From geological to soil parent material maps - a random forest-supported analysis of geological map units and topography to support soil survey in South Tyrol

Fabian E. Gruber[a,*], Jasmin Baruck[a], Volkmar Mair[b], Clemens Geitner[a]

[a] *Institute of Geography, University of Innsbruck, Innrain 52f, 6020 Innsbruck, Austria*
[b] *Office for Geology and building materials testing, Eggentaler Straße 48, 39053 Kardaun, Autonomous Province Bolzano – South Tyrol, Italy*

Abstract

Parent material is an important factor of soil formation and consequently plays a dominant role in both traditional field soil survey and digital soil mapping. The emergence of a new generation of geological maps at high spatial and thematic resolution in South Tyrol raises the question of how to effectively incorporate these into soil mapping efforts. By comparing the units of these geological maps with the parent material description of soil pits, we evaluate to what extent these can be used as soil parent material maps. Random forest classification and feature selection are applied to highlight those terrain parameters that i) best distinguish between the different surficial geology units, ii) separate soil profile sites with different soil parent material, and iii) can be used together with the geologic map to train a classifier to model the distribution of soil parent material in the study area. The main issue detected by analysing the differences between the geologic map units and the soil parent material information is the dominant role of till, which acts as soil parent material for a large number of soils located on different geological map units. While slope debris is another class with a low producer's reliability, the issues concerning its misclassification are connected to fuzzy categorical transitions between soil parent material classes. Terrain parameters characterizing surface roughness, specifically a combination of vector ruggedness measure (VRM) and topographic rough-

*Corresponding author
Email address: Fabian.Gruber@ages.at (Fabian E. Gruber)

ness index (TRI), were identified as being best suited to join the geological map units in modeling soil parent material and indicate areas where till as soil parent material should be expected. By evaluating these results together with the distribution of soil types, a geologic-topographic characterisation is performed for each geological map unit, with the aim of highlighting specific combinations of geological units and topographical situations which should be in the center of future detailed soil surveys, consequently facilitating the soil mapping procedure.

Keywords: geologic map, soil survey, soil parent material, digital soil mapping, surface roughness, random forest classification, Alpine environment

1. Introduction

Geologic maps have always been an important aid in soil survey as parent material is a decisive factor in soil formation (Jenny, 1941). The importance of this relationship is highlighted by the fact that, vice versa, soil maps have themselves been applied to support and improve geologic mapping (Brevik and Miller, 2015). Providing both the physical structure and the chemical composition of the mineral constituents, parent material plays a fundamental role regarding the direction as well as speed of soil evolution. This is particularly the case in young soils (Schaetzl et al., 2000) such as those predominantly found in the Alps (Geitner et al., 2017). Thus, in order to understand the spatial pattern of soils in the Alps, it is essential to identify the types and origins of parent materials, which are, at least in the lower and medium elevations of the Alpine environment, dominated by quaternary unconsolidated sediments. These deposits vary considerably in thickness; they are often multi-layered and exposed to recent morphodynamics, all of which control soil horizon development and properties (Phillips and Lorz, 2008). In this context, it is indicated to also include characteristics of the subsolum as often as possible, mainly in order to make soil information more suitable for a wide range of environmental issues, as discussed in detail by Juilleret et al. (2016). Consequently, geological maps at various scales have been used as an environmental variable in digital soil mapping (DSM), representing the soil forming factor parent material, or simply 'p'. In their study which presents the 'scorpan' framework of inferring soil information, McBratney et al. (2003) present a table which gives an overview of studies applying DSM and indicates in which of these studies the parent material was involved as an independent variable, thereby highlighting its value and role in digital soil mapping. Fracek and Mosimann (2013), who modelled the carbonate-free depth of soils in Switzerland, attribute a large part of the modelling error to discrepancies between the map units of geologic maps and the actual soil parent material, highlighting the significance of this specific environmental variable in soil modelling. How this important variable is classified, however, will vary greatly depending on the available data, the soil classification system used, the specific mapping guidelines applied, and most importantly the particular geologic and geomorphological setting of the investigated area.

In its guidelines for soil description, the Food and Agriculture Organization of the United Nations promotes a hierarchical system for describing lithologies that constitute the soil parent material, based on the major classes igneous rock, metamorphic rock, consolidated and unconsolidated sedimentary rock (FAO, 2006).

While the classification system used by the surveyors employed by the Forestry Services of the Autonomous Province of Bolzano - South Tyrol as well as North Tyrol (Englisch and Kilian, 1999; Amt der Tiroler Landesregierung, Abteilung Forstplanung, unpublished) is similar to that of the FAO when bedrock is regarded as being the parent material, the former is more closely adapted to the Alpine environment. Specifically, the major class of unconsolidated sedimentary rocks has a far greater number of types in order to satisfy the demands posed by the diversity of the glacial deposits, but also the more recent deposits which are driven mainly by the high relief and steep slopes found in Alpine regions. While such an adaptation of the classes and types of soil parent material to the given circumstances is certainly necessary, communication between soil scientists regarding soil parent materials and comparability is hindered by the multitude of classifications. Juilleret et al. (2016), who stress the importance of describing the subsolum in soil survey, propose a morphogenetic procedure for characterising and classifying subsolum material applying a structure similar to that of the WRB.

A number of studies have compared the information from soil surveys with geologic maps. Juilleret et al. (2012) for instance compared C-horizon data from soil profiles with parent material as derived from a geologic map, concluding that surficial geologic maps can be improved with available soil profile data. They also highlight the necessity of improved communication and exchange between the two sciences. Miller and Lee Burras (2015) compared surficial geology maps produced by a Geological Survey with comparable maps produced using Soil Survey maps of higher spatial resolution, reporting an agreement of 81%. In their review covering the subject of improving geologic maps with soil maps, Brevik and Miller (2015) argue that information from both mapping approaches should be integrated, but always under consideration of the limitations and differences between the disciplines. Whereas most of the previously mentioned studies analyse the possibility of using soil survey information, which generally comes at a higher spatial resolution, for mapping surficial geology, the aim of the presented study is to analyse the application of detailed geologic maps as parent material maps, highlighting those geologic units of the study area where the derivation of soil parent material from geology is not as straightforward as assumed.

A second important soil forming factor is topography or relief. It is considered in traditional soil survey, for instance by mapping landscape position (in the sense of soil catenas) and local slope and curvature (FAO, 2006), and also DSM, where the variables which represent topography in a given model can be chosen from a wide set of available parameters. Examples of such terrain parameters can be found, amongst others, in Böhner and Antonić (2009), Gallant and Wilson (2000) and Olaya (2009). Regarding the geomorphometric characterisation of geologic or soil parent material units, a number of considerations should be taken into account when choosing which parameter groups to investigate. While regional parameters well describe the hydrologically relevant, relative position in the landscape, they, as well as absolute and relative height-related parameters, are strongly correlated to the underlying geological structure of a given region. Local parameters such as slope and curvature on the other hand are often used to infer soil properties and give insight into local morphodynamics, but may also vary strongly within a map unit. To characterise parent material units, especially with regard to topographic - and as a result, soil - variability, an intermediate terrain parameter set describing a unit's land surface is of particular interest. For instance, researchers have long investigated ways to quantify the roughness or ruggedness of terrain, reaching from the analysis of field data and topographic maps to computing roughness indices on raster grids. Geology, geomorphology as well as habitat modelling and wildlife management have been the main scientific research areas in which such investigations were performed on land surfaces. Hobson (1972) presented three different roughness values and applied them to field measurements, correlating them to rock type. Beasom et al. (1983) and Nellemann and Thomsen (1994) applied contour line analysis as a way of characterizing terrain ruggedness. Grid-cell based roughness measures were proposed as the topographic roughness index (TRI) by Riley et al. (1999) or the vector ruggedness measure (VRM) by Sappington et al. (2007), who expanded on the work of Hobson (1972). Similar approaches based on eigenvalue rations of an orientation matrix were proposed for application in geology for instance by Woodcock (1977) and Coblentz et al. (2014).

The objective of the presented study is to evaluate how to make best use of available high (both spatial and thematic) resolution geologic and topographic information for soil survey. This is of special interest as, while the quaternary deposits were often neglected in older geologic surveys or aggregated in a single unit, the new generation of geologic maps created within the CARG (Geologic and geothematic cartography of Italy) framework exhibit high detail regarding these surficial deposits. The study area in the Überetsch/Oltradige region of the Autonomous Province Bolzano – South Tyrol was therefore chosen due to the availability of detailed geologic maps, but also the diversity of soil forming factors, especially geology and topography. By applying random forest classification and feature selection, we investigate which terrain parameters, with emphasis on roughness measures, are best suited to produce a parent material map based on a combination of an available geological map and topographical information. Additionally, the same method is applied to distinguish terrain parameters that, for each soil parent material class, best separate those profile site points that are correctly classified in the geological map from those of the same class that are misclassified. Based on this analysis and a similar investigation into characteristic terrain parameters of the surficial geology map units, each of these units is characterised with regard to topography and soil. We emphasize those units which are often confused or show overlap, and should therefore be surveyed with greater effort and detail, always under consideration of relevant topographic information. The main aim of the random forest classification is to identify the topographic characteristics of the parent material and geologic units in order to facilitate future detailed soil surveys.

2. Study area and data

2.1. Overview

The study area includes the wide vale of Eppan-Kaltern, the Überetsch, which is a paleovalley of the Etsch River located just south-west of Bozen in the Autonomous Province of Bolzano - South Tyrol, and extends in the north to the debris fan of Andrian in the Etsch Valley and the adjacent hill slope on the orographic right of the Etsch River. The western border of the study area is the steep slope of the Mendola-Roèn-Ridge (max. 2116 m a.s.l.), whereas the eastern border of the Überetsch and the study area is represented by the the Mitterberg, a ridge of Permic Vulcanites from which steep slopes descend to the Etsch Valley (approx. 200 m a.s.l.). The Kalterer Lake represents the southern limit of the investigated area. The land use of the paleovalley and its debris cones as well as the Etsch valley is dominantly apple orchards and vineyards, whereas the slope of the Mendola-Roèn-Ridge and the hilly outcrops of Vulcanites are covered by forests. Some pastures are located mainly on the flat areas of the Mendola-Roèn-Ridge. Figure 1 shows the study area as well as the surficial geologic units and soil profile locations. The study area is characterised by the mild and humid transition zone between the mediterranean and continental climate zone. Montiggl, located in the southeastern quadrant of the study area, is characterised by a longtime mean annual precipitation of 802 mm,and a longtime mean annual temperature of 11.9 °C (Thalheimer, 2006).

2.2. Surficial geologic units

A detailed description of the geologic situation can be found in the commentary to the new geologic map of Eppan (Avanzini et al., 2006). The paleovalley of Überetsch is described by Scholz et al. (2005) as a complex system of gravelly lateral moraines and large kame terraces, the result of the 'Kaltern lobe', a Pleniglacial tongue of the Etsch valley glacier. Additionally, laterally eroded remainders of debris fans that were deposited against the recessing glacier can be found along the lower slopes of the Mendola-Roèn-Ridge, as well as recent debris flow deposits, often composed of mainly limestone and dolomite fragments. The vale bottom itself is filled with Pleistocene sediments and contains a number of valleys carved into the gravels by fossil melt-water. At the eastern and western borders of the Pleistocene sediments, outcrops of Permic igneous rhyolite and Lapilli-Tuff are responsible for a hilly topography. A prominent example is located at the eastern border of the study area where the Überetsch is separated from the Etsch valley by a steep slope down from the Mitterberg with an elevation difference of approximately 400 m. In the western part of the study area, the steep slopes of the Mendola-Roèn-Ridge are dominated by various Dolomite units, with intermittent layers of sand- and siltstones. Except for the very steep

