Centre for Geo-Information

Thesis Report GIRS-2007-13

DEVELOPMENT OF AN IMAGING SPECTROSCOPY BASED METHOD FOR MAPPING AND MONITORING PLANT FUNCTIONAL TYPES IN RIVER FLOODPLAINS

A case study using Spectral Mixture Analysis in the Millingerwaard

Lucía Sánchez Prieto

June 2007

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Development of an imaging spectroscopy based method for mapping and monitoring plant functional types in river floodplains

Application of Spectral Mixture Analysis

Lucía Sánchez Prieto

Registration number 80 03 20 725 100

Supervisors:

Dr. Ir. Lammert Kooistra
Drs. Harm M. Bartholomeus

A thesis submitted in partial fulfillment of the degree of Master of Science
at Wageningen University and Research Centre,
The Netherlands.

June, 2007
Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2007-13
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Acknowledgements

I would like to thank all the people that have contributed in some way or another to the development of this thesis.

First of all I would like to thank my supervisors Lammert Kooistra and Harm Bartholomeus for their time, guidance and critical advice. Special thanks to Lammert Kooistra for his instruction through the main concepts of this study, also for his dedication, patience, charm and optimism. Thanks to Harm Bartholomeus for his energy, technical support and for giving up part of his lunch breaks.

I would also like to thank Raul Zurita Milla for his help with the application of fully constrained unmixing and for his inspiring philosophical statements, and to Aldo Bergsma for his availability and dedication when trying to decipher the ins and outs of ESRI ArcGIS. I am also thankful to the students of prof. Karle Sykora and Ark Association for their field data facilitation, as well as to the staff involved into the HyEco campaign, whose work made this study achievable. Also, thanks to Gertjan Geerling for contributing with extra information and to Claudia Naom for her support with my English writing. Not to forget Willy ten Haaf and Antoinette Stoffers who solved efficiently all administrative and material details.

I would like to thank too all the people who helped me overcome all these months of hard work by making everyday something special in Wageningen: starting from my roommates Bast, Annemiek, Sander, Melle, Valerie and Aurelian, my salsa mates Janneke, Pentcho, Baro, Roelof, etc., my friends Walter, Sandra, Richy, Sancho and Stijn, the neighbors from 99, and so many others.

Last but not least, my deepest thanks to my family, who always sent me their love despite the distance and to Imanuel for his patience, love, care and existence.

This work is dedicated to my mother and to the memory of my father, who made it all possible.

Abstract

Due to flood risk increase in the Netherlands during the last decades, new river management strategies are being developed in areas such as river floodplains. As floodplains of the Rhine River are part of the National Ecological Network of the Netherlands, these strategies are meant to safeguard both flood protection as well as nature rehabilitation objectives. Nature rehabilitation implies that former agricultural land is transformed into natural areas, but vegetation is also an important component which influences the hydraulic roughness of the floodplains which increases peak discharge. In order to prevent catastrophic flood events river managers need to forecast vegetation dynamics. Dynamic Vegetation Models (DVM) are used to predict vegetation succession, but to initialize and validate these models, information about the spatial distribution of vegetation is required.

This research is focused on establishing a methodology for mapping and monitoring floodplain vegetation by the application of imaging spectroscopy techniques. Vegetation classes are defined according to the concept of Plant Functional Types (PFTs), because of its appropriateness when being used by DVMs. PFTs are defined in a wide variety of terms, but in this study they were defined as vegetation clusters that have a similar response to water flow impact. This response was measured by what is known as hydraulic resistance which was characterized by quantifying specific plant traits: height, density and flexibility. Since the heterogeneity of PFTs leads to intimate mixture of classes in a sub-pixel scale Spectral Mixture Analysis (SMA) was considered an appropriate technique. SMA was used to classify into PFTs imaging spectroscopy data acquired by HyMap and CASI in 2004 and 2001, respectively.

The methodology was developed using HyMap as base image. In this first approach, nine PFTs based on plant size were defined. Linear spectral unmixing was applied to the first 23 MNF bands from HyMap using ten endmembers (nine PFT classes and soil). These were extracted from the image based on ground truth knowledge derived from field data. Three methods were applied: unconstrained, semi-constrained and fully constrained linear unmixing. Finally, a temporal analysis was performed in which both HyMap (2004) and CASI (2001) were subject to previous methodology. In this second approach, five PFTs were redefined based on plant species and six endmembers were used.

Overall classification accuracy of HyMap improved in second approach with respect to first from 66 % to 40 %. Classification of woody PFTs showed better performance (68 %) than herbaceous PFTs (47 %). Dwarf shrubs (*Crataegus monogyna*) were classified with 57 % accuracy, shrubs (*Sambucus nigra*) with 81 % and pioneer trees (*Salix sp.*) with 64 % in the second approach. Performance of CASI image showed worse results with an overall accuracy of 21 %. The results from this research reveal that it is possible to map and monitor PFTs in river floodplains by applying Spectral Mixture Analysis to hyperspectral images.

Table of contents

List of figures

List of tables

Abbreviations

B	Broad leaf
CASI	Compact Airborne Spectrographic Imager
CCA	Convex Cone Analysis
D	Density
d	Stem diameter
DVM	Dynamic Vegetation Model
EVI	Environmental Vulnerability Index
f	Flexibility
h	Height
IR	Infrared
L	Long leaf
LAI	Leaf Area Index
LiDAR	Light Detection and Ranging
MESMA	Multiple Endmember Spectral Mixture Analysis
MLC	Maximum Likelihood Classifier
MNF	Minimum Noise Fraction
MODIS	Moderate Resolution Imaging Spectroradiometer
MWS	Medium Woody vegetation with Shrub structure
MWT	Medium Woody vegetation with Tree structure
N-D	Multidimensional
NDVI	Normalized Difference Vegetation Index
NIR	Near-Infrared
OSA	Overall Sub-pixel Accuracy
PCA	Principal Components Analysis
PFT	Plant Functional Type
PPI	Pixel Purity Index
RD	Rijksdriehoeksmeting
RMSE	Root Mean Square Error
RNB	Robust Non-woody vegetation with Broad leave
RNL	Robust Non-woody vegetation with Long leave
ROI	Region of Interest
RWS	Robust Woody vegetation with Shrub structure
RWSC	Robust Woody vegetation with Shrub structure with *Crataegus monogyna*
RWT	Robust Woody vegetation with Tree structure
S	Shrub
SNB	Soft Non-woody vegetation with Broad leave
SNL	Soft Non-woody vegetation with Long leave
SWS	Soft Woody vegetation with Shrub structure
SWT	Soft Woody vegetation with Tree structure
T	Tree
UTM	Universal Transverse Mercator
VCNNC	Vegetation Community based Neural Network Classifier

1 INTRODUCTION

1.1 Context

About half of the Netherlands' surface area is less than one meter above sea level and one third of it is actually below sea level. If there were no dikes or dunes, two thirds of the Netherlands would be flooded on a regular basis. Therefore flood risk[1] is a continuous threat and it is expected to increase in the coming decades along the river Rhine. The two main reasons for this are that climate change will cause a significant increase in the probability of extreme floods and that the potential damage of floods (the level of investments in areas at risk) is doubling every three decades (Hooijer, 2002). Moreover, the room available for improved flood risk management in the future is rapidly decreasing due to urbanization along the river banks (Hooijer, 2002). This affirmation promotes the potential damage of floods. And that is why flood risk management strategies should be developed in anticipation of higher peak discharges in the future (Hooijer, 2002).

New flood management strategies are intended to cover both flood protection as well as nature rehabilitation objectives (Baptist et al., 2004). Naturalized vegetation, in floodplains, plays a principal role in nature rehabilitation objectives, but it also determines flood evolvement because of its influence on hydraulic resistance (Anderson 2006). Due to management and restructuring in these areas, vegetation is subject to periodical changes. Accurate and up-to-date information on this dynamic vegetation is of vital importance to river managers because the maximum discharge capacity depends on it through its hydraulic resistance (Geerling et al., 2006). Especially of importance is to monitor shrubs since they experience fast development and are the vegetation type that show highest roughness (Geerling et al., 2005b).

Dynamic vegetation models are increasingly being used to forecast vegetation evolvement and hence can be used to support river management strategies. Many dynamic vegetation models use an approach based on plant functional types (PFTs) as classification units (Adams et al., 2004; Peng, 2000; Bonan et al., 2003; Epstein et al., 2001; Pausas, 2003; Jauffret, 2003; Pan, 2002). In contrast to ecotopes, traditionally used as classification units in river floodplains (Hartmann and Knotters, 2006), PFTs present characteristics that make them more adequate to be used by river manager

[1] Flood risk is a potential negative impact to an asset or value that may arise from a flooding event.

nowadays. Ecotopes are the smallest ecologically-distinct landscape features in a landscape. They are complex vegetation units with varying structural and functional plant types. PFTs are defined in a wide variety of terms; it depends on the scale (global, national, local, etc) or the field of application (plant physiology, ecology, etc). In ecological modeling they are defined as vegetation clusters that have similar response to environmental factors (Epstein at al., 2001; Gondard et al., 2003) making them more suitable than ecotopes to be applied into DVM.

Mapping and monitoring PFTs in river floodplains by yearly field work campaigns would take large amounts of time and would be too expensive. Currently, aerial photography has been used to discriminate vegetation types or ecotopes within the Dutch river floodplains. But this is a time consuming method with limited spectral possibilities that does not allow to document within-class variation of vegetation roughness (Straatsma, 2006). However, new digital remote sensing techniques, such as advanced multispectral or hyperspectral radiometers, detect multiple narrow spectral bands facilitating fine discrimination between different targets and it is less time consuming than aerial photography. It offers the possibility to record cyclic vegetation changes in distribution since there are images available in a periodical basis. Despite of it, little research has been done at present to derive PFTs from imaging spectrometry and no satisfactory methodology exists; the difficulty lies much in the heterogeneity of PFTs on the land cover (Sun et al., 2006).

There are different methods to approach classification of PFTs based on imaging spectroscopy techniques. Due to the high heterogeneity of the PFT classes, pixels covering more than one class are likely to be found. This will lead to pixels containing mixed spectral information of different PFTs. These mixed classes often results in poor classification accuracy when conventional algorithms such as the maximum likelihood classifier (MLC) are used (Borisova, 2005). It is possible to obtain better results if the mixed pixels are decomposed into different spectra. Spectral mixture analysis (SMA) is one of the most often used methods for handling the spectral mixture problem (Rosso et al., 2005; Li et al., 2005; Schmid et al., 2005; Borishova, 2005; Ustin et al., 1999). It is a physically based model that provides quantitative estimates of the distribution of materials within the image scene at a sub-pixel scale (Malenovsky et al., 2006; Borisova, 2005). Spectral mixture analysis (SMA) has been applied successfully before to derive vegetation classes (Rosso et al., 2005; Schmid et al., 2005; Li et al., 2005).

1.2 Objective

The objective of this thesis is to develop an imaging spectroscopy based method, using spatial mixture analysis (SMA), for mapping and monitoring plant functional types (PFTs) in river floodplains with especial emphasis on monitoring the process of shrub encroachment.

1.3 Research questions

The following research questions are addressed in this work:

1. How to discriminate plant functional types in a floodplain using imaging spectroscopy techniques?
2. Which methodology can be developed for mapping plant functional types using remote sensing techniques (with especial emphasis on shrubs)? Is Spectral Mixture Analysis (SMA) a good method?
3. Is it feasible to monitor shrub encroachment by comparing two images from two different sensors, while each of these images is analyzed by the previously developed methodology?

1.4 Outline of the report

Chapter 1 is an introduction to the study, first context, importance and problem definition are described and then objective and research questions are presented. Chapter 2 is a summary of the main contents gathered during the literature review. Chapter 3 consists of a description of all materials and methods used during this research, starting from the study area, then the data used and a detailed description of the methodology applied. In chapter 4 results are presented and in chapter 5, these results are discussed. Chapter 6 is a final conclusion of this study and chapter 7 gives some recommendations. Then, the list of all literature consulted for this study is shown in the bibliography in chapter 8. Finally, more detailed information can be found in the appendices in chapter 9, where tables, maps and figures are presented.

2 LITERATURE REVIEW

2.1 Floodplains

2.1.1 Floodplain characteristics

Flood affects periodically areas known as floodplains along rivers. A floodplain is a flat or nearly flat land adjacent to a stream or river that experiences occasional flooding (Wikipedia, 2006). Flooding process can vary in time, distribution and strength. Floodplains can support particularly rich ecosystems, both in productivity and diversity. These are termed riparian zones or systems. Riparian land is usually more fertile than the adjoining hillslopes, having better access to water and other nutrients. As a result, riparian zones support a higher density and diversity of plant species that tend to grow larger than their terrestrial counterparts. Riparian communities are subject to powerful fluvial forces, principally floods and the associated geomorphological processes of erosion and deposition. The action of these forces (both their magnitude and frequency) causes spatial diversification amongst plant assemblages, and a typical riparian zone is composed of an unstable mosaic of patches having distinct floristic compositions or existing at different successional stages (Anderson 2006). The variation of riparian plant species and associations with distance from the main river is linked to floodplain inundation duration and frequency, and to height above the water table that depend on the variability in floodplain topography (Florsheim & Mount, 2001). This results in a lateral structure of vegetation that can be divided into four functional ensembles on the basis of mean annual flood durations (Anderson 2006). Such a relationship is shown conceptually in figure 1, where the flow duration curve (right side) divides the lateral zonation (left side) into *dry*, *damp*, *marginal* and *aquatic zones*.

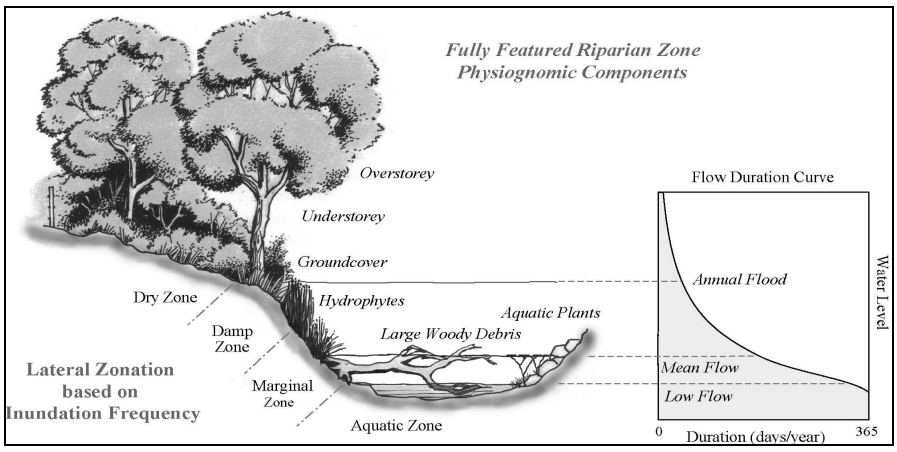


Figure 1: Lateral profile of a fully featured riparian zone including a lateral zonation based on inundation frequency (Anderson 2006)

2.1.2 Flooding events

Along the Rhine river tributaries, the main embankments, so called winter dikes, provide protection against flooding. Flood events will occur when the 'design discharge' of the embankment is exceeded (Baptist et al., 2004). Due to climate change the 'design discharge' needed is expected to increase. This will demand additional flood reduction measures. In the past these problems would have been addressed by heightening the dikes. But nowadays new measures are introduced which should result in a more sustainable approach to flood protection, such as lowering floodplains, removing groynes, removing hydraulic obstacles, widening floodplains by dike replacement and excavating secondary channels (Baptist et al., 2004). These measures provide opportunities for ecological rehabilitation. Floodplains in the river Rhine are part of the National Ecological Network of the Netherlands and many efforts are made to rehabilitate floodplain nature (Baptist et al., 2004).

One approach to reducing the risk of flooding is by increasing the hydrologic network capacity. Capacity is principally limited by the amount of resistance in the channels that make up the network. Resistance can be thought of the friction that slows down flow; greater resistance reduces the flow of water that a channel can convey before overflowing. High resistance is also caused by vegetation, which occupies space in the channel and presents obstacles that slow flow down. Plants have a pronounced effect on flow patterns and morphological development in rivers (van den Bosch, 2003). It also happens that vegetation is relatively easy to remove. Therefore, the removal of vegetation from the area in and near stream channels (the riparian zone) has been practiced as measure for flood mitigation (Anderson 2006). But a trade-off exists between the increase in flow depth caused by reduced channel capacity due to the presence of vegetation and the potential decrease in peak discharge of the flood wave that a densely vegetated channel network produces (Anderson 2006). However the work from Anderson (2006)[2] proves that the impact of riparian restoration is limited to changes in peak depth in the order of 10-20%, while a full cover of dense vegetation may increase overbank duration by between 40-100%. This demonstrates that

[2] This results were obtained from the application of the ROVER (a model to quantify vegetation roughness) to the upper Murrumbidgee catchment in New South Wales, Australia. This river supports large irrigated agricultural and irrigated pastoral developments and an increase in urban development. It is at an altitude above 370 m. Wetland forest as well as grassland are composed of different species with the ones we encounter in the Waal river in Netherlands, such as *Eucalyptus sp*., *Acacia sp*., *Dichanthium sp*., *Calotis sp*., etc.

vegetation can be used as main tool for river restoration and at the same time as tool for flood management.

2.1.3 Hydraulic roughness of vegetation

Floodplain vegetation relates to the discharge capacity through its hydraulic resistance. This influence depends very much on the different vegetation types, because each of them has a distinct hydraulic roughness (van den Bosch, 2003). The hydraulic resistance is the magnitude of the turbulence[3] resulting from water in laminar[4] regime passing through an obstacle. Its value affects the velocity of the regime[5] and depends on the shape and characteristics of the obstacle (Paruelo & Lauenroth, 1996; Laurent et al., 2004). Figure 2 shows the hydraulic resistance values for different ecotopes, which are present in the floodplain of the river Waal. Hydraulic resistance has been defined in terms of the *Nikuradse equivalent roughness*[6] (*K)* constant.

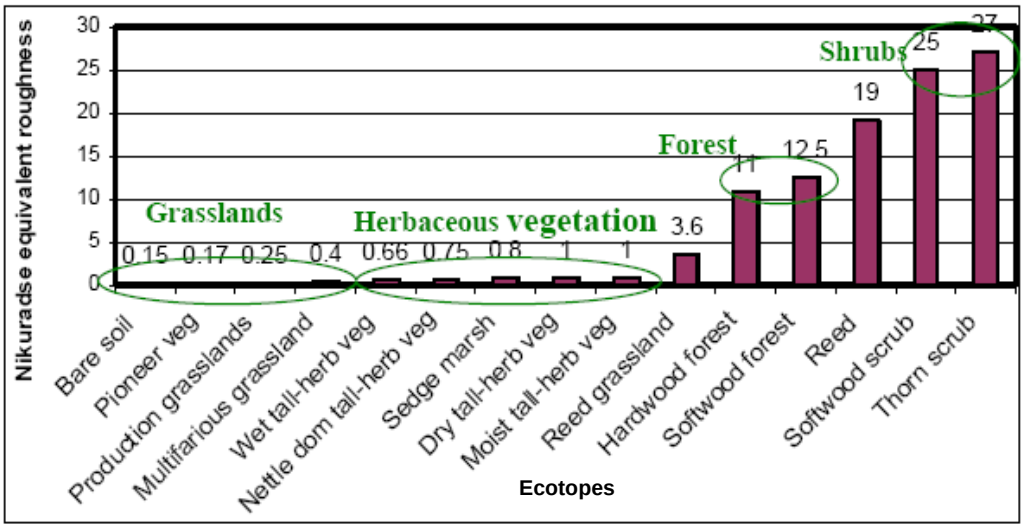


Figure 2: Hydraulic resistance by ecotopes in the river floodplain measured in terms of Nikuradse equivalent roughness (m) (Geerling et al., 2005b)

As can be observed, bushes have the highest hydraulic resistance of the presented ecotopes. This is because they have dense structure beginning at the soil level. Forested areas leave more freedom to the water as branching starts at a higher level. Herbaceous

[3] Turbulence is a flow regime characterized by chaotic, stochastic property changes.
[4] Laminar flow occurs when a fluid flows in parallel layers, with no disruption between the layers.
[5] The water regime of a freshwater ecosystem is the prevailing pattern of water flow over a given time. It refers to the duration and timing of flooding resulting from surface water (overland flow), precipitation, and ground water inflow.
[6] For a given surface roughness is often described in terms of a characteristic roughness height *k*. This characteristic height may be taken as the 'equivalent sandgrain roughness' height *ks* defined by Nikuradse (1933) (Schockling et al., 2006).

vegetation and grasslands are the ecotopes that are almost not obstacle for the water to flow (Suarez-Barranco, 2006). Although compared to agricultural use area (bare soil, production grasslands and multifarious grasslands), with *k* ranging from 0.15 to 0.40, herbaceous vegetation has increased resistance, 0.66 to 1.00.

All ecotopes present hydraulic resistance, hence all of them are of interest to derive flood behavior using dynamic vegetation models. However, a special attention will be given to shrubs, since this is the ecotope that mostly affects hydrological roughness of floodplains.

2.2 Plant Functional Types

2.2.1 Dynamic Vegetation Models

As vegetation is an important component of hydraulic roughness in river floodplains, dynamic modelling approaches have been developed to accurately simulate the influence of vegetation on flow and morphology of rivers (Baptist et al., 2005). Forecasting vegetation dynamics is a basic component to be considered by river managers when trying to predict flood risk. Predicting response of vegetation or simple biomass to a changing environment can be done by so called Dynamic Vegetation Models (DVMs). These models incorporate explicit representation of key ecological processes to forecast vegetation dynamics (Wamelink et al., 2003; Peng 2000).

Traditionally, ecotopes have been used to characterize vegetation types for the Dutch river floodplains. But using ecotopes in dynamic modelling has some limitations because one ecotope contains high heterogeneity regarding vegetation structure, response to environment, role in the ecosystem, etc. A new concept has gained terrain in literature (Epstein at al., 2001; Gondard et al., 2003; Paruelo & Lauenroth, 1996; Laurent et al., 2004; Kleyer, 2002): Plant Functional Type (PFT) that is more adequate as input for DVM. This will be described in the next chapter.

An often used DVM in the Netherlands to model vegetation at the regional to national scale is the model chain SMART2-SUMO2 (Wamelink et al., 2005). It is used to derive plant functional types' development (Schouwenberg et al., 2000; Wamelink et al., 2005). The SMART2 model simulates soil processes; it describes linked biotic and abiotic processes in the soil solution as well as in the solid phase. SMART2 delivers the nitrogen availability to SUMO2 as the sum of external N input and mineralization. SUMO2 is a process oriented model that provides estimates for nutrient uptake and litterfall (including roots, branches and stem wood) and models vegetation succession

and biomass production for time steps of one year; the biomass production in root, shoot and leaf is simulated for five functional types: (1) herbs and grasses; (2) dwarf shrubs; (3) shrubs; (4) pioneer tree; and (5) climax tree (Kooistra et al., 2006b; Wamelink et al., 2006; Wamelink, 2005). Results from Kooistra et al. (2006b) demonstrate that imaging spectrometer derived products can be used for validation and initialization of DVMs. PFTs mapping using remote sensing techniques can be useful when initializing or validating the application of this model chain.

2.2.2 Vegetation classes

As we have seen before, there is a need of knowing how flood risk evolves so that river managers can derive measures for flood control. To derive flood risk evolvement, river managers require monitoring the spatial distribution of vegetation and in a configuration which can be used by, the previously mentioned, dynamic vegetation models.

Monitoring vegetation is done on a vegetation classes' basis (Laurent et al., 2004; Geerling et al., 2006; Zhang et al., 1997). Vegetation may be classified using different criteria: plant species, plant communities, vegetation types, vegetation structures, ecotopes, plant functional types, etc. Plant species are the basic units of biological classification and they are assigned according to common ancestors (Wikipedia, 2007). Plant community is the vegetated portion of all the interacting organisms living together in a specific habitat (Wikipedia, 2007).

In current literature many authors use the term vegetation type (Lindstro & Jaenson, 2003; Nangendo et al., 2005; Song et al., 2005) but most of them do not outline its definition. Vegetation type responds to a classification, but this classification can be done according to different criteria. The criteria might be one or the combination of some of the next: climate pattern, plant habitat, phenology, growth form and/or dominant species. In North America, for instance, vegetation types are based on a combination of the criteria, while in Europe classification often relies much more heavily on floristic (species) composition alone. (Wikipedia, 2006)

Vegetation structure, however, refers to one specific trait of vegetation. It is characterized primarily by the horizontal and vertical distribution of plant biomass, particularly foliage biomass. And it is determined by an interacting combination of environmental, historical factors and species composition (Wikipedia, 2006).

According to AEM (Antropogenic Ecotope Mapping & classification system), ecotopes are the smallest ecologically-distinct landscape features in a landscape mapping and classification system (The Ecotope Mapping Working Group, 2005). As such, they represent relatively homogeneous, spatially-explicit landscape units that are useful for stratifying landscapes into ecologically distinct features for the measurement and mapping of landscape structure, function and change (e.g. softwood forest). They are identified using flexible criteria that depend on the specific application involved. These are criteria defined within a specific ecological mapping and classification system and are defined by the interaction of biotic and abiotic components, including vegetation, soils, hydrology, and other factors (The Ecotope Mapping Working Group, 2005). According to Geerling et al. (2005) an ecotope is a spatial unit of a certain extension (usually 0.25–1.5 ha), which is homogenous as to vegetation structure and the main abiotic factors on site. Geerling et al. (2005) define the following ecotopes in floodplains: Forest cultivated, Agriculture, Water (main channel), Bare soil, Pioneer vegetation, Grassland vegetation, Herbaceous vegetation, Bush (shrubs and trees <5 m): Open canopy (20–60% coverage) or Closed canopy (>60% coverage), Forest (>5 m): Open canopy (20–60% coverage) or Closed canopy (>60% coverage) and water (side channel).

2.2.3 Plant Functional Types

Classification of vegetation can also be done by so called Plant Functional Types (PFTs). The definition of PFT is currently receiving a lot of attention in ecology (Epstein at al., 2001; Gondard et al., 2003; Paruelo & Lauenroth, 1996; Laurent et al., 2004; Kleyer, 2002). Some methods have been developed for objective groupings of species into functional types (Chapin et al. 1996, Hodgson et al. 1999) but functional classification is context dependent and we should not expect to find a useful, universal classification into functional groups (Noble & Gitay, 1996). The purpose of developing a functional classification is to find some generally applicable simplification of the diversity of life while retaining information about the most important processes and interactions for the purpose in hand (Noble & Gitay, 1996). Therefore defining functional types is an operational procedure, often related to a specific scale.

Paruelo & Lauenroth (1996) define it as a group of species that share traits (morphological and physiological attributes) and play a similar role in an ecosystem. They use five PFT for grassland and shrubland in central North America; shrubs, C3

grasses, C4 grasses, succulents and forbs (nongraminoid herbs). Plant Functional types (PFT) provide a logical link between physiological and life history[7] strategies at the plant level and processes at the ecosystem level (Chapin 1993). Laurent et al. (2004) describe it as a group of plants defined by their height, leaf form, phenology and climatic requirements. They describe eleven PFTs at global scale: tropical broad-leaved evergreen, rain-green trees, temperate needle-leaved evergreen and summergreen trees, temperate broad-leaved evergreen and summer-green trees, boreal needle-leaved summer-green and evergreen trees, boreal broadleaved summer-green trees, C3 grasses and C4 grasses. Kleyer (2002) establishes that plant functional types (PFTs) are groups of plant species with similar plant traits and similar realized niches[8] with respect to multiple environmental factors. Traits such as plant architecture and height, seed bank longevity, attributes of the dispersal or seed weight and number are fundamental to relevant aspects of plant life history, e.g. storage effects, competitive ability or colonization (Kleyer, 2002). Niche patterns should result from syndromes, i.e. suites of attributes of traits or trait states, which represent functional adaptations of plants to the environment. When comparing syndromes along environmental gradients, general principles about functional plant-environment relations may be identified (Kleyer, 2002). Kleyer (2002) distinguishes PTFs according to attributes for dispersal, seed production, seed persistence, sapling, resource supply and disturbance intensity. Epstein at al. (2001) and Gondard et al. (2003) define functional types as a classification based on species responses to environmental factors. And Peng (2000) define PFT as a set of plant species (e.g. tropical evergreen broad-leaf rainforest trees) characterized by their physiognomic and morphological traits and response to climate.

Reading these previous definitions two basic patterns can be observed. On the one hand there is a difference in the term PFT depending on the scale. On a global or national scale (Laurent et al., 2004; Paruelo & Lauenroth, 1996) PFTs are defined by morphological and physiological attributes (plant height, leaf form, phenology, etc.) and climatic requirements. While in a local scale more specific traits are used to characterize PFTs, such as attributes for dispersal, seed production, seed persistence, sapling,

[7] Plant life-history traits are seed size, seed weight, plant height, leaf weight, leaf area, SLA (mean leaf area/dry weight) and number of seeds per plant (Austrheim & Eriksson, 2003).

[8] A niche is a term describing the relational position of a species or population in an ecosystem. It includes how a population responds to the abundance of its resources and enemies and how it affects those same factors. The abiotic or physical environment is also part of the niche. It may include descriptions of the organism's life history, habitat, and place in the food chain.

resource supply and disturbance intensity (Kleyer, 2002). However, for both scales, PFTs are considered to play a similar role in the ecosystem.

On the other hand definitions are markedly different if they come from the field of plant physiology (Paruelo & Lauenroth, 1996; Laurent et al., 2004; Kleyer, 2002) or the field of modelling (Epstein at al., 2001; Gondard et al., 2003). In modelling, definitions of PFTs are focused on their response to environmental factors and not so much on plant morphological or physiological traits.

PFTs are relevant to help understand and predict species distribution in present and future environments with changing disturbance regimes including climate (Gondard et al., 2003). Hence, many dynamic vegetation models use plant functional types (PFTs) as classification unit (Adams et al., 2004; Peng, 2000; Bonan et al., 2003; Epstein et al., 2001; Pausas, 2003; Jauffret, 2003; Pan, 2002). The advantages that PFTs have, when used to represent vegetation patches, compared to other biome-based classification systems, are three. First, PFTs are linked to leaf-level physiology measurements, more adequate for setting parameters in land models; since these models are expanding beyond their traditional biogeophysical roots to include biogeochemistry (photosynthesis and carbon cycle), they require specification of leaf-level and whole-plant physiological parameters. Second, PFTs allow modellers to represent land surface more accurately, since composition and structure of PFTs can be separately specified in a grid cell. And third, representing vegetation in terms of PFTs allows land models to better interface with ecosystem dynamic models, because the latter typically simulate vegetation change in terms of the abundance of PFTs (Sun et al., 2006). But multiple pathways or trajectories are possible in these models as a result of varying inputs or processes (Gondard et al., 2003) and a concerted effort to refine and standardize the concept of PFTs should be made (Adams et al., 2004).

In contrast to ecotopes, traditionally used as classification units in river floodplains (Hartmann and Knotters, 2006), PFTs present all these characteristics that make them more adequate to be used by river managers nowadays. Ecotopes are complex units which consist of vegetation with a varying structure and functional types. These complex units are simplified in such a way that apparently they represent spatial units of certain extension, homogenous as to vegetation structure and the main abiotic factors (e.g. shrubs, grassland, hardwood, etc.). PFTs are more realistic; they do not form such compact units, but instead they appear in a tangle disposition throughout the whole area and same PFTs are likely present in different ecotopes. Land cover in terms

of PFTs is characterized by a great heterogeneity. Figure 3 shows an example of ecotopes and PFTs in a landscape.

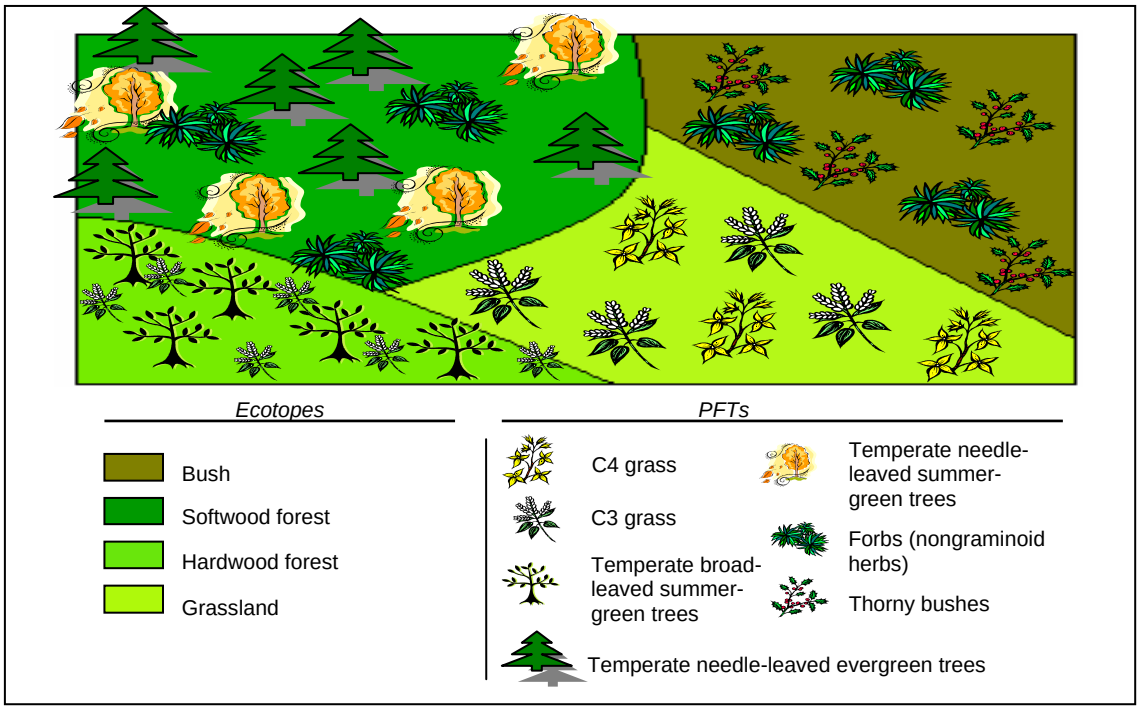


Figure 3: Example of relationship between ecotopes and plant functional types (PFT)

2.3 Spectral Mixture Analysis

2.3.1 Applicability

Evolvement of plant functional types in space is a relatively fast process which must be monitored regularly. Performing field work yearly would be too expensive and would take large amounts of time, therefore is not feasible. Currently, mainly aerial photograph has been used to discriminate vegetation types or ecotopes within the Dutch river floodplains. Time is required for each of the several steps of processing negatives and prints and interpretation of photos. In addition successful interpretation depends upon the skill and experiences of the individual interpreter, for instance the boundaries between different classes are settled manually which implies subjectivity. It works with either black-and-white, true color or color infrared -depending on the source-, which limits the spectral possibilities. Hence this is a time consuming method that moreover does not allow to document within-class variation of vegetation roughness (Straatsma, 2006).

However, new digital remote sensing techniques, such as advanced multispectral or hyperspectral radiometers, detect multiple narrow spectral bands facilitating fine

discrimination between different targets. This sensor's very high-spectral resolution facilitates fine discrimination between different targets based on their spectral response in each of the narrow bands. The advantage of space platform remote sensing is the broad scale, practical, repeatable collection of data which makes it applicable to real world problems (Pengra et al., 2006). Previous experiences have been successfully developed to monitor vegetation dynamics by remote sensing (Zhang et al., 1997; Rosso et al., 2005; Schmid et al., 2005; Geerling et al., 2006). Mapping marshland vegetation in California (Rosso et al., 2005) was carried out in 2005. Marshes were mapped using Spectral Mixture Analysis (SMA) and multiple endmember spectral mixture analysis (MESMA). These two methods were compared to investigate their appropriateness to characterize marshes species. They suggest that MESMA could be more successful if (a) four or five endmember models are tested and (b) more than one endmember per plant species or class is included, to accommodate for variations in spectral characteristics of each component across the scene. A multisensor approach was investigated by Schmid et al. (2005) to determine changes over time of wetland characteristics in semiarid environments in Spain. Linear spectral unmixing was carried out successfully in a highly dynamic and anthropogenic-affected wetland area.

Floodplain vegetation has been classified using remote sensing techniques. Geerling et al. (2006) proved that using data fusion of spectral (CASI) and LIDAR data gave better results than results from separate data, specially for those vegetation classes which are important to predict hydraulic roughness, i.e. bush and forest (Geerling et al., 2006). Their study, in the Waal River, consisted of a five-class set of vegetation types, which was considered to be the minimum set to estimate hydraulic resistance for river management purposes. The paper from Straatsma (2006) described a new method to derive hydrodynamic roughness of floodplain surface. He used a combination of multispectral data (CASI), and airborne laser scanning data (ALS). Roughness in vegetated areas was assumed to be a function of vegetation height and density. The result was three vegetation maps with distribution of herbaceous height, herbaceous density and forest density.

To date research on how to utilize satellite observations to map PFTs remains limited. It appears that at present no satisfactory methodology exists for the extraction of PFTs from satellite observations (Sun et al., 2006). The difficulty lies much in the heterogeneity of PFTs on the land cover. The Moderate Resolution Imaging Spectroradiometer (MODIS) Land Team is producing a global PFT map for use in the

Community Land Model (CLM). This MODIS PFT is the only global PFT data set available now, but errors and uncertainties of this data set are noticeable enough to compromise credibility of global change research (Sun et al., 2006).

There are different methods to approach classification of PFTs based on spectroscopic techniques. Due to the high heterogeneity of the PFT classes, pixels covering more than one class are likely to be found. This will lead to pixels containing mixed spectral information of different PFTs. These mixed classes often results in poor classification accuracy when conventional algorithms such as the maximum likelihood classifier (MLC) are used (Borisova, 2005). Traditional classification approaches assign only one possible value or category per pixel, in most cases, misrepresenting some components or oversimplifying mixtures of components (Rosso et al., 2005). It is possible to obtain better results if the mixed pixels are decomposed into different spectra. In order to solve the mixed pixel problem, scientists have developed different models to unmix the pixels into different proportions of the endmembers (Mishev, 1991; Ishoku, 1996). Spectral mixture analysis (SMA) is one of the most often used methods for handling the spectral mixture problem (Rosso et al., 2005; Li et al., 2005; Schmid et al., 2005; Borishova, 2005; Ustin et al., 1999).

The fact that classes can be mapped and quantitatively represented as a separate entity in an area that otherwise is classified as one class by traditional mapping approaches, underlines the potential of SMA to follow changes in vegetation cover. Because the same endmembers can be used to analyze a time sequence, SMA has the capability to estimate changes in abundance. The potential to estimate the spatial distribution and abundance of a species, rather than thematic classes, has great value in monitoring a biological invasion, because changes can be detected and quantified (Rosso et al., 2005).

Monitoring vegetation with remote sensing techniques is more complex than only mapping. Monitoring implies to keep track of vegetation changes in a spatial and temporal scale. Commonly different sensors will be used for it. Therefore, a methodology to monitor vegetation has to be adapted to different sensors while dealing with different spatial resolution, spectral bands, image extension, geo-reference, radiometric and geometric calibrations, etc. Schmid et al. (2005) demonstrated the complementary use of remotely sensed data together with the SMA technique to incorporate results derived from hyperspectral data into a temporal series of multispectral data. They proved the capacity to use endmembers derived from

hyperspectral information in the analysis of existing multispectral data from different sensor.

2.3.2 Description

SMA is a physically based model that provides quantitative estimates of the distribution of materials within the image scene (Malenovsky et al., 2006). It assumes that the spectrum measured by a sensor is a linear combination of the spectra of all components within the pixel (Borisova, 2005; Rosso et al., 2005). In the process of 'unmixing', SMA estimates the fit of selected endmembers to the observed value of a pixel in order to estimate its composition. As a result, each pixel carries information about the predicted abundance of each of the endmembers (Rosso et al., 2005).

The mathematical notation of the linear spectral unmixing procedure is:

$$R_i = \sum_{j=1}^{j=m} R_{ij} f_{ij} + e_i \qquad i = 1...n \qquad j = 1...m$$

Where R_i is the reflectance value of the pixel in band i, R_{ji} is the reflectance value of endmember j in band i, f_{ji} is the fraction of endmember j in band i, e_i is the residual error per band i, n is the number of bands and m is the number of endmembers.

In order to guarantee a physical interpretation of the results, two constraints are usually added to the unmixing problem. The first one makes sure that the fractions will be positive and below 1 while the second constraint ensures that the sum of all the recovered fractions adds to unity (Zurita-Milla et al, 2005):

$$\sum_{j=1}^{j=m} f_j = 1 \quad \text{and} \quad 0 \le f_{ji} \le 1$$

When these two constraints are not taken into account, the process is denominated unconstrained linear unmixing. When only the first constraint is considered it is a semi-constrained linear unmixing. This constraint may be given more or less weight by adding this equation to the whole set of function as many times as required. And finally when both constraints are added to the calculation, the method is called fully constrained linear unmixing.

The root mean square error (RMSE) is the parameter used to know the fit of the model and it is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{i=n} e_i}{n}}$$

The higher the maximum RMSE value, the worse the model fit in terms of determining the distribution of the abundance values (Schmid et al., 2005). The RMS image is an important indicator of surface features that were not included in the linear spectral unmixing (Schmid et al., 2005).

2.3.3 Endmember selection

The selection of endmembers can be performed in two ways (Plaza et al., 2004): 1) by deriving them directly from the image (image endmembers); or 2) from field or laboratory spectra of known target materials (library endmembers). These two methods are discussed below.

1) *Deriving endmembers directly from the image (image endmembers)*

The advantage of this method is that selected endmembers are under similar atmospheric conditions and spectral/radiometric biases with respect to the image (Malenovsky et al., 2006). In the case of a temporal analysis it is still an advantage provided that the endmembers are selected from the same image where the SMA is going to be performed. In case the endmembers come from a different image, inaccuracies originated from those biases must be considered.

The endmembers can be derived in two different ways:

- *Directly from the purest pixels in the image***.** Pure pixels are areas in the image in which unique endmembers are represented. This technique requires reliable field information in order to locate precisely the pure pixels. The area selected can be a single pixel or the average of several pixels in a homogeneous area. But for low spatial resolution this can be difficult to achieve since vegetation distribution can be heterogeneous at the pixel scale, thus leading to highly-mixed pixels (Wang et al., 2006).
- *With automated procedures***.** The endmembers are found with automatic methods such as PPI (pixel purity index), N-FINDR or Convex Cone Analysis (CCA) (Plaza et al., 2004). These classic approaches select endmembers based on the search for spectral convexities in the N-D[9] space. They can be partial or fully automated. Another method, IEA, is based on an iterative process in which those pixels that reduce the error obtained in constrained spectral unmixing operations are used as endmembers.

[9] Multidimensional.

2) *Deriving endmembers from field or laboratory spectra of known target materials (library endmembers)*

Spectral libraries are collections of laboratory spectra or spectra measured in the field (Malenovsky et al., 2006; Kooistra et al., 2005). They can be collected for different purposes and are measured using various types of instruments (portable non-imaging spectrometers) of often unknown quality or precision (Malenovsky et al., 2006). Library endmembers are not recommendable because of possible wavelength shifts, unreliable instrument calibration or significant degrees of noise, which are not uncommon to these data (Malenovsky et al., 2006). Besides field spectra are not easy to scale up to match spectra measured by the sensor because it cannot adequately capture the multiple-scattering environment of a canopy (Powell et al., 2006).

Another consideration is the number of endmembers, which is limited by the dimensionality of the data. The maximum amount of endmembers cannot be more than the number of bands minus the number of constraints. The constraint is one (Malenovsky et al., 2006) in the case of semi-constrained unmixing and three (second constraint accounts for two) in the case of fully-constrained unmixing.

2.3.4 MNF transformation

While hyperspectral imagery are capable of providing a continuous spectrum ranging from 0.44 to 2.48 microns (in the case of HyMap) for a given pixel, it also generates a vast amount of data required for processing and analysis. Due to the nature of hyperspectral imagery (i.e. narrow wavebands), much of the data is redundant (Green et al., 1988). A minimum noise fraction (MNF) transformation is used to reduce the dimensionality of the hyperspectral data by segregating the noise in the data (Green et al., 1988). The MNF transform is a linear transformation which is essentially two cascaded Principal Components Analysis (PCA) transformations. The first transformation decorrelates and rescales the noise in the data. This results in transformed data in which the noise has unit variance and no band to band correlations. The second transformation is a standard PCA of the noise-whitened data (Green et al., 1988).

Before applying spectral unmixing to the image, it is recommended to compress data with MNF transformation in order to improve the subsequent spectral processing results. It reduces redundancy, minimizes the influence of systematic sensor noise

during image analysis and helps the endmembers to be orthogonal and uncorrelated. (Pengra et al., 2006)

2.3.5 Validation

Validation of results may be done using three different approaches (Kooistra et al., 2007). First an analysis of the spatial continuous map for the RMSE is used as indicator of surface features that are not included as endmembers in the linear spectral unmixing classification. Secondly the validation focuses on the comparison between modeled and observed coverage. This comparison is done by the representation of modeled vs. observed by means of scatter plots. This gives a general trend of the model accuracy. In addition, it is done by overlaying field data with the abundance maps and easing the visualization of trends. The third method, only applicable to fully constrained linear unmixing, is a sub-pixel accuracy assessment called OSA (overall sub-pixel accuracy), which gives an estimation of classification accuracy by adding the values of the diagonal from the confusion matrix[10] per pixel. Assuming that the estimated fractions are correctly positioned within the pixel, the OSA is computed as follows (Zurita-Milla at al., 2005):

$$OSA = \frac{\sum_{j=1}^{j=n} \min\left\{f_j^{\,p}, f_j^{\,o}\right\}}{\sum_{j=1}^{j=n} f_j^{\,o}} = \sum_{j=1}^{j=n} \min\left\{f_j^{\,p}, f_j^{\,o}\right\} \qquad \text{j=1,2,…,n}$$

In this equation f represent fraction, p is the predicted and o the observed, while j are the different endmembers. In this expression the OSA is calculated per pixel. The minimum fractions of the observed vs. predicted per pixel are added for all classes and weighted to the total predicted fraction, which is 1 as consequence of the first constraint.

[10] A confusion matrix is a plot used to evaluate the performance of a classifier during supervised learning. Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. Values in the diagonal of the matrix count the correct predictions.

3 MATERIALS AND METHODS

3.1 Study area

This thesis work will focus on the study area Millingerwaard. The Millingerwaard is a floodplain that is located along the Waal, the main branch of the river Rhine in the Netherlands. It covers around 700 ha and it is part of the Gelderse Poort, a nature reserve with a surface of 6700 ha located in between the cities Arnhem, Nijmegen and Emmerich, near the Dutch-German border. Before 1990 the main function of the floodplain was agricultural production but in the period of 1990-1993 it was gradually changed into nature. The floodplain was allowed to undergo natural succession and a regime of natural grazing by cattle and horses in low densities was introduced (Kooistra et al., 2006b).

The current vegetation of the floodplain consists of mixed patches and ecotones, i.e., transitions between communities with a dominance of grass, herbaceous vegetation, shrub, and a large softwood forest (Kooistra et al., 2006b). Softwood forest in the Millingerwaard is dominated by willow trees (*Salix fragilis* and *Salix alba*). The forest canopy has an open structure with dense undergrowth (*Urtica dioica*, *Arctium lappa*, *Galium aparine*). Some open water bodies exist as well as high ground water levels due to the low elevation. The non-forest vegetation is characterized by a heterogeneous patchy structure of different vegetation succession stages. Dominant species are *Urtica dioica*, *Calamagrostis epigejos*, and *Rubus caesius* (Kooistra et al., 2006). During high floods, the whole floodplain except for the higher parts of the river dunes is flooded.

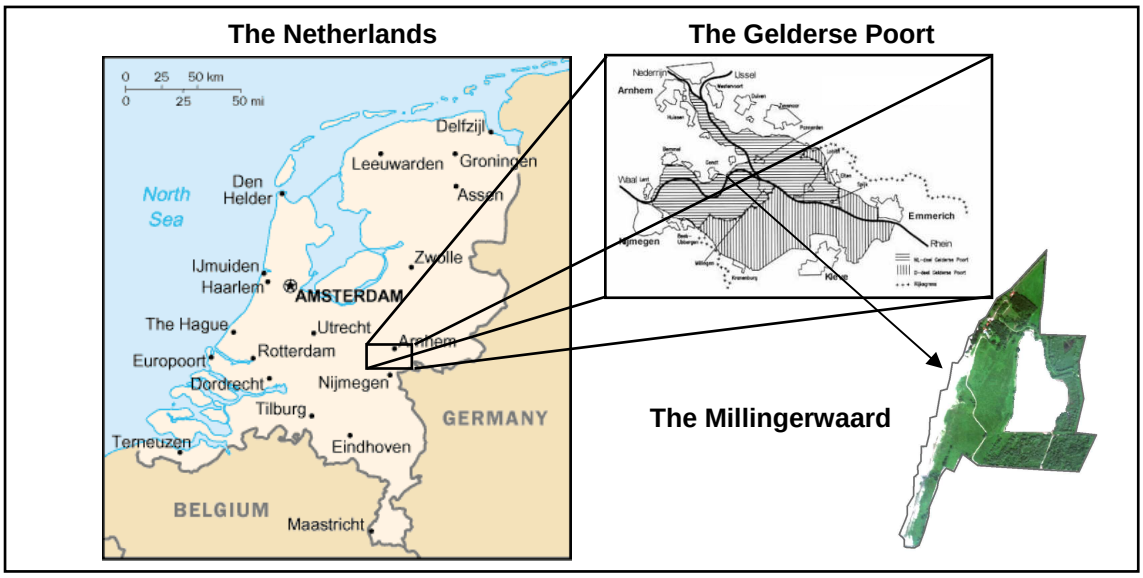


Figure 4: Study area

3.2 Data

Several fieldwork campaigns have been carried out in the Millingerwaard that resulted in a wide range of data sets available for the development of this work. In September 2001 a hyperspectral image was acquired by the Compact Airborne Spectrographic Imager (CASI) operated by NERC (UK) and financed by the Ministry of Traffic and Public Water. In summer 2002 fieldwork was performed by students of prof. Karle Sykora (Plant Ecology, Wageningen University) to gather vegetation data. The Geo-Information Centre of Wageningen University (CGI) developed an airborne imaging spectroscopy campaign in the Millingerwaard in 2004 (HyEco'04 campaign) where field data was gathered and hyperspectral images were acquired using the HyMap sensor. Another campaign was held by the CGI in 2005 (HyEco'05 campaign) where field data was also gathered. During the years 2005-2006, the Ark Association (www.arknature.nl) made also an inventory of woody species in the floodplain.

