systems (WRB, Soil Taxonomy) and other well-known soil classification systems (German, French, Canadian). Soil types (great soil groups) are a central taxonomic category, which involves soil being characterised by a specific sequence of diagnostic horizons and/or diagnostic features. Subtypes as subdivisions of the soil types include modal (typical), inter-grade representatives, extreme debasification degrees and extremes of the soil texture (Němeček et al., 2011). Subtypes serve the same role as the qualifiers in WRB, but their criteria vary slightly in some soil types in CTCSS (e.g. the Arenic subtype has different limits in Černozem and in Kambizem).

The CTCSS soil types and their WRB counterparts are presented in Table 4.7. Only a few soil types could be directly correlated to the RSGs of WRB. Semantic correlation was used in case of complete agreement between the classification criteria in WRB and CTCSS. This condition is fulfilled mostly in soils defined by morphological properties. Directly correlated soil types include Fluvisols, Gleysols, Stagnosols and Vertisols. Another example of possible direct correlation are soils defined by clay migration in the profile (Hnědozem and Šedozem/Luvisols). Minimum values required for luvic/argic horizon in CTCSS are higher (clay ratio between Bt- and E-horizons 1.3) than in WRB (clay ratio 1.2) and presence of clay coatings is compulsory. This implies an automatic conversion of the luvic horizon in CTCSS to an argic horizon in WRB. A similar approach was used in case of Luvic Cambisols in CTCSS. In CTCSS, the subtype Luvic is defined by presence of clay coatings in the B horizon. This criterion is sufficient for identification of an Argic horizon in WRB. All Luvic Cambisols in CTCSS could be then correlated as Luvisols in WRB.

Other soil types were correlated using numeric limits required in WRB. Available profile analytical data were sufficient to distinguish Argic, Mollic, Spodic and Calcic horizons. Expert judgement was applied in case of lack of data.

The following WRB RSGs were identified: Albeluvisol, Arenosols, Cambisols, Fluvisols, Gleysols, Chernozems, Leptosols, Luvisols, Phaeozems, Planosols, Podzols, Regosols, Stagnosols, Vertisols. According to simplifications of the e-SOTER legend resulting from WP2 methodology, some of the RSGs were grouped: Fluvisols, Gleysols, Stagnosols and Planosols were grouped as Hydromorphic soils, Albeluvisols and Luvisols were grouped as Luvisols and Luvisols/Alisols.

Table 4.7. CTCSS soil types and their WRB counterparts according to Němeček et al. (2011).

CTCSS soil group	WRB06 RSG (qualifier)
Litozem	Lithic Leptosol
Ranker	Leptosol (Skeletal)
Rendzina	Rendzic Leptosol
Pararendzina	Leptosol (Calcaric)
Regozem	Regosol, Arenosol
Fluvizem	Fluvisol
Koluvizem	-
Smonice	Vertisol
Černozem	Chernozem
Černice	Gleyic Chernozem
Šedozem	Greyic Phaeozem
Hnědozem	Luvisol
Luvizem	Albeluvisol
Kambizem	Cambisol
Pelozem	Cambisol (Clayic)
Kryptopodzol	Entic Podzol
Podzol	Podzol
Pseudoglej	Stagnosol
Stagnoglej	Stagnosol
Glej	Gleysol
Organozem	Histosol
Kultizem	Anthrosol
Antropozem	Technosol

In addition, the effectiveness of a semantic (direct) approach and quantitative approach for soil correlation was evaluated. Table 4.8 presents the level of correlability which is the percentage of profiles of a certain CTCSS soil type that falls into a certain WRB soil group after correlation. This may or may not be the analogical soil group. Correlation at the lower level included only subtypes/qualifiers that occurred in more than five soil profiles in one or both of the classification systems.

Average correlability at soil type level was 88%. Overall accuracy of the correlation is high but there is considerable variation (from 50% to 100%) between the soil types. Ten Czech soil types have high (80–100%) correlability, seven soil types have intermediate (60–80%) correlability and one soil type has low (less than 60%) correlability. The conversion of some soil types, for instance Rankers, Rendzinas, Pararendzinas, Černice, Černozems, Podzols or Luvizems, has low accuracy and requires analytical and morphological data of corresponding profiles. Other soil types such as Glejs, Fluvizems or Hnědozems can be correlated with a high probability of accurate assignment. The analysis showed high incompatibility at the level of soil subtypes/qualifiers (Table 4.9). Correlation at the lower taxonomical level should be subject to analytical processing of quantitative soil data. Three main causes of problems in correlating the WRB and the CTCSS are different concepts or criteria of the soil unit, different limit values and indistinct criteria in CTCSS. The most frequent problem is the different setting of limit values. Zádorová and Penížek (2011) provide further details on this analysis.

Table 4.8. Correlability of soil types in CTCSS and WRB 2006 (soil types and Reference Soil Groups + main qualifiers) .

CTCSS	WRB 2006	Correlability (%)	Accessory units
Ranker	Haplic Leptosol	67	Leptic Cambisol
Rendzina	Leptosol (Calcaric)	75	Cambisol (Eutric)
Pararendzina	Leptosol (Calcaric)	92	Cambisol (Eutric)
Regozem	Regosol	80	Chernozem
Regozem Arenic	Arenosol	100	-
Fluvizem	Fluvisol	100	-
Smonice	Vertisol	100	-
Černozem	Chernozem	91	Phaeozem, Regosol
Černice	Gleyic Chernozem	67	Gleyic Phaeozem
Šedozem	Luvic Gleyic Phaeozem	100	-
Hnědozem	Luvisol	100	-
Luvizem	Albeluvisol	63	Albic Luvisol
Kambizem	Cambisol	94	Luvisol, Arenosol
Pelozem	Cambisol (Clayic)	50	Cambisol (Siltic)
Kryptopodzol	Entic Podzol	100	-
Podzol	Podzol	67	Cambisol (Dystric)
Pseudoglej	Stagnosol	74	Planosol, Stagnic Albeluvisol
Glej	Gleysol	100	-
total		88	

Table 4.9. Correlation at the lower taxonomical level.

CTCSS subtype	Count	Criteria CTCSS	WRB qualifier	Count	Criteria WRB
arenic	19	sand, loamy sand or sand, loamy sand, sandy loam	Arenic	1	sand, loamy sand in a layer 30 cm or more thick
pelic	111	clay, sandy clay, silty clay, sandy clay loam, clay loam, silty clay loam	Clayic	29	clay in a layer 30 cm or more thick
dystric	22	BS less than 30 %	Dystric	81	BS less than 50 %
luvic	52	Variable criteria in different soil types	Luvic	28	having a luvic horizon
stagnic	47	medium redoximorphic features within 60 cm from soil surface	Stagnic	110	25 % or more of soil volume stagnic colour pattern within 100 cm from soil surface
gleyic	15	strong reducing conditions below 60 cm from soil surface	Gleyic	36	25 % or more of soil volume gleyic colour pattern within 100 cm from soil surface

### Germany

The information of the German soil profile data set is mainly restricted to the soil type, substrate, parent material, texture, pH (CaCl<sub>2</sub>), and organic carbon. Information about cation exchange capacity, exchangeable cations, base saturation is quite incomplete. The correlation from German soil classification in WRB (RSG and qualifiers) is therefore mainly based on expert knowledge and recommendations according to Ad-hoc-AG Boden (2005). As the concept of the German soil classification is based on soil genesis as expressed by characteristic horizon sequences (Ad-hoc-AG

Boden 2005), it is possible to derive WRB 2006 RSGs from the German soil type under consideration of parent material and local climate.

Table 4.10. Correlation between WRB and German soil classification.

KA5 soil type	WRB 2006	Prefix	Suffix	N
AB-GG	Fluvisol	Gleyic	Siltic	3
AB-GG	Fluvisol	Gleyic	Skeletal	3
AB-GG	Fluvisol	Gleyic		4
ABn	Fluvisol	Haplic	Skeletal	1
ABn	Fluvisol	Haplic		2
BBn	Cambisol	Haplic		49
BBn	Cambisol	Haplic	Skeletal	4
BB-PP	Podzol	Haplic		9
BB-PP	Podzol	Haplic	Skeletal	15
BB-SS	Luvisol	Stagnic		1
BB-SS	Cambisol	Stagnic		1
GGa	Fluvisol	Gleyic	Skeletal	3
GGa	Fluvisol	Gleyic		1
GGa	Fluvisol	Gleyic		10
GGa	Gleysol	Haplic	Siltic	1
GGa	Fluvisol	Gleyic	Siltic	1
GGn	Gleysol	Haplic		10
GGn	Gleysol	Haplic	Siltic	2
GGn	Gleysol	Haplic	Skeletal	1
GGn	Gleysol	Histic		1
LF-SS	Stagnosol	Luvic	Albic	3
LF-SS	Luvisol	Stagnic		1
LLn	Luvisol	Haplic	Abruptic	1
LLn	Luvisol	Haplic		2
LL-SS	Stagnosol	Luvic	Albic, Dystric	1
LL-SS	Stagnosol	Luvic		14
pBB\PP	Podzol	Haplic	Skeletal	1
pBBn	Cambisol	Haplic	Dystric, Skeletal	1
pBBn	Cambisol	Haplic	Dystric	1
pBBn	Cambisol	Haplic		2
PPn	Podzol	Haplic	Skeletal	12
PPn	Podzol	Haplic		8
pSS-LF	Luvisol	Albic		1
RQn	Leptosol	Hyperskeletal		1
RQn	Regosol	Haplic		3
RQn	Regosol	Haplic	Skeletal	1
RQn	Regosol	Haplic	Skeletal, Transportic	4
RQn	Regosol	Haplic	Transportic	4
SGn	Stagnosol	Haplic	Albic, Dystric	1
SS-LF	Luvisol	Albic		1
SS-LL	Luvisol	Stagnic		4
SS-LL	Luvisol	Stagnic	Hypereutric	1
SS-LL	Luvisol	Stagnic		1
SSn	Stagnosol	Haplic		7
SSn	Stagnosol	Luvic		1
sYK-GG	Gleysol	Mollic		1
YKn	Cambisol	Haplic	Colluvic, Dystric	1
YKn	Cambisol	Haplic	Colluvic	6

The German “Gley-Vega” has both a gleyic colour pattern and fluvic material, suggesting Gleyic Fluvisols. The German “Vega” has fluvic material starting within 25 cm from the soil surface and no other diagnostic horizons and properties. Consequently, it fulfills the requirements of Haplic Fluvisols. “Braunerde-Podzols” have a spodic horizon and fulfill the requirements of Haplic Podzols. Haplic Podzols are quite common at highest elevations on the different granites of the Ore mountains. “Braunerde-Pseudogleys” show a stagnic colour pattern, sometimes clay enriched subsoils and sometimes a Cambic horizon. Soils without clay illuviation are Stagnic Cambisols, while soils which fulfill the criteria of an Argic horizon can be classified as Stagnic Luvisols. Luvisols require a base saturation of more than 50% in the major part between 50 and 100 cm soil depth. This condition is commonly fulfilled where the parent material contains carbonates (e.g. loess) or where the parent material is not too siliceous or the pH values are rather high. The “Gleye” soils have a gleyic colour pattern and can be classified as Gleysols. Those with fluvic material within 25 cm from the soil surface are classified as Gleyic Fluvisols. “Parabraunerden” have an Argic horizon. Parabraunerden are widespread in loess areas, suggesting a base saturation of more than 50% and can be classified as Haplic Luvisols. “Braunerden” have a horizon sequence matching the requirements of a Cambic horizon and can be translated as Haplic Cambisols. “Regosole” are soils at the initial state of soil formation and can be translated as Haplic Regosols. “Fahlerden” show in addition to “Parabraunerden” an Albic horizon and can be classified as Haplic Luvisols (Albic). Table 4.10 summarizes soil correlation for the German part of the Central European window.

#### 4.3.3 Validation of the soil component of the e-SOTER units

Figure 4.5 shows the e-SOTER soil map. Here the dominant soil group in the soil component associated to an individual e-SOTER unit is used for mapping.

##### *Stringent mode, full legend*

Table 4.11 presents the sample error matrix of mapped soil group, which corresponds to the dominant soil group of the e-SOTER units that contain the validation locations (Figure 4.6, bottom), against the observed soil group (Figure 4.6, top). The overall purity is 32.1% (30.1% in the Czech part and 32.3% in the German part of the window). The rather low purity can be assigned to high variability of the soil cover and often low dominance of the dominant soil unit in the SOTER units. The difference between the overall purity in the CE and WE windows is remarkable (32.1% for the CE and 51% for the WE window, see section 4.2). The source of this discrepancy can be explained by the fact that the validation criteria for several RSGs were relaxed in the WE window. This flexibility refers to soil groups with stagnic and gleyic properties (section 4.2.3). By contrast, the dataset available for the CE window was considered of sufficient quality and completeness for more strict validation criteria. This means that the location where e.g. Stagnic Luvisol was observed, was considered correct only if Luvisol or Luvisol/Alisol was mapped. When taking into account the large area of soils influenced by water stagnation in the CE window, the use of strict or relaxed validation can result in significantly different overall purities.

Figure 4.7 (top) shows the validation data set with correct and incorrect classification. The errors are fairly evenly distributed in the window which corresponds with the low overall purity. Nevertheless, the most remarkable error accumulations can be found in three parts of the window: two lowland areas in Central Bohemia and Southern Moravia, represented by loess-derived loamy soils, and SW Germany where a large number of Podzols are incorrectly mapped as Cambisols.

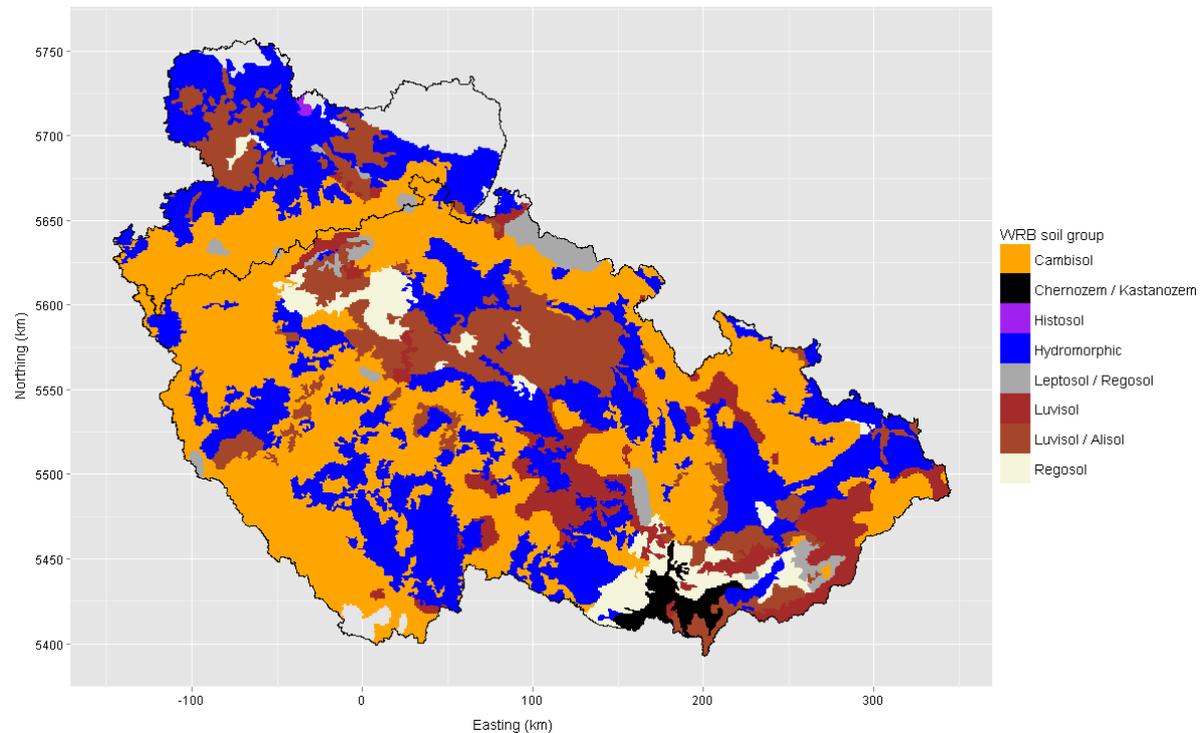


Figure 4.5. The 1:1M e-SOTER soil map where the dominant soil group of the e-SOTER units is used for mapping. Largest part of the area is in the Czech Republic, while the north-western corner lies in Germany.

The soil groups show different levels of accuracy. The most problematic soil units are Chernozems, Podzols, Cambisols and Luvisols. Chernozems and Phaeozems were represented by 63 soil profiles in the validation dataset (almost exclusively occurring in the Czech part of the window). However, the class representation of the two RSGs is very low (4% for the Chernozems and 0% for Phaeozems which means that the latter is never the dominant RSG in e-SOTER soil map). This result indicates that the area with mollic soils is heavily under-represented on the e-SOTER soil map (Figure x), which is confirmed when this map is compared to Czech soil maps. The majority of the observed Chernozems and Phaeozems are mapped as Luvisols and Regosols. Soil pedons dominated by mollic soils are adjacent to the Luvisol areas and both soil units are complementary in a large area of flat and slightly undulating terrain of lowlands and plains. The association of Chernozems with Regosols is due to the long-term evolution of the agricultural land. Most of the area is situated in the undulating relief underlain by silty sediments. Chernozems are extremely vulnerable to the soil erosion and nowadays the former Chernozem region is a mosaic of Chernozems and its eroded form, the Regosol, on the steep slopes and its accumulated forms, deep colluvial soils. The process of the material redistribution was accelerated in last 60 years so that the situation in the 1960s, when the

soil survey was implemented, was different and the area of Regosols has expanded ever since. The high map unit purity (66.7%) of the Mollic soils has to be interpreted with care because there were only three validation locations situated within a mapped mollic soil group.

