University of Zurich[UZH]

Department of Geography

Mapping and analysis of bare soil in the Swiss agricultural area using Landsat-8 data and Google Earth Engine

GEO 511 Master's Thesis

Author
Fabio Fornallaz
fabio.f@gmx.ch
11-711-801

Supervised by
Dr. Rogier de Jong
Sanne Diek

Faculty representative
Prof. Dr. Michael Schaepman

29.09.2017
Department of Geography, University of Zurich

Abstract

Soil is an important resource for many ecosystem services and is crucial for our everyday life. However, the soil functions are under increasing pressure due to environmental changes as well as the continuous population growth. Because of the importance of the resource, there is a high interest in up-to-date and consistent soil information from local to global scale. Traditional soil sampling methods to produce this information are not adequate anymore. Though, newer technologies like remote sensing can help to close this gap of information.

Previous soil remote sensing studies were often facing the problem that most agricultural fields were covered by vegetation and the amount of bare soil in the image was limited. In this thesis, we developed a new method that avoids this problem by processing a longer time series of Landsat-8 satellite data with the help of the Google Earth Engine (GEE). First, we calculated the so-called brownest pixel composite (BPC) over the agricultural area of the Swiss Plateau. This composite shows the "brownest" moment of every pixel in the time series. Afterwards, a threshold was defined to distinguish this brownest moment into either bare soil or any other land cover type. The spectral reflectance of all these bare soil pixels were then used together with soil samples from the harmonized soil database (HSD) to create multiple linear regression (MLR) models. Finally, the models were used to predict the percentages of sand, silt, clay, and soil organic matter (SOM) in the topsoil of all the bare soil pixels in the study area.

The results showed that the developed method can successfully be used to maximize the bare soil coverage over a large study area. The MLR models showed valuable results for the prediction of the soil properties clay, sand, and SOM. For silt, the performance of the MLR model was not sufficient. The predicted soil property maps have a higher spatial resolution than any other existing soil maps and the predicted spatial patterns of the soil properties are feasible. The methods developed in this thesis can be used for several applications and we show the potential for upscaling the products to a European or even worldwide level.

List of Figures

List of Tables

List of Abbreviations

APEX	Airborne Prism Experiment
BI	Bare Soil Index
BPC	Brownest Pixel Composite
BPI	Brownest Pixel Index
BSC	Bare Soil Composite
DSM	Digital Soil Mapping
GEE	Google Earth Engine
HSD	Harmonized Soil Database
HYSOMA	Hyperspectral Soil Mapper
MLR	Multiple Linear Regression
NABO	Swiss Soil Monitoring Network
NDSI	Normalized Difference Snow Index
NDVI	Normalized Difference Vegetation Index
NIR	Near-Infrared
OLI	Operational Land Imager
R^2	Coefficient of Determination
RMSE	Root-Mean-Square Error
SR	Surface Reflectance
SOM	Soil Organic Matter
SWIR	Short-Wave-Infrared
TOA	Top-Of-Atmosphere

1. Introduction

The world's ecosystem provides us every day with crucial goods and services worth as much as the gross national products of all the world's economies. More than half of it arises on land, where soil plays a major role (Brady & Weil, 2017). Soil contributes to ecosystem services like the food production, the purifying and storage of water, the sequestration of carbon, or the prevention of land degradation (Sanchez et al., 2009; Amundson et al., 2015; Ballabio et al., 2016). However, the global environmental changes increase the pressure on soil and lead to an alteration and reduction of its provided services (Mulder, 2013). This is problematic, as for example, the agricultural system need to produce more and more food to meet the requirements of the increase in population and consumption (Foley et al., 2011). Based on the importance of these services, and the fact that degraded soil is not renewed easily, there is a high interest in having up-to-date and consistent soil information from a local to global scale (Dewitte et al., 2012). Unfortunately, soil information is in many regions either missing, not available in the required resolution, not up-to-date, or not providing the required soil property information (Nussbaum, 2017).

1.1 Available Soil Maps

Omuto et al. (2013) provides an overview about the available soil products from a global to regional and local scale. These global and regional products were mainly produced with soil data collected between the 1960s and 1990s. Thus, since then there is a gap of soil information and the recent products are not up-to-date. There are new projects to produce updated global soil maps with recent technology like the GlobalSoilMap. This project has the aim to produce a standardized, high-resolution digital soil information system that shows selected soil properties for the entire world (Arrouays et al., 2014). However, these global products are not useful for gaining information in a heterogeneous country like Switzerland, as the spatial resolution is too coarse. Therefore, it is relevant to produce own soil maps that are suitable for the diverse conditions. In Switzerland, different soil maps have been produced, from a national one on a scale of 1:200'000, to regional and local ones on a scale of 1:25'000 and 1:5000 (Bonnard, 1999). However, the latter, large-scale soil maps are just available in a few cantons of Switzerland. The national soil map derives information about the soil type, but also about soil functions like the water storage capacity, the root penetration depth, or the nutrient storage capacity (Bundesamt für Landwirtschaft, 2012). These soil functions are interesting for farmers to improve their decisions about crop management practices (Ge et al., 2011). However, as already stated above, the resolution of the national soil maps is too coarse to provide information of the different soil properties in an in-field resolution. Besides, most soil maps in Switzerland were produced before the 1990s and are therefore, depending on the desired soil property, partially not up-to-date (Wulf et al., 2015). For farmers and other stakeholders, an up-to-date and high resolution soil

map would be of high interest. Additionally, these fine scale and recent soil products could also contribute to land management models and improve the output the models provide (Gomez Giménez et al., 2016).

The traditional soil sampling methods used for the first products are not appropriate for updated soil maps, as they are very time consuming, costly, and often not standardized (Wulf et al., 2015). Newer technologies like remote sensing can help to close the gap of information. It can provide up-to-date soil information over extended areas in a high spatial resolution and help to increase the soil property products in Switzerland.

1.2 Remote Sensing and Soil Maps

Analysis of soil with remote sensing techniques started in the 1970s. The first analyses were mainly done with field-spectrometers and airborne images (Mulder et al., 2011). These techniques are limited to smaller areas. To provide soil information on a larger area, the use of satellite remote sensing techniques is crucial. Several studies to predict soil parameters by satellite images have already been conducted and have shown promising results.

Demattê et al. (2007) predicted different physical and chemical soil attributes in an area in Brazil using a single Landsat-7 image together with soil samples. The multiple linear regression (MLR) model between the spectral reflectance and the soil properties showed best results in indicating clay and sand. Nanni et al. (2012) was using similar methods to discriminate soil pixels into the different soil classes in a study area in Brazil. The results showed that 14 out of 16 soil classes could get predicted with a success rate of > 40%. Similar studies with satellite sensors were also conducted by Shabou et al. (2015) and Fiorio & Demattê (2009).

All these studies show promising results, but they all have the issue that acquiring soil reflectance spectra from the air and space is limited by the coverage of vegetation. Around 56% of the global land areas are covered by green vegetation. The remaining areas are covered by dry vegetation, snow, urban, and bare soil (Ben-Dor et al., 1999). Hence, just a very limited amount of bare soil is visible during one acquisition by remote sensing sensors. The calculation of soil properties is therefore limited to the amount of bare soil visible in the image. This reduces the usability of remote sensing techniques in soil science for generating large and continuous soil maps that show the spatial variation of different soil properties.

To avoid this problem, recent studies have started to use multi-temporal images from the same region and fuse the bare soil areas together. Demattê et al. (2016a) analyzed five satellite images from five consecutive years over the same area and captured in the same season. With this method they could increase the total bare soil area from 36% of a single image to 85% for the fused image. Diek et al. (2016) took a similar approach with airborne imaging data. They combined three airborne images from the same spot on different dates and could double the amount of bare soil pixels compared to

one acquisition. These studies showed that the use of multi-temporal remote sensing images increase the amount of captured bare soil information. However, images from airborne systems are limited in temporal resolution and a further increase to the maximum amount of bare soil in a larger area would be very costly.

This thesis pursues the approach of multi-temporal images and tries to maximize the bare soil coverage over a large area. The study area is the whole Swiss Plateau (around 11'000 km^2), an intensely used agricultural area of Switzerland. This will also give the opportunity to calculate different soil properties over a large and continuous area in a high spatial resolution. The chosen soil properties for the prediction are the three soil textures sand, silt, and clay, as well as the soil organic matter (SOM) content. These properties play a major role in the behavior of soil and have a big influence on soil characteristics like the water storage capacity (Ben-Dor et al., 1999). In addition, the prediction of these properties by satellite images have already shown promising results, as described above.

To analyze multi-temporal satellite images over this large area, the Google cloud platform "Google Earth Engine" (GEE) was used. GEE gives the possibility to process hundreds of images over large areas in a reasonable computation time (Gorelick et al., 2017). GEE stores the complete archive of all the satellite images from the Landsat program, as well as from other satellite and airborne missions. Therefore, it is well suited for the purpose of this thesis.

1.3 Research Questions

The thesis is structured mainly in three parts. Firstly, a so-called "brownest pixel composite" (BPC) is calculated for the study area. This product will be the counterpart to the well-established greenest pixel composite. The greenest pixel composite shows the maximum of vegetation coverage over a certain area. In comparison, the aim of the BPC is to show the minimum amount of vegetation coverage over a certain area. Hence, over agricultural areas it should uncover the reflectance of bare soil. The main question to be answered in the first part is:

- Which method is suitable to generate a brownest pixel composite over a large area?

The second part of the thesis is about defining a threshold to distinguish the BPC into bare soil and other land cover types. The threshold is calculated with additional information from airborne imaging spectrometer data and from fieldwork. The questions to be answered in the second part are:

- What is the optimal threshold to distinguish bare soil from other land-cover types?
- What is the maximum extent of bare soil in the study area of the Swiss Plateau during the three years of investigation?

In the final part of the thesis, the reflectance values from the bare soil pixels are used together with soil sample data to create MLR models. The models are then used to predict soil properties over the bare soil area in a spatial resolution of 30 m. The question to be answered is:

- Can remote sensing techniques help to improve the available soil products and predict structural soil properties over the agricultural areas?

2. Study area & Data

2.1 Study area

The study area (Figure 1) consists of the Swiss Plateau, one of the three major regions of Switzerland. The Swiss Plateau is located between two other regions consisting of the Jura Mountains in the north and the Swiss Alps in the south. With a length of around 300 km and a width between 30 and 70 km the Swiss Plateau covers around 27% of the area of Switzerland. The Swiss Plateau is delimited in the southwest by the Lake Geneva and in the northeast by the Lake Constance. Other big lakes within the Swiss Plateau are in the northwest: Lake Neuchâtel, Biel, and Murten and in the east mainly the Lake of Zurich.

The Figure 1 shows, beside the three major regions, also the distribution of the suitability for cropland. The areas in red show the highest suitability, followed by the areas in orange. Nearly all the good production area for cropland lies inside the Swiss Plateau, which is the main agricultural area in Switzerland. Nearly 50% of the area is used for agricultural activities. For these reasons, it has been chosen as the study area, as the agricultural area shows bare soil between the cultivation. The main crop types in the study area are the winter crops barley, wheat, and triticale, and the summer crops maize silage and corn, sugar beet, potatoes, and rape (Franzen et al., in preparation). Additionally, temporary grassland is commonly included in the crop rotation (Gomez Giménez et al., 2016).

The altitude of the Swiss Plateau varies between 500 and 1500 m a.s.l. with an average of 580 m. The climate in the study area varies due to the large extent of the Swiss Plateau and due to the differences in altitude. The climate in the area is classified as Maritime Temperate or Oceanic climate. The average annual precipitation is 1000 mm with regional variations between 800 and 1400 mm. The mean temperature lies at around 0 °C in January and reaches up to around 20 °C in July. The western part of the Swiss Plateau shows slightly higher temperatures and less annual variation in precipitation. In addition, there are also less snow days in the western than the eastern part (MeteoSchweiz, 2013).

Geologically, the Swiss Plateau shows a crystalline basement that is covered by Mesozoic sediments. On top of these sediments lies the important Molasse layer that was formed during the Alpine origin and consists in accumulated materials from the Alps. The present landscape was afterwards formed

and shaped by the ice age glaciers which lead to the typical landforms throughout the Swiss Plateau (Gnägi & Labhart, 2015).

The soils in the Swiss Plateau are mainly classified as different kinds of Cambisols. These soils are very well suited for agricultural activities. Towards the Alps, the poorly developed Regosols occur more often. In the western part of the Swiss Plateau, as well as around the lakes and rivers, the soils are also often classified as Luvisols, Fluvisols, or Histosols (Bundesamt für Landwirtschaft, 2012; IUSS Working Group WRB, 2015).

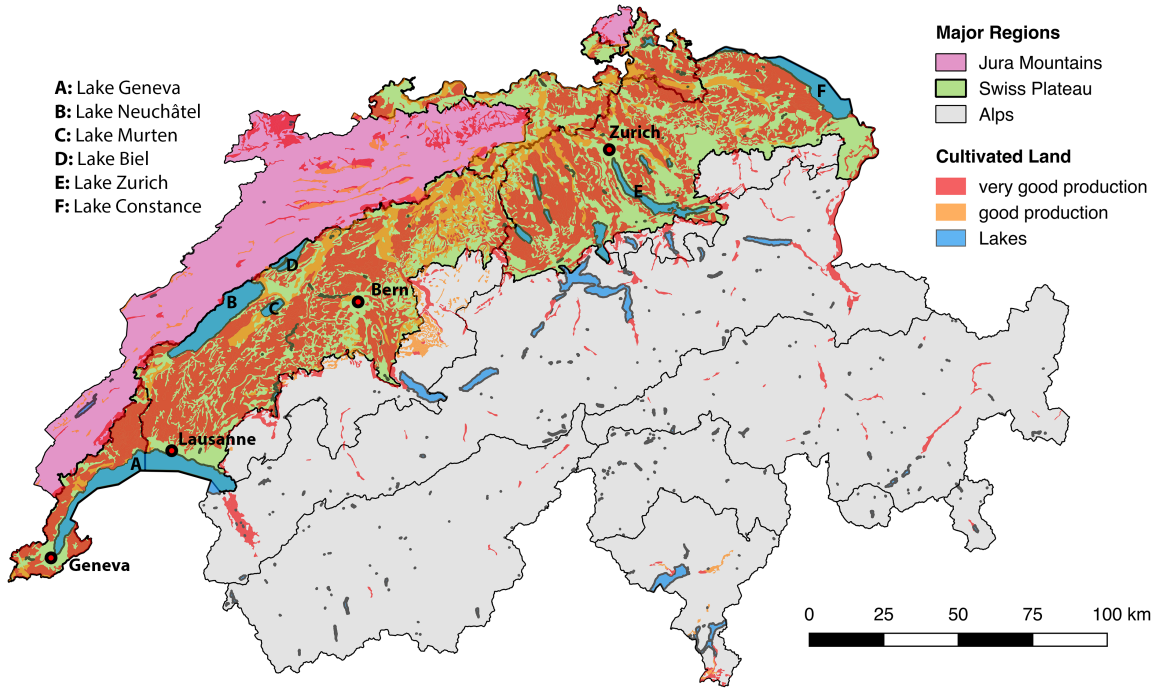

Figure 1: Overview of Switzerland showing the three major regions and the suitability for cropland. The study area is shown in green together with the labels of important lakes and cities.

2.2 Satellite Data

For this study, the USGS Landsat-8 Operational Land Imager (OLI) data was used. Landsat-8 is an optical multispectral satellite launched in February 2013 with the first images available from April 2013. Table 1 shows the different OLI Bands, the wavelength region, and the spatial resolution (USGS, 2016). One Landsat-8 scene covers an area of approximately 170 km north-south by 185 km east-west. The study area of the Swiss Plateau gets covered by five flight strips (Path/Row: 194/27, 195/27, 195/28, 196/27, 196/28). The entire earth surface is recorded once every 16 days. For regions with overlapping scenes, the temporal resolution is higher. Figure 2 shows the amount of images taken over the Swiss Plateau in the study period of 01/01/2014 until 01/01/2017. The different flight

strips can be clearly seen. Some regions show overlapping areas with double the amount of images, others show just one image every 16 days. In the western and eastern end of the Swiss Plateau, as well as in the middle we have an area with just 50-56 images from the three years of the study period (bright blue). Between these areas we have overlapping flight strips with around 99 to 106 images. In the small dark blue spot 160 images are available.

Table 1: Overview of the Landsat-8 OLI bands with the corresponding wavelength region and spatial resolution

Landsat-8 OLI Band No.	Band Name	Wavelength [µm]	Spatial Resolution [m]
1	Coastal / Aerosol	0.44 – 0.45	30
2	Blue	0.45 – 0.51	30
3	Green	0.53 – 0.59	30
4	Red	0.64 – 0.67	30
5	NIR	0.85 – 0.88	30
6	SWIR1	1.57 – 1.65	30
7	SWIR2	2.10 – 2.29	30
8	Pan	0.50 – 0.68	15
9	Cirrus	1.36 – 1.38	30

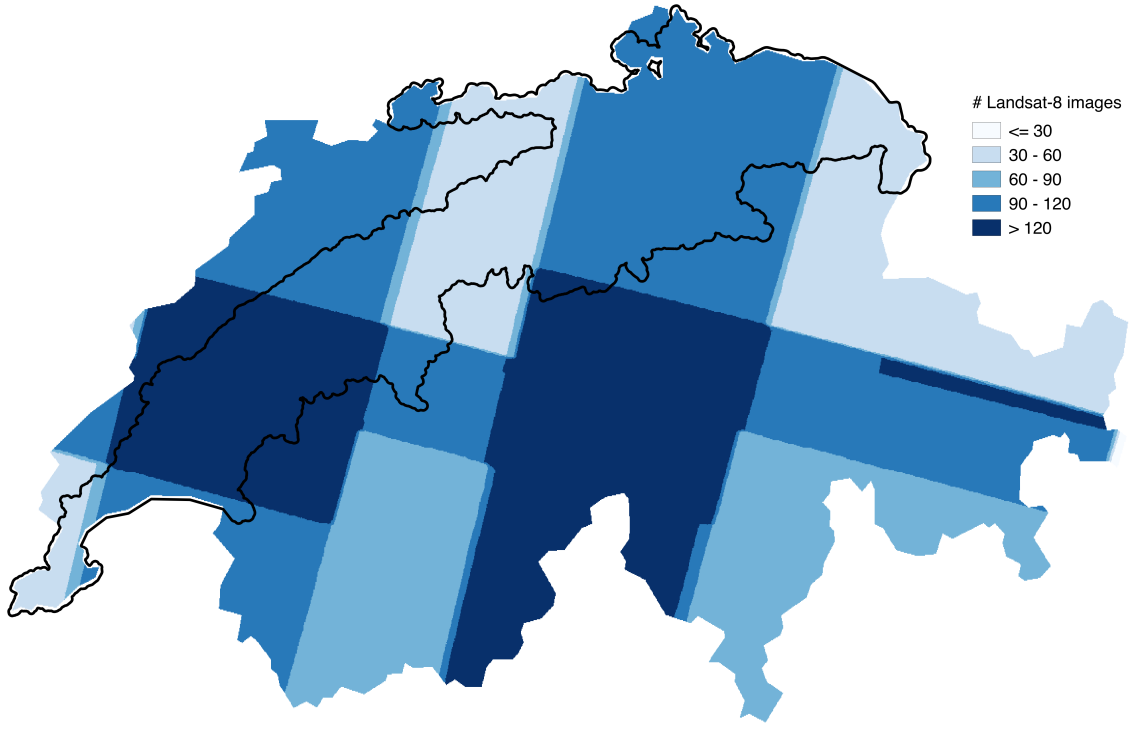

Figure 2: Number of Landsat-8 images available in the study period of 01/01/2014 to 01/01/2017. The black line represents the study area of the Swiss Plateau.

2.3 Airborne Imaging Spectroscopy

To calculate the threshold, different images from the Airborne Prism Experiment (APEX) were used. APEX is an airborne imaging spectrometer that covers the spectral range from 372 – 2540 nm in up to 532 spectral bands (Schaepman et al., 2015). Table 2 shows the different flight strips used for the threshold calculation. All scenes were orthorectified and atmospherically corrected surface reflectance scenes with a spatial resolution of 2 m (Hueni et al., 2009). The APEX scenes were further processed using the default thresholds of the Hyperspectral Soil Mapper (HYSOMA) software (Chabrillat et al., 2011). The software selects pixels that are spectrally bare soil by masking water and (active and residual) vegetation pixels. These pixels are detected by different spectral indices like the Normalized Difference Vegetation Index (NDVI), Normalized Difference Red Blue Index, or Normalized Cellulose Absorption Index (Chabrillat et al., 2011). The result is a binary image that classifies every pixel into bare soil, or any other land cover types. Built-up areas like streets or houses are also classified as bare soil, as they were not excluded in the water and vegetation mask. These misclassified pixels were later excluded in GEE by drawing polygons around the bare soil fields. Additionally, for some scenes also the corresponding RGB image was analyzed (Band 39 for Red, 17 for Green, 06 for Blue) as a reference for calculating the threshold.

Table 2: APEX scenes used in this thesis.

Location	Flight Date	Binary bare soil scene	RGB composite
Eschikon	18/07/2014 10/04/2015	Available Available	Not Available Available
Oensingen	21/04/2015 24/06/2015	Available Available	Available Available
Greifensee	18/07/2014 10/04/2015 10/07/2015	Available Available Available	Not Available Available Not Available

2.4 Soil Data

For calculating the soil properties, the Harmonized Soil Database (HSD) from the Swiss Soil Monitoring Network (NABO) was used (Rehbein et al., 2011; Walthert et al., 2016). The database contains evaluated soil samples from seven different projects in Switzerland. Four of these projects have been carried out in the canton of Zurich, two in the canton of Berne, and one in the forest area of Switzerland (Figure 3). The samples were taken between the year 1960 and 2014, with the most in the 1990s (see Figure 4). During the generation of the HSD, all data from the different sources were combined and samples with insufficient quality have been removed (Walthert et al., 2016).

