

Modelling spatial distribution of soil types and characteristics in a high Alpine valley (Val di Sole, Trentino, Italy)

Isabelle ABEREGG¹, Markus EGLI^{1*}, Giacomo SARTORI² & Ross PURVES¹

¹Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

²Museo Tridentino di Scienze Naturali, Via Calepina 14, 38122 Trento, Italy

*Corresponding author e-mail: markus.egli@geo.uzh.ch

SUMMARY - *Modelling spatial distribution of soil types and characteristics in a high Alpine valley (Val di Sole, Trentino, Italy)* - Detailed soil maps in Alpine areas are often not available due to the high variability of the topography, the inaccessibility of parts of the area and consequently high production costs. In the context of growing demand for high-resolution spatial information for environmental planning and modelling, fast and accurate methods are needed to provide high-quality digital soil maps. We performed a spatial analysis to model several characteristics of Alpine soils in Val di Sole, Val di Peio and Val di Rabbi (in total 374 km²). Soil modelling was performed using a non-parametric classification and decision tree analysis (CART: Classification and Regression Tree Analysis). The classification and decision tree analysis used forced splitting rules (according to expert knowledge). Soil type modelling was done using 15 end nodes. Spatial modelling of humus forms could be achieved with 9 terminal nodes. Field and chemical data (115 sites) served as a basis for modelling. In addition, conventional soil mapping was performed on three relatively small test areas. The modelling results could therefore be tested using these maps. Modelling of soils and humus forms was performed successfully with an accuracy of about 65% for soil types and higher values (up to 78%) for the humus forms. The main soil type in the investigation area is a ranker (WRB: *Umbric Leptosol*). The other soil groups (including *Cambisols*, *Umbric Podzols*) each covered about 11-15% of the investigation area. Around 66% of the area was dominated by the humus form moder.

RIASSUNTO - *Distribuzione del modello spaziale dei tipi e delle caratteristiche del suolo in un'alta valle alpina (Val di Sole, Trentino)* - Per le zone alpine non sono in genere disponibili carte pedologiche a scale di dettaglio, a causa della complessità della topografia, dei problemi di accesso a certe zone e degli alti costi che comporta la loro stesura. In un contesto di crescente bisogno di informazioni ad alta risoluzione per la gestione dell'ambiente e per la messa a punto di modelli ambientali, si rendono però necessari metodi per produrre in modo speditivo ed economico carte pedologiche digitali di alta qualità. Abbiamo dunque condotto un'analisi spaziale finalizzata a modellizzare vari caratteri di suoli alpini in Val di Sole, Val di Peio e Val di Rabbi (in totale 374 km²). La modellizzazione del suolo è stata realizzata utilizzando la procedura non parametrica di classificazione e di regressione ad albero (CART: Classification and Regression Tree Analysis), in base ai dati di campagna e chimici di 115 siti. La classificazione e regressione ad albero ha impiegato criteri di split forzati (basati su conoscenze di esperto). La modellizzazione del tipo di suolo è stata eseguita mediante un albero con 15 nodi terminali, quella della forma di humus con un albero con 9 nodi terminali. I risultati dei modelli elaborati sono stati testati tramite il confronto con tre carte pedologiche tradizionali di altrettante zone campione di dimensioni relativamente ridotte. Tale confronto ha permesso di evidenziare un'alta capacità predittiva dei modelli, con un'accuratezza del 65% per il tipo di suolo e valori più alti (fino al 78%) per la forma di humus. Il principale tipo di suolo presente nell'area di studio è il ranker (WRB: *Umbric Leptosol*). Gli altri tipo di suolo (*Cambisols*, *Umbric Podzols*) occupano ciascuno circa l'11-15% dell'area indagata. La forma di humus moder è presente nel 66% dell'area.

Key words: Alpine area, soil modelling, humus forms, Alpine soils, classification and decision tree analysis

Parole chiave: area alpina, modellizzazione dei suoli, forme di humus, suoli alpini, classificazione e regressione ad albero

1. INTRODUCTION

Previous investigations in Val di Sole and neighbouring areas (Sartori *et al.* 2005; Egli *et al.* 2006a) identified the main soil types for this central Alpine region. The soils have predominantly developed on siliceous parent material. Rankers, podzolic soils and cambisols are the main types.

There are, however, little informations available about the precise distribution of different soil types and their characteristics and, typically, detailed soil maps are not available in Alpine areas.

The production of conventional soil maps in Alpine areas is extremely laborious and therefore expensive. A major problem is the high variability of landforms with very distinct changes within short distances (steep valleys, ridges, rough or even slopes etc.). The changing topography affects also soil types and their properties. An additional problem is the inaccessibility or problematic accessibility of many sites.

In the context of growing demand for high-resolution spatial information for environmental planning and modelling, fast and accurate methods are needed to provide high-

quality digital soil maps (Rahmann *et al.* 1997; Tognina 2004; Behrens *et al.* 2005). In this context, data-mining methods may provide solutions. The term “data-mining” comprises various methods and techniques from statistics, mathematics and information theory (e.g. artificial neural networks, decision trees etc.; see Scull *et al.* 2003; McBratney *et al.* 2003) aiming to automatically extract hidden predictive information from existing datasets (Behrens *et al.* 2005). Over the last 10 years, *Digital Soil Mapping* (DSM) has emerged as a credible alternative to traditional soil mapping. However, DSM should not be seen as an end in itself, but rather as a technique for providing data and information for a new framework for soil assessment (Carré *et al.* 2007).

Basic digital data describing a landscape such as digital elevation models (DEM), geological maps, precipitation information and vegetation maps are often available. These datasets form the basis for soil modelling (see also soil forming factors as defined in Jenny (1980)). According to Scull *et al.* (2003), Geographic Information Systems (GIS) can be used to predict soil properties on the basis of such environmental variables, which are much easier to measure than the actual soil distribution. This idea is based on the paradigm of Jenny’s soil-forming factors according to which the soil (type) at a specific location is the result of the soil forming factors climate, organisms, relief, parent material and time. Advances in mathematical theories and statistical methods (through enhanced calculating capacity) have stimulated research activities in the field of predictive soil mapping and the solution of Jenny’s equation (Scull *et al.* 2003).

The review papers of Scull *et al.* (2003) and McBratney *et al.* (2003) give an overview on predictive soil-modelling techniques and their utilisation. Inductive models are used to derive and quantify the relationships between soil types and environmental variables (e.g. Lagacherie & Holmes 1997; Behrens *et al.* 2005; Carré *et al.* 2007). Other models are based more on expert knowledge, where existing knowledge is encoded in clear decision rules to spatially deduce the distribution of different soil types and characteristics (e.g. Zhu *et al.* 2001; Wilemaker *et al.* 2001; Egli *et al.* 2005, 2006b).

Soil, however, can only be measured at a finite number of sites and times with small supports, and any statement concerning the soil at other sites or times involves prediction. Spatial variation in soil characteristics is so complex that no description of it can be complete, and so prediction is inevitably uncertain.

The main aim of this work is to model the distribution of soils and their properties in a rugged, Alpine topography and to test the suitability of a GIS-based inductive model that can be combined with expert knowledge.

2. STUDY AREA

The study area is located in the north-western part of the Trentino Province (Fig. 1) and comprises Val di Sole and the two adjacent lateral valleys, namely Val di Rabbi and Val di Peio. The region is characterised by a large altitudinal gradient ranging from 700 m a.s.l. at Malè to glaciated peaks at 3769 m a.s.l. (Cima Cevedale). The climate is humid and temperatures are moderate: at Peio (1580 m