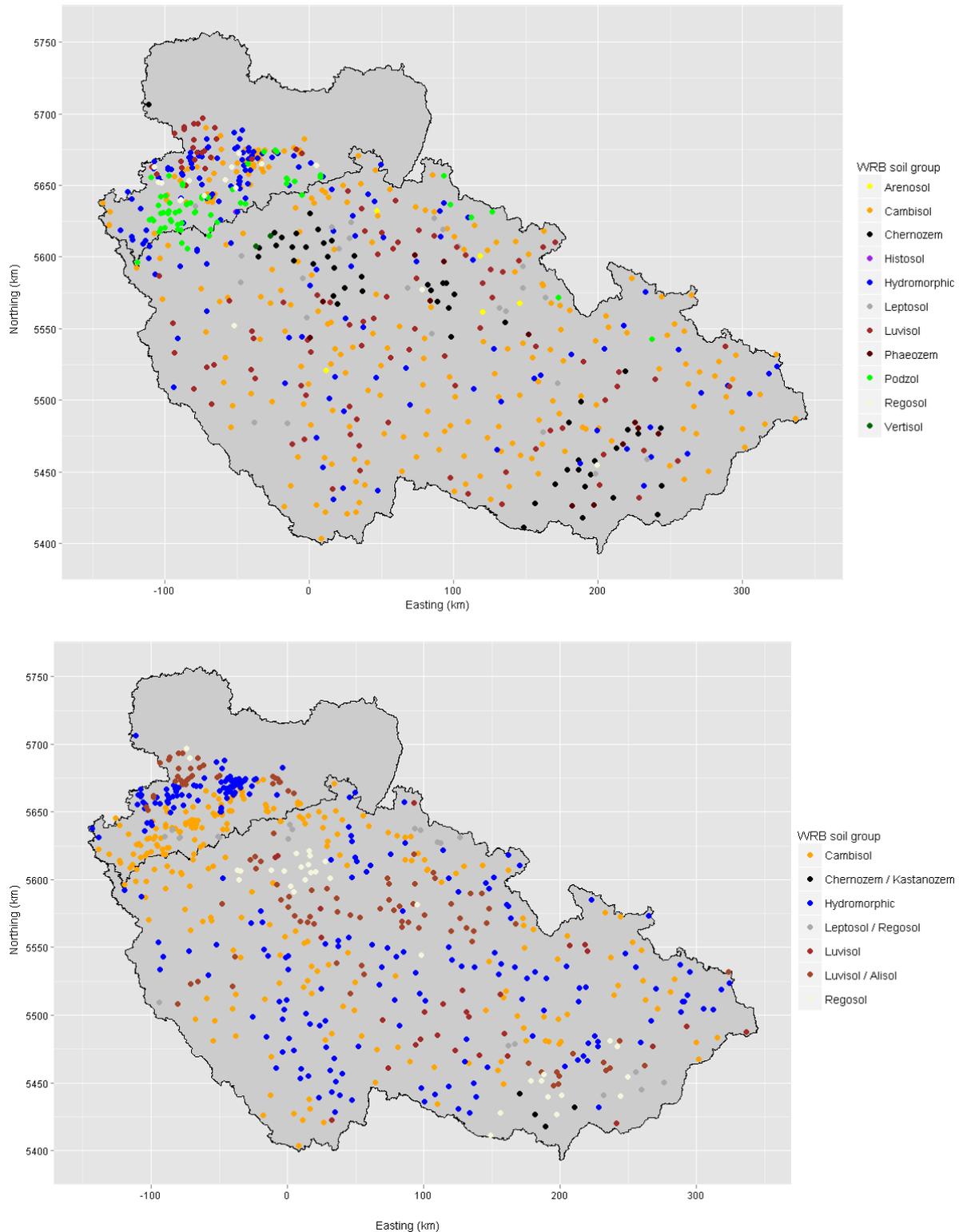


Figure 4.6. Observed WRB soil group (top) and dominant WRB soil group in the e-SOTER soil component (bottom) at the Central European validation sites.

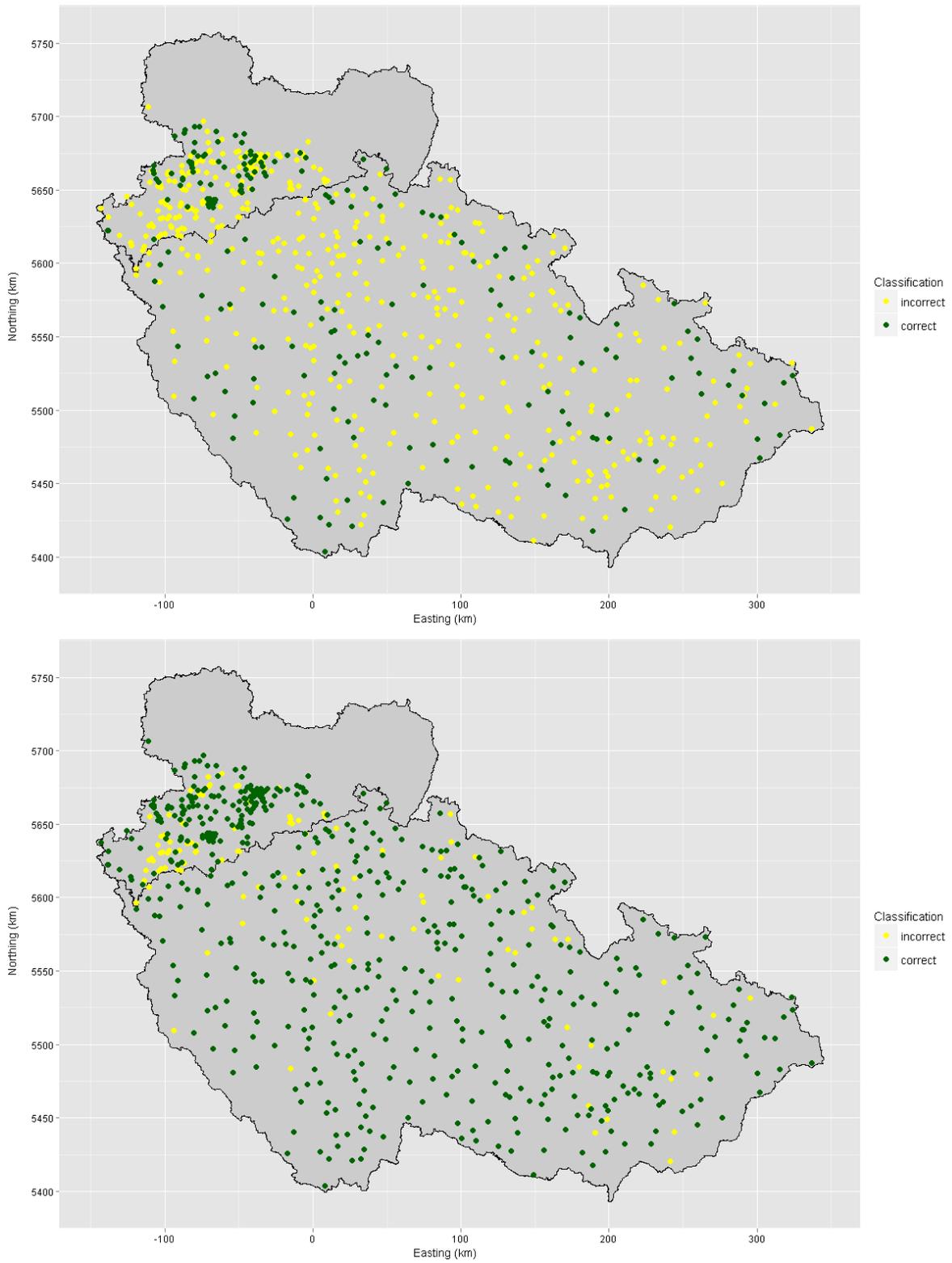


Figure 4.7. Validation results for the stringent mode (top) and flexible mode (bottom) and full legend.

Podzols dominate large parts of SW Germany and neighbouring Czech regions. However, these soils were not indicated as a dominant soil group in any of the e-SOTER soil components, while these were represented by 63 profiles in the validation dataset (52 in the German part). Podzols are prominent mostly in the windward granite part of the Ore mountains, where the acid parent material and high precipitation amount imply the podzolization process of variable intensity. However, they are often associated with Cambisols, Hydromorphic and undeveloped soils so the confusion with these units is comprehensible. The fact that there were only twelve Podzol profiles available in the CE training dataset (Table 1 of Deliverable D5) might have hampered mapping of this soil group and might therefore contribute to the poor validation results for the Podzols.

Cambisols, Luvisols, Hydromorphic soils and Leptosols form the majority of the soil cover in highland regions of the Czech Republic. The four groups have different intergrades represented by transitional soil units and often can be found in associations. Medium or low map unit purity and class representation of these soils and fact that a large part of for example Cambisols (map unit purity 46.4%) falls in the soil component dominated by Luvisols (map unit purity 25.2%) or Hydromorphic soils (map unit purity 26.2%) and vice versa can be then assigned to the rather low dominance of any of these units and their complementarity. The most significant difference between map unit purity (26.2%) and class representation (40%) occurs in the Hydromorphic soil unit; it was mapped as a dominant unit at a large part of the area (Figure 4.5) and it seems over-represented in the e-SOTER map. However, a majority of the soil units forming the soil cover in these regions are more or less influenced by reducing conditions. Soils meeting the criteria limits for classification as Hydromorphic and soils having only stagnic/gleyic qualifier vary at short distances and form a complex mosaic. Moreover, especially Stagnosols are naturally associated with clay illuviated soils which explains a partial misclassification between Luvisols and Hydromorphic soils. Leptosols have very low both map unit purity and class representation (both around 4%). These are confused mostly with Cambisols, Luvisols and Hydromorphic soils.

#### *Stringent mode, generalized legend*

Table 4.13 shows the error matrix of mapped against observed soil groups for the generalized legend. The overall purity is 31.2% (29% in the Czech part and 34.9% in the German part of the CE window). The overall purity for the generalized legend is very similar to the overall purity for the full legend. The effect on map unit purity and class representation varied between soil groups (Table 4.14). Argic soils were better represented, while Cambisols had smaller map unit purity and class representation for the generalized legend.

#### *Flexible mode*

The overall purity for the flexible mode is 83% for the full legend and 86% for the generalized legend (87% and 92% in the Czech part and 76% and 76% in the German part). This means that the e-SOTER soil component includes the observed soil group in more than 80% of the validation area. The large difference between overall purities in stringent and flexible modes is due to a high spatial variation

of the soil units and generally low dominance of any soil unit in the soilscape. Figure 4.7 (bottom) shows the validation locations correctly and incorrectly classified. The only area of accumulated incorrect validation is in the SW part of the Ore mountains where Podzols are observed but not included in the e-SOTER soil components.

Table 4.11. Error matrix of mapped (dominant) soil group against observed soil group for the full legend. Bold type indicate the number of validation locations for which the mapped soil group corresponds to the observed soil group.

Mapped RSG	Observed RSG											Total
	AR	CM	HS	HY	CK	LR	LV	PH	PZ	RG	VT	
AR	<b>0</b>	0	0	0	0	0	0	0	0	0	0	0
CM	1	<b>121</b>	0	53	2	5	26	0	46	7	0	261
HS	0	0	<b>0</b>	0	0	0	0	0	0	0	0	0
HY	2	83	0	<b>60</b>	7	7	52	5	8	5	0	229
CK	0	0	0	0	<b>2</b>	0	0	1	0	0	0	3
LR	0	2	1	5	1	<b>1</b>	4	0	8	0	0	22
LV	0	15	0	6	3	0	<b>4</b>	0	1	0	0	29
LA	2	16	0	22	18	9	<b>30</b>	4	0	4	1	106
PH	0	0	0	0	0	0	0	<b>0</b>	0	0	0	0
PZ	0	0	0	0	0	0	0	0	<b>0</b>	0	0	0
RG	0	3	0	4	18	3	2	2	0	<b>0</b>	1	33
VT	0	0	0	0	0	0	0	0	0	0	<b>0</b>	0
Total	5	240	1	150	51	25	118	12	63	16	2	683

Table 4.12. Estimated map unit purities of and class representations based on an e-SOTER soil map with the full legend that depicts the dominant soil group of the soil components.

Soil group	Map unit purity (%)	Class representation (%)
Arenosol	- <sup>a</sup>	0
Cambisol	46.4	50.4
Histosol	- <sup>a</sup>	0
Hydromorphic	26.2	40
Chernozem / Kastanozem	66.7	3.9
Leptosol / Regosol	4.5	4
Luvisol	13.8	28.8
Luvisol / Alisol	28.3	- <sup>b</sup>
Phaeozem	- <sup>a</sup>	0
Podzol	- <sup>a</sup>	0
Regosol	0	0
Vertisol	- <sup>a</sup>	0

<sup>a</sup> Not mapped as dominant soil group in CE validation area.

<sup>b</sup> Not observed in the validation dataset.

Table 4.13. Error matrix of mapped (dominant) soil group against observed soil group for the generalized legend.

Mapped RSG	Observed RSG									Total
	AG	AR	CM	DK	HS	HY	PZ	SH	VT	
AG	<b>41</b>	2	51	25	0	28	7	15	2	171
AR	0	<b>0</b>	0	0	0	0	0	0	0	0
CM	19	1	<b>88</b>	2	0	51	39	11	0	211
DK	0	0	0	<b>4</b>	0	1	0	0	0	5
HS	0	0	0	0	<b>0</b>	0	0	0	0	0
HY	48	2	81	12	0	<b>57</b>	7	11	0	218
PZ	0	0	0	0	0	0	<b>0</b>	0	0	0
SH	10	0	20	20	1	13	10	<b>4</b>	0	78
VT	0	0	0	0	0	0	0	0	<b>0</b>	0
Total	118	5	240	63	1	150	63	41	2	683

Table 4.14. Estimated map unit purities of and class representations based on a SOTER soil map with the generalized legend that depicts the dominant soil group of the soil components.

Soil group	Map unit purity (%)	Class representation (%)
AG	23.9	34.7
AR	- <sup>a</sup>	0
CM	41.7	36.7
DK	6.3	80
HS	- <sup>a</sup>	0
HY	26.1	38
PZ	- <sup>a</sup>	0
SH	5.1	9.8
VT	- <sup>a</sup>	0

<sup>a</sup> Not mapped as dominant soil group in CE validation area.

## 5. Uncertainty Propagation Analysis

### 5.1 Methodology

#### 5.1.1 Introduction

Uncertainty propagation analysis analyses the effect of uncertain model inputs on model output (Heuvelink, 1998). In case of the e-SOTER methodology the model input is the SRTM DEM and the model output the physiographic unit (PU) map.

Other inputs in the e-SOTER procedure are parent material maps (surface condition, genetics, texture and carbonate) and the soil component database. However, these are categorical inputs and the uncertainty associated to such inputs is not easily quantified (as opposed to continuous inputs such as elevation). Besides, these are not really used as model inputs in the e-SOTER procedure in a sense that the model output is a function of these inputs. The parent material maps are merely used to derive terrain units from the physiographic units by spatial overlay. The soil component database is then linked to terrain units by a common identifier. Therefore, the uncertainty propagation analysis focused on the effect of DEM uncertainty on the PU map.

Uncertainty propagation analysis of DEM uncertainty is done for the WE and CE pilot areas.

#### 5.1.2 Quantifying DEM uncertainty

The e-SOTER methodology uses the 90 m SRTM DEM to derive the PU map. This DEM is not error-free: there will be a difference between SRTM elevation and true elevation. Therefore the first step is to quantify the DEM error. This is done as follows (Temme et al., 2008).

Let us denote the true elevation as  $Z$ , about which we are uncertain. True elevation can be represented as the sum of the SRTM elevation  $z^*$  and an unknown error  $\varepsilon$ :

$$Z(s) = z^*(s) + \varepsilon(s), \quad (5.1)$$

where  $s$  is a spatial location. The DEM error  $\varepsilon$  at location  $s$  is unknown (unless we visit the location and measure the true elevation), which means that the error must be quantified in terms of a probability distribution. A sensible and convenient distribution for the error is the normal distribution. This distribution is fully characterized by two parameters: the mean  $\mu$  and standard deviation  $\sigma$ . If accurate measurements (preferable with negligible error) of the true elevation at a set of control points are available ( $z(s_i)$ ,  $i=1, \dots, n$ ), then the  $\varepsilon(s_i)$  can be computed at these points (i.e. the difference between true elevation and SRTM elevation). From the obtained errors the mean and standard deviation can be estimated and thus the probability distribution of the error derived. For this, a simplifying stationarity assumption is needed, which is elaborated in the next section.

