Hyperspectral remote sensing of biochemical and biophysical parameters:

The derivative red-edge “double-peak feature”, a nuisance or an opportunity?

Moses Azong Cho

Promotor:

Prof. Dr. A.K. Skidmore
Professor of Vegetation and Agricultural Land Use Survey
International Institute for Geoinformation Science and Earth Observation (ITC), Enschede and Wageningen University, The Netherlands

Co-Promotor:

Dr. S.E. van Wieren
Assistant professor, Resource Ecology Group
Wageningen University, The Netherlands

Examining Committee:

Prof. Dr. M.E. Schaepman
Wageningen University, The Netherlands

Prof. Dr. S.M. de Jong
Utrecht University, The Netherlands

Prof. Dr. C.M. Lambi
University of Buea, Cameroon

Prof. Dr. F.D. van der Meer
Utrecht University and ITC, Enschede
The Netherlands

This research is carried out within the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) in Wageningen University, the Netherlands

Hyperspectral remote sensing of biochemical and biophysical parameters:

The derivative red-edge "double-peak feature", a nuisance or an opportunity?

Moses Azong Cho

Thesis

To fulfil the requirements for the degree of Doctor
on the authority of the Rector Magnificus of Wageningen University
Prof. Dr. M.J. Kropff
to be publicly defended on Thursday 12th of April, 2007 at 15:00 hrs
in the auditorium at ITC, Enschede, The Netherlands

Hyperspectral remote sensing of biochemical and biophysical parameters

ISBN: 978-90-8504-622-6
International Institute for Geo-information Science & Earth Observation,
Enschede, the Netherlands (ITC)
ITC Dissertation Number: 142

Table of contents

Abstract

Improved quantification and monitoring of biophysical and biochemical attributes is required to predict the response of ecosystems to climate change and acquire deeper understanding of the carbon cycle. Remote sensing is widely viewed as a time- and cost-efficient way to proceed with large-scale monitoring of vegetation parameters. For over thirty years, use has been made of broadband sensors such as Landsat TM/ETM+. The advent of hyperspectral remote sensing or imaging spectrometry enlarged the number of available bands within the visible, near-infrared (NIR) and shortwave infrared (SWIR). Hyperspectrally detectable variables associated with leaf chlorophyll content, phenological state and vegetation stress such as the spectral shift of the red-edge (670-780 nm) slope and its inflection point termed the red-edge position (REP), are not accessible with broadband sensors. State of the art indices and analytical techniques applied for broad-band remote sensing are not always suitable for information extraction from high dimensional hyperspectral data. This study aimed to develop new hyperspectral indices and propose innovative ways for empirically estimating biochemical and biophysical parameters from hyperspectral data.

The red edge position is estimated using the first derivative of the spectral curve. Existing curve fitting approaches localise the REP while assuming a derivative curve with a single peak. The proposed linear extrapolation method localises the red edge position while explicitly considering two peaks in the derivative curve. The major contribution of this study is that the linear extrapolation method allows optimised estimates of leaf chlorophyll or nitrogen content while minimising the confounding effects of background and the structure of leaves and canopy. By minimising these canopy effects, the linear extrapolation may be useful for detecting early physiological stresses associated with changes in leaf chlorophyll/nitrogen levels. The linear extrapolation method also shows high potential for discriminating tree and shrub species at both the leaf and canopy scales. Lastly, it could be used as a more stable predictor for monitoring green grass biomass in the Majella National park, Italy compared with two-band vegetation indices. The method is simple to implement, but sensitive to spectral noise. Spectral smoothing is recommended when noise is a problem.

The study also highlights the utility of partial least squares (PLS) regression based on airborne hyperspectral imagery (HyMap) for estimating grass biomass and beech (*Fagus sylvatica* L.) forest mean diameter-at-breast height (DBH) in the Majella National Park, Italy. PLS regression produced lower prediction errors for grass biomass and beech forest mean DBH compared with univariate regression involving vegetation indices such as NDVI. NDVI may be simple to implement but could be lacking in terms of exploiting the information content inherent in several narrow bands.

In a nutshell, this study makes a contribution in the domain of information extraction from hyperspectral data for estimating vegetation parameters such as leaf chlorophyll/nitrogen concentration, grass biomass and forest structural parameter using empirical models. Other studies are focused on developing physically based methods given the lack of robustness and portability of empirical models for varying environmental conditions. However, empirical models that are less sensitive to environmental conditions such as models based on the linear extrapolation REP could be used to support the development of physically based models, particularly to estimate the values of the model parameters, or to refine the underlying concepts on which the model is constructed. The future of hyperspectral remote sensing could hinge on enhancing the link between empirical and physically based approaches.

Samenvatting

Verbetering van de kwantificering en monitoring van biofysische en biochemische attributen is noodzakelijk voor het voorspellen van de invloed van klimaatsverandering op ecosystemen en het verkrijgen van een beter begrip in de koolstof cyclus. Aardobservatie door middel van remote sensing wordt algemeen beschouwd als de meest tijd- en kost-efficiënte wijze voor grootschalige monitoring van vegetatie parameters. Traditioneel wordt hiervoor gebruik gemaakt van breed band sensoren zoals Landsat TM/ETM+. De komst van hyperspectrale remote sensing of beeldvormende spectrometrie vergrootte het aantal beschikbare banden in het zichtbare, nabije en korte golf infrarood bereik. Hyperspectraal detecteerbare variabelen geassocieerd met het chlorophyl gehalte, de fenologie en stress van bladeren zoals de verschuiving van de hellingshoek van de spectrale curve op de rand van rood en infrarood (de zogenaamde red edge, gelegen tussen 670-780 nm) en het bij deze curve behorend inflectie punt (red edge position, REP), zijn niet waarneembaar met breed band sensoren. Gangbare indices en analyse technieken gebruikt voor breed band remote sensing zijn niet altijd geschikt voor het extraheren van informatie uit hoog dimesionele hyperspectrale data. Deze studie beoogt de ontwikkeling van nieuwe hyperspectrale indices en stelt innovatieve wegen voor voor het empirisch schatten van biochemische en biophysische parameters uit hyperspectrale data.

De eerste afgeleide van de spectrale curve wordt gebruikt voor het schatten van de red edge positie. Bestaande curve fitting methodes veronderstellen voor het bepalen van de red edge positie een enkele piek in deze eerste afgeleide. De ontwikkelde lineaire extrapolatie methode maakt het mogelijk de positie van de red edge te schatten in geval van een dubbele piek in de eerste afgeleide. De belangrijkste bijdrage van deze studie is dat de lineaire extrapolatie methode een optimale schatting van het cholorophyl en stikstof gehalte van bladeren mogelijk maakt door minimalisatie van verstoring door de achtergrond en de structuur van blad en bladerdek. Door het minimaliseren van de effecten van blad en bladerdek zou de lineaire extrapolatie techniek nuttig kunnen zijn voor het in een vroeg stadium detecteren van fysiologische stress geassocieerd met veranderingen in het gehalte van chlorophyll en stikstof in bladeren. De lineaire extrapolatie techniek bleek ook grote potentie te hebben voor het onderscheiden van boom- en stuiksoorten, zowel op het niveau van blad en bladerdek. Bij toepassing in Majella Nationaal park in Italië

bleek het ook een stabielere voorspeller voor het monitoren van de biomassa van groen gras te zijn dan een twee band vegetatie index. De methode is simpel toe te passen, maar gevoelig voor spectrale ruis. Gladstrijken van spectra wordt aanbevolen waar ruis een probleem is.

De studie toonde ook het nut aan van partiële kleinste-kwadraten regressie voor het schatten van de biomassa van gras en de diameter op borst hoogte (DBH) in beukenbos, in een studie in Majella Nationaal Park waarbij gebruik werd gemaakt van airborne hyperstrale beelden (HyMap). Partiële kleinste-kwadraten regressie reduceerde de voorspellingsfouten voor gras biomassa en de DBH in beukenbos ten opzichte van de fouten verkregen met univariate regressie technieken die gebruik maakten van vegetatie indices zoals NDVI. NDVI mag dan wel simpel toepasbaar zijn, maar kan tekort schieten in het gebruik maken van het rijkere informatie gehalte inherent aan meerdere nauwe banden.

Kort samengevat draagt deze studie bij aan het domein van de informatie extractie uit hyperspectrale data voor empirische schatting van vegetatie karakteristieken zoals chlorophyll en stikstof gehalte van blad, gras biomassa en de structuur van bos. Andere studies concentreren zich, om tegemoet te komen aan het door variatie in milieu omstandigheden veroorzaakte gebrekkige robustheid en overdraagbaarheid van empirische modellen, op het ontwikkelen van generieke op fysische processen gebaseerde modellen. Empirische modellen die minder gevoelig zijn voor milieu omstandigheden, zoals de lineaire extrapolatie methode, zouden kunnen worden gebruikt voor het ondersteunen van het verder ontwikkelen van fysisch gebaseerde modellen, in het bijzonder voor het inschatten van model parameters of het verfijnen van de concepten op basis waarvan deze modellen worden geconstrueerd. De toekomst van hyperspectrale aardobservatie ligt wellicht in een sterkere link tussen empirische en fysisch gebaseerde benaderingen.

To my parents John and Christina Azong

Acknowledgements

Inspired by the luxuriant Bamenda montane forest with its dripping mosses and lichens, flashing springs and waterfalls, and harmonious blend of bird songs I dreamt, all the way, all along. It certainly feels like a dream come true at this point in time. I am all full of admiration and gratitude to Prof. Andrew Skidmore, my promotor for deciphering my burning desire to be part of the scientific community and let alone for playing a central role in its fulfilment. Together with my co-promotor, Dr. Sip van Wieren, your scientific guidance and moral support were simply impeccable.

This research would not have been undertaken without the financial support of ITC, to whom I am deeply indebted. I am equally humbled by the assistance and support I got from the research coordination office, especially from Prof. Martin Hale and Ms Loes Colenbrander. My stay at ITC was also facilitated by the dynamic support of the Education Affairs Department, with special thanks to Theresa van den Boogaard, Marie-Chantal Metz and Bettine Geerdink. To Esther Hondebrink and Ard Kosters, I say thanks for your wonderful support.

I would like to thank other members of staff in the Department of Natural Resources at ITC who helped me in one way or the other. My special thanks go to Dr. Fabio Corsi, Dr. Jan de Leeuw and Dr. Clement Atzberger who guided me through some of the scientific papers that make up this thesis. To Dr. Jan de Leeuw, I say thank you for putting me through some of the statistical methods and for translating the abstract into Dutch. Dr. Iris van Duren, Dr. Bert Toxopeus, Prof. Alfred de Gier and Henk van Oosten also helped me with useful ideas. Prof. Herbert Prins and Dr. Jan Clevers of Wageningen University were equally very instrumental. The support from the staff of UNIFARM experimental station and the Laboratory of the Department of Tropical ecology at Wageningen University, notably Andre Maasen, Bertus and Jan Walsem cannot be underestimated. My gratitude also go to the management of the Majella National Park, Italy, and particularly to Dr. Theodoro Andrisano.

The support of fellow PhD students was very important. I certainly cannot forget the live discussion during the end-of-month PhD student sittings, tea breaks, Wednesday soccer rendez-vous and particularly the forth-nightly discussion among PhD students in the Department of

Natural resources. To begin with, my very special thanks go to Istiak Sobhan, his wife and daughter. You have been more than a brother and family to me. I would also like to record my particular thanks to Dr. Martin Yemefack, Dr. Onisimo Mutanga, Dr. Amon Murwira, Dr. Jelle Ferwerda, Dr. Alfred Duker, Dr. Chaichoke Vaiphasa, Dr. Etien Koua, Dr. Pravesh Debba, Dr. Javier Morales, Peter Minang, Marleen Noomen, Jane Bemigisha, Roshanak Darvishzadeh, Richard Onchaga, Ulanbek Turdukulov, Pieter Beck, Jamshid Farifteh, Emmanuel Owusu, Diana Chavarro, Tyas Basuki, Obolokile Obakeng, Kate Lance, Mhosisi Masocha, Nicky Knox, Tiejun Wang, Claudia Pittiglio, Guofeng Wu, Farhang Sargordi, Trias Aditya, Prasun Gangopadhyay, Graciela Peters Guarin, Veronica Botero and Enrique Castellanos

My sincere thanks go to former teachers, colleagues and friends, the late Paul Muma, Henry Wanye, Bauke de Vries, Adri Mulder, Karsten Feuerriegel, Konneh Ursula, Ian Chebe, George and Bridget Shungafor, Helen Eyong, Vitalis Fru, Ernest Ngwana, Lawrence and Evelyn Moto, Yemefack Berthe, Lawrence Nsoyuni, Forbacha Charles, Chi Napoleon, Delphine Ebanga and Emmanuel Kejem. Dutch families and friends, Mrs Eva Skidmore, Pauline van Rijckevorsel and Rocco, Titia and Marion Tank, and the Cameroonian community in Enschede, namely, Alfred, Courage, Azise, Boris, Sylvester, Mierelle, Stanley Acham and Madam Pauline.

Finally, it would have been impossible without the devout support of my entire family. The love and practical support I got and continue to enjoy from my wife Matilda Azong have always given me reason to strive for higher heights. It would difficult to ever understand and even to find words to describe the sacrifices my beloved children; Mokom, Kuko and Fri-Ngwe have had to endure in our absence. I am full of gratitude to Mr/Mrs David Muchoh, Mr/Mrs Peter Acham, Mr/Mrs Joseph Ngam, Mrs Naomi Tamufor, Mrs Elisabeth Acham; the Azongs, namely, Ni Azaah, Tifu, Mokom, Kyen, Tembei, Achiri, Bei, Honorine and Anyim; Dr Divine Muchoh, Gerald Fru, Theresse Ngam, Dr Irene Asanga, Fri Asanga, Forkum Augustin and Sama Forkum.

Chapter 1

General introduction

1.1 Remote sensing of biochemical and biophysical parameters

Ecological studies require quantification and monitoring of biochemical and biophysical attributes of ecosystems (Asner, 1998). Estimates of foliar biochemicals such as the levels of chlorophyll and nitrogen provide us with indicators of plant productivity, stress and the availability of nutrients (Knipling, 1970; Curran, 1989). Meanwhile, biophysical parameters such as leaf area index (LAI) and biomass are important for quantifying primary production or carbon cycle of terrestrial ecosystems (Mooney, 1986; Bonan, 1993). Direct field techniques for estimating these vegetation attributes require frequent destructive harvesting (Gower et al., 1999). Such techniques are difficult, extremely labour intensive, and costly in terms of time and money. They can hardly be extended to cover large areas (Scurlock and Prince, 1993). However, estimates of biochemical and biophysical parameters over large areas may be obtained using remote sensing data acquired from air or space platforms. Remote sensing techniques have been most successful for quantifying and monitoring biophysical parameters (Daughtry et al., 1992; Price, 1992; Clevers, 1997; Wylie et al., 2002; Cohen et al., 2003; Colombo et al., 2003). There are fewer applications of air or spaceborne remote sensing for estimating biochemical properties (Wessman et al., 1989; Johnson et al., 1994; Boegh et al., 2002), primarily because of the low spectral resolution of most existing sensors.

1.2 Conventional remote sensing of vegetation parameters

Most of the remote sensing products for quantifying biochemical and biophysical parameters are derived from broadband sensors such as NOAA advanced very high resolution radiometer (AVHRR), SPOT and Landsat TM/ETM+ with three to seven spectral bands (Richardson et al., 1983; Wiegand et al., 1991; Anderson et al., 1993; Duchemin, 1999; Van Wagtendonk and Root, 2003). Most studies have focused on developing empirical relationships between ground-measured vegetation parameters and spectral indices commonly known as vegetation indices. Vegetation indices are mathematical transformations of vegetation reflectance into dimensionless measures that function as predictors of vegetation parameters. The most known and widely used vegetation index is the

normalised difference vegetation index (NDVI) developed by Rouse et al. (1974). It is based on the contrast between the maximum absorption in the red due to chlorophyll pigments and the maximum reflectance in the near infrared (NIR) caused by leaf cellular structure (Eq. 1.1.).

$$NDVI = (NIR-red) / (NIR + red) \quad (1.1)$$

However, there are major limitations with the NDVI despite its wide application in ecological remote sensing. Several studies show that broadband NDVI can be unstable, varying with soil colour, canopy structure, leaf optical properties and atmospheric conditions (Huete and Jackson, 1988, Middleton, 1991, Kaufman and Tanré, 1992; Qi et al., 1995, Todd et al., 1998). Furthermore, broadband NDVIs asymptotically approach a saturation level after a certain biomass or LAI (Sellers, 1985, Gao et al., 2000). Other variants of NDVI such as the soil adjusted vegetation index (SAVI) and atmospherically resistant vegetation index (ARVI) have been developed to correct for soil and atmospheric effects. Despite the developments of the improved variants of NDVI, it has been demonstrated that empirical models derived from vegetation indices are highly site and sensor specific and therefore unsuitable for application to large areas or in different seasons (Curran, 1994; Gobron et al., 1997).

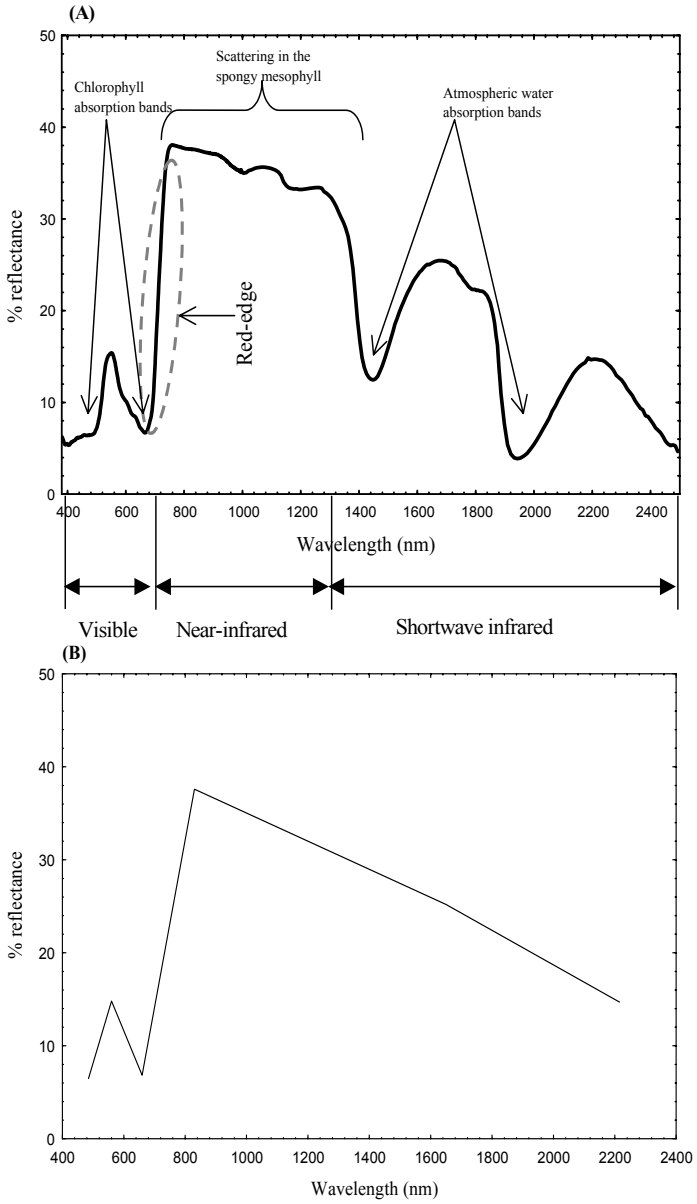


Fig. 1.1 (A) Contiguous spectrum of healthy green vegetation using handheld GER 3700 spectrometer and (B) the same spectrum re-sampled to 6 bands of Landsat TM band-setting.

1.3 Hyperspectral remote sensing of vegetation parameters

A major limitation of broadband remote sensing products is that they use average spectral information over broadband widths resulting in loss of critical information available in specific narrowbands (Blackburn, 1998, Thenkabail *et al.* 2000). Recent developments in hyperspectral remote sensing or imaging spectrometry have provided additional bands within the visible, NIR and shortwave infrared (SWIR) (Fig.1.1.). Most hyperspectral sensors acquire radiance information in less than 10 nm bandwidths from the visible to the SWIR (400-2500 nm) (Asner, 1998). For example, the spectral shift of the red-edge (670-780 nm) slope associated with leaf chlorophyll content, phenological state and vegetation stress, is not accessible with broadband sensors (Collins, 1978; Horler, et. al., 1983).

Empirical techniques continue to dominate hyperspectral remote sensing studies of vegetation parameters. Two main empirical techniques are investigated with hyperspectral data; namely, univariate regression involving hyperspectral or narrowband indices, and multivariate techniques based on more than two wavebands. Recent applications of hyperspectral remote sensing show improvements not only in the estimation of biophysical parameters (Broge and Leblanc, 2000; Mutanga and Skidmore, 2004; Thenkabail et al., 2004) but also in the estimation specific biochemical compounds such as leaf pigments (e.g. chlorophyll) and nutrient (e.g. nitrogen) concentrations with spectral information of specific band(s) and narrowband vegetation indices (Blackburn, 1998; Filella and Peñuelas, 1994; Yoder and Pettigrew-Crosby, 1995).

Since statistical techniques lack robustness and portability, some studies have focused on the development and application of deductive or physical-based approaches involving leaf and canopy radiative transfer models (Jacquemoud and Baret, 1990; Kuusk, 1991; Verstraete et al., 1996). However, these models are computationally demanding and require a large number of leaf and canopy variables, which are often difficult to estimate (Fang et al., 2003). The focus of this study shall be on empirical methods for estimating vegetation parameters.

1.3.1 The red-edge position (REP)

Prominent among new hyperspectral remote sensing products is the wavelength of maximum slope in the red-NIR transition or red-edge (670-780 nm). This wavelength point is known as the red-edge position (REP) (Horler et al., 1983; Clevers et al., 2002). Increases in the amount of chlorophyll causes a broadening of the major chlorophyll absorption feature centred around 680 nm (Buschmann and Nagel, 1993; Dawson and Curran, 1998), causing a shift in the red edge slope and REP towards longer wavelengths, i.e. the "red shift" (Gates et al., 1965; Collins et al., 1977; Horler et al., 1980; Horler et al., 1983; Hare et al., 1984; Boochs et al., 1990; Clevers et al., 2002). Low leaf chlorophyll concentrations cause shifts of the red-edge slope and REP towards the shorter wavelengths i.e. the "blue shift". These characteristic shifts in the REP have been used as a means to estimate foliar chlorophyll concentration/content and also as an indicator of vegetation stress (Chang and Collins, 1983; Horler et al., 1983; Curran et al., 1995; Clevers et al., 2002; Lamb et al., 2002; Smith et al., 2004). An advantage of the REP over the NDVI is that it is less sensitive to varying soil and atmospheric conditions, and sensor view angle (Curran et al. 1995, Blackburn and Pitman, 1999, Clevers et al. 2001).

However, one critical problem has limited the application of the REP for estimating leaf chlorophyll concentration or its correlates e.g. leaf nitrogen concentration. Derivative analysis of contiguous spectra, the simplest method for locating the REP usually reveals the presence of two dominant peaks in the red-edge around 700 and 725 nm causing a bimodal distribution of REP data around these peaks and a discontinuity in the REP/chlorophyll relationship (Horler et al. 1983). Experimental studies show that low leaf chlorophyll concentration is associated with REP values near 700 nm, while high chlorophyll concentration in combination with leaf internal scattering influence REP values near 725 nm (Horler et al., 1980; Boochs et al., 1990; Lamb et al., 2002). Zarco-Tejada et al. (2003) demonstrated in an experimental study that the existence of the double peak is due to chlorophyll fluorescence emission at about 690 nm and 730 nm. Earlier studies had shown that the chlorophyll fluorescence ratio (F690/F730) decreases with increasing chlorophyll content of developing leaves (Hák et al., 1990; Babani et al., 1996) and increases during autumnal chlorophyll breakdown of various tree leaves (D'Ambrosio et al., 1992).

Several other techniques for locating the REP have been developed. These include model-fitting techniques such as the simple linear four-point interpolation method (Guyot and Baret, 1988) and computational complex procedures including fitting a high-order polynomial function (Pu et al., 2003) or an inverted Gaussian function (Bonham-Carter, 1988) to the reflectance spectrum. The above techniques do mitigate the discontinuity in the REP data caused by the double-peak feature (Clevers et al., 2002), but do not particularly target changes around the dominant chlorophyll peaks (around 700 and 725 nm). Thus, could accurate estimates of chlorophyll or nitrogen concentration be obtained with REP derived by tracking changes near the dominant red-edge peaks?

1.3.2 Multivariate regression techniques

Most vegetation indices are computed from two bands. One of the disadvantages of this method is that it utilises a limited amount of the total spectral information available in an image (Foody et al., 2003). Fewer studies have evaluated statistical techniques that integrate information from several spectral bands. A commonly used technique that involves several hyperspectral bands is multiple stepwise regression (Kokaly and Clark, 1999; Curran et al., 2001; De Jong et al., 2003). However, Curran (1989) and De Jong et al. (2003) point out that this method might be affected by multicollinearity among hyperspectral bands. Increasingly, remote sensing studies of vegetation are making use of multivariate techniques such artificial neural networks (ANN) (Miller et al., 1995; Skidmore et al., 1997: Mutanga et al. 2004) and partial least squares regression (Hansen and Schjoerring, 2003; Huang et al., 2004) to avoid collinearity and having to deal with a large number of predictor variables. But their applications to various ecosystems remain limited. Thus, could the use of multiple hyperspectral bands improve the estimation of biophysical parameters compared to hyperspectral indices?

1.4 Research Objectives

(i) To develop a technique for locating the REP that mitigates the destabilising effect of the double-peak feature on the REP data and tracks changes around the dominant chlorophyll sensitive peaks (700 and 725 nm)?

(ii) To test the performance of the new method vis-à-vis other REP techniques for estimating nitrogen concentrations, estimating

foliar chlorophyll for a wide range of canopy and environmental conditions, discriminating species at leaf and canopy scales and estimating grass biomass and forest structural attributes.

(iii) To investigate the performance of empirical techniques based on multiple hyperspectral bands for estimating grass biomass and forest structural attributes.

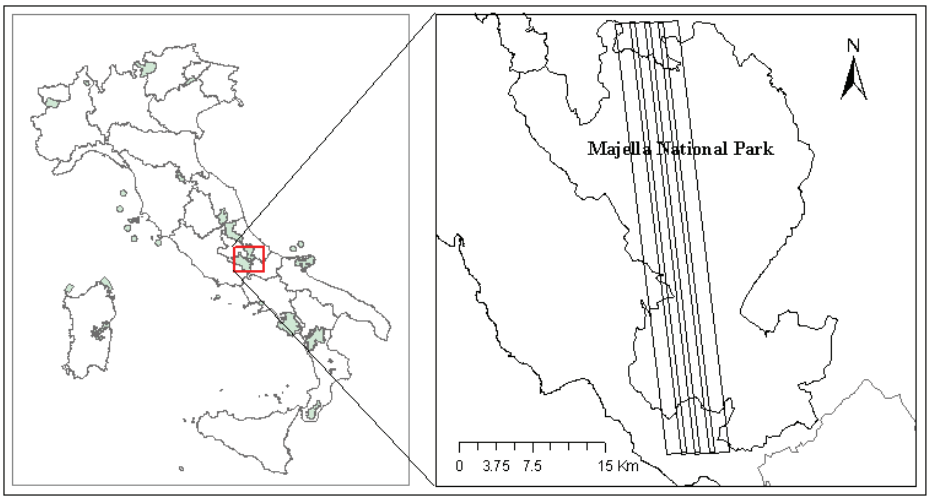


Fig.1.2 Location of the Majella National Park, Italy and HyMap flight lines.

1.6 General methods

The investigations were conducted using leaf and canopy spectra of plants that were grown in the greenhouse, data simulated from leaf and canopy reflectance models and field data from the Majella National Park, Italy (Fig.1.2). Different remote sensing devices were adopted. These include field spectrometers and an airborne hyperspectral sensor (i.e. HyMap).

The Majella National Park, Italy (latitude 41°52' to 42°14'N, longitude 13°50' to 13°14'E) covers an area of 74095 ha. The park extends into the southern part of Abruzzo, at a distance of 40 km from the Adriatic Sea. This region is situated in the massifs of the Apennines (Conti, 1998). The park is characterised by several mountain peaks, the highest being Mount Amaro (2794 m). More specifically, the study site (latitude 41°49' to 42°14'N, longitude 13°57' to 14°6'E) is situated between Mounts Majella and Morrone to the east and west, respectively. It covers an area of

40 km by 5.5 km. Gallego Fernández *et al*. (2004) argue that plant community dynamics in Mediterranean basin ecosystems are driven mainly by alternating episodes of human intervention and land abandonment. For example, abandoned settlement and agricultural areas in Majella are returning to oak (*Quercus pubescens*) woodlands at the lower altitude (400 m to 600 m) and beech (*Fagus sylvatica*) forest at the higher altitude (1200 m to 1800 m). Between these two formations is a landscape composed of shrubby bushes, patches of grass/herb vegetation, and bare rock outcrops.

1.7 Thesis outline

The main body of this thesis starts with chapter two, entirely focused on the development of a new REP technique and its application for estimating leaf nitrogen concentration. Leaf and canopy spectra of plants grown in the greenhouse and the spectra of leaf stacks consisting of mixed grass/herb species collected from the Majella National Park have been used to test the new REP model.

In the subsequent chapters (3 to 5), the performance of the new REP technique is compared with that of conventional techniques for (i) estimating leaf chlorophyll content under a wide range of perturbing factors such as varying soil background, leaf area and mass and sun view angles using simulated data from leaf and canopy reflectance models, (ii) discriminating shrub and tree species at the leaf and canopy scales, (iii) estimating green grass biomass in the Majella Park for two consecutive years (summers of 2004 and 2005) using airborne HyMap images.

Chapters 6 and 7 are focused on testing other empirical techniques based on multiple HyMap bands such as partial least squares regression and artificial neural networks for estimating green grass biomass and beech forest structural parameters.

Chapter 2

A new technique for extracting the red-edge position from hyperspectral data: The linear extrapolation method

This chapter is based on
Cho, M.A. and Skidmore, A.K. (2006), A new technique for extracting the red edge position from hyperspectral data: the linear extrapolation method. Remote sensing of Environment, 101:181-193.

Abstract

The position of the inflexion point in the red edge region (680 nm to 780 nm) of the spectral reflectance signature, termed the red edge position (REP), is affected by biochemical and biophysical parameters and has been used as a means to estimate foliar chlorophyll or nitrogen content. In this paper, we report on a new technique for extracting the REP from hyperspectral data that aims to mitigate the discontinuity in the relationship between the REP and the nitrogen content caused by the existence of a double-peak feature on the derivative spectrum. It is based on a linear extrapolation of straight lines on the far-red (680 nm to 700 nm) and NIR (725 nm to 760 nm) flanks of the first derivative reflectance spectrum. The REP is then defined by the wavelength value at the intersection of the two lines. The output is a REP equation, $REP = - (c_1 - c_2) / (m_1 - m_2)$, where c_1 and c_2, and m_1 and m_2 represent the intercepts and slopes of the far-red and NIR lines, respectively. Far-red wavebands at 680 nm and 694 nm in combination with NIR wavebands at 732 nm and 760 nm or at 724 nm and 760 nm were identified as the optimal combinations for calculating nitrogen-sensitive REPs for three spectral data sets (rye canopy, and maize leaf and mixed grass/herb leaf stack spectra). REPs extracted using this new technique (linear extrapolation method) showed high correlations with a wide range of foliar nitrogen concentrations for both narrow and wider bandwidth spectra, being comparable with results obtained using the traditional linear interpolation, polynomial and inverted Gaussian fitting techniques. In addition, the new technique is simple as is the case with the linear interpolation method, but performed better than the latter method in the case of maize leaves at different stages of development and mixed grass/herb leaf stacks with a low nitrogen concentration.

Keywords: hyperspectral data; red edge position; linear extrapolation technique; foliar nitrogen concentration

2.1 Introduction

The region of the red-near infrared (NIR) transition has been shown to have high information content for vegetation spectra (Collins et al., 1977; Collins, 1978; Horler et al., 1983). This region is generally referred to as the "red edge". It represents the region of abrupt change in leaf reflectance between 680 nm and 780 nm caused by the combined effects of strong chlorophyll absorption in the red wavelengths and high reflectance in the NIR wavelengths due to leaf internal scattering (Gates et al., 1965; Horler et al., 1983). Increases in the amount of chlorophyll, for example, results in a broadening of the major chlorophyll absorption feature centred around 680 nm (Buschmann and Nagel, 1993; Dawson and Curran, 1998), causing a shift in the red edge slope and wavelength of maximum slope (or inflection point) towards longer wavelengths (Gates et al., 1965; Collins et al., 1977; Horler et al., 1980; Horler et al., 1983; Hare et al., 1984; Boochs et al., 1990; Clevers et al., 2002). The latter is termed the red edge position (REP). Shifts in the REP to longer or shorter wavelengths have been used as a means to estimate changes in foliar chlorophyll content and also as an indicator of vegetation stress (Chang and Collins, 1983; Horler et al., 1983; Curran et al., 1995; Clevers et al., 2002; Lamb et al., 2002; Smith et al., 2004).

Since the REP is defined as the inflection point of the red-NIR slope, an accurate determination of the REP requires a number of spectral measurements in narrow bands in this region (Clevers et al. 2002). Fortunately, recent developments in imaging spectrometry have provided additional bands (contiguous spectra of less than 10 nm bandwidths) within the red edge region compared to broadband imagery such as Landsat Thematic Mapper (Asner, 1998). Subsequently, the REP is defined by the maximum first derivative of the reflectance spectrum. However, the limitation of this approach is that the maximum first derivatives of contiguous spectra have been well documented to occur within two principal spectral regions (around 700 and 725 nm) causing a bimodal distribution of REP data around 700 and 725 nm and a discontinuity in the REP/chlorophyll relationship (Horler et al. 1983). Several other studies have revealed the existence of this double-peak feature in the first derivative of contiguous spectra. Boochs et al. (1990) identified two peaks in winter wheat at 703 and 735 nm. Smith et al (2004) also found peaks in canopy spectra of grass near 702 and 725 nm. Clevers et al. (2004) used the Analytical Spectral Devices (ASD)

FieldSpec FR spectroradiometer with a 1 nm spectral resolution and observed two peaks in canopy spectra of grass near 700 and 720 nm. Zarco-Tejada et al. (2003) observed the double-peak feature at 690-710 nm and found out that it is a function of natural fluorescence emission at 690 and 730 nm.

Experimental studies have shown that low leaf chlorophyll concentration is associated with REP values near 700 nm, while high chlorophyll concentration in combination with leaf internal scattering influence REP values near 725 nm (Horler et al., 1980; Boochs et al., 1990; Lamb et al., 2002). Model fitting techniques such as the simple linear four-point interpolation method (Guyot and Baret, 1988) and computational complex procedures including fitting a high-order polynomial function (Pu et al., 2003) or an inverted Gaussian function (Bonham-Carter, 1988) to the reflectance spectrum somewhat generate continuous REP data (Clevers et al., 2002). In other words, these techniques mitigate the discontinuity in REP data caused by the double-peak feature. But the first question we pose is whether these techniques adequately track variations in spectral reflectance near the low and high chlorophyll sensitive peaks (near 700 and 725 nm). Dawson and Curran (1998) proposed a three-point Lagrangian interpolation approach. But Clevers et al. (2002) argue that this approach is suitable for coarsely sampled spectra and is not capable of resolving the destabilising effect of the double-peak feature when determining the REP. In this study, we hypothesise that first, the discontinuity in the REP/chlorophyll relationship caused by the existence of a double-peak feature on the derivative spectrum could be mitigated and secondly, spectral changes near the low and high chlorophyll sensitive peaks could be adequately tracked if the REP is determined as an intersection of two straight lines extrapolated through two points on the far-red (680 nm to 700 nm) and two points on the NIR (725 nm to 760 nm) flanks of the first derivative reflectance spectrum. The second research question is whether the proposed linear extrapolation method yields similar results to conventional methods such as the linear four-point interpolation, high-order polynomial and inverted Gaussian fitting techniques in explaining variations in foliar nitrogen concentration.

Plant nitrogen status is often related to chlorophyll content (Everitt et al., 1985; Boochs et al., 1990; Yoder and Pettigrew-Crosby, 1995). But such a relationship depends on the physiological status of the plant (Mooney, 1986; Boochs et al., 1990), e.g. changing from low to high positive

correlations with increasing leaf age (Wenjiang et al., 2004). Nitrogen is used for the formation of components such as chlorophyll, the carbon fixing enzyme ribulose biphosphate carboxylase (Rubisco) and inert structural components in cell tissue (Mooney, 1986; Jongschaap and Booij, 2004). Reflectance measurements in the visible and red edge wavelengths (400-700 nm) have been used to determine foliar nitrogen concentration (Thomas and Oerther, 1972; Bausch and Duke, 1996; Sullivan et al., 2004). But the results rely on the close correlation between nitrogen and chlorophyll pigments (Yoder and Pettigrew-Crosby, 1995; Hansen and Schjoerring, 2003; Haboudane et al., 2004) because pigments (chlorophyll, carotenoids and xanthophylls) predominantly determine most spectral features between 400 and 700 nm (Gates et al., 1965; Gausman, 1977; Yoder and Pettigrew-Crosby, 1995; Blackburn, 1998; Carter and Knapp, 2001; Merzlyak et al., 2003). Therefore, the estimation of foliar nitrogen concentration based on the REP as the predictor (Lamb et al., 2002; Jongschaap and Booij, 2004) indirectly depends on shifts in the REP mainly attributed to changes in chlorophyll concentration.

