Masterthesis

Assessment of Soil erosion parameters in Costa Rica using reflectance hyperspectral and simulated EnMAP imagery

Sarah Sophia Malec

Global Change Ecology (M.Sc.)
University of Bayreuth, Germany
✉ sasoma@posteo.de

June 25, 2015

1. Supervisor
Dr. Martin Wegmann
University of Würzburg
German Aerospace Center

2. Supervisor
Prof. Dr. Thomas Köllner
University of Bayreuth

1. External Supervisor
Dr. Derek Rogge
German Aerospace Center

2. External Supervisor
Dr. Uta Heiden
German Aerospace Center

Contents

List of Figures

1 | Background and Motivation

1.1 Scope of Thesis

In preparation of the launch of EnMAP in 2018 there is a need to assess the capabilities of this satellite to provide useful information for various types of environmental challenges related to human activity and climate change. One such challenges is the prevention of land degradation. In 2014 the United Nations Food and Agriculture Organization (FAO) was calling for "urgent action to improve the health of the world's limited soil resources to ensure that future generations have enough supply of food, water and energy" (FAO, 2014). Population growth and the accompanying increase in land use (e.g. demand of land for food production, urban growth) on large spatial scales, highlights the importance of utilizing remote sensing as instrument for providing high quality information on the spatial extent of land degradation and making it available to decision makers, thus meeting the obligations claimed by FAO. One major aspect of land degradation is the loss of top soils owing to erosional processes, which have multiple environmental and socio-economic consequences. According to the FAO 2014 the definition for soil degradation follows largely the definition of the term land degradation, which is defined as the reduction in the capacity of land to provide ecosystem goods and services securing its function for all beneficiaries of the land. As such, the key objective of this study is to assess the capability of using simulated EnMAP imagery to derive essential landcover information, specifically photosynthetic active vegetation (PV), non-photosynthetic active (NPV, inlcuding standing dry vegetation and litter) and bare soil (including stones and rock) fractional cover, that can be used in combination with terrain information (e.g slope) to estimate the potential of soil erosion in a regional context. The region of interest in this study is located west of San Jose, Costa Rica, and comprises a highly fragmented landscape characterized by urban, agricultural areas, and forests in a mountainous terrain.

The study makes use of a large scale airborne HyMap hyperspectral survey

flown over portions of Costa Rica in 2005 and the EnMAP End-to-End Simulation tool (EeteS) developed at the Deutsches GeoForschungsZentrum (GFZ) (Segl et al., 2012) to generate a EnMAP simulation. Specific parameters are assessed that can be linked to soil erosion potential, specifically relative surface fractional cover of PV, NPV and bare soil that can be derived directly from hyperspectral imagery using existing automated approaches without the need for ground based data. Combining fractional cover estimates with terrain info a subsequent assessment of soil erosion parameters is conducted, such as the USEL C-factor or slope length and steepness factor (LS factor) to underpin the potential of hyperspectral data to assess soil erosion on a large spatial scale without the need of cost and labour intensive field surveys. Fractional cover estimates are derived from the reflectance imagery using a spectral mixture analysis approach, specifically a Multiple End-member Spectral Mixture Analysis (MESMA) approach. The geomorphological properties are derived from a digital elevation model (DEM). The analysis will be done independently for the airborne HyMap, which will be treated as reference data as no ground truth is available, and simulated spaceborne EnMAP imagery. The use of airborne data as reference was used by (Roberts et al., 2003) to assess the capacity of Hyperion data for fire danger analysis and in Rogge et al. (2014) airborne data was used to assess EnMAP for geological mapping. Subsequently, from a detailed qualitative and quantitative comparison of the EnMAP results with HyMap as reference, insights will be given into the capacities of the upcoming EnMAP satellite for the task of providing useful information related to estimating soil erosion potential at a regional scale.

1.2   State of the Art

The drivers of soil erosion can be described as a function of the biophysical environment determining the type of the degradation process. These parameters are soil properties, climate variables, the terrain, vegetation, management and time (Renard et al., 2011; Lal, 2003). Shoshany et al. (2013) describes the degradation process of soil loss considering the Universal Soil Loss Equation (USLE). The factors of USLE include, next to climate variables and erosion parameters, topographic indices and the C factor. The C factor is a surface cover and management factor, capturing the effects of vegetation canopy and ground covers in reducing soil loss (Shoshany et al., 2013; De Asis and Omasa, 2007). Many other modelling approaches exist to detect and quantify soil erosion processes and rates that utilize similar input parameters as the USLE and the Revised USLE, yet there implementations differ. For example, the European Soil Erosion Model (EUROSEM)

has been applied for modelling event-based soil erosion in Costa Rica, Nicaragua and Mexico (Veihe et al., 2001) and Bakimchandra (2011) introduces a Fuzzy-GIS based approach to model regional soil erosion risks in the Upper Awash Basin, Ethiopia.

From these studies and other studies it has been shown that optical Earth Observation (EO) can be a useful tool to measure landscape components influencing the condition of the land on a large geographical scale. For example, Asner et al. (2005) have used this concept for estimating canopy gap fractions and fractional photosynthetic cover to assess land-cover change, ecological variability, and biogeochemical processes across the Amazon and bordering cerrado regions of Brazil. Guerschman et al. (2015) provide an overview about a wide range of studies targeting NPV, PV and bare soil fractional cover mapping using multispectral data like Landsat, ASTER and MODIS. This review has shown the great importance of such inventories due to the long-term availability of such sensors and the accessibility of data. Bayer (2012) made use of spectral unmixing for fraction cover estimation of vegetation and different soil chemicals for modeling soil organic matter as indicators for the assessment of ecosystem degradation in South Africa. The study used airborne hyperspectral images and field observation data such as ground reflectance, vegetation cover and density as well as information about surface conditions. Delegido et al. (2015) have discriminated brown from green vegetation using a modification of the Leaf Area Index (LAI). All studies make use of spectral features related to the different classes. In the visible region the spectral signal of NPV and PV can be clearly distinguished based on the shape of the reflectance curve (Delegido et al., 2015). In contrast the detection of NPV and exposed soil is more challenging. However, specific absorption feature of cellulose and lignin occurring in the SWIR region of the spectrum help to differentiate between NPV and bare soil (Delegido et al., 2015). These studies show, that a reliable differentiation between NPV and soil, and between various types of soils is difficult when spectral information is limited in the short-wave infrared (SWIR) region, as it is the case for existing multi-spectral EO systems (Asner and Lobell, 2000; Elvidge, 1990; Okin, 2007).

The capabilities of EnMAP for environmental and geological applications have been demonstrated by a few studies. For example, Rogge et al. (2014) considered airborne AISA and simulated EnMAP data for mapping geology in a subarctic terrain. The study demonstrated that 2 m AISA imagery better discriminated between different rock types compared to the 30 m simulated EnMAP imagery, but the EnMAP data was still able to distinguish broad scale lithological units. The key differences were attributed to the large difference in

spatial resolution. Another study by Okujeni et al. (2015) demonstrate that EnMAP imagery is well suited for mapping impervious, vegetation and soil surface types and even allows a more detailed discrimination of sub classes. Similar to the results of Rogge et al. (2014) limitations for mapping classes of interest arise from the 30 m spatial resolution of EnMAP. Schwieder et al. (2014) compared the performance of different machine learning regression techniques, namely Support Vector Regression, Random Forest Regression, and and Partial Least Squares Regression, for the assessment of fractional shrub cover based on EnMAP imagery. The results again underpin the potential of EnMAP data for monitoring changes in ecosystem functioning.

The present thesis aims to quantify the composition of landscape components (PV, NPV and soil) to subsequently derive the C factor to assess soil erosion. A number of studies presently exist which have also aimed to quantify this parameter by remote sensing approaches (e.g. De Asis and Omasa (2007); Kefi et al. (2011); Förster et al. (2014b). Kefi et al. (2011) successfully estimated the C factor based on an existing vegetation index the Transformed Adjusted Vegetation Index (TSAVI). Other approaches have estimated the fractional cover of PV, NPV and bare soil by spectral unmixing techniques. De Asis and Omasa (2007) successfully applied spectral mixture analysis to evaluate the fractional cover NPV, PV and bare soil in different ecosystems. The inclusion of NPV contributed to mapping potential erosion in these areas. However, most of the studies have been conducted in arid and semi-arid ecosystems with only a few studies dealing with the potential of hyperspectral data to assess soil erosion parameters in the tropics. As such, there is a need to assess the capability of using simulated EnMAP imagery to supply quality information of soil erosion parameters at regional scales that are relevant for land management practices in tropical ecosystems.

1.3   Soil erosion by water: Parameters and drivers

The humid tropics are well known as important global carbon sinks and hotspots for biodiversity (Tomich et al., 2005). The rapid loss of rainforest is a significant environmental issue, as it negatively impact various ecosystem functions that result in land degradation. As healthy soils are fundamental for many ecosystem services (e.g water regulation and purification or biodiversity support) a solid understanding of soil degradation is necessary to maintain and support these services. In tropical ecosystems the loss of vegetation in conjunction with severe rainfall events common for these climate zone can result in more dramatic levels and rates of soil erosion (Labrière et al., 2015) compared to other climatic zones.

Where soil erosion occur several off- and on-side effects of the degradation processes can be observed. On-side effects are commonly attributed to changes in soil quality, structure, stability and texture, which can lead to the removal of entire layers of soil or organic matter or increase the susceptible of those areas to extreme conditions (e.g. drought) (Roose, 1996). In agricultural regions the loss of fertile top soils negatively affects crop or biomass yields forcing farmers to expand agricultural areas or increase the use of fertilizers. The off-side effects of soil erosion are coupled to the different on-side effects. For example, the eroded nutrients and sediments travel to areas at lower elevations causing several off-site impacts such as the sedimentation of reservoirs or riverbed filling. This can result in decreased water quality or an increased flood risk (Lal, 2003). Additionally, the increase used of pesticides or fertilizers in location where soil erosion occurs, can result in a contamination or pollution of downstream water sources by frequently being transported to lower elevations (Roose, 1996).

An analysis of the linkage between land use type and soil erosion reveals that the degree and type of vegetation cover from non-existing (e.g. bare soil), to ground cover (e.g. living and dead grass) to forests is of major importance in assessment of the potential for soil erosion. Of particular importance are the canopy, surface cover and below surface effects that vegetation has on limiting surface erosion (Wischmeier and Smith, 1978). Soils covered by dense vegetation are often characterized by higher occurrence of organic carbon, moisture and different soil structure/aggregates resulting in better growth conditions for other plants resulting in minimal soil erosion. The role of NPV on soil erosion has been widely analyzed, where occurrence of NPV can significantly reduce the potential for soil erosion owing to well developed root systems and protection from flowing water and wind through physical binding of soil particles (Gray and Sotir, 1996). Soil erosion potential is highest where soils are bare of PV or NPV. In such cases, soil type becomes an important factor as different soil types are less or more susceptible to erosion, for example clay rich soils tend to be more resistant to erosion (Mirsal, 2008).

1.4 Motivation

Imaging spectroscopy sensors, commonly referred to as hyperspectral sensors, are characterized by recording the Earth surface by nearly continuous narrow spectral bands that can capture absorption features related to biochemical and geochemical Earth surface characteristics. The importance of imaging spectroscopic data and techniques are considered essential for monitoring of ecosystem function and

value (Ustin et al., 2004) and have been used for various soil mapping and soil degradation studies owing to the ability to identify surface materials and quantify surface properties (Ben-Dor et al., 2009). For example, the physical based differentiation of NPV and exposed soil can be carried out due to specific absorption features of cellulose and lignin occurring in the SWIR region of the spectrum (Asner and Lobell, 2000; Daughtry et al., 2006; Delegido et al., 2015). Additionally, the SWIR region also provides useful information on soil types, in particular the occurrence of clay minerals (Mulder et al., 2011). The potential of these reflectance characteristics have been successfully used in several studies using airborne hyperspectral data (e.g. Förster et al. (2014a); Asner and Lobell (2000); Bayer (2012)) and spectral mixture analysis, which deconvolves individual pixel spectra into their constituent components, has been proven to be a useful tool to estimate subpixel information on vegetation cover and soil components (Okin et al., 2001; Ustin et al., 2004).

However, airborne campaigns are not suitable for long-term and large scale operational monitoring of degradation processes because they are expensive and time-consuming. The first experimental hyperspectral satellite, Hyperion, does provide high spectral resolution data and spatial coverage, and has be used in numerous studies related to vegetation and soils. However, the system has been plagued by issues related to reduced instrumental performance (Kruse et al., 2003; Roberts et al., 2003), and is nearing the end of its lifespan. Thus, moving forward the ability to study soil and landscape degradation processes such as soil erosion, new operational systems that can provide large spatial coverage with multi-temporal and high spectral resolution capabilities are necessary. Upcoming hyperspectral satellite systems such as EnMAP, HYSPIRI (Roberts et al., 2012), HISUI (Matsunaga et al., 2013) and others should be able to provide the required data on an operational basis.

2 | Assessment of Soil erosion parameters in Costa Rica using reflectance hyperspectral and simulated EnMAP imagery

Article Info

Authors information:

Sarah Sophia Malec

Global Change Ecology

University of Bayreuth

✉sasoma@posteo.de

Keywords:

Soil erosion

Costa Rica

Spectral Unmixing

HyMap

EnMAP

Abstract

Soil erosion can be linked to relative fractional cover of photosynthetic-active vegetation, non-photosynthetic-active vegetation and bare soil, and terrain characteristics such as the C and LS factors. This study investigates the capacity of EnMAP imagery to map areas prone to soil erosion in a region near San Jose, Costa Rica, characterized by spatially extensive coffee plantations and grazing in a mountainous terrain. Simulated EnMAP imagery is based on airborne hyperspectral HyMap data. Fractional cover estimates are derived using a Multiple End-member Spectral Mixture Analysis approach. The C and LS factors are derived from a digital elevation model. The analysis was done independently for EnMAP and HyMap, the latter treated as reference data as no ground truth is available. Results demonstrate that with EnMAP imagery it is possible to extract quality endmember classes with important spectral features related to PV, NPV and soil, and be able to estimate relative cover fractions. This spectral information is critical to separate soil and NPV which greatly can impact integrated modelling with terrain characteristics. From a regional perspective EnMAP can be used to highlight specific areas that may be prone to erosion where this information can be extract directly from the imagery using automated processes.

2.1   Introduction

As the primary aim of the upcoming German hyperspectral satellite mission Environmental Mapping and Analysis Program (EnMAP) is to monitor and characterize the evolution of various terrestrial and aquatic ecosystems on a global scale (Kaufmann et al., 2006), one such challenge is the assessment of land degradation. Thus, the thesis aims to quantify parameters related to soil erosion as one aspect of land degradation. At the beginning of the 21st century

about 19.65 percent of the terrestrial surface worldwide were defined as degraded land (Eswaran et al., 2001). According to the FAO 2014 the definition for soil degradation follows largely the definition of the term land degradation, which is defined as the reduction in the capacity of land to provide ecosystem goods and services securing its function for all beneficiaries of the land. Thus, soil degradation is a change in the soil health status resulting in a diminished capacity of the ecosystem to provide goods and services for its beneficiaries. One major aspect of soil degradation is the loss of top soils due to erosional processes, causing multiple environmental and socio-economical consequences. In areas where erosion takes place soil loss lowers the soil fertility and modifies natural landscape configurations. This negatively effects crop yields forcing farmers to expand agricultural areas or increase the use of fertilizers. The eroded nutrients and sediments traveling to areas with lower elevations causing several off-site effects, like the sedimentation of reservoirs or riverbed filling resulting in decreasing water quality or an increased flood risk (Lal, 2003; Millennium Ecosystem Assessment, 2005). Labrière et al. (2015) underpinned that the type of land use itself does not automatically allow inferences about the erosion potential of a site, more importantly land uses or anthropogenic activities creating bare soil patches in the landscape should receive greater attention. Considering these findings and with the purpose to utilize hyperspectral remote sensing as an instrument for understanding the actual and potential spatial extent of soil degradation on large spatial scales, without the need of cost and labor intense field surveys, this study determines soil erosion parameters, as a crucial factor of soil degradation, in the Province of San Jose, Costa Rica.

The prevailing climate regime, which causes severe rainfall events, agriculture, cultivation of coffee and the topographical conditions in mountainous regions determine a high risk of soil erosion in some areas. Hyperspectral data has the potential to identify parameters indicating soil erosion, as specific absorption features allow the performance of an sub-pixel classification with respect to the classes PV, NPV and bare soil. By an detailed comparison of the results for HyMap data and simulated EnMAP data new insights will be given into the capacities of the upcoming EnMAP satellite for environmental studies dealing with spectral unmixing of PV, NPV and soil. The analysis of the the different endmembers extracted for both data sets as well as the spatial use of these EMs for image unmixing will help to understand new advantages and challenges for future EnMAP applications from mono- and ,more importantly, multitemporal imageries. Future studies may apply the methodology and more importantly show its validity and ability to transfer it from the regional to the national

scale. The methodological approach used in this study comprises two main parts and is applied independently on the two hyperspectral data sets. The first is the derivation of fractional cover estimates for PV, NPV and soil from the hyperspectral imagery. The second is the integrated soil erosion modelling which combines the derived fractional cover maps with landform information. An assessment of the EnMAP results is given by comparing the generated outputs with those derived from the reference HyMAP imagery as no ground truth data is available for this study area.

2.2 Study area

The study was mainly conducted in the province of San Jose and partially in the provinces Heredia and Alajuela located on the western side of the country in the Central Pacific (Fig. 1, upper left corner 10°2′, 59.71″ N 84°9′, 6.7″ W and lower right corner 9°46′, 51.96″ N 84°25′, 38.82″ W). The 900 km^2 area is delimited by the extent of the simulated EnMAP imagery and was chosen because of knowledge based information about areas of extensive coffee plantations prone to soil erosion. The largest urban area in the study area is San Jose, located in the north-east part of the imagery, but several smaller villages and settlements exist, including Santiago de Puriscal and Mercedes. The strong anthropogenic influences and mountainous conditions causing an highly fragmented landscape characterized by urban and agricultural areas, infrastructures and forested areas including two privately owned areas protected under the "payments for ecosystem services" (PES) programme (Porras et al., 2013). Coffee plantations span over the whole study site with relatively dense occurrences in areas with moderate to high elevations.

Several rivers cross the study area, with the largest river Tarcoles emptying into the pacific ocean after the mainstreams Rosales, Ciruelas, Jaris and Picagrces confluence and become Tarcoles in the west part of the imagery. The elevations of the study site range from 200 to 2200 m above sea level with the steepest slopes (54°) occurring in the southeast in the Cerro Cedral, the second highest mountain of the Cerros de Escazú.

The climate is dominated by a tropical wet and dry climate (Köppen-Geiger climate classification) with a pronounced dry season in the low-sun months between December and April and a rainy season from May to November in the high-sun months. Due to the mountainous relief and the large extent of the study area the prevailing microclimate causes local deviations in the presented climate classification. This is mirrored in the bioclimatic classification after the

Holdridge life zones system, which divides the site into 11 biomes (Appendix: Fig. A) (Holdridge, 1967). The total annual precipitation within the site averages between 2000 and 4000 mm with mean annual temperature around 20.5 ° Celsius at the low-lands, but ranges between 12-18° Celsius in mountainous areas.

The soil in the study area can be divided into 5 major groups. The major soil type in the south of the region around the Cerro Cedral and westwards are Ultisols, also known as red clay soils, which are defined as mineral soils without calcareous material, less then 10% weatherable minerals in the A-horizon, with a clayey B horizon and less that 50% base saturation (Soil Survey Staff, 1999). These soils are acidic and leached, thus, they are more suitable for tree crop plantations than for the cultivation of food crops (FAO, 1992). Along the main river valley of the study site the dominating soil type are Entisols, soils that show no development of an pedogenic horizon and comprise only an A-horizon. These soils are characteristic for landscapes where the soil material is not in place long enough for pedogenic processes to form distinctive horizons (Soil Survey Staff, 1999), which is the case in the river valleys where sediment depositions occur in frequent intervals followed by being washed out in the rainy season. The north of the study site is dominated by two major soil types, namely Alfisols and Inceptisols. Alfisols are characterized by a clay-enriched subhorizon. In contrast to Ultisols, these soils show relatively high natural fertility and are considered as a relatively important soil type regarding food and fiber productions (Soil Survey Staff, 1999). In the Central Valley typically volcanic Alfisols feature a high content of volcanic glass and/or amorphous or poorly crystalline iron and aluminum minerals and high content of soil organic matter, resulting in loose soils, which are good for cultivation and relatively resistant to water erosion on moderate slopes (Tiessen et al., 2010). Inceptisols are more developed that Entisols, but they still show only a minimal horizon development, not sufficient to be characterized as a different soil order. In the southern part of the study site Vertisols occur as patches in two areas. Vertisols are characterized as clayey soils, that develop deep cracks during parts of the year owing to dry and wet conditions causing swelling and shrinking. These changing soil properties makes this soil type relatively inappropriate for commercial agricultural use as wet and dry seasons require special cultivation practices (Soil Survey Staff, 1999). The variations in landscape configurations, climate and soil types have allowed the development of a high diversity of habitats and species richness, not only at the study site, also throughout the country.

2.3   Data

2.3.1   HyMap data

For the given study a hyperspectral dataset acquired by an HyMAP II sensor, which was provided by the Centre for Earth Observation, University of Alberta, Canada, was used. The total dataset consist of 232 hyperspectral flight lines acquired during a flight period between 1st March and 6th April 2005 covering approximately 85% of the country. HyMap comprises 125 contiguous spectral bands across the reflective solar wavelength region from 0.45- 2.5 μm. For this survey the spatial resolution varies between 15.5 and 16.9 m due to the high variability in topography across the country. The data was originally georeferenced and atmospherically corrected by HyVista Corporation. With the purpose to get an impression of data distribution and to facilitate the data handling kmz-files were generated for all flight lines. The kmz-files where created using R-G-B Look up images and the corresponding GLT-file also delivered by HyVista Corp.