### 5.1.3 Modelling DEM uncertainty

Once DEM uncertainty is quantified on the basis of a sample of elevation measurements at control points, the DEM uncertainty can be modelled. Hereby spatial correlation must be taken into account. After all, it is very likely that errors at closely located locations show similarities. To characterize the spatial correlation structure of the errors and to model the error in space, geostatistical methods are used. These methods typically assume that the mean and variance of the error distribution are constant in space. For the mean this is plausible but for the variance this might be an unrealistic assumption. It is perhaps more realistic to assume that the variance of the error is proportional to the ruggedness of the terrain. Therefore we model the spatially correlated residual  $\varepsilon$  in equation (6) as (Beekhuizen et al., 2011):

$$\varepsilon(s) = r(s) \times \varepsilon_s(s), \quad (5.2)$$

where  $r(s)$  is the terrain ruggedness at location  $s$ , and where  $\varepsilon_s(s)$  is a spatially correlated, normally distributed residual with constant mean and variance. The spatial structure of  $\varepsilon_s(s)$  is characterized by the semivariogram. The semivariogram is a function that describes the degree of spatial dependence between two spatial locations:

$$\gamma(h) = \frac{1}{2} E[\varepsilon_s(s) - \varepsilon_s(s+h)]^2, \quad (5.3)$$

where  $\gamma(h)$  is called the semivariance and  $h$  is the lag distance, i.e. the distance between two spatial locations. The semivariance typically is small when  $h$  is small, i.e. when the separation distance between two locations is small. The semivariance is computed for all pairs of control points. These are then plotted against the lag distance, yielding the experimental semivariogram. Next a semivariogram model is fitted to the experimental semivariogram that describes the spatial dependence for all possible point distances, not only those in the sample. Once the spatial dependence structure of the standardized error is quantified, we can model DEM uncertainty at every location in the study area.

We followed the approach suggested by Beekhuizen et al. (2011) to calculate the terrain ruggedness  $r(s)$ . Terrain ruggedness  $r(s)$  is defined as the standard deviation of all DEM heights in a kernel of  $p \times p$  pixels centred around  $s$ . The optimum value for the kernel width  $p$  can be found by calculating the correlation between error and ruggedness for different kernel widths. We used widths of 3, 5, 7, 9 and 10 pixels. The kernel width with the greatest correlation coefficient is selected.

### 5.1.4 Monte Carlo simulation for DEM uncertainty propagation analysis

The effect of DEM uncertainty on the PUs is analysed with Monte Carlo stochastic simulation. This works as follows (Temme et al., 2008; Beekhuizen et al., 2011).

A location (a 90 m raster cell) in the pilot area is selected at random. At this location a value from the probability distribution of the standardized error  $\varepsilon_s(s)$  is drawn and assigned to that location. At the second randomly selected location the conditional probability distribution of  $\varepsilon_s(s)$  is computed

by conditioning the probability distribution of  $\varepsilon_s(s)$  to the value that was drawn at the first location. Conditioning is done using the semivariogram and kriging. This ensures that the simulated values have the specified spatial dependence structure. Next a value is drawn from the conditional probability distribution and added to the dataset with simulated values (which at this point contains the value simulated at the first location). This procedure is repeated until all raster cells have been visited. The method of visiting locations in random order and the drawing of values of the conditional probability distribution at each location, taking into account the values drawn at previously visited locations, is called unconditional sequential Gaussian simulation (Goovaerts, 1997).

The raster with simulations of  $\varepsilon_s$  is transformed to simulations of  $\varepsilon$  by multiplying it with the terrain ruggedness  $r$ . Next, the residuals  $\varepsilon$  are added to the SRTM elevation measurements  $z^*$ . The resulting raster represents one realization (out of many) of true elevation  $Z$  given the specified uncertainty model. The simulation procedure is repeated  $N$  times. Each simulated raster is a realization of the true elevation. The differences between these realizations represent the uncertainty about the true DEM.

The final step in the uncertainty propagation analysis is to run the e-SOTER procedure for the  $N$  simulated DEMs, resulting in  $N$  physiographic unit (PU) maps. The PU maps and the landform attribute (LFA) maps that can be derived from the PU maps are then compared to analyse the effect of DEM uncertainty on these e-SOTER products.

### 5.1.5 Analysis of DEM uncertainty on e-SOTER terrain products

A map of each of the four LFAs was derived from each of the simulated PU maps, giving  $N$  maps for each LFA. From these maps the probability distribution of the LFA classes was computed for each location in the pilot area (a 990 m x 990 m raster cell; the output resolution of the e-SOTER procedure) by dividing the number of simulations of each LFA class by  $N$ . Next, from the location-specific probability distributions the dominant LFA class was determined, which simply is the class with the largest probability. A map depicting the dominant LFA class was subsequently compared to the default LFA map. This default map was obtained by running the e-SOTER procedure with the original SRTM DEM of the pilot area as input.

The Shannon entropy was used as a measure of uncertainty of the LFA classes, which is defined as:

$$H = - \sum_{i=1}^{n_c} \pi(c_i) \log(\pi(c_i)), \quad (5.4)$$

where  $\pi(c_i)$  is the probability that LFA  $C$  takes class  $c_i$  and  $n_c$  is the number of LFA classes (Brus et al., 2008). We used a logarithm with base  $n_c$  so that the maximum entropy is 1, which occurs when all LFA classes have equal probability. The minimum value for the entropy is 0, which occurs when there is no uncertainty and one of the LFA classes has probability 1, and all others 0.

## 5.2 Application to the Western European pilot area

### 5.2.1 Ground control points and DEM error

Field measured ground control points were not available to quantify DEM error in the WE pilot area. We therefore followed the approach suggested by Beekhuizen et al. (2011) and simulated a collection of control points by randomly sampling the NEXTmap digital terrain model (DTM) of Great Britain with 10-m spatial resolution. The vertical accuracy of this DTM is  $\pm 1$  m. The elevation at the control points does therefore not have negligible error which is preferred for error propagation analysis. Nevertheless it is considerably smaller than the vertical accuracy of the SRTM DEM which has an absolute vertical accuracy requirement of  $\pm 16$  m for 90% of the area and a relative vertical accuracy (the error in a local 225x225 km area) requirement of  $\pm 6$  m (Rabus et al., 2003).

A random sample of 2,000 points was generated within a rectangle covering the WE pilot area. Of these 1,581 points remained after excluding points outside the pilot area (e.g. in the sea). Figure 5.1 shows the SRTM DEM of the WE pilot area, the location of the control points and the DEM error at these points. The mean error is 0.11 m, the variance  $3.8 \text{ m}^2$ . The maximum absolute error is 16.6 m. Figure 5.2 (left plot) shows a ruggedness map based on a  $3 \times 3$  kernel, thus showing the ruggedness within a 270 by 270 m window. The  $3 \times 3$  kernel ruggedness map had the strongest correlation with DEM error ( $r = 0.296$ ) and was therefore used to compute the standardized error  $\varepsilon_s(s)$ . Figure 5.2 (right plot) shows the scatterplot of the ruggedness against the absolute DEM error.

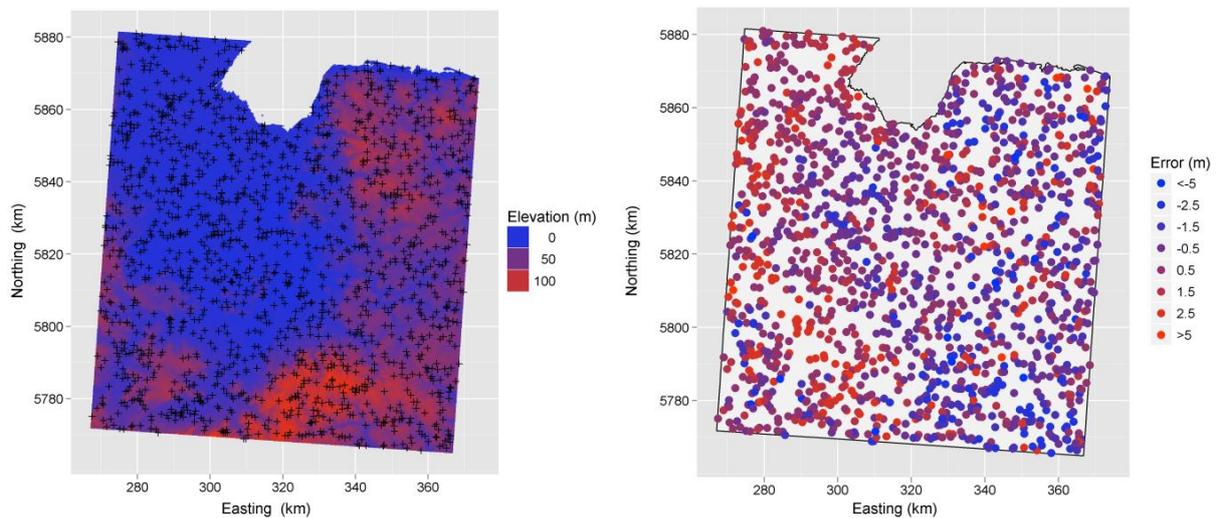


Figure 5.1. SRTM DEM with the location of the control points (left) and the DEM error at these points (right).

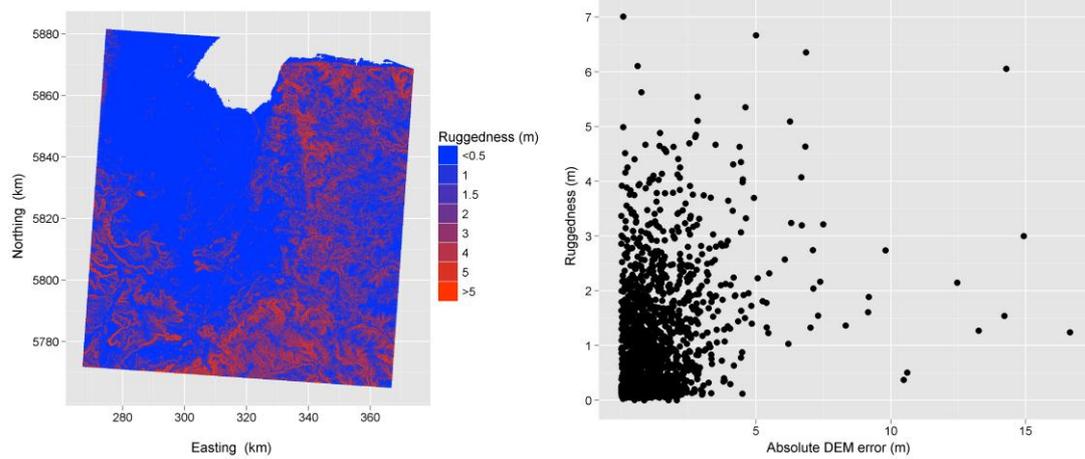


Figure 5.2. Ruggedness map created with a 3 x 3 cell kernel (left) and correlation between 3 x 3 kernel ruggedness and be absolute DEM error (right).

### 5.2.2 Modelling DEM uncertainty

Figure 5.3 shows the histogram and boxplot of the standardized error. Both plots show that the standardized errors approximately follow the normal distribution which is convenient for modelling the spatial structure of the error. There are five outliers with absolute standard errors around 20 m. These points were excluded from the dataset for semivariogram modelling as these heavily influence the experimental semivariances. The experimental semivariogram was computed from the standardized errors at the control points. A theoretical semivariogram model was fitted from the experimental semivariogram with an exponential structure, a nugget of  $4.39 \text{ m}^2$ , a partial sill of  $3.44 \text{ m}^2$  and a range of 12,051 m (Figure 5.4). The semivariogram and standardized errors at the control points were used to generate 1,000 Monte Carlo simulations of the DEM error of which Figure 5.5 shows four examples. Next, the simulated fields of DEM error were added to the SRTM DEM to create 1,000 simulated DEMs.

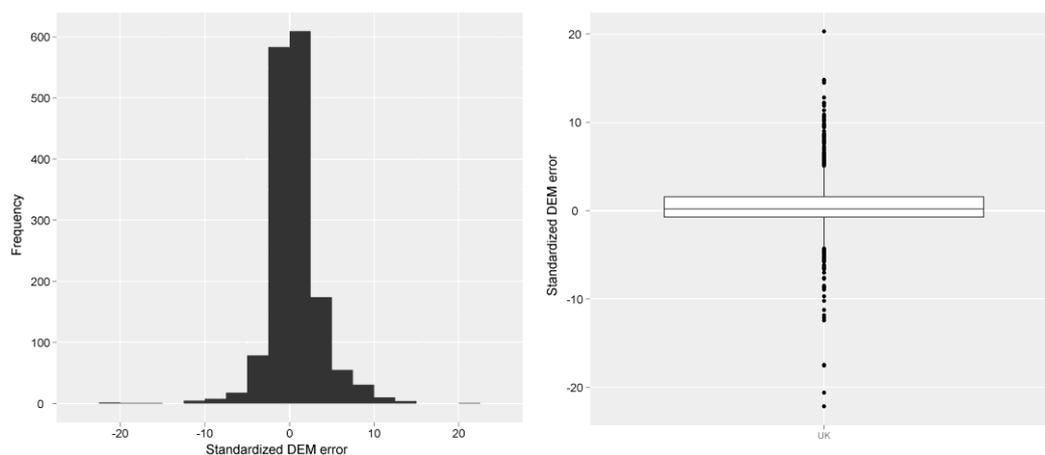


Figure 5.3. Histogram and boxplot of the standardized DEM error.

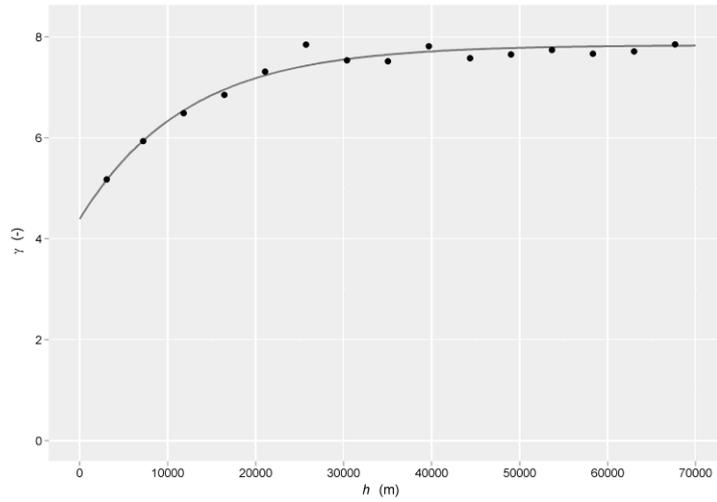


Figure 5.4. Experimental semivariogram of the standardized DEM error at the control points (dots) and fitted semivariogram model (line).

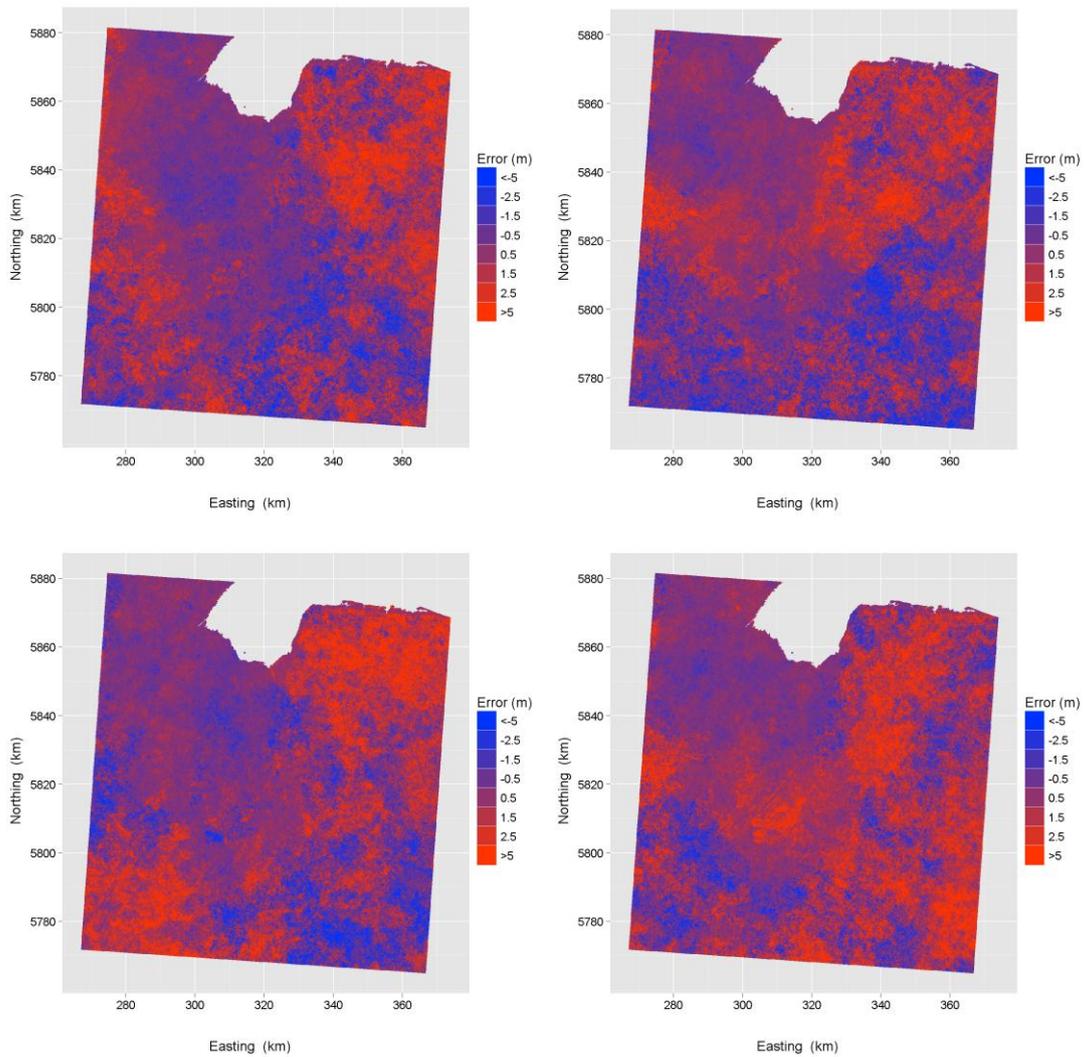


Figure 5.5. Four equiprobable simulations of DEM error.

