

# Algorithms vs. surveyors: a comparison of automated landform delineations and surveyed topographic positions from soil mapping in an Alpine environment

Fabian E. Gruber<sup>a,\*</sup>, Jasmin Baruck<sup>a</sup>, Clemens Geitner<sup>a</sup>

<sup>a</sup>*Institute of Geography, University of Innsbruck, Innrain 52f, 6020 Innsbruck, Austria*

---

## Abstract

Landform delineation has long been used in digital soil mapping to infer soil-relevant information. While its potential as an environmental variable in soil parameter modeling has been investigated for various automated landform delineations, little research has been invested into the relationship between the delineation of landforms by algorithms based on digital terrain models (DTM) and the perception of landforms by the soil surveyor during field work. Five Open Source automated landform classification algorithms and a support vector machine classifier based on single terrain parameters are investigated with regard to their ability to replicate topographic position, at two different scales, as described by surveyors for soil profile sites in the Alpine environment of South Tyrol. We also analyse how the variation of parameters and cell size affects the distribution of the computed landforms. While a clear trend regarding grid cell size and window size can be observed with regard to the difference between macro and meso scale topographic positions, the overall classification accuracy regarding the different topographic position classes was less promising. Although some automated classifications partly resemble the surveyor's classification, a considerable number of issues remain to be investigated in order to explain the lack of reproducibility of surveyor position, some of which are linked to the Alpine environment of the study area. These include the dominance of the backslope position, the objectivity of the surveyor in rugged terrain under forest, and the fuzzy nature of classifying topographic position, especially in steep terrain. By applying

---

\*Corresponding author

Email address: [Fabian.Gruber@uibk.ac.at](mailto:Fabian.Gruber@uibk.ac.at) (Fabian E. Gruber)

a forward stepwise feature selection procedure for a model based on single terrain parameters, we show that at macro scale a regional terrain parameter (topographic wetness index) and curvatures at a coarse DTM resolution of 50 m are the most influential in distinguishing topographic position, whereas at meso scale it is the topographic position index (TPI) with a search radius of just 70 m combined with slope gradient. This study is an important first step towards consolidating topographic perception during field survey and digital terrain analysis, which, at least in Alpine terrain, still requires more investigations.

*Keywords:* topographic position, automated landform classification, soil profile site description, soil surveyor, support vector classification

---

## 1. Introduction

- Topography has always been acknowledged as an important control on the formation and, hence, distribution of soil. Schaetzl (2013) notes that by solely discussing a soil based on the description of a pit face, a soil surveyor
- 5 disregards the possibly most influential factor in its formation - the landscape. Consequently, soil description guidelines for soil classification schemes require the characterization of landform and topography of soil profile sites. For instance, when following the FAO guidelines for soil description, topography is described using the four categories major landform, relative position of the
  - 10 site within the landscape, slope form and slope angle (FAO, 2006). Similarly, the Austrian (Nestroy et al., 2011) as well as the German soil classification and mapping manuals (Ad-hoc-Arbeitsgruppe Boden, 2006) require the measurement of slope angle and a description of the landform on which the soil profile site is located at three different scales, specifically macro, meso and
  - 15 micro, relative to the surrounding 100 to 500 m, 50 to 100 m and 5 to 10 m, respectively (Englisch and Kilian, 1999). The various landform and slope position descriptions are usually performed by the surveyor while at location, supported by topographic maps and possibly aerial photographs, and based on expert rule sets as well as the surveyor's mental soil landscape model.
  - 20 Readily available digital terrain models (DTMs) of increasing resolution have led to research into landform modeling and the segmentation or stratification of DTMs into landform units. Approaches vary from expert-based rulesets to completely automated landform classifications, from supervised to unsupervised classifications, and include classifications with crisp as well

25 as fuzzy borders. The output units may be attributed with the names of landforms as mapped by a surveyor, but can also represent elementary land units that adhere to certain geometric constraints. Similar to the information required of the soil surveyor (slope angle and landform), the input variables for DTM-based landform classifications range from local terrain variables,  
30 such as slope and curvature, to regional variables like catchment area, which further describe a profile site's position in the landscape.

An early example of the local terrain variable approach was proposed by Dikau (1988), who combined plan and profile curvature as well as the radius of curvature to create a map of form elements. Pennock et al. (1987)  
35 describe a similar landform element classification based on plan and profile curvatures. Addressing the problem of appropriate scale, Wood (1996) presented an approach based on slope and multiple curvature calculations to model 6 morphometric features, and added the possibility to calculate the terrain parameters at different scales, i.e. at different window sizes. A similar  
40 analysis with regard to curvature and slope was performed by Blaszczyński (1997). Minàr and Evans (2008) proposed a concept of elementary landforms on the basis of homogeneous areas of altitude and its derivatives, separated by lines of discontinuity.

Another branch of landform classifications applies not only local, but also  
45 regional terrain attributes (Gallant and Wilson, 2000) that include information on the surrounding area of a central pixel. Schmidt and Hewitt (2004) extendet Dikau's form elements by means of fuzzy classification and introduced landscape context into the classification scheme by implementation of the TOP HAT approach (Rodriguez et al., 2002) to additionally distinguish  
50 between valleys and hills. Similarly, MacMillan et al. (2000) proposed a landform element classification based on heuristic rules and fuzzy logic. Therein, the landform elements of Pennock et al. (1987) are classified based on a semantic input model, and landscape context is added via terrain parameters that describe each cells slope position relative to its watershed. Klingseisen  
55 et al. (2008) present a landform classification based on classic local terrain attributes as well as elevation-related regional attributes. The slope class is further segmented using breakpoints in the slope profile. In a similar approach, Matsuura and Aniya (2012) also applied breakpoint detection to create subdivisions of the slope class. Hollingsworth et al. (2006) is an example  
60 for an application of regional terrain parameters in landform classification that can be linked to the hydrological regime, as they use the Static Wetness Index in their decision tree-based land unit mapping approach. Weiss

(2000) proposed an automated landform classification by combining the Topographic Position Index (TPI), which compares the elevation of a pixel to that of surrounding pixels, at large and small scale. Involving the surrounding landscape at a given search radius, r.geomorphons (Jasiewicz and Stepinski, 2013) represents a classification based on line-of-sight calculations and pattern recognition.

The previously mentioned landform classification approaches have in common that the resulting classes are attributed names, which have specific implications regarding the description or characterisation of each landform element. A different approach is unsupervised classification, wherein an algorithm separates the grid cells into classes that may or may not be afterwards provided with a name attribute related to landforms. Instead of organising a map according to certain heuristic rules, unsupervised classifications create groups of grid cells that are similar with regard to certain terrain parameters, but the group boundaries are not constrained by any existing classification system. Irvin et al. (1997) compared the results of a crisp and a continuous clustering algorithm with regard to manually delineated landforms. While Adediran et al. (2004), Arrell et al. (2007) and Burrough et al. (2000) similarly applied clustering of terrain derivatives with regard to landforms, Moravej et al. (2012) based their clusters not on the terrain attributes but on their first principle components. A different unsupervised approach is that of Iwashashi and Pike (2007), who applied a nested-means partitioning algorithm to delineate terrain types or surface-form classes.

Yet another approach to landform delineation is derived from object-based image analysis (OBIA), its principles applied to geographic information science are described by Blaschke et al. (2014). It differs to the previously described approaches, as in a first step homogenous areas with regard to certain terrain parameters are segmented, which can later be aggregated and classified into landform elements. Drăgut and Blaschke (2006) classified landform elements similar to the concepts of Dikau (1988) and Pennock et al. (1987) by applying OBIA to a group of terrain parameters which has been used in many landform classification attempts, i.e. slope, plan and profile curvature, and elevation. Gerçek et al. (2011), Mashimbye et al. (2014) and Kringer et al. (2009) provide further examples of the application of OBIA to landform delineation, the latter for use in soil mapping procedures.

Geomorphologic questions may be the main subject of interest for automated landform classification, nonetheless the relationship between landforms and soil has long been of strong scientific interest, consequently leading

to research into classifying landforms for soil mapping purposes (Schmidt and Hewitt, 2004; Herbst et al., 2012; Hughes et al., 2009; Barringer et al., 2008). MacMillan et al. (2000) well illustrate how different landforms can be interpreted in terms of soil formation and erosion, and the catena concept in  
105 soil science (Schaetzl, 2013) shows the importance of topographic position in pedogenesis. Many have highlighted the importance of topography as a soil forming factor especially in mountainous areas (Geitner et al., 2011b; Herbst et al., 2012). Consequently, it seems important, especially for future research, to consolidate the perception of topographic position in field soil survey on  
110 the one side and digital terrain analysis on the other, in order to advance the understanding of the interdependencies between soil and topography.

The aim of this study is to increase the understanding of the perception of landscape by soil surveyors. This is done by testing various landform classification algorithms with regard to their suitability to emulate the mental  
115 soil-landscape model of soil surveyors, especially with regard to the concept of topographic position. A resulting map of topographic position shall serve as an additional support for soil surveyors during field work. For each classification method we also give an overview of the influence of parameter thresholds and window sizes on the number and distribution of computed  
120 landform classes to allow ~~for~~ better understanding of how these algorithms classify DTMs. Whereas past publications presenting new landform classification approaches commonly measure the performance of their classification by visual or statistical comparison with existing thematic maps (or classifications performed by human interpretation of photo- or topographic maps  
125 for validation purposes) or by correlation with metric parameters such as soil depth, the authors deem it important to analyse the classifications performed during field work. Consequently, in this study we compare computed classifications to the topographic positions of numerous soil profile sites in South Tyrol (Italy) as mapped by soil surveyors. The Alpine environment  
130 of this research area presents an additional challenge to automated landform classification, as most previous studies have concentrated on areas suited for agriculture, thus having significantly less distinguished elevation differences as well as slope gradients.

The investigated algorithms were restricted to supervised classifications  
135 implemented in the readily available Open Source geographic information systems (GIS) GRASS GIS (GRASS Development Team, 2017) and SAGA GIS (Conrad et al., 2015). Although some classifications have been compared with the results of different algorithms, and Barka et al. (2011) compared

a number of automated landform classification algorithms with regard to  
140 correlation with soil and forest units, to our knowledge there has been no systematic comparison to point classifications by surveyors. Furthermore, in this study we investigate the topographic position not only at one, often not clearly defined, scale, but specifically at meso as well as macro scale as defined in the survey manual used by the surveyors (Englisch and Kilian,  
145 1999).

## 2. Material and Methods

### 2.1. Data

#### 2.1.1. Forestry Service point data set

The point data set used in this study was provided by the Department of  
150 Forestry Planning of the Autonomous Province Bolzano - South Tyrol (Italy). The data set includes 1,968 points with known coordinates (Figure 1), spread over all of South Tyrol, ranging from the colline to the subalpine elevation zone. Nevertheless, the flat valley bottoms are under-represented in the data set, as they are used for agriculture (mainly orchards and vineyards) rather  
155 than forestry, which is restricted to slopes and higher altitudes. Given the nature of the survey, the data set also lacks points above the treeline. This point data set is the result of various projects of the Forestry Departement, the most substantial part coming from the Forestry Survey (APB, 2006). Additionally to the silvicultural description of each site, the data set contains  
160 a soil profile description for most of the points. Part of the soil profile site information is a description of the topographic position. Figure 2 shows the distribution of these profile sites' topographic positions at different levels of generalisation for macro and meso scale.

Of the 1968 data points, 1668 profile descriptions include the topographic  
165 position at macro scale, 1468 at meso scale and only 108 additionally provide information on the micro scale topographic characteristics. Consequently,  
1721 profile sites were involved in the analysis of meso and macro scale topographic position, of which 1313 data points included information regarding both scales. The survey was performed adhering to the Forestry Survey  
170 Manual (Englisch and Kilian, 1999), where the classes for the topographic position at macro and meso scale are mostly the same, with the difference that the former is attributed to the area within 100 to 500 m of the soil profile site, while the latter is judged based on the topography of the area within 50 to 100 m. For this study, to simplify analysis, the topographic position classes

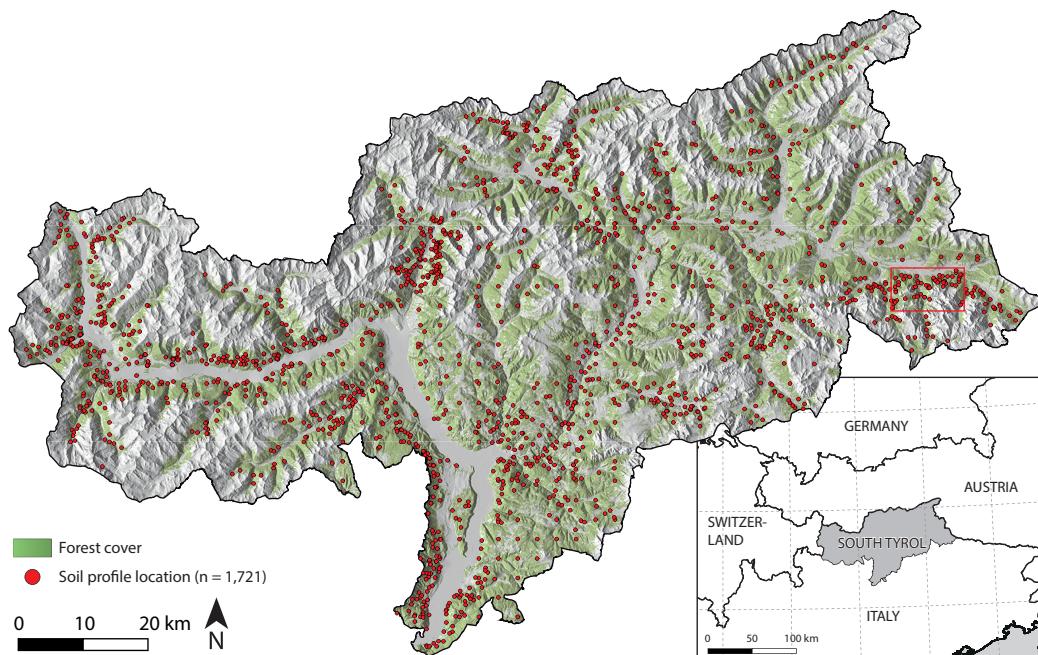


Figure 1: Hillshade of South Tyrol (Italy) showing the locations of the soil profile sites of the Forestry Service data set. The red rectangle indicates the area for which die the landform classifications are exemplified in Figure 6. Forest cover information provided by the Forestry Service of the Autonomous Province Bolzano - South Tyrol, elevation data retrieved from APB (2016).