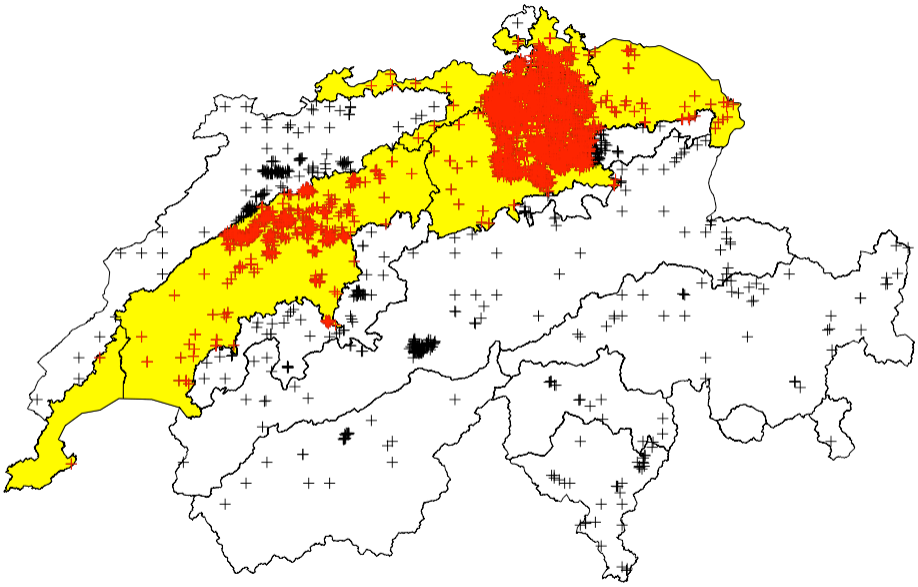

Figure 3: Soil sample locations from the HSD. Every sample is represented by a "plus" sign. The study area is shown in yellow and soil samples inside the study area are shown in red.

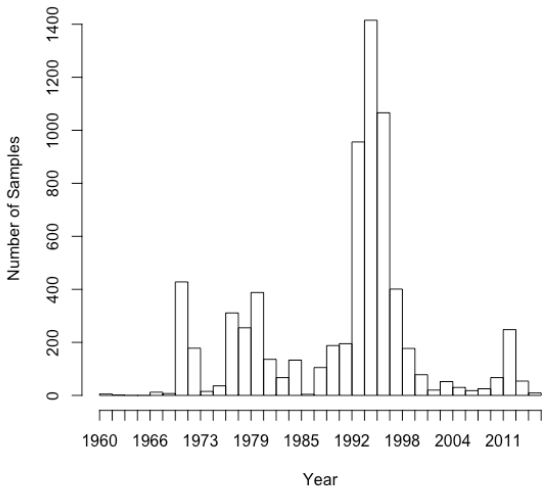

Figure 4: Year of sampling of the soil samples in the HSD.

To fit the need of this study, the HSD was further filtered for a) soil properties, b) topsoil, c) lab measurements, d) missing values, and e) geographic area.

a) The database was filtered for the soil textures sand (2 - 0.2 mm), silt (0.2 - 0.002 mm), clay (<0.002 mm), as well as for SOM. All other soil parameters were removed.
b) Optical remote sensing satellites will just receive reflectance from the top part of the soil, as the sun penetrates the soil just to 50 µm depth (Ben-Dor et al., 1999). Thus, the database needs to get filtered for topsoil. The definition of topsoil varies between references. Most commonly, in other soil remote sensing studies, it is defined as the top 20 cm of soil (Demattê et al., 2016b). Therefore, the database got filtered for soil samples in the range from 0 – 20 cm depth. A smaller topsoil range would also strongly decrease the amount of available samples.
c) Walthert et al. (2016) showed that there is a difference in the database from the estimation of the soil properties in the field and the lab. The field values were estimated by expert judgments and are less accurate than lab measurements. Analysis of the differences showed root mean squared error (RMSE) values between 7.14 - 15.42 %. To reduce this uncertainty, lab measurements were preferred.
d) Many soil samples in the HSD did not contain the percentage of all the required soil textures. For example, in the canton of Zurich, the percentage of sand was not included. As the added percentage of sand, silt, and clay should always be 100%, the missing values were calculated out of the remaining two soil textures.
e) Furthermore, the database was filtered to the geographic area of the Swiss Plateau.

After these preprocessing steps a total of 4783 sample points, each from a different location, were left (see Table 3). For 40% of these samples, the amount of soil textures as well as the amount of SOM is available. For 5% of the samples we have only information about the soil textures. The remaining 55% contain just information about the percentage of SOM in the soil.

Table 3: Amount of soil samples left after the filtering steps.

No. of samples	Soil Texture + SOM	Only Soil Texture	Only SOM
4783	1927 (40%)	235 (5%)	2621 (55%)

The proportion of the three soil textures, sand, silt, and clay are used to classify the soil into different texture classes. These classes again can be used to characterize the soil's behavior in water storage capability, porosity, or fertility, which are high important parameters for the plant growth (Ben-Dor & Dematthê, 2016; Shabou et al., 2015). The soil texture classes of the samples of the Swiss Plateau can be seen in the soil triangle of Figure 5. Most of the soil samples are classified as loam, clay loam, or sandy loam.

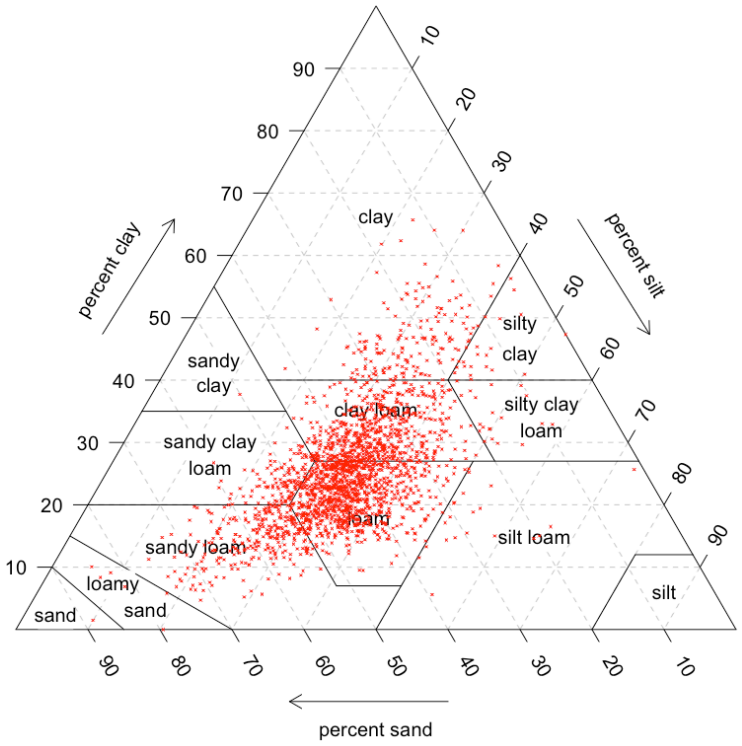

Figure 5: Soil texture triangle of the remaining soil samples. Each side of the triangle represents the percentage of one of the three textures sand, silt, and clay. Each sample is represented by a red dot and the location of the dots represents the soil textural class.

3. Methods

3.1 Brownest Pixel Composite for Analysis

The BPC was calculated in two different ways. First, as the main product, the BPC for the study area was calculated in a spatial resolution of 30 m with the Landsat-8 surface reflectance (SR) images. This product is later being used for the calculation of the threshold and the different soil properties. Second, the BPC was calculated for entire Switzerland in a spatial resolution of 15 m with the Landsat-8 Top-of-Atmosphere (TOA) images. This product was calculated for visualization purposes only.

The reasons for those two different products are the following. For the calculation of soil properties we are just interested in the reflectance values of bare soil. The soil data from the HSD is mainly available in the area of the Swiss Plateau. Hence, we limited the calculation of soil properties to this area to reduce uncertainties in the prediction. Additionally, the atmosphere has an influence on the reflectance of the ground. To reduce this effect, SR images were preferred. The SR images in GEE do not contain the panchromatic band with a spatial resolution of 15 m. Thus, the product is calculated in a spatial resolution of 30 m.

Aside of the calculation of the soil properties we are also interested to show the maximum extend of bare soil in Switzerland as a counter product to the well established greenest pixel composite. For this reason, the second product was calculated for entire Switzerland. To increase the visual experience, the product was calculated in the highest possible spatial resolution of 15 m. This was just possible with the use of the panchromatic band that is just available in TOA images. Thus, for each of these two products different satellite data were used and the working steps slightly differ. The differences can be seen in the work flow diagram in Figure 6. The green part explains the working steps for the main product, the BPC of the Swiss Plateau. It also describes the index that is used for the calculation of the BPC. The same index is also used in the second product but is not explained again. The second product for visualization purposes is shown in blue in the work flow diagram. The different working steps for the second product are described afterwards, starting in section 3.2.

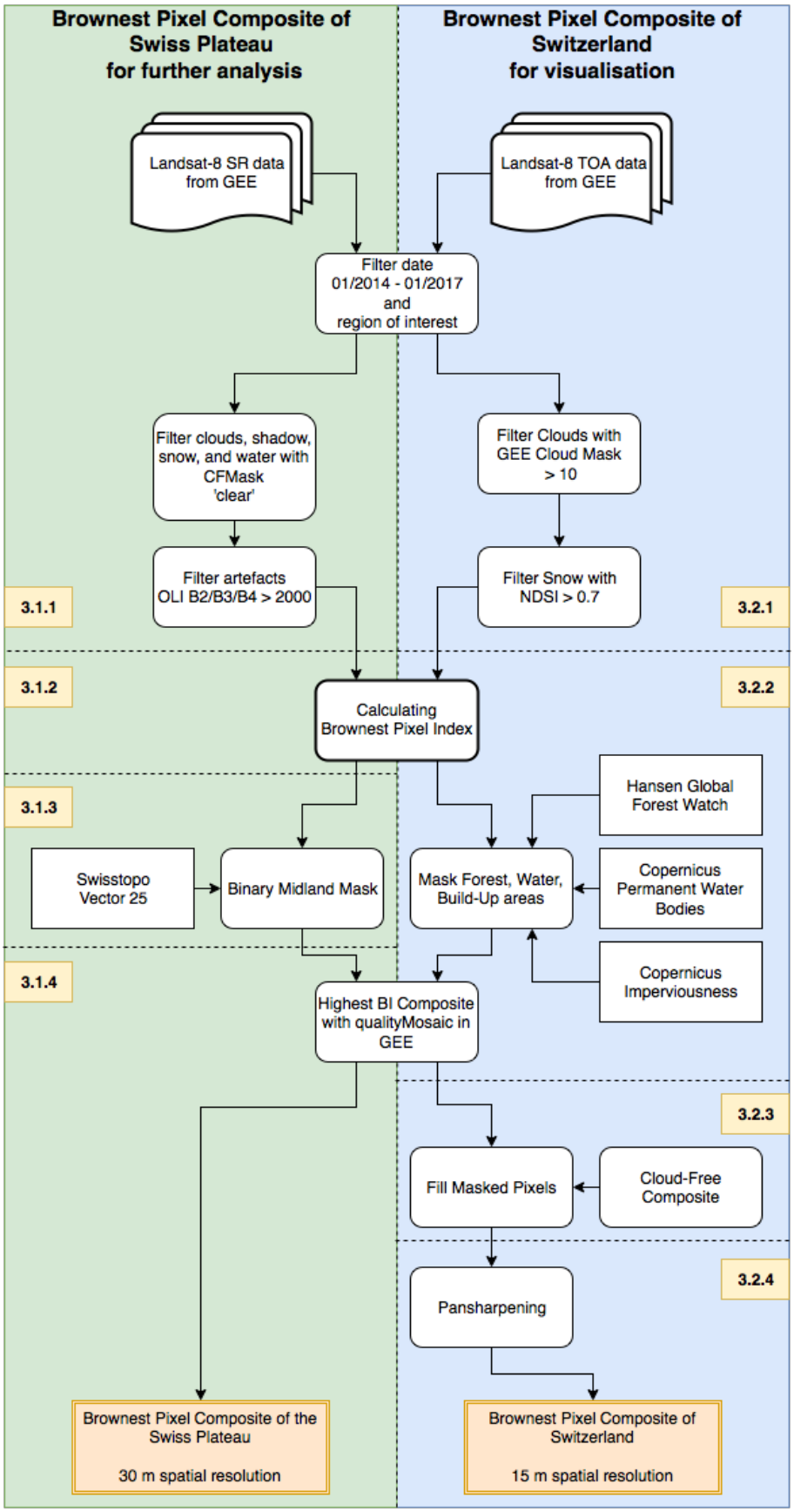

Figure 6: Working steps for the generation of the BPC. The left, green side shows the working steps for the BPC for analysis. The right, blue side shows the working steps for the BPC for visualization.

3.1.1 Preprocessing

The Google Earth Engine platform was used for the whole process of generating the BPC. Landsat-8 OLI SR product was chosen for the analysis. To reduce the amount of data, the Landsat-8 SR collection was filtered by date (01/01/2014 until 01/01/2017) and location (boundary of Swiss Plateau). In Switzerland it is mandatory to perform crop rotation. This means that the planted type of crop changes each year and also that different fields are fallow on different times. The longer the time series of satellite data, the higher the chance to detect each field on a fallow moment. With a time span of three consecutive years the different crop rotation periods should get covered.

The SR product was initially calculated from the TOA reflectance data. During those processing steps, a new band called *CFMask* was added. CFMask is an algorithm that classifies every pixel into one of the following five classes: clear, water, cloud shadow, clouds, and snow (USGS, 2017). We are only interested in clear pixels over land. Thus, all pixels except the class 'clear' were masked and removed.

Visual analysis of the remaining images showed some outlier pixels. These showed very high reflectance values in all bands. In the USGS Product Guide (2017) this problem is described. The algorithm has some issues when there is a high temperature difference between the ground and the cloud. In addition, challenging are also thin clouds and very bright targets. To reduce these artifacts, another filter was applied. This filter detected all the pixels with extraordinary high reflectance values over 2'000 in the OLI RGB bands 2, 3, and 4. The value of 2'000 for this filter was chosen after analysis of the outlier pixels. After these preprocessing steps, the image collection was ready for further analysis.

3.1.2 Brownest Pixel Index

A spectral index was chosen for the calculation of the BPC. An index has the advantage that it is calculating the proportion between different wavelength regions and is therefore not sensitive to the absolute amount of reflectance. This is crucial when analyzing a longer time-series of satellite data as the ground is exposed to different meteorological conditions. For example, higher moisture on the ground leads to lower reflectance values in the whole spectra that would have a strong influence on a calculated value (Nocita et al., 2013). With the use of an index this and also other influences can get reduced.

The brownest pixel index (BPI) should be sensitive for bare soil. Several indices exist already to discriminate bare soil from other land cover types. These indices face often the challenge that built-up areas show similar spectral response as bare soil. Thus, they are often created to be able to discriminate between these two land cover types. In this thesis, the built-up area will get removed with additional data, so that just agricultural area is left for analysis. Therefore, the BPI must not be able to discriminate between built-up and bare soil area. Piyoosh & Ghosh (2016) give an overview about existing bare soil indices. The BPI should be an easy to apply index that is applicable on

Landsat-8 OLI data. Moreover, the calculation should be possible over larger area and in a longer time-series. Two different indices were selected as BPI and are described below. The results were compared and the most suitable one was selected as the BPI for the final product.

Normalized Difference Vegetation Index

The NDVI is widely used for the monitoring and assessment of vegetation activity (Lillesand et al., 2015). It is also used for calculating the greenest pixel composite, showing the highest amount of vegetation activity over a given time period (e.g. Landsat 8 Annual Greenest-Pixel TOA Reflectance Composite in GEE). NDVI values below 0.2 correspond to water, snow, or non-vegetated areas like bare soil. Thus, the lowest NDVI composite should be able to detect the bare soil area over a given time period. Equation 1 shows the calculation of the NDVI.

$$NDVI = \frac{R_{NIR} - R_{red}}{R_{NIR} + R_{red}}$$

Equation 1: Calculation of the NDVI where R is the reflectance of the NIR and red spectral region.

Bare Soil Index

The Bare Soil Index (BI) is mainly used in the field of land cover classification. It is used to differentiate between bare soil and other land cover types and was introduced by Rikimaru et al. (2002). The BI enhances the NDVI with adding the blue and SWIR wavelength region, which can be seen in Equation 2. The BI was already successfully used in forest research (Jamalabad & Akbar, 2004), as well as to map bare soil areas (Zhao & Chen, 2005). For the SWIR region, the OLI band 7 (SWIR2) was chosen, as it lies in the region of the clay absorption features (Ben-Dor & Demattê, 2016).

$$BI = \frac{(R_{SWIR2} + R_{red}) - (R_{NIR} + R_{blue})}{(R_{SWIR2} + R_{red}) + (R_{NIR} + R_{blue})}$$

Equation 2: Calculation of the BI where R is the reflectance of the SWIR2, NIR, red and blue spectral region.

3.1.3 Selecting Agricultural Areas

The calculation of the BPI is just feasible over agricultural areas, as forest, water, and built-up areas never show bare soil. To mask these land cover types, a binary mask of the Swiss Plateau was used. This binary mask was created with the national map data from swisstopo (Bundesamt für Landestopografie swisstopo, 2008). All land covers except the following were masked and removed: cropland, (permanent) grassland, orchards, vineyards, and tree nurseries.

3.1.4 Generation of the Composite

After the masking steps, the BPC of the remaining pixels was calculated using the *qualityMosaic* function in GEE. This function chooses in the pixelwise time series the date with the highest value in the chosen band. For the BPC, the BPI value was selected for the *qualityMosaic*. The result is an image that shows in each pixel the highest BPI value from the whole study period. Thus, it shows the desired BPC for the Swiss Plateau in a spatial resolution of 30 m.

3.2 Brownest Pixel Composite for Visualization

The next steps describe the calculation of the BPC for the entire Switzerland. As described in the beginning of the method section, this product is calculated for visualization purposes, to get an overview about the maximum bare soil area in Switzerland. These steps are also adaptive for every other part of Europe by just changing the region of interest. It is not used for the analysis of the soil properties as it uses TOA satellite data which are not corrected for the different atmospherical conditions.

3.2.1 Preprocessing

The Landsat-8 TOA product was chosen for the calculation. This product does not contain the SR CFMask band. Therefore, the cloud filtering was done using the GEE cloud score algorithm. The algorithm computes for every pixel a cloud likelihood score from 0 – 100 using the reflectance from the different bands. All pixels with a cloud score higher than ten were removed. The threshold of ten was also chosen by Huang et al. (2017) and showed feasible results.

To remove pixels with snow coverage, the Normalized Difference Snow Index (NDSI) was calculated (see Equation 3). In this study, the threshold of 0.7 was chosen to exclude pixels that contain snow. The threshold was chosen by visual analysis.

$$NDSI = \frac{R_{green} - R_{SWIR1}}{R_{green} + R_{SWIR1}}$$

Equation 3: Calculation of the NDSI where R is the reflectance of the green and SWIR1 spectral region.

3.2.2 Selecting Agricultural Areas

To be able to receive mainly the agricultural area in Switzerland, forest, water, and built-up areas were masked using different freely available products. These products have the advantage that they are available for whole Europe and could also be used to calculate the BPC in other regions.

Forest

To mask the forest area in Switzerland, the Hansen et al. (2013) *Hansen Global Forest Change* dataset was used. This dataset is already implemented in GEE and represents the global forest change since the year 2000 in a resolution of 30 m. The dataset was calculated from Landsat images and shows beside the forest change also the forest distribution in the band *treecover2000.* This band classifies every pixel in its percentage of tree cover inside the pixel, with values from 0 to 100. All pixels with a value greater than zero have been masked and removed.

Water bodies

For the water bodies, the *Permanent Water Bodies* dataset from the Copernicus Land Monitoring Service was used (Langanke et al., 2016). This dataset contains the European water bodies in a resolution or 20 m. A pixel is classified as water, when more than 90% of the pixel contains open water. All water pixels have been masked.

Built-Up

For the built-up area, the *Imperviousness* dataset from the Copernicus Land Monitoring Service was used (Langanke et al, 2016). The spatial resolution is 100 m. For every pixel the degree of imperviousness from 0 to 100 is calculated. The calculation is based on the NDVI. All pixels with a value higher than zero were removed.

3.2.3 Generation of the Continuous Composite

After the masking step, the BPC was calculated similar to section 3.1.4. The aim of this product is to create an image of entire Switzerland without any gaps of the masked pixels. Thus, the masked pixels from above need to get filled. For this, an additional cloud-free composite was created. This composite was calculated from all the remaining cloud-free images after the preprocessing steps from section 3.2.1. From the remaining images the mean reflectance per pixel was calculated. This was done using the *reduceToMean* function in GEE. Afterwards, all pixels that were masked in section 3.2.2 were then filled with the corresponding pixel from the cloud-free composite. The result is a gapless, complete and continuous image of Switzerland that shows the BPC over the agricultural area.

3.2.4 Enhancement of the Spatial Resolution

To increase the spatial resolution of the final product, the image was pansharpened. This was done using the Landsat-8 panchromatic band with a resolution of 15 m, together with the RGB bands 2,3, and 4. First, the blue, green, and red bands were converted from the RGB color space into the HSV color space. Afterwards, the ‘hue’ and ‘saturation’ from the HSV product was concatenated together with the panchromatic band. Finally, this product was converted back to the RGB color space. The

result is a natural color image with 15 m spatial resolution for the entire Switzerland that can get used for visualization.

3.3 Threshold

The BPC of the Swiss Plateau contains for each pixel the reflectance of the moment with the highest BPI value from the three years time period. In the agricultural area, the pixels will contain either the reflectance from the bare soil or from (permanent) grassland. The BPI value gives us an indication if a pixel contains rather bare soil or grassland. To be able to distinguish between those two classes, a threshold needs to be defined. This threshold should classify each pixel into either bare soil or any other land cover type. For defining this threshold, a combination of two approaches was chosen; analysis with APEX data as well as a field investigation. Figure 7 shows the workflow diagram for defining the threshold.

3.3.1 Airborne Imaging Spectroscopy

The seven binary bare soil images from APEX were loaded into the GEE. The bare soil fields were selected manually and everything else was excluded in the analysis as it contains misclassified pixels. To get the corresponding BPI value for each bare soil pixel, the Landsat-8 scene with the least time difference to the APEX scene was considered. Table 4 shows the corresponding Landsat-8 scene for each APEX scene and the time gap in days.

Table 4: APEX scenes and the corresponding Landsat-8 scene with the least time difference.

Location	APEX Date	Landsat Date	Landsat Scene	Time gap in days
Eschikon	18/07/2014	19/07/2014	LC81940272014200	+1
	10/04/2015	08/04/2015	LC81950272015098	-2
Oensingen	21/04/2015	24/04/2015	LC81950272015114	+3
	24/06/2015	27/06/2015	LC81950272015178	+3
Greifensee	18/07/2014	19/07/2014	LC81940272014200	+1
	10/04/2015	08/04/2015	LC81950272015098	-2
	10/07/2015	06/07/2015	LC81940272015187	-4

Every APEX scene was later analyzed separately. First, the 2 m resolution of the APEX scene was reduced to match the 30 m resolution of Landsat-8 images in GEE. The APEX pixels were reduced with a mean function to fit the Landsat pixel size. The new 30 m APEX pixel was classified as bare when more than 50% of the primary pixels have been classified as bare. Afterwards, the projection was changed to the one from Landsat-8 so that the pixels fit exactly. In addition, each Landsat-8 scene

was preprocessed as described in section 3.1 and for every pixel the BPI was calculated. For all the newly calculated 30 m APEX bare soil pixels, the corresponding Landsat-8 BPI value was extracted, stored in a list, and exported.