Preliminary analysis of the data showed processing artefacts and inconsistencies between the flight lines which couldn't be explained without knowledge about the processing algorithms applied to it. Therefor the whole data set was reprocessed including data quality assessment (Bachmann et al., 2007), spectral wavelength calibration (Richter and Schläpfer, 2015), georectification (Krauss et al., 2013), rugged terrain atmospheric correction (Richter and Schläpfer, 2015), BRDF correction (Schläpfer et al., 2015), as well as the generation of an image mosaic
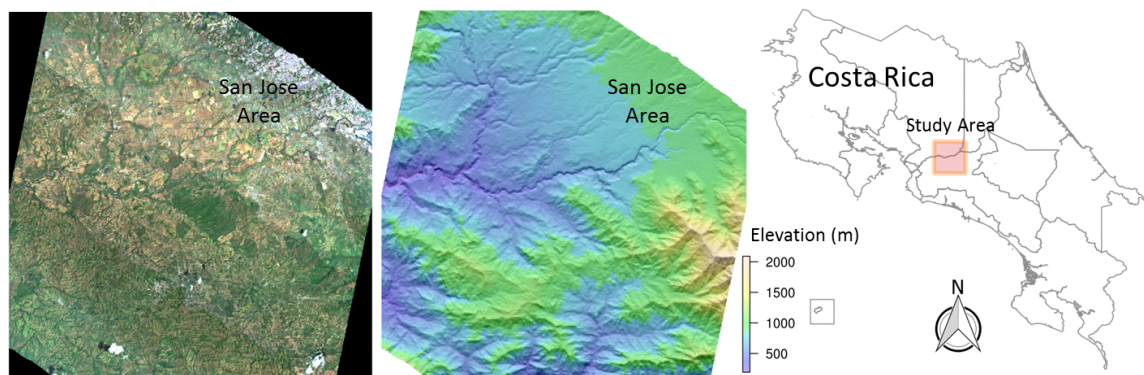

Figure 1: *HyMap RGB true color mosaic of study area (30 x 30 km) located in central Costa Rica near San Jose (left), associated DEM (middle) and location within the country (right).*

for the study area (Fig. 1) following the standardized processing chain developed by the German Aerospace Center (DLR). The processing chain is integrated into the FP7 EUFAR (European Fleet for Airborne Research) project and a detailed overview about the processing steps is given in Figure 13 in the chapter "Supplementary Material". The data delivered by HyVista Corp. was reprocessed with respect to the processing levels 1 (radiometric corrected), 2geo (orthorectified) and 2atm (atmospheric corrected) based on the raw radiance and GPS data. The spatial resolution of the reprocessed data was set to a uniform 15 m pixel size. For further analysis 16 spectral bands in the main water absorption regions near 1.4. and 1.9 μm were removed resulting in 109 bands between 0.454-1.310, 1.504-1.785 and 2.010-2.445 μm. All of the processing steps used were done to ensure the generation of an high quality reflectance mosaic, which subsequently served as input to generate a high quality EnMAP simulation.

2.3.2   EnMAP simulation

The EnMAP satellite will cover the spectral range of 420-2450 nm with 242 bands with an width of each band varying between 6.5 - 10 nm. The ground sampling distance will be 30 m with a swath width of 30 km. The purpose of the EnMAP End-to-End Simulation tool (EeteS) developed at GFZ is to provide an tool for accurately simulating EnMAP imagery, including calibration (Rogge et al., 2014). Eetes is further used to support the overall system design, to optimize the instrument parameters and to develop as well as evaluate the prepocessing chain and algorithms serving scientific applications (Segl et al., 2012; Guanter et al., 2009; Segl et al., 2010). The high quality reflectance mosaic generated from the HyMap data is a prerequisite for the simulation as it served as input to EeteS to generate the EnMAP-like imagery. EeteS first transformed the HyMap data into raw EnMAP data by predicting atmospheric conditions for the default value July, 20th and accounting for spatial, spectral, and radiometric differences between the two sensors. Subsequently the EnMAP image processing chain is applied, including onboard calibration, atmospheric correction and orthorectification to resulting in an geocoded at-surface reflectance image (Appendix, Fig. B). As the simulation follows the prospective orbital path of EnMAP the 30 x 30 km scene is cropped on the west and east sides and the cropped areas in the north and south are caused by the size of the input data (Fig. 1). The final EnMAP simulation used for subsequent analysis was reduced by 51 bands based on the main water absorption regions and within the overlapping region of EnMAPs two spectrometers between 0.905-0.980 μm resulting in 191 bands between 0.450-1.311,

1.499-1.783 and 2.007-2.439 μm that match the spectral ranges used with HyMap.

2.3.3 Digital Elevation Model and auxiliary data

The auxiliary data consists of several thematic vector layers covering the entire country of Costa Rica giving information about climate, vegetation, soil types, land cover, conservation status or the distribution of coffee plantations. The data set was mainly used to provide an detailed description of the study region and allow a descriptive analysis of relationships between the distribution of areas prone to soil erosion and the cultivation of coffee. The digital elevation model (DEM) covers with a spatial resolution of 90 m whole of Costa Rica and it was used to extract geomorphological information relevant for soil erosion parameter assessment. Both data sets were provided by the Centre for Earth Observation of the University Alberta,Canada.

2.3.4 Global Urban Footprint

The Global Urban Footprint (GUF) for the entire country of Costa Rica was made available by the DLR and it has a spatial resolution of 12 m. The GUF project maps settlements and urban areas by the application of an fully automated processing chain from TerraSAR-X and TanDEM-X radar images acquired between 2010 and 2013 on a global scale (Esch et al., 2014). In this study the GUF was used to masked urban areas from the HyMap and EnMAP data as the analysis focuses on the assessment of soil erosion potential in natural ecosystems and those under cultivation. As the temporal resolution of the GUF mask and the HyMap data differs, slight overestimation of urban areas are expected, but these are expected to not influence the analysis of this study.

2.4 Methods

2.4.1 Spectral Unmixing

Hyperspectral remote sensing extracts information from the earth surface, using the radiance measured by airborne or space borne sensors. These spectroscopes detect the radiance with continuous narrow bands at varying pixel sizes (e.g. between 1 and 30 m resolution). Each pixel measured by the sensor detects the spectral signature of the different materials occurring in the pixel. Thus, the resulting spectral signature of the pixel is a combination of several pure spectral

signature. Finding the pure spectral signature of the image components, the image endmembers, is crucial for the identification of subpixel constitutents by spectral unmixing. Spectral Mixture Analysis approaches aiming to utilize the physical features of the pixel components represented by the endmembers to model the components and their fractional abundances based on an inverse operation to yield the spectrum measured at the sensor. In general, the model used is either a linear or a non-linear unmixing algorithm. The latter is applied when the material in a pixel is heterogeneous distributed, leading to physical interactions of scattered light with the different materials, resulting in multiple scattering effects. In literature two types of interaction are distinguished, the classical and microscopic (intimate) interaction level (Bioucas-Dias et al., 2012). If the light detected by the sensor is first scattered from one or more objects and then reflected from another object, this case is called a classical interaction and it is usually solved by bi-linear model (Bioucas-Dias et al., 2012). The second interaction case assumes that a photon is emitted from one material and absorbed by another material leading to the higher emissions of the latter material, than this material would usually show (Hapke, 1993). In this case, different approaches aiming to model the radiance pathway exist. For detailed information dealing with non-linear unmixing algorithms please see Hapke (1993). In this study a linear mixture model was used, which assumes that a photon interacts with only one material. Despite this assumption, it has been shown in many geological and ecological applications, that the approximation of the light scattering mechanisms is valid (van der Meer et al., 2012; Bioucas-Dias et al., 2012; Bachmann, 2007). A linear unmixing model occurs when the incoming radiation interacts just with one material, thus the spectral mixing takes place within the sensor (Bachmann, 2007; Bioucas-Dias et al., 2012).

Based on this assumption the associated model calculated per image pixel can be expressed as the summation of the individual material reflectance functions ρ j,e multiplied by their surface fraction fe within the effective ground-projected instantaneous field of view resulting in the sensed reflectance value ρ:

$$\rho j = \rho j, e * fe + \delta j \tag{2.1}$$

where δ j denotes the residual error term. Furthermore, all fractions should be greater or equal than zero, and smaller or equal to one, plus should sum to one (Bachmann, 2007). This constrained linear equation system is then solved in a least-squares fashion using either singular value decomposition or the pseudoinverse.

This shows, the necessary input parameters to spectral unmixing are the

material reflectance functions ρ j,e of all components of a pixel. Thus, the collection of these endmembers is the first step before deconvolving the pixel spectra into the fractional abundances of endmembers.

In nature it is unlikely that one green, dry or soil component can adequately represent the spectral variability of its ground cover class, because there is usually a high variability in plant species and soil types occurring in the area of interest. (Bachmann, 2007). Therefore, the multiple endmember spectral mixture analysis (MESMA) approach was developed where the main concept is to optimize the endmember set for each pixel independently based on the inter and intra class variability contained in a multiple endmember set. One way to deal with the variety of possible scene components and their mixture within one pixel is to include all possible endmembers in one mixture model. Disadvantages of this method are that the problem with linear dependent endmembers is aggravated, resulting in wrong abundances (Bachmann, 2007). Moreover, even though the linear mixture model can be solved as long as the number of endmembers is less than the number of bands plus one, the "effective" dimensionality of hyperspectral data is smaller due to the high degree of correlation between bands, preventing this method (Bachmann, 2007). A far better approach is to optimize the endmember set for each pixel independently. This was recently done for semi-arid and arid environments by Mulder et al. (2011) for the mapping of chaparral, Bayer (2012) who applied a Monte Carlo unmixing for the SWIR2 region based on endmember "bundles" for the mapping of desertification and Förster et al. (2014b) who used MESMA derived cover fractions and vegetation height obtained from LiDAR data to derive bi-temporal C factor maps.

Spatial Spectral Endmember Selection

The collection of spectral endmembers aims to find the most spectrally pure, or spectrally unique, pixels within the hyperspectral data set. Two ways for the collection of endmembers exist, where one can either be purely based on the reflectace spectrum of a pixel component collected in the field or laboratory, or on image based approaches (Veganzones and Graña, 2008; Rogge et al., 2007). The latter option is seen to be more adequate when working in heterogeneous landscapes. When endmembers are derived from an image the spectral signature is recorded under the same lightning conditions as the image, whereas spectral signatures of a material derived in the laboratory are taken under optimized conditions, which are not representing the physical conditions of the image acquisition (Rogge et al., 2007). The derivation of the endmembers from the

image can be addressed by different methods. The most common methods rely on the geometric properties of hyperspectral data in N-Dimensional space. Such approaches use the fact that the convex hull of pure endmembers in the data forms the probability simplex, where endmembers represents the vertices of the simplex. Most classical approaches assume at least one pure pixel for every image component (Akhtar et al., 2015). The most popular example is the pixel purity index (Boardman, 1993; Boardman et al., 1995). The existence of pure spectra in the data is also a condition for other automated methods such as the Sequential Maximum Angle Convex Cone (SMACC) (Gruninger et al., 2004) or the Spatial Spectral Endmember extraction Tool (Rogge et al., 2007).

As field data for the given cover types is not available for the study region and would commonly not be available when doing large regional surveys, this study uses the updated spatial spectral endmember extraction tool (SSEE) to directly extract the image endmembers from the mosaic (Rogge et al., 2012, 2007). The updated SSEE method was specially adapted to work with large multi-flight line datasets. In contrast to methods extracting the endmembers independent of spatial information SSEE extracts endmembers that are spectrally similar, but spatially independent to allow a good representation of the intra-class variability as well as the integration of expert knowledge in the selection process to select a final endmember set appropriate for a specific application. In a first step SSEE divides the radiance image into equal sized non-overlapping subset regions for which a set of eigenvectors is calculated that explain the majority of spectral variance for each subset via singular value decomposition. This way the spectral contrast within specific spatial neighborhoods can be addressed (Rogge et al., 2007). The foundation is that the local eigenvectors can be applied at the full scale of the image to improve the extraction of globally low contrast endmembers that are commonly missed using purely spectral based methods. Subsequently the image data are projected onto the local eigenvectors compiled from all subset regions and pixels with extreme eigenvectors are chosen as endmembers. As a result SSEE extracts a fairly large set of candidate endmembers containing numerous endmembers which are spetrally similar, but spatially independent. After carefully organizing the resulting endmembers into four broad main classes PV, NPV, soil and others (e.g. urban or artificial materials, water), the final selection of endmembers was automatically performed based on statistical measures.

These measures are the Class Average RMSE (CAR) (Dennison and Roberts, 2003b) and the Endmember Average RMSE (EAR) (Dennison and Roberts, 2003a). Both measures have been suggested by Dennison and Roberts (2003a,b) for assessing suitable endmembers for MESMA approaches and are based on the

assumption that a suitable endmember can be modelled with an low RMSE by a linear combination of an endmember of the same class plus a shade component. As EAR is calculated as the average RSME produced by a spectrum when it is used to model all other members of the same class low EAR scores are representing spectra having the lowest average RMSE. CAR is an important indicator if an endmember decreases the separability of two classes (Bachmann, 2007). The measure calculates the average RMSE produced by a each endmember of a class when it is used to modelle a different endmember class. Thus, endmember with low RMSE indicate low seperability between endmember classes (Bachmann, 2007).

μ**MESMA**

The resulting multi-representative class spectra are input into the μMultiple Endmember Spectral Mixture Analysis (μMESMA) algorithm of (Bachmann et al., 2005; Bachmann, 2007) to estimate the cover fractions of PV, NPV and bare soil across the study site for the HyMap and EnMAP data. The program implements the general MESMA approach (Roberts et al., 1998) and was developed for the quantification of fractional ground coverage in semi-arid environments (Bachmann, 2007; Bachmann et al., 2009), and also applied in context of soil and mineral mapping (Zabcic et al., 2014; Bayer, 2012). The program requires the input of unique spectral libraries, were each library represents one of the three classes and consists of multiple spectra accounting for intra-class spectral variability across the imagery.

μMESMA is performed by linearly unmixing each image spectrum using all possible combination of all spectra within the libraries plus an optional shade component, with one spectrum of each endmember group being used at a time. The optimized combination and fraction of endmembers for pixel unmixing is defined by the lowest root mean squared error (RMSE), predefined abundance constraints and by the consideration of class specific absorption features of the spectra (e.g. chlorophyll absorption, clay-OH-absorption, or ligno-cellulose). The approach first performs an Iterative Error Analysis (IEA), where after a first unmixing iteration pixel with very high RMSE are identified and tested if they represent spectra not included in the input libraries (Bachmann et al., 2005). This way the user can enhance the predefined libraries to ensure that all pure spectra representing the components of interest are included, and thus increase the accuracy of unmixing results. For this study, the final input libraries extracted by SSEE remained untouched and were used for a second unmixing iteration. Afterwards the measured spectrum and the residual (i.e. the difference between

the measured and the modelled spectra of a pixel) are checked for class specific absorption features (Bachmann et al., 2005) to identify characteristic divergences. The underlying principle of this step is that the mixture model for a pixel is adjusted if an underestimation of clay in an clay-rich soil is observed. By checking certain wavelength regions featuring the absorption of clay minerals (around 2300 nm) divergences between the modelled and the measured spectrum are minimized by using a soil spectrum with higher clay content (Bachmann et al., 2005). Additionally, μMESMA accounts for the effects of changes in the overall albedo of the soil endmembers to avoid changes in the soil endmembers between adjacent pixels. In the nature small-scale scattering (pixel level) of different soil types is rarely observed. Therefore, μMESMA also allows an slight increase of the error score if a pixel can be modelled with the dominant soil endmember in the neighborhood to reduce the occurrence of strong spatial fragmentation of endmember abundances. Furthermore, μMESMA extents the three class endmemberss by a shade component to model the fractional cover. The approximation of the shade component is based on a methodology proposed by Adler-Golden et al. (2001) which is calculated considering the average image reflectance, atmospheric composition as well as illumination- and viewing geometries. These parameters are derived from the atmospheric correction, flight position data as well as from the image itself using ATCOR. For a detailed description of the methodology see Bachmann (2007). Bachmann (2007) has shown by using using synthetic test scenes and field measurements as reference, that the typical mean error in ground cover fractions was found to be 12%, with Pearson correlation coefficients around 0.8 (R^2 values, significance p < 0.005), thereby being comparable to other MESMA approaches.

In the context of this study, a first iteration of μMESMA was carried out using the SSEE derived endmember classes for PV, NPV and soils. Subsequently μMESMA proposed spectra of all pixels having a high modeling error as possible additional endmembers, these were then interactively checked, but no spectrum was included in the final library defined by SSEE. In addition to the endmembers for PV, NPV and soil, a shade component was included. The final abundances where then adjusted for the shade component, and rescaled accordingly.

2.4.2   Soil erosion parameters and integrated modelling

The assessment of soil erosion commonly utilizes predictive models such as the Universal Soil Loss Equation (USLE) or the Revised Soil Loss Equation (RUSLE) (Mitra et al., 1998; Bakimchandra, 2011). The USLE was originally developed at

the United States Department of Agriculture (USDA) based on numerous soil erosion studies as a tool for conservation and first published by (Wischmeier and Smith, 1968). Nearly 30 years later the model was updated with the purpose to transfer USEL into a physically-based model with improved methodologies for parameter estimations and to computerize the predecessor model (Renard et al., 1997, 2011). Until today the model is widely used to predict annual soil loss caused by sheet and rill erosion and modified for specific application purposes (Renard et al., 1991).As the inertial application purpose of USEL was to provide conservationists with a tool supporting decisions on sustainable cropping or management strategies its applications are limited to long-term estimations on annual soil loss and, thus erosional processes linked to specific rain events or on specific time steps can rarely be conducted. Also, the application of the model in tropical regions is seldomly used compared to semi-arid or temperate biomes, due to rather poor parameter calibration for these regions caused by leaking physical data (e.g. land use, soil erodibility information). The final outcome from USLE the soil loss per unit area A is calculated as:

$$A = R * K * L * S * C * P \qquad (2.2)$$

where R is defined as the rainfall-erosivity index, a statistically derived index based on the sum of the annual rainfall energy in every storm event multiplied by the maximum rainfall intensity within 30-minute intervals. The R-factor quantifies the potential of rain-drops or splashes to erode soil particles and its strong correlation with rainfall events explains its geographical dependency. The soil erodibility factor K defines soil properties related to the susceptibility of a soil type to be eroded by rain-splash or runoff, like the cohesivity or bulk-density of the soil. The L and S parameters together are defined as the topographic factor LS a product of slope length and slope steepness. The intensity of management and land cover on soil loss is expressed by the crop and management factor C. This parameter is estimated based on a ratio between erosion rates on bare soil and under different cropping systems. The consideration of plant cover, type of tillage management and organic matter production makes the derivation of this factor rather challenging and to some extend subjective. Similar to the C-factor, the conservation practice factor accounts for management practices considering contouring, stripe cropping or terracing (Wischmeier and Smith, 1978).

Other approaches aiming to quantify soil loss, also make use of the RUSEL parameters in order to apply machine learning algorithms, e.g. fuzzy logic or neuronal networks (Bakimchandra, 2011; Kim and Gilley, 2008; Mitra et al., 1998).

These approaches tend to use the concept of USLE and investigate parameterization with respect to expert knowledge, case studies and from literature to then develop a set of constraints and rules ranking the erosion potential of the area (Bakimchandra, 2011). Similar considerations and the successful application of RUSEL in tropical ecosystem, as nicely reviewed by Labrière et al. (2015), reinforce the decision to focus on the capabilities of hyperspectral data to derive the C factor and to combine the fractional cover information with topographic feature, expressed by the LS factor, to estimate areas potentially prone to soil erosion. In the literature in-depth analysis of the relationships between slope and erosion risk have been conducted and the results show the validity for rating specific slope ranges for the quantification of potential soil erosion severity(e.g. Bakimchandra (2011)). C-factor derivation followed the methodology proposed by De Asis and Omasa (2007):

$$C = fC_{soil} / (1 + fC_{PV} + fC_{NPV})$$ (2.3)

where fractional cover of soil (fC_{soil}) is divided by the degree of vegetative surface cover defined as the sum of abundance values received for PV and NPV (fC_{PV} + fC_{NPV}). The factor is normalized by the addition of 1 in the denominator. Values close to 0 indicate areas with very high abundances of vegetation cover and high values represent areas with exposed soils. The LS factor was derived using the watershed hydrological model implemented in GRASS GIS which uses for parameter calculation the algorithms introduced in Weltz et al. (1987). Even though the calculation of the LS factor is controversy discussed in literature and algorithms have been developed for slope length calculation, this study did not further investigate these approaches (Hickey et al., 1994; Desmet and Govers, 1996). In GRASS GIS the LS factor at a point r=(x,y) on a hillslope is computed using the algorithm introduced by Mitasova et al. (1996) which is based on the unit stream theory and can be expressed as:

$$LS(r) = [A(r)/a0]^m * [sinb(r)/b0]^n * (m + 1)$$ (2.4)

The slope steepness (S) is defined as the ratio of sol loss from the field slope gradient to soil loss from slope under a certain slope with otherwise identical conditions. The slope length (L) is generally refereed to as "the horizontal distance from the origin of overland flow to the point where either 1) the slope gradient decreases enough that deposit begins or 2) runoff becomes concentrated in a defined channel" (Renard et al., 1997). In equation 2.4 the first term of the formula represents L where A[m] is the upslope contributing area per unit contour width

and a0 the standard slope length (22.1 m = 72.6 ft) empirically derived based on the calibration of RUSLE. The exponents "m" and "n" are constants that are also empirical derived as part of RUSEL calibration, but they can be interdependently calibrated if the data are available for a specific prevailing type of flow and soil conditions. The second term describes S where b is the slope in degree and b0 have been defined based on the standard USEL plot (integrated in RUSEL) as 9% = 5.16°.