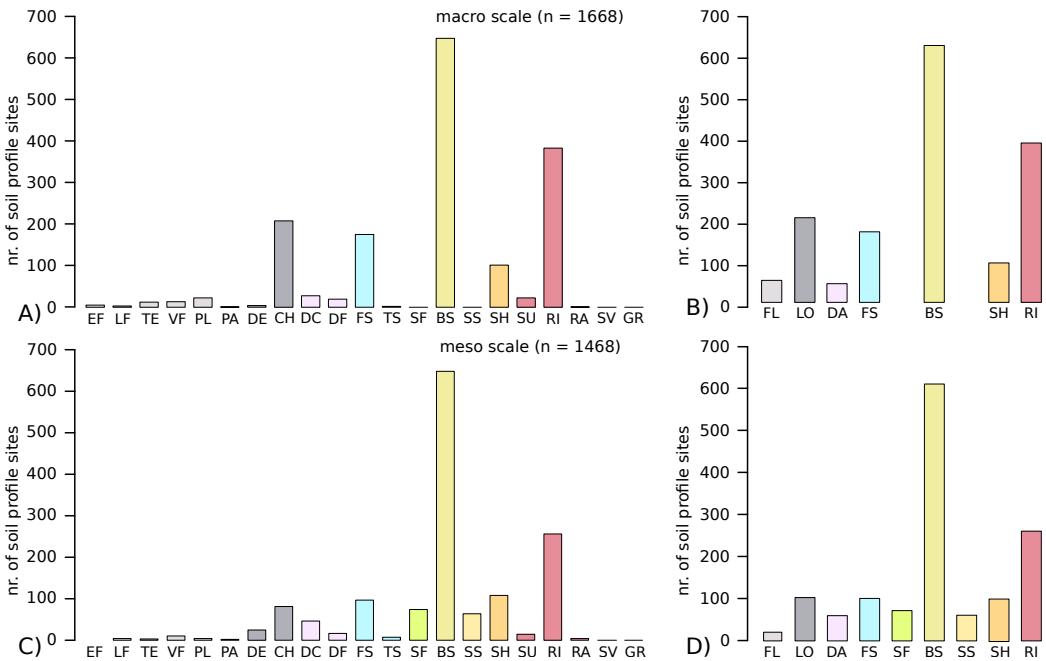


Figure 2: Distribution of the topographic positions of the profile site data points at macro (A, B) and meso (C, D) scale. The histograms on the left show the detailed topographic positions from the survey, while B and D show the generalisations applied in the study at macro and meso scale, respectively. See Table 1 for the abbreviations of topographic positions.

175 were generalised into fewer classes and classes with very small numbers of  
members were removed. Table 1 gives an overview of the original classes of  
the Forestry Survey and describes their ~~generalisation applied for this study~~  
reclassification into the generalised topographic positions FL (flat), LO (hy-  
drologically low area), DA (debris accumulation), FS (footslope), SF (slope  
180 flattening), BS (backslope), SS (slope steepening), SH (shoulder) and RI  
(ridge). The classes slope flattening and slope steepening were only analysed  
at meso scale.

### 2.1.2. Digital elevation data

185 The base input for the landform classification is a DTM with a grid size  
of 2.5 m, the result of an airborne laser scanning mission. Wack and Stelzl  
(2005) report an average achieved last-pulse point density (after preprocessing  
and filtering) of 1.3 pts/m<sup>2</sup> for valleys and areas with relevant infrastruc-  
ture, and 0.8 pts/m<sup>2</sup> for the rest of the Autonomous Province Bolzano –  
South Tyrol. The average standard deviation of heights is reported as 6.7 cm  
190 based on 8 independent check sites, each consisting of about 60 ground points.  
It is Small areas are freely available for download at the homepage of the Au-  
tonomous Province Bolzano - South Tyrol (APB, 2016), whereas the whole  
data set can be acquired for free by inquiry by email. For use in this study  
the DTM was resampled to 10, 50, 100 and 150 m to consider the influence of  
195 scale. Resampling to coarser grids was performed with average aggregation  
using the GRASS GIS 7 module 'r.resamp.stats'. The presented methodology  
was not applied to the original DTM due to computational limits regarding  
such a regional approach.

## 2.2. Methods

### 200 2.2.1. General Workflow

Figure 3 gives an overview of the workflow that was applied to each of the  
investigated landform classification approaches. For each automated land-  
form classification algorithm, a number of landform maps are computed ap-  
plying a wide range of parameter settings. The Forestry data points are then  
205 attributed with the various landform units that were computed for their spe-  
cific locations with the different settings. For each one In the next step of the  
workflow, a stepwise forward feature selection procedure using support vector  
machine (SVM) classification Cortes and Vapnik (1995) is applied for each  
automated landform algorithm in order to derive the parameter combination  
210 of each automated landform classification algorithm best suited to classify

Table 1: Overview of the topographic positions and their description for the surveyor (according to Englisch and Kilian (1999)) as well as the generalised classification (FL = flat, LO = hydrologically low area, DA = debris accumulation, FS = footslope, SF = slope flattening, BS = backslope, SS = slope steepening, SH = shoulder and RI = ridge) applied in the study. '-' signifies that the class members were removed from the analysis.

Topographic position	Abbrev.	description for surveyor	generalised class
plane	EF	expanded flat area	FL
locally flat	LF	small flat area	FL
valley floor	VF	flat area bounded by upward slopes	FL
terrace	TE	flat area bounded by one upward and one downward slope	FL
plateau	PL	flat area bounded by downward slopes	FL
depression	DE	circular concave landform	LO
pan	PA	oval concave landform	LO
channel	CH	elongated concave landform	LO
shoulder	SH	convex landform with erosion dominating	SH
footslope	FS	concave landform with accumulation dominating	FS
backslope	BS	erosion equals accumulation	BS
slope steepening	SS	slope bounded by areas with lesser slope gradient	SS
slope flattening	SF	slope bounded by areas with larger slope gradient	SF
summit	SU	circular convex landform	RI
ridge	RI	oval convex landform	RI
rampart	RA	elongated convex landform	RI
toeslope	TS	concave transition from footslope to flat	FS
debris fan	DF	gently sloping mound from gulley accumulation	DA
debris cone	DC	steep sloping mound from gulley accumulation	DA
interchanging gulleys and ridges	GR	interchanging gulleys and ridges	-
secondary valley flanks	SV	secondary valley flanks	-

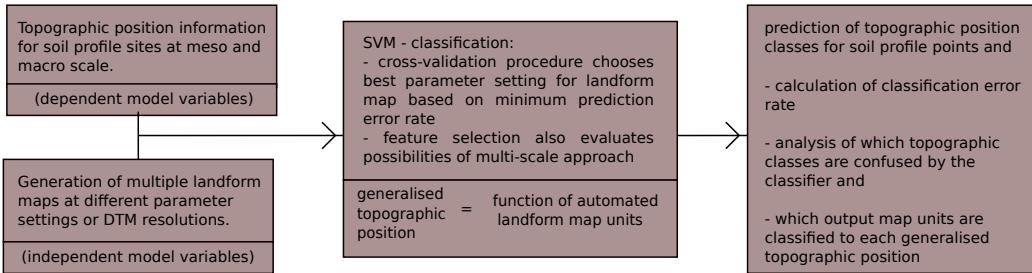


Figure 3: Overview of the workflow that was applied to each automated landform classification as well as the group of single terrain parameters. ~~The plot of cross-validated error versus number of parameters (topright) demonstrates how the addition of features (different parameters settings or terrain derivatives) may affect the prediction error. The thin colored lines represent the cross-validated feature selection for each of the 10 test samples (each consisting of 9/10 of the data points) in 10-fold cross-validation, whereas the thick black line represents the cross-validated error of the entire parameter selection procedure.~~

the landscape (represented by the DTM) according to the surveyor's mental soil-landscape model (represented by the surveyor's generalised topographic position). This means that for each automated landform classification, the parameter setting is chosen with which the SVM-classifier can classify the map units of the automated landforms into the generalised topographic positions of the surveyors with the highest accuracy. The SVM-classifier based on the landform map produced with the best parameter setting is then used to predict the generalised topographic positions for each Forestry Service point. With this model output, it is determined which map units of the automated landform delineations are mapped, or linked, to which generalised topographic positions by the SVM-classifier. This procedure is performed for the generalised topographic positions at macro scale as well as the generalised topographic positions at meso scale.

Additionally, we examine how variation of the individual model parameters affect the distribution of landform classes and how they are attributed to the various topographic positions. Applying a feature selection procedure further allows to investigate whether a multi-scale approach is beneficial for any of the different landform classifications. By also applying this procedure to single terrain derivatives, we investigate how combinations of these derivatives compare to automated classifications and highlight those terrain parameters and landform classifications that contribute most to the description and separation of landforms and slope positions.

### *2.2.2. Automated landform classification algorithms*

The following automated landform classification algorithms were performed on the DTMs applying a wide range of parameter settings. They were chosen for their open-source implementation in SAGA or GRASS GIS. Furthermore, the investigated algorithms were restricted to supervised classifications.<sup>235</sup>

*Dikau's curvature classification* (Dikau, 1988). Dikau's proposition of a geomorphographic system of classifying landforms includes a classification of landform elements based on horizontal and vertical curvature. Depending on whether a grid cell is classified as convex, concave, or planar with regard to these two directions, it is given membership to one of nine classes of landform element types. The resulting landform element map contains the classes V/V, GE/V, X/V, V/GR, GE/GR, X/GR, V/X, GE/X, X/X, where the first letter describes the profile curvature as being convex (X), concave (V) or elongated (GE), while the second letter of the class name describes the planar curvature as GR (planar), X or V. This classification is implemented in the SAGA Tool 'Curvature Classification'. The threshold value for distinguishing between planar and convex or concave is the only parameter open to variation. Hoersch et al. (2002) used this approach to correlate topography with the spatial distribution of vegetation. Jasiewicz and Stepinski (2013) evaluated and compared their landform classification algorithm 'r.geomorphon' with Dikau's curvature classification.<sup>240</sup><sup>245</sup>

*Wood's morphometric features* (Wood, 1996). The calculation of this classification was performed using the algorithm implemented in the GRASS GIS 7 add-on 'r.param.scale'. For a chosen window size it computes a number of terrain parameters by fitting bivariate quadratic polynomials. The classification of each grid cell into the classes planar, pit, channel, pass, ridge and peak is performed based on the parameters slope and curvature, specifically cross-sectional, longitudinal, maximum and minimum curvature. A variation of the slope and curvature thresholds is possible for the classification into the six morphometric features. Bolongaro-Crevenna et al. (2005) applied Wood's morphometric features and compared the distribution of the morphometric features for traditionally mapped landforms in Morelos State, Mexico, using double ternary diagrams. Ehsani and Quiel (2008) compared the results of Wood's classification algorithm with landform elements extracted using an unsupervised Artificial Neural Networks algorithms based<sup>250</sup><sup>255</sup><sup>260</sup><sup>265</sup>

on the same slope and curvature parameter maps. Both aforementioned  
270 studies discuss the choice of appropriate thresholds for slope and curvature. Ehsani and Quiel (2009) extended their previous study to include Landsat ETM+ bands in their analysis.

*Fuzzy landform elements* (Schmidt and Hewitt, 2004). While Dikau's curvature classification is based on local terrain parameters, Schmidt and Hewitt  
275 (2004) extended this approach by evaluating the resulting form elements with regard to their landscape context. Instead of horizontal curvature, tangential curvature is used for the classification of form elements in slopeing areas, analogous to Dikau's procedure, and maximum and minimum curvature are included for the classification of flat areas analogous to Wood's  
280 parametrization (Wood, 1996). Fuzzy classification is applied for differentiating sloping and flat areas as well as the different form elements. This results in a map where each grid cell is attributed one of the classes termed 'foot hollow', 'hollow', 'shoulder hollow', 'footslope', 'back slope', 'shoulder slope', 'foot spur', 'spur', 'shoulder spur', 'plain', 'channel', 'pit', 'ridge',  
285 'saddle', and 'peak'. The fuzzy classification of these landform elements is implemented in the SAGA GIS tool 'Fuzzy landform classification'. The parameters varied in the present study are the fuzzy classifiers for slope and curvatures. Schmidt and Hewitt (2004) further present an extension of this approach involving landscape context by including a TOP HAT ridge and  
290 valley detection (Rodriguez et al., 2002) to differentiate the hillslope position of the landform elements into hill, hillslope and valley. These landform elements have been used in modeling rooting depth (Schmidt and Hewitt, 2004) and in soil-landscape modeling (Schmidt et al., 2005). Hughes et al.  
295 (2009) applied the landform elements for modeling the distribution of loess in North Otago, New Zealand, with promising results regarding primary loess. Mokkaram et al. (2005) compare this fuzzy classification of slope and curvatures to TPI-based landforms, with regard to their relationship to the units of a geological map.

*Topographic position index (TPI) and TPI-based landform classification* (Weiss,  
300 2000). The TPI is defined as the difference between the elevation of a central grid cell and the mean elevation of the surrounding grid cells with any given search radius and thus defines the relative position of the grid cell along a topographic gradient (Guisan et al., 1999). Weiss (2000) developed a system that classifies a grid cell into one of the ten landform classes 'canyons', 'midslope drainages', 'upland drainages', 'U-shape valleys', 'plains', 'open slopes',  
305

'upper slopes', 'local ridges', 'midslope ridges', and 'mountain tops'. This classification table is based on the combination of a large scale TPI and a small scale TPI, i.e. one with a large and one with a small search radius. Additionally, slope is incorporated to distinguish 'plains' from 'open slopes'.  
310 On its own, the TPI has been applied in different fields, for instance Gerçek and Zeydanli (2010) used the TPI in an object-based approach to delineate landform elements for mapping land management, while Reu et al. (2013) analyse its application in an geoarcheological setting. While Mokarram et al. (2015) compared TPI-based landforms to Schmidt's fuzzy landforms, Barka  
315 et al. (2011) analysed the TPI-based landform classification as well as other landform classifications with regards to soil and forest units.

