# SPATIAL ASSESSMENT OF EROSION AND ITS IMPACT ON SOIL FERTILITY IN THE TAJIK FOOTHILLS

*Bettina Wolfgramm[1], Bruno Seiler[2], Mathias Kneubühler[2] and Hanspeter Liniger[1]*

1. University of Berne, Centre for Development and Environment (CDE), Institute of Geography, Berne, Switzerland; bettina.wolfgramm(at)cde.unibe.ch
2. University of Zurich, Remote Sensing Laboratories (RSL), Department of Geography, Zurich, Switzerland

## ABSTRACT

Efficient planning of soil conservation measures requires, first, to understand the impact of soil erosion on soil fertility with regard to local land cover classes; and second, to identify hot spots of soil erosion and bright spots of soil conservation in a spatially explicit manner. Soil organic carbon (SOC) is an important indicator of soil fertility. The aim of this study was to conduct a spatial assessment of erosion and its impact on SOC for specific land cover classes. Input data consisted of extensive ground truth, a digital elevation model and Landsat 7 imagery from two different seasons. Soil spectral reflectance readings were taken from soil samples in the laboratory and calibrated with results of SOC chemical analysis using regression tree modelling. The resulting model statistics for soil degradation assessments are promising ($R^2$=0.71, *RMSEV*=0.32). Since the area includes rugged terrain and small agricultural plots, the decision tree models allowed mapping of land cover classes, soil erosion incidence and SOC content classes at an acceptable level of accuracy for preliminary studies. The various datasets were linked in the hot-bright spot matrix, which was developed to combine soil erosion incidence information and SOC content levels (for uniform land cover classes) in a scatter plot. The quarters of the plot show different stages of degradation, from well conserved land to hot spots of soil degradation. The approach helps to gain a better understanding of the impact of soil erosion on soil fertility and to identify hot and bright spots in a spatially explicit manner. The results show distinctly lower SOC content levels on large parts of the test areas, where annual crop cultivation was dominant in the 1990s and where cultivation has now been abandoned. On the other hand, there are strong indications that afforestations and fruit orchards established in the 1980s have been successful in conserving soil resources.

**Keywords:** soil erosion, soil organic carbon, reflectance spectra, remote sensing, Tajikistan.

## INTRODUCTION

The foothills of central Tajikistan consist mainly of easily erodable loess deposits. In the 1990s, increasing poverty triggered by the civil war and the transformation of the economy led to widespread cultivation of steep slopes, formerly used as grazing land. In these areas water erosion is considered to be the fastest and most widespread soil degradation process (1), it also has a highly negative impact on soil fertility. When the surface soil is removed through erosion, organic matter and clay particles are lost, resulting in reduced fertility, biological activity, aggregation and rooting depth (2). Examples of areas, where conservation measures have been successfully implemented, give an idea of the potential of the area with regard to sustainable land management (3). In the topic of land degradation assessments the term “hot spots” refers to areas where degradation and degradation risk are high; and the term “bright spots” refers to areas where degradation is being arrested and even reduced (4). In the context of this study hot spots indicate areas with high prevalence of soil erosion and low soil fertility. Bright spots on the other hand indicate areas with low prevalence of soil erosion and high soil fertility. Efficient planning of soil conservation measures necessitates, first, to understand the impact of soil erosion on soil fertility with regard to local land cover classes; and second, to identify hot spots of soil erosion and bright spots of soil conservation in a spatially explicit manner. Hot spots make it possible to focus soil conservation efforts, while bright spots provide an idea of the potential of the land resources and may serve as examples for successfully implemented conservation measures.

For spatial assessments, satellite imagery provides readily available information, which permits land cover classification as well as soil erosion detection. Hence, such data sets have been the base of many erosion studies in the past 30 years (5). Particularly in areas with (seasonal) low vegetation cover, the signal received by satellites is dominated by soil spectral properties and can thus be interpreted in terms of varying soil surface conditions permitting soil degradation assessments (6). Direct calibrations between satellite imagery and soil erosion incidence have advantages in areas where the application of widely used erosion models is difficult as the necessary parameters for local conditions might not have been determined. Classification tree models have been successfully applied to link ordinal classes of soil erosion from field observations and Landsat 7 (ETM+) data (7).

An important indicator for assessing soil fertility in dry ecosystems is soil organic carbon (SOC). Hill & Schütt (8) found that SOC is positively correlated to growth conditions for cereal crops in dryland agriculture and shows a strong correlation with qualitative erosion indicators. Various studies have aimed at mapping SOC contents based on satellite imagery (8,9,10). Soil spectral reflectance measured under standard conditions in the laboratory permits rapid prediction of SOC at low cost (11,12).

The aim of this study was to conduct a spatial assessment of erosion and its impact on SOC for specific land cover classes. The results provide the bases for preliminary land degradation assessments.

**Study area**

The landforms in central Tajikistan include valley floors (mainly irrigated agriculture), foothills (rainfed areas and grazing land) and mountainous areas (mainly grazing land), which succeed each other from South to North (Figure 1).

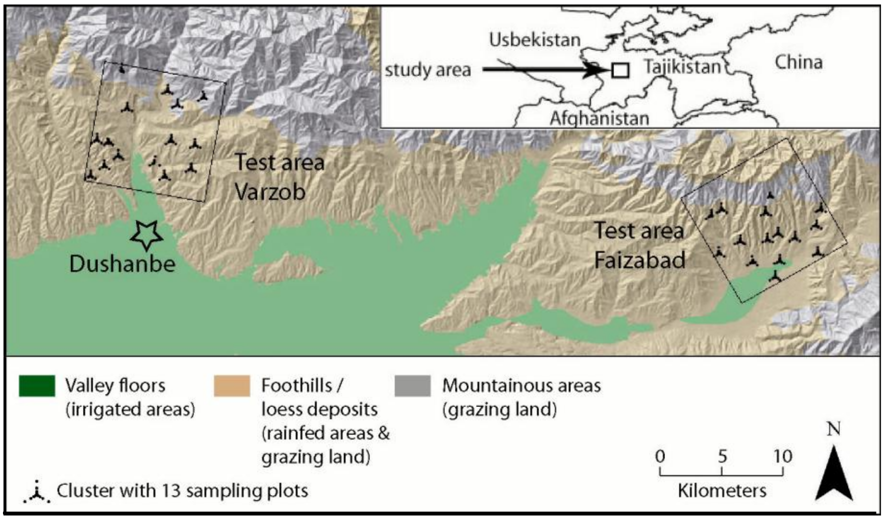


*Figure 1: Overview of Central Tajikistan with the capital Dushanbe and the test areas of Faizabad in the East and Varzob in the West, consisting of 10 km by 10 km area each. Each of these test areas contains 15 clusters, which in turn contain 13 sampling plots each.*

Within the foothills two test areas (Faizabad and Varzob) having loess deposits and similar climatic regions were selected. Soils are defined as brown carbonate by the local Tajik definition system (13). Soils dominated by granodiorit mother rock are also situated in the study area. Typical values for SOC are 1-2%. $CaCO_3$ contents vary between 2-30%, depending on the mother rock, and also the state of erosion (13). Rainfall distribution is similar in the whole area and rainfall is concentrated in the period from November through to April. Total rainfall is around 900 mm per year. Highest rainfall intensities are observed in late March and early April. The main crop is winter wheat, both on the land of state farms on the valley floor; and on hill slopes, where subsistence farmers cultivate land. The fields are prepared (ploughed or harrowed either by animal traction or with machinery) in November when the rain starts. The young wheat plants provide minimal vegetation cover during the heavy rains in March and April. Crop rotation (every 2-4 years) is carried out with flax, chickpeas and beans. These crops are sown in late March and thus do not provide any vegetation cover during the spring rains. Thus, it is in March and April when inter-rill and rill erosion processes most affect the cropland. Grazing land is the predominant land use type in the foothills. It is often common land and its use is not regulated, leading to heavy degradation of grazing lands in the vicinity of the villages.

## MATERIALS

### Raster data

The input data for spatial assessment consisted of a digital elevation model (DEM) calculated from Russian topographic maps (scale 1:50 000, contour distance 10 m) and ETM+ imagery from two different seasons. ETM+ imagery was the only readily available satellite imagery covering the study area. Despite its rather low spatial resolution, it was chosen for mapping land cover in an area dominated by small agricultural fields and a rugged terrain, in order to test its usefulness for preliminary studies. The state of maximum vegetation activity is represented by an image recorded on 24 May 2002 and the situation during the dry season (best visibility of bare ground) is represented by an image recorded on 22 August 2000. Image rectification was performed using GPS ground control points measured in the field and additional control points extracted from Russian topographic maps. Residuals in *x* direction were for both images on average 53 m, *y*-residuals amounted to 11 m for the image from 2002 and 20 m for the 2000 image data set (14). The atmospheric correction of the geo-referenced images was conducted using ATCOR3 (15). The full dataset consisted of bands 1, 2, 3, 4, 5 and 7, band indices 3/1 (iron oxide index) (16), the optimised soil adjusted vegetation index (OSAVI) (17), tasselled cap layers for brightness, greenness and wetness (18), each for both the ETM+ scenes, as well as the DEM products slope and curvature.

### Ground truth

Ground truth data for the Faizabad test area were collected in early June 2004 and for the Varzob test area in early June 2005. For efficient sampling of the full variation of land cover and soil characteristics over the study area, a ground survey campaign was conducted using a randomised systematic sampling scheme. Each test area covered an area of 10 km by 10 km and included 15 clusters; where each individual cluster included 13 observation plots (see Figure 1). In this way, a total of 400 plots were sampled. The size of an observation plot was approximately 30 m by 30 m. The extent of the area with uniform land cover surrounding the observation plot was also recorded.

The existence of visible signs of erosion processes in the field was noted as described by Stocking & Murnaghan (19). The following indicators for rill and inter-rill erosion were included: rills, pedestals, armour layer, signs of splash effects caused by raindrops, plant/tree root exposure and tree mounds. As the severity of erosion resulted in poor separability, sample locations were separated into two classes only: sites with visible signs of erosion and those without signs of erosion. Topsoil (0-20 cm depth) and subsoil (20-50 cm depth) samples were collected from each sampling plot as composite or separate samples from two sampling pits, resulting in 1040 soil samples.

Land cover characteristics, especially, fractional vegetation cover (FVC), were recorded at the stage of maximum vegetation activity after the spring rains, and in accordance with the FAO land cover classification system (LCCS) (20).

## METHODS

Various statistical methods were used to: (i) Select the calibration data set with regard to spatial independence of observations, (ii) Predict SOC content from spectral readings by building a soil spectral library and (iii) Map field observations (soil erosion incidence, SOC and land cover types). To determine state and processes of soil degradation and conservation a hot-bright spot matrix was developed.

**Building a soil spectral library for prediction of soil organic carbon (SOC)**

The soil spectral library for the prediction of SOC was established following the procedure described by Shepherd & Walsh (11) that includes the following steps: (a) Sampling of soil variability within the target area, (b) Measuring of soil spectral reflectance, (c) Determination of a reference dataset based on the spectral data space, (d) Acquiring soil property data from soil chemical analysis for the reference dataset and (e) Calibrating of soil property data to spectra applying multivariate calibration methods. Soil spectral reflectance is measured under standard conditions in the laboratory. Air-dried and ground soil samples (2 mm) were filled into Duran glass petri-dishes and illuminated by a mug-light with a tungsten quartz halogen light source. Spectral reflectance data were measured with a FieldSpec PRO FR spectroradiometer (Analytical Spectral Devices Inc.). Measurements were recorded at wavelengths of 350 to 2500 nm at an interval of 1 nm. Pre-processing of spectra included selection of wavelengths at every 10$^{th}$ nanometre and omitting bands with low signal to noise ratio. For scatter correction caused by differences in grain size, continuum removal was applied to the spectra (21). Based on representation in a principle component space, 254 samples were selected for soil chemical analysis. For SOC analysis, the samples were pre-treated with dilute HCl and subsequently analyzed by dry combustion using a Roboprep automatic C/N analyzer (Europa Scientific, Crewe, UK). The resulting soil property data was then calibrated to the continuum removed spectra. For spectral libraries including inhomogeneous soil types, regression tree models, being nonparametric, proved to be superior to linear multivariate calibration approaches (22). Regression tree models were elaborated using the CART 5.0 software (23). In order to combine trees, bootstrap aggregation (bagging) was applied with the numbers of trees to be grown set to 100 as recommended by Steinberg and Colla (24). A total of 193 samples were used as calibration samples and 61 samples (30%) were selected as holdout samples and used for validation.

**Decision tree modelling for classification of satellite imagery**

In order to determine the minimal sampling distance at which spatial independence of observations can be expected, a semivariogram analysis of the predicted SOC content, of slope measured in the field and of OSAVI (May 2002 imagery) data extracted for each sampling point, was conducted. Independence was determined for sample points at a distance of 230 and more meters. This result is highly congruent for all three datasets (SOC content, slope and OSAVI values). Subsequently, 10 samples per cluster were used as model input, and the remaining samples were used for validation.

The cultivated plots on the hill slopes are often extremely small. During field sampling, 25% of the sample plots on cultivated land showed a uniform area around the sample plot that was smaller than 30 m $\times$ 30 m. The area of 65% of the sample plots on cultivated land was between 30 m $\times$ 30 m and 100 m $\times$ 100 m, and only 10% were recorded as being larger than 100 m $\times$ 100 m (1 ha). Sample plots smaller than 30 m $\times$ 30 m were excluded from land cover classification modelling, which improved modelling results considerably. Additional point information was extracted from the satellite image for the land cover types "aquatic" ($n$=33) and "settlement area" ($n$=43). These cover types can be easily distinguished visually on the image.

Decision tree modelling, using machine learning algorithms for classification of remote sensing imagery, is a promising approach, because it is a nonparametric and hence does not require normal distribution of the histograms of raster information, in contrast to the widely applied maximum likelihood algorithm. Decision trees yield a set of rules which are easy to interpret and suitable for deriving a physical understanding of the classification process (25). Furthermore, rule-based classification allows ancillary data (e.g. topographic information) for incorporating, which can increase classification accuracy and precision (26).

To establish relationships between reference data sets collected in the field (SOC content classes, soil erosion incidence and land cover types), and raster data (satellite imagery and topographic information as described above), information from raster data was extracted for each sampling point and pixel based calibrations were elaborated. Modelling was done using the software CART 5.0 (23), which is based on binary recursive partitioning. In a decision tree estimation algorithm, the most important component is the method used to estimate splits at each internal node of the tree (25). Results of Zambon et al. (27) indicated that the gini and the class probability splitting rules are the most appropriate for image classification. In this study the gini rule was applied. The best decision tree was determined using both cross validation and the validation sample sets as test sets. The resulting models were used for classification of the raster datasets and were implemented using the knowledge classifier of ERDAS Imagine.

Soil erosion classes “incidence of erosion present” versus “no incidence of erosion”; and SOC content classes “low” (≤1.1%) and “high” (>1.1%) were modelled. According to an assessment carried out in the same areas comparing SOC contents of soils not affected by soil erosion and soils showing various states of erosion, the threshold of 1.1% SOC corresponds with the threshold determined between lightly and moderately eroded soils (28). For the SOC content model, the information of band number 3 of ETM+ imagery was excluded from modelling, to avoid unwanted influences of red coloured soils from the granodiorite mother rock. The eight land cover types identified for modelling were: Aquatic area (rivers, streams); Settlement area; Annual cropland, Cropland with permanent grass and forbs (this type includes mainly fallow cropland); Land with dominant tree and shrub cover, Grazing land with herbaceous vegetation and FVC <30%, 30-75% and >75%, respectively. Decision trees often produce several final nodes, representing sub-categories of one single land cover type. These final nodes are expected to contain detailed land cover information. The final nodes will subsequently be called land cover classes.

All models were validated by determining producer’s and user’s accuracy as well as the percentage of correct classified samples for the validation sample set as described by Foody (29).

**Hot-bright spot matrix**

A simple approach, the hot-bright spot matrix, was developed to link the soil erosion incidence and the state of soil fertility for uniform land cover classes using spatial assessments. This approach also allows the mapping of different states of soil degradation and conservation.

Land cover classes resulting from decision tree modelling were assumed to be uniform with regard to the state of the soil resources. In an initial step, the land cover classes were characterised by the field data - for each land cover class, percentages of samples with incidence of soil erosion (*x*-axis) and percentages of samples with high SOC content (*y*-axis) were plotted against each other in a scatter plot. In a second step, the same characterisation was conducted using the raster data. A GIS analysis was conducted to determine the overlaying areas between the elaborated land cover map and the soil incidence map on the one hand, and between the land cover map and the SOC content map on the other hand. The percentage of pixels per land cover class, classified as “erosion observed” and as “SOC content high” (SOC >1.1%), were derived (zonal statistic in ArcMap (ESRI Inc.)). For each land cover class, these percentages were then plotted against one another.

If a specific land cover class shows incidence of soil erosion for less than 50% of the samples (or the area) only limited erosion is expected for this land cover class, whereas land cover classes with incidence of soil erosion in more than 50% of the samples (or the area) are considered to be subject to widespread erosion processes. The same approach was applied with regard to SOC content classes. The quarters of the scatter plot thus represent: (A) Bright spots of well conserved land characterised by a non-degraded state of soil resources (high SOC contents) and limited soil erosion processes, (B) Stable areas of land subject to other degradation processes (e.g. soil nutrients exploitation) since SOC content is low but incidence of soil erosion is limited, (C) Degrading areas of land that may have been subject to land use changes and is characterised by a non-(i.e. not yet) degraded state of soil resources and widespread soil erosion processes and (D) Hot spots of degraded land, where SOC contents are already low and which is degrading further since erosion processes are widespread.

To produce hot-bright spot maps showing the four classes, namely, “hot spots”, “stable areas”, “degrading areas” and “bright spots”, the soil erosion incidence map and the SOC content map were simply combined using the combine function in ArcMap (ESRI Inc.).

Variation of SOC contents is expected to be high. 53 sample pairs collected from a single sampling plot separated by a distance of around 7 m showed that the mean coefficient of variation (*CV*) within fields is 23% for grazing land and 14% for annual and permanent cropland. To test whether differences in soil erosion incidence and SOC content between the four states of soil resources described by the hot-bright spot matrix were significant, the nonparametric Kruskal-Wallis test and the all-pairwise Connover test were carried out. Land cover classes belonging to different quarters of the hot-bright spot matrix were expected to show significant differences. Differences between the land cover classes were analysed based on field data (percentages of samples with soil erosion incidence and with high SOC content). The same tests were also carried out for the hot-bright spot maps: soil erosion incidence data and continuous SOC content values of samples attributed to a specific class of the hot-bright spot map were tested for differences. For all tests the statistical significance level was defined as $p \leq 0.05$.

## RESULTS AND DISCUSSION

### Soil spectral library for determining soil organic carbon (SOC) content

Results from the calibration of soil spectral reflectance data and reference data from SOC chemical analysis are presented in Figure 2.

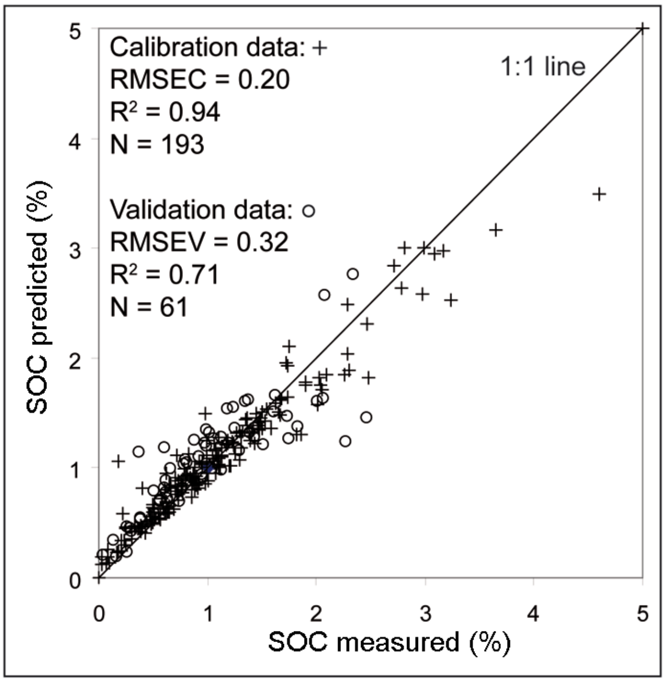


*Figure 2: Calibration of SOC to continuum removed spectra using regression tree modelling. Measured SOC contents plotted against predicted values for SOC. (Root mean square error for calibration (*RMSEC*) and validation (*RMSEV*) data, as well as coefficients of determination (*$R^2$*) and numbers of samples (*N*) for calibration and validation data are displayed.)*

The coefficient of determination for the calibration sample set was $R^2$=0.94 and for the validation sample set $R^2$=0.71. The root mean square error for the calibration set (*RMSEC*) was 0.20 and for the validation set (*RMSEV*) it was 0.32. The rather big differences in $R^2$ and root mean square error between the calibration and validation sets indicate some instability in the model. Additional calibration samples could help to improve model stability. Nevertheless, the results are promising with regard to applications of soil spectral libraries for land degradation assessments in the future.

### Soil erosion incidence map and SOC content map

A very simple decision tree model was employed to distinguish between areas with and without incidence of erosion. From the 22 raster layers used as input variables, information derived from the slope raster layer, the OSAVI of May 2002 and the August 2000 image, were most effective in separating the two states of erosion (see Figure 3, left side). For each variable, a splitting rule (a threshold), is determined by the model. A sample goes left if its value is below the threshold and to the right if its value is above. The most accurate model is the decision tree model with four final nodes. Three of the nodes classify samples showing no signs of erosion and one classifies samples with incidence of erosion. This shows that it is mainly factors controlling erosion that determine the soil erosion incidence model. The calibration data yielded 76% correctly classified samples, producer's and user's accuracy being 80 and 82% for erosion, and 69 and 67% for no erosion. The validation data set yielded 73% correctly classified samples (producer's accuracy and user's accuracy being 80 and 76% for erosion, and 61 and 67% for no erosion).

The SOC content class model is determined by 6 final nodes (Figure 3, right side). High OSAVI values on the ETM+ May 2002 imagery match with the well conserved soil resources, a fact that was observed in the soil erosion incidence model also. Tasselled cap brightness and wetness information derived from the ETM+ August 2000 imagery is subsequently decisive. It is not surprising that it is the image representing the dry season when sparse vegetation cover and high fractions of barren soil prevail, which provides most valuable information for the SOC content model. The calibration data yielded 76% correctly classified samples; producer's and user's accuracy for low SOC content are 64 and 70%, and for high SOC content are 83 and 79%. The validation data set yielded 75% correctly classified samples (producer's accuracy and user's accuracy being 62 and 72% for low SOC content, and 83 and 76% for high SOC content).

Forty-six percent of the Faizabad test area and 68% of the Varzob test area are classified as showing incidence of erosion; 42% and 18% respectively, are classified as "without incidence of erosion". The rest of the area (12% in Faizabad and 13% in Varzob) is aquatic or settlement area and is thus not included in the statistics. The difference in the level of erosion between the two test areas can be explained by a difference in their landforms: while around 20% of the area in Faizabad is characterised by a wide and flat valley floor with slopes of 5-10%, almost the entire test area in Varzob is situated on slopes >10% (see also Figure 1). Analogous to the different erosion levels in the two test areas, there is a difference in area classified as having "high" SOC content (67% in Faizabad, and only 55% in Varzob). However, large areas in Varzob, showing a low SOC content are situated along the higher mountain ranges in the North, where stony soils prevail, while in Faizabad low SOC content areas are close to the villages at medium altitude, where it can be assumed that the original fertile brown soils have been depleted.

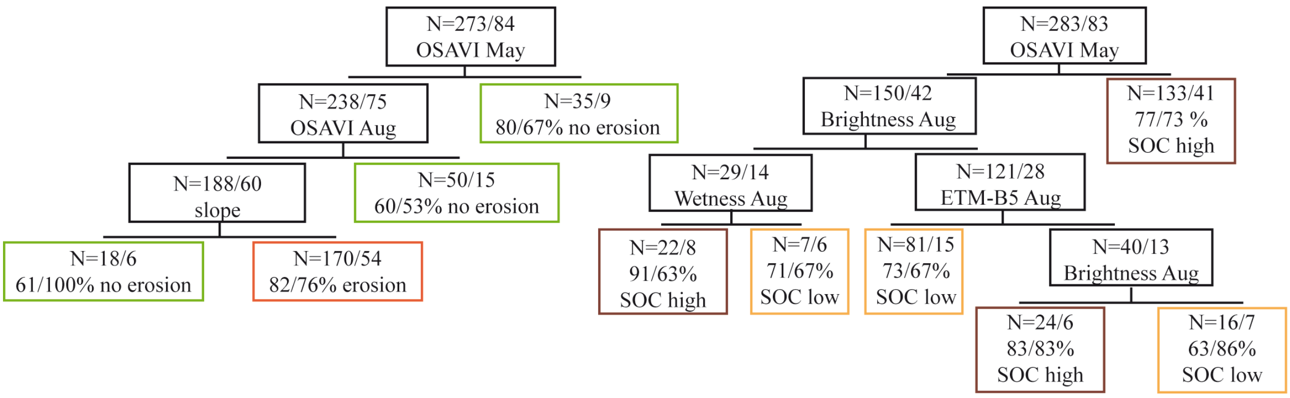


*Figure 3: Tree models for mapping eroson incidence (left side) and SOC content classes low (SOC ≤1.1%) and high (SOC >1.1%) (right side). For each splitting node, number (N) of calibration / validation samples and ideal splitter are indicated. For each final node, number (N) of calibration / validation samples and proportion of correctly classified samples for the calibration / validation sets are provided. The abbreviations for input variables used are: ETM May (Landsat 7 scene from May 2002 imagery), ETM Aug (Landsat 7 scene from August 2000 imagery), B5 (band 5), OSAVI (OSAVI), brightness (tasselled cap band 1), wetness (tasselled cap band 3), and slope (slope raster information).*

## Land cover classification

The result of the land cover classification is a decision tree with 19 final nodes. Band 4 of the May 2002 image allows splitting off aquatic area and grazing land with low fractional vegetation cover (FVC), showing river sediments and soils with high amounts of iron oxides. The iron oxide index from the August 2000 image singles out the settlement area, which is dominated by tin roofs. The rest of the samples are classified by the OSAVI information derived from the August 2000 and the May 2002 imagery into groups characterised by seasonality of FVC. On a more detailed level of the decision tree, slope steepness is also a critical parameter. Three slope categories were identified by the land cover decision tree model: ≤ 16%, 17-36% and >36%. Further, band 3 of the May 2002 image is ideal to distinguish between areas showing different percentages of bare soil. Table 1 provides details on accuracies determined from the validation data set. Accuracy levels for predicted land cover types are low, but overall land use types such as, "cropland", "grazing land", "settlement" and "aquatic area", were identified at an acceptable level of accuracy, with producer's accuracies ranging from 65% for grazing land to 89% for aquatic areas, and user's accuracies ranging from 53% for cropland to 100% for aquatic areas. 72% of the validation samples were correctly classified.

Furthermore, the final nodes of the decision tree model provide useful information on seasonality and FVC. These final nodes, the land cover classes, are used for further analysis. Validation of these land cover classes is difficult; the size of the validation data set should be considerably large to validate the large number of land cover classes (19 in this case).

*Table 1: Accuracy assessment of land cover classification based on the validation set, containing samples from test areas Faizabad and Varzob. Columns represent reference data and rows represent classification data.*

| Land cover types<br>Overall accuracy: 51% | Aquatic area | C annual | C perennial | C trees / shrubs | G FVC <30% | G FVC 30-75% | G FVC >75% | Settlement | Row total | Producer's accuracy | User's accuracy |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Aquatic areas | 16 | 1 | 1 | | | | | | 18 | 94% | 89% |
| Cropland annual | | 5 | 1 | 3 | | 2 | 2 | | 13 | 50% | 38% |
| Cropland perennial | | 3 | 3 | 1 | 1 | 2 | | | 10 | 25% | 10% |
| Cropland trees and shrubs | | | 1 | 1 | | | | | 2 | 8% | 50% |
| Grazing land FVC <30% | | | | | 1 | | | | 1 | 17% | 100% |
| Grazing land FVC 30-75% | | | 5 | 2 | 4 | 5 | 2 | | 18 | 42% | 28% |
| Grazing land FVC >75% | | 1 | 1 | 2 | | 3 | 5 | | 12 | 56% | 42% |
| Settlement | 1 | | | 3 | | | | 8 | 12 | 100% | 67% |
| Column total | 17 | 10 | 12 | 12 | 6 | 12 | 9 | 8 | 86 | | |

| Major land use types<br>Overall accuracy: 72% | Aquatic areas | Cropland | Grazing Land | Settlement | Row total | Prod. acc. | User's acc. |
|---|---|---|---|---|---|---|---|
| Aquatic areas | 16 | 2 | | | 18 | 94% | 89% |
| Cropland | | 18 | 7 | | 25 | 53% | 72% |
| Grazing land | | 11 | 20 | | 31 | 74% | 65% |
| Settlement | 1 | 3 | | 8 | 12 | 100% | 67% |
| Column total | 17 | 34 | 27 | 8 | 86 | | |

*Abbreviations are defined as follows: C = cropland, G = grazing land, FVC= fractional vegetation cover*

**Hot-bright spot matrix - state and processes of soil degradation for specific land cover classes**

In order to understand the impact of soil erosion on soil fertility with regard to local land use / land cover types the hot-bright spot matrix was assessed (Figure 4). The scatter plot on the left displays the land cover classes as characterised by the field data, while the scatter plot on the right shows information derived from the elaborated raster layers (extent of analysis being the test areas of Faizabad and Varzob). The comparison of the two scatter plots, together with the results of the Kruskal-Wallis and all-pairwise Connover test, makes it possible to identify land cover classes, from which substantial conclusions can be derived.

Depending on whether the assessment is based on field or raster data, five land cover classes have distinct differences in the soil erosion incidence and SOC content characteristics (nodes 8, 12, 13, 17, 18, circled red in Figure 4 below). The results of the Kruskal-Wallis and subsequent all-pairwise Connover test also show that for these five classes, p-values in most cases are greater than 0.05; hence the land cover classes cannot be distinguished from one another. These classes were excluded from further analysis. However, the area coverage of the five land cover classes together adds up to only 16%. Therefore the effect on the succeeding assessment is relatively small.

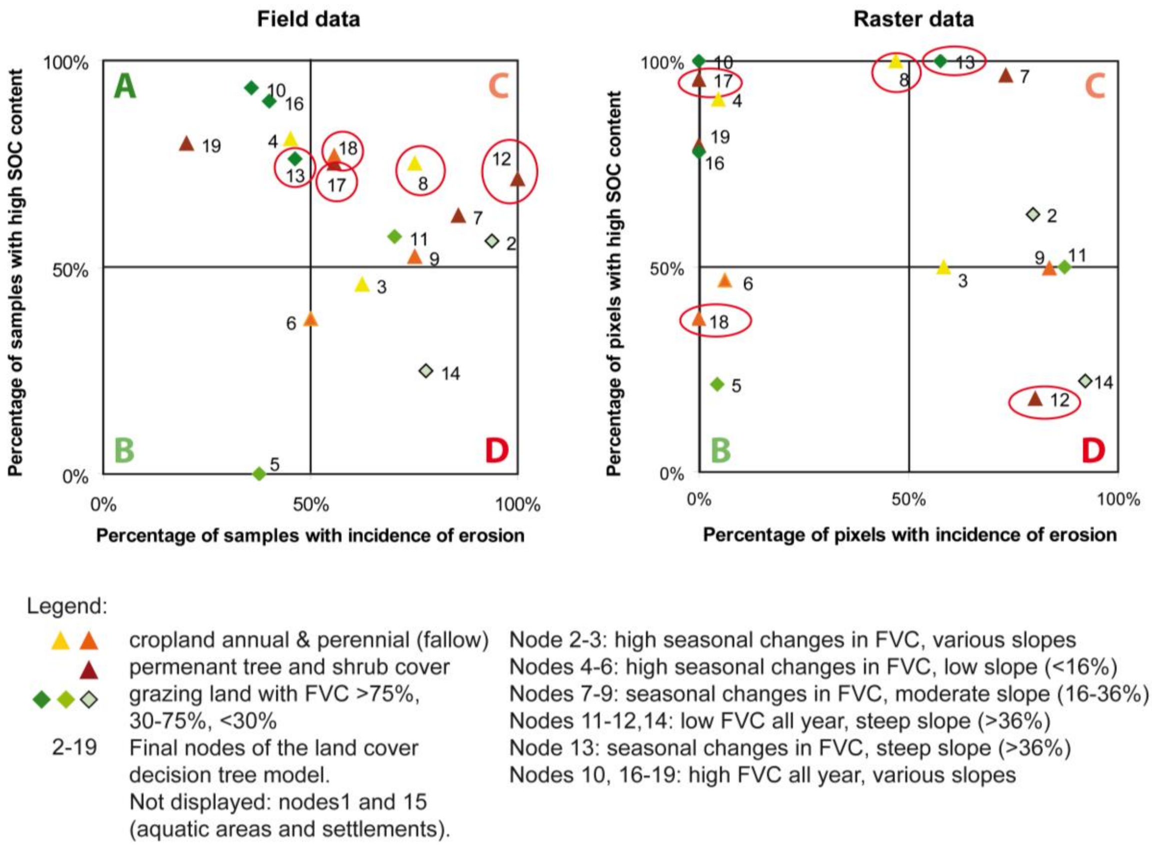


*Figure 4: Hot-bright spot matrix for land cover classes characterised by field data (left side) and by raster data (right side). For each land cover class, percentage of samples / pixels showing incidence of erosion was plotted versus percentage of samples / pixels classified as "high SOC content". Land cover classes falling into different quarters are circled red. A: bright spots, B: stable areas, C: degrading areas, and D: hot spots.*

For most other land cover classes, the application of the hot-bright spot matrix did allow for definitive conclusions; this is confirmed by the results of the Kruskal-Wallis and the subsequently conducted all-pairwise Conover tests. The land cover classes situated in quarter A (bright spots) are well distinguished, as their SOC content is significantly different from that of land cover classes situated in quarter B (stable areas) or D (hot spots). With regard to erosion, bright spots differ significantly from the degrading areas and also from the land cover class "grazing land with FVC

<30%" classified as hot spot (node 14). On the other hand, the difference from annual cropland also classified as hot spot (node 3) is not pronounced, but still identifiable for nodes 19 and 10 ($p \leq 0.09$). In this study, hot spots are generally not very well distinguishable; neither does their SOC content differ significantly from that of land cover classes classified as degrading areas ($p = 0.06$-$0.46$), nor is the difference in soil erosion incidence significant between hot pots and stable areas. Test results for the land cover classes identified as stable areas show that grazing land with FVC 30-75% (node 5) differs significantly from the other classes, but perennial cropland (node 6) does not.

In the scatter plot based on raster data, the different land cover classes are concentrated to a high degree in the respective corners of the matrix, and are therefore well distinguishable. This can be attributed to the fact that OSAVI information plays an important role in all 3 models (soil erosion incidence, SOC content level and land cover). Thus tendencies with regard to FVC in the field data are more pronounced in the raster data.

Seasonal vegetation characteristics, together with slope, provide strong explanatory factors for erosion incidence and SOC content. Land cover classes with high FVC during the time of high vegetation activity (May), and moderate FVC during the dry season (August), show the lowest percentage of area with erosion and highest with high SOC content (nodes 10, 16-19). Also, cropland situated on flat areas in the valley floor showing no erosion incidence is classified as conserved land (node 4). Of most concern are, on the one hand, areas with perennial vegetation but year round low FVC, often situated on steep slopes (nodes 11 and 14); and on the other hand, annual cropland with high seasonal changes in FVC (node 3) and perennial cropland, showing seasonal changes in FVC (node 9). Included in node 9 are degraded cropland areas now left fallow, but with no vegetation cover developing and thus being subjected to further degradation by erosion. In contrast, croplands with perennial vegetation showing at least seasonally high FVC are classified as stable areas (node 6). The fact that prevalence of erosion is low indicates that the perennial vegetation is effective in reducing erosion processes. Land cover classes, where soil fertility is high but where widespread soil erosion may strongly affect soil fertility, are grazing areas on steep slopes with moderate FVC (node 11). Also classified as degrading areas, are areas with permanent tree and shrub cover with seasonal changes in FVC (node 7), which include heavily grazed rangelands and possibly intercropping systems. Among the grazing land classes, classes with FVC<75% (nodes 2, 5, 11 and 14) are at least as much or more prone to erosion and show significantly lower SOC content than classes with FVC>75% (nodes 10 and 16). Steepness is a crucial factor, especially on grazing land, as the differences between nodes 5 and 11 show.

The most important land cover class showing high erosion and low SOC content is perennial cropland (fallow land), described by node 9. According to the zonal statistics, this class accounts for 29% land cover in the Varzob test area and for 13% in Faizabad. This reflects the often mentioned degradation of soil resources by cultivation of lands susceptible to erosion, and the subsequent abandonment of degraded fields. With regard to areas showing low erosion incidence and high SOC content, the permanent tree and shrub cover (node 19) and grazing land with FVC > 75% (nodes 10 and 16) dominate. These land cover types account for 10% of the test areas. The relatively high percentage of areas with permanent tree and shrub cover (node 19 covers 6% of the area) shows that in many cases afforestations and implementation of orchards has been very successful.

Thus, with regard to the impact of erosion on soil fertility the following major results have been obtained for the two test areas: There are indications that on cropland, SOC content may fall drastically if erosion passes a certain level, which can be concluded by comparing node 4 with nodes 3 and 9. This is most likely if seasonal fluctuations in FVC are high, indicating that FVC is not sufficient to prevent serious impact of erosion on soil fertility, especially in periods of highly erosive rainstorms (node 3). Reclamation of land is a slow process. Even if erosion processes are not observed any more on abandoned fields, SOC content remains low (node 6). For grazing land, slope seems to be an important factor for erosion incidence, but not for differences in SOC content. If a high FVC (FVC>75%) can be maintained on grazing land, SOC content is likely to remain high even if there is some incidence of erosion. While afforestation and establishment of fruit orchards transformed many dry areas into well conserved land, especially on steep slopes, these projects have in some cases not been successfully implemented or maintained and may even have caused adverse effects (node 7). Thus, plot specific soil conservation technologies and continuous efforts are crucial for such projects.

**Hot-bright spot maps**

Kruskal-Wallis tests based on field data (the validation sample set consisting of 84 samples) confirmed significant differences between the different states of degradation, but they also revealed classes which do not allow differentiation. The SOC content of samples from different classes of the hot-bright spot matrix were significantly different, with $p \leq 0.05$ for all comparisons, except for the classes "degrading" and "stable" ($p=0.067$). P-values for soil erosion incidence were between $p \leq 0.0001$ for comparison of classes "bright" and "degrading", $p=0.023$ for "bright" and "hot", $p=0.310$ for "stable" and "degrading" and $p=0.927$ for "hot" and "stable". From these results, one can conclude that the simple approach presented here allows clear differentiation between the hot and bright spots. More detailed class distinction requires more reliable information on soil erosion. Since topographic information is available at a more appropriate resolution than satellite imagery, improved results are likely to be achieved by increasing efforts in modelling topographic factors.

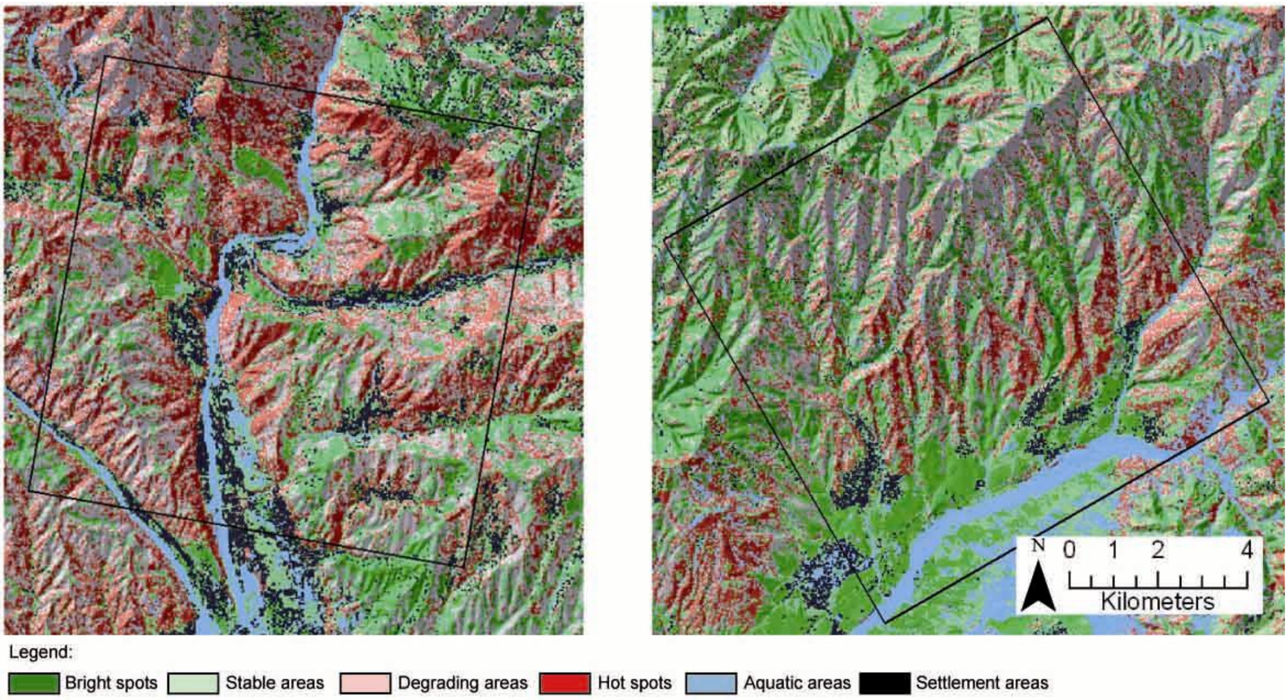


*Figure 5: The combination of the soil erosion incidence map and the SOC content class map showing bright spots (dark green), stable areas (light green), degrading areas (light red) and hot spots (red) for test areas Varzob (left) and Faizabad (right). (Blue areas on the ridges in the Northern parts of Faizabad test area, are grazing land misclassified as aquatic areas).*

Distinct patterns with regard to the distribution of hot and bright spots can be distinguished from Figure 5. Hot spots and degrading areas are wide-spread over the whole Varzob test area. The strip of land classified as hot spot along the Northern boundary of the test area is, however, determined by natural conditions, which are characterised by a mountainous area and stony soils. On the other hand, well distinguishable larger patches of bright spots, representing the afforestation mentioned above, can also be identified in various parts of the test area, showing that this conservation measure has been successful on different slopes and in different expositions. In the Faizabad test area, the well conserved annual cropland along the river in the valley floor is well recognisable. Villages are situated at the foot of the hill slopes. Many of the hill slopes in the vicinity of the villages are hot spots which can be linked to the intensive use of these areas during the 1990s.

However, there are some hill ranges which are classified as bright spots, representing afforestations and fruit orchards in the Faizabad area.

## CONCLUSIONS

SOC content information available at a relatively high spatial density is crucial for establishing a model, which allows the mapping of SOC content classes, based on satellite imagery and ancillary raster data. Prediction of SOC for a high number of samples in a cost effective way was achieved by establishing a soil spectral library for the study area. The coefficient of determination for the SOC model is $R^2$=0.71, and the root mean square error of validation RMSEV=0.32, which can be considered a sufficient level of accuracy for soil degradation assessments. For future assessments in the Tajik foothills, the spectral library elaborated would facilitate rapid and cost effective predictions of SOC from soil reflectance readings.

This study showed that in an area where difficult terrain and small cultivated plots prevailed, a spatial assessment of incidence of soil erosion, SOC content and land cover, based on a multi-date composite of Landsat ETM+ imagery and topographic information is possible at a rather low level of accuracy, acceptable for preliminary studies. Even though land cover types at a high level of detail were not detected accurately, the final nodes of the decision tree model for land cover classification contributed very useful information on seasonality and fractional vegetation cover (FVC) for the sites reflected by the specific node. Thus, machine learning algorithms for building decision trees have been very helpful in extracting information from satellite imagery, especially for areas where land cover types may be highly in-homogenous. Especially in such conditions, further research on this alternative approach for classification of satellite imagery should be conducted.

The hot-bright spot matrix developed for this study is a simple approach, allowing soil erosion information to be linked with soil fertility indicators in a flexible manner, since it may be used at various spatial resolutions and for either raster derived data (as in this study) or for field data. The approach helps to gain a better understanding of the impact of soil erosion on soil fertility for land cover classes present in a study area and to identify hot spots and bright spots in a spatially explicit manner.

The preliminary assessment of the impact of erosion on soil fertility in the study area has proved helpful for focusing conservation efforts to specific land cover types. The results show distinctly lower SOC content levels on large parts of the test areas (around 20% of the area), where annual crop cultivation was dominant in the 1990s and where cultivation has now been abandoned. On the other hand, there are strong indications that afforestations and fruit orchards established in the 1980s have been successful in conserving soil resources.

## ACKNOWLEDGEMENTS

We thank the two anonymous referees for their helpful comments on the manuscript. Keith Shepherd from the World Agroforestry Centre (ICRAF) in Nairobi provided decisive methodological inputs for the study and he and his team were very supportive with regard to any open questions. The Soil Science Research Institute of the Tajik Academy of Agricultural Science contributed to the study by responding to scientific, administrative and organizational questions. The Remote Sensing Laboratories (RSL), Institute of Geography of the University of Zurich, provided the spectroradiometer. The investigations were supported by the National Centre for Competence in Research North-South (NCCR) / Work package 4, financed by the Swiss National Science Foundation (SNSF) and the Swiss Agency for Development and Cooperation (SDC).

## REFERENCES

1 Sadikov K R, 1999. Mountain Regions of Central Asia – Sustainable Development Issues (Tajik Technical University) 298 pp.

2 Ditzler C, 2002. Quality and Erosion. In: Encyclopedia of Soil Science, edited by R Lal (Marcel Dekker, Inc., New York) 1066-1068

3 Ergashev M, G Nekushoeva & B Wolfgramm, 2007. Conversion of grazing land to fruit and fodder plots. In: Where the Land is Greener - Case Studies and Analysis of Soil and Water Conservation Initiatives Worldwide, edited by Liniger HP, Critchley W (CTA, Wageningen) 189-196

4 Ponce-Hernandez R & P Koohafkan, 2004. Methodological Framework for Land Degradation Assessment in Drylands (LADA) (Draft) (FAO, Land and Water Development Division, Rome) 48 pp.

5 Vrieling A, 2006. Satellite remote sensing for water assessment: A review. Catena, 65: 2-18

6 Haboudane D, F Bonn, A Royer, S Sommer & W Mehl, 2002. Land degradation and erosion risk mapping by fusion of spectrally-based information and digital geomorphometric attributes. International Journal of Remote Sensing, 23(18): 3795-3820

7 Cohen M J, KD Shepherd, & MG Walsh, 2005. Empirical reformulation of the universal soil loss equation for erosion risk assessment in a tropical watershed. Geoderma, 124: 235-252

8 Hill J & B Schütt, 2000. Mapping complex patterns of erosion and stability in dry Mediterranean ecosystems. Remote Sensing of Environment, 74: 557-569

9 Udelhoven T, C Emmerling, & T Jarmer, 2003. Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: A feasibility study. Plant and Soil, 251: 319-329.

10 Palacios-Orueta A & SL Ustin, 1998. Remote Sensing of Soil Properties in the Santa Monica Mountains I. Spectral Analysis. Remote Sensing of Environment, 65: 170-183

11 Shepherd KD & MG Walsh, 2002. Development of Reflectance Spectral Libraries for Characterization of Soil Properties. Soil Science Society of America Journal, 66: 988-998

12 Vagen T, KD Shepherd & MG Walsh, 2006. Sensing landscape level change in soil fertility following deforestation and conversion in the highlands of Madagascar using VIS-NIR spectroscopy. Geoderma, 133: 281-294

13 Kuteminskij V & R Leonteva, 1966. Potshvi Tajikistana [Soils of Tajikistan] (IRFON publisher)

14 Guntli D, 2006. Classification of Land Cover and Land Use - An Object-oriented Approach in Western Tajikistan (Diploma Thesis, RSL, Department of Geography, University of Zurich) 157 pp.

15 Richter R, 2005. Atmospheric/topographic correction for satellite imagery (DLR, IB 565-01/05)

16 ERDAS, 2003. ERDAS Field Guide (Leica Geosystems GIS & Mapping LLC, Atlanta) 672 pp.

17 Rondeaux G, M Steven & F Baret, 1996. Optimisation of soil-adjusted vegetation indices. Remote Sensing of Environment, 55: 95-107

18 Crist EP & RC Cicone, 1984. A physically based transformation of Thematic Mapper Data – The TM tasseled cap. IEEE transaction on Geoscience and Remote Sensing, 22: 256-263

19 Stocking MA & N Murnaghan, 2001. Handbook for the Field Assessment of Land Degradation (Earthscan) 121 pp.

20 Di Gregorio A & L Jansen, 1998. Land Cover Classification System (LCCS): Classification Concepts and User Manual (FAO) 91 pp.

21 Seiler B, 2007. Quantitative Assessment of Soil Parameters in Western Tajikistan using a Soil Spectral Library Approach (Diploma Thesis, RSL, Department of Geography, University of Zurich) 169 pp.

22 Brown D J, K D Shepherd & M G Walsh, 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma, 132: 273-290

23 Breiman L, J Friedman, R Olshen & C Stone, 1984. Classification and Regression Trees (Salford Systems) 359 pp.

24 Steinberg D & P Colla, 1997. CART: Tree-structered non-parametric data analysis. (Salford Systems). 336pp.

25 DeFries RS & JC Chan, 2000. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sensing of Environment, 74: 503-515

26 Lawrence R & A Wright, 2001. Rule-based classification systems using classification and regression tree (CART) analysis. Photogrammetric Engineering & Remote Sensing, 67: 1137-1142

27 Zambon M, R Lawrence, A Bunn & S Powell, 2006. Effect of alternative splitting rules on image processing using classifcation tree analysis. Photogrammetric Engineering & Remote Sensing, 72(1): 25-30

28 Jakutilov MR, AM Burikin, AA Sadriddinov et al., 1963. Potshvi Tajikistana - erosia potsh i borba s neji [Soils of Tajikistan - soil erosion and the fight against it] (TADJIKGOSISDAT)

29 Foody MG, 2002. Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80: 185-201