As an integrated erosion modeling approach the LS factor is integrated into equation 2.3 by multiplying the C factor with the LS parameter. Both factors are normalized to have equal weight. Two erosion potential maps were produced. The first case is base on equation 2.3 where PV and NPV are combined in the denominator. In the second case NPV is combined with soil in the numerator. The rational for these two cases are to highlight the importance of being able to distinguish and separate fractional cover for soil and NPV, which is possible with hyperspectral data, but more problematic with multispectral data.

2.4.3   Evaluation

As noted the objective of this work is to evaluate the capabilities of EnMAP for deriving essential landcover information and in combination with landform information to estimate the potential of soil erosion for a given region. As ground data is not available the capabilities of EnMAP are assessed by comparing the independent results with the reference data (HyMap). The assessment will first consider the endmembers derived from SSEE for both data sets. After, the class distributions between HyMap and EnMAP are compared and mixing relationships of each class between the sensors are visualized as scatter plots. Subsequently, besides a visual qualitative comparison of the fractional abundance maps for HyMap and EnMAP data, correlations between the abundance values have been tested by carrying out linear regression in RStudio. The fractional cover values of the reference data served as predictor variable to estimate the fractional cover values of EnMAP. To identify areas with significant differences between the modeled abundance values the methodology of Januchowski et al. (2010), implemented as SigDiff function in the R package SDMTools, was used. The significance of the difference between modelled fractional covers is pixel-wise calculated as the probability of the difference detected in each cell relative to the mean and variance of all pixel-specific differences. The comparison is carried out class-wise. Furthermore, qualitative and quantitative comparisons of spatial and statistical distributions of the fractional estimates for PV, NPV and

soil with respect to the integrated erosion modelling output was conducted by the aforementioned methodologies.

2.5   Results

2.5.1   Endmember detection

For the implementation of SSEE the key input parameters were used: window size = 24 pixels; SVD threshold = 99% of variance. SSEE delivered in total 700 candidate endmember pixel for the HyMap image that represent 0.0175 percent of all non-zero pixels within the reflectance image. 587 pixels represented urban/artificial surface components, water bodies, clouds and 1 endmember even smoke, 49 endmembers where detected for PV, 51 for NPV and the smalles amount, 13 pixels represent exposed soils endmembers. The final libraries, reduced by the statistical measures ERA and CAR, consist of 18 spectra for the class PV, 17 for NPV as well as 10 spectra for soil (Fig. 2). Furthermore, the 18 PV spectra could have been grouped based on the shape of the specra into six sub classes, for NPV 4 classes have been identified and the soil endmembers indicate the occurrence of 5 differing soil types. For the purposes of clarity figure 3 shows the mean spectra for each sub class. In the case of EnMAP, having a coarser spatial, but spectrally finer resolution than HyMap, urban or artificial image components are represented with 260 candidate endmembers , PV with 110, NPV with 105 and soil with 13 endmembers each, in total making up 0.0488 percent of all non-zero pixels within the simulated image. After finalizing the endmember libraries PV is supported with 20 spectra and NPV with 16 spectra each class divided into 6 sub classes, the 13 soil endmembers identified only three different soil sub classes (Fig. 2). Only for soil the number of endmember sub classes that could have been distinguished is less than for HyMap.

Looking at the soil sub classes from both HyMap and EnMAP one can highlight a number of key features related to the known soils in the area that distinguish them from NPV. Most notably is the diagnostic narrow Al-OH clay feature seen at 2200 nm in a number of the soil sub classes. The depth of this absorption features does vary for the sub classes indicating the relative abundance of clay minerals is variable in the soils detected. Additionally, a broad absorption features appears in the visible portion of the spectra, which can be attributed to iron oxide and iron hydroxide minerals such as goethite (FeOOH) and hematite (Fe_2O_3), which is consistent with the oxidized red coloured soils in the region. One sub class does show a clay feature, but has very low reflectance in the visible portion of the
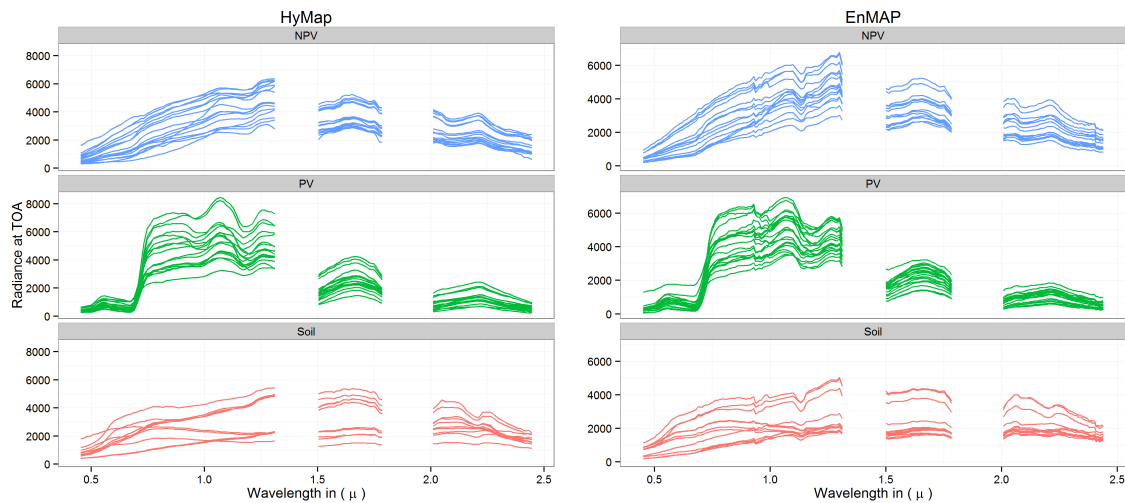
Figure 2: *Endmember libraries defined based on the endmembers detected by the SSEE tool and statistically reduced.*

spectra, which may indicate a darker organic rich top soil. The two additional sub classes detected by HyMap appear to show similar features, but in different magnitudes. With respect to the NPV class spectra the diagnostic lignin feature found at approximately 2105 nm that is useful to distinguish from PV in the SWIR does show up. A detailed assessment of the endmembers is not given here, but the results do indicate that even at the 30 m pixel resolution of EnMAP it is possible to extract relatively pure soil spectra with characteristics that are expected for the soils in the area, and also, effectively discriminate NPV and PV.

2.5.2 μ**MESMA**

Visual comparison of μMESMA results (Fig. 4) for HyMap and EnMAP show high overall similarities of fractional abundance pattern for all classes, nevertheless results from HyMap display a more local-scale spatial distribution of abundances. This result is expected as one pixel of EnMAP is covered by four HyMap pixels. For a more detailed analysis of the differences of the fractional abundance values an more quantitative assessment is conducted. Looking at the class distributions for both sensor systems the histogram of soil class show the best match between both sensors, while the fractional cover distribution of PV exhibits the greatest deviation between HyMap and EnMAP (Fig. 5). For the soil class EnMAP values show a slight unidirectional offset towards higher fractional abundances compared
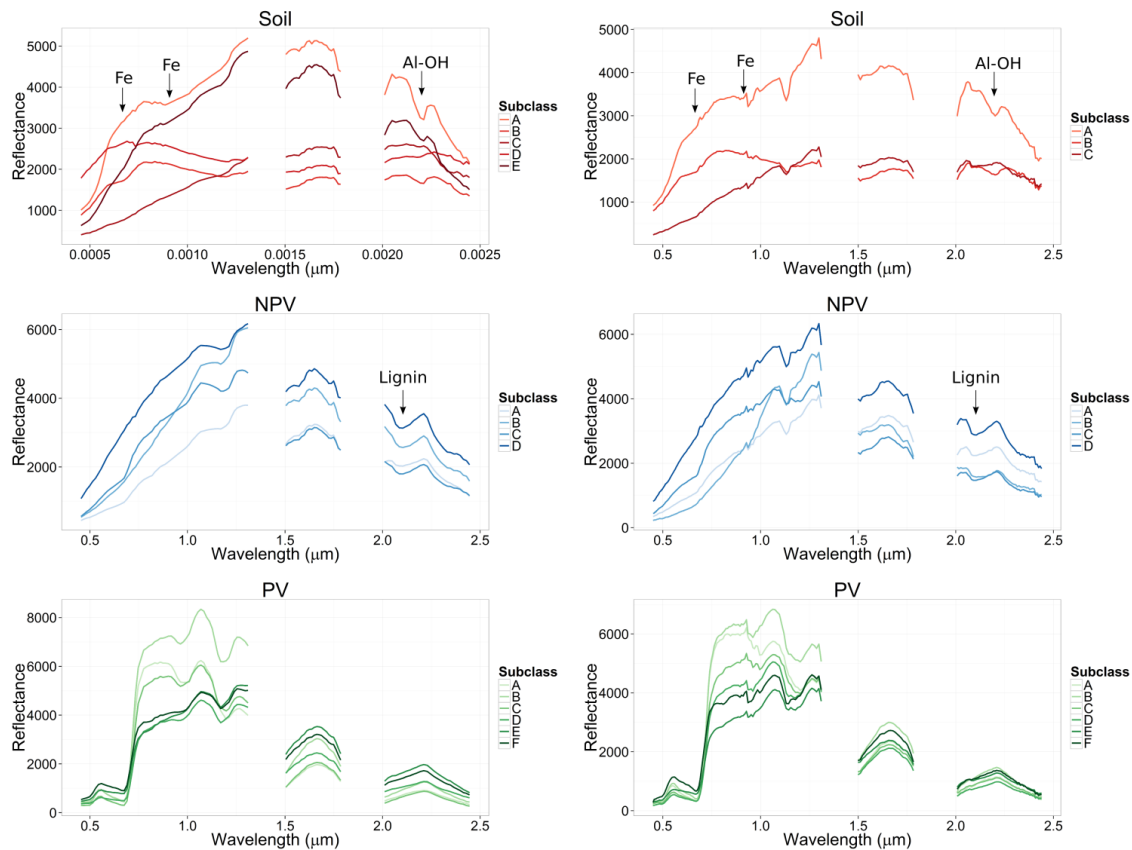
Figure 3: *PV, NPV and soil classes and sub classes (for clarity only the mean spectra per sub class are shown) derived independently from the extracted SSEE candidate endmembers for the original HyMap (left column) and EnMAP (right column) imagery. Arrows highlight key spectral features related to soils (Fe, Al-OH) and NPV (lignin).*
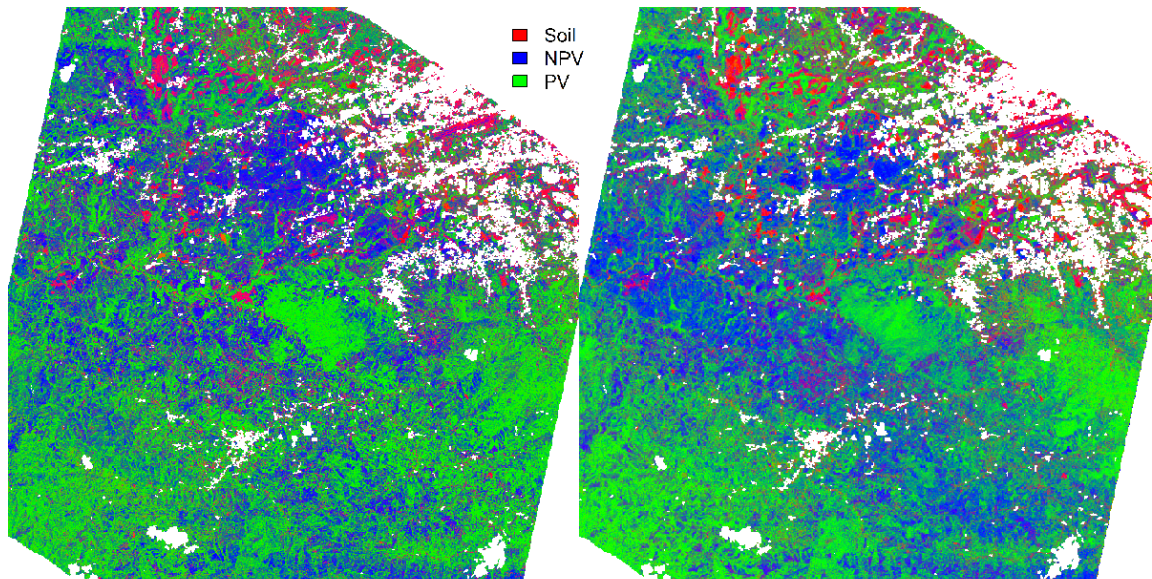
Figure 4: *Spectral unmixing results from MESMA displayed as 3 band RGB composite of abundances for HyMap (left) and EnMAP (right). White areas are masked based on the GUF and cloud mask.*

to the fractional abundances measured by HyMap. In Case of NPV, considering HyMap values as reference, low to moderate abundance values (25% - 50%) are overestimated in the EnMAP imagery, while the opposite case is observed for high abundances (70% - 100%). Fractional cover values of PV for EnMAP show a negative skewed distribution in contrast to HyMap where the abundance values are almost uniform distributed with an slight increase in number of pixels for high abundance values (90% - 97%). These pattern can also be visually observed in the fractional abundance maps for the single classes (Appendix, Fig. C). This increase above 90% may be attributed to higher sub pixel mixing in the 30 m EnMAP pixels. For example, small clusters of trees or even individual trees surrounded by NPV and/or soil may be common in some areas. These may be detectable at the 15 m HyMap data, but mixed within the EnMAP data. However, the cause of this anomalous feature has yet to be determined.

Next, the abundance maps of HyMap and EnMAP are class-wise compared on pixel level by plotting the pixels of the classes against each other as density scatterplots. The mixing relationship show a similar pattern among the data sets (Fig. 6). As also seen in the visual comparision the high density areas for the various mixtures between classes for EnMAP are more spread out compared with
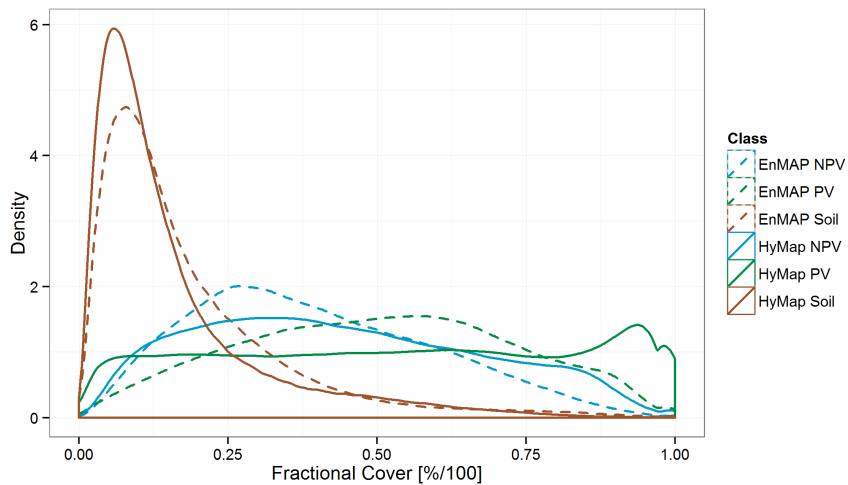
Figure 5: *MESMA fractional cover histograms for HyMap and EnMAP imagery.*

HyMap. Low fractional cover values for soil are associated with fractional cover values of PV and NPV covering the whole abundance range. Higher soil fractions (50-60% HyMap; 40-80% EnMAP) do show up as a thin slightly more dense areas with very low PV (0-10%) fractions. These represent agricultural fields with limited vegetative growth. Mixing between PV and NPV is strongly negatively correlated for both data sets. For HyMap there is a small high density area associated with higher (40-50%) NPV with very low PV fractions. Additionally, we also see a small high density area of high NPV fractions (0-90%) with low PV fractions (0-10%). These features are not seen in the EnMAP data. Overall the mixing relationships show that PV and NPV dominate most pixels and are interchangeable with neither having a stronger preference to mix with bare soil, which is very limited in the region. The relationship between NPV and soil follows the same pattern than shown for PV and soil. The strongest relation exist between green and dry vegetation (HyMap: R^2=0.67; EnMAP: R^2=0.54). To account for the relationship between the geomorphological parameter slope derived from the DEM and class abundance, it was examined to assess whether patterns exist and indicate potential correlations. Linear regression analysis could not reveal significant correlations and, from additional visual assessments it can be concluded that all three surface components occur with different fractions at all slope degrees (Appendix, Fig. D).

In a subsequent step the similarity between the calculated fractions of HyMap and EnMAP was quantitatively assessed by plotting the pixels of each class from
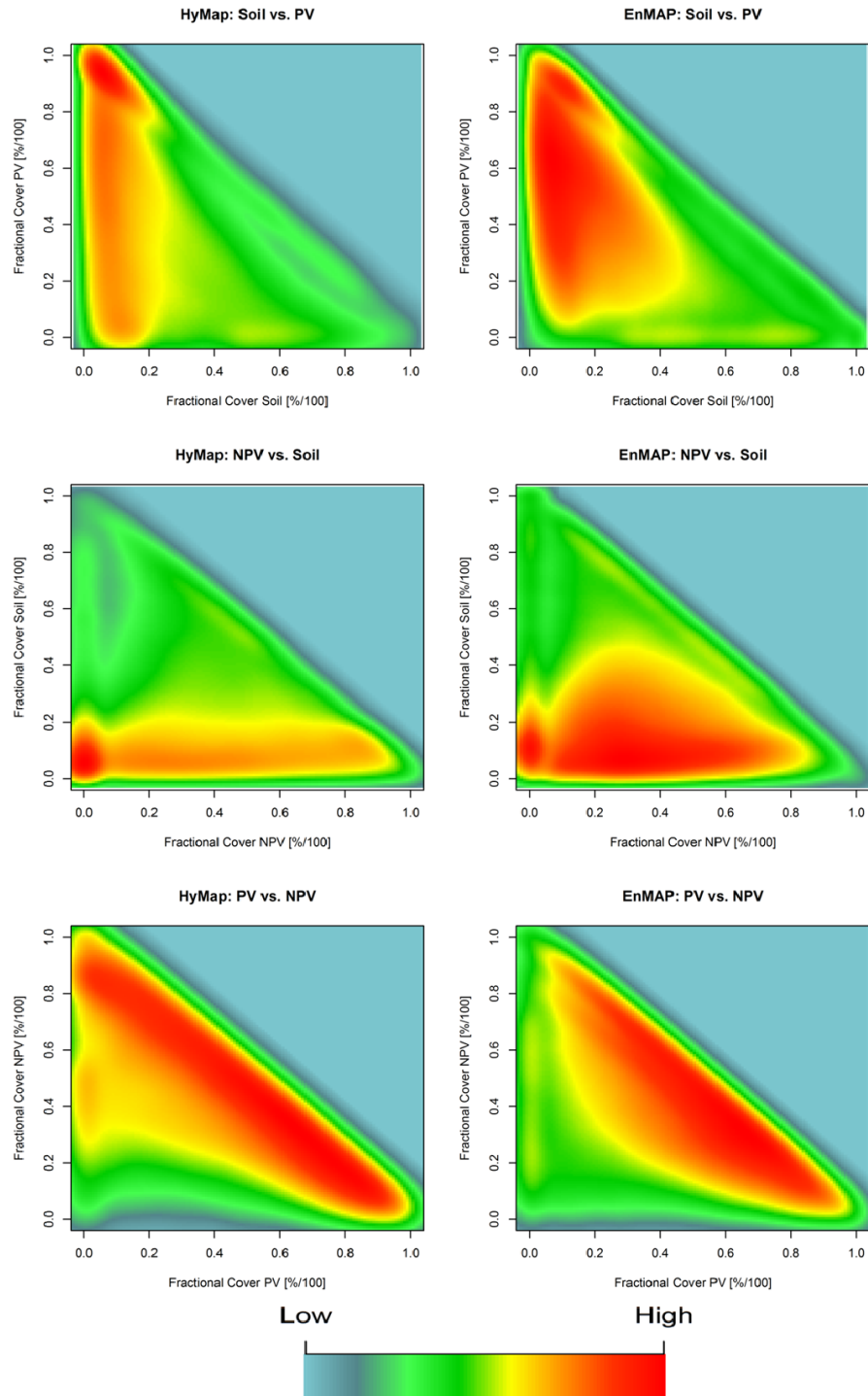
Figure 6: *Density scatter plots showing relationship between fractional cover classes for HyMap (column 1) and EnMAP (column 2).*
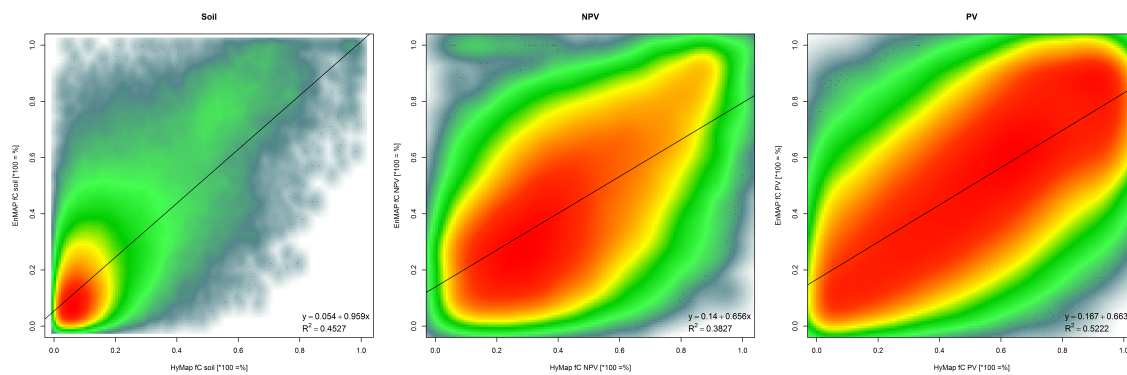
Figure 7: *Scatterplots of soil, NPV and PV fractional abundance results for EnMAP compared to the reference imagery (HyMap) with associated slope, intercepts and calculated R^2 values.*

HyMap separately against the corresponding class of EnMAP. Homogeneous regions were defined to reduce effects of the difference on pixel size and remove edge effects along spatial boundaries between different classes. Figure 7 show the resulting scatter plots with associated slope, intercept and Pearsons R^2 values. For all 3 classes one can see that there is a overall correlation for all fractional values. However, it is also evident that there is also a high degree of variance along the correlation axis, indicated also by the low R^2 values. This may indicate that even within homogeneous regions actual fractional abundances can be quite variable. It is evident from the scatter plots that EnMAP commonly over estimates soil fractions, whereas NPV overestimates at low fractional values and underestimates at high values. A similar case can be seen for PV, although not as strong as NPV. The results further demonstrate that the EnMAP and HyMAP results are consistent overall, but local differences can result in significant differences in calculated fractions, even in homogeneous areas.