*Geomorphon-based landform classification (Jasiewicz and Stepinski, 2013).* This pattern recognition-based landform classification is implemented through the GRASS GIS 7 add-on 'r.geomorphon'. It performs line-of-sight calculations in the direction of the 8 neighboring grid cells. Depending on whether the line-of-sight ends downwards, upwards, or in the flat, this direction is attributed '-' '+' or '0', respectively. Based on this ternary pattern, each grid cell can be classified as one of the ten landform elements 'flat', 'peak', 'ridge', 'shoulder', 'spur', 'slope', 'hollow', 'footslope', 'valley' and 'pit'. The  
320 two parameters considered in this research are the search distance ( $L$ ), up to which the line-of-sight computations are performed as well as the slope angle (flat) that constitutes the threshold up to which a direction is classified as flat. The method has been applied to generate a general purpose geomorphometric map of Poland (Jasiewicz and Stepinski, 2013) as well as a  
325 physiographic map of Poland based on landscape similarity (Jasiewicz et al., 2014). A soil information system developed within the same project as the presented study also applies landform elements derived with this approach (Geitner et al., 2017).

*Terrain parameters SVM-based landform delineations with local and regional terrain parameters.* The last classification approach is a data mining approach in which a number of terrain parameters are combined in a SVM classifier in the same way that the multiscale-approach is tested for the automated classifications mentioned above. The classifier can choose from a wide range of local and regional terrain parameters calculated at different  
335 resolutions and window sizes with the GRASS GIS tool 'r.param.scale' and the SAGA GIS tool library 'Terrain Analysis'. The local parameters are  
340

slope and various curvatures and the regional parameters include wetness indices, topographic position indices, relative heights as well as valley and ridge detection algorithms implemented in SAGA GIS.

345 *2.2.3. Support vector machine classification*

The extraction of the parameters for each classification that result in a landform classifications that best corresponds to that of the surveyor, was performed by applying SVM classification in a forward stepwise feature selection procedure. SVM classification was introduced by Cortes and Vapnik 350 (1995) as a binary classifier. The basic idea is to find a hyperplane that best separates the points of two classes by maximizing the margin between those points of both classes that are closest to each other, called support vectors. A tuning parameter  $C$  allows for certain violations of this margin, thus increasing the number of support vectors by those points that are on the 'wrong' side of this (soft) margin. This simple approach for linear boundaries is extended to the nonlinear case using kernels to increase the feature 355 space. For this study, SVM classification was performed using the 'e1071' package (Meyer et al., 2014) of the statistical computing environment R (R Core Team, 2014) with a radial kernel and 10-fold cross-validation. Variation 360 of the tuning parameters gave no reason to defer from the default parameter setting. Multiclass-classification applies a one-against-one approach, thus fitting each class against each other and then attributing the appropriate class to each point. Examples for consideration and application of SVMs with regard to soil mapping are Ballabio (2009), Behrens and Scholten (2006) and 365 Rossel and Behrens (2010).

*Parameter selection procedure.* For each group of automated landform classifications as well as one group of single local and regional terrain parameters, the best fitting parameter combination was investigated by applying a stepwise forward selection approach. In a first step, the parameter 370 ~~ecombination~~setting that, applied as the only predictor variable in the SVM-classifier, best replicates ~~on its own best fits~~ the surveyor's classification was chosen based on minimum classification error. In the next step each of the remaining parameter combinations were added to the model to find the combination that best improved the original (single parameter ~~ecombination~~setting) 375 model. Ten-fold cross-validation was performed for each step as well as for the entire parameter selection procedure. The meaningfulness of this feature addition was evaluated using the 'one-standard-error-rule' (James et al., 2013)

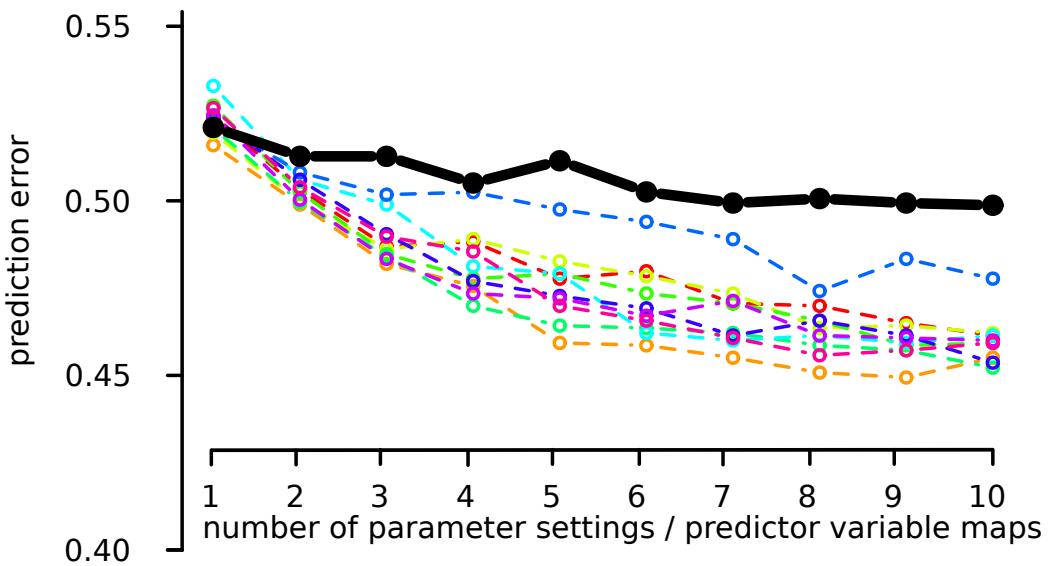


Figure 4: This plot of cross-validated error versus number of parameters (settings or predictor variable maps) demonstrates how the addition of features (different parameters settings or terrain derivatives) may affect the prediction error. The thin colored lines represent the cross-validated feature selection for each of the 10 test samples (each consisting of 9/10 of the data points) in 10-fold cross-validation, whereas the thick black line represents the cross-validated error of the entire parameter selection procedure.

as well as visual inspection of plots relating the cross-validated error to the number of parameters, as exemplified in ~~the workflow diagram in~~ Figure 34, in  
380 order to choose a minimum of parameters to simplify interpretation. Overall prediction error was chosen as accuracy measure for the parameter selection process, because the application of the kappa index requires simple random sampling (Congalton, 1991), whereas the profile sites in the data set of this study were~~s~~ sampled mostly for practical reasons. If not specified as 'cross-  
385 validated', the accuracy values in the result table of each approach are based on a model supported by the entire dataset.

### 3. Results

For each automated landform classification algorithm a SVM-classification was performed to evaluate which parameter setting produces a landform map  
390 that best reproduces the surveyor's description of topographic position at meso and macro scale. Table 2 gives an overview of which map units are attributed to which generalised topographic position by the SVM-classifier based on the best parameter setting of each automated landform classification algorithm. See Table 1 for how the generalised topographic positions FL (flat), LO (hydrologically low area), DA (debris accumulation), FS (footslope), SF (slope flattening), BS (backslope), SS (slope steepening), SH (shoulder) and RI (ridge) are composed of the classes according to Englisch  
395 and Kilian (1999).

#### 3.1. Dikau's curvature classification

400 ~~Dikau's landform element classification scheme produces a map containing the classes V/V, GE/V, X/V, V/GR, GE/GR, X/GR, V/X, GE/X, X/X, where the first letter describes the profile curvature as being convex (X), concave (V) or elongated (GE), while the second letter of the class name describes the planar curvature as GR (planar), X or V.~~

405 At macro scale, the curvature classification map that best represents the surveyors topographic position is based on the 100 m DTM and applies a curvature threshold for planar areas of 0.006. Table 2 shows which map units, computed with this specific parameter setting, are attributed to which generalised topographic position by the the SVM classifier. This map results in a  
410 cross-validated overall accuracy of 45 % ~~with the landform elements GE/GR being mapped to the topographic position BS, while the classes GE/V and VV are attributed to the position LO, V/GR to FS and the classes GE/X,~~

Table 2: Overview of how the map units of the different automated landform classification algorithms are attributed to the generalised topographic positions found in Table 1 by the SVM-classifier based on the best parameter setting of each algorithm. FL = flat, LO = hydrologically low area, DA = debris accumulation, FS = footslope, SF = slope flattening, BS = backslope, SS = slope steepening, SH = shoulder and RI = ridge

Dikau's curvature classification		Wood's features	Schmidt's fuzzy elements	TPI-based landforms	Geomorphon-based forms
macro scale					
FL	-	-	plain	plains	flat
LO	GE/V, V/V	channel	foot hollow, hollow	canyons, midslope-drainages, U-shape valleys	hollow, valley
DA	-	-	-	-	footslope
FS	V/GR	-	foot slope	-	-
BS	GE/GR	planar	back slope	open slopes	slope
SH	-	-	-	-	-
RI	GE/X, X/GR, X/X	ridge	spur, shoulder spur, shoulder slope, peak, pit	midslope ridges, mountain tops, upper slopes	peak, ridge, shoulder, spur
meso scale					
FL	-	-	plain	plains	-
LO	V/V	channel	plain, foot hollow, hollow	canyons	flat, valley
DA	-	-	-	-	-
FS	-	-	foot slope	U-shaped valleys	-
SF	-	-	-	-	-
BS	GE/GR, GE/V, GE/X, V/GR, V/X, X/GR, X/V	planar	back slope	open slopes	hollow, slope
SS	-	-	-	-	-
SH	-	-	-	-	-
RI	X/X	ridge	spur, shoulder spur, shoulder slope, peak, pit	midslope ridges, mountain tops, upper slopes	footslope, ridge, shoulder, spur

Table 3: Accuracy values (%), based on the entire dataset, of Dikau's curvature classification maps computed with the best parameter setting for each topographic position. '-' indicates that a topographic position class does not exist at the given scale. res = DTM resolution, tc = curvature threshold for plane, Ov = Overall accuracy, FL = flat, LO = hydrologically low area, DA = debris accumulation, FS = footslope, SF = slope flattening, BS = backslope, SS = slope steepening, SH = shoulder and RI = ridge

macro scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
res = 100 m & tc = 0.006	0	27	0	16	-	87	-	0	23	45
meso scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
res = 50 m & tc = 0.007	0	26	0	0	0	93	0	0	21	47

415 ~~X/GR and X/X to the topographic position RI.~~ The positions FL, DA and SH are not mapped in this parameter setting and most of the members of the classes LO, FS, SH and RI are misclassified as BS, which dominates the resulting landform classification. When DTM resolution is increased to 10 m, the position BS is distributed over all of the classes of the curvature classifica-

tion, leading to all of these classes being attributed to the class BS. With increasing grid cell size the dominance of the curvature class GE/GR increases  
420 and the overall number of mapped curvature classes decreases, resulting in the misclassification of most topographic positions as BS. An increase of plane threshold leads to a reduction of mapped curvature classes, resulting in a decline in correctly classified RI positions, while a decreasing threshold leads to a more balanced distribution of curvature classes, increasing  
425 the misclassification of the less dominant topographic position classes to BS. The addition of a further predictor mapset with a different parameter setting leads to no substantial increase in the correct classification rate of surveyed topographic position. At meso scale topographic position, a similar plane threshold (0.007) returned the best results, however at a higher DTM resolution of 50 m. ~~The curvature classes GE/GR, GE/V, GE/X, V/GR, V/X, X/GR and X/V are all mapped to the topographic position BS.~~ Only the  
430 more extreme curvature combinations V/V and X/X are mapped to ~~different~~  
~~LO and RI respectively other than BS (Table 2)~~. As is the case at macro scale, an increase in DTM resolution leads to a more diverse distribution of the points with a backslope position amongst the curvature classes, and consequently decreases the landform classification's ability  
435 to distinguish the backslope position from the other topographic positions. Table 3 gives an overview of the best parameter settings for both scales.

### 3.2. Wood's morphometric features

440 ~~Wood's landform classification algorithm yields a map containing the six morphometric features planar, pit, channel, pass, ridge and peak.~~ Table 4 gives an overview of the best parameter settings for ~~w~~Wood's morphometric features. At macro scale, the parameter setting that returns the classification which best corresponds to the surveyed classes is based on the 50 m DTM, a  
445 window size of 250 m, a curvature threshold of 0.002 and a slope threshold of 14°. This however results in a map containing only the morphometric features channel, planar and ridge, which are consequently mapped to the surveyed topographic positions LO, BS, and RI, while all other positions are not recreated. An increase in window size only leads to further misclassification to the position BS, decreasing the number of points mapped to the position LO and eliminating all RI positions. A reduction of window size or, similarly, an increase in DTM resolution leads to an increased number of morphometric feature classes. As a consequence, points with the topographic position BS are more evenly distributed on the different morphometric fea-

455 tures, diminishing the maps ability to distinguish between BS and the other  
 topographic positions, resulting in misclassification to the BS class. A re-  
 duction of the planar curvature threshold leads to similar results, while an  
 increase leads to patterns similar to those associated with an increase in win-  
 dows size. While raising the slope threshold increases the number of points  
 460 classified as BS at the cost of the remaining two classes LO and RI, a reduc-  
 tion leads to slightly less points classified as LO and a small increase of the  
 proportion of the morphometric feature ridge.

Table 4: Accuracy values (%), based on the entire dataset, of Wood’s parametric feature maps computed with the best parameter setting, for each topographic position at macro and meso scale. res = grid cell size, ws = window size, ts = slope threshold, tc = curvature threshold, Ov = Overall accuracy, FL = flat, LO = hydrologically low area, DA = debris accumulation, FS = footslope, SF = slope flattening, BS = backslope, SS = slope steepening, SH = shoulder and RI = ridge

macro scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
res = 50 m & ws = 250 m & ts = 14° & tc = 0.002	0	39	0	0	-	80	-	0	36	46
meso scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
res = 10 m & ws = 110 m & ts = 11° & tc = 0.006	0	39	0	0	0	90	0	0	25	47

At meso scale, two parameter settings tied with regard to best overall cross-validated accuracy at 47 %. While both only predict the topographic  
 465 positions LO, BS and RI, the setting with the larger window size (150 m compared to 110 m), higher slope threshold (13° compared to 11°), and lower curvature threshold (0.004 compared to 0.006), performed slightly better for the position RI, whereas the other setting performed slightly better for LO and BS. The effects of varying the parameters window size as well as slope and curvature thresholds are comparable to those at macro scale. Additional  
 470 parameter sets yield no significant improvement.

