A Fuzzy Logic Based Method for Modeling the Spatial Distribution of Indicators of Decomposition in a High Mountain Environment

Authors: Hellwig, Niels, Anschlag, Kerstin, and Broll, Gabriele

Source: Arctic, Antarctic, and Alpine Research, 48(4) : 623-635

Published By: Institute of Arctic and Alpine Research (INSTAAR), University of Colorado

URL: https://doi.org/10.1657/AAAR0015-073

Arctic, Antarctic, and Alpine Research, Vol. 48, No. 4, 2016, pp. 623–635
DOI: http://dx.doi.org/10.1657/AAAR0015-073

A fuzzy logic based method for modeling the spatial distribution of indicators of decomposition in a high mountain environment

Niels Hellwig[1,*], Kerstin Anschlag[1], and Gabriele Broll[1]

[1]Institute of Geography, University of Osnabrück, Seminarstraße 19ab, 49074 Osnabrück, Germany
[*]Corresponding author's email: niels.hellwig@uni-osnabrueck.de

ABSTRACT

Upscaling of sample data on indicators of decomposition to the landscape scale is often necessary for extensive ecological assessments. The amount of such data is mostly scarce even with high sampling efforts. Moreover, environmental conditions are very heterogeneous in high mountain regions. Therefore, the aim was to find a suitable technique for spatial modeling under these circumstances.

A method combining decision tree analysis and the construction of fuzzy membership functions is introduced for a GIS-based mapping of decomposition indicating parameters. It is compared with an approach solely based on decision trees. Within a case study in the Italian Alps the spatial distribution of humus forms, classified by the occurrence of an OH (humified residues) horizon, is examined. There appears to be a strong relationship with elevation and a minor correlation with slope exposition.

The fuzzy logic-based approach proves to be suitable for modeling the spatial distribution of indicators of decomposition. Mapping fuzzy values allows for the representation of small-scale variability and uncertainty of data due to a relatively low sample size in a very heterogeneous environment.

Introduction

Decomposition processes are of high significance for the functioning of terrestrial ecosystems. As part of various material cycles, these processes ensure the survival not only of decomposing but also of producing and consuming organisms (Swift et al., 1979). A prominent indicator with regard to decomposition are humus forms (Andreetta et al., 2012; Ascher et al., 2012; Graefe and Beylich, 2006; Ponge, 2013). They can be defined as manifestations of dead organic matter at different stages of decomposition in the topsoil, which in forest ecosystems consist of organic layers (OL = litter, OF = fragmented residues, OH = humified residues) and the uppermost mineral horizon.

Analyzing and assessing the impacts of ecological processes and interactions are required at the landscape scale for numerous purposes. In contrast, ecological field data at the landscape scale is often scarce due to high costs and low accessibility, especially in high mountain environments. Upscaling by means of spatial modeling allows for bridging the gap between the local study scale and the target landscape scale.

In terms of soil ecology, there is a wide range of such modeling methods associated with the concept of digital soil mapping (McBratney et al., 2003), comprising pedotransfer functions, geostatistical techniques, and factor-based approaches (Behrens and Scholten, 2006). Pedotransfer functions allow the derivation of soil variables from other factors easier to determine by means of mathematical formulas (Bouma, 1989; Wösten et al., 2001); thus a detailed quantitative comprehension of the correlations between environmental factors is prerequisite

(McBratney et al., 2002). The application of geostatistical techniques (e.g., kriging, co-kriging) is particularly critical in areas with a high heterogeneity of environmental covariates. For these techniques an accordingly higher density of samples is indispensable (Heuvelink and Webster, 2001). Factor-based approaches are based on considering the soil properties as a system state, whose configuration is determined by the soil-forming factors (e.g., clorpt model, Jenny, 1941; scorpan model, McBratney et al., 2003). Different methods have been applied implementing the factor-based approach (Behrens and Scholten, 2007; McBratney et al., 2003), including linear regression and classification models, artificial neural networks, tree-based regression and classification models, support vector machines, and fuzzy logic models.

This study aims at refining a spatial knowledge-based modeling technique and establishing it for the prediction of indicators of decomposition processes and properties under a highly heterogeneous topography and a relatively small sample size. Decomposition processes are influenced by various environmental factors. In a high mountain environment these are in a large part mediated by the elevation and the slope exposition, but for some factors in a nonlinear way (such as vegetation, where thresholds for different zones exist depending on the topography). Therefore a fuzzy logic approach based on a data mining decision tree algorithm accounting for nonlinearities is hypothesized to fit the situation.

The first part describes the methodological approach proposed in this study. It is followed by the presentation of a case study conducted in a study area in the Italian Alps. In this case study, the approach utilizing fuzzy logic is applied for modeling the occurrence of humus forms showing an OH horizon and comparing it with an approach solely utilizing decision tree analysis.

Methodological Framework

Construction of a Decision Tree

Binary decision trees for data mining (i.e., classification and regression trees [CART]; Breiman et al., 1984) serve as a practicable and simply interpretable tool to statistically model complex and nonlinear dependencies between (environmental) influencing factors and a target variable on the basis of sample data (Aberegg et al., 2009; De'ath and Fabricius, 2000; McKenzie and Ryan, 1999; Mertens et al., 2002). Decision trees are constructed by recursively partitioning the sample set into pairwise disjoint subsets that show a higher rate of homogeneity with respect to the target variable. The rules for partitioning have the form $x_i \leq c, c \in V_i \subseteq \mathbb{R}$, if the influencing factor $X_i \in \{X_1, \ldots, X_a\}$ taking values in V_i is interval or ratio scaled, and the form $x_i \in S$ otherwise, with S covering a subset of the property values $M_i = \{m_1, \ldots, m_z\}$ of X_i (x_i is evaluated for every sample as the value of X_i) (Fig. 1). As part of the partitioning procedure, an inhomogeneity measure ι is calculated in order to establish an appropriate decision rule. Each time, the difference of the inhomogeneity of

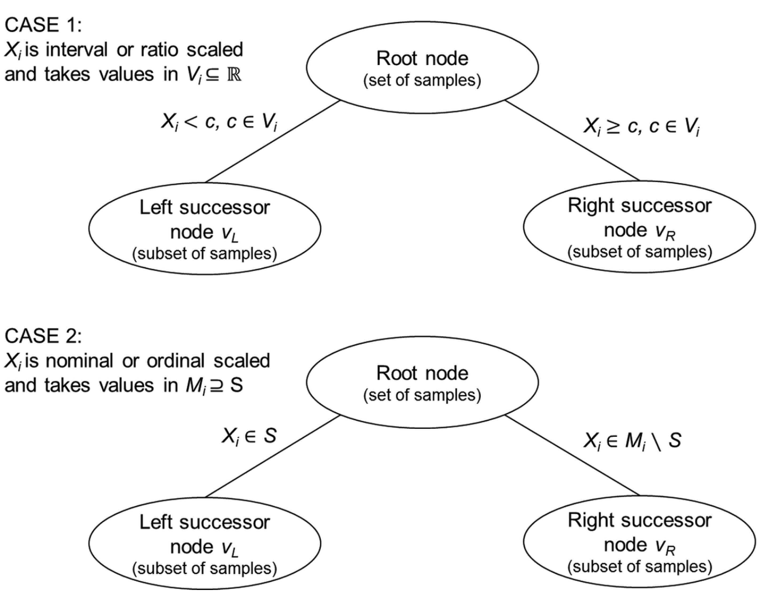

FIGURE 1. Illustration of a node split in a decision tree depending on the scale of measurement of the influencing variable $X_i \in \{X_1, \ldots, X_a\}$.