### 5.2.3 Analysis of DEM uncertainty on the e-SOTER physiographic units

Figures 5.6 to 5.9 show the results of the uncertainty propagation analysis for the four LFAs. The LFAs with the largest differences between the dominant and the default LFA class are 'flatness' and 'slope'. Although the dominant and default class for 'flatness' correspond for 90% of the pilot area, the default flat area is much larger than the dominant flat area. Figure 5.10 (left plot) shows a detail of the low-relief area in the north-western part of the pilot area and provides an explanation for this difference. The part of the pilot that is designated as 'flat' in the default map and as 'non-flat' in the dominant map corresponds to the northern part of the study area for which the relief is somewhat more pronounced. The simulated error adds extra relief to the SRTM DEM. This has the largest effect on the northern part of the study area and makes that this area cannot be classified as flat anymore. Not surprisingly, the entropy map shows that the transition zone from relatively low to relatively high areas, indicated by the green colours in Figure 5.10 (left plot), has the largest uncertainty about the prevailing flatness class.

The default and dominant slope class map correspond for only 51%. The area with slope class 2 on the dominant map is almost five times larger than the area on the default map (class 1 equals a slope between 0 and 2%, class 2 between 2 and 5%). Again the explanation for this difference is the extra relief added to the SRTM DEM by the simulated error. The eastern and south-western parts of the pilot area are most affected by DEM uncertainty. These parts are the relatively high parts with the most rugged terrain (Figures 5.1 and 5.2). The range of simulated DEM errors in this part of the pilot area is larger than for the flatter parts (Figure 5.5). An increase in relief in this area by adding a simulated error field to the DEM will therefore have a larger effect on slope than in the flatter areas. Furthermore, a large part of the more rugged terrain has a slope between 1 and 2%, which corresponds to the upper half of slope class 1 (Figure 5.10, right plot). An increase in slope therefore results in a change in slope class more quickly than in parts with smaller slope percentages. The confusion between slope classes is largest around the boundary between the classes as indicated by the entropy map. In addition, relatively large entropy values are found in two small areas with somewhat less rugged terrain in the eastern part of the pilot area and in a small area with somewhat more rugged terrain in the large flat area.

For LFAs 'relief intensity' and 'hypsometry' DEM error has little effect on the outcome LFA class. For relief intensity the default and dominant maps correspond for 100% of the area, for hypsometry this is 96%. The entropy for relief intensity is 0 for almost the complete pilot area which means that there is no uncertainty about the prevailing class. For hypsometry the uncertainty about the prevailing elevation class is largest around the class boundaries.

When considering all four LFAs together then the correspondence between default output of the e-SOTER procedure and dominant output of the procedure on basis of simulations is 39%. Leaving LFA 'slope' out of this comparison then the correspondence is 86%.

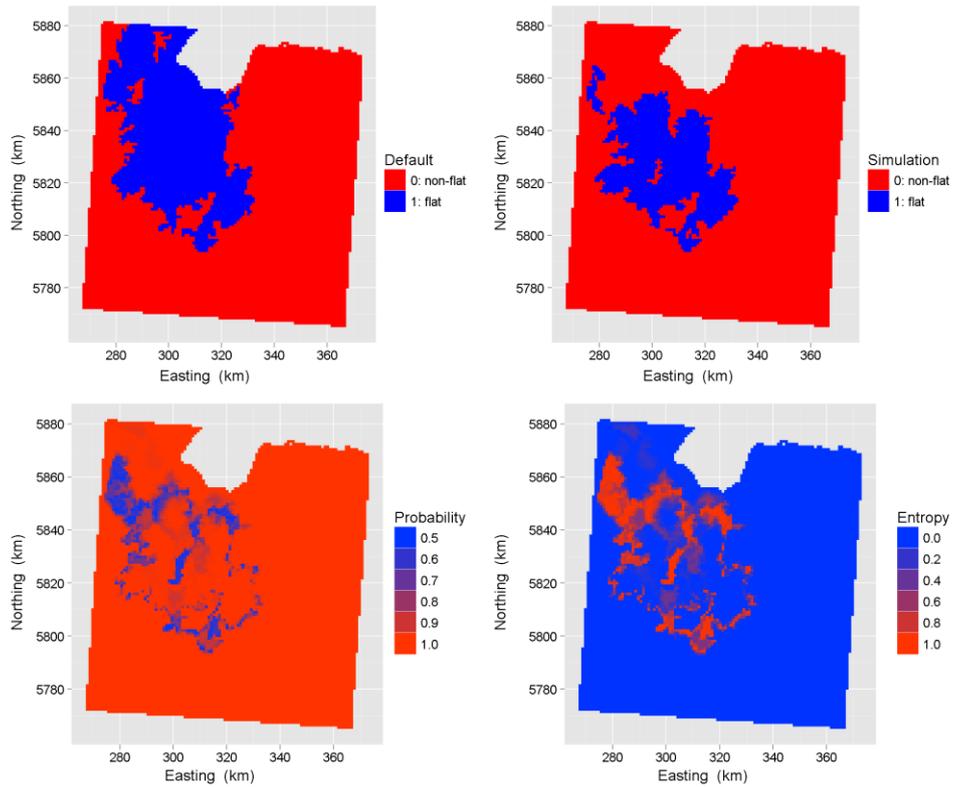


Figure 5.6. Results error propagation analysis for LFA 'Flatness'.

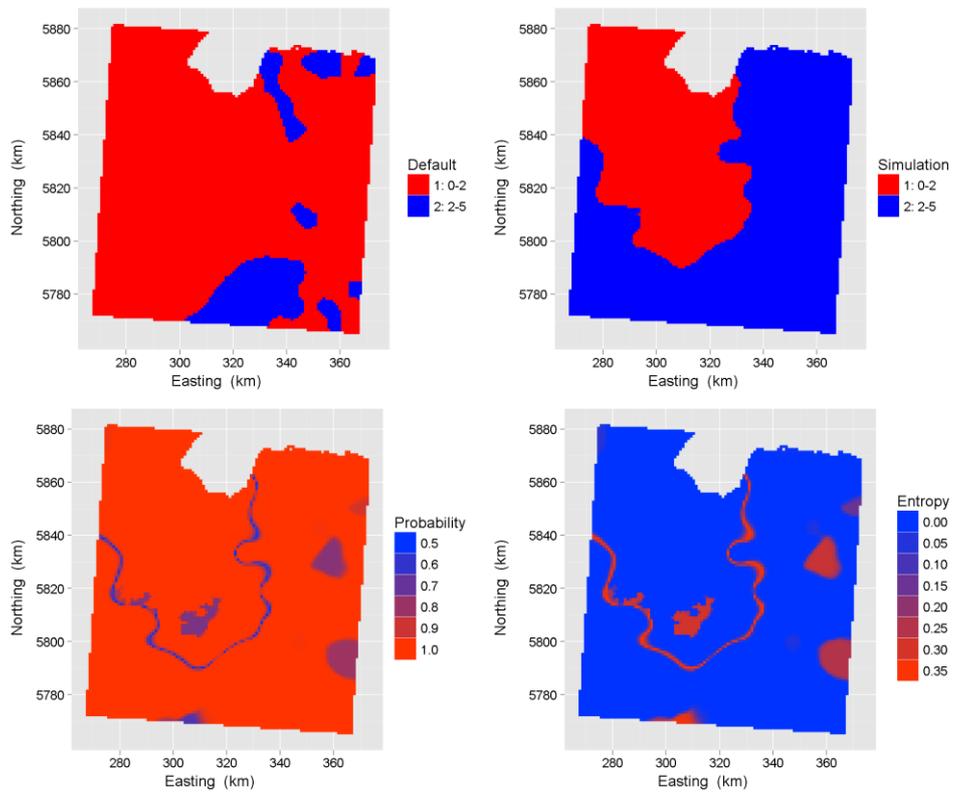


Figure 5.7. Results error propagation analysis for LFA 'Slope'.

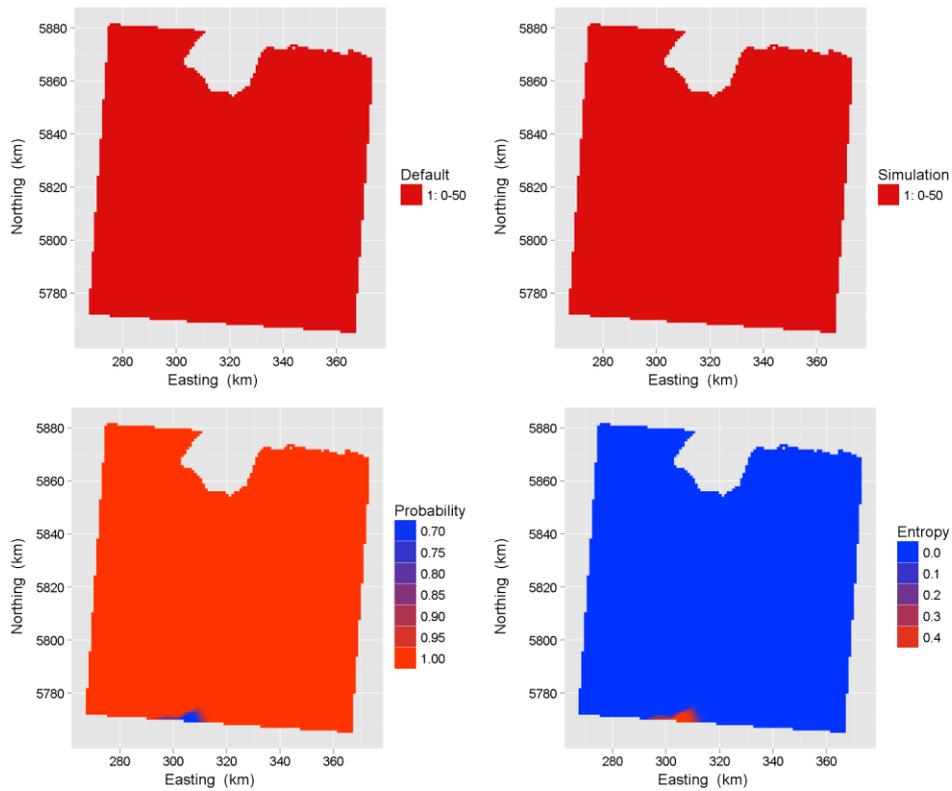


Figure 5.8. Results error propagation analysis for LFA 'Relief intensity'.

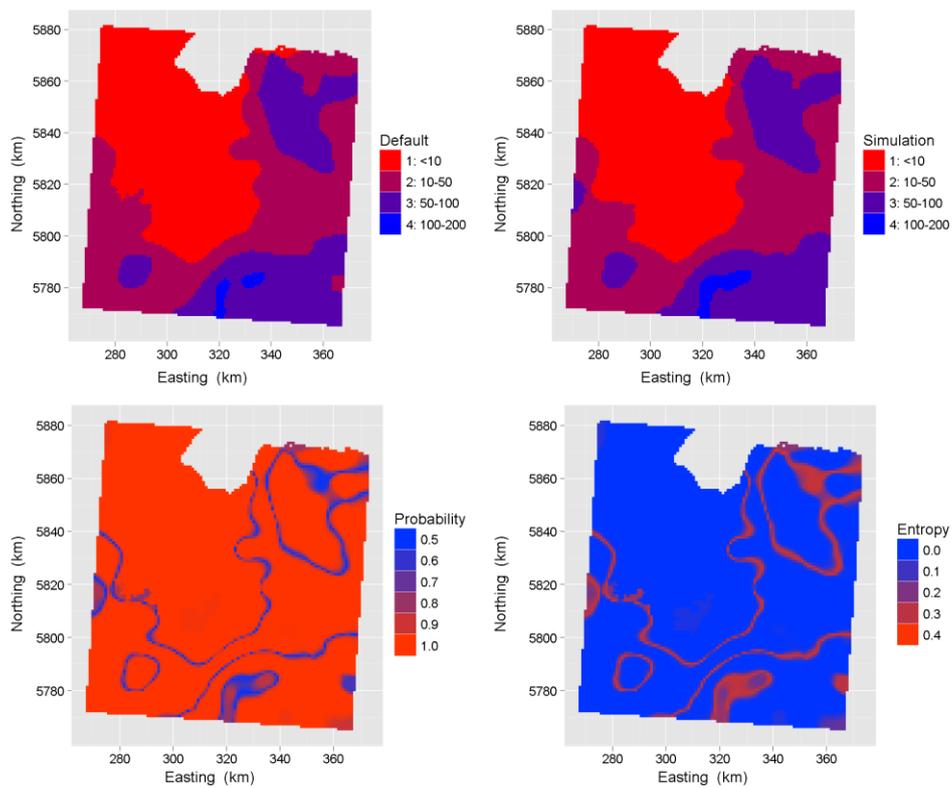


Figure 5.9. Results error propagation analysis for LFA 'Hypsometry'.

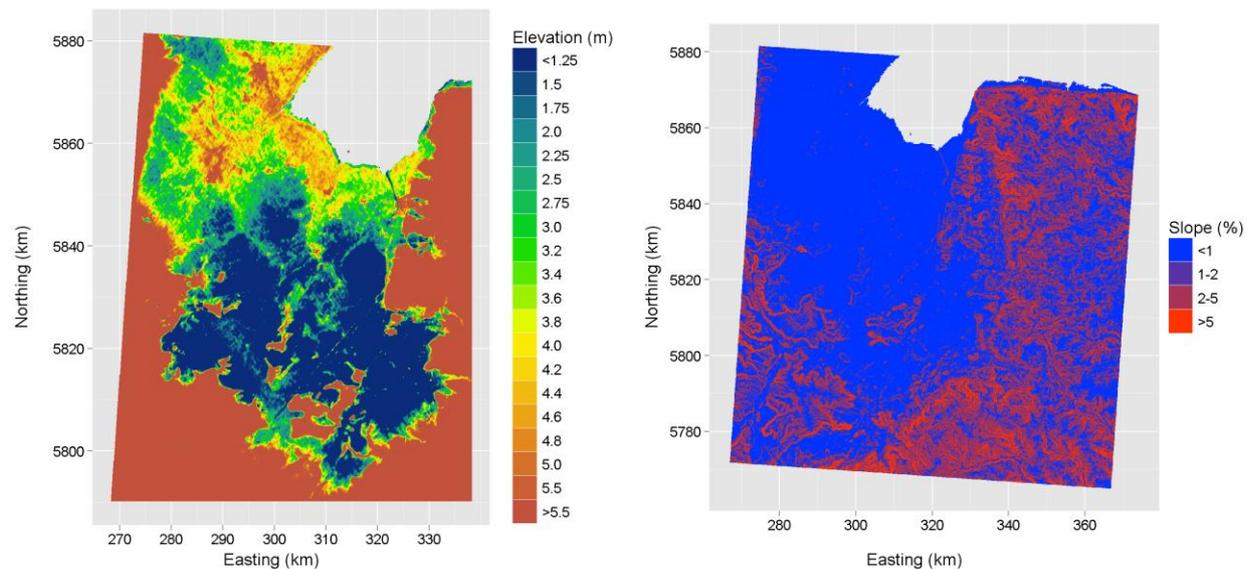


Figure 5.10. Detail of the SRTM DEM of the WE pilot area, highlighting elevation differences in the north-western part of the area (left) and the slope map derived from the SRTM DEM (right).

### 5.3 Application to Central European pilot area

#### 5.3.1 Ground control points and DEM error

Like for the WE pilot, field measured ground control points were not available to quantify DEM error. Again we used a more accurate local DEM to obtain the ‘true’ elevation at sampling points. For the Czech part of the pilot area a 1:25,000 topographical map was sampled at 330 locations. For the German part a DTM with 25-m resolution (DGM25; vertical accuracy  $\pm 1\text{-}5$  m) was randomly sampled at 5,000 locations. Of these 500 locations were randomly selected to achieve a sampling density comparable to that of the Czech part of the pilot area. Thus in total 830 control points were used for DEM uncertainty analysis in the CE pilot area. Figure 5.11 shows the SRTM DEM of the CE pilot area, the location of the control points and the DEM error at these points. The mean error is  $-2$  m, the variance  $242$  m<sup>2</sup>. The minimum and maximum absolute errors are  $-207$  and  $79$  m. Figure 5.12 (left plot) shows a ruggedness map based on a  $3 \times 3$  kernel. Like for the WE pilot the  $3 \times 3$  kernel ruggedness map had the strongest correlation with DEM error ( $r = 0.638$ ) and was therefore used to compute the standardized error  $\varepsilon_s(s)$ . Figure 5.12 (right plot) shows the scatterplot of the ruggedness against the absolute DEM error. This plot clearly shows an increasing variation in absolute error when the ruggedness increases, indicating that the assumption of constant variance of the error that underlies geostatistical modelling is indeed unlikely to be valid.