### 3.3. Fuzzy landform elements

475 Schmidt’s fuzzy landform classification expands Dikau’s curvature classification by introducing the terrain parameters minimum and maximum curvature as well as fuzzy membership functions. This results in a map where each grid

cell is attributed one of the classes termed 'foot hollow', 'hollow', 'shoulder hollow', 'footslope', 'back slope', 'shoulder slope', 'foot spur', 'spur', 'shoulder spur', 'plain', 'channel', 'pit', 'ridge', 'saddle', and 'peak'. Regarding the topographic position at macro scale, the parameter setting that produced the map that best corresponds to the surveyed position relied on the 150 m DTM and lower and upper slope thresholds for the semantic input model of 6° and 12° (Table 5). The thresholds for curvature that establish fuzzy membership to the classes straight and curved were chosen to be 0.0001 and 0.003. This parameter setting failed to reproduce correctly any of the data points of the classes DA and SH and achieved an overall accuracy of 49 %. The resulting map attributes the fuzzy class 'plain' to the topographic position FL and the classes 'foot hollow' and 'hollow' to LO. The classes 'foot slope' and 'back slope' are mapped to the positions FS and BS, respectively, while the classes 'spur', 'shoulder spur', 'shoulder slope', 'peak' and 'pit' are assigned the topographic position RI. Table 2 shows how the SVM classifier attributes the fuzzy elements to the generalised topographic positions. Reducing the DTM grid size to 50 m, while maintaining the rest of the chosen parameters, increases the number of fuzzy classes eventually mapped to BS from one to nine, which decreases the topographic positions depicted in the resulting map to the positions FL, BS and RI. An increased grid cell size of 250 m further exaggerates the membership of data points to the class BS. Changes of plus or minus 3° to the slope thresholds lead to only slightly different results. A reduction of the lower curvature thresholds results in a slight shift of classification from BS to RI, whereas an increase thereof decreases the attribution of points to the classes RI, LO and FS due to misclassification to BS. Landform maps based on a larger upper curvature threshold hardly contain any landform classes except for BS, and to a much smaller extent FL and RI.

At meso scale, there is no dominant forerunner amongst the parameter setting, with several maps achieving similar overall accuracies of 49 %. The most frequently chosen setting of the cross-validation process is based on the 50 m DTM, a slope membership function that is based on the threshold values 3° and 21°, and relatively large curvature thresholds of 0.004 and 0.006, which are quite consistent in the cross-validation process. The mapping of the fuzzy landforms to the topographic positions is comparable to that at macro scale, except that 'plain' is mapped to LO instead of FL (Table 2). As is the case in all classification approaches, the topographic position classes SF and SS are not attributed any fuzzy landform classes. A smaller grid cell size of 10 m leads to the BS points being distributed over all fuzzy landform classes,

Table 5: Accuracy values (%), based on the entire dataset, of Schmidt's fuzzy elements maps computed with the best parameter setting for each topographic position at macro and meso scale. res = DTM resolution, ts = fuzzy slope thresholds, tc = fuzzy curvature thresholds, Ov = Overall accuracy, FL = flat, LO = hydrologically low area, DA = debris accumulation, FS = footslope, SF = slope flattening, BS = backslope, SS = slope steepening, SH = shoulder and RI = ridge

macro scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
res = 150 m & ts = 6° - 12° & tc = 0.0001 - 0.003	38	36	0	32	-	81	-	0	29	49
meso scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
res = 50 m & ts = 3° - 21° & tc = 0.004 - 0.006	0	48	0	12	0	90	0	0	26	49

resulting in all of these classes being mapped to the class BS. The reduction of DTM resolution leads to similar results, however, in this setting the reason is an output map that only contains a small number of different classes, dominantly 'back slope' and, to a much smaller degree, 'plain'. Variation of curvature thresholds leads to comparable changes in landform distribution as at macro scale. Additional parameter settings lead to no improvement of overall accuracy.

### 3.4. TPI-based landform classification

The TPI-based landform classification based on the classification table by Weiss (2000) results in the map units 'canyons', 'midslope drainages', 'upland drainages', 'U-shape valleys', 'plains', 'open slopes', 'upper slopes', 'local ridges', 'midslope ridges', and 'mountain tops'. The cross-validation process at macro scale does not produce a single parameter setting decisivley better than others, but the best settings all involved an inner search radius between 60 and 100m and an outer radius between 250 and 350m. All settings preferred the 10m DTM. The single parameter combination that was most frequently at first place applied search radii of 70 and 250m. This results in a cross-validated accuracy of 48%, where the landform classes 'canyons', 'midslope drainages' and 'U-shape valleys' are mapped to the topographic position LO, 'plains' is mapped to the position FL and 'open slopes' to BS. While the topographic positions DA, FS and SH are not recreated

535 using a single TPI-based landform classification, the surveyed position 'RI' is represented by the landform classes 'midslope ridges', 'mountain tops' as well as 'upper slopes' where the landform classes are attributed to the generalised topographic positions according to Table 2. Both an increase and  
 540 a decrease of the inner search radius lead to increased misclassification of points to BS. The latter case leads to less landform classes mapped to the topographic position LO, despite more points of the landform class 'U-shape valley', the number of which strongly decreases with increasing inner search radius. Higher values for the outer TPI radius lead to more points being assigned to the landform class 'U-shape valleys', however, in this setting this  
 545 landform class is mapped to the topographic position FS instead of LO. The cross-validation process does not imply a significant increase in overall accuracy through addition of a second landform classification based on a different parameter setting.

Table 6: Accuracy values (%), based on the entire dataset, of TPI-based landform maps computed with the best parameter setting for each topographic position. res = DTM resolution, R1 = inner search radius, R2 = outer search radius, Ov = Overall accuracy, FL = flat, LO = hydrologically low area, DA = debris accumulation, FS = footslope, SF = slope flattening, BS = backslope, SS = slope steepening, SH = shoulder and RI = ridge

macro scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
res = 10 m & R1 = 70 m & R2 = 250 m	26	43	0	0	-	88	-	0	27	48
meso scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
res = 10 m & R1 = 50 m & R2 = 90 m	30	32	0	13	0	93	0	0	30	50

550 At meso scale, the feature selection procedure leads to the choice of the TPI-based landform classification map based on the 10 m DTM and inner and outer search radii of 50 and 90 m. Table 2 shows how the SVM classifier distributes the landforms computed with these parameters among the surveyed topographic position classes. In this setting, the members of the topographic position RI belong to the landform classes 'midslope ridges', 'mountain tops', and 'upper slopes'. The positions BS, FS, LO and FL are  
 555 represented by the landform classes 'open slopes', 'U-shape valleys', 'canyons' and 'plains', respectively. This setting leads to a cross-validated accuracy of 50 %, with the topographic positions SS, SH, SF, and DA not being attributed any data points. Increasing the outer radius leads to an increasing

560 number of landform classes, for instance involving data points classified as 'local ridges', and a slightly more balanced distribution of the data points amongst the different landform classes. This however leads to more landform classes being incorrectly assigned to the topographic position BS.

### 3.5. Geomorphon-based landform classification

565 The automated classification of DTMs with 'r.geomorphon' returns a map of maximum 10 landform classes, i.e. 'flat', 'peak', 'ridge', 'shoulder', 'spur', 'slope', 'hollow', 'footslope', 'valley' and 'pit'. Similarly to other classification methods, all or only part of the 10 available landform elements are mapped, depending on the research area's topography and the chosen algorithm parameters. Table 7 shows the accuracies of the best performing maps. Regarding macro scale topographic position, the parameter and input combination that, on its own, best recreates the topographic positions from field survey, is based on the 50 m DTM, a search window (L) of 400 m and a flatness threshold of 10°. The resulting landform element map can correctly  
570 classify 49 % of the points. A model of topographic position based on this parameter setting attributes the geomorphon-based landforms to the generalised position classes according to Table 2 maps the landform elements 'flat' to the surveyor's FL, 'footslope' to DA, 'hollow' and 'valley' to LO, 'slope' to BS, and the elements 'peak', 'ridge', 'shoulder' and 'spur' to RI. No points  
575 are mapped to FS and SH, as the generated map mostly incorporates the FS points into the BS class, and to a lesser degree, to LO, while SH points are dominantly classified as BS and RI. Beside FS and SH, RI also shows a strong overlap with BS. Regarding the effects of parameter variation, while a lower flatness threshold increases the amount of points classified as RI rather  
580 than BS, an increase in search window has the opposite effect and increases the misclassification of other topographic positions as BS. The addition of the results of a further classification based on a different set of parameters does not significantly improve overall accuracy.  
585

590 At meso scale, the map with the lowest overall classification error rate is based on the 10 m DTM, a search radius of 80 m and a flatness threshold of 8°. Whereas at macro scale the threshold of 10° was dominant amongst the maps with the highest correct classification rate, at meso scale 8° are preferred. In this setting the landform elements 'flat' and 'valley' are mapped to the topographic position LO, 'hollow' and 'slope' are mapped to BS, and 'footslope', 'ridge', 'shoulder' and 'spur' are mapped to RI. Again the dominant BS falsely incorporates a great amount of other classes, and FL, DA,  
595

Table 7: Accuracy values (%), based on the entire dataset, of geomorphon-based landform maps computed with the best parameter setting for each topographic position and scale. res = DTM resolution, L = search radius, fl = flatness threshold, Ov = Overall accuracy, FL = flat, LO = hydrologically low area, DA = debris accumulation, FS = footslope, SF = slope flattening, BS = backslope, SS = slope steepening, SH = shoulder and RI = ridge

macro scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
res = 50 m & fl = 10° & L = 400 m	38	49	20	0	-	81	-	0	37	49
meso scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
res = 10 m & fl = 8° & L = 80 m	0	17	0	0	0	92	0	0	38	49

FS, SF, SS and SH are not attributed any points by the model based on this map (Table 2). The overall classification rate is comparable to the one at macro scale.

600 *3.6. Terrain parameters SVM-based landform delineations with local and regional terrain parameters*

The same procedure that was applied for different parameter settings of each classification algorithm was performed for numerous local and regional terrain parameters, similarly varying parameters such as grid size and search 605 window. At macro scale, the terrain parameter that, on its own, best reproduced the surveyed topographic positions using SVM classification was the topographic wetness index, a regional terrain parameter relating a grid cell's catchment area to its slope. The resulting model correctly classifies 32 % of the topographic positions described as LO, 85% of BS and 43% of RI, which 610 leads to an overall cross-validated accuracy of 48 %. The remaining positions FL, DA, FS, and SH are not attributed any data points in the resulting map, their members being misclassified as one of the other three positions. Similar to the results of the investigated classification methods, misclassification to the topographic position BS was dominant. The addition of the terrain 615 parameter profile curvature, calculated with the 50 m DTM and a window size of 350 m, improves the cross-validated overall accuracy to 49 % by correctly classifying 38 % of the data points attributed to the topographic position FS. This parameter combination also introduces the missing topographic positions FL, DA and SH into the resulting maps. The addition of the terrain 620 parameter minimum curvature, based on the 50 m DTM and a window size

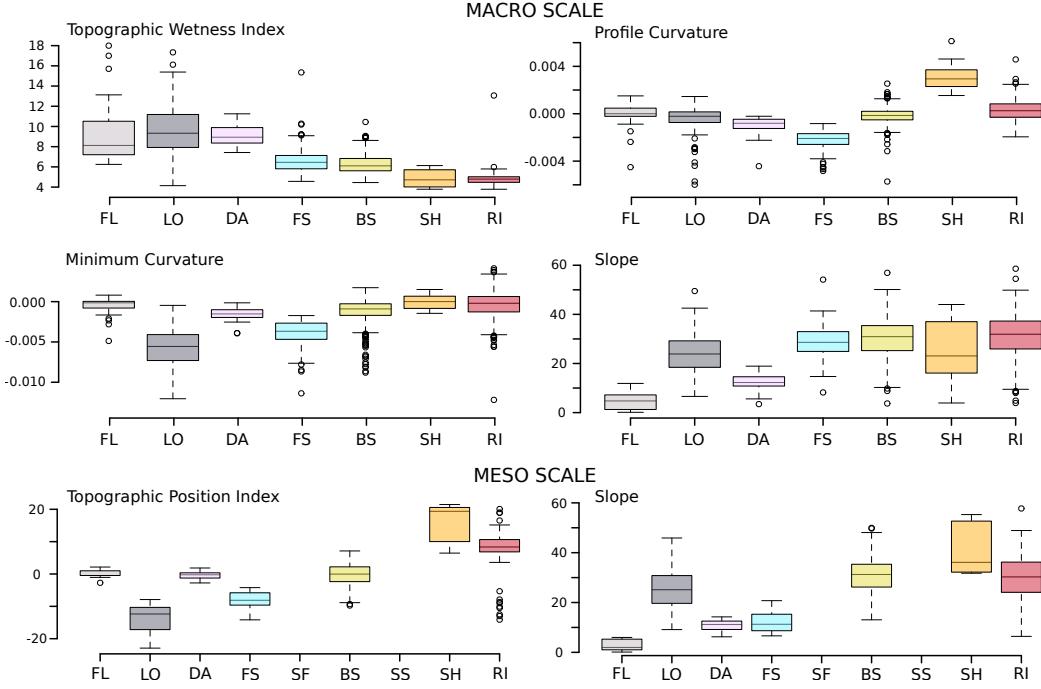


Figure 5: Boxplots of the terrain derivatives involved in the SVM classifier. They show the range and focal statistics of the terrain parameters of the points attributed to each position class by the models at macro (four parameters) and meso scale (two parameters). The horizontal line inside the boxes (interquartile range) represents the median value while the whiskers extend to 1.5 times the interquartile range.

of 250 m, further strengthens the less common classes FL, LO and DA, increasing the overall cross-validated accuracy to 51 %. Slope computed on the 10 m DTM with a window size of 150 m then leads to an increase in accuracy of FL and DA. Further features did not improve the overall accuracy of the model. Figure 5 illustrates the distribution of the terrain parameters of the data points as classified by the terrain parameter-based models for both scales, thus showing how the SVM-classifier assigns generalised topographic positions based on these parameters. For instance, while the topographic wetness index serves well to separate FL and DA from SH and RI at macro scale, profile curvature aids in distinguishing between FL and DA as well as between SH and RI.

