Geoderma 247–248 (2015) 151–160

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

# Effective map scales for soil transport processes and related process domains — Statistical and spatial characterization of their scale-specific inaccuracies

Markus Möller [a,*], Martin Volk [b]

[a] Martin Luther University Halle-Wittenberg, Institute of Geosciences and Geography, Department of Remote Sensing and Cartography, Von-Seckendorff-Platz 4, 06120 Halle (Saale), Germany
[b] UFZ — Helmholtz Center for Environmental Research, Department of Computational Landscape Ecology, Permoserstr. 15, 04318 Leipzig, Germany

ARTICLE INFO

*Article history:*
Received 22 March 2014
Received in revised form 6 February 2015
Accepted 9 February 2015

*Keywords:*
Digital soil mapping
Geodata accuracy
Mass balance index
Scale
Tillage erosion

ABSTRACT

Digital Soil Mapping (DSM) aims at the creation of reliable, reproducible and dynamic spatial soil information according to specific users' requests and demands. Positional and temporal inaccuracies as well as the question of an optimal resolution of digital elevation models (DEMs) indicate scale-related issues which represent typical challenges for DSM.

In this study, the *effective map scale* (EMS) approach is presented which enables the detection of operational scales where soil-related processes take place, the localization of corresponding process domains, as well as the statistical and spatial visualization of their scale-specific inaccuracies. The underlying algorithm can be considered as a test procedure for predictive efficiency where measurements, characterizing a soil-related process as well as a proxy variable and its scale-specific variation, are optimized. In doing so, positional and semantic inaccuracies of legacy data can be detected.

The EMS approach is applied to the example of an agricultural parcel where soil erosion by tillage is assumed. Auger samples have been taken in order to quantify the amount of soil loss and accumulation during the last 80 years in a German landscape with a complex topography and dominating loess parent material. The measurements have been related to the terrain attribute *Mass Balance Index* (*MBI*), which acts as an indicator for tillage erosion and has been varied according to both scale and soil surface complexity. The indicator *MBI* is derived from a high resolution digital elevation model and combines the basic terrain attributes *slope*, *curvature* and *vertical distance to channel network* due to their importance for tillage erosion processes. Different scale levels have been created by a region-growing segmentation algorithm. Each scale level contains discrete soil-terrain objects, represented by polygons.

The scale-related analysis of *MBI* variations and measurements has revealed a range of EMSs where process domains are visible. Their accuracies are characterized from various perspectives: (1.) The analysis of single *MBI* variants of scale and complexity by *linear regression* expresses the spatial and statistical variance of EMSs. (2.) The application of the data mining algorithm *random forest* on all the *MBI* variants of complexity per scale level leads to a spatial and statistical suppression of uncertain process domains and an emphasis of process domains, which could be predicted with a higher reliability. In this study, the process domains of accumulation could be identified on a range of operational scale levels. Due to the positional inaccuracies of auger samples and temporal inaccuracies based on overlaying long-term soil transport processes, the process domains of soil loss could not be sufficiently located.



## 1. Introduction

Spatial information about soils and their functions is mostly stored and provided by legacy soil maps of different scales. Various uncertainties are related to these maps. Apart from the fact that soil boundaries represent transition zones of soil properties (Lagacherie et al., 1996), there are misfits between original paper maps and actual, more accurate, soil-related information, like digital elevation models (DEM) or remote sensing data. In addition, soil map boundaries must often be considered as the result of a subjective — and therefore not reproducible — delineation (Carré et al., 2007a; Möller et al., 2012; Finke, 2012).

Legacy soil maps result from traditional soil sampling carried out in an empirical manner without statistical considerations (Carré et al., 2007b). The location of legacy soil samples is also often concerned by an unknown positional accuracy which can cause incorrect co-variate assignments (Finke, 2012).

Temporal-related inaccuracies of soil-related information occur due to soil profile changes over the years. Especially in intensively-used and hilly regions, land management leads to tillage erosion (Lobb, 2008),

* Corresponding author.
*E-mail address:* markus.moeller@geo.uni-halle.de (M. Möller).
*URL:* http://www.geo.uni-halle.de/geofern/mitglieder/moeller/ (M. Möller).

http://dx.doi.org/10.1016/j.geoderma.2015.02.003
0016-7061/

and thus to an impairment or leveling of soil horizons. According to Van Oost et al. (2005), contemporary soil movement is mainly caused by tillage operations due to the increased mechanization of agriculture during recent decades. This redistribution of soil within agricultural regions substantially accelerates soil profile truncation and sediment burial in specific terrain positions, and has a strong impact on soil profile evolution (DeAlba et al., 2004). Auerswald (2006) has pointed out that, especially in former studies conducted in Germany, soil loss was often wrongly assigned to water erosion while tillage or other forms of erosion were overlooked. There are still not many studies on tillage erosion in Germany, based on tracers or soil truncation mapping, but it might comprise a significant percentage on total erosion processes in landscapes with complex and hilly topography (Auerswald, 2006; Deumlich et al., 2006; Sommer et al., 2008).

Among other factors, the slope gradient and its variation (terrain curvature) strongly affect the extent of tillage erosion. While net erosion is related to convex terrain positions, deposition occurs in concave terrain positions and depressions (Van Oost et al., 2006; Heckrath et al., 2006; Papiernik et al., 2007; Lobb, 2008; Vieira and Dabney, 2009). This is especially true in landscapes with a complex topography.

Information about soil-related terrain topographies can be derived from DEMs, which are available in different spatial resolutions and accuracies (Hengl and MacMillan, 2009). Over recent years, high resolution DEMs based on airborne laser scanning have been made increasingly available. Such hrDEMs are characterized by both a high density of sampling and high vertical accuracy (Nelson et al., 2009). In Germany, hrDEMs are provided by state surveying authorities. Nowadays, they are available nation-wide. Because of their scale dependency, landscape-related processes are not always characterized properly by hrDEM-based terrain features (MacMillan and Shary, 2009). Furthermore, "[…] the choice of DEM resolution […] should always be adapted to the context of the analysis being used" (Zirlewagen and Wilpert, 2010). Thus, "the question of an optimal resolution remains to be answered" (Sørensen and Seibert, 2007).

*Digital Soil Mapping* (DSM) aims at the creation of reliable, reproducible and dynamic spatial soil information according to specific users' requests and demands (Lagacherie and McBratney, 2006). Positional and temporal inaccuracies, as well as the question of an optimal resolution, indicate scale-related issues which represent typical challenges for DSM (Carré et al., 2007a; Finke, 2012). Bishop et al. (2012) distinguish different perspectives on spatial scales. Positional and temporal inaccuracies are related to the *measurement scale* "representing the smallest area over which data can be […] represented to maintain distinguishable parts of an object." At *operational scales*, processes emerge and take place. "The scale at which data are analyzed" is considered as the *computational scale* which is related to the *geographic scale* "representing […] the size of objects".

In this study, an approach is presented which relates measurement and computational scales. Following Finke (2012), we term the resulting relations as *effective map scales* (EMS). They enable the detection of operational scales at which soil-related processes take place, the localization of corresponding process domains, as well as the statistical and spatial visualization of their scale-specific inaccuracies.

The EMS approach has been exemplified on a single agricultural parcel (Section 2.1). There, the measured changes of topsoil thickness could not be explained by results of water erosion modeling (Schmidt et al., 2009). Thus, the hypothesis should be tested that such changes in top soil thickness are mainly caused by tillage erosion. For this purpose, measurements of soil loss and accumulation have been carried out by auger sampling (Section 2.3.3). The measurements have been related to the hrDEM-based indicator *Mass Balance Index* (*MBI*) as well as different scale levels which are considered as optimizing variables. The *MBI* combines the basic terrain attributes *slope* and *curvature* due to their importance for tillage erosion processes (Section 2.3.1). Hierarchically coupled scale levels have been created by the application of a region-growing segmentation algorithm which aggregates raster cells of terrain attributes according to their neighborhood in a feature space and a raster grid (Section 2.3.2). Both variables have been statistically optimized against the observations in which linear regression analysis and a regression- and ensemble-based decision tree algorithm are applied (Section 2.3.4).