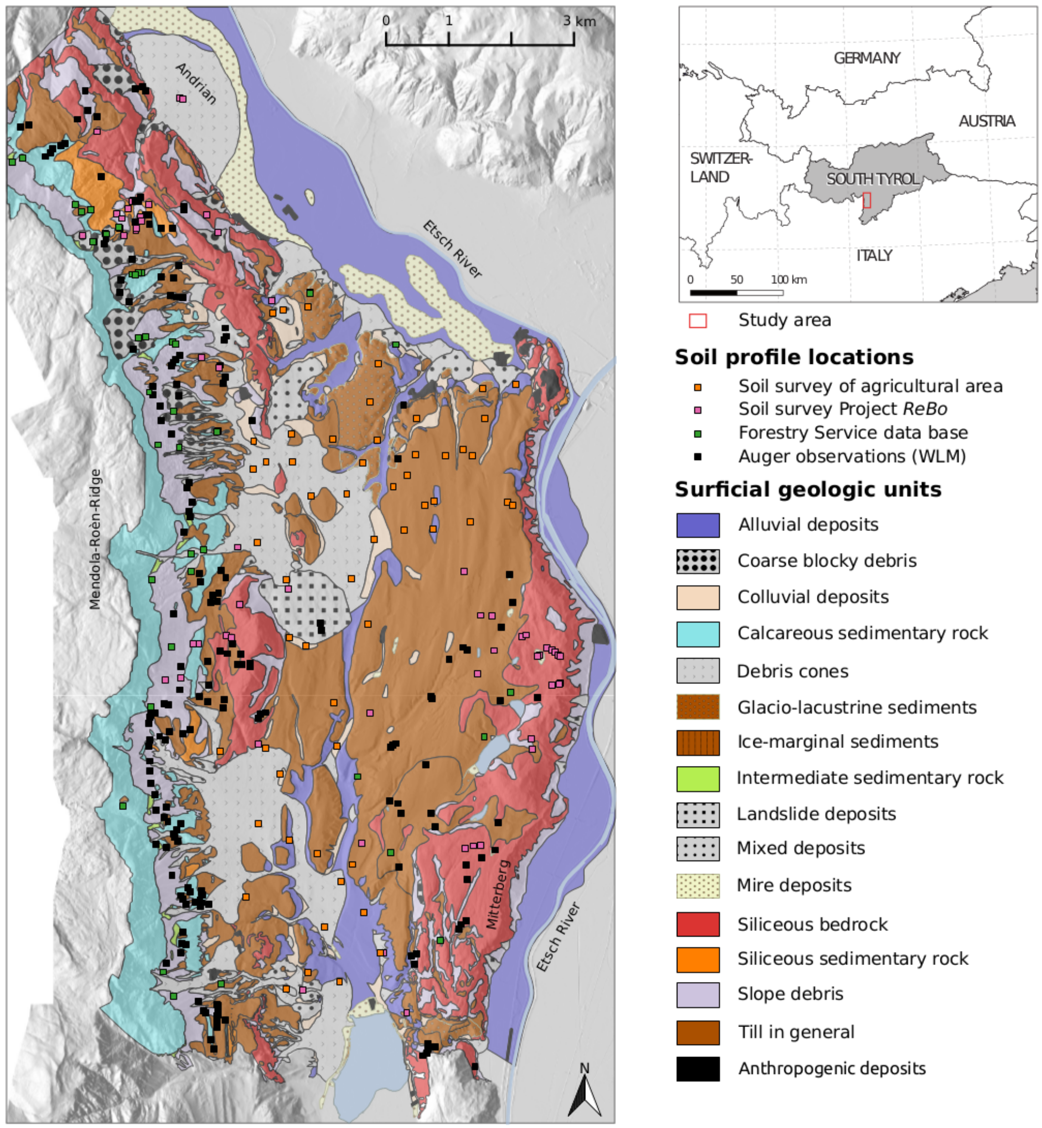

Figure 1: Overview of the study area showing the surficial geology as well as the locations of the soil profile sites analysed in this study. Computations were performed only for the areas covered by surficial geologic units.

Dolomite walls of the ridge, the rarely occurring outcrops of these formations are surrounded and mostly covered by Pleistocene and Holocene slope debris, and in locally flatter areas by till.

The study area comprises two map sheets of the new geologic map of Italy, sheet Eppan, which covers the northern and major part of the area, as well as sheet Mezzo-Lombardo for the southern part. The sheets were published at a scale of 1 : 50,000 in 2007 and 2012, respectively. Mapping was performed at a scale of 1 : 10,000, this information was kindly provided by the Department of Geology and Building Material Tests of the Autonomous Province Bolzano – South Tyrol, in shapefile format and used for the analysis performed in this study.

For simplification of the analysis and data harmonisation, the original geologic map units were grouped and generalized with regard to soil formation into the 16 surficial geology units (SGUs) described in Table 1. This permits a comparison with the parent material units encountered and identified by the soil surveyors in the field.

SGU	abbrev.	short description	% area
alluvial deposits	AD	Holocene and Pleistocene deposits of silt, sand and gravels	14.9
coarse blocky debris	CBD	Holocene and Pleistocene blocky deposits of mass movements	1.8
colluvial deposits	CD	footslope deposits	2.4
calcareous sedimentary rock	CSR	limestones and dolomites	8.4
debris cones	DC	Holocene conic deposits from debris flows and torrents	12.7
glacio-lacustrine and lacustrine deposits	GLD	(fine) sand deposits (with dropstones)	2.5
ice-marginal sediments	IMS	clast-supported gravels	0.2
intermediate sedimentary rock	ISR	silt- and sandstones	0.2
landslide deposits	LD	large landslide deposits	1.2
mire deposits	MrD	Holocene and Pleistocene silt and peat deposits	3.3
mixed deposits	MxD	Pleistocene deposits from debris flows, torrents and avalanches	2.1
siliceous bedrock	SB	rhyolite and rhyodazite tuffs and ignimbrites	13.0
slope debris	SD	Holocene and Pleistocene debris on slopes (gravitational)	10.3
siliceous sedimentary rock	SSR	silt- and sandstones	1.1
till in general	TG	undifferentiated glacial deposits	25.9

Table 1: Generalized surficial geologic units in alphabetic order with abbreviations and short descriptions. Additionally, the proportion of the study area covered by each unit is given. Anthropogenic deposits are not included in the analysis.

2.3. Soil data sets and soil classification

Figure 2 gives an overview of the distribution of the soil types amongst the soil profiles incorporated into the analysis and located within the study area (Figure 1). The soil classification scheme applied in this study is the Austrian system (Nestroy et al., 2000, 2011), as most of the soil profile descriptions available for this study apply this system which is also well adapted to the Alpine environment (Baruck et al., 2016). Furthermore, it is generally recommendable to use local systems for large-scale mapping as they are well adapted to the local circumstances, and not all available soil profile data, especially those from points investigated only with augers, included sufficient

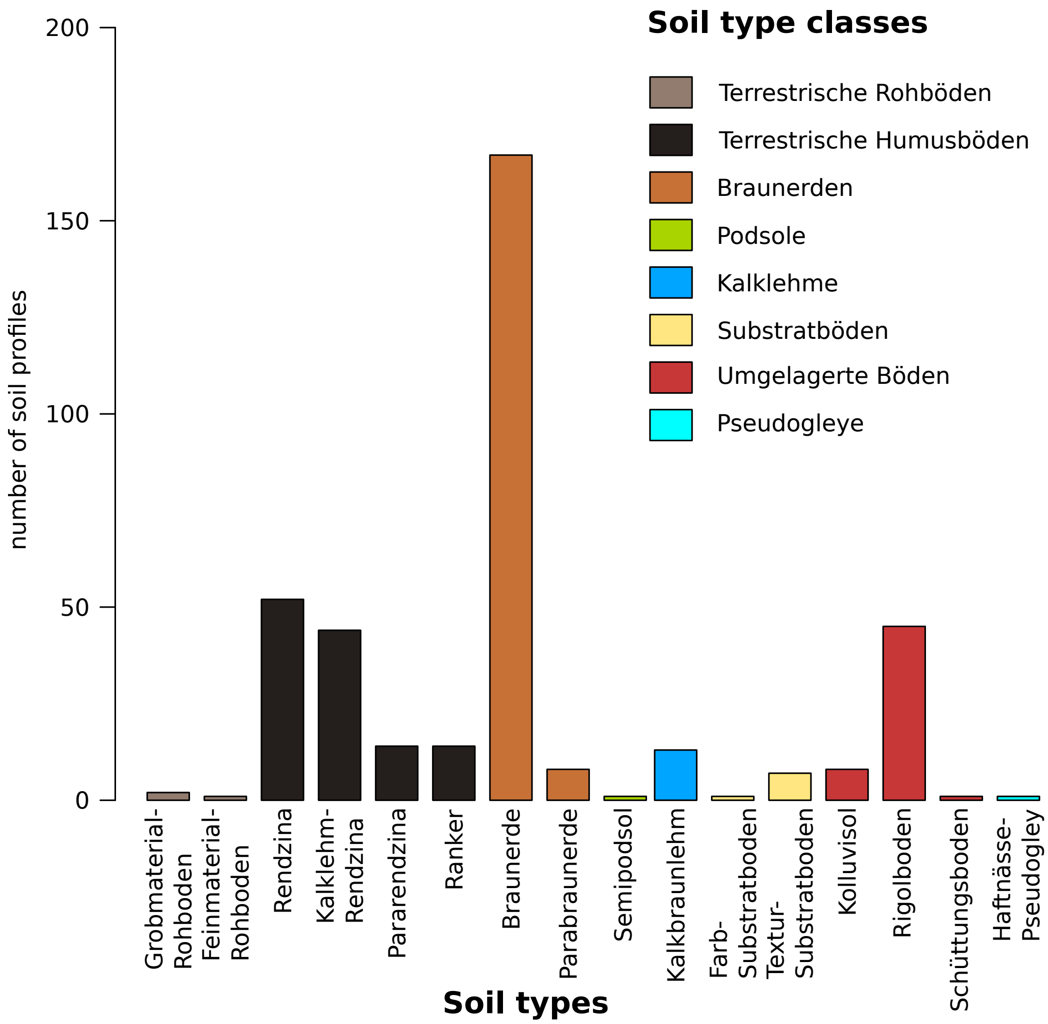

Figure 2: Barplot showing the number of soil profiles (total n=379) for each of the soil types identified at least for one of the data points in the study area. Coloring indicates the soil type classes. Classification was performed according to the Austrian soil classification (Nestroy et al., 2000, 2011). Correlations with WRB reference groups can be found in Table 2.

information for deriving the reference soil group according to the World Reference Base for soil resources (IUSS Working Group WRB, 2015). Table 2 gives an overview of which reference soil groups are best correlated to the relevant soil types in the Austrian classification. Classification of the soil profiles was performed at the subtype level, a basic overview of the Austrian soil classification system can be found in Baruck et al. (2016).

soil type	corresponding WRB groups	short description	soil class
Feinmaterial-Rohboden	Leptosol, Regosol, Histosol, Arenosol	only initial soil formation (A_i-horizon) on parent material with less than 40 V.-% coarse fraction	Terrestrische Rohböden
Grobmaterial-Rohboden	Leptosol, Regosol, Histosol	same as Feinmaterial-Rohboden but with more than 40 V.-% coarse fraction	Terrestrische Rohböden
Rendzina	Leptosol, Histosol	only an A- or organic horizon on calcareous bedrock	Terrestrische Humusböden
Kalklehm-Rendzina	Leptosol	soils with a loamy A-horizon on calcareous bedrock	Terrestrische Humusböden
Pararendzina	Leptosol, Regosol, Umbrisol, Histosol	only an A-horizon on carbonate-containing unconsolidated sediment or bedrock	Terrestrische Humusböden
Ranker	Leptosol, Umbrisol, Regosol	only an A- or organic horizon on siliceous bedrock	Terrestrische Humusböden
Braunerde	Cambisol, Fluvisol, Luvisol, Umbrisol, Regosol	with brown B-horizon owing to weathering and (re-)formation of clay minerals.	Braunerden
Parabraunerde	Luvisol, Albiluvisol, Cambisol	with eluvial A-horizon over clay-enrichened B-horizon	Braunerden
Semipodsol	Podzol, Regosol	characterized by moderate podzolidation	Podsole
Kalkbraunlehm	Cambisol, Luvisol	with a yellow- to redbrown cohesive B-horizon on calcareous bedrock, often fossil soils	Kalklehme
Farb-Substratboden	Regosol, Luvisol	strong influence of color of parent material, overprinting horizon differentiation	Substratböden
Textur-Substratboden	Regosol	strong influence of texture of parent material, overprinting horizon differentiation	Substratböden
Kolluvisol	Anthrosol	developed from fine soil material relocated by (often human-induced) erosion	Umgelagerte Böden
Rigolboden	Anthrosol	influenced by deep, homogenizing human cultivation	Umgelagerte Böden
Haftnässe-Pseudogley	Stagnosol, Planosol	influenced by shallow, capillary stagnation phases	Pseudogleye

Table 2: Corresponding WRB reference groups for the Austrian soil types encountered in the study area, along with a simplified description based on Nestroy et al. (2000, 2011).

Soil data base 1: soil survey of agricultural areas in the Überetsch/Oltradige region. From 1993 to 1995 a soil survey of the farmlands in the region Überetsch was conducted by Thalheimer (2006). Soil types were classified according to Soil Taxonomy, resulting in a soil map with 18 different soil series. 58 detailed soil pit descriptions were incorporated into the presented study, all located either in vineyards or apple orchards. Using the horizon descriptions, chemical properties as well as photographs of the pit face, theses soil profiles were reclassified applying the Austrian System (Nestroy et al., 2000, 2011).

Soil data base 2: soil survey project 'ReBo - Terrain Classification of ALS Data to support Digital Soil Mapping'. During this project with the aim to investigate optimal cooperation between soil survey and terrain classification, 55 soil pit profiles were described in the presented study area based on field work in 2014 and 2015. In addition to detailed soil profile descriptions, grain size distribution, pH analyses and soil organic matter were measured in the lab. Soil classification was performed following Kilian (2015), based on Nestroy et al. (2000, 2011).

Soil data base 3: data base of the Forestry Service of the Autonomous Province Bolzano – South Tyrol. From the point data base provided by the Forestry Service of the Autonomous Province Bolzano – South Tyrol, 42 soil pit locations are situated in the study area discussed in this article. For each of these sites, which either belong to the Forestry monitoring scheme or were mapped during the course of the Forestry type survey (APB, 2006), soil type, parent material and a detailed soil pit description are provided.

Soil data base 4. While performing field work for the Forestry type survey (APB, 2006), the engineering office 'Wald Landschaft Mensch' (WLM) also produced a model of the forest substrate for South Tyrol for use in modeling forest types. For validation, a large number of augur points were described and geolocated, of which 227 are located within the study area and were kindly provided by the company for use in this study. The information for each point includes the soil type as well as a description of the geologic situation, including information on cover beds, if present.

2.4. Digital terrain model

The digital terrain model applied in this study has a grid cell size of 2.5 m and was computed as the result of an airborne laser scanning mission. For the flatter valley bottoms and densely inhabited areas, the average achieved last-pulse point density is reported as 1.3 pts/m^2 by Wack and Stelzl (2005), whereas this measure is lower for the remaining, less populated areas and mountain slopes (0.8 pts/m^2). Using 8 independent sites with 60 ground points for elevation accuracy asessment, the average standard deviation of heights was established as 6.7 cm. While all tiles were previously available for download at the homepage the Autonomous Province of Bolzano – South Tyrol (APB, 2016), the entire data set is now provided freely upon request per email, while smaller areas can still be downloaded. In addition to being used at this high resolution, the DTM was also resampled to 10 and 50 m using average aggregation with the tool 'r.resamp.stats' implemented in GRASS GIS (GRASS Development Team, 2016), in order to consider the issue of scale in the analyses.

3. Methods

3.1. *Terrain parameters with emphasis on roughness measures*

3.1.1. *Local and regional terrain parameters*

The terrain parameters used in this study were computed using the open source GIS GRASS (GRASS Development Team, 2016) and SAGA (Conrad et al., 2015). The computations were performed at different spatial resolutions (2.5, 10 and 50 m) and, if applicable, at different moving window sizes. Local terrain parameters, including slope and minimal, maximal,profile, longitudinal as well as cross-sectional curvature were computed based on the algorithms of Wood (1996) as implemented in the tool 'r.param.scale' in GRASS, which allows varying the computational window size. Regional terrain parameters, amongst them catchment area, topographic wetness index, height above channel network and others, were computed with SAGA both at high (2.5 m) and low (50 m) spatial resolution.

3.1.2. *Surface roughness-related terrain parameters*

Topographic roughness index. The topographic roughness index was presented by Riley et al. (1999) for use in habitat analyses. It compares the elevation of a central pixel to the elevations of cells within a given search window by calculating the square root of the sum of the squared differences.

Terrain texture. In their automated landform classification algorithm, Iwahashi and Pike (2007) present terrain texture, or feature frequency, which can also be interpreted as a measure for surface roughness. This measure is quantified by the number of features, i.e. pits or peaks, in a specified search window. These features are extracted by computing the difference between a DTM and its median-filtered version.

Vector-based roughness measures. In an attempt to decorrelate roughness from slope, Sappington et al. (2007) expanded on the work of Hobson (1972) to introduce vector ruggedness measure (VRM), which is calculated based on the orientation of vectors normal to the surface in a given area. Its implementation as the tool 'r.vector.ruggedness' in GRASS was used for computation in this study. Grohmann et al. (2010), who also authored a GRASS implementation ('r.rougness.vector'), analysed several roughness measures at different resolutions and window sizes with regard to their ability to depict terrain features. They highlight the ability of VRM to detect fine-scale roughness features and attribute low roughness values to steep but smooth slopes, but also acknowledge its inability to delimit slope breaks and identify regional relief.

Landform diversity. Landform diversity can also be interpreted as a form of surface roughness, i.e. a high number of landforms in a given search window signifies high terrain roughness. With regard to landform delineation, Jasiewicz and Stepinski (2013) presented an automated landform classification based on a pattern recognition approach using line-of-site calculations. The two main parameters that can be tweaked for landform segmentation are the search radius up to which the surroundings of a central raster cell are scanned and the flatness threshold angle, below which a surface is considered as flat. In this study, the resulting map, which features a maximum of ten landform classes, is used as the input for the calculation of patch indices such as edge density or mean patch size (mps) and diversity indices like dominance or richness (Figure 3). Computation of the different diversity measures was performed with the GRASS module for landscape structure analysis, 'r.li'. The landform maps were calculated at micro and meso scale (search radii of 7.5 and 50 m, respectively) applying flatness thresholds of 1 and 10°, based on Gruber et al. (2017), who analysed 'r.geomorphon' and other automated landform classifications with regard to optimal parameter values for representing topographic position in soil survey.

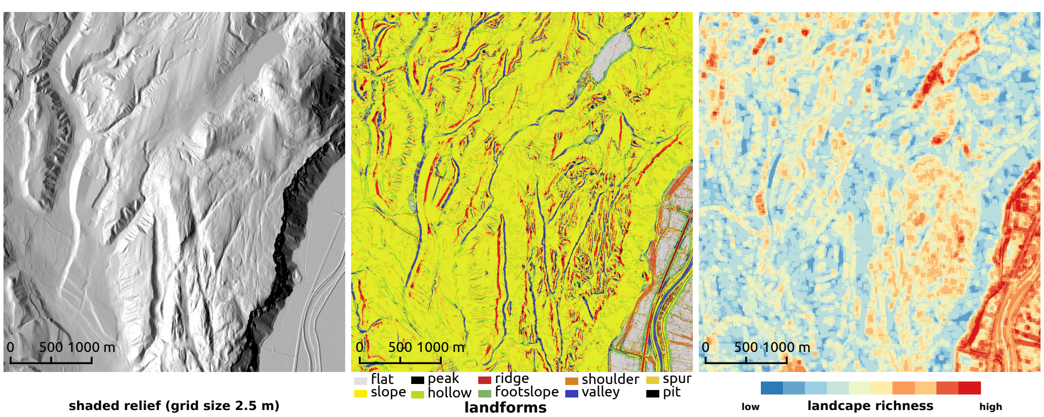

Figure 3: Example of a landscape diversity measure based on a landform map computed for the southeastern quadrant of the study area (Figure 1) with a search radius of 50 m and a flatness threshold of 1°. **A** shows the shaded relief of the 2.5 m DTM on which the calculations are based and **B** the resulting landform map. Based on this, landform richness (**C**) is computed as the number of different landforms in a search window of 20 cells.

3.2. Random Forest classification

Random forests (Breiman, 2001) belong to the family of machine learning algorithms, meaning that based on test data, statistical patterns between explanatory variables and dependent variables are sought and then applied to predict the latter based on new data sets comprised of the dependent variables. A comparison of machine learning approaches for classification in soil mapping can be found in Heung et al. (2016). Contrary to single classification trees, random forests use an ensemble of classification trees to predict class affiliation, each based on a bootstrap sample of the test data and a limited number of randomly chosen explanatory variables. Random forest classification, as implemented in the R (R Core Team, 2014) package 'randomForest', also has the advantage of computing measures for variable importance, such as 'mean decrease accuracy' and 'mean decrease Gini', which provide insight into, and quantify, the importance of each predictor variable in a specific model. Additionally, the out-of-bag error, for which each model based on the bootstrap sample is validated with the remaining unused data, provides a valuable estimate of the prediction error of random forests. As the presented study emphasized the use of random forest classification as an investigative tool over its use for predictive mapping, the authors refrained from optimizing the parameters of the classification algorithm.

3.3. General workflow

In a first step we investigate how well the units of the geologic map correspond to the parent material identified by the soil surveyor at the soil pit locations, thus evaluating the performance and reliability of available geologic maps to support soil survey in South Tyrol. After generalisation of the geologic map units into surficial geology units (SGUs) which can be compared to the parent material units used in the various surveys, a confusion matrix is produced which shows to which extent the geologic units are in accordance with the parent material mapped by the surveyor.

The next step, a data mining approach using random forest classification, is performed to better understand the topographic characteristics of the surficial geology and parent material units. By applying a five-fold cross-validated forward step-wise feature selection and evaluating the 'mean decrease accuracy' of the predictor variables, we identify which terrain parameters are best suited to perform the tasks of:

1. Expanding the geological map to a parent material map by modeling parent material based on the profile site descriptions, the topography of their locations and the information provided by the new geologic map with its detailed information with regard to surficial geology.
2. Discriminating for each SGU between those points which where correctly classified to a parent material class by the geological map and those where this information differed to the parent material identified by the field surveyors.
3. Distinguishing between adjacent soil parent material classes as identified by the soil surveyors during field survey.
4. Separating and consequently characterizing the surficial geology units as described by the geological map.

In contrast to using all possible terrain parameters at once, this procedure produces only a limited number of explanatory variables, which has the advantage of simplifying interpretability, as the main aim is to identify, highlight, and understand those situations were the most confusion between parent material units as documented by surveyors on the one side, and the geologic map on the other side, occur. The whole feature selection procedure for each of the tasks was itself performed in five-fold cross-validation with four parameter sets: 1) all terrain parameters, 2) local terrain parameters, 3) regional terrain parameters, 4) surface roughness-related parameters. Regarding task 1, emphasis was soon set on the roughness measures, which are deemed best suited to enhance the geologic map, as the geological units already implicitly contain information with regard to regional and local terrain parameters.

The connection between the two important soil forming factors, parent material and topography, on the one hand, and soil as the result of theses factors on the other, is then investigated by analyzing the diversity and distribution of soils for each geologic unit. This is performed from two points of view: the soil type distribution is done for profile sites per geologic unit, but also per parent material unit as attributed by the soil surveyor. This gives insight into how the surveyors' mental soil landscape model relates specific parent material units to specific soil types, especially when applying a morphologic-genetic classification such as the Austrian soil classification (Nestroy et al., 2000, 2011). The synthesis of this information regarding topography, parent materials, and soils then leads to a geologic-topographic characterization (GTC) that describes each surficial geology unit.

4. Results and discussion

4.1. Comparison of soil parent material at soil profile sites with geologic map units

For the following results and discussion, the confusion matrices and the accuracy measures are calculated and interpreted as if the field observations described by the surveyors had been planned to validate the use of the geologic map as a map of soil parent material. That is, the columns represent the reference classes while the rows are the classified classes. The producer's reliability is therefore calculated by considering the values and sum per column of the confusion matrix, whereas user's accuracy, or its counterpart, the error of commission, is obtained by aggregating across the rows. A comparison of the soil parent material class identified in the field survey and the SGU from the geological map for the same position shows that while these match for 186 of the profile sites, leading to an overall accuracy, or correct classification rate, of 49 %, there is a high rate of misclassification.

A first view at the confusion matrix (Table 3) indicates that a considerable amount of these misclassifications involve the classes till (TG) and slope debris (SD).

4.2. Geologic-topographic characterization of the predominant SGUs

Through synthesis of i) results of the comparison of the soil point data and the geologic map, ii) the random forest analysis of terrain parameters and iii) the distribution of soil types in the study area, a geologic-topographic characterization (GTC) of the SGUs for use in soil survey is performed. Only those units with substantial areal extent and sufficient soil profile points are described in detail.

4.2.1. Alluvial deposits

The SGU alluvial deposits (AD) occupies 16.3 km^2, amounting to 14.9 % of the study area. It incorporates alluvial deposits in the paleovalley but also in the current Etsch valley (see Figure 1). The existing soil profile points are however limited to the paleovalley. The alluvial deposits share very long borders with the SGUs till, debris cones, mire deposits and colluvial deposits, due to the units bifurcated north-south transection of the study area. Long borders can therefore also be found with the units glacio-lacustrine deposits, mixed deposits, slope debris and till. There is agreement between the geological map and the soil profile description in 7 (54 %) of the 13 profiles

	AD	CBD	CD	CSR	DC	GLD	IMS	ISR	LD	MrD	MxD	SB	SD	SSR	TG	sum
AD	**7**	0	0	0	0	1	0	0	0	0	0	0	1	0	3	12
CBD	0	**4**	0	0	0	0	0	0	0	0	0	0	8	0	2	14
CD	0	0	**3**	0	0	0	0	1	0	0	0	0	1	0	0	5
CSR	0	0	1	**2**	0	0	0	7	0	0	1	0	13	0	12	36
DC	4	0	3	0	**5**	0	0	2	0	0	0	0	20	0	2	36
GLD	0	0	0	0	0	**5**	0	0	0	0	0	0	0	0	2	7
IMS	1	0	0	0	0	0	**0**	0	0	0	0	0	0	0	0	1
ISR	0	0	0	0	0	0	0	**0**	0	0	0	0	3	0	0	3
LD	0	1	0	0	0	0	0	0	**0**	0	0	2	0	0	0	3
MrD	0	0	0	0	0	2	0	0	0	**0**	0	0	0	0	1	3
MxD	0	0	1	0	0	0	0	0	0	0	**0**	0	6	0	1	8
SB	0	1	1	0	0	2	0	1	0	0	2	**14**	3	1	24	49
SD	0	3	0	2	1	0	0	4	0	0	0	4	**55**	0	15	84
SSR	0	0	0	0	0	0	0	2	0	0	1	0	3	**3**	1	10
TG	1	0	0	1	0	1	0	3	0	1	5	0	9	0	**88**	109
sum	13	9	9	5	6	11	0	20	0	1	9	20	122	4	151	**186**

Table 3: Tabular comparison of parent material geounits as observed by soil surveyor (columns) and in the geologic map (rows). Abbreviations: AD = alluvial deposits, CBD = coarse blocky debris, CD = colluvial deposits, CSR = calcareous sedimentary rock, DC = debris cones, GLD = glacio-lacustrine and lacustrine deposits, IMS = ice-marginal deposits, ISR = intermediate sedimentary rock, LD = landslide deposits, MrD = mire deposits, MxD = mixed deposits, SB = silceous bedrock, SD = slope debris, SSR = siliceous sedimentary rock, TG = till in general

for which the soil surveyors identified this SGU as the soil parent material. Some confusion with the SGUs debris cones and till can be observed, the former from the viewpoint of producer's reliability and the latter as an error of commission.

An investigation of the soil profile sites that were identified by surveyors as having alluvial deposits as parent material but are located on a different unit on the geological map, shows that these locations are characterized by consistently lower vector strength at 112.5 m window size, indicating rougher terrain. This is confirmed by higher landform richness values at meso scale along with higher vector ruggedness for the incorrectly classified data points (Table 4). Addressing the confusion of alluvial deposits with till, the random forest investigation shows that besides obvious channel-level-related terrain parameters, which well separate the units, convexity contributes strongly to

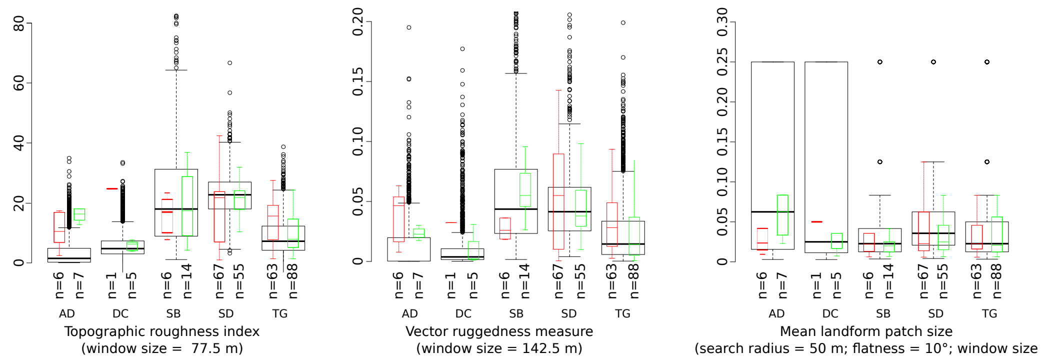

Figure 4: Box plots of three different roughness measures for the most common and confused parent material classes. The large black box plots characterize the topography of the respective units of the geologic map, while the smaller box plots inside the larger ones represent the values of the soil profile sites for which the surveyors identified the corresponding parent material class. Green boxes indicate the distribution of the terrain parameters for those profile sites for which surveyor and SGU map agreed regarding the parent material, whereas red boxes represent sites located on a different SGU.

distinguishing between alluvial deposits and till on the geologic map, with the former characterised by lower values. It must however be taken into account that the SGU in the study area also incorporates parts of the wide Etsch valley floor, whereas the soil profile points are situated in the paleovalley of the Überetsch. The same investigation, but performed from the viewpoint of the surveyor (i.e. using the field data points and their parent material information), identifies a very local (small window size) version of the TRI as best suited to distinguish between the two parent material groups, with AD interestingly characterised by higher values indicating higher ruggedness. This can be attributed to the relative smooth surface of the till unit on the one hand, and the often entrenched location of alluvial deposits in the paleovalley. The separation of the profile sites of the parent material classes AD and debris cones, the second most common confusion for AD, is less clear than for AD and TG. High resolution roughness values are less important than in the case of till, with texture computed with 50 m grid cell size being the most decisive parameter. A comparison of the geological map units' topography also highlights the importance of this roughness parameter, which is surpassed only by the regional terrain parameter catchment slope, computed at the same low DTM resolution. The choice of these parameters seems closely linked to the transitional landscape position of debris cones, situated between the steep slopes of the Mendola-Roèn-Ridge and the flatter valley bottoms.

Of the 11 soil profile or auger points located on areas covered by alluvial deposits according to the reclassified geological map, the soil type Braunerde is predominant, with some occasional anthrosols, represented by the Austrian class Rigolboden. The brown soils (Braunerde) are to be expected as the more or less pronounced stability of the flat alluvial deposits of the paleovalley have allowed a certain degree of pedogenetic processes to occur. The anthrosols on the other hand are typical for the orchards and vineyards, commonly found on alluvial deposits in the region, where landscape as well as soil have seen strong anthropogenic influence. The distribution of the soil types of the profile points identified to have alluvial deposits as parent material by the surveyors is comparable regarding the dominance of brown soils. Additionally, the soil types Kalklehm-Rendzina and Pararendzina were encountered, both characterised by A-horizons rich with organic material, on more or less unweathered carbonate-containing parent material, indicating that alluvial deposits were also identified at places lacking the prolonged stable conditions necessary for the development of a B-horizon, especially on calcareous material. So while the alluvial deposits unit is characterized on the geological map by the lowest mean slope aside from mire deposits, the soil data as well as the topographic analysis indicate that it is nevertheless necessary for future surveys to also investigate the less typical, rougher and sloping areas at the border or in proximity of alluvial deposit units. Furthermore, the carbonate components suggest that local material from the western slopes is also incorporated in alluvial deposits of the paleovalley.

4.2.2. Siliceous bedrock

The SGU siliceous bedrock in the study area is characterised by outcrops of rhyolitic ignimbrite in the Überetsch paleovalley, for instance forming the Mitterberg which seperates the paleovalley from the current Etsch valley. The unit represents 13 % of the study area, and shares a long border with the SGU till, but also the units coarse blocky debris and debris cones. Shorter borders exist with almost all other SGUs, especially mixed deposits and slope debris. Regarding the comparison of the unit on the geological map with the parent material as reported by the surveyors, there is an interesting discrepancy between user's accuracy and and producer's reliability when attempting to use the geological map as a parent material map. While the parent material of only 14 of the 49 profile sites located on siliceous bedrock according to the geological map was identified as siliceous bedrock by the surveyors, 70 % of all soil pits with this parent material were actually on the correct unit of the geological map. This means that when investigating soils on siliceous bedrock, the probability that they are encountered in the unit SB is higher than for other units, however it is similarly probable to encounter other parent materials in this unit. The greatest confusion occurs with the unit till, which was found to be the parent material of almost every second soil profile located on the SGU siliceous bedrock. Further parent materials identified on the SB unit include mixed and slope debris, as well as glacio-lacustrine deposits (Table 3). Confusion with slope debris or mixed deposits is understandable due to the often fuzzy transition from weathered bedrock to slope debris, or the fact that mixed deposits may very well contain exclusively siliceous material from bedrock units at higher elevations in the catchment, or underneath the debris cover layer. Additionally, the soils resulting from such parent material must not necessarily be very different to those on siliceous bedrock. Till, as a parent material, is a different case, as the material may be derived from catchments with a very different geology and the grain size distribution is not similar to that of slope debris or siliceous bedrock. When evaluating the misclassification of the parent materials till and siliceous bedrock, it is essential to consider possible differences regarding the mapping procedures in geology and pedology. Furthermore, the parent material layer must not necessarily be very thick to act as such. So, whereas the soil surveyor is particularly interested in the parent material, i.e. the material which through pedogenetic processes and weathering slowly becomes the solum no matter the thickness of this specific layer, the geologist's main concern is the underlying material, and may consider drapes of till only once their thickness reaches a certain threshold, for instance 1 m, as conveyed by the surveyors of the geologic survey of South Tyrol.

The random forest-based analysis of how to topographically seperate the profile sites with siliceous bedrock as parent material that are situated on the SB unit of the geological map, from those with the same parent material but on different SGUs, showed that the latter group is situated in areas with lower convexity. Regarding roughness parameters, the vector ruggedness measure based on the 2.5 m grid and a large search window of approx. 130 m performed best. Soil profiles that evolved from siliceous bedrock and are also located on the SGU SB are characterised by higher vector ruggedness values. Given the strong confusion with the unit till, as well as the fact that these two units share a long border, it is of interest to examine the topographical differences between points that are attributed to one of these two parent material units. The TRI at a window size of approx. 100 m leads the terrain parameters with a 5-fold cross-validated classification error of approx. 15 %. The minimum curvature based on the 10 m grid performs best from the group of local terrain parameters. In general, soil profile sites with SB as parent material are characterized by rougher topography (increased TRI values) and slightly positive minimal curvature values when compared with data points were till was identified as parent material. This is well in line with the increased appearance of this parent material on convex, ridge-like structures in the SGU siliceous bedrock. As Table 3 shows, confusion with slope debris is present from both the user's and the producer's point of view. The random forest-based data mining shows that, regarding the parent material as indicated by the surveyors, the terrain parameters that are most characteristic of the difference between the two units are based on landform diversity. In the case of SB and SD, the patch density of the micro scale landform map helpa separate the two parent units due to the latter groups lower values, implying that areas of slope debris have a slightly more homogeneous landform distribution, in this specific case large polygons of the landform slope, apparently a good indicator for slope debris. Furthermore, the same TRI applied to separate SB and TG can also be a very useful addition to the parameter set when distinguishing SB from SD.

An investigation into the soil types classes of the soil profile sites located on the SGU siliceous bedrock shows that brown soils are the dominating soil type, represented by 37 of the 49 locations (75 %). The class Terrestrische Rohböden, represented by the soil type Ranker, is second most common on this SGU, on which single examples of the soil types Semipodsol, Kalkbraunlehm, Farb-Substratböden and Kolluvisol were also encountered. Taking into account only the soil profile sites for which the surveyors identified SB as parent material, whether or not actually located on this SGU, the soil type brown soils similarly dominates the soil type Ranker, with 13 and 6 profile sites, respectively. All of the other soil types found on this unit of the geological map were identified by the surveyors to not have SB parent material, however one soil profile site with this parent material was attributed the soil type Grobmaterial-Rohboden. This soil type however does fit well for soils on siliceous bedrock outcrops, where only initial soil formation is possible, whereas the other soil types located on the SGU of the geological map are less plausible on SB, especially Kalkbraunlehm. In addition, the vast majority of the soil profile sites located on this SGU but identified to have a different

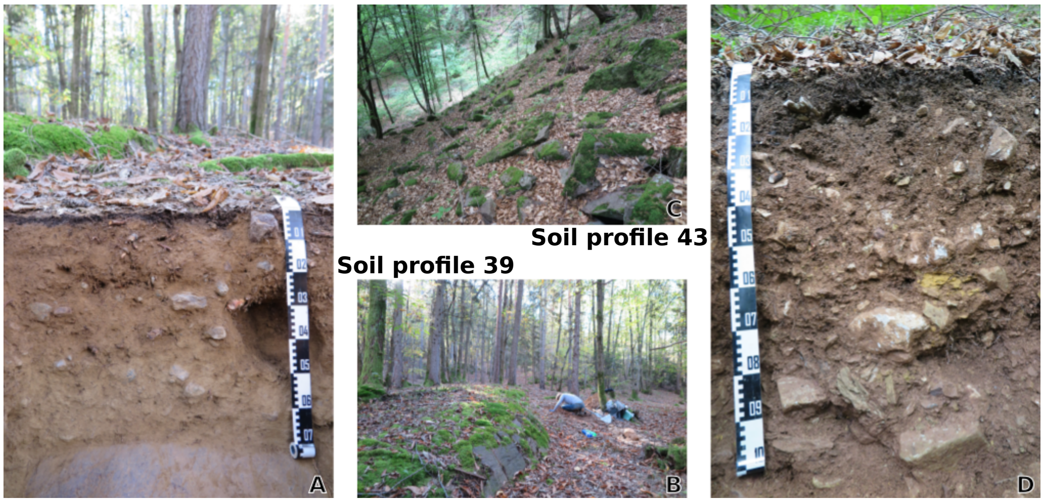

Figure 5: A) Soil profile 39 on the SGU siliceous bedrock with parent material identified as till, representative of the typical situation of till on siliceous bedrock. B) Soil profile site 39, with visible rhyolite outcrops. C) Soil profile site 43. D) Soil profile 43, located above the debris cone of Andrian on the SB map unit. The yellow components are from up-slope ISR layers, while calcareous components from even higher layers can also be identified.

parent material class (mostly till) by the surveyors, also feature the soil type Braunerde, thereby decreasing the severity of misclassifying SB and TG, at least from a soil type perspective. Figure 5A shows an example of such a situation, consisting of a brown soil developed from a cover layer of till on siliceous bedrock. From this perspective, the soil type Ranker is more characteristic of the SB unit, as it is less common on TG with its characteristically more developed soils and the increased influence of calcareous materials. Furthermore, an examination of the points where parent material classes from observation and SGU map coincide shows a decrease in dominance of brown soils compared to Rankers, with proportions of 57 and 35 % for both soil types, respectively. This is also well in line with the fact that SB sites are characterized by higher roughness values and situated on convex, ridge-like structures where decreased soil formation can be assumed. In conclusion, the soil types Braunerde and Ranker are characteristic of this unit, with the former representing locations were weathering and clay-reformation has led to horizon differentiation not present in the less developed Ranker soils which are found on siliceous bedrock outcrops. Figure 6 shows the distribution of soil types on siliceous bedrock as well as other SGUs from different points of view, i.e. according to the SGU map, the soil survey, and a union of both.

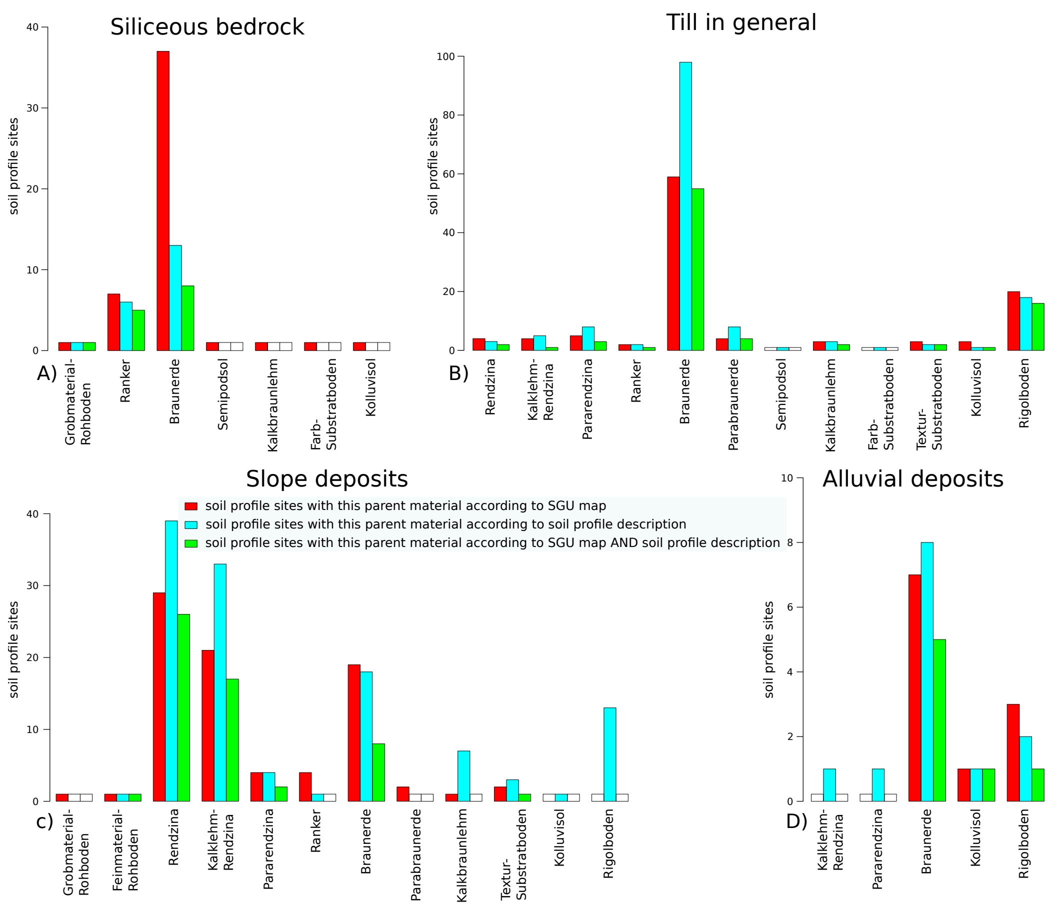

Figure 6: Distribution of soil types for the most common SGUs or parent material units from different points of view. Red bars indicate the number of profile sites with this soil type according to the SGU map while the turquoise bar represents the number of profile sites for which the surveyors identified this parent material, regardless of the SGU it is located on. The green bar shows the number of profile sites for which both points of view agree.

4.2.3. Till in general

The unit till comprises lodgement and subglacial till, as well as other, undifferentiated till materials. It is found in the paleovalley as well as on flat terraces of the the Mendola-Roèn-Ridge. Covering more than 25 % of the study area, it is the most common SGU on the geological map, sharing borders with every single SGU, the longest, each with a length of at least 30 km, being those with the SGUs alluvial deposits, siliceous bedrock and slope debris. According to the geologic map, 109 of the soil data points are located on the SGU till, accounting for 29 % of the points. In 88 of these locations, the surveyors agreed with regard to the parent material being till, leading to a user's accuracy of 80 %. The majority of the other soil data points on this SGU were attributed with the parent materials slope debris or mixed deposits, with some intermediary sedimentary rocks also identified as parent material. While the user's accuracy for till is the best amongst the SGUs, the producer's reliability is not comparable, as the surveyors also identified till as the parent material at 63 further locations on different SGUs. The relative majority of these locations are found on the unit siliceous bedrock, already discussed above, but a large number of soil profiles with this parent material are also located on the units calcareous sedimentary rock and slope debris. Furthermore, till was identified as the parent material of soils on the units SSR, MxD, MrD, GLD, DC, CBD and AD. An important takeaway point from these results is that while, at least for this study area, the SGU unit till is a good indicator for where to reliably encounter soils evolved from till, it is of great importance to expect this parent material also on various other SGUs.

A topographic comparison of those points with till as parent material located on the SGU till with points with the same parent material but on different SGUs (Figure 4), shows that the misclassified, latter group is characterized by a higher roughness, as characterized by a higher TRI and also a steeper slope (mean value of 21 compared to 12°). Regarding confusion with the soil parent material unit calcareous sedimentary rock, data mining the terrain parameters of the relevant soil profile sites again highlights the role of the parameter TRI, in this case dominantly at a window size of 50 m, in topographically separating the unit till from other parent material units. CSR profile sites are characterized by TRI values in the upper quartile of the values characteristically displayed by TG profile sites. Of the local terrain parameters, slope, together with minimal curvature, performs well in separating the points of both groups, but not as clearly as it separates both units on the geological map, as TG profile points in the study area can be found at slope values higher than indicated in the geological map. To characterize the difference of the TG sites with slope debris, the TRI with a window size in the range of 100 m performs best of all local, regional and roughness-related terrain parameters in the forward stepwise feature selection. Similar to the situation with CSR, slope is the best performing local terrain parameter, but as before, the separation is clearer on the geological map, signifying that till can be found at steeper angles than expected from the geological map. Till as the parent material of soil profiles on the surficial geological map units CBD, DC and GLD is not as common as for the above discussed units and additionally these units do not have as many overall members, nevertheless some interesting observations can be derived from the random forest-based investigation. For instance when compared with till, debris cones, which have a considerably uniform topography at the analyzed grid cell size of 2.5 m, are characterized by lower local roughness (VRM with a window size of 50 m or TRI with a window size of only 7.5 m) and a low landform diversity as represented by a high mean patch size of landforms calculated at meso scale and a flatness threshold of 10°. The profile site topography of the unit coarse blocky debris, on the other hand, is distinguishable from till by a higher TRI at a window size of 125 m as well as steeper slope. Other aspects that may account for the high number of till parent material sites on units other than till have been discussed in the siliceous bedrock section.

Typical for the study area, and especially characteristic of the till parent material unit, brown soils dominate the soil types of the profile sites on the corresponding map unit with 59 %, but even more so when considering only the sites for which the surveyors identified till as parent material (70 % of 151 profile sites). This proportion is comparable to that of brown soils for the data subset for which till is indicated as parent material by both the surveyors and the geological map. While brown soils also dominate on SB, the variety of subtypes on till is greater. While SB shows mainly typical brown soils and some with signs of beginning podzolisation, these subtypes are accompanied on till by brown soils containing calcareous components as well as brown soils characterized by higher clay contents, a strong influence of the clay on horizon differentiation (soil type Textur-Substratboden), and also vertical translocation of these clay particles (soil type Parabraunerde, which develops only on this parent material in the study area). The latter soil types and subtypes are common on till parent material, especially when developed on sub-glacial or lodgement till. The presence of soils of the class of Umgelagerte Böden, i.e. soils that have been rearranged vertically or horizontally either directly or indirectly by human influence, can be accounted for by the fact that most vineyards or apple orchards are located on this unit. Of the class of less developed soils with only an A-horizon rich with organic matter (Terrestrische Humusböden), accounting for 13% of the soils with the parent material till, the majority is influenced by the calcareous components of the till, clearly separating these soils from those on SB. The role played by the calcareous till material, brought into the till mainly locally from the slope of the Mendola-Roèn-Ridge, is also highlighted by the presence of soils of the class Kalklehme. In summary, the soils of the SGU till are characterized by well developed soils, often with calcareous properties. This is well in line with the topography of this unit, distinguished by low roughness and gentle slopes.

4.2.4. Slope debris

Slope debris, as a SGU of the geologic map, occupies 10% of the study area. Its by far longest border is shared with the unit siliceous bedrock, other important borders are with the units till, calcareous sedimentary rock and debris cones, the first three units greatly influencing the distribution of components in the slope debris units. When comparing the parent material of soil profile sites with the map SGUs, the slope debris unit has slightly better user's accuracy than producer's reliability, as surveyors established slope debris as the parent material of 55 of the 84 profile points on this SGU, but also for 69 soil data points on other SGUs. Similar to the SGU siliceous bedrock, the most confusion regarding parent material on the SGU slope debris occurred with the unit till, which was identified at 15 soil profiles. As is the case with the SB unit, some of this confusion may be attributed to thin layers or punctual deposits of till. Considering slope debris, another important aspect is that the debris in question may very well be composed of till material that has been transported gravitationally. The same explanations may hold for other parent materials which were identified on the slope debris unit, especially for the bedrock units SB, ISR and CSR. While some isolated outcrops are possible, the most likely cause is that the constituents of the slope debris are so dominated by transported material of one of these bedrock classes, that the surveyors determine this unit as the parent material of the examined soil profile. On the other hand, misclassification between the units coarse blocky debris and slope debris can be attributed to the fuzzy border between these units, ultimately linked to the grain size distribution, and the subjective interpretation thereof, especially during field survey. Contrary to most other SGUs (with the exception of CBD, LD and MxD), this unit itself does not provide information regarding the mineralogy of its component, which can only partially be derived through interpretation of its location in the catchment and the uphill geologic situation. Consequently, this unit is much better described by its topography than its material, as the latter may be highly diverse, as demonstrated in Figure 5D.

Those profile sites for which the geological map correctly proposes slope debris as the soil parent material differ from the slope debris points on other SGUs mainly by higher slope heights and also larger slope angles, i.e. the geological map tends to under-represent slope debris situations in lower regions of the catchments and at less steep positions. 20 of the 122 soil profiles with slope debris as parent material are located on debris cones according to the geological map, which are good examples of the just mentioned situation. This is presumably closely related to the furrowed and disturbed character of debris cones in the southern half of the study area, where the situation may not present itself clearly during field survey. Additionally, the material composition as well as its origin is basically the same for slope debris and debris cones in this area. Topographically, roughness measures do well in separating the slope debris points from the debris cones points of the field survey data set. Slope debris is characterized by higher vector ruggedness values at a window size of 52.5 m, which are very low for debris cones. Slope angles also differ considerably, implying that the surveyors tend to annotate slope debris if the topography does not show the low roughness and slope values characteristic of debris cones. This difference of roughness and slope with regard to the conceptual difference of slope debris and debris cones is also apparent when comparing the topography of both units as delineated by the geological map, however the difference is more distinct in this case. Similarly, calcareous sedimentary rock as parent material seems connected to roughness values slightly higher than those present for the average slope debris site, a relationship which is again more pronounced in the geological map. Consequently, it is in these transitional zones were the misclassification occurs. This topographical proximity of the units is of course reinforced by the fuzziness of distinguishing between the two parent materials as such. As is the case with siliceous bedrock and slope debris, the border between the source material, in this case the calcareous bedrock, and its derivative, the slope debris, is often not clear when considering the parent material of a soil profile. The degree of weathering of the source material certainly adds to the confusion. The issues are comparable for the parent material units intermediate and siliceous sedimentary rock. While the gradual transition from source material, in these cases bedrock of differing mineralogical composition, to slope debris poses various issues that cannot be explained only by topography, this is even more the case when considering the possibility of confusion between slope debris and coarse blocky debris, the distinction between which raises a number of questions with regard to the objectivity of estimating block or grain size of soil profile parent material. Misclassification between the units slope debris and mixed deposits are also predominantly a question of definition and the explanatory power of topography is very limited.

Given its confounding nature due to its dynamic topography and great variability of components, this unit shows a high diversity of soil types, encompassed only by the parent material till due to its variety of anthrosols and well developed soils. With regard to soil type classes, the class incorporating soils with only A-horizons with high organic matter content on top of unweathered parent material (Terrestrische Humusböden) stands out, both with respect to the soil profiles on the map unit (70%) as well as those for which surveyors specified the parent material slope debris (64%). While both of these approaches of analyzing the distribution of soil classes have brown soils (25 and 15%, respectively) as the most important secondary soil class, the main differences are the classes Kalklehme and Umgelagerte Böden, of which the latter is found only from the viewpoint of the soil profile descriptions. The presence of multiple soil profiles of the Kalklehme class shows that soil surveyors are more likely to identify slope debris as the parent material also for locations which are generally more favourable to soil formation than the areas indicated as this SGU on the geologic map. This is indeed consistent with the topographic analysis of the profile points. Similarly, these transitional locations between steeper slopes and flatter regions are also indicative of the confusion between slope debris and debris cones, exemplified by the relatively large proportion (12%) of the soil type Rigolboden, which also highlights that soil surveyors tend to emphasize the components, in this case slope debris, over the geomorphological form as presented by debris cones. In summary, soils with only organic material-rich A-horizons on C-horizons are the most common representatives of soils on this parent material, with those related to calcareous parent material (Rendzina and Kalklehm-Rendzina) more prevalent than their siliceous counterparts, especially when considering only those soil profile locations for which surveyors identified slope debris as parent material. Nevertheless, soils developed from slope debris can show signs of advanced soil formation in less sloping, smoother topography, for instance on neighboring debris cones or other deposits.

4.2.5. Calcareous sedimentary rock

The calcareous sedimentary rocks in the study area are predominantly responsible for the steep walls of the upper elevations of the Mendola-Roèn-Ridge (dolomite and limestone), but also represent some thin layers in lower parts of that slope, interchanging with intermediate and siliceous sedimentary rock layers. Despite covering only 8.4% of the area, this unit has long borders due to its layers that span from north to south of the study area. Slope debris units often occupy locations downslope of the CSR units, accounting for the long border length of 40 km with this unit. The fuzzy transition from weathered bedrock to slope debris is a major issue also for this SGU. The confusion matrix representing the comparison of the parent material as indicated by the geologic map and the parent material identified in field survey (Table 3) shows that, consequently, this is the unit responsible for the most misclassification, contributing to the very low user's accuracy of this SGU. Additionally, due to its intermittent layering with ISR, it is not surprising that this unit was found to be the parent material for seven investigated soil profiles in the calcareous sedimentary rock unit. The producer's reliability is slightly better, however, all in all only five of the soil profiles were attributed the parent material unit calcareous sedimentary rock, of which two were in fact situated on the SGU slope debris. This again accentuates the problem with differentiating these two units when adjacent, as one is often the result of weathering and gravitational transport of the other unit. Similarly to the other bedrock units in the study area, till was also reported as the soil parent material at twelve sites in this unit, accounting for a third of the soil profile locations on this SGU and highlighting the necessity to expect thin layers of till on areas mapped as bedrock. Due to the small number of soil profile points actually identified as having CSR as parent material, the random forest classification approach of identifying important terrain parameter differences between the correctly classified points and those on other units is not very meaningful. Topographical separation from intermediate sedimentary rock, a major cause of misclassification, is also not very adequate, as the most important parameter in this case is slope height. This is to be expected, since the main difference is the vertical position in the overall geological structure of the study area, a difference which is however diminished by the above mentioned intermittent layering and gravitational movement down-slope, in addition to the above discussed problem regarding the definition of slope debris versus weathered bedrock.

The Rendzina is the typical proponent of soil types on calcareous sedimentary rock, as exemplified by the two soil profile sites for which both the geologic map and the surveyors indicate this unit as the soil's parent material. The situation described in the topographical description of the unit leads to a high number of soil profiles along with a variety of different soil types. Although these include soils where soil formation processes have lead to B-horizon differentiation such as Braunerde and Kalklehme, the class Terrestrische Humusböden nevertheless constitutes the majority. However, when considering only the soil profile sites that qualify for CSR parent material from the surveyor's point of view, this is the only class, represented by the soil types Rendzina and Kalklehm-Rendzina. Given the topographic situation and the complex geologic setting, it is necessary to assume the possibility of soils with signs of increased pedogenetic development on this map unit, however a different parent material such as till or slope debris of varying composition must be considered for any of these soil types.

4.2.6. Debris cones

The unit debris cones is located west of the center of the paleovalley, between the slope debris of the Mendola-Roèn-Ridge slope in the west, which are often the source area of these deposits, and the till deposited in the paleovalley in the east. An additional, pronounced debris cone is located at Andrian in the north of the study area in the main Etsch valley. Units with which long borders exist are mixed, colluvial, and alluvial deposits. The number of profile sites is comparably small with 36 profile sites, considering that the unit occupies almost 13% of the study area. A reason is that a large part of the debris cones are covered by settlements. Although the soil surveyors noted debris cones as parent material for only five soil profile sites, it must be considered that 20 of the misclassified parent materials on the debris cones unit were identified as slope debris. Given the long mutual border and the fact that the source of the debris cones material is mainly the slope debris and the calcareous bedrock units which themselves are the origin of the slope debris, this misclassification may seem acceptable. In fact, it rather points out a difference in the point of view of the soil and the geologic surveyors, where the first group is more interested in the material while latter emphasizes the landform presented by the debris cone unit. The remaining misclassifications are all with units that border this SGU, or, in the case of intermediate sedimentary rock, are located in the source region of the the debris cones.

The topographic issues regarding the confusion with other parent materials such as AD, SD and TG have been discussed in the unit-specific subsections, highlighting this units landscape position as the main distinguishing characteristic.

As discussed, compared to the number of soil pit locations for which the surveyors identified the parent material class debris cones, the number of sites located on this unit of the geological map is much higher, which also leads to a greater soil type variability, especially at subtype level. While surveyors identified this unit only once for more developed soil types such as Braunerde or Kalklehme, they constitute 42% of the sites indicated on the map. Given the strong anthropogenic influence on debris cones, a large proportion of the profiles have experienced vertical anthropogenic rearrangement in an effort to enhance the general conditions for viticulture and horticulture. Soils with organic matter-rich A layers can also be found in areas with less intensive anthropogenic influence, though more commonly on areas where only the geologic map indicates debris cones, whereas the surveyors more commonly classified the parent material as slope debris, highlighting the issue of defining the difference, as discussed in section 4.2.4.

4.2.7. Intermediate sedimentary rock and siliceous sedimentary rock

The units intermediate and siliceous sedimentary rock are situated in the slope of the Mendola-Roèn-Ridge and are characterized by more or less thin layers, often intermittent with layers of calcareous sedimentary rock, and only a limited amount of outcrops. Another common characteristic is that none of the soil profile points for which the surveyors identified one of these units as the soil's parent material is actually situated on the corresponding geological map unit. The parent material of all three soil profile sites on ISR was classified as slope debris, which is understandable considering the issues discussed in the previous sections regarding slope debris and weathered bedrock. As further evidence of this issue, ISR was detected as the parent material of four soil profiles on the map unit slope debris. Additionally, intermediate sedimentary rock was identified as the parent material of 7 soils on the CSR map unit and two on the SSR unit. There is of course no question that the layers are correctly located on the geological map with regard to the underlying geological structure of the study area. These soil profile sites rather highlight the possibility of finding any or all of the materials of these intermittent layers on these units as well as any potential down-slope slope debris units due to the prevailing geomorphological dynamics in the study area. Gravitational transportation leads to these materials serving as soil parent material not where the layer is indicated on the geological map, but in fact down-slope of these layers. Additionally, these materials are seldom found as homogeneous units due to mixing and multi-layering caused by gravitational transport. For these and other reasons discussed in the CSR section, topographical separation is limited for these units.

As the situation described above does not seem to favor soil formation processes to take place in a constant manner, it is surprising that while two out of the three soils on this geological map unit belong to the class Terrestrische Humusböden, the majority of the soil profile sites specified as having ISR or SSR as parent material by the soil surveyors are of the classes Braunerde or Kalklehme, i.e. soils with A- and B-horizons. More expected is the presence of the soil type Kolluvisol on the SSR units (both on the map and regarding field survey). In summary, it appears that in such a dynamic environment as on the SGUs ISR, SSR and slope debris, soil surveyors tend to annotate slope debris as the parent material of less developed soils such as those of the Terrestrische Humusböden class, while placing more emphasis on the the mineralogical properties of the components when describing soils that experienced more pedogenetic processes. Nevertheless, this analysis implies that soils developed from any of these intermittent layers are more likely to be encountered in down-slope slope debris units, especially for ISR. Regardless of the present dynamics, more evolved soils are possible but may present very heterogeneous parent material due to multi-layering. Figure 5D shows an example of such a profile site in steep topography with a typical slope debris parent material mixture containing ISR material, which has however been deposited on a SB unit located down-slope of these layers.

4.2.8. Colluvial, mixed, and landslide deposits

All three of these SGUs are characterized as material that has been transported down-slope by various processes. Regarding colluvial deposits, the surveyors indeed identified this parent material unit for three of the five profile sites on this map unit. The parent material of the remaining two sites was mapped as ISR and SD. While the latter can be attributed to the similarity of the material and the uncertainty of the definition of the difference between slope debris and colluvial deposits, the former could indicate that soil surveyors may value the information regarding mineralogy as being of greater importance to explaining a soil profile than the information regarding the morphodynamic history of the material. The definition of mixed deposits as deposits from debris flows, torrents and avalanches is similarly dominated by the means of transportation rather than the material components. Interestingly, none of the soil profiles located on this map unit were linked to this parent material by the surveyors, but five of the profiles on till were. With regard to soil parent material, an important issue is whether this material is in fact composed of the same material as till and has been simply transported by one of the aforementioned processes. The parent material of the majority of the soil profiles on the MxD unit of the geological map was identified as slope debris, highlighting the definitional proximity of these units, at least concerning the view of surveyors focusing on soil. Compared to these two units, the issue regarding landslide deposits is a different one. This unit was never used by the soil surveyors to describe the parent material of a soil profile site. Of the three data points on this map unit, the surveyors explicitly chose the parent material siliceous bedrock for two locations, which is in fact the material of the landslide body in the study area. Coarse blocky debris was also used at one soil profile site, which well describes the structure of the material involved. This example provides insight into how the different emphasis and focus of soil and geologic surveyors influences their classification, each highlighting a different aspect of the same unit.

Topographically, these three units are not easy to differentiate, especially compared to the unit slope debris, which is very close with regard to both definition as well as morphometry and discussed in the slope debris subsection.

The soil types of the three soil profile sites with parent material which both geological map and surveyors identify as colluvial deposits, are of the classes Umgelagerte Böden and Brown soils and are representative of both the map unit and the soils as attributed by the surveyors, although the latter also link this parent material unit to a less developed soil type like Ranker. In general, the colluvial deposits are characterized, or understood, as the accumulation of eroded slope material, which can also be A-horizon-material, for instance transported away from agricultural lands. Consequently, disturbed soils but also better developed soils on older accumulations are conceivable examples of soils on colluvial deposits. Given the similar genesis of mixed deposits, the soil classes distributed amongst the soil profiles with this parent material are comparable, despite some of the points being located on TG according to the geologic map. But as this parent material can actually consist of till or calcareous sediments transported from up-slope locations, the soil types Braunerde and Kalklehme are good fits and show that confusion between the units does not lead to severe misclassifications, at least from a soil surveyor's point of view. As discussed in the topography section, the landslide mass in the study area is composed of siliceous bedrock, with Ranker and Braunerde being the anticipated soil types for soil developed from such parent material.

4.2.9. Glacio-lacustrine deposits, ice-marginal sediments and mire deposits

Even when taking into account that only seven soil profile sites are located on this map unit, the geological map has a relatively good user's accuracy for the soil parent material unit GLD. It comprises both glaciolacustrine and lacustrine deposits which are concentrated at the northern border of the paleovalley. The producer's reliability is below 50 %, as surveyors indicate this parent material also for soil profile sites on the SGUs alluvial deposits, mire deposits, siliceous bedrock and till. Confusion with the first two SGUs seems acceptable due to the common fluviatile or hydromorphic history, whereas the other misclassifications do indicate the possibility of local glaciolacustrine deposits as the parent material on these units. The topographical analysis of those points on other SGUs but with GLD as the surveyed parent material hints that these are located in regions where landform diversity is partly dominated by a single class at meso scale, indicating a smoother surface than for the GLD map unit, as well as being characterized by a more pronounced negative minimal curvature suggesting concave topography. The random forest-based investigation into the topographical difference between soil profile points with SB and with GLD as parent material characterize the latter with lower roughness values regarding TRI and VRM at window sizes between 100-130 m, indicating similar regions on SB as possible GLD areas. Regarding soil type, investigations from both viewpoints, i.e. the geological map and the soil profile data, show the soils on this unit as belonging to either the soil type Braunerde or Rigolboden, the latter representing agriculturally used land. Due to the lack of profile points in mire deposits, GLD is the only SGU on which surveyors mapped a hydromorphic soil type, in this case Pseudogley.

Ice-marginal sediments are closely associated with till units and are mostly located in close proximity, consequently sharing long borders. Given their often thin, elongated form and the fact that they cover only 0.2% of the study area, only one soil profile site is located on this SGU on the geologic map, of which the parent material however is classified as alluvial deposits. This situation again highlights the problem with fuzzy definitions that may overlap between closely related parent material units. Similarly, it may not be as important to the soil surveyors that the gravelly parent material is of ice-marginal origin rather than alluvial, whereas this tends to be the case for geologic surveyors specialized on quaternary deposits.

Similarly, on the map SGU mire deposits the surveyors identified the soil parent material glacio-lacustrine deposits twice and till once, with the only soil profile with mire deposits as parent material being an anthrosol (filled mire deposit) located on the SGU till. The influence of stagnant water on both GLD and MrD can help understand the confusion between both units, and local mire deposits due to reduced permeability on clay-rich sub-glacial till are also rather likely. Given these local phenomena as well as the small number of sample profile sites, the topographical separation of the units IMS and MrD provides little relevant information, neither does an analysis of soil type distribution.

4.3. Random forest classification of parent material based on geologic map data and topography

The result of the feature selection procedure for predicting parent material classes was that, without exception, the SGU information from the geologic map was chosen first in every single feature selection run. This is important as it also contains information on the confusion between certain classes implicitly in the sample point data. Contrary to when investigating which terrain parameters best highlight the differences between two specific parent material units, local and regional terrain parameters including relative elevations, were not identified as the most important predictor variables. Instead, the roughness-related measures dominated the selection process, even with the parameter set including all available terrain parameters. In this case, the SGU information as well as the TRI at a window size of 77.5 m and the VRM at a window size of 142.5 m were selected to produce a model of soil parent material. The addition of further explanatory variables was not seen as an improvement with regard to both the model performance as well as its interpretability after evaluating some possible expanded models both with visual inspection of the results as well as the one-standard-error rule (James et al., 2013). The reason for the decreased importance of regional, but also local, terrain parameters is that this information is already implicitly accounted for by the geologic map, as the various SGU classes are closely linked to certain relative elevations, especially the vertical distance to channel base level, due to the general geologic structure of the study area.

Additionally, at least to a certain degree, some units like till, slope debris or alluvial deposits also provide implicit information regarding slope angles. Furthermore, a certain amount of slope information is also included in some roughness measures, especially the TRI, which increases with slope as well as with increasing roughness. The vector ruggedness measure however well compliments the TRI and the SGU information as it provides information on surface roughness independent of the slope gradient. Figure 7 shows the original SGU map (A) as well as the parent material map based on SGUs, TRI and VRM (B). This modeled parent material map leads to an overall

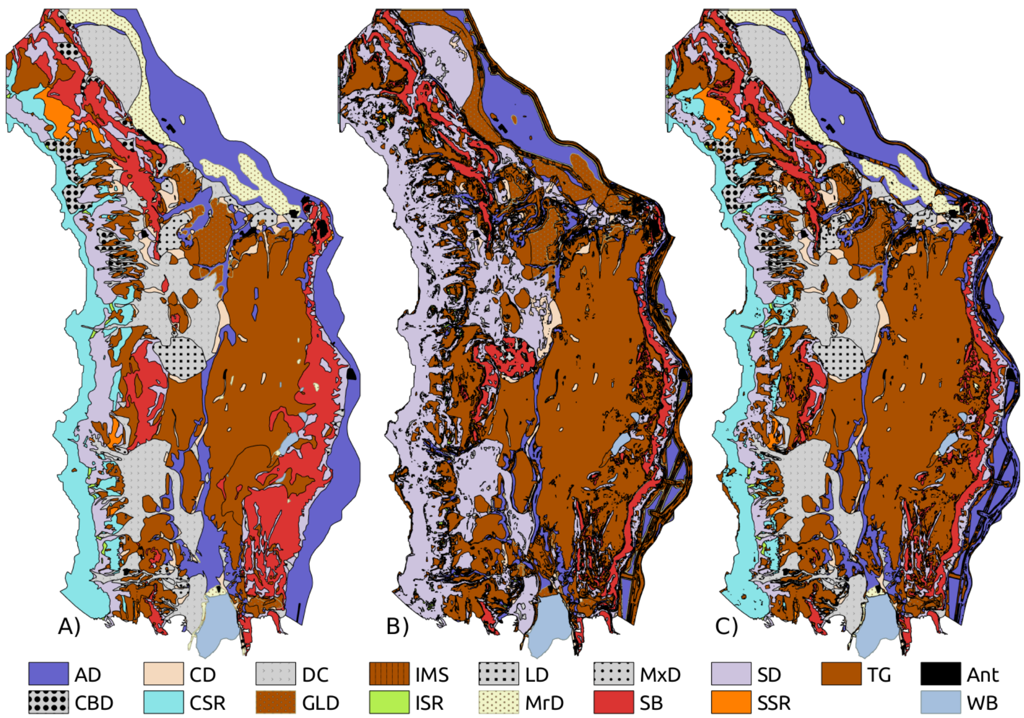

Figure 7: A) Original reclassified SGUs. B) Random forest-based parent material map modified based on VRM at a window size of 142.5 m and TRI at a window size of 77.5 m. C) parent material map with only till modeled and overlayed onto the original SGU map. Abbreviations: AD = alluvial deposits, CBD = coarse blocky debris, CD = colluvial deposits, CSR = calcareous sedimentary rock, DC = debris cones, GLD = glaciolacustrine and lacustrine deposits, IMS = ice-marginal deposits, ISR = intermediate sedimentary rock, LD = landslide deposits, MrD = mire deposits, MxD = mixed deposits, SB = silceous bedrock, SD = slope debris, SSR = siliceous sedimentary rock, TG = till in general

accuracy of approx. 90 %, however the out-of-bag correct classification rate amounts to only 65 %, implying quite substantial over-fitting, especially for the parent material classes with small sample sizes. Consequently, the predicted map must be evaluated with caution. A comparison of the resulting parent material map, adjusted only with two roughness measures, with the original SGU map, reveals a considerable areal increase of the TG unit. This is perfectly understandable given the low producer's reliability revealed in Table 3 by the high number of soil profile sites located on different SGUs but with TG as the surveyed parent material class. The most obvious errors linked to this unit are the highways in the Etsch valley as well as the mire deposits surrounding the large debris cone in the north of the study area, which are modeled as TG. Similarly conspicuous is the fact that the areas originally covered by the SGUs debris cones and mixed deposits now seem to be covered by the slope debris unit, which seems to contradict the good prediction results of these units. Detailed investigation into this issue shows that over-fitting occurred in these cases, with very small, local DC and MxD polygons leading to the false impression of correct classification. On the other hand, the severity of these misleading results may appear less grave when considering the similarity of these three units. Given that they all represent material that has been transported down-slope by various processes and that their composition may in fact be the same in many cases, this misclassification may seem acceptable if one keeps in mind the information provided by the original map with regard to the geomorphological history of these units. Small units like CSR, ISR and SSR are taken over by slope debris and/or till in a similar manner. An interesting unit in the predicted map is the former landslide deposit in the center of the study area. As the surveyors put more emphasis on describing the deposited material itself than its geomorphological history, the unit is now modeled as siliceous bedrock with coarse blocky debris, which well represents the view of the soil surveyors. Another insightful aspect of the predicted parent materials are new, additional polygons which suggest small glaciolacustrine deposits inside the largest till unit east of the Mitterberg. Based on the experience from previous field work in this area, the authors deem these predictions worthy of future investigation.

After considering and evaluating the issues that arise with regard to the individual SGUs as well as from modeling parent material for the study area, it was concluded that while the modeled parent material map is helpful in giving insights into issues that arise when using geologic maps and topography as a proxy for parent material, too much of the information provided by the original SGU map,i.e. the detailed geologic map, is lost. The under-representation of the class till, however, which had been identified as the main single issue when applying the geologic map for soil survey, seems to have been dealt with appropriately by the random forest classifier. Hence, a new approach was chosen by modeling only the till class and then overlaying this information onto the original SGU map. The forward step-wise feature selection procedure for the units till vs. not-till again chose one TRI and one VRM map as model input features, but with slightly different window sizes (100 m and 102.5 m, respectively) than for the parent material model. This approach (Figure 7C) combines the model's good performance with regard to modeling the often thin till layer, while retaining most of the information from the geologic map, which now, keeping in mind the geologic-topographic characterization, can help in understanding the composition of slope debris and other transported materials. Compared to the original confusion matrix (Table 3), the overall correct classification rate is improved form 49 % to 65 % with this approach.

Considering the process of feature selection for the various applications of random forest classification in this study, it is important to have in mind that a large number of different versions were computed for most of the terrain parameters and the classifier was allowed to choose freely amongst them. As the change in computational window size is often not large enough to lead to significantly different results, the slightly different versions of one terrain parameter will be highly correlated. Similarly, landform diversity measures such as edge-density or mean patch size are also closely correlated and represent similar patterns in the data. Consequently, it is not possible to state that one parameter combination is better than another for separating parent material classes, only that the specific feature selection procedure chose this combination over others. The differences in overall accuracy between the best parameters are often minimal. However, the application of five-fold cross-validation to the entire feature selection procedure sheds light onto how constant the choice of the top ranked parameters is. When comparing the features selected by the forward stepwise selection and those with the highest importance measures 'mean decrease accuracy', the results were mostly consistent with regard to the types of parameters chosen, for instance a version of TRI or profile curvature. The main difference is that while the forward-stepwise feature selection procedure chose two different roughness measures (TRI and VRM) to go with the geologic information for modeling soil parent material, the top ranked parameters based on 'mean decrease accuracy' were often the same parameter (e.g. TRI) but at slightly different window sizes. In our view this demonstrates the advantage of the forward stepwise feature selection procedure, as opposed to the ranking based on importance measures.

When analyzing the results of the soil parent material model or box plots such as in Figure 4, the different class sizes of the soil pit or auger descriptions must be considered. As this study is based on legacy soil data already available, a sampling scheme could not be applied. The over-representation of large classes like till or slope debris are however not as great of a problem as the small sample size of classes such as mire deposits or siliceous sedimentary rock, as for instance Congalton (1991) do in fact advise over-sampling for classes with high variability. For future studies, a sampling scheme that leads to a more balanced distribution of the samples per parent material unit should however be applied in order to decrease the problem of over-fitting classes with small sample sizes and high confusion rates.

4.4. Mapping approaches and soil parent material classification

The geologic-topographic characterization of surficial geology units and the application of random forest classification to model soil parent material gives insight into a number of issues which arise when using geologic maps for soil parent material information and require more detailed discussion. While the units till and landslide deposits highlight the differences in mapping approaches between soil science and geology, the slope debris unit points out the difficulties of classifying soil parent material units due to fuzzy transitions between classes. This unit is also exemplary of problems that arise specifically in an Alpine environment and its high morphodynamics.

Between the two different frameworks of mapping, geology on the one hand and soil on the other, it is important to acknowledge the main focus of attention of each branch of research. Miller and Lee Burras (2015) note that the resulting maps of the two sciences try to communicate different aspects. While geology refers to geologic materials and general landform regions, soil science is concerned more with soil properties with regard to land use and management decisions. A typical example is the unit landslide debris, which seems only of interest to the mappers with focus on geology, whereas the soil surveyors pay attention to the chemical and physical properties of the material. This leads to the latter group classifying the parent material of soil profile sites on this unit as either siliceous bedrock (emphasis on mineralogical properties) or coarse blocky debris (emphasis on physical properties such as as grain size distribution). In the presented study, we identify the till unit as a major challenge and reason for misclassification when using geologic maps as soil parent material maps. While the geologists map a number of different glacial deposit units, soil or forestry surveyors seldom differentiated between lodgement till and other glacial deposits, only when a decisively higher clay content and bulk density was determined. On the other hand, the chemical properties, especially with regard to acidic or calcareous properties of the till material are of greater interest to the soil surveyors, who further differentiate till with regard to this aspect. The geologic map on the other hand also contains information with regard to the age and the stratigraphic unit to which certain moraines belong, whereas this is only of interest to the soil surveyor if it features additional information with regard to the mineralogical content of the different constituents of the moraine. The rather low producer's reliability of the till unit leads to an additional issue linked to different mapping purposes or approaches, as there may exist a difference with regard to how pronounced a certain feature or characteristic must be in order to be considered for mapping. While a soil surveyor does note underlying layers in his profile description, the main focus is on the material from which the soil developed and which consequently plays an important role in characterizing it and its location. The geologist on the other hand may not place as much emphasis on the thin layer that actually acts as the soil parent material, but more on the thick layers of different origin beneath the solum, as its role in describing the geological setting is much greater. The case of a thin till layer on siliceous bedrock, as illustrated in Figure 5A, exemplifies such a situation. Consultations with the geologic mappers showed that for their purpose, only regolith cover layers thicker than approximately 1 m are usually mapped. This is an important insight that must always be considered when using geologic maps to infer information regarding soil parent material, especially in areas like the study area characterized by younger soils. It must however also be kept in mind that natural differences can be expected between any two mappers (Miller and Lee Burras, 2015), even of the same scientific field.

As Olofsson et al. (2013) point out with regard to remote sensing applications, uncertainty can be related to geolocation and interpreter variability. This statement is also applicable regarding the uncertainties involved in the comparison of parent material information from field data on the one hand and geologic maps on the other. While geolocation errors may play a role in densely forested slopes as found in the study error, the fuzzy transition from one surficial geology unit to another are the main cause for confusion at categorical transitions. This could even be aggravated by the fact that multiple interpreters were part of the various surveys analyzed in this study, though they all adhered to the same survey manual Englisch and Kilian (1999). As Congalton (1991) state regarding accuracy assessments, classifications should be mutually exclusive, but regarding soil parent material this is not always as clear as expected. The issue of slope debris and the bedrock units from which it originates is an example for a case where both classes can be in fact considered correct, especially in border areas between the units. Consequently, for future surveys it would be helpful to follow the guidelines of Olofsson et al. (2013) with regard to reference labeling. It would be beneficial to note difficult cases regarding the classification of soil parent material for future reference and consensus development, not only within soil science but also as an effort for improved communication with geologists.

The question of what is more relevant from a pedological point of view, the transportational history and morphodynamic background of deposits or the information regarding its composition, addresses both the differences in mapping purposes as well as nomenclatural difficulties. In the study area, mixed deposits from mass movements and torrents often have the same components as till or slope debris, which themselves may possibly be composed of the same material due to a mutual source area. Consequently, one could argue that a single class for this material, which is characterized by downslope transportation, may suffice. However, while there may exist situations were the slope debris is composed of till material, the general difference for soil science is that slope debris contains coarser components whereas with till the proportion of fine particles is much higher. While this leads to a lower density of the slope debris material, till is often also characterized by over-consolidation. From a pedological point of view, another very important difference is that soils from slope debris can be understood with knowledge of local geology and that they may not be so different from soils that evolved from bedrock at higher elevations in the catchment, whereas till can consist of chemically very different components. So while it seems acceptable, as performed by the soil parent material model described in this study, to include mixed deposits and slope debris in one class, the differentiation between till and slope debris units is essential for the understanding of the spatial association of soils. An important result of the presented study is therefore, at least in this study area, that till can be found at steeper slopes than expected presumably due to lateral consolidation of till material by the glaciers, and also as a thin cover layer on bedrock units, especially the one composed of rhyolitic tuffs and ignimbrites.

A further issue that arises specifically in this study area is the influence of the Alpine environment on the interpretability of geologic units as parent material units. Heung et al. (2014) note that while traditional geologic maps focusing on bed rock are a valuable input for digital soil mapping when the residual materials form the soil parent material, but less so in areas distinguished by glaciation and high geomorphodynamics. Similarly, in their comparison of surficial geology maps derived from Soil Survey maps on the one hand and the Geologic Survey on the other, Miller and Lee Burras (2015) point out that the level of agreement was lower for areas with complicated geologic histories. In the presented study area, the issue of increased morphodynamics in high relief areas and, as a consequence, multi-layered soils, is best demonstrated by the intermediate and siliceous sedimentary rock units. These units are often composed of soft and rather thin layers. As a consequence they are easily eroded and covered by slope debris from up-slope bedrock or slope debris units, and therefore under-represented in the soil pit data. Nevertheless, components from the ISR and SSR units are possible in multi-layered soil profiles down-slope of these units (Figure 5D).

5. Conclusion

The comparison of soil parent material as noted by soil surveyors in their description of soil pits with the surficial geology units of new, detailed geological maps, shows that while there is a confusion rate of over 50 %, these maps are immensely helpful in understanding the distribution of soil parent material. The main issue encountered was the abundance of till acting as soil parent material on other, different surficial geology units. On the one hand, a often very thin layer of till was found to be the soil parent material of many soils on the unit siliceous bedrock, and on the other hand it was encountered at steeper slope angles and at higher elevations than expected. The presented study study shows that roughness measures, specifically a combination of the topographic roughness index and vector ruggedness measures, present a tool to further differentiate surficial geology units, highlighting regions were the probability of encountering till as soil parent material is high and thereby characterizing the variability of these units. In this study we also show that the fuzziness of the categorical transitions between two parent material classes, for instance between slope debris and any bedrock class which acts as the slope debris‘ source area, represents a further cause for misclassification. By incorporating the soil types found on the various units into the analyses, the presented approach attempts to point out which misclassifications of soil parent material lead to incorrect assumptions regarding soil distribution, but also emphasizes those cases where misclassification does not negatively affect the derivation of soil units.

The geologic-topographic characterization of the surficial geology units, representing a synthesis of geologic, topographic and soil-related information, is considered to be an appropriate approach to make best use of available geologic information. It sheds light onto how the mapping approaches between geologists and soil scientists vary, and at the same time provides a possible link between the two sciences. By performing such an analysis prior to future detailed field soil surveys, the surveyor can make best use of available information and concentrate the time and money consuming task of field work, involving soil pits and auguring, on units identified as highly variable and uncertain regarding soil and its parent material. This will strengthen the understanding of the pedologic relevance of these units and prove helpful for devising future sampling procedures and the regionalisation of point information. While the presented approach incorporating soil type distribution into the geological-topographic characterization of geological units certainly helps understand the pedological influence of these, future studies may want to include specific soil properties, where available, such as bulk density and hydrological properties, into the characterizations.

The authors are of the opinion that the new generation of geological maps with their detailed description of quaternary deposits are an important tool for future soil survey and signify a step towards closing the gap between geology and soil science.

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

Amt der Tiroler Landesregierung, Abteilung Forstplanung, unpublished. Waldtypisierung Tirol - System zur Substratansprache.

APB, 2006. Waldtypisierung Südtirol. Autonome Provinz Bozen - Südtirol. URL: http://www.provinz.bz.it/forst/studien-projekte/waldtypisierung.asp.

APB, 2016. Download landeskartographie - autonomous province bolzano, south tyrol. URL: http://www.provinz.bz.it/natur-raum/themen/landeskartografie-download.asp.

Avanzini, M., Bargossi, G., Borsato, A., Castiglioni, G., Cucato, M., Morelli, C., Prosser, G., Sapelza, A., 2006. Erläuterungen zur geologischen Karte von Italien im Maßstab 1:50000 Blatt 026 Eppan. Technical Report. Agenzia per la protezione dell'ambiente e per i servizi tecnici.

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, 312 – 331. URL: http://www.sciencedirect.com/science/article/pii/S0016706115300343, doi:http://dx.doi.org/10.1016/j.geoderma.2015.08.005. soil mapping, classification, and modelling: history and future directions.

Beasom, S.L., Wiggers, E.P., Giardino, J.R., 1983. A technique for assessing land surface ruggedness. The Journal of Wildlife Management 47, 1163–1166. doi:10.2307/3808184.

Böhner, J., Antonić, O., 2009. Chapter 8 land-surface parameters specific to topo-climatology, in: Hengl, T., Reuter, H.I. (Eds.), Geomorphometry Concepts, Software, Applications. Elsevier. volume 33 of *Developments in Soil Science*, pp. 195 – 226. URL: http://www.sciencedirect.com/science/article/pii/S0166248108000081, doi:http://dx.doi.org/10.1016/S0166-2481(08)00008-1.

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., Miller, B.A., 2015. The use of soil surveys to aid in geologic mapping with an emphasis on the eastern and midwestern united states.

Soil Horizons 56. URL: http://dx.doi.org/10.2136/sh15-01-0001, doi:10.2136/sh15-01-0001.

Coblentz, D., Pabian, F., Prasad, L., 2014. Quantitative geomorphometrics for terrain characterization. International Journal of Geosciences 5, 247–266. URL: http://dx.doi.org/10.4236/ijg.2014.53026.

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: https://www.geosci-model-dev.net/8/1991/2015/, doi:10.5194/gmd-8-1991-2015.

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.

Fracek, K., Mosimann, T., 2013. Wissensbasierte Modellierung der Mächtigkeit des kalkfreien Bodenbereiches in den Waldböden des Kantons Basel-Landschaft (Nordwestschweiz). Waldökologie, Landschaftsforschung und Naturschutz 13, 5–16.

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 Alps - Challenges 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.

GRASS Development Team, 2016. Geographic Resources Analysis Support System (GRASS GIS) Software, Version 7.0. Open Source Geospatial Foundation. URL: http://grass.osgeo.org.

Grohmann, C.H., Smith, M.J., Riccomini, C., 2010. Multiscale analysis of topographic surface roughness in the midland valley, scotland. IEEE Transactions on Geoscience and Remote Sensing PP, 1 – 14. doi:10.1109/TGRS.2010.2053546.

Gruber, F.E., Baruck, J., Geitner, C., 2017. Algorithms vs. surveyors: A comparison of automated landform delineations and surveyed topographic positions from soil mapping in an alpine environment. Geoderma 308, 9 – 25. URL: http://www.sciencedirect.com/science/article/pii/S001670611730664X, doi:http://dx.doi.org/10.1016/j.geoderma.2017.08.017.

Heung, B., Bulmer, C.E., Schmidt, M.G., 2014. Predictive soil parent material mapping at a regional-scale: A random forest approach. Geoderma 214215, 141 – 154. URL: http://www.sciencedirect.com/science/article/pii/S0016706113003443, doi:http://dx.doi.org/10.1016/j.geoderma.2013.09.016.

Heung, B., Ho, H.C., Zhang, J., Knudby, A., Bulmer, C.E., Schmidt, M.G., 2016. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. Geoderma 265, 62 – 77. URL: http://www.sciencedirect.com/science/article/pii/S0016706115301300, doi:https://doi.org/10.1016/j.geoderma.2015.11.014.

Hobson, R.D., 1972. Spatial analysis in geomorphology. Harper and Row, New York. chapter Surface roughness in topography: quantitative approach. pp. 221–245.

IUSS Working Group WRB, 2015. World Reference Base for Soil Resources 2014, update 2015 International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports 106. FAO, Rome. Rome.

Iwahashi, J., Pike, R.J., 2007. Automated classifications of topography 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 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., Stepinski, T.F., 2013. Geomorphons a pattern recognition 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.

Jenny, H., 1941. Factors of Soil Formation - A System of Quantitative Pedology. Dover Publications, Inc. New York.

Juilleret, J., Dondeyne, S., Vancampenhout, K., Deckers, J., Hissler, C., 2016. Mind the gap: A classification system for integrating the subsolum into soil surveys. Geoderma 264, Part B, 332 – 339. URL: http://www.sciencedirect.com/science/article/pii/S0016706115300604, doi:http://dx.doi.org/10.1016/j.geoderma.2015.08.031.

Juilleret, J., Iffly, J.F., Hoffmann, L., Hissler, C., 2012. The potential of soil survey as a tool for surface geological mapping: a case study in a hydrological experimental catchment (huewelerbach, grand-duchy of luxembourg). Geologica Belgica 15. URL: http://popups.ulg.ac.be/1374-8505/index.php?id=3496.

Kilian, W., 2015. Schlüssel zur Bestimmung der Böden Österreichs. volume 81. 2 ed., Österr. Bodenkundl. Ges.

McBratney, A.B., Mendonca Santos, M., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3 – 52.

Miller, B.A., Lee Burras, C., 2015. Comparison of surficial geology maps based on soil survey and in depth geological survey. Soil Horizons 56. URL: http://dx.doi.org/10.2136/sh14-05-0005, doi:10.2136/sh14-05-0005.

Nellemann, C., Thomsen, M.G., 1994. Terrain ruggedness and caribou forage availability during snowmelt on the arctic coastal plain, alaska. Arctic 47, 361–367. URL: `http://www.jstor.org/stable/40511597`.

Nestroy, O., Aust, G., Blum, W., Englisch, M., Hager, H., Herzberger, E., 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. Bodenkdl. Ges. 79.

Nestroy, O., Danneberg, O., Englisch, M., Geßl, A., Hager, H., Herzberger, E., Kilian, W., Nelhiebel, P., Pecina, E., Pehamberger, A., Schneider, W., Wagner, J., 2000. Systematische Gliederung der Böden Österreichischs (Österreichische Bodensystematik 2000). Mitt. Österr. Bodenkdl. Ges. 60.

Olaya, V., 2009. Chapter 6 basic land-surface parameters, in: Hengl, T., Reuter, H.I. (Eds.), Geomorphometry Concepts, Software, Applications. Elsevier. volume 33 of *Developments in Soil Science*, pp. 141 – 169. URL: `http://www.sciencedirect.com/science/article/pii/S0166248108000068`, doi:`http://dx.doi.org/10.1016/S0166-2481(08)00006-8`.

Olofsson, P., Foody, G.M., Stehman, S.V., Woodcock, C.E., 2013. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sensing of Environment 129, 122 – 131. URL: `http://www.sciencedirect.com/science/article/pii/S0034425712004191`, doi:`https://doi.org/10.1016/j.rse.2012.10.031`.

Phillips, J.D., Lorz, C., 2008. Origins and implications of soil layering. Earth-Science Reviews 89, 144 – 155. URL: `http://www.sciencedirect.com/science/article/pii/S001282520800055X`, doi:`https://doi.org/10.1016/j.earscirev.2008.04.003`.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: `http://www.R-project.org/`.

Riley, S.J., DeGloria, S.D., Elliot, R., 1999. A terrain ruggedness index that quantifies topographic heterogeneity. Intermountain Journal of Sciences 5, 23 – 27.

Sappington, J.M., Longshore, K.M., Thompson, D.B., 2007. Quantifying landscape ruggedness for animal habitat analysis: A case study using bighorn sheep in the mojave desert. Journal of Wildlife Management 75. URL: http://dx.doi.org/10.2193/2005-723, doi:10.2193/2005-723.

Schaetzl, R.J., Krist Jr., F.J., Rindfleisch, P.R., Liebens, J., Williams, T.E., 2000. Postglacial landscape evolution of northeastern lower michigan, interpreted from soils and sediments. Annals of the Association of American Geographers 90, 443–466. URL: http://dx.doi.org/10.1111/0004-5608.00204, doi:10.1111/0004-5608.00204.

Scholz, H., Bestle, K.H., Willerich, S., 2005. Quartärgeologische Untersuchungen im Überetsch. Geo.Alp 2, 1–23.

Thalheimer, M., 2006. Kartierung der landwirtschaftlich genutzten Böden des Überetsch in Südtirol (Italien). Laimburg Journal 3, 135–177. URL: http://www.laimburg.it/de/projekte-publikationen/blickpunkte.asp?news_action=300&news_image_id=828290.

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".

Wood, J.D., 1996. The geomorphological characterisation of digital elevation models. Ph.D. thesis. University of Leicester, UK.

Woodcock, N.H., 1977. Specification of fabric shapes using an eigenvalue method. Geological Society of America Bulletin 88, 1231–1236.