The terrain parameter that best represents the topographic positions at meso scale is the topographic position index based on the 10 m DTM and a search radius of 70 m, which results in mapping the topographic positions

Table 8: Accuracy (%), based on the entire dataset, for maps computed with different predictor sets of terrain parameters for each topographic position at macro and meso scale. The first row shows the results using only the best terrain parameter, whereas the terrain parameters in the rows below are subsequently added to the predictor set, indicated by 'plus', consequently increasing the overall accuracy. res = DTM resolution, ws = window size, Ov = Overall accuracy, TPI = Topographic Position Index, FL = flat, LO = hydrologically low area, DA = debris accumulation, FS = footslope, SF = slope flattening, BS = backslope, SS = slope steepening, SH = shoulder and RI = ridge

macro scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
topographic wetness index res=50m	0	32	0	0	-	85	-	0	43	48
plus profile curvature res=50m ws=350m	6	26	2	39	-	84	-	12	40	51
plus minimum curvature res=50m ws=250m	21	37	13	38	-	84	-	13	40	53
plus slope res=10m ws=150m	42	36	33	35	-	85	-	12	42	55
meso scale	FL	LO	DA	FS	SF	BS	SS	SH	RI	Ov
TPI res=10m r=70m	0	31	0	0	0	94	0	5	35	51
plus Slope res=50m ws=150m	45	33	25	8	0	93	0	3	35	52

635 LO, BS, SH, and RI at an overall cross-validated accuracy of 50 %. The best performing local terrain parameter is cross-sectional curvature at a window size of 50 m, which has a similar accuracy but does not map the class SH. The cross-validated feature selection process implies that one additional terrain parameter should be added to the topographic position index. The parameter  
 640 slope based on the 50 m DTM and a window size of 150 m introduces the topographic positions FL, DA and FS to the output map of topographic positions. Table 8 shows how the addition of terrain parameters lead to the mapping of additional topographic position classes at both scales.

### *3.7. Comparison of the best parameter settings*

645 Table 9 gives an overview of how the different classification algorithms performed at different scales.

Table 9: Cross-validated overall accuracy of the best representatives of each classification algorithm as well as the best performing SVM model based on a combination of single terrain parameters.

Cross-validated overall accuracy	macro	meso
Dikau's curvature classification	0.45	0.47
Wood's morphometric features	0.46	0.47
Schmidt's fuzzy landform elements	0.48	0.48
TPI-based landforms	0.48	0.50
Geomorphon-based forms	0.48	0.48
Terrain parameters	0.51	0.52

650 To compare the similarity of the best parameter settings of the different classifications, the percentage of datapoints that were classified as the same topographic position was calculated for each pair of algorithms (Table 10) at each scale. Figure 6 gives an overview of how the different classification approaches model the topographic position at macro scale for a small area in the east of South Tyrol.

### *3.8. Data point analysis*

655 The results presented above called for an investigation into whether certain points were consistently classified wrongly in all classification approaches or if the correct classification depends on the applied method. Figure 7 shows the distribution of how often the various data points were classified correctly. It shows that points that were either always classified correctly or always

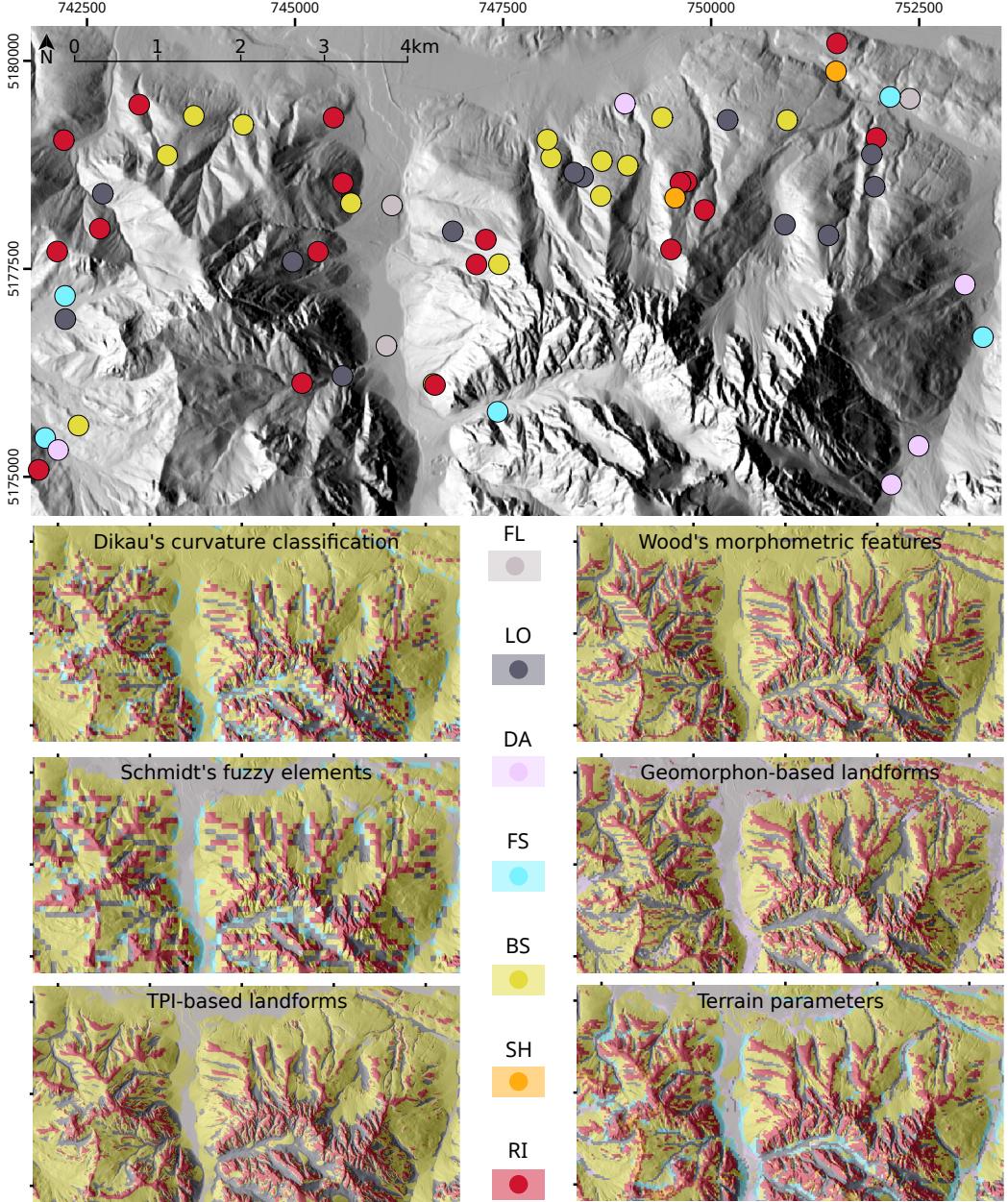


Figure 6: Comparison of the modeled topographic positions of the six evaluated classification methods at macro scale. The region displayed is indicated by the red rectangle in Figure 1. The topmost map shows the position of the sample points in this area, the colors indicate the topographic position as described by the surveyor. The smaller maps represent the topographic positions as modeled with the best fitting parameter setting for each classification approach. The coordinate reference system of the maps is WGS 84 / UTM zone 32N. FL = flat, LO = hydrologically low area, DA = debris accumulation, FS = footslope, SF = slope flattening, BS = backslope, SS = slope steepening, SH = shoulder and RI = ridge

Table 10: Similarity of the classification methods as described by the percentage of data points classified as the same topographic position. The values below the diagonal refer to meso scale, while those above represent macro scale. DCC = Dikau's curvature classification, WMF = Woods morphometric features, SFL = Schmidt's fuzzy landform elements, TBL = TPI-based landforms, GBF = geomorphon-based forms, TP = Terrain parameters

%	DCC	WMF	SFL	TBL	GBF	TP
DCC	1	0.76	0.70	0.77	0.70	0.68
WMF	0.85	1	0.68	0.74	0.74	0.69
SFL	0.83	0.89	1	0.70	0.67	0.70
TBL	0.83	0.82	0.81	1	0.74	0.71
GBF	0.81	0.82	0.80	0.82	1	0.66
TP	0.79	0.79	0.80	0.90	0.82	1

misclassified, dominate. [Figure 7](#) also underlines the dominant role of the backslope position with regard to overall classification accuracy. To investigate which terrain parameters best distinguish between the points always misclassified and always classified correctly, a SVM classification approach, comparable to the one used to distinguish topographic position, was applied.

Figure 8 gives an overview of the most influential terrain parameters regarding classification success, identified for macro and meso scale. At macro scale, these parameters are cross-sectional curvature and texture (Iwashashi and Pike, 2007), with the latter being calculated by subtracting the DTM from a median-filtered version of itself, thus representing a measure of DTM roughness. At meso scale, it is the parameters slope and topographic position which help to characterize those points that are seldomly classified correctly. Furthermore, Figure 8 provides a comparison of the slope gradient noted by the surveyor in his profile site description and that derived from the DTM.

#### 4. Discussion

All in all, the described methodology searches for parameter settings or terrain derivatives that best distinguish the topographic positions at two scales, leading to a trade-off between 1) landform maps with a large number of diverse classes, where a dominant topographic position appropriates all modelled landform classes and 2) landform maps with very few modelled landform classes which are able to largely encompass the dominant topographic position but fail to differentiate the DTM enough to consider the less dominant topographic positions. This study shows that, in the Alpine

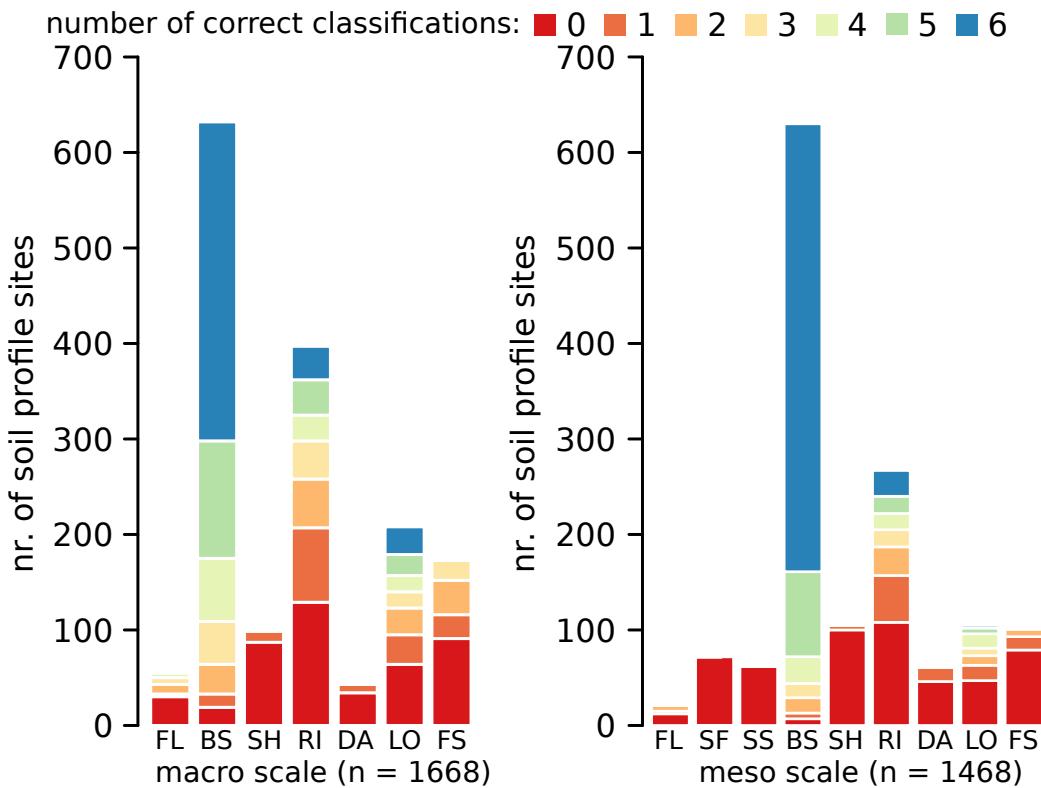


Figure 7: At each scale and generalisation level, a stacked barplot shows the distribution of the number of times a specific data point was classified correctly using the six different methods. FL = flat, LO = hydrologically low area, DA = debris accumulation, FS = footslope, SF = slope flattening, BS = backslope, SS = slope steepening, SH = shoulder and RI = ridge

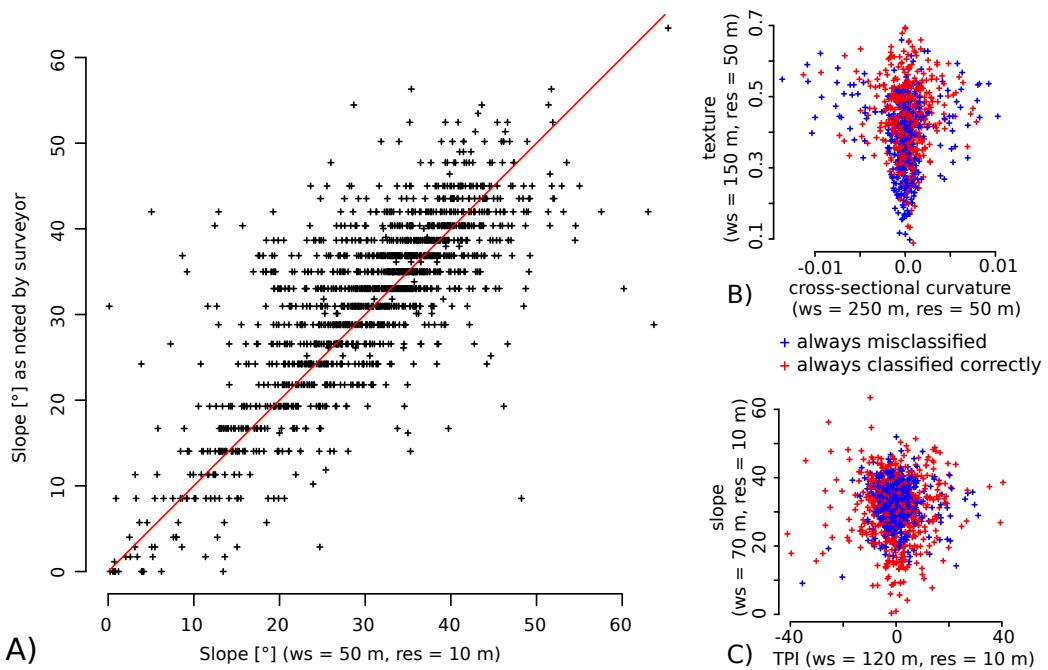


Figure 8: A) comparison of slope values of the soil pit site descriptions and the values derived from the 10 m resolution DTM. The red line indicates were slope values should be if equal. B) and C) show scatter plots of the two terrain parameters that best differentiate the data points that were always classified correctly from those always misclassified, for macro and meso scale, respectively. The parameter texture represents the number of peaks and pits in a search window computed by subtracting the DTM from its own median filtered version. Abbreviations: ws = window size, res = DTM resolution, TPI = Topographic position index

context, especially with the main land use being forestry, the reconstruction of topographic positions described by surveyors using terrain derivatives and classification schemes based on these, is possible only to a certain degree. The various investigated classification algorithms performed similarly at macro scale with 7 classes and at meso scale with 9 classes, yielding overall accuracies of just around 50 %.