To further evaluate the HyMap and EnMAP results a comparison of the different endmember sub classes from both sets of spectral libraries used in each pixel to unmix the spectra was conducted. The comparison is focusing only on the predefined homogeneous areas and is shown in figure 8 as a confusion matrix. For none of the three classes soil, NPV and PV significant patterns can be observed. Soil represented by 3 sub classes in the satellite data and 5 classes in the airborne data. With the exception of HyMap sub class E all endmembers are more or less "randomly" substitute with the three EnMAP endmember sub classes with the majority of pixels been modeled by MESMA with endmembers belonging to sub
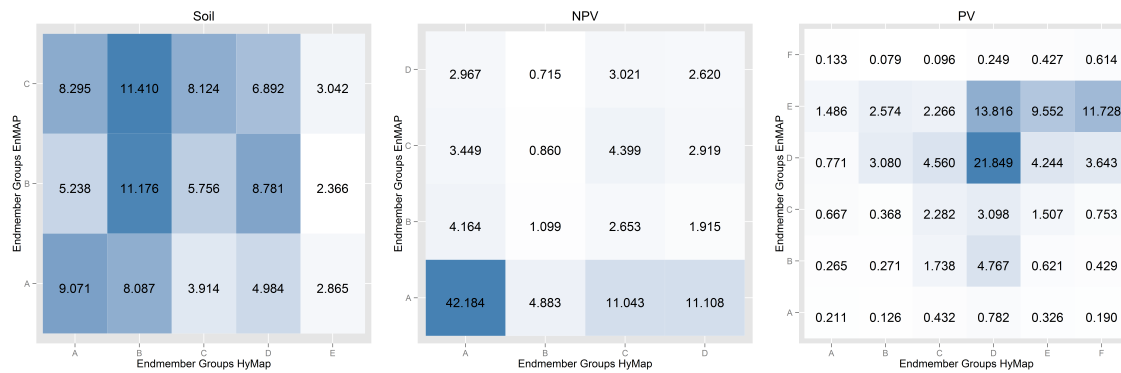
Figure 8: *Percentage of endmembers used to unmix HyMap and EnMAP data according to different endmember groups independently derived with SSEE for both data sets.*

class C. A significant relation between endmember sub class A from class NPV (42%) for both data sets can be observed underpinning also the dominance of spectra belonging to this sub class across the entire study site for both images. The fact that just for HyMap a higher diversity of different sub classes have been used by MESMA (group C 11% and group D 11%) might indicate, that MESMA is sensitive to the spatial resolution of input images. Both data sets delivered the same corresponding vegetation sub classes, but, similar to NPV, one dominate sub class can be identified (D with 22%) and most pixel have been modeled for EnMAP with spectra from sub class D and E (around 80%). Compared to HyMap these two sub classes are used for approximately 61% of all pixel with fractional cover values for PV and the third most used group is F (around 17%).

A further assessment of fractional cover values estimated by MESMA for both images was accomplished by looking location-wise at statistically significant differences in pixel values. The results for the soil class are displayed in figure 9. The distribution map shows the same pattern that soil fractional cover maps estimated by MESMA. Nearly all pixel with an soil endmember identified in images are either overestimated in the EnMAP data or underestimated for HyMap. Assuming HyMap as "truth" data, one can see that more pixels are underestimated by EnMAP (yellow). An visual investigation of these results reveal, that the coarser pixel size of EnMAP changes pixel surface composition compared to HyMap. This results in a use of different endmember combinations for modelling abundances in these locations and the satellite image tend to use higher fractions of the soil endmember if this is present in the pixel, but smaller soil patches predicted

with higher abundances from airborne data can not be adequately captured with EnMAP. Looking at the significant difference between the fractional abundance maps of PV and NPV a correlation appears between MESMA's modelling RMS (Appendix, Fig. G) and over- or underestimations. For the fractional cover of PV significant differences between the results of both sensors appear mainly within two flight lines which have been used for the mosaic creation. Both flight lines show high RMSE for the unmixing results based on the HyMap mosaic, whereas the EnMAP unmixing results show a rather low RMSE at these locations. The significant difference of PV fractional cover maps indicate, that the modelling RMSE is associated with PV, as EnMAP tends to underestimate the abundance of PV. Instead the significant differences between NPV fractional cover results shows, that MESMA based on EnMAP data assigns higher fraction cover values in these locations than HyMap does. As the modelled shade component does only slightly differs between HyMap and EnMAP as well as the detected endmembers for both classes NPV and PV pick up the same sub classes for both sensor systems, this result might be associated with the differences in spectral resolution of both systems. Comparing the endmember sub classes for PV subclass D and E have been used to unmix most of the pixel for EnMAP and both mean spectra of the sub classes show a slight drop around 2200 nm, which does not appear for the corresponding HyMap classes. A similar case can be observed for NPV, where the mean spectrum of sub class A shows a stronger lignin feature than the HyMap spectrum of class A does. These slight differences in the spectra shape in the SWIR region of the main sub classes used for unmixing might cause the shown differences between the resulting fractional cover maps for both data sets. This result emphasizes that μMESMA is sensitive to key spectral features features that are needed to discriminate the classes PV, and more importantly NPV and soil.

2.5.3   Soil erosion parameters

Consistent with MESMA's soil fractional cover maps for HyMap and EnMAP the C-factor map calculated based on equation 2.3 picks up large and connected patches, whereas the scattered effected observed in the soil abundance maps get blurred in the C-factor map (Appendix, Fig. E (A)). This is expected, because the C value is sensitive to abundances of PV and NPV, thus relatively low absolute fractions observed in soil abundance maps are less sensitive to soil cover as the C values expresses the relation between all fractions observed in a location. The C values are scaled between 0 and 1, where 0 relates to dense vegetated (PV and NPV) areas and 1 to bare soils. Looking at the relative distribution of C
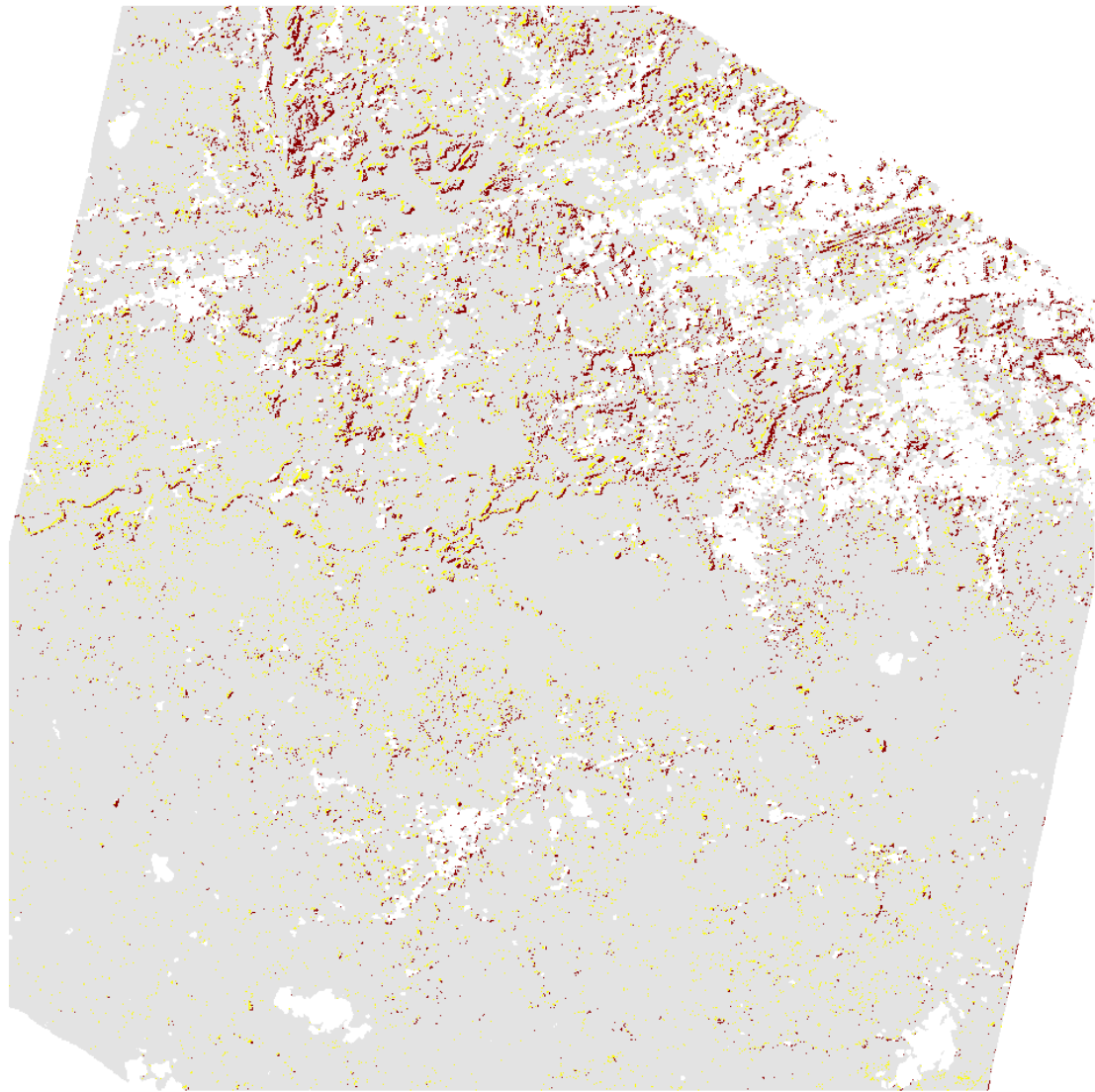
Figure 9: *Distribution maps of significant difference between modelled soil fractional covers of HyMap and EnMAP. Dark red areas represent pixel where soil fractional cover values of EnMAP are significantly higher than for HyMap, yellow areas indicate areas where soil fractional cover values are significantly lower for EnMAP than predicted for HyMap data. Grey areas are locations where no significant changes of modelled soil abundances are detected.*

factor values for HyMap most pixels indicate very low C values between 0 and 0.15. Compared to soil fractional cover values the distribution is slightly shift toward lower values with higher frequencies of C values of 0 (Fig. 10). The same shift in the distribution of C values can observed for EnMAP. The second case of C-factor maps for HyMap and EnMAP, where NPV and soil are combined in the nominator show much higher C values for the entire study area with far larger patches (Appendix, Fig. E (B)). This is expected as this case is not solely sensitive to soil, but also to NPV, thus pixel with high abundances of NPV are categorized as potentially prone to soil erosion by high C values. This result clearly highlight the importance of being able to distinguish and separate fractional cover for soil and NPV.

LS factor values for the study area vary between 0.03 and 36.40 with very high values (>35) in the south and southeast of the study site in the Cerro Cedral where steep slopes about 30 degrees dominate (Appendix, Fig. F). Very low values are observed in the urban areas and in the northwest part of the study region were agricultural activities occur and along the rivers (0 to <5). In these locations also the slope is relatively low. Moderate LS values (around 20) dominate hilly terrains. The range of the LS value is consistent with study of Ghosh et al. (2013) who calculated the LS factor for the tropical Dahalai river basin in India (LS value between 0 and >35). Therefor we do assume, that the LS factor calculation in GRASS GIS does fit the topographic features of Costa Rica.

2.5.4 Integrated soil erosion modelling

For the integrated erosion modelling approach used here the LS factor is integrated into equation 2.3 by multiplying the C factor, for both cases (PV+NPV and Soil+NPV), with the LS parameter. The same LS factor is used for both HyMAP and EnMAP results as the same DEM is used. Similar to the fractional cover estimates Figure 11 show the same overall patterns between HyMAP and EnMAP for each case. However, with respect to the two cases one can see notable differences. Specifically, there is an overall increased erosion potential on the steep slopes in the more mountainous regions and decreased potential in the flatter agricultural areas in the north part of the study area. Additionally, much of the high erosion potential in the southern more mountainous area is reduced in the first case, but with some local areas still remaining high potential (e.g. arrow (i) in figure 11). In other cases erosion potential almost completely disappears (e.g. arrow (ii) in 11). These differences are directly attributed to NPV being combined with soil in the second case, which would reflect the inability to separate these
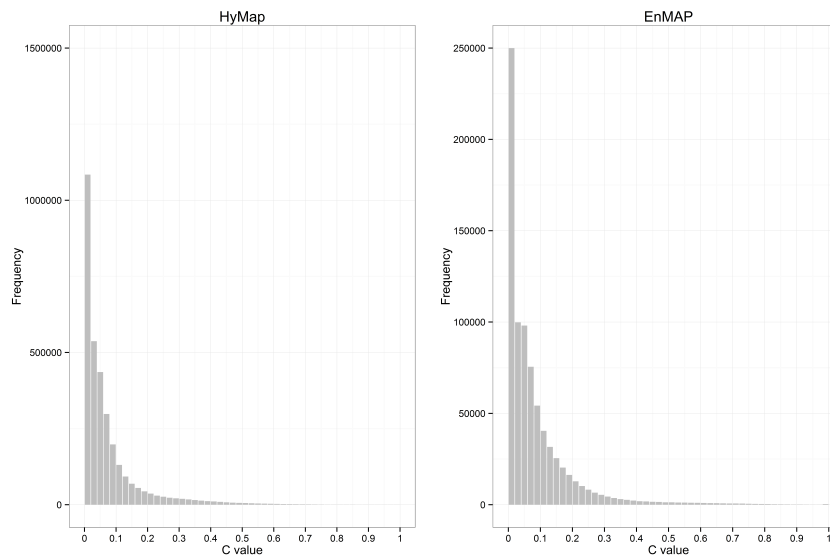
Figure 10: *Histogram showing the distribution of C factor values for HyMap and EnMAP.*

two classes based on spectral information. This highlights the importance of NPV on reducing erosion potential.

With respect to the more optimal case of separating soil and NPV we can see that the key differences between HyMap and EnMAP are again at a local scale (Fig. 11 (A)). To highlight these local differences Figure 12 shows 3 different areas within the full study area. The first area (A) is located in an area characterized by large agricultural fields with various degrees of soil exposure. A number of rivers also cross cut through the area. Closer examination of the results show that areas more prone to erosion occur proximal to the rivers and not necessarily associated with high soil exposure in the fields. Also notable differences do occur between the outputs, such as a large high potential area in the EnMAP but missing in the HyMap, as well as the loss of a river feature in the EnMAP results (e.g. arrow (i) and (ii) in Figure 12 (A))). The second zoom area (B) is characterized by a more rugged hilly terrain with extensive plantations with dense vegetative cover mainly occurring in the valleys. The distribution of soil erosion prone areas appear to occur in well defined zones and do not visually correlate with areas of exposed soil and NVP. Arrows (iii) and (iv) in figure 12 (B) highlight cases where erosion potential occurs very local at a small spatial scale in HyMap, but are lost entirely in the EnMAP results. The last zoom area (C) occurs in the highest most rugged mountain areas where there is limited agricultural activity and more extensive
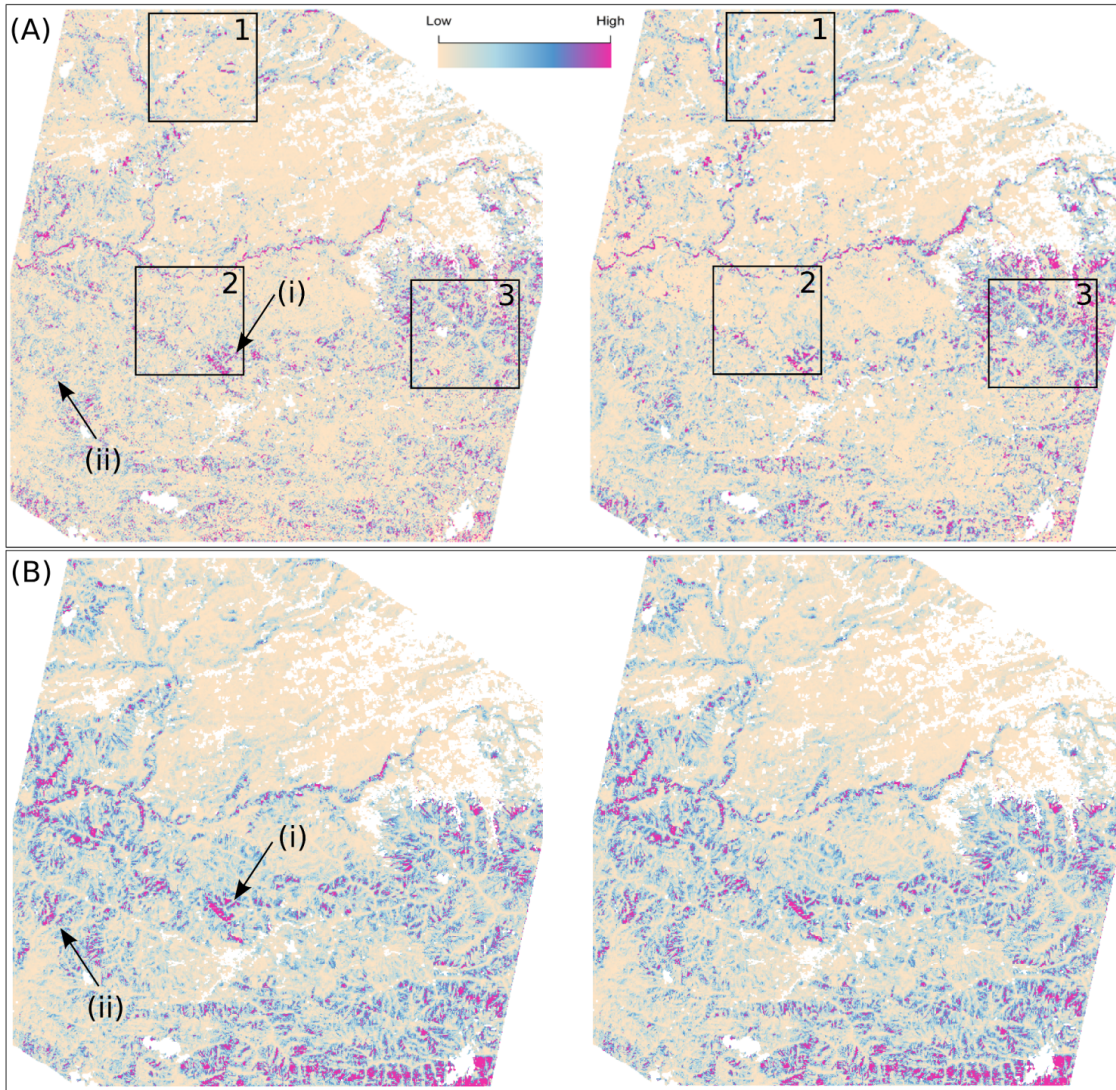
Figure 11: (A) Soil erosion potential maps for the full study area derived from C (case 1) and LS factor information for HyMap (left) and EnMAP (right) imagery; (B) Erosion potential maps for the full study area derived from C (case 2) and LS factor information. See text for details on arrow descriptions. Boxes 1, 2 and 3 represent zoom windows A, B and C in figure 12.

PV. In this latter case high erosion prone areas are more abundant than in the other two zoom areas. In this case HyMap and EnMAP results are very similar with EnMAP only showing higher overall values that are slightly more spatially extensive.

2.6 Discussion

2.6.1 EnMAP Simulation Quality

The output from the Eetes simulation software is well tested and understood (Segl et al., 2012), and as such, the critical factor for the EnMAP simulation is the input data. The generation of the EnMAP simulation included a great deal of care in building the highest quality mosaic of the airborne HyMap data. Special attention was paid to improving radiometric quality through the assessment of wavelength shifts for individual lines and during the extent of the flight campaign. Through improvements in radiometric quality important spectra features in the SWIR near 2100-2200 associated with lignin allowed for improved discrimination of representative NPV and soil spectra from the image data directly. Even though HyMap has only half the overall spectral bands as EnMAP the difference does not impact the ability to discriminate important features of interest (e.g. clay and lignin), which means that this information is transferred to the simulation. The high quality of the resulting mosaic was also evident in the fact that the resulting fractional cover estimates are highly consistent across the HyMap flight lines which can be attributed to improved radiometric consistency and BRDF correction, although some inconsistency does remain. As such, the differences observed between HyMap and EnMAP results can be attributed to the differences in sensor characteristics rather than issues related to the simulation generation.