a certain node v and the sum of the inhomogeneities of the successor nodes v_L and v_R are maximized: $\max\{\iota[v] - [\iota(v_L) + \iota(v_R)]\}$. For classification trees the gini index ι_{Class} (Equation 1) serves as an estimator of the inhomogeneity of a node v, with the conditional probability $p(y_i|v)$, $i = 1, \ldots, z$, of a value of the target variable $y_i \in Y$ in v. For regression trees the resubstitution error ι_{Reg} (Equation 2) is used, where the number of all samples is denoted as N and the number of samples in v is denoted as $N_v \leq N$, with the related values of the target variable y_n, $n = 1, \ldots, N_v$ (Breiman et al., 1984).

$$\iota_{Class}(v) = \sum\nolimits_{k \neq \iota} p(y_k \mid v) \cdot p(y_\iota \mid v), y_k, y_\iota \in Y \qquad (1)$$

$$\iota_{Reg}(v) = (1/N) \sum\nolimits_{n=1}^{N_v} (y_n - \overline{y}(v))^2 \qquad (2)$$

The recursive partitioning procedure is performed on all subsets. It terminates as soon as no higher degree of homogeneity can be reached or a threshold of the number of elements in a node is undercut.

Fuzzy Logic Model

The fuzzy logic model is based on the concept of fuzzy sets (Zadeh, 1965). Unlike ordinary sets, fuzzy sets enable their elements to show a partial degree of membership in the range from 0 (no membership) to 1 (full membership). In this way, fuzzy logic models are capable of representing continuous graduations from one class to another class (e.g., soil types, humus forms), which has been applied in soil science on numerous occasions (de Gruijter et al., 2011; McBratney and Odeh, 1997; Qi and Zhu, 2011; Zhu, 2006; Zhu et al., 1996, 2001). In the context of digital soil mapping, fuzzy logic models have been designed and applied for predicting various soil properties (Ashtekar and Owens, 2013; de Menezes et al., 2013). By means of fuzzy membership functions μ: $E_1 \times \ldots \times E_n \rightarrow [0,1]$, $(x_1,\ldots,x_n) \mapsto \mu(x_1,\ldots,x_n)$, dependencies of a target variable on environmental covariates (with the domains $E_1, \ldots, E_n$) can be described. These functions refer membership degrees of the target variable to different environmental circumstances (McBratney and Odeh, 1997).

In order to derive suitable fuzzy membership functions from an existing decision tree, tuples containing the value that is inhered in a distinct tree node and the related values of the covariates are used to approximate a general function rule. In case of the fuzzy membership $s_{ij,k,a}$ of a modeled variable k as a function of a single environmental variable a with the value $z_{ij,a}$ at location (i,j), a two-dimensional rule is needed, such as the bell-shaped function provided by Shi et al. (2009) and Shi (2013) (Equation 3), with the maximal membership $max \in [0,1]$, the central values of the function v_1 and v_2, the inflection points w_1 and w_2, and with r_1 and r_2 determining the steepness of the function parts.

$$\begin{cases} s_{ij,k,a} = max * \exp\{[(z_{ij,a} - v_1)/w_1]^{r_1} \ln[0.5]\} & if\ z_{ij,a} < v_1, \\ s_{ij,k,a} = max & if\ v_1 \leq z_{ij,a} \leq v_2, \\ s_{ij,k,a} = max * \exp\{[(z_{ij,a} - v_2)/w_2]^{r_2} \ln[0.5]\} & if\ z_{ij,a} > v_2 \end{cases} \qquad (3)$$

In case of a continuously increasing or decreasing behavior of the modeled variable k along the gradient of an environmental variable a, a sigmoidal function can be derived from Equation 3 by utilizing only the increasing or decreasing function parts, respectively. Membership values between 0 and 1 can be derived from the fuzzy membership functions depending on the values of the environmental covariates. If information about these covariates is extensively available, membership values can be modeled for every site across the study area and for every class of the target variable.

Similar approaches based on a combination of decision trees and fuzzy logic have been described elsewhere (Chiang and Hsu, 2002; Suárez and Lutsko, 1999) and utilized for soil scientific purposes (Ai et al., 2013; Ribeiro et al., 2014; Qi and Zhu, 2011). To our knowledge, though, this is the first time that a model combining decision tree analysis and fuzzy logic is applied for mapping indicators of decomposition.

Case Study: Modeling of Humus Forms at Val di Rabbi (Trentino, Italy)

Study Area

The study area (522.7 km^2) belongs to the Autonomous Province of Trentino in northern Italy and encompasses the two Alpine valleys Val di Sole

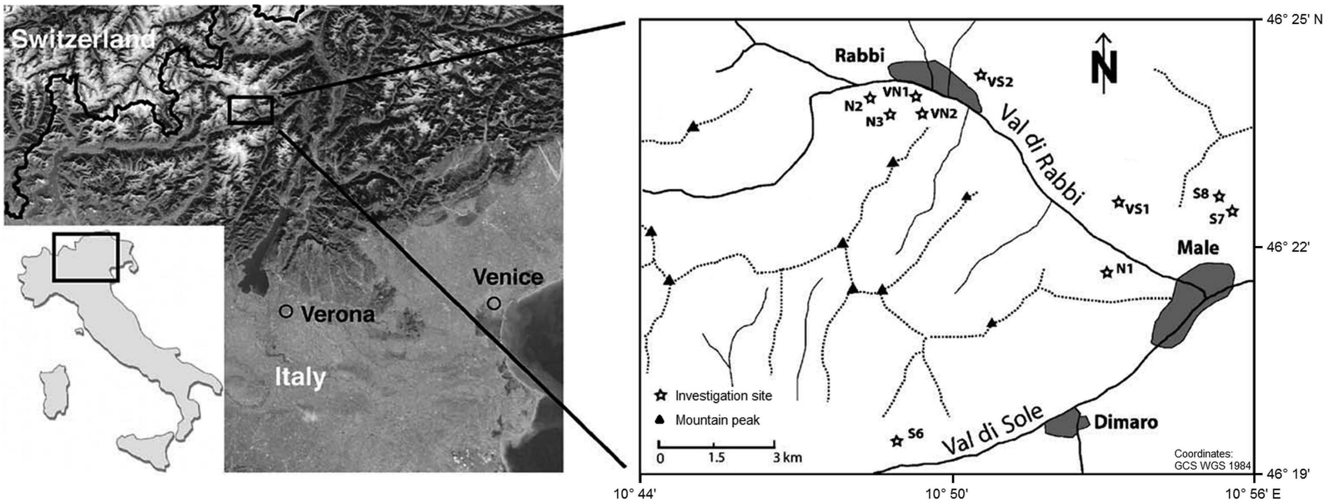

FIGURE 2. Location of the study area and investigation sites in the Autonomous Province of Trentino (Italy) (modified from Egli et al., 2006).

and Val di Rabbi (Fig. 2). The entire area is characterized by siliceous parent material, with paragneiss, mica schists and phyllites prevailing, and with some lenses of orthogneiss (Aberegg et al., 2009). In terms of soil classes, the range up to 1900 m a.s.l. is dominated by Haplic Cambisols (Dystric) and Umbric Podzols. At higher elevations, the prevalent classes are Entic Podzols, Albic Podzols, and Umbric Podzols (Sartori and Mancabelli, 2009).