#### 4.1. Performance of the different classification approaches

The overall accuracies of the best parameter setting of each classification approach (Table 9) do not differ as much as might be expected, presumably because all are calibrated to the same data set. Due to the overwhelming variability of the backslope class as assigned by the surveyors, all approaches had difficulties with distinguishing smaller classes, especially separating shoulder and footslope from backslope positions. From a pedologic point of view this seems problematic, as these positions are deemed to provide very different environments for soil formation due to the relation between erosion and accumulation. This problem however seems to be partly related to the specific data set and the semantics of defining topographic positions. At both scales, the results of Dikau's curvature classification belonged to the maps with the smallest overall accuracy. The main reason for this is that only curvatures are involved in classification, as implied by the method's name. Consequently, there is no possibility to distinguish between planar slopes and planar valley bottoms, leading to the omission of the class FL in the resulting maps of topographic position (Figure 6). DA and SH were also omitted at both scales. The two pure meso scale classes SF and SS were also not replicated, an omission that is however shared with all other five classification approaches. Wood's morphometric features result in an overall accuracy comparable to that of Dikau's classification. These two classifications also have the most similar output of all classification approaches regarding meso scale and the second most similar at macro scale (Table 10). Schmidt's fuzzy elements, which builds on Dikau's classification, can differentiate between flat and sloping areas and introduces fuzzy membership, leading to slightly better overall performance at both scales. Its cross-validated accuracy of 48 % at both scales is comparable to the results of classification with 'r.geomorphon'. A major difference between the two classifications at macro scale is the handling of the classes DA and FS. While the geomorphon based approach correctly identifies some DA positions and has a better classification rate for RI and LO, the class FS is not attributed to any points. This class, however, is

handled quite well by Schmidt's algorithm. Both approaches fail to correctly  
720 identify any points belonging to the SH position. At meso scale, the geo-  
morphon based classification maps only the classes LO, BS and RI. While  
Schmidt's approach is outperformed at the RI position, it classifies more LO  
positions correctly and introduces the FS class to its output map. The TPI-  
based classification has a slightly smaller overall accuracy value and one class  
725 less in its model of topographic position at macro scale, compared to the two  
previously discussed approaches, with neither DA nor FS positions mapped.  
At meso scale however, it performs best among the automated classifications,  
where it maps one more class than Schmidt's approach, introducing FL to  
the output map.

730 The approach with the best performance, the SVM classifier based on  
a combination of single terrain parameters, also gives the most insight into  
what features appear to have the greatest influence on the surveyors perception  
of his position in the landscape. At macro scale, the regional parameter  
topographic wetness index computed with the 50 m DTM best replicates  
735 topographic position as surveyed during field work. This seems a suitable  
choice as the macro scale position is very much linked to the profile site's position  
within its wider surroundings, for instance the relative position in the  
catchment. The model of macro scale topographic position based on this terrain  
parameter is then refined by adding curvature information, specifically  
740 profile and minimum curvature, as well as slope gradient. Figure 5 shows  
how the modeled topographic positions are attributed to different ranges of  
terrain parameters for both scales. The choice of these parameters also comes  
as no surprise, as most of the automated classifications also incorporate  
curvature values, and the survey manual similarly uses curvature terms such as  
745 convex and concave to describe the differences between the topographic positions.  
At meso scale, the terrain parameter-based model requires two terrain  
parameters and is the only investigated approach to incorporate all topographic  
positions except for slope steepening and flattening. In this model,  
topographic position index, computed at a DTM resolution of 10 m and a  
750 search radius of 70 m, is combined with the slope, computed with the 50 m  
DTM and a window size of 150 m. As the meso scale position is intended to  
describe topographic position with an extent of 50 - 100 m, the combination  
of these terrain parameters is understandable, as the TPI relates a center  
pixel to its surrounding, and the slope further differentiates relatively flat  
755 from sloping areas. The choice of the TPI is also in line with the strong  
performance of the TPI-based classification approach at meso scale.

#### *4.2. Overall assessment of results*

Considering the often mentioned importance of topography in pedogenesis and, consequently, in the soil surveyors mental soil landscape model, it  
760 is surprising that the topographic positions cannot be better replicated using simple terrain parameters or existing landform classification algorithms. What is more, the rule set on which surveyors classify topographic position is based on morphometric characteristics (Englisch and Kilian, 1999). This raises several questions with regard to the reason for the presented results.  
765 The main issues that require discussion are (a) the quality of the sample points, which is the reliability of point data regarding their location, (b) the field classification scheme, influenced by subjectivity and interpretability of the topographic position description performed by the soil surveyor, and (c) the properties of the point sample. A further issue that necessitates exploring  
770 is the adequacy of the examined classification algorithms and their default parameter settings with regard to their application in an Alpine environment, or in other words, the impact of this environment, specifically the high relief, on the perception of topographic position.

*Reliability of individual profile points.* Figure 7 shows that at macro scale a substantial amount of points, consisting of almost a third of the data set,  
775 were misclassified in every single classification approach. On the contrary, approximately the same number of data points were classified correctly in all six classification approaches. This indicates that only a third of the dataset is prone to changes in class due to the different approaches. So in summary,  
780 a third of the data points could not be correctly included into a single classification attempt. Possible reasons are explored in the following sections. Wrong coordinates due to limitations of global positioning system (GPS) accuracy under tree cover could be an explanation, however the use of coarser, resampled DTMs should take into account a large part of the uncertainty  
785 of maximum 50 m as conveyed by discussions with surveyors, especially at macro scale. The consistent effect of grid cell size regarding the difference between topographic position at macro and meso scale reported farther below, indicates that, while individual points may be incorrectly localised, there is no systematic error in the localisation of the profile sites. Otherwise the  
790 DTM resolution or chosen window size should show no correlation with the extents that are used in the survey manual to distinguish meso and macro scale. Consequently, other issues at least contribute to the large number of points that cannot be classified correctly. Figure 8 B) and C) show that

those points always misclassified at macro scale are often characterized by  
795 high texture values, ~~calculated by subtraction of the DTM from a median filtered version of itself~~, indicating rough terrain. At meso scale, the consistently misclassified points surround the correctly classified points which center around a TPI of 0 and a slope of around 30°, which well characterizes the backslope position (see Figure 5). These scatterplots are indicative  
800 of the issues that arise in Alpine terrain and are discussed in the following sections.

*Objectivity of topographic position description and classification.* Congalton (1991) notes that a classification system should be mutually exclusive and exhaustive, however with topographic positions and landforms there are not  
805 clear and crisp boundaries. The differences and transitions between different positions are open to interpretation by the individual surveyor. A remarkable aspect of the surveyed topographic positions of the soil profile sites (Figure 2) is the dominance of the backslope class, which is present at macro scale and even more at meso scale. However, given the Alpine environment in which the  
810 survey was performed, this is not unreasonable. The low overall classification rate can be partly attributed to the generally high rate of misclassification of other positions as backslope due to the wide range of the terrain parameters found at positions classified as such. Brevik and Arnold (2015) note the important influence landscape has on what the surveyor sees when describing  
815 a soil profile. Although this statement presumably refers to the undeniable influence of topography on soil formation, it can also be (mis)interpreted as describing the effect of the soil surveyors personal mental soil-landscape model on which soil, or horizon, characteristics are observed in situ. One hypothesis is that this relationship goes in both directions, meaning that  
820 what the surveyor observes in the soil pit may also reflect on his perception of topographic position due to his expectation of where such soils can be found. For instance, many soils in an Alpine environment are formed from multi-layered deposits (Baruck et al., 2016; Geitner et al., 2011a), hence the backslope position with its dynamic equilibrium of erosion and accumulation  
825 would seem an appropriate topographic position for such soils. A further plausible situation leading to an abundance of backslopes in the profile site data set, is that the backslope position is the 'go-to' class that is attributed to a profile site which does not convincingly fit the descriptions of any of the other topographic positions available in the survey manual. The possibility  
830 of any of these two situations to arise is, in our opinion, further increased

by limited visibility of the surrounding landscape as a consequence of the profile sites' location in partly densely vegetated forests. It must further be noted that the data point description stems from various surveyors, each with his own mental landscape model. Matsuura and Aniya (2012) compared  
835 the results of their breakline detection-based slope segmentation model with maps prepared by manual interpretation of slope contours. They attribute the reported deviations among the manual delineations to differences in the human interpreters definitions of slope segments, especially regarding lower, mid and upper slopes. If these differences are present even in situations  
840 were the human interpreter has a good overview of the larger topographic situation, as is the case when using a slope contour map, it is likely that these differences are similar or even greater in field survey situations, where visibility is possibly additionally impaired as mentioned above. Two classes  
845 used in our study, where we particularly estimate the subjectivity and non-exclusivity of topographic positions to be an issue, are slope flattening and slope steepening, which are only included at meso scale. However, none of the six approaches exemplified in this study managed to replicate these topographic positions, presumably due to the very fuzzy definition of these positions, especially when competing with the backslope position and its wide  
850 terrain parameter ranges.

*The point sample regarding topographic position.* Although the authors were not part of one of the various survey teams and thus not involved in the sampling scheme, this issue must also be addressed. Parts of the data can be seen as part of a stratified sampling scheme, however not with regard to  
855 topographic position but forest types. It cannot be termed random stratified sampling, as the profile points were selected on practical criteria, especially the proximity to forestry service roads. Due to the focus on forest types, certain topographic positions, for instance valley floor and debris fans or cones, were undersampled, as these positions are dominated by agriculture  
860 and not forests. An additional field campaign, based on random stratified sampling and including land uses other than forestry, would help to better capture the topographic variability of these positions and strengthen their position in the model. As discussed above, the backslope position dominates the sample and leads to an unbalanced data set, however Congalton (1991)  
865 advises increased sample sizes for classes with high variability.

#### *4.3. Differences in meso and macro scale*

While the overall performance of the classification of topographic position was rather poor, the applied methodology does a considerably better job at highlighting the differences between topographic positions at macro and meso scale with regard to DTM resolution. Regarding the algorithms where, as is the case with Dikau's curvature classification and Schmidt's fuzzy elements, grid cell size controls the scale of the area that is considered for calculating the relevant terrain parameters, the classifications that best fit the topographic positions at macro scale were all chosen from a coarser DTM than those for meso scale. While the macro scale classification using Dikau's curvature classification is based on the 100 m grid cell size DTM, Schmidt's fuzzy elements preferred an even coarser resolution of 150 m. Considering that the calculation of the local terrain attributes regards a neighborhood of 3 grid cells, the involved area has an extent of 300 m and 450 m for Dikau's and Schmidt's classifications, respectively. According to the survey manual (Englisch and Kilian, 1999), an extent of 100 - 500 m should be regarded for macro scale topographic positions, which is well in line with the 300 and 450 m. The manual describes meso scale topographic positions as having an extent of 50 to 100 m, which agrees with the 50 m grid DTM chosen by Dikau's and Schmidt's algorithms, although the considered 3 grid cells amount to a slightly larger extent.

The other set of classification approaches investigated have different options to incorporate scale into their computation. While Wood's classification can change the window size for which the local terrain parameters are calculated, the TPI-based landform classification uses two search radii for which that regional parameter is calculated, whereas the geomorphon-based approach has one search radius for which the line-of-sight calculations are performed. The three approaches could furthermore choose between the 10 and the 50 m DTM. For macro scale topographic positions, Wood's and the geomorphon approach went with the 50 m DTM while the TPI-based approach used the 10 m grid cell size. The window size, or outer search radius, for Wood's classification, the TPI-based, and the geomorphon-based approaches were 250, 250 and 400 m, respectively. As with the two other grid cell size-driven approaches, these values are well within the range of the surveyor manuals macro scale definition. For the DTM resolution used for meso scale positions, all three search window-driven approaches went with the 10 m DTM and outer search radii or window sizes of 110, 90, and 80 m, respectively. The change in chosen DTM resolution from macro to meso scale

is present in all but one method, and the extent of the meso scale topographic positions, as defined for the surveyors, is well reproduced in the search radii and grid cell sizes.

Considering the overall classification result, there is no significant difference between meso and macro scale, although it must be noted that at meso scale two additional classes, slope steepening and slope flattening, exist but were never replicated in the resulting maps.

Regarding the various maps of topographic position resulting from the six classification approaches, the similarity matrix in Table 10 indicates that the respective output maps at meso scale are more similar to each other than at macro scale. It must however be acknowledged that, in general, the macro scale maps were able to replicate a larger number of topographic position classes, whereas for meso scale models the output maps often contained only three classes, presumably leading to better values in the similarity matrix. This is supported by the fact that the terrain parameter based model (TP), which models all classes except for slope steeping and flattening, shows the smallest similarity to the other approaches.

#### 4.4. The effect of the Alpine environment on the chosen parameter values

Most of the evaluated automated landform classification algorithms include a slope threshold or fuzzy membership function that determines whether a grid cell is flat or sloped. Similarly, a large number include a curvature threshold or fuzzy membership function responsible for the classification as planar. The slope values chosen for the various classification methods in this study are consistently much higher than those found in literature or the default values of the software used. While most of the studies performing landform classification (for instance Barka et al. (2011), Bocco et al. (2001), Ehsani and Quiel (2008), Jasiewicz and Stepinski (2013), MacMillan et al. (2000)) apply slope thresholds of 2 or 3°, or fuzzy membership functions approximately centering on these values, the presented methodology chose values between 8° and 10° for the geomorphon-based classification and between 11° and 15° for Wood's features. The same trend was noted regarding the choice of curvature, where much higher values were allowed for planar grid cells than expected from published studies. Only Schmidt and Hewitt (2004) apply fuzzy membership functions for slope and curvature that show overlap with the values chosen in this study. However, most of the studies applying one or more of the discussed landforms do not concentrate on

940 areas with high relief such as South Tyrol. The authors of this study attribute the choice of high slope and curvature values to the steep terrain in the study area and the under-representation of valley floor and similarly flat topographic positions in the data set. Higher slope thresholds allow for better differentiation of sloping areas, which is where the majority of the profile sites are situated. Some studies that include areas with high relief (Bocco et al., 2001; Herbst et al., 2012) deal with this issue by adjusting their flatness threshold for plateaus. This underlines the necessity of adapting slope and curvature thresholds or membership functions to the topography in the area of interest, especially in an Alpine environment. Considering the discussion  
945 950 on the objectivity of topographic position descriptions above, what is considered to be sloping in hummocky terrain could be considered flat regarding morphodynamics in an Alpine context, especially in the perception of a soil surveyor while performing field work. A comparison of the slope gradient as described during field survey with the slope gradient derived from a DTM at  
955 various window sizes showed the best fit with a window size of 50 m based on the 10 m grid. A scatterplot of the two slope gradients values (Figure 8) shows that, while an overall agreeing trend can be observed, there is often a substantial difference between the slope gradient perceived by the surveyor and the one derived digitally. Whereas points with an exceedingly large difference in slope indicate issues regarding the uncertainty of the profile site location, the scatterplot does suggest larger deviations in steeper terrain, further illustrating the influence the high relief environment has on how the human surveyor perceives and evaluates his position in the landscape. The often, with respect to the DTM, increased slope value mapped by the surveyor may also be an issue of scale, as the slope measured in field is commonly at micro rather than meso scale. Additionally, in steeper terrain it is often easier to dig soil pits in micro-scale slope steepenings, making use of the decreased material cover as well as the adjacent relative flattening from which the excavation is performed.

960 965

#### 970 4.5. Classifier choice and parameter selection

While there exist simpler statistics to compare two maps, for instance those based on the  $\chi^2$  coefficient, the machine learning approach was chosen in order to investigate the most influential terrain parameters for replicating the surveyor's perception of topographic position by constructing a model  
975 using SVM classification. To reduce the number of methods applied in the study and for the sake of comparability, this approach was also applied for

selecting the best fitting parameter settings for each of the applied automated classifications. Next to advancing the understanding of the surveyors perception of landscape with regard to his mental soil landscape model, the  
980 creation of a map of modeled topographic position to aid soil surveyors in the field was an additional goal, especially in light of the results presented in this study. With regard to different machine learning approaches, for instance random forest classification (Breiman, 2001), SVM classification was chosen for its ability to create smooth prediction surfaces, for, as noted by  
985 Steger et al. (2016), these improve the readability and hence utility of maps for practical purposes, for instance soil survey.

As most landform classification algorithms in themselves already involve the classification and generalisation of a number of terrain parameters, classification power was rarely increased by the combination of result maps of  
990 the same classification scheme applying different parameter settings or scale, for instance combining two geomorphon-based landform maps computed with different search radii. The assumption is, that this is the case because the aim of the procedure is to recreate topographic position at a very specific scale, namely meso or macro scale as defined in the survey manual. When building  
995 a model through forward selection of single terrain parameters however, the classification results did improve. Generally, this improvement abated after a feature set of three to four terrain parameters had been selected. While the addition of several parameters may have slightly increased the classification accuracy, it must be kept in mind that the aim of this procedure is  
1000 to find some basic knowledge regarding the perception of topography by the surveyor, and not to overfit our model to the dataset by incorporating a vast number of features.

## 5. Conclusion

In this study, we investigated five landform classification approaches implemented in Open Source GIS with regard to their ability to emulate topographic position as perceived by soil surveyors in the field. While the overall accuracies of the investigated algorithms were relatively similar, the results of Schmidt's fuzzy elements classification and r.geomorphon seem best suited at macro scale, whereas the TPI-based landform classification performed best  
1005 at meso scale. All classification approaches showed that macro scale topographic positions require DTMs with a coarser resolution, or larger search radii, depending on the method, than at meso scale. The extents analysed in  
1010

the various classifications coincide well with the extents attributed to meso and macro scale in the survey manual. To better understand the surveyors' soil-landscape model, a support vector machine classifier was trained by applying a step-wise forward feature selection procedure to a wide range of terrain derivatives, both local and regional. At meso scale the topographic position index with a search radius of 70 m, combined with slope, best replicated the surveyors' topographic position. The topographic wetness index computed with a 50 m DTM together with various curvatures proved to be the most useful<sup>1</sup> terrain derivatives to emulate the surveyors' view of macro scale topographic position. The presented workflow highlights how landform classifications change with varying algorithm parameters and which thresholds and resolutions best correspond to human perception. Together with the described selection procedure of terrain parameters and their specific DTM resolutions, such observations aid in giving insight into the assessment of topographic position by soil surveyors and at which scales these are performed.

The dominance of the backslope position in profile site data from an Alpine Environment was identified as one of the major issues of replicating topographic position, due to the high variability of terrain at sites to which this position is attributed by the soil surveyors. This, together with overall decreased visibility of the surrounding area in forested slopes, and the fuzzy semantics of topographic position especially in steep, rugged terrain, may explain a large part of the difficulty of bringing together the topographic position as perceived by the surveyor during field work, and derivatives of digital elevation models. Nevertheless, shoulder and footslope are topographic positions with high relevance with regard to soil formation and consequently more effort is necessary to better differentiate these positions in digital terrain analysis. The same argumentation is valid for slope flattening and steepening at meso scale. Further research into whether these topographic positions, which represent transitional areas, function at a different scale than the other positions, is necessary. A possible approach could be to discriminate these positions only from areas already classified as backslope in a previous, coarser landform delineation attempt.

Despite these results, we are not of the opinion that the topographic positions defined in soil surveys are not reliable, but rather that the segmentation of digital elevation data into landforms or topographic positions is not always as straightforward as often perceived. Furthermore, the way a soil surveyor segments the landscape into positions relevant to pedogenesis may very well differ from how the same landscape is perceived by researchers and

mappers from other fields of science, such as geomorphology or, more abstract, geographic information science. An important takeaway point from this study is that the specific topographic setting, in this case the Alpine environment, strongly impacts how topographic position is perceived and understood, and this factor should always be kept in mind when working with digital data to support human surveyors in the field. This especially applies to the slope and curvature thresholds, which differ greatly from those applied in other studies or implemented as default values in the GIS used. Even if only to attain an estimate of the deviation between automated and field survey classification, we propose that DTM-based classifications should be validated and consolidated with real field data, not just the results of cartographic or photogrammetric interpretation. It is our opinion that this study is a valuable step in promoting an understanding of the relationship between the human surveyors' perception of topographic position related to soil and its development, and DTM derivatives. It shows that in order to bridge the gap between field soil survey and digital terrain analysis, further research is necessary on both sides. On the side of field survey, additional differentiation of the backslope class seems a constructive goal, with the aim of increased specificity of this class, making it less arbitrary. The issue of scale should also be considered when determining the slope angle for profile site descriptions. Digital terrain analysis, on the other side, is tasked with better detecting the classes footslope and shoulder, as well as slope steepening and flattening, due to their significance to soil formation. This may include research into the proper scale at which these classes can be distinguished better, especially for the latter classes, and also by relating them to adjacent topographic positions in an object- or search radius-based manner.

To ensure a successful and effective collaboration between digital terrain analysis and field survey, we encourage other researchers to perform similar analyses, especially within different topographic settings. This shall enhance cooperation and bring to surface similar issues relating to the perception of landscape by the soil surveyor.

## Acknowledgements

This research was performed within the project 'Terrain Classification of ALS Data to support Digital Soil Mapping', funded by the Autonomous Province Bolzano – South Tyrol (15/40.3).

## References

- Ad-hoc-Arbeitsgruppe Boden, 2006. Bodenkundliche Kartieranleitung. KA5. Schweizerbart Science Publishers, Stuttgart, Germany. URL: [http://www.schweizerbart.de//publications/detail/isbn/9783510959204/Bodenkundliche\\_Kartieranleitung\\_5\\_Aufl](http://www.schweizerbart.de//publications/detail/isbn/9783510959204/Bodenkundliche_Kartieranleitung_5_Aufl).  
1090
- Adediran, A.O., Parcharidis, I., Poscolieri, M., Pavlopoulos, K., 2004. Computer-assisted discrimination of morphological units on north-central Crete (Greece) by applying multivariate statistics to local relief gradients. *Geomorphology* 58, 357–370. URL: <http://dx.doi.org/10.1016/j.geomorph.2003.07.024>.  
1095
- APB, 2006. Waldtypisierung Südtirol. Autonome Provinz Bozen - Südtirol. URL: <http://www.provinz.bz.it/forst/studien-projekte/waldtypisierung.asp>.
- APB, 2016. Download Landeskartographie - Autonome Provinz Bozen - Südtirol. URL: <http://www.provinz.bz.it/natur-raum/themen/landeskartografie-download.asp>.  
1100
- Arrell, K., Fisher, P., Tate, N., Bastin, L., 2007. A fuzzy c-means classification of elevation derivatives to extract the morphometric classification of landforms in Snowdonia, Wales. *Computers & Geosciences* 33, 13661381. URL: <http://dx.doi.org/10.1016/j.cageo.2007.05.005>, doi:10.1016/j.cageo.2007.05.005.  
1105
- Ballabio, C., 2009. Spatial prediction of soil properties in temperate mountain regions using support vector regression. *Geoderma* 151, 338 – 350. URL: <http://www.sciencedirect.com/science/article/pii/S0016706109001499>, doi:<http://dx.doi.org/10.1016/j.geoderma.2009.04.022>.  
1110
- Barka, I., Vladovic, J., Malis, F., 2011. Landform classification and its application in predictive mapping of soil and forest units. *GIS Ostrava* 1, 23–26.
- Barringer, J.R.F., Hewitt, A.E., Lynn, I.H., Schmidt, J., 2008. National Mapping of Landform Elements in Support of S-Map, A New Zealand Soils Database, in: Zhou, Q., Lees, B., Tang, G.a. (Eds.), *Advances in Digital*  
1115

- Terrain Analysis. Springer Berlin Heidelberg. Lecture Notes in Geoinformation and Cartography, pp. 443–458. URL: [http://dx.doi.org/10.1007/978-3-540-77800-4\\_24](http://dx.doi.org/10.1007/978-3-540-77800-4_24), doi:10.1007/978-3-540-77800-4\_24.
- 1120 Baruck, J., Nestroy, O., Sartori, G., Baize, D., Traidl, R., Vraj, B., Brm, E., Gruber, F.E., Heinrich, K., Geitner, C., 2016. Soil classification and mapping in the Alps: The current state and future challenges . Geoderma 264, Part B, 312 – 331. URL: <http://www.sciencedirect.com/science/article/pii/S0016706115300343>, doi:<http://doi.org/10.1016/j.geoderma.2015.08.005>. soil mapping, classification, and modelling: history and future directions.
- 1125 Behrens, T., Scholten, T., 2006. Chapter 25 A Comparison of Data-Mining Techniques in Predictive Soil Mapping, in: P. Lagacherie, A.M., Voltz, M. (Eds.), Digital Soil Mapping An Introductory Perspective. Elsevier. volume 31 of *Developments in Soil Science*, pp. 353 – 617. URL: <http://www.sciencedirect.com/science/article/pii/S0166248106310252>, doi:[http://dx.doi.org/10.1016/S0166-2481\(06\)31025-2](http://dx.doi.org/10.1016/S0166-2481(06)31025-2).
- 1130 1135 Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., Tiede, D., 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS Journal of Photogrammetry and Remote Sensing 87, 180–191. URL: <http://dx.doi.org/10.1016/j.isprsjprs.2013.09.014>, doi:10.1016/j.isprsjprs.2013.09.014.
- 1140 Blaszczyński, J.S., 1997. Landform Characterization with Geographic Information Systems. Photogrammetric Engineering and Remote Sensing 63, 183–191.
- Bocco, G., Mendoza, M., Velzquez, A., 2001. Remote sensing and GIS-based regional geomorphological mapping – a tool for land use planning in developing countries. Geomorphology 39, 211 – 219. URL: <http://www.sciencedirect.com/science/article/pii/S0169555X01000277>, doi:[http://dx.doi.org/10.1016/S0169-555X\(01\)00027-7](http://dx.doi.org/10.1016/S0169-555X(01)00027-7).
- 1150 Bolongaro-Crevenna, A., Torres-Rodrguez, V., Sorani, V., Frame, D., Arturo Ortiz, M., 2005. Geomorphometric analysis for characteriz-

- ing landforms in Morelos State, Mexico. *Geomorphology* 67, 407–422. URL: <http://www.sciencedirect.com/science/article/pii/S0169555X04002661>.
- 1155 Breiman, L., 2001. Random forests. *Machine Learning* 45, 5–32. URL: <http://dx.doi.org/10.1023/A%3A1010933404324>, doi:10.1023/A:1010933404324.
- Brevik, E.C., Arnold, R.W., 2015. Is the Traditional Pedologic Definition of Soil Meaningful in the Modern Context? *Soil Horizons* 56. URL: <http://dx.doi.org/10.2136/sh15-01-0002>, doi:10.2136/sh15-01-0002.
- 1160 Burrough, P.A., van Gaans, P.F., MacMillan, R., 2000. High-resolution landform classification using fuzzy k-means. *Fuzzy Sets and Systems* 113, 37 – 52. URL: <http://www.sciencedirect.com/science/article/pii/S0165011499000111>, doi:10.1016/S0165-0114(99)00011-1.
- 1165 Congalton, R.G., 1991. A review of assessing the accuracy of classification of remotely sensed data. *Remote Sensing of Environment* 37, 35 – 46. doi:10.1016/0034-4257(91)90048-B.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1. 4. Geoscientific Model Development 8, 1991–2007. URL: <http://www.geosci-model-dev.net/8/1991/2015/gmd-8-1991-2015.html>, doi:10.5194/gmd-8-1991-2015.
- 1170 Cortes, C., Vapnik, V., 1995. Support-vector networks. *Machine Learning* 20, 273–297. doi:10.1007/BF00994018.
- Dikau, R., 1988. Entwurf einer geomorphographisch - analytischen Systematik von Reliefeinheiten. Technical Report. Selbstverlag des Geographischen Instituts der Universität Heidelberg. Arbeitsbericht im Rahmen des DFG-Schwerpunkteprogramms "Digitale Geowissenschaftliche Kartenwerke".
- 1175 Drăgut, L., Blaschke, T., 2006. Automated classification of landform elements using object-based image analysis. *Geomorphology* 81, 330344. URL: <http://dx.doi.org/10.1016/j.geomorph.2006.04.013>, doi:10.1016/j.geomorph.2006.04.013.