2.6.2 μMESMA

The distribution of significant difference of fractional abundance values of classes between HyMAp and EnMAP data show an bias in predicted abundance values (e.g. over or under estimation of values). However, qualitative comparisons clearly indicate similar overall distribution of fractional estimations (e.g. spatial patterns, mixing relationships). This leads to the question, if these differences are related to spectral and/or spatial differences. In Thorp et al. (2013) Airborne Visbile/Infrared Imaging Spectrometer (AVIRIS) data was used to investigate the effect of spatial and spectral characteristics on MESMA performances for PV,
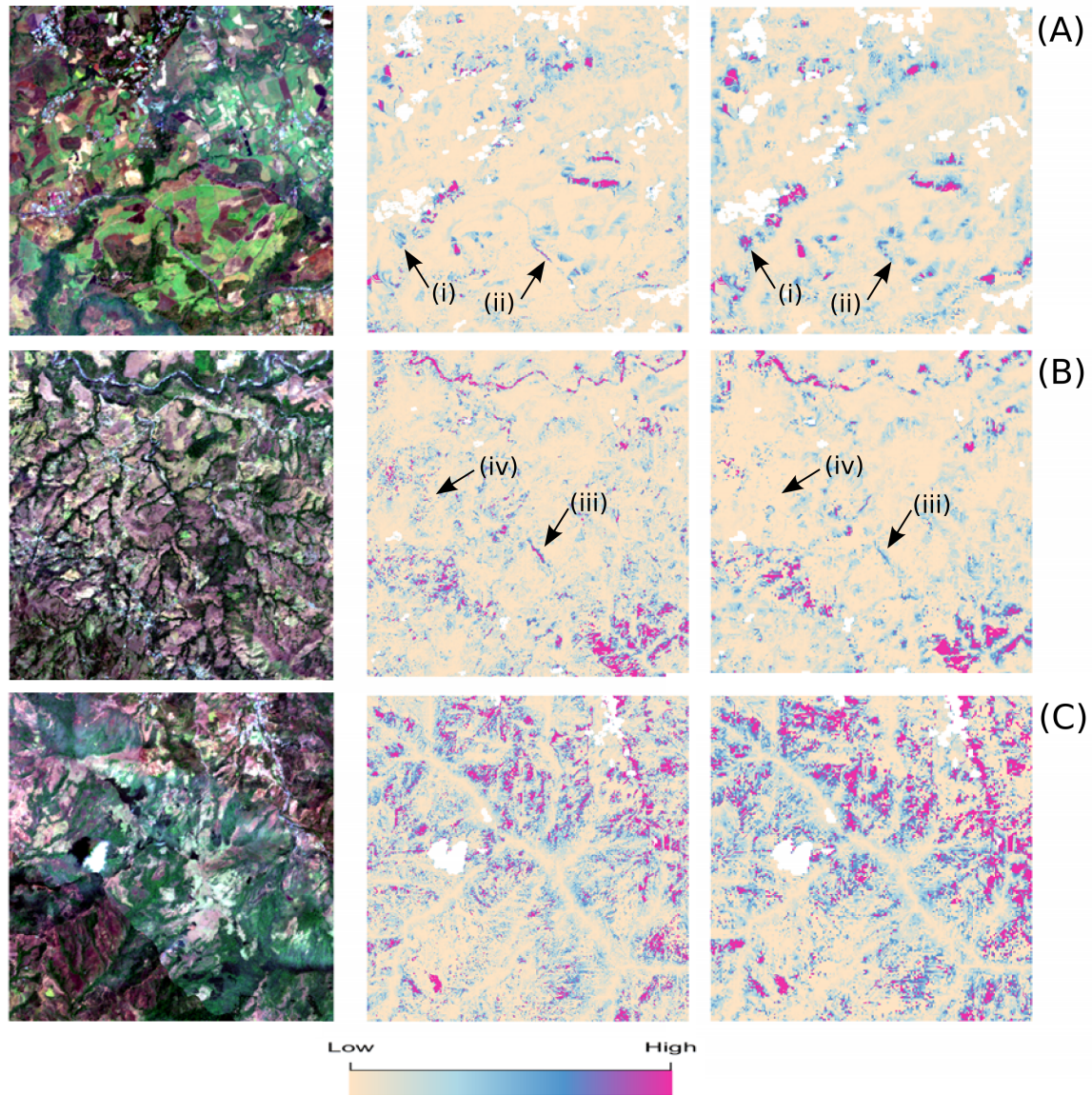
Figure 12: *Zoom areas **(A)**, **(B)** and **(C)** highlighting local differences of full soil erosion potential map from figure 11. Left column shows RGB true color image for each zoom area. HyMap is shown in the middle column with EnMAP results shown in the right column. See text for details on zoom areas.*

NPV, soil in semi-arid rangelands. Results showed that MESMA is more sensitive to spectral degradation than spatial degradation. Testing different degradation scenarios revealed the effect of image endmember spectra on MESMA analysis. The findings of Thorp et al. (2013) pointed out a higher sensitivity of MESMA with spectral rather than spatial resolution, which was based on the removal of key spectral ranges (e.g. VIS/NIR, SWIR). In this study the same spectral wavelength ranges were used for both sets of imagery, with only the number of bands changed. Additionally, the spectral features visible in EnMAP are dependent on the original spectral features in the HyMap data, with the exception of noise specific features added during the simulation process. For the results from this study one can see that the higher spectral resolution of EnMAP (242 bands compared to HyMap with 125 bands) on the derivation of relative fractional cover appears to be negligible. It is shown, that the applied endmember extraction method results in similar endmember groups, except from soil, for both data sets. The key absorption features characterizing the classes and their associated sub classes were detectable in both EnMAP and HyMap. The applied endmember extraction method derived similar endmember classes for both data sets, with the exception of two additional sub classes for soil extracted from the HyMap data. This indicates that for both the HyMap and EnMAP spatial resolutions the majority endmember representatives of PV, NPV and soil occur locally in reasonable large homogeneous regions. Whereas for the two extra soil sub classes derived from HyMap pixel size likely influenced their ability to be extracted. This is likely related to the coarser spatial resolution of EnMAP, where smaller soil patches that are still large enough to appear as pure pixels in HyMAP are mixtures with other surface components in the 30 m EnMAP pixels. Similar conclusions were drawn by Rogge et al. (2014), whom considered airborne AISA and simulated EnMAP data for mapping geology, pointing out that the different spatial resolutions (2 and 30 m pixels) was the key influence on important differences in the extracted endmembers.

Results derived for both airborne HyMap and simulated satellite EnMAP imagery produce similar overall spatial patterns and qualitative values of fractional coverage. The key differences were found at a local scale, where the distribution of surface materials commonly occurs below that of the 30 m pixel size of EnMAP. This was evident from the class to class scatter plots that highlight slightly different distribution densities and mixing characteristics between classes (Fig. 6). For homogeneous regions one would expect the fractional values to be better correlated. However, in Figure 7 and 9 one can see that even for homogeneous regions there exists a high degree of variance of absolute fractions between the EnMAP results and reference data. Thus, not only local differences in spatial

distribution and values occur, there is also an absolute difference between the two data sets.

With increasing pixel size there is a higher degree of surface mixtures within a given pixel. This result is more critical for local studies, where mapping results have to be very precise in respect to the location. As seen from the perspective of large scale studies (e.g. regional or national scale) the significant differences in the fractional cover values of HyMAP and EnMAP appear as spatially related patterns, thus the areas where high or low values appear are just shifted by a a few meters, but the overall distributions are still very similar. Similar results were also shown in Okujeni et al. (2015) who used hyperspectral remote sensing data can retrieve vegetation cover, vegetation type, and soil type in areas of low vegetation cover using MESMA.

Another interesting outcome is the use of specific sub classes used within MESMA. With respect to this study there is not a strong correlation between which sub class were used compared to the reference data. Additionally, specific sub classes appear to dominate. The latter case may just indicate an overall spatial dominance of a specific sub class, but it may also indicate that specific spectral features in a given sub class dominate the unmixing process. In either case, this presents a problem if the future purpose is to further build the integrated soil erosion modelling based on different soil and/or vegetation types.

2.6.3   Integrated soil erosion modelling

As ground data is not available the integrated modelling results can only quantitatively assessed. Combining NPV and PV in the C factor is important, as it influences the weighted effect on LS in specific locations. For example, in areas in the image where it appears that soil and NPV are dominant, the resulting erosion factor is low. This indicates a strong influence of the NPV that helps to bind the soil together. On the other hand if NPV and soil are combined in the C factor calculation, these areas may be shown to be prone to erosion as shown in Figure 11. Thus, additional spectral information provided by EnMAP is critical to better separate soil and NPV. From a regional perspective EnMAP can be used to highlight specific areas that may be prone to erosion. This is evident from the results as the prone areas highlight specific regions and are not randomly distributed. This result is consistent with the statement of De Asis and Omasa (2007) that traditional methods to quantify vegetation information to map the C factor are inaccurate. Since vegetation indexes like the NDVI relay on the detection of healthy or saturated chlorophyll content of the plants which is not

necessarily representing the protection function of the vegetation for soil erosion. It is also shown, that different class distributions of PV and NPV observed between HyMap and EnMAP are removed by the calculation of the C factor based on the purposed approach, because the cover fractions of those classes are summed and set in relation to soil fractions. The abundances for soil are spatially very similar between air borne and satellite imagery and overestimation detected by EnMAP only effects the scaling of the calculated erosion risks. Considering the three zoom areas one can see that areas potentially prone to soil erosion are located in distinctive patterns that reflect the use of integrated C and LS factors. As the same LS factor is used for HyMAP and EnMAP, it can be assumed that differences are primarily controlled by the C factor. Thus, discriminating and accurately mapping fractional cover of PV, NPV and soil using automated image based methods is essential, which has been successfully demonstrated in this study.

As the calculation of the LS factor is controversial discussed in literature and numerous studies have been conducted on the parameterization of the equation to calculate slope length, bringing up many questions and concerns about the quality of estimations Renard et al. (2011); McCool et al. (1995), the suitability of the LS factor algorithm used in this study might be increased by detailed parameter investigation for the Costa Rican area. Yet, McCool et al. (1995) pointed out that the LS calculation from RUSLE does estimate soil loss in terrains with steep slopes better that the LS methodology implemented in USLE, because soil loss is much more sensitive to the slope steepness than to slope length. Additional considerations should concern the fact, that RUSLE was developed for soil loss estimations on semi-arid rangelands, thus parameter adjustment might enhance the accuracy of the proposed methodology for tropical ecosystems. Similar deliberations should be made regarding C factor calculation. Even though, De Asis and Omasa (2007) have shown, that the correlation between field measured C factors and the proposed C factor calculation based on spectral unmixing is very high (Pearsons R^2=0.94), other studies also claim the importance of considering the type of management. For example, Tiessen et al. (2010) showed that only-downslope cultivation is a highly erosive tillage practice with potential erosion rates more than double of those for slopes being up-slope cultivated. Looking at the erosion potential of plantations (Fernández-Moya et al., 2014) pointed out that plantations are not directly causing high erosion rates, in fact the severity of soil loss is determined by the presence or absence of an understory vegetation layer. Even though a leaf layer may be present, studies have confirmed that missing understory vegetation can increase soil loss compared to bare soil plots (e.g. Calder (2001); Labrière et al. (2015)). This is because the raindrops are collected by

the leafs in a first step, but as they get to heavy much bigger drops are released to the surface having a much higher kinetic energy than the actual raindrops. Thus the absence of protective ground cover increases erosion rates. This is consistent with the depletion of soil organic carbon pools detected by Häger (2012) in coffee agroforestry systems in Costa Rica. The comparison of carbon storage in the different farm types showed that highest storage rates are related to the average tree density per hectar, tree height and wood density. Against this backdrop, C factor calculation should also consider, in addition to the fractional cover of soil, PV and NPV, the landuse type and it's associated erosion risk.

Based on expert knowledge correlations between C values and coffee plantations can be identified. Considering newest research and the fact that the ratio of soil loss between bare soils and cropland can be 20:1 in the humid tropics (10:1 in the Mediterranean) (Labrière et al., 2015) attention has to be drawn in this locations to secure early detection of depleting vegetative cover and dispersal of bare soils. Integrated modelling using soley two RUSEL parameters does provide first impressions of the distribution of areas prone to erosion. Labrière et al. (2015) demonstrated, that sufficient vegetative cover is crucial for soil erosion control in tropical ecosystems. Structural changes in the composition of PV, NPV and soil triggered by agricultural management practices or other activities harming fragile ecosystems can abruptly change the susceptibility of such location for erosion processes. Special attention should be given to agricultural areas and in general to anthropogenic influenced landscape to mitigate and prevent soil erosion processes.

As noted before (see chapter "Motivation"), this study does not aim to model annual or event-based soil loss rates, it seeks to spatially identify areas which are potential prone to erosion ensuing from landscape configurations and surface cover. Tropical climates experiencing severe rainfall events, where most of the annual precipitation falls only during a few rainfall events (Calvo-Alvarado et al., 2014). This fact does highly underpin, that actual soil loss rates especially in the tropics, can only be estimated if the high rainfall erosivity is integrated in modelling approaches. Microclimate variations in Costa Rica making the raifall erosivity index extremely variable in space (Calvo-Alvarado et al., 2014), meaning, that areas being identified by the underlying approach as highly susceptible to soil erosion may actually show lower soil loss rates that other areas, because the rain energy to detach soil particles is far less in those locations.

3 | Conclusion and Outlook

As no ground truth data is available for this study, the assumption was made, that the HyMap results are more accurate and can be considered as reference data. Given HyMap as reference it has been demonstrated that at the 30 m pixel resolution of EnMAP it is possible to extract both quality endmember classes that contain important spectral features related to PV, NPV and soil, and also, to be able to estimate relative cover fractions for a large regional study area. Problems lie in a bias of either over or under estimations of fractional cover compared with the reference data. Moreover, differences in the fractional cover maps of NPV and PV between EnMAP and HyMap might indicate, that the detected endmembers need further refinement. Regarding the fractional cover estimation it has been shown, that this information can be extract directly from the imagery using automated processes.

If longer term monitoring of soil erosion potential is of interest then airborne data acquisitions are not optimal. EnMAP, on the other hand, is in a position to promote this temporal data at low cost. Additionally, the results demonstrate that the 30 m spatial resolution can be overcome using spectral mixture analysis approaches that can reasonable estimate sub pixel cover fractions, as long as good spectral representations of ground cover classes can be derived.

This work does also demonstrate the potential of EnMAP to be used to assess some drivers of soil degradation (e.g. erosion) over vast regions (potentially 30x5000 km of imagery per day) that can be made available to the scientific community to derive useful information about Earth Surface processes. Moving forward one goal should be to expand the existing methodologies to other regions of the HyMap Costa Rica survey data and to enhance the integrated modelling approach by parameter calibration for tropical ecosystems as well as to take into account other RUSEL parameters which can be derived on a large scale based on available data such as the R factor as it has been shown by (Calvo-Alvarado et al., 2014). With the integration of recent high quality remote sensing products such as the 12 m DEM processed from TerraSAR-X and TanDEM-X or the Global

urban footprint produced by the German Aerospace Center in combination with future EnMAP time series, soil erosion estimates can be accomplished regional or nationally with high quality spatial precision. The approach introduced in this study does show, that the launch of EnMAP offers enormous potential for the development of an fast and cost effective tool to accomplish the task of continuous monitoring the structural conditions of the landscape over time. Thereby the temporal component of erosion can be considered to account for seasonally soil losses coupled with crop rotations or plant phenology.

Bibliography

Adler-Golden, S. M., Levine, R. Y., Matthew, M. W., Richtsmeier, S. C., Bernstein, L. S., Gruninger, J. H., Felde, G. W., Hoke, M. L., Anderson, G. P., and Ratkowski, A. (2001). Shadow-insensitive material detection/classification with atmospherically corrected hyperspectral imagery.

Akhtar, N., Shafait, F., and Mian, A. (2015). Futuristic greedy approach to sparse unmixing of hyperspectral data. *IEE Transactions on Geoscience and Remote Sensing*, 53(4):2157–2173.

Asner, G. and Lobell, D. S. E. . (2000). A biogeophysical approach for automated swir unmixing of soils and vegetation. *Remote Sensing of Environment*, 74(1):99–112.

Asner, G. P., Knapp, D., Cooper, A., Bustamante, M., and Olander, L. (2005). Ecosystem structure throughout the brazilian amazon from landsat observations and automated spectral unmixing. *Earth Interactions*, 9:1–31.

Bachmann, M. (2007). *Automated estimation of ground cover fractions using MESMA unmixing*. PhD thesis, University of Wuerzburg, Germany.

Bachmann, M., Habermeyer, M., Holzwarth, S., Richter, R., and Müller, A. (2007). Including quality measures in an automated processing chain for airborne hyperspectral data. In *Proceedings 5th EARSel Workshop on Imaging Spectroscopy, Bruges, Belgium*.

Bachmann, M., Müller, A., and Dech, S. (2009). Increasing and evaluating the reliability of multiple endmember spectral mixture analysis (mesma). In *Proceedings 6th EARSeL-SIG-IS*, Tel Aviv, Israel.

Bachmann, M., Müller, A., Habermeyer, M., Schmidt, M., and Dech, S. (2005). Iterative mesma unmixing for fractional cover estimatesŰ evaluating the portability. In Zagajewski, B., Sobczak, M., and Wrzesien, M., editors, *Proceedings of 4th*

EARSeL Workshop on Imaging Spectroscopy. New quality in environmental studies., EARSeL and Warsaw University, Warsaw 2005.

Bakimchandra, O. (2011). *Integrated fuzzy-GIS approach for assessing regional soil erosion risks.* PhD thesis, University of Stuttgart, Germany.

Bayer, A. (2012). *Methodological developments for mapping soil constituents using imaging spectroscopy.* PhD thesis, University of Potsdam.

Ben-Dor, E. and. Chabrillat, S., Demattê, J., Taylor, G., Hille, J., Whiting, M., and Sommer, S. (2009). Using imaging spectroscopy for soil properties. *Remote Sensing of Environment*, 113(Supplement 1):S38–S55.

Bioucas-Dias, J., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., and Chanussot, J. (2012). Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 5(2):354–379.

Boardman, J. W. (1993). Automating spectral unmixing of aviris data using convex geometry concepts. *Summaries of the Fourth Annual JPL Airborne Geoscience Workshop, JPL Publication 93-26*, 1:11–14.

Boardman, J. W., Kruse, F. A., and Green, R. O. (1995). Mapping target signatures via partial unmixing of aviris data. *Summaries, fifth JPL airborne earth science workshop. JPL Publication 95-1*, 1:23–26.

Calder, I. R. (2001). Canopy processes: implications for transpiration, interception and splash induced erosion, ultimately for forest management and water resources. *Plant Ecology*, 153(1-2):203–214.

Calvo-Alvarado, J. C., Jiménez-Rodríguez, C. D., and Jiménez-Salazarr, V. (2014). Determining rainfall erosivity in costa rica: A practical approach. *Mountain Research and Development*, 34(1):48–55.

Daughtry, C., Doraiswamy, P., Hunt Jr, E., Stern, A., McMurtrey III, J., and Prueger, J. (2006). Remote sensing of crop residue cover and soil tillage intensity. *Soil and Tillage Research*, 91(Issue 1-2):101–108.

De Asis, A. M. and Omasa, K. (2007). Estimation of vegetation parameter for modeling soil erosion using spectral mixture analysis of landsat etm data. *ISPRS Journal of Photogrammetry & Remote Sensing*, 62:309 Ű 324.

Delegido, J., Verrelst, J., Rivera, J. P., Ruiz-Verdu, A., and Moreno, J. (2015). Brown and green lai mapping through spectral indices. *International Journal of Applied Earth Observation and Geoinformation*, 35:350–358.

Dennison, P. E. and Roberts, D. A. (2003a). The effects of vegetation phenology on endmember selection and species mapping in southern california chaparral. *Remote Sensing of Environment*, 87(2-3):295 – 309.

Dennison, P. E. and Roberts, D. A. (2003b). Endmember selection for multiple endmember spectral mixture analysis using endmember average {RMSE}. *Remote Sensing of Environment*, 87(2-3):123 – 135.

Desmet, P. and Govers, G. (1996). A gis procedure for automatically calculating the usle ls factor on topographically complex landscape units. *Journal of Soil and Water Conservation*, 51:427–433.

Elvidge, C. (1990). Visible and near infrared reflectance characteristics of dry plant materials. *International Journal of Remote Sensing*, 11:1775–1795.

Esch, T., Elsayed, S., Marconcini, M., Marmanis, D., Zeidler, J., and Dech, S. (2014). Dimensioning the degree of urbanization Űa technical framework for the large - scale characterization of human settlement forms and patterns based on spatial network analysis. *Submitted to Journal of Applied Geography*.

Eswaran, H., Lal, R., and Reich, P. F. (2001). Land degradation: an overview. In Bridges, E., Hannam, I., Oldeman, L., Pening de Vries, F., Scherr, S., and Sompatpanit, S., editors, *Responses to Land Degradation. Proc. 2nd. International Conference on Land Degradation and Desertification*. Oxford Press, New Dehli, India.

FAO (1992). *The AFNETA alley farming training manual - Volume 2: Source book for alley farming research*. FAO.

FAO, U. N. N. C. (2014). Un agency calls for urgent action to protect global soil from depletion, degradation. Accessed 02.06.2015.

Fernández-Moya, J., Alvarado, A., Forsythe, W., Ramírez, L., Algeet-Abarquero, N., and Marchamalo-Sacristán, M. (2014). Soil erosion under teak (tectona grandis l.f.) plantations: General patterns, assumptions and controversies. *CATENA*, 123(0):236 – 242.

Förster, M., Bruzual-Alfonzo, H., Clasen, A., and Kleinschmit, B. (2014a). The utilization of geobia for a hyperspectral mixture analysis of tree crown components. volume 3, pages 363–367.