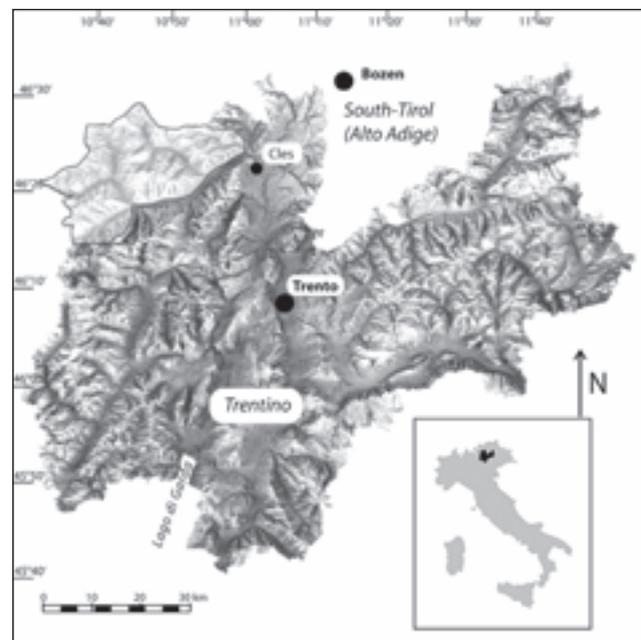
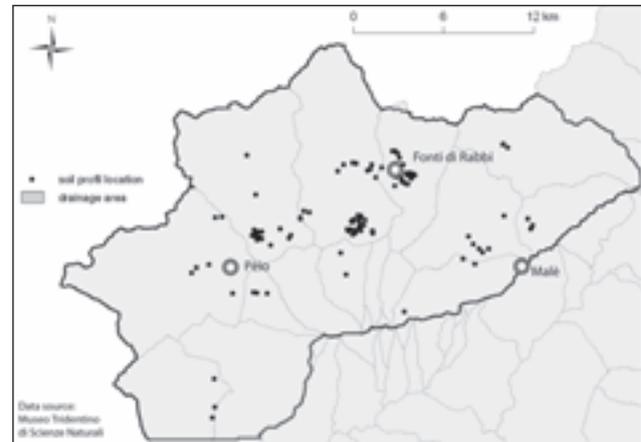


Fig. 1 - Study area (Val di Sole, Val di Rabbi and Val di Peio) and distribution of soil profiles.

Fig. 1 - Area di studio (Val di Sole, Val di Rabbi e Val di Peio), e localizzazione dei profili di suolo.

a.s.l.) the mean annual air temperature is around 6.8 °C and precipitation around 855 mm yr⁻¹ (Ufficio Previsioni e Organizzazione, Provincia Autonoma di Trento). With higher altitudes temperature decreases and precipitation increases (to about 0 °C and 1300 mm yr⁻¹ at 2400 m a.s.l.).

The geology of the study area is dominated by siliceous metamorphic rocks belonging to the Austroalpine lithostratigraphic units (Seidlein 2000, see Fig. 2). Only a very small part of the study area can be attributed to the calcareous Dinaric Alps and the Adamello granite intrusion (southern Alps). The Austroalpine region and the southern Alps are separated by the Insubric line.

The Austroalpine lithostratigraphic units between the Insubric and Peio (a minor geological fault) lines consist mainly of paragneiss and to a lesser extent of orthogneiss, whereas north of the Peio line the dominant materials for soil development are schists and phyllites.

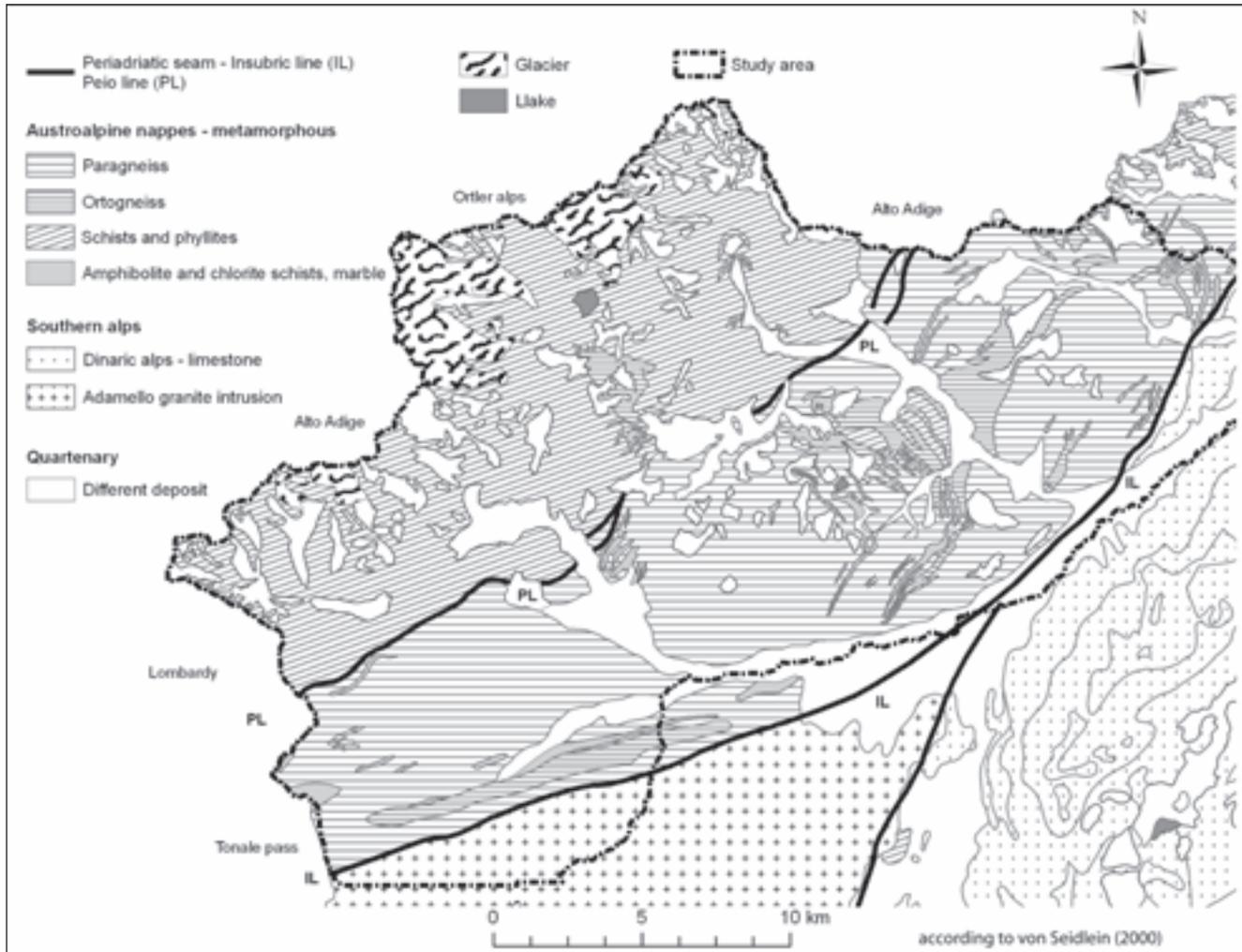


Fig. 2 - Geological situation in the study area.
 Fig. 2 - Geolitologia dell'area di studio.

Besides gneiss, schists and phyllites, other geological materials like amphibolite, chlorite schists or marble occur in some very small areas. The whole area was affected by glaciation and large parts of the soils have developed on morainic materials.

In the humid moderate climate of the region the main soil processes on siliceous material are podsolisation and at lower altitudes brunification. The vegetation in Val di Sole is typical for the Central Alps. The subalpine belt with spruce fir starts at a lower limit compared to the average in the Alps. In addition, beech is completely missing in the lower colline and montane belt (Tab. 1, Landolt 1992). Depending on solar radiation, the suprasubalpine belt forms the timberline at an altitude of 1900 to 2100 m a.s.l. Larch and the Swiss stone pine are the dominant species of the central Alpine vegetation. At higher altitudes, dwarf-shrubs and alpine meadows follow. The zonation of the vegetation is also shown in figure 3.

In some very small areas of the southern boundary of the study area, broad-leaved species dominate on calcareous areas at lower altitudes. These areas were excluded from the investigation.

3. MATERIAL AND METHODS

3.1. Soil classification system

The various soil types are differentiated according to the traditional French nomenclature (Duchaufour 2006) and the WRB (FAO 1998). In the investigation area, a total of 115 soil profiles were examined during 2003-2007. Chemical and physical analyses are available for 24 profiles and physical analyses only for 8 profiles. For the remaining 83 sites, field observations and measurements were taken. The field measurements included the determination of the soil type, humus form, soil depth, soil thickness, Munsell-Color, pH, volumetric content of soil skeleton and the estimation of the texture. The sites are shown in figure 1. The classification of the humus forms is according to BGS & FAL (2002). The differentiation of humus forms is based on the sequence of horizons and their development. Three main types were distinguished for modelling: mull, moder and nor. Briefly, mull has an O1 and A horizon (Of is only weakly developed), whereas a moder has in general the sequence O1-Of-(Oh)-A. Mor has the horizons O1-Of-Oh and no humic A-horizon.

Tab. 1 - Altitudinal zonation of the vegetation. The belts in Val di Sole correspond mainly to the general distribution of the central Alps.
 Tab. 1 - Zonazione altitudinale della vegetazione. Le fasce vegetazionali in Val di Sole corrispondono in linea di massima con quelle generali delle Alpi centrali.