Data Basis and Preprocessing

Six investigation sites were located inside the closed coniferous forest along different elevations between 1180 and 1660 m a.s.l., three each at north-exposed slopes (N1, N2, N3) and south-exposed slopes (S6, S7, S8). They are comparable with respect to geology (all sites have paragneiss as parent material) and local topographical position (all sites are located on uniform slopes). The north-exposed sites are dominated by Norway spruce (*Picea abies*) and the south-exposed sites by European larch (*Larix decidua*). The site S6 was located in a former coppice.

Every investigation site (~25 m^2, depending on the local variability of site conditions) comprised three or six sampling plots, respectively (humus profiles at a length of up to 1 m), according to the number of different soil cover types (grass, moss, litter, fern). These plots also included different local slope dynamics (i.e., erosive and accumulative characteristics) at every investigation site. The total number of sampling plots amounted to $n = 30$. The humus profiles were dominated by moder conditions (humus forms showing a continuous OH horizon), with transitions to mull-like conditions (OH horizon missing), primarily at the lowest north-facing site N1 as well as at the south-facing sites (Table 1).

The model of humus forms addresses the occurrence of a humus form with an OH horizon. Data from the sampling plots were obtained, indicating the occurrence of an OH horizon with values from the interval [0,1] (Table 2). The value 1 was assigned to a sample where a continuous OH horizon was present. If an OH horizon did not exist, the value 0 was assigned. For discontinuous OH horizons the small-scale presence and absence of an OH horizon often changed abruptly and irregularly (thus without the possibility to trace it back clearly to other factors). Therefore, the intermediate value 0.5 was used for this situation.

In order to examine the combined effects of elevation and slope exposition, the particular elevation and slope exposition values were aggregated to three levels of elevation and two levels of exposition. For elevation, the intermediate values 1200 m, 1400 m, and 1630 m were derived from the elevation values of the related sampling plots. For slope exposition, north-exposed sites were assigned a value of 360°N (equal to 0°N), and south-exposed sites were assigned a value of 180°N (Table

TABLE 1

Soil cover types, dominating humus forms, and humus profiles at the investigation sites (N1–N3, northern slope exposition; S6–S8, southern slope exposition) (according to Egli et al., 2006; personal communication, D. Tatti and G. Sartori).

			Dominating humus form according to		
Site	Elevation (m a.s.l.)	Soil cover types	German classification (Ad-hoc-AG Boden, 2005)	Classification from Switzerland (Gobat et al., 2014)	Typical humus profile
N1	1180–1195	moss (90 %), fern (10 %)	Mullartiger Moder	Hémimoder/ Eumoder/Dysmoder	OL-OF-(OH-)AE
N2	1395–1410	moss (100 %)	Typischer Moder	Dysmoder	OL-OF-OH-AE
N3	1595–1605	grass (80 %), moss (20 %)	Typischer Moder	Dysmoder	OL-OF-OH-E
S6	1200–1220	litter (90 %), grass (10 %)	Mullartiger Moder	Eumésoamphi	OL-OF-(OH-)A
S7	1380–1395	grass (100 %)	Mullartiger Moder	Hémimoder	OL-OF-(OH-)AE
S8	1650–1660	litter (80 %), grass (20 %)	Mullartiger Moder/ Typischer Moder	Hémimoder	OL-OF-(OH-)AE

TABLE 2

Data basis for modeling. Values of topography and humus forms have been aggregated from all sampling plots per investigation site.

Site	Number of sampling plots	Elevation (m a.s.l.)[1]	Slope exposition	Percentage of humus forms showing an OH horizon
N1	6	1200	north (360°N)	15.00%
N2	3	1400	north (360°N)	66.67%
N3	6	1630	north (360°N)	60.00%
S6	6	1200	south (180°N)	6.67%
S7	3	1400	south (180°N)	50.00%
S8	6	1630	south (180°N)	53.33%

[1]From DTM, aggregated by plot elevations.

2). The humus form data of each sampling plot was weighted according to the estimated percentage of its soil cover type in relation to the overall area at this site (Table 1).

Four additional sites for model validation were studied with a reduced number of sampling plots ($n = 8$). Two sites each were located at north-facing (VN1, VN2) and south-facing (VS1, VS2) slopes (Table 3).

Information about the elevation in the study area was taken directly from a bare ground digital terrain model (DTM) with a grid width of 10 m (Aberegg et al., 2009; compiled by the Provincia Autonoma di Trento on the basis of the topographic map with the scale of 1:10,000). A model representing slope exposition values was derived from this DTM with the slope method by Horn (1981).

Modeling

Two modeling approaches were juxtaposed. The first approach used results from decision tree analysis without any fuzzification process, and the second one combined decision trees and fuzzy logic. Decision trees were built using the statistical software R (R Core Team, 2015) and the R package rpart (Therneau et al., 2015). The routine can be found in the online appendix (file DecTreeAnalysis.R together with the data file hf_data.txt).

Bell-shaped fuzzy membership functions describing the occurrence of OH horizons in dependence on the elevation and slope exposition were formulated according to Equation 3. As the effects of two influencing variables (elevation and slope exposition) were examined, a two-step fuzzification procedure needed to be applied. With the first step, fuzzy membership functions were used to build submodels. These submodels correspond to the subtrees originating from the secondary split in the tree, thus they depend on the less influencing variable. The fuzzy membership functions were fitted by using values of the influencing variable

TABLE 3

Validation sites: topographic position and percentage values (observed and modeled) of humus forms showing an OH horizon.

				Decision tree model		Fuzzified decision tree model	
Site	Elevation (m a.s.l.)	Slope exposition	Observed value	Predicted value	Deviation	Predicted value	Deviation
VN1	1210	north	45.0%	10.8%	34.2%	11.2%	33.8%
VN2	1380	north	100.0%	51.7%	48.3%	62.0%	38.0%
VS1	1340	south	0.0%	63.3%	–63.3%	46.0%	–46.0%
VS2	1570	south	50.0%	51.7%	–1.7%	52.1%	–2.1%

together with the related percentages of a humus form with an OH horizon. With the second step, the submodels were combined by means of weighting functions that thereby realize fuzzification of the primary influencing variable. The resulting fuzzy membership functions were realized with the ArcGIS extension tool ArcSIE (Shi, 2013). This tool processed them to build maps that spatially predict the occurrence of OH horizons.

