Placing soil-genesis and transport processes into a landscape context: A multiscale terrain-analysis approach

Markus Möller[1]*, Martin Volk[2], Klaus Friedrich[3], and **Leo Lymburner[4]**

[1] Geoflux Gbr, Paracelsusstr. 6, 06114 Halle (Saale), Germany
[2] UFZ, Helmholtz Centre for Environmental Research, Department of Computational Landscape Ecology, Permoserstr. 15, 04318 Leipzig, Germany
[3] Hessian Agency for the Environment and Geology (HLUG), Rheingaustr. 186, 65203 Wiesbaden, Germany
[4] Geosciences Australia, Remote Sensing & Strategy, Symonston, ACT, 2609, Australia

Abstract

Landforms and landscape context are of particular importance in understanding the processes of soil genesis and soil formation in the spatial domain. Consequently, many approaches for soil generation are based on classifications of commonly available digital elevation models (DEM). However, their application is often restricted by the lack of transferability to other, more heterogeneous, landscapes. Part of the problem is the lack of broadly accepted definitions of topographic location based on landscape context. These issues arise because of: (1) the scale dependencies of landscape pattern and processes, (2) different DEM qualities, and (3) different expert perceptions. To address these problems, we suggest a hierarchical terrain-classification procedure for defining landscape context. The classification algorithm described in this paper handles object detection and classification separately. Landscape objects are defined at multiple scales using a region-based segmentation algorithm which allows each object to be placed into a hierarchical landscape context. The classification is carried out using the terrain attribute mass-balance index across a range of scales. Soil genesis and transport processes at established field sites were used to guide the classification process. The method was tested in Saxony-Anhalt (Germany), an area that contains heterogeneous land surfaces and soil substrates. The resulting maps represent adaptation degrees between classifications and 191 semantically identified random samples. The map with the best adaptation has an overall accuracy of 89%.

Key words: landforms / terrain analysis / landform semantics / segmentation / mass-balance index

Accepted March 26, 2007

1 Introduction

Landforms are an important controlling boundary condition for current geomorphic processes (*Dehn* et al., 2001). Soil-related processes such as soil erosion and accumulation occur at multiple spatial and temporal scales, in each case controlled by different factors and with different intensities (*Steinhardt* and *Volk*, 2003). The development of effective soil-protection measures, such as those provided by soil-erosion models, requires the availability of scale-specific soil information (*Kirkby* et al., 1996; *Helming* and *Frielinghaus*, 1999). However, high-resolution soil data are often not available (*Steinhardt* and *Volk*, 2002; *Möller* and *Helbig*, 2005; *Behrens* et al., 2005). In contrast to soil data, digital elevation models (DEM) are usually available on different scales and typically have higher spatial resolution than soil maps. It is well known that strong relationships exist between the spatial distribution of soils and the topography of a given landscape (*Conacher* and *Dalrymple*, 1977; *Speight*, 1988). The use of digital terrain analysis can help to reduce the need for costly conventional survey methodologies by establishing a relationship between terrain attributes, soil genesis/transport processes, and different soil types. This process when combined with field validation can be used to provide high-resolution soil information. This has resulted in the increasing use of topography in many digital–soil mapping (DSM) projects (*McBratney* et al., 2003; *Behrens* and *Scholten*, 2006; *Lagacherie* et al., 2006). There are three key factors to consider when performing a DEM-based landform classification:

(1) Landforms occur on different scales (*Schmidt* and *Dikau*, 1999; *Evans*, 2003). Several approaches tackle the scale problem by using different window sizes—representing scales of interest—for the derivation of multiscale terrain attributes (*Gallant* and *Dowling*, 2003; *Fisher* et al., 2004; *Schmidt* and *Hewitt*, 2004; *Jenness*, 2005). The attributes show scale-specific alterations (*Gallant* and *Hutchinson*, 1997; *Thompson* et al., 2001; *Shary* et al., 2005). Their classification enables consideration of spatial context and uncertainty. The main disadvantage in this approach is that the large moving window sizes reduce the resulting output coverage. Coverage is defined here as the spatial extent of the input and resulting data set (*Bierkens* et al., 2000).

* Correspondence: M. Möller; e-mail: moeller@geoflux.de

(2) Common landform semantics do not exist because of their dependence on user's perception which reflects the user's discipline paradigm (*Bishr*, 1998; *Dehn* et al., 2001). For example, an ecologist and an engineer may define "floodplains" using completely different criteria. Semantics denote the relationships between computer representations and the corresponding real-world feature within a certain context (*Bishr*, 1998). The semantic issue is often counteracted by either

(a) using fuzzy rules in the classification process (*Burrough* et al., 2000; *MacMillan* et al., 2000; *Fisher* et al., 2004; *MacMillan* et al., 2004; *Schmidt* and *Hewitt*, 2004; *Drâgut* and *Blaschke*, 2006) or

(b) using an expert knowledge base that considers the geometrical and topological features as well as object and semantic hierarchies (*de Bruin* et al., 1999; *Wielemaker* et al., 2001; *Drâgut* and *Blaschke*, 2006).

The heuristic classification approach is subjective, but enables better inclusion of expert knowledge (*MacMillan* et al., 2000; *Drâgut* and *Blaschke*, 2006) whereas automatic classifications have the advantage of greater objectivity. However, problems may arise from the semantic interpretation of the automatically defined classes (*Burrough* et al., 2001).

(3) Landform-classification approaches are generally difficult to transfer to heterogeneous landscapes because of the aforementioned scale and definition issues (*Schmidt* and *Hewitt*, 2004; *MacMillan* et al., 2004). This is of particular concern for statistically based approaches. Because of their rigid thresholds, heuristic approaches are unable to take into account specific landscape conditions in large study areas. The implementation of fuzzy rules and class definitions with relative values and relative positions to neighboring objects can increase the transferability of heuristic approaches (*Drâgut* and *Blaschke*, 2006).

This paper focuses on the development of an automatic procedure of terrain-object delineation and classification which

(1) takes into consideration landscape heterogeneity and scale without coverage reduction and

(2) allows the adaptation of landform definitions to user's perception.

Our method aims to classify four simple landforms: floodplain, depression, plain, and slope. The classification algorithm treats terrain segmentation and classification separately. The terrain-segmentation process generates discrete landscape units, represented by polygons, at multiple scales. These polygons are related *via* hierarchy, *i.e.*, a larger-scale "parent" polygon may contain a series of smaller "children" polygons, where each child polygon may be unique, but each child polygon also "inherits" attributes from its parent. This hierarchy can also be established across multiple scales ("grandparents" and "great grandparents"), thereby enabling the definition of hierarchical multiscale terrain-object structures (*cf.*, section 2.3). The classification of these polygons is carried out by means of the terrain attribute "mass-balance index" (*cf.*, section 2.2) across a range of spatial scales using a multihierarchical query procedure, a statistically and probability-based operator (*cf.*, section 2.4).

2 Material and methods

2.1 Site description

A study area with heterogeneous soil and relief conditions was selected to demonstrate the applicability of our new methodology. The study area of Könnern, which represents such conditions, is situated in the S of the German State of Saxony-Anhalt near the city of Halle (Fig. 1). The area of 100 km^2 corresponds to the land area equivalent to the offi-

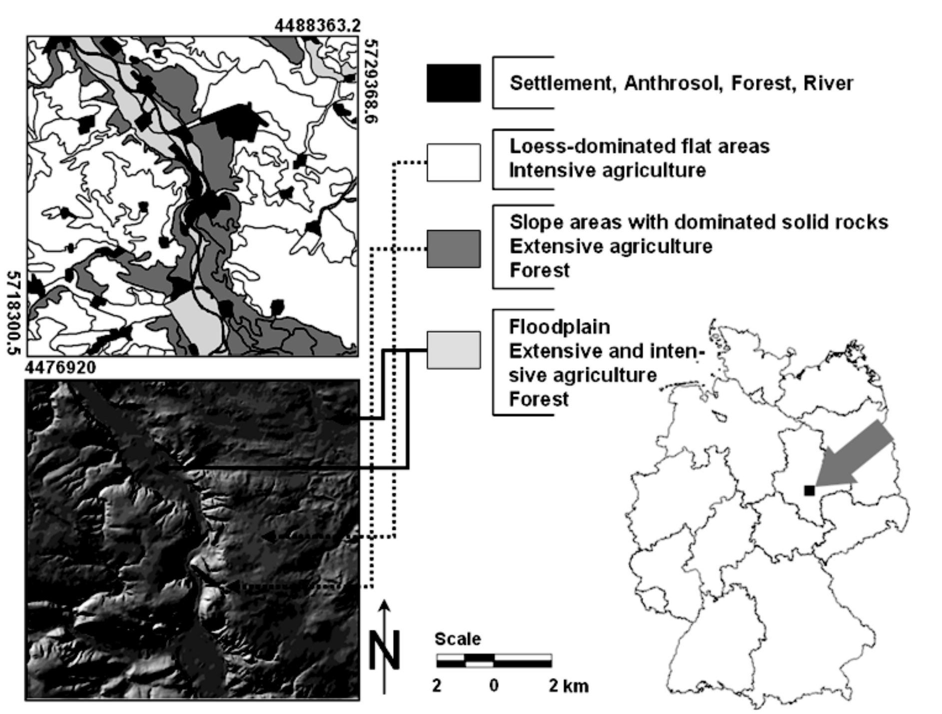


Figure 1: Study area Könnern: Shaded relief and soil-related nature units according to mesoscale agricultural site-mapping program MMK (http://www.lagb.sachsen-anhalt.de [soil data] and http://www.lvermgeo.sachsen-anhalt.de [DEM]).

cial topographic map of Könnern at a scale of 1:25,000. The region is among the driest regions of Germany with a mean annual precipitation <500 mm (*Trefflich*, 1997).

Soil and relief formation of the study area was dominated by processes under glacial and periglacial conditions during the Saalian and Weichselian glacial stages (*Altermann*, 1970). The plateau unit, mainly covered by Chernozem soils developed in the Weichselian Loess and Saalian moraine material, forms the largest and central part of the study area. It is divided by the valley of the Saale river, extending in a S–N direction. In the N and S part, the valley floor was enlarged by solution subsidence processes (*Kunert*, 1970). Neighboring relief units of the Loess plateau are the Fuhne floodplain in the NE and the Schlenze valley in the SW. Fluvisols and Gleysols have developed in the floodplain sediments. The plateau margins are characterized by sharp and shallow valleys (*Möller*, 2005). Depending on landforms and substrate, Leptosols or Regosols have developed in the sandstones or claystones of Palaeozoic ages.

Due to the fertile soils of the study area (Chernozems), the landscape is influenced by an intensive agriculture. The soils are at strong risk of erosion because of heterogeneous landscape morphology, the erodibility of the dominant loess substrate and the intense summer rainstorm events.

2.2 Data base and preparation

The study was carried out using publicly available elevation data with a resolution of 10 m and vertical and horizontal accuracy of approx. 0.5 m (see www.lvermgeo.sachsen-anhalt.de/de/main.htm). The DEM was originally generated *via* the digitization of elevation contours. Structure elements (*e.g.*, dams) or lakes are not included. The ANUDEM algorithm by *Hutchinson* (1989) was applied in order to create a hydrological sound DEM.

The following geomorphometric attributes listed in Tab. 1 were derived from the DEM. The variables *h, n, k,* and *ht* were used as input to the terrain structuring (*cf.*, section 3.1). Attributes *ra* and *hd* refer to neighboring, sub- and superobjects (*cf.*, section 2.3) and enter into the floodplain-classification process (*cf.*, section 2.4.1).

Table 1: Inputs for terrain-classification algorithm.

Terrain attribute		Program
Elevation	*h*	ANUDEM 5.2[a]
Slope	*n*	Landserf 2.2[b]
Mean curvature	*k*	
Vertical distance to channel network	*ht*	SAGA1.1[c]
Mass-balance index	MBI	own application
Vertical distance to neighboring objects	*hd*	eCognition3.0[d]
Vertical distance between neighboring objects within a superobject	*ra*	

[a] http://cres.anu.edu.au/outputs/anudem.php
[b] http://www.landserf.org/
[c] http://www.saga-gis.uni-goettingen.de/html/index.php
[d] http://www.definiens-imaging.com

Previous landform-classification approaches have used process-based terrain attributes (*e.g., Blaszczynski*, 1997; *Park* et al., 2001). In this study, we used the mass-balance index MBI (*Friedrich*, 1996, 1998) based on the assumption that different soil-related landforms can be identified based on their MBI values. We assume that negative MBI values represent areas of net deposition such as depressions and floodplains; positiveMBI values represent areas of net erosion such as hill slopes, and MBI values close to zero indicate areas where there is a balance between erosion and deposition such as low slopes and plain areas.

The mass-balance index is derived from transformed $f(k, ht, n)$ values (Eq. 1). As shown in Fig. 2a, high positive MBI values occur at convex terrain forms, like upper slopes and crests,

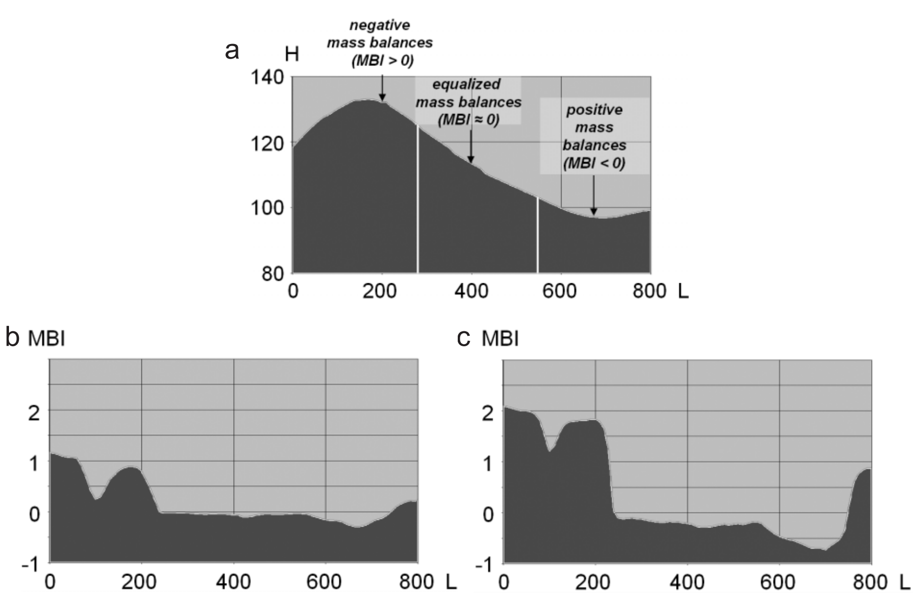


Figure 2: Relation between slope cross section and mass-balance index (MBI) (H, height; L, length of cross section); a) cross section and positions of negative, positive, and equalized mass balances, b) MBI with $T_{ht,n}$ = 15 and T_k = 0.067, c) MBI with $T_{ht,n}$ = 15 and T_k = 0.67.

while lower MBI values are associated with valley areas and concave zones at lower slopes. Balanced MBI values close to zero can be found in midslope zones and mean a location of no net loss or net accumulation of material. Figure 2 also demonstrates how MBI values provide information about relatively balanced states of potential material transport but do not quantify the volume of material in flux.

$$\mathrm{MBI} = \begin{cases} f(k) \times [1 - f(n)] \times [1 - f(ht)] & \text{for } f(k) < 0 \\ f(k) \times [1 + f(n)] \times [1 + f(ht)] & \text{for } f(k) > 0 \end{cases} \quad \text{with MBI} \in [-1,\ 3]. \tag{1}$$

The attributes were transformed according to Eq. 2 (*Friedrich*, 1996, 1998):

$$f(x) = \frac{x}{(|x| + T_x)}$$

$$\text{with } x = k,\ n,\ ht,\ h,\ f(k) \in [-1,\ 1],\ f(n,\ ht,\ h) \in [0,\ 1]. \tag{2}$$

This reciprocal operation is extended by the transfer constant T_x which allows different value ranges to be stretched or smoothed: the smaller T_x, the more the value range in the histogram is stretched. This has a large effect on the curvature attribute k which is considered most significant for changing both soil conditions (*Friedrich*, 1996; *Ad-hoc-AG Boden*, 2005) and MBI value range. The comparison of the two MBI versions makes the outcome of the T_k values for MBI characteristic clear: the lower T_k, the greater the relative difference within the value range (Fig. 2b and c).