Altitudinal zonation (Landolt 1992)	Altitude (m a.s.l.)	Description
Colline belt (oak-beech-belt)	600-1000 (depending on aspect and radiation)	Downy oak (<i>Quercus pubescens</i>); only in the most southern part of the study area
Montane belt (european silver fir-beech-belt)	1000-1600	European silver fir (<i>Abies alba</i>); only in the most southern part of the study area, mostly together with spruce fir (<i>Picea excelsa</i>)
Subalpine belt (spruce fir-belt)	1600-1900	Spruce fir (<i>Picea excelsa</i>) with scots pine (<i>Pinus sylvestris</i>), at higher altitudes with larch (<i>Larix decidua</i>) and swiss stone pine (<i>Pinus cembra</i>)
Suprasubalpine belt (swiss stone pine-belt)	1900-2200 (timberline)	Swiss stone pine (<i>Pinus cembra</i>) and larch (<i>Larix decidua</i>) on shallow mountain soils; with Alpine rose (<i>Rhododendron ferrugineum</i>) and juniper (<i>Juniperus communis</i>) as shrub
Alpine belt (alpine meadow-belt)	2200-2700	Upper limit defined by coherent meadow; low grass with sedge (<i>Carex sempervirens</i> and <i>carex curvula</i>); taller habits with dwarf-shrub (<i>Rhododendron fer.</i>)
Subnival belt (cushion plant-belt)	2700-3000	Individually growing herbaceous plants or low cushion-like habitats

3.2. Soil mapping

A conventional soil map with a scale of 1:10,000 was produced for 3 test areas to obtain more information about the

soils in the region and to increase the existing soil database.

The 3 test areas are located at different expositions and in varying altitudinal zones. They served, furthermore, as a validation of the model's output.

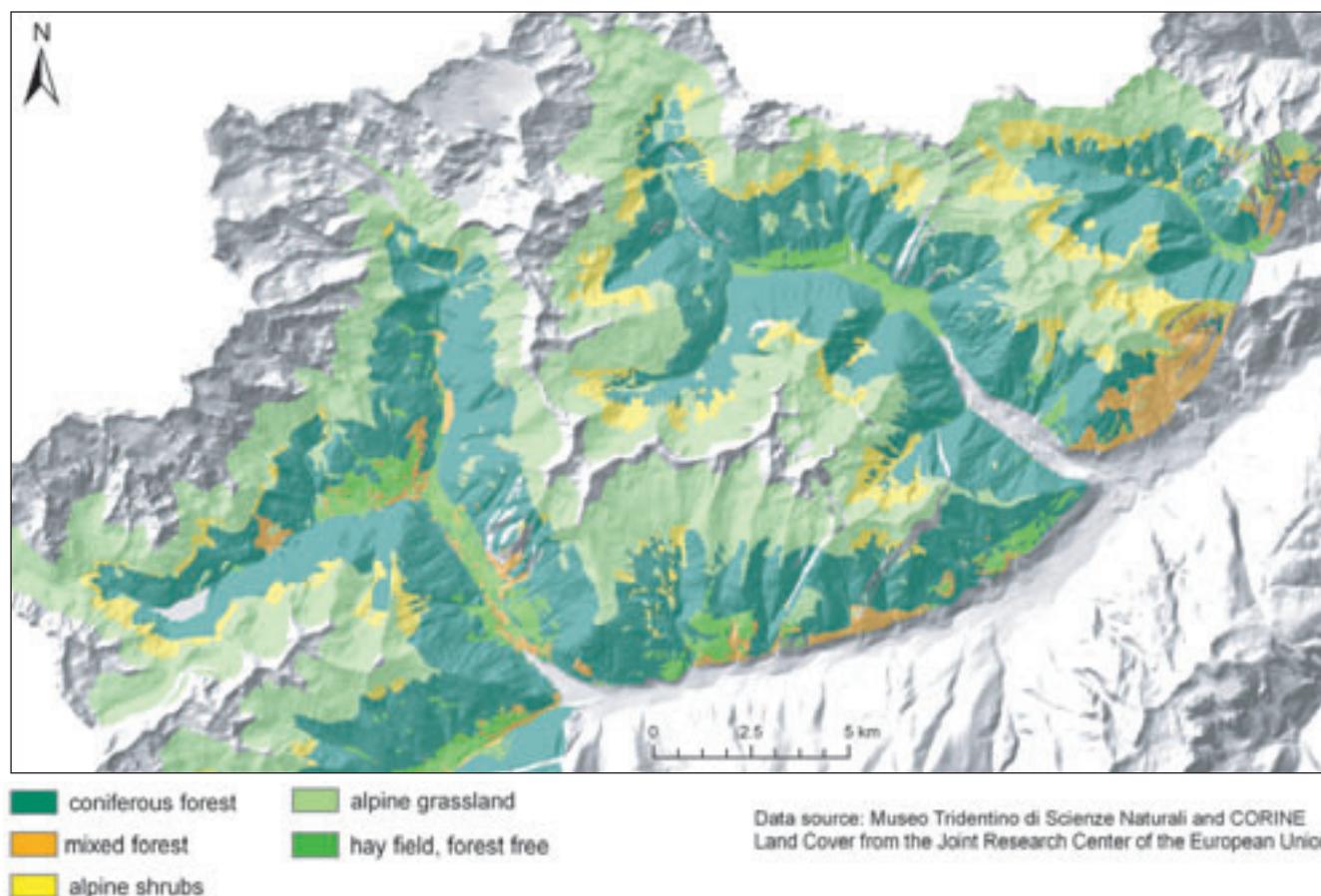


Fig. 3 - Vegetation types in the study area.

Fig. 3 - Tipi vegetazionali nell'area di studio.

3.3. Modelling approach

The available digital datasets related to environmental factors are listed in table 2. Apart from the CORINE Land Cover Data (Nuñez de Lima 2005), the Provincia autonoma di Trento provided all datasets. The digital elevation model (DEM) with a resolution of 10 m and the thematic vector data (vegetation and geology) were converted to raster datasets of the same resolution and projection.

These raster datasets, together with the soil profiles, were the basis for the statistical analyses to build the predictive soil map model.

The DEM provides climatic and topographic information. Altitude and exposure are directly linked to factor climate. Exposure (north and south exposition) and the slope angle were directly derived from the DEM. The profile curvature of the DEM enabled the identification of several landforms. The following landform elements were defined: accumulation areas, erosion areas and regions in equilibrium. Accumulative landforms were characterised by concave, equilibrium landforms by flat and the erosive landforms by convex curvatures.

The geological map provided 36 different categories. These categories had to be reclassified into 5 main pedologically relevant parent materials: 1) granite and gneiss, 2) schists and phyllites, 3) siliceous deposits, 4) amphibolites and chlorite schists and 5) limestones.

The vegetation map had three key vegetation types: 1) unproductive areas in the Alpine belt and summit areas, 2) Alpine meadows and 3) forests. Forests were, furthermore, subdivided into deciduous and coniferous forests. As the available vegetation map does not have the class "Alpine shrubs" and the designation of some forest types was imprecise, CORINE Land Cover data was integrated to overcome this restriction. Figure 3 displays this combined vegetation map.

Soil modelling was performed using a non-parametric classification and decision tree analysis (CART:

Classification and Regression Tree Analysis; see Mertens *et al.* 2002). According to Breiman *et al.* (1984), CART is a hierarchical classification which aims to group elements of a sample in relation to a dependent variable (target variable). Regarding this target variable, the generated groups should be as homogenous as possible – optimally all group members have the same value for the target variable. The grouping or classification (if the target is continuous, then a regression is used) is done by using the independent variables (environmental variables), which can be continuous or discrete. This leads to a binary decision tree with branches, splitting nodes and final leaves (terminal nodes). As CART is an automated statistical method, not all used environmental variables will appear in the dendrogram. They are not used in the resulting dendrogram if they are of no significance. The CART algorithm chooses automatically the values of input-variables which produce a subset of the highest-possible uniformity of a target variable. The so-called split based on the specificity of the input variable with which separation into branches occurred. With this procedure, a decision tree will be formed which corresponds to a classification rule. Every end-node receives a specific class *j* of the target variable. It may happen that the end node has not only one but several classes. In such a case the dominant class (or value) is chosen. The optimally pruned subtrees have to be chosen in that way that the misclassification rate *r(t)* is minimised for the splitting rule *j(t)*. The misclassification rate *r(t)* is given by

$$(1) \quad r(t) = \min_j \sum_i C(i/j)p(j/t)$$

where *C(i/j)* corresponds to the misclassification of an object with the class value *j* as *i*. The probability that an object falls into an end-note *t* and class *j* is given by *p(j/t)*. The allocated class has to be chosen in the way that the expression

$$(2) \quad C(i/j)p(j/t)$$

Tab. 2 - Available digital information about environmental variables used to model soil properties of the region. Projected coordinate system: UTM (Monte Mario, Rome – Italy). *European Commission, Joint Research Centre, Institute for Environment and Sustainability
 Tab. 2 - Informazioni digitali riguardanti le variabili ambientali disponibili per la modellizzazione dei caratteri dei suoli della zona di studio. Sistema di coordinate geografiche: UTM (Monte Mario, Roma). *Commissione Europea, Joint Research Centre, Institute for Environment and Sustainability