Modeling was conducted for the central part of Val di Rabbi. According to our field experience, the prediction area was selected to include the coniferous forest zone at the valley sides between 1100 m and 1800 m a.s.l.

The results from modeling were assessed with the mean error (Equation 4) and the root mean squared error (RMSE) (Equation 5) of the predictions at the validation sites:

$$ME = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \qquad (4)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \qquad (5)$$

The number of samples for validation is n, y_i are the observed values at the validation sites, and $\hat{y}_i$ are the related values predicted by the model.

Results

The sample data on the spatial distribution of humus forms showing an OH horizon and the related data on elevation and slope exposition was used for the construction of a decision tree. This yielded a tree with a primary partition induced by the factor elevation at 1300 m a.s.l. (Fig. 3). The left subtree consisted of a single node, representing the relatively similar percentages of humus forms with an OH horizon at north-exposed and south-exposed slopes below 1300 m. The right subtree applied to elevations from 1300 m upwards and included another partition to distinguish between northern and southern slope expositions.

The nodes of the decision tree were obtained by recursive partitioning of the sample set. Each of the three leaf nodes held a subset with a prediction about the occurrence of OH horizons for the study area, which was specific to the related elevation and exposition range (Fig. 3):

- below 1300 m: 10.83% of the area exhibited a humus form with an OH horizon (based on 12 samples)
- at south-exposed slopes from 1300 m upwards: 51.67% of the area exhibited a humus form with an OH horizon (based on 9 samples)
- at north-exposed slopes from 1300 m upwards: 63.33% of the area exhibited a humus form with an OH horizon (based on 9 samples)

The first model variant simply used this separation and mapped the values from the leaf nodes to the related elevation and exposition ranges inside the study area. This caused a steplike behavior of the distribution function (Fig. 4).

The second model variant also originated from the values of the leaf nodes, but includes a fuzzification at the transitions between different elevation and exposition ranges. As there are two variables

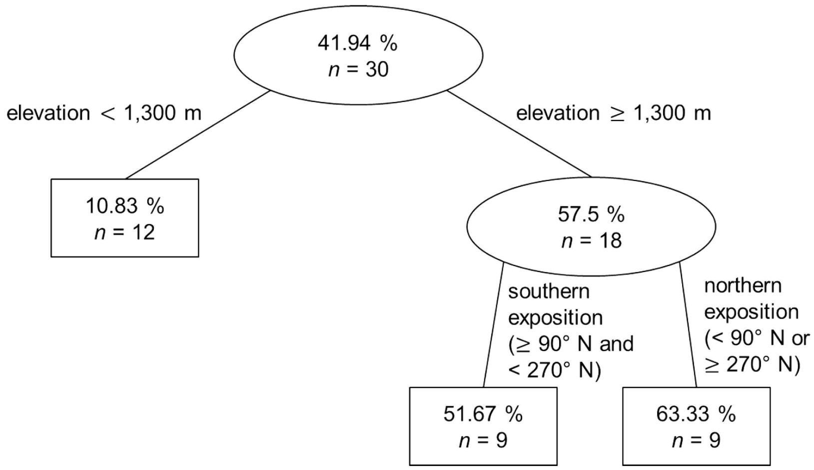

FIGURE 3. Decision tree for the distribution model of OH horizons. The upper value inside the tree nodes represents the projected percentage of the area with humus forms showing an OH horizon in relation to the overall area at this elevation and slope exposition. The lower value n indicates the number of related samples.

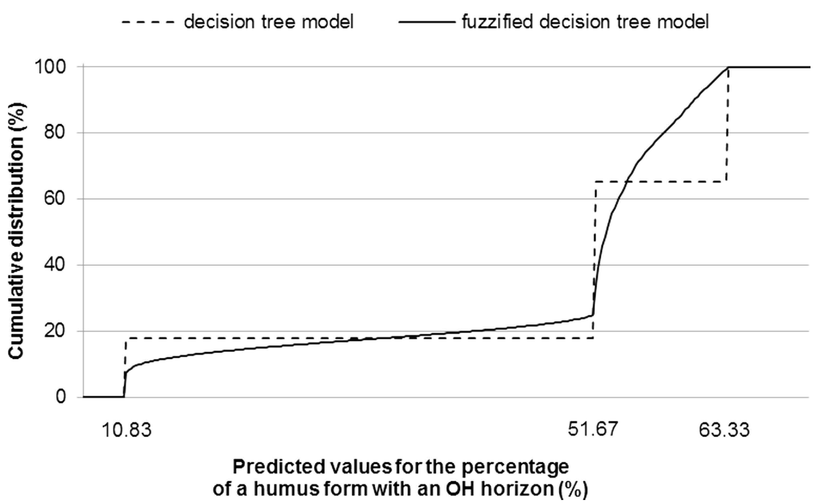

FIGURE 4. Cumulative distribution functions for the values predicting the percentage of a humus form with an OH horizon in the two models that are compared.

influencing the occurrence of OH horizons, two steps of fuzzification need to be realized. With the first step, the two subtrees at tree level 2 were fuzzified, both being independent on elevation (Fig. 3). For elevations below 1300 m a.s.l. (left subtree) the constant value 10.83% was used for all slope expositions (constant function $s_{ij,k,a}$ = 0.1083). At higher elevations (right subtree) fuzzy membership functions were constructed, which covered the pairs ($z_{ij,a}$, $s_{ij,k,a}$) referring to Equation 3 ($z_{ij,a}$ = slope exposition in °N, $s_{ij,k,a}$ = percentage of humus forms showing an OH horizon): (0, 0.6333), (180, 0.5167), (360, 0.6333). Figure 5 illustrates possible functions for different parameters r_1 / r_2 (assuming $r_1 = r_2$). In this case the function with $r_1 = r_2 = 2$ was chosen for further modeling, as a function with a median trend in the increasing occurrence of OH horizons from south to north exposition (Equation 6).

$$\begin{cases} s_{ij,k,a} = 1 - 0.4833 * \exp\{[(z_{ij,a} - 180)/285]^2 \ln[0.5]\} & \text{if } z_{ij,a} < 180, \\ s_{ij,k,a} = 0.5167 & \text{if } z_{ij,a} = 180, \\ s_{ij,k,a} = 1 - 0.4833 * \exp\{[(z_{ij,a} - 180)/285]^2 \ln[0.5]\} & \text{if } z_{ij,a} > 180 \end{cases} \quad (6)$$

Under the assumption $r_1 = r_2$ both of the functions for western and eastern slope expositions behaved equally ($v_1 = v_2$ and $w_1 = w_2$), so the system of equations Equation 6 reduced to a single equation (Equation 7).