3.2.1 Image data

The images that were used in this research are HyMap image (2004) and CASI image (2001). HyMap image was acquired on July 28th 2004 by the HyMap sensor. This sensor uses 126 spectral bands (30 VIR, 32 NIR, 32 SWIR1 and 32 SWIR2) ranging from 400 to 2500 nm (bandwidth 15 - 20 nm) and a spatial resolution of 5m (Kooistra et al., 2005). CASI image was acquired on August 15th 2001 using the Compact Airborne Spectrographic Imager (CASI) sensor. Spectral information is measured in 10 bands (1 blue, 1green, 4 red and 4 NIR) (Verrelst 2004).

In Table 1 a summary of the characteristics of both images is presented. A more detailed table can be found in Appendix 1.

Table 1: Image data characteristics

Image	*Year*	*Number of bands*	*Projection System*	*Data type*	*Sampling unit*
HyMap	2004	126 bands	UTM	ENVI standard	Cells of 5 x 5m
CASI	2001	16 bands	RD Dutch coordinate system	ERDAS IMAGINE	Cells of 2 x 2m

3.2.2 Field data

Four different sources of field data were used in this study: Sykora group, the two HyEco campaigns from the CGI and Ark Association. The available field data and its characteristics are summarized in Table 2.

Table 2: Field data characteristics

Source	*Year*	*Field information*	*Vegetation type*	*Projection system*	*Data format*	*Sampling unit*
Sykora group	2002	316 releveés: location, community name & height range. 88 releveés: location & height range	All VTs	RD Dutch coordinate system	Shapefile (points)	Area of 2 x 2 m
HyEco'04 campaign	2004	21 releveés: location, plant community & sps abundance per community	Herbaceous	RD and UTM	Shapefile (points)	Area of 2 x 2 m
HyEco'05 campaign	2005	18 releveés: location, height per VT, plant community & genus abundance per community	All VTs	RD and UTM	Shapefile (points)	Area of 5 x 5 m
Ark Association	2005-2006	3896 individual plants: location, date, elevation, sp name, height, volume, stem thickness, appearance, full grown, damage, eaten by cow, couple, dear, beaver, rabbit, mouse, litter, dead or alive	Shrubs and trees	RD Dutch coordinate system	Shapefile (points)	Individual plants

3.3 Methodology

3.3.1 Overview of methodology

The methodology was developed using exclusively HyMap (2004) as base image. At the end, a temporal analysis was performed (Section 3.3.5) where both images (HyMap (2004) and CASI (2001)) were classified. The aim of this section was to compare the PFTs distribution of both images to derive information about vegetation dynamics in a time step of three years. The methodological approach developed for HyMap was used to classify both images. But such methodology was subject to some modifications: PFT classes, number of endmembers and images extension. On the one hand, the characteristics of CASI added new conditions and limitations to previous defined steps, such as number of endmembers and extension of the area. These new limitations had to be taken into account into a second HyMap classification, so that results could be comparable. On the other hand some lessons learned from preceding results were considered in this second approach in order to apply for improvements.

The methodology consisted first of a conceptual model definition of PFTs (Section 3.3.2). Second, images and field data were object of a series of preprocessing steps (Section 3.3.3). HyMap image was masked and resized. CASI image was subject to a projection transformation, an atmospheric correction, masking and resizing. Field data was adapted to the PFTs model and to the HyMap image characteristics. This

implied a classification into PFTs, a format conversion in some cases, a transformation of the projection system and the adaptation of the sampling units. Third, spectral mixture analysis was applied to the image HyMap (Section 3.3.4) by a selection of endmembers, a MNF transformation, a linear spectral analysis and a validation process. And finally the temporal analysis was performed (Section 3.3.5).

3.3.2 Conceptual model of PFTs classification

3.3.2.1 Conceptual model

As was mentioned in section 2.2, PFTs are defined in literature in a wide variety of terms. In this work, more in accordance with ecological modeling at local scale, PFTs were defined as:

"*Vegetation clusters that have a similar response to water flow impact*"

This response, that originates a great proportion of hydraulic roughness in floodplains, is measured by the hydraulic resistance. Hydraulic resistance can be characterized by the evaluation of specific plant traits under the assumption that similar response to water flow is induced by similar plant traits. Therefore by measuring specific plant traits, classification of vegetation into PFTs can be carried out (Figure 5).

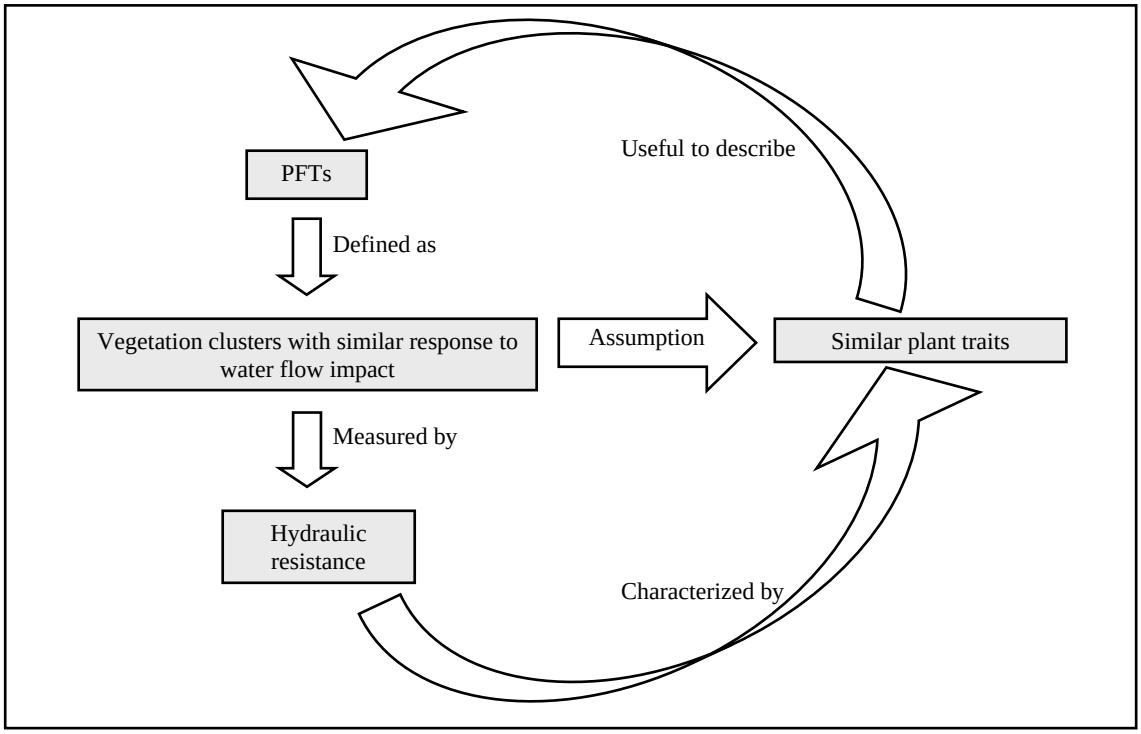


Figure 5: Conceptual model of PFTs. Relationship between PFTs, hydraulic resistance and plant traits

The selection of plant traits to establish this classification was done considering mainly, relevance to hydraulic resistance, but existence of field data was also taken into account as well as the separability of spectral signatures:

- *Relevance to hydraulic resistance.* Traits that have an effect on hydraulic resistance are density of foliage, branch structure, vertical extent, stem diameter and rigidity (Anderson et al., 2006; Naden et al., 2006).
- *Existence of field data for training and validation.* The existing data varies depending on the different datasets. From Ark data -which consist of woody vegetation- the information that we can derive is species, height, stem diameter and crown projection per individual plant. The Sykora dataset -which consists mainly of herbaceous vegetation-, has information on a plot basis of 2x2 m and it consists of existing species and Braun Blanquet[11] index per species. Field information from the HyEco campaign consists of the number of individuals per existing specie in plots of 2x2 m for 2004 and plots of 5x5 m for 2005.
- *Difference between and within spectral signatures*. The chosen traits must have spectral significance so that it is possible to classify the PFTs using spectral unmixing techniques. This aspect was initially taken into account only for herbaceous PFT classes, for woody vegetation it was evaluated in a different section (3.3.4.1). In this section a final definition of the endmembers is proposed.

According to relevance of hydraulic resistance and data availability, the selected traits were:

- ***Height***. This trait refers to the height (h) of the plant. This height refers to actual height of the plant measured in field but, when this information is not available, the potential height of the plant derived from literature is acceptable.
- ***Density***. This trait refers to the density of the branch-leaf complex per individual plant. Thicker foliages cause an increase in Manning's coefficient[12] along floodplains (Anderson, 2006) and therefore an increase in hydraulic resistance. Due to limited field information, horizontal density was not possible to derive,

[11] Species abundance index.
[12] Manning's roughness coefficient (n) values are used in the Manning's formula for flow calculation in open flow channels. It represents the resistance to flood flows in channels and floodplains. Higher values indicate higher hydraulic resistance.

therefore only vertical density was considered. This trait is measured in two different ways: one for woody[13] vegetation and another one for herbaceous[14].

a) For woody vegetation it is measured by the vertical distribution of branching in the stem. Two basic structures can be found: shrubs (S), with branches growing from the bottom part of the stem, so a higher density is assumed, and trees (T), with branching starting at a certain distance from the bottom, in this case density is assumed to be lower than for shrubs.

b) For herbaceous vegetation this parameter was measured by the leaf type. Two general leaf types are considered which belong to two different herbaceous groups: graminoids and forbs. Graminoids[15] are grass or grass-like plants in which the length dimension of leaves is predominant with respect to the width dimension (Figure 6.a). Forbs are non-graminoids herbs in which the opposite occurs; the width dimension is predominant (Figure 6.b). Due to this, the leaf type for graminoids was termed long (L) and the leaf type from forbs was termed broad (B). Results from Liras-Laitas (2005) show that monocotyledonous[16] species (*Calamagrostis epigejos, Agrostis stolonifera)* have different spectral properties than dicotyledonous[17] species (*Cirsium arvense, Rubus caesius, Urtica dioica*) in the Millingerwaard. Assuming that species with broad leaf type have thicker foliage than species with long leaf type, we conclude that L type vegetation present less hydraulic resistance than B type vegetation.

a) Long leaf type **b) Broad leaf type**

Figure 6: Herbaceous leaf types

[13] A woody plant is a plant having hard lignified tissues or woody parts especially stems. Woody vegetation, as considered in this research, includes subshrubs, shrubs and trees.
[14] Vascular plants without significant woody tissue above or at the ground
[15] Includes grasses (*Poaceae*), sedges (*Cyperaceae*), rushes (*Juncaceae*), arrow-grasses (*Juncaginaceae*), and quillworts (*Isoetes*)
[16] Include graminoid families *Poaceae*, *Cyperaceae*, *Juncaceae* and *Juncaginaceae*
[17] Include most abundant forbs.

- ***Flexibility***. This trait is described as the ease to bend with water impact. It is measured also in different ways depending on the woodiness.

a) For all woody vegetation it is measured by the stem diameter (d).

b) For herbaceous vegetation stem diameter (d) is not a relevant parameter since it does not experience big changes compared to the ones in the woody case. So it is assumed that stem thickness is always under 1cm for herbaceous vegetation and hence only one high flexibility type is applied for all.

3.3.2.2 PFTs classification

Subject to the PFTs conceptual model defined in section 3.3.2.1, the PFT classes were established. Two different approaches were followed, one for woody vegetation and one for herbaceous, which are described after Table 3. The final classification of PFT according to this conceptual model is shown in Table 3.

Table 3: PFTs classes defined according to conceptual model

PFTs	*Resistance*	*Clusters*	*Height(Height (m))*	*Flexibility(Stem thickness (cm))*	*Density(Tree, Shrub)*
			Herbaceous		
SNL	1	H1D2F1	≤ 0.5	<1	L
SNB	2	H1D1F1	≤ 0.5	<1	B
RNL	3	H2D2F1	> 0.5	<1	L
RNB	4	H2D1F1	> 0.5	<1	B
			Shrubs		
SWS	2	H1D2F1	0-0.5	0-5	S
		H2D2F1	0.5-4	0-5	
MWS	4	H1D2F2	0-0.5	5-15.0	S
		H1D2F3	0-0.5	15-130	
		H2D2F2	0.5-4	5-15.0	
		H3D2F1	4-35.0	0-5	
		H3D2F2		5-15.0	
RWS	6	H2D2F3	0.5-4	15-130	S
		H3D2F3	4-35.0	15-130	
			Trees		
SWT	1	H1D1F1	0-0.5	0-5	T
		H1D1F2	0-0.5	5-15.0	
		H1D1F3	0-0.5	15-130	
MWT	3	H2D1F1	0.5-4	0-5	T
		H2D1F2	0.5-4	5-15.0	
		H3D1F1	4-35.0	0-5	
		H3D1F2	4-35.0	5-15.0	
RWT	5	H2D1F3	0.5-4	15-130	T
		H3D1F3	4-35.0	15-130	

a) Woody vegetation

Woody vegetation classification into PFTs was done by two subsequent aggregations. The first led to eighteen clusters based on plant traits measurements, the second and final one, led to six classes of hydraulic resistance. By these two subsequent aggregations a PFTs classification was obtained that associates hydraulic resistance with plant traits.

For the creation of the first clustering, the first step was to define groups within each plant trait. Three groups were created for height (H), two groups for flexibility (F) and another three for density (D):

- 3 groups of height: H1 ($0<h\leq0.5$ m), H2 ($0.5<h\leq4$ m] and H3 ($4<h\leq35$ m].
- 2 groups of density: D1 (Trees) and D2 (Shrubs).
- 3 groups of flexibility: F1 ($1<d\leq5$ cm], F2 ($5<d\leq15$ cm] and F3 ($15<d\leq130$ cm].

All possible combinations of these groups were carried out so that eighteen clusters were derived (Figure 7). These eighteen clusters, which are based on specific plant trait combinations, were regrouped according to six PFTs woody classes (SWT, MWT, etc.) associated to six levels of vegetation hydraulic resistance, ranging from lowest (Resistance 1) to highest resistance (Resistance 6) (Figure 7):

- Resistance 1: SWT (Soft Woody vegetation with Tree structure)
- Resistance 2: SWS (Soft Woody vegetation with Shrub structure)
- Resistance 3: MWT (Medium Woody vegetation with Tree structure)
- Resistance 4: MWS (Medium Woody vegetation with Shrub structure)
- Resistance 5: RWT (Robust Woody vegetation with Tree structure)
- Resistance 6: RWS (Robust Woody vegetation with Shrub structure)

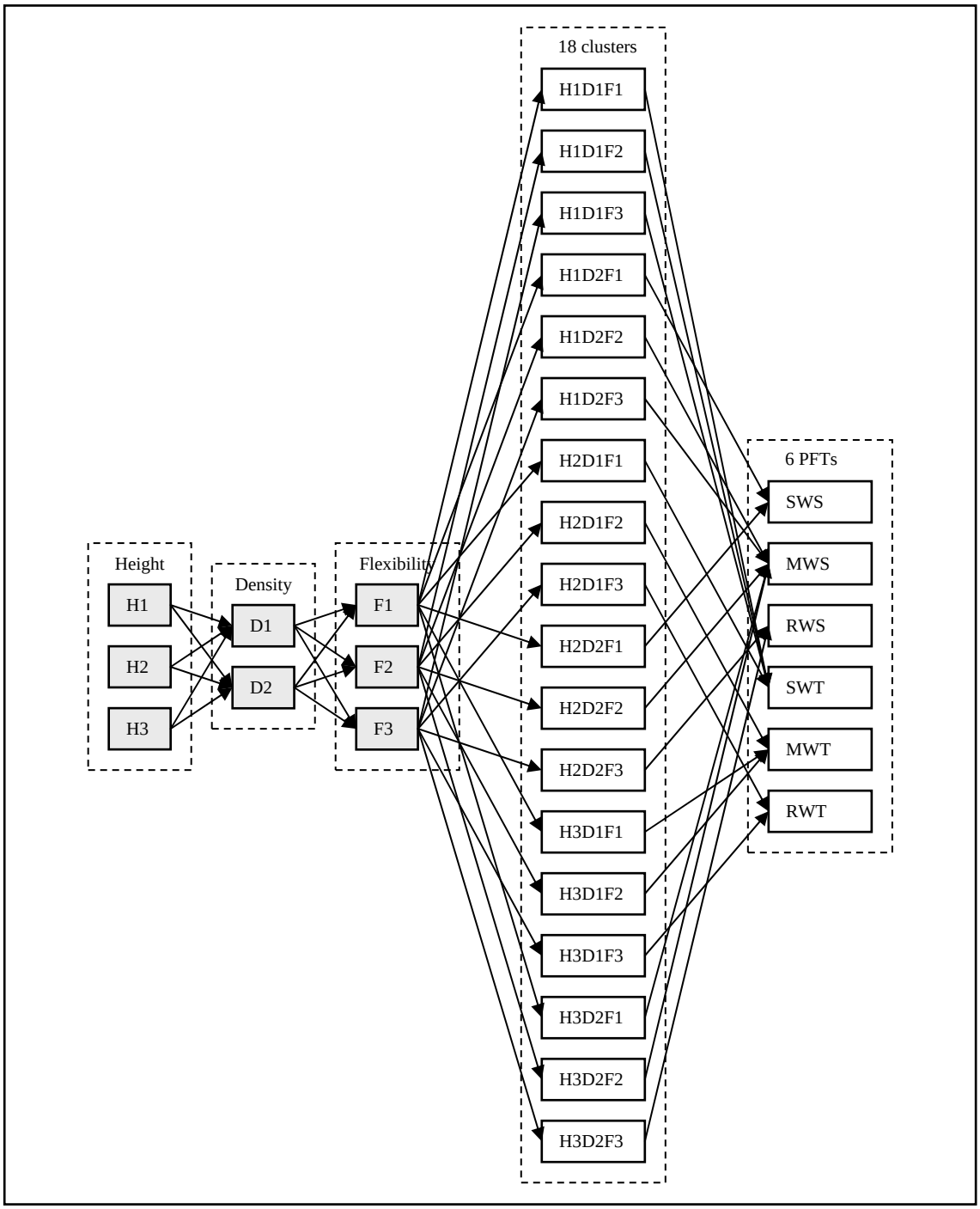


Figure 7: Classification of woody vegetation into PFTs

b) Herbaceous vegetation

Herbaceous vegetation classification into PFTs was done by only one aggregation. As for woody vegetation, the first step was to create groups within each plant trait. Two groups were created for height (H), one group for flexibility (F) and another two for density (D).

- 2 groups of height: H1 ($h \leq 0.5$ m) and H2 ($h > 0.5$ m).
- 2 groups of density: D1 (Long leave) and D2 (Broad leave).

- 1 group of flexibility: F1 (d<1 cm).

All possible combinations of these groups were carried out so that four clusters were derived (Figure 8). They were named according to four PFT herbaceous classes ranging from lowest (1) to highest (4) hydraulic resistance (Figure 8):

- Resistance 1: SNL (Soft Non-woody vegetation with Long leave)
- Resistance 2: SNB (Soft Non-woody vegetation with Broad leave)
- Resistance 3: RNL (Robust Non-woody vegetation with Long leave)
- Resistance 4: RNB (Robust Non-woody vegetation with Broad leave)

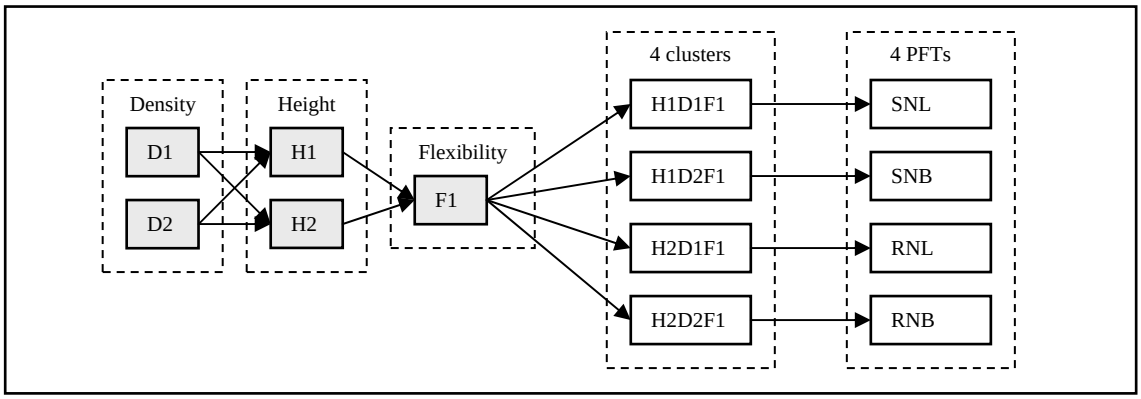


Figure 8: Classification of herbaceous vegetation into PFTs

3.3.3 Preprocessing

3.3.3.1 Image Data

Various aspects have to be taken into account before processing a hyperspectral image. These are internal calibration -correction from sensor noise-, external calibration -atmospheric and irradiation correction-, geo-referencing, masking of water and clouds if necessary and resizing to reduce computational time.

a) HyMap image

HyMap image had been radiometrically, atmospherically and geometrically corrected before (Kooistra et al., 2005; Kooistra et al., 2007). Since this image, in UTM coordinates, was the base map and all data was transformed to UTM, no projection transformation was required. The existence of water bodies, such as the river Waal and small lakes, and also the existing of built o agricultural areas could lead to confusion as these land cover types are not object of the unmixing classification. Therefore areas with these land cover types were masked out of the area of study.

Two masks were successively applied. First a mask was applied to exclude water bodies from the image. This was done by calculating the Normalized Difference

Vegetation Index or NDVI for the whole image. Vegetation NDVI typically ranges from 0.1 up to 0.6, with higher values associated with greater density and greenness of the plant canopy. Surrounding soil and rock values are close to zero while the differential for water bodies such as rivers and dams have the opposite trend to vegetation and the index is negative (Lillesand & Kiefer, 1994). In order to exclude water, NDVI values lower or equal to 0 were masked out. Values between 0 and 0.1 were kept because sand was selected as one of the endmembers (4.2.1 Table 7).

NDVI is defined as (Orlovsky, 2002):

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

NIR and RED are the reflectivity over the NIR regions (700-900 nm) and the red (630-740 nm). This corresponds to bands 19 to 32 and band 14 to17 respectively. Better results, understood as bigger contrast between vegetated and non-vegetated areas, was given by the bands 19 in the NIR and 15 in the red. This was evaluated by applying NDVI transformation to various combinations of RED and NIR bands. The resulting NDVI images were compared in between. Random locations of vegetated areas showed highest NDVI values, while random locations of non-vegetated areas showed lowest NDVI values, when using bands 19 and 15.

The second mask was meant to select only the area of interest. It implied the exclusion of most important urban and agricultural areas. This was done by creating a new vector file in ArcGIS in which such area was drawn. This file was imported to ENVI as a vector file and exported to ROI. As a ROI it could be used to build a mask. This mask was applied to the previously masked image with all pixels with NDVI values above 0.

Clouds were not observed in the study area, although presence of clouds is indicated by NDVI negative values (Lillesand & Kiefer, 1994) and hence were already excluded by applying the first mask.

The image was also resized to reduce computational time. The resizing consisted of suppressing the non-value area resulted from the masking. Resizing was only done after MNF transformation was applied due to the incoherence explained in chapter 3. 3. 4. 2. A second resizing was done to HyMap when facing the temporal analysis (chapter 3.3.5). As the original extension of CASI image is different than that from HyMap and the initial study area defined for HyMap was cut in CASI, the area of interest had to be

reduced during the temporal analysis in order to perform the unmixing to the same area in both images.

b) CASI image

CASI image was subject first to a projection transformation from RD to UTM coordinate system. Then, an atmospheric correction was applied to it; since only radiometric and geometric corrections were originally done. Finally, masking and resizing were applied.

Projection transformation was performed in ArcGIS. After this projection transformation, some spatial inaccuracies were observed. CASI image was translated some meters with respect to HyMap and the field data (already transformed into UTM coordinates). This was deduced by looking at specific features in the image. A translation of 10 m east was performed to correct this shift. It was exported into ERDAS IMAGINE and by the function "Set drop point" was shifted 10 m East to the correct position. This image was subject to the "Modeler" in ERDAS before it could be introduced in ENVI software; to assure the proper image format.

Atmospheric correction was performed in ENVI using the "Empirical line correction". The application of this function demanded pairs of spectral data from locations in the image and the field. As no field measurements were taken during the day the image was taken, data from field locations were gathered from the HyMap image -which had been atmospherically corrected- assuming that these locations did not changed much in time. Taking into account too that at least one dark and bright point are recommended four locations were chosen: river, lake, sandy area and a roof. Then, the empirical line correction was performed.

Finally the image was masked and resized. As well as HyMap, two subsequent masks were applied, one with NDVI values greater than zero and another one by setting the area of interest.

3.3.3.2 Field Data

Before spectral unmixing can be carried out, field data must be adapted so that they can fulfill the concept of PFTs as explained in the section 3.3.2.2 and can be of use for the unmixing purposes: endmember selection and validation. But field data is not present in a way that fits the needs of the classification procedure, because it was not gathered based on any specific application. Therefore projection system, format,

sampling unit and information have to be modified to fulfill specific requisites. Different procedures were applied to each dataset due to their diverse characteristics (Table 2). These modifications may lead to loss of accuracy because assumptions have to be made.

Field data preprocessing consisted of the following steps. First, the data about vegetation had to be re-classified into PFTs so that field information could be adopted for the image classification. Second, the projection system of all field data sets was transformed from the Dutch RD coordinate system (Rijksdriehoeksmeting) to the Universal Transversal Mercator (UTM) system. Third, the format had to be converted in some cases. And fourth the sampling units of all sets had to be adapted so that they could be compared to the image format.

In the next sections the preprocessing of the different data sets is described.

a) Ark Data (2005-2006)

Classification on PFTs (Ark dataset)

This data set consists only of woody vegetation. It is the most detailed one because many variables were measured in the field per individual plant, such as RD coordinates, height, stem diameter, crown projection, damages, etc. Based on some of these measures, classifying each individual plant into all six PFT woody classes (Table 3) is immediate because there is information about all three traits per plant: height, stem diameter and shrub/tree structure[18].

However during the implementation of these PFTs to the unmixing process, classes were rearranged, because, even though all classes were represented, some were very scarce. This made it impossible to extract sets from these classes for endmember selection or validation assessment. This was the case for classes SWT, MWT and SWS. Only 72 and 34 plants were classified as SWT and MWT respectively (in contrast, i.e. 1614 plants were classified as MWS). These plants presented a low crown projection with averages of 0.96 m^2 and 3.97 m^2 for each class respectively. Class SWS was more abundant (1888 plants) but with a small average crown projection also (1.55 m^2). This, together with the fact that the pixel surface is 25 m^2, made impossible an approximation to pure pixels for these three classes.

When applying unmixing to an image all existing features in the field have to be considered because lack of endmembers lead to less accurate results (Malenovsky et al.,

[18] This character of the plant was based on the specie. Some species have typically a shrub structure while others have a tree structure.

2006). Therefore these classes poorly represented could not be simply ignored, especially because some (SWS) are very abundant. Hence a rearrangement was done to the original six PFT woody classes. Only three of them were kept:

- Resistance 1: MWS (Medium Woody vegetation with Shrub structure)
- Resistance 2: RWT (Robust Woody vegetation with Tree structure)
- Resistance 3: RWS (Robust Woody vegetation with Shrub structure)

All individuals previously classified as SWS were assumed to join herbaceous classes if their height was under 0.5 m, and re-classified as MWS if it was above 0.5 m. Individuals classified as SWT and MWT were assumed to join any of the shrub classes.

Projection transformation (Ark dataset)

The projection system for the Ark dataset was transformed from RD coordinates to UTM. For some unknown reason transformation of coordinates in ArcGIS software was constantly resulting in a shift of a couple of hundred meters from the expected position. This led to the use of alternative coordinate calculator software for the Netherlands[19].

For the use of this software, RD coordinates were exported to a text file, after being converted into UTM, this text file was imported into Arc GIS. Then, the coordinates were displayed and exported to a vector file that was subsequently joined to the original Ark data set so that all the information was contained in the new projected set (Figure 9). The quality of this transformation was examined by overlaying the plots to the HyMap image and checking the fit of paths, borders and water areas.

[19] Coordinate calculator 41downloaded from the Web. Available from http://www.rdnap.nl/download/download.html

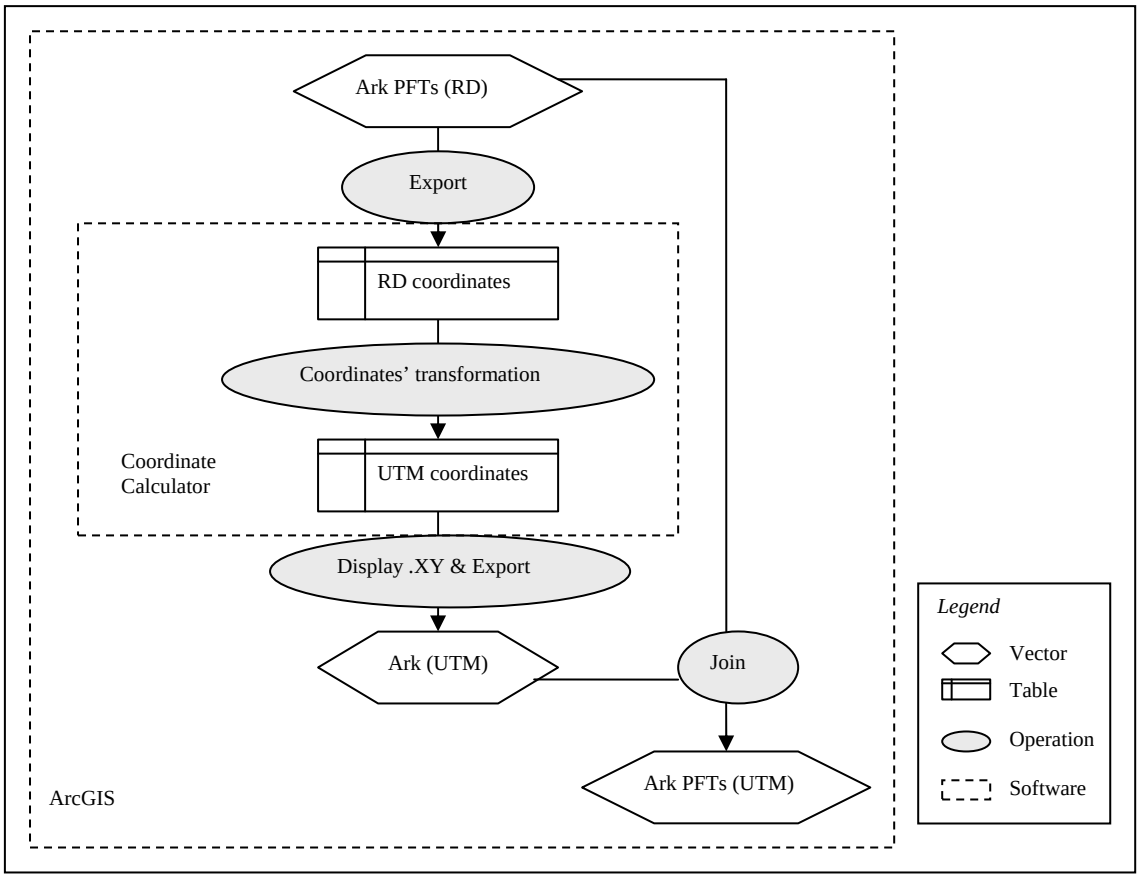


Figure 9: Ark data, transformation of projection system

Sampling unit adaptation (Ark dataset)

Once the classification of the data was done and the projection system was transformed, the sampling unit of the data was checked. In the Ark dataset the sampling units are points representing individual plants, but the pixel size of the image is 5x5 m, and field information has to be comparable to the image. This demands the sampling unit to be modified. By knowing the number of plants belonging to each PFT per pixel and the crown projection per individual plant, the total crown projection of each PFT class per pixel of 5x5 m can be derived. After this conversion, field knowledge is immediately comparable to the image information.

This conversion was carried out using the ArcGIS software. The next steps were followed (Table 10):

- All individuals from each class were extracted from the Ark dataset (shapefile) into separate datasets (shapefiles) (s1)
- Each dataset (s1) was converted to raster using a pixel size of 1x1 m (r1). The value assigned to each pixel was the crown projection. The output extent was

specified equal to the image to avoid a shift between this raster's origin and the image's.

- Each raster (r1) was aggregated in a new raster of pixel size 5x5 m. The value given to each pixel in this new raster (r2) was the sum of the values (crown projection area) of all the pixels of 1x1 existing into each pixel of 5x5. Therefore the total crown projection per pixel of 5x5 m is known.
- Parallel to this part, Ark vector data was converted to raster data of 5x5 keeping the same extent of the image again, and then converted to points (p1); each point representing the center of the pixels. Points different to zero were selected so that the remaining points were the center of those pixel in which one or more Ark measures had been taken (p2).
- All aggregated raster sets of 5x5 (r2) were sampled using these points (p2) and a table with all volumes per class and pixel was derived.
- This table was displayed according to the p2 coordinates and exported as a vector file (v1).
- All points from the vector file (v1) with coverage greater than 12.5 m^2 out of 25 m^2 (area of the whole pixel of 5x5 m) were selected for each class separately because only these points were used for setting up the endmembers (the ones with the highest values and purest content) and for validation (the rest).

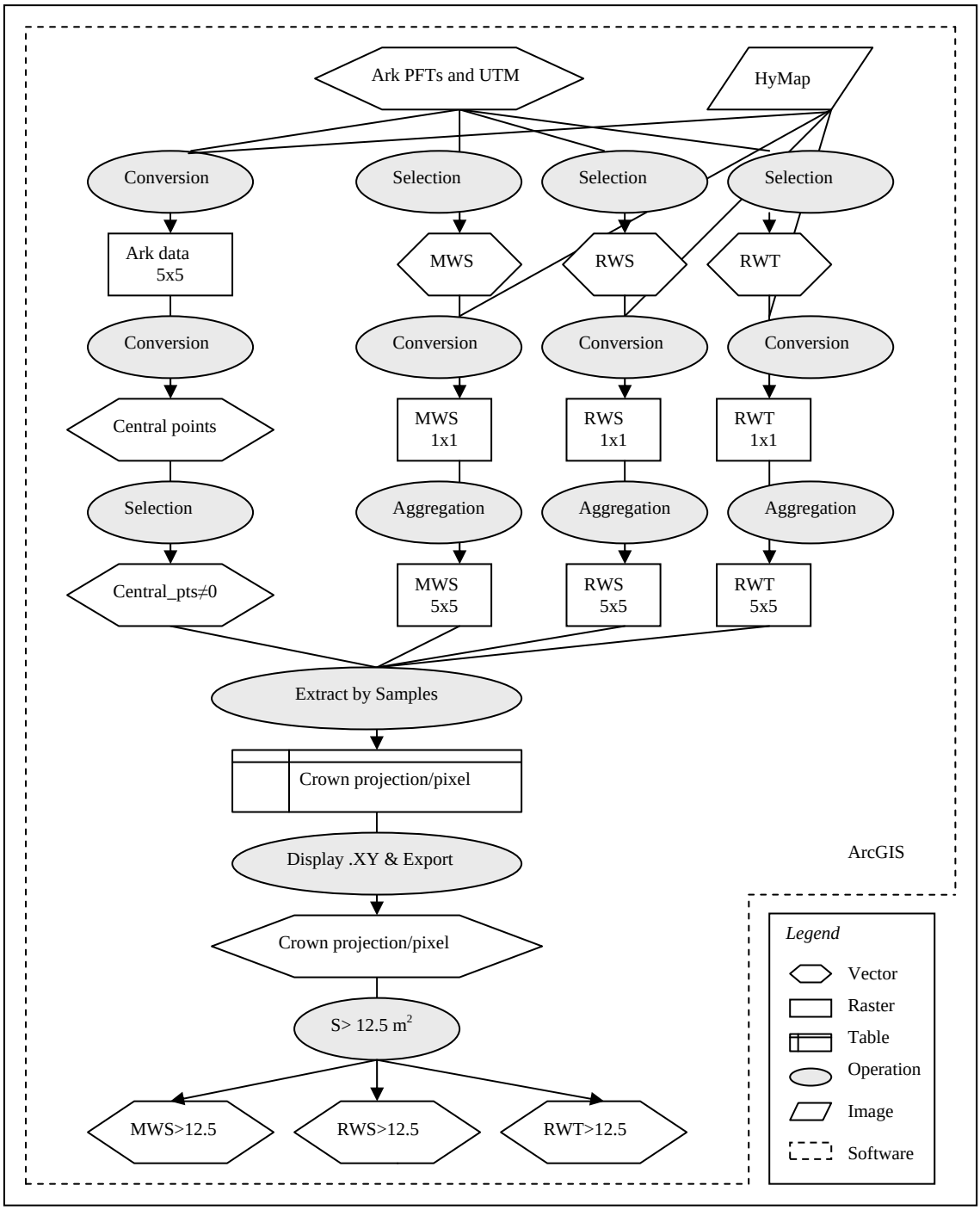


Figure 10: Ark data, conversion of sampling units

During the endmembers selection a new modification was done to these three woody classes. Due to its high spectral variability, RWS was divided into two classes (Section 3.3.4.1): RWS and RWSC. The first one (RWS) dominated by *Salix fragilis*, and the second one (RWSC) dominated by *Crataegus monogyna*. Therefore a fourth validation set was created from Ark dataset in order to validate RWSC (Figure 11).

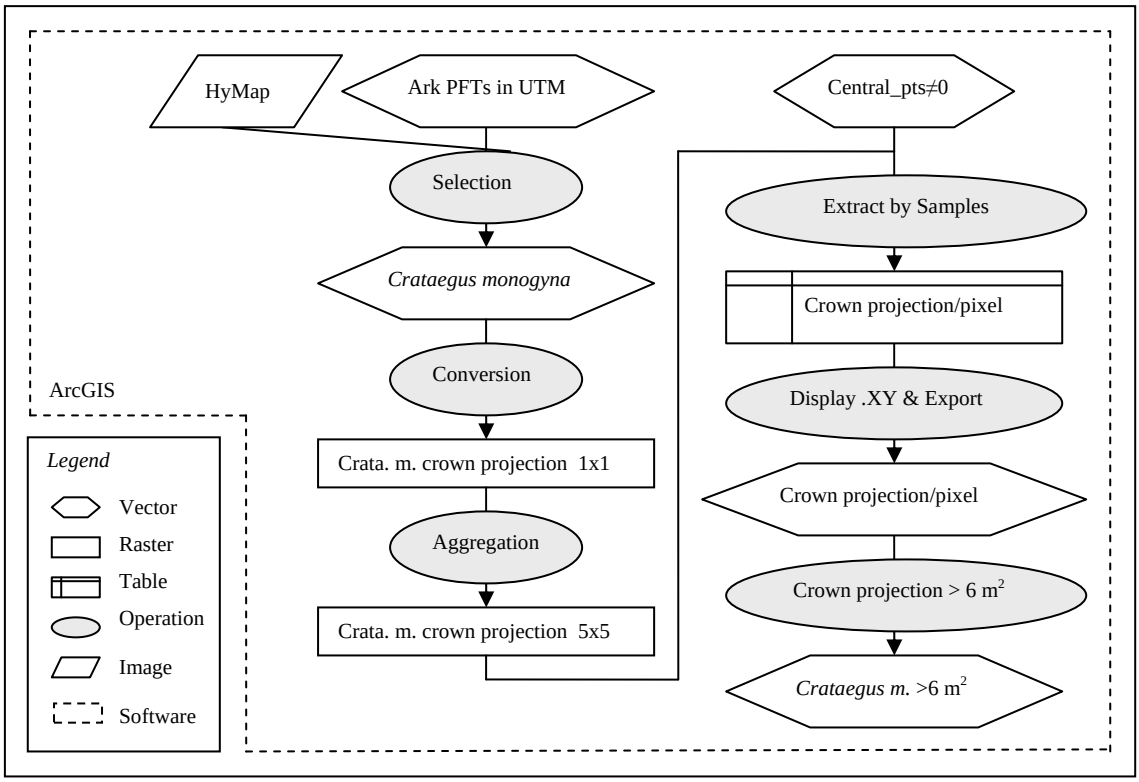


Figure 11: Creation of validation set for class RWSC

b) Sykora Data (2002)

Classification on PFTs (Sykora dataset)

This data set consists mainly of herbaceous vegetation type, but shrub vegetation is also present and tree species are seldom found. Data was not gathered by individual plant but in plots of 2x2. In these plots an exhaustive registration of species and species abundance in terms of the Braun Blanquet index[20] was collected. Also information on plant community and height range per plot was available. This information however, was not of use for this purpose since height range was an average of the whole plot and plant communities were described by syntaxonomy nomenclature which is not related to plant abundances but to botanical singularity. Therefore only species and species abundance per plot was considered to be applicable.

The classification of herbaceous vegetation is based in two characteristics: height and leaf type, as explained in chapter 3.3.2.1. Precise height per individual plant or specie from the field did not exist for these plots, so height per specie was estimated according to literature. All sources consulted were found in internet. They are listed in the end of chapter 9 (bibliography). To derive each species' height more than two

[20] Index to express species' abundance in terms of coverage. Each value is related to a range of cover.

sources were consulted. Consistency between these sources was found for most of the species. Considering that actual height of plants was likely under potential height, height was estimated as the average of the values found in literature minus certain quantity[21]. Under the assumptions that height could be estimated using literature and that graminoid species have L leaf type while forbs have B leaf type (Figure 6), all herbaceous species were classified as SNB, SNL, RNB or RNB (Table 3).

The classification of woody vegetation is done considering: height, stem diameter and shrub/tree structure. No height nor stem diameter information per plant was available in the dataset, therefore an accurate classification of these data was not viable, specially considering the previous classification of Ark Data. But a rough classification was done only considering estimated height of species according to literature (same process explained for herbaceous vegetation in previous paragraph). These woody classes resulted: SWS, MWS, MWT and RWT. All of them were considered in the final percentages of PFTs per plots but, except for SWS, no set for training or validation was extracted from them. On the one hand their presence is low, on the other hand, Ark dataset –which represents the other source for woody vegetation- has special characteristics that do not make it recommendable for mixing or testing with this other data set. In the Sykora dataset the real height and stem diameter per plant is not available so height was estimated per specie according to literature. This means that when in Ark data classification, two individuals from the same specie may be in different classes -because they have a different size- or two individuals from different species may be in the same class -because they have similar size-, in the Sykora dataset classification, it is assumed that all individual from one specie belong to the same class.

In Sykora dataset there was a group of 58 tree locations. The information available about these trees was that their height was greater than 5 m, and that most of them are *Salix alba*. These trees were classified as RWT, assuming that the stem diameter was greater than 15 cm, and used as validation and training set during the temporal analysis (3.3.5.3).

Individuals classified as SWS class however were considered for training and validation. These individuals were solely belonging to *Rubus sp.* This subshrub specie was not present in Ark data set but it is of great importance for river managers since it is

[21] This quantity was assumed as 0.1 m when the plant potential height was less than 1 m, 0.2 m when it was between 1-2 m, 0.5 m when it was between 2-4 m and 1 m is it was higher than 4 m.

abundant and rapidly developed in the field. Moreover there is not incompatibility with Ark SWS class, because this class was disintegrated in Ark dataset.

Projection transformation (Sykora dataset)

The projection transformation of this data set comprises exactly the same steps as the transformation executed for Ark data (Figure 10).

Sampling unit adaptation (Sykora dataset)

The sampling units for the Sykora dataset are points representing plots of 2x2 m. Two considerations have to be taken into account regarding these plots. First that the center of the plots is not necessarily coincident with the center of the pixels and second that the pixel size (5x5 m) is greater than the plot size (2x2 m).

For the first consideration a moderate distance between the center of the plots and the center of the pixel is assumed to be acceptable because the plot size (2x2 m) is smaller than the pixel size (5x5 m) and if the distance is small enough the plot still falls into the pixel. The first threshold distance chosen was 0.7 m, but the number of plots that fell into this group was too small and it was raised to 1.5 m (Figure 12).

For the second consideration –the fact that the plot size is smaller than the pixel size- it is assumed that the composition of the plot is extensive to the whole pixel in the same proportion.

As described before, information used from this data set was species type and species' abundance per plot. Species have been associated so far to PFTs, so now we know PFTs' abundance per plot. But abundance per plot came expressed in terms of Braun Blanquet (B.B.) index. This index refers to a range of cover and for simplification ranges were converted to mean.

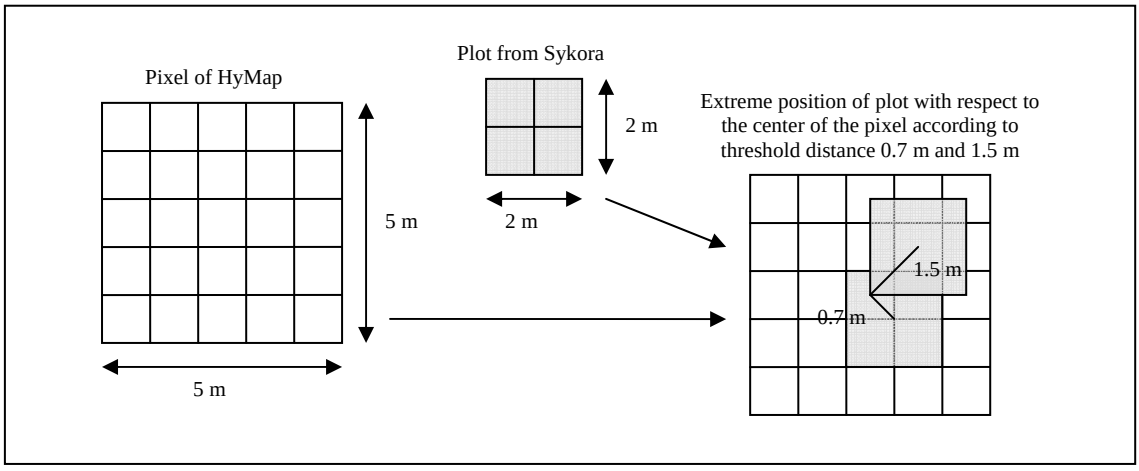


Figure 12: Relationship between Sykora plots size-position and HyMap pixel size-position

Table 4: Conversion of Braun-Blanquet cover-abundance scale to mean cover

Braun-Blanquet scale	*Range of cover (%)*	*Mean cover (%)*
5	75-100	87.5
4	50-75	62.5
3	25-50	37.5
2	5-25	17.5
1	<5	2.5
+	<5	2.5

These mean covers are expressed for various vegetation strata so although they are expressed in terms of percentages they sum up a different quantity than one hundred. For simplification, for this study it is assumed that abundances are in one unique stratum. Therefore they were summed and weighted up to one hundred. This was considered to be accurate enough for validation. Nevertheless, for the search of endmembers it was used just as a first approximation to the purest plots, but B.B. abundance was carefully studied on those plots to take final decisions. An extra complication was that some BB abundance addition per plot was below one hundred, for simplification these plots were deleted, because in the processing, while weighted up to one hundred, they could be confused with pure pixels, when actually they contain important proportions of soil.

A flowchart with all steps followed apart from the projection transformation is presented in Figure 13.

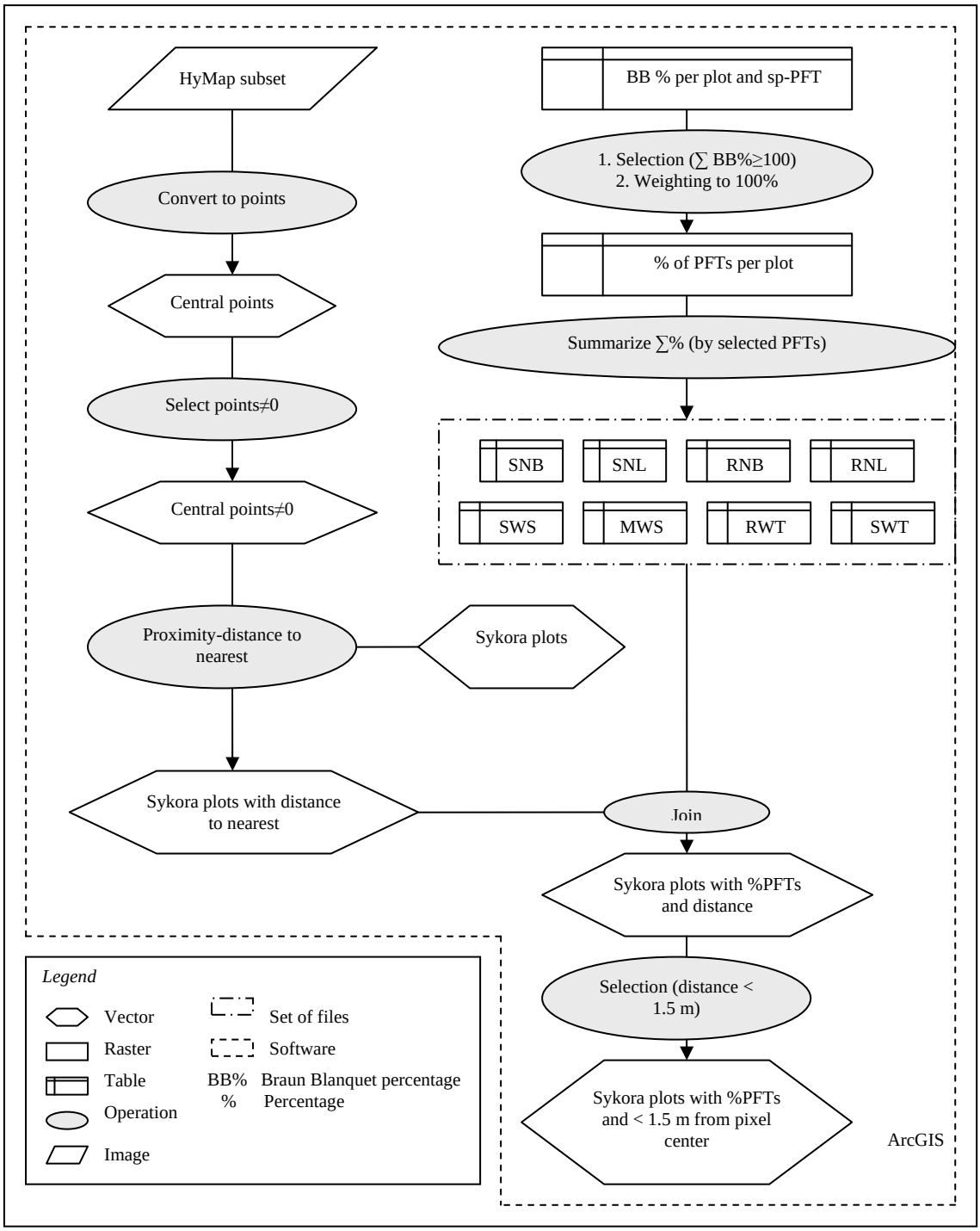


Figure 13: Sykora data, conversion of sampling units

c) HyEco'05 campaign Data (2005)

Classification on PFTs (HyEco'05dataset)

Available information from this data set is similar to that of the previous data set. It consists of existing species and species' abundance per plot (5x5 m). Although in this case abundance is expressed in terms of quantity of plants per specie. It consists mainly of herbaceous vegetation; woody vegetation is scarce. Therefore it has been classified following the same principle as for Sykora dataset: estimating height from literature.

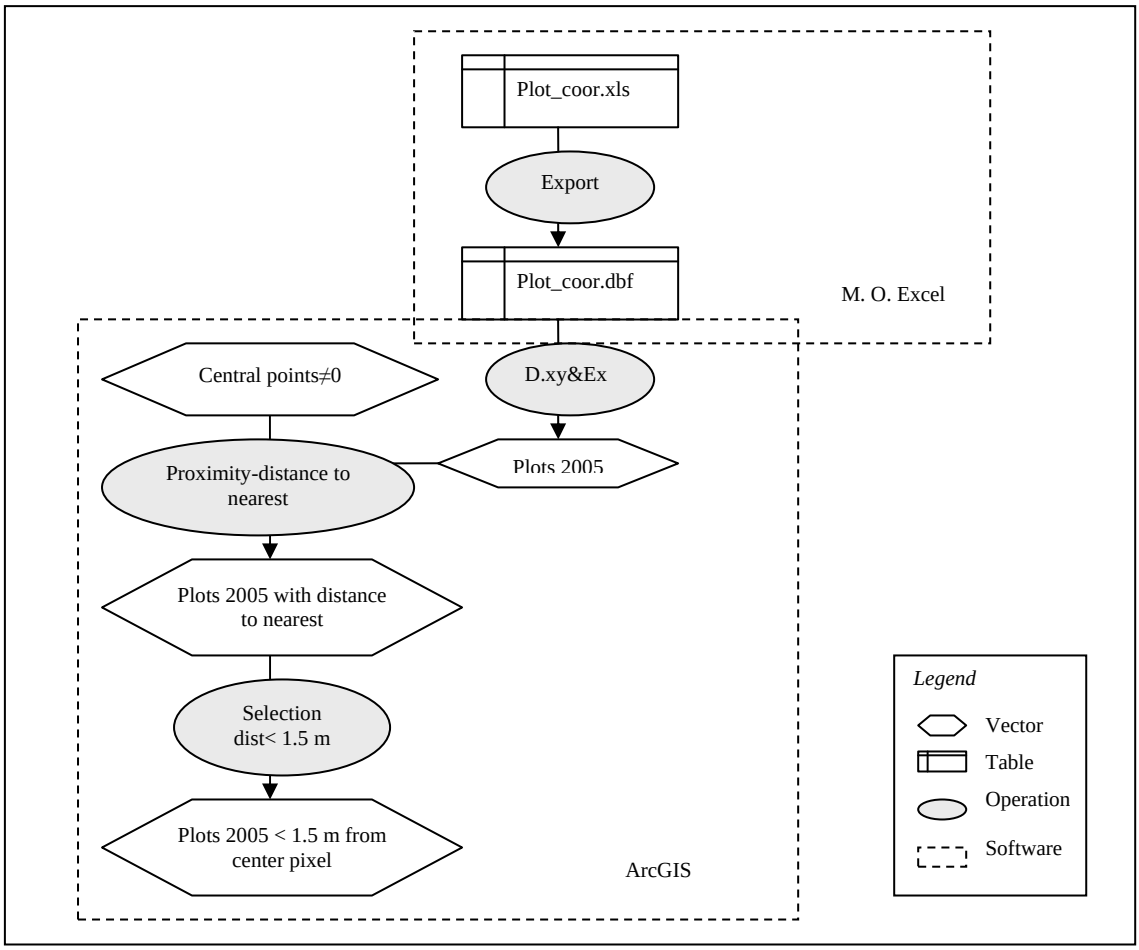


Figure 14: HyECO'05 data, format conversion

Projection transformation (HyEco'05dataset)

In this case projection transformation was not required since each plot was georeferenced using RD coordinates as well as UTM.

Format conversion (HyEco'05dataset)

This information was available in Excel format. But the spatial representation in ArcGIS was needed for the consecution of this work. Therefore, the coordinates in Excel format were saved as "dBASE IV" and imported to ArcGIS, here they were displayed in the view and exported as a shapefile (Figure 14).

Sampling unit adaptation (HyEco'05dataset)

There is also a shift between the pixel center and the center of the plots. Therefore a selection of the nearest plots to the center of the pixel using a threshold of 1.5 m was done.

Again after the association between species and PFTs we know the abundance per plot in terms of PFTs. This abundance, however, is in terms of number of plants of

every PFT class per plot. It was assumed that each plant covered the same area and a coverage percentage of PFTs per plot was calculated based on plant number.

d) HyEco'04 campaign Data (2004)

The preprocessing of this data follows the same steps as the data from HyEco'05, since the datasets are similar. The only difference is that plots in this case represent an area of 2x2 m and a new assumption has to be made: PFTs coverage percentages per plot of 2x2 m are extensible to the pixel area of 5x5 m.

3.3.4 Spectral Mixture Analysis

3.3.4.1 Selection of endmembers

Once image and field data have been adapted to the needs of the unmixing classification purpose the conditions for the selection of endmembers have been prepared. This task requires wariness because the primary and most important key to success for spectral unmixing is the selection of the right endmembers.

As described in chapter 2.3.3 several methods can be used for the selection of endmembers. There is a spectral library available from other studies carried by the Center of Geo-Information (CGI) of Wageningen University that may be considered, but these library endmembers are not recommended because of difficulties to scale up to match spectra measured by the sensor. Deriving endmembers with automatic procedures is not recommendable since spectra from different PFTs vegetation classes are very similar. Therefore, in this study, endmembers were derived directly from the image. This method has the advantage that selected endmembers are under similar atmospheric conditions and spectral/radiometric biases with respect to the image (see chapter 2). In the case of the HyMap image this is not of importance since it has been internally and externally calibrated. Moreover the great amount of field knowledge facilitates its application.

A second consideration is the number of endmembers, which is limited by the number of bands. In the case of a hyperspectral image such as HyMap, with 126 bands, this is not an important limiting fact.

For the determination of the endmembers, a selection of ROIs[22] was made. These regions are meant pure pixels. The following steps were followed:

- First, a group of candidate pixels that were considered 'pure' or nearly 'pure' were chosen from Sykora and Ark datasets for each class (Table 5).

[22] Regions Of Interest: field points or polygons used for classification training and accuracy assessment.

Table 5: First selection of ROIs for ten PFT endmembers. Each ROI consists of one pixel

ROIs	*Plot/ID*	*Year*	*Source*	*Cover*	*Species (BB%[23])*	*OtherPFT*	*Dist(m)*
SNB1	Plot 122	2002	Sykora	79%	50-75Geranium, others.	12% SNL 10% RNB	1.41
SNB2	Plot 232	2002	Sykora	68%	75-100Trifolium,12-25 Cirsium	20% SNL 6% RNB	1.41
SNL1	Plot 63	2002	Sykora	67%	25-50Carex-hirta,25-50Carex-arenaria,25-50Festuca	19% SNB	1
SNL2	Plot 139	2002	Sykora	60%	25-50Carex-hirta,25-50Cynodon,5-12Festuca	24%SNB 20%RNB 8% RNL	1.41
RNB1	Plot 103	2002	Sykora	83%	75-100Urtica,12-25thalictrum, etc.	9% SNB 7% RNL	1
RNB2	Plot 114	2002	Sykora	78%	50-75Urtica,12-50Rumex,25-50Heracleum	11% SNB 7% RNL	1.41
RNB3	Plot 110	2002	Sykora	77%	75-100Urtica	16% SNB	1.41
RNB4	Plot 251	2002	Sykora	74%	75-100Urtica	11% SNB 9% SNL	1.41
RNL1	Plot 219	2002	Sykora	53%	50-75Calamagrostis	20% SNB 17% SNL 8% RNB	1.41
RNL2	Plot 82	2002	Sykora	50%	50-75Calamagrostis	22% SNB 10% SNL 20% RNB	1
SWS	Plot 224	2002	Sykora	63%	75-100Rubus	36% RNB	1.41
MWS1	ID 91	Ene06	Ark	86%	5Sambucus	Herbs	
MWS2	ID 104	Ene06	Ark	84%	4.5Sambucus,1Rosa	Herbs	
MWS3	ID 106	Ene06	Ark	80%	5Sambucus	Herbs	
MWS4	ID 784	Nov05	Ark	72%	2.5Crataegus	Herbs	
MWS5	ID 265	Ene06	Ark	76%	6Sambucus	16% RWS	
RWS1	ID 1409	Mar06	Ark	97%	2.5Sambucus,0.5 Salixfra[24]	3% MWS	
RWS2	ID 1220	Mar06	Ark	99%	1Salix fragilis	1% MWS	
RWS3	ID 1271	Mar06	Ark	89%	1Salix fragilis	11% MWS	
RWS4	ID 1114	Mar06	Ark	100%	3Crataegus	Herbs	
RWS5	ID 1190	Mar06	Ark	80%	2Crataegus,1Sambucus	Herbs	
RWT1	ID 2167	Feb06	Ark	100%	2Salix alba	Herbs	
RWT2	ID 2097	Feb06	Ark	100%	2Salix alba	Herbs	
RWT3	ID 2232	Feb06	Ark	100%	2Salix alba	Herbs	

[23] Braun Blanquet abundance.
[24] *Salix fragilis*

- Then, a shapefile was created in ArcGIS containing all these points (Figure 15).
- After, this shapefile was imported into ENVI where each point was converted into a ROI (Figure 15).
- The signatures of each ROI were plotted in ENVI (Appendix 2) and studied comparatively to the rest, with the intention of avoiding high variability within classes and favouring the difference between separate classes. This discussion is described after Table 5.
- Based on the former discussion a selection of definitive vegetation ROIs was made (Table 8) in ENVI (Figure 15). These nine pixels represented the nine PFTs.
- Finally, a sandy ROI was added as tenth endmember to the former set. The spatial location if the 10 endmembers can be found in Figure 37.

Between the two pixels regarding class SNB there is a high variability (Figure 30). This is understandable once we consider that the species representing each pixel are completely different in between: SNB1 consists of *Geranium sp*. and SNB2 consists of *Trifolium sp*. and *Circium sp*. Because these two last species are more significant due to their abundance in the area, pixel SNB2 was selected.

For class SNL there is not such a high variability (Figure 30). Due to the important presence of *Festuca sp*., pixel SNL1 is preferable. Moreover, this spectra is more differentiable with SNB1 regarding both patron and magnitude of the signature (Figure 31).

Variability within class RNB is not high (Figure 31). This is explained by the fact that all pixels have an important proportion of *Urtica dioica*. Pixels with lowest proportion of class RNL (Table 5), which are RNB3 and RNB4, are considered more representatives. However, RNB4 has the most outstanding signature of all, so RNB3 is considered to be purer and it is selected as ROI.

The two pixels for class RNL have a low inter-variability (Figure 31). This is a sign that, although *Calamagrostis epigeos* does not cover the whole area -as can be seen in the coverage percentages (Table 5)-, it determines highly the signature of the pixel. As RNL2 has a larger proportion of class RNB (Table 5) than RNL1, RNL1 is considered to be more appropriate as ROI.

Only one pixel was selected for class SWS because it was the unique acceptable pixel, hence it was considered valid as ROI for this class. Its signature can be found in Figure 32.

Five candidate pixels were chosen to represent MWS pure ROI (Table 5). The three first ROIs -MWS1, MWS2 and MWS3- composed of *Sambucus nigra* presented a similar signature (Figure 33), while there was a noticeable difference with the other two –MWS4 and MWS5-. The inter-variability of this last two was higher. MWS1 is considered the best option for ROI since it is represented by *Sambucus nigra* without presence of any other woody specie in a proportion of 86% (Table 5). *Sambucus nigra* is one of the most abundant woody species in the area.

In reference to the five candidates for class RWS, there was also a high variability (Figure 34), except for those with presence of *Salix fragilis* (RWS2 and RWS3). Pixel RWS2 is considered the best option because it is the purest (99% of class RWS).

Pixels RWT2 and RWT3 representing class RWT were very similar (Figure 35) and had the typical tree spectrum, with lower reflectance in the NIR, while RWT1 had an anomalous high reflectance. When looking at the spatial location of the pixels, RWT2 is more appropriate because pixel RWT3 is located at the edge of the study area and the reflectance may be influenced by inconvenient objects in the border. This ROI is composed by *Salix alba*.

After these decisions were made, it could be seen that *Crataegus monogyna* was not represented in any ROI. But this specie is very important for river managers because of its abundance and fast development, so it was decided to include it as a separate ROI: RWSC. The problem with this class is that there was a high variability within the three pixels representing this specie –MWS4, RWS4 and RWS5- and this variability was not size dependent (Figure 36). Therefore a decision had to be taken to choose one of the three different spectra and other sources were consulted.

First the field spectrometry measurement of *Crataegus monogyna* spectrum from the internship report of Catherine Teerhuis (2006) was consulted. These spectra were also gathered in the Millingerwaard. The same high variability within spectra could be observed. This specie has a specific structure which makes it more variable on reflectance in the NIR because of high scattering.

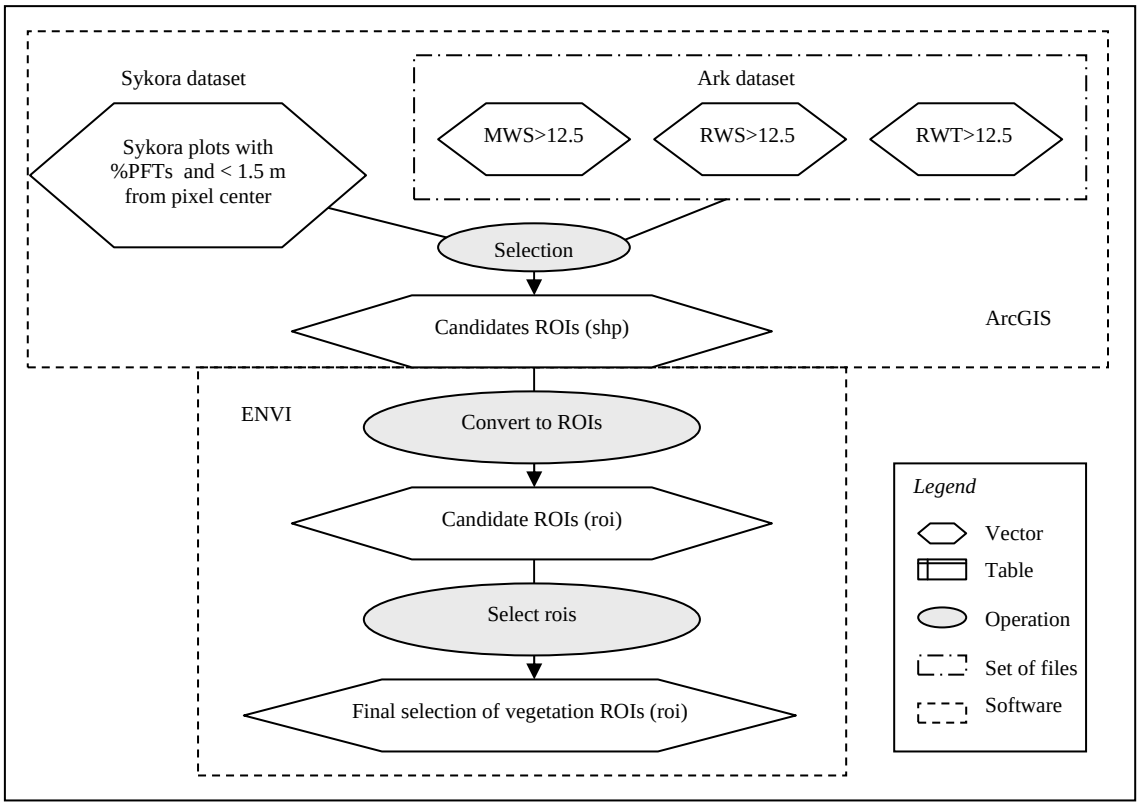


Figure 15: Implementation of ROIs

Then the aerial photograph from 2003 was consulted. MWS4 was not noticeable while RWS4 and RWS5 were clearly seen, so MWS4 was suppressed. RWS4 was chosen against RWS5 because it is more pure representative of *Crataegus monogyna*, RWS5 also contains *Sambucus nigra* (Table 5).

Due to the above reasoning nine ROIs representing nine vegetation endmembers were selected. To these nine ROIs (Table 8) another ROI for sand was added, finally ten ROIs were used for the unmixing.

3.3.4.2 MNF transformation

MNF transformation was applied to the images before the unmixing was executed (Figure 16). This was intended to find data's inherent dimensionality by avoiding information redundancy, minimize systematic sensor noise and get the endmembers to be orthogonal and uncorrelated.

Two important considerations were taken into account while running MNF transformation with ENVI. The first one concerns the masking of the image. Image and mask must be input to the transformation separately. If the image is input with

previously applied mask, during the PCA transformation the abundance of "0" DN[25] causes a strong translation to the new axis towards 0.

The second consideration concerns the extent of mask-image. When a mask is introduced into the MNF transformation, the new axes are spatially located according to the multidimensional distribution of the information contained on the mask area of the image; therefore, the result changes if different masks are applied while it should be equal for different extents if the mask is exactly the same. The incoherence in ENVI MNF transformation was that for different extents and same mask-image, different results were derived. In order to manage this incoherence a final decision was taken: to use the complete image.

Once the MNF transformation has been performed, only some of the new bands are significant, the others are mainly noisy. The noise bands can be identified by different approaches. First, by looking at the eigenvalues[26]; eigenvalues less than 1 are usually excluded as noise because eigenimages[27] with near-unity eigenvalues are normally noise-dominated (Fang, 2006). But also by registering important breaks in the slope of the eigenvalues graphical representation or by looking at the eigenimages; noise images can be identify by high sparkling and absence of features. The first 23 bands were chosen to perform the unmixing in HyMap by combining this information (Figure 16).

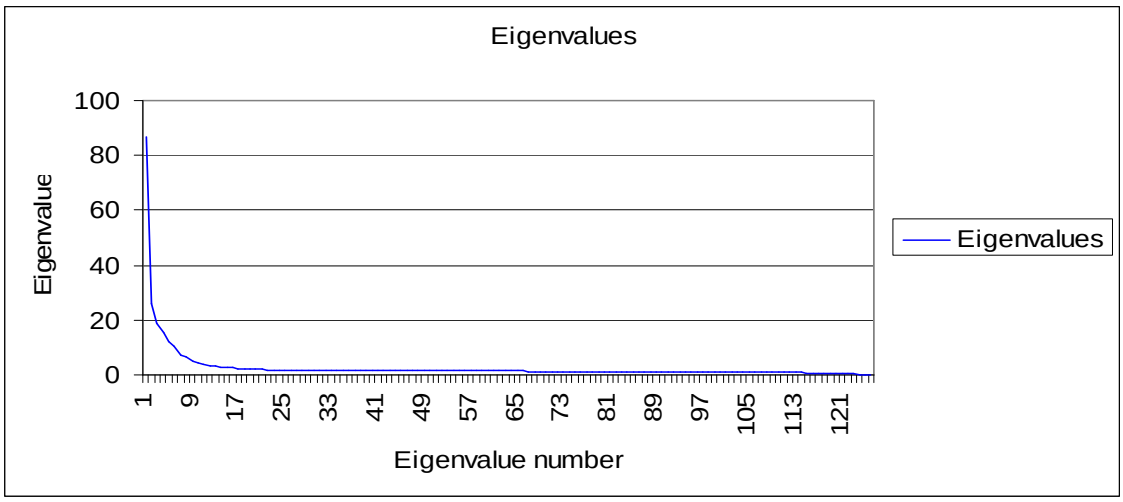


Figure 16: Eigenvalues of MNF transformation of HyMap image

[25] Digital number that expresses the radiation value of a pixel. When an image is masked, all DN out of the masked region are "0".
[26] Special set of scalars associated with a linear system of equations determined by matrix diagonalization (these equations are the linear orthogonal transformation from the original bands to the new bands).
[27] Refers to each of the new bands after the MNF transformations.

3.3.4.3 Linear spectral mixture analysis

Linear spectral unmixing was applied to HyMap image using the ten endmembers defined previously (Figure 16). It was performed in three different ways, each time including different number of constraints. These three methods were designed as: unconstrained unmixing, semi-constrained unmixing and fully constrained unmixing as explained in section 2.3.2. The unconstrained and semi-constrained versions were performed in ENVI. But this software does not include fully constrained unmixing, so this was implemented using MATLAB software.

The 23 band image resulting from the MNF transformation was unmixed using the ten ROIs previously defined (Table 8). As a result we obtained a set of ten abundance images (one per endmember) and another image with the RMSE value per pixel (Appendix 3 Figures 38-48). When applying the method with MATLAB other images were also produced (Figure 16, red outline): a hard classified[28] image and an exitflag[29] image. The hard classified image can be found in Appendix 3 (Figure 49). All pixel values from the exitflag image were 1, meaning that the function converged to a solution.

[28] The class with the highest fraction in each pixel is assigned to that pixel.
[29] An integer is assigned to each pixel identifying the reason why the algorithm terminated. The exitflag indicates whether the function converged to a solution (1), the number of interactions was exceeded (0) or the function was unfeasible (-1).

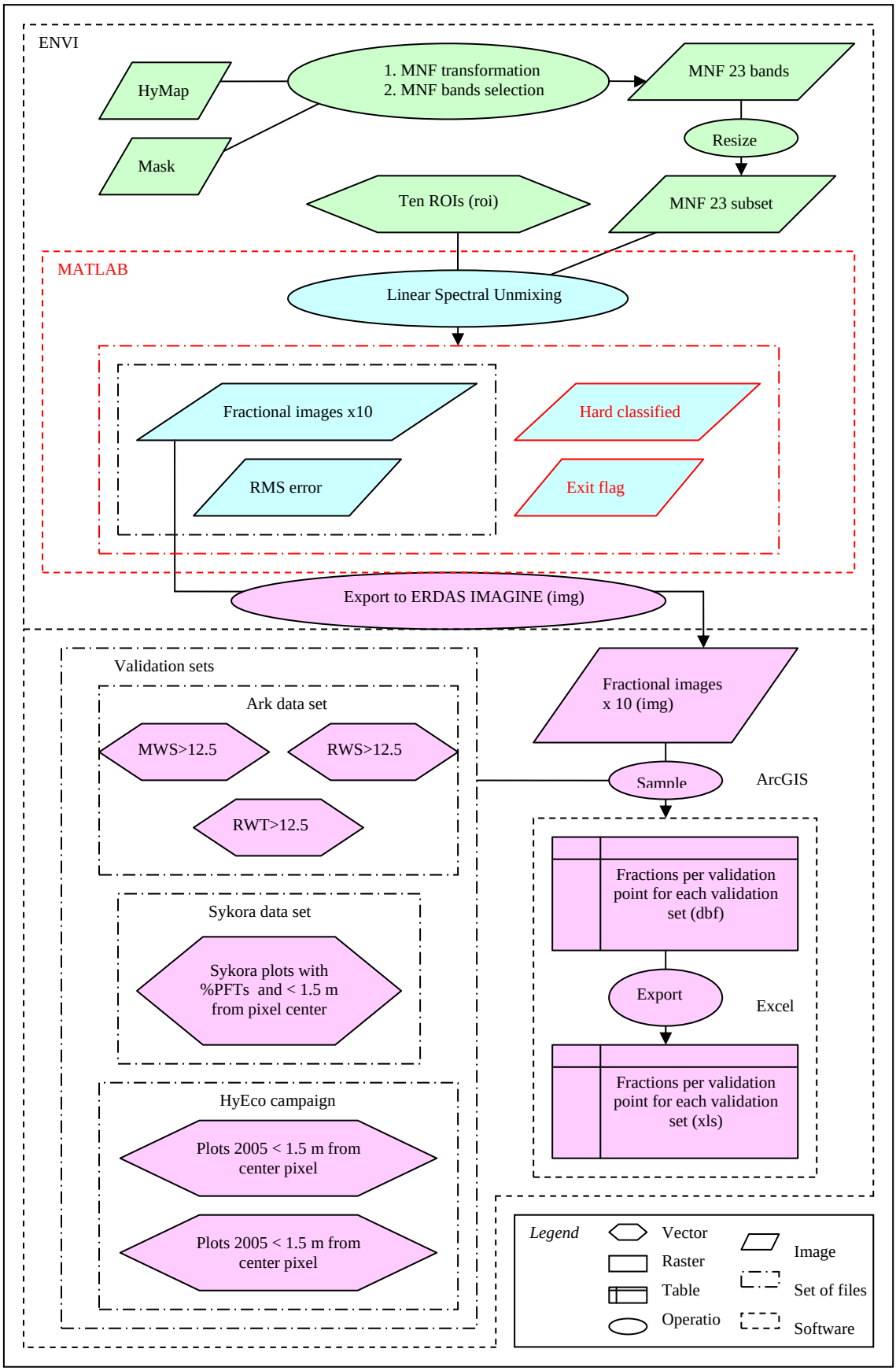


Figure 17: Application of Linear Spectral Unmixing to HyMap. In green, MNF transformation (ENVI), in blue, application of liner spectral unmixing (ENVI and MATLAB (red outline)) and in pink, sampling of abundance values using field plot locations for validation purposes (ArcGIS)

3.3.4.4 Validation

Validation of results was done using the four different approaches described in section 2.3.5: an analysis of the spatial continuous map for the RMSE, the comparison between modeled vs. observed coverage (by means of scatter plots and overlaying field plots to abundance maps) and a subpixel accuracy assessment (OSA). The last method could only be applied to fully constrained unmixing. This is because fully constrained unmixing is the unique method whose numerical results are immediately significant in the physical space; fractions are between zero and one, and the addition of all fractions in one pixel equals to one.

The description of the RMSE image can be found in Figure 20.

The comparison between modeled and observed by means of scatter plots can be found in Appendix 4. It was done by the following steps (Figure 17):

- The abundance maps (in ENVI format) were exported to ERDAS IMAGING format. In this way they could be used in ArcGIS.
- The preprocessed datasets from Ark, Sykora and the HyEco campaigns were assigned as validation sets. For this the points used as endmembers were extracted.
- This validation sets were used, in ArGIS, to sample values from the abundances maps with the function "Sample". The output was a table with the modeled values for each point from the validation sets.
- The output table was exported to Excel. In Excel the modeled values were confronted to the observed[30] values using scatter plots.

The comparison between modeled and observed by overlying plots to maps can be found in Appendix 5. Plots in Figures 61, 62 and 63 are areas of 5x5 m with more than 50% coverage of that specific PFT. They were created as shown in Figure 10.

The sub-pixel accuracy assessment was applied only to the fully constrained unmixing. The first thing we have to take into account when we calculate OSA is that it is a multidimensional index. It is a unique value to estimate how well all cover types are modeled into a pixel. Although, when the abundance of one class is particularly high in certain plots, the OSA value for those plots can be understood as an estimation of accuracy for such class. Always keeping in mind that such value is influenced by the agreement of all other classes existing in that pixel.

[30] The observed values are the PFTs proportions derived from the validation plots, for the years 2002 (Sykora), 2004 (HyEco'04), 2005(HyEco'05) and 2005-06 (Ark).

Therefore, the ideal situation to calculate OSA would be by sampling all PFT classes together using unique validation points for all. However, the construction of the validation sets in this study makes this impossible. Herbaceous classes (SNL, SNB, RNL and RNB) were validated with plots from 2002, as well as SWS, and each woody class (MWS, RWS, RWSC and RWT) was validated with separate plots from 2005-06.

Consequently, OSA was calculated in the next way:

- Herbaceous classes i (SNB, SNL, RNB, RNL) were calculated together. The expression for OSA was:

$$OSA = \frac{\sum_{i=1}^{i=4} \min\left\{f_i^{\,p}, f_i^{\,o}\right\}}{\sum_{i=1}^{i=4} f_i^{\,o}} \times 100 \qquad \text{i=1,2,3,4}$$

- SWS (*Rubus sp.*) was calculated together with herbs (SNB, SNL, RNB, RNL). As SWS is in all cases conforming the plots in more than 37% (usually more than 62%) this index can be considered as indicative of class SWS. The equation was:

$$OSA = \frac{\sum_{l=1}^{l=5} \min\left\{f_l^{\,p}, f_l^{\,o}\right\}}{\sum_{l=1}^{l=5} f_l^{\,o}} \times 100 \qquad \text{l=1,2,3,4,5}$$

- Woody classes k (MWS, RWS and RWT), except for RWSC, could be joint into one dataset and hence they were estimated together. The equation was:

$$OSA = \frac{\sum_{k=1}^{k=3} \min\left\{f_k^{\,p}, f_k^{\,o}\right\}}{\sum_{k=1}^{k=3} f_k^{\,o}} \times 100 \qquad \text{k = 1,2,3}$$

- The OSA for RWSC was estimated independently since its validation set was independent (Figure 11). The equation was:

$$OSA = \frac{\sum_{t=1}^{t=1} \min\left\{f_k^{\,p}, f_k^{\,o}\right\}}{\sum_{t=1}^{t=1} f_k^{\,o}} \times 100 \qquad \text{t = 1}$$

OSA was also calculated in two different ways for all woody sets. First all plots from the sets were used. Then, the overall accuracy was calculated using exclusively those plots that contained only the species existing in the final pixels selected as ROIs for the unmixing procedure (Table 8). These are: *Rubus sp.* for class SWS, *Sambucus nigra* for class MWS, *Salix fragilis* for class RWS, *Crataegus monogyna* for class RWSC and *Salix alba* for class RWT.

The OSA values can be found in Table 10 and OSA values spatial distribution can be found in Appendix 6 (Figures 69, 70 and 71).

3.3.5 Temporal analysis

The temporal analysis was based on the comparison of the PFTs distribution in the images CASI from 2001 and HyMap from 2004. These two images were object of a spectral mixture analysis by adapting the methodology described before.

Some aspects from the previous methodology were modified in this second analysis. First, the characteristics of CASI image, with 10 bands (HyMap had 126 bands) required a reduced number of endmembers for SMA to be applicable. Second, results derived from previous SMA classification applied to HyMap provided key information to redefine PFTs in order to improve spectral separability of classes. And finally, in order to be more in accordance with future applications (Section 1.1), the new PFTs were redefined taking into account the PFTs as conceived by the DVM SMART2-SUMO2. These are herbs, dwarf shrubs, shrubs, pioneer trees and climax trees.

These modifications had also to take place in a second classification of HyMap image, so that both images could be comparable. The subset extension was also slightly different in this section, because in the complete extension of the CASI image, previous subset was cut.

3.3.5.1 PFT conceptual model

The new PFTs classification was based on the PFTs classification of the SMART2-SUMO2 model with some modifications. The model defines five classes: herbs, dwarf shrubs, shrubs, pioneer trees and climax trees. Concerning adaptation of such classes to the area of study the last class was excluded. Due to the early successional stage of the area, “climax trees” are scarce. Based on conclusions derived from previous results, classes are better classified by spectral techniques when they are species-based (Section 5.1). Therefore, PFTs were defined in this second approach by

giving primary importance to plant species and not to physical measures (such as height or stem diameter).

These new considerations for the PFTs definition can be assimilated into previously defined PFTs model (Section 3.3.2.1) because in a way, all three traits (height, density and flexibility) are species related. Average plant height[31] of *Crataegus monogyna* is 2.16 m, *Sambucus nigra* 3.88 m, *Salix fragilis* 6.98 m and *Salix alba* 8.00m. Density, understood as foliage thickness, as well as flexibility, understood as stiffness, is specific for different species (Anderson, 2006). A detailed study of these two last traits for the species present in the Millingerwaard is out of the scope of this thesis. However, average height can give an approximation of hydraulic resistance. Higher average height of the individuals is assumed to correspond to higher resistance.

Therefore, the new woody PFTs were represented by individuals of the species defined in previous paragraph (the most abundant woody species in the Millingerwaard). Only individuals greater than 0.5 m (previously classified as MWS/T and RWS/T) were chosen as field datasets.

Herbs, however, were not represented by specific species. This would be non sense considering their high heterogeneity. But they were divided into two groups: "Grasses" and "Forbs" based on the spectral variability observed in such class (Appendix 7 Figure 73) and the difference in hydraulic resistance (3.3.2.1). Average height was also calculated for herbs according to literature, 0.56 m for "Grasses" and 0.64 m for "Forbs".

Therefore, five new PFTs classes were defined (Table 11):

- Resistance 1: **Grasses**. Graminoid herbs, including high grasses like *Calamagrostis epigeos*.
- Resistance 2: **Forbs**. All other herbs, including *Rubus sp.*, *Urtica dioica*, etc.
- Resistance 3: **Dwarf shrubs**. *Crataegus monogyna* individuals with medium and robust resistance (MWS and RWS)
- Resistance 4: **Shrubs**. *Sambucus nigra* individuals with medium and robust resistance (MWS and RWS)
- Resistance 5: **Pioneer trees**. *Salix sp.* individuals, including *Salix fragilis* with medium and robust resistance (MWS and RWS) and *Salix alba* with medium and robust resistance (MWT and RWT)

[31] Calculated as the average of all individual plant greater than 0.5 m from Ark dataset

3.3.5.2 Preprocessing

The preprocessing of the images is defined in section 3.3.3.1. As the extension from the two images is not coincident, a new common subset was defined, so that SMA could be applied to the same area.

Regarding the preprocessing of the field data, the new characteristics of the CASI image, with a smaller pixel size than HyMap, and the selection of new PFT classes demanded a new datasets preprocessing following the steps defined in section 3.3.3.2 for Sykora and Ark datasets. Data from the HyEco campaigns 2004 and 2005 was not used.

Concerning woody vegetation two datasets were used: Ark and Sykora. For the preprocessing of the Ark dataset the steps shown in Figure 10 were followed for both images. Three sets for endmember selection and validation were derived for each image: “Dwarf shrubs”, “Shrubs” and “Pioneer trees”. In addition, trees from Sykora datasets were used, but no preprocessing was needed.

Concerning herbs (Sykora dataset), the preprocessing consisted of merging classes SNL and RNL to derive “Grasses”, as well as SNB and RNB to derive “Forbs”. Distance of herbaceous plots to the center of the pixels was considered for HyMap with a threshold of 1.5 m again. For CASI however it was not taken into account since the pixel size is smaller. The furthest the plot could be from the pixel center is 1.4 m and, considering the inaccuracies resulting from the assumptions made before, taking this caution was not considered necessary.

3.3.5.3 Spectral mixture analysis

a) Endmember selection

The new endmembers consist of the five PFT classes defined above and a sand endmember. The endmember selection for each PFT was done following the steps described in Figure 15. The first candidate ROIs were the purest pixels from each PFT. They were picked from different regions in the area, trying to avoid the local spectral effect of specific grouping structures. Details of these ROIs are shown in Table 6. ROIs for HyMap were not completely covered in the case of woody vegetation, while for CASI they were all covered 100% due to the small pixel size (Table 6). The spectral variability within herbs can be observed in Appendix 7 Figure 73 (CASI) and Figure 76 (HyMap). The spectral variability within Pioneer trees (those from Sykora and Ark

dataset) is in Appendix 7 Figure 74 (CASI) and 77 (HyMap). The spatial location of the ROIs in Table 6 can be found in Appendix 7 Figure 78.

The 20 ROIs (HyMap) and 19 ROIs (CASI), defined in Table 6, were merged per PFTs. In the case of HyMap, spectra within classes were similar in the MIR and visible, except for forbs which presented more variability. In the case of CASI, they were closer in the visible, except for dwarf shrubs and the two different sources of pioneer trees. Therefore, five ROIs resulted, each ROI composed of a group of pixels and representing each PFT (in previous approach they were unique pixels). To these 5 endmembers, an endmember for soil was added. This endmember was selected from the sandy area next to the river shore, southwest of the area of study. The mean spectra of these ROIs can be found in Appendix 7 Figure 72 (CASI) and 75 (HyMap).

Table 6: Selection of ROIs for six PFT endmembers (HyMap and CASI). Each ROI consists of various pixels

	HyMap (Pixel of 25 m^2)			***CASI (Pixel of 4 m^2)***		
PFTs	Id	Species & Year	Crown proj. (m^2)	Id	Species & Year	Crown proj. (m^2)
Grasses	263	Carex+Festuca 2002	-	263	Carex+Festuca 2002	-
	2219	Calamagrostis epig. 2002	-	2210	Calamagrostis epig. 2002	-
	282	Calamagrostis epig. 2002	-	282	Calamagrostis epig. 2002	-
	2139	Carex+Fest+Cynodon 2002	-	2139	Carex+Fest+Cynodon 2002	-
Forbs	2122	Geranium+others 2002	-	2122	Geranium+others 2002	-
	2103	Urtica+Thalictrum+others2002	-	2103	Urtica+Thalictrum+others2002	-
	2114	Urtica+Rumex+Heracleum 2002	-	2114	Urtica+Rumex+Heracleum2002	-
	2251	Urtica dioica2002	-	2251	Urtica dioica 2002	-
	2110	Urtica dioica2002	-	2110	Urtica dioica 2002	-
	2232	Trifolium+Circium 2002	-	2224	Rubus sp. 2002	-
	2224	Rubus sp. 2002	-	2232	Trifolium+Circium 2002	-
Dwarf shrubs	784	Crataegus monogyna 2005-06	18	1665	Crataegus monogyna 2005-06	4
	1176		16	1225		4
Shrubs	265	Sambucus nigra 2005-06	23	175	Sambucus nigra 2005-06	4
	91		21	183		4
	560		15	1740		4
	822		15	1163		4
Pioneer trees	2097	Salix alba & Salix fragilis 2005-06	25	3090	Salix alba & Salix fragilis 2005-06	4
	2096		20	3017		4
	362	Salix alba 2002	unknown	359	Salix alba 2002	unknown
	333		unknown	332		unknown
	2236	Salix alba & Salix fragilis 2005-06	16	-	-	-

b) MNF transformation

HyMap was unmixed using 23 MNF bands, as it was before. CASI was also transformed at first. Although CASI image is only consisting of ten bands, they appear in a narrow area of the electromagnetic spectrum and data redundancy is likely to be found. The first eight MNF bands (minimum number to apply unconstrained unmixing) were chosen according to the same criteria used for HyMap in section 3.3.3.2. The eigenvalues from CASI are shown in Figure 18.

However, the result from linear unmixing applied to the eight MNF bands was compared to the unmixing applied to the ten original bands and no improvement was observed. This comparison was done by means of R^2 and adjustment to the one-to-one line. Therefore, the final decision was to use the ten original bands for all three methods: unconstrained, semi-constrained and fully constrained unmixing.

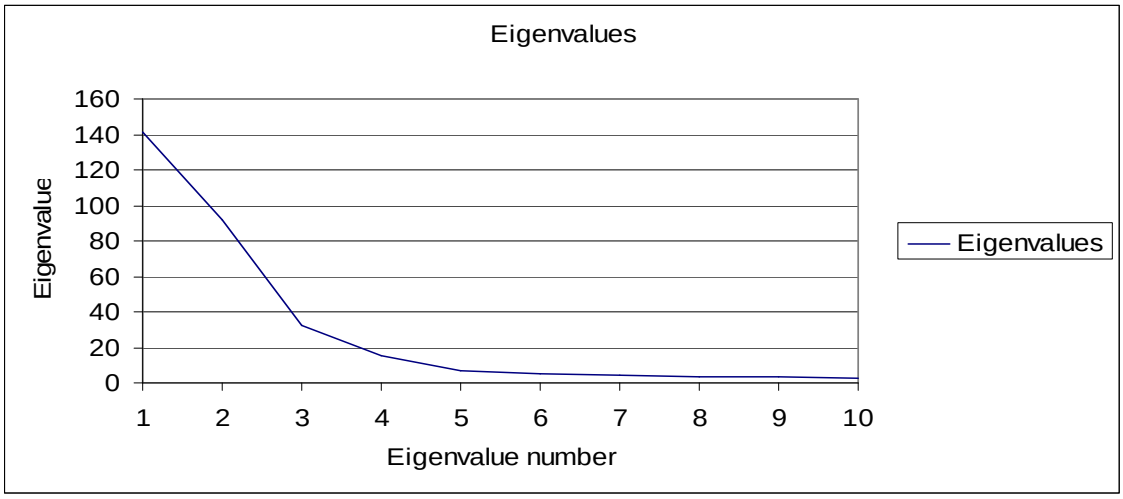


Figure 18: Eigenvalues from MNF transformation of CASI image

c) Linear Spectral Unmixing

HyMap first 23 MNF bands and CASI 10 original bands were subject to a linear spectral unmixing using the six endmembers previously defined. The three methods were again applied: unconstrained, semi-constrained and fully constrained unmixing. The steps followed are shown in Figure 17.

Validation

Validation of results was done following similar methodology as in section 3.3.4.4: analysis of the RMSE image, comparison between predicted and observed coverage (by means of plots and overlaying field data to the abundance maps) and sub-pixel accuracy assessment (OSA).

OSA was calculated using two validation sets, one from 2002 for herbaceous PFTs (grasses and forbs together) and another one, from 2005-06, for woody PFTs (dwarf shrubs, shrubs and pioneer trees together). It was calculated first as the average of all plots and then separately per PFT. This was done by using only the plots covered by such PFT by 50 % or more. Results are shown in Table 12.

4 RESULTS

This chapter presents a summary of the most important results. Many of these results are located in the Appendices. This will be specified in each section.

4.1 PFT classes

The conceptual definition of the PFT classes as defined in Table 3 went through two consecutive modifications during next steps in the methodology. These were first, when classifying Ark field dataset into PFTs (Section 3.3.3.2) and second, when dealing with the spectral characterization of the endmembers (Section 3.3.4.1). The final PFTs defined are detailed in Table 7. Resistance is expressed in separate scales for herbaceous and woody vegetation, being 1 the lowest resistance. The corresponding traits related to each PFT, as defined in section 3.3.2.2, are specified in Table 7.

Table 7: PFTs classes

PFTs	*Resistance*	*Height (Height (m))*	*Flexibility (Stem thickness (cm))*	*Density (Tree, Shrub, Herbaceous)*	*Species*
			Herbaceous		
SNL	1	H1D2F1	≤ 0.5	<1	Graminoids
SNB	2	H1D1F1	≤ 0.5	<1	
RNL	3	H2D2F1	> 0.5	<1	Forbs
RNB	4	H2D1F1	> 0.5	<1	
			Shrubs		
SWS	1	-	-	S	*Rubus sp.*
	2	0-0.5	0-35		*Sambucus nigra*
MWS		0.5-35	0-15		*Crataegus monogyna* *Rosa canina* *Cornus sanguinea*
RWSC	3	0-35	15-130		*Crataegus monogyna*
RWS	5	0-35	15-130		*Salix fragilis* *Sambucus nigra*
			Trees		
RWT	4	0.5-35	15-130	T	*Salix alba*

4.2 Spectral Mixture Analysis

4.2.1 Endmembers

The previous nine PFT classes and soil were used as endmembers for the spectral unmixing of the HyMap image. The final nine vegetation ROIs are described in Table 8. Spectral signatures for all the different ROIs, including sand, are presented in Figure 19.

Table 8: Final selection of ROIs for nine PFTs

PFT	*ROI*	*Plot/ID*	*Year*	*Sour .*	*Cover*	*Species (BB%[32])*	*OtherPFT*	*Dist*
SNB	SNB2	Plot 232	2002	Syk	68%	75-100Trifolium,12-25 Cirsium	20% SNL 6% RNB	1.41
SNL	SNL1	Plot 63	2002	Syk	67%	25-50Carex-hirta,25-50Carex-arenaria,25-50Festuca	19% SNB	1
RNB	RNB3	Plot 110	2002	Syk	77%	75-100Urtica	16% SNB	1.41
RNL	RNL1	Plot 219	2002	Syk	53%	50-75Calamagrostis	20% SNB 17% SNL 8% RNB	1.41
SWS	SWS	Plot 224	2002	Syk	63%	75-100Rubus	36% RNB	1.41
MWS	MWS1	ID 2041	Ene06	Ark	86%	5Sambucus	Herbs	
RWS	RWS2	ID 746	Mar06	Ark	99%	1Salix fragilis	1% MWS	
RWSC	RWS4	ID 694	Mar06	Ark	100%	3Crataegus	Herbs	
RWT	RWT2	ID 206	Feb06	Ark	100%	2Salix alba	Herbs	

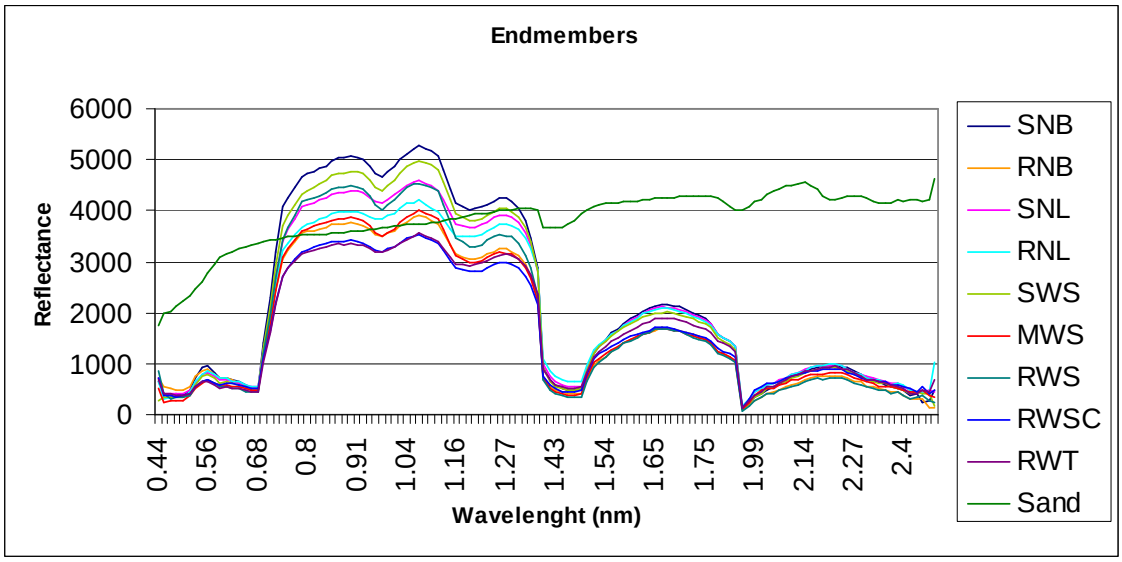


Figure 19: Endmember spectra for sand and nine PFTs identified in the Millingerwaard

4.2.2 Abundance maps

Linear unmixing was applied using the ten identified endmember spectra (Figure 19) and abundance maps for the PFTs in the Millingerwaard were produced. The abundance maps are presented in Appendix 3. For the unconstrained and semi-constrained unmixing, eleven maps were derived. For the fully constrained unmixing, thirteen maps were derived (including exitflag and hard classified map). The first ten

[32] Braun Blanquet abundance

maps in both cases are abundance maps, each fractional map is related to one endmember, and it contains the modeled fraction[33] per pixel of such endmember (Figures 38-47). Bright areas represent higher fractions and dark areas lower fractions. The eleventh map is the RMSE, which is given per pixel (Figure 48). Brighter areas represent areas with higher RMS error. The other two maps, exclusive outcome from the fully unmixing algorithm, are the exitflag map and the hard classified map (Figure 49). The exitflag map has a value of 1 for all pixels which means that the function converged to a solution in all pixels.

4.2.3 Validation

4.2.3.1 RMSE analysis

The RMSE images for both, the unconstrained and constrained unmixing approach, presented comparable spatial error patterns (Appendix 3 Figure 48). Slight variations can be observed, but the highest RMSE values are always found in the same regions. In Figure 20 the lighter regions with a high RMSE are indicated.

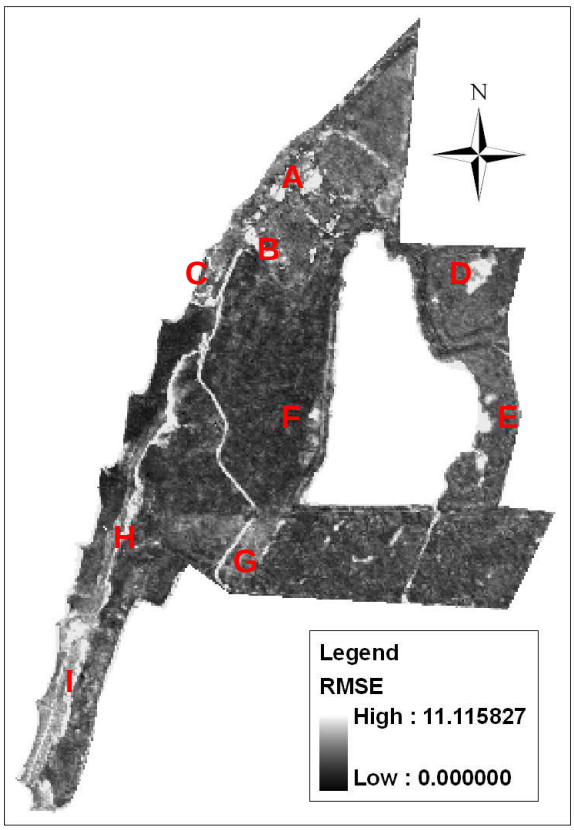


Figure 20: RMSE image from HyMap unconstrained unmixing (using 10 endmembers)

[33] The outcome fractions are expressed in a proportion out of one.

Table 9: Cover type in regions with high RMSE

Region	*Cover type in regions with high RMSE*
A	Area with patches of forest which relates to bigger shadows, also some buildings and areas richer in uncovered soil. This forest patches have a different structure than the forest in region J from where the forest endmember (RWT) was selected and also other species such as *Populus sp*
B	Soil and buildings
C	Presence of forest and shrubs combined with sand
D	Wet soil and water
E	Wet soil and water
F	Presence of herbaceous species such as *Mentha aquatica* and *Lycopus europaeus*
G	Patchy area of *Galium*, *Urtica*, *Ranunculus*, *Alopecurus*, *Trifolium*, etc
H	Area rich in species such as *Lolium*, *Bromus*, *Potentilla*, *Eryngium*, *Medicago* and *Avena*
I	Aquatic vegetation (*Polygonium Amphibiun*) and *Chenopodium rubra* strongly mixed over sand dunes
Other	Paths and borders

The lighter regions with high RMSE in Figure 20 were compared with the vegetation map of the Millingerwaard (Van Geloof & de Ronde 2002) and other sources such as an aerial photograph from 2003. The cover types existing in them are presented in Table 9.

4.2.3.2 Comparison between modeled and observed cover

a) Presentation in scatter plots

Scatter plots representing modeled vs. predicted coverage can be found in Appendix 4. This appendix is divided into three sections: unconstrained (Figures 50-53), semi-constrained (Figures 54-57) and fully constrained unmixing (Figures 57-61). In each section four figures are found. In the first figure, herbaceous PFTs modeled abundances are compared to observed abundances from 2002. In the second figure, woody PFTs modeled abundances are compared to observed abundances from 2002 (SWS) and 2005-06 (the rest). In the third figure herbaceous modeled abundances are compared to observed abundances from 2004 and in fourth figure herbaceous modeled abundances are compared to observed abundances from 2005.

For the representation of the observed vs. modeled coverage by means of scatter plots there are two different aspects of interest. First, the adjustment of the points to the regression line which is measured by R^2 and secondly, the adjustment of the regression line to the one to one line. The one to one line is the ideal line because it represents

modeled coverage equals to observed coverage. This line has its origin in zero and has a slope of forty-five degrees. This equation is y = x.

b) Overlaying field data to the abundance maps

Field data overlaid to abundance maps can be found in Appendix 5. Figures 62-64 present all PFT abundance maps overlaid to the validation plots. Figures 65-68 present abundance maps of MWS, RWS, RWSC and RWT, overlaid first with all individual plants that belong to such PFT class, and then, overlaid to individual plants from the species used as ROI for the unmixing procedure (Table 8). This is, for instance, MWS overlaid first to plants from class MWS, and then, to individuals from *Sambucus nigra*.

By overlaying field data to the abundance maps we can deduce general trends of agreement or disagreement using a visual approach. As the constraining did not make a noticeable difference, this overlaying was done only over the unconstrained unmixing fractional maps.

4.2.3.3 Sub-pixel accuracy assessment by OSA

The average OSA value and its standard deviation are shown in Table 10 for herbaceous vegetation and woody vegetation separately and finally together. These two groups are subdivided into PFTs as was explained in section 3.3.4.4.

Table 10: OSA value for HyMap fully constrained unmixing with 10 endmembers

Vegetation types	*Most abundant PFTs*	*Year of validation plots*	*Selection of validation plots*	*OSA Mean*	*OSA Stdev*
Herbaceous vegetation	SNL,SNB,RNL & RNB	Plots 02	All	39.12	21.44
		Plots 04	All	35.61	28.99
		Plots 05	All	36.42	27.26
			Average all plots	**38.79**	**21.85**
Woody vegetation	SWS	Plots 02	All	52.72	14.19
			≥ 62% *Rubus sp.*	55.25	10.63
			< 62% *Rubus sp.*	49.18	18.51
	MWS	Plots 05-06	All	40.57	39.85
			Only *Sambucus n.*	54.33	41.03
			Mixed *Sambucus n.*	43.23	36.81
			Other species	0.00	0.00
	RWS	Plots 05-06	All	37.07	38.70
			Only *Salix fragilis*	71.29	6.97
			Mixed *Salix f.*	65.22	33.77
			Other species	11.59	28.39
	RWSC	Plots 05-06	RWSC plots	40.90	40.06
	RWT	Plots 05-06	All (*Salix alba*)	35.94	33.22
			Average all plots	**42.24**	**34.31**
			Average purest plots	**44.23**	**37.74**
All			***Average all plots***	**40.06**	**27.16**
			Average purest plots	**41.30**	**30.34**

In the case of woody vegetation the OSA value was calculated first using all plots from such PFT and then selecting only those plots covered exclusively by the species used as ROI for the unmixing procedure (Table 8). Table 10 will be discussed in chapter 5.2.4.

In addition, with the aim to find spatial patterns on OSA values, the OSA values per plot were overlaid to the image of the Millingerwaard (Appendix 6). Figure 69 shows distribution of OSA value for herbaceous plots. Figure 70 shows two images, one with the distribution of OSA for SWS plots and another one for plots of MWS, RWS and RWT together. Figure 71 shows distribution of OSA values for plots of RWSC.

4.3 Temporal analysis

4.3.1 PFT classes

For the temporal analysis, PFT classes were redefined as described in section 3.3.5.1. The new PFTs classification is shown in Table 11. Resistance goes from 1 (lowest resistance) to 5 (highest resistance). Then, we have the plant species and average height per PFT. Regarding density and flexibility the information is incomplete, but as explained in section 3.3.5.1 measurements are not available. Finally the new PFTs created for the temporal analysis are related to the previously defined PFTs in Table 7.

Table 11: PFTs classes for the temporal analysis

PFTs	*Resistance*	*Plant Species*	*Height*	*Density*	*Flexibility*	*Previous PFTs*
Grasses	1	Graminoids	0.56 m	Long leaves	-	All SNL & RNL
Forbs	2	Non-graminoid herbs	0.64 m	Broad leaves	-	SNB, RNB & SWS
Dwarf shrubs	3	*Crataegus monogyna*	2.16 m	--	-	*Crataegus monogyna* from MWS & RWS
Shrubs	4	*Sambucus nigra*	3.88 m	-	-	*Sambucus nigra* from MWS & RWS
Pioneer trees	5	*Salix alba* & *Salix fragilis*	8.00 m and 6.98 m	-	-	*Salix alba & fragilis* from MWS & RWS

4.3.2 Spectral Mixture Analysis

The 6 endmember spectra (5 PFTs and sand) found in Appendix 7 (Figure 72 and 75) were used for the unmixing of HyMap and CASI images respectively. The abundance maps produced by the unconstrained unmixing can be found in Appendix 8. Figure 79 shows abundance maps for all five PFTs for HyMap and CASI. Figure 80 shows abundance map for Sand and RMSE for HyMap and CASI. The abundance maps produced by semi-constrained and fully constrained unmixing are not shown since they did not present relevant improvements. In addition, the hard classified images for both, HyMap and CASI, produced by fully constrained unmixing are shown in Figure 81.

The validation of the unmixing results was done by analyzing the RMSE image, presenting modeled vs. predicted fractions by scatter or linear plots, overlying plots to abundance images and by the OSA value. The results are summarized in next paragraphs.

The RMSE maps can be seen in Figure 21. Concerning the HyMap image (Figure 21.a) the same comments made in 4.2.3.1 Table 9 are applicable, because this error image is very similar to the error image produced by the previous HyMap unmixing using 10 endmembers. Concerning the CASI image (Figure 21.b) we can observe a very high RMSE value. The maximum RMSE is 389, while the mean is 5 (the maximum RMSE value for HyMap was 12 and the mean 0.4). Values greater than 10 are the 19 % of the total, values greater than 50 are the 1 % of the total. In addition to the regions with RMSE values encountered for HyMap in 4.2.3.1 Table 9, that are again present in CASI (Figure 21.b), other important regions show a high error. These are regions A, B and C in Figure 21.b.

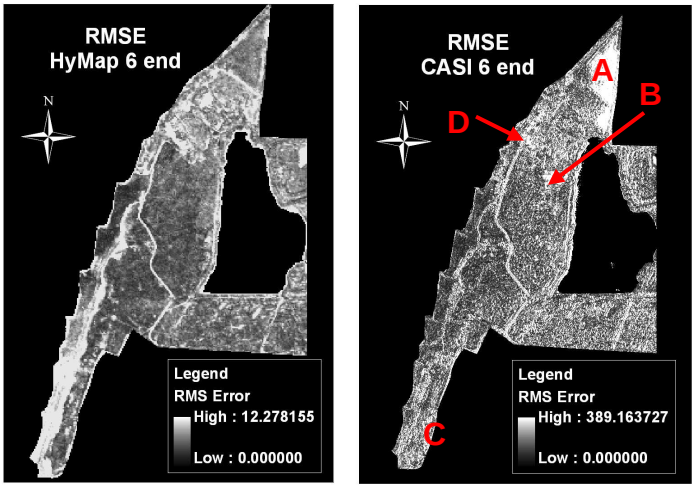


Figure (a) **Figure (b)**

Figure 21: RMSE image from HyMap and CASI unconstrained unmixing (using 6 endmembers)

Region A is a grazing area. Region B and C are regions rich in *Crataegus monogyna*, individuals of *Salix alba* and abundance of woody species not considered as endmembers, such as *Rosa canina*, *Acer pseudoplatanus*, *Salix cinerea*, *Salix viminalis*, etc. In addition, in those region plants are grouped in a more loosely configuration than in other regions. Especially very high errors are found in region D where houses are situated. As a general trend, high errors are found where big trees are located (of any of the species).

The comparison between modeled vs. predicted fractions by means of scatter or linear plots can be found in Appendix 9. Figure 82 shows 5 scatter plots from HyMap unconstrained unmixing, one per PFT. Figure 83 shows 2 scatter plots from CASI herbaceous PFTs (grasses and forbs) and 3 linear plots from CASI woody PFTs (dwarf shrubs, shrubs and pioneer trees). These linear plots were used instead of scatter plots because the crown projection of woody species on CASI validation plots were in all cases greater to 4 m^2 (the pixel area), therefore for all plots the expected modeled coverage was 1. Therefore, if the modeled fractions were equal to the observed fractions (ideal situation) the red line in Figure 83 should be the equation y = 1.

The observed values overlaid to the abundance maps are in Appendix 10. The abundance maps from HyMap unconstrained unmixing are shown in Figures 84 (grasses, forbs and dwarf shrubs) and 85 (shrubs and pioneer trees). The abundance maps from CASI unconstrained unmixing are shown in Figures 86 (grasses, forbs and dwarf shrubs) and 87 (shrubs and pioneer trees).

Finally, OSA was calculated for both images using the results from the fully constrained unmixing (Table 12). It was done separately for herbaceous and woody vegetation as explained in section 3.3.5.3.

Table 12: OSA value for HyMap and CASI fully constrained unmixing with 6 endmembers

Vegetation types	*Most abundant PFTs (≥50%)*	*Year of validation plots*	*OSA Mean*		*OSA Stdev*	
			HyMap	**CASI**	**HyMap**	**CASI**
Herbaceous vegetation	Grasses	Plots 02	48.90	39.10	20.20	28.71
	Forbs	Plots 02	51.29	38.48	27.43	31.24
		Average herbs	**47.91**	**41.25**	**27.09**	**30.16**
Woody vegetation	Dwarf shrubs	Plots 05-06	57.70	4.00	37.05	11.01
	Shrubs	Plots 05-06	81.08	41.64	33.63	38.90
	Pioneer trees	Plots 05-06	64.08	41.86	33.79	71.70
		Average woods	**68.01**	**21.15**	**43.92**	**34.57**
All		**Average all**	**66.98**	**21.87**	**43.44**	**34.62**

5 DISCUSSION

This chapter is organized following the structure of the previous chapter and, at the same time, addressing the research questions posed in section 1.3. Each of the answers to the research questions can be found in this chapter and they will be summarized at the end of each section.

5.1 PFT classes

A first view on the study area of the Millingerwaard indicates the high complexity of its vegetation composition. Intimate mixture of plants from different species, ages, sizes, and structural types may be encountered. This complexity is favored by the presence of cattle and horses in a low density, which allows a highly varied vegetation to develop and creates grazed areas unevenly distributed (Schmidt et al., 2005). When referring to classification into plant functional types as defined in this study, vegetation complexity in floodplains is meant to be simplified into a number of PFT classes with common hydrological roughness. But we must consider that PFTs, as acting and reacting units in vegetation dynamics to changing environmental conditions, may not be easily observed in nature (Skarpe, 1996). Despite of this fact vegetation was classified into PFTs in the Millingerwaard. A group of plant characteristics was considered to be relevant for the definition of these classes (Section 3.3.2.1). At the highest level, the woodiness of the plant was used as main plant trait for a first division into PFTs. At the next level, other characteristics were considered relevant; height, stem diameter, and branching for woody vegetation and height and leaf type for non-woody vegetation. These characters are relevant to determine hydraulic roughness of vegetation according to literature (Anderson et al., 2006; Naden et al., 2006). However, some important aspects are missing, such as inherent species flexibility, distribution of plants in a stand, stand compaction and orientation of the plant with respect to the local flow direction (Anderson 2006). Regarding species flexibility, a deeper study of mechanical properties of each of the species found in the Millingerwaard was considered out of the scope of this thesis. Stand structure is an important factor to account for. The effective drag coefficient for one stem within a group of stems is less than that for an isolated stem due to sheltering effects (Anderson 2006), therefore roughness synergize when plants are grouped. However, stand characteristics are not possible to derive by using spectral information exclusively. But group structure can be mapped by using LIDAR in

combination with spectral methods (Geerling et al., 2006). The orientation of the plants with respect to the local flow direction is not an inherent characteristic of the vegetation, instead it implies the interaction between earth's physiognomy, flow direction and vegetation distribution, therefore it cannot be integrated into the definition of PFTs. Recognizing all previous plant traits as relevant, by combining all of them and assuming the high heterogeneity of the area, a high number of PFTs is expected (Table 7). A smaller set of classes would have also been appropriate. To estimate hydraulic resistance for river management purposes, five classes are considered to be enough (Geerling et al, 2006). This number would have also been enough for the application into the model chain SMART2-SUMO2 in which only five PFTs are considered (Wamelink et al, 2005). However, the above reasoning –relevant plant traits and heterogeneity- together with the application of linear spectral unmixing, which requires an exhaustive registration of all endmembers present in the area (Malenovsky et al., 2006), led finally to the selection of nine PFT classes (Table 7).

Plant functional types were in first instance meant to be composed of a heterogeneous set of species with a similar hydraulic roughness. Hydraulic roughness was characterized by physical characteristics (Section 3.3.2.1): height, stem diameter, branching and leaf type. This characterization of hydraulic roughness and therefore, PFTs, can be summarized as size-based. But facing the spectral characterization of PFTs (Section 3.3.4.1) it became a fact that groups based on plant size were not represented by unique or similar spectral signatures. Furthermore, there was a high variability within signatures representing one unique class, as it may be observed in the ROIs spectra shown in Appendix 2. This could be expected based on the evidence that signatures of vegetation are related to properties as leaf thickness, leaf surface structure, chlorophyll-carotenoid content, dry matter content and leaf internal structure (Kumar et al., 2001), as well as reflectance in the canopy level (Law et al., 1994); canopy density and canopy structure (Rosso et al., 2005). These properties are interrelated to floristic composition (Schmidtlein & Sassin, 2002) and not to plant size. SMA has the potential to estimate the spatial distribution and abundance of a species, rather than thematic classes (Rosso et al., 2005). It has been used by several authors to discriminate species at the sub-pixel level (Rosso et al., 2005; Li et al, 2005). It can also be used to identify clusters at a larger scale; hygrophytic vegetation was mapped by Schmid et al. (2005) and Powell et al. (2006) mapped vegetation present in urban land cover. But in these cases vegetation was mapped against soil or impervious surface. Although separability

by SMA appears to perform better when it is species based, some authors claim that compared to many other vegetation attributes, plant species composition is difficult to detect by remote sensing techniques (Schmidtlein & Sassin, 2002).

Once the endmember selection was carried out it became a fact that PFT classes were represented by ROIs with presence of unique species (Table 8). This was contrary to the first approach of size-based PFTs characterization. It was more in accordance with PFTs as vegetation clusters based on a division by species, species-based PFTs. This was the consequence of the high variability encountered within the spectra of ROIs representing same PFT. This spectra variability was found to be strongly correlated, in most cases, to species because ROIs containing same species were spectrally similar while ROIs containing different species were unlike each other (Appendix 2 and Table 5). So in search for unique spectra representing each PFT and taking into account that some species were especially abundant in the area and, hence, more important to be considered by river managers, a group of ROIs was selected representing PFTs as unique species. This fact became noticeable all throughout the validation results (Section 3.3.4.4). In Figures 65-68 from Appendix 5 we observe that agreement decreases when the abundance map of one PFT (e.g. MWS) is overlaid to individual plants classified as such PFT (e.g. MWS), and increases when it is overlaid to individual plants from specific species (*Sambucus nigra*). Being these species in all cases, those species present in the ROIs used for the endmember definition during the unmixing procedure (Table 8). In Table 10 (Section 4.2.3.3) we can observe also how the OSA value increases in all cases when it is calculated as the average of those plots with pure presence of such species.

A first conclusion from this first part of the study is that trying to map vegetation clusters as closed blocks determined by size-based PFTs using linear spectral unmixing does not give good results. However, species-based PFTs give better results. Therefore we can say that vegetation classification via spectral unmixing should be done in species-based clusters. This may be understood also as an answer to the first research question: “How to discriminate plant functional types in a floodplain using spectroscopy techniques?”

5.2 Spectral Mixture Analysis

The spectral mixture analysis approach for the classification of the HyMap image was developed taking several assumptions into account. The first assumption regards the linearity of the spectral mixing model. When we apply SMA we assume that the spectrum measured by a sensor is a linear combination of the spectra of all components within the pixel (Borisova, 2005). However this linearity is only true when photons interact with one material instead of several materials. It happens in general with materials such as snow or bare soil but not always with vegetation. Only part of the light is directly reflected by vegetation, the rest is absorbed or transmitted to other layers, from which it can be reflected to the sensor again. This results in a reflection that interacts with more than one object and a non-linear spectral mixing (Malenovsky et al, 2006), which is the real situation in the Millingerwaard.

The other three assumptions are related to field data and might have led to uncertainty in the final results. The first one is that field data from the years 2002, 2004, 2005 and 2006 were used as training and validation sets to classify a HyMap image from the year 2004. This assumption seems to be especially problematic when dealing with the endmember selection from the image (Section 3.3.4.1). During this procedure, the evidence of existence of PFT classes at specific sites is based on field knowledge from the years 2002 and 2005-06. This implies that the key step of the image unmixing is base on important inaccuracies. If the coverage of a reference location from 2002 used as endmember (e.g. RNB) was not RNB in 2004 anymore but for example RWT, the whole system of equations is confused and the inaccuracies are extensive to the rest of PFTs results. Regarding medium and robust woody vegetation this assumption is acceptable because this data was derived from the years 2005-06, and we can assume that its distribution has not dramatically changed in a period of one or two years. Nevertheless, herbaceous vegetation as well as small woody vegetation was derived from the year 2002, and taking into account the rapid evolvement of this plant types, an important variation in spatial distribution in a three-year period is expected. These facts are supported by the scatter plots shown in Appendix 4 and Appendix 9. In general woody classes (except for RWT) show a better trend than herbaceous classes. The average OSA value (Table 10) as well, is lower for herbaceous (38%) than for woody vegetation (42%). However, for the case of small woody vegetation (SWS), also from 2002, the OSA value is relatively high, 52%. We must take into account that, in this case, only *Rubus sp*. was into class SWS. This means that the definition of SWS was

already in accordance with the concept of species-based PFTs and a better accuracy can be expected.

The second assumption refers to the purity of the pixels used as ROIs for the endmembers setup. As it may be observed in the final selection of endmembers (Table 8) the purity of the ROIs is rarely 100%, moreover we encounter quite low values as 53%, 63% or 67%, while we are assuming that they are completely pure. This is a fact derived from the available field information in combination with the pixel size of the HyMap image. This could be solved by, either executing field campaigns in accordance with this application, hence gathering accurate information of PFT pure plots, or by making available a reliable PFTs endmembers' library.

The third assumption is done when facing the need to relate the spatial sampling unit of the available field data with the spatial condition of the pixels in the image. The coverage of those plots (Sykora and HyEco'04 datasets) smaller than the pixel size it is assumed to be proportionally extensive to the area of the whole pixel where it is located, although the majority of those plots are not even centered with respect to the pixel's center. Also, the coverage of those plots (HyEco'05 dataset) with the same size of the pixel but not centered with respect to the center of the pixel it is assumed to be extensive to the whole pixel where it falls in. And the same assumption is done when presuming that the entire crown projection of those individuals whose stem coordinates are located into one pixel belong also to that pixel (Ark dataset). There is not a clear correlation between some of these factors and accuracy of results which may indicate that this inaccuracy is not very relevant.

5.2.1 Abundance maps

The results of the spectral unmixing, in the abundance maps, are expressed as abundance values. These values should lie, in theory, in the range between 0 and 1. When the unconstrained model was used, proportions were out of this range as we can see in Appendix 3 (unconstrained and semi-constrained abundance maps). This may be due to one or more of the following reasons: 1) the linear mixture model does not adequately fit the data; 2) the endmembers are badly chosen and do not represent the extremes of the distribution of reflectance values for that endmember; 3) the number of endmembers is not sufficient to describe the data set (Schmid et al., 2005). A combination of these factors may explain the inaccurate results obtained in this study. As argued before, when referring to vegetation, a non-linear mixture model is expected,

while in this study, the unmixing function was assumed to be linear. Non-linear models are still under study. The paper from Nielsen (2001) shows that, although not fully satisfactory, the semi-parametric model gives better estimates of endmember abundances than the linear model (using arbitrary endmembers from AVIRIS data). Due to previous spatial and qualitative inaccuracies (such as pixel coverage) described in previous paragraphs, the final ROIs chosen (Table 8, Figure 19) were not the ideal spectral representation of the endmembers. The number of endmembers, ten, does not seem to be small to represent the area of study. Li et al. (2005) used nine endmembers to map marsh species, good accuracy for some species was found by performing a qualitative comparison. Rosso et al. (2005) defined eight endmembers to map species in marshland vegetation with correlations around 0.50 and Schmid et al. (2005) used five endmembers to map hygrophytic vegetation with low RMSE and fractional values between zero and one. The problem was more into the conceptual model of PFTs (Figure 5) and how this model was applied to the definition of PFTs in the Millingerwaard (Section 3.3.2.2) than into the number itself. Species-based PFTs were more appropriate than size-based PFTs in terms of spectral separability as discussed in section 5.1. In a sense the spectral variability observed within PFT classes is not caused because the number of endmembers is not enough, but because the clustering was not the appropriate. The Jeffreis-Matusita separability measure could have been used to compute ROI spectral separability in ENVI.

In the comparison between the abundance maps (Appendix 3) we can observe how higher fractional values are complementary between maps of different PFTs. This is a result of the linear spectral model itself. This can be clearly observed between classes such as SNB (Figure 38) and SNL (Figure 39), or MWS (Figure 43) and RWSC (Figure 45).

Regarding the hard classified image derived from the fully constrained unmixing (Figure 22) the next things are observed. The patches of herbaceous classes are well located. Nevertheless, the accuracy in the distribution of each of the classes SNB, SNL, RNB and RNL separately is uncertain. SWS (*Rubus sp.)* encountered in the centre of the study area (A) is also feasible. MWS (*Sambucus nigra*) is well mapped locally (B), but some areas were overestimated, like the area (C) west to the forest. The wide forested area (D) in the south of the lake was mapped as RWSC (*Crataegus monogyna*) when it is mainly RWT (*Salix alba*). This confusion was originated during the endmembers spectra definition; we observe in Figure 19 how the RWSC spectrum is

similar to the RWT spectra. This may explain also why RWT class is underestimated. The sandy area (E) southwest of Figure 22 with *Quenopodium rubra* was the only important area mapped ass RWT. The forested area in the north (F) was classified as RWS (*Salix fragilis*) confused with *Salix alba* and other tree species, but it was locally well mapped in the western part of area (E).

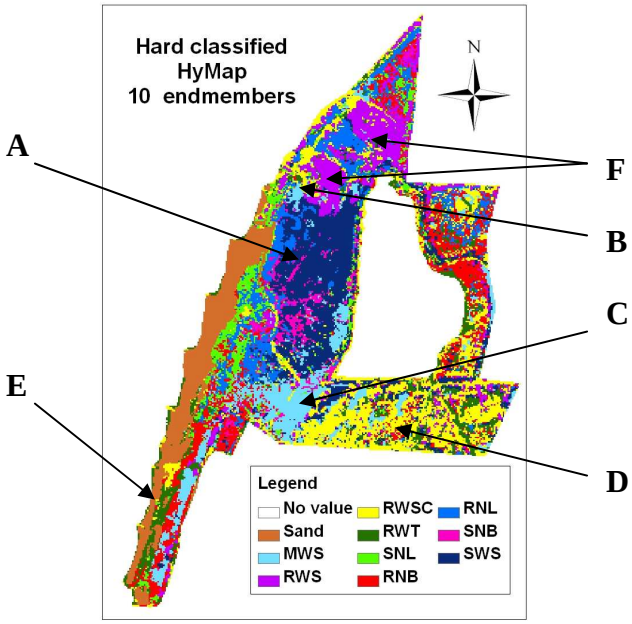


Figure 22: Hard classified image from fully constrained unmixing

5.2.2 RMSE analysis

Regions with high RMSE values are regions that could not be explained by means of the selected endmembers (Figure 20). This is due to the existence of other land covers than those represented by the endmembers, or to the influence of other factors that modify the expected reflectance of the cover types represented by the regions of interest, such as moisture or shadow. Shadow is an important component to account for, Powell et al. (2006) applied MESMA using paired endmembers, shadowed and not, with accurate results. To avoid effect of shadows it is recommended to combine unmixing methods with the use of LIDAR or other IR based images (Geerling et al., 2006).

From the comparison between regions with high root mean square error and field knowledge, possible reasons of endmember absence can be derived. Referring to Figure 20 and Table 10, region A corresponds to an area with patches of forest with a different structure from that of the forest in region J from where the forest endmember (RWT) was selected and also with presence of other species such as *Populus sp*. Areas with

forest are also related to presence of shadow area. The shadow component was not accounted for as endmember; pure pixels representative of shadow areas were not encountered due to the relatively large pixel size (5x5 m). Buildings and areas with uncovered soil found in regions A and B were not considered as endmembers either and hence a high error in these areas was expected. Region C is a sand dune with scarce presence of forest and shrubs; the intimate spectral mixture of these endmembers with shadow might be causing higher errors. The structure of the forest here is also different. Other regions such as D, E and borders are wet areas with soil richer in moisture. Area F, with presence of herbaceous species such as *Mentha aquatica* and *Lycopus europaeus*, can be explained with a high error since these species were not considered as endmembers, due to the high variability of species, the plant functional types were finally represented by a small set of species. The same occurs with region H; species such as *Lolium perenne*, *Bromus hordeaceus*, *Potentilla sp.*, *Eryngium campestre*, *Medicago sp.* and *Avena sp.* are not included as endmembers. Region G is a patchy area of *Galium sp.*, *Urtica dioica*, *Ranunculus sp.*, *Alopecurus sp.*, *Trifolium sp.*, etc. Although some of these species are accounted for as endmembers -such as *Urtica dioica* and *Trifolium sp*.- they are not very abundant -presence of *Urtica dioica* is under 15%[34] in all pixels of this area-. Aquatic vegetation such as *Polygonium Amphibiun* and dune vegetation such as *Chenopodium rubra* were not represented as endmembers either and this can explain the high RMSE existing in region I. All paths in the image have also a high RMSE. Paths are mainly sandy but as the pixel size is quite broad the reflectance in these areas is an intimate mixture of sand and a wide variety of the species existing in the borders of the paths.

5.2.3 Comparison between predicted and observed cover

Presented in scatter plots

The comparison between predicted and observed coverage by means of scatter plots shows a low agreement with low R^2 values and bad adjustment to the one to one line (Appendix 4). There is not a clear improvement while constraining the unmixing method. Appendix 4 is divided in three parts: unconstrained, semi-constrained and fully constrained unmixing where this can be observed. The fact that fractional values are commonly negative or greater than 1 when unconstrained unmixing is applied indicates

[34] This percentage is considering the simplification of the Braun-Blanquet index explained in section 3.3.3.2 for Sykora dataset

that either the endmembers chosen are not good to represent the area or that the linearity of the unmixing process is not possible. This was already discussed in section 5.2.1. When constraints are applied, the unmixing process is still unfit; however fractions are mathematically forced to be between 0-1 and to sum up to 1 in each pixel. Therefore, senseless results can be expected.

Herbaceous observed fractions (Figure 50) show normally lower agreement with modeled fractions than woody vegetation (Figure 51), except for class RWT. As it was established in section 5.2, the assumptions regarding the time period existing between the moments when field data was gathered and when the image was taken was more acceptable for sets containing woody vegetation. On the other hand herbaceous species composition is much more varied than that of woody vegetation leading to a more complex identification in the unmixing process. Going deeper into herbaceous vegetation we can observe that agreement is better in most cases for grass type herbs (SNL and RNL) (Figure 50), this is explained by the fact that forbs present much more heterogeneity regarding leaf type and plant structure. Performance improved when validation sets from the years 2004 and 2005 were used (Figure 52 and 53). This coincides with the year when the image was taken. Specific data about grazed areas would have been useful for a better mapping of the herbaceous types. Grazed areas appear to have a much higher LAI (so apparently a much higher RSR[35]) at a given biomass compared to the ungrazed ones (Schmidt et al., 2005).

Regarding woody vegetation, the best prediction was found for classes MWS, RWSC and RWS (Figure 51). MWS has a positive trend, although R^2 is very low (0.07) showing a high dispersion. Its prediction improves in the area where the endmember was selected from, but is worse for the rest (Figure 23). This may be due to spectral influence of other herbaceous coverage into the endmember chosen (MWS ROI was covered by Sambucus in 86%, Table 8) which might be coincident for all pixels around that ROI with presence of *Sambucus nigra*. But it can be also a consequence of specific grouping structure in different areas (Figure 23).

[35] Remote Sensing Reflectance

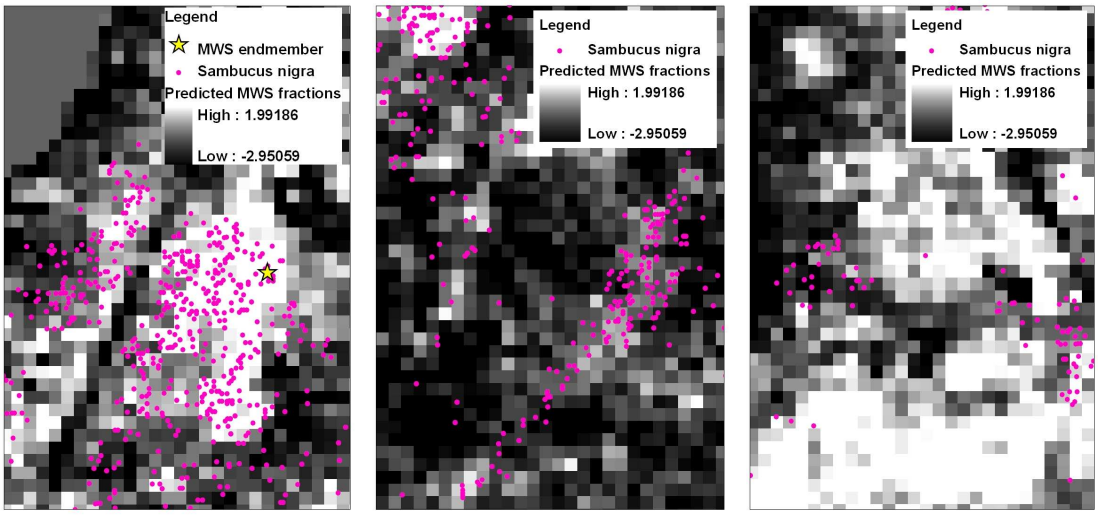


Figure (a) **Figure (b)** **Figure (c)**

Figure 23: Picture (a) details the region where the endmember MWS was selected; we can observe a better trend in the agreement between modeled and observed. In picture (b), which is not very far from the region (a), the agreement is lower, although we can still observe certain trend. In picture (c) there is not agreement and the trend is opposite to what it is expected

RWSC shows a negative trend when validated with RWS plots (Figure 51), composed mainly by *Salix fragilis* and *Sambucus nigra*, but it is positive when it is validated using plots composed exclusively by *Crataegus monogyna* of all sizes (Figure 51). RWS shows also a positive trend (Figure 51) and the R^2 is higher (0.24). This class endmember was represented by a 99% pure pixel of *Salix fragilis* (Table 8) which may explain this improvement.

Class SWS show a low R^2 value in the unconstrained (Figure 51) and fully constrained (Figure 59) but improves in the semi-constrained unmixing (Figure 55) by reaching the value of 0.21. We can observe in the scatter plot (Figure 50) that the SWS fraction is overestimated for low values of coverage and underestimated for high values. However, using other validation methods (Section 5.2.4) SWS distribution appears to be more accurate.

The RWT distribution is badly predicted (Figure 51). As found for class MWS this class is exclusively well mapped in the region where the endmember was selected (Figure 24.a) and into a small region close to it with similar characteristics (Figure 24.c).

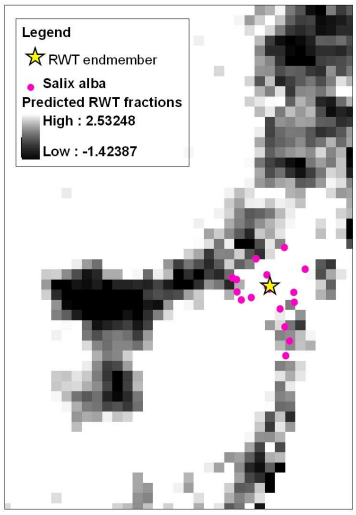


Figure (a)

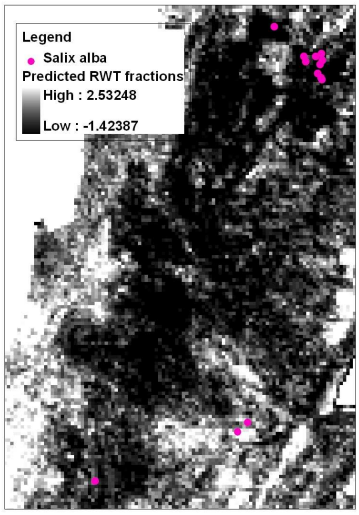


Figure (b)

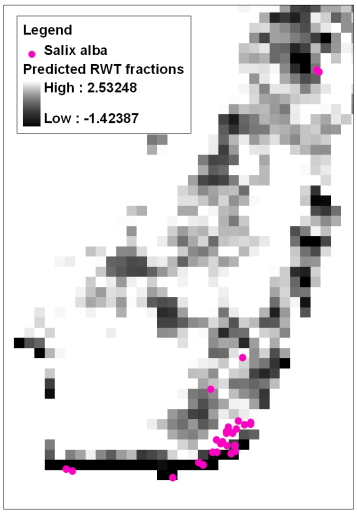


Figure (c)

Figure 24: Picture (a) details the region where the endmember RWT was selected; we can observe a medium agreement. In picture (b), the agreement is low. In picture (c) there is a partial good agreement, we can observe a very bad agreement in the edges, this can be explained by the stronger presence of water in this edges.

But the mapping is negative for many other wide regions in which *Salix alba* is present, this can be explained by the existing understory species that influence the pixel overall spectrum and also the specific stand structure in the different regions. To get a better mapping of forested areas the endmember for RWT class should have been chosen from other areas, like the wide forested area under the lake (Figure 25 A). This could have been done based on field data from 2002 which was gathered from this forested area. This data was not used because it was poorly documented regarding species and plant sizes.

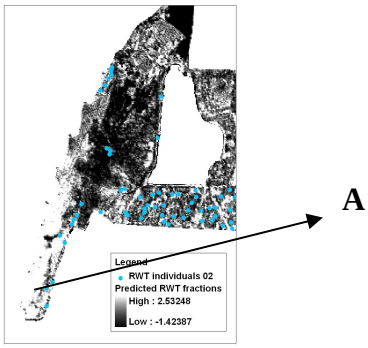


Figure 25: RWT individual from 2002 overlaid to the RWT abundance map

Presented by overlaying plots to maps

From the observation to the overlay between plots and maps we may derive similar trends to what was discussed in the previous section (scatter plots) but also some new aspects are emphasized now.

Regarding non-woody vegetation, pictures showing grass type herbs (SNL and RNL) show a higher agreement than forbs (SNB and RNB) (Figure 62), as observed in the scatter plots. However, small woody vegetation (SWS) seems to have a stronger agreement than that derived from the scatter plots (Figure 63).

Medium and robust woody vegetation overlay was assessed using three different approaches. First, only the plots[36] with high specific PFTs density were overlaid to the abundance maps (Figures 63 and 64). Then, all individuals belonging to that specific PFT class were overlaid (Figures 65.a, 65.b, 66.a, 66.b, 67a, 67.b and 68.a). These plots and individuals were heterogeneous regarding plant species but homogeneous regarding plant size. And finally exclusively individuals from one species were overlaid to the fractional maps (Figures 65.c, 65.d, 66.c, 67.c, 67.d and 68.b).

Plots overlaid over fractional maps from classes MWS and RWS show in general an unsatisfactory agreement in spatial trend but local agreement is found (Figure 63). Plots from RWSC overlaid over RWSC fractional map (Figure 64) shows good agreement. Plots from RWT (Figure 63) show local high agreements in a few small patches in the western side of the southern area (Figure 26 A). Here, the endmember was chosen. The high values existing in the eastern side of B are striking. This area is rich in sand and dune vegetation such as *Chenopodium rubra*. As discussed in section 5.2.2, a high RMSE was found here, which means that the model couldn't really predict the coverage in this region.

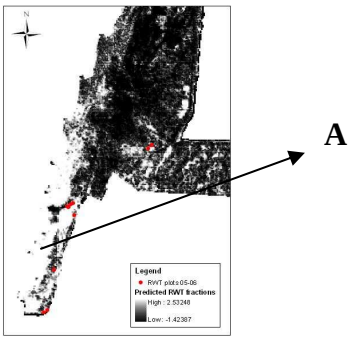


Figure 26: RWT plots overlaid to RWT abundance map.

Regarding individual plants, individuals from class MWS overlaid to the abundance map (Figures 65.a and 65.b) show a very low agreement while if we overlay *Sambucus nigra* individuals the agreement increases (Figures 65.c and 65.d). The same occurs with class RWS, the agreement occurs now more locally since the distribution of

[36] These PFT plots are areas of 5x5 m with more than 50% coverage of that specific PFT.

Salix fragilis is not so broad (Figure 66). When we confront all individuals from class RWS to the abundance map of RWSC the coincidence may be considered almost random (Figures 67.a and 67.b), but if what we overlay is only *Crataegus monogyna* individuals we observe clear trends of agreement (Figure 67.c), excepting the southern region. Here the grouping density appears to be looser (Figure 67.d). It is not observed that the individuals in this region are younger than in the center of the study area. Average height and crown projection are 3.36 m and 3.1 in the southern region, while they are 2.62 m and 2.3 m^2 in the center. The good trend agreement observed in the center of the study area (Figure 67.c) may be explained by the fact that the endmember chosen for RWSC was covered 100% by *Crataegus monogyna*. This is contradictory, although, when we go back to the endmember selection and take up the discussion on the spectra from *Crataegus monogyna* (Section 3.3.4.1), in which we deduced a high within variability. But it supports the conclusion that species-based PFTs are better mapped by using spectral techniques (Section 5.2.1). Regarding RWT class, bad predictions are clear for both RWT and *Salix alba* (Figure 68) as it was discussed in section 5.2.2.

5.2.4 Sub-pixel accuracy assessment (OSA)

The overall sub-pixel accuracy will be discussed from two different view points. First it will be discussed looking at the average values and secondly, looking at its spatial distribution. In general, the results based on OSA show low values with averages lower than 50% (Table 10). The standard deviation is normally high, showing a big range of values and hence the low reliability of the results. The first fact that can be observed by looking at the values is that the overall accuracy for woody vegetation (42 %) is higher than that for herbaceous vegetation (38 %), as derived from previous validation approaches. For herbaceous PFTs, accuracy is higher using validation sets from the years 2002 (39 %) and lower for the years 2004 (35 %) and 2005 (36 %). This is contradictory with results derived from scatter plots (Figures 50, 51 and 52) in which it seems that there is a slight improvement in 2004 and 2005. The small differences observed from both validation methods should not be considered highly significant; especially taking into account that few plots were used for 2004-05 (5 and 8) while more than 100 were used for 2002. Small woody vegetation appears to be better predicted also using this parameter showing one of the largest OSA values; over 50 %.

Another fact that may be observed is that prediction improves for woody vegetation when it is based on species rather than when it is based on PFTs themselves (Table 10). Always OSA values increase when fractional maps are validated with plots pure in species, decrease when these species are mixed with others and decrease even more if these species are not present. For instance, class MWS average OSA is 40 %, if we select only those plots composed of *Sambucus nigra* it increases to 54 %, if those plots contain *Sambucus nigra* mixed with other species the average is 43 %, on the contrary, if the plots are composed of other species the average is 0 %. The difference is smaller though when considering the overall average (40 % all plots and 41 % purest plots), but this is explainable by the fact that there are more herbaceous plots –where no difference between species was made- and hence these have more weight than the woody plots in the final overall accuracy value.

If we take a look now at the spatial distribution of OSA (Appendix 6), we find out as discussed in section 5.2.3 that in all cases the overall accuracy increases in those regions where the endmember was selected.

5.2.5 Description of an optimal methodology to map PFTs in river floodplains

This section is the answer to the second research question: "Which is the best methodology for mapping PFTs using linear spectral unmixing?"

The best methodology to map PFT in a floodplain based on previous results (assuming that the endmembers will be selected from the image) is:

1. To have a hyperspectral image and a set of field data available, both from the same year and period of the year.
2. Field data should be gathered according to the objective of the SMA application. This means that, ideally, plots should be based in a grid with the same pixel size and origin as the image. This could be done by establishing a reference location in the field as origin, and a fixed orientation. Qualitative data should be gathered in the form of percentages of species coverage per plot in a unique stratum. Grazed areas should be specified. The grouping density of such species (number of stems per surface unit) should be considered because of its relevance for hydraulic roughness.
3. The definition of plant functional types should be done based on the conceptual model shown in section 3.3.2.1. Relevant traits to determine vegetation roughness are: height, density and flexibility. The application of this model to

the definition of the PFT classes should be done based on a study of the most relevant species in the area. This study should include: mean height, foliage density and inherent species flexibility of stems. Based on these data, species can be arranged by hydraulic resistance and conceived as PFTs.

4. Spectral mixture analysis should be applied following the steps described in section 3.3.4. Grazed areas and shadow should be considered as separate endmembers. This method should be combined with the use of LiDAR or other IR based images to improve grouping structure separability.

5.3 Temporal analysis

The temporal analysis was designed and performed after the results of SMA on HyMap were obtained. Therefore, some of the conclusions derived from the previous discussion (Sections 5.1 and 5.2) were adopted in the spectral mixture analysis applied in this temporal study. Therefore, two different things will be discussed here: the comparison between HyMap and CASI images classifications and the comparison between the first HyMap image classification (using ten endmembers) and the second HyMap image classification (using six endmembers).

5.3.1 Spectral Mixture Analysis

As it was inferred from previous results low accuracy was not a consequence of the number of endmembers but of the division of vegetation into PFTs (Section 5.1). The introduction, in this second approach, of the species-based (Section 5.1) definition of PFTs raised the accuracy of the HyMap classification. Average OSA value for the HyMap image was 40 % in the previous approach (Table 10) against 66 % in this second approach (Table 12). In the case of herbaceous vegetation it increased from 37 % to 47 %, and in the case of woody vegetation from 42 % to 68 %. This raise was although the number of PFTs was reduced from ten to six.

In this second approach we observe again that herbaceous classes (47 %) show a less accurate distribution than woody classes (68 %). In the case of high coverage of specific PFT this value increases, e.g. pixels covered in more than 50 % by shrubs have 81 % of accuracy. This may be due to a combination of two facts:

1. Field data for herbaceous classes belong to 2002 (two years previous to when the image was taken, 2004) while they have a fast evolvement.
2. Herbaceous vegetation classification into PFTs “grasses” and “forbs” implies high heterogeneity of species and spectral signatures.

OSA results (Table 12) from CASI image (2001), support the first fact because in this case, that herbs field data was gathered closer to the image date (2002), herbs show more accuracy than woody vegetation, which was gathered further (2005-06). OSA value for herbaceous vegetation is 41 % while is 21 % for woody vegetation. OSA value for herbaceous classes is still higher for HyMap (47 %) than for CASI (41 %), which shows the low overall accuracy of the CASI classification. However, while the OSA value is 41 % for shrubs and 41 % for pioneer trees, it drops to 4 % for dwarf shrubs, which might be causing such low average OSA value for woody vegetation. The reason for this might be that in 2001, when the image was taken, the ROI used as endmember for dwarf shrubs (Table 6) did not have the crown projection they showed when the data field was gathered (2005-06).

Another factor that might be causing an improvement in the results from HyMap image is the endmember used for pioneer trees. In previous approach, ROIs for endmember definition of trees were not selected from the wide forested area in the south of the lake (Figure 25 A). This was because of uncertainties regarding the field data gathered in this area in 2002. In this case, because this area was not correctly mapped in previous approach (Figure 22 D and Figure 25 A), data from 2002 was used together with data from 2005-06 to make up the ROI for pioneer trees. OSA value for HyMap improved from 35 % (Table 10) to 64 % (Table 12).

Besides, ROIs representing the endmembers were not unique pixels. In this approach they were conformed by a combination of pixels. These pixels were selected preferably from different regions in the area to reduce exclusion for grouping effects.

Regarding the distribution of OSA values (Appendix 11) we may observe two things. Comparatively, OSA values are higher for HyMap classification than for CASI, as inferred from the average values (Table 12). As a general trend, in the HyMap image, values are higher in those regions were plants are more densely grouped (Figures 89.a and 90.a).

From the abundance maps observed in Appendix 8 we may observe that clear trends and complementary features are shown in the distribution of PFTs in the case of HyMap (Figure 79). CASI abundance maps show, in general, less clear trends and redundancy of high fractions in same areas for different classes, e.g. grasses and shrubs.

HyMap RMSE image shows similar values and features using ten or six endmembers. Therefore, discussion made in section 5.2.1 is applicable here, except for region J which did not show high RMSE. Regarding the CASI image a very high

RMSE values are found. These are specified in section 4.3.2. These high errors are an indication that the endmembers in the CASI unmixing were not correctly addressed. High errors appear in regions with presence of grass. Regions with presence of *Crataegus monogyna* (dwarf shrubs) show high RMSE value as well. High errors are encountered too in regions with other tree species (Section 4.3.2) because they were not considered as endmembers. In general, we may observe that high errors in CASI do not show clear trends as they show in HyMap. Instead, they are more evenly distributed in a sparkling configuration. This indicates low precision in the unmixing model.

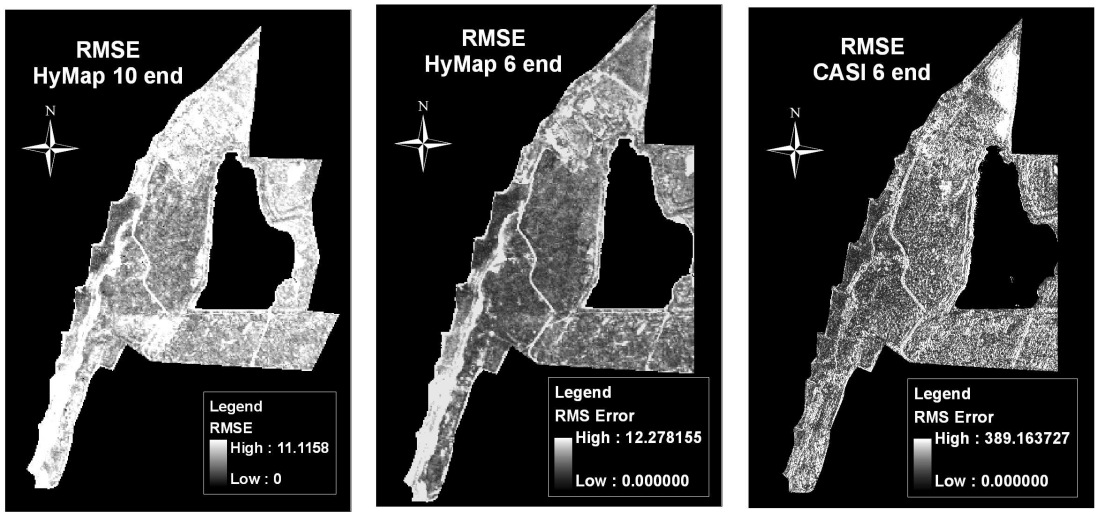


Figure (a) **Figure (b)** **Figure (c)**

Figure 27: RMSE image from the unconstrained unmixing of HyMap image using 10 endmembers (a), using 6 endmembers (b) and CASI image using 6 endmembers (c)

From the scatter plots shown in Appendix 9 (Figure 82) we may derive that classification of all PFTs for HyMap show a positive trend close to the one to one line (100 to 100 in this case). However, the dispersion is still quite high, with low R^2 values. Plots with higher observed coverage show less dispersion than the rest. Figure 83 shows the scatter plots resulting from the CASI classification. We can see that trends are worse for all PFT classes.

In Appendix 10 (Figures 84 and 85) we observe a good trend of agreement between the abundance maps from HyMap and the plots of each PFT. Figures 86 and 87, from CASI, show a lower agreement, except for the class Pioneer trees, which is good mapped in some regions but in others is over estimated. One possible explanation for this is that its ROI spectrum was confused with the one from grass (Figure 72).

Apart from the fact that woody vegetation covered less area in 2001 than in 2005-06, other important factors may be causing also low accurate results in the case of the

CASI image as the number of spectral bands. HyMap image, with 126 bands, shows in this study more adequate characteristics to classify intimate mixtures of vegetation PFTs than the multispectral image CASI, with 10 bands. Another cause of low is that this image was initially in RD coordinates and it was transformed to UTM, as explained in section 3.3.3.1. During this projection transformation, some spatial inaccuracies were observed. A translation of 10 m east was performed to correct this shift as a practical solution. But, likely, the geo-referencing of the CASI image support inaccuracies which might have affected the unmixing procedure. One of the possible reasons behind this shift might the specific geometrical correction this image was object to.

5.3.2 Monitoring PFTs in river floodplains

This section is the answer to the last research question: "Is it feasible to monitor shrubs encroachment by comparing two hyperspectral images from two different sensors, while each of these images is analyzed by the previous methodology?"

Changes in vegetation distribution between the years 2001 and 2004 can be derived from the comparison of the hard classified images presented in Figure 28. Figure 29 shows these changes quantitative. According to Figure 29, grasses and dwarf shrubs increased their extension from 2001 to 2004, and forbs, shrubs, trees and sand decreased.

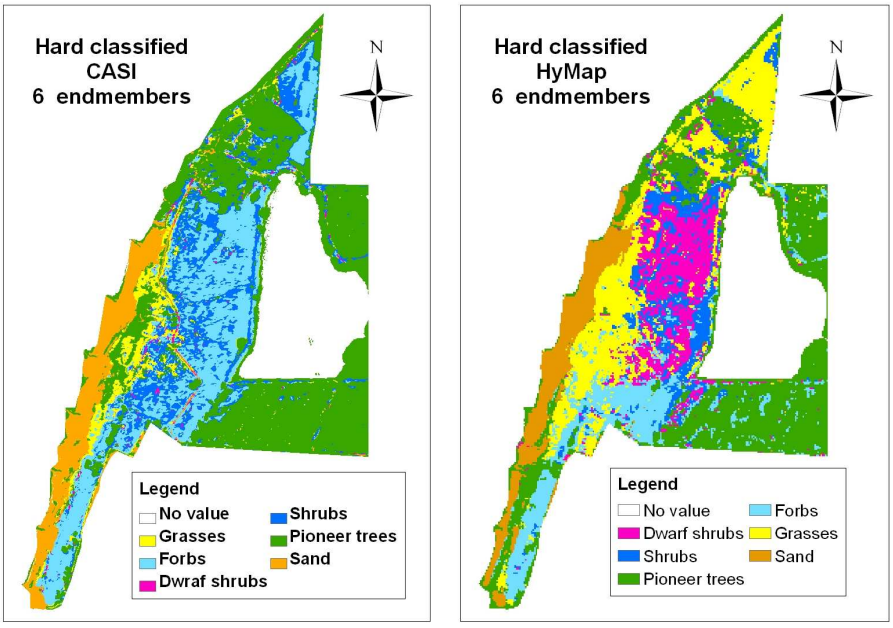


Figure (a) **Figure (b)**

Figure 28: Hard classified maps from fully constrained unmixing, CASI 2001 (a) and HyMap 2004 (b)

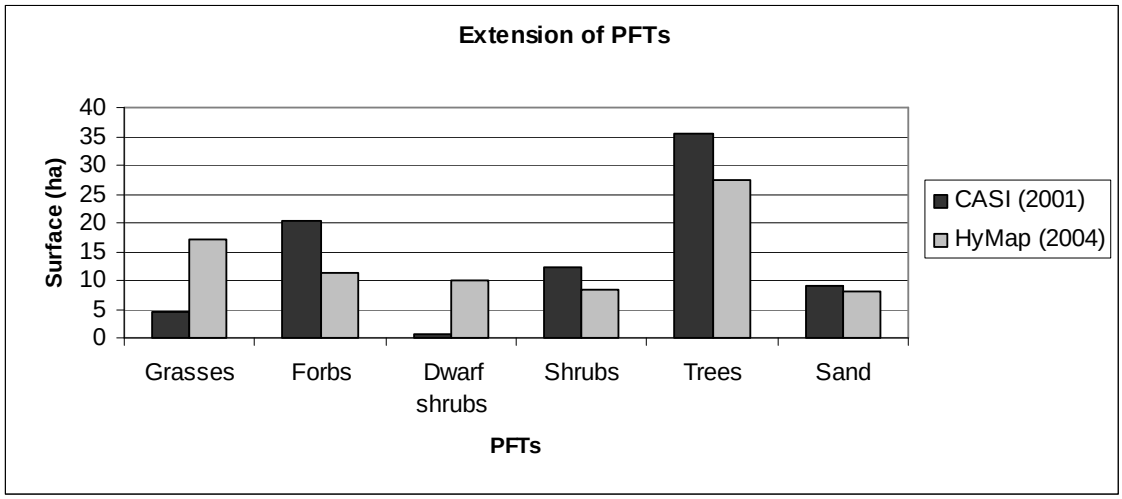


Figure 29: Changes in the extension of PFTs according to the classification of the images CASI (2001) and HyMap (2004).

However, the unreliability of the CASI image classification, as discussed in section 5.3.1, disables these affirmations. After all, monitoring PFTs using spectral techniques is feasible as long as mapping PFTs is feasible. Therefore, the same considerations regarding mapping PFT explained in section 5.2.5 should be applied here.

6 CONCLUSIONS

This thesis was aimed to develop an imaging spectroscopy-based method for mapping and monitoring plant functional types in river floodplains, with special emphasis on monitoring the process of shrub encroachment.

In despite of the high heterogeneity of vegetation in a river floodplain it is possible to map PFTs by using Spectral Mixture Analysis. A high number of PFT classes is not required to reach a high classification accuracy. However, of primary importance is the division of vegetation into PFTs. This division should satisfy two conditions: a) the spectral separability of classes and b) the characterization into classes that show different hydraulic roughness. The woodiness of the plant must lead the division at the highest level. At a second level, most abundant woody species must serve as basis for the division. Shrubs and trees show relatively high classification accuracy with respect to herbs. Herbaceous vegetation show improved results when divided into two groups: graminoids and forbs. Graminoids are more easily classified than forbs. Classification performs better for those pixels covered mainly by unique species. Woody species, due to their different mean height, foliage density and inherent stem flexibility can be used by river managers to map differences on hydraulic roughness of river floodplains.

An optimal methodology should involve the use of a hyperspectral image (HyMap) rather than a multispectral image (CASI). Field data is necessary to support the endmember selection and validation assessment. A temporal and spatial coordination between image and field data is an important requirement. Qualitative data in the field should be gathered, for this application, in the form of percentages of species coverage per plot in a unique stratum. Grazed areas should be specified. Grouping density must be considered because of its importance for hydraulic roughness. LiDAR data may be used to support spectral techniques in the analysis of grouping structures and shadow in future research. The preprocessing of the images must include an atmospheric correction, an accurate geo-referencing and a masking of the area of interest. SMA is applied first by a selection of endmembers (preferably from the image, but it could be also done from a reliable endmembers' library). It is recommended not to use a unique pixel per endmember but instead by merging pixels from different regions in the study area, as accuracy improves in the region where the endmembers are selected. Then, it follows a MNF transformation, which is necessary to reduce the data and exclude noise. After the linear unmixing is applied, a validation analysis should be

performed. Different validation approaches will underline different aspects of the classification accuracy.

In this way, changes of PFTs distribution in time can be derived from a time series of hyperspectral images. But, previously, in order to reduce some uncertainties, previous methodology should be tested using field data and hyperspectral images from same acquisition year and period of the year.

Imaging spectroscopy techniques show in this research to be a promising tool to map vegetation in river floodplains. The developed methodology could improve the efficiency of the tasks of the river managers with respect to monitoring nature rehabilitation and flood risk management. However, some effort has to be done to bring to an agreement the conditions for the application of spectral techniques and the needs of the ecological models with respect to the characterization of PFTs.

7 RECOMMENDATIONS

In this chapter several recommendations for future research are listed:

- PFTs abundance maps obtained by spectral mixture analysis may be used to initiate or validate the dynamic vegetation model SMART2-SUMO2 to address the potential of remote sensing techniques when simulating vegetation models.
- Spectral mixture analysis could be applied to a specific band selection and test whether there is an improvement or not. The response of plant canopies to radiation is a function of the intensity of radiation in various spectral regions, and the absorbing and transmitting properties of the vegetation. Leaves reflect little in the blue and red wavelengths due to absorption by photosynthetic pigments, and reflect strongly in the near-infrared (NIR) due to intra-and inter-leaf scattering. At the canopy level, multiple scattering between layers of vegetation increases NIR reflectance and is related to canopy geometry and leaf optical properties (Law et al., 1994; Gao et al., 2003). The reflectance in the NIR region incorporates also a nonlinear (interaction) term related to multiple scattering (Gilabert et al., 2002).
- Horizontal density of plants, which has an important role in hydraulic roughness (Anderson, 2006), could be mapped using LAI index derived by using spectral techniques.
- Combination of spectral methods, such as SMA, with the use of LiDAR or other IR based images could be tested. This would avoid the effect of shadows and improve the separability between different grouping structures (Geerling et al., 2006).
- The application of advanced unmixing techniques such as MESMA (Multiple Endmember Spectral Mixture Analysis) can give improved results. MESMA is recommended to face the high heterogeneity of vegetation in the Millingerwaard. The spatial and spectral variability, which is a challenge in the Millingerwaard, it is potentially addressed by multiple endmember spectral mixture analysis (MESMA) (Powell et al., 2005). This method allows the number and type of endmembers to vary on a per pixel basis (Powell et al., 2005). MESMA is especially recommended when there are different spectra for the same specie (Rosso et al., 2005). It allows for the simultaneous testing of

more than one endmember per class which appears to produce higher accuracy (Rosso et al., 2005).

- Evidential reasoning should be also tested to map PFTs. This method relates indexes as LAI, EVI, albedo, etc. with PFTs. It was applied by Sun et al. (2006) at a global scale using MODIS images. Past research has shown that evidential reasoning can produce better results compared to traditional classifiers (Sun et al., 2006).
- VCNNC (Vegetation Community based Neural Network Classifier) could also be used in future research. It provides sub-pixel fractional abundance of vegetation species with high accuracy (Wang et al. 2006), which constitutes useful information in these areas.
- Further work using medium-resolution sensors like MODIS and MERIS is recommended to upscale the approach from the floodplain level to the river catchment scale.

8 BIBLIOGRAPHY

Adams, B.; White, A.; Lenton T.M., (2004). An analysis of some diverse approaches to modeling terrestrial net primary productivity. Ecological Modeling 177: 353-391.

Anderson, B.G., (2006). Quantifying the interaction between Riparian Vegetation and Flooding: from Cross-section to Catchment scale. PhD. University of Melbourne.

Anderson, B.G.; Rutherfurd, I.D.; Western, A.W., (2006). An analysis of the influence of riparian vegetation on the propagation of flood waves. Environmental Modelling & Software 21: 1290-1296.

Austrheim, G.; Eriksson, O., (2003). Recruitment and life-history traits of sparse plant species in subalpine grassland. Canadian Journal of Botany 81: 171-182.

Baptist, M.J.; Penning, E.; Duel, H.; Smits, A.J.M.; Geerling, G.W.; Van Der Lee, G.E.M; Van Alphen, J.S.L., (2004). Assessment of the effects of cyclic floodplain rejuvenation on flood levels and biodiversity along the Rhine river. River research and applications 20: 285-297.

Baptist, M.J.; van den Bosch, L.V.; Dijkstra, J.T.; Kapinga, S., (2005). Modelling the effects of vegetation on flow and morphology in rivers. Large Rivers. Vol. 15, No.1-4, Arch. Hydrobiol. Suppl. 155/1-4 : 339-359.

Bonan, G.B.; Levis S.; Sitch S.; Vertenstein, M.; Olenson K.W., (2003). A dynamic global vegetation model for use with climate models; concepts and description of simulated vegetation dynamics. Global Change Biology 9: 1543-1566.

Borisova D.; Nikolov, H.; Danov M., (2005). Spectral mixture analysis for data verification and validation. Solar-Terrestrial Influences Laboratory-BAS Bulgaria.

Bresheards, D.D., (2006). The grassland-forest continuum: trends in ecosystem properties for woody plant mosaics? Frontiers in Ecology Environment 4(2): 96–104.

Teerhuis, C., (2006). Species specific measurements of vegetation leaf reflectance for the development of a field protocol and spectral library. Internship report Department of Environmental Science, Geo-information Science and Remote Sensing. Wageningen University

Chapin, F.S., (1993). Functional role of growth forms in ecosystem and global processes. Ehleringer and C. B. field, editors. Scaling physiological processes. Academic Press, San Diego, California, USA. Pp. 287-312 in J.R.

Chapin III, F. S., Bret-Harte, S. M., Hobbie, S. E. and Zhong, H. 1996. Plant functional types as predictors of transient responses of arctic vegetation to global change. Journal of Vegetation Science. 7: 347–358.

Díaz, S.; Hodgson, J.G.; Thompson, K.; Cabido, M.; Cornelissen, J.H.C.; Jalili, A.; Montserrat-Martí, G.; Grime, J.P.; Zarrinkamar, F.; Asri, Y.; Band, S.R.; Basconcelo, S.; Castro-Díez, P.; Funes, G.; Hamzehee, B.; Khoshnevi, M.; Pérez-Harguindeguy, N.; Pérez-Rontomé, M.C.; Shirvany, F.A.; Vendramini, F.; Yazdani, S.; Abbas-Azimi, R.; Bogaard, A.; Boustani, S.; Charles, M.; Dehghan, M.; de Torres-Espuny, L.; Falczuk,V.; Guerrero-Campo, J.; Hynd, A.; Jones, G.; Kowsary, E.; Kazemi-Saeed, F.; Maestro-Martínez, M.; Romo-Díez, A.; Shaw, S.; Siavash, B.; Villar-Salvador, P. & Zak, M.R., (2004). The plant traits that drive ecosystems: Evidence from three continents. Journal of Vegetation Science 15: 295-304, 2004.

Epstein H.E.; Chapin III s.; Walker M. D.; starfield A.M., (2001). Analyzing the functional type concept in artic plants using a dynamic vegetation model. OIKOS 95: 239-252.

Florsheim, J.L.; Mount, J.F., (2001). Restoration of floodplain topography by sand-splay complex formation in response to intentional levee breaches, Lower Cosumnes River, California. Geomorphology 44: 67–94.

Friedl, M.A.; McIver, D.K.; Hodges, J.C.F.; Zhang, X.Y.; Muchoney, D.; Strahler, A.H.; Woodcock C.E.; Gopal, S.; Schneider, A.; Cooper, A.; Baccini, A.; Gao, F.; Schaaf, C., (2002). Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment 83, 287–302

Gao F.; schaaf C.B.; Strahler A.H.; Jin Y.; Li x., (2003). Detecting vegetation structure using kernel-based BRDF model. Remote Sensing of environment 86 198-205.

Geerling, G.W.; Ragas A.M.J.; Leuven, R.S.E.W; van den Berg, J.H.; Breedveld, M.; Liefhebber, D.; Smits, A., (2005). Succession and rejuvenation in floodplains along the River Allier (France). Hydrobiology 565:71–86.

Geerling, G.W.; Katier, E.; Smits, T., (2005b). Cyclic rejuvenation. Lecture notes. Centre for Water and Society & Environmental Science.

Geerling, G.W.; Labrador-García, M.; Clevers, J.G.P.W.; Ragas, A.M.J.; Smits, A.J.M., (2006) Classification of floodplain vegetation by data-fusion of Spectral (CASI) and LiDAR data. International Journal of Remote Sensing. Unpublished.

Gilabert, M.A.; Gonzáez-Piqueras, J.; García-Haro, F.J.; Meliá J., (2002). A generalized soil-adjusted vegetation index. Remote Sensing of Environment 82 03-310.

Gondard, H.; Jauffret, S.; Aronson, J.; Lavorel, S., (2003). Plant functional types: a promising tool for management and restoration of degraded lands. Applied Vegetation Science 6: 223-234.

Green, A.A.; Berman, M.; Switzer, P.; Craig., M.D., (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Trans. Geoscience and Remote Sensing, 26, pp.65-74.

Hartmann, R.A.; Knotters, A.G., (2006). Monitoringsmogelijkheden vegetatiestructuurtypen. Rijkswaterstaat. Ministerie van Verkeer en Waterstraat.

Hodgson, J. G.; Wilson, P. J.; Hunt, R., 1999. Allocating C-S-R plant functional types: a soft approach to a hard problem. – Oikos 85: 282–294.

Hooijer, A.K.F; Kwadijk, J; Pedroli, B., (2002). Towards Sustainable Flood Risk Management in the Rhine and Meuse River Basins. Main results of the IRMA SPONGE research project. NCR, 18(1568-234X).

Ichoku, C.; Karnieli, A., (1996). A review of mixture modeling techniques for sub-pixel land cover estimation. Remote Sensing Reviews, 13, pp. 161–186.

Jauffret, S.; Lavorel, S., (2003). Are plant functional types relevant to describe degradation in arid, southern Tunisian steppes?. Journal of Vegetation Science 14: 399-408.

Kleyer, M., (2002). Validation of plant functional types across two contrasting landscapes. Journal of Vegetation Science 13: 167-178.

Kooistra, L.; Sterkx, S.; Liras E.; Mengesha, T.; Verbeiren, B.; Batelaan, O.; van Dobben, H.; Schaepman, M.; Schaepman&Strub, G.; Stuiver, J., (2005). HyEco'04: an airborne imaging spectroscopy campaign in the floodplain Millingerwaard, the Netherlands. CGI report, ISSN 1568-1874.

Kooistra, L.; Suarez Barranco, M. D.; van Dobben, H.; Schaepman, M., (2006a). Regional scale monitoring of vegetation biomass in river floodplains using imaging spectroscopy and ecological modelling. Centre for Geo-information and Landscape Centre, Wageningen University and Research Centre, The Netherlands.

Kooistra, L.; Wamelink, W.; Schaepman-Strub, G.; Schaepman, M.; van Dobben, H.; Aduaka, U.; Batelaan, O., (2006b). Assessing and predicting biodiversity in a floodplain ecosystem: assimilation of imaging spectroscopy products into a

dynamic vegetation model. Submitted to Remote Sensing of Environment. Unpublished.

Kooistra, L.; Sanchez-Prieto, L.; Bartholomeus H.M.; Schaepman M.E., (2007). Regional mapping of plant functional types in river floodplain ecosystem using airborne imaging spectroscopy data. Center for Geo-Information. Wageningen University. Unpublished.

Kumar, L., Schmidt, K., Dury, S., Skidmore, A., 2001. Imaging spectrometry and vegetation science. In: Meer, F.D.v.D., Jong, S.M.d. (Eds.), Imaging Spectrometry Basic Principles and Prospective Applications. Kuwer Academic Publishers, Dordrecht/Boston/London, p. 115.

Laurent, J.M.; Bar-Hen, A.; François, L.; Ghislain, M.; Cheddadi, R., (2004). Refining vegetation simulation models: From plant functional types to bioclimatic affinity groups of plants. Journal of Vegetation Science 15(6): 739-746.

Law, B.E.; Waring R.H. (1994). Remote sensing of Leaf Area Index and radiation intercepted by understory vegetation. Ecological applications. 4(2), pp.272-279.

Li, L.; Ustin, S.L.; Lay, M., (2005). Application of multiple endmember spectral mixture analysis (MESMA) to AVIRIS imagery for coastal salt marsh mapping: a case study in china Camp, CA, USA. International Journal of Remote Sensing. Vol. 26, No. 23, 5193-5207.

Lillesand, T.M.; Kiefer, R.W., (1994). Remote Sensing and Image Interpretation. Third edition. Unites States and Canada. John Wiley & Sons, Inc.

Lindstro, A.M.; Jaenson, T.G.T., (2003). Distribution of the Common Tick, *Ixodes ricinus* (Acari: Ixodidae), in Different Vegetation Types in Southern Sweden. Journal of Medical Entomology v. 40, no. 4 p. 375-378.

Liras-Laita, E. (2005). Imaging Spectroscopy for Ecological Monitoring at the Test Site the Millingerwaard: Species Mapping using Spectral Libraries and Soil-Vegetation-Atmosphere-Transfer Models. Thesis Report GIRS-2005-04. Laboratory of Geo-Information Science and Remote Sensing. Wageningen University and Research Centre. The Netherlands

Malenovsky, Z.; Bartholomeus, H. M.; Acerbi-Junior, F.W.; Schopfer J.T.; Painter, T.H.; Epema, G.F.; Bregt, A.K., (2006). Scaling dimensions in spectroscopy of soil and vegetation. International Journal of Applied Earth Observation and Geoinformation. In press.

Mishev D., (1991). Spectral characteristics of mixed classes of natural formations, Acta Astronautica, vol.25, N8/9, pp.443-446.

Naden, P.; Rameshwaran, P.; Mountford, O.; Robertson, C., (2006). The influence of macrophyte growth, typical of eutrophic conditions, on river flow velocities and turbulence production. Hydrological processes 20, 3915-3938.

Nangendo, G.; Stein, A.; Steege, H. ter; Bongers, F., (2005). Changes in woody plant composition of three vegetation types exposed to a similar fire regime for over 46 years. Forest ecology and management v. 217, issue 2-3 p. 351-364.

Nielsen, A.A., (2001). Spectral Mixture Analysis: Linear and Semi-parametric Full and Iterated Partial Unmixing in Multi- and Hyperspectral Image Data. International Journal of Computer Vision 42 (1/2), 17-37.

Pan, Y.; Mc-guire, A.D.; Melillo, J.M.; Kicklighter, D.W.; Sitch, S.; Prentice, I.C., (2002). A biogeochemistry-based dynamic vegetation model and its application along a moisture gradient in the continental United States. Journal of Vegetation Science 13: 369-382.

Paruelo, J.M.; Lauenroth, W.K., (1996). Relative abundance of plant functional types in grasslands and shrublands of North America. Ecological Applications 6(4): 1212-1224.

Pausas, J.G., (2003). The effect of landscape pattern on Mediterranean vegetation dynamics: A modelling approach using functional types. Journal of Vegetation Science 14: 365-374.

Peng, C., (2000). From static biogeographical model to dynamic global vegetation model; a global perspective on modelling vegetation dynamics. Ecological Modelling 135: 33-54.

Pengra B.W.; Johnston C.A.; Loveland T.R., (2006). Mapping an invasive plant, *Phragmites australis*, in coastal wetlands using the EO-1 Hyperion hyperspectral sensor. Remote sensing of environment. In press.

Powell, R.L.; Roberts, D.A.; Denninson, P.E.; Hess, L.L., (2006). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. Remote sensing of environment. In press

Rademacher, J.G.M. ;Wolfert, H.P., (1994). Het Rivier-Ecotopen-Stelsel: een indeling van ecologisch relevante ruimtelijke eenheden ten behoeve van ontwerp en beleids studies in het buitendijkse rivierengebied. Lelystad: Rijswaterstaat.

Rosso, P.H.; Ustin S.L.; Hastings A., (2005). Mapping marshland vegetation of San Francisco Bay, California, using hyperspectral data. International Journal of Remote Sensing Vol. 26, No. 23, 5169–5191.

Schmid, T.; Koch M.; Gumuzzio, J., (2005). Multisensor Approach to Determine Changes of Wetland Characteristics in Semiarid Environments (Central Spain). IEEE transactions on geoscience and remote sensing, vol.43, no. 11.

Schmidt, A.; Van Dobben H.; Wamelink W.; Kooistra L.; Schaepman M., (2005). HyEco'04: Using Hyperspectral reflectance data to initialize ecological models. Proceedings of 4th EARSeL Workshop on Imaging Spectroscopy. New Quality in environmental studies. EARSeL and Warsaw University.

Schmidtlein, S.; Sassin, J., (2004). Mapping of continuous floristic gradients in grasslands using hyperspectral imagery. Remote sensing of environment 92: 126-138.

Shockling, M.A.; Allen, J.J; Smits, A.J., (2006). Roughyness effects in turbulent pipe flow. J. Fluid Mech., vol. 564, pp.267-285.

Schouwenberg, E.P.A.G.; Houweling, H.; Jansen, M.J.W.; Kros J.; Mol-Dijkstra, J.P. (2000). Uncertainty propagatioin in model chains: a case study in nature conservation. Alterra-rapport 001.

Skarpe, C., (1996). PFTs and climate in southern African savanna. Journal of Vegetation Science 7: 397-404.

Song, M.H.; Zhou, C.P.; Ouyang, H., (2005). Simulated distribution of vegetation types in response to climate change on the Tibetan Plateau. Journal of vegetation science official organ of the International Association for Vegetation Science, v. 16, issue 3 p. 341-350.

Straatsma, M., (2006). Floodplain roughness mapping synergy: LIDAR and spectral remote sensing. Utrecht University, Faculty of Geosciences, Department of Physical Geography. The Netherlands. Unpublished.

Suarez-Barranco M.D., (2006). Estimation of the spatial frequency distribution of biomass in river floodplains. Thesis code number: GRS-80436. Laboratory of Geo-Information Science and Remote Sensing. Wageningen University and Research Centre. The Netherlands.

Sun, W.; Liang, S.; Xu, G.; Fang H.; Dickinson R., (2006). "Mapping Plant Functional Types from MODIS Data Using Multisource Evidential Reasoning". Unpublished.

Thomas, v.; Treitz, P.; Jelinski, D.; Miller, J.; Lafleur, P.; McCaughey, H., (2002). Image classification of a northern peatland complex using spectral and plan community data. Remote sensing of environment 84: 83-99.

Ustin, S.L.; Smith, M.O.; Jacquemoud, S.; Verstraete, M.; Govaerts, Y., (1999). Geobotany: vegetation mapping for Earth sciences. In: Remote Sensing for the Earth Sciences: Manual of Remote Sensing, 3d ed., v. 3, ed. A.N. Rencz, John Wiley and Sons, New York, USA, pp. 189–233.

Van den Bosch, L., (2003) Influence of vegetation on flow and morphology in the river Allier, France. Civil Engineering and Geosciences. University of Technology: 119. Delft, The Netherlands.

Van Geloof, I.: de Ronde, I. (2002). De vegetatie van de Millingerwaard na 10 jaar "natuuronwikkeling". Leerstoelgroep Natuurbeheer en Plantenecologie & Laboratorium voor Geo-informatiekunde en Remote Sensing. Wageningen university.

Verrelst, J., (2004). Adopting plant community data for floodplain vegetation mapping using an image fusion of CASI and LIDAR data. Thesis Report GIRS-2004-40. Laboratory of Geo-Information Science and Remote Sensing. Wageningen University and Research Centre. The Netherlands.

Wamelink, G.W.W.; ter Braak C.J.F.; van Dobben H.F., (2003). Changes in large-scale patterns of plant biodiversity predicted from environmental economic scenarios. Landscape Ecology 18: 513–527.

Wamelink, G.W.W.; de Jong, J.J.; Van Dobben, H.F.; Van Wijk, M.N., (2005). Additional costs of nature management caused by deposition. Ecological Economics 52: 437-451.

Wamelink, G.W.W.; van Dobben H.F.; Mol-Dijkstra J.P.; Schouwenberg, E.P.A.G.; Kros, H.; de Vried, W.; Beredse, F., (2006). Effect of nitrogen deposition reduction on biodiversity and carbon sequestration. Department of Environmental Sciences, Nature Conservation and Plant Ecology Group, Wageningen University. Submitted.

Wang, C.; Menenti, M.; Stoll, M.P.; Belluco, E.; Marani, M., (2006). Remote Sensing of Environment. In press.

Zhang, M.; Ustin, S.L.; Rejmankova, E.; Sanderson, E.W., (1997). Monitoring Pacific Coast Salt Marshes Using Remote Sensing. Ecological Applications, Vol. 7, No. 3, pp. 1039-1053.

Web References

Aerial photography military applications. War Department, Washington. Available from http://www.cartome.org/fm30-21/fm30-21.htm#IV (10/06/207)

Bureau of Meteorology, Australian Government. Available from http://www.bom.gov.au/sat/NDVI/NDVI2.shtml (29/01/2007)

Cheseapeake Bay & Mid-Atantic from Space. Towson University, Maryland, USA. Available from http://chesapeake.towson.edu/data/all_ndvi.asp (29/01/2007).

Department of Environment and Climate Change. New South Wales, Australia 2007. Available from http://www.environment.nsw.gov.au/soe/95/9_4.htm and http://threatenedspecies.environment.nsw.gov.au/tsprofile/veg1_list.aspx?name=Grasslands (10/06/2007)

Ferrier, R.C. (1998). SMART2. Available from http://www.macaulay.ac.uk/dynamo/smart2.htm (1/12/2006)

Michigan State University of Engineering, 2002. Available from http://www.egr.msu.edu/~northco2/BE481/Peakdischarge.htm (28/05/2007)

Orlovsky, L.; Orlovsky N., (2002). Dust in the wind: Remote imaging aids researchers studying the dynamics of disastrous dust storms. Spie's Oemagazine. Available from http://oemagazine.com/fromTheMagazine/sep02/pdf/dust.pdf (29/01/2007)

RDNAP. Coordinate calculator download. Available from http://www.rdnap.nl/download/download.html (20/01/2007)

The Barn University of California. Center for spatial technologies and remote sensing. Available from http://www.cstars.ucdavis.edu/classes/mexusenvi/tut13.htm (17/04/2007)

The Ecotope Mapping Working Group, (2005). What are ecotopes? Available from http://www.ecotope.org/about/ecotopes.htm (27/11/2006)

The USA's government official web portal of NASA, 2005. Available from http://science.hq.nasa.gov/earth-sun/technology/passive.html (10/06/207)

United States Department of Agriculture. Natural resources Conservation Service. Growth Habits Codes and Definitions (2007). Available from http://plants.usda.gov/java/profile?symbol=ARLA3# (2006/2007)

University of Texas at Austin. Center for space research. Available from http://www.csr.utexas.edu/projects/rs/hrs/process.html (13/02/2007)

Ustin, S.L. (1997). Geobotany: Vegetation Mapping for Earth Sciences. Available from http://www.cstars.ucdavis.edu/papers/html/ustinetal1998b/ (17/11/2006)

Wikipedia Foundation, Inc, 2001. Available from http://en.wikipedia.org/wiki/Main_Page (2006-07)

Web pages consulted to derive expected height of plant species:

http://www.ppws.vt.edu/~saskew/turfweeds/?do=view&batch=&id=39&referrer=http%3A%2F%2Fturfweeds.contentsrvr.net%2Fplant.php%3Fdo%3Dview%26batch%3D%26id%3D39 (30/01/2007)
http://www.fao.org/AG/AGP/AGPC/doc/Gbase/DATA/Pf000428.htm (30/01/2007)
http://recursos.cnice.mec.es/bancoimagenes2/buscador/imagen.php?idimagen=52288&expresion=docu%3A146 (30/01/2007)
http://www.dijon.inra.fr/hyppa/hyppa-a/crxhi_ah.htm (31/01/2007)
http://www.pfaf.org/database/plants.php?Carex+hirta (31/01/2007)
http://www.botanical-online.com/fotoscirsiumarvense.htm (31/01/2007)
http://ucce.ucdavis.edu/datastore/detailreport.cfm?usernumber=30&surveynumber=182 (31/01/2007)
http://www.unavarra.es/servicio/herbario/htm/Conv_arve.htm (31/01/2007)
http://herbarivirtual.uib.es/cas/especie/4544.html (31/01/2007)
http://www.asturnatura.com/Consultas/Ficha.php?Especie=Dactylis%20glomerata (31/01/2007)
http://tncweeds.ucdavis.edu/esadocs/elytrepe.html (31/01/2007)
http://www.hipernatural.com/es/pltgalio.html (31/01/2007)
http://www.ibiblio.org/pfaf/cgi-bin/arr_html?Glechoma+hederacea (31/01/2007)
http://www.infojardin.net/fichas/plantas-medicinales/herniaria-glabra.htm (31/01/2007)
http://waste.ideal.es/holcuslanatus.htm (31/01/2007)
http://www.dijon.inra.fr/hyppa/hyppa-f/odove_fh.htm (31/01/2007)
http://www.fruitveg.com/sp/hortalizas2.php3 (31/01/2007)
http://plantaardigheden.nl/plant/beschr/gonnve/zilverschoon.htm (01/02/2007)
http://www.hoseito.com/FLORES%20SILVESTRES/Potentilla%20anserina.htm (01/02/2007)
http://erick.dronnet.free.fr/belles_fleurs_de_france/potentilla_reptans1.htm (01/02/2007)
http://davesgarden.com/pf/go/1357/index.html (01/02/2007)
http://es.gardening.eu/plantas/Plantas-perennes/Achillea-ptarmica/1367/ (23/03/2007)
http://www.pflanzen-vielfalt.de/product.php?products_id=909999 (23/03/2007)
http://www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=242415415 (23/03/2007)
http://utc.usu.edu/factsheets/GlyceriaFSF/Glyceria_maxima.htm (23/03/2007)
http://www.npwrc.usgs.gov/resource/plants/vascplnt/species/jart.htm (23/03/2007)
http://www.florelaurentienne.com/flore/Groupes/Spermatophytes/Angiospermes/Dicotyles/103Composees/03_Leontodon/autumnalis.htm (23/03/2007)
http://www.wilde-planten.nl/geoorde%20zuring.htm (24/03/2007)
http://www.soortenbank.nl/soorten.php?soortengroep=flora_nl&id=2775&menuentry=soorten (24/03/2007)
http://www.massnrc.org/pests/pestFAQsheets/hogweed.html (24/03/2007)

9 APPENDIXES

APPENDIX 1: Information on images

Table 13: Characteristics of HyMap image and CASI image

	HyMap	***CASI***
Acquisition date	28th July 2004 (Strip number 1)	15th August 2001
IFOV (Instantaneous Field Of View)	2.5 mrad along track 2.0 mrad across track	
Field of view	61.3 degrees	
Pixels	1538 along track 512 across track	288 along track 512 across track
Swath	2300 m at 5 m GIFOV (along track) 4600 m at 10 GIFOV (along track)	
Spectral configuration		
VIS Spectrometer (1)		
Number of bands	30	6
Band number	1-30	1-6
Spectral range	450-890 nm	437-705 nm
Spectral resolution	8.1-16.2 nm	18 nm
NIR Spectrometer (2)		
Number of bands	32	4
Band numbers	31-62	7-10
Spectral range	890-1350 nm	729-890 nm
Spectral resolution	14.5-16.9 nm	18 nm
SWIR1 Spectrometer (3)		
Number of bands	32	
Band numbers	63-94	
Spectral range	1400-1800 nm	
Spectral resolution	13.1-15.6	
SWIR2 Spectrometer (4)		
Number of bands	32	
Band numbers	95-126	
Spectral range	1950-2480 nm	
Spectral resolution	18.3-21.3 nm	
Data scaling		
Final HyMap units (calibrated at-sensor radiance)	L [µW / cm2 sr nm]	
Data rescaling	L = 1000 DN (bands 1-62) L = 4000 DN (bands 63-126)	
Data formats		
HyEco-1_rad.bsq	Band Sequential (BSQ). Calibrated radiance X,y: pixels [ ] Z: [µW / cm2 sr nm]	
HyEco-1_rad_geo.img	Band Interleaved by Line (BIL)	
Sampling		
Line rate (lines per second)	16 Hz	
Resampling	Bilinear	
Pixel size	5 x 5 m	2 x 2 m
Map projection	UTM, Zone 31 N	Stereographic. RD Dutch Coordinate System
Geodetic Datum	WGS-84	Bessel
Parameters of strip1		
Flight altitude	2300 m (above sea level)	
Flight heading	0 deg	
Solar position	Air mass: 1.192 Zenith: (refracted): 33.050 Azimuth: 178.913 Cos incidence: 0.838 Cos zenith: 0.838	
Solar day	Solar time: 717.495	

	Julian day: 53214.984
Acquisition time calculation	UTC (Universal Time) = GMT (Greenwich Mean Time) GMT = MEST -2 (Middle European Summer Time) MEST = Local time
Acquisition time	11:38 hrs UTC 13:38 hrs MEST (or local time)
Start latitude / start longitude	51.8953 N / 5.9947 E
End latitude / end longitude	51.8525 N / 5.9936 E
Dimensions raw (x = across track, y = along track, z = spectral bands) [pixels]	512, 1538, 126 (198'438'912 bytes = 189 MB)
Dimensions geocoded (x = long., y = lat., z = spectral bands) [pixels]	581, 1416, 126 (207'319'392 bytes = 197 MB)

Source: Kooistra et al. (2005), Verrelst (2004) and Geerling et al. (2006).

APPENDIX 2: Endmembers selection

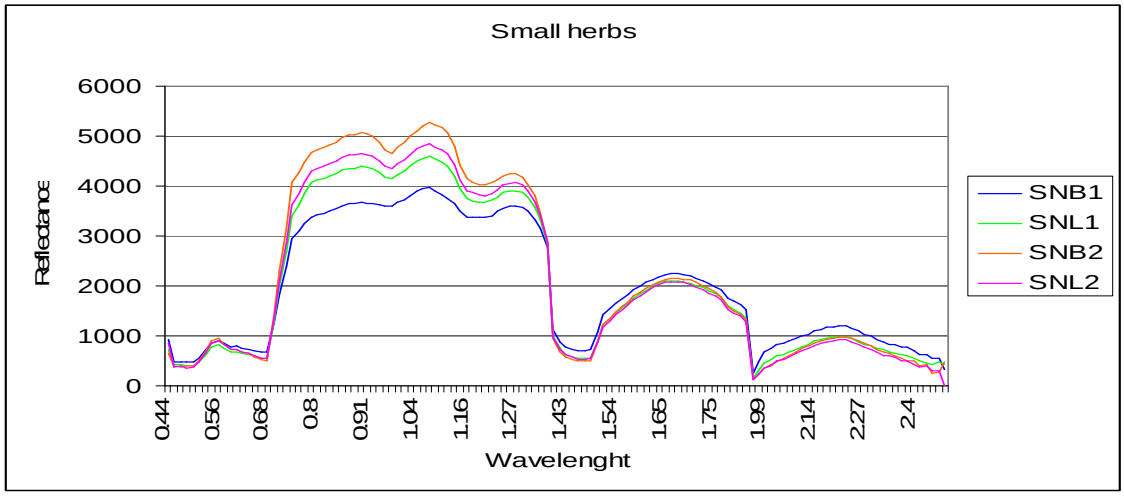


Figure 30: ROIs spectra small herbs

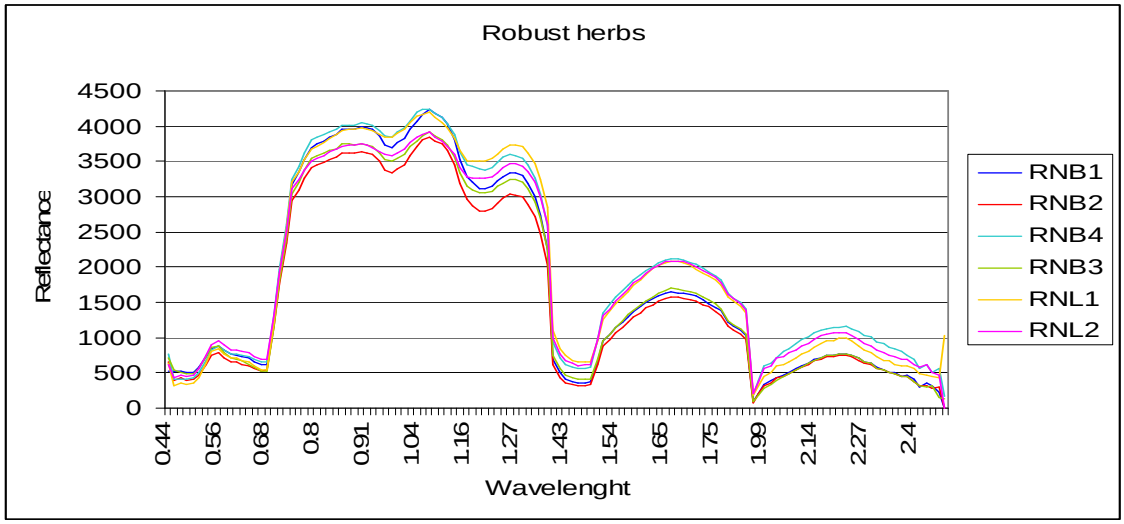


Figure 31: ROIs spectra robust herbs

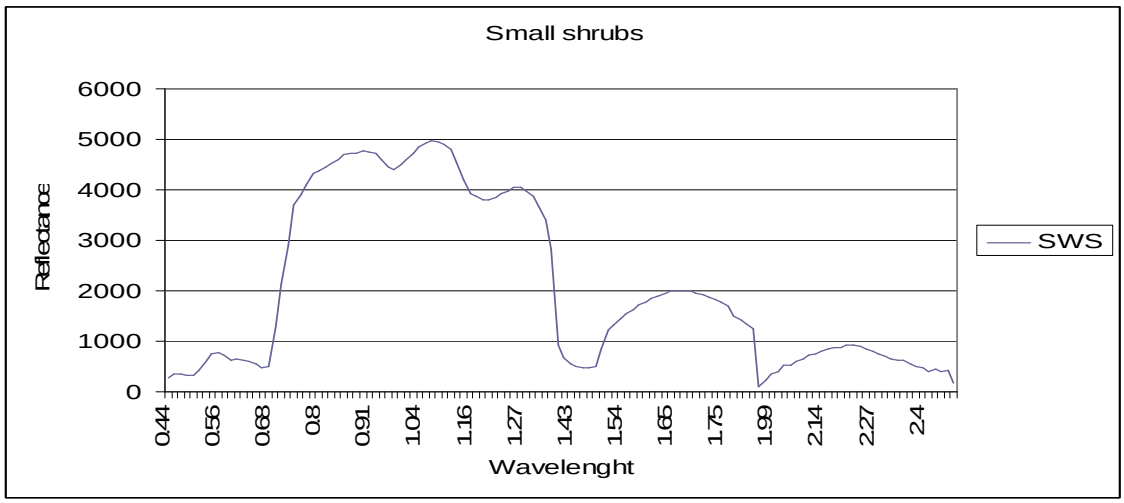


Figure 32: ROIs spectra small shrubs

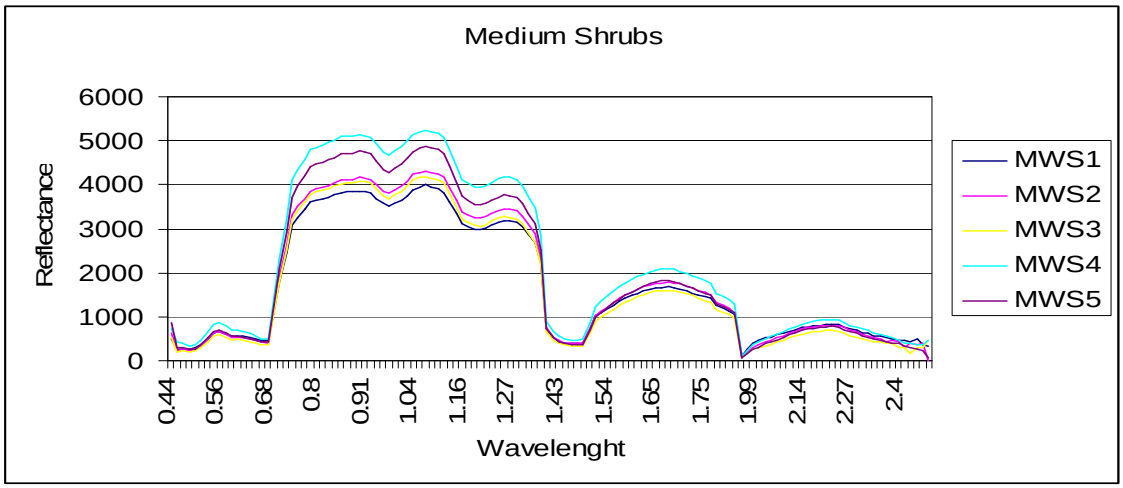


Figure 33: ROIs spectra medium shrubs

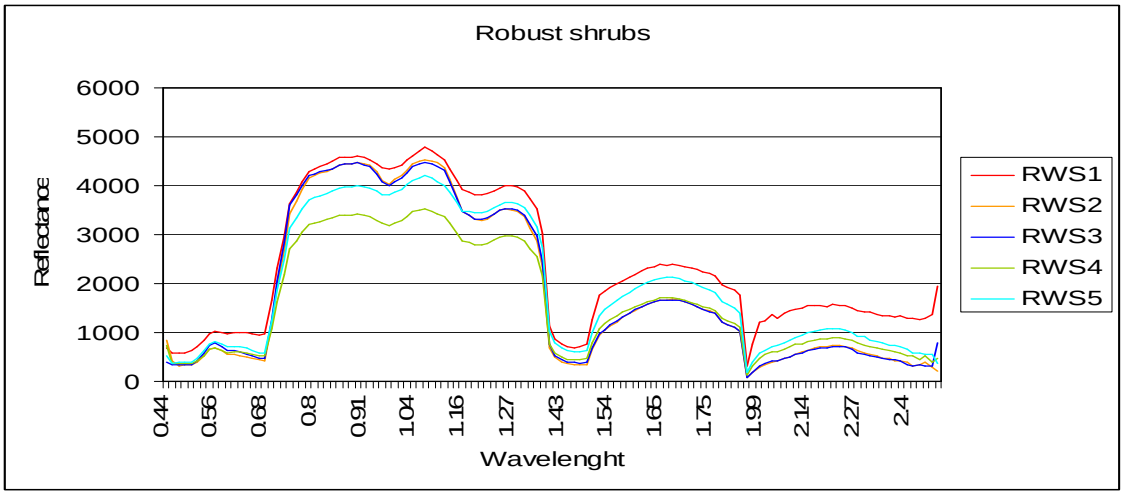


Figure 34: ROIs spectra robust shrubs

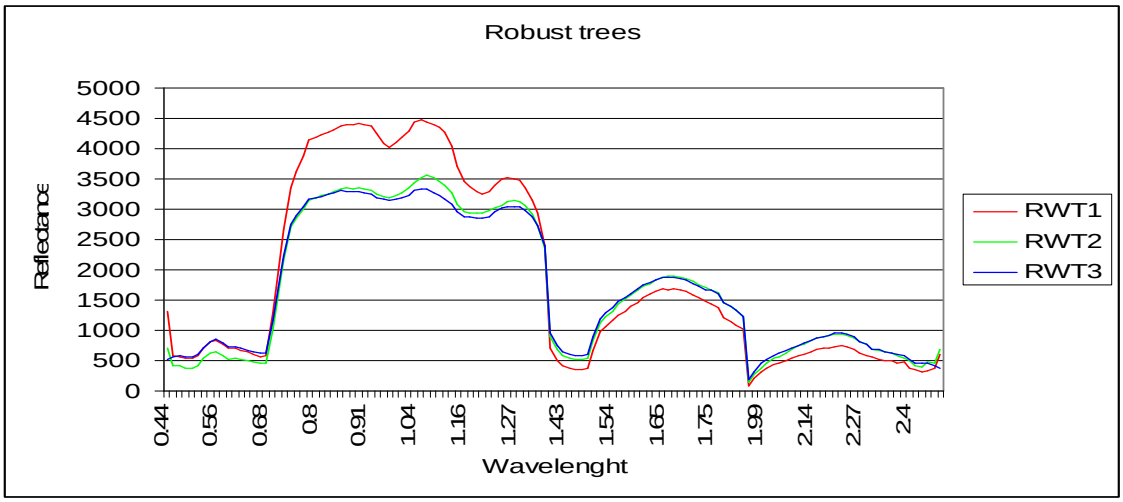


Figure 35: ROIs spectra robust trees

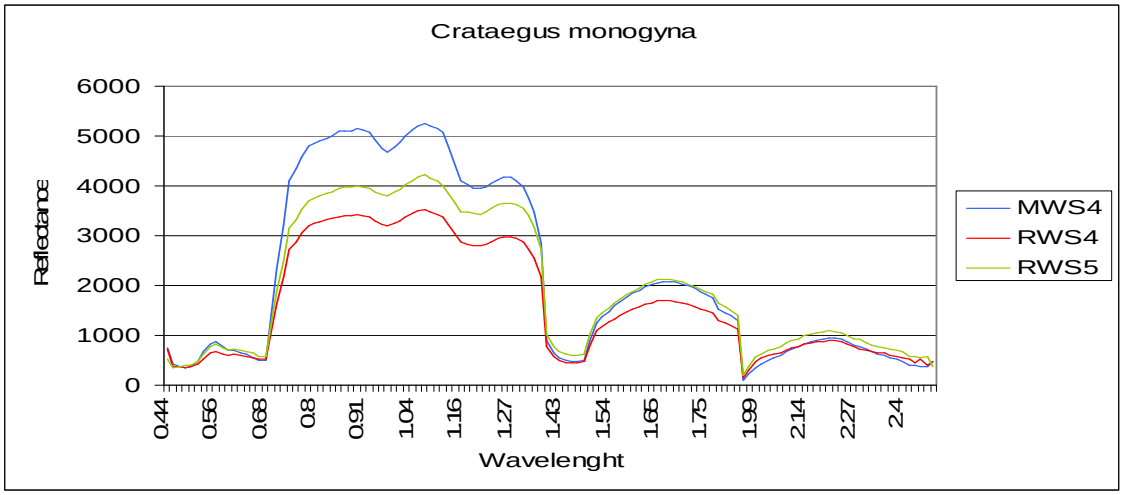


Figure 36: ROIs spectra RWSC

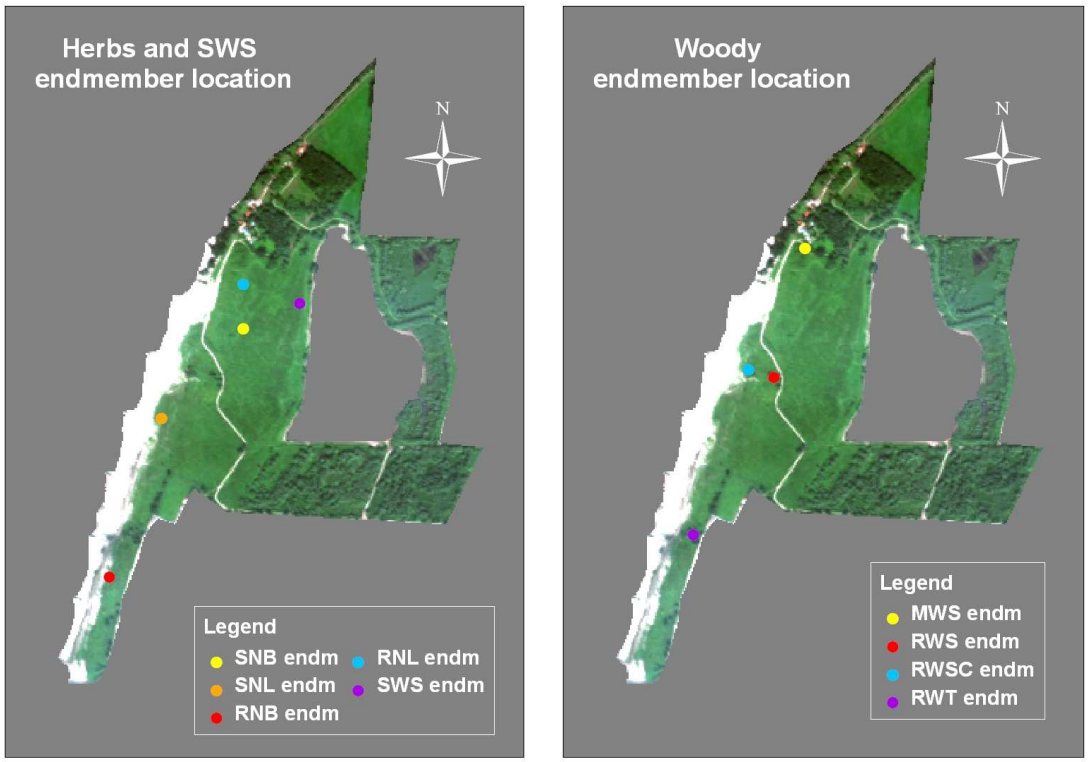


Figure 37: Location of final nine PFT endmembers

APPENDIX 3: Abundance maps and RMSE map.

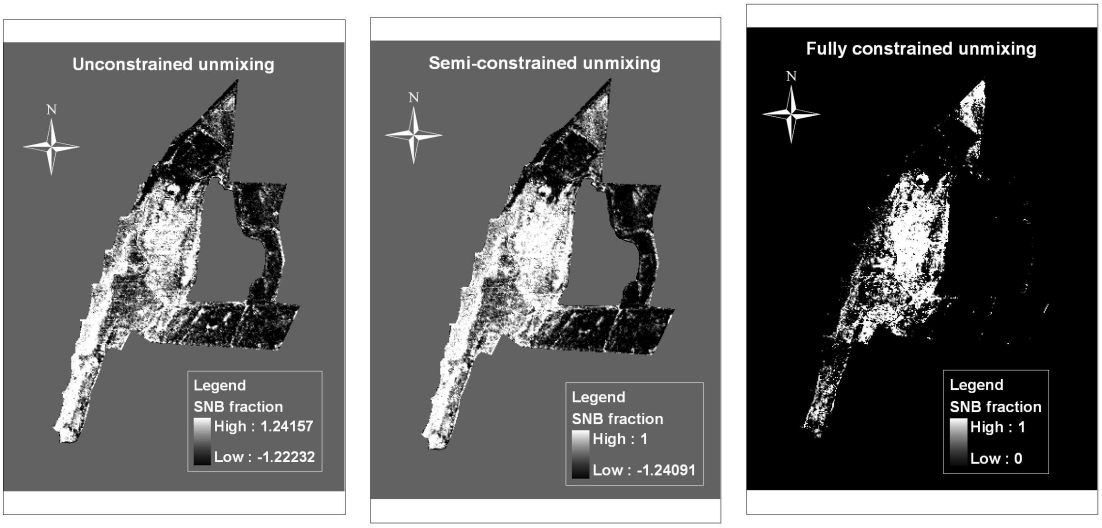


Figure 38: SNB abundance maps

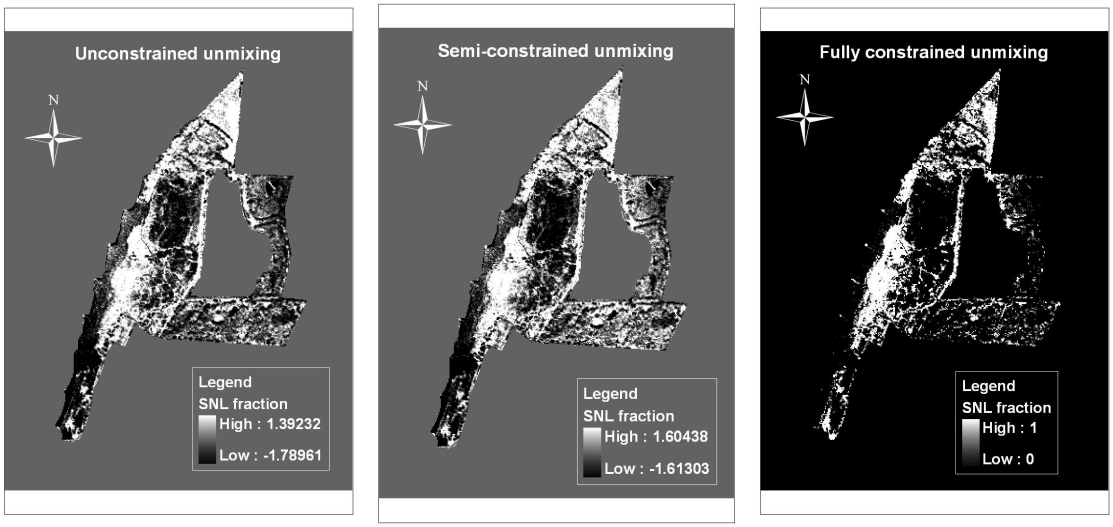


Figure 39: SNL abundance maps

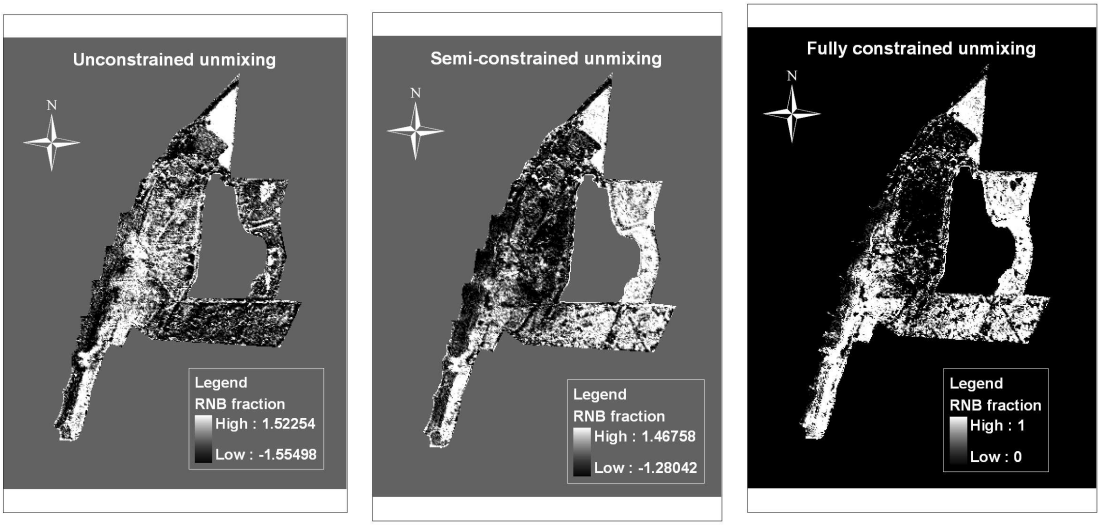


Figure 40: RNB abundance maps

Figure 41: RNL abundance maps

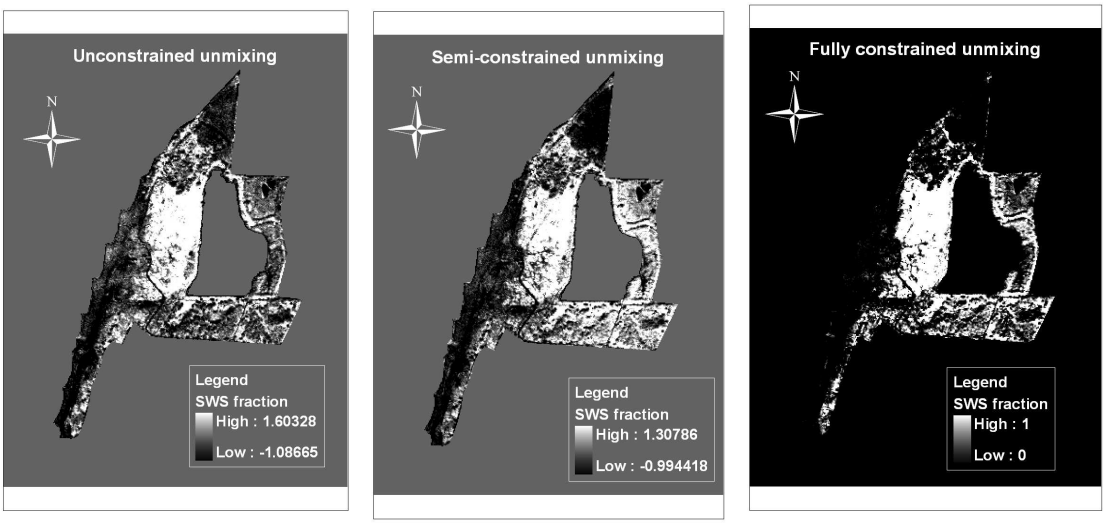


Figure 42: SWS abundance maps

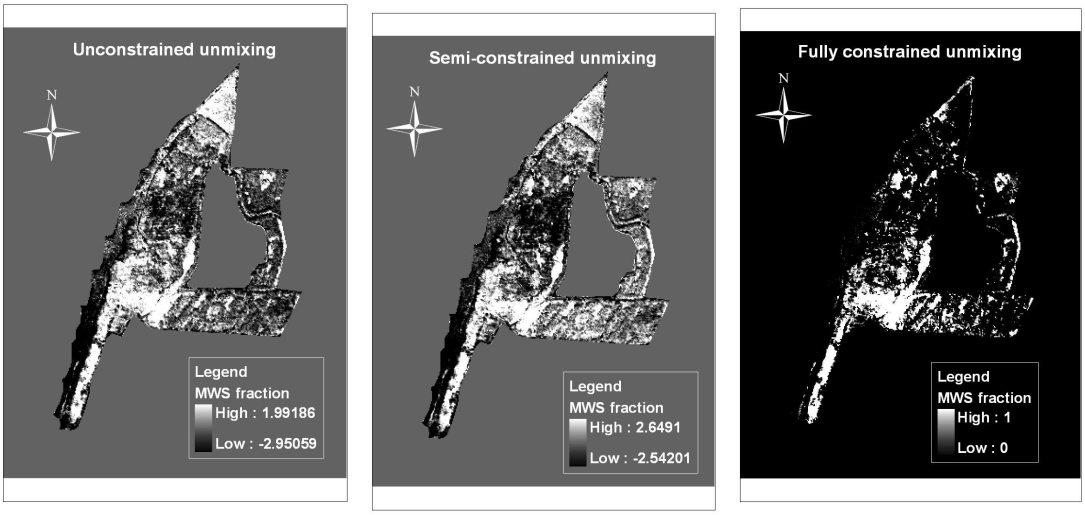


Figure 43: MWS abundance maps

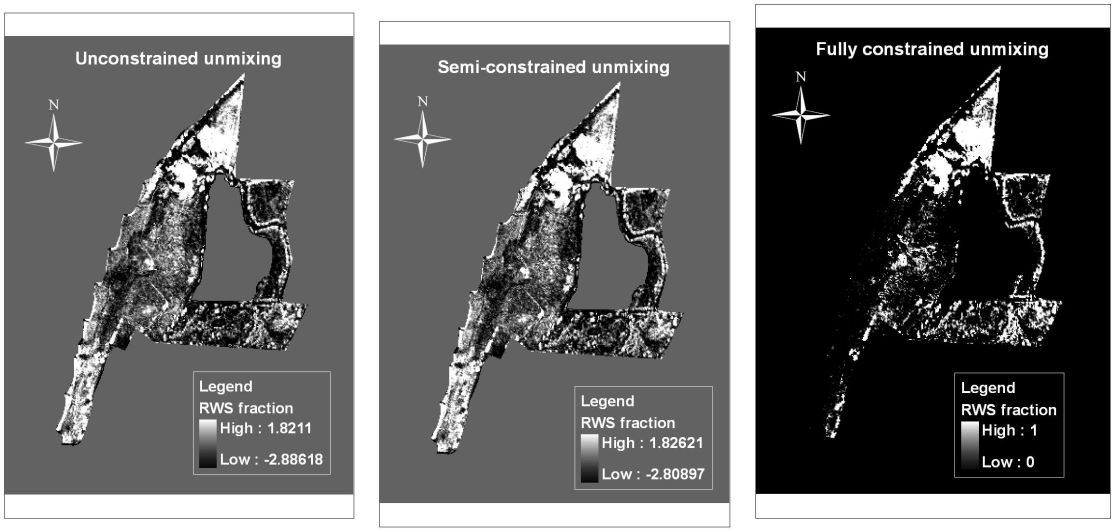


Figure 44: RWS abundance maps

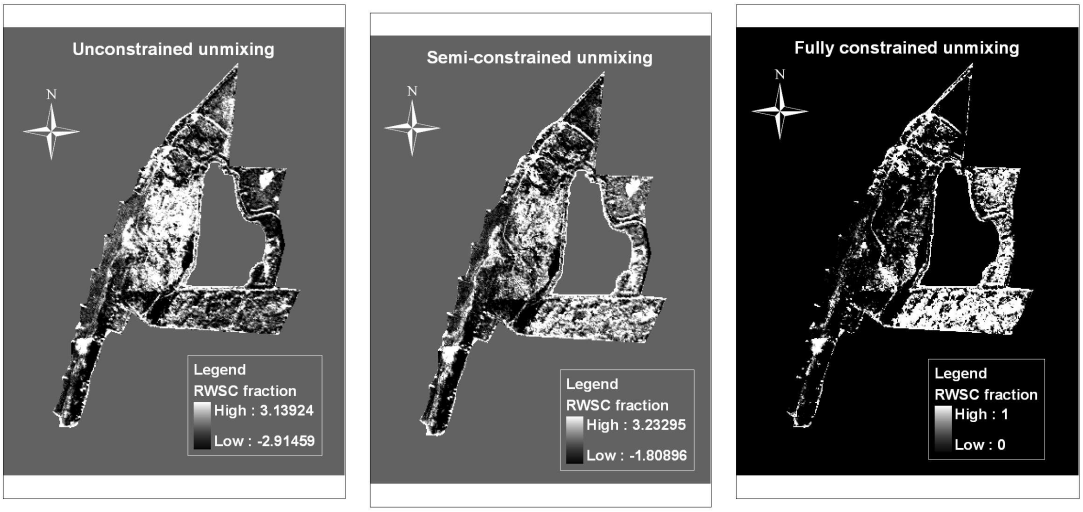


Figure 45: RWSC abundance maps

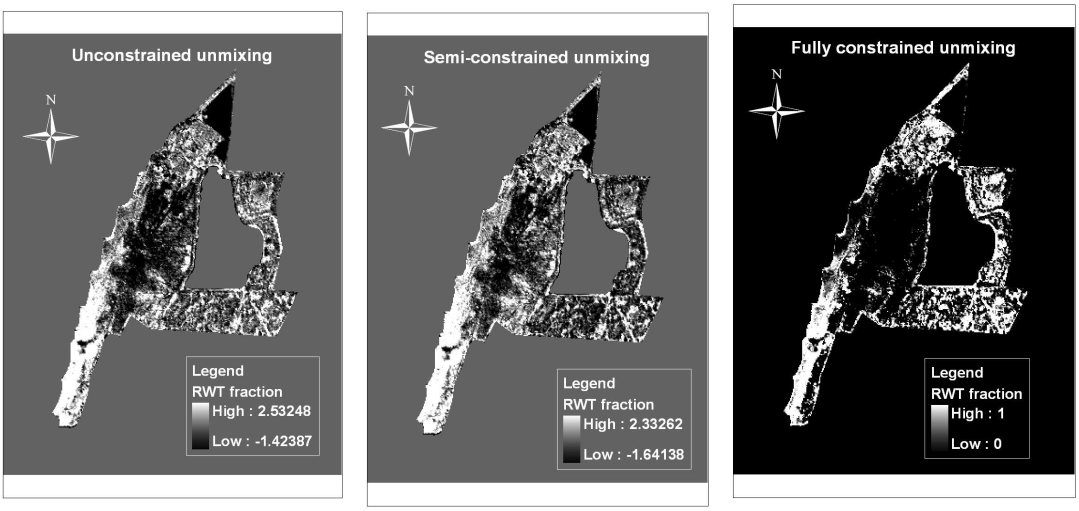


Figure 46: RWT abundance maps

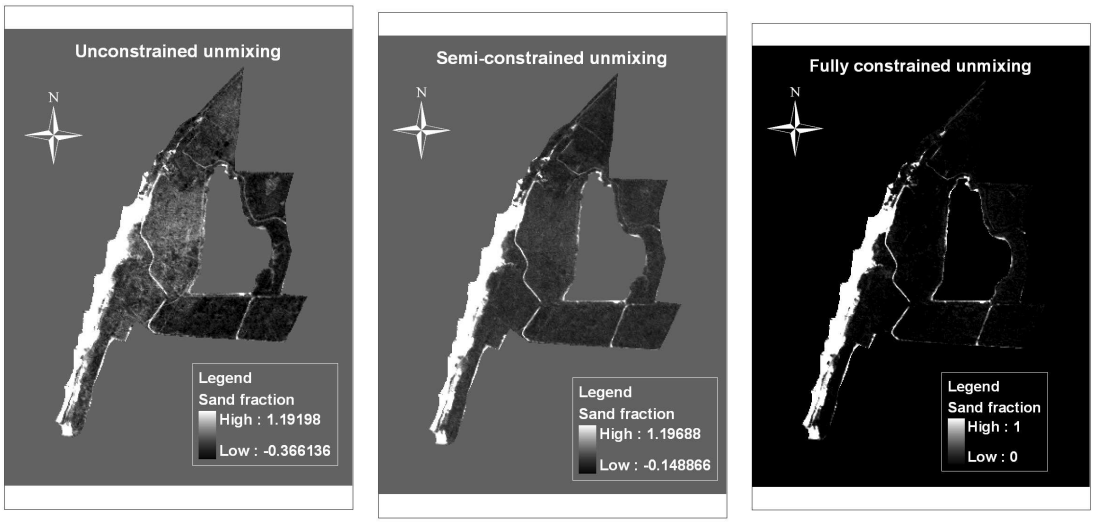


Figure 47: Sand abundance maps

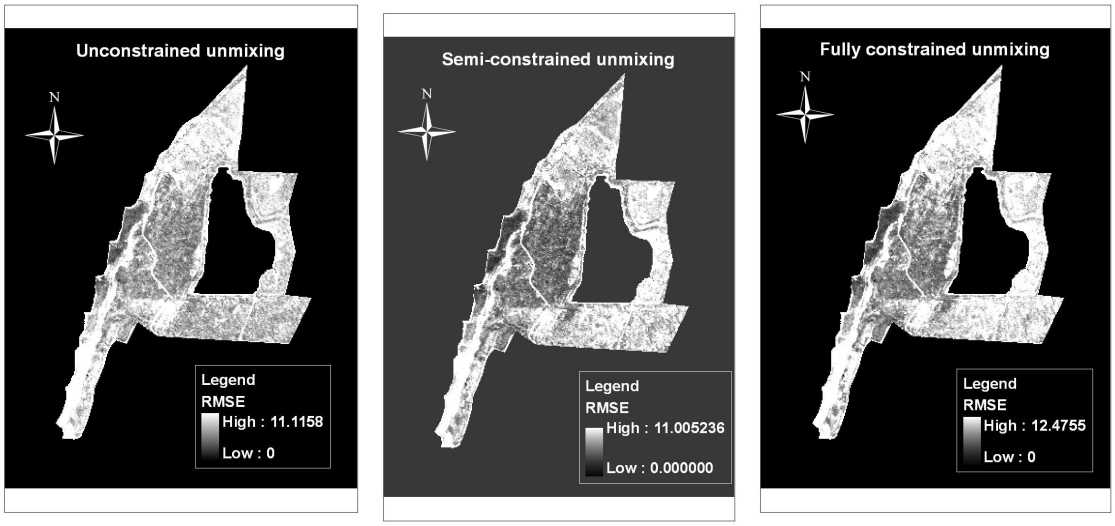


Figure 48: RMSE abundance maps

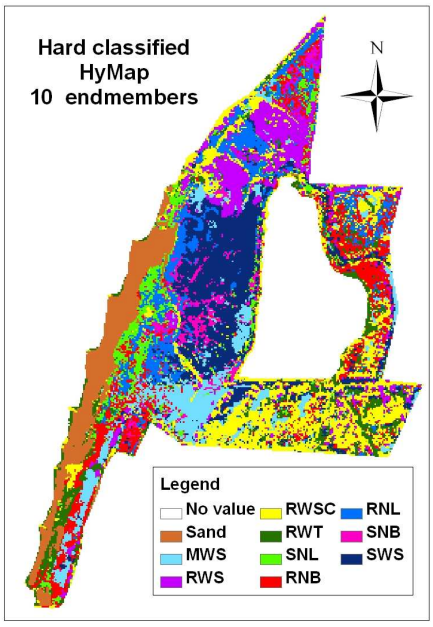


Figure 49: Hard classified image from fully constrained unmixing

APPENDIX 4: Comparison of observed vs. modeled by means of scatter plots

a) Unconstrained linear spectral unmixing

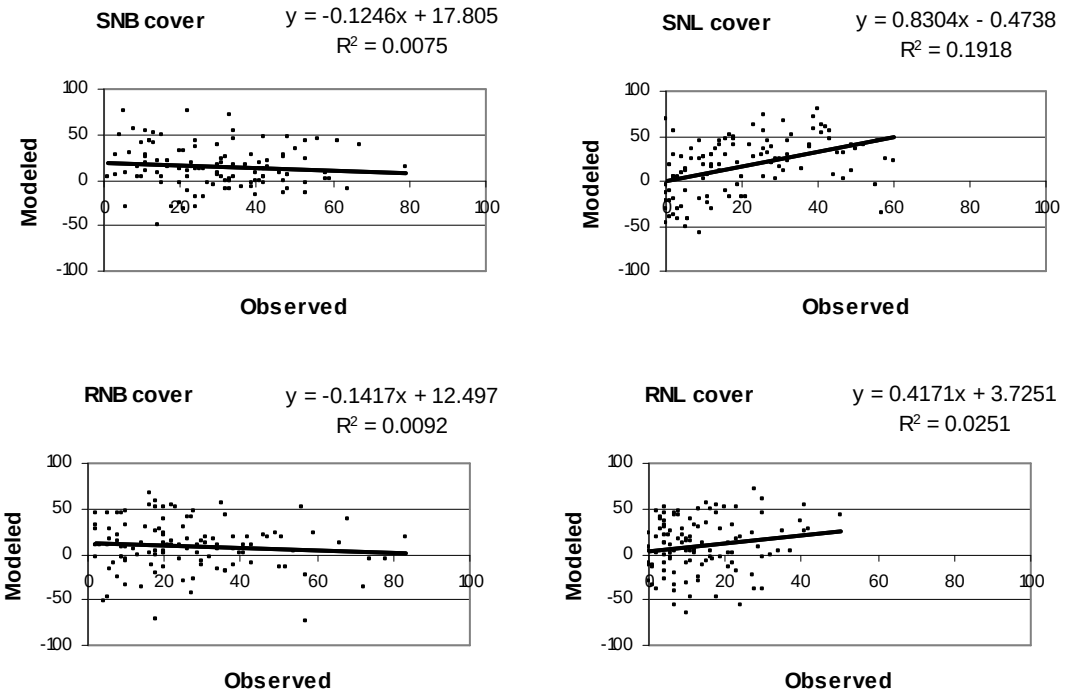
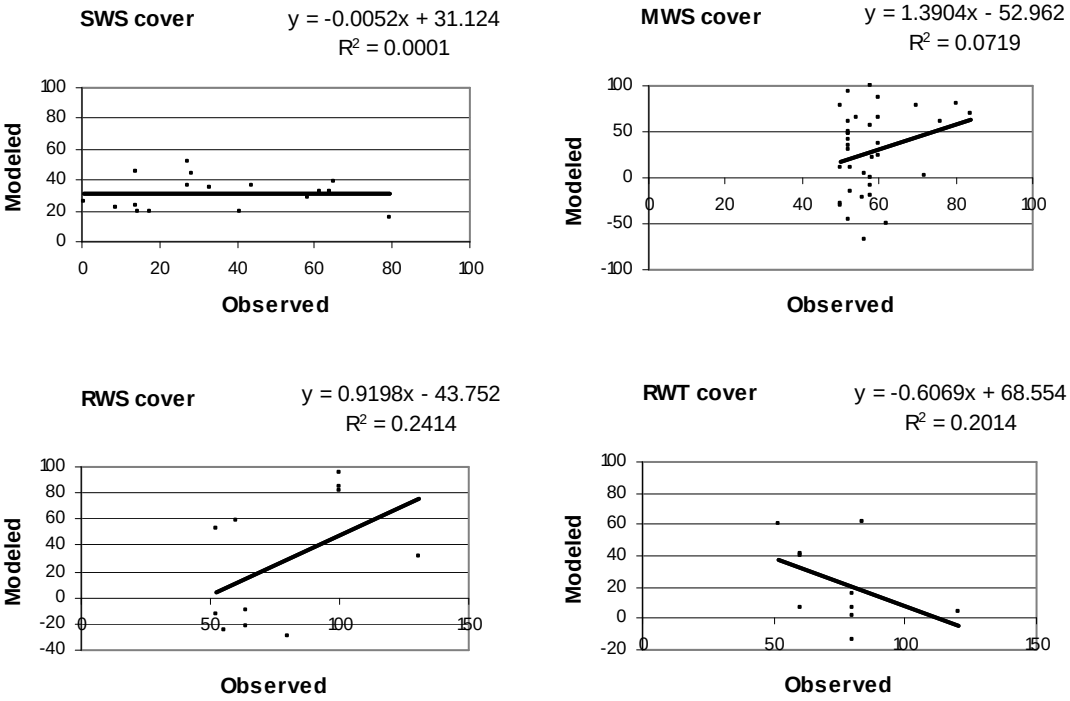


Figure 50: Herbaceous PFTs fractions predicted by unconstrained unmixing compared to fractions observed in the field in 2002



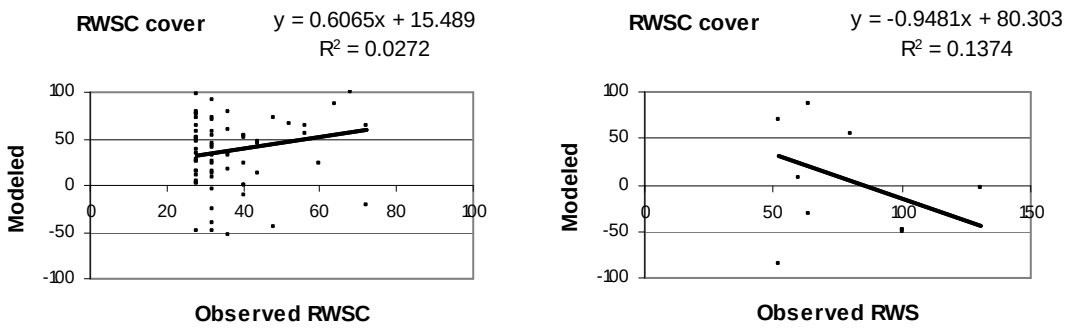


Figure 51: Woody PFTs fractions predicted by unconstrained unmixing compared to fractions observed in the field in 2002 (SWS) and 2005-06 (rest)

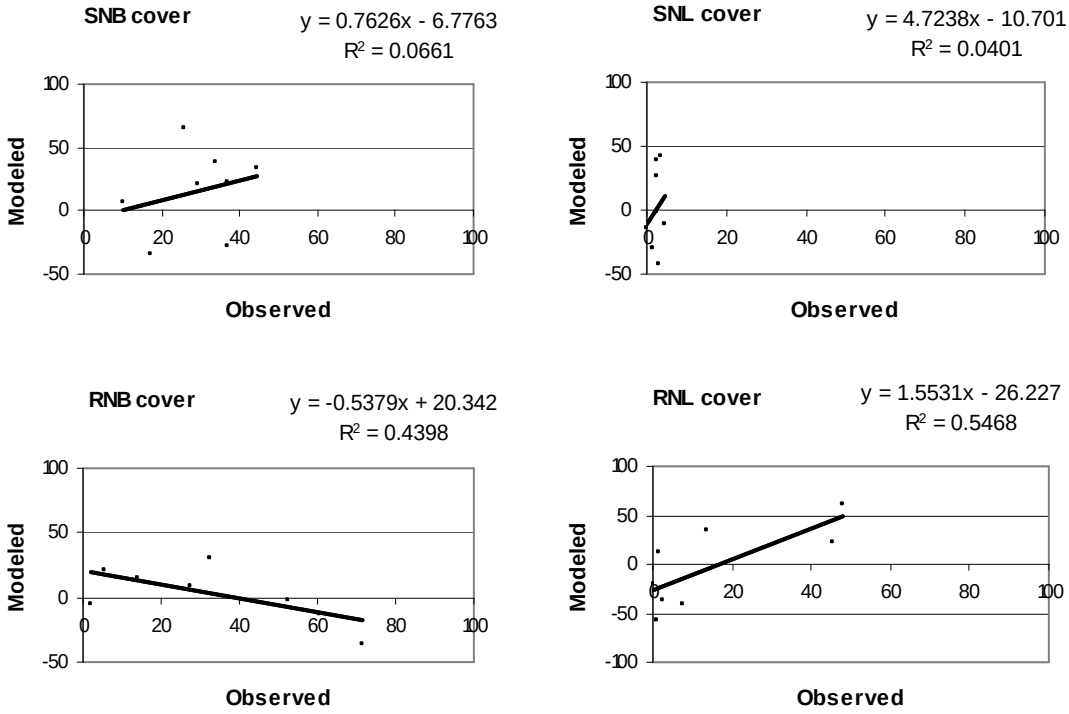
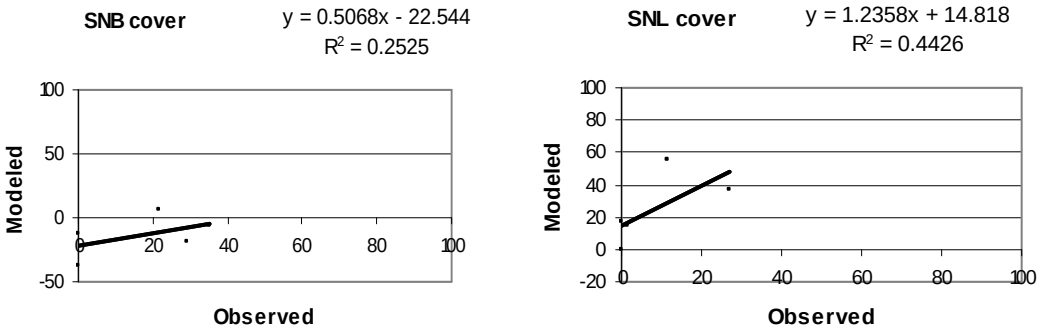


Figure 52: Herbaceous PFTs fractions predicted by unconstrained unmixing compared to fractions observed in the field in 2004



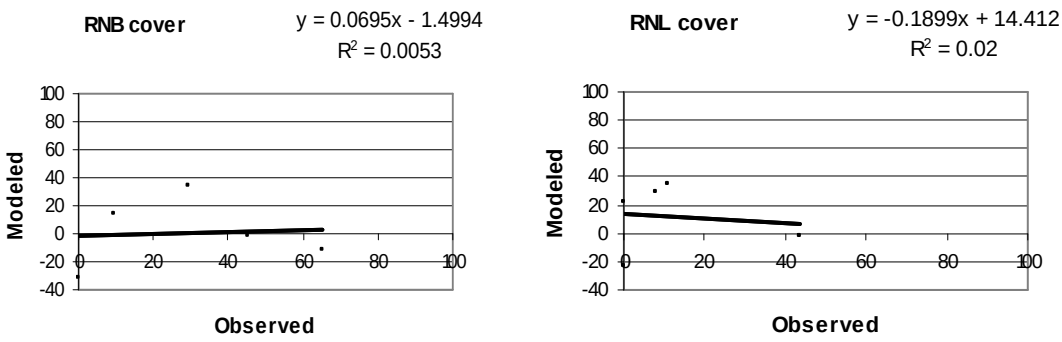


Figure 53: Herbaceous PFTs fractions predicted by unconstrained unmixing compared to fractions observed in the field in 2005

b) Semi-constrained linear spectral unmixing

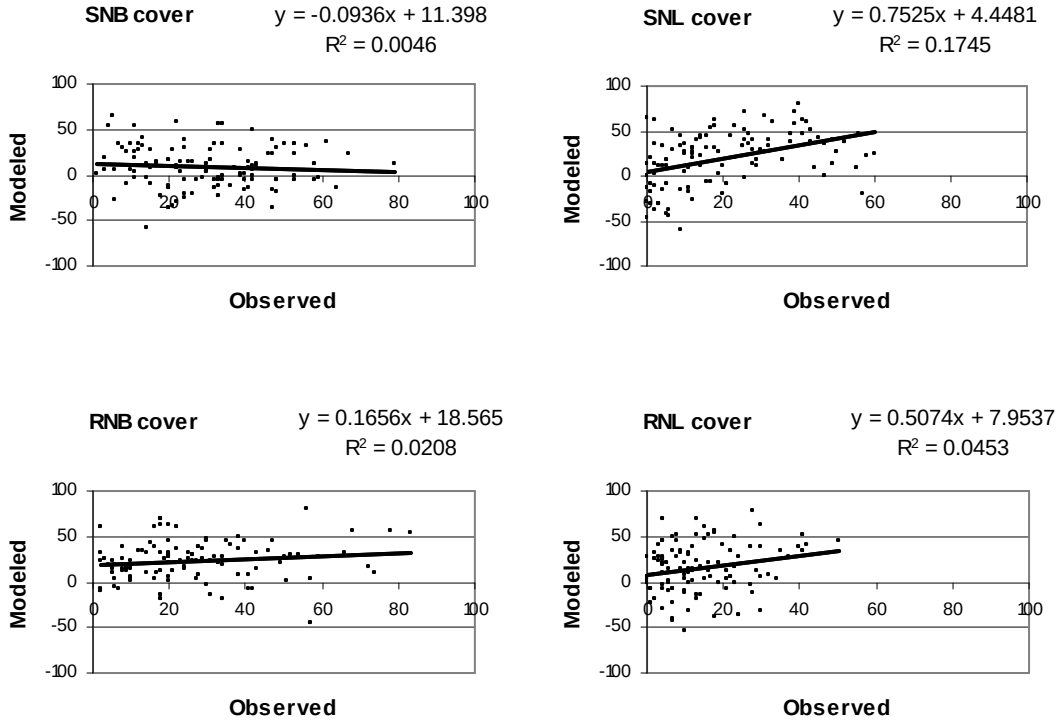
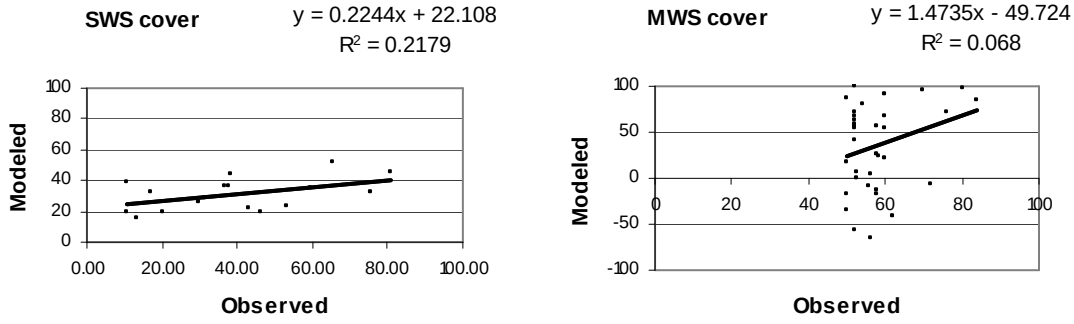


Figure 54: Herbaceous PFTs fractions predicted by semi-constrained unmixing compared to fractions observed in the field in 2002



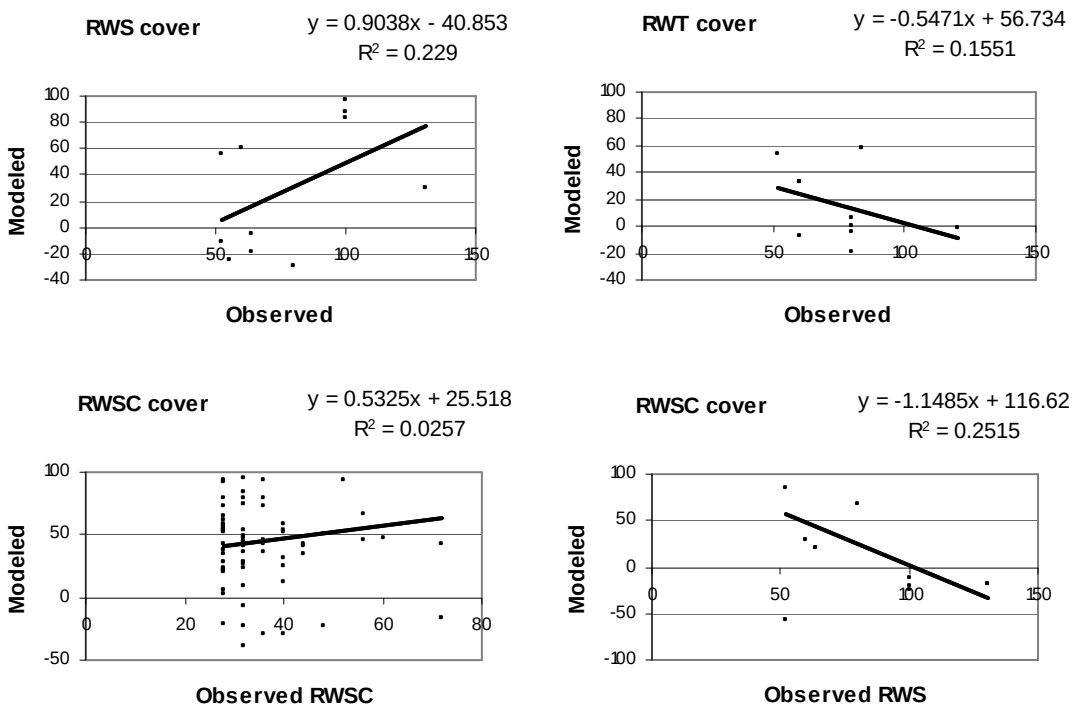


Figure 55: Woody PFTs fractions predicted by semi-constrained unmixing compared to fractions observed in the field in 2002 (SWS) and 2005-06 (rest)

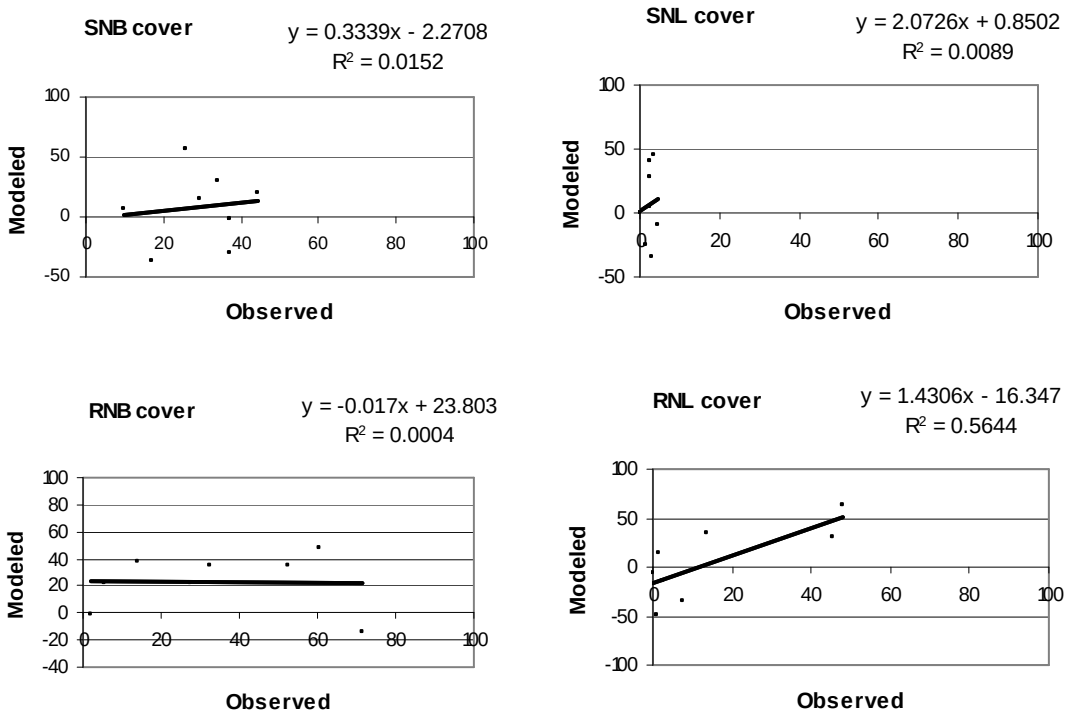


Figure 56: Herbaceous PFTs fractions predicted by semi-constrained unmixing compared to fractions observed in the field in 2004

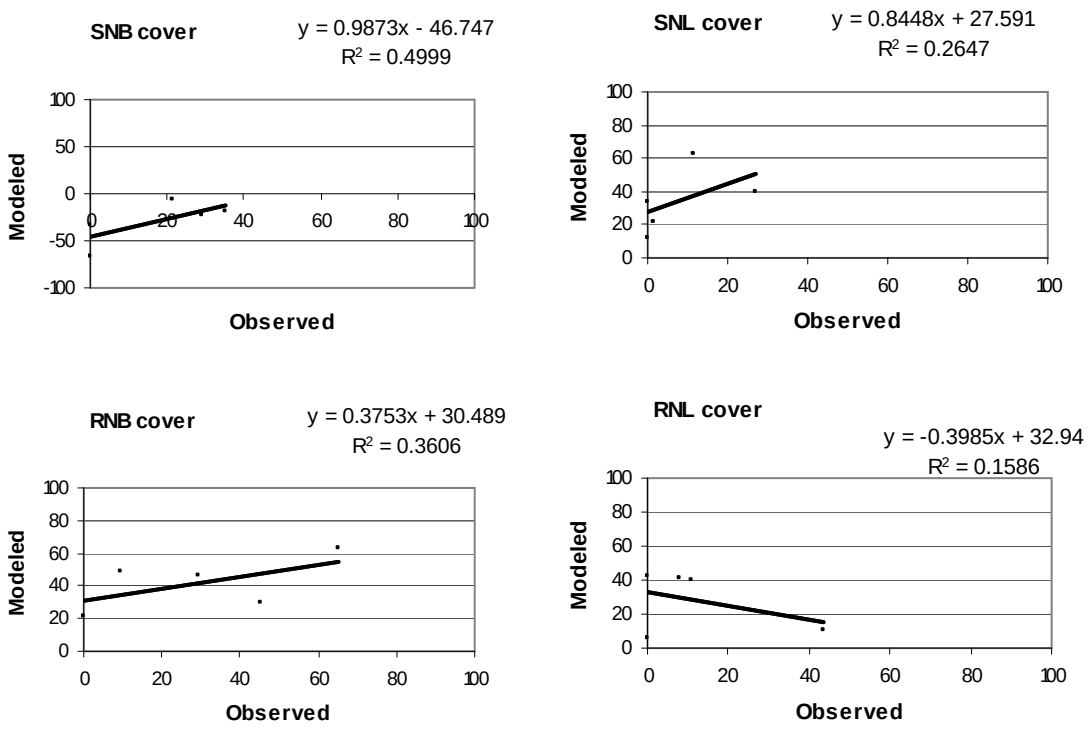


Figure 57: Herbaceous PFTs fractions predicted by semi-constrained unmixing compared to fractions observed in the field in 2005

c) Fully constrained linear spectral unmixing

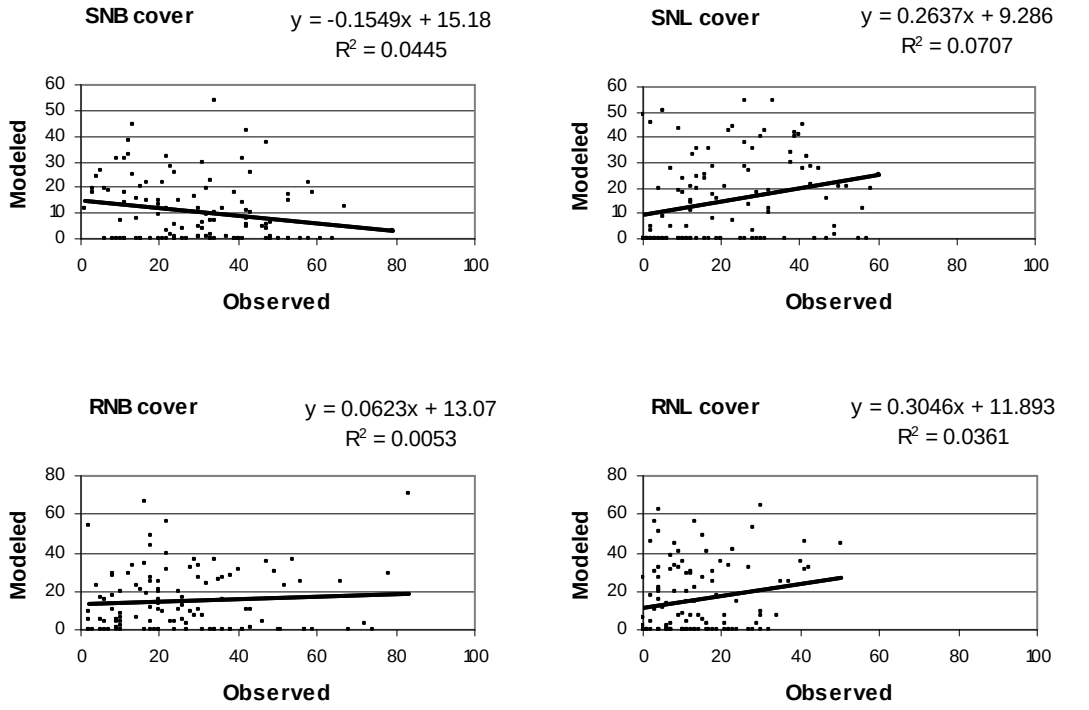


Figure 58: Herbaceous PFTs fractions predicted by fully constrained unmixing compared to fractions observed in the field in 2002

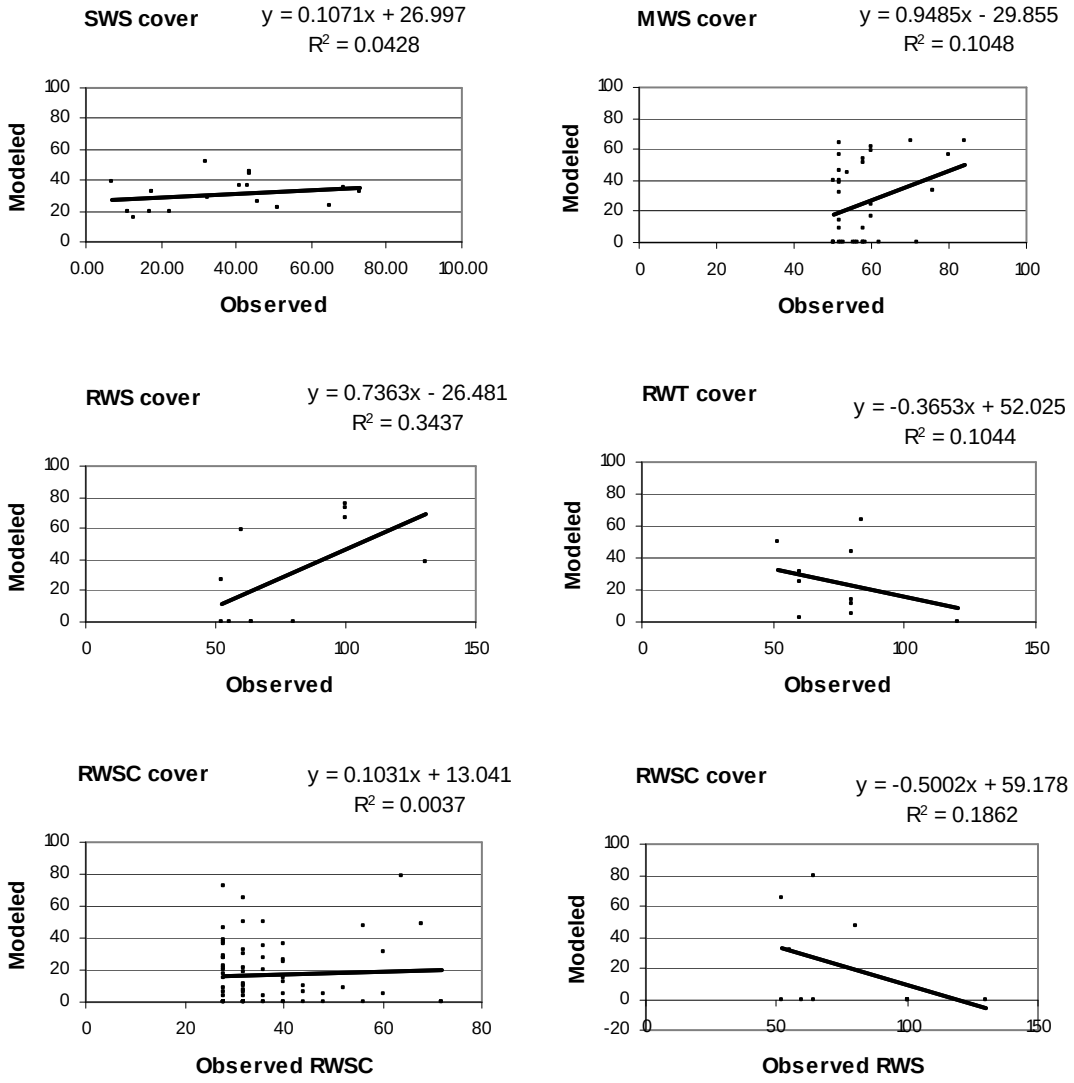
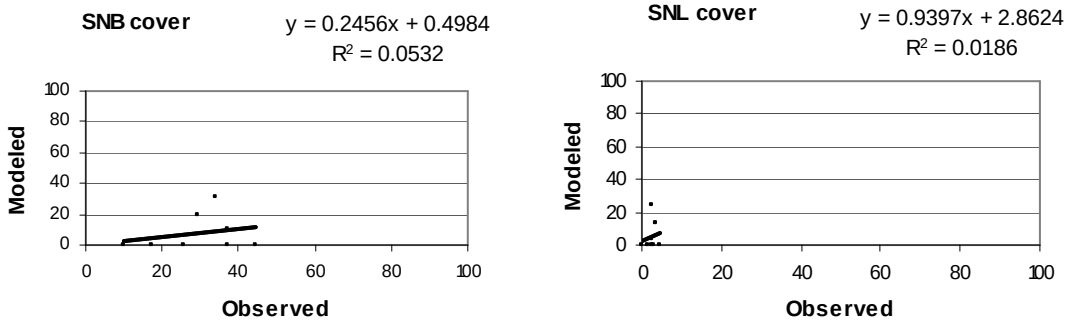


Figure 59: Woody PFTs fractions predicted by fully constrained unmixing compared to fractions observed in the field in 2002 (SWS) and 2005-06 (rest)



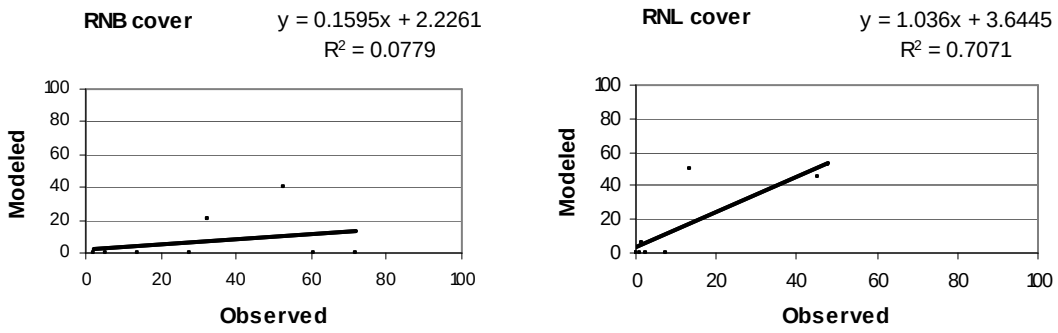


Figure 60: Herbaceous PFTs fractions predicted by fully constrained unmixing compared to fractions observed in the field in 2004

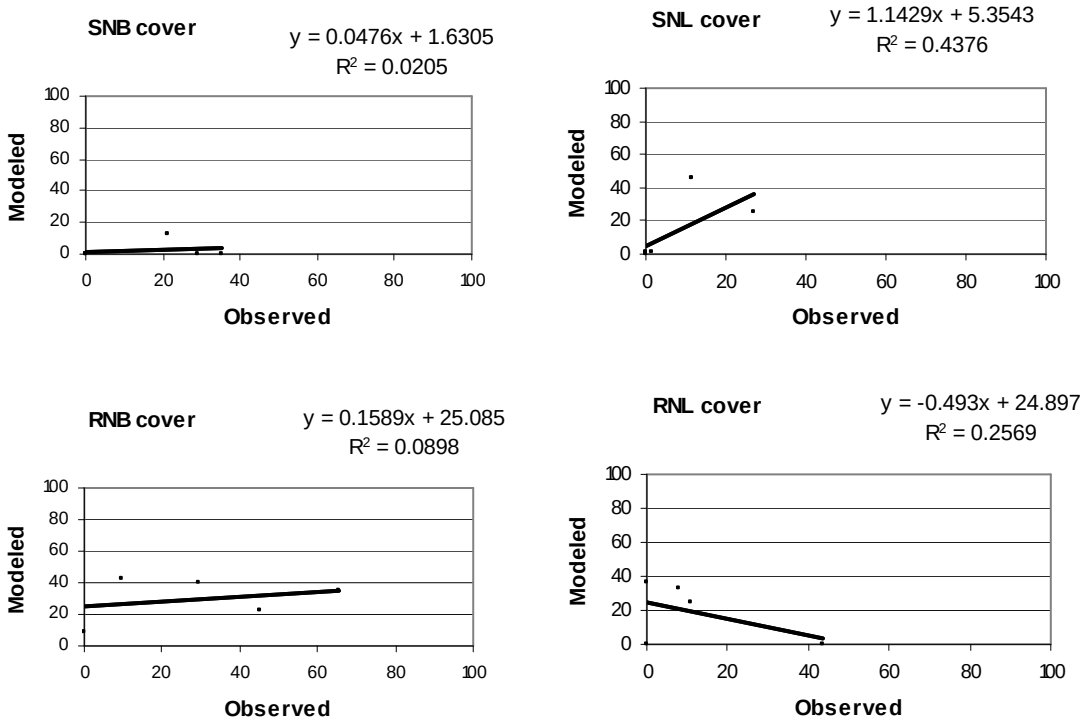


Figure 61: Herbaceous PFTs fractions predicted by fully constrained unmixing compared to fractions observed in the field in 2005

APPENDIX 5: Comparison of observed vs. modeled by overlaying plots to maps

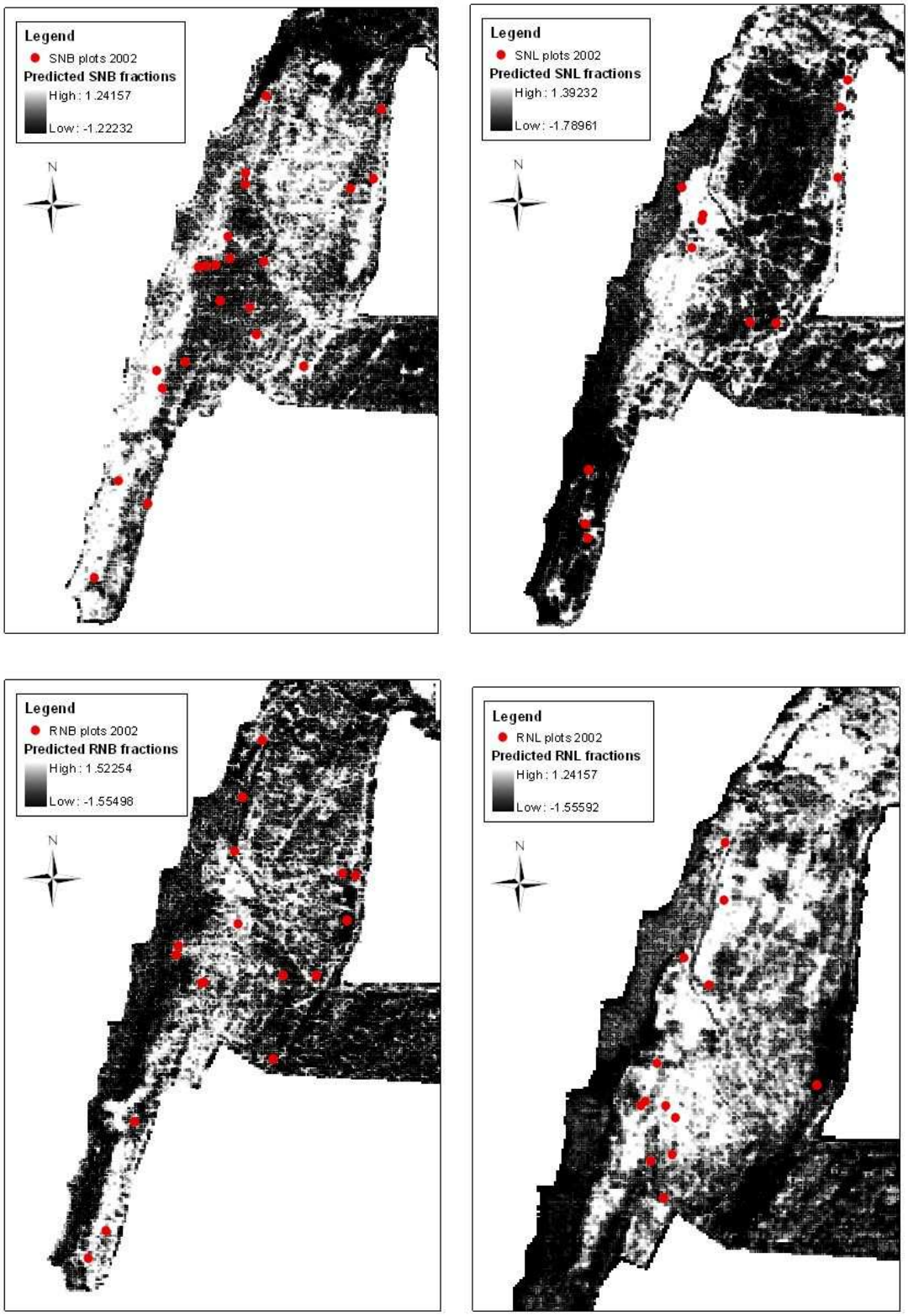


Figure 62: Abundance map of herbaceous PFTs overlaid with validation plots from 2002

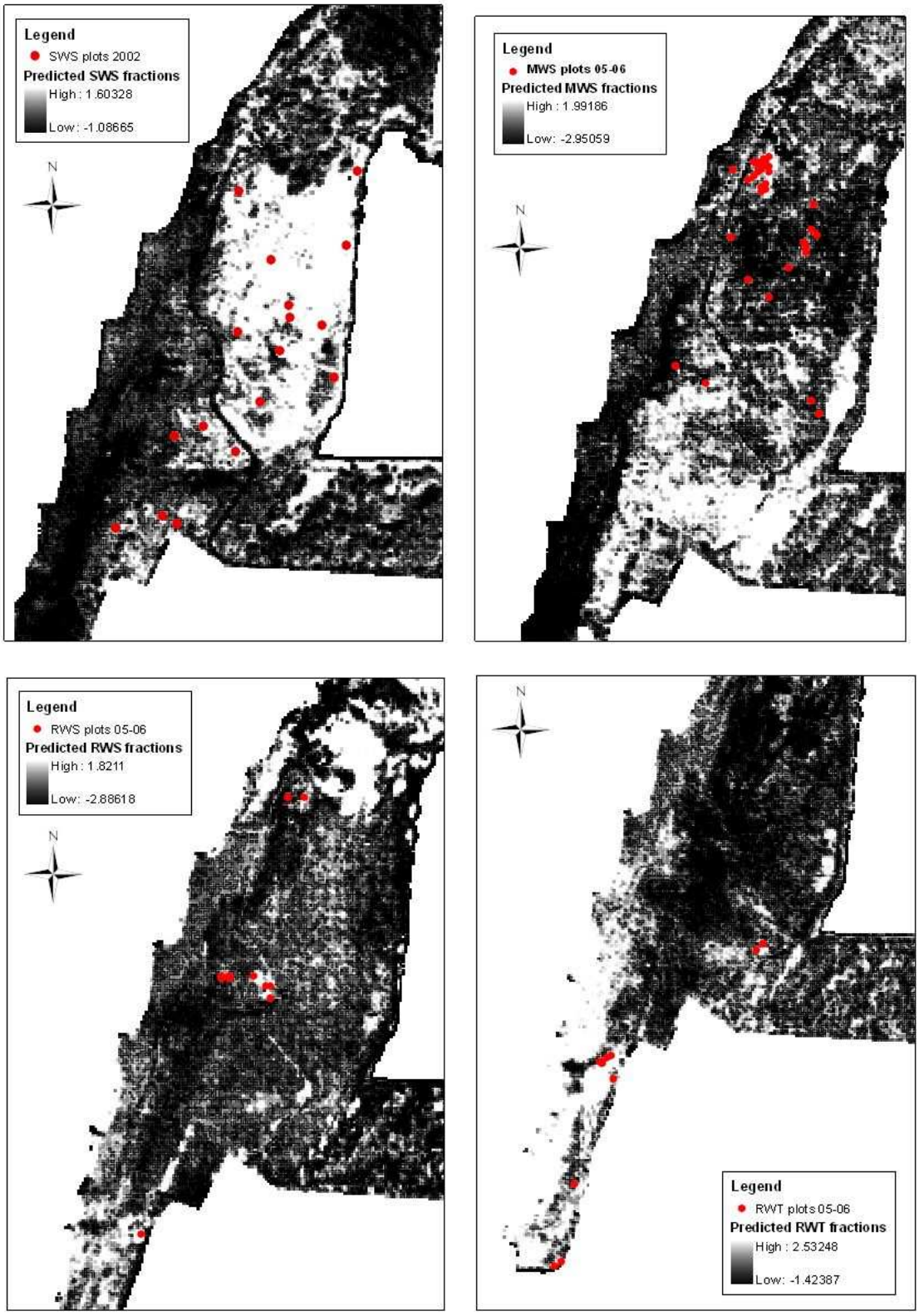


Figure 63: Abundance maps of small woody vegetation overlaid with validation plots from 2002 and medium and robust woody vegetation overlaid with validation plots from 2005-06

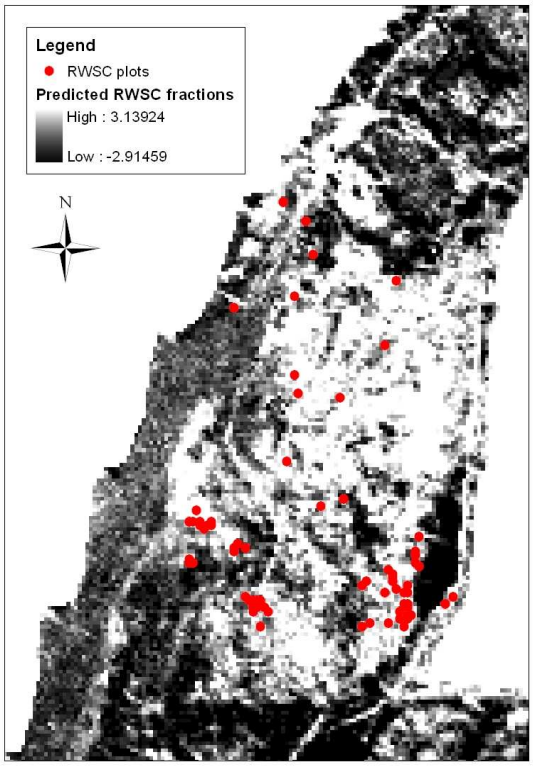


Figure 64: Abundance maps of RWSC overlaid with validation plots from 2005-06

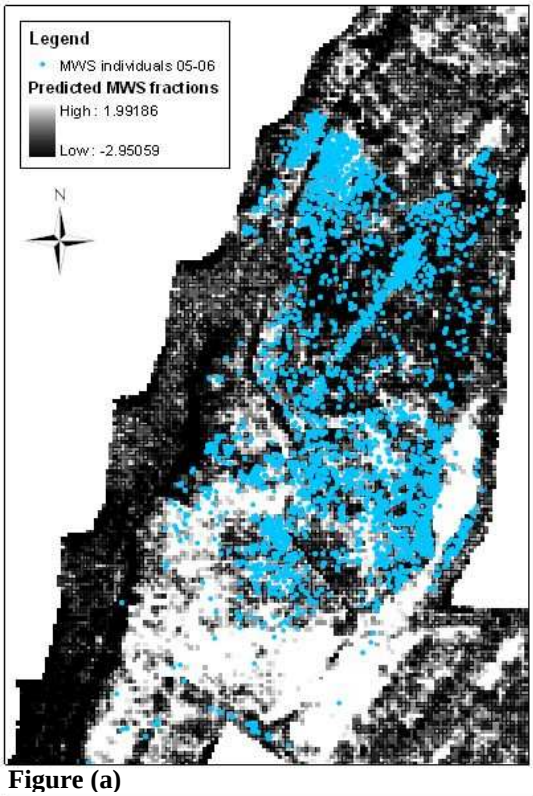


Figure (a)

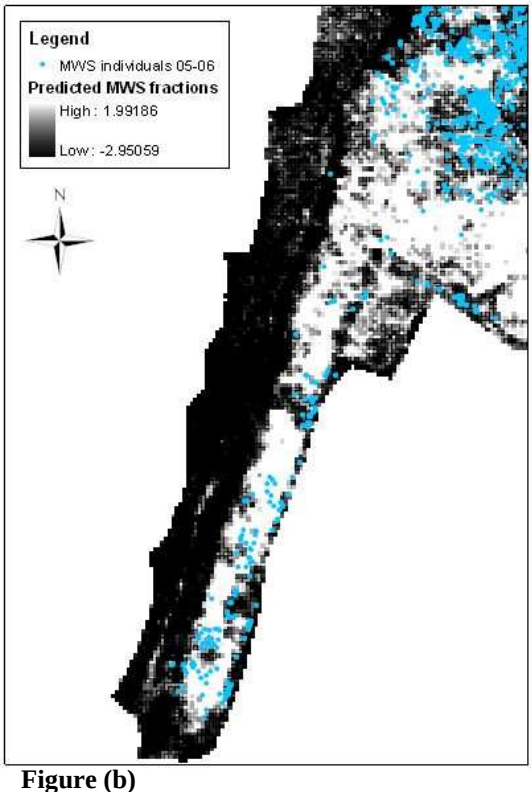


Figure (b)

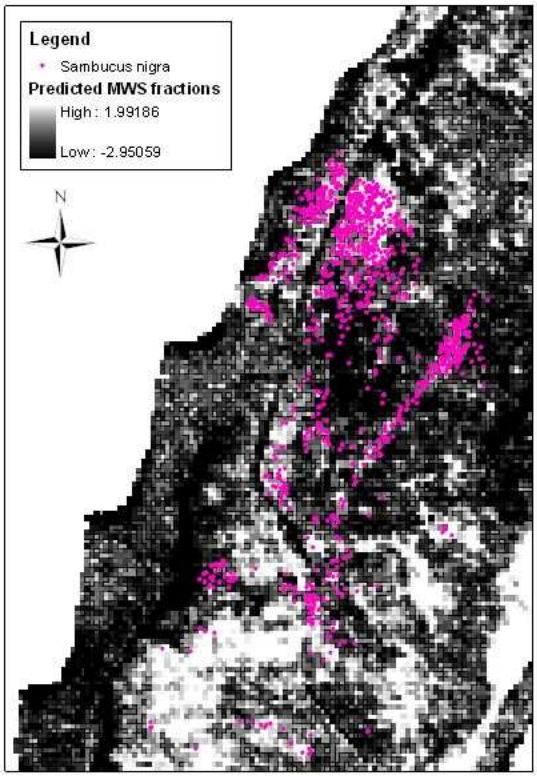


Figure (c)

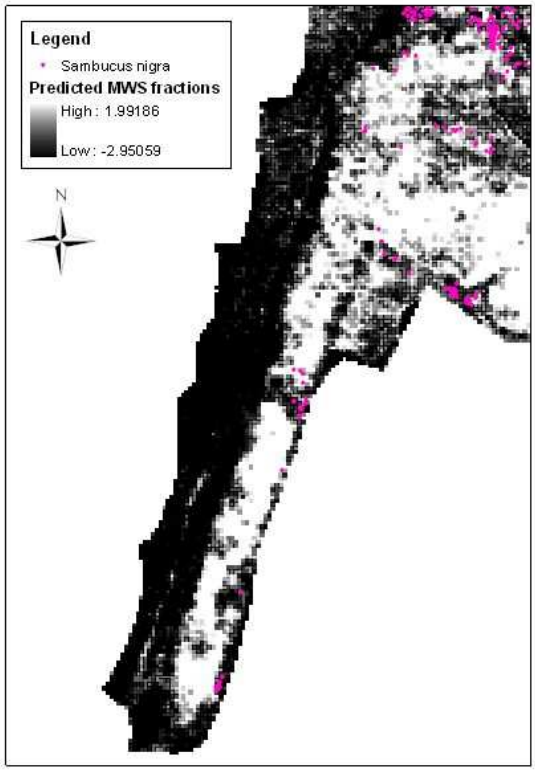


Figure (d)

Figure 65: Abundance map of MWS overlaid with MWS individual plants (a) and (b) and *Sambucus nigra* (c) and (d), field data from 2005-06

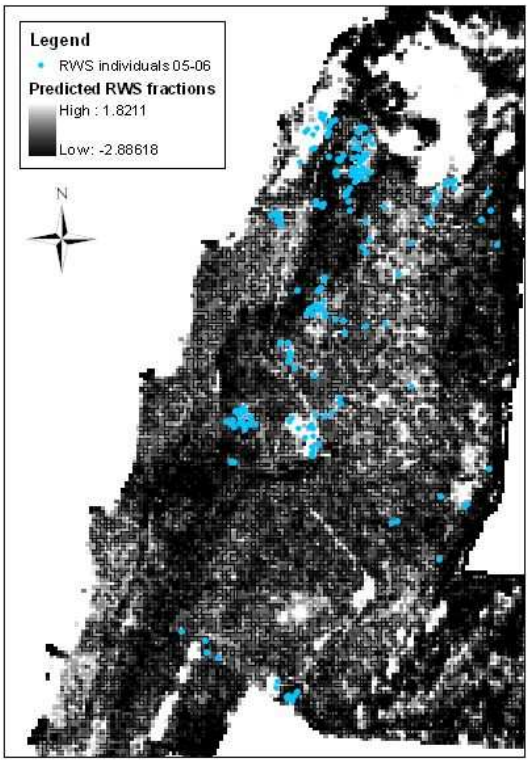


Figure (a)

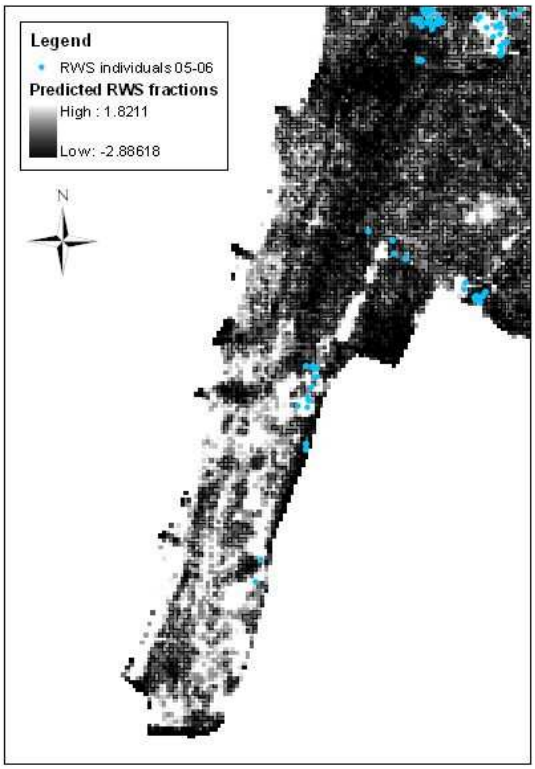


Figure (b)

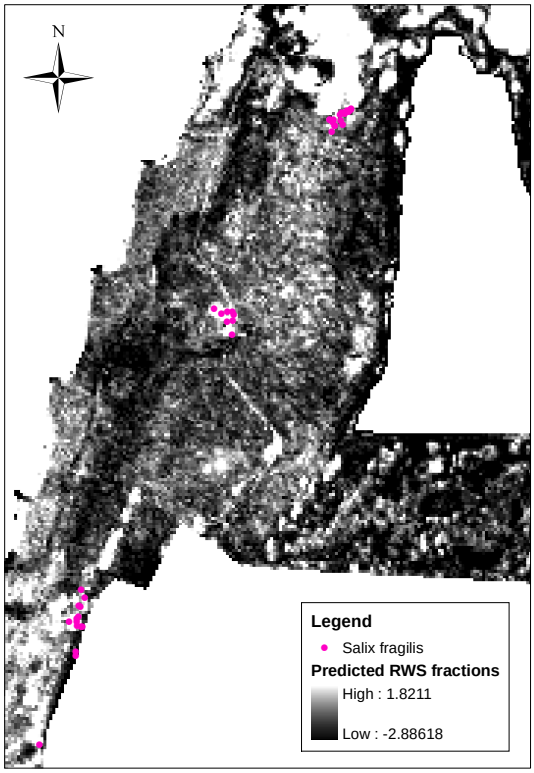


Figure (c)

Figure 66: Abundance map of RWS overlaid with RWS individual plants (a) and (b) and *Salix fragilis* (c), field data from 2005-06

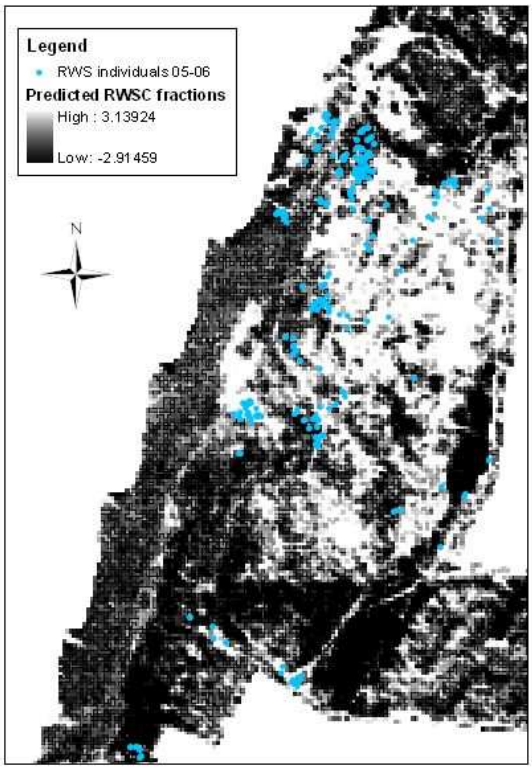


Figure (a)

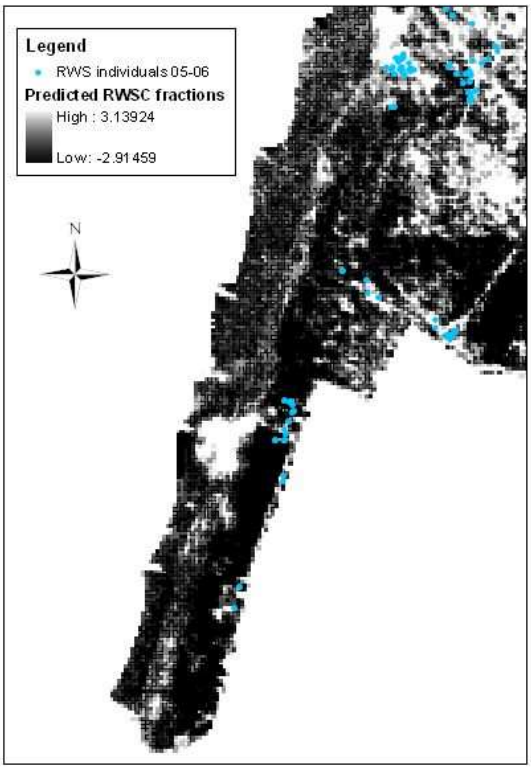


Figure (b)

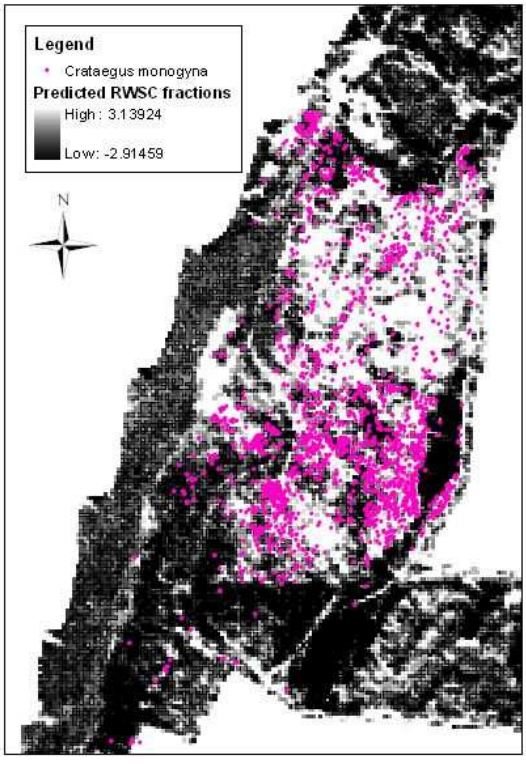


Figure (c)

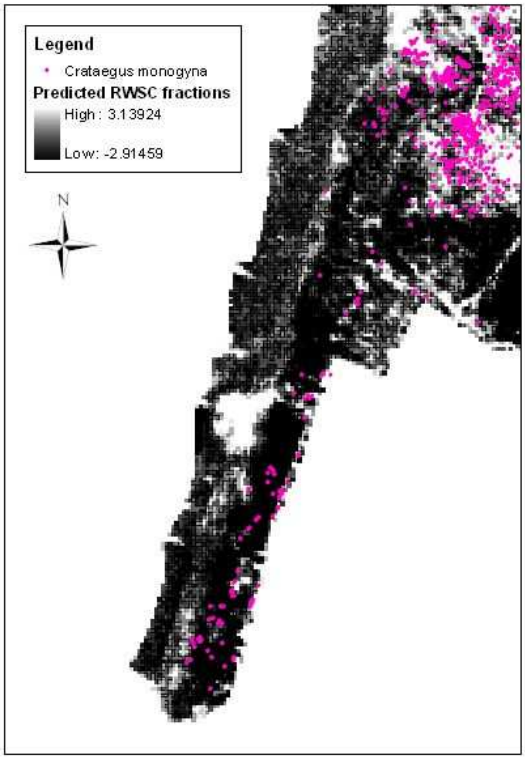


Figure (d)

Figure 67: Abundance map of RWSC overlaid with RWS individual plants (a) and (b) and *Crataegus monogyna* (c) and (d), field data from 2005-06

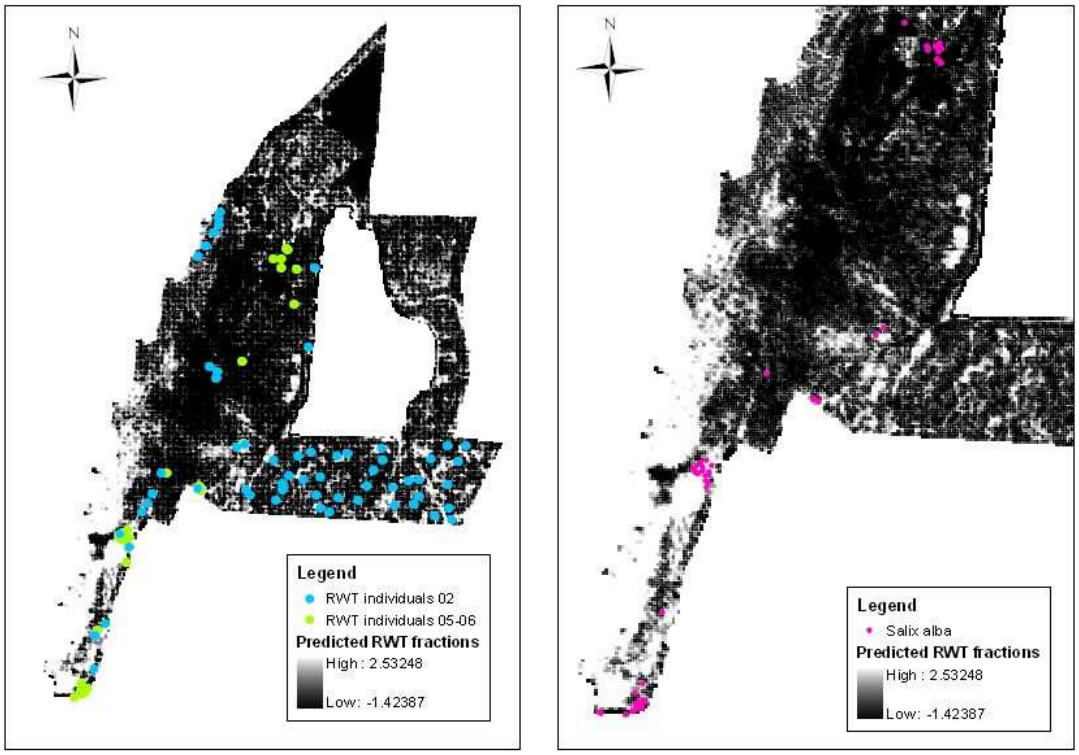


Figure (a) **Figure (b)**

Figure 68: Abundance map of RWT overlaid with RWT individual plants (a) and *Salix alba* (b), field data from 2002 and 2005-06

APPENDIX 6: OSA (Overall Sub-pixel Accuracy) spatial distribution

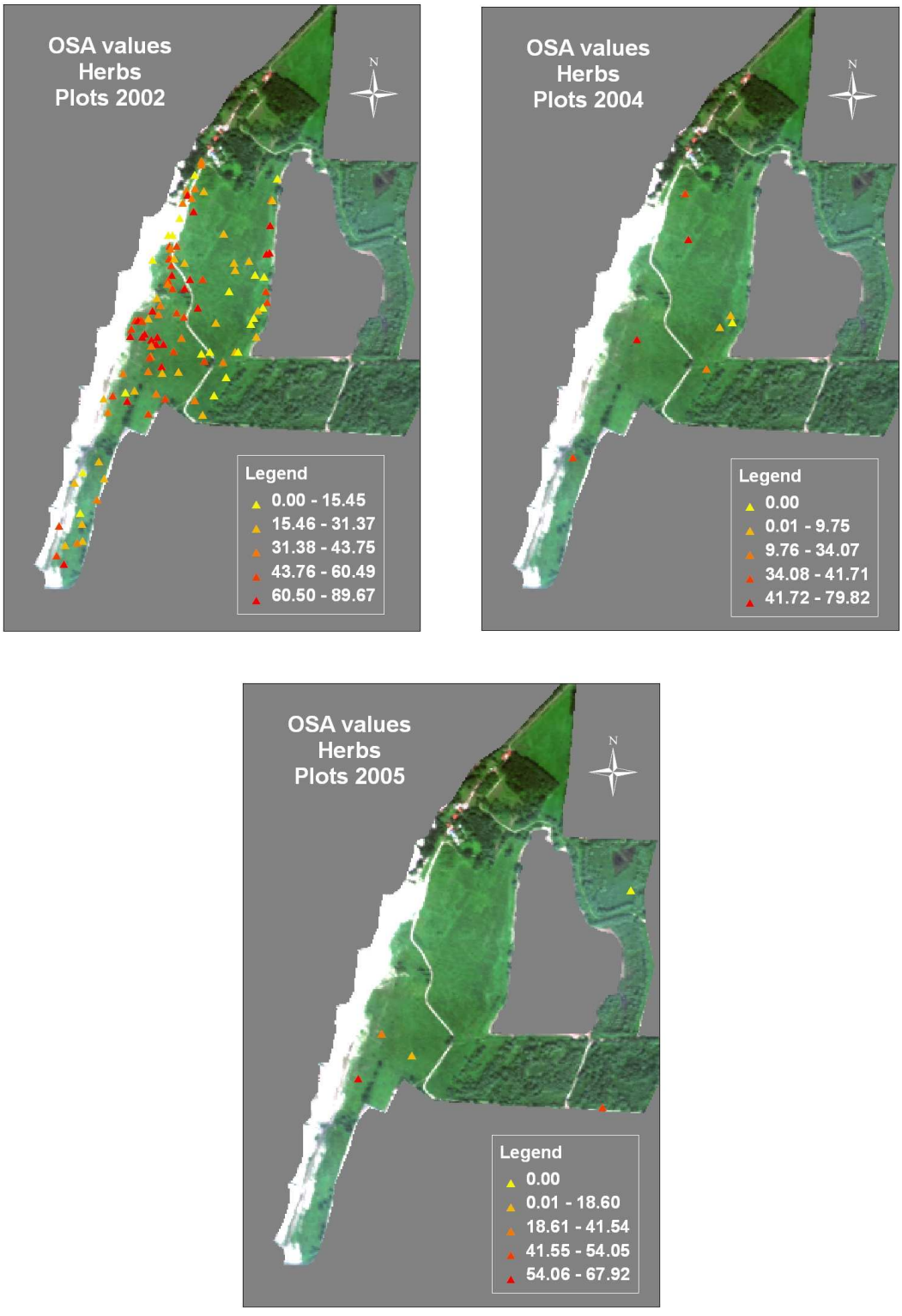


Figure 69: Overall sub-pixel accuracy spatial distribution for herbaceous vegetation (SNB, SNL, RNB and RNL)

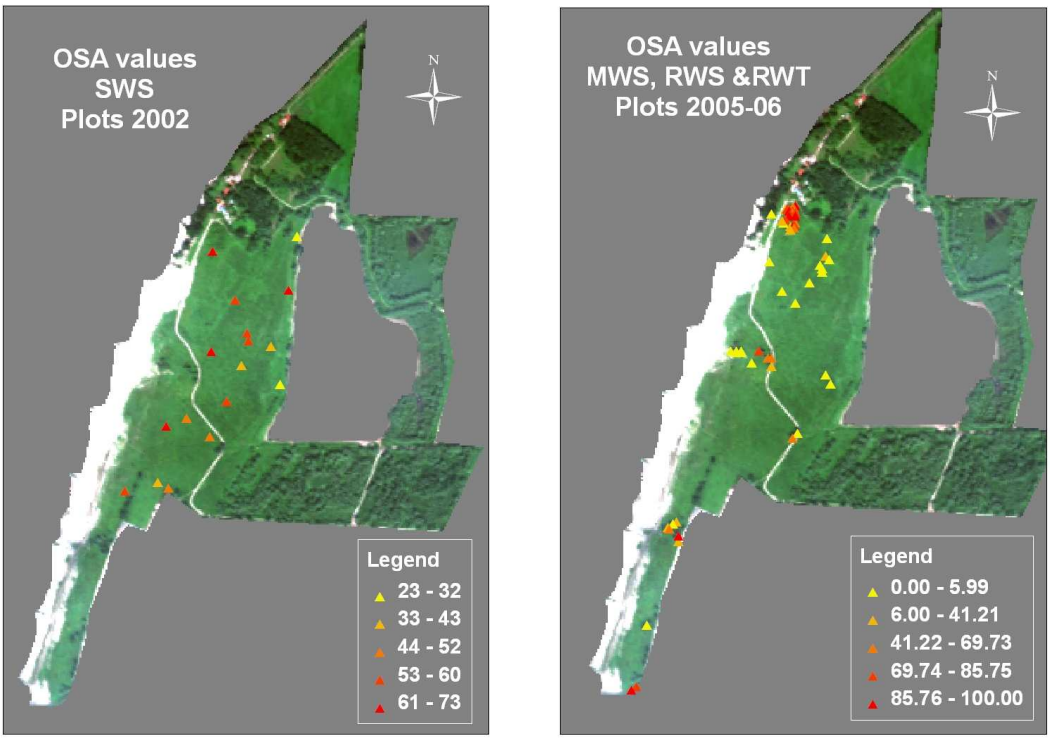


Figure 70: Overall sub-pixel accuracy spatial distribution for small woody vegetation (SWS) and medium and robust woody vegetation (MWS, RWS and RWT)

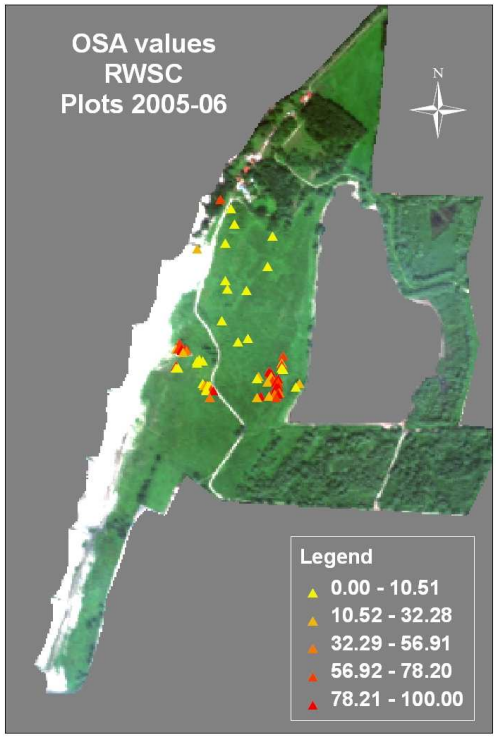


Figure 71: Overall sub-pixel accuracy spatial distribution for RWSC

APPENDIX 7: Temporal analysis. Endmembers

a) CASI image

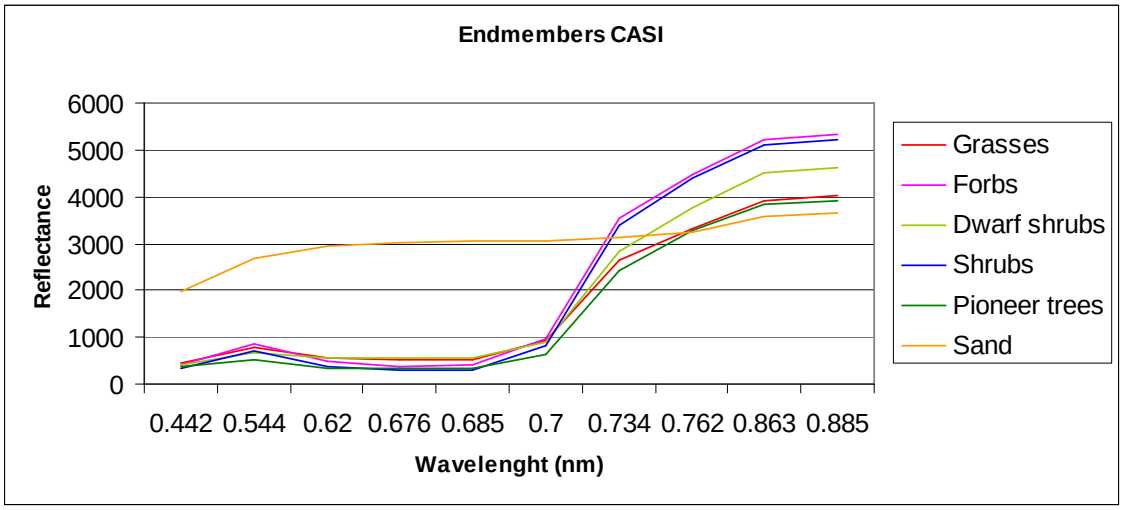


Figure 72: Endmembers spectra from CASI

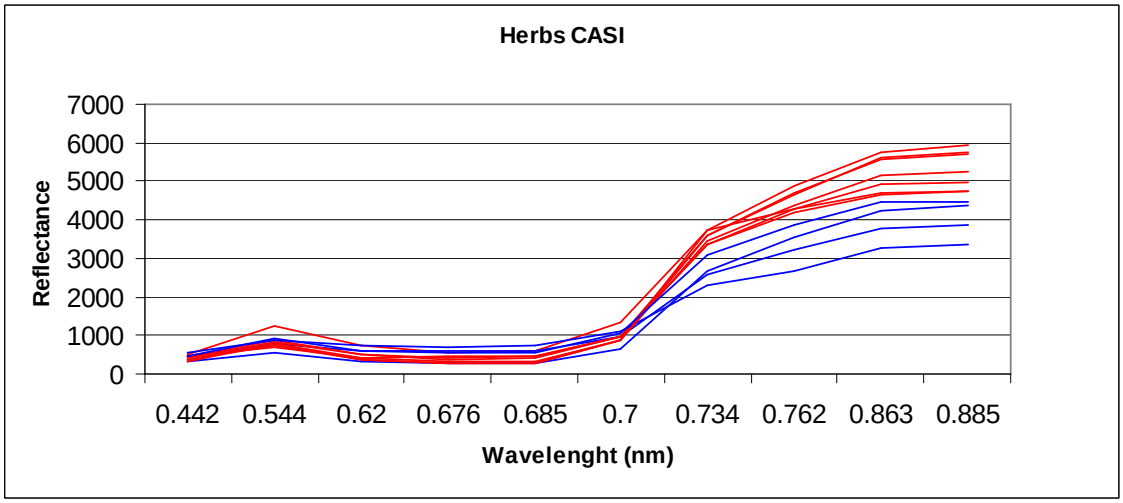


Figure 73: Spectrum of various herbaceous plots (CASI). Red spectra correspond to forbs while blue correspond to grasses

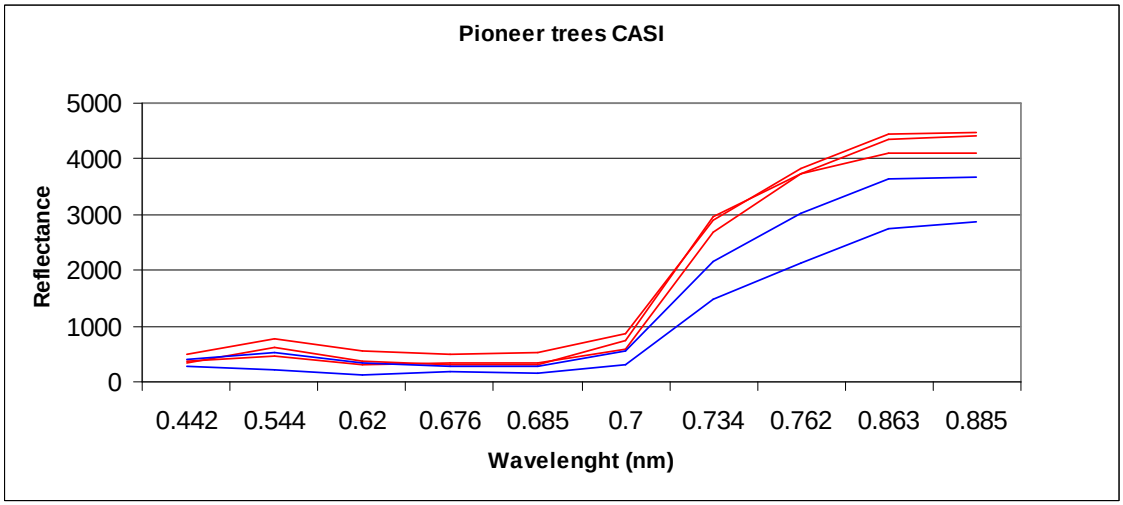


Figure 74: Spectrum of Pioneer trees (*Salix sp.*) (CASI). Red spectra correspond to isolated individuals or small groups in the western area while blue spectra correspond to the forested area on the East

b) HyMap image

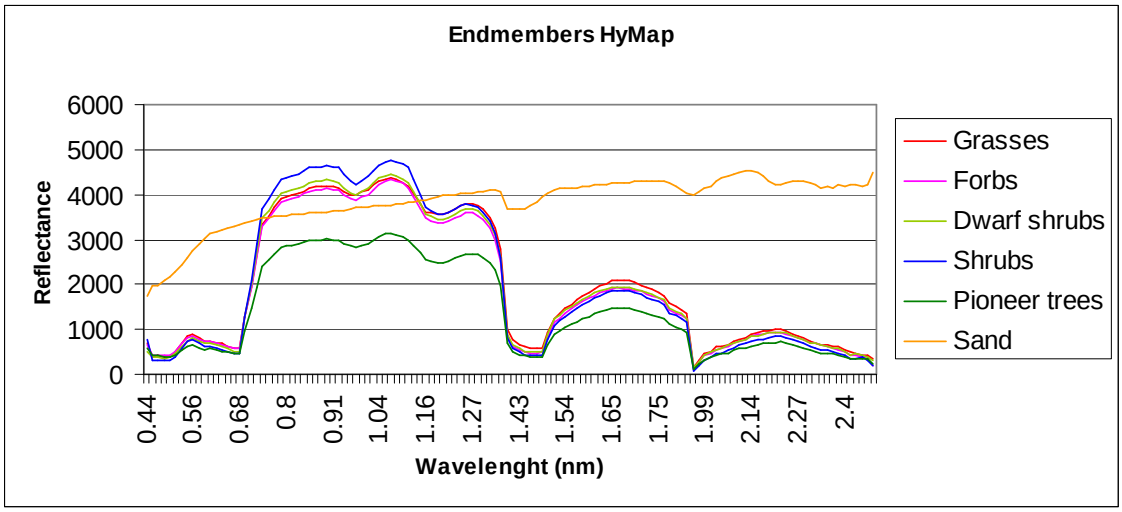


Figure 75: Endmembers spectra from HyMap

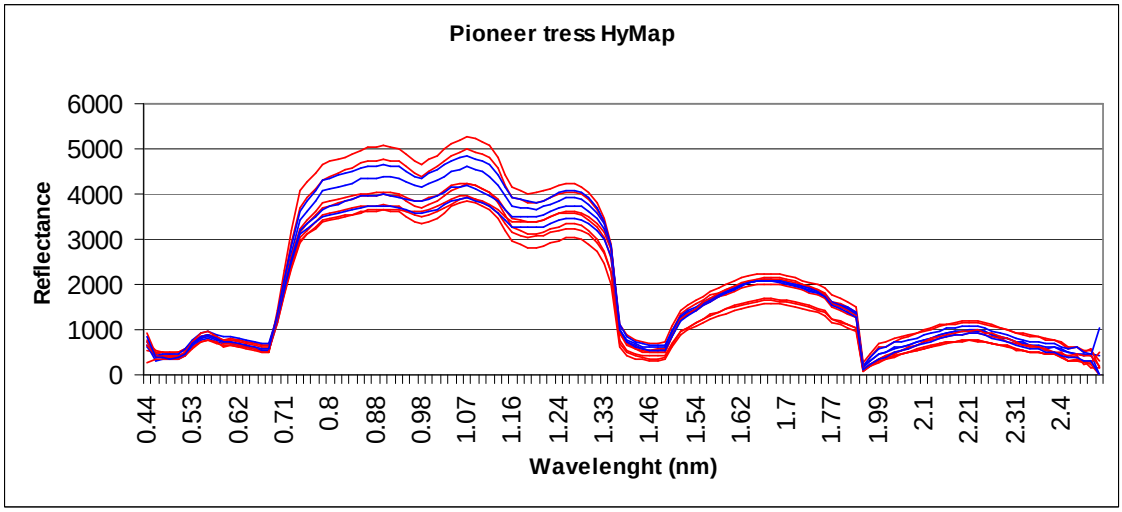


Figure 76: Spectrum of various herbaceous plots (HyMap). Red spectra correspond to forbs while blue correspond to grasses

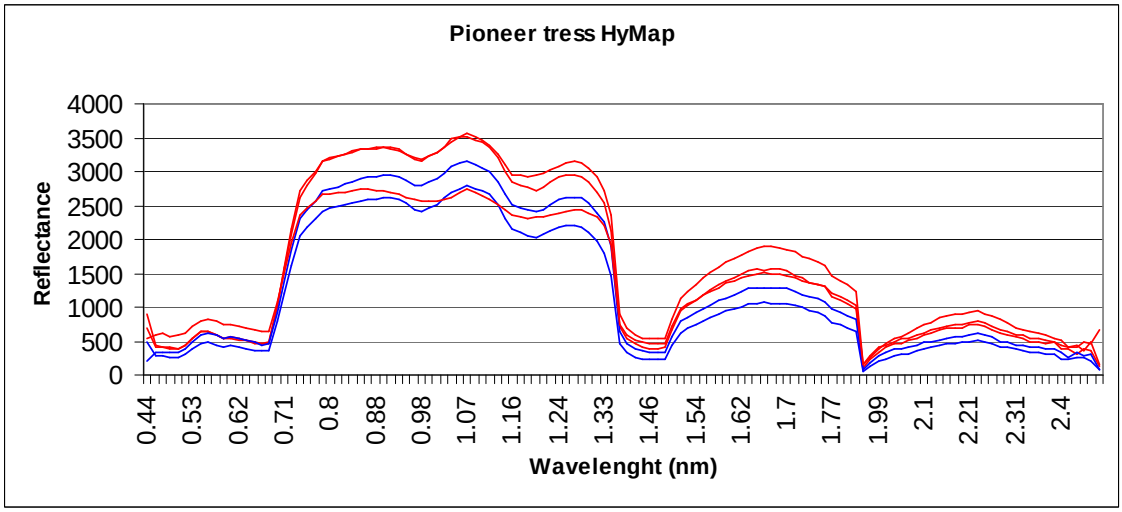


Figure 77: Spectrum of Pioneer trees (*Salix sp.*) (HyMap). Red spectra correspond to isolated individuals or small groups in the western area while blue spectra correspond to the forested area on the East

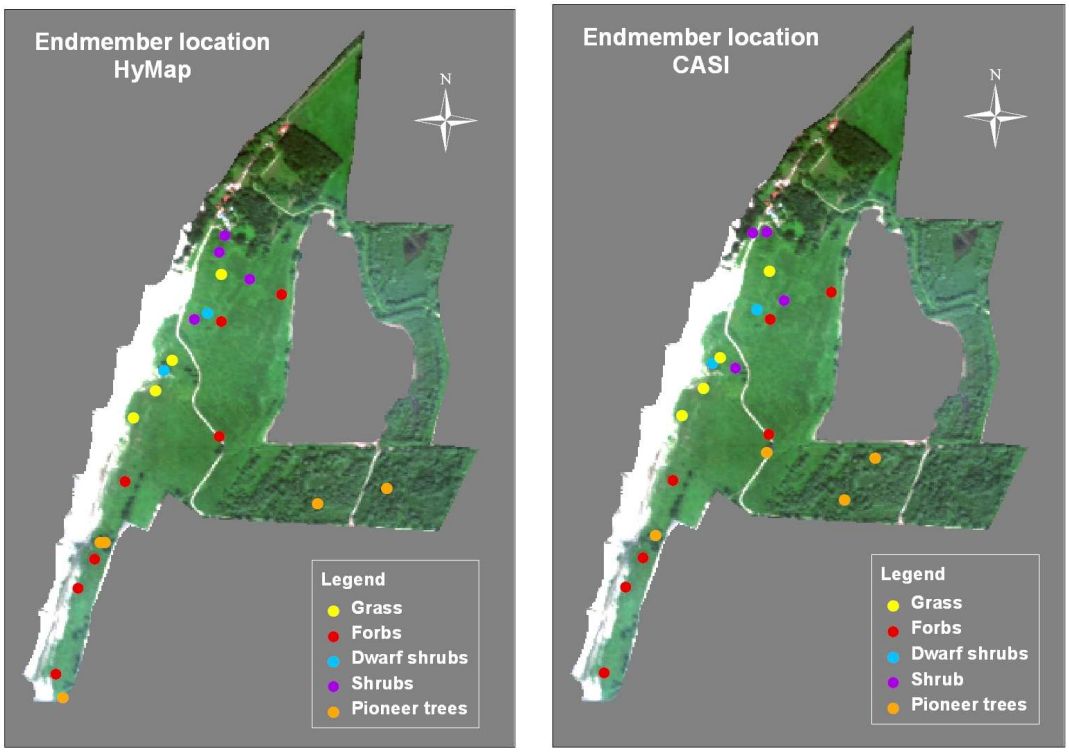


Figure 78: Endmember location for HyMap and CASI

APPENDIX 8: Temporal analysis. Abundance maps and RMSE maps

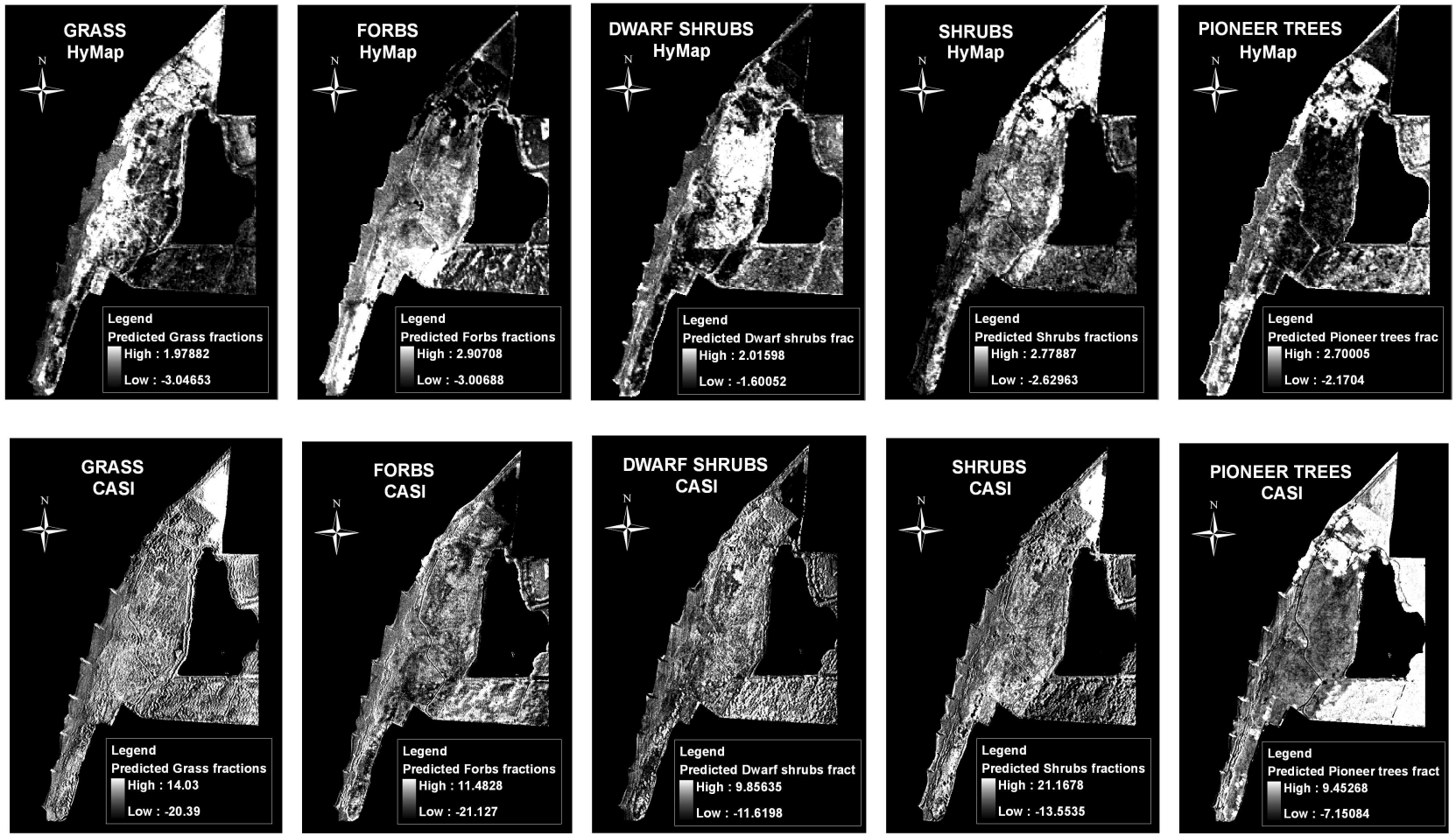


Figure 79: Abundance maps of the PFTs classes Grass, Forbs, Dwarf shrubs, Shrubs and Pioneer tress. First row corresponds to HyMap unconstrained unmixing and second row corresponds to CASI unconstrained unmixing.

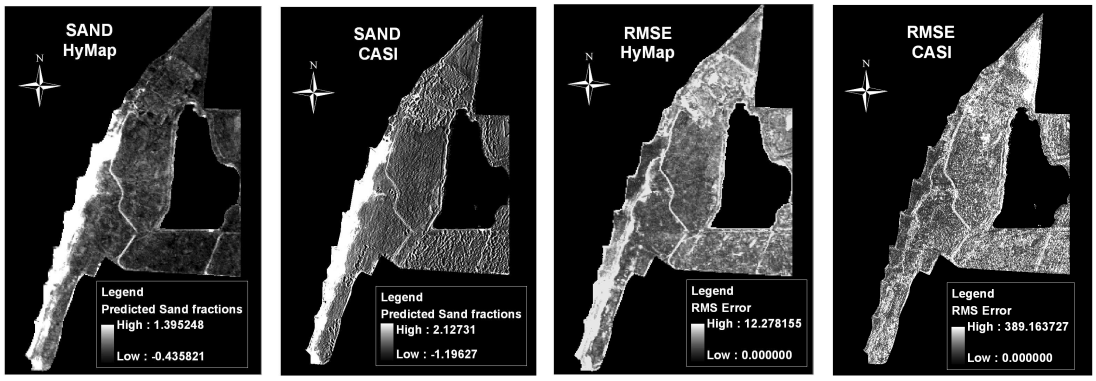


Figure (a) **Figure (b)** **Figure (c)** **Figure (d)**

Figure 80: Abundance maps for Sand and RMSE from unconstrained unmixing, HyMap (a)/(c) and CASI (b)/(d)

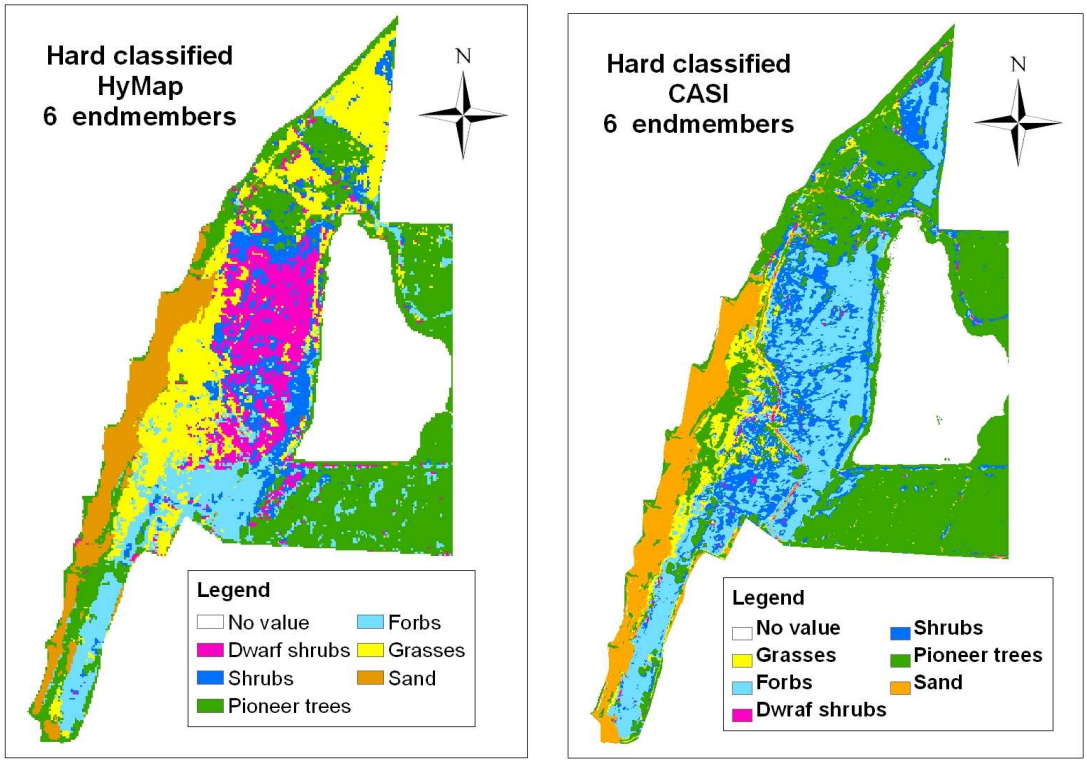


Figure (a) **Figure (b)**

Figure 81: Hard classified maps from fully constrained unmixing, HyMap (a) and CASI (b)

APPENDIX 9: Temporal analysis. Comparison of observed vs. modeled by means of plots

a) HyMap Unconstrained Unmixing

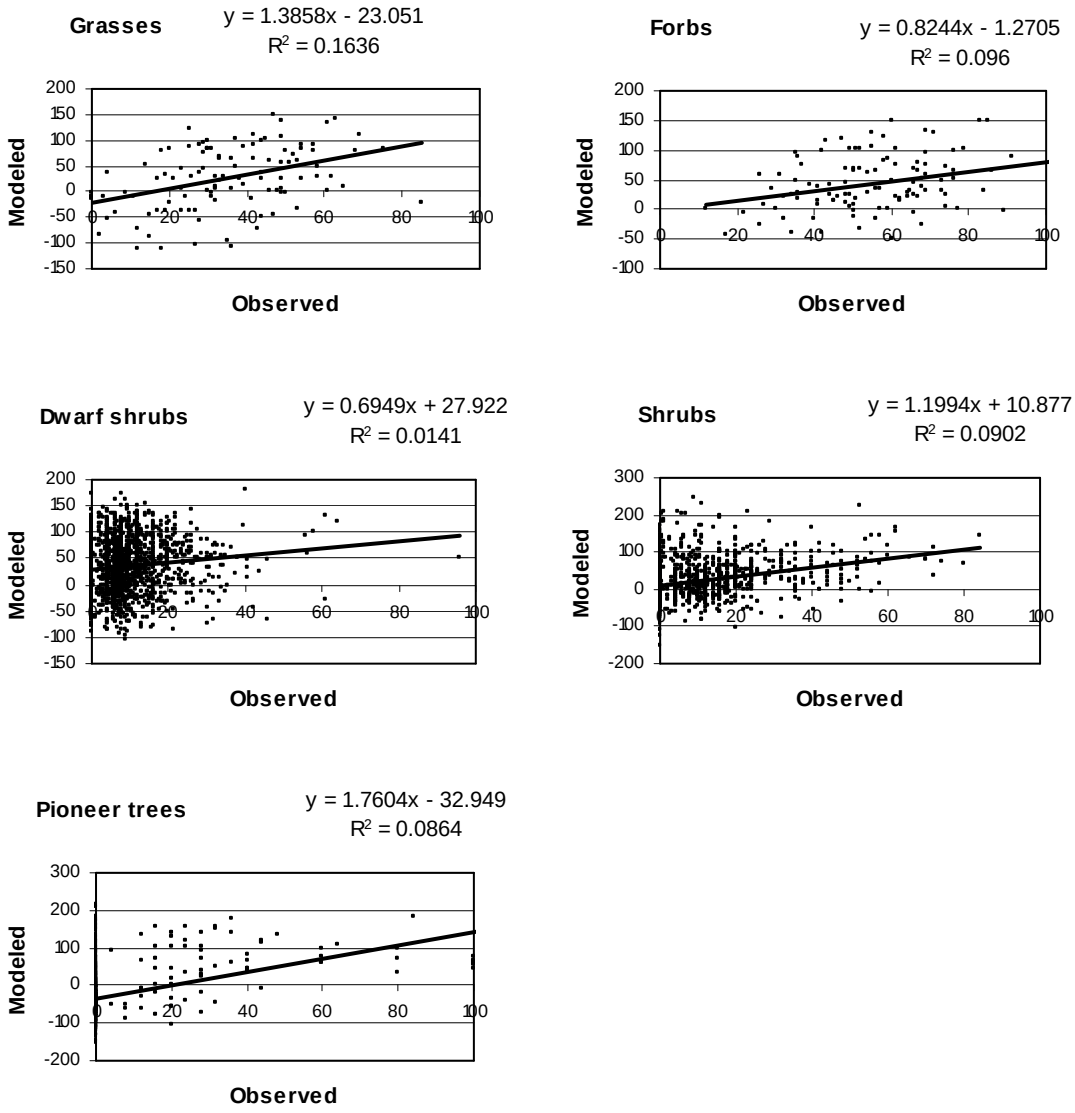


Figure 82: PFTs fractions predicted by HyMap unconstrained unmixing compared to fractions observed in the field in the years 2002 (Grass and Forbs) and 2005-06 (Dwarf shrubs, Shrubs and Pioneer trees)

b) CASI Unconstrained Unmixing

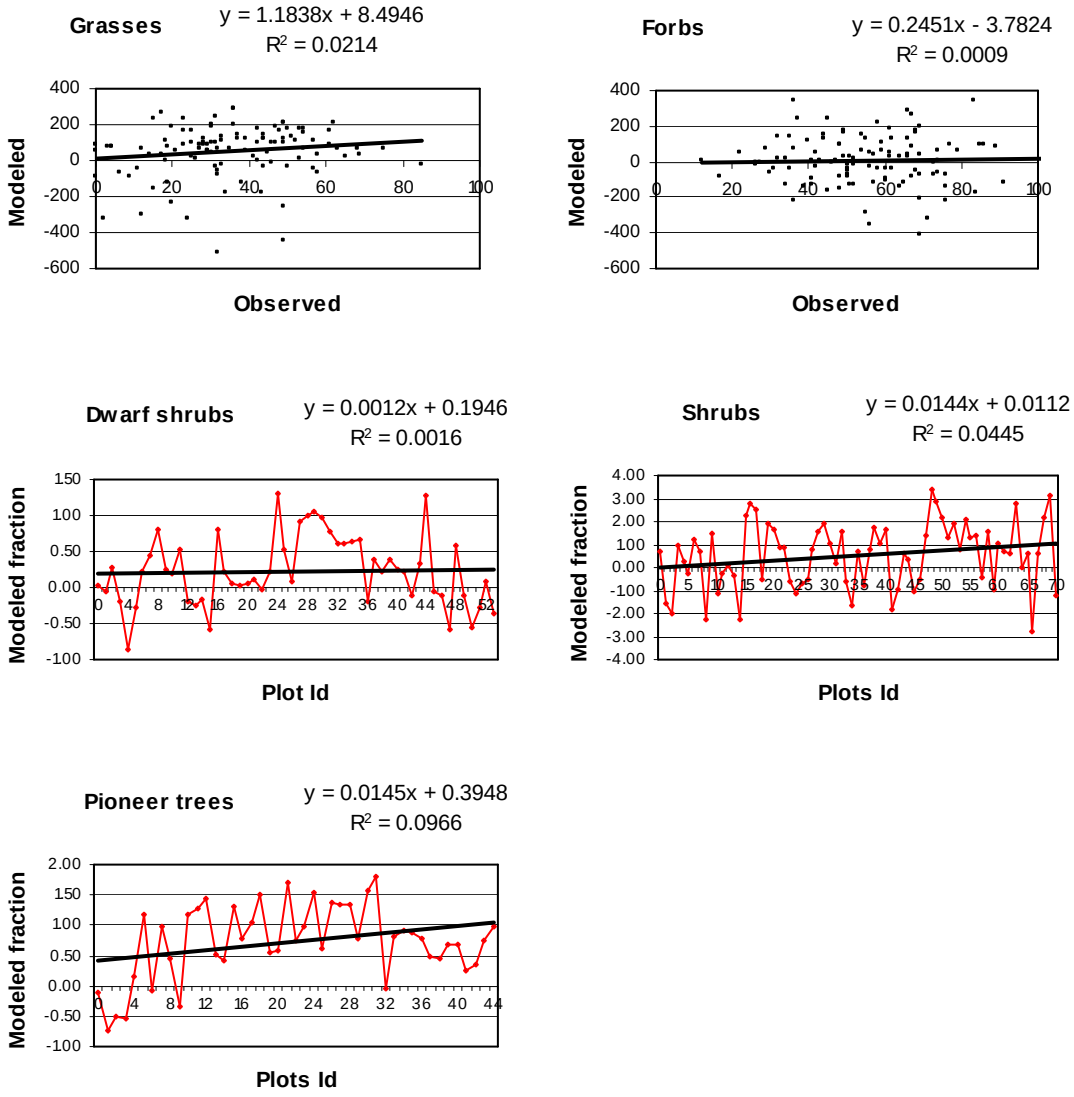


Figure 83: PFTs fractions predicted by CASI unconstrained unmixing compared to fractions observed in the field in the years 2002 (Grass and Forbs) and 2005-06 (Dwarf shrubs, Shrubs and Pioneer trees)

APPENDIX 10: Temporal analysis. Comparison of observed vs. modeled by overlaying plots to maps

a) HyMap Unconstrained Unmixing

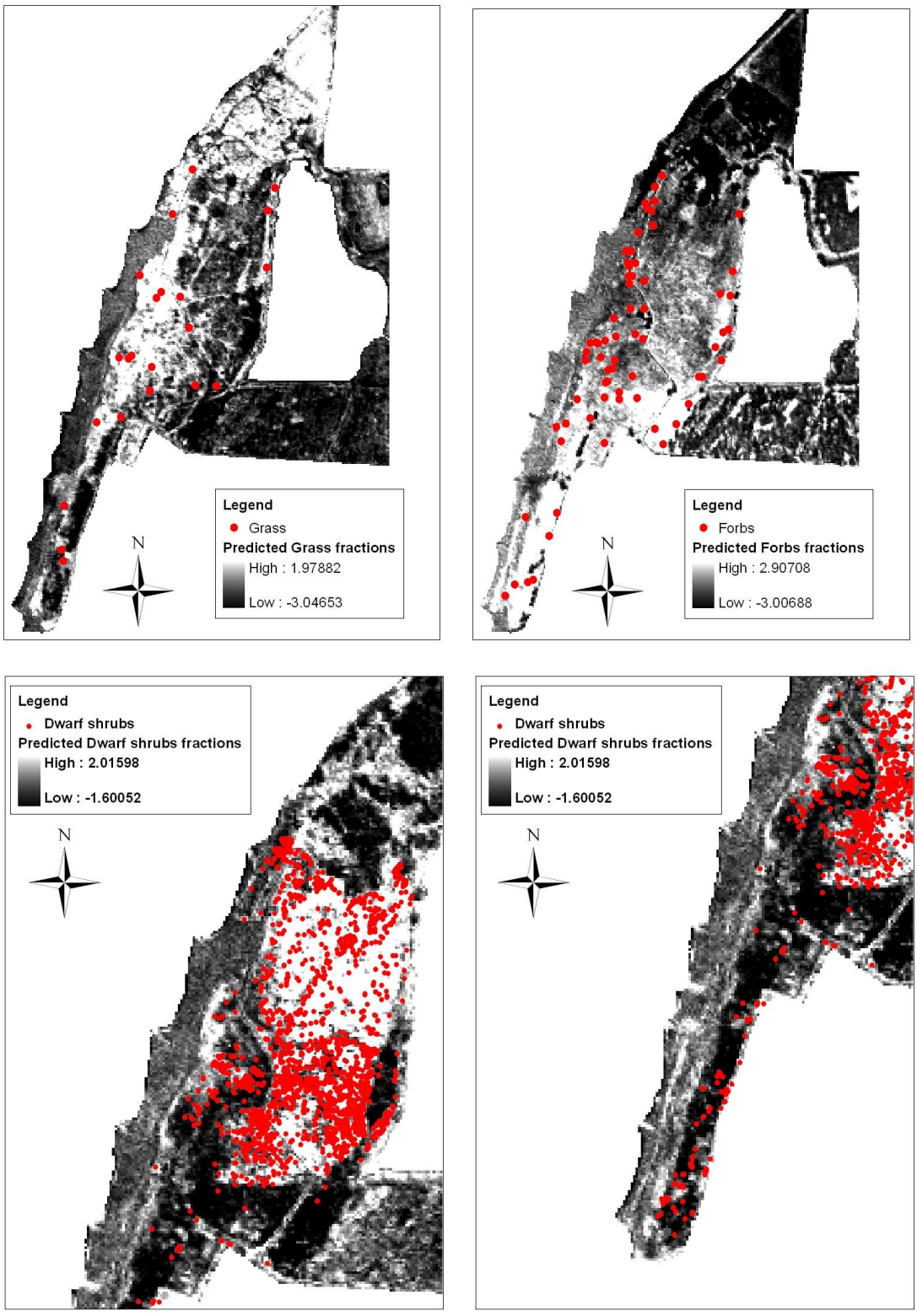


Figure 84: HyMap abundance maps of Grass, Forbs and Dwarf shrubs overlaid with individual plants from each class respectively, field data from 2002 (Grass and Forbs) and 2005-06 (Dwarf shrubs)

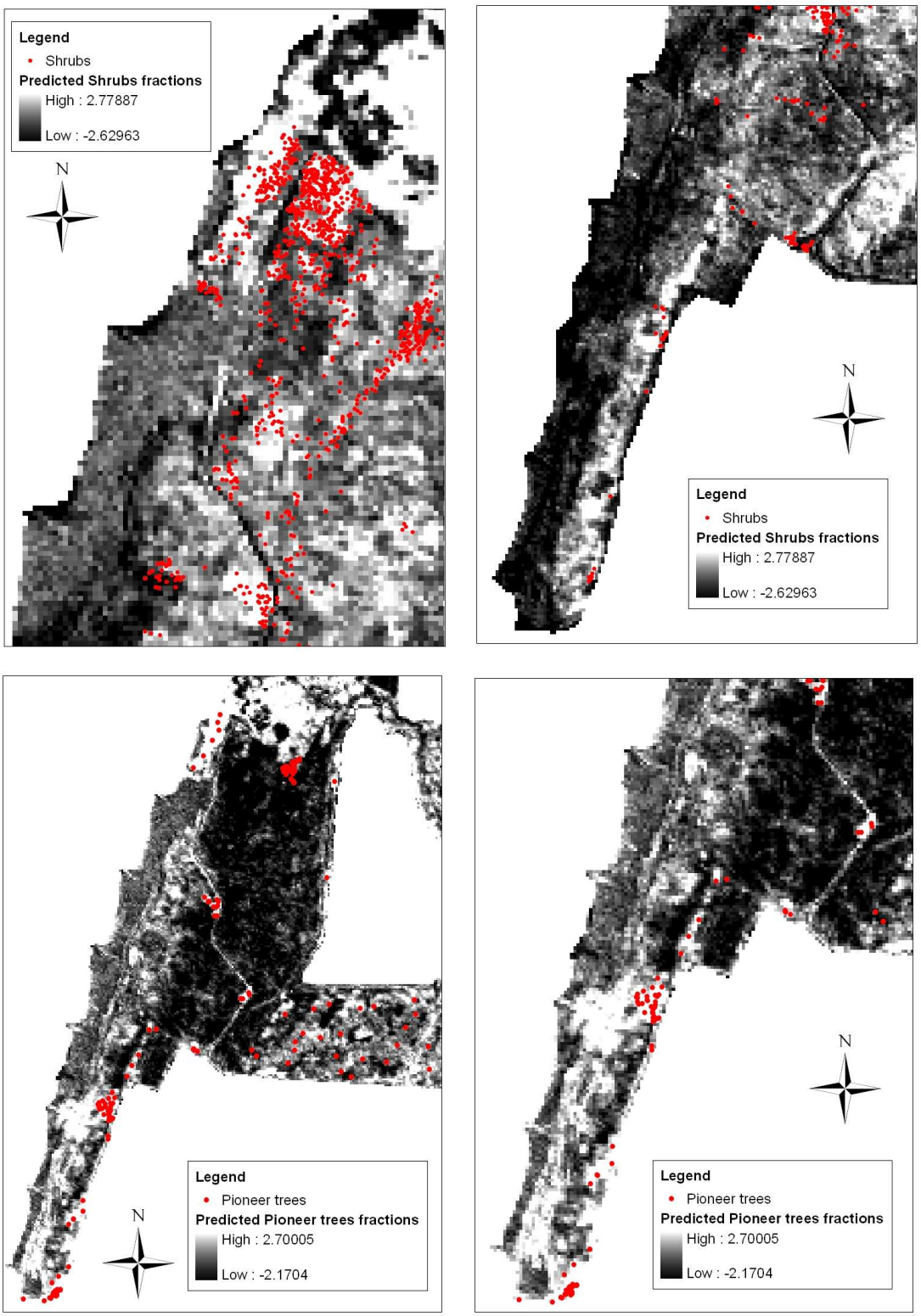


Figure 85: HyMap abundance maps of Shrubs and Pioneer trees overlaid with individual plants from each class respectively, field data from 2005-06

b) CASI Unconstrained Unmixing

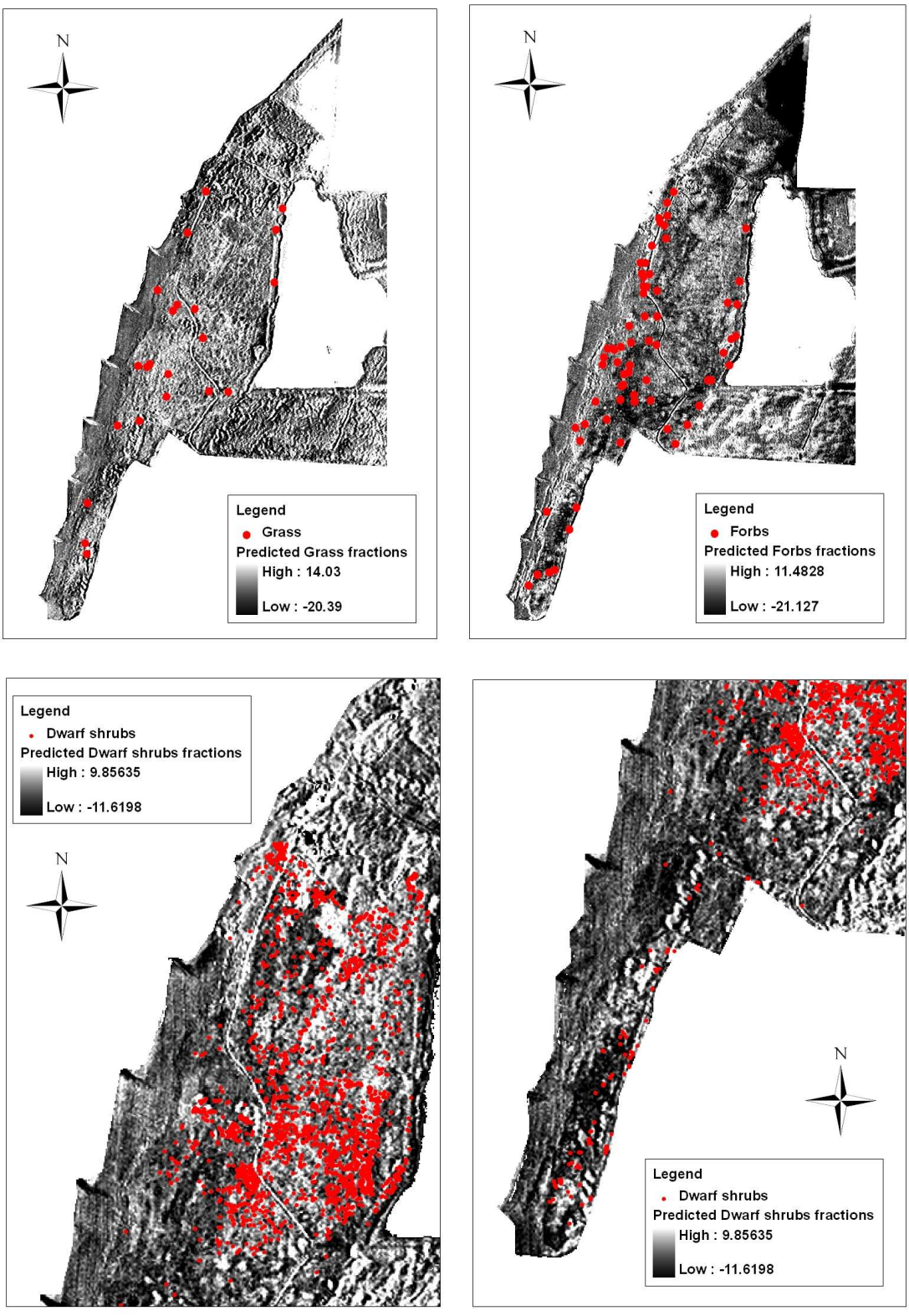


Figure 86: CASI abundance maps of Grass, Forbs and Dwarf shrubs overlaid with individual plants from each class respectively, field data from 2002 (Grass and Forbs) and 2005-06 (Dwarf shrubs)

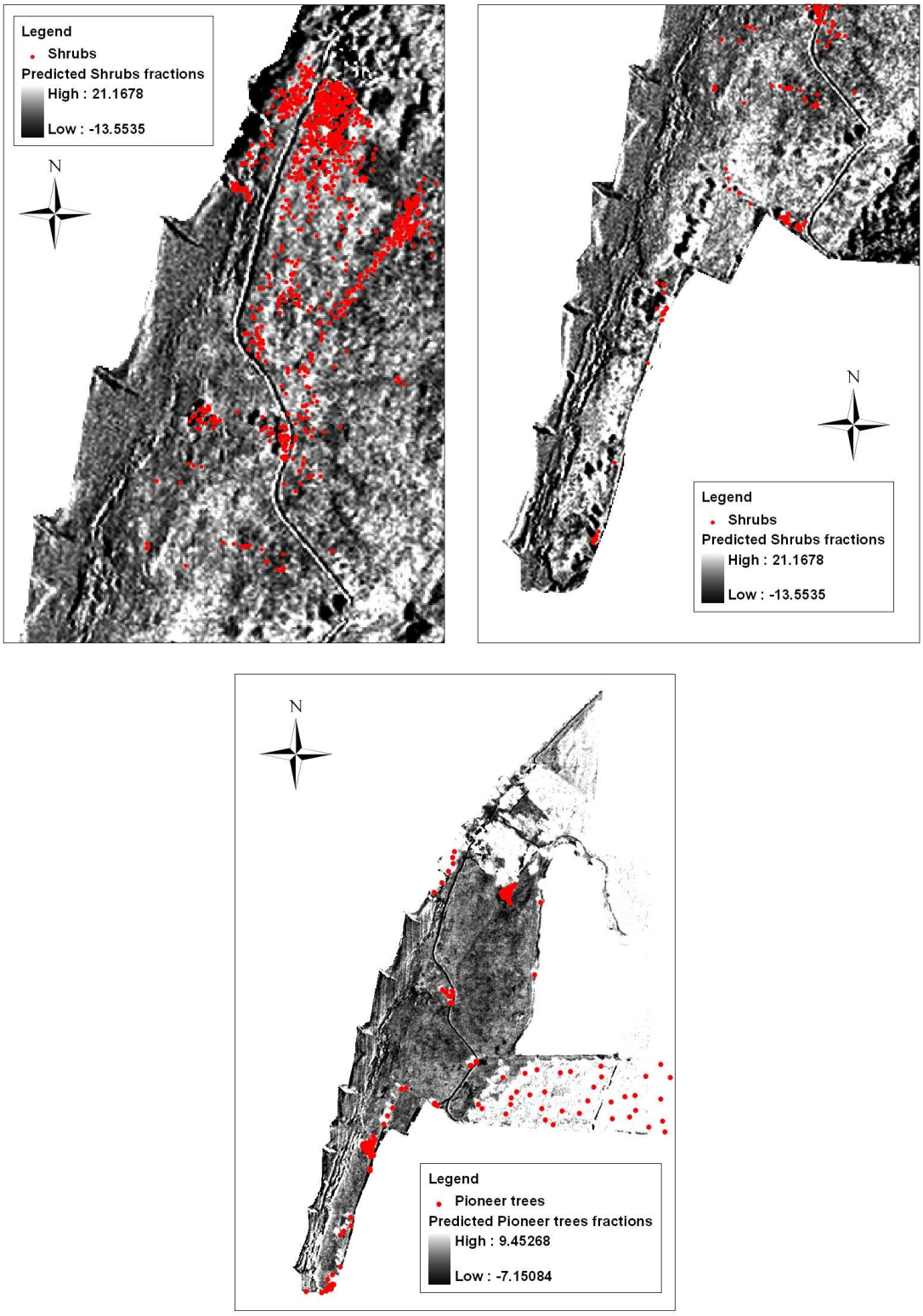


Figure 87: CASI abundance maps of Shrubs and Pioneer trees overlaid with individual plants from each class respectively, field data from 2005-06

APPENDIX 11: Temporal analysis. OSA (Overall Sub-pixel Accuracy) spatial distribution

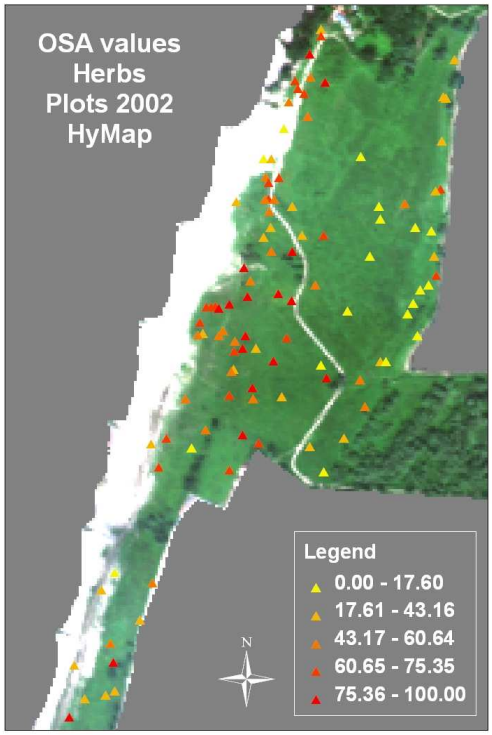


Figure (a)

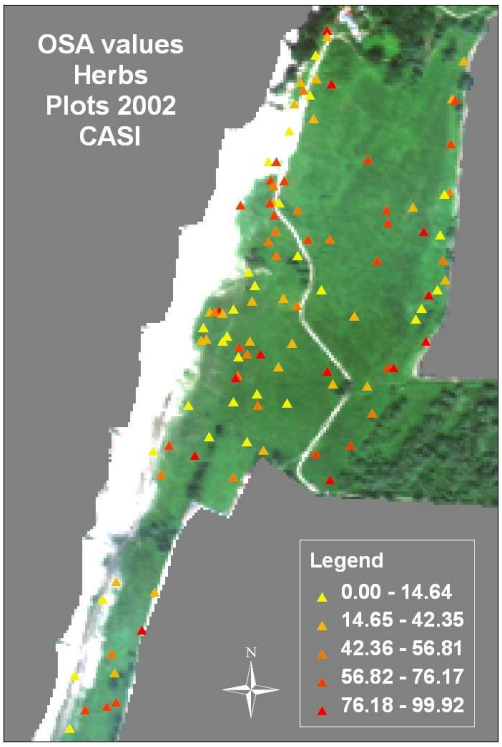


Figure (b)

Figure 88: Overall sub-pixel accuracy spatial distribution for herbs. Results from fully constrained unmixing HyMap (a) and CASI (b)

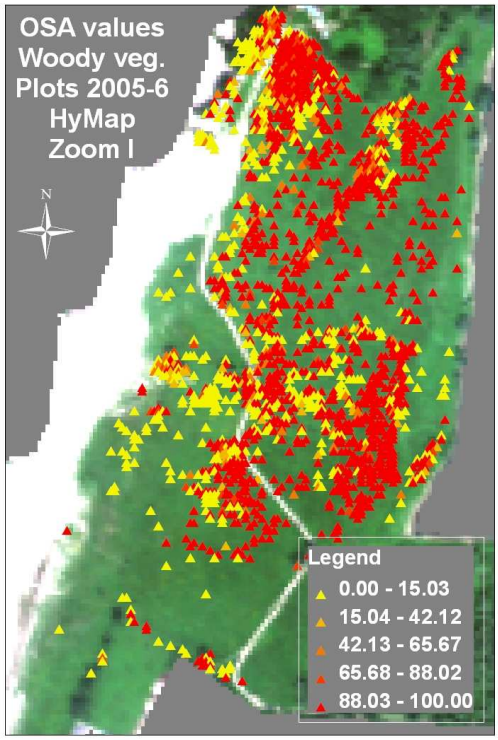


Figure (a)

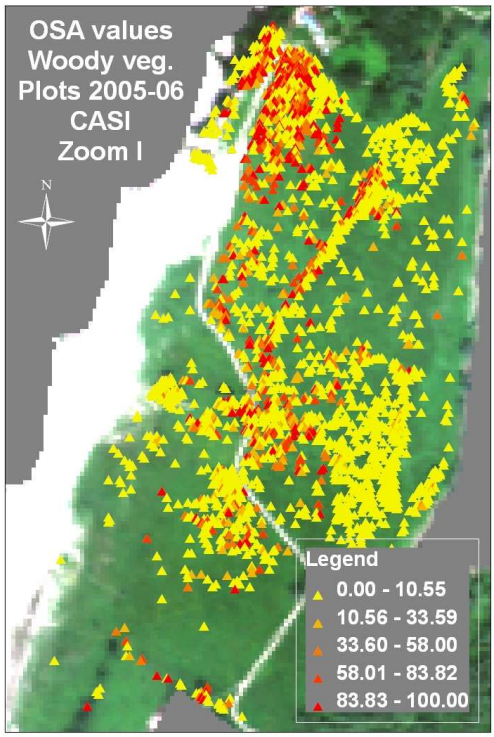


Figure (b)

Figure 89: Overall sub-pixel accuracy spatial distribution for woody vegetation I. Results from fully constrained unmixing HyMap (a) and CASI (b)

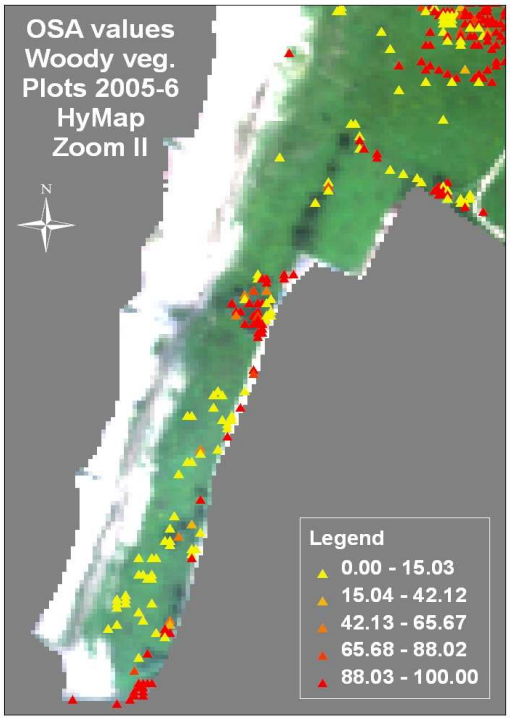


Figure (a)

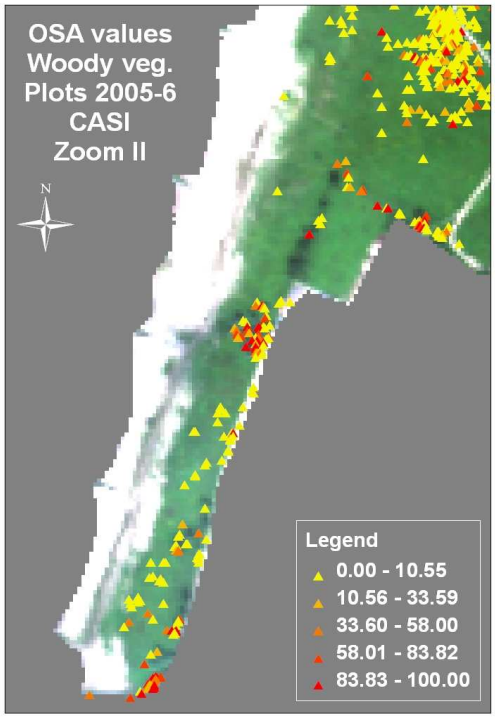


Figure (b)

Figure 90: Overall sub-pixel accuracy spatial distribution for woody vegetation II. Results from fully constrained unmixing HyMap (a) and CASI (b)