Data type	Resolution/Scale	Details
Soil forming factors		
Digital elevation model (DEM)	10x10 m	
Vegetation	1:10,000	Forest types, pasture, unproductive areas
CORINE Land Cover*	1:100,000	15 categories
Geology	1:100,000	36 categories
Catchment area		
Hydrological watersheds	1:10,000	
Soil mapping and orientation		
Hydrology (lakes, rivers)	1:10,000	
Glacier	1:10,000	
Settlement area	1:10,000	
Topographic maps	1:10,000	
Orthofotos	1x1 m	

is minimised. The homogeneity of the node is described by the extended gini-index of diversity with

$$(3) \quad i(t) = \sum_{j,i} C(i/j)p(i/t)p(j/t)$$

CART Pro 6.0 calculates from a sequence of subtrees with varying end-node numbers the optimally pruned subtree. The decision tree structure can, furthermore, manually be influenced by introducing forced splitting rules. Thereby, expert knowledge can be included into the elaboration of the decision and regression tree.

4. RESULTS

4.1. Main soil types and processes

The soils ranged from shallow *Umbric Leptosols* (Duchaufour 2006: rankers) at high altitudes to well-developed *Skeletal Podzols* (Duchaufour 2006: iron-humic podzols) and *Dystric-Chromic Cambisols* (Duchaufour 2006: brown podzolic soils with a clear E horizon, ochric brown soils with an E horizon). Rankers are weakly developed soils with an A-C profile which developed under grass vegetation, initial brunification or podzolisation and a humic A horizon. Enhanced soil development showed the iron-humic podzols and the brown podzolic soils, typically found under forest (coniferous) vegetation. The former have a horizons sequence of E-Bhs-Bs-C and the brown podzolic soil a sequence of AE-E-Bs-C, without any visible illuviation of organic substance into the subsoil. The cryptopodzolic soils, with an OE-Bhs(-Bs)-C horizons sequence, can be considered as a transitional development step between a ranker and a podzol (for a detailed description see also Sartori *et al.* 1997, 2005). *Dystric Cambisols* (acid brown soils) do not show any signs of illuviation.

The analysis of the soil type distribution showed that *Episkeletic Podzols* (iron-humic podzols; Duchaufour 2006) and *Dystric-Chromic Cambisols* (brown podzolic soils with an E horizon) predominantly appear on north facing slopes between 1400 to 1600 m a.s.l. *Enti-Umbric Podzols* (humic ochric brown soils, with a typical AB-horizon) are characteristic for southern exposures at altitudes higher than 1800 m a.s.l. *Enti-Umbric Podzols*, *Skeleti-Entic Podzols* cryptopodzolic soils and brown podzolic soils) were predominantly found in the Alpine dwarf-shrub zone (such as Alpine rose and juniper), just above the timberline.

Rankers dominate in the high-alpine belt. At lower altitudes, they only occur at geomorphically very active sites (e.g. erosion). *Dystric Cambisols* (acid brown soils) are typical for forest-free areas of the montane and subalpine belts. The *Dystric-Chromic Cambisol* (ochric brown soils) is the most widespread soil type in the region. It can predominantly be found in the subalpine and Alpine belts.

Below the limit of 1600 m a.s.l. *Dystric-Chromic Cambisols* (ochric brown soils: AE(A)-Bs-C) and *Dystric Cambisols* (acid brown soils: A-Bw-C) coexist. The ochric brown soils are the most frequent soil type of this zone (39% of the sampled sites). In these soils, podzolisation is weak or completely missing (Sartori *et al.* 1997, see Fig. 4).

Previous studies (Egli *et al.* 2006) showed that the north slopes exhibited higher leaching of elements and consequently a higher weathering intensity. On south-facing sites, intense podzolisation processes were measurable only above 2000 m a.s.l. Furthermore, accumulation of organic substances is greatest close to the timberline (1900-2100 m a.s.l.) regardless of exposition. These measurements agree with the observation of our soil profiles.

The typical range of the most important properties for each soil type and the relative distribution of the values are given in tables 3 and 4. These attributes clearly reflect the stage of soil development. In table 5 the most frequent value range of individual soil properties is assigned to the soil types.

4.2. Soil modelling

Independent variables were used to define the split-criteria of the data into the left and right branches at the nodes. The dendrogram divides the data in groups which are as homogenous as possible regarding the variable "soil type" (Fig. 5). The altitude acts as a splitting criteria at the root node. A part of the sample branches to the right and the other to the left. Subsequent splits occur by other variables such as vegetation, aspect and again altitude. This splitting procedure results in a tree with 15 terminal nodes. Because the geology is quite similar in the whole region, only five of the six variables are used in the dendrogram: altitude, aspect, slope, landform and vegetation.

The constructed algorithm with the detailed split-criteria is then implemented in GIS and results in the predictive distribution of the soil types in the region. Additionally, the modelling of the spatial distribution of the humus form is done in a similar way. The resulting dendrogram for the modelling of humus forms has nine terminal nodes. The humus forms are determined using the variables vegetation, aspect, altitude and soil types. Because the soil type is considered as an important variable for humus modelling, the implementation of the algorithm in the GIS requires the previous modelling of the soil types in the study area.

As modelling of soil types and corresponding characteristics is bound to a likelihood and therefore to errors (see below), the modelled soil map is called the "hypothesis map". The hypothesis map for soil types is given in figure 6. This map shows that the class *Umbric Leptosols* (ranker according to Duchaufour (2006)) is found in about 29% of the whole area (Tab. 6). The *Enti-Umbric Podzols* (humic ochric brown soils) comprise about 19% of the whole area, whereas the other soil groups (except the class "no soil") have a more or less similar distribution with 11-15% of the whole area.

Using the 3 mapped test areas, the accuracy of the model approach for soil types could be measured. The three test sites included 2 subalpine sites (Val di Peio, Val di Rabbi) below the timberline and one close and above the timberline. One of the two subalpine sites (Val di Peio) was subjected to anthropogenic impact (grazing, erosion), while the other (Val di Rabbi) was an almost natural site. By comparing the modelled area of soil types with the mapped ones in the test areas, the accuracy of the model could be calculated. This accuracy was calculated on the base of a modelled value which matches to 100% with a measured one. Minor deviations are for this purpose not taken into

Tab. 3 - Characteristics of the different soil types. The number of observations and the corresponding relative distribution for the characteristics soil depth, skeleton content in the topsoil and the subsoil are shown. *Soil depth relevant for plant growth: soil depth minus skeleton content **BA= acid brown soils; BO= ochric brown soils; BOe= ochric brown soils with an E horizon; BOu= humic ochric brown soils; PU= iron-humic podzols; OP= brown podzolic soils; RA= ranker.

Tab. 3 - Caratteri dei differenti tipi di suolo. Sono indicati il numero totale di osservazioni e la distribuzione relativa alle varie classi, per profondità del suolo, contenuto di scheletro nel topsoil e nel subsoil. *Profondità del suolo rilevante per la crescita della pianta: profondità del suolo meno contenuto di scheletro **BA= suoli bruni acidi; BO= suoli bruni ocrici; BOe= suoli bruni ocrici con orizzonte E; BOu= suoli bruni ocrici umiferi; PU= podzol umoferrici; OP= suoli ocra podzolici, RA= ranker.

Soil types		Observations number (n) distribution (%)	Soil depth (cm)*					Skeleton content topsoil (weight %)					Skeleton content subsoil (weight %)						
WRB (FAO 1998)	Duchaufour 2006 (Sartori et al. 2005)**		< 10	10-30	30-50	50-70	n/%	< 1	1-5	5-15	15-35	35-70	n/%	1-5	5-15	15-35	35-70	> 70	n/%
<i>Dystric Cambisols</i>	Acid brown soils (BA)	n %	0 0	4 27	10 67	1 7	15 100	4 31	5 38	2 15	2 15	0 0	13 100	2 14	2 14	7 50	2 14	1 7	14 100
<i>Dystric-Chromic Cambisols</i>	Ochric brown soils (BO)	n %	0 0	9 31	14 48	6 21	29 100	7 25	8 29	7 25	5 18	1 4	28 100	2 7	2 7	6 21	18 62	1 3	29 100
<i>Dystric-Chromic Cambisols, Episkeletic Podzols</i>	Ochric brown soils with E horizon (BOe), Iron-humic podzols (PU)	n %	0 0	1 9	9 82	1 9	11 100	3 27	2 18	3 27	2 18	1 9	11 100	0 0	1 9	1 9	8 73	1 9	11 100
<i>Enti-Umbic Podzols</i>	Humic ochric brown soils (BOu)	n %	0 0	0 0	1 25	3 75	4 100	1 25	0 0	3 75	0 0	4 100	0 0	0 0	1 25	1 25	2 50	4 100	
<i>Enti-Umbic Podzols, Skeleti-Entic Podzols</i>	Cryptopodzolic soils (RPu) Brown podzolic soils (OP)	n %	0 0	7 70	2 20	1 10	10 100	1 10	4 40	4 40	0 0	1 10	10 100	1 10	1 10	1 10	6 60	1 10	10 100
<i>Umbic Leptosols</i>	Rankers (RA)	n %	3 23	9 69	1 8	0 0	13 100	1 8	1 8	7 54	3 23	1 8	13 100	1 8	1 8	1 8	10 77	0 0	13 100
Total			3	30	37	12	82	17	20	26	12	4	79	6	7	17	45	6	81
			4	37	45	15	100	22	25	33	15	5	100	7	9	21	56	7	100

Umbic Leptosol (ranker) Distri-Chromic Cambisol (ochric brown soils) Chromi-Episkeletic Cambisol (ochric brown soil with E) Episkeletic Podzol (iron humic podzol)

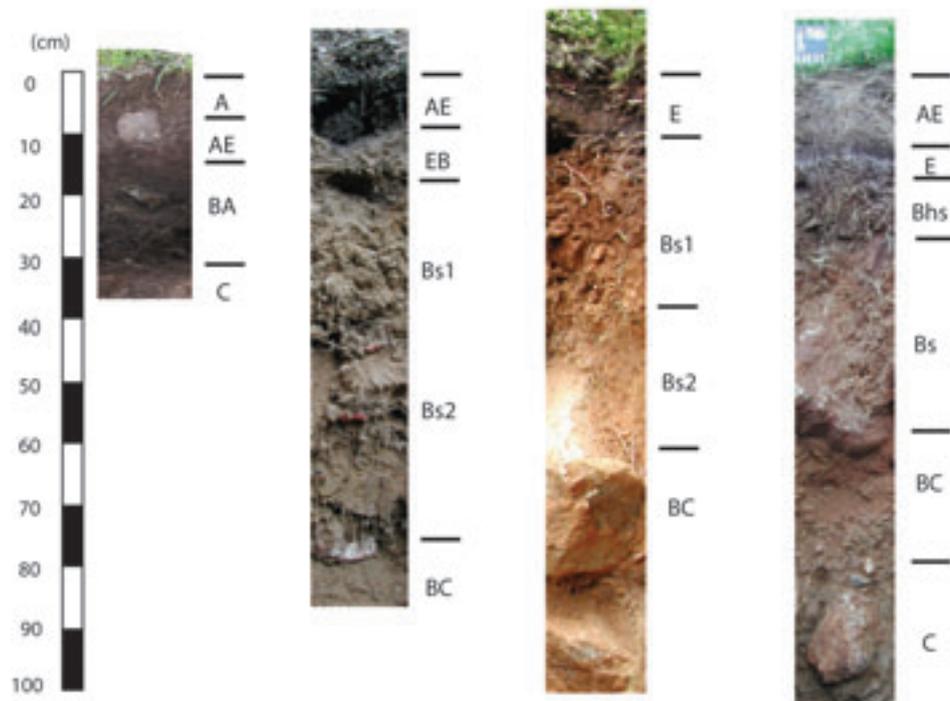


Fig. 4 - Photographs of some selected soil profiles in the investigation area: *Umbic Leptosol* (Lavina Rossa, 2380 m a.s.l.), *Distri-Chromic Cambisol* (Favari, Val di Rabbi, 1180 m a.s.l.), *Chromi-Episkeletic Cambisol* (Fonti di Rabbi, Val di Rabbi, 1620 m a.s.l.), *Episkeletic Podzol* (sotto Malga Tremenesca, Val di Rabbi, 1910 m a.s.l.).

Fig. 4 - Fotografie di alcuni profili tipici dell'area di studio: *Umbic Leptosol* (Lavina Rossa, 2380 m s.l.m.), *Distri-Chromic Cambisol* (Favari, Val di Rabbi, 1180 m s.l.m.), *Chromi-Episkeletic Cambisol* (Fonti di Rabbi, Val di Rabbi, 1620 m s.l.m.), *Episkeletic Podzol* (sotto Malga Tremenesca, Val di Rabbi, 1910 m s.l.m.).

Tab. 4 - Acidity classes (number of observations and relative proportion) of the topsoil and the subsoil as a function of the different soil types. *BA= acid brown soils; BO= ochric brown soils; BOe= ochric brown soils with an E horizon; BOu= humic ochric brown soils; PU= iron-humic podzols; OP= brown podzolic soils; RA= ranker.

Tab. 4 - Classi di acidità (numero totale di osservazioni e proporzione relativa di ogni classe) del topsoil e del subsoil nei differenti tipi di suolo. *BA= suoli bruni acidi; BO= suoli bruni ocrici; BOe= suoli bruni ocrici con orizzonte E; BOu= suoli bruni ocrici umiferi; PU= podzol umoferrici; OP= suoli oca podzolici, RA= ranker.

Soil types WRB (FAO 1998)	Duchaufour 2006 (Sartori <i>et al.</i> 2005)*	Observations number (n) distribution (%)	pH (CaCl ₂) topsoil			pH (CaCl ₂) subsoil		
			< 3.3	3.3-4.2	n _{tot} / %	< 4.3	4.3-5.0	n _{tot} / %
<i>Dystric Cambisols</i>	Acid brown soils (BA)	n %	1 50	1 50	2 100	2 100	0 0	2 100
<i>Dystri-Chromic Cambisols</i>	Ochric brown soils (BO)	n %	4 67	2 33	6 100	5 83	1 17	6 100
<i>Dystri-Chromic Cambisols, Episkeletic Podzols</i>	Ochric brown soils with E horizon (BOe), Iron-humic podzols (PU)	n %	5 56	4 44	9 100	5 56	4 44	9 100
<i>Enti-Umbric Podzols</i>	Humic ochric brown soils (BOu)	n %	0 0	4 100	4 100	0 0	4 100	4 100
<i>Enti-Umbric Podzols, Skeleti-Entic Podzols</i>	Cryptopodzolic soils (RPu), Brown podzolic soils (OP)	n %	1 20	4 80	5 100	3 60	2 40	5 100
<i>Umbric Leptosols</i>	Rankers (RA)	n %	2 33	4 67	6 100	6 100	0 0	6 100
		Total n	13	19	32	21	11	32
		%	41	59	100	66	34	100

consideration and do not contribute to the accuracy. The accuracy for the modelled soil types varied considerably: an overall accuracy of 93.1% was obtained for the subalpine and quasi-natural area, 57.1% for the anthropogenically influenced, subalpine area and only 43.2% for the high-

alpine area. Around 65% of the whole are have been, thus, modelled correctly.

The soil classes BO (*Dystri-Chromic Cambisols* / ochric brown soils) and BOe/PU (*Dystri-Chromic Cambisols, Episkeletic Podzols* / ochric brown soils with E horizon, iron-

Tab. 5 - Soil characteristics from sample data (Tabs 3-4) related to the soil types. The modal values (most frequent) were assigned to the specific soil types. ¹Soil depth relevant for plant growth= profile depth minus skeleton content; ²weight - %; TS= Topsoil (all horizons with characteristics of an A or E); SS= Subsoil (all horizons with characteristics of a B).