Förster, S., Wilczok, C., Brosinsky, A., and Segl, K. (2014b). Assessment of sediment connectivity from vegetation cover and topography using remotely sensed data in a dryland catchment in the spanish pyrenees. *Journal of Soils and Sediments*, 14(12):1982–2000.

Ghosh, K., Kumar De, S., Bandyopadhyay, S., and Saha (2013). Assessment of soil loss of the dhalai river basin, tripura, india using usle. *International Journal of Geosciences*, 4(1):11–23.

Gray, D. H. and Sotir, R. (1996). *Biotechnical and Soil Bioengineering Slope Stabilization: A Practical Guide for Erosion Control*. John Wiley and Sons, Canada.

Gruninger, J. H., Ratkowski, A. J., and Hoke, M. L., editors (2004). *The sequential maximum angle convex cone (SMACC) endmember model*. Proc. SPIE 5425, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X.

Guanter, L., Segl, K., and Kaufmann, H. (2009). Simulation of the optical remote-sensing scenes with application to the enmap hyperspectral mission. *IEEE Transactions on Geoscience and Remote Sensing*, 47:2340–2351.

Guerschman, J.P.and Scarth, P., McVicar, T., Renzullo, L., Malthus, T., Stewart, J., Rickards, J., and Trevithick, R. (2015). Assessing the effects of site heterogeneity and soil properties when unmixing photosynthetic vegetation, non-photosynthetic vegetation and bare soil fractions from landsat and modis data. *Remote Sensing of Environment*, 161:12–26.

Hapke, B. (1993). *Theory of Reflectance and Emittance Spectroscopy*. Cambridge University Press. Cambridge Books Online.

Häger, A. (2012). The effects of management and plant diversity on carbon storage in coffee agroforestry systems in costa rica. *Agroforestry Systems*, 86(2):159–174.

Hickey, R., Smith, A., and Jankowski, P. (1994). Slope length calculations from a dem within arc/info grid. *Computers, Environment and Urban Systems*, 18(5):365 – 380.

Holdridge, L. R. (1967). *Life Zone Ecology*. PhD thesis, Tropical Science Center, San Jose, Costa RIca.

Januchowski, S. R., Pressey, R. L., VanDerWal, J., and Edwards, A. (2010). Characterizing errors in digital elevation models and estimating the financial costs of accuracy. *International Journal of Geographical Information Science*, 24(9):1327–1347.

Kaufmann, H., Segl, K., Chabrillat, S., Hofer, S., Stuffler, T., Müller, A., , Richter, R., Schreier, G., Haydn, R., and Bach, H. (2006). A hyperspectral sensor for environmental mapping and analysis (invited paper). In *IEEE International Geoscience and Remote Sensing Symposium & 27th Canadian Symposium on Remote Sensing, Denver, Colorado, USA*.

Kefi, M., Yoshino, K., Setiawan, Y., Zayani, K., and Boufaroua, M. (2011). Assessment of the effects of vegetation on soil erosion risk by water: a case of study of the batta watershed in tunisia. *Environmental Earth Science*, 64(3):707–719.

Kim, M. and Gilley, J. E. (2008). Artificial neural network estimation of soil erosion and nutrient concentrations in runoff from land application areas. *Computers and Electronics in Agriculture*, 64(2):268 – 275.

Krauss, T., d'Angelo, P., Schneider, M., and Gstaiger, V. (2013). The fully automatic optical processing system catena at dlr. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-1/W1:177–183.

Kruse, F., Boarsmann, J., and Huntingtin, J. (2003). Comparison of airborne hyperspectral data and eo-1 hypersion for mineral mapping. *IEEE Transaction on Geoscience and Remote Sensing*, 41:1388–1400.

Labrière, N., Locatelli, B., Laumonier, Y., Freycon, V., and Bernoux, M. (2015). Soil erosion in the humid tropics: A systematic quantitative review.

Lal, R. (2003). Soil erosion and the global carbon budget. *Environment International*, 29(4):437–450.

Matsunaga, T., Iwasaki, A., Tsuchida, S., Tanii, J., Kashimura, O., Nakamura, R., Yamamoto, H., Tachikawa, T., and Rokugawa, S. (2013). Current status of hyperspectral imager suite (hisui). In *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, pages 3510–3513.

McCool, D., Foster, G., Renard, K., and Weesies, G. (1995). The revised universal soil loss equation. In *Proc. DOD Interagency Workshop on Technologies to Address Soil Erosion on DOD Lands. Dept. of Defense. San Antonio, TX*, pages 199–202.

Millennium Ecosystem Assessment (2005). *Ecosystems and Human Well-being: Synthesis*. Island Press, Washington, D.C.

Mirsal, I. (2008). *Soil Pollution: Origin, Monitoring & Remediation*. Springer.

Mitasova, H., Brown, W. M., Johnston, D. M., and Mitas, L. (1996). Gis tools for erosion/deposition modeling and multidimensional visualization, part ii: Unit stream power-based erosion/deposition modeling, and enhanced dynamic visualization. *Report for USA CERL, Univ. Illinois at Urbana-Champaign, Illinois*, page 38pp.

Mitra, B., , Scotta, H. D., Dixon, J. C., and McKimmey, J. M. (1998). Applications of fuzzy logic to the prediction of soil erosion in a large watershed. *Geoderma*, 86(3 -4).

Mulder, V., de Bruin, S., Schaepman, M., and Mayr, T. (2011). The use of remote sensing in soil and terrain mapping âĂŤ a review. *Geoderma*, 162(1âĂŞ2):1 – 19.

Okin, G. (2007). Relative spectral mixture analysis. a multitemporal index of total vegetation cover. *Remote Sensing of Environment*, 106:467–479.

Okin, G., Roberts, D., Murray, B., and Okin, W. (2001). Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sensing of Environment*, 77:212–225.

Okujeni, A., van der Linden, S., and Hostert, P. (2015). Extending the vegetationŰimperviousŰsoil model using simulated enmap data and machine learning. *Remote Sensing of Environment*, 158:69–80.

Porras, I., Barton, D., Chacón-Cascante, A., and Miranda, M. (2013). Learning from 20 years of payments for ecosystem services in costa rica.

Renard, K.G.and Foster, G., Weesies, G., McCool, D., and Yoder, D. (1997). *Predicting soil erosion by water - a guide to conservation planning with the revised universal soil loss equation (RUSLE)*, volume Agriculture Handbook No. 703. U.S. Department of Agriculture Handbook 703.

Renard, K., Yoder, D., Lightle, D., and Dabney, S. (2011). *Universal Soil Loss Equation and Revised Universal Soil Loss Equation*. Blackwell Publishing Ltd., 1st edition.

Renard, K. G., Foster, G., Weesies, G., and P., P. J. (1991). Rusle revised universal soil loss equation. *Journal of Soil and Water Conservation*.

Richter, R. and Schläpfer, D. (2015). *Atmospheric/topographical correction for airborne imagery. ATCOR4 User Guide*. Wessling, Germany,.

Roberts, D., Dennison, P., Gardner, M., Hetzel, Y., Ustin, S., and Lee, C. (2003). Evaluation of the potential of hyperion for fire danger assessment by comparison to the airborne visible/infrared imaging spectrometer. *IEEE Transaction on Geoscience and Remote Sensing*, 41:1297–1320.

Roberts, D., Gardner, M., Church, R., Ustin, S., Scheer, G., and Green, R. (1998). Mapping chaparral in the santa monica mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment*, 65:267–279.

Roberts, D., Quattrochi, D., Hulley, G., Hook, S., , and Green, R. (2012). Synergies between vswir and tir data for the urban environment: An evaluation of the potential for the hyperspectral infrared imager (hyspiri) decadal survey mission. *Remote Sensing of Environment*, 117:83–101.

Rogge, D., Bachmann, M., Rivard, B., and Feng, J. (2012). Spatial sub-sampling using local endmembers for adapting osp and ssee for large-scale hyperspectral surveys. *IEEE Journal of selected topics in applied earth observation and remote sensing*, 5(1):183–195.

Rogge, D., Rivard, B., Segl, K., Grant, B., and Feng, J. (2014). Mapping of nicu-pge ore hosting ultramafic rocks using airborne and simulated enmap hyperspectral imagery, nunavik, canada. *Remote Sensing of the Environment*, 152:302–317.

Rogge, D., Rivard, B., Zhang, J., Sanchez, A., Harris, J., and Feng, J. (2007). Integration of spatial-spectral information for the improved extraction of endmember. *Remote Sensing of the Environment*, 110:287–303.

Roose, E. (1996). *Land husbandry - Components and strategy*. 70 FAO Soil Bulletin. Rome: Food and Agriculture Organization of the United Nations (FAO).

Schläpfer, D., Richter, R., and Feingersh, T. (2015). Operational brdf effects correction for wide-field-of-view optical scanners (brefcor). *Geoscience and Remote Sensing, IEEE Transactions on*, 53(4):1855–1864.

Schwieder, M., Leitão, P., Suess, S., Senf, C., and Hostert, P. (2014). Estimating fractional shrub cover using simulated enmap data: A comparison of three machine learning regression techniques. *Remote Sensing*, 6(4):3427–3445.

Segl, K., Guanter, L., Kaufmann, H., Schubert, J., Kaiser, S., Sang, B., and Hofer, S. (2010). Simulation of spatial sensor characteristics in the context of the enmap hyperspectral mission. *IEEE Transaction on Geoscience and Remote Sensing*, 48:3046–3054.

Segl, K., Segl, K., Guanter, L., RogaSS, C., Küster, T., Roessner, S., Kaufmann, H., Sang, B., Mogulsky, V., and Hofer, S. (2012). Etes- the enmap end-to-end simulation tool. *IEEE Journal of selected topics in applied earth observation and remote sensing*, 5(2):522–530.

Shoshany, M., Goldshleger, N., and Chudnovsky (2013). Monitoring of agricultural soil degradation by remote sensing methods: a review. *International Journal of Remote Sensing*, 34(17).

Soil Survey Staff (1999). *Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys*. Natural Resources Conservation Service. U.S. Department of Agriculture Handbook 436, 2nd edition.

Thorp, K., French, A., and Rango, A. (2013). Effect of image spatial and spectral characteristics on mapping semi-arid rangeland vegetation using multiple endmember spectral mixture analysis (mesma). *Remote Sensing of Environment*, 132(0):120 – 130.

Tiessen, K. H. D., Sancho, F., Lobb, D., and Mehuys, G. (2010). Assessment of tillage translocation and erosion by the disk plow on steepland andisols in costa rica. *Journal of Soil and Water Conservation*, 65(5):316–328.

Tomich, T., Palm, C., Velarde, S., Geist, H., Gillison, A., Lebel, L., Locatelli, M., Mala, W., van Noordwijk, M., Sebastian, K., Timmer, D., and White, D. (2005). Forest and agroecosystem tradeoffs in the humid tropics. a crosscutting assessment by the alternatives to slash-and-burn consortium conducted as a sub-global component of the millennium ecosystem assessment. alternatives to slash-and-burn programme. Technical report, Nairobi, Kenya.

Ustin, S., Roberts, D., Gamon, J., Asner, G., and Green, R. (2004). Using imaging spectroscopy to study ecosystem processes and properties. *BioScience*, 54(6):523–534.

van der Meer, F. D.and van der Werff, H. M. v. R. F. J., Hecker, C., Bakker, W. H., Noomen, M. F., van der Meijde, M., Carranza, E. J. M., and Smeth, J. B. D.and Woldai, T. (2012). Multi- and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation*, 14:112–128.

Veganzones, M. and Graña, M. (2008). Endmember extraction methods: A short review. In Lovrek, I., Howlett, R., and Jain, L., editors, *Knowledge-Based Intelligent Information and Engineering Systems*, volume 5179 of *Lecture Notes in Computer Science*, pages 400–407. Springer Berlin Heidelberg.

Veihe, A., Rey, J., Quinton, J., Strauss, P., Sancho, F., and M., S. (2001). Modelling of event-based soil erosion in costa rica, nicaragua and mexico: evaluation of the eurosem model. *Catena*, 44:187–203.

Weltz, M. A., Renard, K. G., and Simanton, J. R. (1987). Revised universal soil loss equation for western rangelands. *U.S.A./Mexico Symposium of Strategies for Classification and Management of Native Vegetation for Food Production In Arid Zones*.

Wischmeier, W. H. and Smith, D. D. (1968). A universal soil-loss equation to guide conservation farm planning. *Trans. Int. Congr. Soil Sci.*

Wischmeier, W. H. and Smith, D. D. (1978). Predicting rainfall erosion losses-a guide to conservation planning. *Agriculture Handbook*, (537).

Zabcic, N., Rivard, B., Ong, C., and Müller, A. (2014). Using airborne hyperspectral data to characterize the surface ph and mineralogy of pyrite mine tailings. *International Journal of Applied Earth Observation and Geoinformation*, 32:152–162.
?

4 | Supplementary Material

4.1 Preprocessing Chain

The hyperspectral dataset used in this study was supplied by the Centre of Earth Observation, University of Alberta, Canada and reprocessed within the framework of this thesis. The preprocessing was based on methodologies integrated in the ISOcertified fully automated processing chain for airborne hyperspectral data (Fig. 13) developed by the DLR and integrated in the FP7 EUFAR (European Fleet for Airborne Research) project FP7 EUFAR (European Fleet for Airborne Research) project (Holzwarth et al., 2011). The first processing step consist of an system correction and was already performed to the supplied data. The second processing step of the standardized preprocessing chain is the application of ortho-rectification, followed by atmospheric correction, ending up in a fully calibrated, standardized data product (Habermeyer et al., 2005). The quality of the intermediate and final products is ensured and documented by quality checks (Bachmann, 2007).
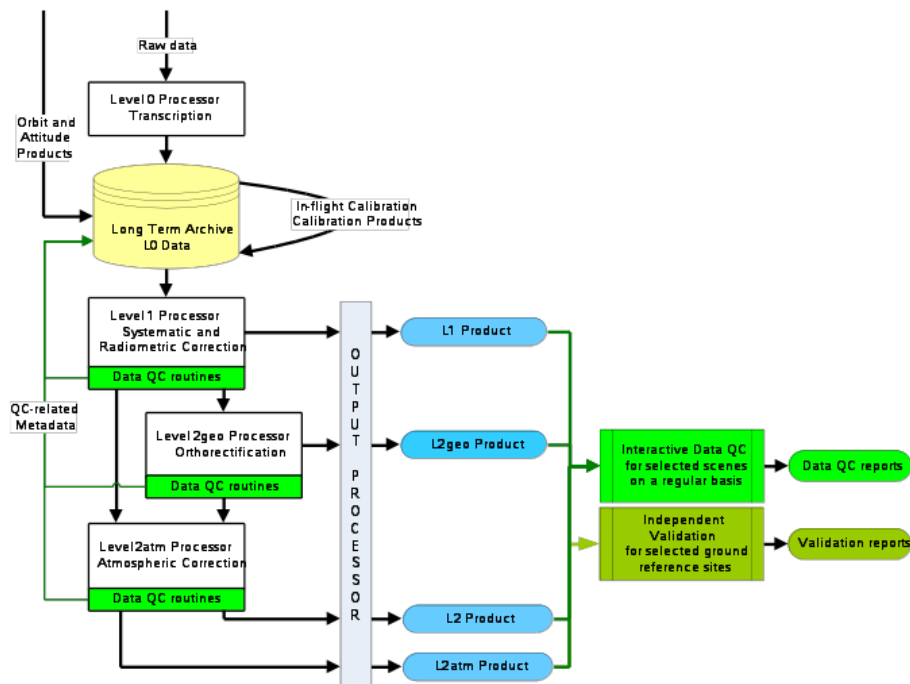
Figure 13: *DLR's standardized processing chain (Source: DLR)*

4.2   Quality Assessment

To provide end users of the data with an transparent and documented processing and assess, an more importantly account for inconsistency in sensor performance during data acquisition or external conditions during the overflight influencing the processing, an automatic quality assessment was performed. Each of the 232 flight lines was analyzed from Level 1 to Level 2 was analyzed by measures based based on scene and sensor characteristics including general information on sensor and software, auxiliary data (e.g. the DEM used for processing), the particular processing methods, as well as scene-specific processing parameters and scene characterization (Bachmann et al., 2007).
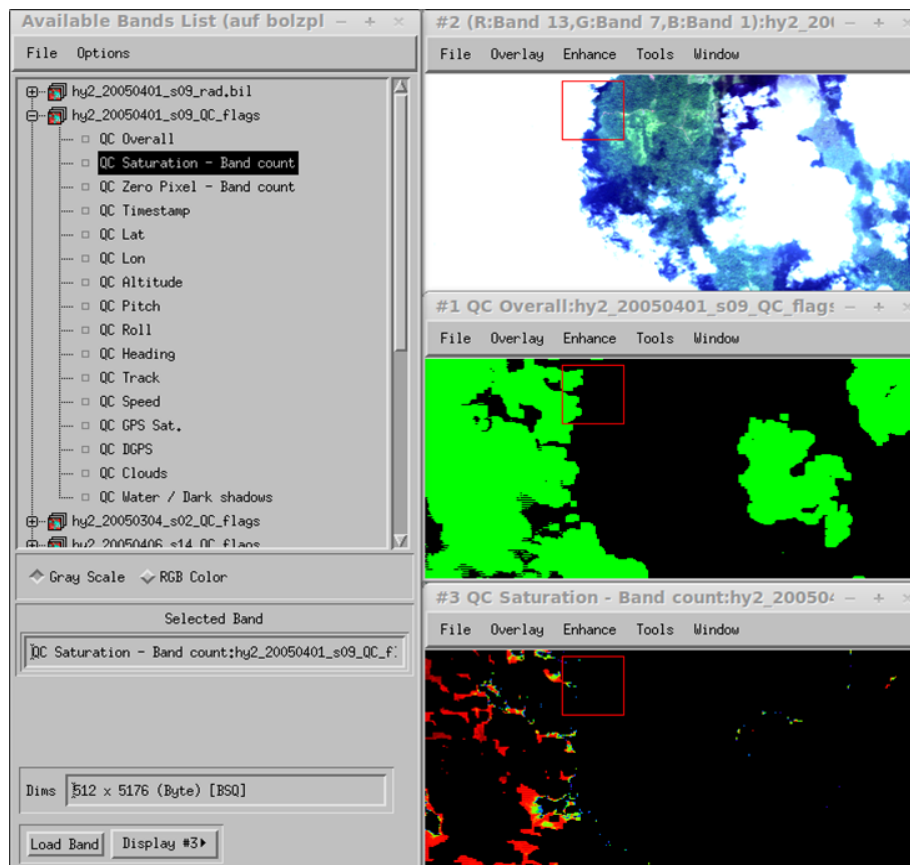
Figure 14: *Example output of Data Quality Flags for cloud cover. True-colour image with a linear 2% stretch over land (top), overall quality with clouds marked in green (note the underestimation) (middle), saturation as band count (botton) (Source:Bachmann (2007).*

The final report and imagery of the data quality assessment summarizes the processor version and level, processing parameters for L2 product generation (e.g. aerosol type, interpolation method, water vapour column), characterization of auxiliary data used for L2 product (e.g. DEM resolution and processing), percentage of "bad" data (e.g., cloud cover, saturated and dead pixels), number of lines with platform instability and DGPS failures, overall scene quality, quicklooks and plots for data characterization (e.g., pitch, roll, heading, cloud and haze masks (Fig. 14), band correlation) as well as general information on the sensor and processing system (Bachmann, 2007).

The output of the overall statistics considering 232 fligh lines, approximatly 1.47 Mio. scan lines, showed for instance a mean cloud coverage of 4% and less

that 1% saturated and no data pixels. Because these results rely on the analysis of L1 data they are most likely underestimated. The analysis of the GPS and IMU data indicated rough flight conditions for approximately 2% of all scan lines, with a mean standard deviation for the parameter roll of approximately 2°with 17 lines having a mean STDEV in roll of >5°. Additional consideration of a STDEV below 15 m of flight altitude the statistics indicate rather shaky flights. The data quality assessment of the GPS and IMU data streams showed small data gaps in the coordinates of pit,roll and heading angles. These data gaps are characteristic for the HyMAP system and typically 2-10 scan lines are affected.