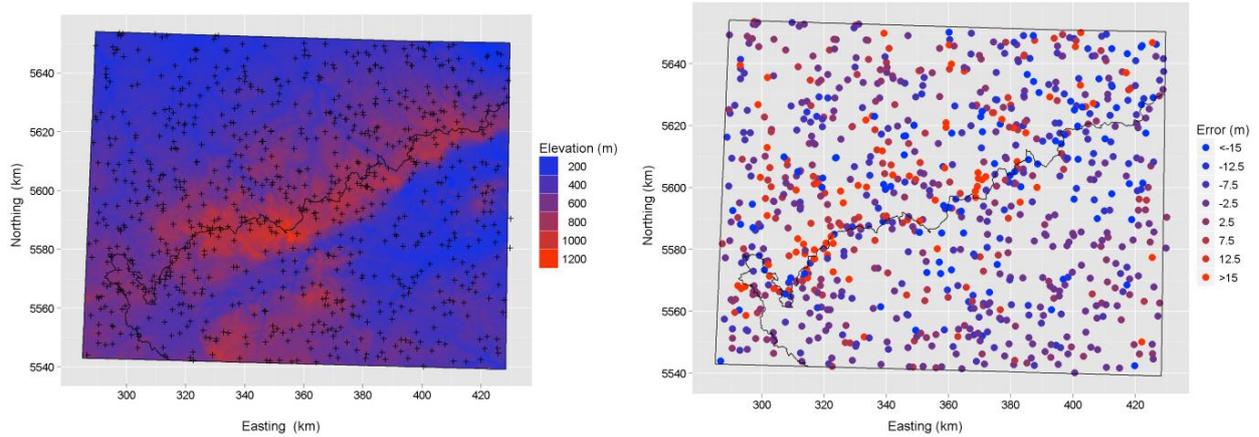


Figure 5.11. SRTM DEM with the location of the control points (left) and the DEM error at these points (right).

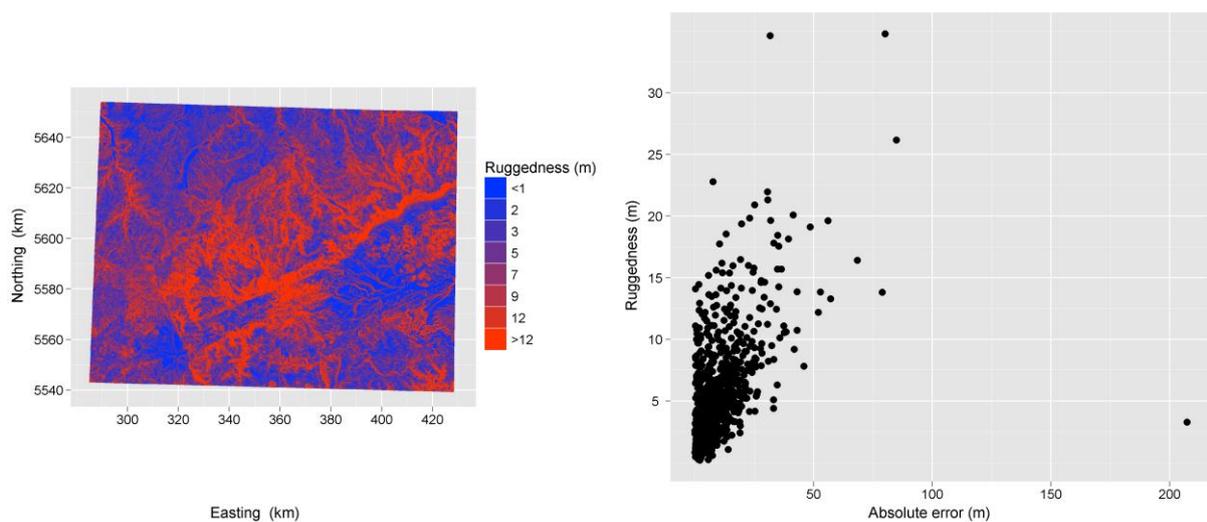


Figure 5.12. Ruggedness map created with a 3 x 3 cell kernel (left) and correlation between 3 x 3 kernel ruggedness and be absolute DEM error (right).

### 5.3.2 Modelling DEM uncertainty

Figure 5.13 shows the histogram and boxplot of the standardized error. Both plots show that the standardized errors approximately follow the normal distribution (for both countries separate and together). There are two outliers with absolute standardized errors of 23 and 63 m. These points were excluded from the dataset for semivariogram modelling. The experimental semivariogram was computed from the standardized errors at the control points. A theoretical semivariogram model was fitted from the experimental semivariogram with an exponential structure, a nugget of 3.19 m<sup>2</sup>, a partial sill of 2.67 m<sup>2</sup> and a range of 1,121 m (Figure 5.14). The semivariogram and standardized errors at the control points were used to generate 1,000 Monte Carlo simulations of DEM error of which Figure 5.15 shows four examples.

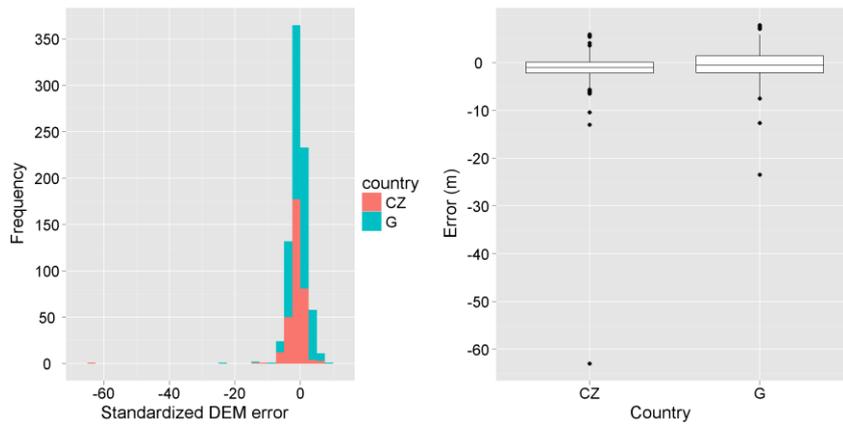


Figure 5.13. Histogram and boxplot of the standardized DEM error.

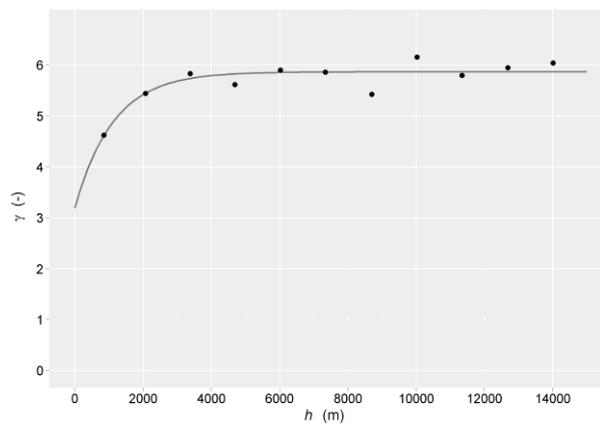


Figure 5.14. Experimental semivariogram of the standardized DEM error at the control points (dots) and fitted semivariogram model (line).

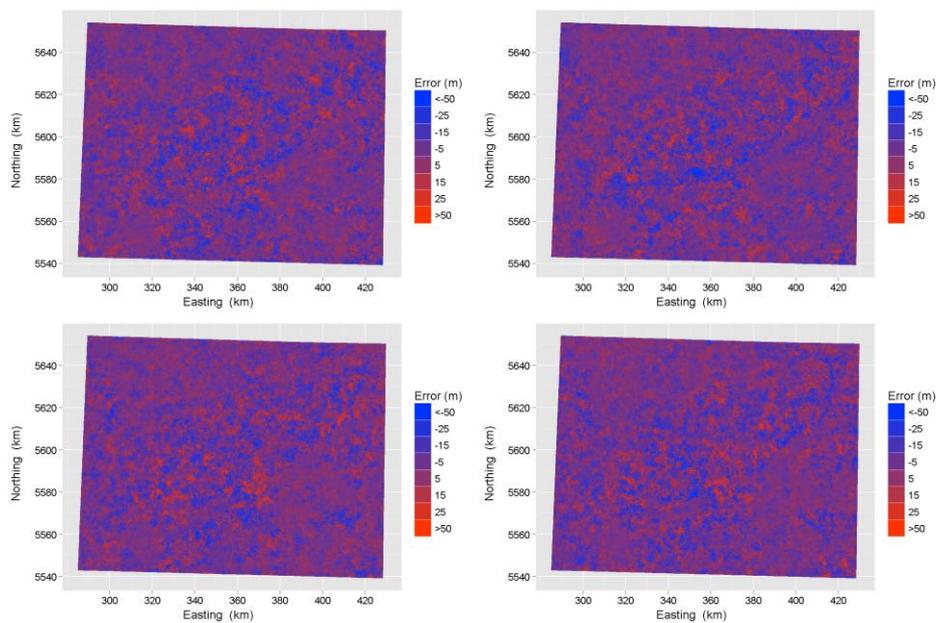


Figure 5.15. Four equiprobable simulations of DEM error.

### 5.2.3 Analysis of DEM uncertainty on the e-SOTER physiographic units

Figures 5.16 to 5.19 show the results of the uncertainty propagation analysis for the four LFAs. The LFAs with the largest differences between the dominant and the default LFA class are 'slope' and 'relief intensity'.

The default and dominant slope class map correspond for only 51%. In general the dominant map shows steeper slopes than the default map. The simulated errors (Figure 5.15) can be enhancing elevation differences substantially and thus steepening the slopes. 80% of the area with slope class 2 on the default map has slope class 3 on the dominant map. This change mainly occurred along the northern border. In the central part of the pilot area with the highest elevations (Figure 5.11) and steepest slopes (Figure 5.20) similar changes are observed. Here a similar shift in slope classes is observed. For slope class 3 the correspondence between the two maps is 61%. 17% of the default area with class 3 has class 4 as the dominant class and 21% has class 5 as the dominant class. For the default slope class 4 a trend is observed. 78% of the default area has class 5 as the dominant class on basis of the simulations. Although the dominant map gives a different representation of the slope classes in the pilot area than the default map, the spatial patterns in this map do seem to be plausible when visually compared to the slope base map derived from the SRTM DEM (Figure 5.20). Based on a visual assessment one might even argue that the dominant map gives a better representation of slope than the default map. For example, slope class 5 on the dominant map better seems to represent the extent of the steep areas (orange- and brown-coloured areas in Figure 5.20) than the default map. Like in the WE pilot confusion between slope classes mainly concentrates around the class boundaries. In addition relatively large entropy values are found in parts with somewhat less rugged terrain such as along the northern border and south of the central mountain ridge.

For LFAs 'relief intensity' the default and dominant maps correspond for 44% of the area. Only 30% of RI class 1 on the default map is depicted as class 1 on the dominant map. 68% is depicted as class 2. On the dominant map class 1 is limited to the river valleys along the northern border and the relatively flat basin-like areas south of the central mountain ridge. The area with class 3 is 3.5 times larger in the dominant map at the expense of class 2 and roughly follows the central and south-central high-elevation areas with the most rugged terrain (Figures 5.11 and 5.12). Like for slope, DEM error adds 'extra' relief to the DEM and because simulated errors can be quite substantial this can result in a change in RI class when the uncertain DEMs are used in the e-SOTER procedure. Especially so because the RI range of classes 1 and 2 is only 50 m, which is quite narrow given the relatively rugged terrain in the CE pilot area. Furthermore, the range of the semivariogram is relatively short which induces spatial variation at relatively short distances when used for geostatistical simulations. This also attributes to an increased relief intensity for the simulated DEMs (to determine the relief intensity a search radius with a 990 m diameter is used (Dobos et al., 2005)). Again the areas with large entropy follow the contours of the class boundaries.

For LFAs 'flatness' and 'hypsometry' DEM error has little effect on the outcome LFA class. For flatness the default and dominant maps correspond for 100% of the area, for hypsometry this is 96%. When considering all four LFAs together then the correspondence between default output of the e-SOTER procedure and dominant output of the procedure on basis of simulations is 28%.

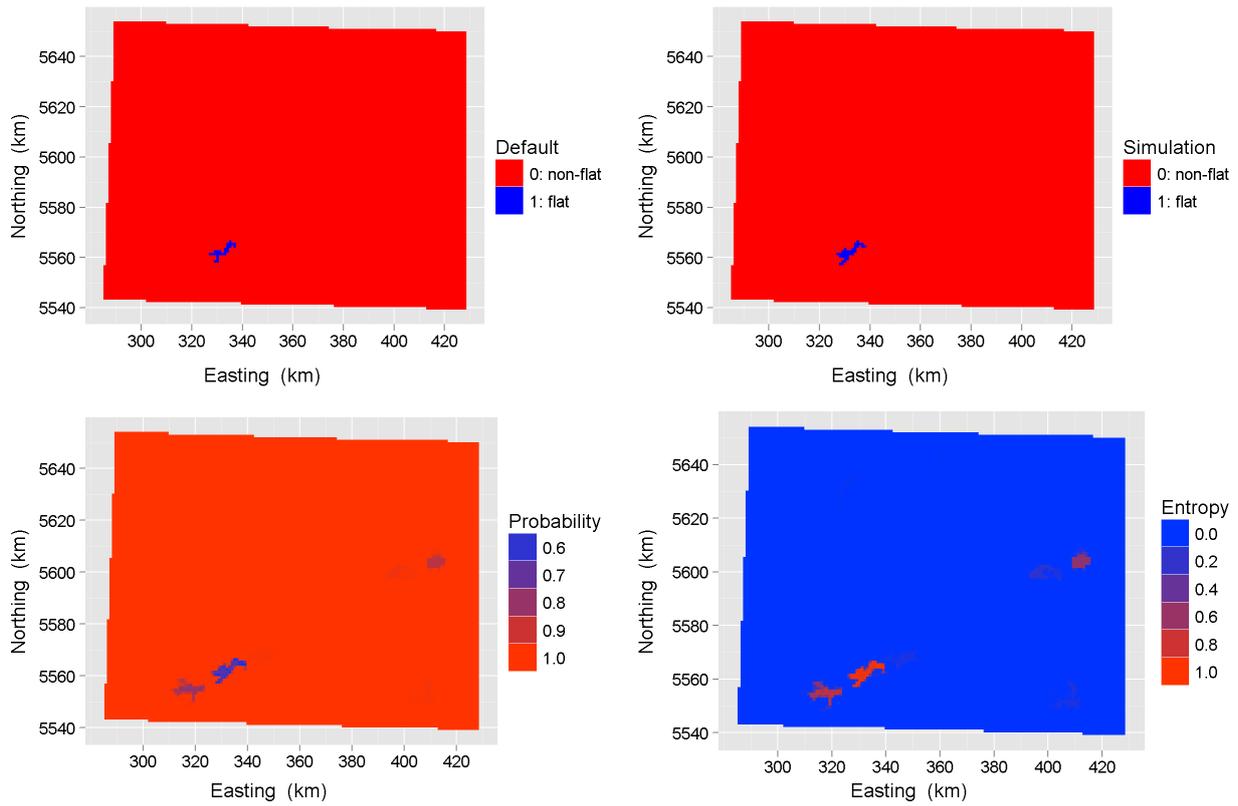


Figure 5.16. Results error propagation analysis for LFA 'Flatness'.

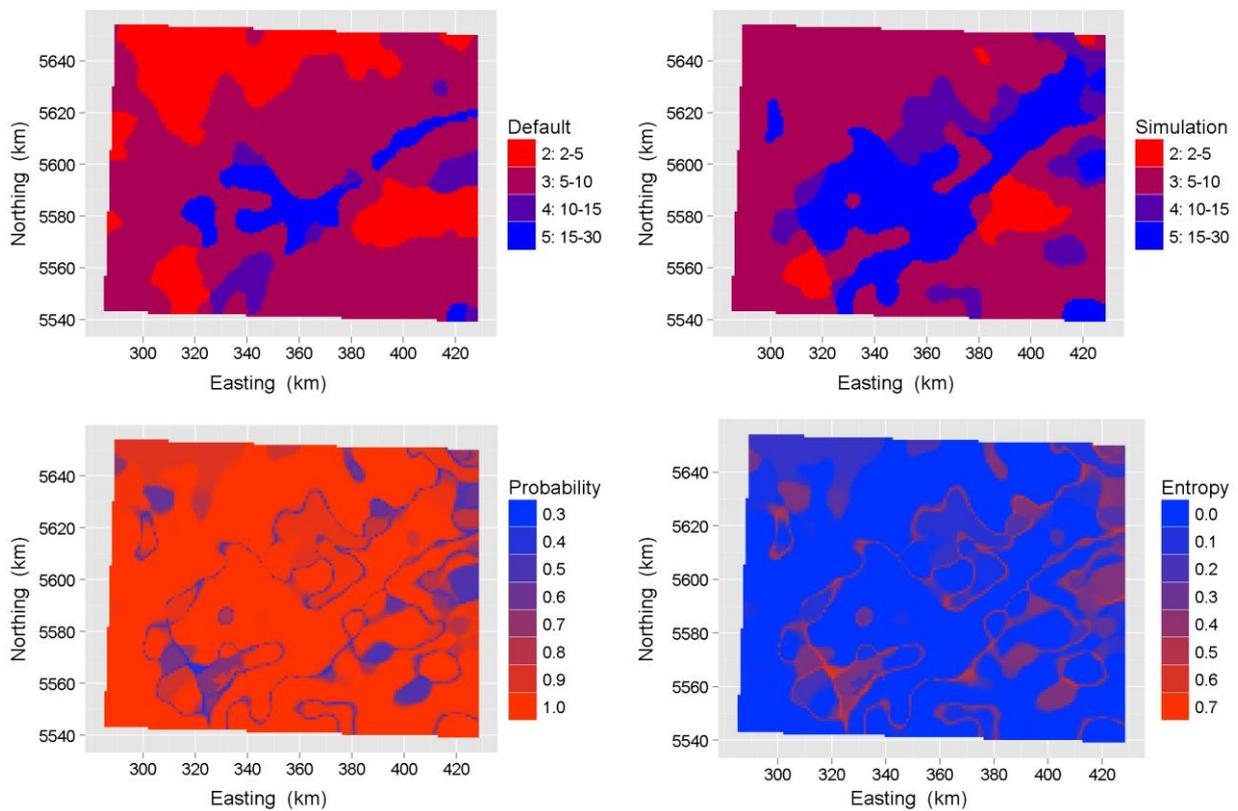


Figure 5.17. Results error propagation analysis for LFA 'slope'.