Tab. 5 - Caratteri dei suoli in relazione al tipo di suolo. A ciascun tipo di suolo sono attribuiti i valori modal. ¹Profondità del suolo rilevante per la crescita della pianta: profondità del suolo meno contenuto di scheletro; ²peso - %; TS= Topsoil (orizzonti A o E); SS= Subsoil (orizzonti B).

Soil types WRB (FAO 1998)	Duchaufour 2006 (Sartori <i>et al.</i> 2005)*	Soil depth ¹ (cm)	Thickness TS (cm)	Skeleton TS ² (%)	Skeleton SS ² (%)	pH TS (CaCl ₂)	pH SS (CaCl ₂)
<i>Dystric Cambisols</i>	Acid brown soils (BA)	30-50	3-6	0-5	15-35	< 3.3-4.2	< 4.3
<i>Dystri-Chromic Cambisols</i>	Ochric brown soils (BO)	30-50	4.5-10	0-15	35-70	< 3.3	< 4.3
<i>Dystri-Chromic Cambisols, Episkeletic Podzols</i>	Ochric brown soils with E horizon (BOe), iron-humic podzols (PU)	30-50	8-12	0-15	35-70	< 3.3-4.2	< 4.3-5.0
<i>Enti-Umbric Podzols</i>	Humic ochric brown soils (BOu)	50-70	7-20	5-15	> 70	3.3-4.2	4.3-5.0
<i>Enti-Umbric Podzols, Skeleti-Entic Podzols</i>	Cryptopodzolic soils (RPu), brown podzolic soils (OP)	10-30	9.5-19	1-15	35-70	3.3-4.2	< 4.3
<i>Umbric Leptosols</i>	Rankers (RA)	0-30	4-9	5-15	35-75	3.3-4.2	< 4.3

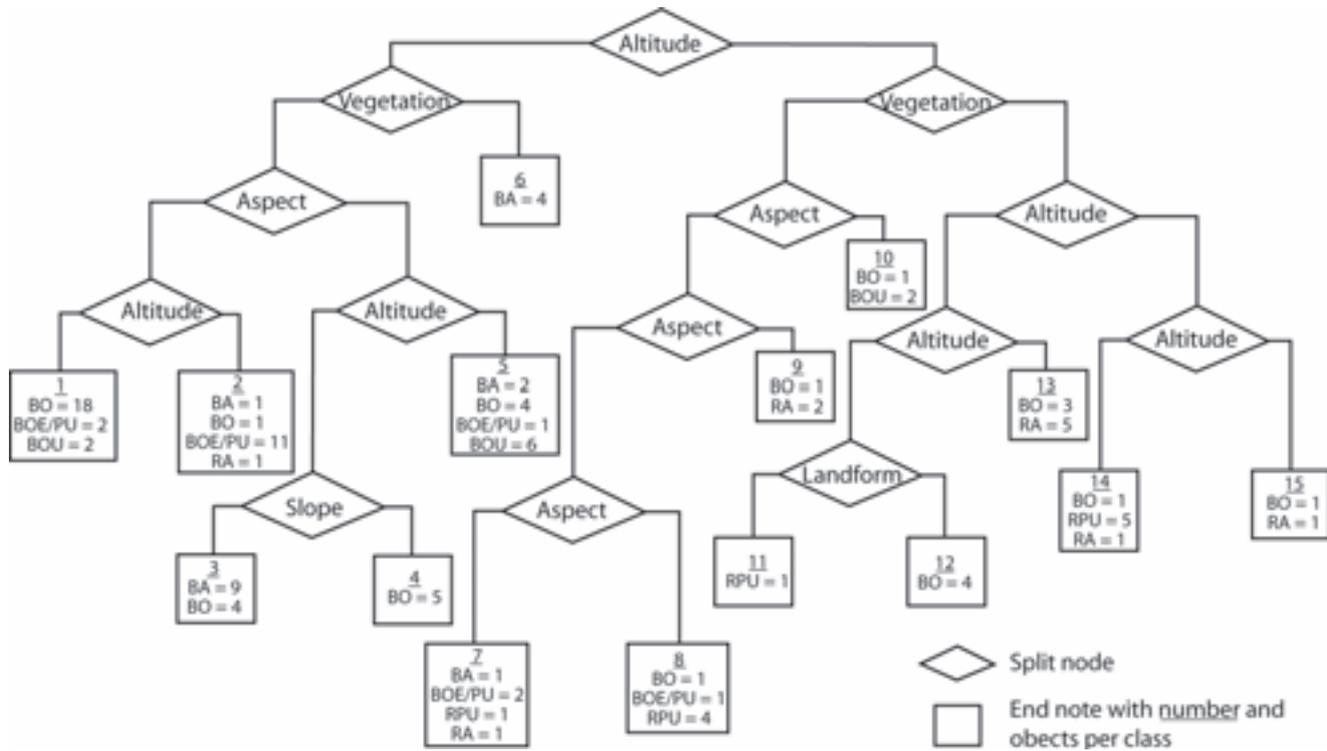


Fig. 5 - Decision tree for modelling the spatial distribution of the soil types.
 Fig. 5 - Albero decisionale per la modellizzazione della distribuzione spaziale dei tipi di suolo.

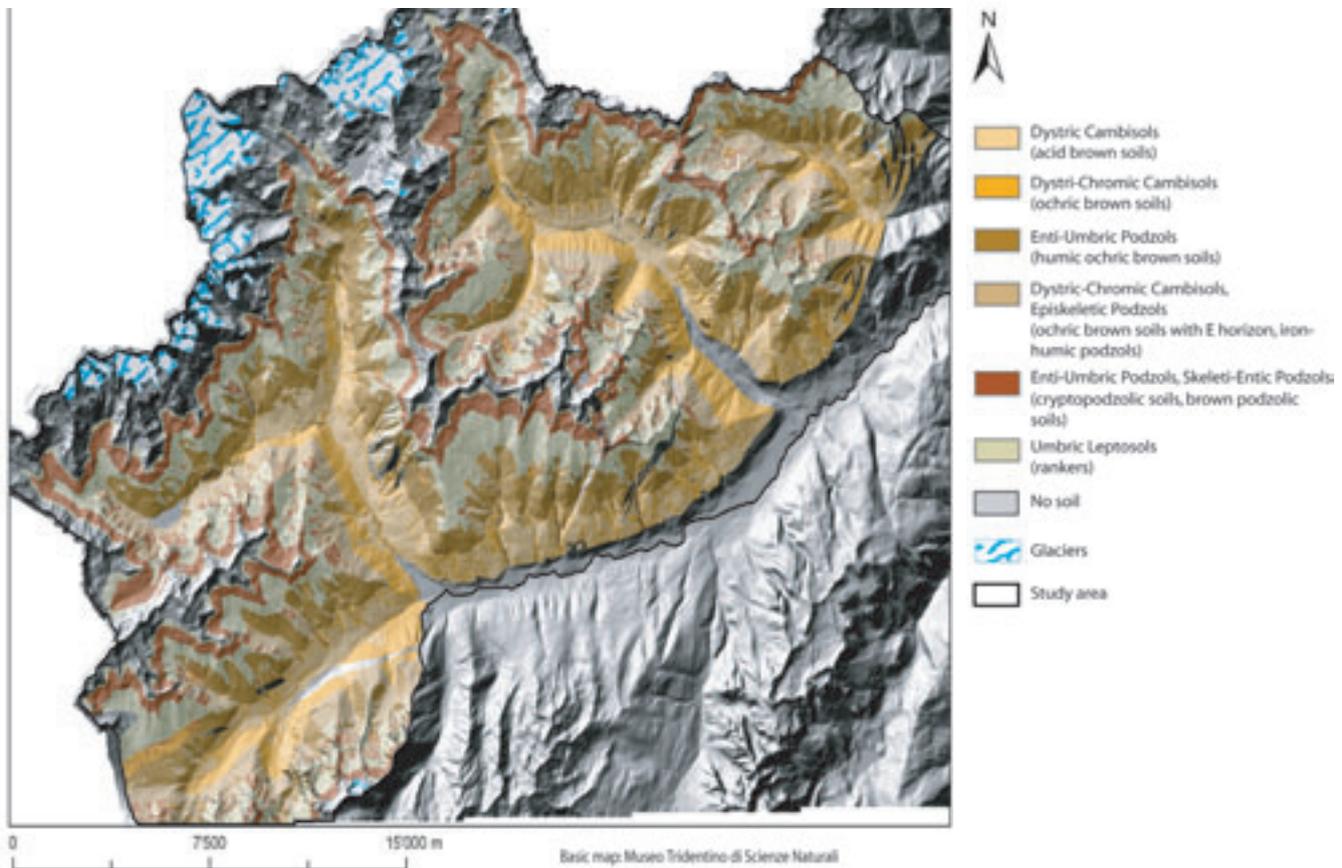


Fig. 6 - Modelled distribution of soil types (hypothesis map) for Val di Sole, Val di Rabbi and Val di Peio using a classification tree having 15 terminal nodes.
 Fig. 6 - Distribuzione spaziale dei tipi di suolo in Val di Sole, Val di Rabbi e Val di Peio, ottenuta un albero decisionale con 15 nodi terminali.

Tab. 6 - Area statistics of the modelled soil types (see Fig. 6). *BA= acid brown soils; BO= ochric brown soils; BOe= ochric brown soils with an E horizon; BOu= humic ochric brown soils; PU= iron-humic podzols; OP= brown podzolic soils; RA= ranker.

Tab. 6 - Statistiche areali della distribuzione spaziale dei differenti tipi di suolo ottenuta dal relativo modello (si veda Fig. 6). *BA = suoli bruni acidi; BO= suoli bruni ochrici; BOe= suoli bruni ochrici con orizzonte E; BOu= suoli bruni ochrici umiferi; PU= podzol umoferrici; OP= suoli oca podzolici, RA= ranker.

Modelled soil type/ soil class		Area in km ²	Area in %
WRB 1998	Duchaufour 2006 (Sartori <i>et al.</i> 2005)*		
<i>Dystric Cambisols</i>	Acid brown soils (BA)	44	11.8
<i>Dystri-Chromic Cambisols</i>	Ochric brown soils (BO)	49.5	13.2
<i>Dystri-Chromic Cambisols, Episkeletic Podzols</i>	Ochric brown soils with E horizon (BOe), Iron-humic podzols (PU)	40.3	10.8
<i>Enti-Umbric Podzols</i>	Humic ochric brown soils (BOu)	70.8	18.9
<i>Enti-Umbric Podzols, Skeleti-Entic Podzols</i>	Cryptopodzolic soils (RPu) Brown podzolic soils (OP)	55.4	14.8
<i>Umbric Leptosols</i>	Rankers (RA)	107.5	28.8
No soil		6.3	1.7
Total		373.8	100

Tab. 7 - Area statistics of the modelled humus types (see also Fig. 7).

Tab. 7 - Statistiche areali della distribuzione spaziale delle differenti forme di humus ottenuta dal relativo modello (si veda Fig. 7).

Modelled humus form	Area in km ²	Area in %
Mull	86.6	23.2
Moder	246.2	65.8
Mor	34.9	9.3
No soil	6.3	1.7
Total	373.8	100.0

humic podzols) were modelled with an accuracy of more than 70% in the high-alpine zone. The matches for the soil class BA (*Dystric Cambisols*) were in the high-alpine zone, however, extremely low. The model, therefore, does not reflect this soil class accurately in high-alpine zones. The high variability of landforms and the patch-wise development of soils in the high-alpine area may be causes for the less accurate modelling in this zone.

The distribution of the modelled humus types show a clear dominance of the moder (Fig. 7, Tab. 7). The moder humus type is found in about two third of the investigation area. The variability of the accuracy of the modelled humus types varies between 35 and 100%. As an average, 78.5% of the soil profiles was correctly modelled. In contrast to the modelled soil types, the lowest accuracy was measured in the anthropogenically influenced area in Val di Peio. The accuracy of the humus model for high-alpine sites is better than for the soil types. Erosion processes in the Val di Peio test area obviously had a major impact on the humus form and consequently on the accuracy of the model.

5. DISCUSSION

Inductive models can have different statistical methods as a basis, depending on the type of the contributing

variables (nominal, ordinal, interval or ratio scale) and the sample size. Most of the methods such as linear regression, linear discriminant analysis and logistic discriminant analysis demands linearity of the relationship between soil and environmental variables and normal distribution of the data, and therefore requires transformation of variables (McBratney *et al.* 2003; Scull *et al.* 2003). Generalised linear models (GLMs), however, do not need such a transformation as they rather intend to transform the model and not the data (McBratney *et al.* 2003). All these methods and models have in common that already existing expert knowledge cannot be integrated (Scull *et al.* 2003) and sample size (number of profile sites) has to be large. Another statistically based method is the non-parametric decision tree analysis (DTA), although called classification and regression trees (CART). Unlike the GLMs and the logistic regression, the results of this approach can be more easily interpreted, a smaller data base is necessary and expert knowledge can be implemented (McBratney *et al.* 2003).

Modelling of soil distribution is a challenging task, especially in mountain areas where rugged topography leads to soil changes within very short distances. Similar attempts include work by Kägi (2006) in the Swiss National Park where soil distribution was modelled using a fuzzy-logic approach. About 60% of the profile sites (regarding soil type, pH etc) were accurately modelled. An accuracy of approximately 70% was obtained using a more heuristic-statistical method as a basis for a decision tree for soil modelling in the Upper Engadine (Egli *et al.* 2005). The model applied in Egli *et al.* (2005) had the disadvantage of not being automated. Methodologically comparable studies include those from Behrens *et al.* (2005) and Lagacherie & Holmes (1997). These studies, however, were not done in an Alpine environment. Within a test area in Rheinland-Palatinate (Germany), covering an area of about 600 km², a digital soil map was predicted (Behrens *et al.* 2005). The overall precision in the training area was 70%. In Languedoc (southern France), Lagacherie & Holmes (1997) also used the CART method for soil modelling. With eight end nodes, they were able to achieve an accuracy of 74%.

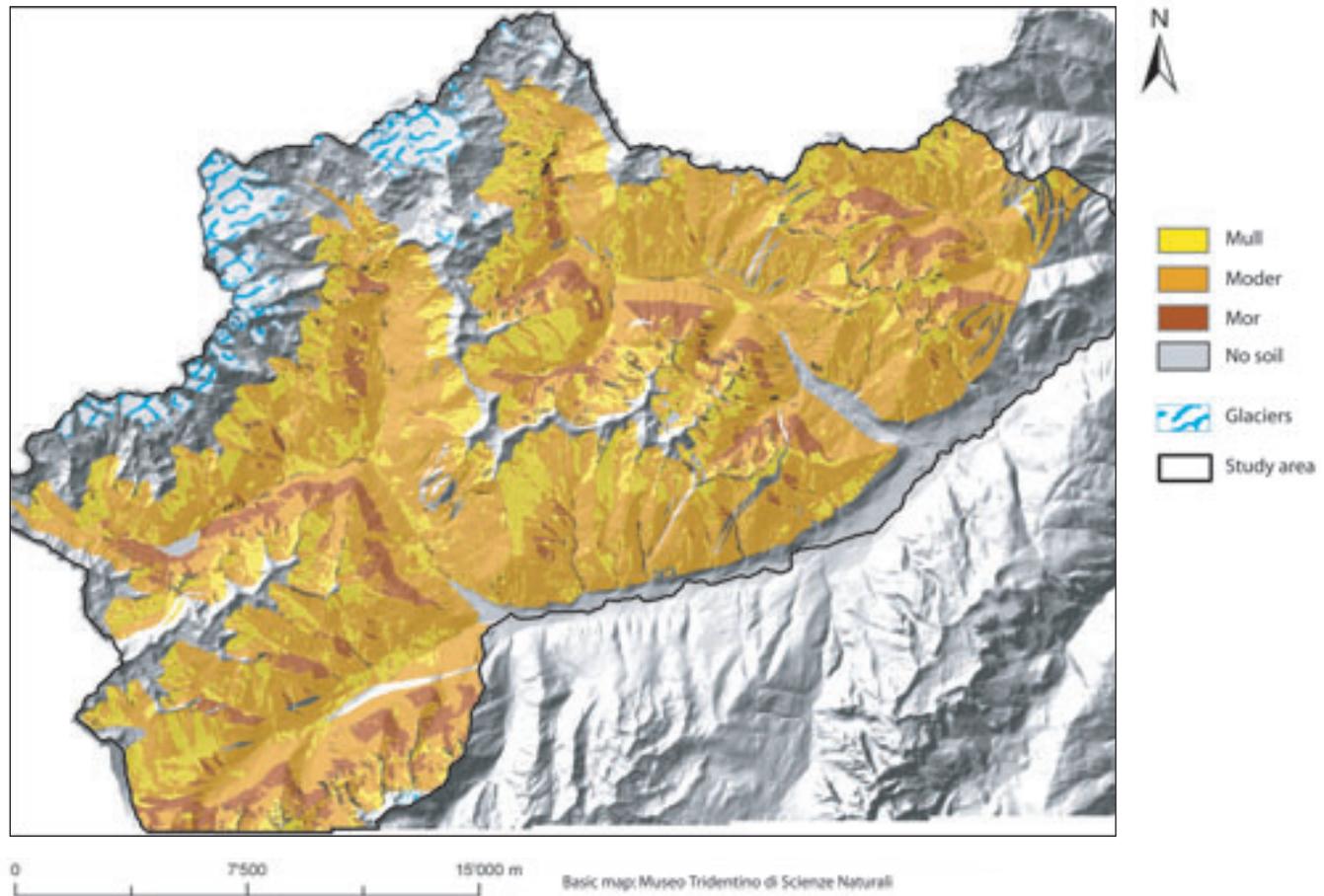


Fig. 7 - Modelled distribution of humus forms in Val di Sole, Val di Rabbi and Val Peio.

Fig. 7 - Distribuzione spaziale delle forme di humus in Val di Sole, Val di Rabbi e Val Peio, ottenuta dal modello (albero decisionale) messo a punto.

The obtained results for the rugged Alpine area in Val di Sole, Val di Rabbi and Val di Peio are therefore comparable with accuracies obtained in other Alpine and non-Alpine areas. Not only soil types, but also other soil properties such as the humus forms can be modelled rather easily. The obtained hypothesis map is appropriate enough to be used also for more general, practical purposes and also for a detailed, local field-based soil mapping.

One general constraint is the underestimation of minor soil types when survey lines are sparse. This constraint influences the direct use of stimulated results when survey data are too sparse and when the minor soil types are of serious importance (see also Li *et al.* 2004).

The soils of the investigated area have developed mostly on acid siliceous materials. The linkages between Alpine soil types developed on these materials and environmental parameters (i.e., altitude, aspect, vegetation) are generally strong (Egli *et al.* 2005; Sartori *et al.* 2005). This could explain the relatively high overall accuracy in our study.

6. CONCLUSIONS

We used the statistically based, non-parametric decision tree analysis. This procedure enabled also the inclusion

of expert knowledge. Using this approach, we obtained the following main findings:

- depending on the feature to be modelled, a mean accuracy of 65% (spatial distribution of soil types) or higher (humus forms) was achieved, which is in a similar range to studies in a less rugged topography;
- the used approach is in large parts automated, can be applied over large area and also allows the application of forced splitting rules (according to expert knowledge);
- the main soil type in the investigation area is ranker (*Umbric Leptosol*), which covers about 28% of the whole area;
- the other classes (*Umbric Podzols*, *Cambisols*) cover each about 11-15% of the area;
- the most frequent humus type is moder, which can be found in about 65% of the area;
- soil modelling does not replace soil mapping in the field. The obtained map is a hypothesis map and serves as a basis for further, local investigations.

ACKNOWLEDGEMENTS

We would like to express our appreciation to the Museo Tridentino di Scienze Naturali and the Dipartimento di

Protezione Civile e Tutela del Territorio (Ufficio Previsioni e Organizzazione, Provincia Autonoma di Trento) for providing basic GIS datasets and B. Kägi for his assistance in the laboratory.

REFERENCES

- Behrens T., Förster H., Scholten T., Steinrücken U., Spies E.-D. & Goldschmitt M., 2005 - Digital soil mapping using artificial neural networks. *J. Plant Nutr. Soil Sci.*, 168: 21-33.
- BGS (Bodenkundliche Gesellschaft der Schweiz) & FAL (Eidgenössische Forschungsanstalt für Agrarökologie und Landbau), 2002 - *Klassifikation der Böden der Schweiz*. FAL, Zürich-Reckenholz: 94 pp.
- Breiman L., Friedman J.H., Ohlsen R.A. & Stone C.J., 1984 - *Classification and Regression Trees*. Pacific Grove, Wadsworth: 358 pp.
- Carré F., McBratney A.B., Mayr T. & Montanarella L., 2007 - Digital soil assessments: Beyond DSM. *Geoderma*, 142: 69-79.
- Duchaufour P., 2006 - *Introduction à la science du sol. Sol, végétation, environnement*. 6th ed. Dunod, Paris: 331 pp.
- Egli M., Margreth M., Vökt U., Fitze P., Tognina G. & Keller F., 2005 - Modellierung von Bodeneigenschaften im Oberengadin mit Hilfe eines GIS. *Geogr. Helv.*, 60: 87-96.
- Egli M., Mirabella A., Sartori G., Zanelli R. & Bischof S., 2006a - Effect of north and south exposure on weathering rates and clay mineral formation in Alpine soils. *Catena*, 67: 155-174.
- Egli M., Wernli M., Kneisel C., Biegger S. & Haerberli W., 2006b - Melting Glaciers and Soil Development in the Proglacial Area Morteratsch (Swiss Alps). II. Modelling the Present and Future Soil State. *Arct. Antarct. Alp. Res.*, 38: 510-521.
- FAO (Food and Agriculture Organization of the United Nations), 1998 - *World Reference Base for Soil Resources. WRB*. FAO (World Soil Resources Reports, 84), Rome: 142 pp.
- Jenny H., 1980 - *The soil resource*. Springer, New York: 377 pp.
- Kägi C., 2006 - *Modellierung von Bodeneigenschaften mittels Fuzzy-Logik im Gebiet des Schweizerischen Nationalparks*. Diploma thesis (unpublished), Department of Geography, University of Zürich: 144 pp.
- Lagacherie P. & Holmes S., 1997 - Addressing geographical data errors in a classification tree for soil unit prediction. *Int. J. Geogr. Inf. Sci.*, 11: 183-198.
- Landolt E., 1992 - *Unsere Alpenflora*. Gustav Fischer Verlag, Stuttgart - Jena: 320 pp.
- Li W., Zhang C., Burt J.E., Zhu A.-X. & Feyen J., 2004 - Two-dimension Markov chain simulation of soil type spatial distribution. *Soil Sci. Soc. Am. J.*, 68: 1479-1490.
- McBratney A.B., Mendocça Santos M.L. & Minasny B., 2003 - On digital soil mapping. *Geoderma*, 117: 3-52.
- Mertens M., Nestler I. & Huwe B., 2002 - GIS-based regionalization of soil profiles with Classification and Regression Trees (CART). *J. Plant Nutr. Soil Sci.*, 165: 39-43.
- Nuñez de Lima M.V., 2005 - *IMAGE2000 and CLC2000 - Products and Methods*. European Commission, Joint Research Centre, Institute for Environment and Sustainability, Ispra, 150 pp.
- Rahmann S., Munn L.C., Vance G.F. & Arneson C., 1997 - Wyoming rocky mountain forest soils: mapping using ARC/INFO Geographic Information System. *Soil Sci. Soc. Am. J.*, 61: 1730-1737.
- Sartori G., Wolf U., Mancabelli A. & Corradini F., 1997 - Principali tipi di suoli forestali nella Provincia di Trento. *Studi Trent. Sci. Nat., Acta Geol.*, 72: 41-54.
- Sartori G., Mancabelli A., Wolf U. & Corradini F., 2005 - *Atlante dei suoli del Parco Naturale Adamello-Brenta. Suoli e paesaggi*. Museo Tridentino di Scienze Naturali, Trento: 239 pp. (Monografie del Museo Tridentino di Scienze Naturali, 2).
- Scull P., Franklin J., Chadwick O.A. & McArthur D., 2003 - Predictive soil mapping: a review. *Prog. Phys. Geog.*, 27: 171-97.
- Seidlein C. von, 2000 - *Petrographie und Struktur des Ostalpinen Altkristallins südlich des Ultentales (Trentino, Nord-Italien)*. Dissertation, Fakultät für Geowissenschaften der Ludwig-Maximilians-Universität München: 122 pp.
- Tognina G., 2004 - Hilfsmittel Bodeninformationssystem und Bodenkarte: Methodik, Realisierbarkeit, Anwendungspotential am Beispiel eines Gebirgskantons. *BGS Bulletin*, 27: 49-56.
- Wielemaker W.G., de Bruin S., Epema G.F. & Veldekamp A., 2001 - Significance and application of the multi-hierarchical land system in soil mapping. *Catena*, 43: 15-34.
- Zhu A.X., Hudson B., Burt J., Lubich K. & Simonson D., 2001 - Soil Mapping Using GIS, Expert Knowledge, and Fuzzy Logic. *Soil Sci. Soc. Am. J.*, 65: 1463-1472.