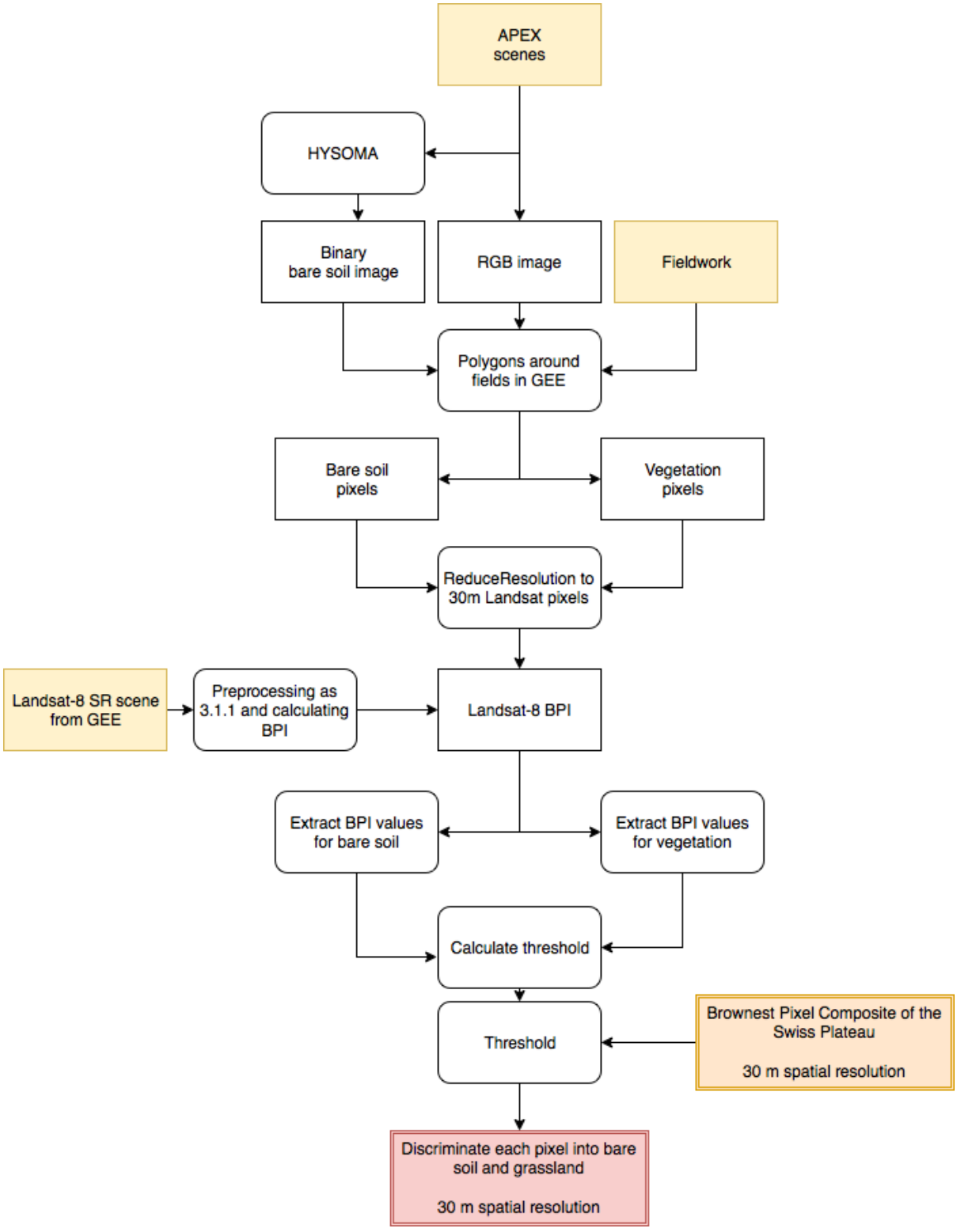

Figure 7: Working steps for defining the BPI threshold.

Additionally, to obtain also a reference of the BPI value from pixels containing vegetation, four APEX natural color (RGB) scenes have been analyzed. Fields that were not classified as bare soil, but visually had a mixture of bare soil and vegetation, or were fully vegetated, were selected with a polygon in GEE. Figure 8 shows an example of the selected bare soil and mixed vegetation fields. Afterwards, the same steps as above were conducted. For every vegetation pixel in the polygons, the corresponding BPI value was calculated, stored and exported.

Figure 8: Scene from Eschikon in April 2015. Left: Binary bare soil image classified by HYSOMA. White pixels are classified as bare and black pixels as any other land cover class. Right: RGB composite of the same scene. The yellow polygons show fields that were selected as bare soil and the red polygons show the fields that were selected as vegetation class (mixture between bare soil and vegetation).

3.3.2 Fieldwork

The APEX analyses above have been done without ground reference data. To obtain also BPI values from pixels that are assured as bare soil, a small fieldwork campaign was conducted. A total of 56 fields were analyzed on a field trip in the northern part of Switzerland on the 16th of March 2017. The fields were classified by eye into the classes bare soil and vegetation. The vegetation class contained fields with different amounts of vegetation cover; from sparse to fully vegetated. A total of 45 bare soil fields and 11 vegetated fields were mapped. An example of three analyzed fields can be seen in Figure 9. The images show the NDVI value of the fields. The first image shows a bare soil field, the second a sparse vegetated one, and the third a fully vegetated one. The brighter the pixel, the higher the NDVI value and thus, the higher the vegetation activity.

Afterwards, the corresponding Landsat-8 scene (Table 5) was analyzed as described above. Each field from the field campaign was drawn with a polygon in GEE. And for every pixel in each class, the corresponding BPI value was calculated, stored in a list, and exported.

Table 5: Date of the fieldwork and the corresponding Landsat-8 scene.

Location	Fieldwork Date	Landsat Date	Landsat Scene	Time gap in days
Fieldwork	16/03/2017	12/03/2017	LC81950272017071	-4

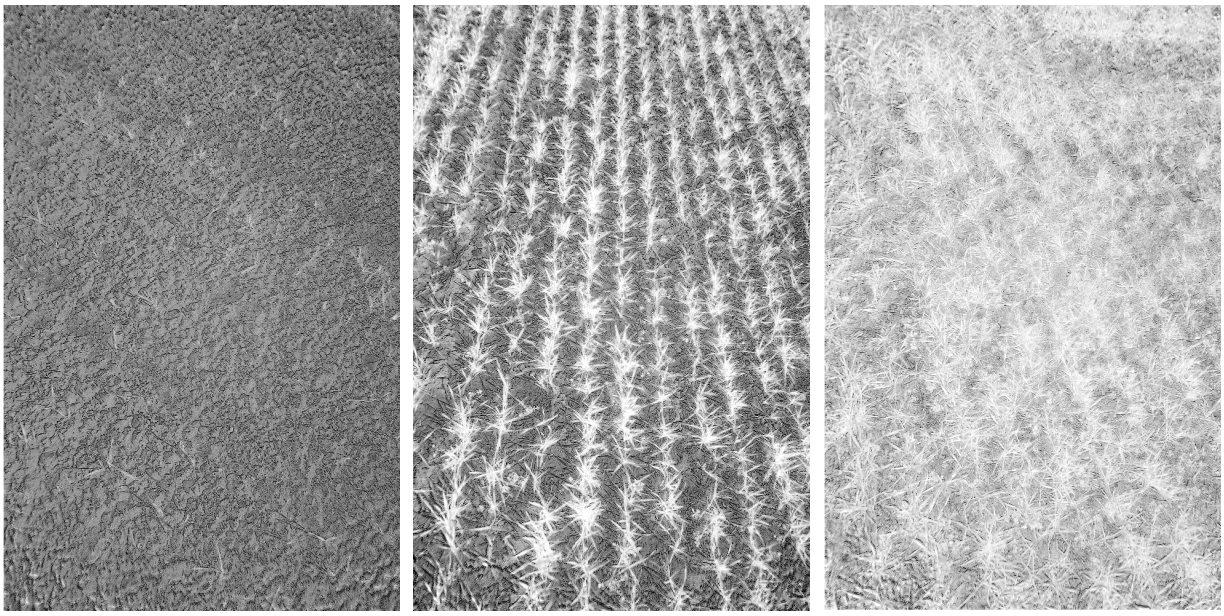

Figure 9: Example of the NDVI value of different fieldwork fields. Left: bare soil field; Middle: sparse vegetated field; Right: vegetated field. Bright color represents high NDVI values (vegetation) and dark color low NDVI values (bare soil).

3.3.3 Calculation of the Threshold

All BPI values from the selected APEX bare soil fields, the APEX vegetation fields, as well as from the fieldwork campaign were combined together with the corresponding classification label (bare soil or vegetation). To receive the optimal threshold for the distinction of those two classes, a sensitivity analysis of different thresholds was conducted. BPI thresholds between -0.2 and 0.2 were used to classify the pixels into either bare soil or vegetation. For every threshold, the producer and user accuracy, as well as the overall kappa was calculated from the classification. The final threshold was chosen with a tradeoff between a high producer and high user accuracy for the class bare soil, together with a high overall kappa value.

3.4 Bare Soil Analysis

The BPC of the Swiss Plateau together with the BPI threshold was used to classify each pixel into either bare soil or into any other land cover class. In the agricultural area of the Swiss Plateau, the other land cover class will be mainly permanent grassland, as that pixel was never bare soil in the three years of investigation as well as located in the agricultural area. This binary bare soil product can be used to get an overview about the geographic distribution of the arable land in the Swiss Plateau.

For the further calculation of the soil properties, we are just interested in pixels that are classified as bare soil. For this, a bare soil composite (BSC) was calculated. For the composite, the maximum bare soil moment from the BPC was chosen. If a pixel was more than one time bare (BPI value higher than the threshold), then the mean reflectance of all the bare soil moments was calculated. This was done using the *ee.Reducer.mean* function in GEE together with the masked bare soil moments. With the mean of several bare moments we can reduce the effect of outliers and should receive a more stable reflectance value of the bare soil pixel.

3.5 Soil Properties Calculation

3.5.1 Extraction of the Spectral Data

The calculation of the amount of sand, silt, clay, and SOM in the bare soil was done using the BSC of the Swiss Plateau together with the HSD (see Figure 10). The HSD sample locations were uploaded as a Google Fusion Table into GEE. For every sample location of the HSD, the corresponding bare soil reflectance was needed. The reflectance of the underlying Landsat-8 pixel of each sample location was extracted. When the field sample was over forest, water, imperviousness, or permanent grassland area, no reflectance data was extracted, as we are just interested in the reflectance of bare soil. Those soil samples could not be used in the further calculation steps. When the field sample lay on the border of two Landsat pixels, the mean reflectance value of both pixels was calculated and extracted. The final list of the soil samples together with the reflectance values was exported and analyzed using R (R Development Core Team, 2008).

3.5.2 Prediction of Soil Properties

The analysis was done for each of the four properties separately. First, all soil sample points without reflectance data were removed. For the calculation of the MLR the OLI bands 2 to 7 were used. This contains the reflectance of the blue, green, red, NIR, SWIR1, and SWIR2 wavelength region.

To get the best MLR model, a cross-validation approach was chosen. Cross-validation has the advantage compared to a simple randomly chosen test and validation data set that there is much less variation in the model performance (Gareth et al., 2017). For each combination of the 6 independent variables (OLI Band 2 to 7) a k-fold cross-validation with k=20 subsets was performed. Fewer subsets

would increase the model variation and more subsets would increase the computational costs, as well as the correlation bias (Gareth et al., 2017). All the different combinations of independent variables were then compared using the RMSE from the cross-validation. For each soil property, the combination with the lowest RMSE value was chosen individually for the calculation. The chosen combination of independent variables was then used to build the MLR model. Afterwards, the model was applied to every pixel of the BSC to predict the soil properties. The coefficient of determination (R^2) as well as the RMSE was calculated to analyze the performance of the MLR model.

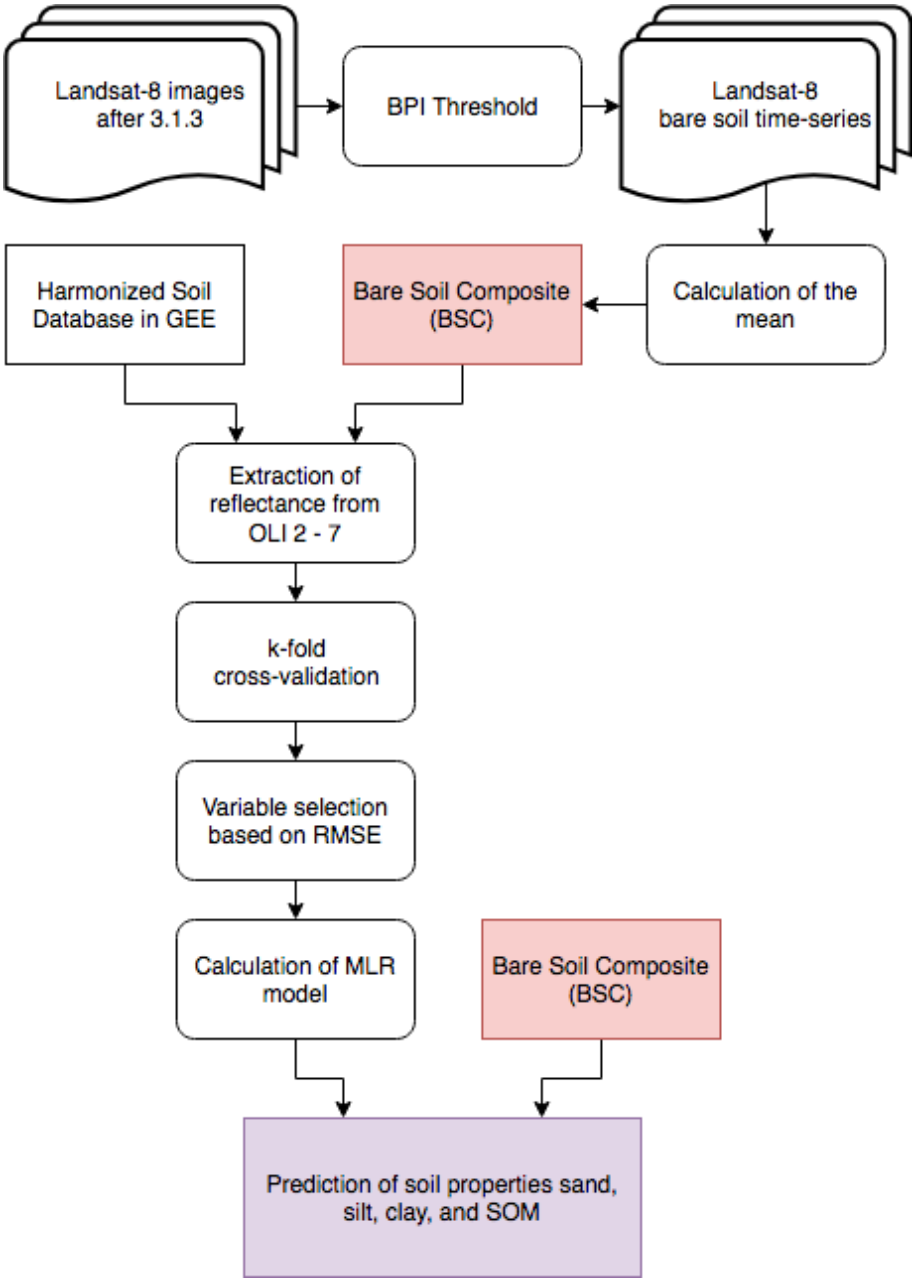

Figure 10: Working steps for the prediction of the different soil properties.

4. Results

4.1 Brownest Pixel Composite

4.1.1 Brownest Pixel Index

Two different indices were compared for the calculation of the BPC. Figure 11 shows an example close to lake Neuchâtel. The left image shows the minimum NDVI composite, the right image the maximum BI composite. White areas are pixels outside the agricultural area. The minimum NDVI composite showed a lot of artifacts all over the image, which are not present in the maximum BI composite. Visual analysis of the original Landsat-8 image showed that these artifacts were mainly not detected snow, fog, or cloud coverage. Therefore, the minimum NDVI is not useful for the selection of the barest soil pixels. The additional band information in the BI reduced this effect and is therefore better suited for calculating the BPC of the agricultural area. The BI was chosen as BPI.

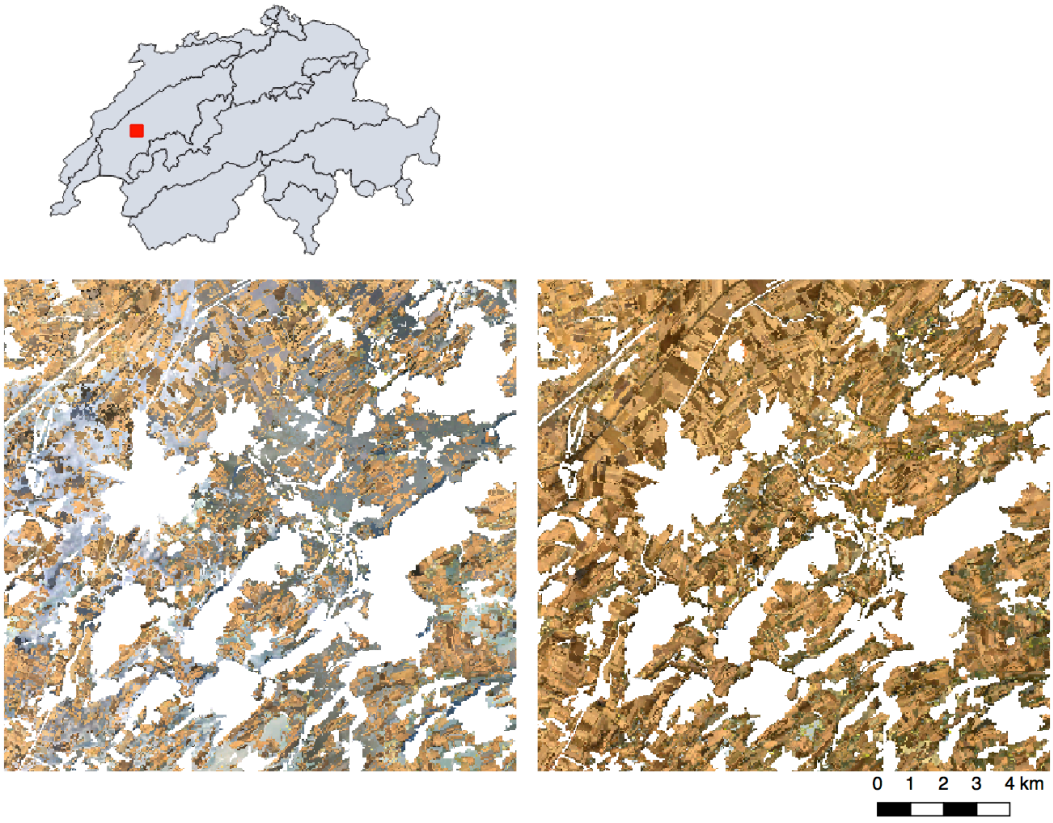

Figure 11: Comparison of the two BPI over the same location. Left: RGB image of the minimum NDVI composite. Right: RGB image of the maximum BI composite. White areas are masked because they do not contain agricultural area.

4.1.2 Brownest Pixel Composite for Analysis

The BPC was calculated for 49.5% of the area of the Swiss Plateau. The remaining 50.5% was masked out in the processing steps as it contained other land cover types than agricultural area. The values of the BI are mainly in a range from -0.4 to 0.4. The histogram in Figure 12 shows two peaks, one at around -0.15 and one peak at 0.15. The lowest part between the two peaks lies around the BI value of zero. The distribution of the values in the map (Figure 13) show that the highest BI values can be found in the western part of the Swiss Plateau, between Geneva and the lake of Biel. Another area with higher BI values can be found in the northeast between the region of Zurich and lake Constance. The middle part of the Swiss Plateau and the southern area shows generally lower BI values. Pixels with a higher BI value are more likely bare soil than pixels with lower BI values.

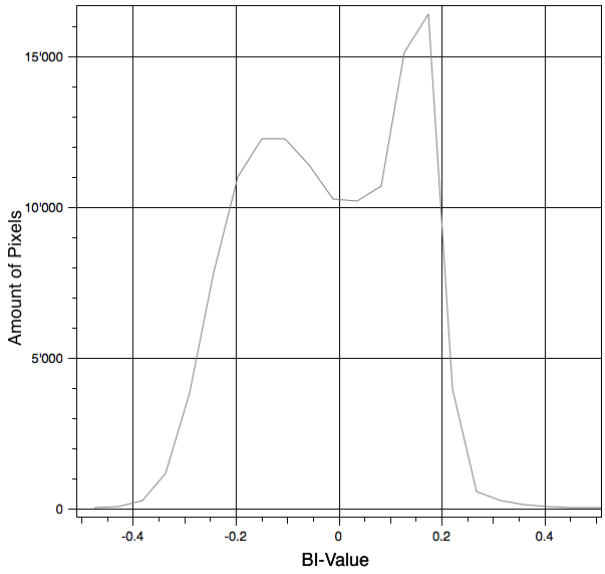

Figure 12: Amount of pixels with the corresponding BI value.

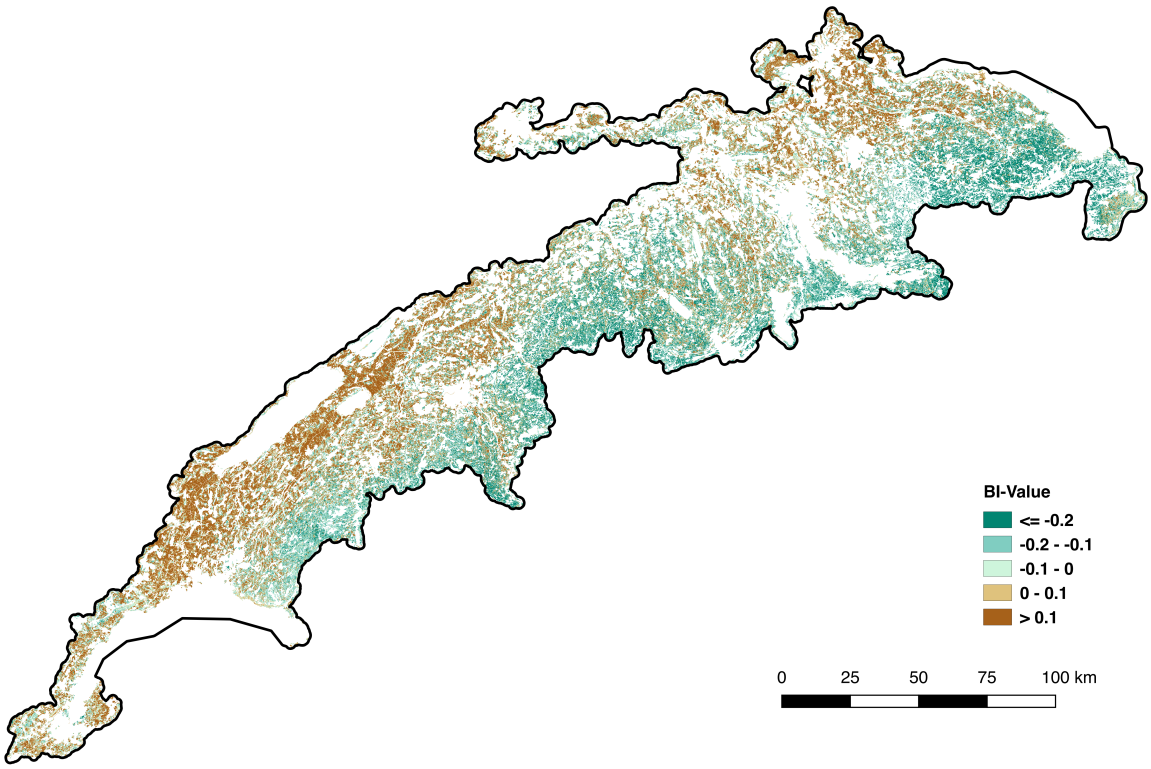

Figure 13: BI values of the BPC in the study area. White areas are masked pixels.

4.2 Brownest Pixel Composite for Visualization

Figure 14 shows the RGB composite of the BPC of Switzerland in a spatial resolution of 15 m. Additionally, two focus areas with different scales are shown to illustrate the high resolution of the product. In the red focus area we can see the lake of Neuchâtel, some forests, settlements, and agricultural area. The agricultural area looks brownish, which shows that many fields are bare and have been used as arable land during the study period. The blue focus area shows a small area in the canton of Berne. Also here, forest, settlements, and agricultural fields can be found.

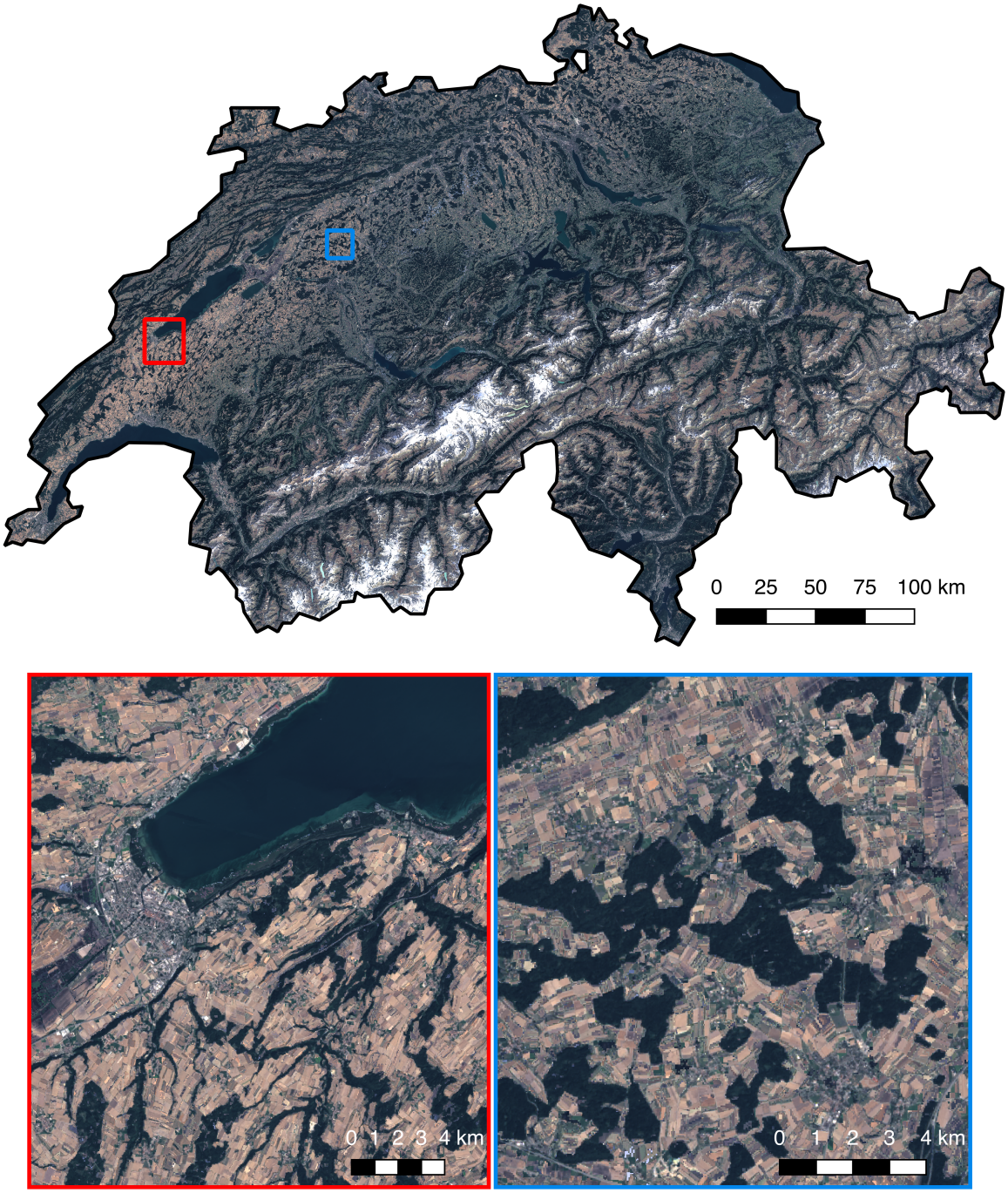

Figure 14: RGB composite of the BPC of Switzerland in 15 m spatial resolution together with two focus areas on bottom.

4.3 Threshold

4.3.1 Airborne Imaging Spectroscopy and Fieldwork

Table 6 shows the amount of pixels for each scene and class that have been selected and analyzed. The BI value for a total of 2531 bare soil pixels and a total of 5075 vegetation pixels were calculated. The amount of pixels in each scene differs a lot. For the scene of Eschikon in July 2014, no more pixels were left after the reduction from the APEX to the Landsat resolution. Therefore, it could not get further used. For other scenes, like the one from Oensingen in June 2015 the BI value of over 2500 pixels was calculated.

Table 6: Amount of bare soil and vegetation pixels analyzed for each APEX scene and the fieldwork campaign.

Location	Flight Date	Bare soil pixels	Vegetation pixels	Total
Eschikon	18/07/14 10/04/15	0 193	0 156	0 349
Oensingen	21/04/15 24/06/15	640 189	723 2356	1363 2545
Greifensee	18/07/14 10/04/15 10/07/15	41 538 62	0 1765 0	41 2303 62
Fieldwork	16/03/17	868	75	943
Total		2531	5075	7606

When comparing the BI value of all the pixels of the two classes (Figure 15), we can clearly see some differences. The bare soil class shows an average of around 0.1 and the vegetation class of around -0.25. Additionally, the variance of the vegetation class is much higher than the one from the bare soil class. The bare soil class shows some outliers to lower BI values.

The boxplots of each of the locations (Figure 16) show again that the bare soil pixels have higher BI values than the vegetation pixels. The bare soil class of the scene of Greifensee in July 2015 shows the lowest average BI values with values below zero. However, in this scene just 62 pixels were analyzed. All other scenes show average BI values higher than zero with the highest values in Eschikon in April 2015. All scenes, except Greifensee July 2014 and Greifensee July 2015, show some outliers with values lower than -0.1. Another interesting fact is that the outliers are nearly always to the lower BI values and none, except Greifensee April 2015, show outliers to larger BI values.

When comparing the boxplots of the two classes in each scene, we can see distinct differences between the classes. Only the values from the fieldwork campaign show less difference in mean BI values. The vegetation pixels show a mean that is higher than zero.

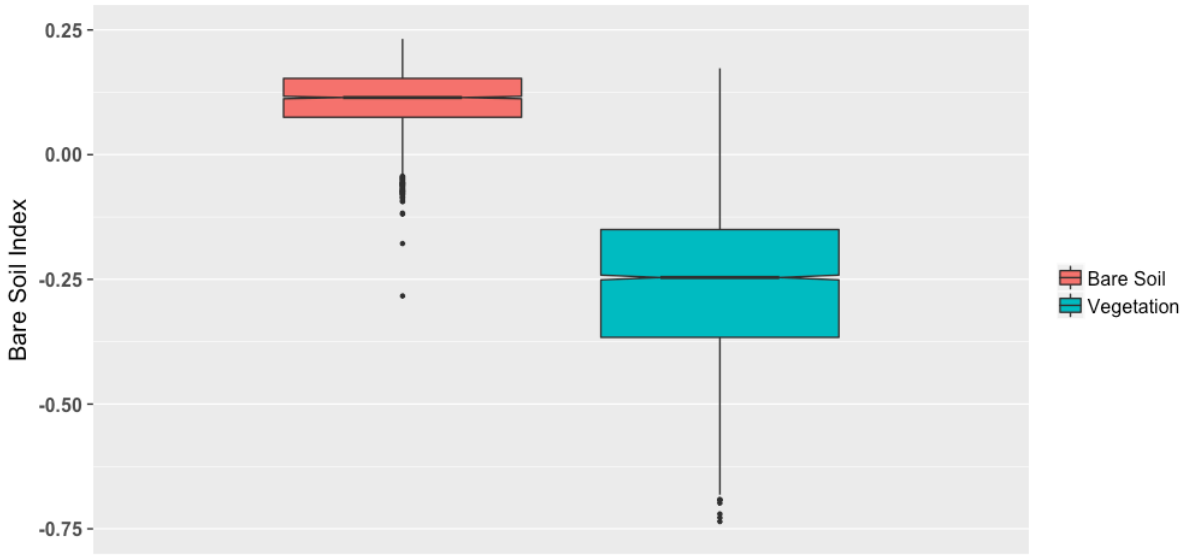

Figure 15: BI value boxplots of all the pixels in the two classes bare soil and vegetation.

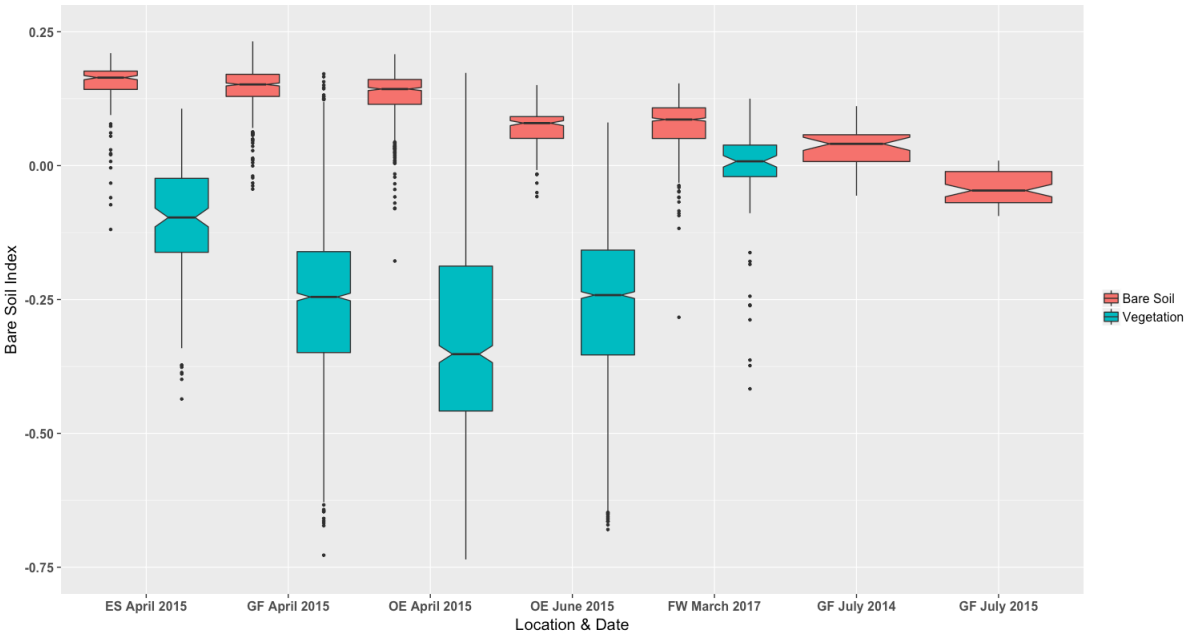

Figure 16: BI value boxplots for the classes bare soil and vegetation for each APEX scene and the fieldwork campaign. ES = Eschikon; GF = Greifensee; OE = Oensingen; FW = Fieldwork. For the last two locations no RGB composite and therefore no vegetation class was calculated.

4.3.2 Calculation of the Threshold

Figure 17 shows the result of the threshold calculation. The overall kappa value is increasing with increasing BI threshold to a maximum of 0.88 and decreases afterwards. It stays relatively stable on a high value over 0.85 between the thresholds of -0.0275 and 0.0425. The user accuracy of the class bare soil is increasing continuously and in the period of high kappa the value goes from 0.87 to 0.94. In the same range the producer accuracy of the class bare soil is decreasing from 0.97 to 0.86. For the optimal threshold, the interface between user and producer accuracy was used. The interface lies at a BI value of 0.0175. At this threshold the kappa value is 0.88 and the user and producer accuracy are both at 0.92. The BI value of 0.0175 is chosen as the threshold for dividing each pixel into either bare soil, or vegetation.

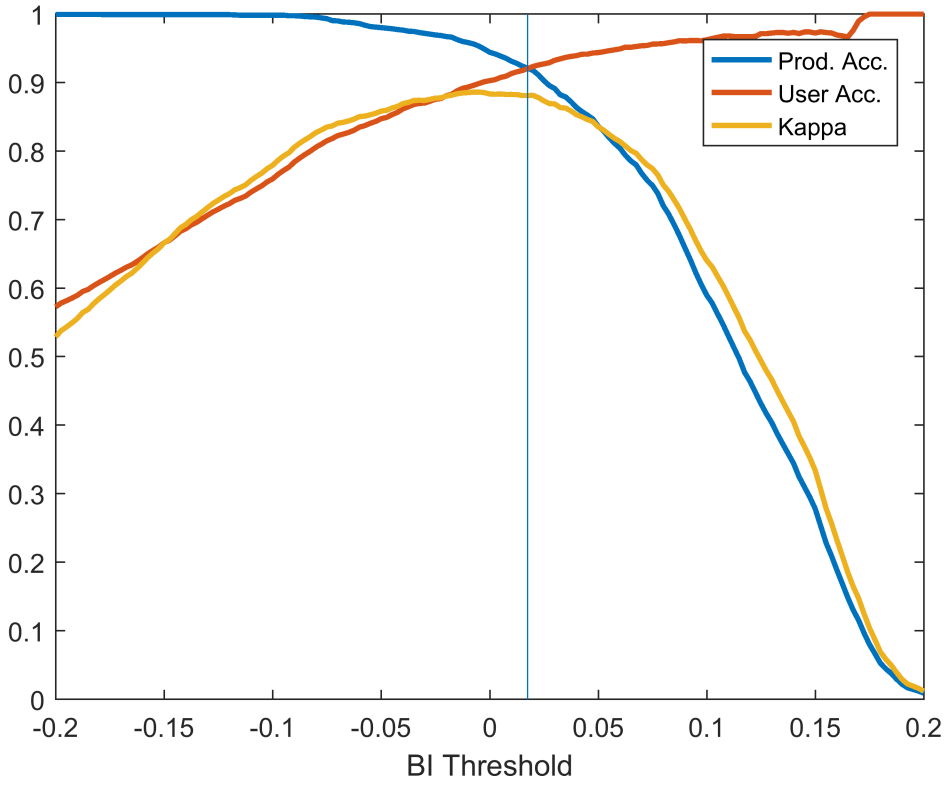

Figure 17: Results of the binary classification with different BI thresholds. The x-axis shows the threshold value from -0.2 to 0.2. The y-axis shows the accuracy of the classification from 0 to 1. The blue line represents the producer accuracy of the class bare soil; the red line the user accuracy of the class bare soil. The yellow line shows the overall kappa value. The chosen threshold is visualized by the vertical blue line and lies at the BI threshold value of 0.0175.

4.4 Bare soil analysis

Figure 18 shows the binary classification of bare soil and grassland. 40.8% of the agricultural area is classified as bare soil and 59.2% of the area as grassland. Different parts of the Swiss Plateau show different distribution of the two classes. The distribution is similar to the results of the BPC. The largest bare soil area occurs in the western part of the Swiss Plateau between the Lake of Geneva and Lake Biel. Another bare soil area is northeast, between Zurich and Lake of Constance. The remaining area of the Swiss Plateau shows mainly grassland with some bare soil fields in between.

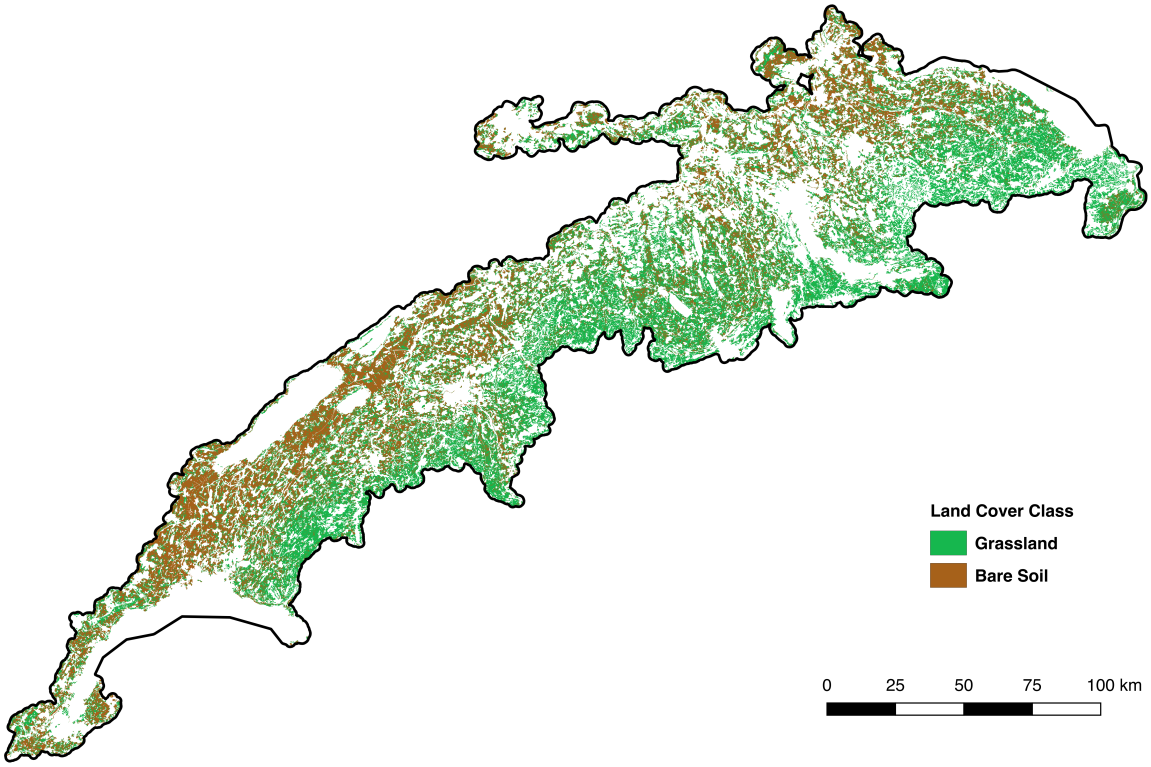

Figure 18: Geographic distribution of the binary classification of the BPC by the BI threshold of 0.0175. All BI values higher than the threshold were classified as bare soil; the lower values as grassland. White areas are masked pixels.

For the further calculation of the soil properties only bare soil pixels are of interest. Figure 19 shows the natural color RGB composite of the BSC. These pixels are used for the further analysis of the soil properties. We can clearly see differences in the RGB soil color, which is an indicator of differences in soil characteristics (Shukla, 2014). In the red focus area we can see in the north, between the lakes, an area with very dark soils, whereas the most other pixels show similar, lighter soil colors. This is an indicator that the soil properties differ in these two areas.

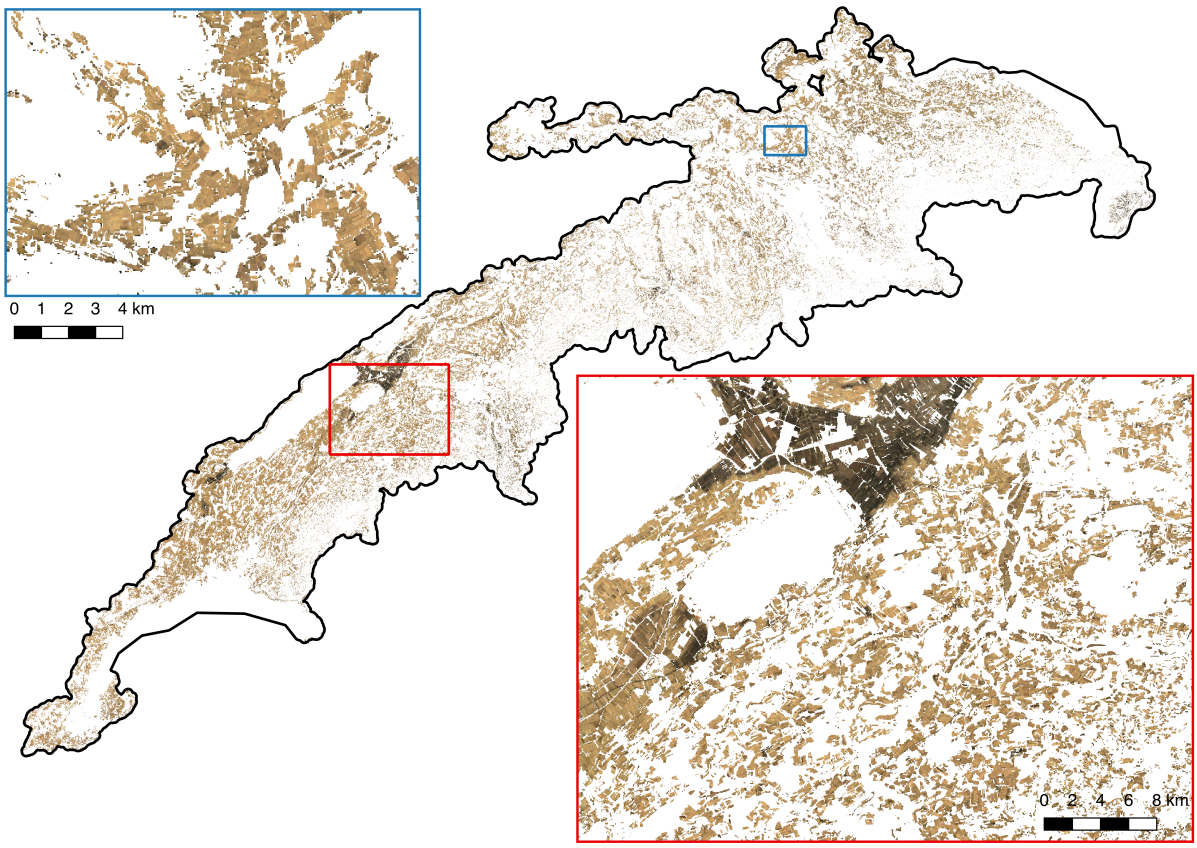

Figure 19: RGB (OLI 4/3/2) composite of the BSC in the study area together with two focus areas.

Figure 20 gives an overview from which months, in the three-year study period, the bare soil pixels were selected for the BPC. A lot of pixels were selected from the months March and April. Afterwards, the amount of chosen pixels decreases strongly. A second peak can be seen in the month of August, from where every fourth pixel in the BPC showed the maximum BI value. Afterwards, the following three months show again a smaller amount of maximum BI value and in the three winter months, December, January, February just a few pixels have been selected for the composite.

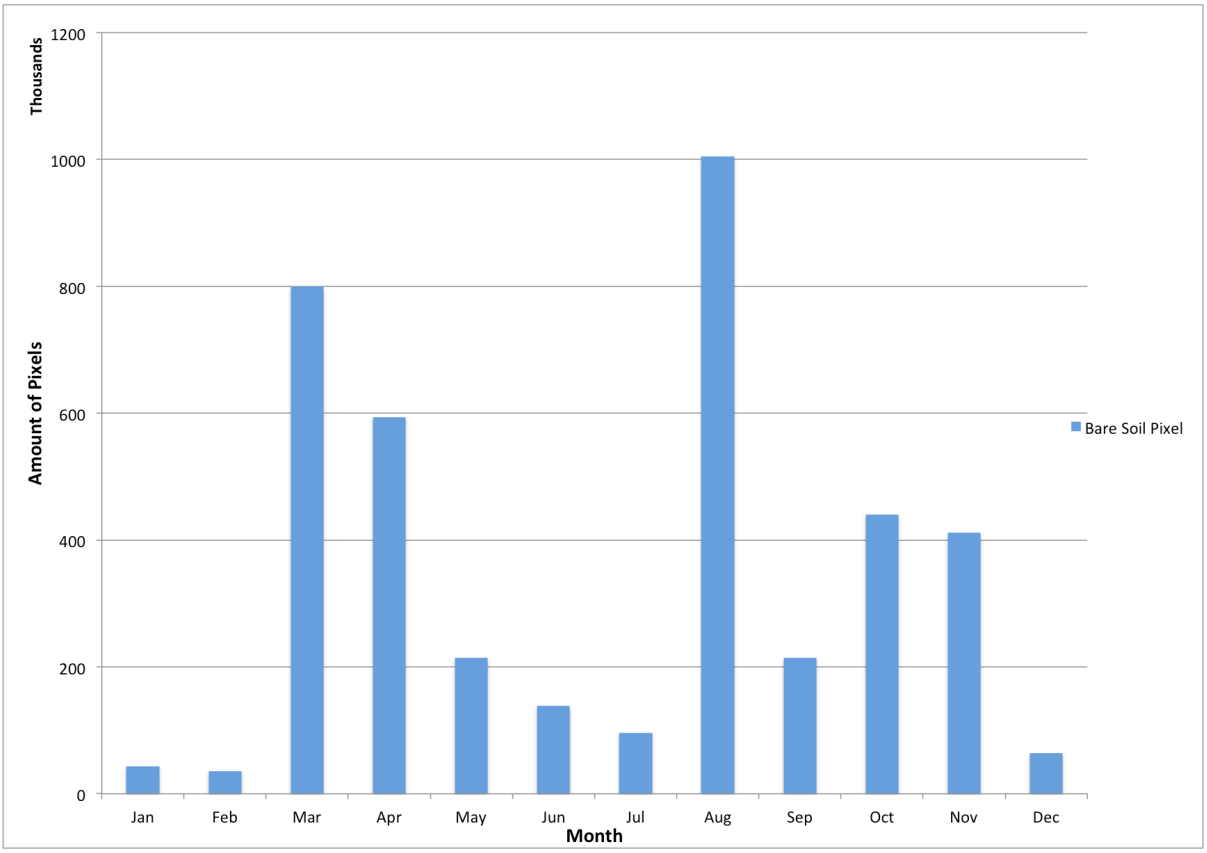

Figure 20: The bar plot shows the amount of pixels chosen from each month for the final BSC. It is an indicator on which month of the year the most fields show bare soil.

4.5 Soil properties calculation

4.5.1 Multiple linear regression with spectral data

From the 4783 different soil samples (see Table 7), around 67% were located in non-bare soil area and are not useable for the soil properties calculation. From the remaining 1597 soil samples, 1580 contain information about the amount of SOM, and 735 contain the percentages of sand, silt, and clay.

Table 7: Statistic about the HSD soil samples used for prediction.

Nr. of samples	Masked	On the BSC	Soil texture	SOM
4783	3186	1597	735	1580

For each of these soil samples, the corresponding reflectance was calculated. The following figures show the mean reflectance of the samples. To analyze the impact of the soil property on the reflectance, the samples were divided into four categories with different amount of the soil property in the soil.

Figure 21 shows clearly that the reflectance is decreasing with an increase in amount of clay in the sample. This trend occurs in all the six OLI bands but seems to be stronger with increasing

wavelength. Samples with a high amount of clay (>45%) show up to 30% decreased reflectance than samples with low amount of clay (<15%).

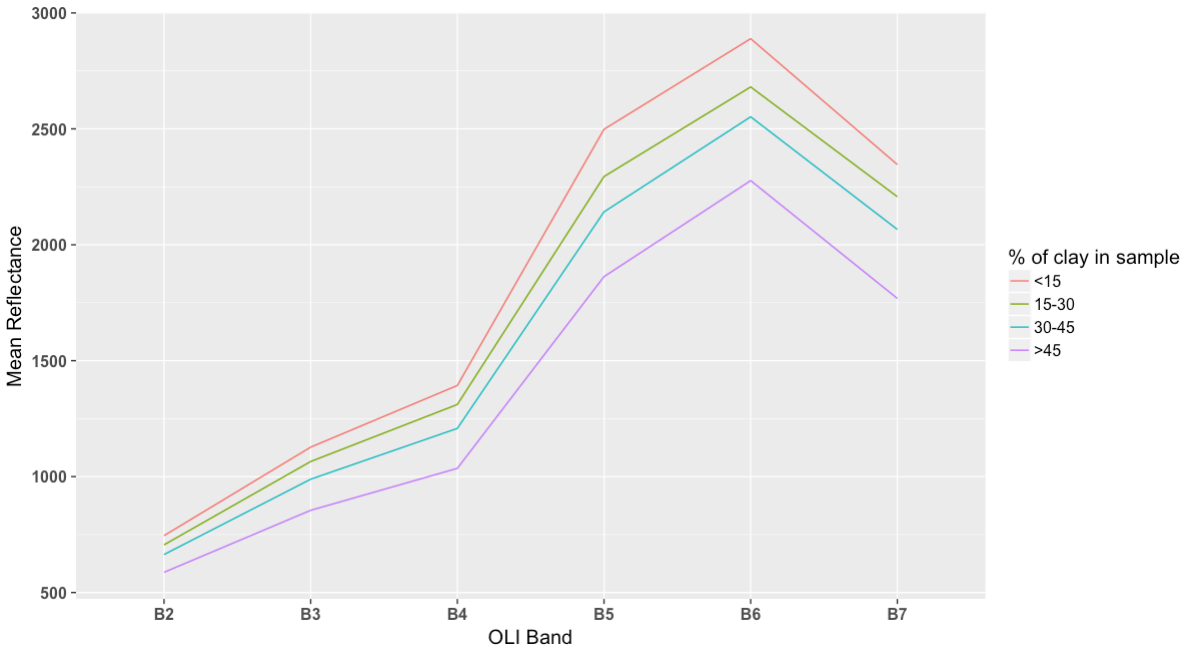

Figure 21: Mean Reflectance of samples with different amount of clay.

Sand shows the opposite trend than clay (see Figure 22). Samples with a high amount of sand show higher reflectance values than samples with lower amount of sand. This trend can again be seen through all six OLI bands but seems to get stronger with increasing wavelength. However, the differences between the classes seem to be smaller than with clay.

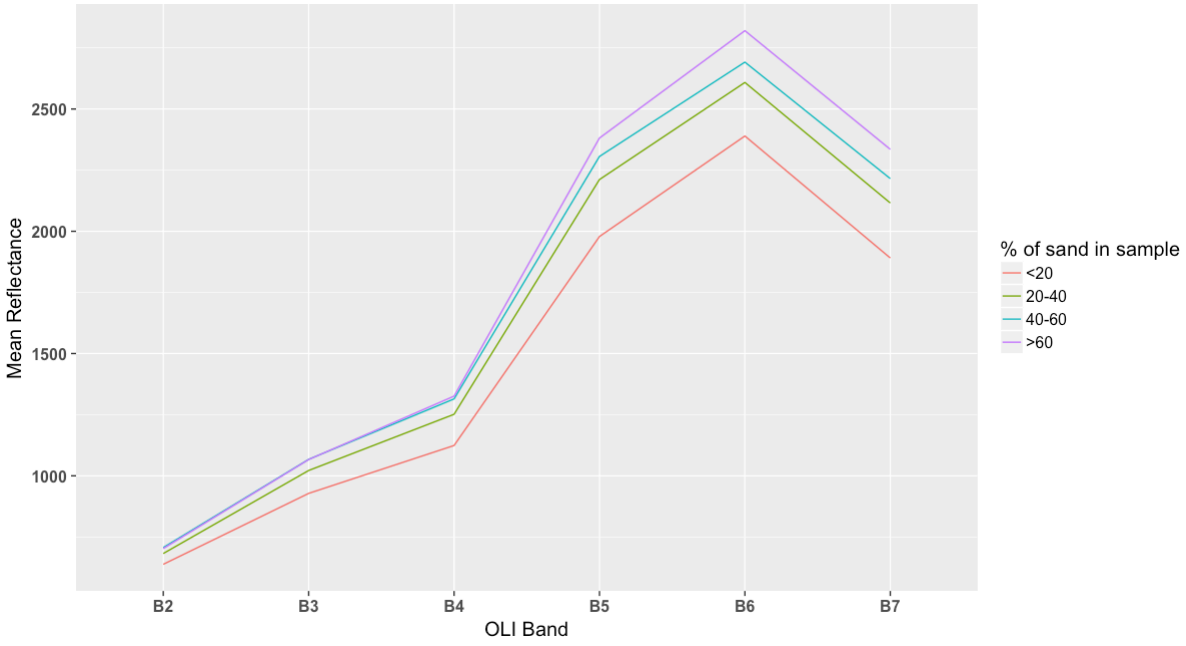

Figure 22: Mean Reflectance of samples with different amount of sand.

For silt, there is no clear trend visible (Figure 23). In the visible wavelength region (OLI bands 2/3/4) samples with a high amount of silt reflect more than sample with a low amount of silt. In the NIR and SWIR region of the wavelength this trend decreases and the samples with the lowest amount of silt (<20%) show more reflectance than the samples with a silt percentage of between 20 and 60%. Generally, the differences are much smaller than for sand and clay.

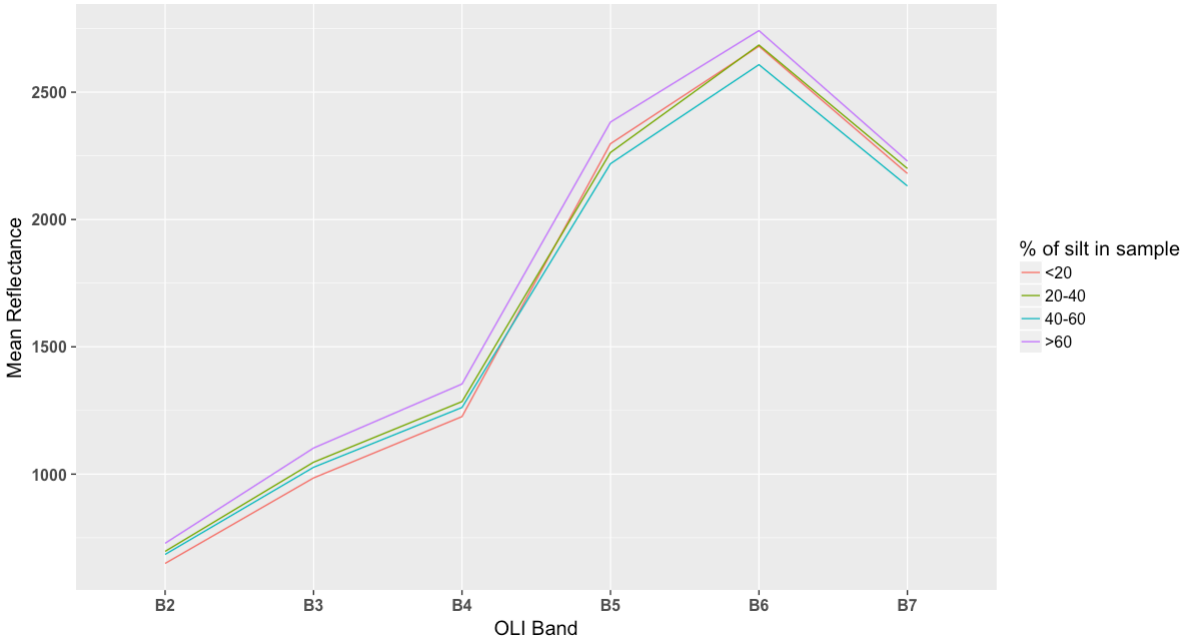

Figure 23: Mean Reflectance of samples with different amount of silt.

SOM (Figure 24) shows again a clear decrease of reflectance with a higher percentage in the sample. This decrease can be seen in all six OLI bands. However, the differences between the visible and the NIR/SWIR region are not that clear than for sand and silt content.

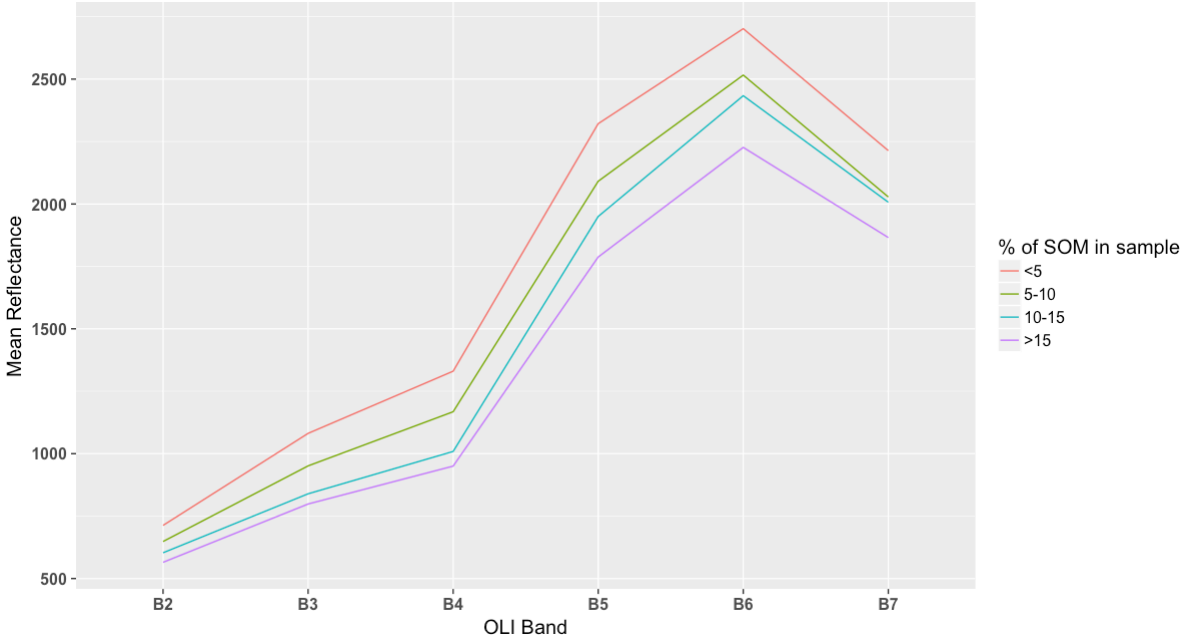

Figure 24: Mean Reflectance of samples with different amount of SOM.

Figure 25 shows the RMSE values of the cross-validation with different amount of independent variables. The shown RMSE value is the one from the best combination of the amount of independent variables. Sand, clay, and SOM show all similar results. The RMSE decreases from the model with just one independent variable to the model with five independent variables. The model containing the information from all six Landsat-8 bands showed thereafter a little increase in RMSE. Therefore, for the calculation of the sand, clay, and SOM content, the best model with five independent variables was selected. In case of silt, the calculation showed a strong decrease to the model with two variables followed by increasing RMSE values. The model containing two independent variables was chosen. However, the differences are just in the second number after the comma. The selected bands for each MLR model together with its equation is shown in Table 8.

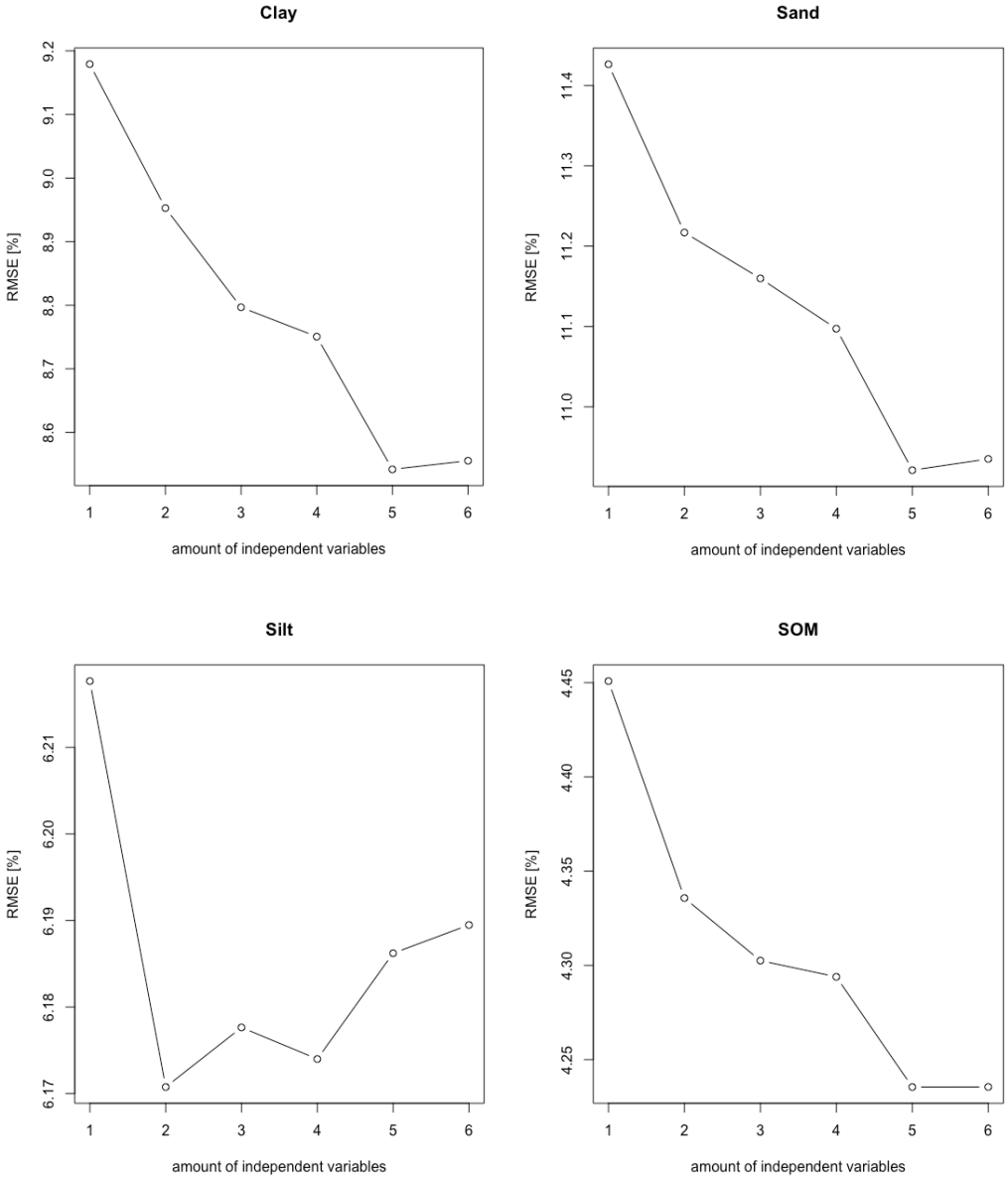

Figure 25: Model selection based on the RMSE values of different models with different amount of independent variables. The RMSE value is always the one from the best model of the amount of independent variables.

For the calculation of the clay and sand content the reflectance from the Landsat-8 green, red, NIR, SWIR1, and SWIR2 wavelength region were used. For the calculation of the silt content the best model contained the information of the green and SWIR2 band. For the SOM model all bands except SWIR1 were used.

Table 8: Multiple linear regression equations for the prediction of clay, sand, silt, and SOM.

Soil property	Intercept	Blue	Green	Red	NIR	SWIR1	SWIR2
Clay	45.7202		+0.0704	-0.0517	-0.0178	+0.0305	-0.0312
Sand	20.9845		-0.0780	+0.0515	+0.0169	-0.0305	+0.0373
Silt	33.3967		+0.0096				-0.0045
SOM	11.2234	+0.0200	-0.0143	-0.0120	-0.0026		+0.0072

The predicted values for each soil property can be seen in Figure 26. Each sample is represented by a black dot and the dashed black line shows the 1:1 line, indicating were each samples should be located if the correlation model was perfect. The analysis of the predicted values compared to the sample values can be seen in Table 9. The model predicted with an R^2 of 0.26 for clay, 0.17 for sand, 0.03 for silt, and 0.28 for SOM. Therefore, the MLR model performs best for clay and SOM, followed by sand. The R^2 for silt is nearly zero and therefore, the correlation model has very limited prediction power. Thus, for silt no prediction map will be calculated. In general, all four plots, as well as Table 9 shows that the models are overestimating the lower sample values and underestimating the higher sample values. For SOM also negative percentages were predicted. However, this is not possible and therefore, the negative values were set to zero in generation of the soil property maps.

Table 9: Basic statistics about the soil samples and the prediction of the soil samples.

Soil Property	Min [%]	Mean [%]	Max [%]	SD [%]	RMSE [%]	R^2
Clay	7.1	25.7	65.7	9.8		
Clay Predicted	11.8	25.7	46.3	5.0	8.4	0.26
Sand	0.0	40.6	75.9	11.9		
Sand Predicted	21.7	40.6	43.8	5.0	10.8	0.17
Silt	12.8	33.4	56.6	6.2		
Silt Predicted	27.0	33.7	38.1	1.0	6.2	0.03
SOM	0.3	4.9	57.4	4.9		
SOM Predicted	-4.1	4.9	14.0	2.6	4.2	0.28

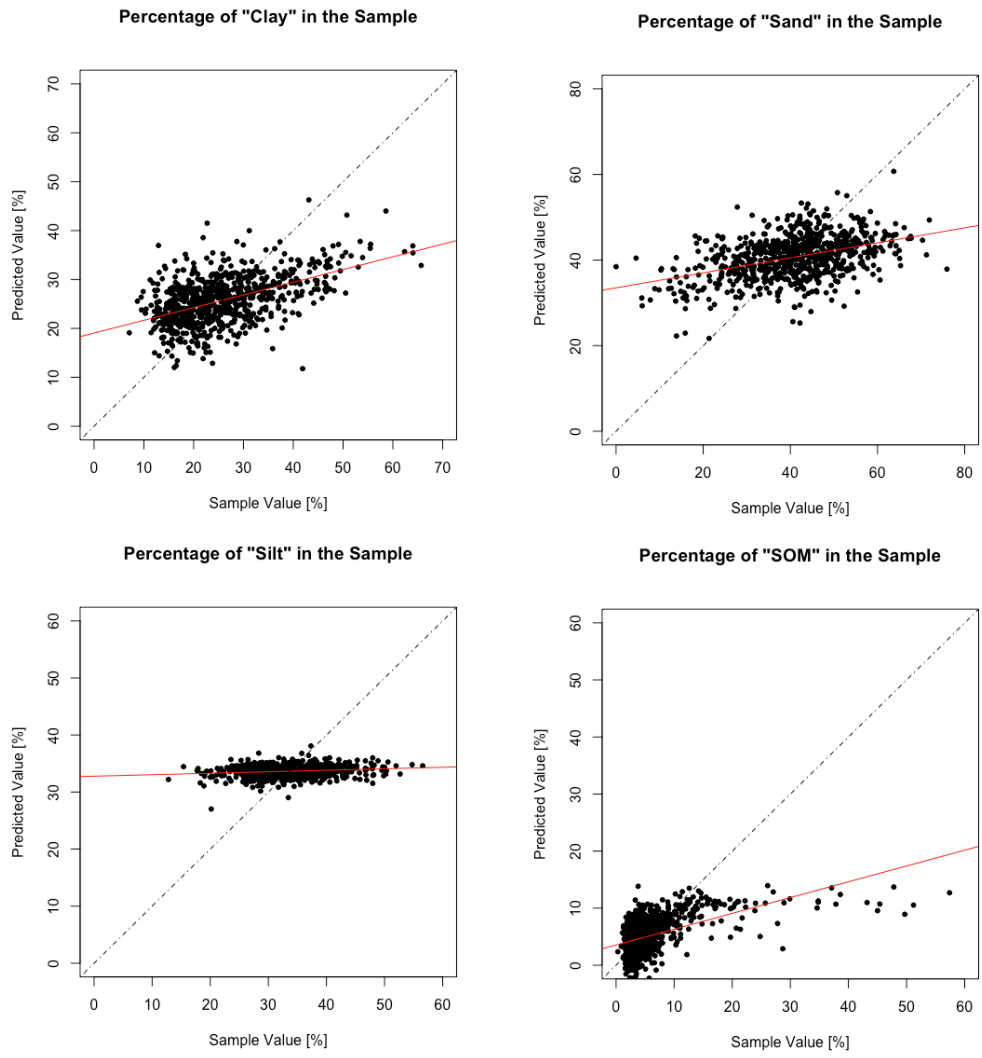

Figure 26: Scatterplots of the MLR models. Black dots represents the soil samples. The x-axes shows the sample value and the y-axis the predicted value. The black dotted line is the 1:1 line and the red line shows the regression line.

4.5.2 Prediction of Soil Properties

In this section the results of the MLR equations for clay, sand, and SOM are shown. The results are plotted for the whole Swiss Plateau as well as for two focus areas. The upper left, blue focus area lies in the canton of Zurich and contains a larger connected bare soil area. The lower right, red focus area is located in the area of high bare soil occurrence, next to the Lake Neuchâtel. This area contains noticeable differences in the soil property prediction.

Clay

The amount of clay in the topsoil of the bare soil fields range from 10% to 40%. The mean of the clay prediction is 25.7%, and the standard deviation is 5.7%.

The overview map of the Swiss Plateau (Figure 27) shows the highest amount of clay in the topsoil in the western part of the Swiss Plateau, next to the lakes Neuchâtel, Biel, Geneva. Also the top eastern part shows larger purple areas with a high amount of clay.

The western focus area shows the big continuous area of high clay percentage of more than 32% next to the lakes. Around this area the clay percentage is generally lower than 27%. Generally it looks like the percentage of clay is higher near the lakes than further away.

The second focus area in the region of Zurich shows generally lower clay values between 15% and 25%. There is a lot of variation between different fields visible. Some show very low values (dark orange) and some fields show higher, purple values.

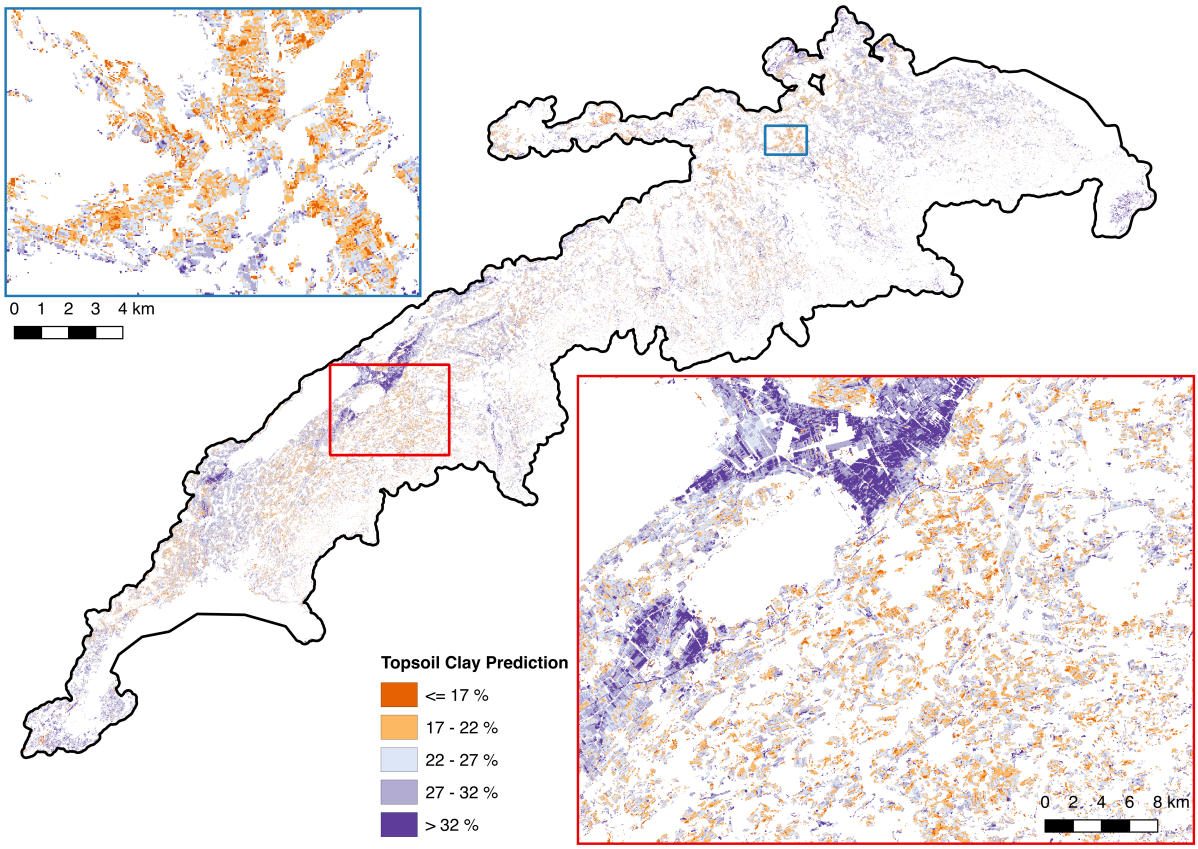

Figure 27: Prediction of topsoil clay for the whole study area and for two focus areas.

Sand

The sand prediction map (Figure 28) shows in general much higher percentage than the clay prediction map with values between 26% and 57%. The mean lies at 40% with a standard deviation of 5.9% Thus, the bare soil fields in the Swiss Plateau contain a higher amount of sand than clay. The distribution seems to be inversely to the clay prediction map. In regions with higher clay values, the

sand percentage is lower and the other way round. The western focus area shows us the same continuous region with about 35% of sand in the topsoil next to the lakes. Around this area the sand percentage increases to values between 40% and 50%.

The second focus area, in the canton of Zurich, shows mainly high percentages of sand with values between 40% and 50%. Again, differences on a field scale are visible.

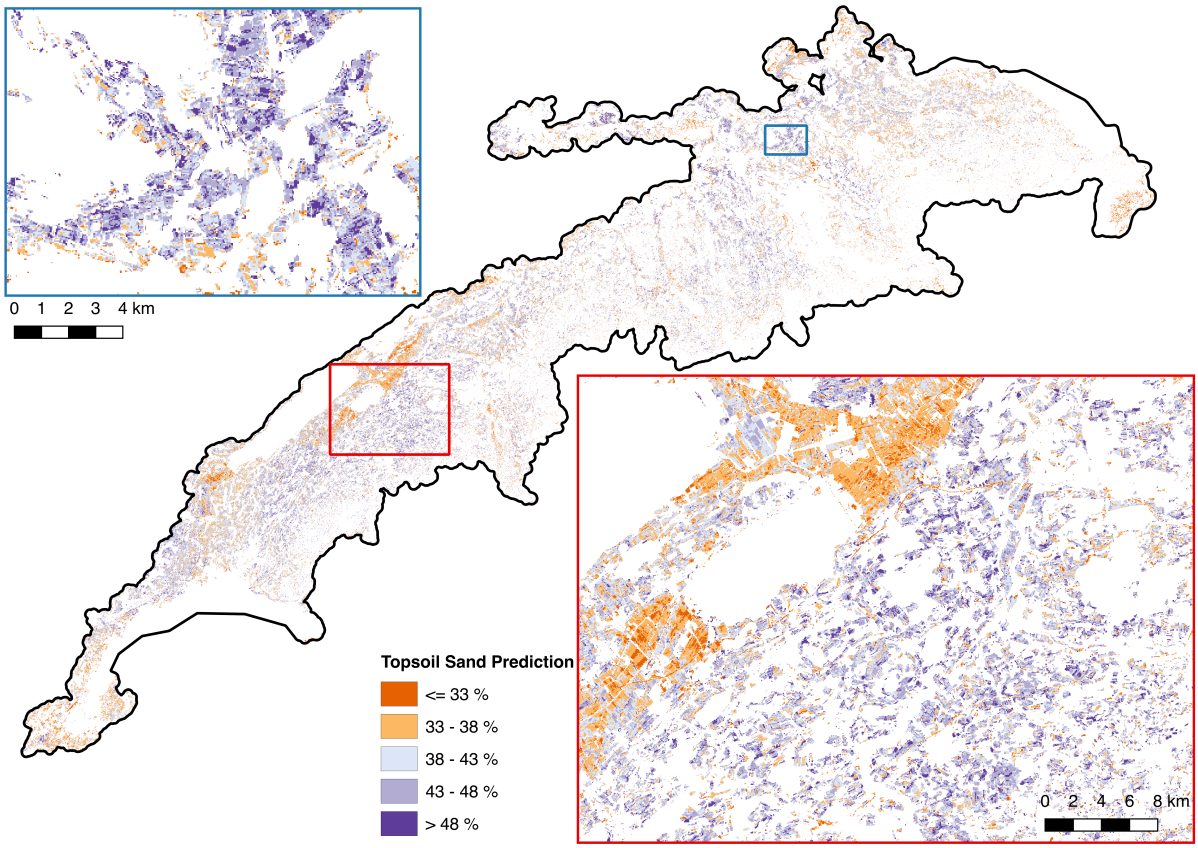

Figure 28: Prediction of topsoil sand for the whole study area and for two focus areas.

SOM

The SOM map (Figure 29) shows values between 0% and 12% with a mean of 4% and a standard deviation of 2.7%. The map of the Swiss Plateau shows generally low values, mainly below 4% SOM with some areas up to 7%. Fields with higher predicted SOM values are just visible next to the Lake of Neuchâtel, which can be seen in the western focus area. The area next to the lake shows the highest SOM percentage with values around 10%. Right next to this area, the values drop to around 5%.

In the second focus area, the values are generally below 5%, and some variations are visible.

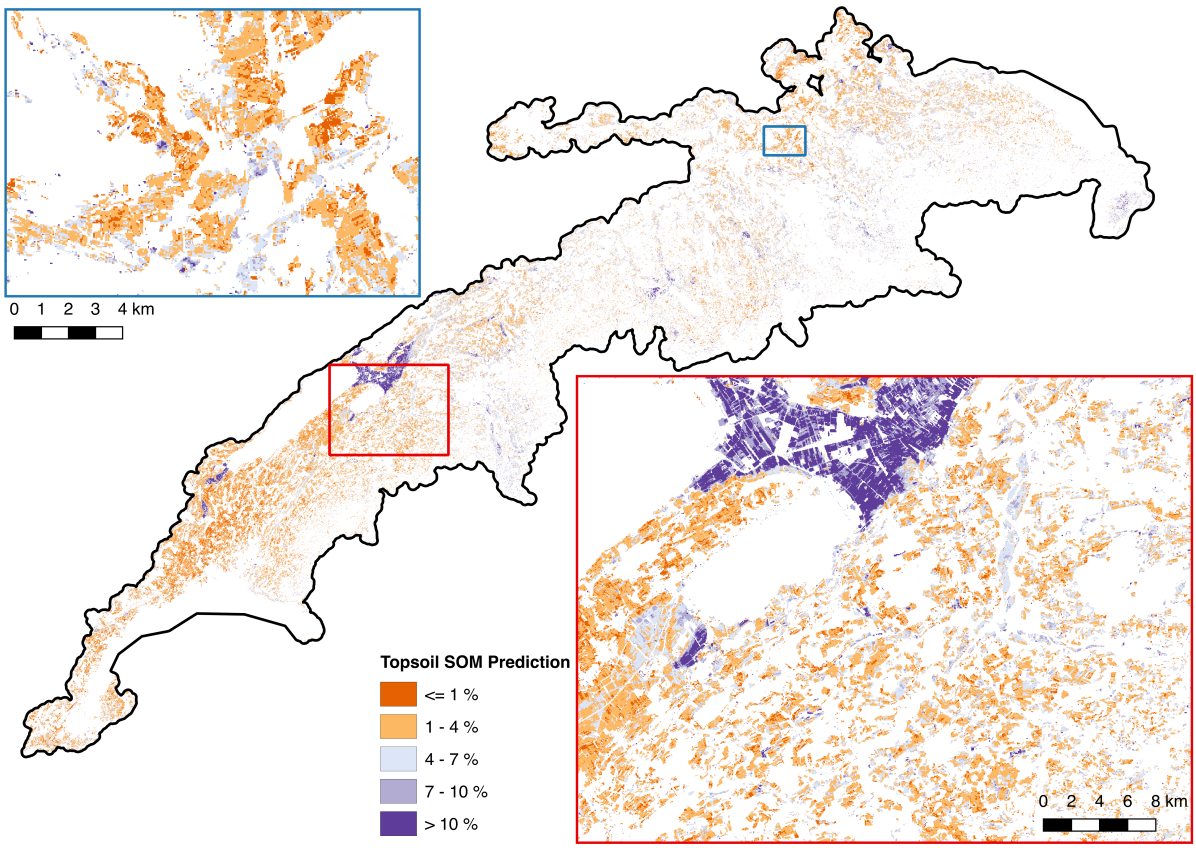

Figure 29: Prediction of topsoil SOM for the whole study area and for two focus areas.

5. Discussion

5.1 Brownest Pixel Composite

5.1.1 Brownest Pixel Index

The composite with the minimum NDVI value did not work as expected. Low NDVI values should correspond to low vegetation coverage and therefore, over agricultural fields, should detect bare soil. This was not always the case, as the result showed. Many artifacts were visible in the result, which seemed to be not detected thin clouds. Figure 30 shows the reflectance of some artifact pixel from the minimum NDVI composite, together with the reflectance from the same pixels in the maximum BI composite, showing a bare soil field. The low spectral difference between the OLI red and NIR in the artifact pixels (blue line) lead to an NDVI of 0.12. The bare soil pixels show an NDVI of 0.21. Hence, the minimum NDVI composite prefers the artifact to the bare soil field. The BI is calculated by the OLI bands blue, red, NIR, and SWIR1. Therefore, it adds the information of two more bands to the index compared to the NDVI. This additional band information changes the decision in this case. The BI value for the artifacts is -0.16, and for the bare soil field 0.13. The maximum BI composite is choosing the bare soil field to the artifact. This shows that the BI is more suitable to detect bare soil than the NDVI. From a visual analysis of the results, the BI seems to be appropriate to detect the "brownest" moment of each pixel in a larger area.

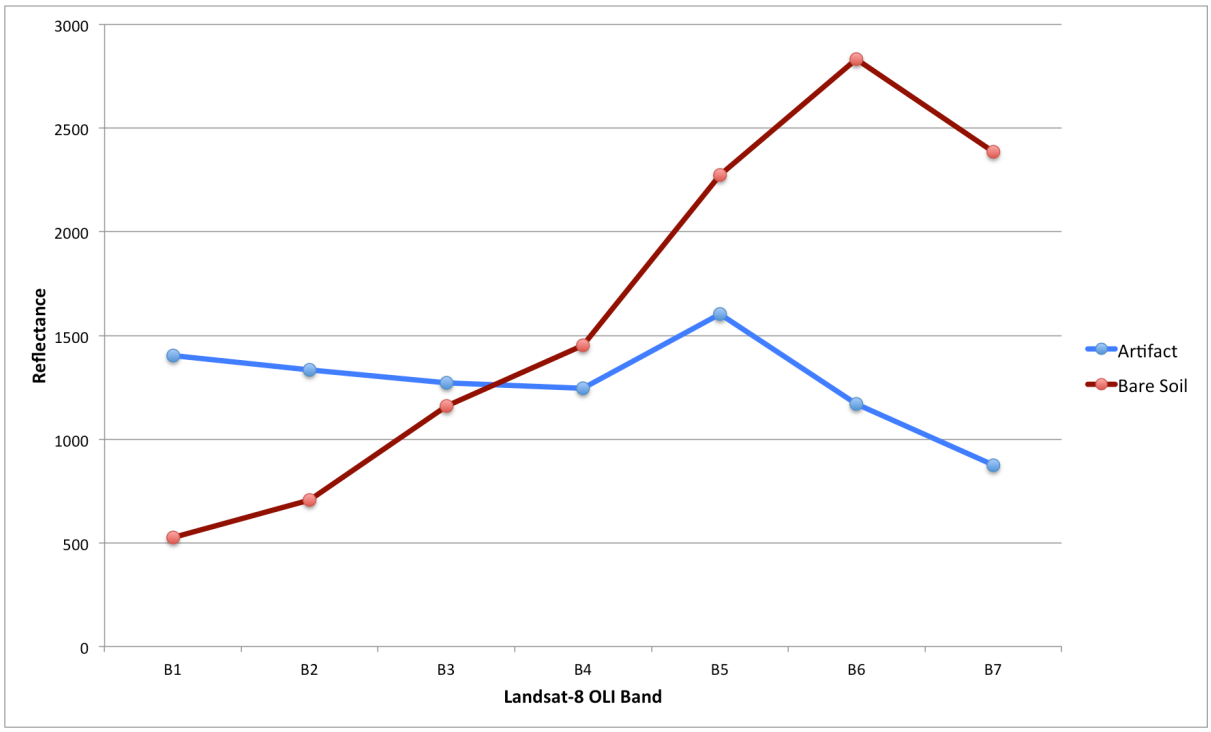

Figure 30: Mean reflectance of some artifact and bare soil pixels.

5.1.2 Brownest Pixel Composite for Analysis

The BPC was calculated using the maximum BI value. The composite was just calculated over agricultural area in the Swiss Plateau. Thus, it should show bare soil over arable land and grassland over pasture. Theoretically, permanent grassland should always show some reflectance of vegetation in the pixel, where arable land should show moments without any vegetation on it. The resulted BI histogram (Figure 12) of all the pixels showed clearly two peaks. The first peak is relatively flat and the pixel values are more scattered, where the second peak is sharper and contains a large amount of pixels. BI values below zero can occur when the added reflectance of SWIR2 and red is lower than the added reflectance of NIR and blue. Vegetation shows normally high reflectance in NIR and lower reflectance in blue, red, and SWIR2. If a pixel contains vegetation, the second part of the BI equation is higher than the first, and the BI value will be below zero. Bare soil shows nearly continuously rising reflectance values with increasing wavelength. SWIR2 together with red will therefore show higher amount of reflectance than NIR together with blue and the BI values will be positive. From this calculation we can already suppose that the threshold should lie somewhere around a BI value of zero. Additionally, we can expect that the two peaks in the histogram are most likely related to the two classes bare soil and grassland.

The distribution of the BI values over the Swiss Plateau seems to be feasible. When comparing the Figure 13 with the Figure 1 from the study area, we can see that the western and eastern part of the Swiss Plateau, which showed high BI values, are also the areas which are characterized by very good suitability for cropland. The middle part of the Swiss Plateau shows less suitability for cropland and shows also lower BI values in Figure 13.

5.2 Brownest Pixel Composite for visualization

This product was calculated for the purpose of high resolution visualization and to show the BPC for entire Switzerland. Compared to the BPC of the Swiss Plateau, we do have a continuous map without missing and masked pixels. This could also be done for the BPC of the Swiss Plateau, but as we are there just interested in the bare soil pixels it is not necessary.

For this product, TOA satellite data were used. This means that the pixel values were not atmospherically corrected and will differ from the SR product. Therefore, the pixel values are contaminated by different processes in the atmosphere and are less appropriate to derive land surface properties (Vermote & Kotchenova, 2008). The difference of TOA and SR can be seen in Figure 31. It shows a single bare soil pixel inside the Swiss Plateau. In the NIR and SWIR region there are just small differences in the reflectance value. In the visible region of the spectra the differences between TOA and SR can be clearly seen. In summary, the effect of the atmosphere is highest in the blue and green band of Landsat-8. The blue band is used for calculating the BI. As a result, also the value of the BI will be different between SR and TOA and potentially, in some pixels, the date of the

maximum BI moment will be a different one compared to the BPC of the Swiss Plateau. The difference can also be seen in the RGB composites (Figure 14 and Figure 19), where the TOA product looks different compared to the SR product. This is the reason why these two products are not comparable. However, it is still useful to show an overview of the "brownest" composite of Switzerland. The improved spatial resolution of 15 m helps to produce a visually nice and continuous map.

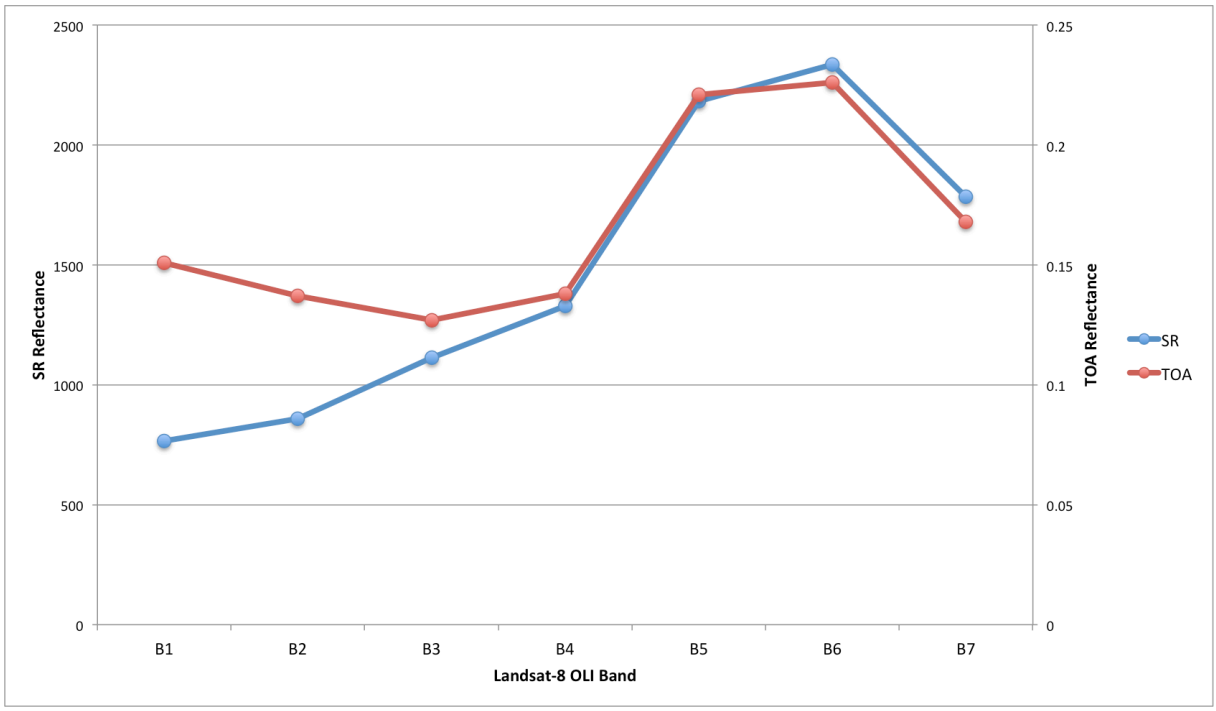

Figure 31: SR and TOA reflectance of the same bare soil pixel.

5.3 Threshold

5.3.1 Airborne Imaging Spectroscopy and Fieldwork

Table 6 showed that the amount of pixels per scene differs a lot. The amount of bare soil pixels in the analysis were limited by the availability of bare soil in the APEX images calculated from the HYSOMA software. Thus, the amount of bare soil pixels is fixed and cannot get increased. In contrast, the chosen fields for the vegetation class were selected manually and the amount of pixels is arbitrary. Not all the fields that show vegetation were selected for the analysis. The selection was based on the idea to choose fields that visually looked like a mixture between bare soil and vegetation. Accordingly, fields with sparse vegetation, very dry vegetation, or visually brown vegetation were chosen. Fully vegetated, green fields were just chosen sporadic, as the BI value would be obviously different to bare soil. With this technique, the vegetation class contains mainly BI values that are just slightly different to the bare soil fields. As a result, the threshold is more sensible for the difference between bare soil and mixed fields.

The overall boxplots (Figure 15) of the class bare soil and vegetation showed that there is a clear difference between the two classes. However, the vegetation class showed much more variance than the bare soil class. The reason for this was explained above. For the vegetation class a lot of mixture fields were selected that show also reflectance of bare soil. Thus, some of the pixels will show higher BI values. The same trend can also be seen in Figure 16 were a boxplot is shown for each scene separately. The vegetation class shows much more variance than the bare soil class.
In the fieldwork campaign, a field was also classified as vegetated when there were just little offshoots on the field. These fields showed a lot of bare soil in between. They got chosen because young vegetation has also an influence on the reflectance in the satellite image and these pixels should also get removed. This explains why the median of the fieldwork vegetation class shows just small differences to the fieldwork bare soil class.
The bare soil class of Greifensee in July 2015 shows with an average below 0 much lower BI values than all the other bare soil boxplots. This APEX scene was captured four days after the Landsat-8 image. It is therefore possible that in these four days some fields got harvested. A selected bare soil field in the APEX scene could therefore still show vegetation in the Landsat image. Hence, we would select the BI value of vegetation instead of bare soil. This would explain why the bare soil class shows values more similar to the vegetation classes.
In summary, we can clearly see that the two classes show different BI values in each scene. The degree of difference as well as the variance within the classes differs among the scenes. This can get explained by the varying time difference between the APEX and Landsat scenes as well as by the different amounts of selected pixels.

5.3.2 Calculation of the Threshold

The threshold was chosen with a tradeoff between high overall kappa and optimized user and producer accuracies of the class bare soil. The chosen threshold of 0.0175 lies around the expected value of zero and additionally in the lower part between the two peaks of Figure 12. At the threshold, the producer accuracy of bare soil is 0.92. This means that we misclassify 8% of the vegetation pixels as bare soil. Therefore, when building the MLR model for predicting the soil properties we have some disturbance reflectance from vegetation pixels. The user accuracy is also 0.92. This means that we are also misclassifying around 8% of the bare soil pixels as vegetation. This has an influence on the maximum amount of bare soil in the study area. When choosing a higher threshold, we could have decreased the amount of vegetation pixels that were misclassified as bare soil. Thus, the disturbance of vegetation pixels in the model building could have been decreased and the prediction accuracy may increase. On the other hand, we would then misclassify a lot of bare soil pixels as vegetation. Hence, the maximum amount of bare soil in the study area would decrease.
It shows that the chosen threshold can have a high influence on the final result. Choosing the right threshold depends on the question to be answered. In this thesis we have created two products, the

maximum bare soil composite as well as the prediction of soil properties. The threshold needs to be chosen so that it fits both tasks equally. For this purpose, the chosen threshold of 0.0175 is appropriate to distinguish between bare soil and other land cover classes.

Demattê et al. (2009) was developing a multi-step way to distinguish the bare soil pixels from others. First, each pixel was evaluated in an RGB composite to see if the pixel color looks like bare soil. The same was also done in a false color composite (SWIR, NIR, red), where bare soil should look purple. Next, a vegetation index was considered, were the pixel should have the value of zero followed by the analysis of a scatter plot of the NIR and red band. The pixel should lay near the soil line to be considered as bare soil (Baret et al., 1993). Finally, the spectral curve of the reflectance from the satellite was compared to spectral laboratory curves of samples from the same region. Only when all these five steps considered the pixel as bare soil, then it was used for further analysis. This method uses way more steps and the pixels were analyzed individually. For the case of a single Landsat image, which was used in many studies, this method seems to be appropriate. It can clearly evaluate more misclassifications than with the simple method used in this thesis. But it is not useful for studying a larger area with a BPC containing millions of pixels. For this aim, the development of an easy to use index together with a threshold, as used in this study, seems to be more adequate.

5.4 Bare Soil Analysis

The threshold enabled to classify each pixel into bare soil or grassland. The statistic showed that nearly 41% of the agricultural area showed bare soil and is therefore used as arable land. These 41% corresponds to a total area of 2488 km^2 of bare soil. For a reference, the area statistic of Switzerland (Bundesamt für Statistik, 2016) was considered. The spatial resolution of this product is 100 m. The land cover gets classified in several categories by expert judgment with the help of aerial images. Unfortunately, for our study area the statistic is not available. To get an approximate overview, the statistics of the cantons that are completely inside the Swiss Plateau (Zurich, Thurgau, and Geneva), together with the ones that are nearly completely covered (Lucerne, Aargau, Solothurn, Berne, Fribourg, and Vaud) were analyzed. The latest data from the year 2009 show that 43% of the agricultural area is classified as arable land (Bundesamt für Statistik, 2016). The percentages are nearly the same, which shows that the calculated bare soil map is feasible and could also be used to improve the area statistic. The much higher spatial resolution of 30 m helps to identify large-scale geographical distribution of bare soil and other agricultural classes. The distinction between grassland and bare soil over larger areas is also of great interest for the integration into land management models (Gomez Giménez et al., 2016).

The methodology used in this thesis is promising for generating a large and continuous bare soil map. Other published studies showed much less continuous soil area as they were often conducted with single Landsat images. For example, Demattê et al. (2016a) was also working with multi-temporal

Landsat images and was able to calculate soil properties for a bare soil area of 120 km^2. Other studies showed even less area of bare soil which is also related to the smaller extent their study was conducted. It shows that the new approach chosen in this thesis, with help of the GEE, can increase the study area and accordingly also the amount of bare soil mapped.

For this thesis, a study period of three consecutive years was chosen to increase the chance to map the field on a bare moment. Figure 32 shows the amount of bare soil pixels that would have been detected with a different study period. When just analyzing Landsat-8 images from one single year, we could receive a total bare soil area of between 2.5 million and 3.2 million pixels. The chosen period of three years was increasing the total bare soil area by 65% compared to the year 2014, 26% to 2015, and 52% to the year 2016. Therefore, a longer study period increases the amount of bare soil detected. An even longer study period may increase the total bare soil area further. For this thesis, the study period was limited by the available Landsat-8 data at the moment of processing. Figure 32 shows also that there are relevant differences between each single year. A reason for this could be the cloud coverage. Depending on the region in the study area there are around 18 to 50 images available per year, as explained in section 2.2. During the winter months, the Swiss Plateau suffers often from high fog and also in the summer months cloud coverage can occur. This could reduce the amount of available cloud-free images in some regions drastically, which would reduce also the chance to detect a bare soil field. This could be an explanation why there is much less bare soil detected in the years of 2014 and 2016. With an increased study period, the data availability per pixel is increased and the chance to detect bare soil, too. With this issue considered, there is a high chance that with an even longer time period than three years the total area of bare soil could get increased further.

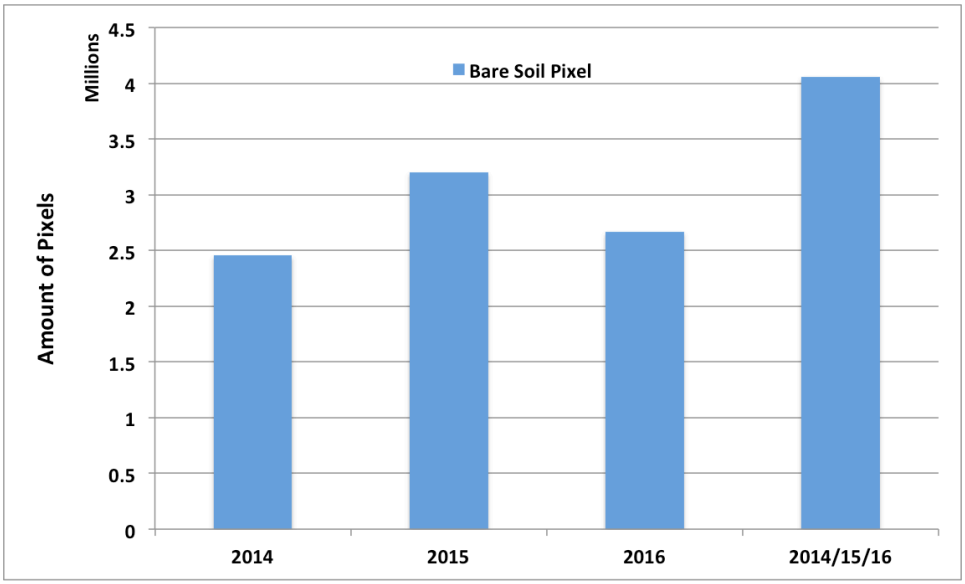

Figure 32: Amount of bare soil pixels captured in the study area with different study periods.

The BSC (Figure 19) shows the natural color composite of all the bare soil pixels. We can clearly see differences in the color of the bare soil. Some fields show very dark soils and other more bright ones. The dark ones are mainly clustered near the lakes (north of red focus area). The soil map of Switzerland classifies the soils in this area as Histosols and Fluvisols (Bundesamt für Landwirtschaft, 2012). These soils are characterized mainly by a high content of SOM (IUSS Working Group WRB, 2015). Chen et al. (2000) showed that dark soils contain a higher amount of SOM than more bright ones. This finding coincide the soil classification. Hence, out of the BSC we can already draw some conclusions about the soil properties. The soil property map will reveal if we can approve this assumption.

The distribution of the chosen bare soil pixels in the BSC (Figure 20) showed that there are months that seem to show more bare soil area than other months. A field normally remains bare in the time after harvesting until the next moment of seeding. During this period, satellite images can detect the bare soil fields. The highest amount of chosen bare soil pixels showed the month of August. In July, the harvesting period for all the winter crops (wheat, barley, and triticale) starts and some of these fields are also harvested in August (Franzen et al., in preparation). Therefore, a lot of fields show bare soil in August. In addition, the chance for low cloud coverage in August is high. The combination of these two factors explain the fact that most bare soil pixels showed the highest BI in August. In September and October, the maize, potatoes, and sugar beet get harvested which will also lead to several bare fields. In the same time winter crops are seeded (Franzen et al., in preparation). From December on, just a few pixels showed the highest BI value. This can mainly be explained with the weather conditions. The study area shows a lot of fog and cloud coverage in these months, and the chance to have a cloud free image is reduced. The second peak of chosen BI pixels can be seen in the months of March and April. The weather starts to get better and the seeding of the summer crops starts in April. Thus, several fields are still bare and can be detected from satellite images. From May on, the winter and summer crops are seeded and most of the fields will show vegetation. This explains the strong decrease in selected bare soil pixels from this time period on.

5.5 Soil properties calculation

5.5.1 Multiple Linear Regression with Spectral Data

Soil Samples

From the adjusted HSD, around 67% of the samples were not on a bare soil pixel and we lost the information for the further analysis. There are several reasons for these amounts of soil samples that were excluded. First, most samples were collected 20 years ago and earlier. Between 1985 and 2009, the urban area was increasing by 23.4% in Switzerland (Bundesamt für Statistik, 2016). In the same period also the forest area increased. The expansion of these areas was mainly at the expense of the agricultural area, which was reduced in total by 5% and for arable land even by 7%. Thus, many soil

samples were collected in regions that are not within agricultural area anymore. Second, some of the samples were collected in the forest area of Switzerland, near roads, or near settlement areas. Thus, they got removed during the masking step. Even though we needed to exclude many soil samples, there are still many samples left: 1580 SOM and 735 soil texture samples.

Reflectance Analysis

The mean reflectance of different amount of soil properties in the sample shows clear evidence that the soil properties and the reflectance values are in a way related. Clay shows a decrease in all bands from low to high content. Sand shows exactly the opposite, with increasing sand content also the reflectance increases. This trend can also be seen in the soil samples; soils with a high content of sand show normally a lower amount of clay, which would also explain this trend. The SOM reflectance shows also a clear decrease of reflectance with an increase in SOM content. In the case of silt, no clear trend is observable.

The same trends are also described in other studies. Dewitte et al. (2012) shows that sandy soils do reflect more than clayey soils. Furthermore, he shows that an increase in SOM leads to darker soil and therefore to less reflectance. Silt doesn't seem to have a strong effect on the reflectance. The reason for this could be that silt is defined as the particle size between sand and clay and doesn't show this strong difference to each of the other two classes. Additionally, when analyzing the soil texture diagram of the HSD in Figure 5 as well as Table 9 we can clearly see that the silt content shows much less variation than the sand and clay content. It seems that the content of clay and sand converse with each other. High sand content seem to lead often to a reduced amount of clay.

Variable Selection

The cross-validation revealed that the best model to predict sand and clay included all the band information except the blue one. However, the performance with all six bands would have been just slightly lower. The mean reflectance plots showed that the influence of sand and clay content can be clearly seen through all six bands. However, we can also see that the influence of the soil property increases absolutely with increasing wavelength. In the NIR and SWIR region the differences between the four classes are much higher than in the visible region. The blue band shows the smallest difference in the reflectance value. Hence, this could explain why the model was not selecting the blue band for the prediction of sand and clay. Additionally, in the SWIR region the clay minerals show some absorption features, especially in the wavelength region of SWIR2 (Ben-Dor & Demattê, 2016). This would explain also why the SWIR1 and SWIR2 bands were chosen for the MLR model. When having a look at the MLR equation (Table 8) of clay and sand, we can see that they are nearly contrary to each other. When a specific band value is added for the clay prediction, then the nearly same band value is subtracted from the sand prediction. This does also confirm the assumption that the sand and clay content in our soils are somehow correlated to each other. The same finding is also

described in the literature by Wetterlind & Stenberg (2010). They show that sand is often predicted as the mirror image of clay because of the featureless quartz and feldspars in the sand.

The cross-validation for the prediction of silt showed that the best model was built out of just two independent variables. Afterwards, the RMSE increases again. Although, it needs to be considered that the differences in RMSE are just very small. The chosen bands for silt prediction are the green and SWIR2 wavelength region. The reflectance plot from Figure 23 does not show a clear trend to a specific wavelength region. Therefore, it stays unclear why the green and SWIR2 was chosen. In the study of Demattê et al. (2007) the SWIR1 band from Landsat-7 was chosen to predict silt. The calculated MLR equation (Table 8) shows that the influence of the two chosen independent variables is very small compared to the MLR equations of the other soil properties. This explains why the predicted silt content shows just very small differences between the samples and probably also why the model performed very poorly.

For the SOM prediction, every band except SWIR1 was chosen. SOM shows spectral activity on the whole spectra but especially in the visible part (Ben-Dor & Demattê, 2016). This could explain why nearly all the six bands were chosen for the predicting model (SWIR1 was not chosen, but the model with all six bands would perform just marginally weaker).

Model Performance

The performances of the MLR models were best for the prediction of clay and SOM, followed by sand. For silt, the MLR model performed very poor. Following, we present some other studies that were using satellite data to predict different soil properties.

The few studies that were conducted with satellite data showed very variable R^2 values. Demattê et al. (2007) was calculating also MLR equations and received for one test-site with 60 soil samples R^2 values of 0.67 for sand, 0.63 for clay, 0.29 for silt, and 0.27 for SOM. Except for SOM their models were performing better than in this thesis. However, for another study area with 50 soil samples the models were performing much worse. The R^2 for sand was 0.08, for clay 0.18, for silt 0.12, and for SOM the model did not show any correlation. Though, with the combination of both study areas they could increase the model performances again to 0.88 for sand, 0.87 for clay, 0.70 for silt, and 0.77 for SOM. The conclusion of this study was that the broader range of data and the increased amount of samples in the combination of both test sites allowed better performance of the MLR models (Demattê et al., 2007). This conclusion increases the value of our work as we were working in a big study area and with several hundreds soil samples. However, it shows also that the performance of the models is very site specific and dependent of the soil types and the variance inside the study area. The most soil samples in our study area show similar percentages of the soil textures. This can be seen in Figure 5 where most of the soils are classified as loam or clay loam. Thus, the variance in our soil samples is very small. This could also be the reason why the MRL models were not performing very well for samples that show property values much higher or much lower than the average.

The study of Nanni & Demattê (2006) using again Landsat data was predicting the soil properties sand with an R^2 of 0.53, clay with 0.68, and SOM with 0.51. For silt, the model showed no correlation. In the study, they were also predicting the soil properties with simulated Landsat bands using laboratory data. These MLR models were predicting the soil properties better, which showed that different factors like the atmosphere or the soil moisture content are having an influence on the spectral response.

There are also other studies that performed worse in predicting soil properties. Liao et al. (2013) was predicting sand with an R^2 of 0.32, silt with 0.21, and clay with 0.36. Their variable selection for the MLR model was just choosing the SWIR2 band to predict all the properties. Although, they have been using TOA satellite images to predict the soil properties. In section 5.2, we showed that this dataset is influenced by the atmosphere.

The few studies that were using satellite data to predict soil properties showed very variable performances depending on the chosen study area and the amount of soil samples. However, in general the results of the prediction models can be seen as quite good when considering that the satellite sensor is far away from the ground and can just collect information from the top part of the topsoil. Our study, as well as the other studies showed that from the three soil textures (clay, sand, silt) the MLR model of clay performs generally best. The models for sand showed also feasible correlations. Though, silt cannot get predicted accurately. Clay shows normally the best prediction most likely because of the spectral signatures of the clay minerals in the NIR and SWIR part of the wavelength spectra (Stenberg et al., 2000). The bad performance from silt was already discussed earlier. Most likely it is because of the particle size that can be assumed to contain a mixture of sand and clay minerals and is therefore more difficult to get distinguished (Wetterlind & Stenberg, 2010).

In summary, we see that most other studies showed generally better performing prediction models than in this thesis. But it needs to get considered that we are using a way more generalized method that can be used to predict soil properties over a large area and for a long time series. In addition, the soil samples used in this study were mainly collected between 20 and 50 years ago and from different projects with different sampling methods. In contrast, all the described studies were collecting the soil samples during their study period and with sampling methods that fit exactly the need of their study. In our study we were also using the Landsat-8 reflectance data from a time span of three years that show different meteorological conditions. Therefore, it is showing way more variability than by just using one single scene. By considering these facts, the model performances in this study can be seen as good and could most likely get improved by further enhancing the methods of this approach.

5.5.2 Prediction of Soil Properties

The prediction of the clay content shows us clear differences over the study area. Conspicuous is the area around the lake in the western part of the Swiss Plateau. It shows clearly a bigger connected area with the highest clay content. This area is classified as Fluvisol or Histosol (Bundesamt für

Landwirtschaft, 2012) and is known to have been flooded for several times in the past. Afterwards, the area got drained during the "Jura water correction" and is now a prominent used agricultural area (Brändle et al., 2015). Nancy et al. (2011) showed that soils, which were flooded from time to time, show higher clay contents than soils that were not flooded. Therefore, it seems to be feasible that the prediction maps show in this area higher amount of clay than in the areas around.

The prediction of the sand content shows the contrary to the clay prediction map. The soils with a high amount of clay show lower amount of sand and the other way round. This distribution corresponds to the assumptions we did before. The MLR equation of sand showed already that it is nearly contrary to the MLR equation of clay. Therefore, high clay content leads of necessity to a lower sand content. The assumptions we did before, that the amount of clay and sand is somehow related, can be clearly seen in the two prediction maps and was also shown in literature (Wetterlind & Stenberg, 2010).

The prediction of the SOM shows even more conspicuous the two areas in the western part of the Swiss Plateau with the highest SOM content. Exactly these two areas are described by Brändle et al. (2015) as soils with a high amount of SOM. The reason for this is again the flooding which occurred regularly in the past. Additionally, it confirms the assumption from the bare soil RGB analysis (section 5.4) that the dark soils refer to high SOM content. On contrary to the other two prediction maps, the SOM prediction showed also percentages below zero. This is naturally not possible but can get explained by literature. Al-Abbas et al. (1972) describe that the SOM has just an influence on the soil reflectance when the content is higher than 2%. The reflectance of soils with a lower SOM content is dominated by other soil properties. This could explain why the model was not working well for low SOM values and why some pixels got predicted with values below zero.

The three soil properties clay, sand, and SOM could get successfully predicted over the whole study area of the Swiss Plateau. The spatial distribution of the properties is in general suitable and especially in the area around the lakes the predicted soil properties correspond to the literature and to the existing soil map of Switzerland. Due to the limited model performances, the predicted soil property maps should be used carefully for determining the exact amount of a specific property in the soil. Nevertheless, the maps can get used to get an overview about the geographic distribution and to illustrate in which areas rather higher or lower amounts occur. Furthermore, the soil property maps or the BSC could also get used as one of the many covariates in a digital soil mapping (DSM) approach (Nussbaum, 2017). The large area of bare soil reflectance from the BSC could increase the output of the DSM models.

The results show a much higher spatial resolution than any other available soil maps in Switzerland and shows differences even in an in-field scale. This information is valuable for different stakeholders. For example, the farmers can use this information to increase their decisions about crop management practices (Ge et al., 2011). Additionally, the soil property maps can also get used to calculate different soil functions like the water storage capacity. Another improvement is that the soil

property map is calculated for the whole Swiss Plateau, which contains parts of 14 cantons of Switzerland. Normally, soil maps are calculated for every canton individually with different methods and on different resolutions. Our method creates a coherent and borderless map that shows the amount of soil properties nearly through entire Switzerland. Therefore, it is also valuable for comparison.

5.6 Limitations and Outlook

We showed the high potential of our methods to calculate the BPC, the BSC, as well as the prediction of the soil properties. Aside the great value, these products show also some limitations but also great potential, too.

Soil moisture has a high effect on the amount of reflectance of a bare soil pixel. The reflectance decreases in the entire spectra with increasing soil moisture (Bogrekci & Lee, 2004). This can have an effect on the prediction of the of soil properties. For example, an increase in SOM or clay content is also decreasing the reflectance. The BSC in this study was composed from different pixels of a three year time period. Hence, the pixels were selected from different meteorological conditions. Therefore, some pixels will be chosen from dry and other from wet conditions and the amount of soil moisture in the bare soil varies strongly. To reduce the effect of soil moisture, other multi-temporal studies were conducted just in dry regions (Shabou et al., 2015), were using images from always the same month but different years (Demattê et al., 2016a) or were correcting the effect of soil moisture (Diek et al., 2016). However, the correction by Diek et al. (2016) was done using imaging spectroscopy data and the method cannot be adapted for this thesis. Hence, several methods exist to predict the soil moisture content from Landsat data (Zhan et al., 2007; Ahmad et al., 2011). However, to apply one of these techniques to a composite consisting of 360 Landsat images like the BSC remains challenging, but would most likely improve the performance of the prediction models.

The method in this thesis shows also some potential for further increasing its scale. The focus of this thesis lied on the agricultural area of Switzerland. Although, the methods for the calculation of the BPC (working with GEE and the use of an easy to calculate BI) was chosen so that it is also scalable for different study sizes and regions. This gives the opportunity to upscale the product to a European or even Worldwide level. Although, it works just for agricultural fields that show bare soil after harvesting. To reduce artifacts in the final product, it is crucial to exclude non-agricultural areas from the calculation of the BPC. In section 3.2.2 different products were introduced that are available on a European (Permanent water bodies and Imperviousness dataset) or even worldwide level (Hansen Global Forest Watch). With the use of these products it would be an easy step to upscale it to a larger area. Although, the BPC can be easily calculated, the threshold to distinguish between bare soil and other land cover type may differ from region to region. To receive a more or less accurate bare soil product, the threshold needs to be evaluated in different regions and needs to be adapted potentially.

The distinction between bare soil and grassland can also get used in a longer time series to receive information about the crop rotation. For every pixel we would then know in which months the pixel showed bare soil and how often it showed bare soil in a specific time period. This would give us valuable information about how extensive specific fields get used. The date of the bare soil moment would give us also information about if winter or summer crops were planted. However, the potential of such a time series analysis still needs to get evaluated.

Another improvement of the BPC as well as for the prediction of the soil properties could be done with the use of the new Sentinel-2 satellites. The two Sentinel-2 satellites show a higher spatial resolution of up to 10 m, a higher spectral resolution of 13 bands, and in combination also a much higher temporal resolution of 5 days (Gatti et al., 2016). The much higher temporal resolution would help to increase the maximum amount of bare soil captured in the BSC. In addition, the more spectral bands would most likely increase the performance of the MLR models.

We showed that our product has the potential for new applications as well as for upscaling the study area. With the use of newer satellite technologies, an increasing predicting accuracy may also be possible.

6. Conclusion

Soil as a resource contributes to many ecosystem services and is crucial for our everyday life. However, the resource is continuously under pressure as a result of global environmental changes as well as the increase in population. A large amount of soils is already highly degraded which leads to hampered soil functions. Due to the importance of this resource we have a high interest in up-to-date soil information on local to global scale. Traditional soil sampling methods for producing this information are not appropriate anymore, as they are very time consuming, costly, and often not standardized. With the use of newer technologies like remote sensing we can close this gap of information. In this thesis we used satellite remote sensing images to calculate soil properties over a large area.

First, we developed a new and simple method to map the “brownest” moment of a pixel in a longer time series. We calculated this for the area of the Swiss Plateau as well as for the whole Switzerland for visualization purposes.

Second, we used additional airborne remote sensing images to define a threshold. This threshold gave us the opportunity to classify each pixel into bare soil or any other land cover type. We then received an overview about the maximum extent of bare soil in the study area during the given study period.

Third, the spectral reflectance of these bare soil pixels together with information from soil samples were used to predict different soil properties for every bare soil pixel. We were able to predict the

amount of clay, sand, and SOM in the topsoil with an adequate accuracy and in a high spatial resolution of 30 m. The research questions of this thesis can get answered the following.

- Which method is suitable to generate a brownest pixel composite over a large area?

The chosen method should be able to get applied over a large area and in a longer time series. We have chosen two easy to apply indices that fulfill these prerequisites. The followed analysis of the minimum NDVI and maximum BI composites showed clear differences. A lot of artifacts were visible in the minimum NDVI composite that were not present in the maximum BI composite. The discussion showed that the additional two spectral bands in the calculation of the BI changes the chosen pixel and fewer artifacts were selected for the final product. The calculation of the brownest pixel composite in the study area was done using GEE and took just few minutes. To mask the non-agricultural area, auxiliary data were used. Some of these datasets are also available on a European or even worldwide level. Therefore, the brownest pixel composite can get easily calculated for a large study area. In summary, the chosen methods of the BI together with auxiliary masking datasets as well as the GEE are suitable to generate a brownest pixel composite over a large area.

- What is the optimal threshold to distinguish bare soil from other land cover types?

For our study area, the optimal threshold to distinguish bare soil from other land cover types is the BI value of 0.0175. At this threshold, the user and producer accuracy of the bare soil class is equal. This means that we misclassify the same percentage of bare soil pixels as vegetation as the other way round. For the calculation of the soil properties, we want to reduce the amount of vegetation pixels that are misclassified as bare soil to reduce the amount of spectral response from vegetation in the MLR models. Although, when reducing this misclassification then we would also increase the amount of bare soil pixels that get misclassified as vegetation. This would reduce the maximum bare soil extent in the study area and would reduce the profit of the bare soil composite. Therefore, the chosen threshold has a high influence on the result. In our study, the threshold is chosen as a tradeoff between these two products. Hence, the threshold of 0.0175 is the optimal threshold to distinguish bare soil from other land cover types in the case of our study. Although, it needs to get considered that the calculation of the threshold is highly dependent on the chosen data and methods.

- What is the maximum extent of bare soil in the study area of the Swiss Plateau during the three years of investigation?

We were able to map a total area of 2488 km^2 of bare soil in the study area. This corresponds to 40.8% of the agricultural area of the Swiss Plateau. Comparison with the area statistic of Switzerland showed that this amount is feasible. This is the maximum extent in the study period of three years. To get the maximum bare soil extent in the study area, the study period has to get increased.

- Can remote sensing techniques help to improve the available soil products and predict structural soil properties over the agricultural area?

We were able to predict the soil properties clay, sand, and SOM for the bare soil composite in the whole study area. The model for silt did not perform well enough and therefore, no map was created. The maps were calculated in a high spatial resolution of 30 m which is more detailed than other existing soil maps in Switzerland. This makes it possible to see differences even on an in-field level, which can be interesting for farmers. As the model performances were limited, the maps are not useful to receive accurate information about the total amount of a soil property in the topsoil. However, the general distribution of the soil properties seem to be feasible and therefore, the maps can get used to gain an overview about geographic differences. In summary, remote sensing techniques can be used to create coherent and borderless soil products for large and even remote areas for fewer costs.

This thesis showed the great value of remote sensing technologies for the soil science community. Further research must be done to avoid some of the limitations. This would help to increase the value of the results and enables the possibility for new applications and upscaling of the study area.

Acknowledgements

After one year of work on this thesis, I would like to thank the people who supported me during this time. First and foremost I would like to thank my supervisor Rogier de Jong for always supporting me throughout my thesis and providing me with many helpful inputs and information. A special thanks goes also to my co-supervisor Sanne Diek for the interesting discussions as well as the help in various topics and all the answers to my questions.
Furthermore, I would like to thank everyone who worked with me in one of the master thesis rooms for the funny times and all the nice coffee and lunch breaks. In the end, I would also like to thank my parents for the endless support throughout my studies.

References

Ahmed, Z., Iqbal, J., 2014. Evaluation of Landsat TM5 multispectral data for automated mapping of surface soil texture and organic matter in GIS. Eur. J. Remote Sens. 47, 557–573.

Al-Abbas, A.H., Swain, P.H., Baumgardner, M.F., 1972. Relating Organic Matter and Clay Content to the Multispectral Radiance of Soils. Soil Sci. 114, 477–485.

Amundson, R., Berhe, A.A., Hopmans, J.W., Olson, C., Sztein, A.E., Sparks, D.L., 2015. Soil and human security in the 21st century. Science. 348, 1261071.

Arrouays, D., McKenzie, N., Hempel, J., Richer de Forges, A., McBratney, A.B., 2014. GlobalSoilMap: Basis of the global soil information system. CRC press.

Ballabio, C., Panagos, P., Monatanarella, L., 2016. Mapping topsoil physical properties at European scale using the LUCAS database. Geoderma 261, 110–123.

Baret, F., Jacquemoud, S., Hanocq, J.F., 1993. About the soil line concept in remote sensing. Adv. Sp. Res. 13, 281–284.

Ben-Dor, E., Irons, J.A., Eperma, A., 1999. Soil Reflectance, in: Remote Sensing of the Earth Science. John Wiley & Sons, pp. 111–189.

Ben-Dor, E., Demattê, J.A.M., 2016. Remote Sensing of Soil in the Optical Domains, in: Remote Sensing Handbook. CRC press, pp. 733–787.

Bogrekci, I., Lee, W.S., 2004. The effects of soil moisture content on reflectance spectra of soils using UV-VIS-NIR spectroscopy. Proc. 7th Int. Conference Precis. Agric. Other Precis. Resour. Manag. 1307–1317.

Bonnard, L.F., 1999. Soil Survey in Switzerland, European Soil Bureau Research Report.

Brady, N.C., Weil, R.R., 2017. The Nature and Properties of Soils.

Brändle, G., Zihlmann, U., Chervet, A., 2015. Der Moorboden, Bodenkundliche Gesellschaft der Schweiz.

Bundesamt für Landestopographie swisstopo, 2008. Landeskarte 1:25000.

Bundesamt für Landwirtschaft, 2012. Minimales Geodatenmodell. 77.2 Digitale Bodeneignungskarte der Schweiz.

Bundesamt für Statistik, 2016. Arealstatistik 2004/09.

Chabrillat, S., Eisele, A., Guillaso, S., Rogaß, C., Ben-Dor, E., Kaufmann, H., 2011. HYSOMA: An easy-to-use software interface for soil mapping applications of hyperspectral imagery. Proc. 7th EARSeL SIG Imaging Spectrosc. Work. 1–7.

Chen, F., Kissel, D.E., West, L.T., Adkins, W., 2000. Field-Scale Mapping of Surface Soil Organic Carbon Using Remotely Sensed Imagery. Soil Sci. Soc. Am. J. 64, 746.

Demattê, J.A.M., Galdos, M. V., Guimarães, R. V., Genú, A.M., Nanni, M.R., Zullo, J., 2007. Quantification of tropical soil attributes from ETM+/LANDSAT-7 data. Int. J. Remote Sens. 28, 3813–3829.

Demattê, J.A.M., Huete, A.R., Ferreira Jr., L.G., Nanni, M.R., Alves, M.C., Fiorio, P.R., 2009. Methodology for Bare Soil Detection and Discrimination by Landsat TM Image. Open Remote Sens. J. 2, 24–35.

Demattê, J.A.M., Alves, M.R., Terra, F. da S., Bosquilia, R.W.D., Fongaro, C.T., Barros, P.P. da S., 2016a. Is it possible to classify topsoil texture using a sensor located 800 km away from the surface? Rev. Bras. Cienc. do Solo 40, 1–13.

Demattê, J.A.M., Morgan, C.L.S., Chabrillat, S., Rizzo, R., 2016b. Spectral Sensing from Ground to Space in Soil Science: State of the Art, Applications, Potential, and Perspectives, in: Remote Sensing Handbook. CRC press, pp. 661–732.

Dewitte, O., Jones, A., Elbelrhiti, H., Horion, S., Montanarella, L., 2012. Satellite remote sensing for soil mapping in Africa: An overview. Prog. Phys. Geogr. 36, 514–538.

Diek, S., Schaepman, M.E., de Jong, R., 2016. Creating multi-temporal composites of airborne imaging spectroscopy data in support of digital soil mapping. Remote Sens. 8.

Fiorio, P.R., Demattê, J.A.M., 2009. Orbital and Laboratory Spectral Data To Optimize Soil Analysis. Sci. Agric. 66, 250–257.

Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P.M., O'Connell, C., 2011. Solutions for a cultivated planet. Nature 478, 337–42.

Franzen, J., Müller, M., Keller, A., n.d. in prep. Stoffbilanzen für Parzellen der Nationalen Bodenbeobachtung - Nähr- und Schadstoffe 1985 - 2014. Agroscope Sci.

Gatti, A., Bertolini, A., Nasuti, C., Carriero, F., 2016. Sentinel-2 Products Specification.

Gareth, J., Witten, D., Hastie, T., Tibshirani, R., 2017. An Introduction to Statistical Learning, Springer.

Ge, Y., Thomasson, J.A., Sui, R., 2011. Remote sensing of soil properties in precision agriculture: A review. Front. Earth Sci. 5, 229–238.

Gnägi, C., Labhart, T.P., 2015. Geologie der Schweiz, 9th ed. Ott Verlag.

Gómez Giménez, M., Della Peruta, R., De Jong, R., Keller, A., Schaepman, M.E., 2016. Spatial Differentiation of Arable Land and Permanent Grassland to Improve a Land Management Model for Nutrient Balancing. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9, 5655–5665.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ.

Hansen, M.C.C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S.A. a, Tyukavina, A., Thau, D., Stehman, S.V. V, Goetz, S.J.J., Loveland, T.R.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O.O., Townshend, J.R.G.R.G., Patapov, P.V., Moore, R., Hancher, M., Turubanova, S.A. a, Tyukavina, A., Thau, D., Stehman, S.V. V, Goetz, S.J.J., Loveland, T.R.R., Kommaredy, A., Egorov, A., Chini, L., Justice, C.O.O., Townshend, J.R.G.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science. 342, 850–854.

Huang, H., Chen, Y., Clinton, N., Wang, J., Wang, X., Liu, C., Gong, P., Yang, J., Bai, Y., Zheng, Y., Zhu, Z., 2017. Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine. Remote Sens. Environ.

Hueni, A., Biesemans, J., Meuleman, K., Dell'Endice, F., Schlapfer, D., Odermatt, D., Kneubuehler, M., Adriaensen, S., Kempenaers, S., Nieke, J., Itten, K.I., 2009. Structure, components, and interfaces of the airborne prism experiment (APEX) processing and archiving facility. IEEE Trans. Geosci. Remote Sens. 47, 29–43.

IUSS Working Group WRB, 2015. World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps, World Soil Resources Reports No. 106.

Jamalabad, M., Abkar, A.A., 2004. Forest canopy density monitoring, using satellite images. XXth ISPRS Congr. 12–23.

Langanke, T., 2016. GIO land High Resolution Layers - summary of product specifications.

Liao, K., Xu, S., Wu, J., Zhu, Q., 2013. Spatial estimation of surface soil texture using remote sensing data. Soil Sci. Plant Nutr. 59, 488–500.

Lillesand, T., Kiefer, R., Chipman, J., 2008. Remote Sensing and Image Interpretation, 6th ed. John Wiley & Sons.

MeteoSchweiz, 2013. Klimaszenarien Schweiz – eine regionale Übersicht. Fachbericht Meteoschweiz 243, 36.

Mulder, V.L., de Bruin, S., Schaepman, M.E., Mayr, T.R., 2011. The use of remote sensing in soil and terrain mapping - A review. Geoderma 162, 1–19.

Mulder, V.L., 2013. Spectroscopy-supported digital soil mapping.

Nancy, M.W., Njue, A.M., Abaya, S.G., Vuai Said, A.H., Lbembe, J.D., 2011. Periodic flooding and land use effects on soil properties in Lake Victoria basin. African J. Agric. Res. 6, 4613–4623.

Nanni, M.R., Demattê, J.A.M., 2006. Spectral Reflectance Methodology in Comparison to Traditional Soil Analysis. Soil Sci. Soc. Am. J. 70, 393.

Nanni, M.R., Demattê, J.A.M., Chicati, M.L., Fiorio, P.R., Cézar, E., Oliveira, R.B., 2012. Soil surface spectral data from Landsat imagery for soil class discrimination. Acta Sci. Agron. 34, 103–112.

Nocita, M., Stevens, A., Noon, C., Van Wesemael, B., 2013. Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy. Geoderma 199, 37–42.

Nussbaum, M., 2017. Digital Soil Mapping for Switzerland. Evaluation of Statistical Approaches and Mapping of Soil Properties.

Omuto, C., Nachtergaele, F., Rojas, R., 2013. State of the Art Report on Global and Regional Soil Information: Where are we? Where to go? Food and Agriculture Organization of the United Nations.

Piyoosh, A.K., Ghosh, S.K., Piyoosh, A.K., Ghosh, S.K., 2016. Development of a modified bare-soil and urban index for Landsat 8 satellite data. Geocarto Int. 6049, 0–1.

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing.

Rehbein, K., van der Meer, M., Grob, U., Wegmann, F., Keller, A., 2011. Das Nationale Bodeninformationssystem NABODAT in der Schweiz. Dtsch. Bodenkundliche Gesellschaft Jahrestagung.

Rikimaru, A., Roy, P.S., Miyatake, S., 2002. Tropical forest cover density mapping. Trop. Ecol. 43, 39–47.

Sanchez, P. a, Ahamed, S., Carré, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., Mcbratney, A.B., Mckenzie, N.J., De, M., Mendonça-santos, L., Minasny, B., Montanarella, L., Okoth, P., Palm, C. a, Sachs, J.D., Shepherd, K.D., Vågen, T., Vanlauwe, B., Walsh, M.G., Winowiecki, L. a, Zhang, G., 2009. Digital Soil Map of the World. Science. 325, 6–7.

Schaepman, M.E., Jehle, M., Hueni, A., D'Odorico, P., Damma, A., Weyermann, J., Schneider, F.D., Laurent, V., Popp, C., Seidel, F.C., Lenhard, K., Gege, P., Küchler, C., Brazile, J., Kohler, P., De Vos, L.,

Meuleman, K., Meynart, R., Schläpfer, D., Kneubühler, M., Itten, K.I., 2015. Advanced radiometry measurements and Earth science applications with the Airborne Prism Experiment (APEX). Remote Sens. Environ. 158, 207–219.

Shabou, M., Mougenot, B., Chabaane, Z., Walter, C., Boulet, G., Aissa, N., Zribi, M., 2015. Soil Clay Content Mapping Using a Time Series of Landsat TM Data in Semi-Arid Lands. Remote Sens. 7, 6059–6078.

Shukla, M., 2014. Soil Physics: An Introduction. CRC Press.

Stenberg, B., Viscarra Rossel, R.A., Mouazen, M., Wetterlind, J., 2010. Visible and Near Infrared Spectroscopy in Soil Science. Adv. Agron. Vol 107 107, 163–215.

USGS, 2016. Landsat 8 (L8) Data Users Handbook.

USGS, 2017. Product Guide - Landsat 8 Surface Reflectance Code (LaSRC) Product.

Vermote, E.F., Kotchenova, S., 2008. Atmospheric correction for the monitoring of land surfaces. J. Geophys. Res. 113, 1–12.

Walthert, L., Bridler, L., Keller, A., Lussi, M., Grob, U., 2016. Harmonisierung von Bodendaten. Schlussbericht von PMSoiL.

Wetterlind, J., Stenberg, B., 2010. Near-infrared spectroscopy for within-field soil characterization: Small local calibrations compared with national libraries spiked with local samples. Eur. J. Soil Sci. 61, 823–843.

Wulf, H., Mulder, T., Schaepman, M.E., Keller, A., Jörg, P.C., 2014. Remote Sensing of Soils, 1–71.

Zhan, Z., Qin, Q., Ghulan, A., Wang, D., 2007. NIR-red spectral space based new method for soil moisture monitoring. Sci. China Ser. D Earth Sci. 50, 283–289.

Zhao, H., Chen, X., 2005. Use of normalized difference bareness index in quickly mapping bare areas from TM/ETM+. Proceedings. 2005 IEEE Int. Geosci. Remote Sens. Symp. 2005. IGARSS '05. 3, 1666–1668.

Personal Declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Fabio Fornallaz
Zurich, September 2017