- Ehsani, A.H., Quiel, F., 2008. Geomorphometric feature analysis using morphometric parameterization and artificial neural networks. *Geomorphology* 99, 1– 12.
- Ehsani, A.H., Quiel, F., 2009. A semi-automatic method for analysis of landscape elements using Shuttle Radar Topography Mission and Landsat ETM+ data. *Computers & Geosciences* 35, 373 – 389. URL: <http://www.sciencedirect.com/science/article/pii/S0098300408000320>, doi:<http://dx.doi.org/10.1016/j.cageo.2007.09.019>.
- Englisch, M., Kilian, W., 1999. Anleitung zur Forstlichen Standortskartierung. 2 ed. FBVA.
- FAO, 2006. Guidelines for soil description. Food and Agricultural Organisation of the United Nations. Rome.
- Gallant, J.C., Wilson, J.P., 2000. Terrain Analysis - Principles and Applications. John Wiley & Sons, Inc.. chapter Primary Topographic Attributes. pp. 51–85.
- Geitner, C., Baruck, J., Freppaz, M., Godone, D., Grashey-Jansen, S., Gruber, F.E., Heinrich, K., Papritz, A., Simon, A., Stanchi, S., Traidl, R., von Albertini, N., Vraj, B., 2017. Chapter 8 - soil and land use in the alpschallenges and examples of soil-survey and soil-data use to support sustainable development, in: Pereira, P., Brevik, E.C., Muoz-Rojas, M., Miller, B.A. (Eds.), Soil Mapping and Process Modeling for Sustainable Land Use Management. Elsevier, pp. 221 – 292. URL: <http://www.sciencedirect.com/science/article/pii/B9780128052006000086>, doi:<http://doi.org/10.1016/B978-0-12-805200-6.00008-6>.
- Geitner, C., Bussemer, S., Ehrmann, O., Ikinger, A., Schfer, D., Traidl, R., Tscherko, D., 2011a. Mensch und Umwelt im Holozän Tirols. Innsbruck. volume 1. chapter Bodenkundlich-stratigraphische Befunde am Ullafelsen im hinteren Fotschertal sowie ihre landschaftsgeschichtliche Interpretation. pp. 109–151.
- Geitner, C., Tusch, M., Meißl, G., Kringer, K., Wiegand, C., 2011b. Effects of topography on the spatial distribution of soils: basic considerations on interdependencies and data sources with examples from the Eastern

- Alps. Zeitschrift für Geomorphologie 55, 127–146. URL: <http://dx.doi.org/10.1127/0372-8854/2011/0055S3-0055>, doi:10.1127/0372-8854/2011/0055S3-0055.
- 1220 Gerçek, D., Toprak, V., Strobl, J., 2011. Object-based classification of landforms based on their local geometry and geomorphometric context. International Journal of Geographical Information Science 25, 1011 – 1023. URL: <http://dx.doi.org/10.1080/13658816.2011.558845>, doi:10.1080/13658816.2011.558845.
- 1225 Gerçek, D., Zeydanli, U., 2010. Object-based classification of landscape into land management units (LMUs), in: GEOBIA 2010: Geographic Object-Based Image Analysis, 29 June - 2 July, 2010, Ghent, Belgium.
- 1230 GRASS Development Team, 2017. Geographic Resources Analysis Support System (GRASS GIS) Software, Version 7.2. Open Source Geospatial Foundation. URL: <http://grass.osgeo.org>.
- Guisan, A., Weiss, S., Weiss, A., 1999. GLM versus CCA spatial modeling of plant species distribution. Plant Ecology 143, 107–122. URL: <http://dx.doi.org/10.1023/A%3A1009841519580>, doi:10.1023/A:1009841519580.
- 1235 Herbst, P., Gross, J., Meer, U., Mosimann, T., 2012. Geomorphographic terrain classification for predicting forest soil properties in Northwestern Switzerland. Zeitschrift für Geomorphologie 56, 1–22. doi:{10.1127/0372-8854/2012/0069}.
- 1240 Hoersch, B., Braun, G., Schmidt, U., 2002. Relation between landform and vegetation in alpine regions of Wallis, Switzerland. A multiscale remote sensing and GIS approach. Computers, Environment and Urban Systems 26, 113–139. URL: <http://www.sciencedirect.com/science/article/pii/S0198971501000394>.
- 1245 Hollingsworth, I.D., Bui, E.N., Odeh, I.O., McLeod, P., 2006. Chapter 29 Rule-Based Land Unit Mapping of the Tiwi Islands, Northern Territory, Australia, in: P. Lagacherie, A.M., Voltz, M. (Eds.), Digital Soil Mapping An Introductory Perspective. Elsevier. volume 31 of *Developments in Soil Science*, pp. 401 – 621. URL: [http://dx.doi.org/10.1016/S0166-2481\(06\)31029-X](http://dx.doi.org/10.1016/S0166-2481(06)31029-X), doi:10.1016/S0166-2481(06)31029-X.

- Hughes, M.W., Schmidt, J., Almond, P.C., 2009. Automatic landform  
1250 stratification and environmental correlation for modelling loess land-  
scapes in North Otago, South Island, New Zealand. *Geoderma* 149,  
92–100. URL: <http://dx.doi.org/10.1016/j.geoderma.2008.11.024>,  
doi:10.1016/j.geoderma.2008.11.024.
- Irvin, B.J., Ventura, S.J., Slater, B.K., 1997. Fuzzy and isodata clas-  
1255 sification of landform elements from digital terrain data in Pleasant  
Valley, Wisconsin. *Geoderma* 77, 137–154. URL: <http://www.sciencedirect.com/science/article/pii/S001670619700019>, doi:10.1016/S0016-7061(97)00019-0.
- Iwahashi, J., Pike, R.J., 2007. Automated classifications of topography  
1260 from DEMs by an unsupervised nested-means algorithm and a three-  
part geometric signature. *Geomorphology* 86, 409 – 440. doi:10.1016/j.geomorph.2006.09.012.
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to  
1265 Statistical Learning. Springer-Verlag New York. URL: <http://www-bcf.usc.edu/~gareth/ISL/index.html>, doi:10.1007/978-1-4614-7138-7.
- Jasiewicz, J., Netzel, P., Stepinski, T.F., 2014. Landscape similarity,  
retrieval, and machine mapping of physiographic units. *Geomorphology* 221,  
104 – 112. URL: <http://www.sciencedirect.com/science/article/pii/S0169555X14003110>, doi:<http://dx.doi.org/10.1016/j.geomorph.2014.06.011>.
- Jasiewicz, J., Stepinski, T.F., 2013. Geomorphons – a pattern recognition  
1270 approach to classification and mapping of landforms. *Geomorphology* 182,  
147 – 156. URL: <http://www.sciencedirect.com/science/article/pii/S0169555X12005028>, doi:10.1016/j.geomorph.2012.11.005.
- Klingseisen, B., Metternicht, G., Paulus, G., 2008. Geomorphometric  
1275 landscape analysis using a semi-automated GIS-approach. *Environmental Modelling & Software* 23, 109 – 121. URL: <http://www.sciencedirect.com/science/article/pii/S1364815207000953>,  
doi:<http://dx.doi.org/10.1016/j.envsoft.2007.05.007>.
- Kringer, K., Tusch, M., Geitner, C., Rutzinger, M., Wiegand, C., Meissl,  
1280 G., 2009. Geomorphometric Analyses of LiDAR Digital Terrain Models as

Input for Digital Soil Mapping, in: Proceedings of Geomorphometry 2009. Zurich, Switzerland, 31 August - 2 September, 2009, pp. 74 – 81.

- 1285 MacMillan, R., Pettapiece, W., Nolan, S., Goddard, T., 2000. A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. *Fuzzy Sets and Systems* 113, 81–109. URL: <http://www.sciencedirect.com/science/article/pii/S0165011499000147>, doi:10.1016/S0165-0114(99)00014-7.
- 1290 Mashimbye, Z.E., de Clercq, W.P., Niekerk, A.V., 2014. An evaluation of digital elevation models (DEMs) for delineating land components. *Geoderma* 213, 312 – 319. URL: <http://dx.doi.org/10.1016/j.geoderma.2013.08.023>.
- 1295 Matsuura, T., Aniya, M., 2012. Automated segmentation of hillslope profiles across ridges and valleys using a digital elevation model. *Geomorphology* 177178, 167–177. URL: <http://www.sciencedirect.com/science/article/pii/S0169555X1200373X>.
- 1300 Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2014. e1071: Misc Functions of the Department of Statistics (e1071), TU Wien. URL: <http://CRAN.R-project.org/package=e1071>. r package version 1.6-4.
- Minàr, J., Evans, I.S., 2008. Elementary forms for land surface segmentation: The theoretical basis of terrain analysis and geomorphological mapping. *Geomorphology* 95, 236259. URL: <http://dx.doi.org/10.1016/j.geomorph.2007.06.003>.
- 1305 Mokarram, M., Seif, A., Sathyamoorthy, D., 2015. Landform classification via fuzzy classification of morphometric parameters computed from digital elevation models: case study on Zagros Mountains. *Arabian Journal of Geosciences* 8, 4921–4937. URL: <http://dx.doi.org/10.1007/s12517-014-1556-y>, doi:10.1007/s12517-014-1556-y.
- 1310 Moravej, K., Eghbal, M.K., Toomanian, N., Mahmoodi, S., 2012. Comparison of automated and manual landform delineation in semi detailed soil survey procedure. *African Journal of Agricultural Research* 7, 2592–2600. doi:10.5897/AJAR11.728.

- Nestroy, O., Aust, G., Blum, W., Englisch, M., Hager, H., Herzberger, E.,  
1315 Kilian, W., Nelhiebel, P., G. Ortner and, E.P., und J. Wagner, A.P.W.S.,  
2011. Systematische Gliederung der Böden Österreichs. Österreichische  
Bodensystematik 2000 in der revidierten Fassung von 2011. Mitt. Österr.  
Bodenndl. Ges. 79.
- Pennock, D., ZebARTH, B., Jong, E.D., 1987. Landform classification and soil  
1320 distribution in hummocky terrain, Saskatchewan, Canada. Geoderma 40,  
297 – 315. URL: <http://www.sciencedirect.com/science/article/pii/0016706187900401>, doi:10.1016/0016-7061(87)90040-1.
- R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL:  
1325 <http://www.R-project.org/>.
- Reu, J.D., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V.,  
Smedt, P.D., Chu, W., Antrop, M., Maeyer, P.D., Finke, P., Meirvenne,  
1330 M.V., Verniers, J., Cromb, P., 2013. Application of the topographic position index to heterogeneous landscapes. Geomorphology 186, 39 – 49. URL: <http://www.sciencedirect.com/science/article/pii/S0169555X12005739>, doi:<http://dx.doi.org/10.1016/j.geomorph.2012.12.015>.
- Rodriguez, F., Maire, E., Courjault-Radé, P., Darrozes, J., 2002. The Black  
Top Hat function applied to a DEM: A tool to estimate recent incision in  
1335 a mountainous watershed (Estibre Watershed, Central Pyrenees). Geophysical Research Letters 29, 9–1–9–4. URL: [10.1029/2001GL014412](https://doi.org/10.1029/2001GL014412), doi:[10.1029/2001GL014412](https://doi.org/10.1029/2001GL014412).
- Rossel, R.V., Behrens, T., 2010. Using data mining to model and interpret  
soil diffuse reflectance spectra. Geoderma 158, 46 – 54. URL: <http://www.sciencedirect.com/science/article/pii/S0016706109004315>,  
1340 doi:<http://dx.doi.org/10.1016/j.geoderma.2009.12.025>.
- Schaetzl, R., 2013. Catenas and Soils, in: Shroder, J.F. (Ed.), Treatise  
on Geomorphology. Academic Press, San Diego, pp. 145–158. URL:  
<http://dx.doi.org/10.1016/B978-0-12-374739-6.00074-9>, doi:[10.1016/B978-0-12-374739-6.00074-9](http://dx.doi.org/10.1016/B978-0-12-374739-6.00074-9).

- Schmidt, J., Hewitt, A., 2004. Fuzzy land element classification from DTMs based on geometry and terrain position. *Geoderma* 121, 243 – 256. URL: <http://dx.doi.org/10.1016/j.geoderma.2003.10.008>, doi:10.1016/j.geoderma.2003.10.008.
- 1350 Schmidt, J., Tonkin, P., Hewitt, A.E., 2005. Quantitative soillandscape models for the Haldon and Hurunui soil sets, New Zealand. *Australian Journal of Soil Research* 43. doi:10.1071/SR04074.
- 1355 Steger, S., Brenning, A., Bell, R., Petschko, H., Glade, T., 2016. Exploring discrepancies between quantitative validation results and the geomorphic plausibility of statistical landslide susceptibility maps. *Geomorphology* 262, 8 – 23. URL: <http://www.sciencedirect.com/science/article/pii/S0169555X16301052>, doi:<http://dx.doi.org/10.1016/j.geomorph.2016.03.015>.
- 1360 Wack, R., Stelzl, H., 2005. Laser DTM Generation for South-Tyrol and 3D-Visualization, in: ISPRS WG III/3, III/4, V/3 Workshop "Laser scanning 2005".
- Weiss, A., 2000. Topographic Position and Landform Analysis. Poster Presentation, ESRI Users Conference, San Diego, CA (2001).
- 1365 Wood, J.D., 1996. The geomorphological characterisation of digital elevation models. Ph.D. thesis. University of Leicester, UK.