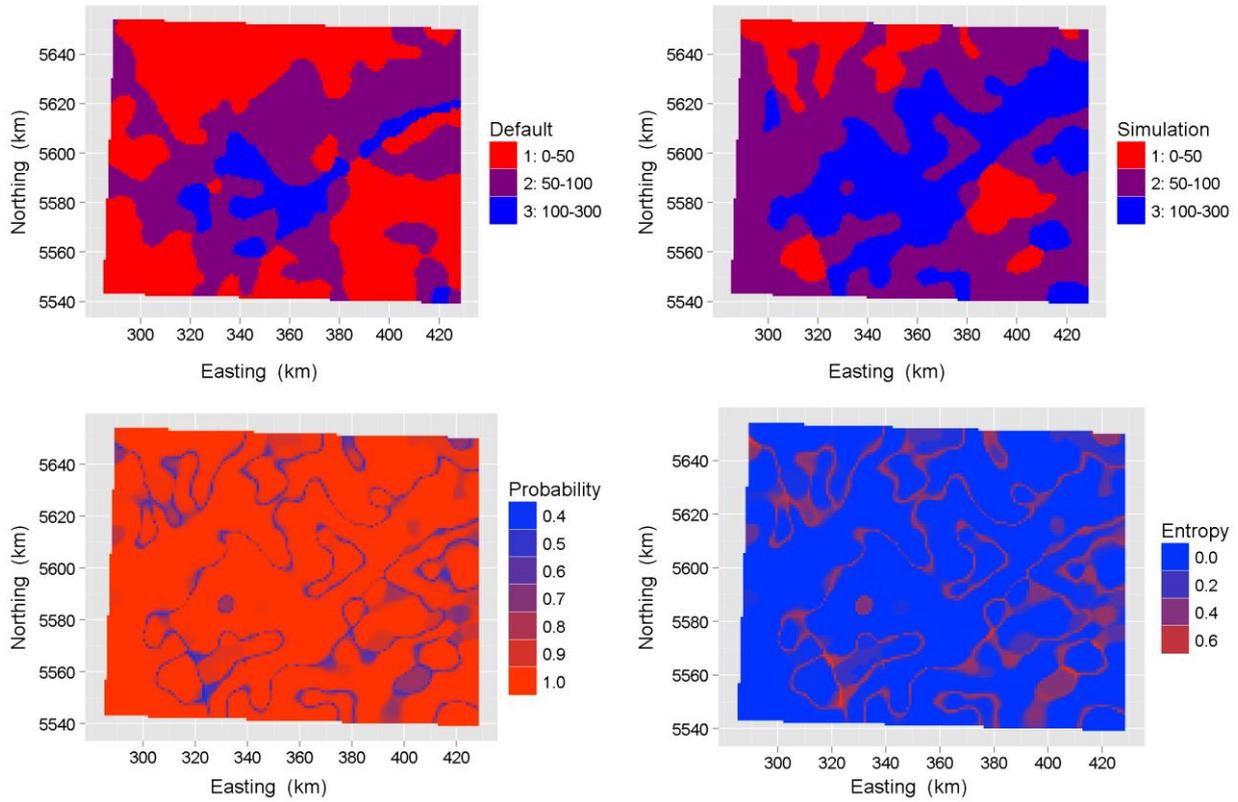


Figure 5.18. Results error propagation analysis for LFA 'Relief intensity'.

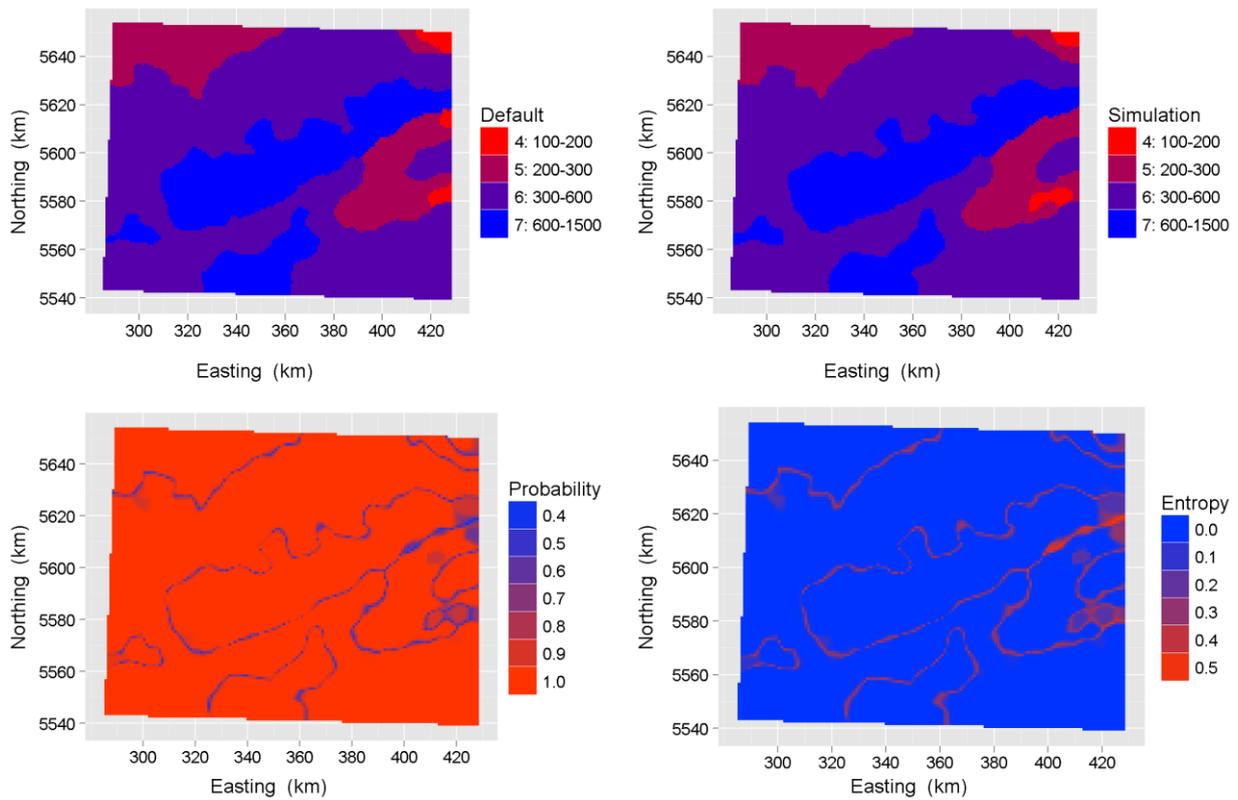


Figure 5.19. Results error propagation analysis for LFA 'Hypsometry'.

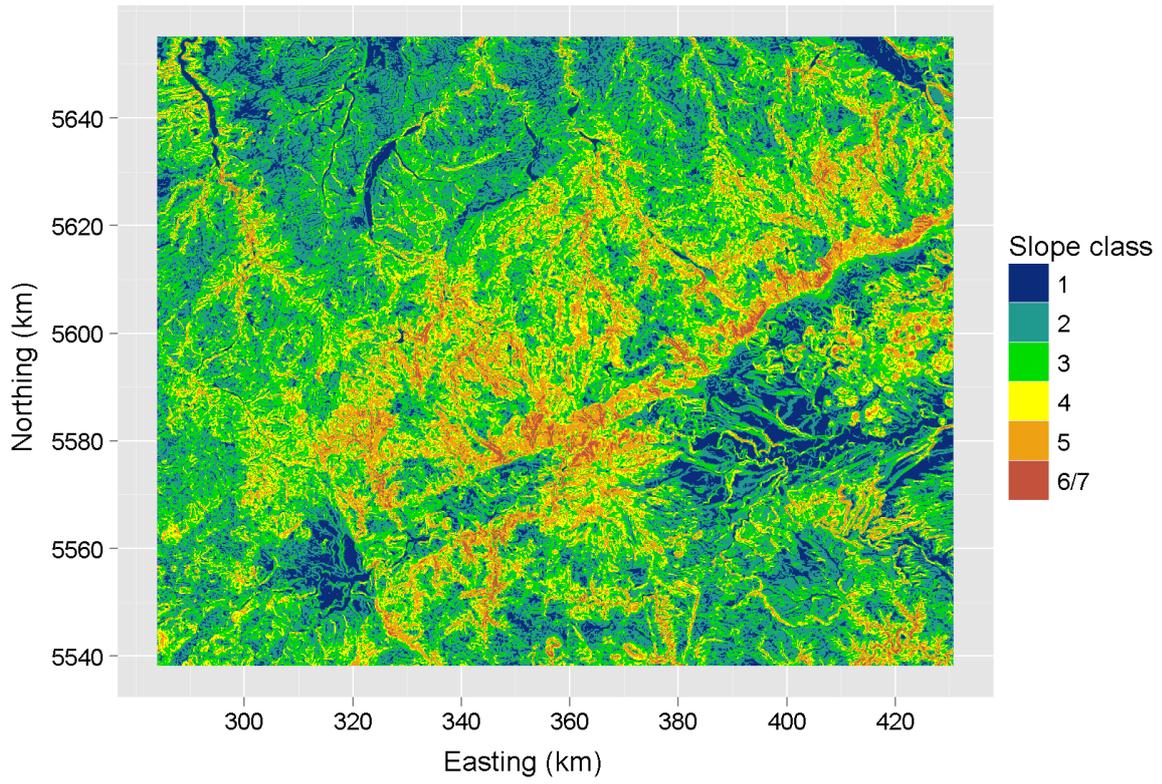


Figure 5.20. Slope map of the CE pilot area as derived from the original 90-m SRTM DEM.

## 6. Validation of the WP3 landform map

### 6.1 Methodology

In work package 3 (WP3) state-of-the-art methods for DEM analysis and soil and parent material mapping were developed to derive digital equivalents of the SOTER units for the 1:250 000 scale pilot areas. For validation we focused on the resulting landform maps since these were the only products that were available by early 2012. These maps were derived by hillshed analysis and the object-oriented approach (see Deliverable 3.1).

Validation of the WP1 landform map involved an assessment of the effect of the various aggregation and generalization steps in the e-SOTER procedure on the output (section 3.1). This is a methodological validation which cannot be directly applied to the WP3 landform maps since these were created with different methods. Therefore, validation of the WP3 maps focused on their ability to predict the spatial distribution of WRB soil groups. The hypothesis behind this assessment was that the WP3 maps are better able to predict the soil spatial distribution than the WP1 map since the WP3 maps were created with more advanced methods and at a higher level of detail. It is therefore expected that the WP3 landform units are more homogeneous in terms of soil variation than the WP1 landform units (the underlying assumption here is that soil can be predicted from terrain attributes or landforms).

Because validation involves two categorical variables (landform type and soil group) we require statistical measures that test whether there is a relationship between the two variables, and, if yes, how strong this relationship is. A well-known test is the Pearson's chi-square ( $\chi^2$ ) test of independence, which is based on analysis of the contingency table or cross tabulation (Table 3.1) (Ott and Longnecker, 2001). The null hypothesis of the  $\chi^2$ -test is independence. This means that any perceived dependence in a contingency table is attributed to chance. (We only have a sample of the population; the apparent dependence may be the result of random variation). From the contingency table the  $\chi^2$ -statistic is computed. Under the assumption that the null hypothesis (i.e. no relationship) is true, this statistic follows a  $\chi^2$ -distribution with degrees of freedom equal to the number of landform classes minus one and the number of soil group classes minus one. By comparing the statistic to the  $\chi^2$ -distribution a  $p$ -value can be computed on the basis of which the null hypothesis is rejected or accepted.

The  $\chi^2$ -test only tells us if there is an association between the two variables. It does not quantify the strength of the association. To assess the strength of an association Ott and Longnecker (2001) propose to use Goodman & Kruskal's lambda ( $\lambda$ ), which is a measure of predictability. The rationale behind this measure is that the stronger the relationship between the two variables (in our case landform and soil group), the better one variable can be predicted from the other. Goodman & Kruskal's lambda is calculated as:

$$\lambda = \frac{e_1 - e_2}{e_1}. \quad (6.1)$$

In the context of predicting soil with landform units,  $e_1$  is the number of errors one makes when there is no information about landform. In this case one predicts the most common observed soil group (the mode) in the dataset at any location;  $e_1$  thus equals the total number of observations in the dataset minus the frequency of the mode.  $e_2$  is the number of errors one makes when using landform unit as predictor of soil group and is computed as follows. For each landform unit the most common observed soil group (the mode) is used to predict the soil at any location within the landform unit. Next for each landform unit the number of errors (i.e. wrong predictions) is determined. This equals the total number of soil observations minus the landform unit mode. Finally, the total number of prediction errors is determined for the landform map. A value  $\lambda=0.05$  means that 5% fewer errors are made when a landform map is used for prediction compared to use of the modal soil group as predictor, i.e.  $\lambda$  is the proportionate reduction in errors. The  $\lambda$  measure can also be used to compare the predictive capabilities of two landform maps. In that case  $e_1$  is the number of errors one makes when predicting soil group with the WP1 landform map and  $e_2$  is the number of errors one makes when predicting soil group with the WP3 landform map.

In addition to predictability we computed the entropy for each landform unit (Equation 5.4; with base equal to the number of soil groups, such that the entropy is always a value between 0 (perfect association) and 1 (no association)). The global entropy was then computed as a weighted mean of the landform unit entropies with weights based on the number of soil sampling sites within each unit. Finally, we also computed the overall purity, that is the areal proportion for which the soil group is correctly predicted by the dominant soil groups of the landform units.

The WP3 maps were validated for the UK part of the western European window with soil data from 2,180 sampling sites (Figure 4.2, top). For the UK two WP3 terrain maps were available: one created with the object-oriented approach and one created with hillshed analysis. The landform classification system used for both methods is the hierarchical Hammond's landform classification (see Deliverable D3.1). The object-oriented map has three legend levels: code (4 entries), class (9) and subclass (18). In addition to these levels with 3, 5, and 7 entries, respectively, the hillshed map has a fourth level that combines subclass with slope position (21 entries). The WP1 map has 21 unique combinations of the four landform attributes (see also chapter 3).

## 6.2 Results

Table 6.1 shows the contingency tables of landform unit and WRB soil group for the WP1 map, Table 6.2 for the slope and hypsometry attributes of the WP1 map, Table 6.2 for the WP3 object map and Table 6.3 for the WP3 hillshed maps. The table entries are the percentage of observations of each soil group within a landform unit (i.e. the percentages are conditioned on the row totals).

For WP1 most landform units are dominated by either Cambisols or Hydromorphic soils (Table 6.1). Luvisols dominate two units. These three soil groups make up 86% of the UK soil dataset. Other soil groups cover only minor parts of the landform units and are never the dominant soil. Hydromorphic soils dominate the relatively flat areas (slope class 1; 46%) followed by Cambisols (22.5%). The gently sloping areas are also dominated by Hydromorphic soils (32%), closely followed by Cambisols (31%). In areas with moderate slope (class 3) Cambisols dominate (32%), closely followed by Luvisols (31%).

Generally, soil distribution shifts from Hydromorphic soils and Cambisols to Cambisols and Luvisols with increasing slope class.

For hypsometry we observe a similar trend. Hydromorphic soils dominate areas with low elevation while higher areas are dominated by Cambisols, although Luvisols and Hydromorphic soils are also frequently found. Not surprisingly, the area within a landform unit that is covered with the shallow leptosols increases with increasing slope and hypsometry class.

Considering slope and hypsometry together, Hydromorphic soils typically dominate areas with low elevation (hypsometry classes 1,2) and relatively flat areas (slope class 1) while Cambisols dominate the high, flat areas. In gently sloping (slope class 2) and moderately sloping (class 3) terrain, Cambisols typically dominate the high areas, while Hydromorphic soils dominate the low areas. Luvisols are the dominant soil group in two landform units.

Table 6.1. Contingency table of WP1 landform unit and WRB soil group. The table shows percentages that are conditioned on the number of observations within each landform unit. The row totals indicates the number of soil observations within the landform unit, the column totals the number of soil observations. The mode, i.e. most common soil group, of the unit is indicated in bold type.

Landform <sup>a</sup>	WRB reference soil group									Total
	anthrosol	arenosol	cambisol	histosol	hydromorphic	leptosol	luvisol	podzol	regosol	
0111	0.6	1.1	16.2	2.8	<b>67</b>	0.6	7.3	1.7	2.8	179
1111	0	0	4.1	13.5	<b>78.4</b>	0	1.4	1.4	1.4	74
1112	0	0	0	0	<b>50</b>	0	50	0	0	2
0112	2.4	10.9	25.5	0.8	<b>34.8</b>	0.8	20.2	1.9	2.7	376
1113	0	0	<b>100</b>	0	0	0	0	0	0	1
0113	0.9	8.3	<b>35.2</b>	0.9	27.8	0	26.9	0	0	108
0114	0	0	0	0	<b>100</b>	0	0	0	0	2
0212	1.1	5.6	17.4	0	<b>44.4</b>	1.1	25.8	3.9	0.6	178
0213	0.3	2.9	<b>33.2</b>	0.3	30.1	3.2	23.4	2.6	3.9	585
0214	0	1.2	<b>33.1</b>	0.2	30.2	8	23.1	0.7	3.4	411
0312	0	4.8	9.5	4.8	<b>57.1</b>	0	23.8	0	0	21
0322	0	0	<b>50</b>	0	<b>50</b>	0	0	0	0	4
0313	0	0	<b>33.3</b>	0	30.8	5.1	25.6	5.1	0	39
0323	0	8.3	25	0	0	25	<b>41.7</b>	0	0	12
0324	0	4.4	<b>40</b>	0	11.1	6.7	35.6	0	2.2	45
0314	0	0.9	31.3	0	16.5	13.9	<b>34.8</b>	2.6	0	115
0315	0	0	<b>50</b>	0	0	33.3	16.7	0	0	6
0325	0	0	<b>46.7</b>	0	26.7	13.3	13.3	0	0	15
0326	0	0	0	0	<b>100</b>	0	0	0	0	2
0423	0	0	<b>100</b>	0	0	0	0	0	0	2
0424	0	0	<b>66.7</b>	0	0	33.3	0	0	0	3
Total	15	89	616	23	777	87	477	41	55	2180

<sup>a</sup> The four digits indicate the flatness class, slope class, relief intensity class and hypsometry class. The landform units are ordered by slope and then by hypsometry.

Table 6.2. Contingency table of WP1 landform unit and the WP1 landform attributes slope and hypsometry. The table shows percentages that are conditioned on the number of observations within each landform unit. The row totals indicates the number of soil observations within the landform unit, the column totals the number of soil observations. The mode, i.e. most common soil group, of the LFA class is indicated in bold type.

LFA	WRB reference soil group									Total
	anthrosol	arenosol	cambisol	histosol	hydromorphic	leptosol	luvisol	podzol	regosol	
<i>Slope</i>										
1	1.5	7	22.5	2.6	<b>46.1</b>	0.5	16.2	1.5	2.2	742
2	0.3	2.7	30.7	0.3	<b>32.3</b>	4.6	23.7	2.1	3.2	1174
3	0	1.9	<b>32.4</b>	0.4	21.6	10.8	30.5	1.9	0.4	259
4	0	0	<b>80</b>	0	0	20	0	0	0	5
<i>Hyps.</i>										
1	0.4	0.8	12.6	5.9	<b>70.4</b>	0.4	5.5	1.6	2.4	253
2	1.9	9	22.5	0.7	<b>38.7</b>	0.9	22	2.4	1.9	581
3	0.4	3.6	<b>33.6</b>	0.4	29.2	3.2	24.2	2.3	3.1	747
4	0	1.4	<b>33.3</b>	0.2	26	9.2	26.2	1	2.6	576
5	0	0	<b>47.6</b>	0	19	19	14.3	0	0	21
6	0	0	0	0	<b>100</b>	0	0	0	0	2
Total	15	89	616	23	777	87	477	41	55	2180

Results for the WP3 object-oriented approach (Table 6.3) show similar relationships between landform unit and soil group as the WP1 approach. Hydromorphic soils dominate the low-lying, relatively flat areas (A\*a; see Appendix 2). Cambisols dominate gently sloping areas of the plains (PLA) and the more undulating parts of plains with high hills (PHH). The WP3 object landform units seem somewhat better able to identify Luvisols. These are the dominant soils in four, however minor, landform units. Again the area covered with leptosols increases as the ruggedness of the terrain increases (A4b, B4a, B4b).

The WP3-hillshed approach shows relationships between landform unit and soil group that are similar to the WP3 object and WP1 approaches. Hydromorphic soils clearly dominate the flat areas (A\*a). These soils cover between 70 and 85% of the landform units. The hillshed approach seems to capture the Hydromorphic soils better than the object-oriented approach. Cambisols dominate the higher, upslope parts of the plains with hills and mountains (PHM) and the tablelands with considerable relief (TCR) although Luvisols and Hydromorphic soils also cover substantial parts of the landform units dominated by Cambisols. Luvisols dominate only three minor landform units. Again the area covered with leptosols increases as terrain ruggedness increases which is clearly shown for the tablelands units.

Table 6.3. Contingency table of WP3 landform unit created with the object-oriented approach and WRB soil group. The table shows percentages that are conditioned on the number of observations within each landform unit. The row totals indicates the number of soil observations within the landform unit, the column totals the number of soil observations. The mode, i.e. most common soil group, of the unit is indicated in bold type. Colors indicate the hierarchical legend entries of the landform classification system.

Code <sup>a</sup>	Landform		WRB reference soil group									Total
	Class	Subclass	AN	AR	CM	HS	HY	LP	LV	PZ	RG	
PLA	PF	A1a	0.7	0	3.3	8.6	<b>82.9</b>	0	2.6	0.7	1.3	152
PLA	PSL	A2a	1	7.4	24.8	1	<b>39.9</b>	0.7	18.5	2.7	4	298
PLA	PSL	A2b	1.1	8.1	<b>33.1</b>	0.4	<b>33.1</b>	0.4	20.8	1.4	1.8	284
PLA	PSL	A2c	3.2	14.4	<b>35.2</b>	1.6	22.4	0.8	19.2	1.6	1.6	125
PHM	PH	A3a	0.8	2.5	26.6	0.6	<b>33</b>	4.1	26	3.7	2.7	485
PHM	PH	A3b	0	6.1	16.3	0	30.6	2	<b>44.9</b>	0	0	49
PHM	PHH	A4a	0	2.9	26.1	0	<b>47.8</b>	1.4	18.1	1.4	2.2	138
PHM	PHH	A4b	0	2.5	22.2	0	19.8	7.4	<b>43.2</b>	0	4.9	81
PHM	PHH	B4a	0	1.9	<b>33.5</b>	0	19.3	15.5	28	1.9	0	161
PHM	PHH	B4b	0	1.4	<b>40.8</b>	0	9.9	23.9	22.5	0	1.4	71
PHM	PLM	A5a	0	0	<b>39.5</b>	0.7	<b>39.5</b>	1.3	13.8	0.7	4.6	152
PHM	PLM	B5a	0	0	21.7	0	<b>43.5</b>	4.3	21.7	8.7	0	23
TAB	TMR	A3c	0	2.1	27.1	0	<b>35.4</b>	4.2	31.2	0	0	48
TAB	TCR	A4c	0	0	<b>42.9</b>	0	31	6	14.3	0	6	84
TAB	TCR	B4c	0	0	40	0	0	0	<b>53.3</b>	0	6.7	15
OPM	OHH	C4a	0	0	0	0	<b>100</b>	0	0	0	0	1
OPM	OHH	C4b	0	0	22.2	0	11.1	22.2	<b>44.4</b>	0	0	9
OPM	OLM	C5a	0	0	<b>75</b>	0	0	0	25	0	0	4
Total			15	89	616	23	777	87	477	41	55	2180

<sup>a</sup> Appendix 2 gives a description of the landform classification levels of the WP3 maps.

Table 6.5 shows the validation measures. The  $\chi^2$ -statistic shows that for each landform map there is very strong evidence that soil group depends on the landform unit and that the evidence increases with a more detailed classification legend. The predictability  $\lambda$  shows that predicting soil with landform all maps give better results than simply predicting the most common (modal) soil class (Hydromorphic soils,  $n=777$ ; the default method) at each location. Using the WP1 map for predicting soil 6.3% fewer errors are made than predicting with the modal soil class. For WP3-object this is 7.9% for the most detailed classification level and for WP3-hillshed it is 9.1%.

The WP3 object-oriented approach differentiates soils less well than the WP1 map at the code and class levels. More errors are made when predicting soil (negative values for  $\lambda$ -WP1), there is more confusion within the landform units (entropy is larger) and the purity is smaller. Only at subclass level, for which the number of landform units is comparable to the number of landform units distinguished on the WP1 map, the WP3 object oriented approach performs somewhat better than the WP1 approach.

The WP3 hillshed approach predicts the spatial distribution of WRB soil groups as well as the WP1 approach at the code, class and subclass levels. Only the entropy is somewhat larger. The results at code level are especially striking since only three landform units are distinguished (plains with hills and mountains, tablelands, open hill and mountains) that differentiate soils as well as the 21 units of the WP1 map. At the subclass-slope level the WP3 hillshed method predicts the spatial distribution of soils better than the WP1 map: 3.4% fewer errors are made and the purity is 2% larger. Only the entropy is similar. This means that there is a similar amount of confusion within the landform units.

Summarizing, validation of the WP3 landform maps based on predictability of the WRB reference soil groups indicates that the hillshed analysis gives the best overall results. Both the hillshed and object-oriented approach give better results than the WP1 map at subclass level, although differences in predictability and purity are modest. The entropy is equal for all three maps, indicating that the WP3 landform units are not internally more homogeneous with respect to soil distribution than the WP1 units.

Table 6.4. Contingency table of WP3 landform unit created with hill shed analysis and WRB soil group. The table shows percentages that are conditioned on the number of observations within each landform unit. The row totals indicates the number of soil observations within the landform unit, the column totals the number of soil observations. The mode, i.e. most common soil group, of the unit is indicated in bold type. Colors indicate the hierarchical legend entries of the landform classification system.

Code <sup>a</sup>	Landform			WRB reference soil group									Total
	Class	Subclass	Subclass_slope	AN	AR	CM	HS	HY	LP	LV	PZ	RG	
PHM	PH	A3a	A3a_lower	0.8	0	9.3	6.2	<b>72.1</b>	0	7.8	2.3	1.6	129
PHM	PH	A3a	A3a_mid	0	1.9	12.5	5.8	<b>71.2</b>	1	5.8	0	1.9	104
PHM	PH	A3a	A3a_upper	4.3	0	8.7	0	<b>82.6</b>	0	0	0	4.3	23
PHM	PH	A3b	A3b_lower	1.9	7.2	19.8	1.9	<b>42</b>	1	19.8	2.9	3.4	207
PHM	PH	A3b	A3b_mid	1.7	13.1	26.9	0.6	<b>34.3</b>	1.1	18.9	2.3	1.1	175
PHM	PH	A3b	A3b_upper	0	10	17.1	0	<b>38.6</b>	0	28.6	4.3	1.4	70
PHM	PHH	A4b	A4b_lower	0	2.9	25.1	1.4	<b>43</b>	1.1	20.1	2.2	4.3	279
PHM	PHH	A4b	A4b_mid	1.2	4.1	<b>32.2</b>	0	27.2	3.8	26.4	1.4	3.8	345
PHM	PHH	A4b	A4b_upper	1.9	2.9	<b>42.7</b>	0	26.2	0	23.3	1	1.9	103
TAB	TCR	A4c	A4c_lower	0	1.9	<b>38.7</b>	0	29.2	5.7	17	1.9	5.7	106
TAB	TCR	A4c	A4c_mid	0	1.5	<b>35.4</b>	0	30.8	7.3	23.1	0.4	1.5	260
TAB	TCR	A4c	A4c_upper	0	2	32.3	0	21.2	7.1	<b>33.3</b>	4	0	99
TAB	TCR	B4c	B4c_lower	0	2.8	<b>38.9</b>	0	27.8	11.1	19.4	0	0	36
TAB	TCR	B4c	B4c_mid	0	2.9	<b>37.5</b>	0	8.7	12.5	36.5	0	1.9	104
TAB	TCR	B4c	B4c_upper	0	3.8	<b>33.3</b>	0	24.4	5.1	28.2	5.1	0	78
TAB	THR	B5c	B5c_lower	0	0	0	0	<b>100</b>	0	0	0	0	1
TAB	THR	B5c	B5c_mid	0	0	<b>42.9</b>	0	14.3	28.6	14.3	0	0	7
TAB	THR	B5c	B5c_upper	0	0	<b>25</b>	0	25	0	<b>25</b>	0	25	4
OPM	OLM	C5c	C5c_lower	0	33.3	<b>66.7</b>	0	0	0	0	0	0	3
OPM	OLM	C5c	C5c_mid	0	0	<b>34.8</b>	0	8.7	26.1	21.7	8.7	0	23
OPM	OLM	C5c	C5c_upper	0	4.2	25	0	4.2	20.8	<b>45.8</b>	0	0	24
Total				15	89	616	23	777	87	477	41	55	2180

<sup>a</sup> Appendix 2 gives a description of the landform classification levels of the WP3 maps.

Table 6.5. Validation measures.

Landform map	No. units	$\lambda$ -WP1 <sup>a</sup>	$\lambda$ -mode <sup>b</sup>	$\chi^2$ ( $p$ )	Entropy	Purity
<i>None<sup>c</sup></i>						
WP1	21	-	0.063	619 (<0.000)	0.65	0.397
<i>WP3-object</i>						
Code	4	-0.056	0.011	165 (<0.000)	0.69	0.363
Class	9	-0.039	0.027	493 (<0.000)	0.66	0.374
Subclass	18	0.167	0.079	703 (<0.000)	0.64	0.407
<i>WP3-hill shed</i>						
Code	3	-0.001	0.063	207 (<0.000)	0.69	0.397
Class	5	-0.001	0.063	307 (<0.000)	0.68	0.397
Subclass	7	-0.001	0.063	464 (<0.000)	0.67	0.397
Subclass_slope	21	0.034	0.091	628 (<0.000)	0.65	0.416

<sup>a</sup> Proportionate reduction in errors compared to the WP1 map ( $[(WP1-WP3)/WP1]$ ). Negative numbers indicate better prediction by the WP1 map.

<sup>b</sup> Proportionate reduction in errors compared to predicting the modal soil class ( $[(mode-WP3)/modal]$ ), which are the Hydromorphic soils.

<sup>c</sup> Predicting the dominant observed soil group (Hydromorphic soils,  $n=777$ ) at each location.

This does not mean, however, that the merit of using more advanced methods for landform classification is limited. The ability to predict soil is only one aspect of landform classification. For example, a visual inspection of landform units shows that the WP1 landform map captures the major variation in terrain in the UK area of the western European window although delineations are quite crude (Fig 6.1, top). The WP3-object map shows a much more detailed delineation of the landscape (Fig 6.2, bottom) than the WP1 map. We therefore expect that the landform units of the WP3 maps are more homogeneous in terms of landform attributes than the WP1 map.

The three dominant soil groups in the UK, the Hydromorphic soils, Cambisols and Luvisols, can occur under a great variety of conditions and are not typical or unique for certain landform units. This explains that these three soils cover considerable areas within almost all landform units, although the dominant soil group may differ. Furthermore, we have seen in section 4.2.2 that the correlation between the soil classification system for England and Wales and WRB is not error-free and that there is considerable confusion between the three dominant soil groups.

Finally, parent material, another important soil forming factor in addition to terrain, is not considered here. We expect that the more detailed delineations of the landform units will result in a more detailed differentiation of parent materials. This might improve predictability of the soil groups. For example, Podzols and Arenosols (sand) and Histosols (organic) might be better differentiated from the other soil groups.