4.3   GPS check and Georectification

The inconsistencies of the gps data detected in the data quality assessment was addressed and corrected by applying an automatic GPS correction for all flight lines based on an algorithm developed by Bachmann. The data is first checked for errors in the coordinates of pit, roll and heading angles. These gps failures usually exhibit from breakdowns of a few seconds of the sensor system and result in frequencies of one pair of coordinates or one value for pit, roll and heading angles for several pixels. After these frequencies are detected, the first and last locations of each frequencies is used for a linear interpolation of the data gap. Due to further existing errors in the three variables, caused by very rough flight conditions which resulted in extreme values in the gps-file, which could not be captured by the existing methodology an extended version of the algorithm was applied. The modified algorithm applied an moving window of 11 vertical pixels to average the values to each of the three variables. This allowed an intensification of anomalies in the data. In a second step values outside the range of two standard derivations were interpolated between the values previous to the aberration and after it. Figure 15 illustrates an recorded data error of the variable roll for one flight line at the gps locations for the pixels, the roll angles after the moving window was applied, the absolute difference and the resulting correction values.
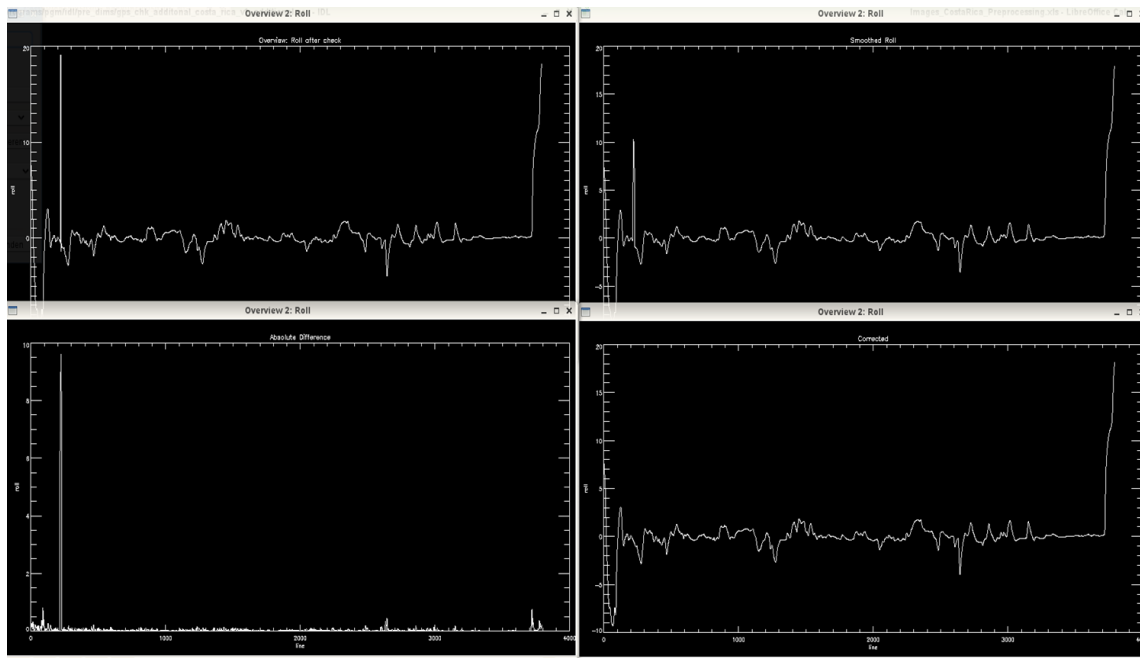
Figure 15: *GPS correction of roll angles for one flight line . Sensor detected changes in roll angles after check (top left), smoothed roll angles (top right), absolute difference between original and smoothed roll angles (bottom left), corrected angles (bottom right).*

The corrected gps-files where then used as an input for the parametric gecorrection with the software modul ORTHO (Muller et al., 2002) , which is part of the standardized processing chain. The position and view angle of the sensor system is reconstructed based on the gps data for each image pixel. As target coordinate reference system UTM, Zone 16 or 17 Nord, Word Geodetic System 1984 was choosen. Additional to the gps data, the boresight angle between sensor and IMU and an Digital Elevation Model (DEM) with 90m * 90m pixel size is used by ORTHO.

4.4 Wavelenght Calibration

The long data acquisition period of six weeks under various atmospheric and illumination conditions as well as uncertainties of the sensor calibration during the flight campaign made an verification of the center wavelength positions against atmospheric absorption bands necessary. In order to ensure the radiometric quality and consistency within and across the flight lines the centered wavelength shifts were assessed for 12 scenes, one at the start and one the end of the flight

campaign and 10 intermediate scenes in equal time intervals to also evaluate the stability of the wavelength shifts over the period of the flight campaign. Therefore, for each image multiple target spectra of homogeneous pixel from beach sands, large concrete areas and open soils were collected and served as inputs for the automatic assessment of the wavelength shifts in ATCOR using the SPECTRAL_ CAL program (Richter and Schläpfer, 2015). The underlying concept of the program is that surface reflectance spectra retrieved from narrow-band hyperspectral imagery often contain spikes and dips in spectral absorption regions of atmospheric gases, like oxygen absorption around 760 nm or water vapor absorption around 940 nm. In most cases these effects can be related to spectral miscalibration and can be removed by an appropriate shift of the center wavelengths of the channels . This is done by an optimization procedure that minimizes the deviation between the surface reflectance spectrum and the corresponding smoothed spectrum for each of the spectrometer of a sensor. In case of the HyMap sensor the wavelength shift is assessed for the four detectors in the VIS (450 - 890 nm), NIR (890 - 1350nm), SWIR1 (1400 - 1800nm) and SWIR2 (1950 - 2480nm) spectral range (Cocks et al., 1998). For further details on the method see Guanter et al. (2006).

The most severe wavelength shift of 5 nm can be observed at the first day in the NIR spectrometer, whereby all detectors show the highest variability in shifts at the first day of the flight campaign. The most stable sensor calibration is observed for the VIS spectral range. However, the subsequent flight days show are relatively stable sensor calibration for each detector varying in the VIS range between -3.3 to -3.75 nm, in the NIR between -0.3 to 1.25 nm (day 1 excluded), in the SWIR1 detector from 0.7 to 1.6 nm (day 1 excluded) and in the SWIR2 between 2 and 3.5 nm (day 1 excluded) (Fig. 16). The mean spectral shift of all days was calculated(except day one) and as a final step in spectral wavelength calibration the wavelength centers of each band where shifted according to these values.
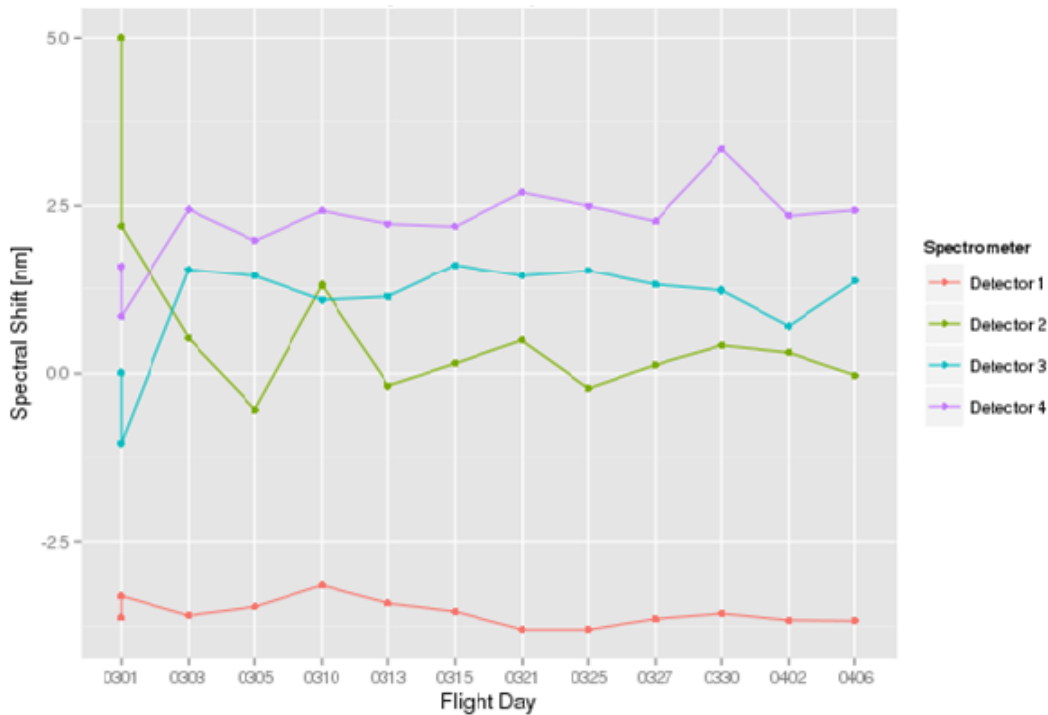
Figure 16: *Detected wavelength shift of the four HyMap spectrometers for single flight line s taken from different days equally distributed over the duration of the flight campaign.*

4.5   Evaluation of Wavelength Calibration

A comparison of the spectral properties of the original and reprocessed data proved the enhanced radiometric consistency within and across flight lines which is a crucial requirement when analyzing the data with respect to spectral characteristics, as it is done in spectral unmixing approaches. The analyzes of the spectra in overlap regions of adjacent image lines at the different spectral ranges show a very similar vertical offset, even though both, the new processed and the original processed data show a distinctive clay feature at 2200 nm the overall shape of the signatures are different (Fig. 17).

Figure 18 shows a high consistency for the shape of the signatures for the original and preprocessed dataset, whereas the clay absorption feature differs. The original data shows for the same pixel in two overlapping flight lines one signature holding the feature the other one doesn't. This severe radiometric inconsistency disappeared in the reprocessed data.

The last example given in Figure 18 shows that the radiometric consistency is
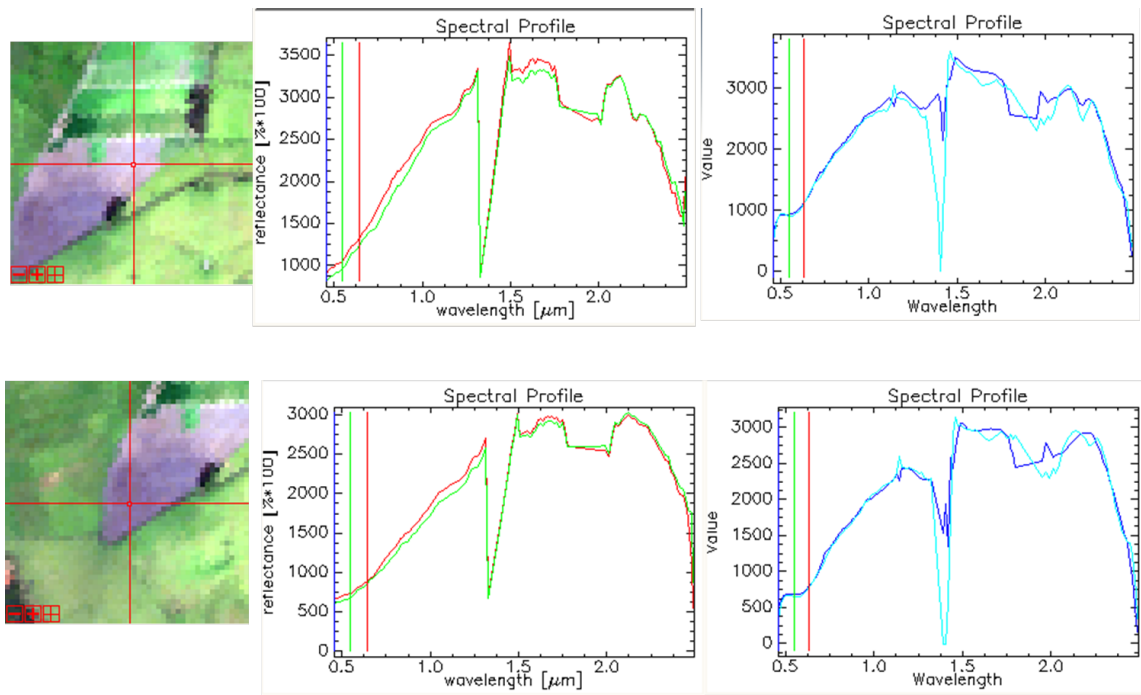
Figure 17: *Spectral signatures of soil for original and DLR processed data. Flight line subset showing pixel location (left), spectral signal of location for overlapping flight lines processed by the DLR (middle), spectral signal of location for overlapping flight lines processed by HyVista Corp. (right). Different colours indicate two different adjacent flight lines.*
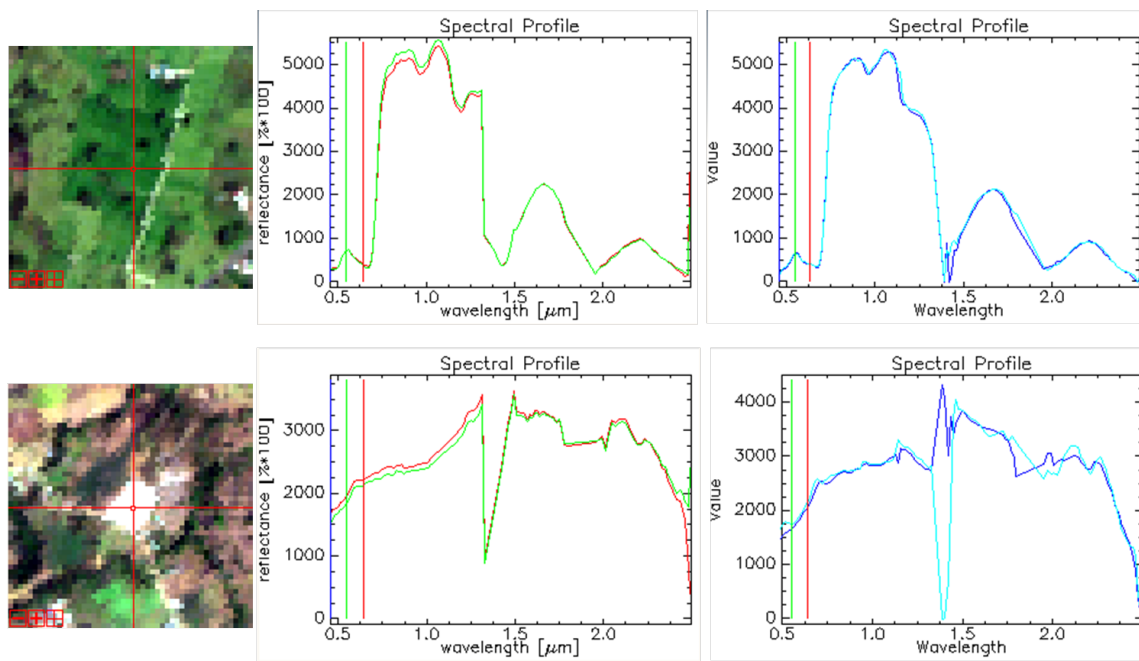
Figure 18: *Spectral signatures of PV and soil for original and DLR processed data. Flight line subset showing pixel location (left), spectral signal of location for overlapping flight lines processed by the DLR (middle), spectral signal of location for overlapping flight lines processed by HyVista Corp. (right). Different colours indicate two different adjacent flight lines.*
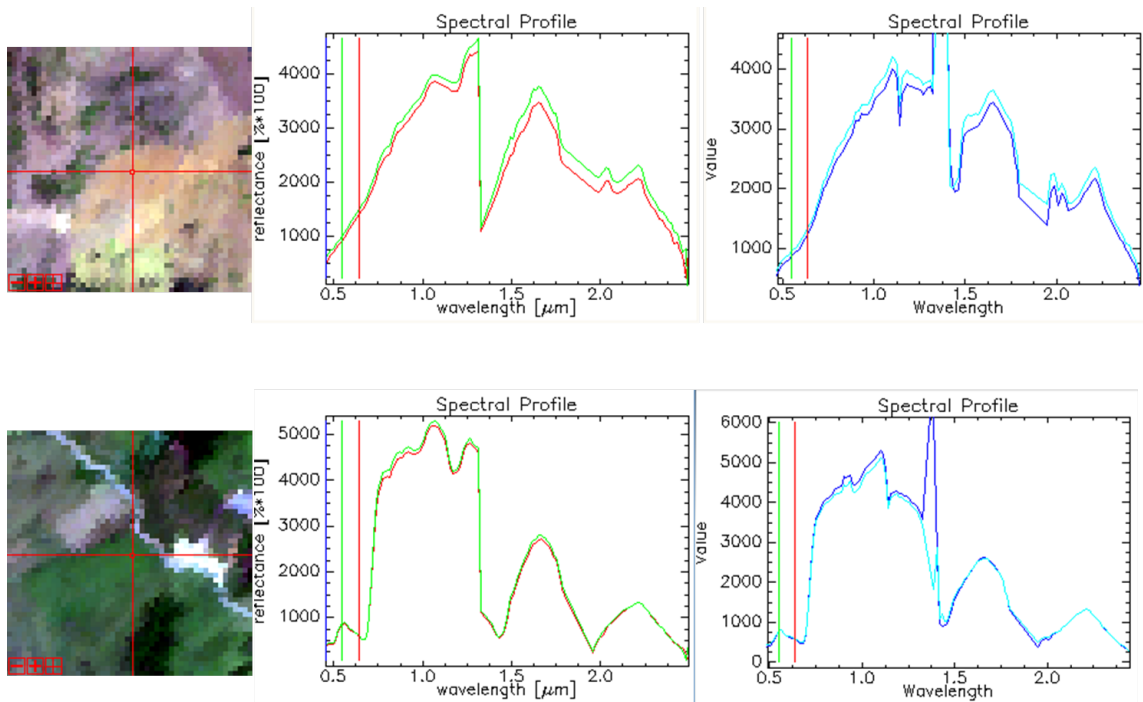
Figure 19: *Spectral signatures of PV and soil for original and DLR processed data. Flight line subset showing pixel location (left), spectral signal of location for overlapping flight lines processed by the DLR (middle), spectral signal of location for overlapping flight lines processed by HyVista Corp. (right). Different colours indicate two different adjacent flight lines.*

not constant when looking at different land cover types. The original data present consistent spectral signatures between flight lines for vegetation. Moreover, the variation between soil spectra appears less severe than in the example before. The radiometric consistency of the original data regarding shape of the signature and center wavelength position depends on the target but the level of consistency is also depended on the flight line. These radiometric inconsistencies have been overcome with the reprocessing of the data set. An significant enhancement of the radiometric quality have been observed for the flight lines which are of interest for the underlying study. Future studies using other flight lines of this data set will yield more vulnerable insight for the evaluation of wavelength calibration.

4.6   Atmospheric Correction

As part of the standardized processing chain atmospheric correction was conducted using ATCOR (Richter and Schläpfer, 2015). This physical approach is based on the MODTRAN radiative transfer model and calculates based on the radiance measured at the sensor the object-specific reflectance values at Earth's Surface. These include a correction of atmospheric absorption- and scattering processes, the consideration of viewing and irradiation geometry (absolute sun and sensor angle, path length through the atmosphere) as well as the correction of relief introduced illumination effects (terrain correction) and radiation effects of the neighborhood (adjacency effect). The latter correction step is based on the 90 * 90 m spatial resolution DEM. After performing some tests with different aerosol models the rural-maritime aerosol model was chosen for all scenes. The visibility was estimated for each image separately based on the values suggested by ATCOR and a subsequent manual verification in the SPECTRA-Modul of AT-COR. The scene dependent water columns were assessed based on the image data. As the terrain correction tends to be over- and underestimated in mountainous regions, additionally an empirical Bidirectional Reflectance Distribution Function (BRDF) correction was applied. Therefore a model for all surface components is calibrated and an object-specific correction is performed. Even though usually just a few pixel per flight line are effected by the BRDF correction (Bachmann, 2007), the correction has significant influence on the EM detection using SSEE. This is because uncorrected pixel show extreme spectra with very high albedo or saturation in different spectral bands.

4.7   Cosine BRDF Correction and Mosaic creation

For the creation of the mosaic covering the study area 7 flight lines have been used which were flown on 3 different days. The data acquisition on different days and time of the day show strong variations of illumination effects especially in the mountainous regions of the site. Therefor, previous to the mosaic creation a surface dependent BRDF effect correction was applied to the flight lines. This was accomplished by using a novel correction method, called BREFCOR, which is implemented within the framework of the ATCOR atmospheric correction process (Richter and Schläpfer, 2015; Schläpfer et al., 2015). BREFCOR, short for bidirectional reflectance distribution function (BRDF) effects correction, uses the Ross-Li sparse reciprocal BRDF model which is tuned based on a preclassification of the imagery and a continuous class index (Schläpfer et al., 2015). This is done
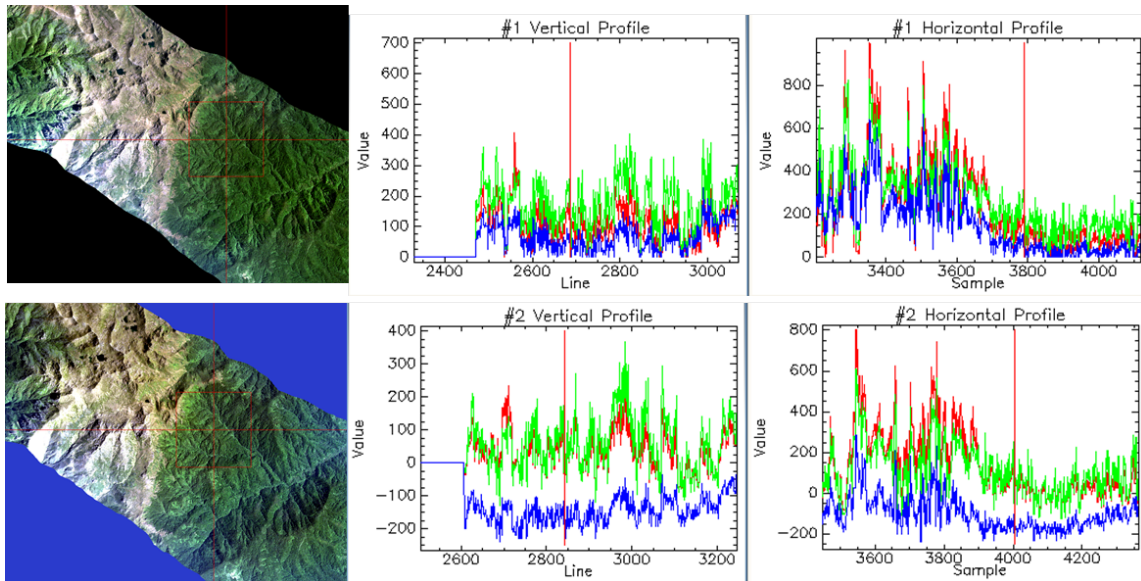
Figure 20: *Images with a BRDF correction applied (top), original processed data probably without BRDF correction (botton).*

by the calculation of an BRDF correction index (BCI), a unified continuous fuzzy classification index for all surface types from water to asphalt and concrete, soils, sparse vegetation and dense vegetation for the whole image (Schläpfer et al., 2015).