2.3 Terrain structuring

Landscapes are hierarchically structured. In concepts of hierarchical landscape structuring (*cf., Steinhardt* and *Volk*, 2001, 2003), the delineation of the largest spatial units (hereafter referred to as terrain superobjects) arises from the significant alteration of landscape-related attributes on the one hand and the arrangement of subordinate units or subobjects within hierarchical superobjects on the other hand (Fig. 3).

An automatic implementation of the hierarchical–landscape structuring concept can be achieved by using a region-growing segmentation algorithm applied to continuous digital spatial data like remote-sensing data or DEMs (*Woodcock* and *Harward*, 1992; *Burnett* and *Blaschke*, 2003; *Hay* et al., 2003; *Drâgut* and *Blaschke*, 2006). Here, the fractal–net evolution approach (FNEA) was executed which is described in detail by *Baatz* and *Schäpe* (2000) and *Benz* et al. (2004). Using a hierarchical and bottom-up region-growing algorithm, the FNE algorithm merges single pixel elements (terrain attributes) to terrain objects on different spatial scales building up a hierarchical network of terrain-object levels (Fig. 3). This means that all objects are surrounded by neighboring objects and each object is related to larger and smaller scales *via* parent–children relationships (*cf.*, section 1). As a consequence, each object carries a data set of information including attributes of statistics, neighboring and hierarchical relationships (*e.g.*, attributes *hd* and *ra* in Tab. 1). These data make it possible to implement a multiscale classification algorithm based on hierarchical features.

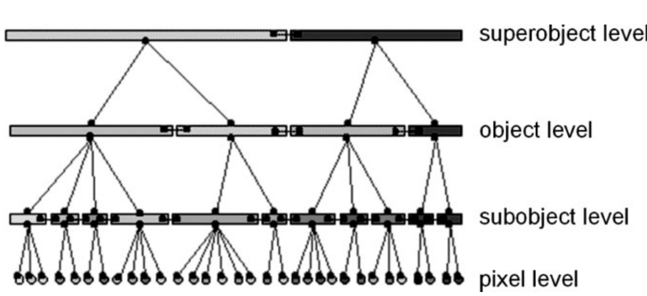


Figure 3: Four-level hierarchical network of terrain objects (*Benz* et al., 2004).

The FNE segmentation algorithm can be considered as an optimization process which minimizes the heterogeneity H of each spatial object for a given resolution over the entire continuous data set with constraints based on local and global conditions. The user-defined heterogeneity H refers to both heterogeneity of pixel values h_{color} and shape heterogeneity h_{shape} according to Eq. 3:

$$H = w_{color}\Delta h_{color} + w_{shape}\Delta h_{shape}$$
$$\text{with } w_{color,shape} \in [0,\ 1],\ w_{color} + w_{shape} = 1. \tag{3}$$

While h_{color} results from the difference between object parameters like object variance, h_{shape} arises from the balance of the object shape features smoothness h_{smooth} and compactness h_{compt} (Eq. 4):

$$\Delta h_{shape} = w_{compt}\Delta h_{compt} + w_{smooth}\Delta h_{smooth}. \tag{4}$$

The parameters w_{color}, w_{shape}, w_{smooth}, and w_{compt} allow finally the weighting of the heterogeneity factors in order to achieve an application-related adaptation of the segmentation results.

2.4 Landform classification

For the purposes of this study, landforms are defined using the following semantics:

- Floodplains are low and flat relative to their surroundings and occur on different scales (*Gallant* and *Dowling*, 2003).

- Depressions and floodplains represent fluvial landforms (*Friedrich*, 1996). They are different in size (floodplains are larger than depressions). Depressions are also low relative to their surroundings but they need not to be flat.

- Slopes, plains, and depressions represent specific scales (*Fisher* et al., 2004; *Jenness*, 2005).

- Slopes, plains, depressions, and floodplains can be differentiated according to their mass balances. Depressions and floodplains show positive mass balances (areas of net deposition), slopes are characterized by negative mass balances (areas of net erosion), and plains are equilibrated. Potential sediment accumulation is therefore more likely to take place in flat than in steeper depression areas. Accumulation reaches a maximum at intense concave curvatures and in a small distance from areas where erosion is occurring. The potential for soil erosion increases with more convex curvature with increasing distance from the channel network (*Friedrich*, 1996; *Möller*, 2005).

This means for our approach that floodplains have to be classified in a multihierarchical manner (*cf.*, section 2.4.1) whereas for the classification of the remaining landforms, specific scales need to be defined (*cf.*, section 2.4.2).

2.4.1 Floodplains

The detection of floodplains is based on a multihierarchical query procedure (Fig. 4a). For each considered scale level (n) resulting from multihierarchical segmentation, a query according to Eq. 5 is performed. The levels that do not fulfil the conditions of the query are transferred to the segmentation level ($n-1$). The procedure is reiterated until no segmentation level is available anymore or the user sets the termination manually, since depression areas appeared from this level onwards.

$$\text{Floodplain} = (hd < 0) \cup \min(\text{MBI}) \cup ra$$
$$\textit{with } ra \in [0,\ x],\ y_n \neq y_{n+1}, y = hd, \text{MBI},\ ra. \qquad (5)$$

The term $hd < 0$ means that floodplains on each hierarchy level are located lower than their surroundings. A terrain object with min(MBI) has the smallest positive mass balance within the corresponding superobject. The variable ra means that the objects with a defined mean change in relief are recorded, whereby x represents a maximum of the relief amplitude that has to be determined by the user. In accordance to *Bernhardt* et al. (1991), a value of $ra = 2$ has been used here. The criterion $y_n \neq y_{n+1}$ is applied to avoid a scenario in which objects are classified that have not experienced a spatial differentiation with the transition to the segmentation levels n to n-1))

2.4.2 Slopes, plains, and depressions

The classification procedure combines a statistic structuring method (k means-cluster analysis) with a probability-based approach (maximum-likelihood algorithm) (*cf., McGarigal* et al., 2002). In order to take into account the landscape heterogeneity of the study area, the classification follows a hierarchical approach by which all subordinated (sub-)objects are classified separately according to the spatial extent of the superior (super-)objects (Fig. 4b).

Samples were selected by the following criteria which correspond to context-based landform definitions:

– Minimal MBI values represent depressions, and maximal MBI values indicate slopes.

– Samples for the plain class occur in a cluster where the values lie in the positive and negative value range close to a value of zero (neutral mass balance).

The following two variables influence the classification results, and these parameters can be adjusted so that the outputs are consistent with reference information:

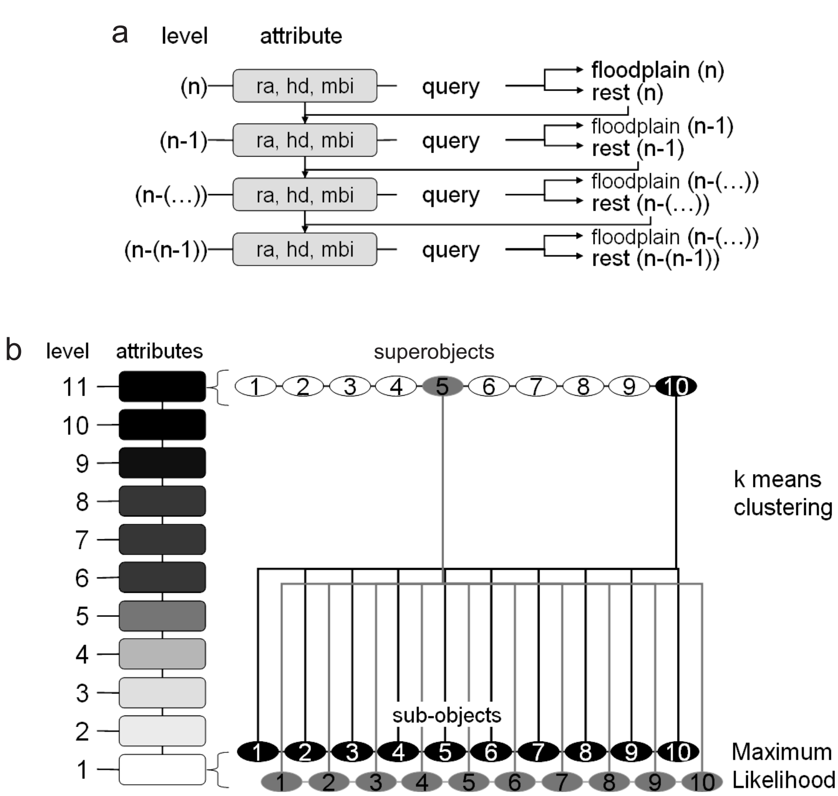


Figure 4: Landform-classification scheme; a) floodplain detection, b) classification of depressions, slopes, and plains.

1) The transfer constant T_k affects the value range of the attribute curvature *k* which is the crucial attribute for MBI calculation (*cf.*, section 2.2).

2) The hierarchy variable determines from which hierarchical level the superobjects are used for the classification procedure. Their modification alters the number of the resulting samples.

2.4.3 Validation

Based on 191 random samples in the study area, elevation cross profiles were set for each point using the Erdas Imagine 8.4 spatial profile tool. From these profiles, we carried out an on-screen determination of the particular landforms considering the local landscape conditions (digital manual mapping, *cf., Möller*, 2005). The sample definition represents the expert knowledge of the user but may also reflect certain class definitions used in a scientific discipline or institution (*e.g.*, soil survey). Figure 5a exemplifies the methodology for a random sample which is situated in a depression landform. The reference information was used to determine the accuracy with which the classification results matched with semantically identified random samples. As adaptation measures the overall accuracy (OA), user's accuracy (UA) and producer's accuracy (PA) were calculated for each landform class deriving from confusion matrix (Fig. 5b; *Stehmann*, 1997; *Foody*, 2002; *Zhan* et al., 2005). The highlighted elements are the main diagonal and contain the cases where the labels depicted in the classification and reference data set agree. The off-diagonal elements represent the cases of label disagreement. *Thomlinson* et al. (1999) stated as a target of a minimum overall accuracy of 85% with no class <70% accuracy (*cf., Foody*, 2002).

Overall accuracy belongs to the most popular measures and is the percentage of all cases correctly allocated to classification. With UA and PA, two class-specific views on confusion matrix can be distinguished depending on whether the calculations are based upon the matrix's row or column marginals (*Foody*, 2002). Producer's accuracy indicates thereby the real hit rate of the classification regarding the reference information (sum of columns). User's accuracy results on the other hand from the "used" classification product. The information content of the classification product is assigned to the reference points (sum of rows).

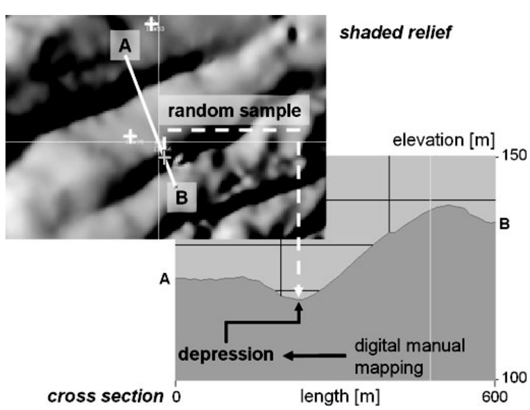
3 Results

3.1 Terrain objects

The transformed attributes *f*(*h*), *f*(*n*), *f*(*k*), and *f*(*ht*) determine the segmentation and the object generation. Their selection was based on two factors:

(1) the published relationships between terrain attributes and their influence on the soil formation and transport processes (*McBratney* et al., 2003; *Ad-hoc-AG Boden*, 2005) and

a



b

Actual Class

Predicted Class	A	B	C	D	Σ
A	n_{AA}	n_{AB}	n_{AC}	n_{AD}	n_{A+}
B	n_{BA}	n_{BB}	n_{BC}	n_{BD}	n_{B+}
C	n_{CA}	n_{CB}	n_{CC}	n_{CD}	n_{C+}
D	n_{DA}	n_{DB}	n_{DC}	n_{DD}	n_{D+}
Σ	n_{+A}	n_{+B}	n_{+C}	n_{+D}	n

$$\text{Percentage correct} = \frac{\sum_{k=1}^{q} n_{kk}}{n} \times 100$$

$$\text{User's accuracy} = \frac{n_{ii}}{n_{i+}}$$

$$\text{Producer's accuracy} = \frac{n_{ii}}{n_{+i}}$$

Figure 5: Validation scheme; a) methodology for sampling of reference data, b) the confusion matrix and some common measures of classification accuracy that may be derived from it (*Foody*, 2002).

(2) the generation of color composites and the related visualization of landforms (Fig. 6a).

The Figs. 6 b–l show the different segmentation levels that represent objects in a multiscale and multidimensional context. The average object sizes OS identify the particular scale area of the segmentation level. The term "multiscale" means that all terrain objects (*e.g.*, of level 7) are both constituted by terrain subobjects (*e.g.*, of level 1) and elements of superobjects (*e.g.*, of level 11; Fig. 6b, h, and l). Multidimensional objects correspond to the classic idea of landform elements or landform facets (*Friedrich*, 1998; *Blaschke* and *Strobl*, 2003; *Drâgut* and *Blaschke*, 2006, *cf.*, section 2.3). However, this only considers objects of the level 1 or 2 because of their low heterogeneity (Fig. 6b and c).

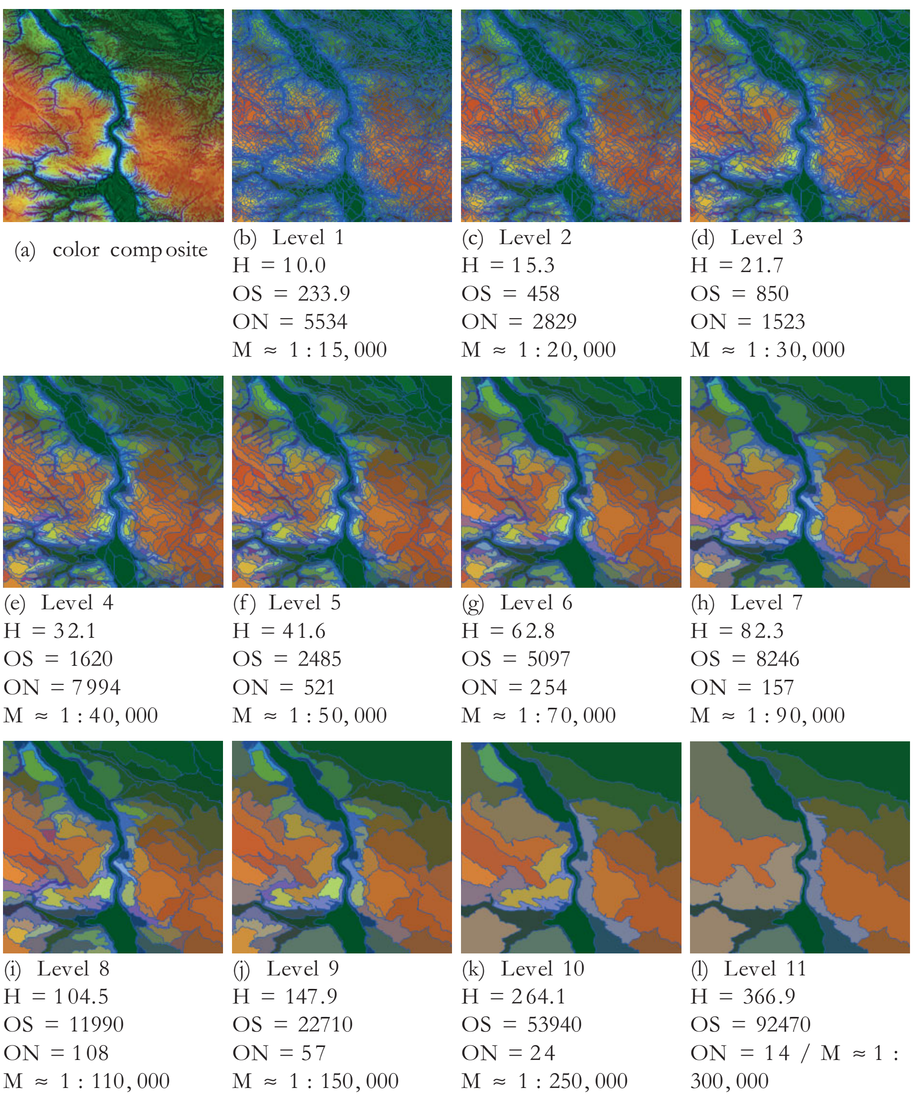
Multidimensionality is exhibited by the fact that on certain aggregation levels, terrain objects emerge or recede. For instance, in the segmentation level 11 (Fig. 6l), the Saale River floodplain appears as a single object, whereas depression areas—which dominate in the segmentation levels shown in the Figs. 6 b to f—are merged in into terrain superobjects.

3.2 Floodplains

The query results and the statistical values of the used attributes are presented in Tab. 2. Six floodplain objects were detected on six different hierarchical levels which are shown in column "Level" (*cf.*, Fig. 6). On level 6 (Fig. 6g), the query was terminated.

3.3 Depressions, slopes, and plains

The segmentation level 1 corresponds to the average object scale of 1:15,000 (Fig. 6b). The modification of (1) the attri-



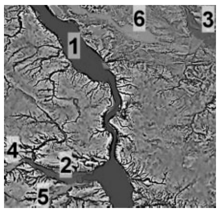
parameter settings: $w_{color} = 0.8$ / $w_{shape} = 0.2$ / $w_{smooth} = 0.9$ / $w_{compt} = 0.1$

color composite: f (h), f (ht) = red / f (k) = green / f (n) = blue

H = scale parameter / ON = objects number/ OS = object size [m^2] / M = mean scale

Figure 6: Scale levels based on the segmentation of transformed terrain attributes *f*(*h*), *f*(*ht*), *f*(*n*), and *f*(*k*).

Table 2: Multihierarchical query results of floodplain detection (*cf.*, Fig. 4).

Floodplains	No	Level	MBI[a]	*hd*[a]	*ra*[a]
	1	11	–0.03	–30.16	0.88
	2	9	–0.15	–8.22	1.93
	3	8	–0.04	–5.25	0.53
	4	7	–0.25	–0.85	1.85
	5	6	–0.06	–0.85	1.42
	6	6	–0.15	–3.21	0.77

[a] *cf.*, Tab. 1

bute variable T_k (*cf.*, Eq. 1) and (2) the hierarchy variable (object number ON) causes different proportions of the resulting classes slope, depression, and plain. The starting point of the classification procedure is the classification variant with the parameter adjustments $T_k = 0.0067$, $T_{n,ht} = 15$, and ON = 14 (Fig. 6l and 7e).

3.3.1 Modification of the attribute variable

Figures 7 a–d and 8a clarify the effects of T_k modifications. According to Fig. 2, low T_k values emphasize both terrain forms in flat and sloped areas (Fig. 7a) whereas high T_k values only highlight terrain forms in sloped areas (Fig. 7d). One consequence is that the same landform is described by different MBI values depending on the used T_k values. Thus, an increase of the T_k values is associated with a distinct increase of flat areas and a decrease of the slope areas at the same time. In contrast, the area proportions of the depression class remain stable.

3.3.2 Modification of the hierarchy variable

All classifications of level 1 subobjects (Fig. 6b) refer to a different level of superobjects (*cf.*, Fig. 4b). Apart from level 9 (ON = 57), level 10 (ON = 45), and level 11 (ON = 14; Fig. 6j, k, and l), an additional level was created (ON = 30). The lar-

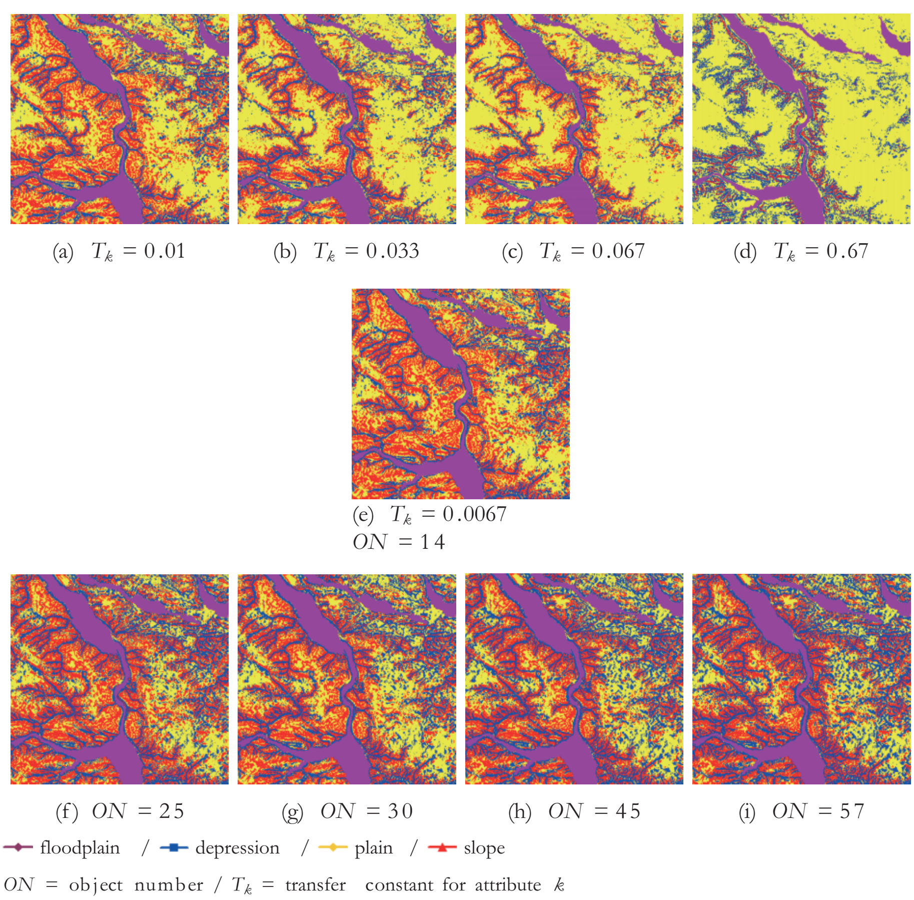


Figure 7: Landform-classification results.

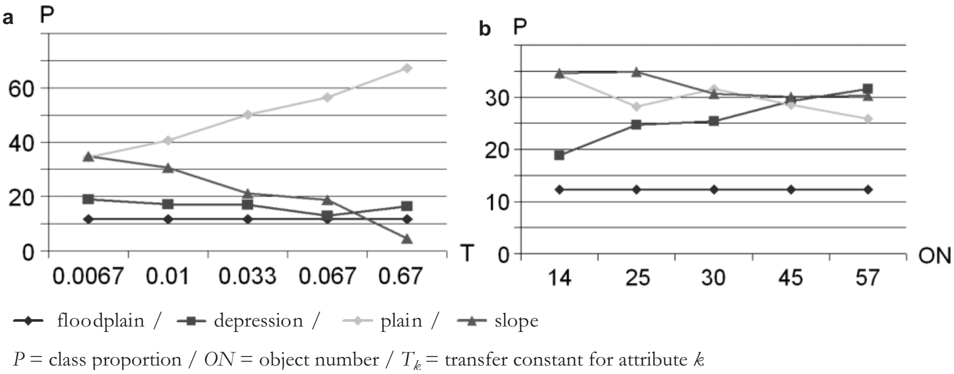


Figure 8: Relations between variable variants of terrain classification and proportions of landforms. (a) Attribute variable with ON = 14, (b) hierarchie variable with T_k = 0.0067.

ger the number of superobjects, the more samples can enter into the classification procedure. Following the sampling strategy (*cf.*, section 2.4.2), the number of samples results from four times the number of superobjects.

As shown in Figs. 7 f–i and 8b, the effect of ON modifications is less sensitive than using the attribute variable. However, as opposed to the attribute variable, an increase of the sample size leads to an increase of the proportion of depression class whereas the proportion of slope areas declines. The resulting proportion of the flat areas varies.

3.4 Validation

Figure 9 summarizes the overall, producer's, and user's accuracies (OA, PA and UA). Accordingly, the attribute variable is the decisive factor for affecting the classification accuracy. This is also shown in Fig. 8. The highest accuracy or the best adaptation between classification results and reference base is achieved with the transfer constant T_k = 0.033 and a number of superobjects ON = 14 (Fig. 9a and b). In terms of classification accuracy, an overall accuracy of 89% was achieved. The classification accuracy of all single classes exceeds the minimum accuracy of 70% (*cf.*, section 2.4.3).

4 Discussion and conclusions

We present a new innovative procedure for the mapping landforms on a soil-genesis and transport basis. The procedure considers multiple spatial scales and can be applied in heterogeneous landscapes. The classified landforms do not inevitably represent soil units, since other factors influence the soil distribution, too. However, the results indicate that this approach will improve existing digital soil-mapping (DSM) methodologies.

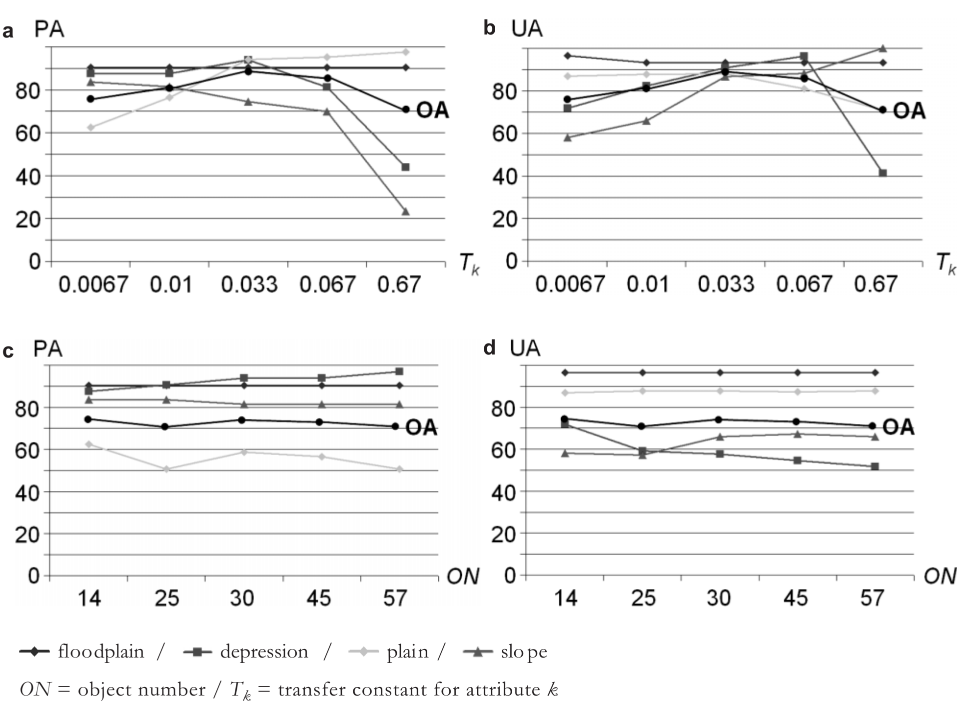


Figure 9: Relations between overall accuracy (OA), user's accuracy (UA), and producer's accuracy (PA) as well as variable variants of landform classification. (a) PA and OA of attribute variable variants with *ON* = 14, (b) UA and OA of attribute variable variants with *ON* = 14, (c) PA and OA of hierarchy variable variants with T_k = 0.0067, (d) UA and OA of hierarchy variable variants with T_k = 0.0067.

4.1 Landscape heterogeneity, scale, and coverage

Our terrain-classification procedure treats object generation and classification separately:

(1) Object generation is based on soil-relevant terrain attributes that are transferred by a region-based segmentation procedure to multihierarchical object structures. As shown by our results, multihierarchical object structures can avoid the limitation of present approaches, like hierarchical moving-window classification procedures (*Gallant* and *Dowling*, 2003; *Fisher* et al., 2004; *Schmidt* and *Hewitt*, 2004; *Jenness*, 2005), which lead to the loss of the resulting coverage by the identification of fluvial landforms (depression, floodplains) on different scales.

(2) The classification procedure distinguished between floodplain detection on the one hand and classification of depressions, slopes, and plains on the other hand:

- Floodplain detection was realized by a multihierarchical query procedure.
- The classification of depressions, slopes, and plains was carried out on a specific target scale (here: approx. 1:15,000) and combined a statistical structuring method (cluster analysis) with a probability-based operator (Maximum Likelihood) and is based on the terrain attribute MBI. The hierarchical relations to superior objects enabled a landscape-specific selection of training areas. This enables landscape heterogeneity to be considered (*cf., MacMillan* et al., 2004; *Schmidt* and *Hewitt*, 2004).

4.2 Landform definition

A key advantage of our classification procedure is the option to modify the area assigned to classes using two different variables, (1) the hierarchical variable and (2) the attribute variable. These two variables allow the adaptation of classification results to reference information and specific class definitions:

(1) The hierarchical variable enables the alteration of sample size and their spatial distribution depending on superobjects which are used for the classification procedure.

(2) The attribute variable affects the value range of MBI by changing of a transfer constant T_x. The MBI has proved to be easily interpretable regarding landform definitions. In this study, all landforms were described by relative values (*e.g.*, maximum MBI value = slope). Thus, the landform definitions are transferable (*cf., Drâgut* and *Blaschke*, 2006).

4.3 Validation

In the majority of soil-related landform-classification approaches, classification quality was deduced from statistical relations between soil and terrain properties (*Pennock* et al., 1987; *Zhu* et al., 1997; *Park* et al., 2001; *Park* and *Vlek*, 2002; *Pennock*, 2003; *Park* and *van de Giesen*, 2004; *Ryan* et al., 2000; *Schmidt* and *Hewitt*, 2004; *MacMillan* et al., 2004). In this study, each classification is labeled by a specific accuracy metric (here: overall, user's, and producer's accuracy). Thus, our approach enables additional applications such as the revision of existing soil maps (*Friedrich*, 1998; *Möller*, 2005). Finally, while reference information is usually mapped during soil survey, our approach realizes the mapping of landforms by an efficient on-screen mapping according to *Möller* (2005).

4.4 Further research

An unsolved problem is the determination of classification-relevant hierarchy levels, for instance, for the delineation of floodplains. One possibility is the identification of landscape-scale thresholds (*Hay* et al., 2001; *Hall* et al., 2004). In this context, the observed relations between hierarchy variable variants and alternating area proportions of landforms require additional research. Furthermore, it remains unclear which parameter adjustments of the segmentation algorithm best represents the underlying terrain units. One option is an object validation based on object-related reference information (*Möller* et al., 2006). This also includes the validation process of landforms itself. Finally, the number of clusters is chosen subjectively. Thus, algorithms have to be included enabling an optimum number of clusters (*e.g., de Bruin* and *Stein*, 1998).

Further possible applications exist in connection with the integration of the MBI attribute in qualitative soil-erosion assessments. Work being undertaken by us aims at the modification of length-slope factor in the universal soil-loss equation (USLE, *e.g., Moore* and *Burch*, 1986; *Hickey*, 2000). This could help to overcome limits of existing USLE-based erosion assessment methods namely the classification of accumulation areas (*Merritt* et al., 2003). Finally, the use of natural system units, such as watershed hierarchies, will be used as a basis for the classification to enable a linkage to hydrological models with the objective to improve their spatial process description (*Volk* et al., 2007).

Acknowlegdments

The authors are very grateful to Dr. *Matthias Stöck* (University of Lausanne, Switzerland) for his helpful comments on an earlier version of this manuscript.

References

Ad-hoc-AG Boden (2005): Bodenkundliche Kartieranleitung. 5th edn., E. Schweizerbart'sche Verlagsbuchhandlung, Hannover.

Altermann, M. (1970): Böden, in Kunert, R. (ed.): Erläuterungen zur Geologischen Spezialkarte 1:25.000. Blatt Könnern. Zentrales Geologisches Institut, Berlin.

Baatz, M., Schäpe, A. (2000): Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. in Strobl, J., Blaschke, T. (eds.): Angewandte Geographische Informationsverarbeitung – Beiträge zum AGIT-Syposium 12, Salzburg. Wichmann-Verlag, Heidelberg, pp. 12–23, http://www.definiens.com/binary_secure/405_baatz_fp_12.pdf

Behrens, T., Scholten, T. (2006): Digital soil mapping in Germany—A review. *J. Plant Nutr. Soil Sci.* 169, 434–443.

Behrens, T., Förster, H., Scholten, T., Steinrücken, U., Spies, E.-D., Goldschmitt, M. (2005): Digital soil mapping using artificial neural networks. *J. Plant Nutr. Soil Sci.* 168, 21–33.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M. (2004): Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *J. Photogram. Rem. Sens.* 58, 239–258.

Bernhardt, A., Schwanecke, W., Diemann, R., Kopp, D., Weiße, R., Barsch, H., Kugler, H., Haase, G. (1991): Mesorelief-Mosaik-Typen, in Haase, G., Barsch, H., Hubrich, H., Mannsfeld, K., Schmidt, R. (eds.): Naturraumerkundung und Landnutzung – geochorologische Verfahren zur Analyse, Kartierung und Bewertung von Naturräumen, Beiträge zur Geographie, Akademie-Verlag, Berlin.

Bierkens, M., Finke, P., de Willigen, P. (2000): Upscaling and downscaling – Methods for environmental research. Developments in Plant and Soil Sciences. Kluwer Academic Publishers, Dordrecht.

Bishr, Y. (1998): Overcoming the semantic and other barriers to GIS interoperability. *Int. J. Geogr. Inform. Sci.* 12, 299–314.

Blaschke, T., Strobl, J. (2003): Defining landscape units through integrated morphometric characteristics, in Buhmann, E., Ervin, S. (eds.): Landscape Modelling: Digital Techniques for Landscape Architecture. Wichmann-Verlag, Heidelberg, pp. 104–113.

Blaszczynski, J. (1997): Landform characterization with geographical information systems. *Photogramm. Eng. Rem. Sens.* 63, 183–191.

Burnett, C., Blaschke, T. (2003): A multi-scale segmentation/object relationship modelling methodology for landscape analysis. *Ecol. Model.* 168, 233–249.

Burrough, P., van Gaansa, P., MacMillan, R. (2000): High-resolution landform classification using fuzzy k means. *Fuzzy Set. Syst.* 113, 37–52.

Burrough, P., Wilson, J., van Gaans, P., Hansen, A. (2001): Fuzzy k-means classification of topo-climate data as an aid to forest mapping in the greater Yellowstone area, USA. *Landsc. Ecol.* 16, 523–546.

Conacher, A., Dalrymple, J. (1977): The nine unit landsurface model: an approach to pedogeomorphic research. *Geoderma* 18, 1–153.

de Bruin, S., Stein, A. (1998): Soil-landscape modelling using fuzzy k means clustering of attribute data derived from a Digital Elevation Model (DEM): *Geoderma* 83, 17–33.

de Bruin, S., Wielemaker, W., Molenaar, M. (1999): Formalisation of soil-landscape knowledge through interactive hierarchical disaggregation. *Geoderma* 91, 151–172.

Dehn, M., Gärtner, H., Dikau, R. (2001): Principles of semantic modelling of landform structures. *Comput. Geosci.* 27, 1005–1010.

Drâgut, L., Blaschke, T. (2006): Automated classification of landform elements using object-based image analysis. *Geomorphology* 81, 330–344.

Evans, I. (2003): Scale-specific landforms and aspects of the land surface, in Evans, I., Dikau, R., Tokunaga, E., Ohmori, H., Hirano, M. (eds.): Concepts and Modelling in Geomorphology: International Perspectives. Terrapub, Tokyo, pp. 61–86.

Fisher, P., Wood, J., Cheng, T. (2004): Where is helvellyn? Fuzziness of multiscale landscape morphometry. *Trans. Inst. Br. Geogr.* 29, 106–128.

Foody, G. (2002): Status of land cover classification accuracy assessment. *Rem. Sens. Environ.* 80, 185–201.

Friedrich, K. (1996): Digitale Reliefgliederungsverfahren zur Ableitung bodenkundlich relevanter Flächeneinheiten. Frankfurter Geowissenschaftliche Arbeiten D 21, Frankfurt/M.

Friedrich, K. (1998): Multivariate distance methods for geomorphographic relief classification, in Heinecke, H., Eckelmann, W., Thomasson, A., Jones, J., Montanarella, L., Buckley, B. (eds.): Land Information Systems – Developments for planning the sustainable use of land resources. European Soil Bureau – Research Report 4, EUR 17729 EN, Office for official publications of the European Communities, Ispra, pp. 259–266.

Gallant, J., Dowling, T. (2003): A multi-resolution index of valley bottom flatness for mapping dispositional areas. *Water Resour. Res.* 39, 1347–1360.

Gallant, J., Hutchinson, M. (1997): Scale dependence in terrain analysis. *Math. Comput. Simulat.* 43, 313–321.

Hall, O., Hay, G., Bouchard, A., Marceau, J. (2004): Detecting dominant landscape objects through multiple scales: An integration of object-specific methods and watershed segmentation. *Landsc. Ecol.* 19, 59–76.

Hay, G., Marceau, J., Dube, P., Bouchard, A. (2001): A multi-scale framework for landscape analysis: Object-specific analysis and upscaling segmentation. *Landsc. Ecol.* 16, 471–490.

Hay, G., Blaschke, T., Marceau, D., Bouchard, A. (2003): A comparison of three image-object methods for the multiscale analysis of landscape structure. *Photogram. Rem. Sens.* 57, 327–345.

Helming, K., Frielinghaus, M. (1999): Skalenaspekte der Bodenerosion, in Steinhardt, U., Volk, M. (eds.): Regionalsierung in der Landschaftsökologie: Forschung – Planung – Praxis. Teubner, Leipzig, pp. 221–232.

Hickey, R. (2000): Slope angle and slope length solution for GIS. *Cartography* 29, 1–8.

Hutchinson, M. (1989): A new procedure for gridding elevation and stream line data extracted from digital elevation models. *Hydrol. Proc.* 5, 31–44.

Jenness, J. (2005): Topographic Position Index extension for ArcView 3.x. Technical report, Jenness Enterprises. http://www.jennessent.com/arcview/tpi.htm

Kirkby, M., Imeson, A., Bergkamp, G. (1996): Scaling up processes and models from the field plot to the watershed and regional areas. *J. Soil Water Conservat.* 51, 391–396.

Kunert, R. (1970): Geologische Spezialkarte 1:25,000 und Erläuterungen zur Geologischen Spezialkarte 1:25.000, Blatt Könnern 4336. Zentrales Geologisches Institut, Berlin.

Lagacherie, P., McBratney, A., Voltz, M. (2006): Digital Soil Mapping: An Introductory Perspective. Developments in Soil Science, Elsevier, Amsterdam.

MacMillan, R., Pettapiece, W., Nolanc, S., Goddard, T. (2000): A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. *Fuzzy Set. Syst.* 81, 81–109.

MacMillan, R., Jones, R., McNabb, D. (2004): Defining a hierarchy of spatial entities for environmental analysis and modelling using digital elevation models. *Comput. Environ. Urban Syst.* 28, 175–200.

McBratney, A., Mendonc, M., Minasny, B. (2003): On digital soil mapping. *Geoderma* 117, 352–352.

McGarigal, K., Cushman, S., Stufford, S. (2002): Multivariate statistics for wildlife and ecology research. New York, Berlin, Springer, Heidelberg.

Merritt, W., Letcher, R., Jakeman, A. (2003): A review of erosion and sediment transport models. *Environ. Model. Software* 18, 761–799.

Möller, M. (2005): Disaggregierung von Bodeninformationen auf der Grundlage digitaler Reliefdaten, in Möller, M., Helbig, H. (eds.): GIS-gestützte Bewertung von Bodenfunktionen – Datengrundlagen und Lösungsansätze. Wichmann-Verlag, Heidelberg, pp. 67–89.

Möller, M., Helbig, H. (2005): GIS-gestützte Bewertung von Bodenfunktionen – Datengrundlagen und Lösungsansätze. Wichmann-Verlag, Heidelberg.

Möller, M., Lymburner, L., Volk, M. (2006): The comparison index: A tool for assessing the accuracy of image segmentation. *Int. J. Appl. Earth Obser. Geoinform.* 9, 311–321, doi: 10.1016/j.jag.2006.10.002.

Moore, I., Burch, G. (1986): Physical basis of the length-slope factor in the universal soil loss equation. *Soil Sci. Soc. Am. J.* 50, 1295–1298.

Park, S., van de Giesen, N. (2004): Soil-landscape delineation to define spatial sampling domains for hillslope hydrology. *J. Hydrol.* 295, 28–46.

Park, S., Vlek, P. (2002): Environmental correlation of three-dimensional soil spatial variability: A comparison of three adaptive techniques. *Geoderma* 109, 117–140.

Park, S., McSweeney, K., Lowery, B. (2001): Identification of the spatial distribution of soils using a process-based terrain characterization. *Geoderma* 103, 249–272.

Pennock, D. (2003): Terrain attributes, landform segmentation and soil redistribution. *Soil Tillage Res.* 69, 15–26.

Pennock, D., Zebarth, B., de Jong, E. (1987): Landform classification and soil distribution in Hummocky terrain, Saskatchewan, Canada. *Geoderma* 40, 297–315.

Ryan, P., Mckenzie, N., O'Connell, D., Loughhead, A., Leppert, P., Jacquier, D., Ashton, L. (2000): Integrating forest soils information across scales: spatial prediction of soil properties under Australian forests. *Forest. Ecol. Manage.* 138, 139–157.

Schmidt, J., Dikau, R. (1999): Extracting geomorphometric attributes and objects from digital elevation models – semantics, methods, future needs, in Dikau, R., Saurer, H. (eds.): GIS for Earth Surface Systems. Gebr. Borntraeger, Berlin, pp. 153–172.

Schmidt, J., Hewitt, A. (2004): Fuzzy land element classification from DTMs based on geometry and terrain position. *Geoderma* 121, 243–256.

Shary, P., Sharaya, L., Mitusov, A. (2005): The problem of scale-specific and scale-free approaches in geomorphometry. *Geografia Fisica e Dinamica Quaternaria* 28, 81–101.

Speight, J. (1988): Land classification, in Gunn, R., Beattie, J., Reid, R., van der Graaff, R. (eds.): Australian soil land survey handbook – guidelines for conducting surveys. Inkata Press, Melbourne, pp. 38–59.

Stehmann, S. (1997): Selecting and interpretation measures of thematic classification accuracy. *Rem. Sens. Environ.* 62, 77–89.

Steinhardt, U., Volk, M. (2001): Scales and spatio-temporal dimensions in landscape research, in Krönert, R., Steinhardt, U., Volk, M. (eds.): Developments and perspectives of landscape ecology. Springer, Berlin, Heidelberg, pp. 137–162.

Steinhardt, U., Volk, M. (2002): An investigation of water and matter balance on the meso-landscape scale: A hierarchical approach for landscape research. *Landsc. Ecol.* 17, 1–12.

Steinhardt, U., Volk, M. (2003): Meso-scale landscape analysis based on landscape balance investigations: problems and hierarchical approaches for their resolution. *Ecol. Model.* 168, 251–265.

Thomlinson, J., Bolstad, P., Cohen, W. (1999): Coordinating methodologies for scaling landcover classifications from site-specific to global: steps toward validating global map products. *Rem. Sens. Environ.* 70, 16–28.

Thompson, J., Bell, J., Butler, C. (2001): Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil landscape. *Geoderma* 100, 67–89.

Trefflich, A. (1997): Die Region Leipzig-Halle-Bitterfeld: Klimatologische Kennzeichnung des Gebietes, in Feldmann, R., Henle, K., Auge, H., Flachowsky, J., Klotz, S., Krönert, R. (eds.): Regeneration und nachhaltige Landnutzung – Konzepte für belastete Regionen. Borntraeger, Berlin, Stuttgart.

Volk, M., Arnold, J., Bosch, D., Allen, P., Green, C. (2007): Watershed Configuration and Simulation of Landscape Processes with the SWAT Model, in Oxley, L., Kulasiri, D. (eds.): MODSIM 2007 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, Christchurch, pp. 2383–2389, http://www.mssanz.org.au/modsim07/papers/43_s47/watersheads47_Volk_.pdf

Wielemaker, W., de Bruin, S., Epema, G., Veldkamp, A. (2001): Significance and application of the multi-hierarchical landsystem in soil mapping. *Catena* 43, 15–34.

Woodcock, C., Harward, V. (1992): Nested-hierarchical scene models and image segmentation. *Int. J. Rem. Sens.* 13, 3167–3187.

Zhan, Q., Molenaar, M., Tempfli, K., Shi, W. (2005): Quality assessment for geo-spatial objects derived from remotely sensed data. *Int. J. Rem. Sens.* 26, 2953–2974.

Zhu, A., Band, L., Vertessy, R., Dutton, B. (1997): Derivation of soil properties using a soil land interence model (SoLIM). *Soil Sci. Soc. Am. J.* 61(2), 523–533