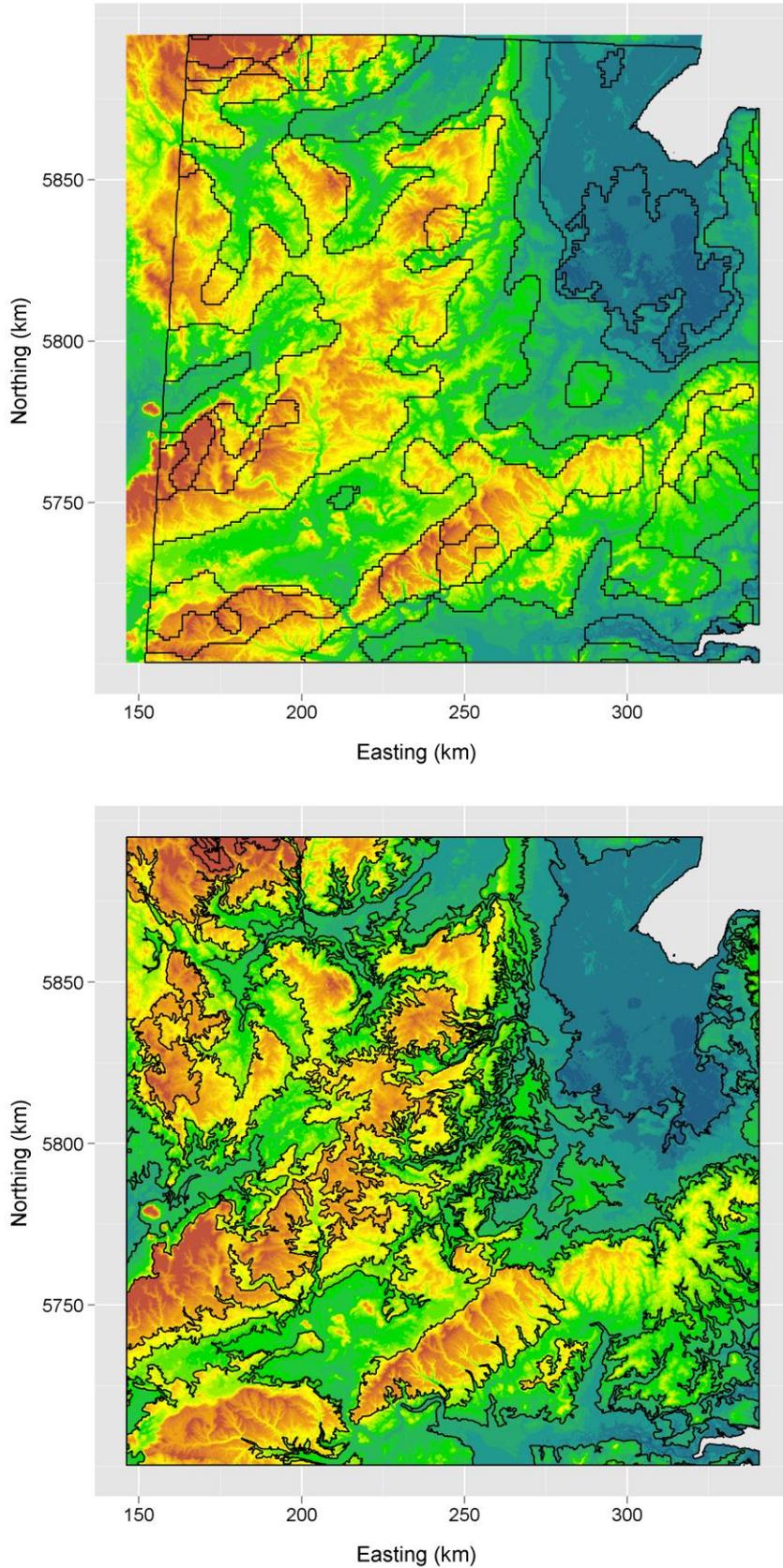


Figure 6.1. Landform classification based on the SRTM DEM by the WP1 procedure (top) and WP3-object oriented approach (bottom) for the central UK area.

## 7. Conclusions and recommendations

### *Landform validation*

- Validation generally shows large purities for both the Western (WE) and Central European (CE) windows. This means that the negative effect of the generalization steps in the e-SOTER landform classification procedure developed in WP1 on the mapped landform attributes is limited.
- Validation results are comparable for WE and CE windows.
- Slope is the landform attribute (LFA) that is most affected by the generalization steps. Overall purity of slope for WE is 45% and for CE 51%. Map units purities and class representations roughly vary between 20 and 80%. The fact that slope is the most heavily affected can be explained by the fragmented appearance of slope classes at 90-m resolution. Slope classes are clustered in small 'islands'. Generalizing such a fragmented map induces more errors in the outcome than generalizing a 90-m LFA map with spatially more contiguous classes.
- Based on the validation results for slope — relatively low purities and large variation within mapped slope classes — we conclude that a slope legend with seven entries might be too detailed for use at scale 1:1M. To reduce the effect of generalization steps and to improve the representation of slope classes we suggest to develop a slope legend with fewer and more broadly defined classes.
- Hypsometry is the least affected LFA. Overall purity, map unit purities and class representations generally vary between 70 and 90%. These results can be explained by the fact that elevation classes cover relatively large, contiguous areas so that the effect of the generalization steps on the outcome classes is small.
- Relief is an important factor that determines the effect of the generalization steps on the PU map. The WE map (less pronounced relief, larger contiguous flat areas) better represents the flat areas than the CE map, while the latter (more pronounced relief, larger contiguous RI classes) better represents relief intensity.

### *Soil validation*

- The 51% purity for the UK validation area, based on validation of the dominant soil group of a soil component, is fairly large, especially for a map created with digital soil mapping methods. The e-SOTER soil map shows the general soil spatial patterns, which is the purpose of a 1:1M soil map.
- Despite an acceptable validation result, Figure 4.1 clearly shows large differences between the 1:1M e-SOTER map and the 1:250,000 soil map of England and Wales. Important error sources in the UK area are the over-representation of Histosols and Podzols on the e-SOTER map and the absence of Leptsols as dominant soil group. We recommend to focus on these groups to improve soil mapping. A more 'local' calibration of the soil predictive relationships derived from remote sensing images might improve mapping.

- Overall purity for the German/Czech validation area is a moderate 32%. The difference with the UK area can be partly explained by stricter validation criteria.
- Important error sources in the German/Czech validation area are the under-representation of Chernozems and Podzols on the e-SOTER soil map. We therefore recommend to focus on these groups to improve soil mapping. Another main error source is the confusion between Hydromorphic soils, Cambisols and Luvisols. Improving the spatial representation of these soils, however, might be difficult at 1:1M scale since these soils often occur in associations, forming complex mosaic soil covers.
- The stringent and flexible mode represent two 'extremes'. The stringent case is very strict. For interpretation of the e-SOTER soil map, however, soil data users will have to choose a soil component from the database to assign to each e-SOTER unit and then the dominant soil type is the most logical choice. The flexible case is the other extreme. Overall purity of the flexible case is much larger than that of the stringent case but is also less informative since the larger the number of soil components associated to an e-SOTER unit the less information one has about the actual soil component.
- Generalizing the soil legend from 17 to 12 entries based on diagnostic properties did not improve validation results.
- The soil maps were validated with legacy soil data obtained by purposive sampling. However, when possible we recommend to use validation data obtained by independent probability sampling. Only then can unbiased estimates of the map quality measures and their standard errors be obtained.

#### *Error propagation analysis*

- In the UK pilot area DEM error has the largest effect on slope class. The dominant slope on the basis of 1,000 simulations is typically one class higher than the default classes (classes based on the 90-m SRTM DEM). Correspondence between the mapped dominant and default classes is 51%. Hypsometry is hardly affected by DEM error. This is because classes are wide so that changes in elevation when DEM error is taken into account do not easily result in different outcome classes.
- In the more rugged CE pilot area slope and relief intensity are most affected by DEM error. Correspondence between the mapped dominant and default classes is 51% for slope and 44% for relief intensity. Hypsometry and flatness are hardly affected by DEM error.
- Generally, in relatively flat areas such as the WE pilot area the flatness attribute is sensitive to DEM error while in areas with more intense relief, such as in the CE pilot area, the relief intensity attribute is sensitive to errors in the DEM.
- Uncertainty about the prevailing LFA class, quantified by the entropy, is generally small. The largest uncertainties are found in zones along the class boundaries.
- Taking into account DEM error seems to improve representation of LFA classes, based on a visual assessment of the maps with the dominant LFA classes as derived from the 1,000 simulations and 90-m SRTM DEM derived base maps.

### *Validation of WP3 landform maps*

- Validation of the WP3 landform maps in terms of predictability of the WRB reference soil groups indicates that the hillshed analysis gives the best overall results.
- Both the hillshed and the object-oriented approach give better results than the WP1 map at subclass level, although differences in predictability and purity are modest. The entropy is equal for all three maps, indicating that the WP3 landform units are not internally more homogeneous with respect to soil distribution than the WP1 units.
- Validation of the landform maps was only carried out for the UK part of the WE window. One should be careful with generalizing these results to other areas. We therefore recommend to carry out a similar assessment for other areas with different soils to obtain a better insight in the added value of state-of-the-art methods for landform classification.
- We recommend to expand the validation procedure by assessing other aspects of landform classification such as an assessment of the variation of landform attributes within the landform map units. We expect that the landform units of the WP3 maps are more homogeneous in terms of landform attributes than the WP1 map, given the more detailed landform delineations compared to the WP1 map as judged from a visual inspection. However, independent validation is required to test these expectations.

## 8. References

- Ad-hoc-AG Boden, 2005. *Bodenkundliche Kartieranleitung*. 5th edition, Schweizerbart'sche Verlagsbuch-handlung, Hannover (in German).
- Avery, B.W. 1980. *Soil Classification for England and Wales (Higher Categories)*. Soil Survey Tehnical Monograph No. 14. Harpenden.
- Beekhuizen, J., Heuvelink, G.B.M., Biesemans, J., Reusen, I., 2011. Effect of DEM uncertainty on the positional accuracy of airborne imagery. *IEEE Transactions on Geoscience and Remote Sensing* 49(5), 1567-1577.
- Brus, D.J., Bogaert, P., Heuvelink, G.B.M., 2008. Bayesian Maximum Entropy prediction of soil categories using a traditional soil map as soft information. *European Journal of Soil Science* 59(2), 166-177.
- Brus, D.J., Kempen, B., Heuvelink, G.B.M., 2011. Sampling for validation of digital soil maps. *European Journal of Soil Science* 62(3), 394-407.
- Foody, G.M., 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80(1), 185-201.
- Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York.
- Kempen, B., Brus, D.J., Heuvelink, G.B.M., 2012. Soil type mapping using the generalized linear geostatistical model: a case study in a Dutch cultivated peatland. Submitted to *Geoderma*.
- Kempen, B., Brus, D.J., Heuvelink, G.B.M., Stoorvogel, J.J., 2009. Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. *Geoderma* 151(3-4), 311-326.
- Liu, C., Frazier, P., Kumar, L., 2007. Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment* 107(4), 606-616.
- Ott, R.L., Longnecker, M., 2001. *An Introduction to Statistical Methods and Data Analysis*. 5th ed. Duxbury, Palm Grove.
- R Development Core Team, 2008. *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria.
- Rabus, B., Eineder, M., Roth, A., Bamler, R., 2003. The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing* 57(4), 241-262.
- Stehman, S.V., 1997. Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment* 62(1), 77-89.
- Temme, A.J.A.M., Heuvelink, G.B.M., Schoorl, J.M., Claessens, L., 2008. Geostatistical simulation and error propagation in geomorphometry. In: T. Hengl, H.I. Reuter (Eds.), *Geomorphometry: Concepts, Software, Applications*. Developments in Soil Science. Elsevier, Amsterdam, pp. 121-140.
- Zádorová, T., Penížek, V., 2011. Problems in correlation of Czech national soil classification and World Reference Base 2006. *Geoderma* 167-168, 54-60.

## Appendix 1. Site, soil profile and topsoil data of the NSI dataset (UK)

### Site data.

Property	Example from data	Description
EAST_NSI	396000	National Grid reference easting at 5km resolution with 1km offset from true grid
NORTH_NSI	216000	National Grid reference northing at 5km resolution with 1km offset from true grid
EASTING	396000	National Grid reference easting of the actual location surveyed
NORTHING	216100	National Grid reference northing of the actual location surveyed
SURVEYDATE	220182	Date of observation
SERIES_NAME	DIDMARTON	The name of the soil series, the basic unit of soil taxonomic classification, named after the place where they were first described
VARIANT		Code to indicate that the profile is a variation of the defined soil series classification
SUBGROUP	5.14	Number used to define a subgroup incorporating the major group, group and subgroup i.e.3.11, being the third level of soil taxonomic classification based on features which further define the inherent characteristics of the soil material
LANDUSE	permanent grassland	Classification of Land use in to 17 classes
SLOPE	2	Slope in degrees
SLOPEFORM	straight	Slope form along the direction of the true slope (concaved, straight or convexed)
MADE_GROUND	Made land - Mining	Identification of sites which have been made or reclaimed by man's influence
OUTCROP		Proportion of rock outcrops within 100m of the profile
ALT	750	Altitude (m) above Ordnance Datum (OD)
ASPECT	SSE	Compass point (bearing code e.g. NNE)
EROSION		Classification of surface features formed by erosion of soil material
DEPOSITION		Classification of surface features formed by the accumulation of soil material
ROCKTYPE	calcareous oolitic limestone	Classification of rock based on recent Geological Survey and other modern publications

## Soil profile data.

Property	Example from data	Description
UPPER_DEPTH	15	Horizon upper depth in cm (<0 indicated litter layer)
LOWER_DEPTH	27	Horizon lower depth in cm (999 implies depth below bottom of profile)
TEXTURE	sandy loam	Soil texture class (includes particle size class and peat codes)
VON_POST		Modified version of the Von Post scale for assessing the degree of decomposition of peat
ESTIMATED_CLAY	17	Clay content (%) estimated in the field
ESTIMATED_SILT	20	Silt content (%) estimated in the field
MATRIX_COLOUR	75YR3/2	Colour of soil matrix following the Munsell colour notation
MOTTLE_ABUND	common	Abundance of mottles ranging from none to very many
MOTTLE_SIZE	medium	Classification of the size of the mottles ranging from extremely fine to coarse
MOTTLE_COLOUR	10YR3/3	Colour of the mottles following the Munsell colour notation
SUB_MOTTLE_ABUND	few	Abundance of subsidiary mottles ranging from none to very many
SUB_MOTTLE_SIZE	medium	Classification of the size of the subsidiary mottles ranging from extremely fine to coarse
SUB_MOTTLE_COLOUR	75YR3/2	Colour of the subsidiary mottles following the Munsell colour notation
STRUCTURE	medium moderately developed subangular	The shape, size and degree of development of the aggregation, if any, of the primary soil particles into naturally or artificially formed structural units (peds, clods etc)
STONE_ABUND	few	Class of stone abundance (per cent by volume)
STONE_SIZE	medium	Size class by diameter (cm) of stones
STONE_TYPE	sandstone stones	Classification of stone based on recent Geological Survey and other modern publications
CARBONATE COATING	non-calcareous no coatings	Calcium Carbonate level of soil estimated in the field Concentrations of clay around voids, mineral grains or peds. These are classified by the proportion of the area coated using the scale Few (<10%)< Common (10-50%) and Many (>50%)
NODULES		The shape, size, nature and composition of nodules and other concretions
POROSITY	very porous	Classification based on the percentage of macropores (>60um) per unit soil volume
ROOTS	common fine woody roots	The abundance, size and nature of roots
SOIL_WATER BOUNDARY	moist sharp irregular boundary	Soil water state (DRY, MOIST or WET) The degree and distinctness of an horizon boundary depending partly on the contrast between adjacent horizons and partly on the thickness of any transitional zone

## Topsoil data.

<i>Property</i>	<i>Example from data</i>	<i>Description</i>
pH	6.4	pH of soil
CARBON	1.6	Organic Carbon (% by weight)
AL_ACID	16836	Total Aluminium concentration (mg/kg)
AS_ACID	4	Total Arsenic concentration (mg/kg)
BA_ACID	89	Total Barium concentration (mg/kg)
CA_ACID	2541	Total Calcium concentration (mg/kg)
CD_ACID	0.6	Total Cadmium concentration (mg/kg)
CD_EDTA	0.2	Extractable Cadmium concentration (mg/l)
CO_ACID	6.2	Total Cobalt concentration (mg/kg)
CO_EDTA	0.3	Extractable Cobalt concentration (mg/l)
CR_ACID	27.3	Total Chromium concentration (mg/kg)
CU_ACID	20.3	Total Copper concentration (mg/kg)
CU_EDTA	8.3	Extractable Copper concentration (mg/l)
F_ACID	44.46	Fluoride extracted with 1mol / l sulphuric acid
FE_ACID	31096	Total Iron concentration (mg/kg)
HG_ACID		Total Mercury concentration (mg/kg)
K_ACID	2210	Total Potassium concentration (mg/kg)
K_NITRATE	101	Extractable Potassium concentration (mg/l)
MG_ACID	1703	Total Magnesium concentration (mg/kg)
MG_NITRATE	55	Extractable Magnesium concentration (mg/l)
MN_ACID	297	Total Manganese concentration (mg/kg)
MN_EDTA	27	Extractable Manganese concentration (mg/l)
MO_ACID	0.67	Total Molybdenum concentration (mg/kg)
NA_ACID	147	Total Sodium concentration (mg/kg)
NI_ACID	18.9	Total Nickel concentration (mg/kg)
NI_EDTA	1	Extractable Nickel concentration (mg/l)
P_ACID	885	Total Phosphorus concentration (mg/kg)
P_OLSON	15	Extractable Phosphorus concentration (mg/l)
PB_ACID	36	Total Lead concentration (mg/kg)
PB_EDTA	13.8	Extractable Lead concentration (mg/l)
SE_ACID	0.24	Total Selenium concentration (mg/kg)
SR_ACID	25	Total Strontium concentration (mg/kg)
V_ACID	38.34	Total Vanadium concentration (mg/kg)
ZN_ACID	72	Total Zinc concentration (mg/kg)
ZN_EDTA	3.1	Extractable Zinc concentration (mg/l)

## Appendix 2. Landform classification levels of the WP3 maps

### *Code*

PHM	Plains with hills and mountains
TAB	Table lands
OPM	Open mountains

### *Class*

PF	Flat or nearly flat plains
PSL	Smooth plains with some local relief
PH	Plains with hills
PHH	Plains with high hills
PLM	Plains with low mountains
TCR	Tablelands with moderate relief
TMR	Tablelands with considerable relief
THR	Tablelands with high relief
OHH	Open high hills
OLM	Open low mountains

### *Slope*

A	More than 80% of the area gently sloping
B	50-80% of the area gently sloping
C	20-50% of the area gently sloping
D	Less than 20% of the area gently sloping

### *Local relief*

1	0-30 m
2	30-91 m
3	91-152 m
4	152-305 m
5	305-915 m

### *Profile type*

a	More than 75% of gentle slope is in lowland
b	50-75% of gentle slope is in lowland
c	50-75% of gentle slope is on upland
d	Less than 75% of gentle slope is on upland