Before inputting the images into the BRDF correction process all scenes where masked based on the cloud cover and bad pixel images calculated within the quality assessment procedure. The cloud cover and bad pixel masks were manually extended to the regions where haze and cloud shadows occurred. These cloud and shadow free images served for the model calibration. First all scene-specific kernels are calculated, then all BIC for all images are derived and aggregated into 4 discrete classes, before in a last step the BRDF-model, e. g. the kernel weights, is calibrated for all classes and scenes and used to create a generic model by combining all image scenes. The calibrated model was then applied for the BRDF correction of the original or unmasked images. Figure 20 shows an example for one flight line without BRDF correction and the same image with the BRDF correction applied. The BRDF corrected images were merged using ENVI Classic.

Bibliography

Bachmann, M. (2007). *Automated estimation of ground cover fractions using MESMA unmixing.* PhD thesis, University of Wuerzburg, Germany.

Bachmann, M., Habermeyer, M., Holzwarth, S., Richter, R., and Müller, A. (2007). Including quality measures in an automated processing chain for airborne hyperspectral data. In *Proceedings 5th EARSel Workshop on Imaging Spectroscopy, Bruges, Belgium.*

Cocks, T., R., Jenssen, A., Stewart, I., Wilson, and Shields, T. (1998). The hymap airborne hyperspectral sensor: The system, calibration and performance. In Schaepman, M., Schläpfer, D., and Itten, K., editors, *Proc. 1st EARSeL Workshop on Imaging Spectroscopy*, pages 37–42, Zurich, EARSeL, Paris.

Guanter, L., Richter, R., and Moreno, J. (2006). Spectral calibration of hyperspectral imagery using atmospheric absorption features. *Applied Optics*, 45(10):2360–2370.

Habermeyer, M., Müller, A., Holzwarth, S., Richter, R., Müller, R., heinz Seitz, K., Seifert, P., and Strobl, P. (2005). Implementation of the automatic preocessing chain for ares.

Holzwarth, S., Bachmann, M., Freer, M., and Hofmann, M. (2011). Standards for airborne hyperspectral image data. In *EARSeL 7th SIG-Imaging Spectroscopy.*

Muller, R., Lehner, M., Muller, R., Reinartz, P., Schroeder, M., and Vollmer, B. (2002). A program for direct georeferencing of airborne and spaceborne line scanner images. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(1):148–153.

Richter, R. and Schläpfer, D. (2015). *Atmospheric/topographical correction for airborne imagery. ATCOR4 User Guide.* Wessling, Germany,.

Schläpfer, D., Richter, R., and Feingersh, T. (2015). Operational brdf effects correction for wide-field-of-view optical scanners (brefcor). *Geoscience and Remote Sensing, IEEE Transactions on*, 53(4):1855–1864.
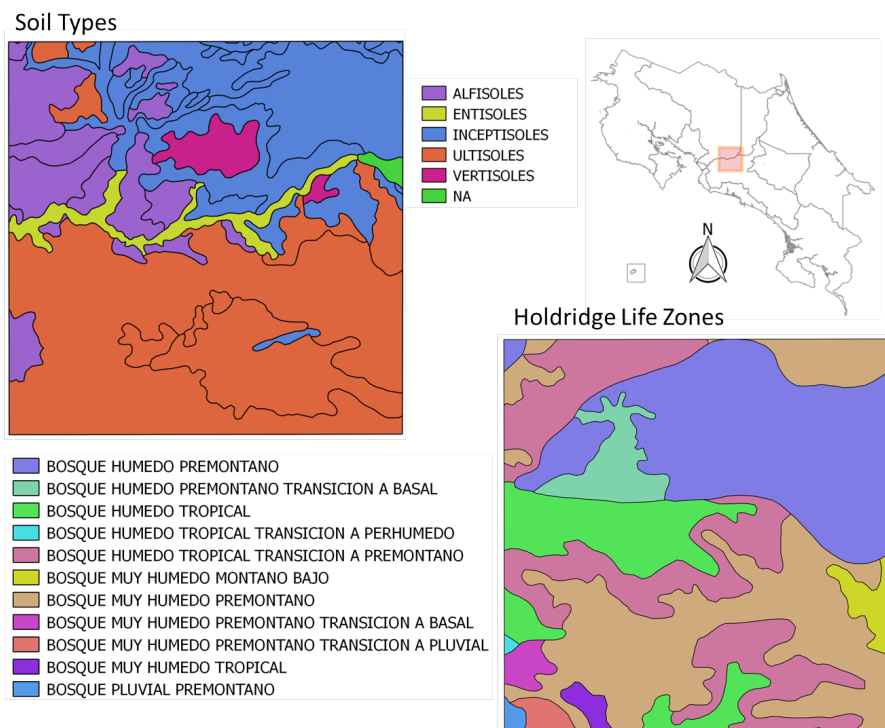
   ?

A | Auxiliary Data

Soil Types

Legend:
- ALFISOLES
- ENTISOLES
- INCEPTISOLES
- ULTISOLES
- VERTISOLES
- NA

Holdridge Life Zones

Legend:
- BOSQUE HUMEDO PREMONTANO
- BOSQUE HUMEDO PREMONTANO TRANSICION A BASAL
- BOSQUE HUMEDO TROPICAL
- BOSQUE HUMEDO TROPICAL TRANSICION A PERHUMEDO
- BOSQUE HUMEDO TROPICAL TRANSICION A PREMONTANO
- BOSQUE MUY HUMEDO MONTANO BAJO
- BOSQUE MUY HUMEDO PREMONTANO
- BOSQUE MUY HUMEDO PREMONTANO TRANSICION A BASAL
- BOSQUE MUY HUMEDO PREMONTANO TRANSICION A PLUVIAL
- BOSQUE MUY HUMEDO TROPICAL
- BOSQUE PLUVIAL PREMONTANO

Figure A: Soil types and Holdridge life zones of study area.

B | EnMAP Simulation

Figure B: *EnMAP imagery simulated using EeTeS-Software by GFZ.*
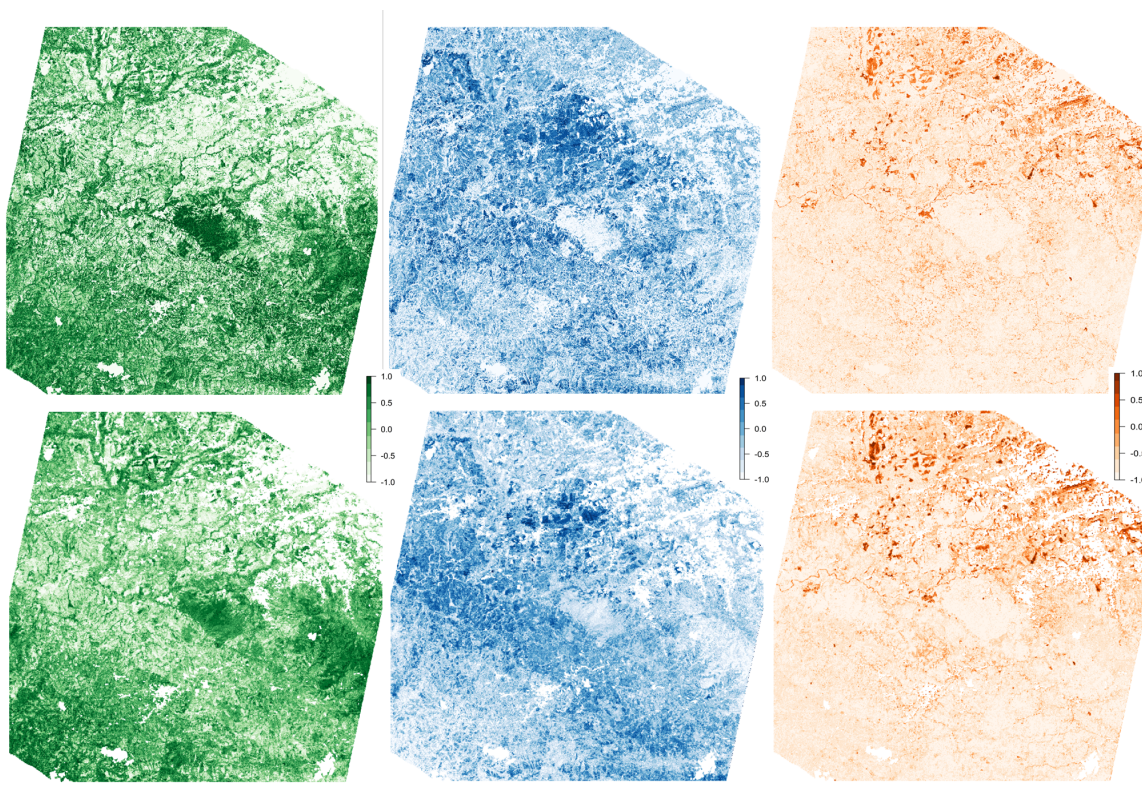
C | Fractional abundance maps

Figure C: *MESMA fractional cover for PV (green), NPV (blue) and Soil (orange)for HyMap (top) and EnMAP(bottom) White areas are masked urban pixels based on the DLR Global Urban Footprint.*
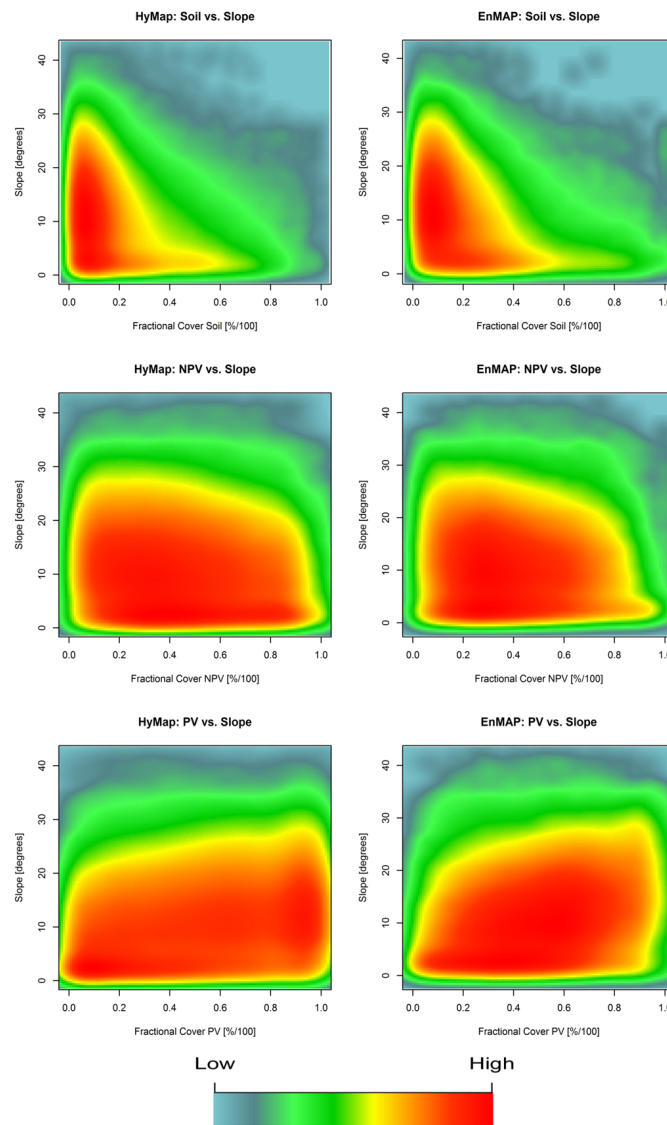
D | Density scatterplots

Figure D: Density scatter plots showing relationship between Soil, PV and NPV fractional cover and slope for HyMap (column 1) and En-MAP (column 2).
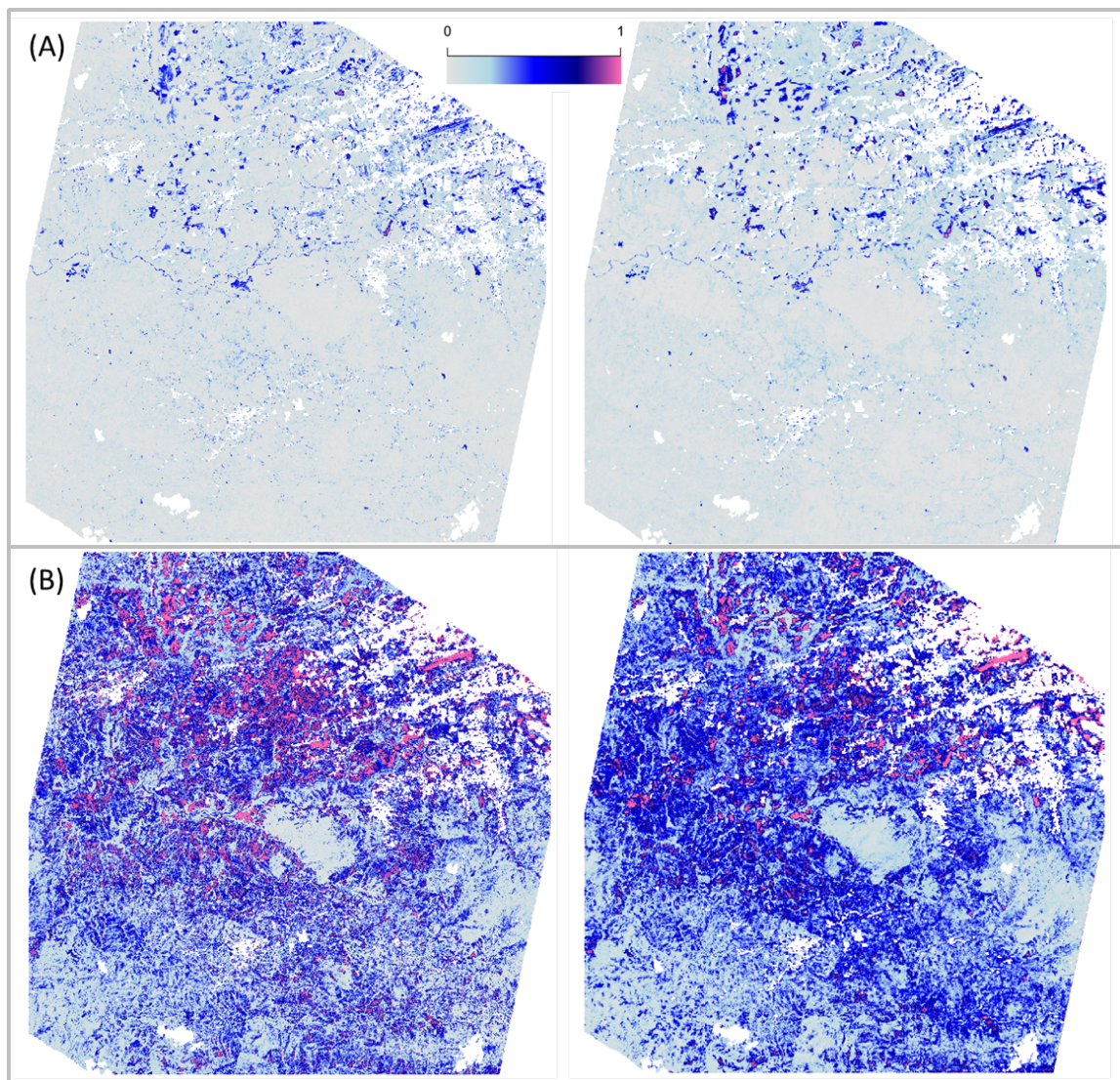
E | C factor

Figure E: *(A) C factor maps for case 1 and (B) case 2 for the full study area for HyMap (left) and EnMAP (right) imagery.*
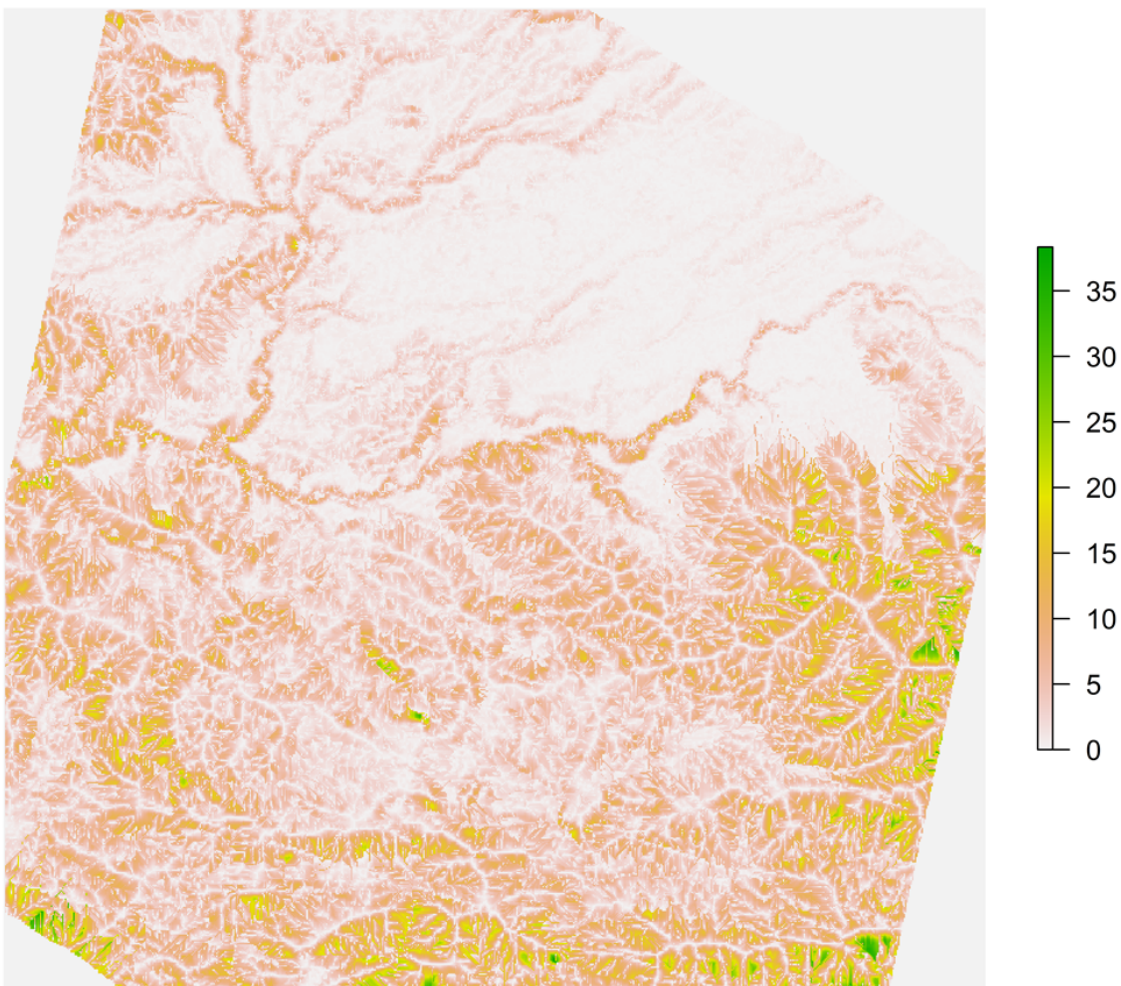
F | LS factor

Figure F: LS factor map for the full study area derived based on the 90 m DEM.

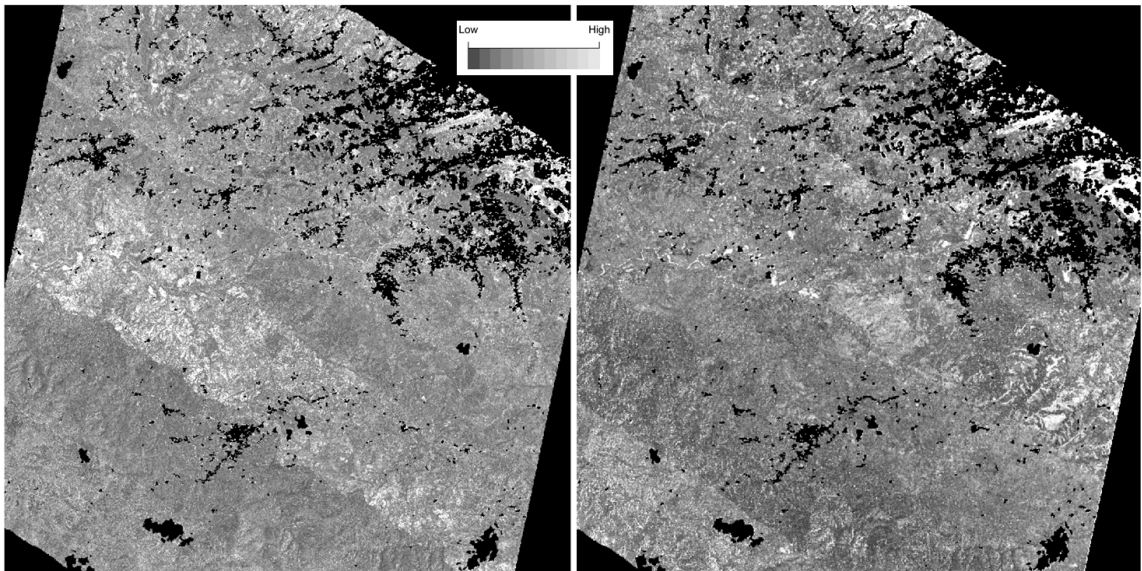# G | Modelling RMS error for MESMA results

Figure G: Modelling RMS error for fractional abundance maps derived with MESMA for HyMap (left) and EnMAP(right).

Acknowledgements

I wish to thank various people for their contribution to this thesis. First, special thanks should be given to my supervisors Dr. Martin Wegmann, Dr. Derek Rogge and Dr. Uta Heiden for their professional guidance, valuable support, and constructive recommendations during my thesis. I would also like to extend my thanks to Dr. Martin Bachmann for his help with various prepocessing challenges and with the application of his unmixing algorithm. Further, I wish to acknowledge the help of Dr. Derek Rogge for proofreading my thesis. Finally, I wish to thank my parents for their support and encouragement throughout my entire study time.

Declaration of originality

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices.