A fuzzy logic slope-form system for predictive soil mapping of a landscape-scale area with strong relief conditions

Citation:

Bui, Le Vinh; Stahr, Karl; Clemens, Gerhard. 2017. A fuzzy logic slope-form system for predictive soil mapping of a landscape-scale area with strong relief conditions. Catena 155: 135-146.

Publisher's DOI:

http://dx.doi.org/10.1016/j.catena.2017.03.001

Access through CIAT Research Online:

http://hdl.handle.net/10568/80475

A fuzzy logic slope-form system for predictive soil mapping of a landscape-scale area with strong relief conditions

Bui Le Vinh[a,b,c*], Karl Stahr[b], Gerhard Clemens[b]

[a] Department of Land Management, Faculty of Land Management, Vietnam National University of Agriculture, Trau Quy, Gia Lam, Hanoi, Vietnam, email: v.bui@cgiar.org, bui_le_vinh@yahoo.com.

[b] Institute of Soil Science and Land Evaluation (310), University of Hohenheim, 70593 Stuttgart, Germany, email: v.bui@cgiar.org, bui_le_vinh@yahoo.com, Karl.Stahr@uni-hohenheim.de, GeCle@gmx.net

[c] International Center for Tropical Agriculture (CIAT), regional office for Asia, Institute of Agricultural Genetics, Km2 Pham Van Dong Street, Tu Liem, Hanoi, Vietnam, email: v.bui@cgiar.org, bui_le_vinh@yahoo.com.

* Corresponding author. International Center for Tropical Agriculture (CIAT), regional office for Asia, Institute of Agricultural Genetics, Km2 Pham Van Dong Street, Tu Liem, Hanoi, Vietnam. Phone: +84 4 375 76969. Email: v.bui@cgiar.org, bui_le_vinh@yahoo.com (L.V. Bui).

Abstract

We studied an improved slope form system using a fuzzy logic method to assess and map soil fertility of a mountain region in northern Vietnam that has strong relief conditions. The lack of good soil mapping techniques in Vietnam has brought about insufficient soil information, which often leads to false recommendations for land use and crop planning. The reviewed literature describes soil-mapping techniques using fuzzy logic method, but all of them are applied for mapping areas that have gentle relief conditions that are unlikely to be applicable to the mountainous soils in our study area. In this paper, we introduce a detailed slope form system that significantly describes the complexity of terrain characteristics of the area to be mapped, and provides more detail about the variability of the soil fertility of the area. Nine basic slopeforms were used to characterize for each of upper-, middle-, and foot slope positions, making the list 27 slopeforms. Together with crest and valley, the total unit number is 29. We investigated soils of the area and classified them into the major soil groups and calculated soil property indices for all of them. We identified four major environmental parameters affecting soil formation and soil quality: geology, elevation, slope inclination and land use. The findings indicated that soil fertility differs at slope positions. Soils located at upper slope positions, where agricultural activity only started recently, are more fertile than those found at middle slope positions. Soils located at foot slope positions, where eroded sediments accumulate, also have high levels of fertility compared to those on the middle slope. The improved slope form system then became an important additional environmental parameter for this soil mapping work. At a same comparable category, i.e. slope position, geology, soil group, elevation, slope gradient, straight slopeforms are an indicator for better soil fertility compared to convex and concave forms. Although the findings could not specify soil fertility variability for all 29 slopeforms, they did emphasize the major differences in soil fertility and soil formation based on three major forms of convex, straight and concave, with other factors taken into account, such as slope inclination, geology and elevation. We expect our results to be used by scientists and local authorities in deriving more effective land use and crop options for land use management strategies for the northern Vietnam's mountain regions.

Keywords: fuzzy logic, relief conditions, slopeform, slope position, soil mapping, soil fertility assessment

1. Introduction

The need for achieving spatial soil information and soil fertility details for research and development purposes has led to the generation of advanced soil mapping techniques (Zhu et al., 1996; Schuler, 2008; Qin et al., 2012). These techniques have been widely used in soil science for studying the spatial distribution of soils (Zhu, A.X., 1997a; Schuler et al., 2010), soil properties (Zhu et al., 1997b; Batjes, 2008; Qin et al., 2013), land evaluation and land-use planning (Herrmann et al., 2001), and soil and land health (Vågen and Winowiecki, 2012; Winowiecki et al., 2015). Fuzzy logic-based mapping techniques have long been applied for mapping spatial soil distribution and soil fertility studies (Burrough, 1989; McBratney et al., 2003; Zhu et al., 2010). The application of these techniques can be implemented in two approaches: data driven and knowledge-based (Zhu et al., 2010). The former is applied when there are sufficient data. The latter is normally used when there are not enough data for mapping soils of a larger area, for example at regional scale. In the latter case, when a large amount of field soil samples cannot be obtained to sufficiently map soils of a landscape-scale area, local soil experts are needed to describe knowledge on soil-environment relationship. To acquire this relationship, fuzzy membership functions defined by soil-environment relationships must be constructed to map spatial continuity of soils. Unique combinations of environmental variables characterize the formation of each soil group are obtained via various purposive soil sampling strategies (Bui et al., 1999; Qi and Zhu, 2003).

One of the much studied environmental variables influencing soil occurrences is transitions among slope positions over landscape: the gradation of positions from crest, upper slope, middle slope, foot slope, to valley along a catena. This knowledge is very important in studying spatial distribution of soils with geomorphology-prone formation and especially soil fertility variability across various slope positions. However, slope gradation has not been studied and quantified thoroughly for all geomorphological types, for example:

- Studies using crisp classification of slope positions (Young, 1972; Conacher and Dalrymple, 1977; Speight, 1990) do not successfully depict the continuity of slope positions because they only assign binary values to map objects.
- Schuler et al. (2010) used a Maximum Likelihood approach to map a Thai region with strong relief conditions, but did not sufficiently capture effects of slopeforms on hillslopes (only identified convex, concave and linear forms) to formation of soils and their fertility variability.
- Qin et al. (2009) used fuzzy logic to derive nine gradual slope positions in studying spatial soil distribution and soil variability. However, this proved successfully only for an area of $60km^2$ with gentle relief conditions (the lowest point of 233.6m asl and the highest point of 352.6m asl and average slope gradient of 2^0).

Qin's slope gradation quantification technique was the inspiration to this study. However, the target area of Yen Chau district in the northern mountain region of Vietnam has stronger relief conditions than those of Qin's work. Yen Chau is characterized by steeper, longer and rougher slopes (elevation difference between the lowest and highest points is over 1400m and 60% total land falls into $9\text{-}35^0$ slope gradient range). Therefore, to fully describe soil-slope gradation relationship in this specific study area required another way of quantifying slope positions.

This study aimed to develop a slopeform system that is detailed enough to study soil-environment relationship in an area, with the focus on the interrelation between soils and terrain morphology. Firstly, five major slope positions for a single hillslope were generated: crest, upper, middle, foot slope, and valley. Peucker and Douglas' (1975) algorithm was applied to generate the highest (crest) and lowest (valley). The three mid-slope positions were generated using Skidmore's (1990)

Relative Position Index (PPI) algorithm. Secondly, the three major forms of convex, straight, and concave were used to characterize the three mid-slope positions, in both horizontal and vertical directions. Since these slope positions occupy most of the slope length, it was hypothesized that different detailed forms could lead to formation of certain soil groups and variability of soil fertility in the area to be mapped. The development of this detailed slopeform system applied fuzzy logic incorporated in SoLIMSolutions 2010 software upgraded by Zhu (2010).

2. Material and methods

2.1. Study area and environmental setting

A total area of 519 km^2 in the Yen Chau district of Son La province (Fig.1a) was mapped. Soils and soil fertility variations were mapped for the 193 km^2 of arable sloped land. Paddy rice fields were found in valleys along major rivers and streams, occupying 17km^2 and forests covered 262 km^2. The remaining 47 km^2 area was not of interest in this research.

The area has strong relief variations represented by a big difference between the highest (1567m asl) and the lowest (153m asl) elevation points (Fig.1b), with most of the long steep slopes found within the range of 9^0-35^0. The monsoonal climate is characterized by two distinctive seasons, a rainy season from May to October and a dry season from November to April of the next year. The statistics of Yen Chau weather station showed an annual average temperature of 24^0C and precipitation of 1257mm (data of 2000-2007).

The geology (Fig.1c) is part of the larger Vạn Yên geology and minerals system (Bao, 2004), which consists of five major geological units: (1) Volcanic Magmatites (VO) including aphyric basalt, magnesium-high basalt, andesitobasalt, andesitodacite, trachyte, agglomerate, and tuffaceous sandstone; (2) Clastic Sediments (SC) including clay shale, marl, sericite schist, agglomerate, polymictic gritstone, sandstone, siltstone, and coal; (3) Yen Chau Formation–Lower Subformation

(K_2yC_1) including four members of conglomerate, gritstone, sandstone, and interbedding of chocolate claystone; (4) Yen Chau Formation–Upper Subformation (K_2yC_2) including two members of sandstone and interbedding of conglomerate; and (5) Limestone (SO). Information about alluvial and colluvial deposit units was initially not available. Their characteristics were only noticed during the soil surveys. Alluvial and colluvial deposits prevail at valleys and lower slopes, where alluvial deposits are found along stream banks at slopes that are 8% or less and within a radius of 100m from the banks. Colluvial deposits follow up to 16% slopes. The spatial delineation of these two new units on the geology map was derived in ArcGIS 9.3.

The major crops on the upland slopes are maize (*Zea mays L.*) and cassava (*Manihot esculenta C.*) (Clemens, et al., 2010; Häring et al., 2010), with the observed effective rooting depths being 0-30cm for cassava and 0-50cm for maize. There is only a small percentage (1.18%) of the natural land used for fruit trees, being located in home gardens and mainly at lower slope positions with moderate slope inclinations (Clemens et al., 2010).

2.2. Sampling and soil characterization

Farmers' knowledge about local soils, their distribution and fertility was studied (Clemens et al., 2010) to plan a field survey. The catena concept was applied to locate soil observations for soil investigation and sampling. Five soil profiles were minimally studied for one single slope covering five major slope positions: crest, upper, middle, foot slope, and valley. The mapping of five major slope positions (see details in 2.3.2) was based on Relative Position Index (RPI), DEM-derived profile and planform curvature parameters. Since soils at forested mountain tops were not studied, the crest position was eliminated and the upper slope position was determined right after the forest-field boundary. 110 soil profiles were investigated and sampled for the five major geological units and the derived alluvial and colluvial deposits. The soil sampling was implemented by PhD, MSc, BSc, and internship students working in the Uplands Program during 2006–2012.

The soil profiles were dug down to the depth of 1.2–1.8m, then field described and classified according to FAO (2006) and the IUSS Working Group WRB (2006). The description and sampling were made for every soil horizon. The soil physical and chemical properties used in this study were the same like Clemens et al. (2010).

For soil fertility control, further calculations of soil properties were carried out for the calibration of the soil fertility mapping model. These calculations were made for the effective rooting space (ERS, dm), which is defined as the maximum depth of water, that can be reached by the roots during years of low rainfall (FAO, 2006). In this study, the ERS was taken down to 70cm and was separated into topsoil (0–30cm) and subsoil (30–70cm) to highlight the higher root densities and soil nutrient stocks in the topsoil to the subsoil (Clemens et al., 2010). These further soil properties computed based on Jahn et al. (2003) for every soil horizon are as follows:

(i) *Physical properties:* Soil volume (l/m^2) was the function of soil thickness (dm) and stone content (%). Soil mass (kg/m^2) is the total weight of soil material in a volumetric unit and was calculated on the function of soil volume and bulk density (BD) multiplied factor 1 for the topsoil and factor 0.5 for the subsoil. It is more compacted in the subsoil, which makes the soil mass of the subsoil a lot higher than that of the topsoil for a same soil depth. The factor 0.5 is used to equalize soil mass of the subsoil with that of the topsoil. Air capacity (AC) and available water capacity (AWC) (l/m^2) were the functions of estimated AC (%) and AWC (%) with soil volume.

(ii) Chemical properties: S-value (mol/m^{-2}) was the function of effective cation exchange capacity (CEC_{eff}) and base saturation (BS). The CEC_{eff} was estimated based on Jahn et al. (2006). S-

value was multiplied with factor 1 for the topsoil and 0.5 for the subsoil. The use of factor 0.5 is the same as explained in (i). The range-standardized sum parameter (N-P-S) proposed by Mausbach and Seybold (1998) was used to quantify soil fertility. This parameter is the sum of the stocks of total nitrogen (N_t), available phosphorus (P_{Bray1}), and S-value and was applied in this research in continuation of the work of Clemens et al (2010). These three stock values were standardized by setting the maximum value to 1 and the minimum value to 0 of each of the three parameters for the whole set of soil profiles and the values in between were calculated. The sum parameter N-P-S was achieved by summing up the 3 standardized values of N, P, and S-value. The parameter N-P-S, therefore, has the value range from 0 to 3.

2.3. Construction of the system of 29 slopeforms

2.3.1. Generation of a digital elevation model (DEM) for the study area

The DEM of the study area was constructed from vector files of contour lines, rivers and streams, the border of the study area, and elevation points using Topo to rater tool in ArcGIS 9.3. The map scale was 1:25,000 and the map resolution was set to be 10m by 10m. The output was a raster file as can be seen in Fig.1b. This DEM file was an important data type that was later used to extract other types of data for deriving slopeforms such as curvatures, relative position index (RPI), slope positions, and those as input parameters for the predictive soil mapping models such as slope inclination and slope aspect.

2.3.2. Generation of five major slope positions

The method for generating five major slope positions was well explained in Skidmore (1990). Firstly, the DEM was used to extract a crest and a valley to identify the highest and lowest points of a single slope applying Peucker and Douglas (1975) algorithm in SimDTA software (Qin et al., 2009). Secondly, the middle slope positions were interpolated applying the Relative Position Index

(RPI) as defined in Skidmore (1990). The value range of RPI is [0,1] with 0 being a valley and 1 being a ridge. This value range is subdivided to describe the five major slope positions for the area: ridge [0.99, 1], upper slope [0.7, 0.99], middle slope [0.3, 0.7], foot slope [0.01, 0.3], and valley [0, 0.01].

2.3.3. Generation of the detailed system of 29 slopeforms

Nine basic slopeforms defined in FAO (2006) were derived for each of upper, middle, and foot slope positions, making up to 27 slopeforms for this long slope range and totally 29 forms from crest to valley. Profile and planform curvature parameters were derived from the DEM using ArcGIS 9.3 to determine vertical and horizontal shapes of the 27 slopeforms. Profile curvature is parallel to the slope and indicates the direction of maximum slope and is the rate of change of gradient. It affects the acceleration and deceleration of flow across the surface and hence influences soil aggradation or degradation. Planform curvature is defined as the rate of change of aspect being perpendicular to the direction of the maximum slope and affects the convergence and divergence of flow across the surface (Odeh et al., 1991).

Three major forms for each of these parameters were: convex, straight, and concave. The parameter values set for these three major forms both vertically and horizontally were: convex > 0.005, straight [-0.005, 0.005], and concave < -0.005. The selection of these values was validated with field check to best describe gradual changes of slopeforms along a slope. The 29 slopeforms were assigned with values of RPI, profile and planform curvatures as indicated in Table 1. Mapping of this slopeform system was achieved by applying SoLIMSolutions 2010 software (Zhu et al., 2010).

2.4. Structural organization of the soil database

The soil database structure was developed based on the characterization of terrain characteristics of the study area. First, terrain units of the area were studied. According to Van Engelen and Wen (1995), terrain units are the general description of physiography and parent material. Three major landforms, level (L, inclination<8%), sloped land (S, 8≤inclination<30%), and steep land (T, inclination>30%), defined by Cong (2011) for a subcatchment scale of Yen Chau were used to classify subdivided landforms in this study. There were three subunits for L: L_1 (0-2%), L_2 (2-4%), L_3 (4-8%); two for S: S_1 (8-16%) and S_2 (16-30%); and six for T: T_1 (30-50%), T_2 (50-60%), T_3 (60-70%), T_4 (70-84%), T_5 (84-100%), and T_6 (>100%), totaling 11 subdivided landforms. The slope parameter was converted from degree to percentage on which the database of this study was built.

These 11 subdivided landforms were merged with the 5 geological units to derive a map of 55 terrain units. Alluvial and colluvial deposits remain as two independent terrain units, totally making up 57 terrain units for the study area. Given the minor impact of slopeforms as well as the latter 2 terrain units to soil formation and fertility at crest and valley, the occurrences of soils and soil fertility variability were studied on the 55 terrain units and 27 slopeforms, i.e. 1485 components. The fact that only 88 soil components from 110 soil profiles were identified would largely influence the quality of this work. To support calibration of the model, we used a reasoning method (section 3.1) built on knowledge from 11 junior and senior soil scientists working in the project to fill in the gaps of missing information. The team commonly agreed that only major slopeforms (straight, concave and convex) in combination with slope gradient had significant influences to soil occurrences and fertility dynamics (section 5.3).

Environmental parameters were collected for each of the soil profiles in which the slopeform and slope position of a profile were taken from the achieved slopeform map. Table 2 is a structural example of the soil database constructed for subdivided landforms T_1 and T_2. The soil database for the other subdivided landforms was organized the same way. Other parameters, such as slope aspect, elevation, were entered in a detailed Excel soil data table.

2.5. Predictive soil mapping under fuzzy logic

The application of fuzzy logic theory in predictive soil mapping techniques have been very well explained in numerous publications and studies (Burrough, 1996; Yang et al., 2007; Zhu et al., 2010). This part only summarizes the fuzzy logic method for predictive mapping of soils and soil fertility indices with the emphasis on a slopeform system relevantly derived for the study area.

2.5.1. Constructions of fuzzy similarity functions

The inference engine in SoLIM is operated using a raster data approach in which fuzzy similarity values (to values of prescribed soils) are calculated for every grid cell. The similarity value ranges from 0 (meaning the soil at a pixel is very different from a prescribed soil) to 1.0 (meaning the soil at a pixel carries exactly the same properties of the prescribed soil). The fuzzy minimum operator is used to calculate fuzzy membership values, or similarity values, (Zhu et al., 1996) for all map pixels of prescribed soils, and is expressed in equation 1. The intersection of sets $\mathrm{A} \in \mathrm{X}$ and $\mathrm{O} \in \mathrm{X}$ which corresponds to the connective "and", and its membership functions is given by:

$$\mu_N(x) = \min\{\mu_\mathrm{A}(x), \mu_\mathrm{O}(x)\}, x \in X \quad (1)$$

where x is an object which belongs to the set of object X, $\mu_\mathrm{A}(x)$ is called the degree of membership of x in A which maps X to the membership space M (when M contains only the two points 0 and 1, and $\mu_\mathrm{A}(x)$ is identical to the characteristic function of a non-fuzzy set).

Environmental variables in one unique combination (soil instance) to a specific soil group are calculated with optimality values using three optimality curves: bell, S, and Z shapes (Zhu, 1999) as can be seen in Fig.2. Equation (1) is applied on these optimality values to derive the similarity value for a given map pixel. The set of environmental variables is used again for other soil instances of this soil type for calculating optimality and similarity values. For example, Luvisols in Yen Chau can either be found on limestone with elevation from 300 m to 900 m a.s.l and with slope inclination below 30% at all slopeforms (instance 1) or limestone with elevation from 900 m to 1000 m a.s.l and slope inclination between 30% and 50% at slopeforms of 6, 15 and 24 (instance 2). After instances of this soil type are studied, the fuzzy maximum operator (equation 2) is applied to finalize the membership value of a map pixel for this soil type from the calculated membership values of the soil instances. For example, the similarity value of instance 1 is greater than that of instance 2, therefore, it better represents the similarity of Luvisols at this specific map pixel.

$$\mu_N(x) = \max\{\mu_A(x), \mu_O(x)\}, x \in X \quad (2)$$

This process goes on to other cells until it completes computing for all of the pixels of the study area. At this point, a thematic similarity map is produced for a single soil group. The engine then moves on to calculate for the rest of the prescribed soil types and more similarity maps are created. Finally, SoLIM can technically merge these similarity maps into one map, which assigns each cell with a value of one prescribed soil group based on the highest similarity value to this group using Hardened Map function in SoLIM. The final map is the soil/soil fertility map that delineates the spatial distribution of all soil/soil fertility groups for the mapping area.

2.5.2. Soil and soil fertility mapping

Reference soil groups (RSGs) need to be described and classified according to FAO (2006) for all geological units and elevation ranges in the area through a collection of visited soil profiles. For all soil profiles studied, information of rock type, elevation, slope gradient, slope position, slopeform, slope aspect, land use type, cultivation period need to be collected to learn the relationship between their formation and fertility variability with the above environmental parameters.

Each of the major RSGs is then further classified into one or more soil subunits using prefix and suffix qualifiers in WRB (2006) to highlight soil property differences within a RSG. Major properties are identified for major RSGs, such as Alisols, Luvisols, Cambisols, etc., for most representative environmental parameters, such as Cutanic for a high clay content, Leptic for high a stone content of soils at slopes greater than 60%, Humic for soils found at upper slope positions where landuse after deforestation is younger and OM contents thus higher, Profondic for soils with an argic horizon having a clay content not decrease by 25 percent or more 150cm from the surface, etc. Information occurring these soil subunits is collected in Table 4 and calibrated in SoLIM to derive a soil subunit map for all major RSGs. This can be easily done based on the described soil profiles in the soil data set made for all RSGs that were contributed by all soil scientists working in the project, which will help provide a more diverse picture of soil distribution in the area.

To prepare for the soil fertility mapping model, soil fertility classes are defined based on calculated N-P-S values (Table 5). These fertility classes are then prescribed based on distinctive combinations of environmental parameters (same with the ones that were used in the RSG mapping model) as soil forming factors (see example in Table 6). The soil fertility classification is made for Alisols, Luvisols, Cambisols, Regosols, Leptosols, Fluvisols, Stagnosols, and Vertisols.

3. Results

3.1. Mapping of the reference soil groups

The soil investigation over the period 2006-2012 resulted in 11 RSGs, in which Alisols and Luvisols were the two most abundant RSGs and found on all rock types with 46 and 36 profiles, respectively. There were 7 Anthrosols, 7 Regosols, 4 Stagnosols, 3 Vertisols, 2 Cambisols, 2 Phaeozems, 1 Leptosol, 1 Gleysol, and 1 Fluvisol. The fact that the total 110 soil profiles did not cover all combinations of relief parameters led to the use a reasoning strategy using expert knowledge of the soil scientists working in the project. Field observations and personal impressions of soil formation in Yen Chau of junior and senior soil scientists working in the SFB 564 project were gathered to estimate soil occurrences in empty relief combinations. This local soil expert then helped fill in the soil formation reasoning table (Table 3).

SoLIM, which is an automated soil inference system that studies the relationship between soils and environmental conditions, was used to generate fuzzy similarity maps (Zhu et al., 2010). Eleven fuzzy similarity maps were derived for the study area with each map representing one RSG. A hardened RSG map for the study area was derived from the fuzzy similarity maps (Fig.3). The statistic results show that Alisols and Luvisols are the most abundant soil groups within the focal mapping area, occupying 47.2% and 38.5% of the total area, respectively. Cambisols are the third abundant soil group with 10.6% of the total area, which prevail at non-forested crest positions, middle slope positions having convex slopeforms, and colluvial deposits at foot slope positions. Leptosols, Regosols, Fluvisols, Stagnosols, and Vertisols have remarkably small areas with 1.8, 1.4, 0.4, 0.1, and 0.02% of the focal mapping area, respectively. The predicted occurrence of Phaeozems spreads over a large area of 54 km^2 and paddy soils (Anthrosols and Gleysols) occupy an area of 17 km^2. However, these three RSGs were not evaluated with accurate assessment of their occurrences and fertility because they are located outside of the area of interest as mentioned before, i.e. arable sloped land or the focal mapping area.

Since the RSGs within the focal mapping area are only major soil groups and each group can have a wide range of soil property values, evaluation of the results needed more detailed information. Indirect assessment (Zhu et al., 2010) of this predictive soil mapping approach was, therefore, necessary. To support this, a soil subunit map and a soil fertility map were derived for more diverse evaluations.

3.2. Soil subunit map

Fig.3 shows the spatial distribution of 18 soil subunits of major RSGs. Alisols, Luvisols, and Cambisols had the most subunits (12/18) and together accounted for 95.8% of the predicted soil map. Cutanic is the most prominent property for Alisols and Luvisols in Yen Chau (Fig.4). Soils found at crest and high upper slope positions had higher amounts of organic matter than those at the lower positions, which resulted in a Humic property for Alisols, Luvisols and Cambisols at these locations. Within the middle slope length, Profondic is the major property at lower upper and higher middle slope positions, Leptic at the middle position having convex forms, equivalent to a high rock content, and Siltic at lower middle slope for both Alisols and Luvisols. Many Luvisols found had a Nitic property at lower middle and higher foot slope positions. Investigated Cambisols had a Dystric property at middle slope positions that have the concave-convex (CV) form, where topsoils were eroded and unconsolidated materials remained right before the bed rock. Cambisols found at foot slope positions mainly had a Luvic property for having deposition of eroded materials from upper slopes right on top of an in-situ argic horizon.

Each of the other RSGs (Regosols, Leptosols, Fluvisols, Stagnosols, and Vertisols) had one most representative subunit.

3.3. Soil fertility map

A detailed classification of 17 soil fertility classes was generated to see whether there would be some influence degree of the environmental parameters to soil fertility at this detail level (Fig.7 a&b). Insignificance degrees resulted in the merging of them into four groups: good, moderate, low, and very low based on comparisons with soil properties calculated for all soil profiles and on information collected at all observation points. The good group included classes 1-4, covering an area of 35.5km^2 or 18.4% of the arable sloped land. The moderate group included classes 5-8, occupying an area of 86.3km^2 or 44.7% of the arable sloped land. The classes 9-12 constituted the low fertility group, which occupied 66.4km^2 or 34.4% of the arable sloped land. The very low group was the smallest one spreading over an area of 4.8km^2 or 2.5% of the arable sloped land and merged from classes 13-17 (Fig.5b).

4. Validation of the soil, soil subunit and soil fertility maps

Fifty extra soil profiles were described and classified using the WRB 2006 to evaluate the match of the observed soil data with the three predicted maps (Fig.7). The sampling strategy was purposive following the catenary sequences for the predictive soil map of Yen Chau focusing on crest, upper, middle-, and foot slope positions. Out of 50 validation points, 10 points were collected for VO, 9 for SO, 10 for SC, 12 for K_2yC_1, and 9 for K_2yC_2.

4.1.Validation of the soil map

Matching the validation points with the predicted RSG map showed an accuracy of 90% (a match of 9/10 points) for VO, 78% (7/9 points) for SO, 60% for SC (6/10 points), 67% (8/12 points) for K_2yC_1, and 88.9% (8/9 points) for K_2yC_2. Overall, 37 out of 50 observation points matched with the inferred RSG map, resulting in an accuracy of 76%. Compared to the results of different predictive soil mapping studies applying different methods, (Zhu et al., 1996; Zhu et al., 2010; Schuler et al., 2010) this accuracy is acceptable. It shows the consideration of slope positions and slopeforms in the soil-landscape relationships is useful in capturing a large portion of RSGs for this area of strong relief conditions in the NW Vietnam.

4.2.Validation of the soil subunit map

The 50 validation soil profiles were classified with soil subunits according to WRB (2006) to validate the results of the predicted soil subunit map. Matching the validation points with the soil subunit map showed an accuracy of 80% (a match of 8/10 points) for VO, 78% (7/9 points) for

SO, 60% for SC (6/10 points), 67% (8/12 points) for K_2yC_1, and 78% (7/9 points) for K_2yC_2. The overall accuracy of the soil subunit map was 72%, i.e. 36 out of 50 observation points matched with the inferred soil subunit map. This accuracy degree is still acceptable, which shows the ability of the model to capture most of the soil subunits identified for the area.

4.3.Validation of the soil fertility map

To validate the predicted soil fertility map (with 17 fertility classes), the same 50 validation data points and two indices applied in Zhu et al. (2010) were used: root mean square error (RMSE) and agreement coefficient (AC). The AC index was defined by Willmott (1984) as follows:

$$AC = 1 - \frac{n * RMSE^2}{PE} \qquad (3)$$

The sum parameter N-P-S was calculated for all of the 50 validation data points, resulting in 50 observed N-P-S values. The locations of these observed values on the soil fertility map were used to extract the 50 corresponding predicted values. The RMSE the 50 data points from the predicted map is 0.58. For the N-P-S data value range from 0 to 3, this RMSE value is rather a big value, which shows quite distant differences between the predicted and observed values. The calculation of the AC index, which is 0.60, also confirmed a medium agreement (or an average match) between the predicted and the observed values at these 50 locations. This medium accuracy value means that the environmental parameters taken were not sufficient to achieve a better match between the predicted soil fertility map and the real fertility variations of soils in the mapping area. This result suggests that more parameters be used in order to improve the certainty of the inference result of the soil fertility map. For example, land-use history must have had lots of impact on soil fertility decline over time and, therefore, the fertility of soils at various locations that have different land-use ages after deforestation. However, this parameter was not incorporated in the model due to the large size of the area to be mapped and the program's limited budget.

5. Discussion

Soil = (Cl, Pm, Og, Tp) t (Jenny, 1941) is a function to conceptualize the soil-environmental relationship, which states that the formation of soils (s) is influenced by different factors, remarkably climate (Cl), parent material (Pm), organisms (Og), and topography (Tp). These factors evolve with factor time (t). In this study, soil formation and soil fertility variability represented by SOM stocks follow the same rule and soil-forming factors are discussed in the order of importance.

5.1. Parent material

The study agrees with that of Schuler (2008), Parton et al. (1987) and Six et al. (2002) on different parent rocks resulting in the formation of different major soils determined by soil compositional properties, such as soil pH, clay content, CEC and BS; and by soil organic matter (SOM) whose stability and dynamics are very much controlled by clay content. In Fig.6, light clay composition in SC, i.e. sandy clay to clay loam in the topsoil and sandy clay loam to light clay in the subsoil, is the reason for having BS smaller than 50% in all soil profiles, which explains the absence of Luvisols and lower stocks of bases. Luvisols were found in all of the other four rock types. VO has highest BS, CEC and S-value values among all rock types due to having high clay contents: clay loam in the topsoil and clay in the subsoil, which is the reason for having Luvisols as the major soil, not Alisols. Both Luvisols and Alisols were found on SO, SC and K_2yC_2, in which average BS, CEC and S-value were highest in SO and lowest in SC.

SOM contents were found to be higher in SC, VO and SO and lowest in K_2yC_1. The finding agrees with that of Clemens et al. (2010) which found the lowest average clay content of soils in K_2yC_1.

However, the SOM content rather depends largely on land-use history. Clemens et al. (2010) and Häring et al. (2010) found that soils developed in valleys of K_2yC_1 and K_2yC_2 had lower average SOM contents. Participatory investigations for land-use history in Yen Chau revealed that K_2yC_1 and K_2yC_2 soils have had longer cultivation periods for better access and denser population.

5.2. Climate in association with elevation

The study coincides with that of Schuler (2008) that Luvisols prevail at elevations below 900m asl. Alisols are the major soil group above this line. The transition of Alisols to Luvisols in rising elevation is explained by elevation-triggered differences in temperature and rainfall. The higher temperature and more seasonal precipitation below 900m asl are more favourable for clay illuviation, which brings higher BS (>50%). Whereas, the lower temperature and higher precipitation at elevations above 900m asl hamper the percolation of water through cracks in the soil, pathway of downward transportation of clay minerals. The lower temperature and more moist soils due to the high rainfall limit the formation of cracks via shrinking process in clay-rich soils.

Similar to many studies (Sevgi and Tecomen, 2009; Chuai et al., 2012; and Vogel and Märker, 2011), positive correlations between SOM and elevation were found below 900m asl in most rock types. These values became negative above this elevation line due to a lower clay content found in soils at this elevation range.

5.3. Relief

Slope inclination, position and form play remarkable roles in the variability of soil properties and occurrences of certain soil types Clemens et al., 2010; Cong, 2011; Qin et al. 2009, 2012). In the monocropping culture of maize in Yen Chau, soils on steep slopes at a same slope position with a similar cultivation period have a remarkably lower topsoil thickness and SOM content than those at gentler slopes. This problem was caused by soil erosion triggered by intensive cultivation activities.

In accordance with the findings of Clemens et al. (2010), soils at crest and upper slope positions tend to have higher SOM stocks commonly represented by Humic property. This is because forest invasion for agricultural land occurred last at these areas, hence younger cultivation periods and more soil carbon retained in the topsoil. Luvisols and Alisols mainly have Vertic, Nitic, Leptic and Profondic properties at the middle slope and Siltic property at the foot slope. The Leptic property of Luvisols and Alisols and Dystric property of Cambisols occur at a location that has a high slope gradient and convex form. This is where the topsoil is shallow because of severe soil erosion and a high stone content is found in the both topsoil and subsoil. When a soil has a shallower soil depth and a remarkably higher stone content, the soil then becomes Regosols or Leptosols. This is where Regosols and Leptosols were mainly found in Yen Chau. Similar to the findings of Schuler (2008), Clemens et al. (2010) and Häring et al. (2010), Siltic for Luvisols and Alisols and Luvic for Cambisols are the major properties in the foot slope position for the deposition of soil particles eroded from upper slope positions.

In the agreement with Schuler (2008) and Clemens et al. (2010), the study found that the straight forms resulted in better soil structure, soil stability, less susceptibility to erosion and better retaining of soil carbon. Thus, soils at locations with straight forms had higher stocks of SOM in the topsoil than those at locations with the other forms. Locations with convex slopeforms at high slope gradients are more prone to erosion and were found to have shallowest topsoil and lowest SOM stocks. Concave forms were found to be either the result of natural formation or small landslides. In the first case, the topsoil can be deep and contain high SOM stocks because of the convergence shape that collects soil particles eroded from higher points, but can, at the same time, result in small landslides due to weaker soil structure. In the latter case, the topsoil was removed and the subsoil is exposed, which resulted in low SOM stocks and poor crop growth. Overall, however, slopeform does not stand out for a very important variable of soil fertility which should rather depend more heavily on other factors such as parent material, elevation and cultivation period or the history of land use.

5.4. Biological activity, human impact and time

Similar to findings of Clemens et al. (2010) and Häring et al. (2010, 2013a and 2013b), macro-organisms were seen most active in the topsoil within the depth of 10-30cm. Major burrowing animals were termites, ants, earthworms, crickets, and beetles whose activities help create space for microorganisms and root penetration, water infiltration, and turn over organic matter in the soil. These activities contributed to the intermediate horizon of AhBt or BtAh in between A and Bt horizons. An E-horizon was rarely found. Microorganisms play an important role in increasing organic matter content in the soil through physical and biochemical processes. As indicated by Häring et al. (2010, 2013a and 2013b), soils in undisturbed forest land of Yen Chau had deeper and darker topsoil than soils that were converted from forest to agriculture. These soils were Phaeozems developed on limestone.

The human impact, especially land-use change from forest to maize, effects the distribution of soils and their properties (Schuler, 2008). A remarkable difference between average SOM stocks in the topsoil of Phaeozems (20-27 kg/m^2) and the other soils (2-14 kg/m^2) well proved this fact. Häring et al. (2010) found a decline in soil organic matter by 66%, N_t by 67%, exchangeable Ca^{2+} by 91%, Mg^{2+} by 94%, K^+ by 73%, available P by 75%, pH values by 2.2 units, and cation exchange capacity by 56%. Häring et al. (2013a) found a higher total SOC loss (6–32%), a lower decomposition (13–40%), and a lower SOC input (14–31%). In terms of the mass of soil loss in

Yen Chau, Tuan et al. (2014) estimated that the loss due to the current maize cropping practice by local farmers reached 174 t ha^{-1} a^{-1}. Soil erosion over years at steep upper and middle slope positions turned Luvisols and/or Alisols into Regosols (Häring et al., 2010), and led to the formation of Cambisols at low to moderate foot slope positions (Clemens et al., 2010).

In regards to the factor time, the initial agricultural activity was mainly slash-and-burn and shifting cultivation in mountainous regions of Vietnam, which was stated to be sustainable (Dao, 2000; Vien et al., 2004) and did not change much the nature of the soils. The population growth over the last decades led to increasing demands on food and forests were tremendously taken for agriculture. Since this time, soils of Yen Chau have been changed forever (Häring et al., 2013a; Häring et al., 2013b; Tuan et al., 2014).

6. Conclusion

This paper applies a fuzzy soil mapping approach to derive fuzzy similarity functions in an effort to develop soil and soil fertility maps for a region in NW Vietnam. These functions are constructed from descriptive knowledge represented by environmental factors that have impacts on soil occurrences and soil fertility variations. To best describe the distinctive extreme relief characteristic of the study area, a detailed 29 fuzzy slopeform system was formulated and constructed on a hypothesis that spatial distribution of soils and their fertility degrees could be better achieved from this parameter. From the results of the study, the following conclusions are made:

1) The system of 29 fuzzy slopeforms was successfully developed from five major slope positions. This parameter well delineates available surface forms of hill slopes from mountain tops to valleys, which allows the possibility to predict the occurrence of a soil and assess its fertility status at any location within the area of 10m by 10m.

2) We were able to formulate the soil-environmental relationship based on the soil data acquired and soil fertility classes calculated for the mapping area. This confirms the ability of this soil mapping approach in obtaining descriptive knowledge for digital soil mapping of an area with limited or no soil information (Zhu et al., 2010).
3) Good map accuracies of the RSG map and the soil subunit map (76% and 72%, respectively) reveal the applicability of the 29 fuzzy slopeform system in digital soil mapping for areas with extreme relief conditions like the one in this research. On the contrary, the validation of the soil fertility map shows just an average accuracy value of 60%. This is because:
4) Soil fertility does not vary strongly with the change in the surface form. Instead, land-use age after deforestation has been found to have a greater impact on soil fertility decline in Yen Chau district (Häring et al., 2010; Häring et al., 2013a; Häring et al., 2013b). For instance, the knowledge of cultivation period acquired from farmer interviews put in comparison with the analytic data for soil fertility control reveals that soils at upper slope positions tend to be more fertile than those at lower slope positions for having a younger period of cultivation due to forest clearance for agriculture happened from the bottom. The quantification of this information in better capturing soil fertility using this fuzzy mapping approach, therefore, should create an interesting research topic for this area in the future.

Acknowledgements

We would like to thank the local authorities of Yen Chau district and communes of Chieng Dong, Chieng On, Chieng Tuong, Chieng Sang, and Yen Son for providing useful information and close cooperation. We thank Andrea Zipp and Dr. Ludger Herrmann for doing and organizing laboratory analyses, and Hannah Kuhfeld for her help in improving the manuscript, and numerous farmers in Yen Chau for their field assistance, especially Lo Van Ha and Hoang Van Vui.

We wish to thank the German Science Foundation (DFG) (SFB 564) for the financial support within the framework of the SFB 564 “Research for sustainable land use and rural development in mountainous regions of Southeast Asia”.

Special thanks go to Prof. A-Xing Zhu, University of Wisconsin, and Prof. Qin Cheng-Zhi, Chinese Academy of Sciences for their free and helpful discussions and kind technical support.

Appendix a. Supplementary data

Supplementary data associated with this article can be found in the online version http://dx.doi.org/10.1016/j.catena.2017.03.01. These data include Google maps of the most important areas described in this article.

References

Bao, N. X. (ed.), 2004. Geology and mineral resources map of Vietnam, (F-48-XXXVII), scale 1:200.000, department of geology and mining, Vietnam, Hanoi. 145 pp.

Batjes, N.H., 2008. Mapping soil carbon stocks of Central Africa using SOTER. Geoderma 146, 58–65.

Bui, E.N. Loughhead, A., Corner, R., 1999. Extracting soil landscape rules from previous soil survey, Sutr. J. Soil Res. 37, 495–508.

Burrough, P.A., 1989. Fuzzy mathematics methods for soil survey and land evaluation. Journal of Soil Science 40, 447–492.

Burrough, P.A., 1996. Natural objects with indeterminate boundaries. In: Burrough, P.A., Frank, A.U. (Eds.), Geographic Objects with Indeterminate Boundaries. Francis and Taylor, London. 365pp.

Chuai, X.W., Huang, X.J., Wang, W.J., Zhang, M., Lai, L. and Liao, Q.L., 2012. Spatial variability of soil organic carbon and related factors in Jiangsu Province, China. *Pedosphere* 22(3), 404–414.

Clemens, G., Fiedler, S., Cong, N.D., Dung, N.D., Schuler, U., Stahr, K., 2010. Soil fertility affected by land use history, relief position and parent material under tropical climate in NW-Vietnam. Catena 81, 87-96.

Conarcher, A.J., Dalrymple, J.B., 1997. The nine unit landsurface model: an approach to pedogeomorphic research. Geoderma 18, 1–154.

Cong, N. D., 2011. SOTER database for land evaluation procedure: A case study in two small catchments of Northwest Vietnam. Hohenheimer Bodenkundliche Hefte, 101, Hohenheim University, Germany.

Douglas, D., Peucker, T., 1975. Detection of Surface-Specific points by Local Parallel Processing of Discrete Terrain Elevation Data, Computer Graphics, Visions and Image Processing. Vol.4, No.2 pp. 375–387 (December 1973).

FAO, 2006. Guidelines for Soil Description, 4th edition. Rome, Italy. 97pp.

Häring, V., Clemens, G., Sauer, D., Stahr, K., 2010. Human-induced soil fertility decline in a mountain region in Northern Vietnam. Die Erde 150, Berlin, Special Issue “Fragile Landscapes”, 235–253.

Häring, V., Fischer, H., Cadisch, G., Stahr, K., 2013a. Implication of erosion on the assessment of decomposition and humification of soil organic carbon after land use change in tropical agricultural systems. European Journal of Soil Science.

Häring, V., Fischer, H., Cadisch, G., Stahr, K., 2013b. Improved $\delta^{13}C$ method to assess soil organic carbon dynamics on sites affected by soil erosion. Soil Biology & Biochemistry 65 (2013) 158–167.

Herrmann, L., Graef, F., Wellwe, U., Stahr, K., 2001. Land use planning on the basis of geomorphic units: experiences with the SOTER approach in Niger and Benin. Z. Geomorph. N. F. 124, 111–123.

IUSS Working Group WRB, 2006. World reference base for soil resources 2006, 2nd ed. World Soil Resources Report, No. 103. FAO, Rome.

Jahn, R., Blume, H.P, Asio, V.B., 2003. Students guide for soil description, Soil classification and site evaluation. Universität Halle. 120 pp.

Jenny, H., 1941. Factors of soil formation: A system of quantitative pedology. McGraw-Hill, New York. 281 pp.

Mausbach, J.M., Seybold, C.A., 1998. Assessment of soil quality. In: Lal, R. (Ed.), Soil Quality and Agricultural Sustainability. Sleeping Bear Press, Chelsea, USA.

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

Odeh, I.O.A., 1991. Elucidation of soil-landform inter-relationships by canonical ordination analysis. Geoderma 49, 1–32.

Parton, W.J., Schimel, D.S., Cole, C.V., Ojima, D.S., 1987. Analysis of factors controlling soil organic matter levels in Great Plains grasslands. Soil Science Society of America Journal 51, 1173–1179.

Qi, F., Zhu, A.X., 2003. Knowledge discovery from soil maps using inductive learning. International Journal of Geographical Information Science 17, 771–795.

Qin, C.Z., Zhu, A.X., Shi, X., Li, B.L., Pei, P., Zhou, C.H., 2009. The quantification of spatial gradation of slope positions. Geomorphology 110, 152–161.

Qin, C.Z., Zhu, A.X., Qiu, W.L., Lu, Y.J., Li, B.L., Pei, T., 2012. Mapping soil organic matter in small low-relief catchments using fuzzy slope position information. Geoderma 171–172, 64–74.

Qin, J.H., Wang, Q., Sung, H., 2013. Changes of organic carbon and its labile fractions in topsoil with altitude in subalpine-alpine area of southwestern China. Acta Ecol. Sin. 33, 5858–5864.

Schuler. U., 2008. Towards regionalisation of soils in northern Thailand and consequences for mapping approaches and upscaling procedure. Hohenheimer Bodenkundliche Hefte, 89, Hohenheim University, Germany.

Schuler, U., Herrmann, L., Ingwersen, J., Erbe, P., Stahr, K., 2010. Comparing mapping approaches at subcatchment scale in northern Thailand with emphasis on the Maximum Likelihood approach. Catena 81, 137–171.

Sevgi, O., Tecomen, H.B., 2009. Physical, chemical and pedogenetical properties of soil in relation with altitude at Kazdagi upland black pine forest. J Environ Biol 30(3): 349–354.

Six, J., Feller, C., Denef, K., Ogle, S.M., De Moraes Sa, J.C., Albrecht, A., 2002. Soil organic matter, biota and aggregation in temperate and tropical soils–effects of no–tillage. Agronomie 22, 755–775.

Skidmore, A.K., 1990. Terrain position as mapped from a gridded digital elevation model. International Journal of Geographical Information Systems 4(1), 33–49.

Speight, J.G., 1990. Landform, In: McDonald, R.C., Isbell, R.F., Speight, J.G., Walker, J., Hopkins, M.S. (Eds.), Australian soil and land survey field handbook, (2nd edition). Inkata Press, Melbourne, pp. 9–57.

Tuan, V.D., Hilger, T., MacDonald, L., Clemens, G., Shiraishi, E., Vien, T.D., Stahr, K., Cadisch, G., 2014. Mitigation potential of soil conservation in maize cropping on steep slopes. Field Crops Research 156 (2014) 91–102.

Vågen T-G., Winowiecky, L.A., 2012. Mapping of soil organic carbon stocks for spatially explicit assessment of climate change mitigation potential. Environ.Res.Lett 8(1), 1–9.

Van Engelen, V.W.P., Wen, T.T., 1995. Global and national soils and terrain digital dabases (SOTER): Procedures manual, ISRIC, Wageningen, the Netherlands. 138pp.

Vien, T.D., Dung, N.V., Dung, P.T., Lam, N.T., 2004. A nutrient balance analysis of the sustainability of a composite swiddening agroecosystem in Vietnam's northern mountain region. Southeast Asian Studies, Vol. 50, No. 4.

Vogel, S., Märker, M., 2011. Characterization of the pre-AD 79 Roman paleosol south of Pompeii (Italy): Correlation between soil parameter values and paleo-topography. Geoderma 160, 548–558.

Willmott, C.J. (1984): On the evaluation of model performances in physical geography. In: Gaile, G.L., Willmott, C.J. (Eds.), Spatial statistics and models. D. Reiden Publi., Dordrecht, the Netherlands, pp. 43–460.

Winowiecki, L.A., Vågen T-G., Massawe, B., Jelinski, N.A., Lyamchai, C., Sayula, G., Msoka, E., 2015. Landscape-scale variability of soil health indicators: effects of cultivation on soil organic carbon in the Usambara Mountains of Tanzania. Nutr Cycl Agroecosyst. DOI 10.1007/s10705-015-9750-1.

Yang, L., Zhu, A.X., Li, B.L., Qin, C.Z., Pei, T., Liu, B.Y., Li, R.K., Cai, Q.G., 2007. Extraction of knowledge about soil-environment relationship for soil mapping using fuzzy c-means (FCM) clustering. Acta Pedol. Sin. 44, 16–23.

Young, A., 1972. Slopes. Oliver & Boyd, Edinburgh, pp. 288.

Zhu, A.X., Band, L.E., Dutton, B., Nimlos, T.J., 1996. Automated soil inference under fuzzy logic. Ecol. Modell 90, 123–145.

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

Zhu, A.X., Band, L., Vertessy, R., Dutton, B., 1997b. Deriving soil property using a soil land inference model (SoLIM). Soil Sci. Soc. Am. J. 61:523–533.

Zhu, A.X., Yang, L., Li, B., Qin, C.Z., Pei, T., Liu, B., 2010. Construction of membership functions for predictive soil mapping under fuzzy logic. Geoderma 155, 164–174.

LIST OF FIGURES

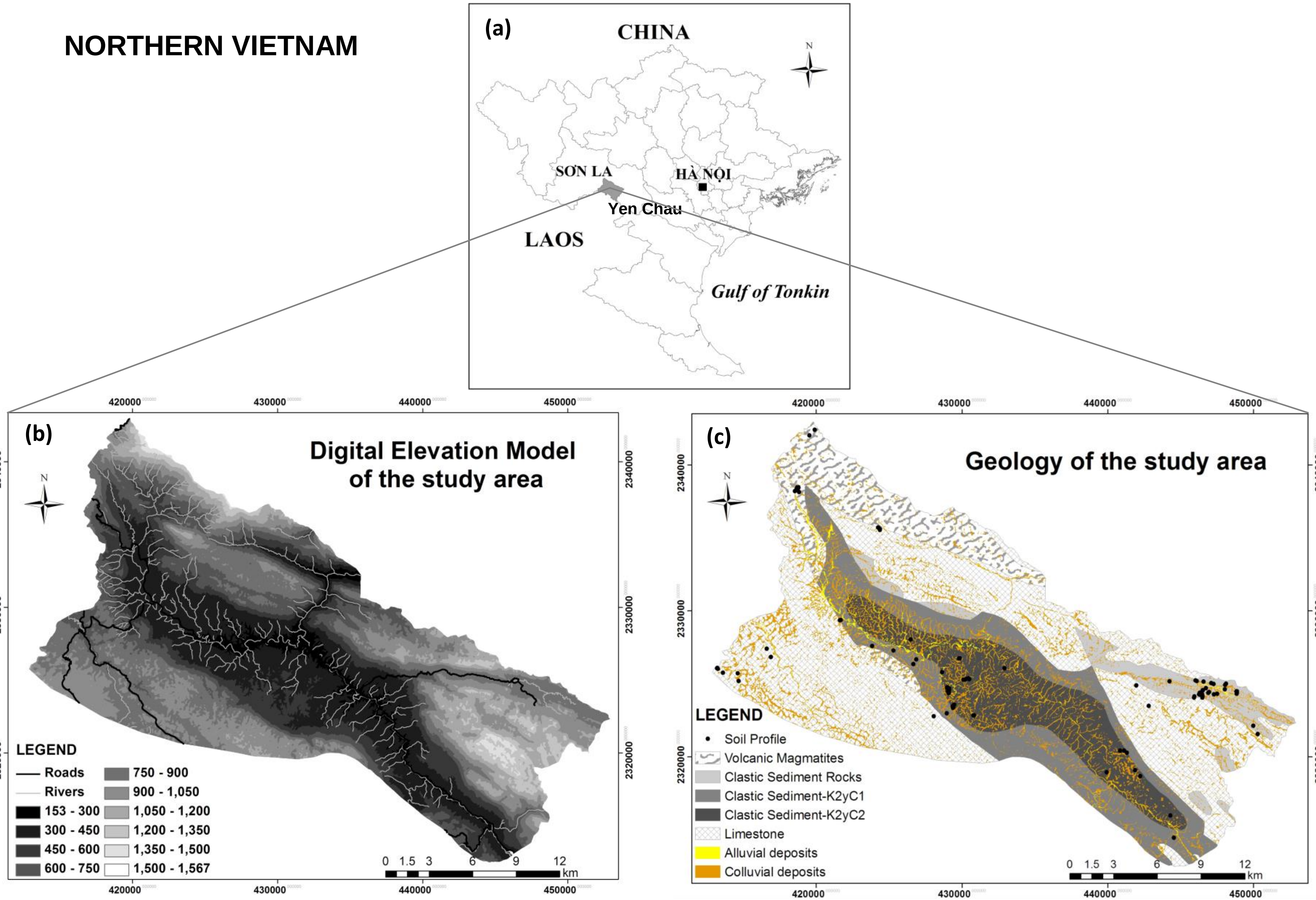

Figure 1. a) Geographic location of the study area in Yen Chau district, Son La province; b) The DEM of Yen Chau district; c) Geology of Yen Chau district and locations of soil profiles

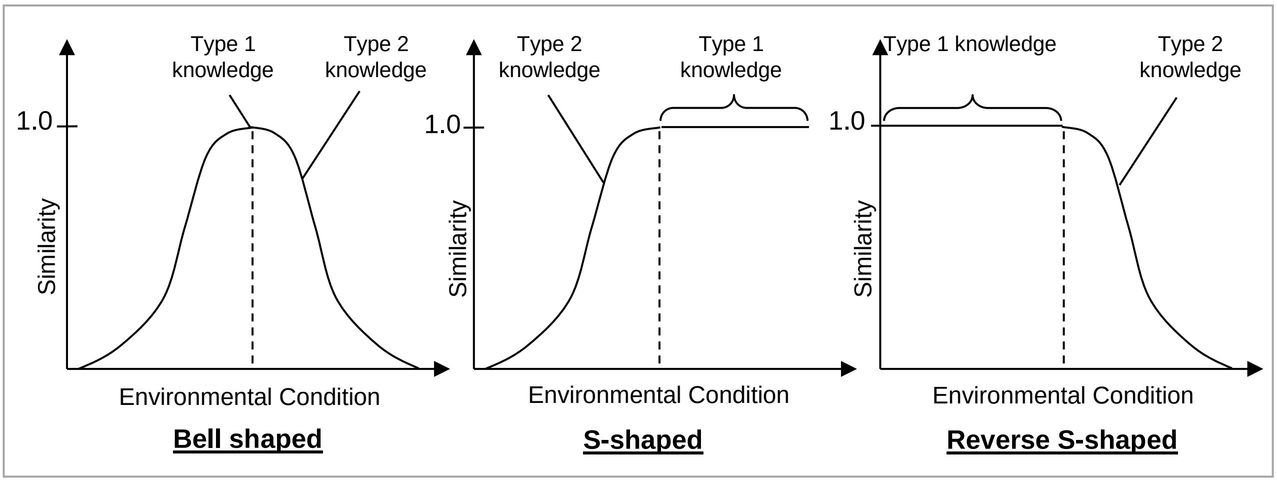

Figure 2. The three basic forms of membership functions (Zhu et al., 1999)

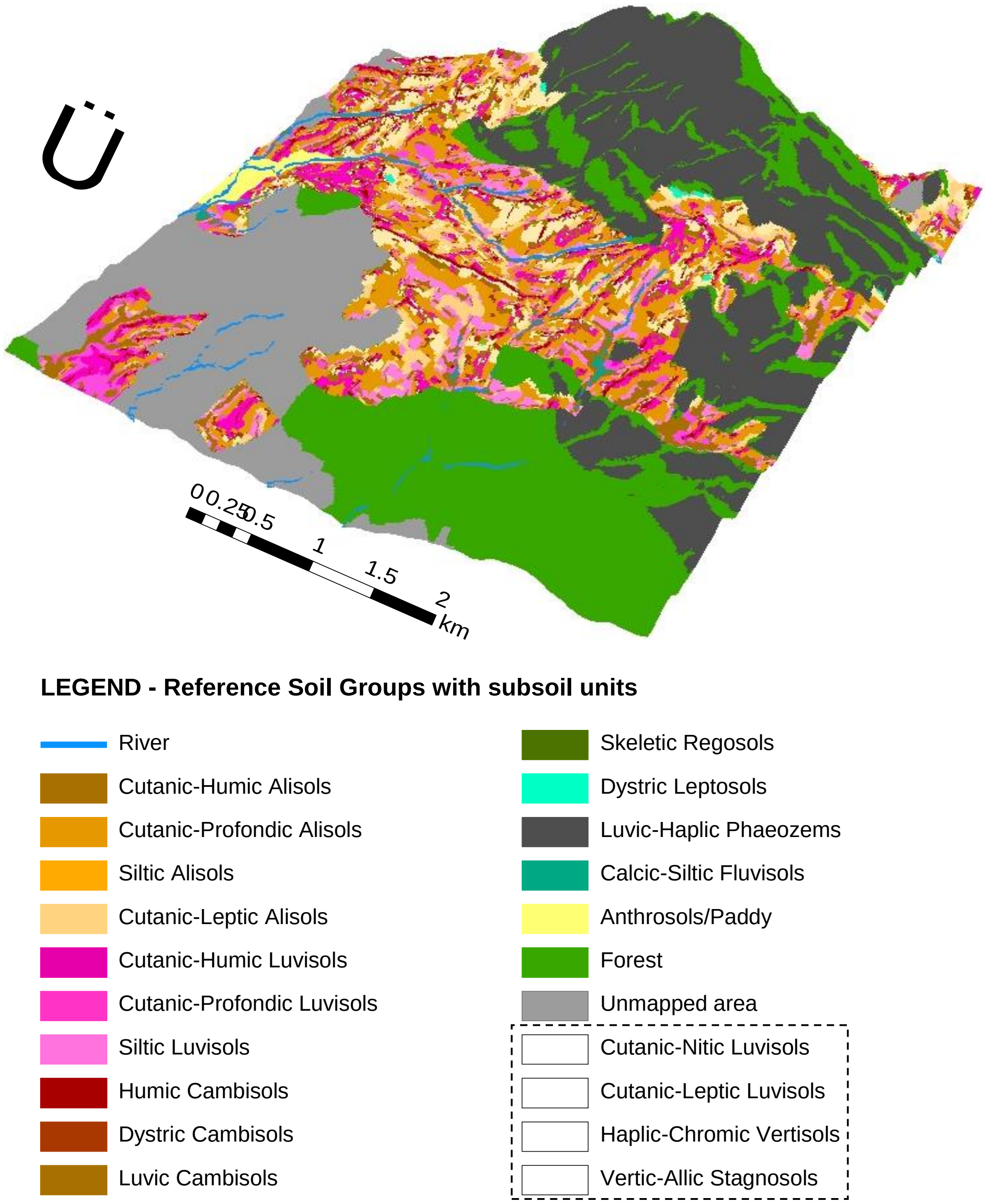

Figure 3. Predicted soil map of a small area of the study area. The last four subsoil units (with blank colour patterns) are not found within this sample area, but characterized for the whole study area

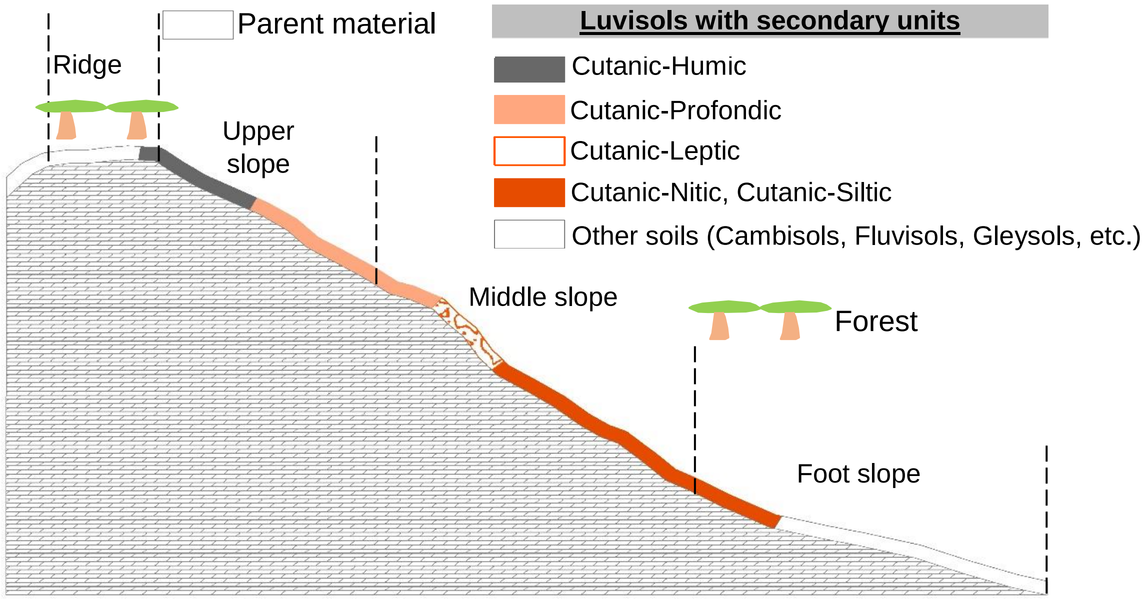

Figure 4. Distribution of subsoil units along a slope for Luvisols

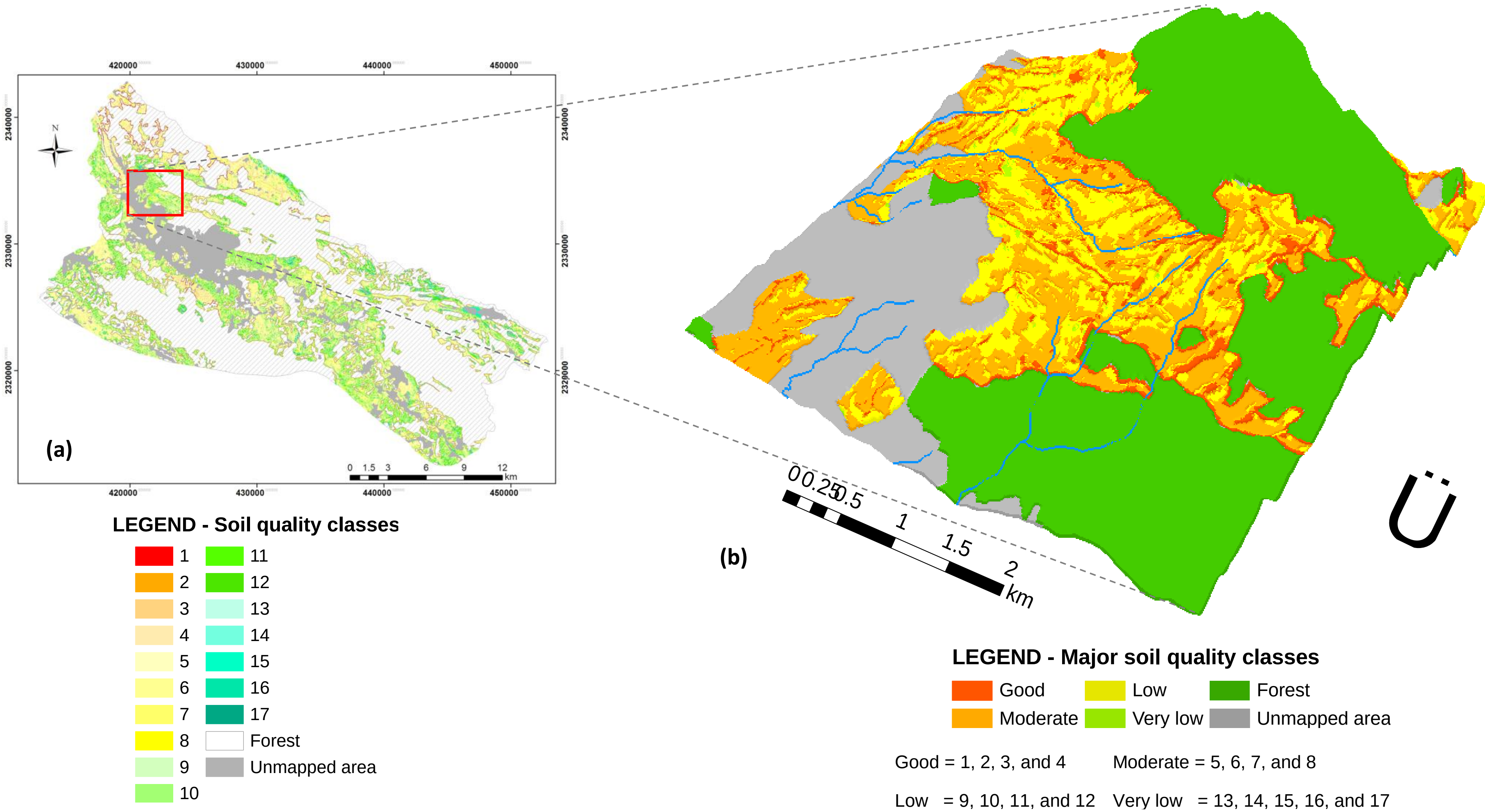

Figure 5. Predicted soil quality map of the study area (a), a 3D representation of major soil quality classes of a smaller are as example (b)

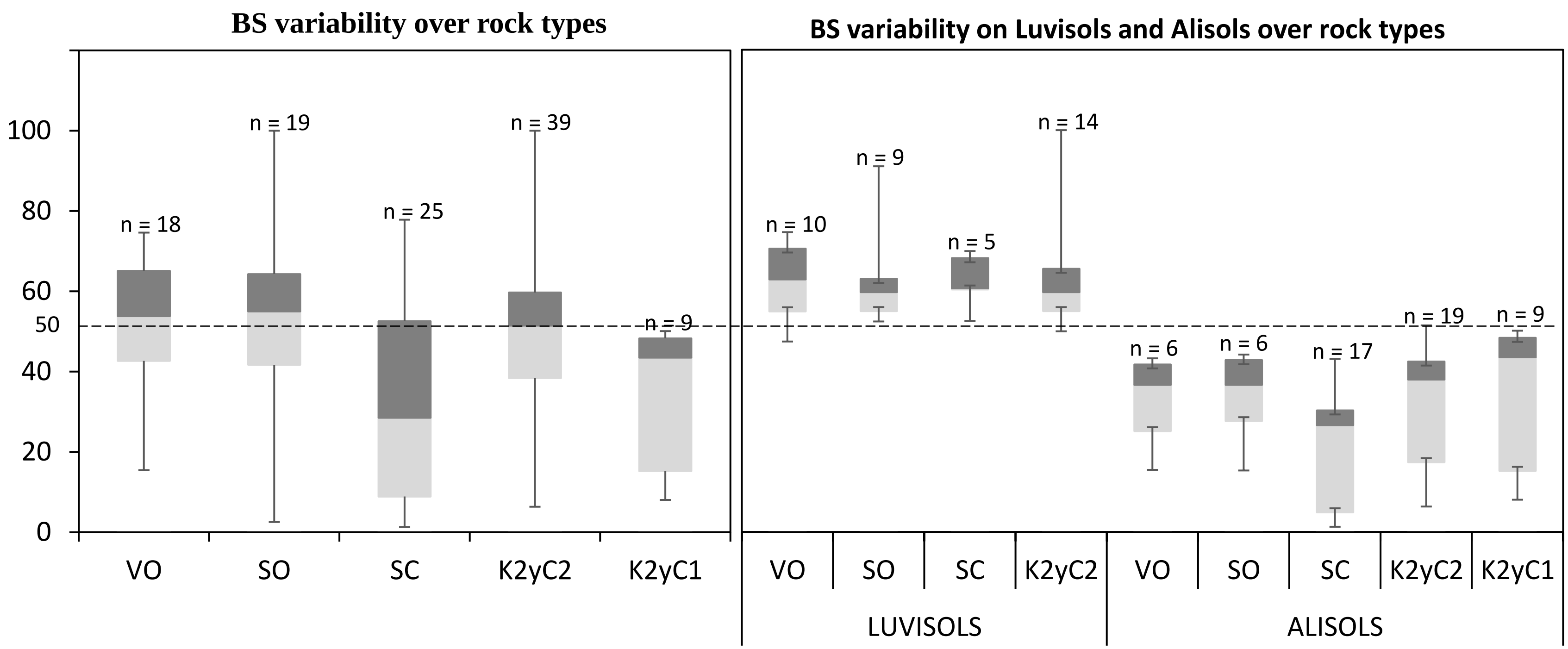

Figure 6. Variability of BS (%) for the ERS over different parent rocks and major reference soil groups

Explanation of the supplementary data in the .csv plot data files:

Luvisols = 1, Alisols = 2, VO = 1, SO = 2, SC = 3, K2yC2 = 4, K2yC1 = 5

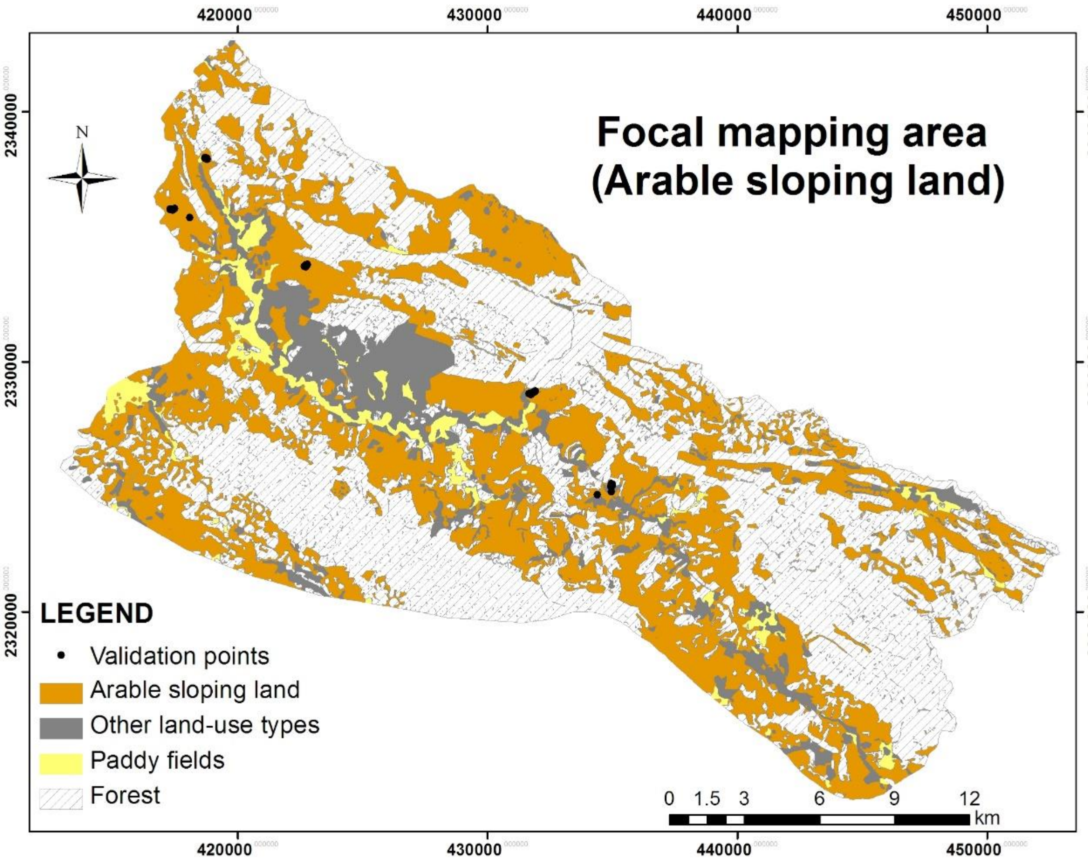

Figure 7. Locations of validation points