$$s_{ij,k,a} = 1 - 0.4833 * \exp\{[(z_{ij,a} - 180)/285]^2 \ln[0.5]\} \quad (7)$$

The submodels for the two elevation ranges (1100–1300 m and 1300–1800 m) represent the (left and right) subtrees of the root node (Fig. 3), which both underwent fuzzification of the slope exposition. As a second fuzzification step the parti-

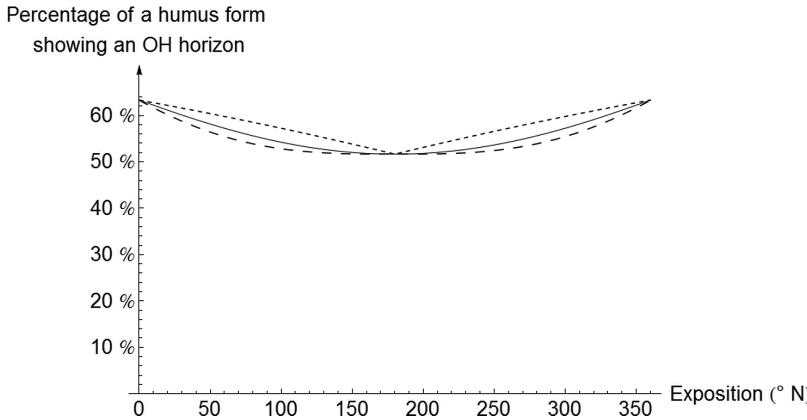

FIGURE 5. Fuzzy membership functions for the distribution model of OH horizons above 1300 m a.s.l., derived from the right subtree in Figure 3. The solid line represents the function with $r_1 = r_2 = 2$, the dashed lines have the parameters $r_1 = r_2 = 1$ (short dashing) and $r_1 = r_2 = 3$ (long dashing).

tion of the root node (where both submodels are connected) needs to be addressed. Fuzzification of the elevation was performed in the range between the examined investigation sites N1/S6 and N2/S7 (1200–1400 m a.s.l.), where predictions tend to be uncertain due to the lack of sampling data. It was accomplished by building functions associating the elevation with a weight between 0 and 1 (Fig. 6). These weighting functions were applied at the fuzzy membership functions of the submodels from the first fuzzification step. As a consequence of fuzzification, the distribution function of the second model showed a continuously increasing behavior without consecutive steps at the thresholds that stem from the decision tree (Fig. 4).

Prediction maps indicating the percentage of humus forms that show an OH horizon in the area selected for modeling were constructed for both model variants using ArcMap 10 and the extension ArcSIE (Fig. 7). According to the values of the fuzzy membership functions, the values of the prediction maps ranged from 10.8% to 63.3%. The lowest values were predicted for low elevation areas between 1100 m and 1200 m. At elevations from 1400 m upwards, where only the fuzzy membership functions for higher elevations was used (in consequence of the second fuzzification step, Fig. 6), there were significantly larger percentages of a humus form with an OH horizon (between 51.67% and 63.33%), with higher prediction values at north-exposed slopes.

Model validation shows notable deviations for both models, which are highest at site VS1 (observed value 0.0 %, predicted value 63.3 % when using the decision tree model and 46.0 % when using the fuzzified decision tree model). The deviations for the sites VN1 and VN2 are moderate to high, the observed value for site VS2 corresponds best with the values predicted by the models (Table 3). When including fuzzification, validation results in a mean error of 30.0 % and an RMSE (root mean squared error) of 34.3 %. When predicting values only based on the decision tree, the mean error is 36.9 % and the RMSE is 43.3 %.

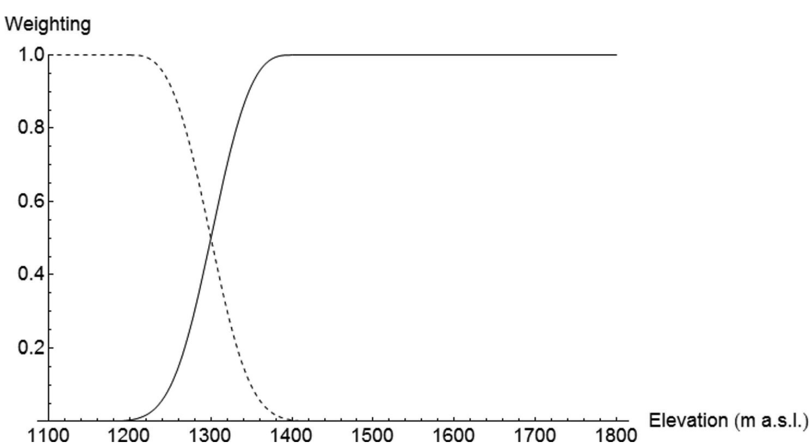

FIGURE 6. Weights for the synthesis of the elevation models. Dashed line: weights for the elevation model below 1300 m a.s.l.; solid line: weights for the elevation model from 1300 m a.s.l. upward. Parameter values of the model for lower elevations (dashed line): $v_2 = 1200$, $w_2 = 100$, $r_2 = 3$. Parameter values of the model for higher elevations (solid line): $v_1 = 1400$, $w_1 = 100$, $r_1 = 3$.

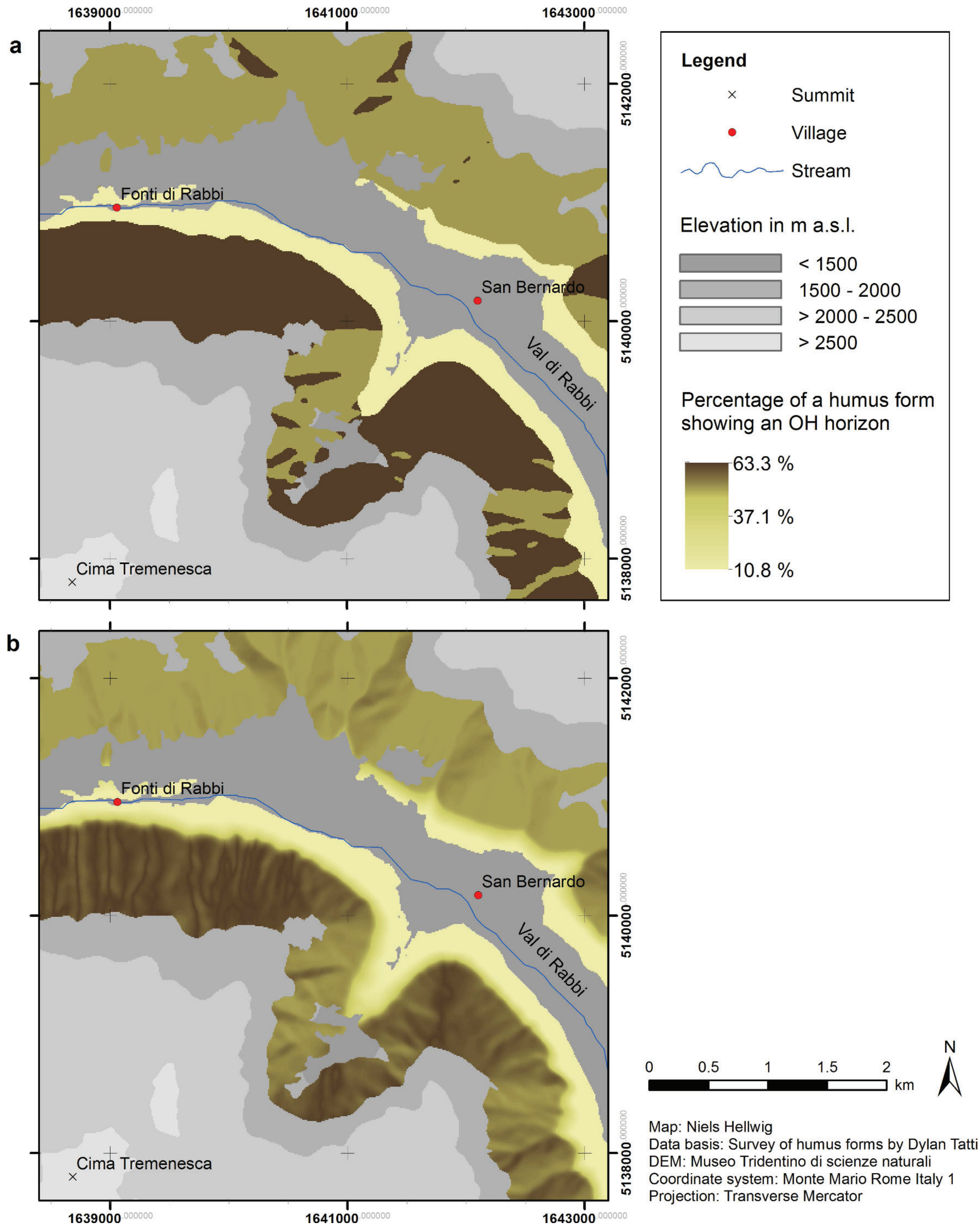

FIGURE 7. Comparison of the predicted spatial distribution of humus forms showing an OH horizon (a) using decision tree analysis without fuzzification procedure, and (b) combining decision tree analysis and fuzzy membership functions. In contrast to the first model (a), which only predicts three different percentage values, the second model (b) incorporates the gradual transitions in humus forms along different elevations and slope expositions. The modeled area includes areas with coniferous forest between 1100 m and 1800 m a.s.l. in the central part of Val di Rabbi (Trentino, Italy).

Discussion

Results from the Case Study

The case study showed predictions for the occurrence of OH horizons that vary in accordance with the modeling approach. The model that includes a fuzzification procedure fitted the real situation potentially better than the one without this procedure, as there were no abrupt changes in the projected values due to any threshold values from the decision tree, which cannot be justified based on the low data amount. Nevertheless, the structure of the decision tree was also reflected in the predictive map of the fuzzy logic model (Fig. 7).

Regarding the occurrence of OH horizons, modeling revealed differences primarily with a changing elevation. At higher elevations within the forest, north-exposed slopes were predicted to exhibit slightly larger percentages of a humus form with an OH horizon than south-exposed slopes. The validation showed that deviations of field observations from the model results were likely to occur in some places. Nevertheless the predicted trends of a higher occurrence of OH horizons with an increasing elevation and also at northern slope expositions were confirmed. The deviations at the validation sites suggested that further thresholds might exist depending on the elevation, which were still not well explained by the model.

This predicted distribution depending on the elevation and slope exposition is in line with the findings of other investigations of forest ecosystems, for example higher accumulation of organic matter at north-facing compared to south-facing slopes (e.g., Aberegg et al., 2009; Ascher et al., 2012; Bernier, 1996; Egli et al., 2009, 2010a, 2010b). The correlation of humus forms and the relief parameters elevation and slope exposition emphasizes the influence of environmental factors such as solar radiation, temperature and vegetation on decomposition processes also in the study area. The effects of other covariates on decomposition such as parent material were not examined in the context of this study, thus modeling results are potentially subject to uncertainties at sites where environmental conditions deviate from the investigation sites (e.g., on mica schists). With a higher number of investigation sites the results could possibly be improved by integrating further potentially influencing variables.

Application of Decision Trees and Fuzzy Logic

Modeling is exerted with a knowledge-based approach, built on the use of decision tree analysis and the concept of fuzzy logic. Decision trees have shown to be well suited for revealing relationships between sample data and environmental factors (Aberegg et al., 2009; De'ath and Fabricius, 2000; Gerlitz, 2015). In the case of a small sample size, decision tree analysis has to be frequently performed without any pruning procedure, since this would eliminate almost every partition of the sample set (for details on the pruning procedure, see Breiman et al., 1984). For that reason the number of sample elements is usually low especially in the leaf nodes, thus the direct use of maximum decision trees, which are not pruned, for prediction is critical. However, even these maximum trees are not generally overfitted, as long as the distinct places of the sampling plots have been determined directly in the field using the knowledge and experience of experts in the sector of decomposition.

The application of fuzzy logic helps to manage the limited predictive capability of the trees, as the divisions of the sample set and the related values of the covariates are fuzzified and do not act as strict thresholds. The use of a nonautomated fuzzification procedure also allows the integration of expert knowledge for the definition of fuzzy membership functions (e.g., when choosing free parameters; see Fig. 5). Because of the relatively low number of samples and the ensuing use of the results from a nonpruned tree, an automated fuzzification procedure does not seem to be appropriate (Gerlitz, 2015; Suárez and Lutsko, 1999).

With the help of fuzzy membership functions, the similarity of the environmental characteristics at a site in comparison to those at sites typical for one specific predicted value (e.g., OH horizon present) can be modeled in the form of a membership (de Menezes et al., 2013; Zhu, 1997). The usage of fuzzy membership values enables also the consideration of a high spatial heterogeneity. This allows, for example, for involving the small-scale variability, which has been shown to be an important characteristic of decomposition processes and properties in the Alps (Bednorz et al., 2000). Consequently, the predicted percentage values, arising from the fuzzy membership values, provide information about the local variability of humus forms, although it is not possible to get spatially differentiated predictions at a resolution higher than 10 × 10 m².

The bell-shape functions used to model the dependency of indicators of decomposition on environmental factors show a high flexibility and can even manage factors that take values in a cyclic range (such as slope exposition). Nevertheless, the application of a bell-shape function to map the relationship between the slope exposition and indicators of decomposition seems to be disadvantageous, as it shows different behavior for sites north exposed (near 0°N and 360°N) and south exposed (around 180°N) (see Fig. 5). A potentially more reasonable way for this would be the implementation of a trigonometrical function.

Uncertainties and Validity of Models

For the application of environmental models it is essential to treat uncertainties, with respect to the acquisition of data on the one hand and to the modeling process on the other hand (Brown, 2010; Keenan et al., 2011). Within the modeled area of the case study, the results are subject to different magnitudes of uncertainty, depending on the similarity of the elevation and slope exposition values to those of the investigation sites. Accordingly, the highest uncertainties of the model results are located at eastern and western expositions and in elevation ranges midway between the investigation sites (e.g., 1300 m a.s.l., between 1200 m and 1400 m). This kind of uncertainty corresponds with the level of accordance of the fuzzy membership functions with the actual relationships in the modeled ecosystem.

Other sources of potential uncertainties of the results are deviations of the elevation values in the DTM from the real values, errors due to the calculation procedure of exposition values and inaccuracies due to the discrete representation of the landscape in the form of a raster (Bocedi et al., 2012; Fisher and Tate, 2006; Wechsler, 2007). In the context of topographical and hydrological analyses, small errors in a DTM can result in major deviations when deriving relief parameters (Holmes et al., 2000; Zhou and Liu, 2004).

The validity of the results from the case study is constrained to the modeled forested areas in the range from 1100 m to 1800 m a.s.l. An explicit validation procedure is essential when having the intention to apply concrete predictions on indicators of decomposition (e.g., in the context of ecosystem management). This validation procedure also needs to consider the effect of additional environmental covariates that might be relevant for decomposition processes at places inside the area selected for prediction showing varying site conditions compared to the sampling locations. When planning to transfer the modeled effects of elevation and slope exposition on decomposition to woodless areas or other regions that exhibit different environmental conditions, a particular assessment of the transferability of the model is required (Wenger and Olden, 2012).

Conclusions

Two modeling approaches for upscaling of sample data on indicators of decomposition from the local scale to the landscape scale have been juxtaposed. This study focused on an area characterized by a highly heterogeneous relief and data from a relatively small number of samples, which have been surveyed in places specified with expert knowledge. Combining decision tree analysis and the use of fuzzy membership functions has shown to serve as a suitable approach. Building decision trees helped to generate information about the influences of the environmental factors elevation and slope exposition. However, direct upscaling of information about decomposition processes from the results of decision tree analysis yielded partially unrealistic predictions that were manifested in abrupt transitions between areas differentiated by the tree.

Continuous and more realistic transitions can be achieved by further processing the results from decision tree analysis through an additional step that comprises the construction of bell-shape fuzzy membership functions. For the parameter slope exposition a future improvement of the mapping could be to leave the functionalities of ArcSIE and use another function type (e.g., trigonometric function). As the modeling approach is based on fuzzy logic, it accounts for small-scale variations in decomposition processes as well as for uncertainties caused by the inference from a few investigation sites to a large study area.

Spatial modeling utilizing the technique presented in this paper is considered to be a useful tool to obtain a detailed insight into decomposition processes in a high mountain environment. Implementing such a model should include a validation procedure and an analysis of uncertainty. Apart from the humus form, this approach could be used to examine a variety of other related parameters on a landscape scale, such as the pH value and the composition of the decomposer community.

Acknowledgments

This study was realized in the context of the D.A.CH. project DecAlp and funded by the German Research Foundation (DFG, grant number BR 1106/23-1). The authors thank all colleagues of the project for an excellent cooperation. We are in particular grateful to Dylan Tatti (University of Neuchâtel) for sharing his data on humus forms and to Giacomo Sartori (Museo Tridentino di Scienze Naturale) for valuable discussions in the field. We also thank the anonymous reviewers for valuable comments on an earlier version of the manuscript.

References Cited

Aberegg, I., Egli, M., Sartori, G., and Purves, R., 2009: Modelling spatial distribution of soil types and characteristics in a high Alpine valley (Val di Sole, Trentino, Italy). *Studi Trentini di Scienze Naturali*, 85: 39–50.

Ad-hoc-AG Boden, 2005: *Bodenkundliche Kartieranleitung*. Fifth edition. Hannover: E. Schweizerbart'sche Verlagsbuchhandlung, 438 pp.

Ai, L., Fang, N. F., Zhang, B., and Shi, Z. H., 2013: Broad area mapping of monthly soil erosion risk using fuzzy decision tree approach: integration of multi-source data within GIS. *International Journal of Geographic Information Science*, 27: 1251–1267.

Andreetta, A., Macci, C., Ceccherini, M. T., Cecchini, G., Masciandaro, G., Pietramellara, G., and Carnicelli, S., 2012: Microbial dynamics in Mediterranean Moder humus. *Biology and Fertility of Soils*, 48: 259–270.

Ascher, J., Sartori, G., Graefe, U., Thornton, B., Ceccherini, M. T., Pietramellara, G., and Egli, M., 2012: Are humus forms, mesofauna and microflora in subalpine forest soils sensitive to thermal conditions? *Biology and Fertility of Soils*, 48: 709–725.

Ashtekar, J. M., and Owens, P. R., 2013: Remembering knowledge: an expert knowledge based approach to digital soil mapping. *Soil Horizons*, 54: doi http://dx.doi.org/10.2136/sh13-01-0007.

Bednorz, F., Reichstein, M., Broll, G., Holtmeier, F.-K., and Urfer, W., 2000: Humus forms in the forest-alpine tundra ecotone at Stillberg (Dischmatal, Switzerland): spatial heterogeneity and classification. *Arctic, Antarctic, and Alpine Research*, 32: 21–29.

Behrens, T., and Scholten, T., 2006: Digital soil mapping in Germany—A review. *Journal of Plant Nutrition and Soil Science*, 169: 434–443.

Behrens, T., and Scholten, T., 2007: A comparison of data-mining techniques in predictive soil mapping. *In* Lagacherie, P., McBratney, A. B., and Voltz, M. (eds.), *Digital Soil Mapping—An Introductory Perspective*. *Developments in Soil Science*, v. 31. Amsterdam: Elsevier, 353–366.

Bernier, N., 1996: Altitudinal changes in humus form dynamics in a spruce forest at the montane level. *Plant and Soil*, 178: 1–28.

Bocedi, G., Pe'er, G., Heikkinen, R. K., Matsinos, Y., and Travis, J. M. J., 2012: Projecting species' range expansion dynamics: sources of systematic biases when scaling up patterns and processes. *Methods in Ecology Evolution*, 3: 1008–1018.

Bouma, J., 1989: Using soil survey data for quantitative land evaluation. *In* Stewart, B. A. (ed.), *Advances in Soil Science*, v. 9. New York: Springer-Verlag, 177–213.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J., 1984: *Classification and Regression Trees*. Belmont: Wadsworth, 358 pp.

Brown, J. D., 2010: Prospects for the open treatment of uncertainty in environmental research. *Progress in Physical Geography*, 34: 75–100.

Chiang, I.-J., and Hsu, J. Y., 2002: Fuzzy classification trees for data analysis. *Fuzzy Sets and Systems*, 130: 87–99.

De'ath, G., and Fabricius, K. E., 2000: Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81: 3178–3192.

de Gruijter, J. J., Walvoort, D. J. J., and Bragato, G., 2011: Application of fuzzy logic to Boolean models for digital soil assessment. *Geoderma*, 166: 15–33.

de Menezes, M. D., Silva, S. H. G., Owens, P. R., and Curi, N., 2013: Digital soil mapping approach based on fuzzy logic and field expert knowledge. *Ciência e Agrotecnologia*, 37: 287–298.

Egli, M., Mirabella, A., Sartori, G., Zanelli, R., and Bischof, S., 2006: Effect of north and south exposure on weathering rates and clay mineral formation in Alpine soils. *Catena*, 67: 155–174.

Egli, M., Sartori, G., Mirabella, A., Favilli, F., Giaccai, D., and Delbos, E., 2009: Effect of north and south exposure on organic matter in high Alpine soils. *Geoderma*, 149: 124–136.

Egli, M., Sartori, G., Mirabella, A., and Giaccai, D., 2010a: The effects of exposure and climate on the weathering of late Pleistocene and Holocene Alpine soils. *Geomorphology*, 114: 466–482.

Egli, M., Sartori, G., Mirabella, A., Giaccai, D., Favilli, F., Scherrer, D., Krebs, R., and Delbos, E., 2010b: The influence of weathering and organic matter on heavy metals lability in silicatic, Alpine soils. *Science of the Total Environment*, 408: 931–946.

Fisher, P. F., and Tate, N. J., 2006: Causes and consequences of error in digital elevation models. *Progress in Physical Geography*, 30: 467–489.

Gerlitz, L., 2015: Using fuzzified regression trees for statistical downscaling and regionalization of near surface temperatures in complex terrain. *Theoretical and Applied Climatology*, 122: 337–352.

Gobat, J.-M., Le Bayon, C., and Tatti, D., 2014: *Clé de Sol—Principaux sols de Suisse*. Neuchâtel: Laboratoire Sol & Végétation, Université de Neuchâtel, Switzerland, 76 pp.

Graefe, U., and Beylich, A., 2006: Humus forms as tool for upscaling soil biodiversity data to landscape level? *Mitteilungen der Deutschen Bodenkundlichen Gesellschaft*, 108: 6–7.

Heuvelink, G. B. M., and Webster, R., 2001: Modelling soil variation: past, present, and future. *Geoderma*, 100: 269–301.

Holmes, K. W., Chadwick, O. A., and Kyriakidis, P. C., 2000: Error in a USGS 30-meter digital elevation model and its impact on terrain modeling. *Journal of Hydrology*, 233: 154–173.

Horn, B. K. P., 1981: Hill shading and the reflectance map. *Proceedings of the IEEE*, 69: 14–47.

Jenny, H., 1941: *Factors of Soil Formation—A System of Quantitative Pedology*. New York: McGraw-Hill, 281 pp.

Keenan, T. F., Carbone, M. S., Reichstein, M., and Richardson, A. D., 2011: The model-data fusion pitfall: assuming certainty in an uncertain world. *Oecologia*, 167: 587–597.

McBratney, A. B., and Odeh, I. O. A., 1997: Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. *Geoderma*, 77: 85–113.

McBratney, A. B., Minasny, B., Cattle, S. R., and Vervoort, R. W., 2002: From pedotransfer functions to soil inference systems. *Geoderma*, 109: 41–73.

McBratney, A. B., Mendonça Santos, M. L., and Minasny, B., 2003: On digital soil mapping. *Geoderma*, 117: 3–52.

McKenzie, N. J., and Ryan, P. J., 1999: Spatial prediction of soil properties using environmental correlation. *Geoderma*, 89: 67–94.

Mertens, M., Nestler, I., and Huwe, B., 2002: GIS-based regionalization of soil profiles with Classification and Regression Trees (CART). *Journal of Plant Nutrition and Soil Science*, 165: 39–43.

Ponge, J.-F., 2013: Plant-soil feedbacks mediated by humus forms: a review. *Soil Biology & Biochemistry*, 57: 1048–1060.

Qi, F., and Zhu, A.-X., 2011: Comparing three methods for modeling the uncertainty in knowledge discovery from area-class soil maps. *Computers & Geosciences*, 37: 1425–1436.

R Core Team, 2015: R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, <http://www.R-project.org/>, accessed 27 October 2015.

Ribeiro, M. V., Cunha, L. M. S., Camargo, H. A., and Rodrigues, L. H. A., 2014: Applying a fuzzy decision tree approach to soil classification. *In* Laurent, A., Strauss, O., Bouchon-Meunier, B., and Yager, R. R. (eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Communications in Computer and Information Science*, v. 442. Cham: Springer-Verlag, 87–96.

Sartori, G., and Mancabelli, A., 2009: *Carta dei suoli del Trentino: scala 1:250.000*. Museo Tridentino di Scienze Naturali di Trento, Centro di Ricerca per l'Agrobiologia e la Pedologia di Firenze.

Shi, X., 2013: ArcSIE, <http://www.arcsie.com/Download/htm>, accessed 11 October 2016.

Shi, X. Long, R., Dekett, R., and Philippe, J., 2009: Integrating different types of knowledge for digital soil mapping. *Soil Science Society of America Journal*, 73: 1682–1692.

Suárez, A., and Lutsko, J. F., 1999: Globally optimal fuzzy decision trees for classification and regression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21: 1297–1311.

Swift, M. J., Heal, O. W., and Anderson, J. M., 1979: *Decomposition in Terrestrial Ecosystems*. Berkeley, Los Angeles: University of California Press, *Studies in Ecology*, v. 5, 372 pp.

Therneau, T. M., Atkinson, E. J., and Foundation, M., 2015: An Introduction to Recursive Partitioning Using the RPART Routines. <http://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf>, accessed 27 October 2015.

Wechsler, S. P., 2007: Uncertainties associated with digital elevation models for hydrologic applications: a review. *Hydrology and Earth System Sciences*, 11: 1481–1500.

Wenger, S. J., and Olden, J. D., 2012: Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods in Ecology and Evolution*, 3: 260–267.

Wösten, J. H. M., Pachepsky, Y. A., and Rawls, W. J., 2001: Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251: 123–150.

Zadeh, L. A., 1965: Fuzzy sets. *Information and Control*, 8: 338–353.

Zhou, Q., and Liu, X., 2004: Analysis of errors of derived slope and aspect related to DEM data properties. *Computers & Geosciences*, 30: 369–378.

Zhu, A.-X., 1997: A similarity model for representing soil spatial information. *Geoderma*, 77: 217–242.

Zhu, A.-X., 2006: Fuzzy logic models. *In* Grunwald, S. (ed.), *Environmental Soil-Landscape Modeling. Geographic Information Technologies and Pedometrics*. New York: CRC Press, 215–239.

Zhu, A.-X., Band, L. E., Dutton, B., and Nimlos, T. J., 1996: Automated soil inference under fuzzy logic. *Ecological Modelling*, 90: 123–145.

Zhu, A. X., Hudson, B., Burt, J., Lubich, K., and Simonson, D., 2001: Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Science Society of America Journal*, 65: 1463–1472.

MS submitted 5 November 2015
MS accepted 28 July 2016