## 2. Database and methodology

### 2.1. Study area

The study area *Strenznauendorf* is situated in the German Federal State of Saxony-Anhalt near the cities of Halle (Saale) and Eisleben (Fig. 1a). The annual precipitation is about 500 mm per year. The relief and soil formation is mainly the result of glacial and peri-glacial conditions where dominant plateaus and floodplain structures and a hilly topography were shaped (Saalian glacial), as well as loess material was deposited (Weichselian peri-glacial; Ruske, 1964). The study site shown in Fig. 1b represents a subset of a plateau margin which goes down to the Saale river floodplain. The color composite in the map combines terrain attributes (see Section 2.3.1). While the floodplain (shown in green) is covered by fluvial sediments and corresponding

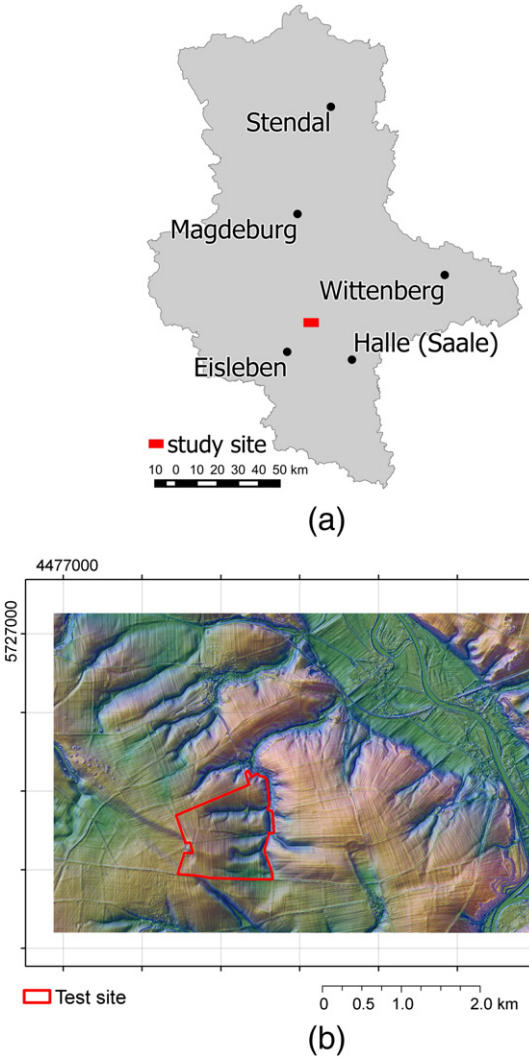



**Fig. 1.** Location of the study area within the German Federal State of Saxony-Anhalt (a) and of the test parcel (b; projection: DHDN/three-degree Gauss–Kruger Zone 4 (EPSG code 31468); see Spatialreference, 2013). The color composite combines the terrain attributes *vertical distance to channel network* (red), *total curvature* (green) and *slope* (blue; see Section 2.3.1).

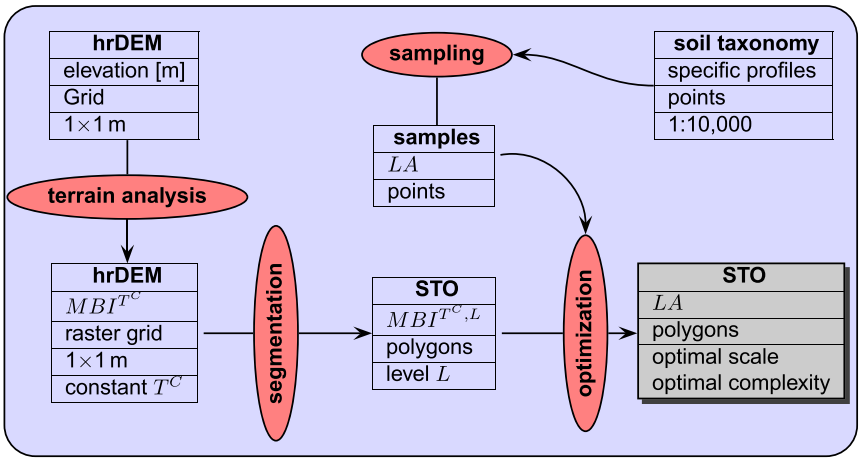



**Fig. 2.** Workflow ($MBI^{T^C,L}$ — *MBI* variants of the transfer constant $T^C$ and the scale level *L*; *LA* — soil loss and accumulation; STO — soil-terrain objects; hrDEM — high resolution digital elevation model).

soils, the plateau (shown in orange) is dominated by Weichselian loess. The plateau margin is structured by valleys (shown in blue) which mainly contain colluvic loess material. Following the nomenclature of the World Reference Base (IUSS Working Group WRB, 2007), haplic regosols and chernozems from loess or colluvic regosols from colluvic loess material are predominant, depending on the terrain positions (Möller et al., 2008, 2012). Within the study area, we selected a parcel with a size of about 1.2 km$^2$, which does not differ in relation to the parent material loess but which is characterized by a complex topography.

### 2.2. Soil database

The German Soil Inventory was introduced in 1937, according to a law in 1934 (in German: "Bodenschätzungsgesetz"). This inventory describes farmland soil properties related to agricultural yields. The demarcation of the soil taxation polygons results from a detailed *Pürckhauer* auger sampling (about 50 × 50 m sample raster). Each polygon is characterized by *one* representative soil profile (in German: "Bestimmender Grablochbeschrieb"). The profiles are mapped according to a strict mapping guide, the profile descriptions can be considered accurate. This especially concerns the detection of soil horizon boundaries which are determined with an accuracy of 5 cm (Altermann et al., 2002).

The horizons and associated attributes of all representative soil profiles are listed in inventory books (in German: "Schätzungsbücher"; Altermann et al., 2002). The German Soil Inventory database describes various soil parameters down to 1 m depth. Today, most of these data are digitized and managed by the German local tax or soil office authorities. In Saxony-Anhalt, the digitization of 400,000 soil taxation polygons and the handwritten attributes of corresponding representative soil profiles were carried out between 1998 and 2003 (Gutteck, 1999; Altermann et al., 2004).

### 2.3. Workflow

The workflow shown in Fig. 2 can be distinguished in four parts. Based on a hrDEM, basic terrain attributes are derived and combined to variants of the *Mass Balance Index* ($MBI^{T^C}$; Section 2.3.1), enabling the detection of tillage erosion-affected process domains. The application of a region-growing segmentation algorithm leads to user-defined aggregation levels of terrain attributes (*L*; Section 2.3.2). The combined index variants and scale levels ($MBI^{T^C,L}$) are related to sampled observations of topsoil loss and accumulation (*LA*; Section 2.3.3) by the application of an optimization algorithm (Section 2.3.4).

The calculation of all the terrain attributes was performed within the SAGA GIS and RSAGA environments (Brenning, 2008; Olaya and Conrad, 2009). The applied segmentation algorithm is implemented within the software eCognition (Trimble, 2012). Statistical operations and visualizations were realized within the statistical environment R (R Development Core Team, 2012).

#### 2.3.1. Terrain analysis: digital elevation model and calculation of terrain attributes

For the German Federal State of Saxony-Anhalt, a laser hrDEM[1] is available. This hrDEM shows a geometric resolution of 1 × 1 m and a vertical accuracy of ±0.15 m. The hrDEM also contains furrows which can affect segmentation results (Fig. 1b). Thus, a multi-directional filter operation has been applied in this study which enables the user-defined smoothing of such elements (Lee, 1980).

The terrain attribute *Mass Balance Index* (*MBI*) was used to characterize the process domains of soil loss and accumulation. Negative *MBI* values represent areas of net deposition, such as depressions, positive *MBI* values indicate areas of net erosion, such as convex hill slopes, and *MBI* values close to 0 refer to areas with a balance between soil loss and accumulation.

Considering the assumptions made in Section 1, the *MBI* results from the combination of the transformed terrain attributes *slope* ($f(S)$) and *total curvature* ($f(C)$) according to Eq. (1). In addition, the terrain attribute *vertical distance to the channel network* ($f(D)$) is included which can be considered as weighting factor. We assume that tillage-based soil erosion risk increases with more convex curvature and slope steepness, as well as with increasing vertical distance from the channel network. Potential sediment accumulation is more likely to be assumed at concave curvatures and flat areas at a small distance from the channel network.

The standard terrain attributes *S* and *C* were calculated according to Zevenbergen and Thorne (1987) using a 3 × 3 window size. The complex terrain attribute *D* is the difference between the original *DEM* and the interpolated channel network base-level ($DEM^B$; Eq. (2)). The reciprocal transformation of the attributes ($f(S, D, C)$) results in a specific value range (Eq. (3)) which enables their controlled combination. Since the effect of $T^D$ and $T^S$ variation on modeling results has been turned out as negligible, the transfer constants $T^D$ and $T^S$ were set to the value 10 resulting in a balanced value distribution (Friedrich, 1996, 1998). Due to the importance of terrain curvature for the extent of tillage erosion

[1] http://www.lvermgeo.sachsen-anhalt.de/de/leistungen/landesaufnahme/dgm/atkis-dgm.htm.

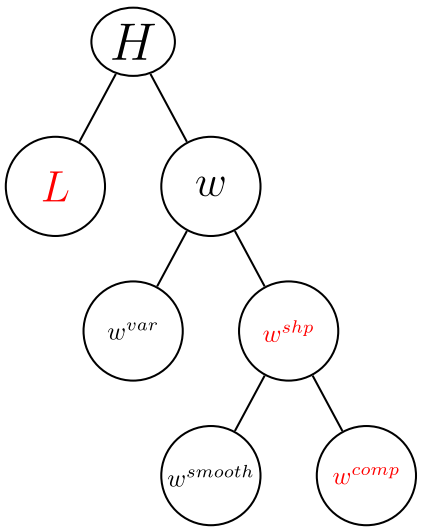



**Fig. 3.** Parameters of the *Fractal Net Evolution Approach* (FNEA). $L$ – *scale level*; $w^{var}$ – variable variance; $w^{shp}$ – shape variance; $w^{shp}$ – smoothness|$w^{comp}$ – compactness;$w^{shp} + w^{spec} = 1$; $w^{comp} + w^{smooth} = 1$. The red-colored parameters are adaptable.

(see Section 1), several $MBI^{T^C}$ variants have been calculated. Low $T^C$ values emphasize both dominant (e.g., floodplains or valleys) and subdominant terrain positions (e.g., small depressions), whereas high $T^C$ values only highlight dominant terrain positions (Möller et al., 2008). In doing so, $MBI^{T^C}$ variants can be created characterizing different states of landscape complexity.

$$MBI = \begin{cases} f(C) \times (1-f(S)) \times (1-f(D)) \text{ for } f(C) < 0 \\ f(C) \times (1+f(S)) \times (1+f(D)) \text{ for } f(C) > 0 \end{cases} \quad (1)$$

$$D = DEM - DEM^{B} \quad (2)$$

$$f(x) = \frac{x}{|x| + T^x} \quad (3)$$

with $x \in \{S, D, C\}$; $f(S, D) \in [0, 1]$; $f(C) \in [-1, 1]$

### 2.3.2. Multi-scale and hierarchical segmentation

Segmentations of the hrDEM-based continuous dataset were carried out in order to create different scale levels of soil-terrain objects (STO; Möller et al., 2012). STOs can be characterized as groups of pixels of terrain attributes which are aggregated to landform elements with scale-specific comparable heterogeneity. In this study, the region-growing segmentation algorithm *Fractal Net Evolution Approach* (FNEA) was applied which has been proven as a suitable algorithm for detecting objects with implications for soil-terrain-related issues (Drăguţ and Blaschke, 2006; Möller et al., 2008, 2012; Drăguţ and Eisank, 2011; Anders et al., 2011). The functionality of the FNEA-algorithm is described in detail by Baatz and Schäpe (2000) and Benz et al. (2004). The algorithm relies on seed pixel groups with both the smallest (here: Euclidean) distance in the pixel raster and in the feature space of the used terrain attributes. Next, the seeds grow as far as the maximum heterogeneity of grid cell values that are reached within the resulting objects.

According to Fig. 3, the objects' heterogeneity $H$ depends on the parameters *scale level* ($L$), *variable variance* ($w^{var}$) and *shape variance* ($w^{shp}$), as well as the objects' smoothness ($w^{smooth}$) and compactness ($w^{comp}$). The parameters $L$, $w^{shp}$ and $w^{comp}$ are adaptable. $w^{shp}$ and $w^{var}$ affect the form of an object. $L$ controls the degree of the pixels' aggregation. The higher the $L$, the more heterogeneous are the resulting objects which are hierarchically and spatially coupled to multi-scale object structures. This means that a superior object of a higher scale level may contain objects of a smaller scale level (Möller et al., 2008).

Already existing boundaries can act as another halting criterion. For instance, Häring et al. (2012) and Möller et al. (2012) used the boundaries of legacy soil maps whose heterogeneous soil units have been geometrically disaggregated. Here, the boundaries of the soil taxation map have not been considered in order to avoid any effect on region-growing.

### 2.3.3. Sampling: measurement of soil loss and accumulation

The *change in surface elevation* technique is a standard method for measuring soil erosion directly, and is best-suited for the hill slope scale (Stroosnijder, 2005). Accordingly, up-to-date changes can be determined by using erosion pins whose top distance to the surface characterizes the short-term extent of soil erosion and sediment deposition. However, this study focuses on long-term changes in surface elevation. Therefore, reference horizons were identified which are not affected by soil erosion processes. Such information can be derived from the German Soil Inventory database which contains for each soil taxation polygon a representative soil profile description (see Section 2.2).

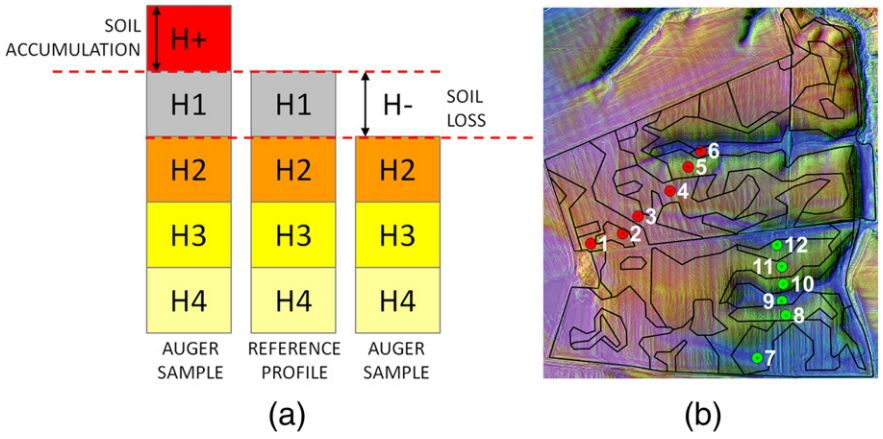
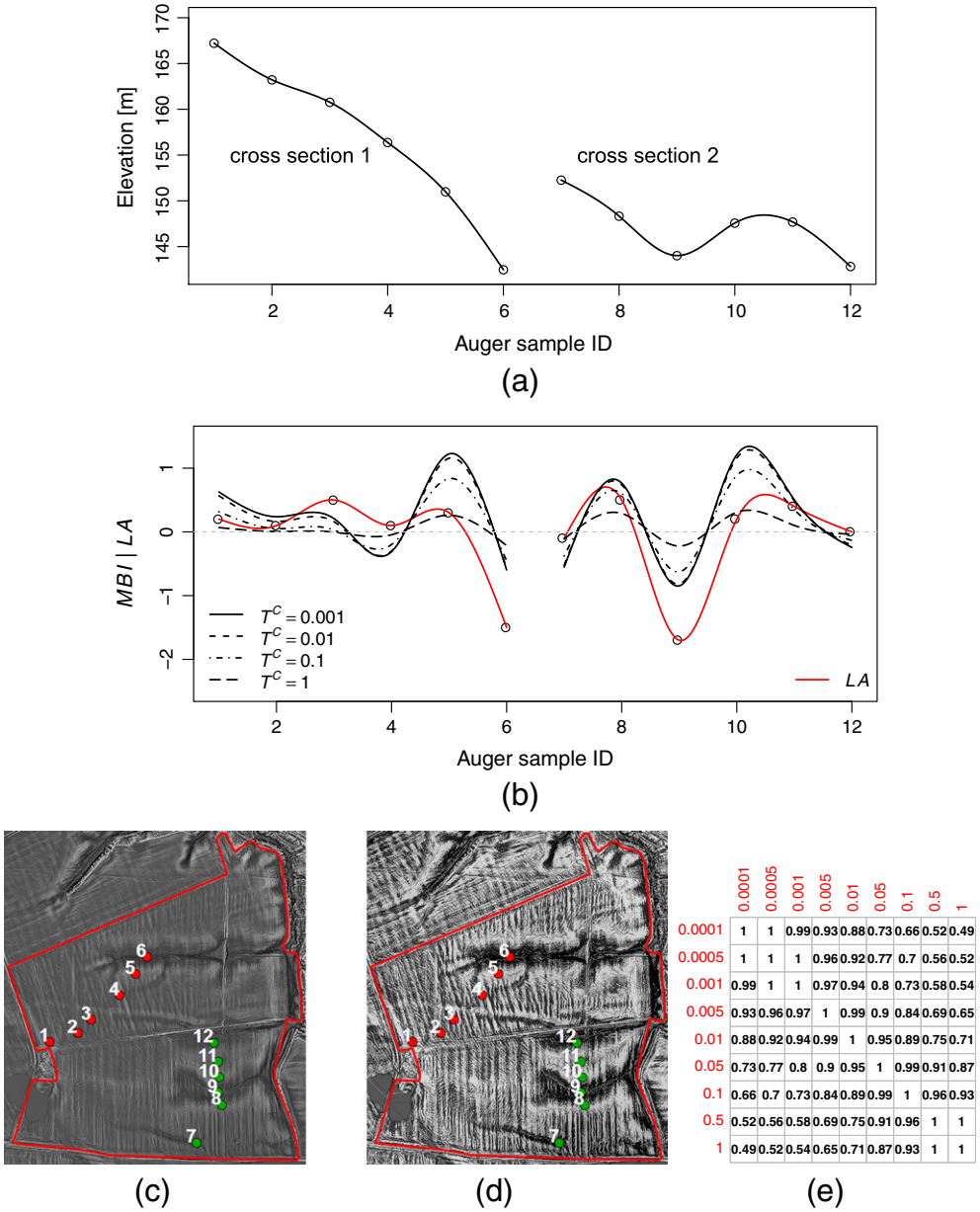



**Fig. 4.** Approach for the comparison of the representative soil profile descriptions, based on original soil taxation, and *Pürckhauer* auger sampling results for the quantification of soil loss and accumulation (a; H1…4 – Horizons 1 to 4 of a soil profile) as well as the positions of two transects of auger samples within the soil taxation polygons of the test site (b) laid over color composites of the terrain attributes *D*, *S* and *MBI* (see Section 2.3.1).

The representative soil profile descriptions were compared with profile descriptions derived from a *Pürckhauer* auger sampling (Fig. 4a). In doing so, the extent of the soil loss and the accumulation should be quantified. The 12 samples were taken along two transects inside the soil taxation polygons where the locations of the original representative soil profiles were assumed (Fig. 4b). Their positions have been set by experienced surveyors with an estimated maximal positional inaccuracy of 10 m (Schmidt et al., 2009).

### 2.3.4. Optimization of scale levels and surface complexity

Based on observations of soil loss and accumulation (see Section 2.3.3), an optimization of $L$ and $MBI^{T^C}$ variants ($= MBI^{T^C,L}$) was performed by applying *linear regression* (LR) and *random forest* (RF). While LR aims at the analysis of each single $MBI^{T^C,L}$ variant, RF is representative of data mining algorithms and stands for a regression- and ensemble-based decision tree algorithm. RF splits the feature space of the explanatory variables (here: scale-specific $MBI^{T^C}$ variants) until the resulting tree shows the best statistical correlation by minimizing the variance. Based on bootstrapped samples, RF generates a large number of independent trees (ensembles). Two thirds of the samples are used for growing trees (*in-bag* data), and one third are randomly drawn with a replacement for the calculation of error estimates by cross-validation (*out-of-bag* data) (Breiman, 2001). The RF algorithm is controlled by two main arguments ($m_{try}$ and $n_{tree}$) which have shown an impact on soil-related predictions (e.g. Grimm et al., 2008; Ließ et al., 2012). While $m_{try}$ defines the number of variables randomly sampled as candidates at each split, $n_{tree}$ determines the number of trees to grow. According to Kuhn and Johnson (2013), $n_{tree}$ is set to 1000 trees, providing stable results. Furthermore, the $m_{try}$ parameter is tuned starting with two and ending with the number of predictors. Due to possible high correlations between $MBI^{T^C}$ variants, a *principle component analysis* (*PCA*) is applied before the actual RF modeling.




| | 0.0001 | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1 | 0.5 | 1 |
|---|---|---|---|---|---|---|---|---|---|
| 0.0001 | 1 | 1 | 0.99 | 0.93 | 0.88 | 0.73 | 0.66 | 0.52 | 0.49 |
| 0.0005 | 1 | 1 | 1 | 0.96 | 0.92 | 0.77 | 0.7 | 0.56 | 0.52 |
| 0.001 | 0.99 | 1 | 1 | 0.97 | 0.94 | 0.8 | 0.73 | 0.58 | 0.54 |
| 0.005 | 0.93 | 0.96 | 0.97 | 1 | 0.99 | 0.9 | 0.84 | 0.69 | 0.65 |
| 0.01 | 0.88 | 0.92 | 0.94 | 0.99 | 1 | 0.95 | 0.89 | 0.75 | 0.71 |
| 0.05 | 0.73 | 0.77 | 0.8 | 0.9 | 0.95 | 1 | 0.99 | 0.91 | 0.87 |
| 0.1 | 0.66 | 0.7 | 0.73 | 0.84 | 0.89 | 0.99 | 1 | 0.96 | 0.93 |
| 0.5 | 0.52 | 0.56 | 0.58 | 0.69 | 0.75 | 0.91 | 0.96 | 1 | 1 |
| 1 | 0.49 | 0.52 | 0.54 | 0.65 | 0.71 | 0.87 | 0.93 | 1 | 1 |

**Fig. 5.** Cross-sections of hrDEM (a), mapped soil loss/accumulation rates and *MBI* variants for the auger profile transects (b), the visualization of two *MBI* raster image variants with $T^C = 1$ (c) and $T^C = 0.001$ (d) and a Pearson's correlation matrix of all $MBI^{T^C}$ variants (e). The smooth lines between the observations were generated by using the spline function of the *stats* package within the statistical environment R (R Development Core Team, 2012).

LR and RF regression models are calculated using the function *train*, which is implemented within the R package *caret* (Kuhn et al., 2014). The *train* function formalizes the training, pre-processing, tuning and performance assessment of a wide variety of spatial modeling techniques. In order to avoid over-fitting, each LR and RF model is based on bootstrapped training samples (25 iterations). The modeling performances are assessed by the *root mean square error* (*RMSE*) and *coefficient of determination* ($R^2$) based on observed and predicted values (Kuhn, 2008; Kuhn and Johnson, 2013).

## 3. Results and discussion

### 3.1. MBI variants and sampling results

In Fig. 5a, hrDEM cross-sections are shown which correspond to the two transects of auger samples presented in Fig. 4b. According to Fig. 5a and b, all the different $MBI^{T^C}$ variants are related to specific terrain positions: for instance, the auger samples with IDs 6 and 9 characterize positions in depressions. In contrast, sample IDs 5, 8, 10 and 11 mark positions of convex terrain positions. Fig. 5b, c and d clarify the effect of using different $T^C$ modifications for the *MBI* calculation on the example of both the $MBI^{T^C}$ cross-sections and raster images. While $T^C = 1$ only emphasizes dominant landforms like valleys and top slopes, a smaller $T^C$ value of 0.001 accentuates smaller convex and concave landforms like ridge-and-furrow patterns. The pair-wise correlation of all $MBI^{T^C}$ variants results partly in high coefficients (Fig. 5e). This concerns especially the variants of the neighboring $T^C$ values.

The *MBI* can be considered as a "terrain attribute on demand" (Möller et al., 2012). Transformed terrain attributes of specific value ranges have been combined according to assumptions about the characteristics of the soil process which should be predicted (see Section 2.3.1). In doing so, expert knowledge has been formalized and transferred to an indicator.

Fig. 5a and b visualize the mapping results of soil loss ($LA > 0$) and accumulation ($LA < 0$), the corresponding raster values of four $MBI^{T^C}$ variants, as well as the corresponding hrDEM cross-sections. Accordingly, there is a strong relation between the positions of soil accumulation and *LA* measurements (sample IDs 6 and 9). A relation also exists between the *LA* measurement of sample ID 8 and the high positive $MBI^{T^C}$ value, representing a terrain position of potential soil loss. This is not true for sample IDs 5 and 10. There, *LA* measurements are smaller than expected. Other *LA* measurements (e.g. sample IDs 3 and 4) are higher than anticipated. The discrepancies between *LA* measurements and $MBI^{T^C}$ can be caused by positional inaccuracies (see Section 2.3.3) or overlaying long-term soil transport processes which is discussed in Section 3.3.

### 3.2. Segmentation

Based on the hrDEM with about 1,500,000 raster cells (level 0), 20 segmentation levels ($L$) have been created from the low correlative terrain attributes *D*, *S* and *MBI*, applying a step-wise scale level increase. The segmentation process leads to a logarithmic decrease of the scale-specific STO number *N* (Fig. 6a). The ongoing STO aggregation results in the emergence and disappearance of specific process domains (Möller et al., 2008). On the scale level $L = 19$, for instance, dominant colluvial process domains appear as single STOs. On the scale level $L = 20$, these STOs begin to merge with surrounding STOs (Fig. 6c).

The resulting multi-scale STOs are hierarchically related which is visualized in Fig. 6b and c on the example of the segmentation levels 20, 19, 17, 8 and 4. There, black-colored STO boundaries represent superior *parent* objects which are disaggregated by smaller red-colored *child* and yellow-colored *grandchild* objects. In doing so, the STOs follow a concept of spatial resolution which is controlled by the scale-specific heterogeneity of objects. Thus, the arbitrariness of the geometric data partitions is minimized which is known as *Modifiable Area Unit Problem* (MAUP; Openshaw, 1984; Burnett and Blaschke, 2003; Drăguţ and Eisank, 2011). The MAUP phenomenon occurs, for instance, by applying simple raster-based re-sampling techniques.

The STOs of small-scale levels (e.g., level $L = 4$ in Fig. 6b) can be considered as elementary topographic landform elements which are "relatively homogeneous" with respect to the input terrain attributes Minár and Evans (2008). Such basic landform elements or facets can be classified as landform types by applying heuristic rules or relative class definitions, building "taxonomical hierarchies of geomorphological landforms" (MacMillan and Shary, 2009). This bottom-up classification approach is different from multi-scale object structures which initially represent geometric hierarchies (Drăguţ and Eisank, 2011). However, multi-scale object structures have also been used for the thematic classification of scale-specific landforms by applying optimization procedures (e.g. Möller et al., 2008; Anders et al., 2011; d'Oleire Oltmanns et al., 2013). This study follows a similar approach. Instead of classifying landforms and using thematic training and validation data, the optimization procedure is based on quantitative measurements which have been related to scale-specific variations of STO-related $MBI^{T^C,L}$ means.

### 3.3. Optimization

The optimization results are summarized in Figs. 7 and 8. Fig. 7 compares the scale-specific accuracy metrics ($R^2$, *RMSE*) of the univariate LR and the corresponding RF variants. In Fig. 8, selected spatial prediction results are visualized which are related to the presented STO boundaries of Fig. 6. Each column in Fig. 8 contains a scale-specific LR and RF prediction as well as the associated scatter plots of mapped and predicted *LA* values. In the following, we address the key findings.

Fig. 7a and b shows the LR accuracy metrics for each single $MBI^{T^C,L}$ variant. The modeling result based on the original raster grid (level $L = 0$) is characterized by poor coefficients of determination with

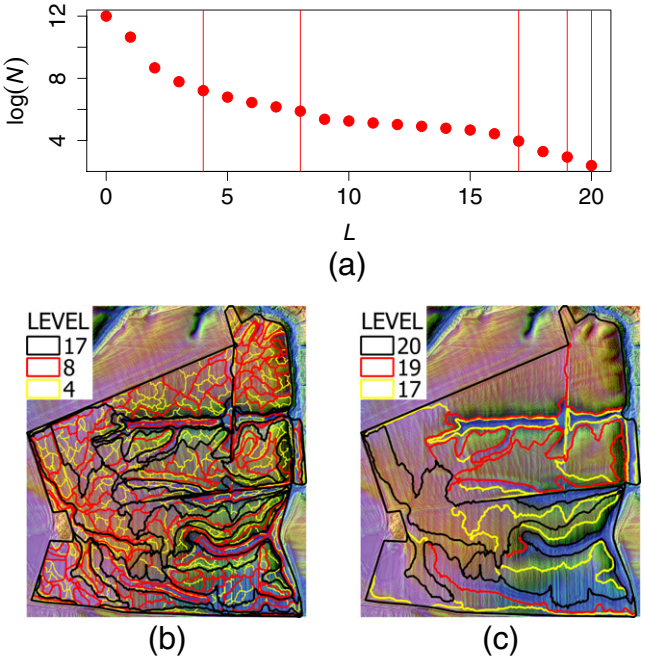



**Fig. 6.** Relation between scale level *L* and STO number *N* (a) as well as examples of different object scale levels laid over color composites of the terrain attributes *D*, *S* and *MBI* (b–c; see Section 2.3.1). The vertical lines in (a) refer to segmentation levels and corresponding scale-specific STO boundaries (b–c). Black-colored STO boundaries represent superior *parent* objects which are disaggregated by smaller red-colored *child* and yellow-colored *grandchild* objects.

$R^2 < 0.40$. Then, the $R^2$ values increase steadily until $L = 17$ and $L = 18$ with a maximal $R^2$ value of 0.68. The following level $L = 19$ is characterized by a slight $R^2$ decrease, before being abruptly dropped to values of $R^2 \leq 0.26$ at level $L = 20$.

Concerning the scale-specific effect of $MBI^{T^C}$ variants, the $R^2$ metrics increase from larger to smaller $T^C$ values at lower scale levels ($L \leq 3$). From level $L = 4$, the trend turns to the opposite direction until level $L = 9$. Between $L = 10$ and $L = 18$, the $R^2$ values are higher, with a tendency towards maximal values around the $T^C$ variants of 0.1 and 0.05. Compared to the $L$ variable, the $T^C$ variable has a lower impact on the $R^2$ value variation.

In Fig. 8a to e, all the mapped and thematically clustered prediction results of five $MBI^{T^C,L}$ variants are based on the variable $T^C = 0.1$, but they differ in the used scale levels ($L = 4$, $L = 8$, $L = 17$, $L = 19$, $L = 20$). The corresponding scatter plots show relatively high residuals. They are especially related to the process domains of soil loss, which are also characterized by high spatial variation (Fig. 8k to o). This phenomenon is generally true for all $MBI^{T^C,L}$ variants. The *RMSE* variations

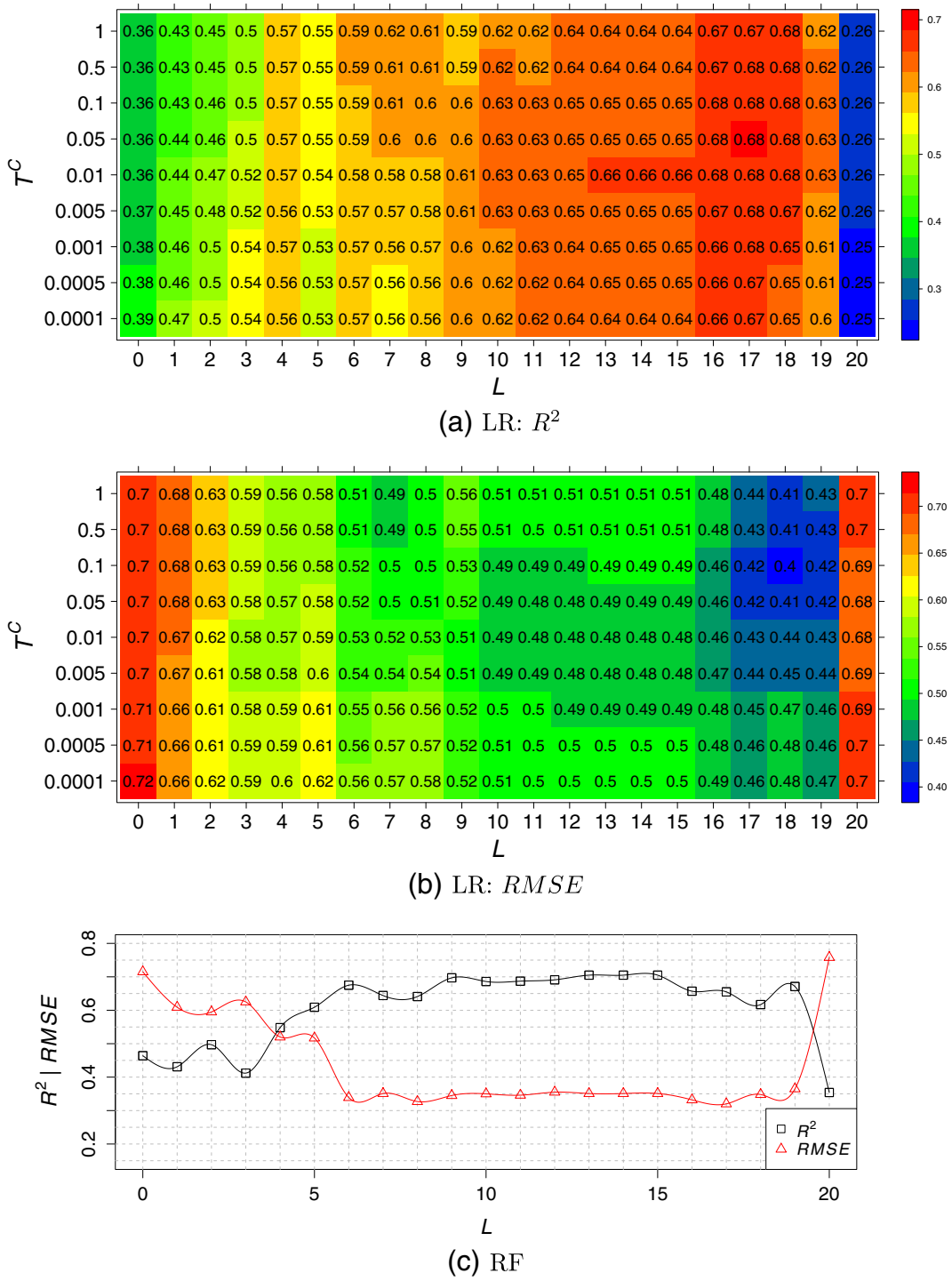



**Fig. 7.** Scale-specific *coefficients of determination* ($R^2$) and *root mean square errors* (*RMSE*) for LR (a, b) as well as RF modeling results (c).

principally follow the $R^2$ pattern in the opposite direction, ranging from 0.72 at level $L = 0$ to 0.4 at level $L = 18$ (Fig. 7b).

The scale-specific consideration of all the $MBI^{T^C}$ variants by the RF data mining algorithm has led to a significant *RMSE* reduction (Fig. 7c). In doing so, the RF algorithm smooths less certain relations. This is visible in both the spatial predictions (Fig. 8f to j) and the corresponding scatter plots (Fig. 8k to o). In contrast to LR, a relatively low soil loss is predicted which is characterized by a small spatial variation. However, the dominant process domains of accumulation have been detected by both algorithms in a similar manner as regards the value range and spatial extent. In other words, the presented optimization procedure generates different perceptions of geodata inaccuracies. While the analysis of single $MBI^{T^C,L}$ relations expresses the spatial and statistical variances of the related geodata, the RF algorithm only emphasizes process domains (here: of soil accumulation) which could be predicted with a higher reliability. This is made possible by the fact that $MBI^{T^C}$ variants represent different degrees of surface complexity regarding convexity and concavity.

Contradictions between terrain positions with expected high soil loss and corresponding measurements (see Section 3.1) have also been observed on other test sites in Saxony-Anhalt by Altermann et al. (2004) and Schmidt et al. (2009). They assume that they might mainly be the result of overlaying long-term soil transport processes caused by water and tillage erosion. The distinction of both types of erosion requires specific tracer-based measurements (Lobb, 2005) and modeling techniques simulating landform evolution (DeAlba et al., 2004; Van Oost et al., 2005; Follain et al., 2006; Vieira and Dabney, 2009). Another reason may be related to the fact that the exact positions of original soil taxation samples cannot be precisely located. Both kinds of possible inaccuracies reflect a common situation in DSM whereby soil scientists have to deal with legacy data of unknown positional and semantic accuracies (Finke, 2012). Here, this issue is addressed by the allocation of samples to STOs of different aggregation levels.

STO boundaries are demarcated in a reproducible manner, which we consider to be a key advantage in comparison to subjective delineations. The boundaries represent scale-specific and statistically significant changes regarding the value distributions of the used terrain attributes. Anders et al. (2011) distinguish STOs of distinct and fuzzy boundaries. Fuzzy boundaries are related to gradual changes of STO value means and thematic accuracies. Distinct boundaries are detectable over a relatively large range of scale levels. At a specific scale level, the merging with neighboring objects leads to an abrupt variation of STO means and to a decreasing prediction accuracy. Such scale jumping can be considered as an indicator for the disappearance of meaningful objects. This refers to concepts of hierarchical landscape structuring according to scale levels, which are related to specific processes and process domains (Steinhardt and Volk, 2002; Volk et al., 2010). In this study, a similar phenomenon could be observed between the levels $L = 19$ and $L = 20$. Here, the $R^2$ value drop is caused by the disappearance of STOs representing process domains of accumulation which can be considered as an expression of "scale dependencies of form and process" (Bishop et al., 2012).

The prediction results represent EMSs where the temporal and positional inaccuracy of observations (measurement scale) is assigned to different aggregation levels of hrDEM-based terrain attributes (computational scale). The scale-related analysis has revealed a range of operational scales where process domains are visible. It is noteworthy

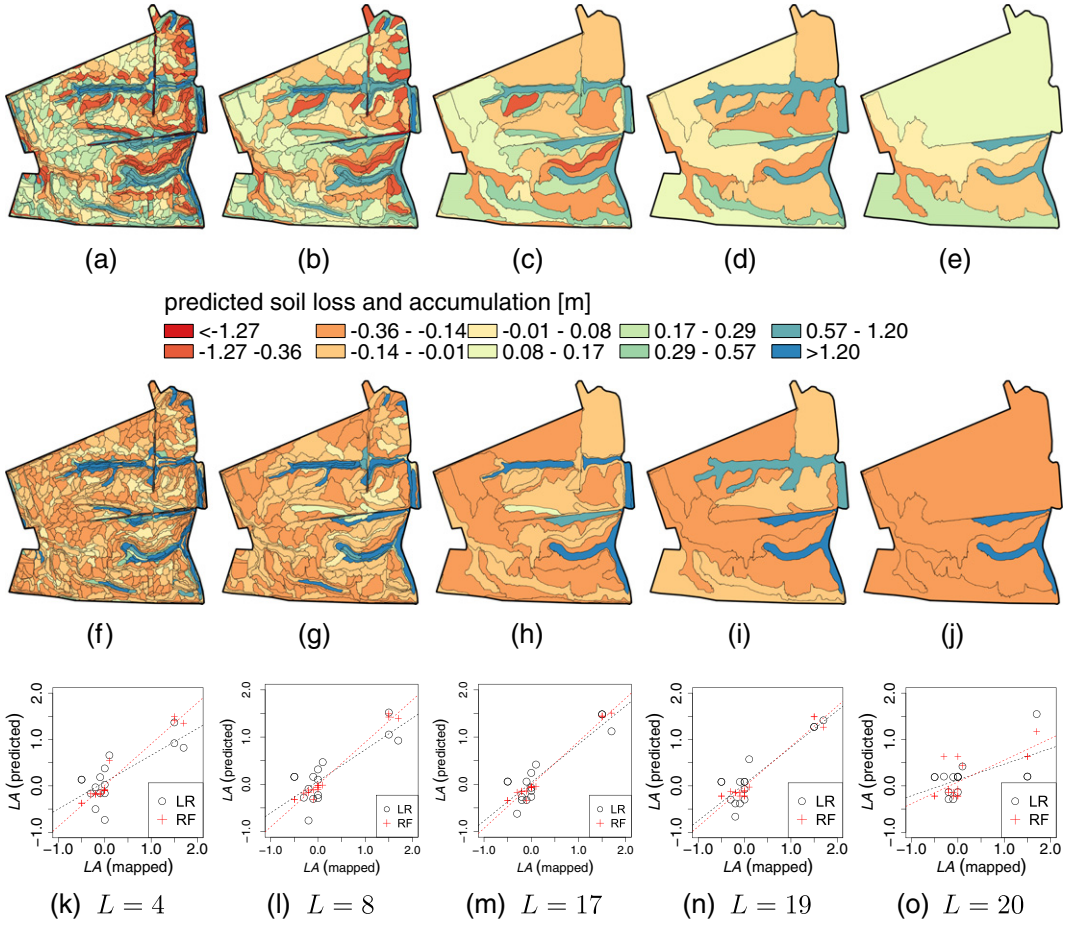



**Fig. 8.** Predicted soil loss and accumulation (*LA*) for scale levels $L = 4$, $L = 8$, $L = 17$, $L = 19$ and $L = 20$ based on *linear regression* (LR) (a-d) and *random forest* (RF) (f-i), as well as corresponding scatter plot of mapped and predicted *LA* values (k-o). The dashed lines stand for linear regression functions based on predicted and mapped *LA* values.

that the original hrDEM and the smallest scale levels are characterized by poor prediction results. This is confirmed by Sørensen and Seibert (2007) and Zirlewagen and Wilpert (2010), who have shown that a higher DEM resolution does not necessarily lead to a more accurate modeling of landscape-related processes. Hence, each modeling requires the determination of most suitable scale levels represented by specific raster grid resolutions (MacMillan and Shary, 2009) or, like in this study, by most effective map scales where operational scale levels emerge.

The scale-specific accuracy metrics place potential users in a position to select the most appropriate EMS. Apart from the desired accuracy, an important criterion is the geographic target scale. As shown in Möller et al. (2008) and Möller et al. (2012), segmentation levels and their related distribution of object sizes can be principally assigned to geographic scale levels. Fig. 8a to j shows classes of the predicted values which result from a thematic clustering. There, the blue colored process domains of accumulation can be considered as stable structures although their spatial coverage varies on different scale levels. Since the applied techniques of region-growing segmentation and thematic clustering are different from the traditional cartographic approaches of semantic and geometric generalization, the question remains how accurate the class boundaries are, considering the cartographic principles of map representations (Bishop et al., 2012).

## 4. Conclusion and outlook

In this study, the *effective map scale* (EMS) approach is presented which enables the detection of operational scales, the localization of process domains, as well as the statistical and spatial visualization of their scale-specific inaccuracies. The underlying algorithm can be considered as a test procedure for "predictive efficiency" (MacMillan and Shary, 2009) where measurements, characterizing a soil-related process, as well as a proxy variable and its scale-specific variation are related and assessed. In doing so, positional and semantic inaccuracies of legacy data can be detected. This information can be used to localize regions, where additional data are needed, and to adapt sampling strategies.

The EMS approach has been exemplified on a single agricultural parcel on which the hypothesis was tested that changes in top soil thickness are mainly caused by tillage erosion. Auger samples have been taken in order to quantify the amount of soil loss and accumulation during the last 80 years. The measurements have been related to the terrain attribute *MBI* which is used as an indicator for tillage erosion, and which has been varied according to both scale and soil surface complexity. Process domains of accumulation could be identified on a range of operational scale levels. Due to positional inaccuracies of auger samples and temporal inaccuracies based on overlaying long-term soil transport processes, process domains of soil loss could not be sufficiently located.

In order to distinguish positional and temporal inaccuracies, we are planning to apply an algorithm by Grimm and Behrens (2010) who have simulated a shifting of sample locations within a fixed neighborhood range around their original location. While in this study the selection of auger samples was mainly driven by expert knowledge, we are now testing algorithms for the automatic detection of representative soil samples and transects considering both terrain attributes and soil taxation polygons (Carré et al., 2007b; Behrens et al., 2008; Brus et al., 2011).

## Acknowledgments

We would like to thank Dr. Michael Steininger and Dr. Oliver Rosche from the Central German Institute for Land Evaluation and Soil Conservation (*in German*: Mitteldeutsches Institut für angewandte Standortkunde und Bodenschutz). They performed the auger sampling and were always open to sharing their "expert knowledge". The study was supported by the German Ministry of Economics and Technology and managed by the German Aerospace Center (Contract no.: FKZ 50EE1230). We are also very grateful to the editor A.B. McBratney and the reviewers who provided valuable advice on how to significantly improve the manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.geoderma.2015.02.003. This data include Google map of the most important areas described in this article (see Fig. 8a).

## References

Altermann, M., Benne, I., Betzer, H.-J., Capelle, A., Diemann, R., Engel, E., Etzkorn, K., Hartmann, K.-J., Keil, B., Niehöster, U., Pfeiffer, E., Sauer, S., Schäfer, K.-P., Vogler, E., Zeitzt, J., 2002. Nutzung der Bodenschätzung zur Bewertzung von Böden. Mitt. Dtsch. Bodenkdl. Ges. 99, 37–97.

Altermann, M., Steininger, M., Rosche, O., 2004. Zur Ableitung der Bodenausbildung (Substrataufbau, Bodentyp) aus den Unterlagen der Bodenschätzung (digitalisierte Grablochbeschriebe einschließlich räumlicher Zuordnung) für das Land Sachsen-Anhalt. Tech. rep., Mitteldeutsches Institut für angewandte Standortkunde und Bodenschutz im Auftrag des Landesamtes für Umweltschutz Sachsen-Anhalt, Halle (Saale).

Anders, N.S., Seijmonsbergen, A.C., Bouten, W., 2011. Segmentation optimization and stratified object-based analysis for semi-automated geomorphological mapping. Remote Sens. Environ. 115, 2976–2985.

Auerswald, K., 2008. Germany. In: Boardman, J., Poesen, J. (Eds.), Soil erosion in Europe. John Wiley & Sons, Chichester, West Sussex, England, pp. 213–230.

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 vol. 12. Wichmann Verlag, Karlsruhe, Germany, pp. 12–23.

Behrens, T., Schmidt, K., Gerber, R., Albrecht, C., Felix-Henningsen, P., Scholten, T., 2008. Transect sampling for digital soil sensing and mapping. Geophys. Res. Abstr. 10, 8087.

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. ISPRS J. Photogramm. Remote Sens. 58, 239–258.

Bishop, M., L.A., J., Shroder, J., Walsh, S., 2012. Geospatial technologies and digital geomorphological mapping: concepts, issues and research. Geomorphology 137, 5–26.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brenning, A., 2008. Statistical geocomputing combining R and SAGA: the example of landslide susceptibility analysis with generalized additive models. In: Böhner, J., Blaschke, T., Montanarella, L. (Eds.), SAGA — Seconds Out. Vol. 19 of Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie, pp. 23–32.

Brus, D., Kempen, B., Heuvelink, G., 2011. Sampling for validation of digital soil maps. Eur. J. Soil Sci. 62, 394–407.

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

Carré, F., McBratney, A.B., Mayr, T., Montanarella, L., 2007a. Digital soil assessments: beyond DSM. Geoderma 142, 69–79.

Carré, F., McBratney, A.B., Minasny, B., 2007b. Estimation and potential improvement of the quality of legacy soil samples for digital soil mapping. Geoderma 141, 1–14.

d'Oleire Oltmanns, S., Eisank, C., Drăguţ, L., Blaschke, T., 2013. An object-based workflow to extract landforms at multiple scales from two distinct data types. IEEE Geosci. Remote Sens. Lett. 10, 947–951.

DeAlba, S., Lindstrom, M., Schumacher, T., Malo, D., 2004. Soil landscape evolution due to soil redistribution by tillage: a new conceptual model of soil catena evolution in agricultural landscapes. Catena 58, 77–100.

Deumlich, D., Funk, R., Frielinghaus, M., Schmidt, W.-A., Nitzsche, O., 2006. Basics of effective erosion control in German agriculture. J. Plant Nutr. Soil Sci. 169, 370–381.

Drăguţ, L., Blaschke, T., 2006. Automated classification of landform elements using object-based image analysis. Geomorphology 81, 330–344.

Drăguţ, L., Eisank, C., 2011. Object representations at multiple scales from digital elevation models. Geomorphology 129, 183–189.

Finke, P.A., 2012. On digital soil assessment with models and the pedometrics agenda. Geoderma 171–172, 3–15.

Follain, S., Minasny, B., McBratney, A.B., Walter, C., 2006. Simulation of soil thickness evolution in a complex agricultural landscape at fine spatial and temporal scales. Geoderma 133, 71–86.

Friedrich, K., 1996. Digitale Reliefgliederungsverfahren zur Ableitung bodenkundlich relevanter Flächeneinheiten. Vol. D 21 of. Frankfurter Geowissenschaftliche Arbeiten. University of Frankfurt, Frankfurt am Main, Germany.

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. Vol. 4 of Research Report. Office for official publications of the European Communities, European Soil Bureau, Ispra, Italy, pp. 259–266 (eUR 17729 EN).

Grimm, R., Behrens, T., 2010. Uncertainty analysis of sample locations within digital soil mapping approaches. Geoderma 155, 154–163.

Grimm, R., Behrens, T., Märker, M., Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island — digital soil mapping using random forests analysis. Geoderma 146, 102–113.
Gutteck, U., 1999. Digitalisierung von Altdaten der Bodenschätzung. Vol. 32 of Berichte des Landesamtes für Umweltschutz Sachsen-Anhalt. Landesamtes für Umweltschutz Sachsen-Anhalt, Halle (Saale), Germany.
Häring, T., Dietz, E., Osenstetter, S., Koschitzki, T., Schröder, B., 2012. Spatial disaggregation of complex soil map units: a decision-tree based approach in Bavarian forest soils. Geoderma 185 (186), 37–47.
Heckrath, G., Halekoh, U., Djurhuus, J., Govers, G., 2006. The effect of tillage direction on soil redistribution by mouldboard ploughing on complex slopes. Soil Tillage Res. 88, 225–241.
Hengl, T., MacMillan, R., 2009. Geomorphometry — a key to landscape mapping and modelling. In: Hengl, T., Reuter, H. (Eds.), Geomorphometry — Concepts, Software, Applications. Vol. 33 of Developments in Soil Science. Elsevier, pp. 433–460.
IUSS Working Group WRB, 2007. World reference base for soil resources. World Soil Resources Reports 103. FAO, Rome.
Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28 (5), 1–26.
Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, New York, Heidelberg, Dordrecht, London.
Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., 2014. Caret: Classification and Regression Training. R package version 6.0-24. http://CRAN.R-project.org/package=caret.
Lagacherie, P., McBratney, A., 2006. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. In: Lagacherie, P., A.M., Voltz, M. (Eds.), Digital Soil Mapping An Introductory Perspective. Vol. 31 of Developments in Soil Science. Elsevier, pp. 3–22.
Lagacherie, P., Andrieux, P., Bouzigues, R., 1996. Fuzziness and uncertainty of soil boundaries: from reality to coding in GIS. In: Burrough, P., Frank, A., Salgé, F. (Eds.), Geographic Objects With Indeterminate Boundaries. Taylor & Francis, London, England, pp. 275–286.
Lee, J.S., 1980. Digital image enhancement and noise filtering by use of local statistics. IEEE Trans. Pattern Anal. Mach. Intell. 2, 165–168.
Ließ, M., Glaser, B., Huwe, B., 2012. Uncertainty in the spatial prediction of soil texture: comparison of regression tree and random forest models. Geoderma 170, 70–79.
Lobb, D.A., 2005. Tillage erosion: measurement techniques. In: Lal, R. (Ed.), Encyclopedia of Soil Science. Marcel Dekker, New York, pp. 1779–1781.
Lobb, D.A., 2008. Soil movement by tillage and other agricultural activities. In: Jorgensen, S.E., Fath, B.D. (Eds.), Encyclopedia of Ecology. Academic Press, Oxford, pp. 3295–3303.
MacMillan, R., Shary, P., 2009. Landforms and landform elements in geomorphometry. In: Hengl, T., Reuter, H. (Eds.), Geomorphometry — Concepts, Software, Applications. Vol. 33 of Developments in Soil Science. Elsevier, pp. 227–254.
Minár, J., Evans, I., 2008. Elementary forms for land surface segmentation: the theoretical basis of terrain analysis and geomorphological mapping. Geomorphology 95, 236–259.
Möller, M., Volk, M., Friedrich, K., Lymburner, L., 2008. Placing soil genesis and transport processes into a landscape context: a multi-scale terrain analysis approach. J. Plant Nutr. Soil Sci. 171, 419–430.
Möller, M., Koschitzki, T., Hartmann, K.-J., Jahn, R., 2012. Plausibility test of conceptual soil maps using relief parameters. Catena 88, 57–67.
Nelson, A., Reuter, H.I., Gessler, P., 2009. DEM production methods and sources. In: Hengl, T., Reuter, H.I. (Eds.), Geomorphometry — Concepts, Software, Applications. Vol. 33 of Developments in Soil Science. Elsevier, pp. 65–85 (Ch. 3).
Olaya, V., Conrad, O., 2009. Geomorphometry in SAGA. In: Hengl, T., Reuter, H. (Eds.), Geomorphometry — Concepts, Software, Applications. Vol. 33 of Developments in Soil Science. Elsevier, pp. 293–308.
Openshaw, S., 1984. Ecological fallacies and the analysis of areal census data. Environ. Plan. A 16, 17–31.
Papiernik, S., Lindstrom, M., Schumacher, T., Schumacher, J., Malo, D., Lobb, D., 2007. Characterization of soil profiles in a landscape affected by long-term tillage. Soil Tillage Res. 93, 335–345.
R. Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria 3-900051-07-0 (http://www.R-project.org/).
Ruske, R., 1964. Das Pleistozän zwischen Halle (Saale), Bernburg und Dessau. Geologie 13, 570–597.
Schmidt, G., Möller, M., Wurbs, D., Rosche, O., Steininger, M., 2009. Qualifizierung von Daten der Bodenschätzung als Grundlage für großmaßstäbige Bodenfunktionsbewertungen. Tech. rep., Ministerium für Landwirtschaft und Umwelt des Landes Sachsen-Anhalt, Halle (Saale), Germany.
Sommer, M., Gerke, H., Deumlich, D., 2008. Modelling soil landscape genesis — a "time split" approach for hummocky agricultural landscapes. Geoderma 145, 480–493.
Sørensen, R., Seibert, J., 2007. Effects of DEM resolution on the calculation of topographical indices: TWI and its components. J. Hydrol. 347, 79–89.
Spatialreference, 2013. Catalogs of Spatial Reference Systems. Accessed on 22nd July 2013. URL. http://spatialreference.org.
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.
Stroosnijder, L., 2005. Measurement of erosion: is it possible? Catena 64, 162–173.
Trimble, 2012. Ecognition developer 8.8. Tech. rep. Trimble Germany GmbH, Munich, Germany (URL http://www.ecognition.com).
Van Oost, K., Van Muysen, W., Govers, G., Deckers, J., Quine, T., 2005. From water to tillage erosion dominated landform evolution. Geomorphology 72, 193–203.
Van Oost, K., Govers, G., De Alba, S., Quine, T., 2006. Tillage erosion: a review of controlling factors and implications for soil quality. Prog. Phys. Geogr. 30, 443–466.
Vieira, D., Dabney, S., 2009. Modelling landscape evolution due to tillage. Trans. ASABE 52, 1505–1521.
Volk, M., Möller, M., Wurbs, D., 2010. A pragmatic approach for soil erosion risk assessment within policy hierarchies. Land Use Policy 27, 997–1009.
Zevenbergen, L.W., Thorne, C.R., 1987. Quantitative analysis of land surface topography. Earth Surf. Process. Landf. 12, 47–56.
Zirlewagen, D., Wilpert, K., 2010. Upscaling of environmental information: support of land-use management decisions by spatio-temporal regionalization approaches. Environ. Manag. 46, 878–893.