Thus, the objectives of this study were to (i) define a simple technique for extracting the REP from hyperspectral data in order to first, mitigate the discontinuity in the REP/nitrogen relationship caused by the double-peak feature on the first derivative spectrum and secondly, track variations in spectral reflectance near the dominant peaks (700 nm and 725 nm) and (ii) compare the performance of the REPs retrieved by the new technique with REPs located by the maximum first derivatives and model fitting techniques such as the linear four-point interpolation, high-order polynomial fitting, and inverted Gaussian fitting methods in explaining variations in foliar nitrogen concentration.

2.2 Experiments and data sets

Three spectral data sets were used in this study. First, greenhouse experiments were carried out to assess the performance of REPs extracted by the proposed linear extrapolation technique for predicting leaf nitrogen concentration with regard to spectral reflectance at (i) leaf scale for maize leaves and (ii) canopy scale for rye grass canopies. Secondly, leaf specimens were collected in the field at the Majella National Park in Italy for applying the same analysis on a stack of mixed grass/herb leaves.

2.2.1 Greenhouse experiment

Maize (*Zea mays*, L.) and rye (*Lolium perenne*, L.) were separately grown in 5-litre pots in a greenhouse chamber for ten weeks. Three soil nitrogen treatments (low, medium and high) were used in order to create different foliar nitrogen concentration levels. The low nitrogen treatment consisted of potting soil. In the medium and high nitrogen treatments, nitrogen was applied as ammonium nitrate fertiliser at the rates of 50 kg N ha $^{-1}$ and 150 kg N ha $^{-1}$, respectively. Four maize and forty rye seeds were sown in each maize and rye pot, respectively. A total of thirty pots were used for each nitrogen treatment. The experiment was carried out under a natural photoperiod (about 11 hours of daylight) with a maximum day temperature of 20ºC and night temperature of 15ºC.

2.2.2 Field experiment

Mixed grass/herb leaves were collected at thirty-four randomly chosen plots (30 m by 30 m) in the Majella National Park in Italy (latitude 41º 52'N, longitude 13º 14'E). The dominant grass species include *Brachypodium genuense*, *Briza media*, *Bromus erectus* and *Festuca sp*. Herbs include *Helichrysum italicum, Galium verum, Trifolium pratense, Plantago lanceolata, Sanguisorba officinalis* and *Ononis spinosa.* Each mixed grass/herb leaf sample consisted of leaf specimens clipped at five subplots (1m by 0.5 m) within each 30 m by 30 m plot. The clipped leaves were placed in transparent polythene sampling bags. The leaf spectra were measured within two hours of collection in the field. The data was collected in September 2004.

2.2.3 Spectral measurements

Leaf and canopy spectral measurements were made using a GER 3700 (Spectra Vista Corporation (NY) USA) spectroradiometer. The GER 3700 is a three dispersion grating spectroradiometer using Si and PbS detectors with a single field of view. The wavelength range is 350 nm to 2500 nm, with a resolution of 1.5 nm in the 350 nm to 1050 nm range, 6.2 nm in the 1050 nm to 1900 nm range, and 9.5 nm in the 1900 nm to 2500 nm range.

(a) Leaf spectral measurements – maize leaves

Leaf spectra were obtained from maize leaves of different ages. The measurements were carried out ten weeks after sowing, before tassel initiation. Leaf specimens along the maize stems (representing different leaf age groups) were collected from 10 pots per treatment for the low and medium nitrogen treatments only. The stems of the low and medium nitrogen treatments had three and four pairs of leaves, respectively. Four corresponding pairs of leaves (i.e. a set of 8 leaves of same age group) collected from the four stems in each pot were taken to represent a leaf sample. We obtained a sample size of n = 67. This was three samples short of an expected n = 70, that is n = 30 and n = 40 for the low and medium nitrogen treatments, respectively, because three pots of the low treatment had an incomplete set of the oldest leaves.

Leaf spectral measurements were made following the method described by Gong et al. (2002) and Vaiphasa et al. (2005). The spectral measurements were made in a dark laboratory room in order to ensure stable atmospheric and uniform illumination conditions (Mutanga et al., 2003; Vaiphasa et al., 2005). The leaves were clipped and laid on a black tray, to reduce the effect of background spectra on sample spectrum (Gong et al., 2002). Spectral measurements were then taken using a fibre optic with a 23º field of view. The sensor was held at about 10 cm above a single leaf blade covering an instantaneous field of view of about 4 cm on the leaf. A 50W halogen lamp positioned next to the sensor was used to illuminate the target. The radiance was converted to reflectance, using scans of a white spectralon reference panel. The spectrum of each sample was determined as the average of 16 spectral measurements (i.e. two measurements per leaf from the eight leaves per sample). The two spectral measurements for each leaf were made at one third and two thirds of the distance from the leaf collar. Two target measurements were made after measuring the reference panel.

(b) Leaf stack spectral measurements – Mixed grass/herb leaves

Spectral measurements of the mixed grass/herb leaf specimens collected from Majella National Park in Italy were made in a dark laboratory room. A stack of mixed grass/herb leaves was placed on a black try in order to control for background reflectance. The measurements were made at about 40 cm at nadir above the leaf stack using a fibre optic with a 23º

field of view, covering an instantaneous field of view of about 16 cm. A 50W halogen lamp positioned next to the sensor was used to illuminate the target. The spectrum of each mixed grass/herb sample was determined as the average of 15 spectral measurements. One target measurement was made after measuring the reference panel. The leaves were re-mixed after each measurement in order to capture the maximum variation in the leaf stack.

(c) Canopy spectral measurements – Rye grass

Canopy spectral measurements were obtained from the potted rye plants following the method described by Mutanga et al. (2003). The potted plants were transferred from the greenhouse to a laboratory room on each day of measurement. A total of 30 pots were measured on the third and fifth weeks after emergence (five pots/treatment/week). The sensor with a field of view of 3^o was mounted on a tripod and positioned 2 m directly above the pot at nadir position, covering an instantaneous field of view of about 11 cm on the target. A 50W halogen lamp, positioned next to the sensor was used to supply illumination on the target. The pots were rotated by about 36^o after every 4th measurement in order to average out differences in canopy orientation on each pot. The radiance was converted to reflectance, using scans of a white spectralon reference panel. Four target measurements were made after measuring the reference panel. The spectrum of each sample was determined as the average of 20 spectral measurements.

2.2.4 Measurement of foliar nitrogen concentration

Following spectral measurements, leaf samples were oven-dried at 70ºC for 24 hours. The dry leaf samples were milled before digestion in tubes with sulphuric acid (H_2SO_4), salicylic acid, hydrogen peroxide (H_2O_2) and selenium (Novozamsky et al., 1983). Afterwards, the nitrogen concentration was measured with a segmented flow analyser at 660 nm. Foliar nitrogen concentration was expressed as a percentage of dry matter. Table 2.1 shows the descriptive statistics of nitrogen concentration for the three data sets (rye, maize and mixed grass/herb).

Table 2.1 Statistics of nitrogen concentration (% dry matter) for rye, maize and mixed grass/herb leaves.

Data set	Mean (%)	95% CI	Minimum	Maximum	SD
Rye (n=30)	4.13	0.63	1.45	6.26	1.69
Maize (n=67)	2.25	0.24	0.59	4.28	1.01
Grass/herb (n=34)	1.23	0.06	0.95	1.69	0.17

CI = confidence interval and SD = standard deviation

2.3 Red edge position algorithms

2.3.1 Maximum first derivative spectrum

The REP is defined by the wavelength of the maximum first derivative of the reflectance spectrum in the region of the red edge. The first derivative was calculated using a first-difference transformation of the reflectance spectrum (Dawson and Curran, 1998) as follows:

$$FDR_{(\lambda i)} = (R_{\lambda(j+1)} - R_{\lambda(j)})/\Delta\,\lambda \tag{2.1}$$

where FDR is the first derivative reflectance at a wavelength i, midpoint between wavebands j and j+1, $R_{\lambda(j)}$ is the reflectance at the j waveband, $R_{\lambda(j+1)}$ is the reflectance at the j+1 waveband, and $\Delta\,\lambda$ is the difference in wavelengths between j and j+1..

2.3.2 Linear four-point interpolation technique

The linear four-point interpolation method (Guyot and Baret, 1988) assumes that the reflectance curve at the red edge can be simplified to a straight line centred near the midpoint between the reflectance in the NIR at about 780 nm and the reflectance minimum of the chlorophyll absorption feature at about 670 nm. It uses four wavebands (670, 700, 740 and 780 nm), and the REP is determined by using a two-step calculation procedure.

(i) Calculation of the reflectance at the inflexion point (Rre):

$$R_{re} = \left(R_{670} + R_{780}\right)/2 \tag{2.2}$$

where R is the reflectance.

(ii) Calculation of the red edge wavelength or red edge position (REP):

$$REP = 700 + 40\left(\frac{R_{re} - R_{700}}{R_{740} - R_{700}}\right) \tag{2.3}$$

700 and 40 are constants resulting from interpolation in the 700-740 nm interval.

2.3.3 Polynomial fitting technique

A fifth-order polynomial (Pu et al., 2003) function (Eq. 2.4) was fitted to the reflectance spectrum between the wavelengths corresponding to the minimum reflectance in the red and the maximum NIR (shoulder) reflectance.

$$R(\lambda) = a_0 + \sum_{i=1}^{5} a_i \lambda^i \tag{2.4}$$

where λ represents 76 bands of GER 3700 from 670 nm to 780 nm. Subsequently, REP is determined from the maximum first derivative spectrum. The first derivative was calculated using a first-difference transformation of the reflectance spectrum obtained from the polynomial fit.

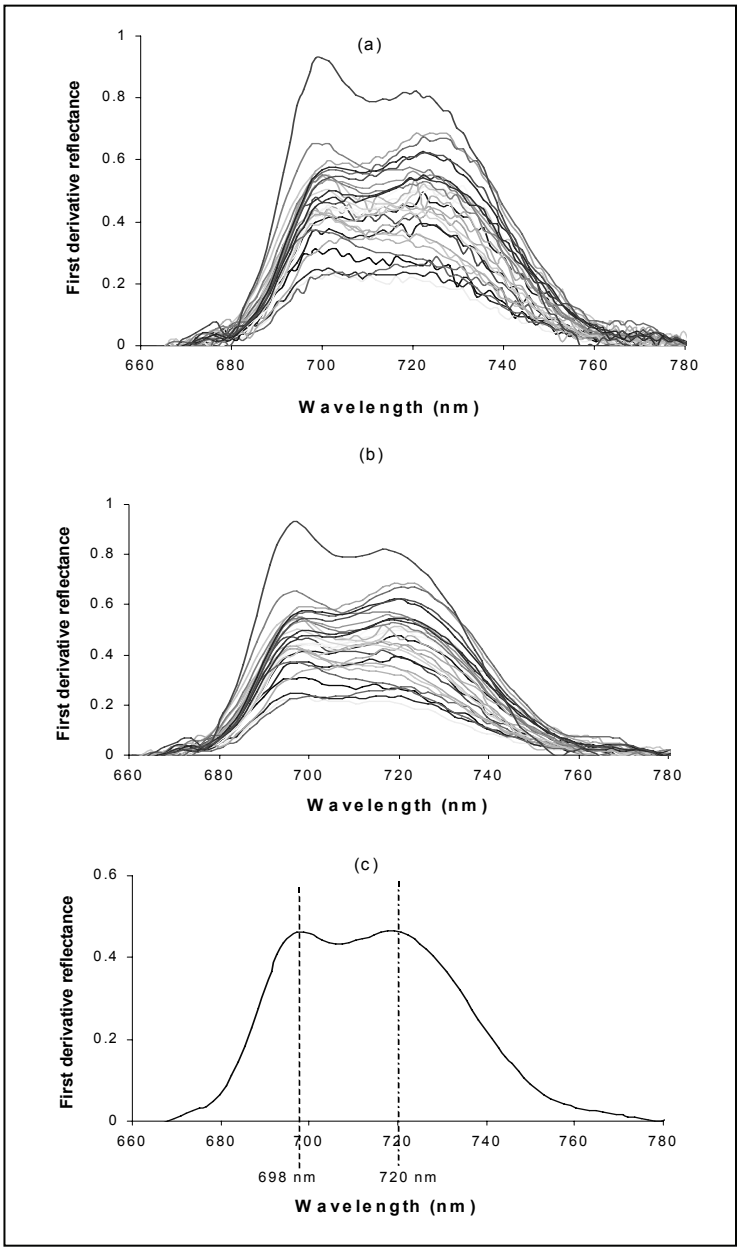


Fig.2.1 Original (a), smooth (b) and mean (c) first spectral derivative curves for rye canopies (n=30), showing the regions of occurrence of two peaks.

2.3.4 Inverted Gaussian fitting technique

An inverted Gaussian (IG) model (Bonham-Carter, 1988; Dawson and Curran, 1998; Miller et al., 1990; Pu et al., 2003) was fitted to the spectral reflectance in the 660-780 nm band range. Accordingly, the IG model (Eq. 2.5) represents the red edge by the reflectance equation:

$$R(\lambda) = R_s - (R_s - R_o)\exp\left(-\frac{(\lambda_0 - \lambda)^2}{2\sigma^2}\right) \quad (2.5)$$

where R_s is the maximum or "shoulder" spectral reflectance, R_o and λ_o are the minimum spectral reflectance and corresponding wavelength, and σ is the Gaussian function variance. The REP is then defined as:

$$REP = \lambda_0 + \sigma \quad (2.6)$$

We used an iterative optimisation fitting procedure to determine parameters of the IG model (Miller et al., 1990). Initial guesses of the model parameters were made after review of each data set. The IG model employs a least-square criterion to fit a normal curve to the reflectance red edge. Typically, λ_o and σ were set at 670 nm and 30 nm, respectively. The values of λ_o, Ro, R_s and σ are then determined by the fitting procedure.

2.3.5 Proposed technique: linear extrapolation technique

2.3.5.1 Model description

It is necessary to understand the nature of the first and second derivative spectral curves of the red edge in order to determine the regions of occurrence of the double-peak feature. We used a Savitzky-Golay second order polynomial least-squares function of five-band window (Savitzky and Golay, 1964) to smooth the first derivative spectra (Fig. 2.1). The second derivative was then calculated using a first-difference transformation of the first derivative reflectance spectrum.

It could be observed from the first derivative curves (Fig. 2.2) that the double-peak feature is located between 690 nm and 740 nm. The multiple peaks correspond to points where the second derivative curves cut the wavelength axis. For example, two maxima can be located in the rye

spectra of the medium and high nitrogen treatments at approximately 698 nm and 720 nm.

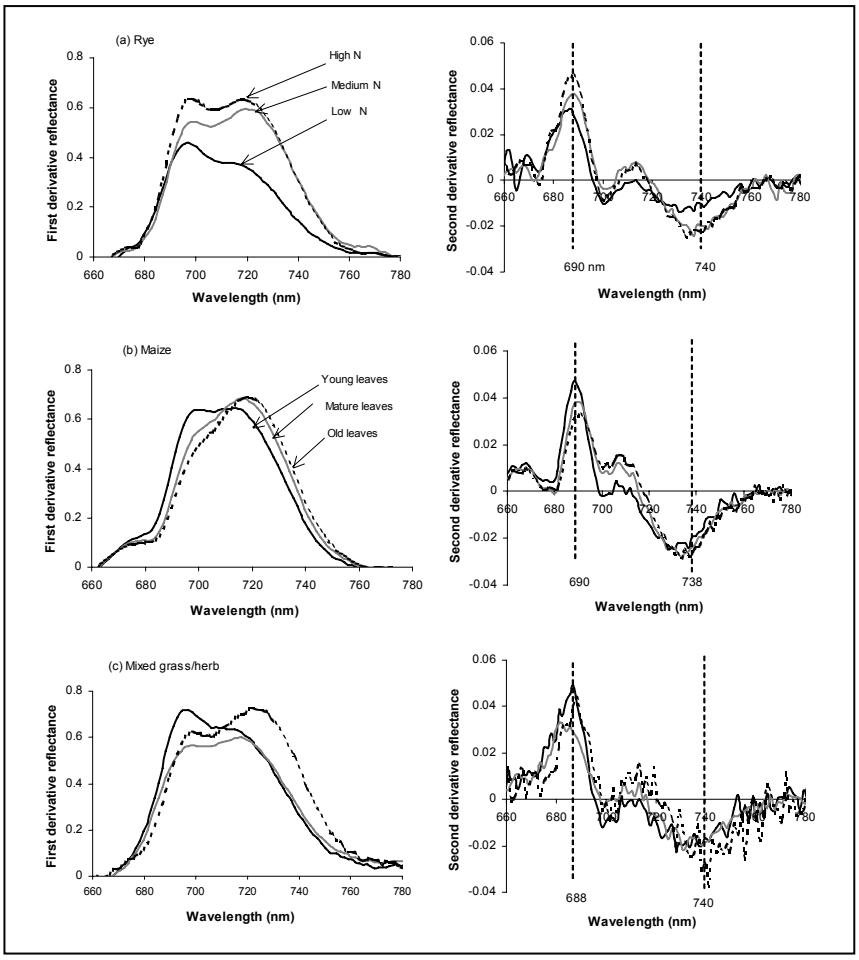


Fig.2.2 Red edge first derivative curves (left) and their corresponding second derivative curves (right) for (a) rye canopies of different nitrogen treatments, (b) maize leaves of different ages and (c) three randomly selected mixed grass/herb samples, showing the presence of one or more peaks. The spectral range delineated by dashed lines shows the region of occurrence of multiple peaks.

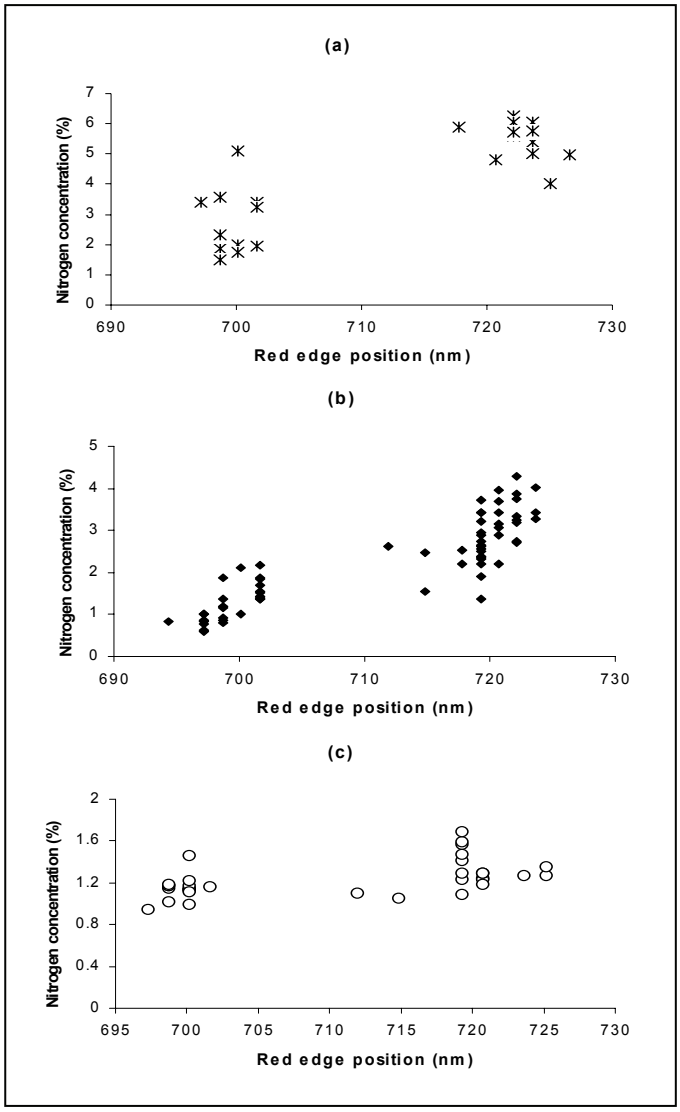


Fig.2.3 Scatter plots between nitrogen concentration and red edge position for (a) rye canopies, (b) maize leaves and (c) mixed grass/herb leaf stacks, extracted from GER spectra (1.5 nm bandwidth) showing discontinuity in the nitrogen/red edge position relationship.

A new technique for determining the REP was therefore designed based on these observations:

(i) to mitigate the destabilising effect of the double-peak on the correlation between nitrogen and REP (determined as the

maximum first derivative). This effect is revealed by a discontinuity in REP/nitrogen relationship (e.g. Fig. 2.3). Fig. 2.3 shows two groups of scatter points centred at around 700 nm and 725 nm for low and high nitrogen concentration values, respectively. Therefore, the second objective was:

(ii) to track variations near the low and high nitrogen sensitive wavebands (700 nm and 725 nm).

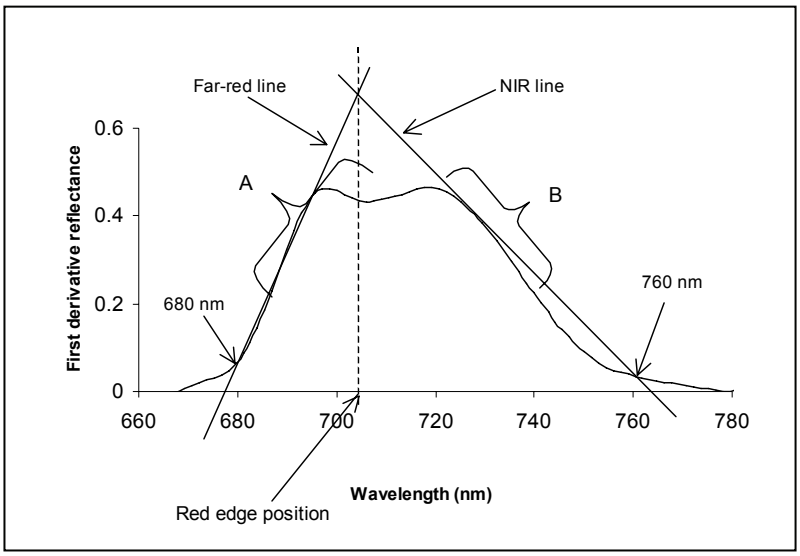


Fig.2.4 Schematic representation of the linear extrapolation technique for extracting the red edge position (REP) – wavelength of the meeting point between two straight lines extrapolated on the far-red and NIR flanks of the first derivative spectrum.

The new technique is based on linear extrapolation of two straight lines (Eqs. 2.7 and 2.8) through two points on the far-red (680 nm to 700 nm) and two points on the NIR (725 nm to 760 nm) flanks of the first derivative reflectance spectrum of the red edge region as illustrated in Fig. 2.4. The REP is then defined by the wavelength value at the intersection of the straight lines (Eq. 2.9).

$$\text{Far-red line: } FDR = m_1\lambda + c_1 \tag{2.7}$$

$$\text{NIR line: } FDR = m_2\lambda + c_2 \tag{2.8}$$

where m and c represent the slope and intercept of the straight lines. At the intersection, the two lines have equal λ (wavelength) and FDR values. Therefore, the REP, which is the λ at the intersection, is given by:

$$REP = \frac{-(c_1 - c_2)}{(m_1 - m_2)} \tag{2.9}$$

In summary, only four coordinate points (or wavebands) are required to calculate the REP by the linear extrapolation method; for instance, two bands near 680 nm and near 700 nm to calculate m_1 and c_1 for the far-red line and two bands near 725 nm and near 760 nm to calculate m_2 and c_2 for the NIR line.

2.3.5.2 Sensitivity analysis

The next step was to determine the four wavebands required to calculate nitrogen-sensitive REPs by the linear extrapolation method. For that reason, we carried out a sensitivity analysis to assess the strength of the correlation between REPs and foliar nitrogen concentration using different waveband combinations. Using fixed wavebands at 680 nm and 760 nm (representing the start and end of the red edge) for the far-red and NIR lines, respectively (Fig. 2.4), the most sensitive wavebands near 700 nm i.e. region A (684-716 nm) for the far-red, and near 725 nm i.e. region B (723-736 nm) for the NIR lines were determined. All combinations (230 in total) of wavebands involving 23 wavebands in region A and 10 wavebands in region B using the GER 3700 instrument were used to calculate 230 (23 x 10) sets of REPs for each spectral data set (rye canopy, maize leaf and mixed grass/herb leaf stack spectra). Subsequently, Pearson correlation coefficients were calculated between REPs for all 230 REP sets and foliar nitrogen concentrations for each data set. The resulting 230 correlation coefficients and their corresponding far-red and NIR band combinations were plotted in a contour plot in order to visualise the most sensitive waveband combinations.

Table 2.2 Red edge spectral coverage for AVIRIS (~10 nm bandwidth), HyMap (15-16 nm bandwidth) and Hyperion sensors (~ 10 nm bandwidth).

AVIRIS Band centre (nm)	HyMap Band centre (nm)	Hyperion Band centre (nm)
665.73	665	671.02
675.62	680	681.20
683.30	695	691.37
692.88	711	701.55
702.46	725	711.72
712.04	740	721.90
721.63	756	732.07
731.21	771	742.25
740.80	786	752.43
750.38	801	762.60
759.97		772.78
769.56		782.95
779.14		793.13
788.73		803.30
798.32		
807.91		

2.3.6 Comparing the performance of various REP extraction techniques for wider bandwidth spectra

The GER spectra in the 350 to 1050 nm range (1.5 nm bandwidth) of all three spectral data sets used in this study were re-sampled to the spectral resolution of two current airborne sensors; the Airborne Visible and Infrared Imaging Spectrometer (AVIRIS) (~10 nm bandwidth), and HyMap (~15 nm bandwidth) instruments, and one spaceborne sensor; the Hyperion (~10 nm bandwidth), in order to assess the utility of the linear extrapolation technique with coarser resolution spectra. The re-sampling was conducted using ENVI (Environment for Visualising Images, Research System, Inc.) software. Table 2.2 shows the "red edge" spectral coverage for each sensor type. The following waveband combinations were used to determine REP values by the linear extrapolation method:

AVIRIS: far-red 683.30 & 702.04 nm and NIR 731.21 & 759.97 nm.
HyMap: far-red 680 & 695 nm and NIR 725 & 756 nm.
Hyperion: far-red 681.20 & 701.55 nm and NIR 732 & 762.60 nm.

The correlation results between foliar nitrogen concentrations and REPs determined by the linear extrapolation method using the broader

bandwidth data were compared to those determined by the linear four-point interpolation, and polynomial and inverted Gaussian fitting methods. The nearest bands to 670, 700, 740 and 780 nm for each sensor's band setting were used to calculate the REP by the linear four-point interpolation technique.

2.4 Results

2.4.1 Sensitivity analysis

The results of the sensitivity analysis of REP extracted by the linear extrapolation method (Fig. 2.5) show that far-red wavebands in the 690-703 nm range in combination with NIR bands in the 723-733 nm range yielded high correlation coefficients with foliar nitrogen concentrations across all three spectral data sets (rye, maize and mixed grass/herb). Note that the REP derived from leaf and canopy spectral data show similar contour patterns, particularly in the case of the maize leaves and rye canopy reflectance data. The high correlation centre for the mixed grass/herb leaf stacks shows a shift towards shorter wavelength lengths, suggesting a less important role of longer NIR wavelengths.

To determine the optimal combinations of wavebands that yielded the highest correlations for all three spectral data sets, the 230 combinations were ranked in ascending order according to decreasing correlation coefficients (r) for each data set (that is, number 1 representing the highest r). The corresponding ranks (across data sets) were then summed. The results of the first ten combinations that yielded the lowest sum of ranks are presented in Table 2.3.

The most sensitive far-red waveband across all three spectral data sets is 694.3 nm; it appears in six of the first ten best combinations. This is followed by the 692.83 nm and 689.9 nm wavebands. The best NIR band is located at 732.46 nm. Certainly, further research is needed to determine the optimal far-red and NIR wavebands for predicting nitrogen for a wide variety of plant species.

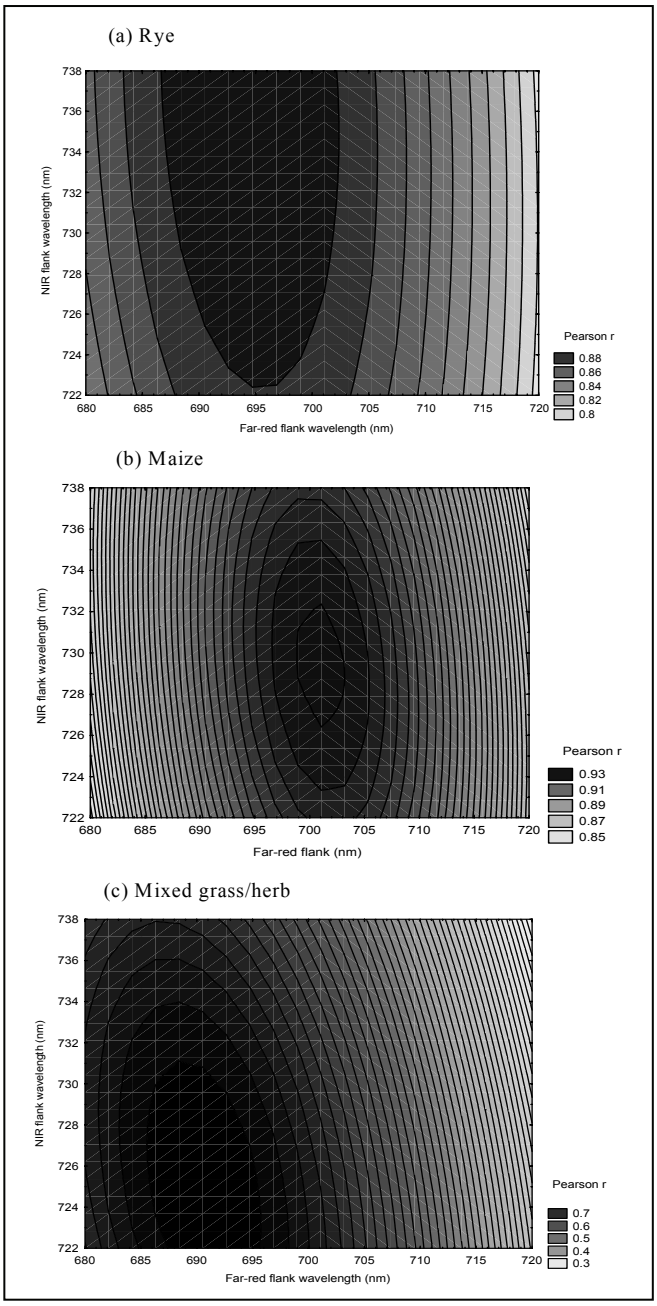


Fig.2.5 Contour plots showing the sensitivity of red edge positions (red edge position/foliar nitrogen concentration correlation – Pearson r) calculated by the linear extrapolation method using fixed wavebands at 680 and 760 nm for the far-red and NIR lines, respectively and different combinations of wavebands on the far-red and NIR flanks of the red edge derivative spectrum.

Table 2.3 The first ten best combinations of far-red and NIR coordinate points used in calculating the red edge position (REP) by the linear extrapolation method. The combinations were ranked in ascending order according to decreasing nitrogen/REP correlation (r) for each data set, followed by the summing of corresponding ranks.

	Combinations		Correlation (r) with leaf nitrogen		
Sum of ranks	Far-red band (nm)	NIR band (nm)	Rye	Maize	Mixed grass/herb
64	694.30	732.46	0.91	0.93	0.69
67	694.30	723.64	0.91	0.93	0.69
74	694.30	729.52	0.91	0.93	0.68
76	694.30	726.58	0.91	0.93	0.68
84	689.90	732.46	0.91	0.92	0.71
84	692.83	732.46	0.90	0.93	0.68
93	692.83	723.64	0.90	0.93	0.68
94	694.30	728.05	0.89	0.93	0.69
95	692.83	726.58	0.90	0.93	0.67
97	694.30	725.11	0.89	0.93	0.69

† All far-red and NIR lines cut the first derivative curve at 680 nm and 760 nm respectively.

2.4.2 Comparing the statistics of red edge positions extracted by different techniques

Table 2.4 shows the statistics for the REPs determined by the linear extrapolation technique (using far-red 679.65 nm and 694.30 nm wavebands and NIR 732.46 and 760.41 nm wavebands), and REPs extracted by the maximum first derivative, linear interpolation, fifth-order polynomial fitting and Gaussian fitting techniques. It is evident from Table 2.4 that the results of REP calculations are dependent upon the choice of method. The REPs retrieved by the four-point interpolation method are biased towards the longer wavelengths, confirming results obtained by Dawson and Curran (1998).

Table 2.4 Statistics of red-edge positions (REP) extracted by various techniques.

REP extraction technique	Mean (nm)	95% CI	Minimum (nm)	Maximum (nm)
Rye (canopy spectra, n=30)				
Maximum first derivative	711.95	4.46	697.23	726.58
Linear interpolation	717.54	0.98	710.86	720.86
Polynomial fitting	714.61	1.41	707.56	720.77
Inverted Gaussian fitting	710.89	1.05	706.39	715.72
Linear extrapolation	707.17	1.99	698.8	715.53
Maize (leaf spectra, n=67)				
Maximum first derivative	711.64	2.54	694.30	723.64
Linear interpolation	717.69	1.64	695.50	724.52
Polynomial fitting	715.38	1.11	701.75	722.22
Inverted Gaussian fitting	708.96	1.19	697.59	715.72
Linear extrapolation	712.58	3.68	685.4	739.05
Grass/herb (leaf stack spectra, n=34)				
Maximum first derivative	710.35	3.68	697.23	725.11
Linear interpolation	717.83	0.46	714.31	720.98
Polynomial fitting	713.06	1.17	704.56	719.24
Inverted Gaussian fitting	710.31	0.71	705.44	714.92
Linear extrapolation	705.63	1.65	694.2	714.76

CI denotes confidence interval.

2.4.3 Performance of five techniques for estimating foliar nitrogen concentration

To compare the strength of the relationship between REPs extracted by different techniques and foliar nitrogen concentration, a bootstrapping resampling technique (McGarigal et al., 2000) was applied to the correlation analyses. Bootstrapping with 1000 replicates was used to provide estimates of the statistics (mean, standard error, confidence interval) for the correlation coefficients between the REP and foliar nitrogen concentration.

REPs extracted by the new technique (linear extrapolation) showed high correlations with a wide range of foliar nitrogen concentrations, being comparable with results obtained using the traditional linear interpolation, polynomial and inverted Gaussian fitting techniques (Table 2.5). But the scatter plots (Fig. 2.6) indicate that the linear interpolation technique and to a lesser extent the polynomial and inverted Gaussian fitting techniques are less sensitive to lower and higher nitrogen

concentrations, because the points tend to saturate at these extremes. For example, when we fitted exponential models to the maize scatter plots, the coefficients of determination (R^2) increased from 0.73 to 0.80 (9.6% increase) for the linear interpolation technique, from 0.78 to 0.84 (7.7% increase) for the polynomial fitting technique, and from 0.83 to 0.86 (3.6% increase) for the inverted Gaussian fitting technique, compared with a slight decrease from 0.86 to 0.85% for the linear extrapolation technique. This could be a further reason why, compared with the other techniques, the REP extracted using the linear extrapolation technique showed a higher R^2 for the mixed grass/herb leaves with a low nitrogen concentration. The results show that the maximum first derivative technique is not an appropriate measure for the red edge position because of the discontinuity it creates in the REP data.

Table 2.5 Bootstrapped correlation coefficients between foliar nitrogen concentrations and red edge positions (REP) extracted using various techniques.

REP extraction technique	Mean	SE mean	LCL 95%	UCL 95 %
Rye (canopy spectra, n=30)				
Maximum first derivative	0.85**	0.05	0.77	0.93
Linear interpolation	0.91**	0.02	0.88	0.96
Polynomial fitting	0.91**	0.03	0.86	0.96
Inverted Gaussian fitting	0.90**	0.03	0.85	0.95
Linear extrapolation (new technique)	0.91**	0.02	0.88	0.94
Maize (leaf spectra, n=67)				
Maximum first derivative	0.86**	0.03	0.81	0.91
Linear interpolation	0.86**	0.02	0.83	0.89
Polynomial fitting	0.88**	0.02	0.85	0.91
Inverted Gaussian fitting	0.91**	0.01	0.90	0.94
Linear extrapolation (new technique)	0.93**	0.01	0.91	0.95
Mixed grass/herb (leaf stack spectra, n=34)				
Maximum first derivative	0.53*	0.10	0.37	0.69
Linear interpolation	0.54*	0.09	0.39	0.69
Polynomial fitting	0.62**	0.08	0.48	0.76
Inverted Gaussian fitting	0.59*	0.07	0.43	0.71
Linear extrapolation (new technique)	0.69**	0.08	0.55	0.83

* = $p < 0.01$, ** $p < 0.001$; LCL and UCL denote lower and upper confidence limits, respectively

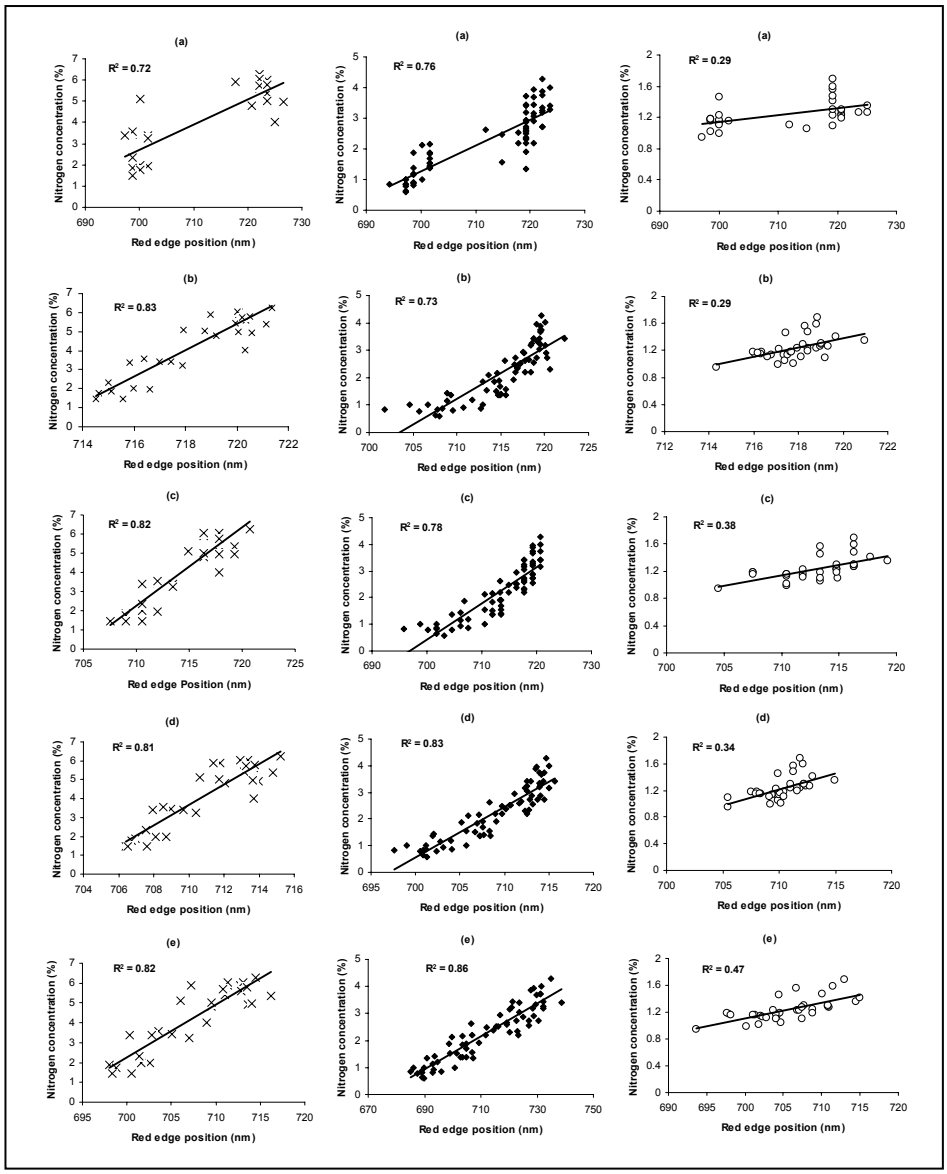


Fig. 2.6 Straight-line model fits for the regression between foliar nitrogen concentration and red edge position (REP) extracted using (a) maximum first derivative, (b) linear interpolation, (c) polynomial fitting, (d) inverted Gaussian modelling and (e) linear extrapolation (new approach) techniques. x = rye canopy, ♦ = maize leaves and ○ = mixed grass/herb leaf stack spectral data sets.

2.4.4 Comparing the performance of various REP extraction techniques for wider bandwidth spectra

The correlations between REPs extracted by the new technique (linear extrapolation) from the three coarser resolution spectra (AVIRIS, HyMap and Hyperion) and foliar nitrogen concentration data also showed comparable results with those obtained using the traditional linear interpolation, high-order polynomial and inverted Gaussian fitting techniques (Table 2.6). The band setting of the HyMap instrument appears to be less favourable to the linear four-point interpolation technique. The results of the REP/nitrogen correlation for the maximum first derivative technique have not been included in Table 2.7 because this technique is not suitable for the small number of bands in the red edge region of the three coarser resolution spectra.

We used one-way ANOVA (analysis of variance) to test if there are significant differences in the mean R^2 between the linear interpolation and linear extrapolation techniques (two simple techniques) for the GER and three coarser resolution (AVIRIS, HyMap and Hyperion) spectra. The test was conducted for all three data sets used in the study (rye canopy, and maize and mixed grass/herb leaves spectra). The results show significant differences between the linear interpolation and linear extrapolation methods only for the maize ($p < 0.01$) and mixed grass/herb ($p < 0.05$) data sets.

Table 2.6 Correlation (R^2) between foliar nitrogen concentration and red edge position derived from AVIRIS, HyMap and Hyperion band settings for the linear interpolation, polynomial fitting, inverted Gaussian fitting and linear extrapolation techniques.

Red edge extraction technique	AVIRIS	HyMap	Hyperion
Rye (canopy spectra, n=30)			
Linear interpolation	0.83	0.82	0.83
Polynomial fitting	0.79	0.81	0.80
Inverted Gaussian fitting	0.81	0.81	0.82
Linear extrapolation	0.82	0.79	0.81
Maize (leaf spectra, n=67)			
Linear interpolation	0.75	0.66	0.74
Polynomial fitting	0.69	0.72	0.61
Inverted Gaussian fitting	0.83	0.83	0.83
Linear extrapolation	0.85	0.86	0.85
Mixed grass/herb (leaf stack spectra, n=34)			
Linear interpolation	0.31	0.17	0.29
Polynomial fitting	0.32	0.36	0.36
Inverted Gaussian fitting	0.30	0.28	0.31
Linear extrapolation	0.43	0.39	0.35

2.5 Discussion

In this study, we hypothesised that first, the discontinuity in the relationship between the REP and nitrogen concentration caused by the double-peak feature on the first derivative spectrum could be mitigated and secondly, spectral changes near the dominant peaks (700 nm and 725 nm) could be adequately tracked if the REP is determined as an intersection of two straight lines extrapolated through two points on the far-red flank (680 nm to 700 nm) and two points on the NIR flank (725 nm to 760 nm) of the first derivative reflectance spectrum. We have demonstrated that the REPs determined by this new approach, termed the linear extrapolation technique from three spectral data sets (rye canopy, maize leaf and mixed grass/herb leaf stack spectra) covered a wide range of wavelengths (from shorter to longer wavelengths) and generated continuous relationships with foliar nitrogen concentrations. More importantly, this study has shown that, by extrapolating straight lines near the first and second peaks in the red edge derivative spectrum, the resulting REPs can account for subtle variations in a wide range of foliar nitrogen concentrations (from low to high). This could be explained by

the fact that the REP values depend on changes in the magnitudes of the first and second peaks. The first and second peaks were shown to be sensitive to low and high foliar nitrogen concentrations, respectively (Fig. 2.3). These results confirm those of Horler et al. (1983) for leaf chlorophyll content. Note that the high positive linear relationship between REPs and nitrogen concentrations obtained in this study depends on the high positive linear relationship between foliar nitrogen and chlorophyll content (Yoder and Pettigrew-Crosby, 1995) because shifts in the REP are mainly attributed to changes in chlorophyll content (Buschmann and Nagel, 1993; Dawson and Curran, 1998).

The sensitivity of the REP using the new method depended on the choice of far-red and NIR bands. Far-red bands at 679.65 nm and 694.30 nm in combination with NIR bands at 732.46 nm and 760.41 nm or at 723.64 nm and 760.41 nm were identified as the optimal combinations for calculating foliar-nitrogen-sensitive REPs for three spectral data sets (rye canopy, maize leaf and mixed grass/herb leaf stack spectra). Further research is needed to determine the optimal far-red and NIR bands for predicting nitrogen for a wide variety of plant species or vegetation communities.

Table 2.7 Summary of the relative performances of five red edge position extraction techniques

REP extraction technique	Complexity	Required spectral type	Suitability for coarse spectra	Correlation with foliar nitrogen	
				Rye and maize	Mixed grass/herb
Maximum first derivative	Easy	Derivative	Poor	High but discont.	Low
Four-point interpolation	Easy	Reflectance	Good	High	Low
Polynomial fitting	Moderate	Derivative	Good	High	Medium
Inverted Gaussian fitting	Difficult	Reflectance	Good	High	Low
Linear extrapolation	Easy	Derivative	Good	High	Medium

Note: discont. – discontinuous

Table 2.7 summarises the relative performances of the five techniques used in this study for extracting the REP. REPs extracted using the new technique show high correlations with a wide range of foliar nitrogen

concentrations, being comparable with the results obtained using the traditional fitting techniques (linear interpolation, high-order polynomial and inverted Gaussian fitting techniques). Though, for all techniques, the REP/nitrogen correlation is low for leaves with a low nitrogen concentration or mixed species foliage (e.g. mixed grass/herb data), a slightly higher correlation result was obtained with the new technique (Table 2.5). The results show that the maximum first derivative technique is not an appropriate measure for the red edge position because of the discontinuity it creates in the REP data and therefore the REP/nitrogen relationship. The differences observed in REP/nitrogen correlations between the five techniques for the three spectral data sets used in this study suggest that the accuracy of different techniques depends on the biological characteristics of the plant material. REP has been shown to be dependent not only on chlorophyll content, but also on additional effects such as leaf developmental stage, leaf layering or stacking and leaf water content (Horler et al. 1983). Although the method used in this study to carry out leaf or leaf stack spectral measurements predominantly captures leaf reflectance, the additional effect of leaf layering (e.g. mixed grass/herb leaves with varying thickness and internal structure) or background material which has been shown to influence NIR reflectance (Horler et al. 1983) may affect the results. Even though (Baranoski and Rokne, 2005) suggest that no single REP extraction method may be considered superior in all cases, the linear extrapolation technique appears to be less sensitive to the effect of leaf stacking or layering.

Degrading the spectral resolution of the GER spectra to the spectral configurations AVIRIS, Hyperion and HyMap, showed a negligible effect on the correlation results for the linear interpolation, higher order polynomial fitting, inverted Gaussian fitting and linear extrapolation techniques. However, the correlation results for the various REP techniques differ among sensors because of different band settings in the region of the red edge. For example, the results show that the HyMap band configuration is not suitable for the linear interpolation method. Thus, the linear extrapolation technique provides an additional choice of method for determining the REP given the increasing number of air-and spaceborne hyperspectral sensors. Another advantage of the linear extrapolation technique is its ease of implementation because only four spectral bands are required for the extrapolation. In addition, it requires first derivative spectrometry. Derivative analysis enhances absorption features and suppresses contributions of non-vegetative reflectance

components (Boochs et al., 1990; Curran et al., 1991). A drawback is that derivative spectra are more sensitive to the inherent spectral noise of the system (Broge and Leblanc, 2000).

2.6 Conclusions

The most important conclusions from this study are:

i. The destabilising effect of the double-peak feature on the REP/nitrogen relationship can be mitigated and spectral changes near the low (700 nm) and high nitrogen (725 nm) sensitive peaks can be tracked by determining the REP as an intersection of two straight lines extrapolated on the far-red (680 nm to 700 nm) and NIR (725 nm to 760 nm) flanks of the first derivative reflectance spectrum. This method has been termed the linear extrapolation method.
ii. Far-red wavebands at 679.65 nm and 694.30 nm in combination with NIR wavebands at 732.46 nm and 760.41 nm or at 723.64 nm and 760.41 nm are the optimal combinations for calculating foliar nitrogen-sensitive REPs by the linear extrapolation method.
iii. The correlation results between REP extracted by the new method and foliar nitrogen concentration are comparable to those of the traditional linear four-point interpolation, and high-order polynomial and inverted Gaussian fitting techniques. However, the results of this study show that the linear extrapolation technique performs better than the linear four-point interpolation technique in the case of maize leaves at different developmental stages and mixed grass/herb leaves with a low nitrogen concentration.
iv. The maximum first derivative technique is not an appropriate measure for the red edge position because of the discontinuity it creates in the REP data for both narrow and wider bandwidth spectra.
v. The linear extrapolation technique is simple to implement.
vi. A drawback of the linear extrapolation technique is that it uses the first derivative spectrum, which is more sensitive to the inherent spectral noise of the system.
vii. The REP could be extracted by the linear extrapolation technique from wider bandwidth spectra e.g. AVIRIS, HyMap and Hyperion.

In summary, the results indicate that the new technique is a practical and suitable technique for extracting REP from hyperspectral data for explaining a wide range of nitrogen concentrations. However, further

research is needed to assess the accuracy of the new technique to predict leaf nitrogen or chlorophyll for a wide variety of plant species. In addition, the efficacy of the technique for predicting other plant parameters such as leaf area index and biomass also needs to be established.

Acknowledgements

The International Institute for Geo-information Science and Earth Observation (ITC) provided financial support for the experiment. The UNIFARM experimental station at Wageningen University contributed to the success of the greenhouse experiment. In this regard, we would like to express our thanks to Dr Sipke E. van Wieren, Andre Maassen, Jan Walsem and Frederik Hengeveld. We also extend our gratitude to the management of Majella National Park in Italy, and particularly to Dr Theodoro Andrisano.

Chapter 3

Towards red-edge positions less sensitive to canopy biophysical parameters for leaf chlorophyll estimation using PROSPECT-SAILH simulated data

This chapter is based on
Cho, M.A., Skidmore, A.K. and Atzberger, C.G. (2006) Towards red - edge positions less sensitive to canopy biophysical parameters using PROSPECT - SAILH simulated data. In: ISPRS 2006: ISPRS mid-term symposium 2006 remote sensing: from pixels to processes, 8-11 May 2006, Enschede, the Netherlands. Enschede: ITC, 2006. 6 p

Cho, M.A., Skidmore, A.K. and Atzberger, C.G. (In review after revision) Towards red - edge positions less sensitive to canopy biophysical parameters for leaf chlorophyll estimation using PROSPECT - SAILH simulated data. International Journal of Remote Sensing

Abstract

Several methods for extracting the chlorophyll sensitive red-edge position (REP) from hyperspectral data are reported in literature. The objective of this study was to test the utility of a recently published approach, the linear extrapolation method under different conditions including variable canopy biophysical parameters, solar zenith angle, sensor noise and spectral bandwidth. REPs were extracted from synthetic canopy spectra that were simulated using PROSPECT and SAILH radiative transfer models. REPs extracted by the linear extrapolation method involving wavebands at 680, 694, 724 and 760 nm produced the highest correlation (R^2 = 0.75) with leaf chlorophyll content with minimal effects of leaf and canopy biophysical confounders (leaf area index, leaf inclination distribution and leaf dry matter content) compared to traditional techniques including the linear interpolation, inverted Gaussian modelling and polynomial fitting techniques. In addition, the new technique is insensitive to changes in solar zenith angle. However, the advantage of using the linear extrapolation method compared to the various alternative methods diminishes with increasing sensor noise and decreasing spectral resolution. In summary, the linear extrapolation technique shows high potential for leaf chlorophyll estimation with radiative transfer models. The efficacy of the technique under field conditions needs to be established.

Keywords: red-edge position; linear extrapolation method; leaf chlorophyll; canopy biophysical parameters; radiative transfer models

3.1 Introduction

Accurate remotely sensed estimates of leaf chlorophyll content can provide valuable information on ecosystem functioning over a wide range of scales e.g. as an indicator of vegetation stress (Horler et al., 1980; Collins et al., 1983; Hare et al., 1984; Daughtry et al., 2000; Clevers et al., 2004) or ecosystem productivity (Mooney, 1986; Peterson et al., 1988; Blackburn, 1998). Commonly used vegetation indices for chlorophyll estimation computed from visible and near infrared (NIR) bands (Gausman, 1977; Gamon et al., 1992; Peñuelas et al., 1995; Lichtenthaler et al., 1996; Gitelson and Merzlyak, 1997; Blackburn, 1998; Haboudane et al., 2002) are also influenced by other leaf and canopy parameters such as carotenoids (yellow pigments), leaf internal structure, mass and stacking, leaf area index (LAI), leaf angle distribution (LAD) and soil reflectance (Huete, 1988; Chappelle et al., 1992; Goward and Huemmrich, 1992; Blackburn, 1998; Daughtry et al., 2000). Research has also demonstrated a strong effect of the observation and illumination geometry on measured canopy spectra (Huete et al., 1992; Qi et al. 1995) that perturb the relation between vegetation indices and leaf chlorophyll content.

A spectral measure that is less sensitive to the effect of variable leaf and canopy biophysical parameters, and environmental conditions on leaf chlorophyll estimation is the wavelength of maximum slope in the region of the red edge (680 to 780 nm), termed the red-edge position (REP) (Horler et al., 1983; Curran et al., 1995; Clevers et al., 2002). The red edge represents the region of abrupt change in leaf reflectance between 680 nm and 780 nm caused by the combined effects of strong chlorophyll absorption in the red and leaf internal scattering in the NIR (Gates et al., 1965; Horler et al., 1983). Increases in the amount of chlorophyll results in a broadening of the major chlorophyll absorption feature centred around 680 nm (Buschmann and Nagel, 1993; Dawson and Curran, 1998) causing a shift in the slope and REP towards longer wavelengths (Gates et al., 1965; Collins et al., 1977; Horler et al., 1980; Horler et al., 1983; Hare et al., 1984; Boochs et al., 1990; Clevers et al. 2002).

A common approach for extracting the REP has been to locate the highest peak in the first derivative spectrum (Horler et al., 1983; Boochs et al., 1990; Buschmann and Nagel, 1993; Filella and Peñuelas, 1994). However, the limitation of this approach is that the first derivative of

contiguous spectra may contain two or more peaks (double-peak feature) near 700 and 725 nm (e.g. Horler et al., 1983; Boochs et al., 1990; Zarco-Tejada et al., 2003; Clevers et al., 2004). The double peak feature causes a peak jump between 700 and 725 nm and a discontinuity in the REP/chlorophyll relationship (Horler et al., 1983).

In our previous study (Cho and Skidmore, 2006), we proposed a new technique based on locating the REP as the point of intersection between two straight lines extrapolated on the far-red and NIR flanks of the first derivative spectrum (the linear extrapolation method). We showed that the linear extrapolation method not only mitigates the destabilising effect of the double peak feature, but also predicts leaf nitrogen concentration with high accuracy. The correlation between REP and nitrogen concentration depends on the close correlation between nitrogen and chlorophyll (Yoder and Pettigrew-Crosby, 1995; Hansen and Schjoerring, 2003; Haboudane et al., 2004; Wenjiang et al., 2004). But the relationship between chlorophyll and nitrogen depends on the physiological status of the plant (Mooney, 1986; Boochs et al., 1990), e.g. changing from low to high positive correlations with increasing leaf age (Wenjiang et al., 2004). It should be noted that nitrogen is not only a major component of leaf chlorophyll, but also forms part of inert structural components of cell tissue (Mooney, 1986; Jongschaap and Booij, 2004). Thus, indices for chlorophyll estimation that are maximally sensitive to chlorophyll with minimal effects of leaf and canopy structure, solar zenith angle, etc. are potentially useful.

The objective of this study was to test whether the linear extrapolation method may be applied under different conditions including variable leaf chlorophyll and canopy biophysical parameters, solar zenith angle, sensor noise and spectral bandwidth. To achieve this objective, we used a numerical experiment involving well-established canopy reflectance models, parameterised to represent a wide range of canopy characteristics. This allowed us to artificially create pseudo measurements that otherwise would have been difficult and expensive to obtain under experimental or field conditions.

3.2 Methods

3.2.1 Radiative transfer models

3.2.1.1 PROSPECT and SAILH models

We simulated synthetic reflectance spectra using the PROSPECT and SAILH radiative transfer models. PROSPECT is a leaf optical properties model developed by Jacquemoud and Baret (1990). It simulates leaf reflectance (ρ_{leaf}) and transmittance spectra (t_{leaf}) between 400 and 2500 nm using four model inputs: leaf chlorophyll content (Cab; µg cm^{-2}), equivalent leaf water thickness (Cw; cm), leaf dry matter content (Cm; g cm^{-2}), and a leaf structure index (N; arbitrary units). Specific absorption and scattering coefficients of leaf components are provided with the model. The model is widely used and well validated (Fourty et al., 1996).

SAILH is a four-stream radiative transfer model developed by Verhoef (1984). It was later modified by Kuusk (1991) to take the hot spot feature into account. For the purpose of this study, SAILH was chosen to simulate bi-directional canopy reflectance (ρ) since it requires only few input variables, while having a predictive power similar to more elaborated reflectance models (Jacquemoud et al., 1995; Jacquemoud et al., 2000; Bacour et al., 2002). SAILH assumes the canopy to be a homogeneous semi-infinite medium with Lambertian leaves characterized by their reflectance and transmittance spectra (ρ_{leaf}, t_{leaf}). Soil reflectance (ρ_{soil}) must be specified at the lower boundary. Canopy structure is characterized by the leaf area index (LAI; m^2 m^{-2}) and the average leaf angle of an ellipsoidal leaf inclination distribution with random azimuth orientation (ALA; degrees). The hot spot effect is modelled using the ratio between leaf size and canopy height (s; m m^{-1}). Further variables characterise the measurement geometry (θz, θv), and the fraction of diffuse illumination (skyl).

Soil reflectance at the lower boundary of the canopy (ρ_{soil}) was modelled using a simple soil parameterization described in Atzberger et al. (2003). In contrast to similar studies, the soil reflectance parameterisation does not only change the overall brightness of a (standard) soil spectrum, but also allows for (small) changes in the spectral shape, for example due to variations in the chemical composition of the soil (here soil carbon content). For a more detailed description and experimental validation of the soil reflectance parameterization, see Atzberger et al. (2003).

Table 3.1 Specification of parameter ranges and distributions for SAILH+PROSPECT reflectance modelling. In all cases, a nadir looking sensor has been assumed. The solar zenith angle was set to 45°. This value was eventually varied by ± 20° to study its effect on the proposed leaf chlorophyll retrieval technique.

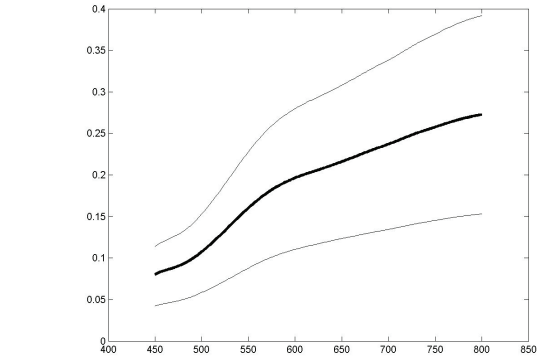
Model parameter	Abbreviation	Units	Distribution	Range[(1)]
Canopy parameter				
Leaf Area Index	LAI	m^2 m^{-2}	uniform	0-10
Average Leaf Angle	ALA	° (degree)	uniform	30-80
Hot spot parameter	hot	no dimension	normal	0.1 ± 0.01
Leaf parameter				
Leaf chlorophyll content	C_{ab}	g cm^{-2}	uniform	20-80
Leaf water content	C_w	cm	uniform	0.004-0.044
Leaf dry matter content[(2)]	C_m	g cm^{-2}	uniform	1.25 x C_w
Leaf structure parameter	N	no dimension	normal	2 ± 0.2
Soil parameter				
Soil brightness	SCALE	no dimension	normal	1 ± 0.14
Carbon content	C_c	g cm^{-3}	uniform	0-6

[(1)]In cases where *distribution* is *normal*, *range* indicates mean ± std.
[(2)]C_m is varied proportional to C_w as proposed by Combal et al. (2002)

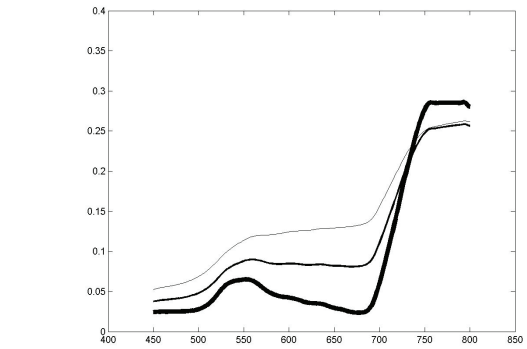
3.2.1.2 Model parameterisation

For a given measurement geometry, the full parameterisation of the radiative transfer models involves nine (structural and biochemical) variables. Their parameter ranges and distributions are described in Table 3.1. Within the distributions of Table 3.1, 1000 parameter sets were randomly chosen to simulate the synthetic canopy reflectance spectra. The wavelength range was restricted between 450 and 800 nm (351 values in 1-nm steps) as the study focuses only on the visible and near infrared (VNIR). The distributions cover a wide range of canopy and leaf properties, including widely varying leaf angles (from planophile to erectophile), different canopy densities (from bare soil to fully developed canopies), different soil albedos and leaf optical properties, etc. For all simulations, a nadir looking sensor was assumed ($\theta v=0°$). The fraction of diffuse illumination (skyl) was fixed to 0.1, independent of wavelengths. For the main part of the study, a solar zenith angle (θz) of 45 degree was

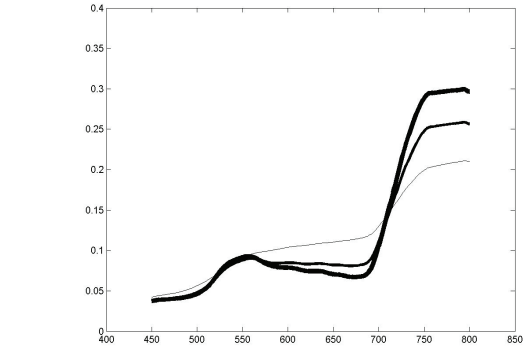
used. This value was varied by ±20° to study its effect on the proposed REP extraction technique. Fig.3.1 illustrates some of the reflectance variability in the synthetic spectra, simulated within the Matlab processing environment (The Mathworks).



(a)



(b)



(c)

Fig.3.1 Illustration of reflectance variability in the noiseless synthetic spectra used in this study. (a) average background soil reflectance (thick line) and range (thin lines). (b) LAI values ranging from 0.5 (thin line), 1.0 to 5.0 (thick line). (c) ALA values ranging from 30° (thick line), 55° to 80° (thin line). For (b) and (c) the average background spectra shown in (a) has been used with remaining SAILH+PROSPECT model parameters fixed to their average values (Table 3.1).

3.2.2 Red-edge position algorithms

We have assessed the correlation between leaf chlorophyll content and REPs determined by the simple maximum derivative, linear interpolation (Guyot and Baret, 1988), inverted Gaussian modelling (Bonham-Carter, 1988; Miller et al., 1990), high order polynomial fitting (Pu et al., 2003) and linear extrapolation (Cho and Skidmore, 2006) techniques. We have not considered the three-point Lagrangian interpolation technique (Dawson and Curran, 1998) because Clevers et al. (2002) show that it is only suitable for coarsely sampled spectra.

(a) Maximum first derivative

The REP is defined by the wavelength of the maximum first derivative of the reflectance spectrum in the region of the red edge. The first derivative was calculated using a first-difference transformation of the reflectance spectrum (Dawson and Curran, 1998) as follows:

$$FDR_{(\lambda i)} = (R_{\lambda(j+1)} - R_{\lambda(j)})/\Delta\lambda \quad (3.1)$$

where FDR = first derivative reflectance at a wavelength i midpoint between wavebands j and j+1
$R_{\lambda(j)}$ = reflectance at the j waveband
$R_{\lambda(j+1)}$ = the reflectance at the j+1 waveband, and
$\Delta\lambda$ = difference in wavelengths between j and j+1.

(b) Linear interpolation technique

The linear interpolation method (Guyot and Baret, 1988) assumes that the reflectance curve at the red edge can be simplified to a straight line centred around the midpoint between the reflectance in the NIR at about 780 nm and the reflectance minimum of the chlorophyll absorption feature at about 670 nm. It uses four wavebands (670, 700, 740 and 780 nm), and the REP is determined by using a two-step calculation procedure.

(i) Calculation of the reflectance at the inflexion point (R_{re})

$$R_{re} = (R_{670} + R_{780})/2 \quad (3.2)$$

where R = reflectance

(ii) Calculation of the red edge wavelength or red edge position (REP)

$$REP = 700 + 40\left(\frac{R_{re} - R_{700}}{R_{740} - R_{700}}\right) \tag{3.3}$$

where 700 and 40 are constants resulting from interpolation in the 700-740 nm interval.

(c) Inverted Gaussian fitting technique

An inverted Gaussian (IG) model (Bonham-Carter, 1988; Miller et al. 1990; Dawson and Curran, 1998; Pu et al. 2003) was fitted to the spectral reflectance in the 660-780 nm band range. Accordingly, the IG model (Eq. 3.4) represents the red edge by the reflectance equation:

$$R(\lambda) = R_s - (R_s - R_o)\exp\left(-\frac{(\lambda_0 - \lambda)^2}{2\sigma^2}\right) \tag{3.4}$$

where R_s = maximum spectral reflectance
R_o = minimum spectral reflectance
λ_o = wavelength of minimum reflectance
σ = Gaussian function variance.

The REP is then defined as:

$$REP = \lambda_0 + \sigma \tag{3.5}$$

We used an iterative optimisation fitting procedure to determine parameters of the IG model (Miller et al., 1990). Initial guesses of the model parameters were made after review of each data set. Typically, R_o was set at 670 and 30 nm was selected for σ. The IG model employs a least-square criterion to fit a normal curve to the reflectance red edge. The values of λ_o, R_o, R_s and σ are then determined by the fitting procedure.

(d) Polynomial fitting technique

A polynomial (Pu et al. 2003) function (e.g. 3rd order polynomial - Eq. 3.6) was fitted to the reflectance spectrum between the wavelengths, corresponding to the minimum reflectance in the red and the maximum NIR (shoulder) reflectance.

$$R(\lambda) = a_0 + \sum_{i=1}^{3} a_i \lambda^i \tag{3.6}$$

where λ = band between 670 nm to 780 nm.

Subsequently, REP was determined from the maximum first derivative spectrum.

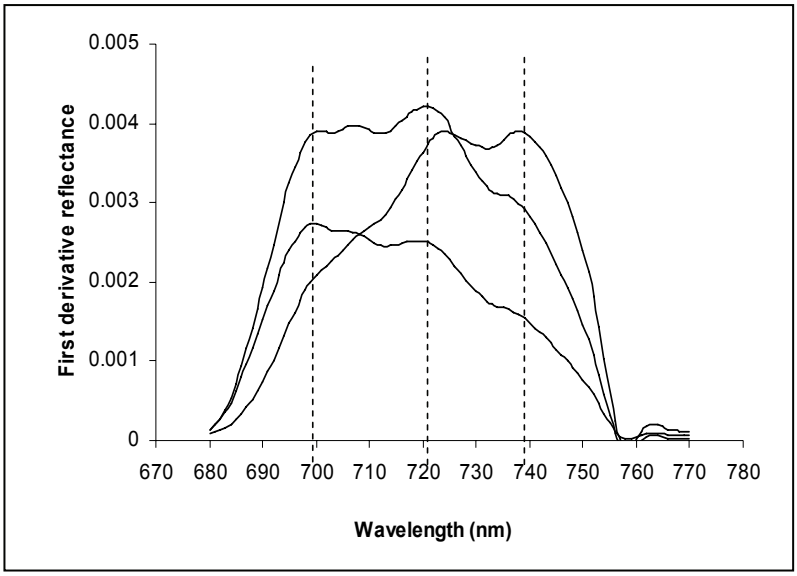


Fig.3.2 First derivative curves for three SAILH+PROSPECT simulated spectra showing multiple peak regions near 700, 720 and 740 nm.

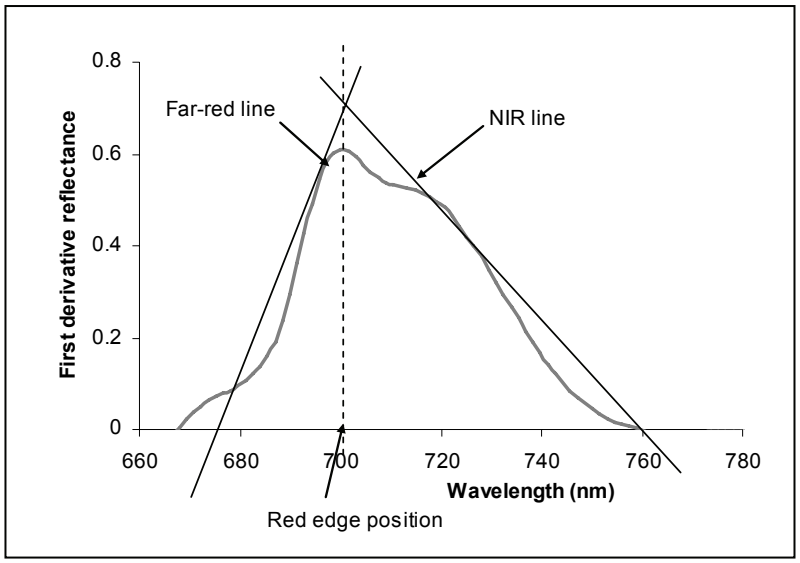


Fig.3.3. Schematic representation of the linear extrapolation technique for extracting the red edge position (REP) – wavelength of the meeting point between two straight lines extrapolated on the far-red and NIR flanks of the first derivative spectrum.

(e) Linear extrapolation technique

The linear extrapolation technique (Cho and Skidmore, 2006) is designed to (i) mitigate the destabilising effect of the double peak feature on the correlation between chlorophyll and REP, and (ii) track changes in slope near 700 nm and 725, where derivative peaks (Fig. 3.2) occur. The REP is calculated as the wavelength at the intersection of two straight lines (Eq. 3.7 & 3.8) extrapolated through two points on the far-red flank and two points on NIR flank of the red edge (680 – 760 nm) first derivative reflectance spectrum (Fig. 3.3).

$$\text{Far-re line: } FDR = m_1\lambda + c_1 \quad (3.7)$$

$$\text{NIR line: } FDR = m_2\lambda + c_2 \quad (3.8)$$

where m and c represent the slope and intercept of the straight lines; c_1 and m_1 for the far-red line and c_2 and m_2 for the NIR line. At the intersection, the two lines have equal λ and FDR values. Therefore, the REP, which is the λ at the intersection, is given by:

$$REP = \frac{-(c_1 - c_2)}{(m_1 - m_2)} \tag{3.9}$$

Cho and Skidmore (2006) identified two combinations of wavebands for calculating leaf nitrogen-sensitive REPs. We shall call them linear extrapolation I involving far-red 680 and 694 nm in combination with NIR 724 and 760 nm, and linear extrapolation II involving far-red 680 and 694 nm in combination with NIR 732 and 760 nm.

3.2.3 Data analysis

3.2.3.1 Relationship between leaf chlorophyll content and red edge position

To evaluate the predictive powers of various REP extraction techniques, the simulated data was divided into training (n = 750) and testing (n = 250) data sets. Linear regression equations (calibration equations) were derived between leaf chlorophyll content and REP using the training data set. The calibration equations were then applied on the REP data from the testing data set in order to predict leaf chlorophyll content. The accuracy of this prediction is reported in terms of the root mean square error (RMSE).

3.2.3.2 Influence of canopy biophysical parameters on red edge positions extracted by various methods

We quantified the main effects and interaction between leaf chlorophyll content and LAI, ALA or leaf dry matter content on REPs derived by the various methods. The contribution of each factor to the total variance in the REP was calculated by dividing its sum of squares by the total sum of squares (Webster, 2000).

3.2.3.3 Effects of solar zenith angle and sensor noise on the linear interpolation method

The effects of solar zenith angle and sensor noise on the linear extrapolation method were assessed by applying the calibration equation obtained from the standard data set (i.e. noise free data and solar zenith angle = 45°) to the corresponding test data for; (i) solar zenith angles at 25° and 65° and (ii) four sensor noise levels (i.e. normally distributed white noise with zero mean and the following standard deviation: 0.05% 0.10%, 0.50% and 1.00% relative to the reflectance). The root mean square errors (RMSE) between the measured chlorophyll and predicted

chlorophyll were used for comparative analyses. The results obtained from the linear extrapolation I and II methods were compared to those of the linear interpolation method because these less complicated methods compared to the inverted Gaussian and polynomial fitting techniques (Pu et al., 2003; Cho and Skidmore 2006).

Table 3.2 Far-red (FR) and near infrared (NIR) wavelengths used for the calculation of the red edge position using the linear extrapolation methods. In case of Hyperion and HyMap spectra, the indicated wavelengths correspond to the central wavelengths.

	Wavelengths (nm) of interest	
Spectral type	Linear extrapolation I	Linear extrapolation II
ASD	FR 680 & 694	FR 680 & 694
	NIR 724 & 760	NIR 732 & 760
Hyperion	FR 681 & 691	FR 681 & 691
	NIR 722 & 763	NIR 732 & 763
HyMap	FR 680 & 695	FR 680 & 695
	NIR 725 & 756	NIR 740 & 756

3.2.3.4 Effects of degrading the bandwidth on the linear extrapolation method

The synthetic 1 nm-data (later called ASD) was re-sampled to the spectral coverage of Hyperion (~10 nm bandwidth) and HyMap (~15 nm bandwidth). The re-sampling was conducted using the ENVI (Environment for Visualising Images, Research System, Inc.) software. REPs were derived by the linear extrapolation method using the wavebands shown in Table 3.2. Subsequently, the strength of the correlations (R^2) between leaf chlorophyll and REP was used to assess the effect of degrading the bandwidth on the linear extrapolation and interpolation methods.

Table 3.3 The relationship between leaf chlorophyll content and red edge position (REP) extracted by various methods.

REP extraction method	R^2 (calibration data set)	Predictive equation	Standard error of prediction (test data set, n = 250)
Max. first derivative	0.50(748 df)	- 647.56 + 0.97*REP	12.75
Linear interpolation	0.60 (748 df)	- 2494.31 + 3.53*REP	10.76
Inverted Gaussian modelling	0.61 (723 df)	- 1707.99 + 2.46*REP	13.87
3rd order polynomial fitting	0.62 (748 df)	- 595.28 + 0.88*REP	10.36
Linear extrapolation I	0.75 (746 df)	- 1111.01 + 1.63*REP	8.98
Linear extrapolation II	0.70 (746 df)	- 866.41 + 1.28*REP	9.77

df = degree of freedom

3.3 Results

3.3.1 Relationship between leaf chlorophyll content and red edge position

The linear regression between leaf chlorophyll content and REP derived by the linear extrapolation method (I and II) yielded higher coefficients of determination (R^2) with the calibration data set and lower standard errors of prediction with the test data compared to the traditional methods (maximum first derivative, linear interpolation, inverted Gaussian and polynomial fitting techniques) (Table 3.3). The REPs located by the inverted Gaussian technique showed several outliers.

Table 3.4 Main and interaction effects (quantified by the coefficient of determination – R^2) between leaf chlorophyll content and canopy biophysical parameters on the red edge position extracted by the wavelength of the maximum slope, linear interpolation, inverted Gaussian, polynomial fitting and linear extrapolation methods.

	Maximum first derivative	Linear interpolat-ion	Inverted Gaussian modelling	3rd order polynomial fitting	Linear extrapola-tion I	Linear extrapola-tion II
C_{ab} and LAI						
C_{ab}	0.490*	0.608*	0.604*	0.625*	0.751*	0.701*
LAI	0.025*	0.015*	0.025*	0.015*	0.008	0.016*
C_{ab}*LAI	0.000	0.008*	0.000	0.000	0.000	0.000
Total R^2	0.515	0.631	0.629	0.641	0.759	0.717
C_{ab} and ALA						
C_{ab}	0.490*	0.608*	0.605*	0.620*	0.750*	0.698*
ALA	0.009*	0.020*	0.013*	0.009*	0.012*	0.011*
C_{ab}*ALA	0.001	0.002*	0.000	0.002*	0.000	0.000
Total R^2		0.630	0.618	0.632	0.767	0.709
C_{ab} and C_m						
C_{ab}	0.490*	0.608*	0.608*	0.621*	0.755*	0.704*
C_m	0.207*	0.250*	0.222*	0.159*	0.142*	0.181*
C_{ab}*C_m	0.004*	0.000	0.004*	0.027*	0.001*	0.001*
Total R^2	0.701	0.858	0.834	0.807	0.898	0.886

* $p < 0.05$, LAI = leaf area index; ALA = average leaf angle of an ellipsoidal leaf inclination distribution with random azimuth orientation; C_m = leaf dry matter content.

Table 3.5 Effect of different solar zenith angles on the prediction accuracy (root mean square error) of leaf chlorophyll by red-edge positions (REP) for the linear extrapolation and interpolation methods.

Solar zenith angle	Linear extrapolation I	Linear extrapolation II	Linear interpolation
25°	8.47	9.40	10.80
45°	8.98	9.77	10.76
65°	8.34	9.01	10.46

3.3.2 Influence of canopy biophysical parameters on red edge positions extracted by various methods

Among the investigated perturbing canopy biophysical variables (LAI, ALA and leaf dry matter content), leaf dry matter content showed the highest influence on the REP (Table 3.4). However, the influence of dry matter content was lowest on the REP derived by the linear extrapolation I method. The contribution of LAI to the total variance of REP was low (R^2 ranges from 0.008 to 0.025) but statistically significant for all the REP techniques with the exception of the linear extrapolation I method. The effect of ALA was equally low but significant for all REP techniques. The interaction effects between the chlorophyll content and biophysical parameters on REP were generally low. We should note that there was no significant ($p>0.05$) correlation between leaf chlorophyll content and LAI or ALA or leaf dry matter.

3.3.3 Effects of solar zenith angle and sensor noise on the linear extrapolation method

The RMSE obtained between the measured and predicted leaf chlorophyll contents for the solar zenith angles at 25°, 45° and 65° were similar (Table 3.5). We used the student's t test to test if there were significant differences between the means of the predicted leaf chlorophyll content for solar zenith angle at 45° (standard data) and 25° or 65°. The results showed that the means are not significantly different ($p>0.05$ in all cases). With respect to noise, the RMSE increased with increasing sensor noise for the linear extrapolation technique (Table 3.6). The linear interpolation technique is not affected by senor noise.

Table 3.6 Effect of different noise levels on the prediction accuracy (root mean square error) of leaf chlorophyll by red-edge positions (REP) for linear extrapolation and interpolation methods.

Noise levels (standard deviation relative to the reflectance - %)	Linear extrapolation I	Linear extrapolation II	Linear interpolation
0.00 (noise free)	8.98	9.77	10.76
0.05	9.01	9.97	10.83
0.10	11.17	16.97	10.75
0.50	74.91	45.40	10.98
1.00	190.7	80.67	11.03

3.3.4 Effects of degrading the bandwidth on the linear extrapolation method

Degrading the bandwidth from 1nm in the original data to the spectral coverage of Hyperion (~10 nm) and HyMap (~15 nm) lowered the correlation between REP and chlorophyll content for the linear extrapolation and interpolation techniques (Table 3.7).

Table 3.7 Correlation (R^2) between leaf chlorophyll content and red edge position derived from ASD, Hyperion and HyMap band settings for the linear extrapolation and interpolation methods.

	Linear extrapolation I	Linear extrapolation II	Linear interpolation
ASD (1 nm)	0.75	0.70	0.61
Hyperion (~10 nm)	0.66	0.59	0.63
HyMap (~15 nm)	0.55	0.50	0.55

3.4 Discussion

The objective of this study was to test whether the linear extrapolation method may be applied under different conditions including variable leaf chlorophyll and canopy biophysical parameters, solar zenith angle, sensor noise and spectral bandwidth.

The results of this study show that REPs derived by linear extrapolation I and II are better predictors of leaf chlorophyll content compared to other traditional methods including the wavelength of maximum first derivative, linear interpolation, inverted Gaussian modelling, and polynomial fitting techniques. In our previous study (Cho and Skidmore, 2006), we obtained the same correlation coefficient between leaf nitrogen

concentration and REP extracted by linear extrapolation I and II for each of the following data sets; rye (*Lolium perenne*) canopy, maize leaf and mixed grass/herb leaf stack spectra. But this study shows that linear extrapolation I performs better than linear extrapolation II for leaf chlorophyll estimation. Note that the correlation between nitrogen and REP depends on the positive relationship between foliar nitrogen and chlorophyll content (Yoder and Pettigrew-Crosby, 1995). But this relationship depends on the physiological status of the plant (Mooney, 1986; Boochs et al., 1990).

The REPs extracted by the various methods are in general less sensitive to leaf and canopy biophysical parameters compared to leaf chlorophyll content. However, the results suggest that REPs derived by the linear extrapolation I technique are least sensitive to LAI and leaf dry matter content compared to the various alternatives. Results reported in literature on the relationship between REP and LAI are mixed. Some studies using one or at most a few closely related species suggest that REP is influenced by both chlorophyll content and LAI (Filella and Peñuelas, 1994; Danson and Plummer, 1995; Pu et al., 2003). On the contrary, Boegh et al. (2002) found no relationship between REP and LAI across eight crop fields consisting of both winter-sown and spring-sown crops but observed a high positive relationship between REP and leaf nitrogen concentration. Broge and Leblanc (2000) using PROSPECT and SAIL simulated data observed that REP poorly relates to LAI.

The results of this study show that changes in the solar zenith angle have no significant effects on the predictive capability of REPs extracted by the linear extrapolation technique. But the method is sensitive to sensor noise. The comparative advantage of the linear extrapolation method over other methods is lost with noise levels greater than 0.05 % standard deviation relative to the reflectance. The sensitivity to spectral noise is attributed to the fact that the linear extrapolation method requires derivative analysis (first difference transformation of the spectrum). Derivative analysis enhances absorption features and suppresses contribution of non-vegetative reflectance components (Boochs et al., 1990), but is sensitive to the inherent spectral noise of the system (Broge and Leblanc, 2000). We have not yet tested the use of smoothing for reducing sensor noise prior to REP calculation. The advantage of using the linear extrapolation I method compared to the linear interpolation

method diminishes with decreasing spectral resolution, confirming results obtained by Cho and Skidmore (2006).

3.5 Conclusions

This study has shown that REPs extracted by the linear extrapolation method involving wavebands at 680, 694, 724 and 760 nm have the potential for maximally explaining variations in leaf chlorophyll content with minimal effects of leaf and canopy biophysical confounders such as LAI, leaf inclination distribution, leaf dry mass content and soil brightness compared to traditional techniques including the linear interpolation, inverted Gaussian and polynomial fitting techniques. In addition, the linear extrapolation method is insensitive to changes in solar zenith angle. However, the advantage of using the linear extrapolation method compared to the various alternative REP algorithms diminishes with increasing sensor noise and decreasing spectral resolution. In summary, the linear extrapolation technique shows high potential for chlorophyll estimation with radiative transfer models. The efficacy of the technique under field conditions needs to be established.

Acknowledgement

We wish to thank the International Institute for Geo-information Science and Earth Observation (ITC) for funding this study. We are equally thankful to Dr. J.G.P.W Clevers for his useful advice.

Chapter 4

Discriminating species using hyperspectral indices at leaf and canopy scales

This chapter is based on
Cho, M.A., Sobhan, I., Skidmore, A.K. and de Leeuw, J., In preparation. Discriminating species using hyperspectral indices at leaf and canopy scales

Abstract

Developments in hyperspectral remote sensing have provided new indices or indicators of biochemical and biophysical properties. Most of the studies involving the novel spectral indices have been conducted at the leaf scale and have been rarely investigated for species discrimination. The objectives of the study were to determine hyperspectral indices that (i) are likely to be influenced by change in spectral measurement from the leaf to the canopy scale and (ii) can discriminate species at both scales. Leaf and canopy reflectance measurements were made from six species (3 shrubs, 3 trees) using an ASD spectroradiometer. The two-sample t test was used to evaluate whether significant differences exist between leaf and canopy indices, while differences between species pairs (15 pairs) were evaluated with ANOVA and pair-wise Bonferroni adjusted t tests. The hyperspectral indices evaluated in this study were, in general, sensitive to the change in spectral measurement scale from the leaf to the canopy. However, among the indices studied, red-edge positions (REP) extracted by the linear extrapolation I method were least sensitive to the change in measurement scale as three out of the six species showed no significant differences between the leaf and canopy indices. With respect to species discrimination, the canopy indices were better discriminators than the leaf indices. This is essential for air- or spaceborne remote sensing of species assemblages. The photochemical reflectance index (PRI) showed the highest potential to discriminate species at the canopy scale (all 15 pairs), while the linear extrapolation REPs showed the highest potential to discriminate the same species pairs (10 pairs) at both scales. Hyperspectral indices might provide new possibilities of differentiating plant species.

4.1 Introduction

Developments in hyperspectral remote sensing have provided more accurate information on structural, biochemical and physiological properties of vegetation (Blackburn, 1998). Most the work on hyperspectral remote sensing of biophysical and biochemical parameters has been achieved through the development of new hyperspectral indices (Chappelle et al., 1992; Vogelmann et al., 1993; Carter, 1994; Gitelson and Merzlyak, 1997). Spectral indices are mathematical transformations of spectral reflectance to enhance vegetation signal (Huete and Jackson, 1988; Qi et al., 1995). Hyperspectral indices might provide new possibilities of differentiating plant species or communities that differ in canopy structure and/or biochemical compositions (Nagendra, 2001).

The normalised difference vegetation index (NDVI) is the most commonly used multi-spectral index of canopy greenness, a correlate of structural aspects such as canopy cover and leaf area index. NDVI calculation is based on the difference in canopy reflectance at red (670-680 nm) and near-infrared (750–850 nm) wavelengths (Rouse et al., 1974; Tucker, 1979). NDVI has been used to differentiate communities consisting of structurally distinct formations e.g. savannah, shrubland and dense forest (Achard and Blasco, 1990; Saney and Elliott, 2002) or phenological distinct formations e.g. deciduous versus evergreen forests (Achard and Estreguil, 1995; Van Wagtendonk and Root, 2003). However, attempts to differentiate plant species have been unsuccessful because they tend to produce overlapping canopy NDVI values (Nagendra, 2001; Pettorelli et al., 2005).

New hyperspectral indices that provide closer correlation with the biochemical and physiological properties of intact leaves or canopies have been developed. Several studies show that narrowband vegetation indices involving visible reflectance near 550 and 700 nm can precisely estimate leaf pigments such as chlorophyll a and b (Carter, 1994; Chappelle et al., 1992; Vogelmann et al., 1993; Gitelson and Merzlyak, 1997) and carotenoids (Gamon et al., 1992; Gamon and Surfus, 1999; Gitelson et al., 2002). These pigments have different roles within the process of photosynthesis and their concentrations can depend on factors such as phenology, degree of canopy development and type of environmental stress (Blackburn and Pitman, 1999).

Another hyperspectral index of interest is the wavelength of maximum slope in the red-edge (670-780 nm) (Curran et al., 1995; Jago et al., 1999). This point is termed the red-edge position (REP). Changes in leaf chlorophyll content cause shifts in the REP to shorter and longer wavelengths for low and high chlorophyll contents, respectively (Horler et al., 1980; Horler et al., 1983; Miller et al., 1990; Clevers et al., 2002). By using data simulated with radiative transfer models (PROSPECT-SAILH), Cho et al. (2006) showed that REPs located by a novel approach, the linear extrapolation method are more sensitive to leaf chlorophyll content than REPs derived by alternative algorithms, including the linear interpolation (Guyot and Baret, 1988), and inverted Gaussian (Bonham-Carter, 1988) and higher order polynomial fitting (Pu et al., 2003) methods.

Most of the studies involving the novel spectral indices have been conducted at the leaf scale and have not been tested for species discrimination. Fewer studies have been carried out at the canopy scale, for example, using plants grown under controlled conditions (Yoder and Pettigrew-Crosby, 1995), natural canopies (Blackburn, 1998) and simulated data (Barton and North, 2001; Cho et al., 2006). It is questionable whether spectral information acquired at the leaf level can be linearly scaled up to understanding the spectral characteristics at the whole plant or community scale. Several experimental and modelling studies show that vegetation reflectance at the latter scale is not only a function of leaf optical properties but also canopy biophysical attributes (foliage clumping, leaf orientation, leaf area, bark, twigs, flowers), soil reflectance, illumination conditions, viewing geometry and atmospheric conditions (Verhoef, 1984; Kuusk, 1991; Jacquemoud et al., 1995; Yoder and Pettigrew-Crosby, 1995; Asner, 1998). Thus, the main objectives of the study were to determine hyperspectral indices that:

i are likely to be influenced by change in spectral measurement from the leaf to the canopy scale and
ii can discriminate species at both scales.

To achieve the above objectives, leaf and top-of-canopy reflectance measurements were made *in situ* from three species of shrubs and three species of trees. Statistical differences between the leaf and canopy indices and between species pairs were examined using the two-sample student t-test.

4.2 Material and methods

4.2.1 Spectral measurements

Leaf and canopy reflectance spectra of three shrub and three tree species (Table 4.1) were collected on clear sky days (30 August and 2 September 2005) using an ASD spectroradiometer (FieldSpec Pro FR, Analytical Spectral Device, Inc, USA.). The ASD covers the spectral range between 350 to 2500 nm. The sampling interval over the 350-1050 nm range is 1.4 nm with a resolution of 3 nm (bandwidth at half maximum). Over the 1050-2500 nm range, the sampling interval is about 2 nm and the spectral resolution is between 10 and 12 nm. The results are then interpolated by the ASD software to produce readings at every 1nm. A 1.2 m long fibre optic cable with a 25˚ field of view was used for the measurements.

Leaf reflectance measurements were made at about 5 cm above sunlit sides of 20 to 30 leaves on the shrub or tree crowns. A crane was used to attain the crowns of tall trees. With respect to the canopy spectra, 20 to 30 measurements were made at different points above the crown at a distance of 1 m to 1.5 m. Measurements were taken on clear sunny days near solar noon (11 am to 2 pm). The radiance data was converted to reflectance using scans of a white spectralon reference panel. At most two target measurements were made after measuring the reference panel.

Table 4.1 Shrub and tree species used in the study.

Plant species	Structural characteristics of the plants
Hedera helix .	Evergreen creeping plant, the adult plants consist of self-support erect stems with unlobed cordate leaves
Rhododendron sp.	Dense shrub, ~1.5 m, evergreen leaves
Prunus spinosa	Dense prickly shrub, ~ 3 m, deciduous
Corylus avellana.	Tree, ~ 4 m, deciduous
Malus domestica	Tree, ~ 4 m, deciduous
Aesculus hippocastanum	Tree, ~ 3 m, deciduous

4.2.2 Spectral indices

Only the leaf and canopy spectra in visible-NIR (VNIR, 400-900 nm) range were considered in this study because the SWIR region showed high noise levels, particularly in the major water absorption bands. The

VNIR spectra for each species were smoothed using a Savitzky-Golay (Savitzky and Golay, 1964) second order polynomial least-squares function with a five-band window. Vegetation indices and REPs were then computed from the leaf and canopy spectra.

4.2.2.1 Vegetation indices

Apart from the traditional NDVI, narrowband indices that are sensitive to chlorophyll and carotenoids were adopted in this study. See Table 2 for the full description of the vegetation indices.

Table 4.2 Vegetation indices selected in the study.

Vegetation index	Formula	Biophysical significance	Reference
Normalised difference vegetation index (NDVI)	$(R_{830}-R_{670})/(R_{830}+R_{670})$	Canopy greenness, Leaf area index (LAI), fraction of photosynthetically active radiation	Rouse et al., 1974; Tucker, 1979
Carter index (CI)	R_{760}/R_{695}	Chlorophyll content	Carter, 1994
Gitelson and Merzylak index (GMI)	R_{750}/R_{700}	Chlorophyll content	Gitelson and Merzlyak, 1997
Vogelman index (VOG)	R_{740}/R_{720}	Chlorophyll content	Vogelmann et al., 1993
Photochemical reflectance index (PRI)	$(R_{531}-R_{570})/(R_{531}+R_{570})$	Conversion of xanthophylls-cycle pigments, photosynthetic light-use efficiency, LAI	Gamon et al., 1992; Peñuelas et al., 1995; Barton and North, 2001
Carotenoid reflectance index (CRI)	$R_{800}(1/R_{520}-1/R_{550})$	Carotenoids (alpha- and beta-xanthophylls), indicator of plant stress	Getilson et al., 2002

Note: R = reflectance

4.2.2.2 Red-edge position (REP)

REPs were derived by the linear four-point interpolation approach (Guyot and Baret, 1988), inverted Gaussian modelling (Bonham-Carter, 1988), polynomial fitting technique (Pu et al., 2003) and the linear extrapolation method (Cho and Skidmore, 2006).

(i) Linear interpolation technique

The linear interpolation method (Guyot and Baret, 1988) assumes that the reflectance curve at the red edge can be simplified to a straight line centred around the midpoint between the reflectance in the NIR at about

780 nm and the reflectance minimum of the chlorophyll absorption feature at about 670 nm. It uses four wavebands (670, 700, 740 and 780 nm), and the REP is determined by using a two-step calculation procedure.

(a) Calculation of the reflectance at the inflexion point (R_{re})

$$R_{re} = (R_{670} + R_{780})/2 \tag{4.1}$$

where R = reflectance

(b) Calculation of the red edge wavelength or red edge position (REP)

$$REP = 700 + 40\left(\frac{R_{re} - R_{700}}{R_{740} - R_{700}}\right) \tag{4.2}$$

where 700 and 40 are constants resulting from interpolation in the 700-740 nm interval.

(ii) Inverted Gaussian fitting technique
An inverted Gaussian (IG) model (Bonham-Carter, 1988; Miller et al., 1990) was fitted to the spectral reflectance in the 660-780 nm band range. Accordingly, the IG model (Eq. 4.3) represents the red edge by the reflectance equation:

$$R(\lambda) = R_s - (R_s - R_o)\exp\left(-\frac{(\lambda_0 - \lambda)^2}{2\sigma^2}\right) \tag{4.3}$$

where R_s = maximum spectral reflectance
R_o = minimum spectral reflectance
λ_o = wavelength of minimum reflectance
σ = Gaussian function variance.

The REP is then defined as:

$$REP = \lambda_0 + \sigma \tag{4.4}$$

We used an iterative optimisation fitting procedure to determine parameters of the IG model (Miller *et al*. 1990). Initial guesses of the

model parameters were made after review of each data set. Typically, R_o was set at 670 and the σ was approximated at about 30 nm. The IG model employs a least-square criterion to fit a normal curve to the reflectance red edge. The values of λ_o, R_o, R_s and σ are then determined by the fitting procedure.

(iii) Polynomial fitting technique
A polynomial (Pu *et al*. 2003) function (e.g. 3rd order polynomial - Eq. 4.5) was fitted to the reflectance spectrum between the wavelengths, corresponding to the minimum reflectance in the red and the maximum NIR (shoulder) reflectance.

$$R(\lambda) = a_0 + \sum_{i=1}^{3} a_i \lambda^i \tag{4.5}$$

where λ = band between 670 nm to 780 nm.

Subsequently, REP was determined from the maximum first derivative spectrum. The first derivative was calculated using a first-difference transformation of the reflectance spectrum (Dawson and Curran, 1998) as follows:

$$FDR_{(\lambda i)} = (R_{\lambda(j+1)} - R_{\lambda(j)})/\Delta\lambda \tag{4.6}$$

where FDR is the first derivative reflectance at a wavelength i, midpoint between wavebands j and j+1, $R_{\lambda(j)}$ is the reflectance at the j waveband, $R_{\lambda(j+1)}$ is the reflectance at the j+1 waveband, and Δ_λ is the difference in wavelengths between j and j+1.

(iv) Linear extrapolation method
The linear extrapolation method (Cho and Skidmore, 2006) is designed to track changes near chlorophyll sensitive peaks in the first derivative of the reflectance red-edge i.e. around 700 and 725 nm (Horler et al., 1983) and to mitigate the destabilising effects of multiple peaks feature on the red-edge derivative spectra on the location of the REP. The spectra of the six species used in this study showed multiple peaks around 700, 720, 730 and 760 nm (Fig.4.1). Smith et al. (2004) observed similar peaks in grass spectra. Zarco-Tejada et al. (2003) observed the multiple-peak feature and showed that it is a function of natural fluorescence emission at 690 and 730 nm.

According to the linear extrapolation approach, the REP is calculated as the wavelength at the intersection of two straight lines (Eq. 4.7 and 4.8) extrapolated through two points on the far-red flank and two points on the NIR flank of the first derivative spectrum. Cho and Skidmore (2006) identified two combinations of wavebands for calculating leaf nitrogen-sensitive REPs. They are hereby termed as linear extrapolation I involving far-red 680 and 694 nm in combination with NIR 724 and 760 nm, and linear extrapolation II involving far-red 680 and 694 nm in combination with NIR 732 and 760 nm.

Far-red line: $FDR = m_1\lambda + c_1$ (4.7)

NIR line: $FDR = m_2\lambda + c_2$ (4.8)

Where, FDR denotes the first derivative value, and m and c represent the slope and intercept of the straight lines. At the intersection, the two lines have equal λ (wavelength) and FDR values. Therefore, the REP, which is the λ at the intersection, is given by:

$$REP = \frac{-(c_1 - c_2)}{(m_1 - m_2)} \quad (4.9)$$

where c_1 and c_2 are the intercepts, and m_1 and m_2 are the slopes of the far-red and NIR lines, respectively.

4.2.3 Data analysis

The two-sample t-test for testing whether differences exist between two population means was adopted in this study to determine spectral indices that are likely to be influenced by the canopy effect. Numerous studies have shown that the two-sample t test is robust to considerable departures from its theoretical assumptions (that both samples come at random from normal populations with equal variances), especially if the sample sizes are equal or nearly equal (Boneau, 1960; Cochran, 1947; Posten et al., 1982; Zar, 1996). We tested the research hypothesis that the means of the leaf and canopy indices for each species were different, i.e., Ho: $\mu_1 = \mu_2$ versus the alternative hypothesis, H_1: $\mu_1 \neq \mu_2$, where μ_1 and μ_2 are the means of leaf and canopy indices, respectively. The test was conducted

for each species using the various spectral indices. The t values were calculated using Eq. 4.10.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{sd_1^2}{n_1} + \frac{sd_2^2}{n_2}}} \tag{4.10}$$

Where, $\bar{X}_1$ and $\bar{X}_1$, sd_1 and sd_2, and n_1 and n_2 represent the means, standard deviations and sample sizes of the leaf and canopy data, respectively.

A two-step procedure was adopted in order to evaluate the potential of the various indices to discriminate between species. First, single factor analysis of variance (ANOVA) was used to test whether differences exist between the species means: the null hypothesis, H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$ versus the alternative hypothesis, H_1: $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq \mu_5 \neq \mu_6$. Secondly, a multiple comparisons test using Bonferroni adjusted t test was carried out in order to determine which pairs of species means differ. Bonferroni adjusted test reduces the chance of committing Type I error (Zar, 1996). We applied the Bonferroni multiple comparisons procedure with $\alpha = 0.05$ to the data. The alpha level was adjusted downwards by dividing 0.05 by 15 (number of species pairs) i.e. 0.05/15 = 0.003. The critical t for this value is 3.26 for a sample size of n = 20 to 30.

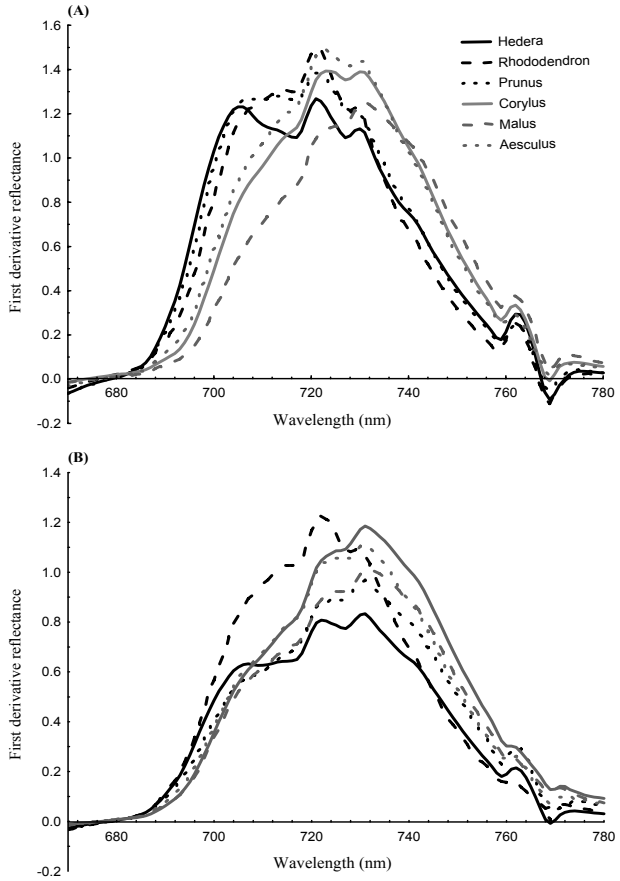


Fig.4.1 Mean first derivative reflectance for (A) leaf and (B) top-of-canopy level measurements for six plant species showing multiple red-edge peaks around 700, 720, 730 and 760 nm.

4.3 Results

4.3.1 Differences between leaf and canopy indices

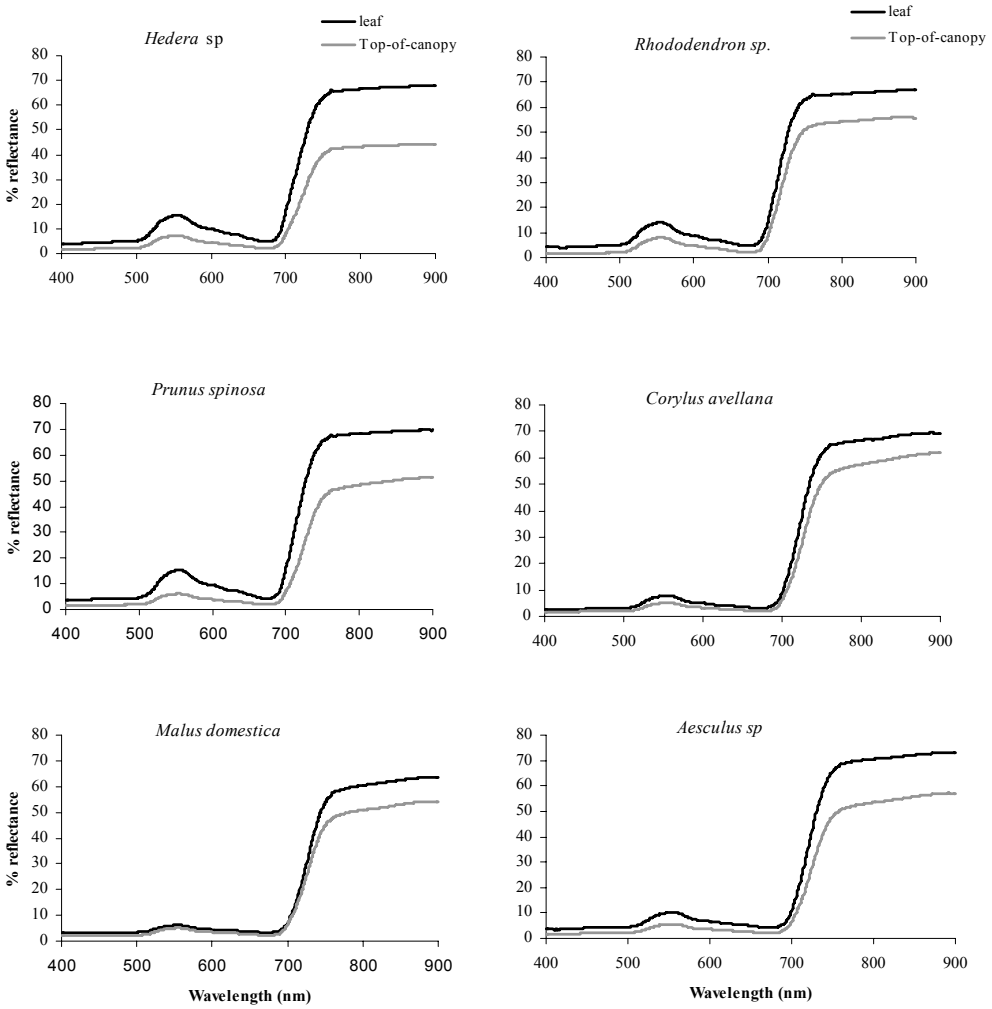


Fig.4.2 Mean leaf and canopy reflectance for six plant species. Spectral measurements were carried out *in situ*, on sunlit sides of the leaves and 1-1.5 m above the canopy for leaf and canopy measurements, respectively.

The leaf VNIR reflectances were higher than canopy reflectances for all six species (Fig. 4.2). The question as to whether the differences were significant for each band was tackled using the two-sample t test. The differences were statistically significant ($p < 0.05$) in all the VNIR bands for all six species, but for *Malus* in the wavelength region between 703-718 nm (Fig.4.3).

The descriptive statistics of the spectral indices have not been presented, but it can be inferred from the negative t values (Tables 4.3 and 4.4) calculated using Eq. 4.10 that the canopy means were higher than the leaf means. These results contradict those of the reflectance data. There were a few exceptions e.g. for most cases of *Malus* where the leaf means were higher that the canopy means. It is unclear why *Malus* showed the odd behaviour. The leaf-scale data showed higher variability compared to the canopy-scale data for each species as illustrated with NDVI and linear extrapolation I REP using *Rhododendron* (Fig. 4.4). The results of the two-sample t test showed that the differences between leaf and canopy means were significant ($p < 0.05$) in 81% and 74% of the cases for vegetation indices and REPs, respectively. However, when the individual indices were compared, the linear extrapolation I REP showed the highest number cases where the differences were not significant (3 species) followed by the linear extrapolation II REP, Carter index, and Getilson and Merzylak index with two cases each.

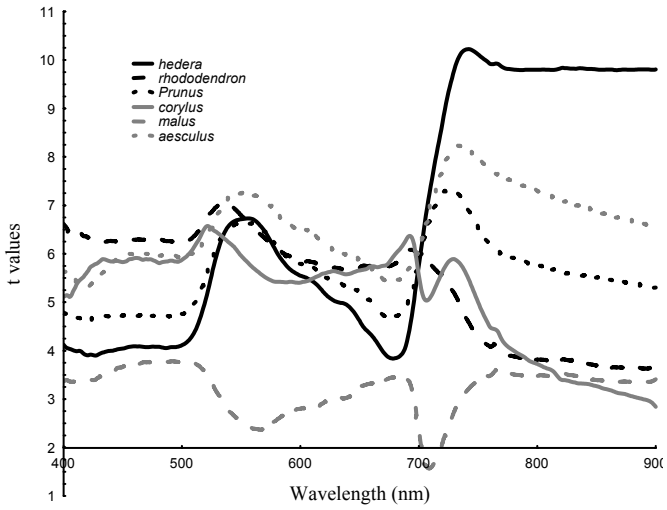


Fig.4.3 Results of two-sample t test for differences between leaf and canopy reflectance for all visible and NIR bands. The wavelength axis cuts the t-values axis at t = 2. Above this critical t value, the difference between the leaf and canopy means is significant ($p < 0.05$).

Table 4.3 Two-sample t-test for differences between leaf and canopy vegetation indices

Species	NDVI	CI	GMI	VOG	PRI	CRI
Hedera	-2.22*	-0.90ns	-1.24ns	-3.38**	-4.95**	-0.44ns
Rhododendron	-7.43**	-8.73**	-7.40**	-5.66**	7.44**	-10.98**
Prunus	-4.11**	-4.00**	-4.45**	-7.94**	-5.26**	-3.39**
Corylus	-8.03**	-8.45**	-4.86**	-3.41**	2.22*	-9.11**
Malus	-2.02*	-0.5ns	0.88ns	2.33*	6.47**	-3.94**
Aesculus	-4.78**	-4.2**	-3.49**	-5.08**	1.34ns	-5.23**

*= $p<0.05$, ** = $p<0.01$, ns = not significant ($p>0.05$)

Table 4.4 Two sample t-test for differences between leaf and canopy red-edge position calculated by various methods.

Species	Linear interpolation	Linear extrapola-tion I	Linear extrapola-tion II	Inverted Gaussian modelling	Polynomial fitting
Hedera	-6.28^{**}	-2.76^{*}	-3.65^{**}	-6.24^{**}	-4.94^{**}
Rhododendron	-6.98^{**}	-1.48^{ns}	-2.17^{*}	-5.64^{**}	-4.46^{**}
Prunus	-11.83^{**}	-4.60^{**}	-5.99^{**}	-11.26^{**}	-10.25^{**}
Corylus	-9.45^{**}	0.73^{ns}	-0.21^{ns}	-7.57^{**}	-7.62^{**}
Malus	1.22^{ns}	6.16^{**}	5.88^{**}	1.64^{ns}	1.83^{ns}
Aesculus	-9.75^{**}	-0.56^{ns}	-1.92^{ns}	-8.56^{**}	-6.84^{**}

* = $p<0.05$, ** = $p<0.01$, ns = not significant ($p>0.05$)

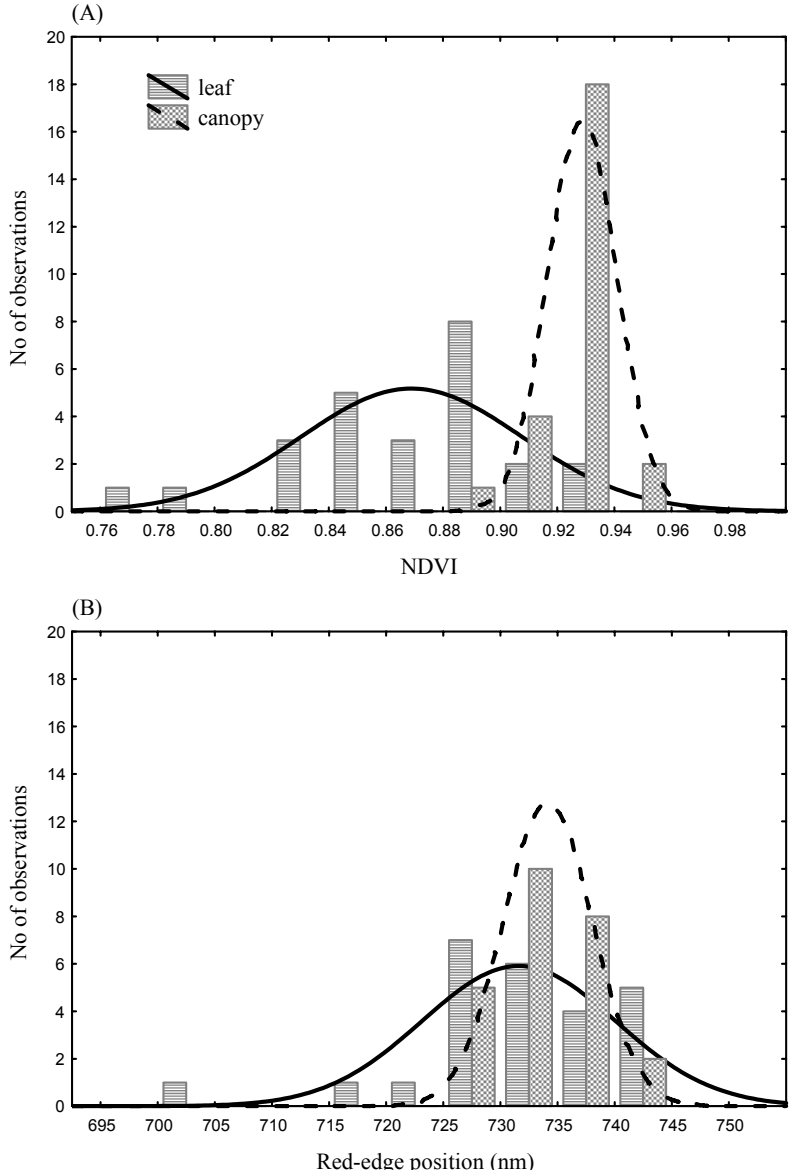


Fig.4.4. Comparing the leaf and canopy distributions for (A) NDVI and (B) red-edge positions derived by the linear extrapolation II method for *Rhododendron.*

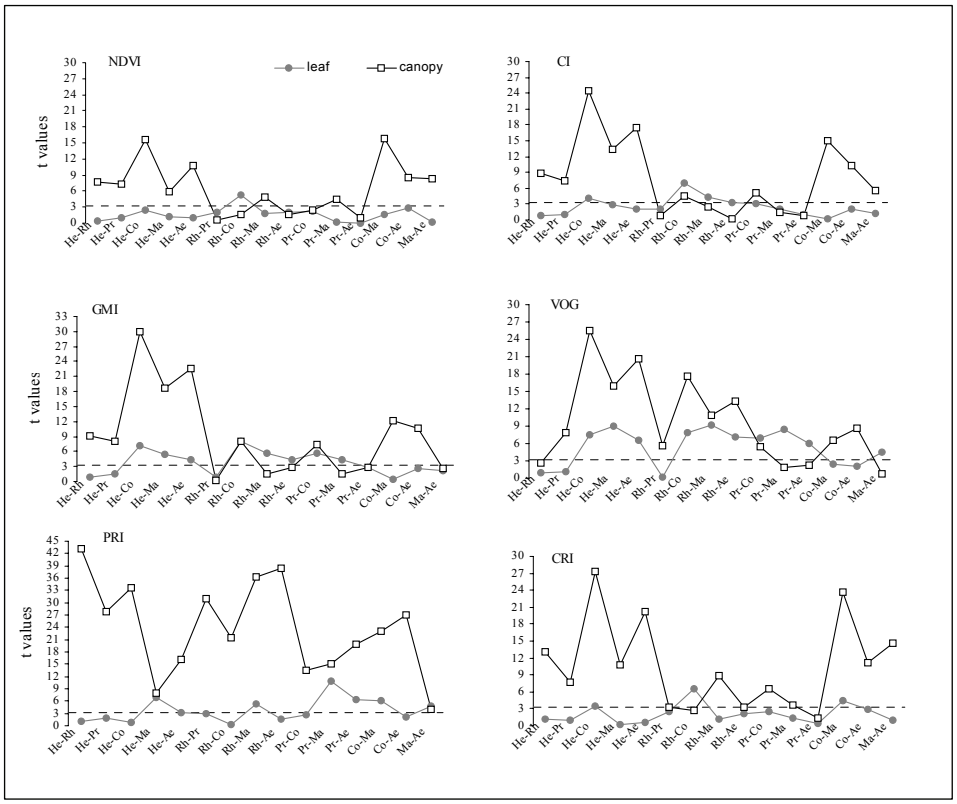


Fig.4.5. Results of two-sample t tests for differences between species (15 pairs) at leaf and canopy scale using NDVI, Carter index (CI), Gitelson and Meryzlak index (GMI), Vogelman index (VOG), photochemical reflectance index (PRI) and carotenoid reflectance index (CRI). Broken lines denote critical t value (t = 3.26) after Bonferroni adjustment above which differences were significant. *Hedera (He), Rhododendron (Rh), Prunus (Pr), Corylus (Co), Malus (Ma) and Aesculus (Ae).*

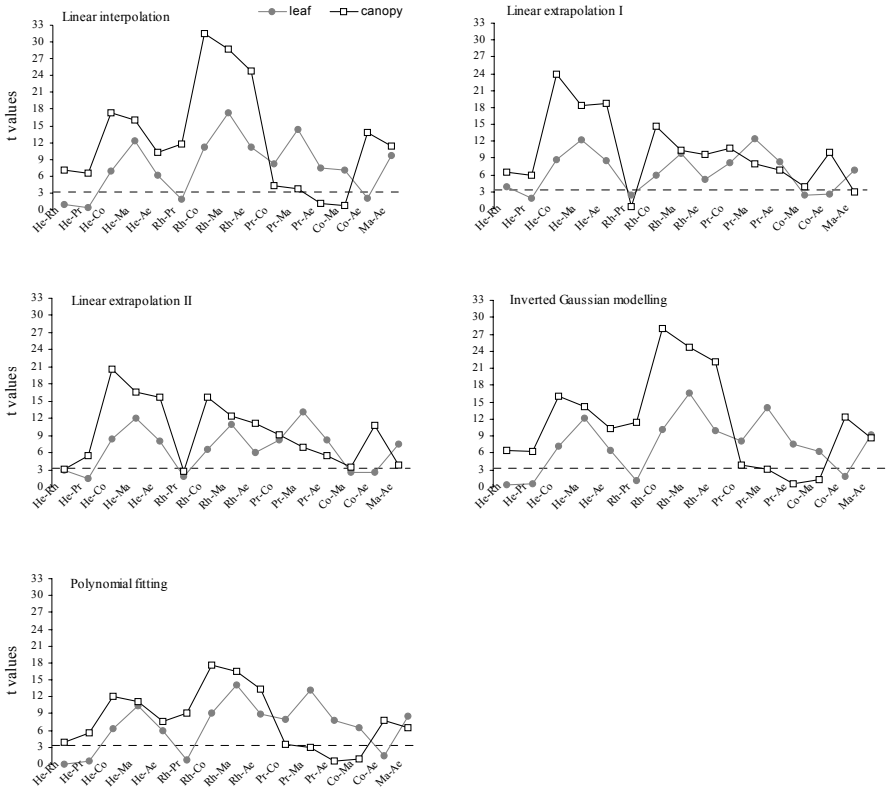


Fig.4.6. Results of two-sample t tests for differences between species (15 pairs) at leaf and canopy scale using red-edge positions extracted using linear interpolation, linear extrapolation I, linear extrapolation II, inverted Gaussian modelling and polynomial fitting methods. Broken lines denote critical t value (t = 3.26) after Bonferroni adjustment above which differences were significant. *Hedera (He), Rhododendron (Rh), Prunus (Pr), Corylus (Co), Malus (Ma) and Aesculus (Ae).*

4.3.2 Discriminating species

The results of the one-way analysis of variance (ANOVA) showed significant differences between the species means for all the spectral indices using the leaf and canopy-scale data, i.e. the null hypothesis, H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$ was rejected for all the indices. P-values less than 0.0001 were obtained for all the tests except for the NDVI using leaf-scale data, which showed a p-value of 0.0139. The results of the multiple comparison test using Bonferroni adjusted t test subsequently showed which pairs of means differ (Fig. 4.5 and 4.6).

More species pairs were differentiated using canopy-scale data than the leaf-scale data (Table 4.5). The potential for NDVI, PRI or CRI to discriminate species was highly biased towards the canopy-scale. The above indices showed the highest differences between the number of separable pairs at the leaf and canopy scales. For example, all 15 species pairs could be differentiated at the canopy level using PRI as against 5 pairs at the leaf scale. The histograms of leaf and canopy PRI in Fig. 4.7(A) provide a visual appreciation of its species discrimination capability at both levels. The NDVI showed the lowest potential to discriminate species at the leaf level, being able to differentiate only a single pair. GMI and VOG were the best vegetation indices at both leaf and canopy scales.

In general, REPs performed better than vegetation indices in discriminating species at both scales. When all indices are compared, REPs extracted by the linear extrapolation I and II showed the highest potential in discriminating the same species pairs at both scales (10 pairs). This is further illustrated with the histograms of the linear extrapolation I REPs in Fig 4.7(B).

A general species separability pattern based on the phenological characteristics appears to emerge at the canopy scale for VOG and REPs. There were lower canopy t values for pairs consisting of species of the same phenology i.e. evergreen vs. evergreen (*Hedera-Rhododendron*) or deciduous vs. deciduous (*Prunus-Corylus*, *Prunus-Malus*, *Prunus-Aesculus* and *Corylus-Malus*) in contrast to species of opposing phenology i.e. evergreen vs. deciduous (*Hedera* or *Rhododendron* – *Prunus, Corylus, Malus* or *Aesculus*). See Fig 4.8. for an illustration of the above phenomenon. There were some few exceptions where species of opposing phenology were weakly discriminated at the canopy scale e.g. *Hedera-Prunus* and *Rhododendron-Prunus*.

Table 4.5 Summary of two-sample t tests for differences between species (15 pairs in total), showing number of pairs of species significantly discriminated ($t > 3.26$, $p < 0.003$) at the leaf, canopy, and at both scales.

Spectral index	Number of significant cases		
	Total at leaf scale	Total at canopy scale	Same species pairs at both scales
Vegetation indices			
NDVI	1	10	0
CI	4	10	2
GMI	8	9	5
VOG	10	11	7
PRI	5	15	5
CRI	3	13	2
Red-edge position			
Linear interpolation	11	13	9
Linear extrapolation I	11	13	10
Linear extrapolation II	10	13	10
Inverted Gaussian modelling	11	12	8
Polynomial fitting	11	12	8

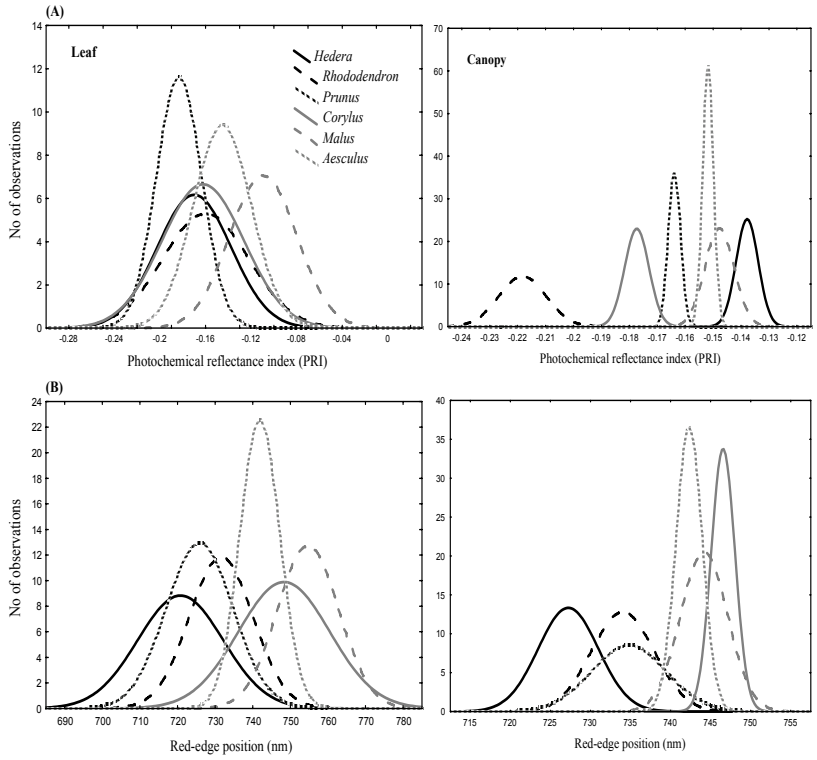


Fig.4.7 Histograms of leaf and canopy indices, namely (A) Photochemical reflectance index (PRI) and (B) linear extrapolation I REP for six species of plants. The histograms illustrate the ability of the indices to differentiate species at the leaf and canopy scales.

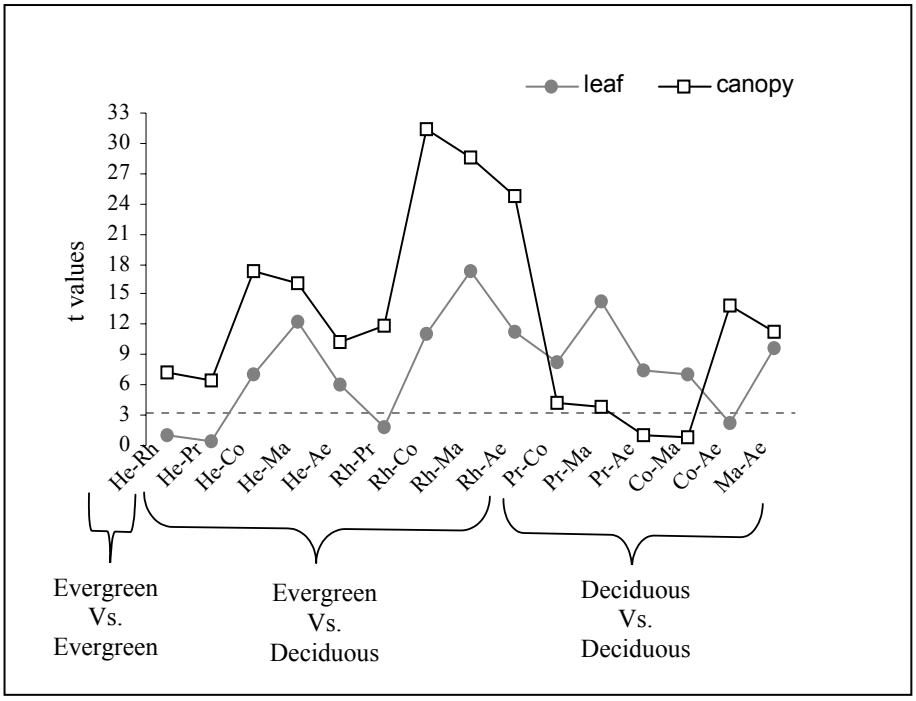


Fig.4.8 Results of two-sample t tests for differences between species (15 pairs) at leaf and canopy scales using red-edge positions extracted using linear interpolation method. Species of opposing phenology (evergreen-deciduous) are better discriminated than species of the same phenology. *Hedera (He), Rhododendron (Rh), Prunus (Pr), Corylus (Co), Malus (Ma) and Aesculus (Ae).*

4.4 Discussion

4.4.1 Differences between leaf and canopy indices

The results of this study revealed systematically higher VNIR reflectances at the leaf scale than at the top-of the canopy. The higher leaf VNIR reflectance may be explained by the effect of multiple scattering caused by leaf stacking since the leaf reflectance were measured *in situ*. Blackburn (1999) showed that the NIR and to a lesser degree, the visible reflectance increases with leaf stacking. He equally argues that the spectral reflectance properties of background materials and areas of shadow can have large influence upon that of the whole canopy even when there is complete canopy. For example, Fig. 4.9 shows

canopy pictures of *Hedera* and *Rhododendron* with dark areas, which may be due to shadow cast by the uppermost leaves.

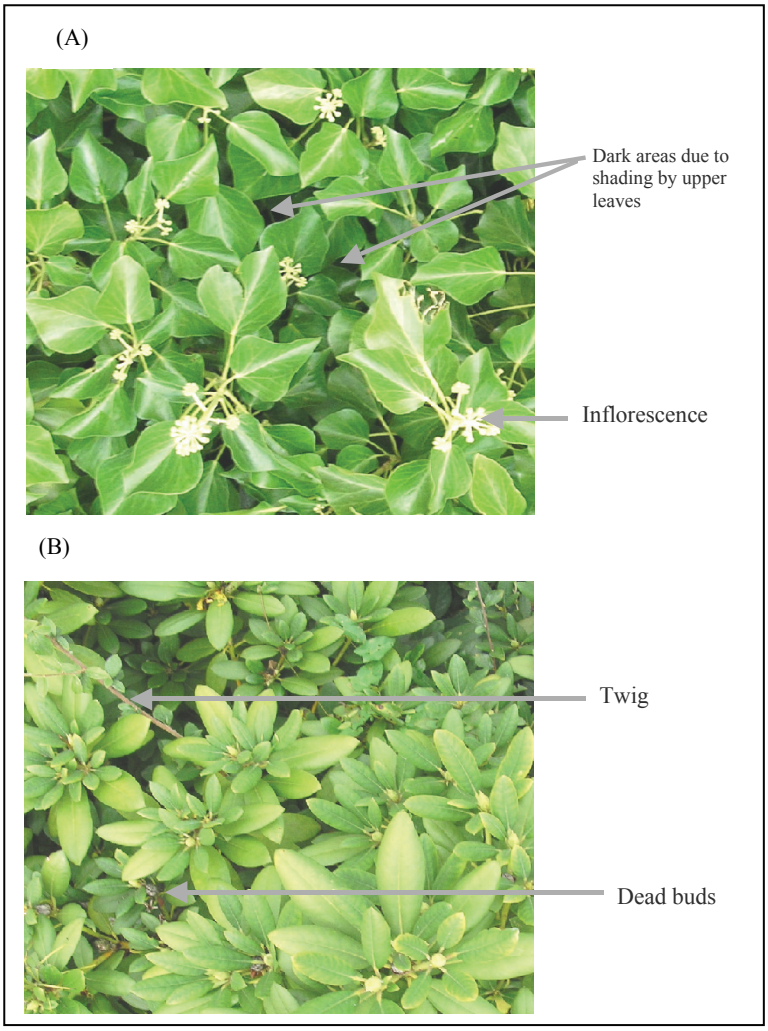


Fig.4.9 Top-of-canopy view of (A) *Hedera* and (B) *Rhododendron.*

The results of this study equally showed significant differences between leaf and canopy indices in 81% and 74% of the cases for vegetation indices and REPs, respectively. Thus, the information contents at both levels are largely different. The change in the spectral information content from the leaf to the canopy scale could be due to differences introduced by the complexity of the canopy, e.g. LAI, foliage clumping and the presence of twigs, flowers and shadow. However, the linear

extrapolation I REP appears to be the least sensitive index to these canopy properties followed by the linear extrapolation II REP, Carter index, and Getilson/Merzylak index. These indices are all chlorophyll content indices. The results of this study support growing evidence that REP extracted by the linear extrapolation method might be less sensitive to canopy structural. For example, by using data simulated with radiative transfer models (PROSPECT-SAILH), Cho et al. (2006) showed that REPs located by the linear extrapolation method are more sensitive to leaf chlorophyll content with minimal effect of LAI and leaf mass compared to REPs located by various alternative algorithms. Cho and Skidmore (2006) in an experimental study using leaf stacks showed that REPs located by the linear extrapolation approach were more sensitive to leaf nitrogen concentration than the various REP alternatives.

Other factors that might have affected the canopy spectra include atmospheric conditions and the bidirectional reflectance (BRDF) effect caused by varying view and solar zenith angles. The ratio or vegetation indices are designed to minimise these effects and to enhance the spectral signal of leaf and canopy biochemical and biophysical properties. The impact of the above perturbing factors on NDVI has long been established (Huete and Jackson, 1988; Kaufman and Tanré, 1992; Qi et al., 1995). Only recently was the impact of the BRDF effect on PRI apparent. Barton and North (2001) using simulated data showed that LAI has a high impact on PRI values followed by changing solar and view zeniths. On the other hand, Clevers et al. (2001) demonstrated that REP are less sensitive to atmospheric conditions and Cho et al. (2006) showed that REPs are not sensitive to varying solar zenith angles.

4.4.2 Discriminating species

In this study, we have shown that species were more easily discriminated at the canopy than at the leaf scale. This conclusion held across a variety of hyperspectral indices. This is essential for air-spaceborne remote sensing of species assemblages. It is possible that the optimum spectral information required to discriminate species at the leaf level was not captured in the leaf samples. This could be explained by the high variability in the leaf indices. Hence, the poorer species separability results at the leaf scale. On the other hand, in addition to the possibility of covering the total spectral information among the leaves, canopy reflectance might provide extra information on the canopy structure (leaf

orientation, leaf clumping, and colour of twigs and flowers), which might enhance the ability of the indices to discriminate between species.

The impact of the canopy characteristics seems to be highest on NDVI, PRI and CRI, which showed the highest differences between the number of separable species pairs at the leaf and canopy scales. Though Gamon et al (1992) proposed the PRI as a sensitive index to xanthophyll cycle pigment contents; Barton and North (2001) showed that it is highly sensitive to canopy structural properties (LAI and leaf angle distribution). This double property of the PRI might have accounted for the high species separability potential at the canopy scale. A drawback of the PRI is that it is strongly influenced by soil background (Barton and North, 2001). However, soil background was not an important factor in this study. Finally, the results of this study show that the REP largely preserves leaf information for discriminating species when the reflectance measurement is scaled up the canopy, with the linear extrapolation REPs having a slight urge over alternative REP algorithms. However, care should be taken when applying the linear extrapolation method because Cho et al. (2006) showed that it is sensitive to spectral noise. We recommend smoothing of the spectrum when noise is a problem.

4.4.3 Implications for upscaling leaf level information to the canopy scale

The results of this study support experimental and modelling studies, which demonstrate poor signal propagation from the leaf to canopy scale (Verhoef, 1984; Kuusk, 1991; Jacquemoud et al., 1995; Yoder and Pettigrew-Crosby, 1995; Asner, 1998). But the significant finding in this study is that canopy indices have a far superior discriminating power than leaf level indices, which is essential for remote sensing of species at the ecosystem level. Moreover, the study shows that the REP provides the best chance for upscaling leaf level information on species discrimination to the canopy scale. Since leaf chlorophyll content was not measured in this study, it remains to be explained why the REP showed a higher ability to discriminate species at both scales than ratio-based vegetation indices.

4.5 Conclusions

This study, although limited in data set, allowed an evaluation of the effects of upscaling reflectance measurements from individual leaves to the top-of-canopy on hyperspectral indices. The conclusion from this study is that spectral indices are generally sensitive to the change in scale of spectral measurement from the leaf to the canopy. However, among the spectral indices studied, the linear extrapolation I REP is least sensitive to the change in measurement scale.

Differences between leaf and canopy indices appear to affect the ability of the spectral indices to differentiate species at both levels. The canopy indices were better discriminators of species than the leaf indices. This is essential for air- or spaceborne remote sensing of species assemblages. The PRI showed the highest potential to discriminate species at the canopy scale. But the REP in general showed the highest potential to discriminate the same species pairs at both scales. Hyperspectral indices might provide new possibilities of differentiating plant species or communities.

Acknowledgments

The International Institute for Geo-information Science and Earth Observation (ITC) provided financial support for this study. We would equally like to thank Mrs Eva Skidmore for logistic support during the field experiments.

Chapter 5

Hyperspectral predictors for monitoring biomass production in Mediterranean mountain grasslands: Majella National Park, Italy

This Chapter is based on
Cho, M.A. and Skidmore, A.K. (*In review)* Hyperspectral predictors for monitoring biomass production in Mediterranean mountain grasslands: Majella National Park, Italy. *International Journal of Remote Sensing.*

and

Cho, M.A., Sobhan, I.M. and Skidmore, A.K. (2006) Estimating fresh grass/ herb biomass from HYMAP data using the red - edge position. In: Remote sensing and modeling of ecosystems for sustainability III conference proceedings / ed. by Wei Gao, Susan L. Ustin. SPIE, 2006. art no. 629805 (Proceedings of SPIE ; 6298) 9 p.

Abstract

The research objective was to determine robust hyperspectral predictors for monitoring grass/herb biomass production on a yearly basis in the Majella National Park, Italy. HyMap images were acquired over the study area in mid and early July of 2004 and 2005, respectively. The robustness of vegetation indices and red-edge positions (REP) were assessed by: (i) comparing the consistency of the relationships between green grass/herb biomass and the spectral predictors for both years and (ii) assessing the predictive capabilities of linear regression models developed for 2004 in predicting the biomass of 2005 and vice versa. Frequently used normalised difference vegetation indices (NDVI) computed from red (665-680 nm) and near-infrared bands, modified soil adjusted index (MSAVI), soil adjusted and atmospherically resistant index (SARVI) and water difference vegetation index (NDWI) were highly correlated with biomass ($R^2 \geq 0.50$) only for 2004 when the vegetation was in the early stages of senescence. Although high correlations ($R^2 \geq 0.50$) were observed for NDVI involving far-red bands at 725 and 786 nm for 2004 and 2005, the predictive regression model for each year produced a high prediction error for the biomass of the other year. Conversely, predictive models derived from REPs computed by the three-point Lagrangian interpolation and linear extrapolation methods for 2004 yielded a lower prediction error for the biomass of 2005, and vice versa, indicating these approaches are more robust than NDVI. The results of this study are important for selecting hyperspectral predictors for monitoring annual changes in grass/herb biomass production in the Mediterranean mountain ecosystems.

5.1 Introduction

Monitoring grass biomass through time can provide important information about the stability of natural ecosystems and whether significant changes are taking place (Jensen, 2000). Remote sensing techniques have been widely used to model the spatial and temporal variability of grass biomass over large areas (Richardson et al., 1982; Everitt et al., 1989; Anderson et al., 1993; Wylie et al., 2002).

The most widely used remote sensing product for biomass estimation is the normalised difference vegetation index (NDVI) developed by Rouse et al. (1974). NDVI is commonly computed from canopy reflectance in the red and near-infrared (NIR) using broadband imagery such as NOAA advanced very high resolution radiometer (AVHRR). However, there are major limitations with the NDVI despite its wide application. Several studies show that broad-band NDVI can be unstable, varying with soil colour, canopy structure, leaf optical properties and atmospheric conditions (Huete and Jackson, 1988; Middleton, 1991; Kaufman and Tanré, 1992; Qi et al., 1995; Todd et al., 1998). It has also been demonstrated that empirical models derived from NDVI are highly site and sensor specific and therefore unsuitable for application to large areas or in different seasons (Curran, 1994; Gobron et al. 1997). Furthermore, broadband NDVIs asymptotically approach a saturation level after a certain biomass or leaf area index (Sellers, 1985, Gao et al. 2000).

A major limitation of broadband NDVI is that, it uses average spectral information over broadband widths resulting in loss of critical information available in specific narrowbands (Blackburn, 1998, Thenkabail et al. 2000). However, recent developments in hyperspectral remote sensing have provided additional bands (narrowbands) within the red-NIR transition that have been utilised to improve grass biomass estimation. For example, Mutanga and Skidmore (2004) show that NDVI computed from 746 and 755 nm solves the saturation problem of estimating grass biomass at high canopy cover. Another hyperspectral predictor that has been assessed for grass biomass estimation is the wavelength of maximum slope in the red-NIR region termed the red-edge position (REP) (e.g. Gilabert et al., 1996; Mutanga and Skidmore, 2004). An advantage of the REP over the NDVI is that it is less sensitive to varying soil and atmospheric conditions, and sensor view angle (Curran et al., 1995; Blackburn and Pitman, 1999; Clevers et al., 2001). Many

recent studies assessing the utility of hyperspectral predictors for estimating grass biomass have focused on single crops or species canopies (Thenkabail et al. 2000, Hansen and Schjoerring 2003, Mutanga and Skidmore 2004). The utility of hyperspectral predictors for estimating or monitoring biomass in natural grass and/or herb communities remains to be established.

The Mediterranean mountain grasslands in the region of Abruzzo, Italy consist of mixed grass/herb communities (Conti, 1998). These systems attain peak biomass in summer. But the hot summer climate, and variable cloud presence and precipitation in the region imply that the vegetation and atmospheric conditions are not stable. Therefore, the monitoring of peak grass/herb biomass on an annual basis would require that the relationship between biomass and the spectral predictor remains stable for different summer atmospheric and vegetation conditions.

Thus, the research objective was to determine stable or robust hyperspectral predictors for estimating biomass production in Mediterranean mountain grasslands on a yearly basis. HyMap data was acquired in the study area, the Majella National Park, Italy in mid and early July 2004 and 2005, respectively. The robustness of vegetation indices and REP for monitoring grass/herb biomass was determined by: (i) comparing the consistency of the linear relationship between biomass and hyperspectral predictors for 2004 and 2005 and (ii) assessing the predictive capabilities of empirical models developed for 2004 in predicting the biomass of 2005 and vice versa.

5.2 Material and methods

5.2.1 The study area

The study site is located in Majella National Park, Italy (latitude 41°52' to 42°14'N, longitude 13°50' to 13°14'E), which covers an area of 74095 ha. The park extends into the southern part of Abruzzo, at a distance of 40 km from the Adriatic Sea. This region is situated in the massifs of the Apennines (Conti, 1998). The park is characterised by several mountain peaks, the highest being Mount Amaro (2794 m).

More specifically, the study site (latitude 41°49’ to 42°14’N, longitude 13°57’ to 14°6’E) is situated between Mounts Majella and Morrone to the

east and west, respectively. It covers an area of 40 km by 5.5 km. Gallego Fernández et al. (2004) argue that plant community dynamics in Mediterranean basin ecosystems are driven mainly by alternating episodes of human intervention and land abandonment. For example, abandoned settlement and agricultural areas in Majella are returning to oak (*Quercus pubescens*) woodlands at the lower altitude (400 m to 600 m) and beech (*Fagus sylvatica*) forest at the higher altitude (1200 m to 1800 m). Between these two formations is a landscape composed of shrubby bushes, patches of grass/herb vegetation, and bare rock outcrops. The dominant grass species include *Brachypodium genuense*, *Briza media*, *Bromus erectus* and *Festuca sp*. Herbs include *Helichrysum italicum, Galium verum, Trifolium pratense, Plantago lanceolata, Sanguisorba officinalis* and *Ononis spinosa.*

5.2.2 Field data collection

Two field campaigns were carried out in the summers of 2004 (28 June to July 16) and 2005 (16 to 29 June). There are four main phytosociological classes (semi-natural/farmlands, grazed/periodically flooded areas, open garrigues and abandoned farmlands) of varying areas within the specific study area. The phytosociological map of the park was provided by the park management. Random sampling with clustering was adopted in the study. Twenty-five coordinates (x, y) were randomly generated in four phytosociological classes using ArcGIS software; eight plots in the semi-natural/farmlands/abandoned farmlands, five plots in the open garrigues and twelve plots in the grazed/periodically flooded areas. The number of samples per vegetation class was proportional to the size of the class. A GPS was used to locate the sample plots in the field. A 30 m by 30 m plot size was adopted in this study. For each of the initially randomly generated plot, an extra plot located at about 150 m away was sampled for fresh green grass/herb biomass. Only plots with more than 20% homogeneous grass/herb cover were sampled, resulting in a total of 47 plots. Above-ground biomass was clipped within five randomly selected subplots (1 m by 0.5 m) from each plot. All dry material was removed from the clipped plants before measuring the green grass/herb biomass. Average green grass/herb biomass per plot was calculated from the five subplot measurements.

5.2.3 Image acquisition and pre-processing

Airborne HyMap data of the study site were obtained on 15 July 2004 and 4 July 2005. The flights were carried out by DLR, Germany's Aerospace Research Centre and Space Agency. The HyMap sensor comprised 126 wavebands, operating over the wavelength range 436 nm to 2485 nm, with average spectral resolutions of 15 nm (436 nm to 1313 nm), 13 nm (1409 nm to 1800 nm) and 17 nm (1953 nm to 2485 nm). The spatial resolution of the data was 4 m. The data was collected at solar noon. The specific study site was covered by four image strips, each covering an area of about 40 km by 2.3 km. The solar zenith and azimuth angles for the image strips ranged between 30-33.7^{o} and 111.5-121^{o}, respectively.

The 2004 and 2005 image strips were atmospherically corrected by DLR. But only the 2005 images were geometrically corrected by DLR. The 2004 images strips were then geometrically corrected from the 2005 images using image-to-image registration. The atmospheric correction was carried out using ATCOR4-r (rugged terrain). ATCOR4 is based on MODTRAN-4 radiative transfer code. However, there were differences between the reflectance of similar pixels in the overlapping sections between image strips for the 2005 image. Spectral calibration between strips was carried out using the empirical line method in Environment for Visualising Images (ENVI 4.2) software (Research System, Inc.) in order to minimize the differences. Ten image spectra collected from a reference strip (e.g. strip 2) and corresponding targets from its overlapping neighbour (strip1 or 3) were used to compute a linear regression function for each channel. Using the regression functions, strips 1 and 3 were then adjusted to have a spectral response similar to that of strip 2. The same process was carried out using corrected strip 3 as the reference image to correct strip 4. The spectra were collected from targets such as roads, agricultural fields, quarry fields, and dense beech forest pixels.

5.2.4 Collecting image spectra for grass/herb plots

Grass/herb areas were extracted from the image strips in order to eliminate mixed grass/shrubs and or tree pixels. First, an NDVI image involving bands at 665 nm and 831 nm was computed for each image strip using the ENVI 4.2 software. A point map of the grass/herb plots was then overlaid on the NDVI images. Pixels of pure grass/herb plots

were used to determine minimum and maximum NDVI threshold values for grass/herbs. Next, a grass/herb region-of-interest map was created using the NDVI threshold values. Subsequently, the region-of-interest map was used to subset grass/herb areas from the image. All other pixels were masked out.

A 7 by 7 pixels window was used to collect grass/herb image spectra from each sample plot in order to avoid including pixels located outside the plot (30 m by 30 m). The spectra were collected and averaged. The spectra of five out of the 47 plots were not extracted from the 2005 image strips because the plots were located in portions covered by clouds.

5.2.5 Data analysis

5.2.5.1 Spectral predictors

Two types of spectral predictors were adopted in this study:

(a) Vegetation indices
Four vegetation indices were used in the study: Narrowband NDVIs calculated from all combinations of red or far-red (600 to 740 nm) and NIR (756 to 1000 nm) bands, Modified soil adjusted vegetation index (MSAVI), soil adjusted and atmospherically resistant vegetation index (SARVI) and normalised difference water index (NDWI). The indices are presented in Table 5.1.

Table 5.1 Summary of vegetation indices analysed in this study. R_{NIR} and R_{red} denote reflectances in the NIR and red, respectively.

Index	Formula	Description	References (e.g.)
NDVI (Normalised difference vegetation index).	$(R_{NIR} - R_{red})/(R_{NIR} + R_{red})$	Related to changes in amount of green biomass and pigment content.	Rouse et al. 1974
MSAVI (Modified soil adjusted vegetation index)	$\frac{2R_{NIR} + 1 - \sqrt{(2R_{NIR}+1)^2 - 8(R_{NIR} - R_{red})}}{2}$	Minimises soil influences on canopy spectra. Red and NIR bands at 831 and 665 nm, respectively.	Huete, 1988; Qi et al., 1994.
SARVI (Soil adjusted and atmospherically resistant vegetation index).	$R_{rb} = R_{red} - \gamma(R_{blue} - R_{red})$ γ = atmospheric aerosol correction function $SARVI = (R_{NIR} - R_{rb})/(R_{NIR} - R_{rb} + L)$ L = soil adjustment factor	Blue and red bands at 466 and 665 nm, respectively. $\gamma = 0.9$, $L = 0.5$	Kaufman & Tanre, 1992; Huete et al., 1994.
NDWI (Normalised difference water index)	$(R_{860\ nm} - R_{1240\ nm}) / (R_{860\ nm} + R_{1240\ nm})$	Sensitive to changes in liquid water content of vegetation canopies. Gao (1996) show that NDWI is less sensitive to atmospheric effects than NDVI	Gao, 1996.

(b) Red-edge position (REP)

Red-edge positions were extracted by three simple methods; the linear four-point interpolation (Guyot and Baret, 1988), three-point Lagrangian interpolation (Dawson and Curran, 1998) and the linear extrapolation (Cho and Skidmore, 2006) methods.

(i) Linear four-point interpolation method

The linear four-point interpolation method (Guyot and Baret, 1988) assumes that the reflectance curve at the red edge can be simplified to a

straight line centred near the midpoint between the reflectance in the NIR at about 780 nm and the reflectance minimum of the chlorophyll absorption feature at about 670 nm. It uses four wavebands, 670, 700, 740 and 780 nm i.e. 665, 695, 740 and 786 for the HyMap spectrum. The REP is then determined by using a two-step calculation procedure.
Calculation of the reflectance at the inflexion point (R_{re}):

$$R_{re} = (R_{665} + R_{786})/2 \tag{5.1}$$

where R is the reflectance at a specified wavelength (e.g. 665 nm).

Calculation of the red edge wavelength or red edge position (REP):

$$REP = 695 + 45\left(\frac{R_{re} - R_{695}}{R_{740} - R_{695}}\right) \tag{5.2}$$

695 and 45 are constants resulting from interpolation in the 695-740 nm interval.

(ii) Three-point Lagrangian interpolation method

The three-point Lagrangian interpolation technique (Dawson and Curran, 1998) is designed to locate REP in spectra that have been sampled coarsely. Lagrangian interpolation is applied to the first derivative of the reflectance spectrum, which is computed as follows:

$$D_{(\lambda i)} = (R_{\lambda(j+1)} - R_{\lambda(j)})/\Delta\lambda \tag{5.3}$$

where $D_{(\lambda i)}$ is the first derivative reflectance at a wavelength i, midpoint between wavebands j and j+1, $R_{\lambda(j)}$ is the reflectance at the j waveband, $R_{\lambda(j+1)}$ is the reflectance at the j+1 waveband, and $\Delta\lambda$ is the difference in wavelengths between j and j+1.
The value of the first derivative at any wavelength (i.e. estimated value) will be D_{λ}. The Lagrangian interpolation technique for three known bands is given by

$$D_\lambda = \frac{(\lambda-\lambda_i)(\lambda-\lambda_{i+1})}{(\lambda_{i-1}-\lambda_i)(\lambda_{i-1}-\lambda_{i+1})} D_{\lambda(i-1)} + \frac{(\lambda-\lambda_{i-1})(\lambda-\lambda_{i+1})}{(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})} D_{\lambda(i)} + \frac{(\lambda-\lambda_{i-1})(\lambda-\lambda_i)}{(\lambda_{i+1}-\lambda_{i-1})(\lambda_{i+1}-\lambda_i)} D_{\lambda(i+1)} \quad (5.4)$$

The band having the maximum first derivative will be λ_i, with λ_{i-1} and λ_{i+1} representing the two bands on either side of the maximum derivative. To determine the REP, a second derivation on Eq. 5.4 is performed and resolved for when the second derivative is zero. i.e.

$$REP = \frac{A(\lambda_i+\lambda_{i+1}) + B(\lambda_{i-1}+\lambda_{i+1}) + C(\lambda_{i-1}+\lambda_i)}{2(A+B+C)} \quad (5.5)$$

where

$$A = \frac{D_{\lambda(i-1)}}{(\lambda_{i-1}-\lambda_i)(\lambda_{i-1}-\lambda_{i+1})}, B = \frac{D_{\lambda(i)}}{(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})}, and \quad C = \frac{D_{\lambda(i+1)}}{(\lambda_{i+1}-\lambda_{i-1})(\lambda_{i+1}-\lambda_i)} \quad (5.6)$$

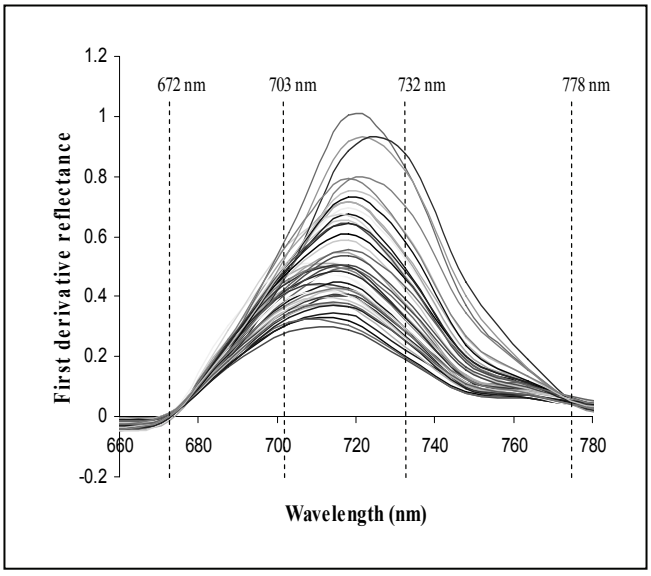


Fig.5.1 First derivative spectra of 2005 sample plots showing bands used in the calculation of red-edge positions by the linear extrapolation method.

(iii) Linear extrapolation technique

The linear extrapolation technique (Cho and Skidmore, 2006) is designed to track changes near chlorophyll sensitive peaks in the first derivative (D) of the red edge i.e. around 700 and 725 nm (Horler et al. 1983). The REP is calculated as the wavelength at the intersection of two straight lines (Eq. 5.7 & 5.8) extrapolated through two points on the far-red flank and two points on NIR flank of first derivative reflectance spectrum. For example, for the HyMap derivative spectra used in this study, the lines were extrapolated through derivative bands at 672 and 703 nm for the far-red line and 732 and 778 nm for the NIR line (Fig. 5.1).

Far-red line: $D = m_1\lambda + c_1$ (5.7)

NIR line: $D = m_2\lambda + c_2$ (5.8)

where m and c represent the slope and intercept of the straight lines; c_1 and m_1 for the far-red line and c_2 and m_2 for the NIR line. At the intersection, the two lines have equal λ and D values. Therefore, the REP, which is the λ at the intersection, is given by:

$$REP = \frac{-(c_1 - c_2)}{(m_1 - m_2)} \quad (5.9)$$

where

$$m_1 = \frac{(D_{703} - D_{672})}{(\lambda_{703} - \lambda_{672})},\ m_2 = \frac{(D_{778} - D_{732})}{(\lambda_{778} - \lambda_{732})},\ c_1 = D_{703} - m_1\lambda_{703},\ c_2 = D_{732} - m_2\lambda_{732} \quad (5.10)$$

5.2.5.2 Assessing the robustness of hyperspectral predictors for monitoring grass/herb biomass

The robustness of the various spectral predictors for monitoring grass biomass was determined in two ways:

(i) the consistency of the linear relations between biomass and the spectral predictors were compared for both 2004 and 2005. The coefficient of determination and prediction errors (the root mean square errors of leave-one-out cross-validation (RMSECV)) were used for the comparison.

(ii) regression models developed for 2004 were used to predict the biomass of 2005 and vice versa. The performances of the various models for predicting either the next or previous years' biomass were compared using the standard errors of prediction (RMSE).

Table 5.2 Green grass/herb biomass data for 2004 and 2005 collected in Majella National Park, Italy.

year	N	Mean (g m^{-2})	SD	Minimum	Maximum
June/July 2004	47	768	366	200	1750
June 2005	42	774	369	210	2010

SD = standard deviation

5.3 Results

5.3.1 Spectral and green grass/herb biomass characteristics for 2004 and 2005

The visible (450-700 nm), NIR (700-1300) and SWIR (1300-2500) reflectances were higher for 2004 than 2005 (Fig. 5.2). These results are consistent with changes that occur when vegetation loses pigmentation and water (Knipling, 1970), e.g. during the early stages of senescence. Furthermore, compared with 2005, the 2004 reflectance spectra showed higher variability (standard deviations) in the chlorophyll (600-700), and leaf/atmospheric water absorption (1450 and 1940 nm) bands (Curran, 1989).

The descriptive statistics for the green grass/herb biomass of 2004 and 2005 are presented in Table 5.2; the data distributions are assumed normal. We used the 2-Sample Student's t-test to compute the confidence interval and perform a hypothesis test for the difference between the means of the biomass of 2004 and 2005. The null hypothesis was H_0: μ_1-$\mu_2 = 0$ versus the alternative hypothesis H_1: μ_1-$\mu_2 \neq 0$, where μ_1 and μ_2 are the mean biomass of 2004 and 2005, respectively. The confidence interval (CI) for the difference in the means at 95% was CI (-161, 149 g m^{-2}). The means were not significantly different at $p<0.05$.

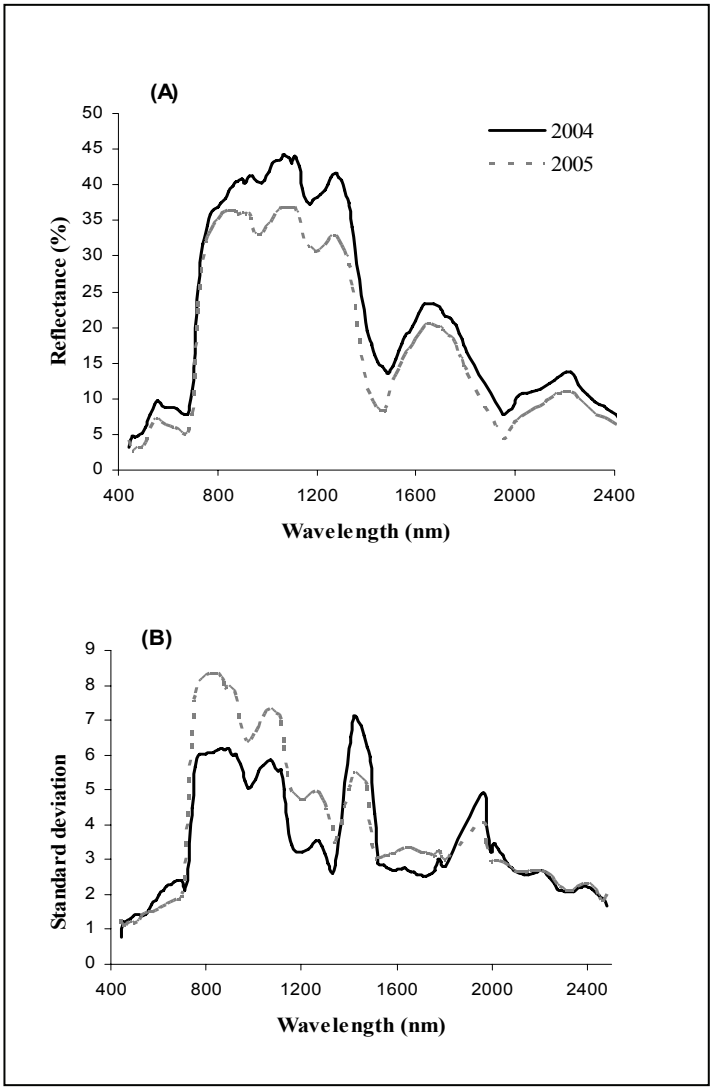


Fig.5.2 Mean reflectance spectra (A) and their corresponding standard deviations (B) for grass/herb plots extracted from HyMap images acquired over Majella National Park, Italy in mid and early July 2004 and 2005, respectively.

Table 5.3 Best NDVI combinations for predicting grass/herb biomass in the Majella National Park, Italy for 2004 and 2005. R^2 = coefficient of determination.

Near-infrared wavelength (nm)	Red or far-red wavelength (nm)	R^2
2004 HyMap image		
786	695	0.56
801	695	0.56
771	695	0.56
756	695	0.56
816	695	0.56
2005 HyMap image		
786	740	0.64
801	740	0.64
771	740	0.62
756	740	0.62
879	725	0.62

Table 5.4 Overall best NDVI combinations for predicting grass/herb biomass in the Majella National Park, Italy for both 2004 and 2005. They are classified according to decreasing difference in the coefficients of determination (R^2) between 2004 and 2005 for combinations that yielded high correlations ($R^2 \geq 50$) for both years.

Near-infrared wavelength (nm)	Red or far-red wavelength (nm)	R^2	
		2004	2005
786	725	0.55	0.58
801	725	0.54	0.59
756	740	0.51	0.62
771	740	0.51	0.62
786	740	0.50	0.64

5.3.2 Predictive performance of vegetation indices

The linear relationships between grass/herb biomass and NDVIs computed from all combinations of wavebands between the NIR (756 to 1000 nm) and red or far-red (600 to 740 nm) produced different patterns for 2004 and 2005 (Fig. 5.3):

(i) in general, more combinations, i.e. 152 out of a total of 180 combinations yielded high coefficients of determination ($R^2 \geq 0.50$) for 2004 compared with 2005 (35 combinations)

(ii) the best five combinations for 2004 involved NIR bands and the red band at 695 nm, while for 2005, the best five combinations involved NIR bands and red-edge bands located at the longer wavelength end between 725 - 740 nm (Table 5.3)

(iii) the best five combinations for both 2004 and 2005 involved NIR bands located at the upper limit of the red edge (786 - 801 nm) and red-edge bands located mid-way along the red-edge slope (725 - 740 nm) (Table 5.4) and

(iv) the more traditional NDVI band combinations involving NIR and red wavelengths around the chlorophyll absorption centre (660-680 nm) performed poorly for 2005 biomass estimation.

A comparative analysis of the predictive performance of the NDVI involving analogous Landsat TM bands (831 & 665 nm), best NDVI for 2004, best NDVI for 2005, overall best NDVI (786 & 725 nm), MSAVI, SARVI and NDWI is presented in Table 5.5. The MSAVI and SARVI provided an insignificant improvement over NDVI computed from red and NIR bands. NDWI produced a higher correlation for 2004 than 2005. Although NDVI (786 & 725 nm) and NDVI (786 & 740 nm) showed high correlations ($R^2 \geq 0.50$) and low RMSECV for both 2004 and 2005, they were poor predictors of the following or previous years' biomass.

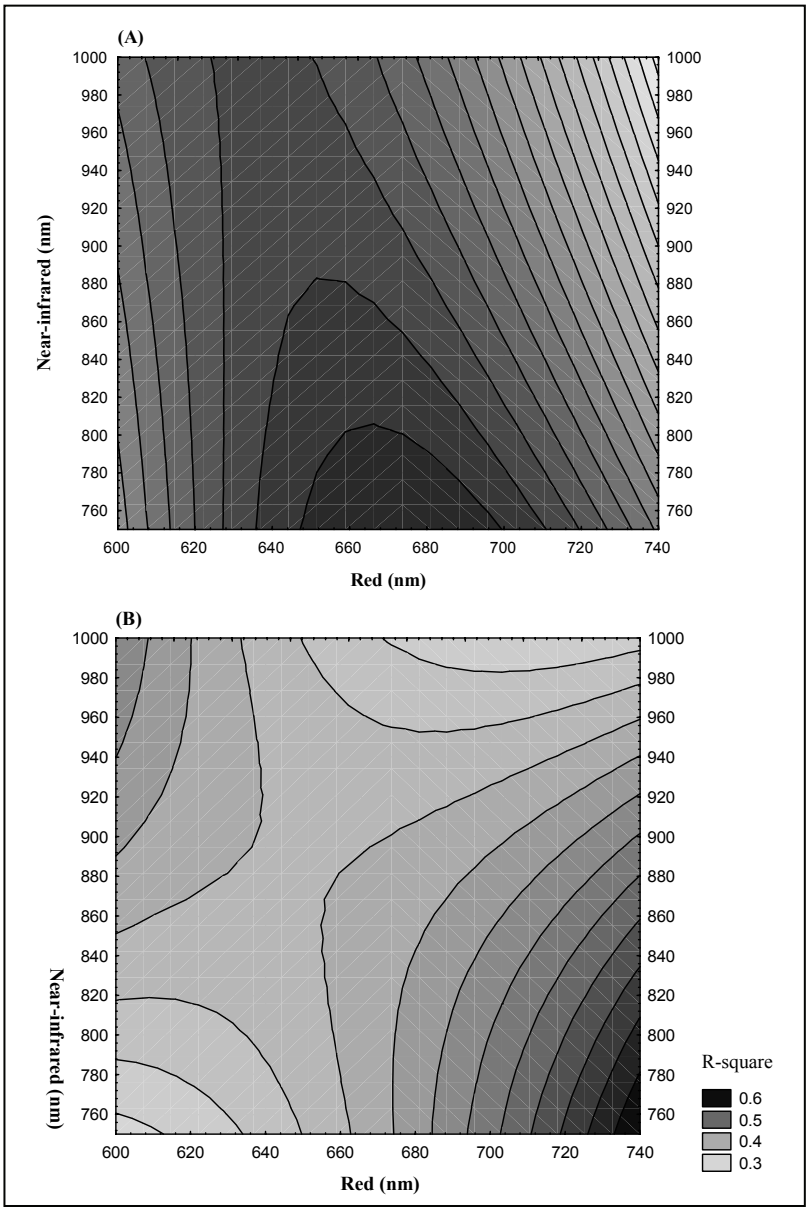


Fig.5.3. Contour plots showing the sensitivity (based on the coefficient of determination i.e. R-square) of the relations between Majella green grass/herb biomass and NDVIs calculated from all combinations of near-infrared (756 to 1000 nm) and red or far-red (600 to 740 nm) bands for (A) 2004 and (B) 2005 HyMap images.

Table 5.5 A comparative analysis of the performance of vegetation indices and red-edge position (REP) extracted by three methods for predicting grass/herb biomass using HyMap images. The images were acquired over Majella National Park, Italy in the summers of 2004 and 2005. R^2 and RMSECV denote the coefficient of determination and the root mean square error of leave-one-out cross validation, respectively.

2004 HyMap image	Linear regression model	R^2	RMSECV ($g\ m^{-2}$)	Prediction error (RMSE) based on 2005 model
NDVI (831 & 665 nm)	- 758.8 + 2328.7 NDVI	0.55	255	301
NDVI(786 & 695 nm)	- 455.7 + 2326.6 NDVI	0.56	252	298
NDVI (786 & 740 nm)	- 425 + 17522 NDVI	0.50	264	273
NDVI (786 & 725 nm)	- 205.1 + 5786.8 NDVI	0.55	252	294
MSAVI	- 1791.5 + 1627.5 MSAVI	0.54	258	304
SARVI	- 283.2 + 1847.6 SARVI	0.55	255	290
NDWI	804.69 + 5729.6 NDWI	0.55	251	389
REP (linear interpolation)	- 146667 + 205131 REP	0.47	272	352
REP (three-point Lagrangian interpolation)	- 52499 + 74475 REP	0.50	265	266
REP (linear extrapolation)	- 27980 + 40498 REP	0.53	258	279
2005 HyMap image	**Linear regression model**	**R^2**	**RMSECV**	**Prediction error (RMSE) based on 2004 model**
NDVI (831 & 665 nm)	- 744 + 2040.4 NDVI	0.32	319	361
NDVI(786 & 695 nm)	- 523 + 2121 NDVI	0.38	306	346
NDVI (786 & 740 nm)	- 697 + 20149 NDVI	0.64	231	349
NDVI (786 & 725 nm)	- 470.2 + 6393.5 NDVI	0.58	253	280
MSAVI	- 1703.4 + 1458.6 MSAVI	0.30	325	365
SARVI	- 270.5 + 1556.4 SARVI	0.31	322	356
NDWI	496.6 + 4866.4 NDWI	0.49	280	444
REP (linear interpolation)	- 181974+ 254570 REP	0.62	239	295
REP (three-point Lagrangian interpolation)	- 51651 + 73184 REP	0.56	258	254
REP (linear extrapolation)	- 33500 + 48110 REP	0.58	252	258

Note: All the relations where statistically significant at $p<0.05$

5.3.3 Predictive performance of the red-edge position

Among the REP methods, only REPs extracted by the Lagrangian and linear extrapolation methods were highly correlated ($R^2 \geq 0.50$) with biomass for 2004 and 2005. Nevertheless, REPs extracted by the linear interpolation method yielded the highest correlation ($R^2 = 0.62$) and the lowest RMSECV (239 g m^{-2}) for 2005 when the vegetation was fresher. Compared with regression models developed using the best overall NDVI (786 & 725), the Lagrangian and linear extrapolation REP models for each year produced higher accuracies for grass/herb biomass prediction for the other year (Table 5.5 and Fig. 5.4). Fig.5.4 shows the predicted grass/herb biomass for a subset area of the 2005 image based on linear regression models derived from the best overall NDVI (786 & 725) and linear extrapolation REP for 2004 and 2005. It could be observed that the predicted maps based on the REP models showed higher similarities compared with the NDVI models.

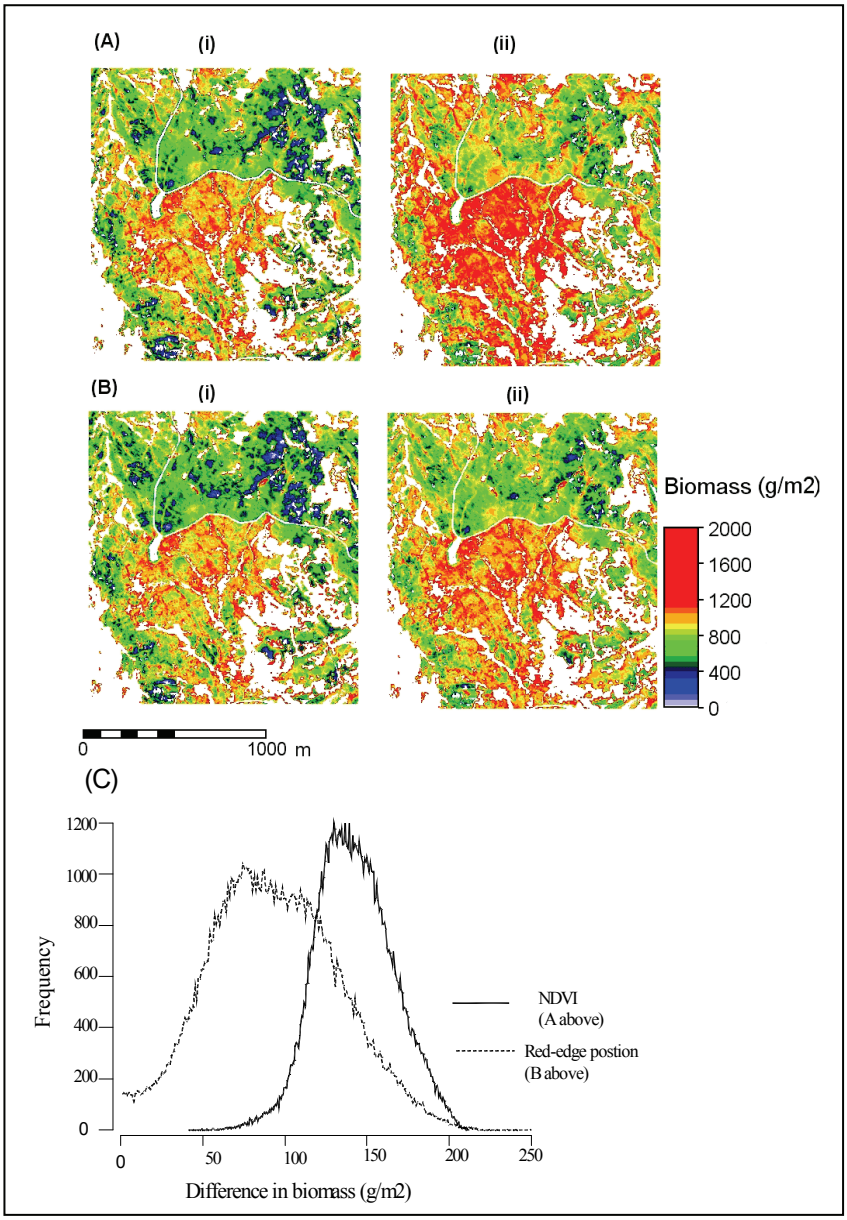


Fig.5.4 Predicted green grass/herb biomass for a subset area of the 2005 HyMap image based on (i) 2005 and (ii) 2004 regression models for (A) NDVI (786 & 725 nm) and (B) red-edge position extracted by the linear extrapolation method. (C) Histogram showing the differences between (i) and (ii), i.e. number of pixels against difference in biomass.

5.4 Discussion

The present study evaluates the robustness or stability of hyperspectral predictors for estimating grass/herb biomass between consecutive yearly hyperspectral images. HyMap images were acquired in the study area, the Majella National Park, Italy on 15 and 4 July 2004 and 2005, respectively. Despite the short period (11 days) between these two consecutive summer acquisition dates, the spectral analyses of grass/herb plots (Fig.5.2) seem to suggest that the vegetation and atmospheric conditions were different. No significant difference was however found between the means of green grass/herb biomass for 2004 and 2005.

This study shows that frequently used NDVIs computed from canopy reflectance in the red (665-680 nm) and near-infrared bands, MSAVI, SARVI and NDWI are not reliable predictors of grass/herb biomass on a yearly basis. The above indices were highly correlated ($R^2 \geq 0.50$) with biomass only for 2004 when the vegetation was in the early stages of senescence. Conversely, the results do support the growing body of evidence which shows that narrow-bands in the red-edge are more consistent predictors of plant biophysical parameters (Thenkabail et al., 2000; Gupta et al., 2003; Hansen and Schjoerring, 2003; Mutanga and Skidmore, 2004). However, the linear regression models derived from the best overall NDVI involving narrow-bands at 725 and 786 nm were year-specific because the models for one year poorly predicted the biomass of another year. Differences in phenological and atmospheric conditions between 2004 and 2005 might have affected the stability or robustness of the empirically derived NDVI models. It has been shown in several other studies that empirical models derived from vegetation indices are highly site and sensor specific and unsuitable for application to large areas or in different seasons (e.g. Curran, 1994; Gobron et al., 1997).

The results of this study show that REPs extracted by the Lagrangian and linear extrapolation methods correlated highly ($R^2 \geq 0.50$) with green grass/herb biomass for both 2004 and 2005. Interestingly, the Lagrangian and linear extrapolation REP models for one year predicted the biomass of the other year with higher accuracies compared with the linear interpolation REP and NDVI (786 & 725 nm) regression models. Differences in phenological and atmospheric conditions might have only a minor effect on the relationship between biomass and the Lagrangian or

linear extrapolation REP compared with the linear interpolation REP. In fact, Clevers et al. (2001) demonstrated that the REPs are least sensitive to atmospheric and soil conditions. This may apply particularly to the Lagrangian and linear extrapolation REPs that are computed from derivative spectra. Derivative analysis enhances absorption features and suppresses contributions of non-vegetative reflectance components (Boochs et al., 1990; Curran et al., 1991). The applicability of the Lagrangian and linear extrapolation REP regression models for different Mediterranean mountain grassland habitats and/or sensor types needs to be established.

In summary, the determination of spectral predictors that produce consistent correlations with peak grass/herb biomass for slightly different phenological and atmospheric conditions could be useful for monitoring annual changes in biomass production. These results are particularly crucial for the Mediterranean mountain landscape because of the unstable summer climate in this region, which makes it difficult to obtain cloud- or haze-free images at a desired phenological stage. Moreover, the robustness of regression models derived from the Lagrangian and linear extrapolation REPs, means that more reliable estimates of biomass can be obtained for a new HyMap image for which field-measured biomass data are unavailable.

5.5 Summary and conclusions

The robustness of hyperspectral predictors for estimating green grass/herb biomass in the Majella National Park, Italy on a yearly basis were assessed in terms of (i) the consistency of the relationships between biomass and the spectral predictors and (ii) the capability of empirical models developed for 2004 to predict the biomass of 2005 and vice versa. We conclude that the relationships between green grass/herb biomass and frequently used NDVIs computed from canopy reflectance in the red (665-680 nm) and near-infrared bands, MSAVI, SARVI and NDWI are not consistent from one year to the other. However, NDVI involving wavebands at 725 and 786 nm, or REPs extracted by the three-point Lagrangian interpolation and linear extrapolation techniques, produced high correlation ($R^2 \geq 0.50$) for both 2004 and 2005. But only the regression models based on REPs extracted by the Lagrangian and linear extrapolation methods for each year produced reliable estimates of biomass for the other year. The results of this study could be useful for

selecting hyperspectral predictors for monitoring annual changes in grass/herb biomass production across other Mediterranean mountain ecosystems.

Acknowledgements

The International Institute for Geo-Information Science and Earth Observation (ITC) provided financial support for this study. We also extend our gratitude to the management of Majella National Park, Italy, and particularly to Dr Theodoro Andrisano.

Chapter 6

Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression

This Chapter is based on
Moses Azong Cho, Andrew Skidmore, Fabio Corsi, Sipke E. van Wieren, Istiak Sobhan (*in press)*. Estimation of green grass/herb biomass from hyperspectral imagery using spectral indices and partial least squares regression. *International Journal of Applied Geoinformation Science and Earth Observation.*

Abstract

The main objective was to determine whether partial least squares (PLS) regression improves grass/herb biomass estimation when compared with hyperspectral indices, that is normalised difference vegetation index (NDVI) and red-edge position (REP). To achieve this objective, fresh green grass/herb biomass and airborne images (HyMap) were collected in the Majella National Park, Italy in the summer of 2005. The predictive performances of hyperspectral indices and PLS regression models were then determined and compared using calibration (n = 30) and test (n = 12) data sets. The regression model derived from NDVI computed from bands at 740 and 771 nm produced a lower standard error of prediction (SEP = 264 g m^{-2}) on the test data compared with the standard NDVI involving bands at 665 and 801 nm (SEP = 331 g m^{-2}), but comparable results with REPs determined by various methods (SEP = 261 to 295 g m^{-2}). PLS regression models based on original, derivative and continuum-removed spectra produced lower prediction errors (SEP = 149 to 256 g m^{-2}) compared with NDVI and REP models. The lowest prediction error (SEP = 149 g m^{-2}, 19% of mean) was obtained with PLS regression involving continuum-removed bands. In conclusion, PLS regression based on airborne hyperspectral imagery provides a better alternative to univariate regression involving hyperspectral indices for estimating grass/herb biomass in the Majella National Park.

Keywords: Green grass/herb biomass; NDVI; red-edge position; HyMap; partial least squares regression; continuum-removal.

6.1 Introduction

Accurate estimates of grass biomass at peak productivity can provide valuable information about the productivity and functioning of rangelands. For example, the quantity of forage influences the grazing patterns of animals (Senft et al., 1985; Pierce et al., 1994). Remote sensing techniques have been widely used to model the spatial variability of grass biomass over large areas (Richardson et al., 1982; Everitt et al., 1989; Anderson et al., 1993; Wylie et al., 2002).

There are major limitations with the normalised difference vegetation index (NDVI) despite its wide application for modelling the spatial variability of biomass. NDVI developed by Rouse et al. (1974) is commonly computed from canopy reflectance in the red and near-infrared (NIR) using broad-band imagery such as NOAA advanced very high resolution radiometer (AVHRR). Several studies show that broad-band NDVI can be unstable, varying with soil colour, canopy structure, leaf optical properties and atmospheric conditions (Huete and Jackson, 1988; Middleton, 1991; Kaufman and Tanré, 1992; Qi et al., 1995). Furthermore, broad-band NDVI asymptotically approaches a saturation level after a certain biomass or leaf area index (LAI) (Sellers, 1985; Gao et al., 2000). Broad-band NDVI uses average spectral information over broad-band widths resulting in loss of critical information available in specific narrow-bands (Blackburn, 1998; Thenkabail et al., 2000).

Recent developments in hyperspectral remote sensing have provided additional bands for vegetation analysis within the visible, NIR and shortwave infrared (SWIR). The results of several studies show that NDVIs computed from specific narrow-bands improve LAI or biomass estimation (Gong et al., 2003; Lee et al., 2004; Mutanga and Skidmore, 2004). Using grass (*Cenchrus ciliaris*) grown in the greenhouse, Mutanga and Skidmore (2004) showed that narrow-band NDVI computed from 740 and 755 nm solves the saturation problem of estimating grass biomass at high canopy cover. Another hyperspectral index that has been assessed for grass biomass estimation is the wavelength of maximum slope in the red-NIR region, termed the red-edge position (REP) (Gilabert et al., 1996; Mutanga and Skidmore, 2004). An advantage of the REP over NDVI is that it is less sensitive to varying soil and atmospheric conditions, and sensor view angle (Curran et al., 1995; Blackburn and Pitman, 1999; Clevers et al., 2001).

NDVI requires only two bands and could be limited in terms of exploiting the rich information content of hyperspectral data. Alternatively, multiple linear regression based on more than two hyperspectral bands has been used to predict vegetation parameters such as foliar nutrient contents and biomass (Curran, 1989; Kokaly and Clark, 1999; Curran et al., 2001; De Jong et al., 2003). However, multiple regression with hyperspectral data is likely to suffer from multicollinearity (Curran, 1989; De Jong et al. 2003). This usually occurs when the number of samples is smaller than the number of wavebands used in the analysis (Curran, 1989; Nguyen and Lee, 2006). In contrast, a multivariate statistical technique that is widely used in chemometrics to deal with this problem is partial least squares (PLS) regression (Feudale and Brown, 2005; Geladi and Kowalski, 1986; Kooistra et al., 2004; van den Broek et al., 1996).

PLS regression is closely related to principal component regression. But instead of first decomposing the spectra into a set of eigenvectors and scores and regressing them against the response variables as a separate step, PLS regression actually uses the response variable information during the decomposition process (Geladi and Kowalski, 1986). Few studies have explored the potential of PLS regression for estimating vegetation parameters using airborne hyperspectral imagery. For example, Schmidtlein and Sassin (2004) used PLS regression for mapping floristic gradients in grasslands and Huang et al. (2004) demonstrated that PLS regression was a better alternative to conventional stepwise regression for estimating foliar nitrogen. Recently, PLS regression was used for predicting wheat biomass using reflectance measurement obtained with a field spectrometer (Hansen, and Schjoerring, 2003). The utility of partial least square regression based on airborne hyperspectral imagery for estimating biomass in natural grass and/or herb communities remains to be established.

Thus, the research objectives were to (i) ascertain the utility of hyperspectral indices derived from airborne hyperspectral imagery for estimating green grass/herb biomass in a semi-natural landscape and (ii) determine whether PLS regression increases the accuracy of green grass/herb biomass estimation when compared to hyperspectral indices. To achieve these objectives, green grass/herb biomass and airborne images (HyMap) were collected in the Majella National Park, Italy in the summer of 2005. The predictive performance of hyperspectral indices

(NDVI and REP) and PLS regression models were then determined and compared using calibration and test data sets.

6.2 Material and methods

6.2.1 The study area

The study site is located in Majella National Park, Italy (latitude 41°52' to 42°14'N, longitude 13°50' to 13°14'E), which covers an area of 74095 ha. The park extends into the southern part of Abruzzo, at a distance of 40 km from the Adriatic Sea. This region is situated in the massifs of the Apennines. The park is characterised by several mountain peaks, the highest being Mount Amaro (2794 m).

More specifically, the study site (latitude 41°49' to 42°14'N, longitude 13°57' to 14°6'E) is situated between Mounts Majella and Morrone to the east and west, respectively. It covers an area of 40 km by 5.5 km. Gallego Fernández et al. (2004) argue that plant community dynamics in Mediterranean basin ecosystems are driven mainly by alternating episodes of human intervention and land abandonment. For example, abandoned settlement and agricultural areas in Majella are returning to oak (*Quercus pubescens*) woodlands at the lower altitude (400 m to 600 m) and beech (*Fagus sylvatica*) forest at the higher altitude (1200 m to 1800 m). Between these two formations is a landscape composed of shrubby bushes, patches of grass/herb vegetation, and bare rock outcrops. The dominant grass species include *Brachypodium genuense*, *Briza media*, *Bromus erectus* and *Festuca sp*. Herbs include *Helichrysum italicum, Galium verum, Trifolium pratense, Plantago lanceolata, Sanguisorba officinalis* and *Ononis spinosa.*

6.2.2 Field data collection

A field campaign was carried out from 16 to 29 June 2005. There are four main phytosociological classes (semi-natural/farmlands, grazed/periodically flooded areas, open garrigues and abandoned farmlands) of varying areas within the specific study area. The phytosociological map of the park was provided by the park management. Random sampling with clustering was adopted in the study. Twenty-five coordinates (x, y) were randomly generated in four phytosociological classes using ArcGIS software; eight plots in the semi-

natural/farmlands/abandoned farmlands, five plots in the open garrigues and twelve plots in the grazed/periodically flooded areas. The number of samples per vegetation class was proportional to the size of the class. A GPS was used to locate the sample plots in the field. A 30 m by 30 m plot size was adopted in this study. For each of the initially randomly generated plot, an extra plot located at about 150 m away was sampled for fresh green grass/herb biomass. Only plots with more than 20% homogeneous grass/herb cover were sampled, resulting in a total of 47 plots. Above-ground biomass was clipped within five randomly selected subplots (1 m by 0.5 m) from each plot. All dry material was removed from the clipped plants before measuring the green grass/herb biomass. Average green grass/herb biomass per plot was calculated from the five subplot measurements.

6.2.3 Image acquisition and pre-processing

Airborne HyMap data of the study site were obtained on 4 July 2005. The flight was carried out by DLR, Germany's Aerospace Research Centre and Space Agency. The HyMap sensor comprised 126 wavebands, operating over the wavelength range 436 nm to 2485 nm, with average spectral resolutions of 15 nm (436 nm to 1313 nm), 13 nm (1409 nm to 1800 nm) and 17 nm (1953 nm to 2485 nm). The spatial resolution of the data was 4 m. The images were collected at solar noon. The specific study site was covered by four image strips, each covering an area of about 40 km by 2.3 km. The solar zenith and azimuth angles for the image strips ranged between 30-33.7^{o} and 111.5-121^{o}, respectively.

The image strips were atmospherically and geometrically corrected by DLR. The atmospheric correction was carried out using ATCOR4-r (rugged terrain). ATCOR4 is based on MODTRAN-4 radiative transfer code. However, there were differences between the reflectance of similar pixels in the overlapping sections between image strips. Spectral calibration between strips was carried out using the empirical line method in Environment for Visualising Images (ENVI 4.2) software (Research System, Inc.) in order to minimize the differences. Ten image spectra collected from a reference strip (e.g. strip 2) and corresponding targets from its overlapping neighbour (strip1 or 3) were used to compute a linear regression function for each channel. Using the regression functions, strips 1 and 3 were then adjusted to have a spectral response

similar to that of strip 2. The same process was carried out using corrected strip 3 as the reference image to correct strip 4. The spectra were collected from targets such as roads, agricultural fields, quarry fields, and dense beech forest pixels.

6.2.4 Collecting image spectra for grass/herb plots

Grass/herb areas were extracted from the image strips in order to eliminate mixed grass/shrubs and or tree pixels. First, an NDVI image involving bands at 665 nm and 831 nm was computed for each image strip using the ENVI 4.2 software. A point map of the grass/herb plots was then overlaid on the NDVI images. Pixels of pure grass/herb plots were used to determine minimum and maximum NDVI threshold values for grass/herbs. Next, a grass/herb region-of-interest map was created using the NDVI threshold values. Subsequently, the region-of-interest map was used to subset grass/herb areas from the image. All other pixels were masked out.

A 7 by 7 pixels window (i.e. 28 m by 28 m) was used to collect grass/herb image spectra from each sample plot in order to avoid including pixels located outside the plot (30 m by 30 m). The spectra were collected and averaged. The spectra of five out of the 47 plots were not extracted from the image strips because the plots were located in portions covered by clouds.

6.2.5 Data analysis

Two main approaches were adopted in this study: (ii) hyperspectral indices (NDVI and REP) and (i) partial least squares regression. The data was randomly split into the training or calibration (n = 30) and test (n = 12) sets. Regression analyses were performed on the calibration set. Empirical validation of the linear regression models for the indices and PLS regression models were carried out using the test set. The performances of the various regression models were compared using the coefficient of determination (R^2) for calibration, the standard error of calibration (SEC, Eq. 6.1) and standard error of prediction based on the independent test data (SEP).

$$SEC = \sqrt{\frac{\sum_{i=1}^{n}(y - y')^2}{n}} \tag{6.1}$$

where y = measured biomass, y′ = estimated biomass for test data and n = number of observations.

6.2.5.1 Hyperspectral indices

(i) Normalised difference vegetation index (NDVI)

Broad-band NDVI is commonly calculated from the red and NIR reflectance (Eq. 6.2). But hyperspectral data provides additional bands in the red (600-700 nm) and red-edge/NIR (700-1300 nm) regions. Narrow-band NDVIs were calculated from all two-band combinations between red or far-red (600 to 740 nm) and NIR (756 to 1000 nm) using the calibration data set (n = 30) in order to determine the best NDVI.

$$NDVI = (R_{NIR} - R_{red})/(R_{NIR} + R_{red}) \tag{6.2}$$

where R = reflectance

(ii) Red-edge position (REP)

Red-edge positions were extracted by three simple methods; the linear four-point interpolation (Guyot and Baret, 1988), three-point Lagrangian interpolation (Dawson and Curran, 1998) and the linear extrapolation (Cho and Skidmore, 2006) methods.

(a) Linear four-point interpolation method

The linear four-point interpolation method (Guyot and Baret, 1988) assumes that the reflectance curve at the red edge can be simplified to a straight line centred near the midpoint between the reflectance in the NIR at about 780 nm and the reflectance minimum of the chlorophyll absorption feature at about 670 nm. It uses four wavebands, 670, 700, 740 and 780 nm i.e. 665, 695, 740 and 786 nm for the HyMap spectrum. The REP is then determined by using a two-step calculation procedure.

Calculation of the reflectance at the inflexion point (R_{re}):

$$R_{re} = \left(R_{665} + R_{786}\right)/2 \tag{6.3}$$

where R is the reflectance at a specified wavelength (e.g. 665 nm).

Calculation of the red edge wavelength or red edge position (REP):

$$REP = 695 + 45\left(\frac{R_{re} - R_{695}}{R_{740} - R_{695}}\right) \tag{6.4}$$

695 and 45 are constants resulting from interpolation in the 695-740 nm interval.

(b) Three-point Lagrangian interpolation method

The three-point Lagrangian interpolation technique (Dawson and Curran, 1998) is designed to locate REP in spectra that have been sampled coarsely. Lagrangian interpolation is applied to the first derivative of the reflectance spectrum which is computed as follows:

$$D_{\lambda(i)} = (R_{\lambda(j+1)} - R_{\lambda(j)})/\Delta\lambda \tag{6.5}$$

where $D_{\lambda(i)}$ is the first derivative reflectance at a wavelength i, midpoint between wavebands j and j+1, $R_{\lambda(j)}$ is the reflectance at the j waveband, $R_{\lambda(j+1)}$ is the reflectance at the j+1 waveband, and Δ_λ is the difference in wavelengths between j and j+1.

The value of the first derivative at any wavelength (i.e. estimated value) will be D_λ. The Lagrangian interpolation technique for three known bands is given by

$$D_\lambda = \frac{(\lambda-\lambda_i)(\lambda-\lambda_{i+1})}{(\lambda_{i-1}-\lambda_i)(\lambda_{i-1}-\lambda_{i+1})}D_{\lambda(i-1)} + \frac{(\lambda-\lambda_{i-1})(\lambda-\lambda_{i+1})}{(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})}D_{\lambda(i)} + \frac{(\lambda-\lambda_{i-1})(\lambda-\lambda_i)}{(\lambda_{i+1}-\lambda_{i-1})(\lambda_{i+1}-\lambda_i)}D_{\lambda(i+1)} \tag{6.6}$$

The band having the maximum first derivative will be λ_i, with λ_{i-1} and λ_{i+1} representing the two bands on either side of the maximum derivative. To determine the REP, a second derivation on Eq. 6.6 is performed and resolved for when the second derivative is zero, i.e.

$$REP = \frac{A(\lambda_i + \lambda_{i+1}) + B(\lambda_{i-1} + \lambda_{i+1}) + C(\lambda_{i-1} + \lambda_i)}{2(A + B + C)} \tag{6.7}$$

where

$$A = \frac{D_{\lambda(i-1)}}{(\lambda_{i-1} - \lambda_i)(\lambda_{i-1} - \lambda_{i+1})}, B = \frac{D_{\lambda(i)}}{(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})}, and \tag{6.8}$$
$$C = \frac{D_{\lambda(i+1)}}{(\lambda_{i+1} - \lambda_{i-1})(\lambda_{i+1} - \lambda_i)}$$

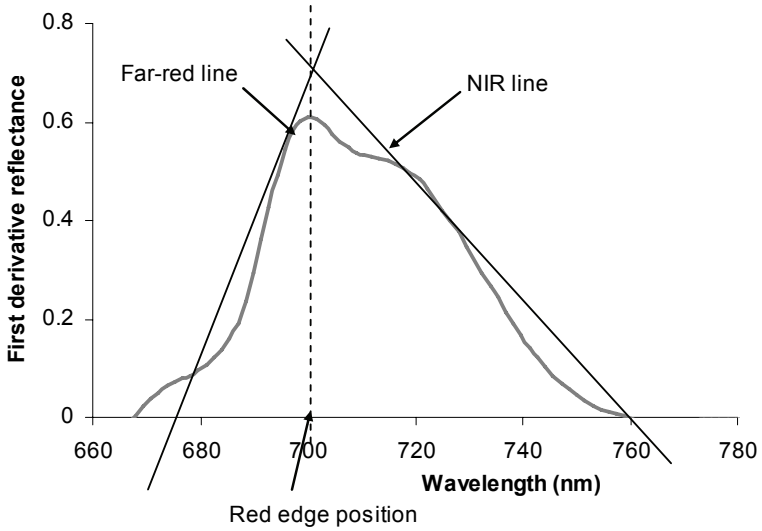


Fig. 6.1 Schematic representation of the linear extrapolation technique for extracting the red-edge position (Cho and Skidmore, 2006).

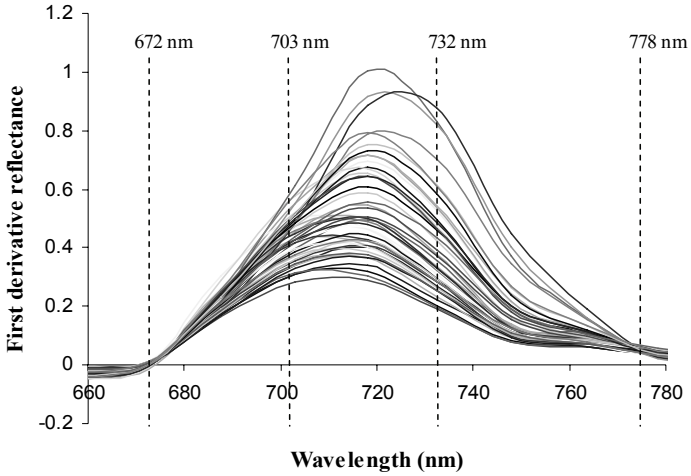


Fig. 6.2 First derivative spectra of grass/herb plots showing bands used in the calculation of red-edge positions by the linear extrapolation method.

(c) Linear extrapolation technique

The linear extrapolation technique (Cho and Skidmore, 2006) is designed to track changes near chlorophyll sensitive peaks in the first derivative (D) of the red edge i.e. around 700 and 725 nm (Horler et al. 1983). The REP is calculated as the wavelength at the intersection of two straight lines (Eq. 6.9 & 6.10) extrapolated through two points on the far-red flank and two points on NIR flank of first derivative reflectance spectrum (Fig. 6.1). For example, for the HyMap derivative spectra used in this study, the lines were extrapolated through derivative bands at 672 and 703 nm for the far-red line and 732 and 778 nm for the NIR line (Fig. 6.2).

$$\text{Far-red line: } D = m_1\lambda + c_1 \tag{6.9}$$

$$\text{NIR line: } D = m_2\lambda + c_2 \tag{6.10}$$

where m_i and c_i (for i = 1, 2) represent the slope and intercept of the straight lines, respectively; c_1 and m_1 for the far-red line and c_2 and m_2 for the NIR line. At the intersection, the two lines have equal λ and D values. Therefore, the REP, which is the λ at the intersection, is given by:

$$REP = \frac{-(c_1 - c_2)}{(m_1 - m_2)} \tag{6.11}$$

where

$$m_1 = \frac{(D_{703} - D_{672})}{(\lambda_{703} - \lambda_{672})}, m_2 = \frac{(D_{778} - D_{732})}{(\lambda_{778} - \lambda_{732})},\ c_1 = D_{703} - m_1\lambda_{703},$$
$$c_2 = D_{732} - m_2\lambda_{732} \tag{6.12}$$

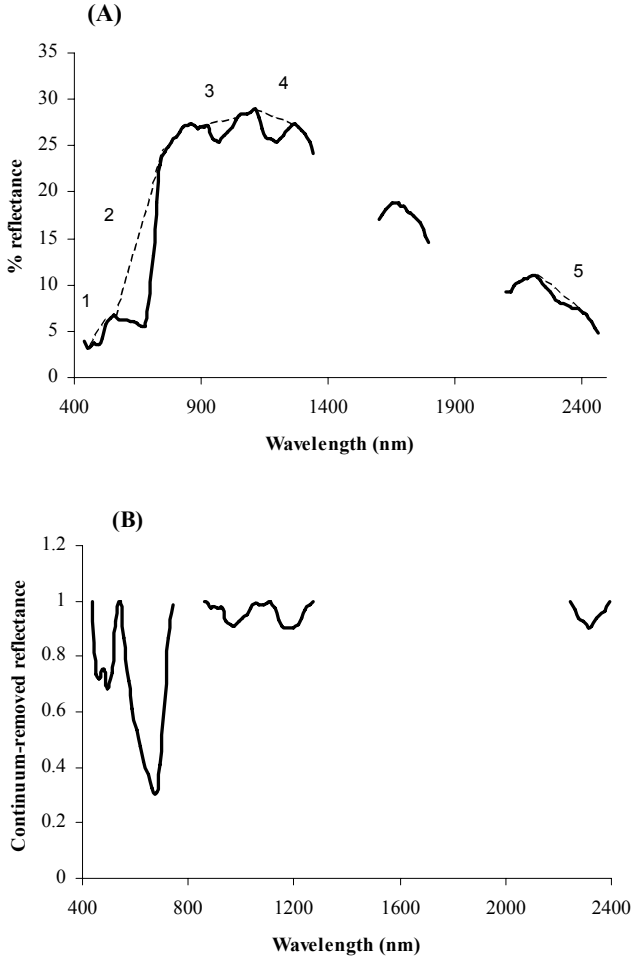


Fig. 6.3 Continuum removal applied to five absorption features (A) and the resulting continuum-removed curve (B).

6.2.5.2 Partial least squares regression

PLS regression is a bilinear calibration method using data compression by reducing the large number of measured collinear spectral variables to a few non-correlated latent variables or factors (Geladi and Kowalski, 1986; Hansen and Schjoerring, 2003). As in multiple regression, the main purpose of PLS regression is to build a linear model (Eq. 6.13),

$$Y = Xb + E \quad (6.13)$$

Where Y: mean-centred matrix containing the response variables (*n* by 1 matrix in this study), X: the mean-centred matrix containing the predictor variables (the spectral bands in this study). b: matrix containing the regression coefficients and E: the matrix of residuals. In PLS regression, the above principle is used on latent variables of X. In this sense, PLS regression is closely related to principal component regression (Geladi and Kowalski, 1986 and Geladi et al., 1999). But instead of first decomposing the spectra into a set of eigenvectors and scores and regressing them against the response variables as a separate step, PLS regression actually uses the response variable information during the decomposition process. Further information on the PLSR model can be obtained in Geladi and Kowalski (1986).

The PLS regression analyses were performed using the original and transformed spectra i.e. first derivative and continuum-removed spectra. The main atmospheric water absorption regions i.e. 1400-1600 nm and 1800-2100 nm were not included in the analyses. Continuum removal was applied to five major absorption troughs in the vegetation spectrum, from the visible to the SWIR (Fig.6.3). The continuum is removed by dividing the original reflectance values in the absorption trough by the corresponding values of the continuum line (Kokaly and Clark, 1999). The output curves have values between zero and one in which the absorption troughs are enhanced (Schmidt and Skidmore, 2001).

Before performing the PLS regression, the spectral data were mean-centred by subtracting their means. The leave-one-out cross-validation method was used to select the optimal number of PLS factors or latent variables to be included in the regression models (Geladi and Kowalski, 1986; Viscarra Rossel, 2005). In order to avoid the collinearity problem and to maintain model parsimony, the criterion to add an additional factor to the model was that it had to reduce the root mean square error of

cross-validation (RMSECV) by > 2% (Kooistra et al., 2004). The RMSECV was determined from the residuals of each cross-validation phase. The performance of PLS models were estimated with the independent test data using the standard error of prediction (Eq. 6.1). The analyses were carried out using STATISTICA software (StatSoft, Inc.)

It has been shown that variable selection enhances PLS regression results (Davies, 2001; Kubinyi, 1996; Martens and Martens, 2000; Schmidtlein and Sassin, 2004). A sub-objective therefore, was to test PLS models based on all HyMap bands with the exception of the main atmospheric water absorption regions and on a small number of selected bands. The selection was based on bands related to vegetation parameters such as leaf chlorophyll, LAI and leaf mass (Table 6.1). The continuum-removed bands were slightly different and involved 466 nm, 495 nm, 679 nm, 695 nm, 725 nm, 710 nm, 740 nm, 895 nm, 975 nm, 1128 nm, 1215 nm, 2260 nm and 2359 nm.

Table 6.1 Wavebands selected for estimating green grass biomass using partial least squares regression.

Waveband centre (nm)	Description	References
466	chlorophyll b	Curran, 1989
695	total chlorophyll	Carter 1994; Gitelson and Merzylak, 1997
725	total chlorophyll, leaf mass	Horler et al, 1983
740	leaf mass and LAI	Horler et al. 1983
786	leaf mass	Guyot and Baret, 1988
846	leaf mass, LAI, chlorophyll	Thenkabail et al. 2004
895	leaf mass, LAI	Thenkabail et al. 2004
1113	leaf mass, LAI	Thenkabail et al. 2004
1215	plant moisture, cellulose, starch	Thenkabail et al. 2004, Curran, 1989
1661	lignin, leaf mass, starch	Thenkabail et al. 2004
2173	protein, nitrogen	Curran, 1989
2359	cellulose, protein, nitrogen	Curran, 1989

6.3 Results

The results of hyperspectral indices are first presented. Subsequently, the results of PLS regression are examined to find out if they provide any improvements over the spectral indices.

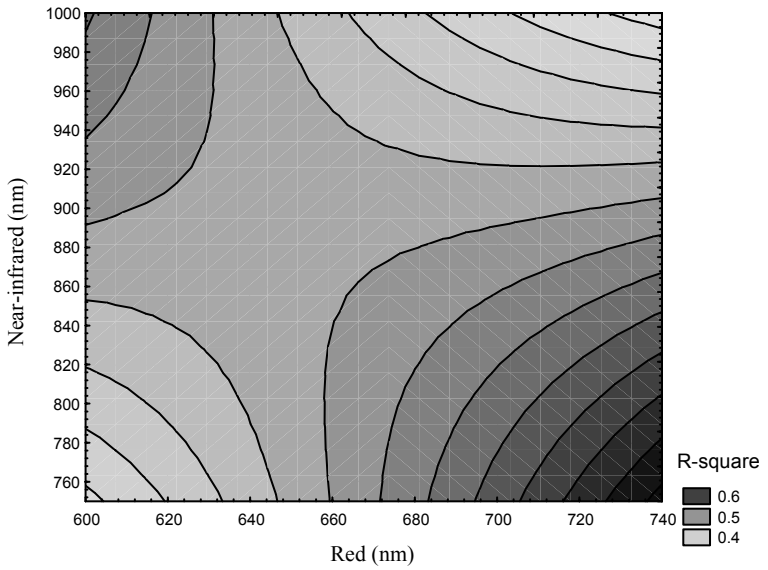


Fig. 6.4 Contour plot showing the sensitivity (R-square) of the relations between green grass/herb biomass and NDVIs calculated from all combinations of near-infrared (756-1000 nm) and red or far-red (600-740 nm) using HyMap bands.

6.3.1 Hyperspectral indices (NDVI and REP)

A comprehensive analysis of the relationships between green grass/herb biomass and NDVIs computed from all two-band combinations between the red or far-red (600-740 nm) and NIR (756-1000 nm) reflectance is presented in Fig. 6.4. The calibration data set (n = 30) was used in this analysis. NDVIs involving far red-edge bands in the 725 nm to 800 nm range produced higher coefficients of determination compared with the traditional NDVIs computed from red and NIR bands. The best five NDVIs are shown in Table 6.2.

Table 6.2 Best NDVI combinations for predicting grass/herb biomass in the Majella National Park, Italy (n = 30). R^2 = coefficient of determination.

Near-infrared band (nm)	Red or far-red band (nm)	R^2
771	740	0.702
756	740	0.700
786	740	0.690
801	740	0.674
862	725	0.670
879	725	0.669
877	725	0.669
846	725	0.665
831	725	0.661
896	725	0.661

Table 6.3 Performance of normalised difference vegetation indices (NDVI) for predicting green grass/herb biomass in Majella National Park, Italy. R^2 = coefficient of determination, SEC = standard error of calibration and SEP = standard error of prediction.

Spectral predictors	Calibration (n = 30)		Independent validation (n = 12)	
	R^2 Actual vs predicted	SEC (g m^{-2})	SEP (g m^{-2})	% of mean
NDVI				
Standard NDVI (801 & 665 nm)	0.42	298	331	43
Best NDVI (771 & 740 nm)	0.70	214	264	34
Red-edge position (REP)				
Linear interpolation	0.70	216	261	34
Lagrangian interpolation	0.66	229	295	38
Linear extrapolation	0.67	223	284	37

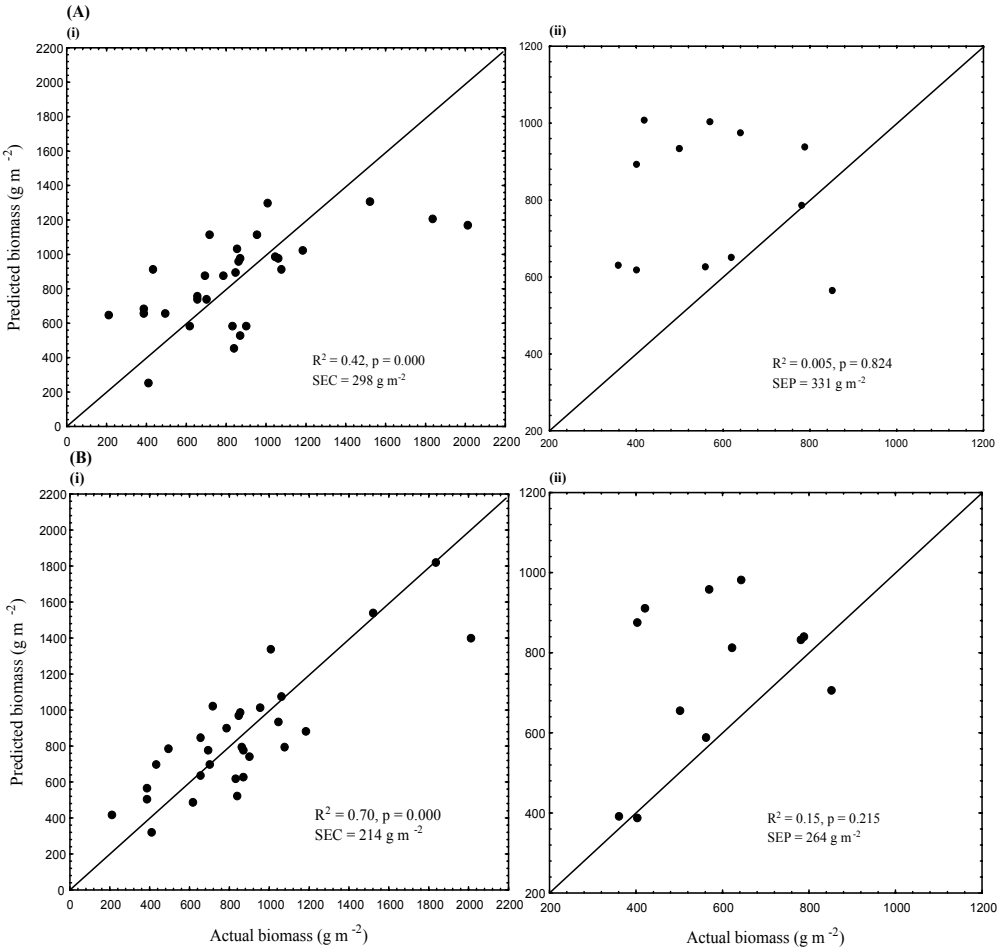


Fig. 6.5 (i) Calibration (n = 30) and (ii) independent validation (n = 12) results for predicting green grass/herb biomass in the Majella National Park using (A) standard NDVI computed from 801 and 665 nm and (B) best NDVI in this study computed from 771 and 740 nm. SEC = standard error of calibration and SEP = standard error of prediction.

A comparative analysis of the predictive performances of the best NDVI involving 740 and 771 nm, standard NDVI involving 665 and 801 nm and REP is presented in Table 3. The regression model involving the best NDVI produced a lower SEC (214 g m^{-2}nd SEP (264 g m^{-2}) compared with the standard NDVI which yielded SEC and SEP values of 298 g m^{-2} and SEP = 331 g m^{-2}, respectively. The calibration model based on the standard NDVI under-estimated biomass values above 1400 g m^{-2}. The high values fall below the diagonal in the plot of predicted versus actual

biomass. The relationship between the actual and predicted biomass in the calibration was highly influenced by the three highest biomass values (Fig. 6.5). These values belong to plots located in seasonally flooded grasslands. The calibration R^2 values dropped from 0.42 to 0.27 and 0.70 to 0.44 for the standard and best NDVI, respectively when the regression was established without the three high biomass values.

REP extracted by the linear interpolation method showed a slightly better performance than REP located by the three-point Lagrangian interpolation and linear extrapolation methods (Table 6.3). In general, the best NDVI and REP models produced similar correlations and prediction errors.

Table 6.4 Performance of partial least squares (PLS) regression for predicting green grass/herb biomass in Majella National Park, Italy. R^2 = coefficient of determination, RMSECV = root mean square error of cross validation, SEC = standard error of calibration and SEP = standard error of prediction.

	Calibration (n = 30)				Independent validation (n = 12)	
	No. of PLS factors	RMSECV (g m^{-2})	R^2 Actual vs. predicted	SEC (g m^{-2})	SEP (g m^{-2})	% of mean
All bands						
Original reflectance	3	251	0.72	207	204	26
First derivative	2	238	0.73	202	228	29
Continuum-removed reflectance	3	279	0.66	230	256	33
Selected bands						
Original reflectance	3	248	0.71	210	199	26
First derivative	2	243	0.71	212	216	28
Continuum-removed reflectance	6	263	0.83	163	149	19

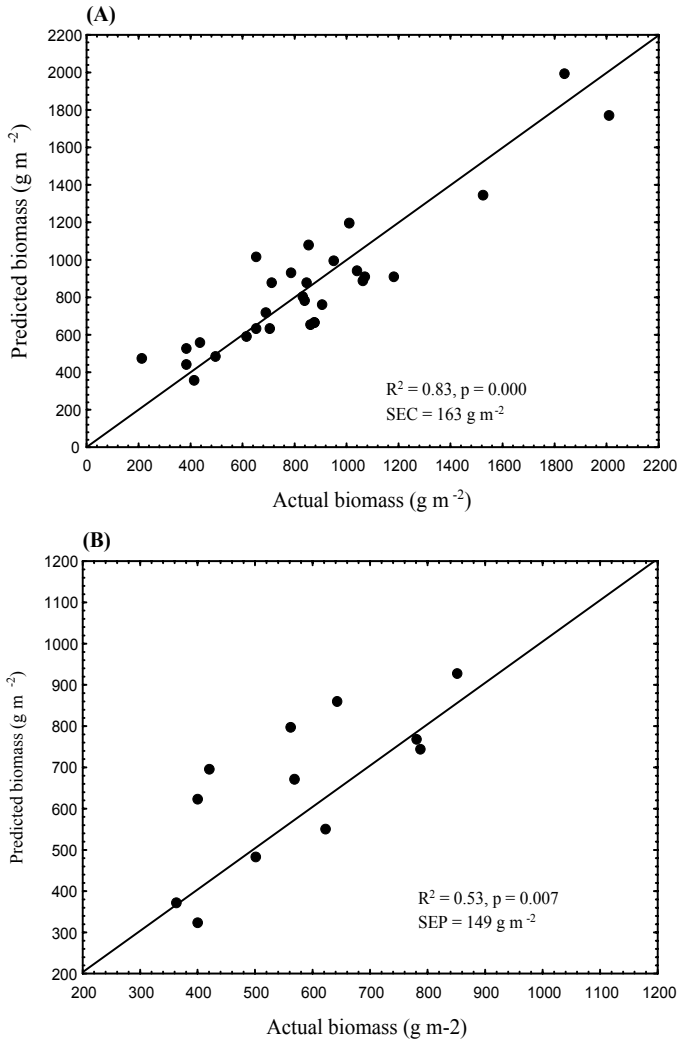


Fig. 6.6 (A) Calibration and (B) independent validation results for predicting grass/herb biomass in the Majella National Park using partial least squares regression based on continuum-removed reflectance bands. SEC = standard error of calibration, SEP = standard error of prediction.

6.3.2 Partial least squares regression

PLS models produced lower ranges of prediction errors for both the calibration (SEC = 163 to 230 g m^{-2}) and validation (SEP = 149 to 256 g m^{-2}) analyses compared to the NDVI and REP models (SEC = 214 to 298 g m^{-2} and SEP = 261 to 295 g m^{-2}).

A comparative analysis of the performances of PLS regression models based on the full HyMap spectrum (with the exception of atmospheric water absorption bands) and selected bands for the original, first derivative and continuum-removed spectra is presented in Table 6.4. Models based on the full spectrum and selected bands for the original and derivative spectra yield similar prediction errors for both the calibration and validation analyses. But the selected continuum-removed bands produced lower prediction errors (SEC = 163 g m^{-2} and SEP = 149 g m^{-2}) compared with models based on all continuum-removed bands (SEC = 230 g m^{-2} and SEP = 256 g m^{-2}). In fact, the selected continuum-removed bands produced the lowest SEP in the study (SEP = 149 g m^{-2}, 19% of the mean). Furthermore, the validation R^2 for the prediction based on the selected continuum-removed bands (Fig. 6.6) was significant ($p < 0.05$) contrary to results obtained from NDVI-based models (Fig. 6.5).

6.4 Discussion and conclusions

The present study demonstrates the utility of partial least squares regression based on hyperspectral airborne imagery for predicting grass/herb biomass in a semi-natural landscape. Predictive models based on partial least squares regression produced higher accuracies compared with hyperspectral indices, namely, vegetation indices and the REP. These results are consistent with those of Hansen and Schjoerring (2003) who in a field experimental study on winter wheat reported a better predictive performance of partial least squares regression analysis compared with NDVI for green biomass estimation.

The results of the study are inclusive with respect to whether band selection improves the predictive performance of PLS regression for grass biomass. With the exception of the model based on continuum-removed bands, bands selection did not improve the predictive performance of PLS regression for predicting the grass/herb biomass of Majella National park. However, the prediction using the selected continuum-removed bands was based on a high number of PLS factors (six in total). A high number PLS factors might compromise the parsimony of the model.

The model based on the standard NDVI under-estimated biomass values above 1400 g m^{-2}. Geladi et al. (1999) argues that such a bias can be caused by random noise or nonlinearities in the true physical relationship.

But this bias was not systematic across the other indices and PLS models, suggesting that it might be caused by the nonlinearity in the standard NDVI/biomass relationship. The standard NDVI has been shown to saturate at high biomass or LAI in several other studies (e.g. Mutanga and Skidmore, 2004; Sellers, 1985; Tucker, 1979). The explanation for the saturation problem is that, as from a LAI of 3, the amount of red light at about 660-680 nm that can be absorbed by leaves reaches a peak while NIR reflectance continues to increase due to the multiple scattering effects (Kumar et al., 2001). This imbalance between the saturation of red light absorption and high NIR reflectance results in a slight change in the NDVI resulting in a poor relationship with biomass (Mutanga et al. 2004). The results of this study showed a better predictive performance of NDVI involving far-red bands compared to the standard NDVI being consistent with those of Mutanga and Skidmore (2004) involving blue buffalo grass (*Cenchrus ciliaris*) grown in a greenhouse.

The predictive performance of the REP derived by various methods was comparable to that of the best NDVI in the study. However several studies show that REP is less sensitive to varying soil brightness, atmospheric condition and sensor view angle compared to NDVI (Curran et al., 1995; Blackburn and Pitman, 1999; Clevers et al., 2001). We showed in a recent study that REPs derived by the Lagrangian and linear extrapolation methods are better than NDVI-based models for monitoring grass/herb biomass between two consecutive yearly HyMap images (summers of 2004 and 2005) in Majella (Cho and Skidmore, in review).

In summary, the study demonstrates the potential utility of PLS regression involving HyMap bands for estimating green grass/herb biomass in the Majella National Park. However, it should be noted that empirical models might be site or sensor specific and unsuitable for application to large areas or in different seasons (Curran 1994; Gobron et al. 1997). In conclusion, PLS regression based on airborne hyperspectral imagery provides a better alternative to univariate regression involving hyperspectral indices for grass/herb biomass estimation in the Majella National Park.

Acknowledgements

The International Institute for Geo-Information Science and Earth Observation (ITC) provided financial support for this study. We also

extend our gratitude to the management of Majella National Park, Italy, and particularly to Dr Theodoro Andrisano.

Chapter 7

Mapping beech (*Fagus sylvatica* L.) forest structure using partial least squares regression on airborne hyperspectral imagery

This chapter is based on
Cho, M.A., Skidmore, A.K., Sobhan, I. (in review) Mapping beech (*Fagus sylvatica* L.) forest structure using partial least squares regression on airborne hyperspectral imagery. Forest Ecology and Management

Abstract

The objective of this study was to investigate whether multivariate statistical models such as partial least squares (PLS) regression and artificial neural networks (ANN) perform better than univariate models based on vegetation indices for estimating and mapping forest structural parameters of a closed canopy beech forest (*Fagus sylvatica* L). Airborne HyMap images and data on forest structural attributes including mean diameter-at-breast-height (DBH), mean tree height and tree density were collected from the Majella National Park, Italy in July 2004. The predictive performances of the various statistical techniques were evaluated using calibration (n = 33) and test (n = 20) data sets. The various calibration models provided higher prediction accuracies for Mean DBH and mean height (standard error of prediction i.e. SEP = 28 to 38% of mean) compared with tree density (SEP = 65 to 70% of mean). The best predictions for DBH (SEP = 5.50 cm, 28% of mean) and height (SEP = 5.61 m, 30% of mean) were provided by PLS regression and ANN, respectively. Soil and atmospherically resistant vegetation index (SARVI) was the best vegetation index for mean DBH (SEP = 6.03 cm, 30% of mean) and mean height (6.25 m, 33% of mean). The predicted map of mean DBH using PLS regression revealed a high heterogeneity of DBH which could be attributed to thinning, the principal management practice in the Park. The results of this study highlight the potential utility of multivariate techniques such as ANN and PLS regression based on several narrowbands for providing improved estimates of beech forest structural attributes compared with univariate methods based on vegetation indices. PLS regression has rarely been applied in this sense and thus offers new possibilities for mapping and monitoring beech forest structure.

Keywords: diameter-at-breast height, tree height, vegetation indices, partial least squares regression, artificial neural networks, HyMap imagery

7.1 Introduction

Information about the distribution of forest structural attributes such as tree diameter, basal area, height and density is essential for forest management. For example, thinning of high-density areas could promote diameter growth (Messina, 1992; Baldwin, et al., 2000; Fuhr et al., 2001). Conventional forest inventory data have been collected by means of field surveys. Such surveys are time consuming, labour intensive and expensive when carried out over broad areas (Gower et al., 1999). Remote sensing, using current or anticipated air-spaceborne sensors is widely viewed as a time- and cost-efficient way to proceed with large-scale estimation of forest structural attributes.

A variety of remote sensing products have been used in forest inventory studies including passive optical and active (radar and lidar) sensors (Nilsson, 1996; Kasischke et al., 1997; Lefsky et al., 1999). By far the most commonly used sensors are multispectral optical sensors such as Landsat TM/ETM+ and SPOT HVR with three to six broad spectral bands covering the visible, near infrared (NIR) and shortwave infrared (SWIR) regions (Woodcock et al., 1997; Franco-Lopez et al., 2001; Ingram et al., 2005). Mathematical transformations of NIR and visible reflectance into ratio indices (vegetation indices) are the most widely used remote sensing predictors of forest parameters. Many vegetation indices have been developed, but the normalised difference vegetation index (NDVI) developed by Rouse et al. (1974) is the most common in use. With increases in leaf area index (LAI), red reflectance decreases as leaf pigments absorb light, while NIR reflectance increases as more leaf layers are present to scatter the radiation (Gates et al., 1965). Thus, the spectral response of forest is directly determined by LAI, foliage mass and canopy biochemical contents and indirectly by the attributes that influence the canopy structure such as density, basal area, mean tree diameter and height (Lefsky et al., 1999; Ingram et al., 2005). However, the major limitation of using NDVI is that it saturates for a certain range of canopy cover or LAI (LAI > 3) (Sellers, 1985; Gao et al., 2000). In addition, NDVI is affected by soil and background reflectance (Huete, 1988), solar zenith and viewing angles (Middleton, 1991; Qi et al., 1995), and atmospheric conditions (Kaufman and Tanré, 1992).

Broadband indices use average spectral information over broad bandwidths, resulting in loss of critical information available in specific

narrowbands (Gong et al., 2003; Thenkabail et al., 2004). The advent of narrowband or hyperspectral sensors has raised new expectations about the possibilities of improving the estimation forest parameters. This is based on the assumption that increased identification of particular spectral features associated with narrowbands could improve estimation of forest attributes compared to broadband sensors (Lefsky et al., 1999; Lee et al., 2004). However, it is difficult to infer from existing literature whether hyperspectral sensors provide an improvement over multispectral sensors for remote sensing of forest structural attributes. For example, Lefsky et al. (2000) observed a slight increase in the ability of Airborne Visible-Infrared Imaging Spectrometer (AVIRIS) to predict forest stand attributes relative to single date Landsat TM data, but a better performance of multitemporal TM data. Gong et al (2003) showed that indices involving NIR and SWIR Hyperion bands were better than NIR-red indices for LAI estimation. Lee et al. (2004) found no improvement of AVIRIS NDVI over ETM+ NDVI for LAI estimation.

An additional disadvantage of using two-band vegetation indices such as NDVI is that it utilises a limited amount of the total spectral information available in an image (Lee et al., 2004; Ingram et al., 2005). Few studies have assessed statistical techniques that integrate information from several spectral bands. For example, multiple stepwise regression was used by Lefsky et al. (2001) and De Jong et al. (2003) to estimate forest parameters. However, this method might be affected by multicollinearity (De Jong et al., 2003). Alternatively, multivariate techniques such as artificial neural networks (ANN) and partial least squares (PLS) regression are capable of mitigating multicollinearity. Foody et al. (2003) and Ingram et al. (2005) using Landsat TM and ETM+, respectively showed that ANN produced higher correlation with forest structural attributes compared to traditional vegetation indices including NDVI. On the other hand, PLS regression has not been assessed for estimating forest parameters. However, Hansen and Schjoerring (2003) and Cho et al. (2006) using a handheld spectrometer and airborne HyMap data, respectively showed that PLS regression improves grass biomass estimation compared to NDVI.

The objective of this study was to investigate whether multivariate statistical models such as partial least squares (PLS) regression and artificial neural networks (ANN) perform better than univariate models based on vegetation indices for estimating and mapping forest structural

parameters of a closed canopy beech forest (*Fagus sylvatica* L). HyMap images and data on forest structural attributes including mean diameter-at-breast-height (DBH), mean tree height and tree density were collected from the Majella National Park, Italy in July 2004. The predictive performances of the various statistical techniques were evaluated using calibration (n = 33) and test (n = 20) data sets.

7.2 Material and methods

7.2.1 Study site

The study site was located in Majella National Park, Italy (latitude 41°52' to 42°14'N, longitude 13°14'E), covering an area of 74095 ha. The Park extends into the southern part of Abruzzo, at a distance of 40 km from the Adriatic Sea. This region is situated in the massifs of the Apennines. The park is characterised by several mountain peaks, the highest being mount Amaro (2794 m). The region is characterised by Mediterranean climate: hot and dry summers and cool and wet winters. The specific study site (latitude 41°49' to 42°14'N, longitude 13°57' to 14°3'E) is situated between mounts Majella and Morrone to the east and west, respectively. It covers an area of 40 by 5.5 km.

The Majella beech forest is located at altitude range of about 1200-1800m. Over the last 60 years, depopulation, changes in the socio-economic conditions and the creation of the National Park in 1995 have led to a pronounced drop in the local demand for small size timber, firewood and charcoal (Ciancio et al., 2006). As a consequence, many coppices are returning to high forest. However, a combination of thinning and the occurrence of avalanches in Majella have given rise to a compound coppice, which is a mixture of coppice and high forest.

7.2.2 Image acquisition and processing

Airborne HyMap data of the study site were obtained on 15 July and 4 July for 2004 and 2005, respectively. The flight was carried out by DLR, Germany's Aerospace Research Centre and Space Agency. The HyMap sensor comprised 126 wavebands, operating over the wavelength range 436 nm to 2485 nm, with average spectral resolutions of 15 nm (436 nm to 1313 nm), 13 nm (1409 nm to 1800 nm) and 17 nm (1953 nm to 2485 nm). The spatial resolution of the data was 4 m. The data were collected

at solar noon. The specific study site was covered by four image strips, each covering an area of about 40 km by 2.3 km. The solar zenith and azimuth angles for the image strips ranged between 30-33.7^{o} and 111.5-121^{o}, respectively. The 2004 and 2005 image strips were atmospherically corrected by DLR. But only the 2005 images were geometrically corrected by DLR. The 2004 images strips were then geometrically corrected from the 2005 images using image-to-image registration. The atmospheric correction was carried out using ATCOR4-r (rugged terrain).

7.2.3 Field measurements of forest stand attributes

Field data for mean DBH, mean height and number of trees were collected from 53 plots within the flight strips. Initially, 20 points were randomly generated using Arc GIS. The plots were located in the field with a GPS. Measurements were made from each randomly selected plot (30 m by 30 m) and from two to three other plots at about 150 and 300 m away. The data were only collected from closed canopy forest only. The DBH of all trees above 7 cm was measured while the tree heights of five to ten trees were measured using a Haga meter. The mean DBH and height were subsequently calculated per plot. Tree density was calculated as the number of trees per hectare.

7.2.4 Data analysis

A 7 by 7 pixels window was used to collect image spectra from each sample plot in order to avoid including pixels located outside the plot (30 m by 30 m). An average spectrum was subsequently calculated for each plot.

7.2.5 Data analysis

The forest parameters were predicted as continuous variables rather than as a set of discrete classes. Lefsky (2001) argues that continuous variable approach offers flexibility because the predictions can be used directly or arranged into multiple sets of classes that match varying purposes.

The predictive capabilities of models based on spectral indices, stepwise regression, PLS regression and ANN were investigated. The data was randomly split into the training or calibration (n = 33) and test (n = 20)

sets. Regression analyses were performed on the calibration set. Empirical validation of the calibration models was carried out using the test set.

Table 7.1 Spectral indices adopted in the study.

Spectral index	Formula	Biological significance	Reference
Simple ratio (SR)	R_{801}/R_{665}	Canopy greenness, Leaf area index (LAI)	Tucker, 1979
Normalised difference vegetation index (NDVI)	$(R_{801}–R_{665})/(R_{801} + R_{665})$	Canopy greenness, Leaf area index (LAI), fraction of photosynthetically active radiation	Rouse et al., 1974; Tucker, 1979
Narrowband normalised difference vegetation index (NNDVI)	$(R_{756}–R_{740})/(R_{756} + R_{740})$	Grass biomass	Mutanga and Skidmore, 2004
Soil adjusted and atmospherically resistant vegetation index (SARVI)	$R_{rb} = R_{red} - \gamma(R_{blue} - R_{red})$ γ = atmospheric aerosol correction function SARVI = $(R_{NIR} - R_{rb})/(R_{NIR} + R_{rb} + L)$ L = soil adjustment factor $\gamma = 1$, L = 0.5	Blue and red bands at 466 and 665 nm, respectively. SARVI minimises atmosphere and soil induced variations	Kaufman and Tanre, 1992; Huete et al., 1994
Modified ratio index (MSR)	$(R_{801}/R_{665} − 1)/(R_{801}/R_{665})1/2 +1)$	Canopy greenness, Leaf area index (LAI)	Chen, 1996
Normalised difference water index (NDWI)	$(R_{860\ nm} − R_{1240\ nm}) / (R_{860\ nm} + R_{1240\ nm})$	NDWI is sensitive to changes in liquid water content of vegetation canopies.	Gao, 1996
Carter index (CI)	R_{760}/R_{695}	Chlorophyll content	Carter, 1994
Vogelman index (VOG)	R_{740}/R_{720}	Chlorophyll content	Vogelmann et al., 1993
Red edge position (REP)	Linear extrapolation method	Chlorophyll or nitrogen content, grass biomass	Cho and Skidmore, 2006

The predictive performances of the various models were estimated and compared using the coefficient of determination (R^2) for calibration and validation, the standard error of calibration (SEC, Eq. 7.1) and standard error of prediction (SEP) based on the independent test data.

$$SEC = \sqrt{\frac{\sum_{i=1}^{n} (y - y')^2}{n}} \quad (7.1)$$

where y = measured DBH, height or density, y′ = predicted DBH, height or density and n = number of observations.

7.2.5.1 Spectral indices

The spectral response of forest is directly determined by LAI, foliage mass, leaf water content and canopy biochemical contents and indirectly by the attributes that influence the canopy structure such as density, basal area, mean tree diameter and height (Ingram et al., 2005; Lefsky et al., 1999). Thus, spectral indices related to LAI, leaf mass leaf water and pigments were investigated in this study (Table 7.1).

7.2.5.2 Stepwise regression

Forward stepwise multiple was used to relate band reflectance and forest structural attributes. The criterion for selecting the optimum number of bands was that the inclusion of an additional band into the regression model should significantly ($p<0.05$) increase the coefficient of determination (R^2).

7.2.5.3. Partial least squares regression (PLS)

PLS regression is a bilinear calibration method using data compression by reducing the large number of measured collinear spectral variables to a few non-correlated latent variables (Geladi and Kowalski, 1986; Geladi et al., 1999; Hansen and Schjoerring, 2003). As in multiple regression, the main purpose of PLS regression is to build a linear model (Eq. 7.2),

$$Y = Xb + E \quad (7.2)$$

where Y: mean-centred matrix containing the response variables (*n* by 1 matrix in this study), X: mean-centred matrix containing the predictor variables (the spectral bands in this study), b: matrix containing the regression coefficients and E: the matrix of residuals. In PLS regression, the above principle is used on latent variables of X. In this sense, PLS regression is closely related to principal component regression (Geladi and Kowalski, 1986 and Geladi et al., 1999). But instead of first decomposing the spectra into a set of eigenvectors and scores and regressing them against the response variables as a separate step, PLS regression actually uses the response variable information during the

decomposition process. Further information on the PLS regression can be obtained in Geladi and Kowalski (1986).

It has been shown that variable selection enhances the predictive performance of PLS regression (Kubinyi, 1996; Martens and Martens, 2000; Davies, 2001; Schmidtlein and Sassin, 2004). A sub-objective therefore, was to test PLS models based on all the HyMap bands and on a small number of selected bands. The selection was based on bands related to leaf chlorophyll, LAI and leaf mass (Table 7.2).

Before the PLS regression models were developed, the spectra and forest parameters were mean-centred, i.e. the average value for each variable was calculated from the calibration set and then subtracted from each corresponding variable. The root mean square error of leave-one-out cross validation (RMSECV) was used as a selection criterion to choose the optimum number of PLS factors or latent variables for predicting the forest structural parameters (Geladi and Kowalski, 1986; Viscarra Rossel, 2005). The RMSECV was determined for each cross-validation phase. The number of factors which yielded the lowest RMSECV was used to develop the calibration equations. The analyses were carried out using STATISTICA software (StatSoft, Inc.) and ParLes software developed by Viscarra Rossel (2005).

Table 7.2 Wavebands selected for estimating green grass biomass using partial least squares regression.

Waveband centre (nm)	Description	References
466	chlorophyll b	Curran, 1989
695	total chlorophyll	Carter, 1994; Gitelson and Merzylak, 1997
725	total chlorophyll, leaf mass	Horler et al, 1983
740	leaf mass and LAI	Horler et al. 1983
786	leaf mass	Guyot and Baret 1988
846	leaf mass, LAI, chlorophyll	Thenkabail et al. 2004
895	leaf mass, LAI	Thenkabail et al. 2004
1113	leaf mass, LAI	Thenkabail et al. 2004
1215	plant moisture, cellulose, starch	Thenkabail et al. 2004, Curran, 1989
1661	lignin, leaf mass, starch	Thenkabail et al. 2004
2173	protein, nitrogen	Curran, 1989
2359	cellulose, protein, nitrogen	Curran, 1989

7.2.5.4 Artificial neural networks (ANN)

ANNs are artificial intelligence techniques based on the function of the human brain (Atkinson and Tatnall, 1997). ANNs are non-parametric and have the advantages over traditional statistical approaches in their ability to handle non-linear relationships and non-Gaussian distribution of data (Mas et al., 2004; Miller et al., 1995; Paola and Schowengerdt, 1995). Although in early studies, ANNs were mostly used to classify land cover or vegetation types (Miller et al., 1995; Bruzzone et al., 1997; Skidmore et al., 1997), the method has shown great potential for predicting continuous forest data (Bacour et al., In press; Schlerf and Atzberger, 2006).

A multi-layer perceptron consisting of an input, one hidden and an output layer was used in this study. The ANN was trained by the back-propagation algorithm. In order to minimise the risk of over-fitting which may results from a large number of input variables (Uno et al., 2005), only the bands of known spectral features for foliar nutrients and structure were applied in the study (Table 7.2). Because of the small number of samples in the calibration data, cross validation was used as a resampling technique to generate a number of networks. The calibration data (n = 33) was divided into approximately ten equal parts. For each experiment, nine tenths of the data were used for the training while one part (n = 3) acted as the test set to determine the best network. Each experiment was performed with the same process parameters (number of epochs or iterations, learning rates and momentum). The best networks were subsequently formed into an ensemble and predictions on the independent test data (n = 20) were made by averaging the re-sampled networks. We also investigated the effect of using an increasing number of epochs on the prediction accuracy of the forest structural parameters. Skidmore et al. (1997) show that the system error decreases as the number of epoch increases.

7.3. Results

Only the results obtained from the 2004 HyMap image are reported in this study, first because parts of image strips 2 and 3 for 2005 were covered by clouds and secondly, none of the regression analysis adopted in the study produced a statistically significant relationship with mean DBH, height or tree density. As illustrated in Fig. 7.1 one difference

between 2004 and 2005 was that the vegetation was drier in 2004 than in 2005 at the time of image acquisition.

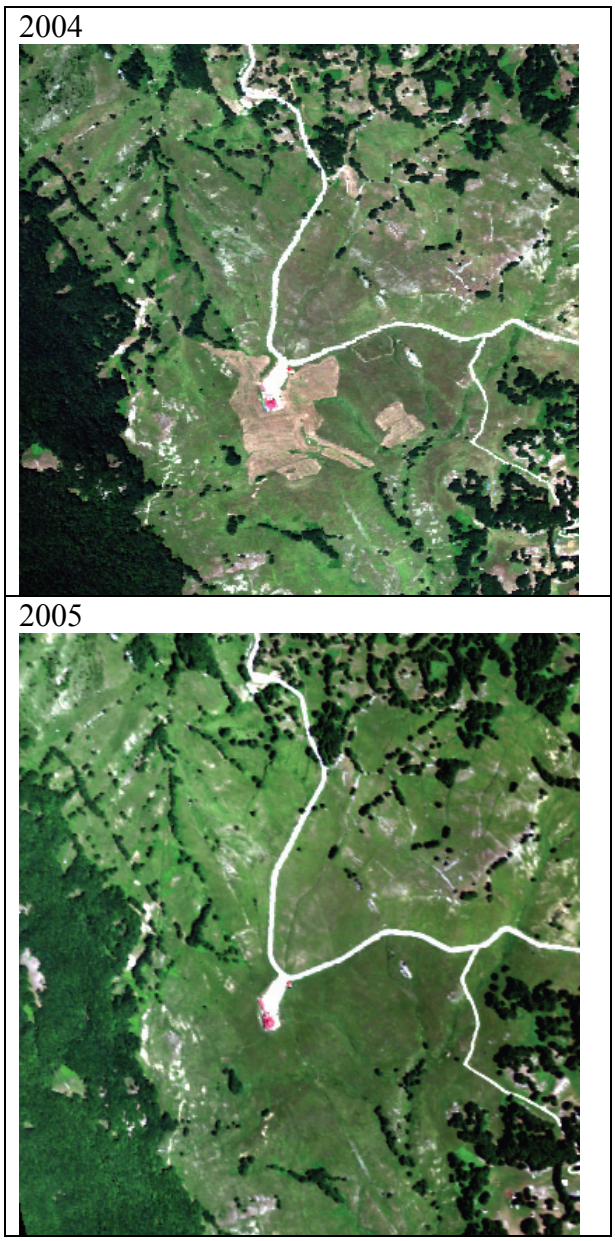


Fig. 7.1 True colour composite images for a subset area of HyMap images acquired on 15 July 2004 and 04 July 2005. The vegetation was drier in 2004.

7.3.1 Structural parameters

Table 7.3 Descriptive statistics of beech forest structural parameters.

Parameter	Mean (n = 53)	Minimum	Maximum	Standard deviation	Skewness
Mean DBH (cm)	19.94	8.00	43.07	8.02	1.01
Mean height (m)	18.70	7.00	45.00	7.23	1.35
Tree density (No. ha^{-1})	1208	222	3089	739	0.77

The descriptive statistics of the beech forest structural parameters are presented in Table 7.3. Each parameter showed a positive skewness indicating a bias of the distribution towards higher values. The Shapiro-Wilk test was used to test the data for normality, the hypotheses were, the null hypothesis (H_o): data follow a normal distribution versus the alternate hypothesis (H_1): the data do not follow a normal distribution. The null hypothesis was rejected in all cases ($p < 0.05$). The relationships among parameters were analysed using Spearman's rank correlation test (a non-parametric test). Mean DBH was positively related to mean height ($r = 0.70$, $p < 0.05$) but negatively related to tree density ($r = -0.91$, $p < 0.05$). Mean height was less highly related to density ($r = -0.60$) than to mean DBH.

7.3.2 Relationship between mean DBH, mean height or tree density and individual band reflectance

The relationships between forest parameters and individual band reflectance were analysed using Spearman's rank correlation test. Statistically significant ($p<0.05$) correlations were predominantly observed in the NIR (Fig. 7.2). The relationships were significant in the following regions:

Mean DBH and tree density: 711-1342 nm
Mean height: 528-589 nm, 725-1405 nm, 1530-1806 nm and 2257 nm.

Mean DBH and mean height on one hand and density on the other showed opposing correlation directions in the NIR. Mean DBH and mean height were negatively correlated with NIR bands, while density was positively correlated with the NIR bands.

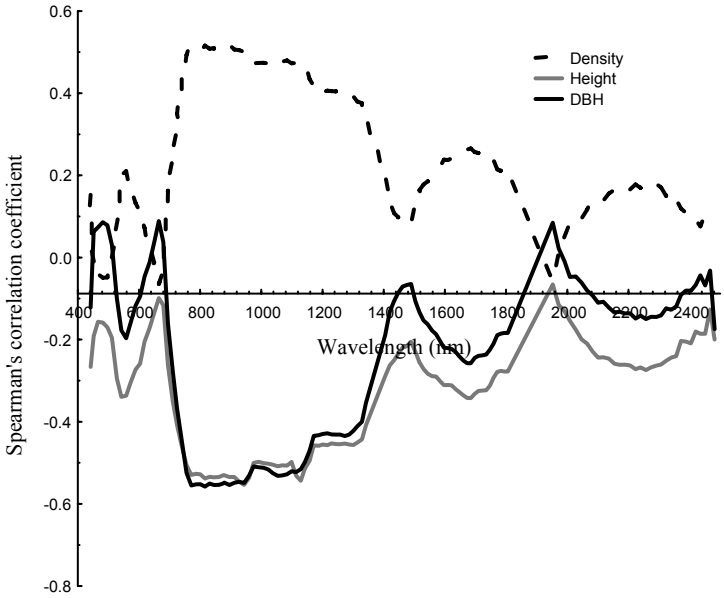


Fig. 7.2 Correlograms of mean tree diameter-at-breast height (DBH), mean height and tree density.

7.3.3 Predicting forest parameters

The predictive capabilities of spectral indices, stepwise regression, PLS regression and ANNs have been assessed. Although the data of mean DBH, mean height and density were not normally distributed; the use of parametric statistical analysis was justified assuming normality under the central limit theorem ($n \geq 30$)

Table 7.4 Predicting forest structural parameters; mean diameter at breast height (DBH), mean height and tree density using spectral indices. SEC = standard error of calibration, SEP = standard error of prediction.

Spectral indices	Calibration (n = 33)		Independent validation (n = 20)	
	R^2 actual vs. predicted	SEC	SEP	% of mean
Mean DBH (cm)				
SR	0.39**	6.39	6.75	34
NDVI	0.39**	6.39	6.64	33
SARVI	0.40**	6.34	6.03	30
MSR	0.39**	6.39	6.72	34
NNDVI	0.15**	7.56	7.32	37
CI	0.26**	7.08	6.84	34
VOG	0.12*	7.71	7.32	37
NDWI	0.43**	6.21	6.11	31
REP	0.07ns			
Mean height (m)				
SR	0.18*	6.29	6.84	37
NDVI	0.21**	6.19	6.76	36
SARVI	0.28**	5.89	6.25	33
MSR	0.19*	6.26	6.81	36
NNDVI	0.11ns			
CI	0.13*	6.48	7.02	38
VOG	0.07ns			
NDWI	0.19*	6.25	6.47	35
REP	0.05ns			
Density (no. trees ha^{-1})				
SR	0.40**	460	848	70
NDVI	0.38**	468	831	69
SARVI	0.40**	459	788	65
MSR	0.39**	461	839	69
NNDVI	0.18*	538	908	75
CI	0.27**	506	863	71
VOG	0.13*	553	903	75
NDWI	0.47**	428	850	70
REP	0.08ns			

* = $p < 0.05$, ** = $p < 0.01$

7.3.3.1 Using spectral indices

Mean DBH was the best-predicted parameter using the spectral indices. SARVI was the best linear predictor of mean DBH (SEP = 6.03 cm, 30% of the mean), mean height (SEP = 6.25, 33% of mean) and density (SEP

= 788 trees ha^{-1}, 65% of the mean) (Table 7.4). In general, the predictions of low mean DBH values were poor for both the calibration and test data (Fig. 7.3). For example, when three classes of mean DBH values were taken into account, namely, low DBH: less than 15 cm, medium DBH: 15-25 cm and high DBH: greater than 25 cm, then 2/8 (25%), 7/9 (78%) and 3/3 (100%) samples for low, medium and high mean DBH classes are correctly predicted on the test data (Fig. 7.3). This analysis revealed an overall prediction accuracy of 60% on the test data.

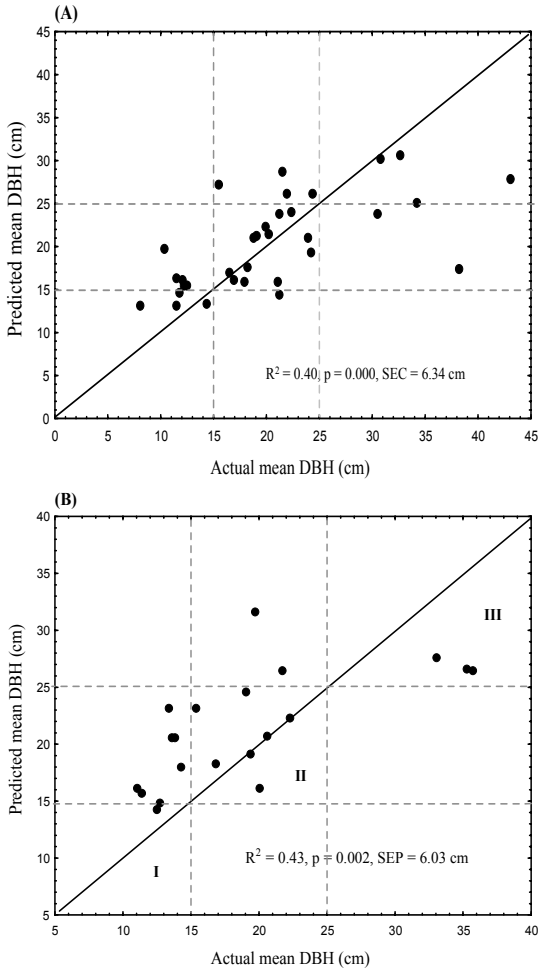


Fig. 7.3 Relationship between predicted and actual tree diameter at breast height (DBH) for (A) calibration and (B) validation analyses using soil adjusted and atmospherically resistant vegetation index (SARVI). SEC = standard error of calibration, SEP = standard error of prediction. I, II and III indicate low medium and high DBH classes.

The graphs of predicted versus actual mean DBH for both the calibration and validation analyses showed that higher values are predicted low and low values are predicted high (Fig 7.3). This phenomenon is known as 'local bias' or contraction (Geladi et al., 1999). Predicted mean DBH values below 15 cm and above 25 cm appear to saturate. Further analyses were carried out with three mean DBH subsets:

(i) Mean DBH less than 25 cm
(ii) Mean DBH greater than 15 cm and
(iii) Mean DBH greater than 15 cm but less than 25 cm

Despite the tendency for saturation at both the low (less than 15 cm) and high mean DBH (greater than 25 cm) values as shown in Fig. 7.3, the results in Fig. 7.4 seem to suggest that the performance of the mean DBH/SARVI regression model is limited by the low and high mean DBH values. The predictive performance of the regression model decreased in the absence of the low or high mean DBH values. The effect was more pronounced in the absence of the low values (Fig. 7.4B). There was no significant relationship between mean DBH and SARVI when both the low and high mean DBH values were kept out of the analysis (Fig. 7.4C).

The NDWI produced the highest calibration R^2 for mean DBH and tree density but produced higher prediction errors on the test data when compared with SARVI. The red-edge indices, namely, carter index, vogelman index and the REP performed poorly for all three structural parameters.

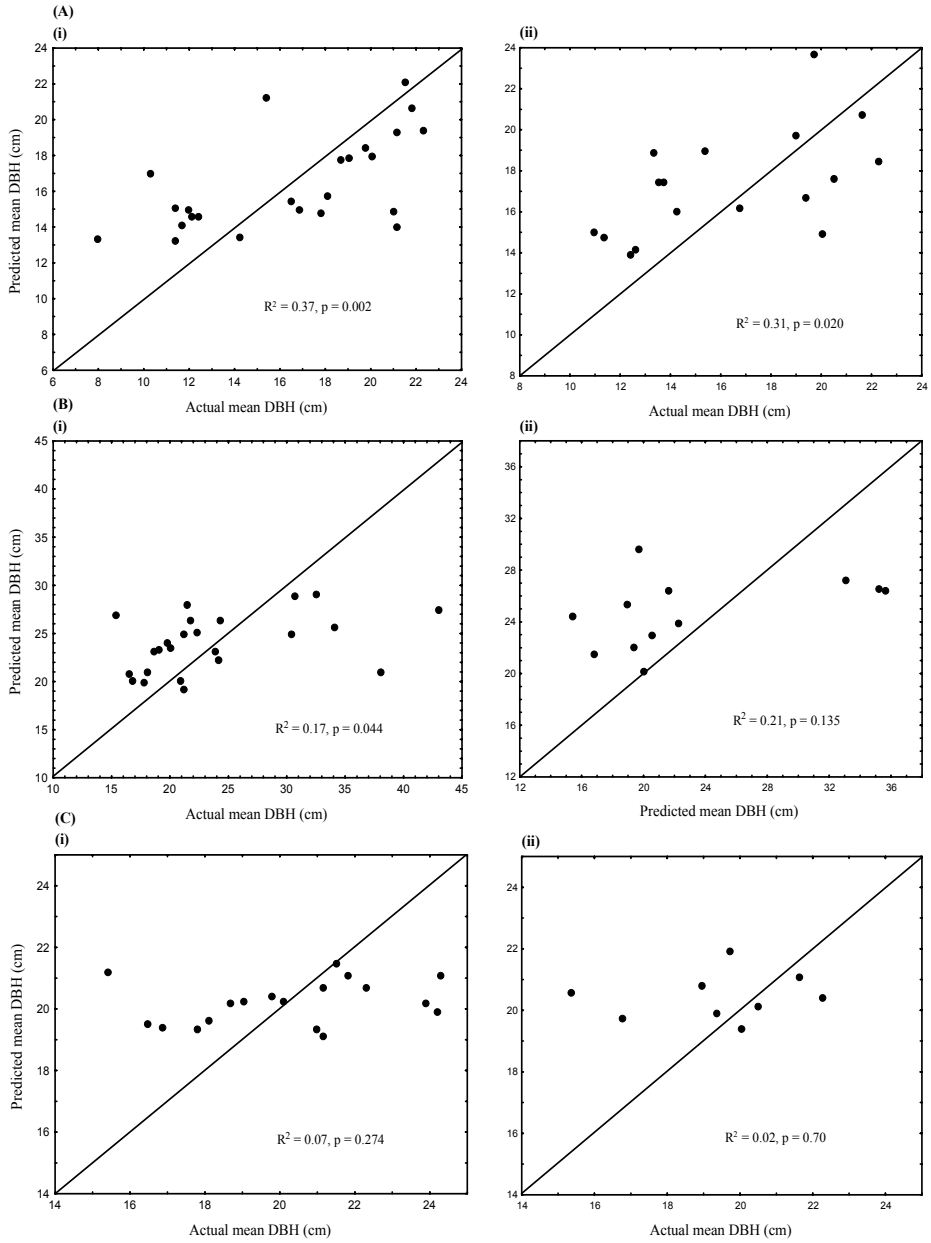


Fig.7.4 Calibration (i) and validation (ii) analyses using soil adjusted and atmospherically resistant vegetation index (SARVI) for (A) mean diameter at breast height (DBH) < 25 cm, (B) mean DBH > 15 cm and (C) 15 cm < mean DBH < 25 cm.

Table 7.5 Predicting forest structural parameters; mean diameter at breast height (DBH), mean height and tree density using forward stepwise regression. SEC = standard error of calibration, SEP = standard error of prediction.

Forest parameter	Bands (nm)	Calibration (n = 33)		Independent validation (n = 20)	
		R^2 actual vs predicted	SEC	SEP	% of mean
Mean DBH (cm)	604, 619, 680, 771, 1158, 1258	0.80**	3.36	7.94	40
Mean Height (m)	695, 1129	0.39**	5.40	6.11	33
Density (no. of trees ha^{-1})	771, 1287	0.61**	370	845	70

** = $p < 0.01$

7.3.3.2 Using stepwise regression

The prediction of mean DBH using multiple stepwise regression (Table 7.5) was poorer (SEP = 7.94 cm) than that of most vegetation indices. The low prediction could be attributed to collinearity amongst the predictor wavebands, i.e. bands at 604, 619, 680, 771, 1158 and 1258 nm. The collinearity is reflected by the high R^2 obtained in the calibration analysis. We used the variance inflation factor (Brauner and Shacham, 1998) computed from Eq. 7.3 to detect the bands causing collinearity.

$$VIF_j = 1/(1-R^2_j) \quad (7.3)$$

where R^2_j is the multiple correlation coefficient of x_j (i.e. the j^{th} predictor band) regressed on the remaining columns of the X matrix (i.e. predictor bands). The value of VIF_j is calculated for $j = 0, 1, \ldots, n$ (all the columns of the matrix X). A high level of collinearity leads the R^2_j value close to 1, which causes VIF_j to attain a large positive value (Brauner and Shacham, 1998). This process showed that bands located at 604, 619, 1158, and 1258 nm contributed to the high collinearity. When the multiple regression analysis was conducted with predictor bands at 680 and 771 nm after eliminating the bands that cause harmful collinearity, the SEP dropped to 6.13 cm. The forward stepwise process yielded only two predictor bands for the estimation of mean tree height or density. The prediction of mean height was comparable to that of SARVI.

7.3.3.3 Using partial least squares regression

The predictive performances of PLS regression based on all the HyMap bands and selected bands were basically similar (Table 7.6). Like in the case of spectral indices, mean DBH was the best-predicted parameter, followed by mean height and lastly density. However, there was a slight improvement in the prediction accuracy of mean DBH (SEP = 5.50 cm, 28% of mean) compared with SARVI (SEP = 6.03 cm, 30% of mean). As was the case with the regression model involving SARVI, the results of the PLS modelling showed a strong 'local bias' in the calibration and test data (Fig. 7.5). The higher values were predicted low and the low values were predicted high.

When three classes of mean DBH were taken into account as suggested in section 3.3.1, 1/8 (25%), 8/9 (89%) and 3/3 (100%) samples for low, medium and high mean DBH classes were correctly predicted on the test data. As with SARVI, this resulted into an overall prediction accuracy of 60%. There was some inconsistency in the prediction of the low mean DBH values between the calibration and test data. As could be observed in Fig. 7.5, the calibration model performed better than the validation model in predicting the low mean DBH values.

When further analyses were carried out using three subsets of the data as suggested in section 7.3.3.1, a different pattern (Fig. 7.6) emerged compared with the results for SARVI: (i) the absence of high mean DBH values enhanced the calibration R^2 but led to a poorer prediction on the test data (ii) while the absence of the low DBH value reduced the calibration R^2 but resulted into a better prediction on the test data and (iii) when both low and high values were left out, the calibration and validation models remained significant. In contrast to SARVI, the calibration R^2 was significant in all three cases. More importantly, leaving out the low mean DBH values reduced the 'local bias' around the 1:1 diagonal line.

Table 7.6 Performance of partial least squares (PLS) regression for predicting mean diameter at breast height (DBH), mean height and tree density in Majella National Park, Italy. R^2 = coefficient of determination, RMSECV = root mean square error of cross validation, SEC = standard error of calibration and SEP = standard error of prediction.

	Calibration (n = 33)				Independent validation (n = 20)	
	No. of PLS factors	RMSECV	R^2 Actual vs predicted	SEC	SEP	% of mean
All bands						
Mean DBH (cm)	3	6.56	53**	5.63	5.66	28
Mean Height (m)	2	6.15	36**	5.54	6.12	33
Density (no. of trees ha^{-1})	2	461	50**	420	828	68
Selected bands						
Mean DBH (cm)	3	6.54	51**	5.74	5.50	28
Mean Height (m)	2	6.16	37**	5.52	6.21	33
Density (no. of trees ha^{-1})	3	451	50**	421	824	67

** = $p < 0.01$

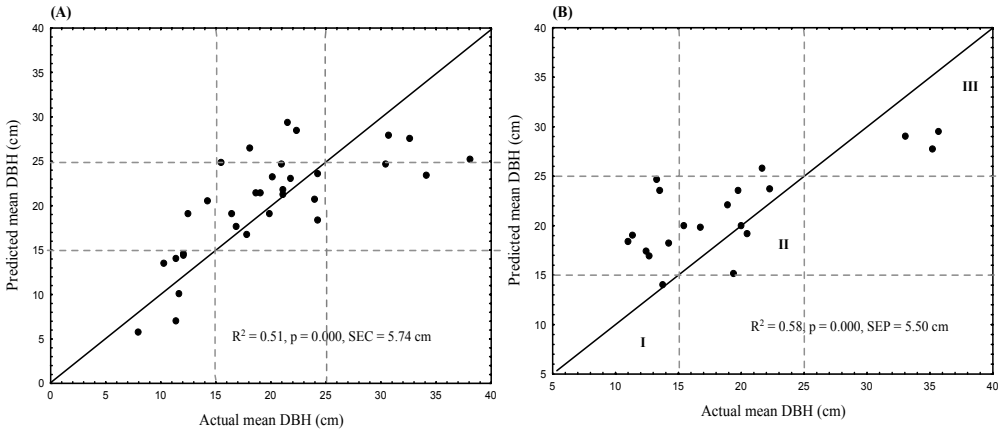


Fig. 7.5 Relationship between predicted and actual tree diameter at breast height (DBH) for (A) calibration and (B) validation analyses using partial least squares regression. SEC = standard error of calibration, SEP = standard error of prediction. I, II and III indicate low medium and high DBH classes.

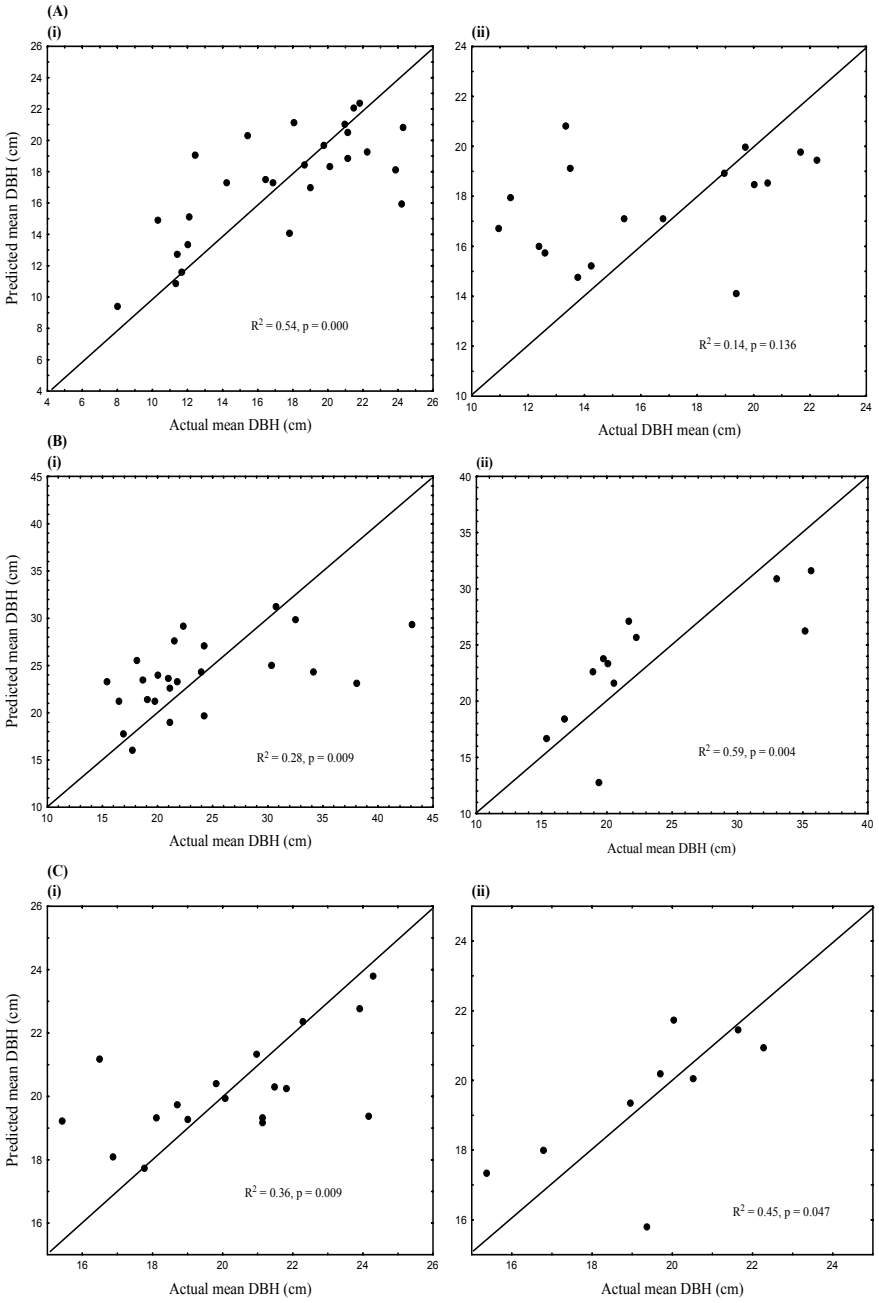


Fig. 7.6 Calibration (i) and validation (ii) analyses using partial least squares for (A) mean diameter at breast height (DBH) < 25 cm, (B) mean DBH > 15 cm and (C) 15 cm < mean DBH < 25 cm.

7.3.3.4 Using artificial neural networks

The predictive performance of the ANN depends on the number of epochs (Fig. 7.7). The SEC decreases with increasing number of epochs up to a point where the data becomes erratic. We used the lowest SEC value corresponding to the lowest number of epochs in the comparative analyses (Table 7.7). The SEP for DBH based on the selected bands (SEP = 5.80 cm, 29% of mean) was slightly higher than that of the PLS regression model (SEP = 5.50 cm, 28% of mean) but lower than that of SARVI (SEP = 6.03 cm, 31% of mean). ANN model was the best predictor of tree height (SEP = 4.61, 30% of mean) among all the statistical methods adopted in this study. However, as was observed for the prediction of DBH using SARVI and PLS models, the graphs of the predicted versus actual height also showed the phenomenon of 'local bias' (Fig. 7.8).

Table 7.7 Artificial neural networks.

	Calibration (n = 33)		Independent validation (n = 20)	
Forest parameter	R^2 actual vs predicted	SEC	SEP	% of mean
DBH (cm)	0.56	5.50	5.80	29
Height (m)	0.56	4.70	5.61	30
Density (no. of trees ha^{-1})	0.74	301	825	68

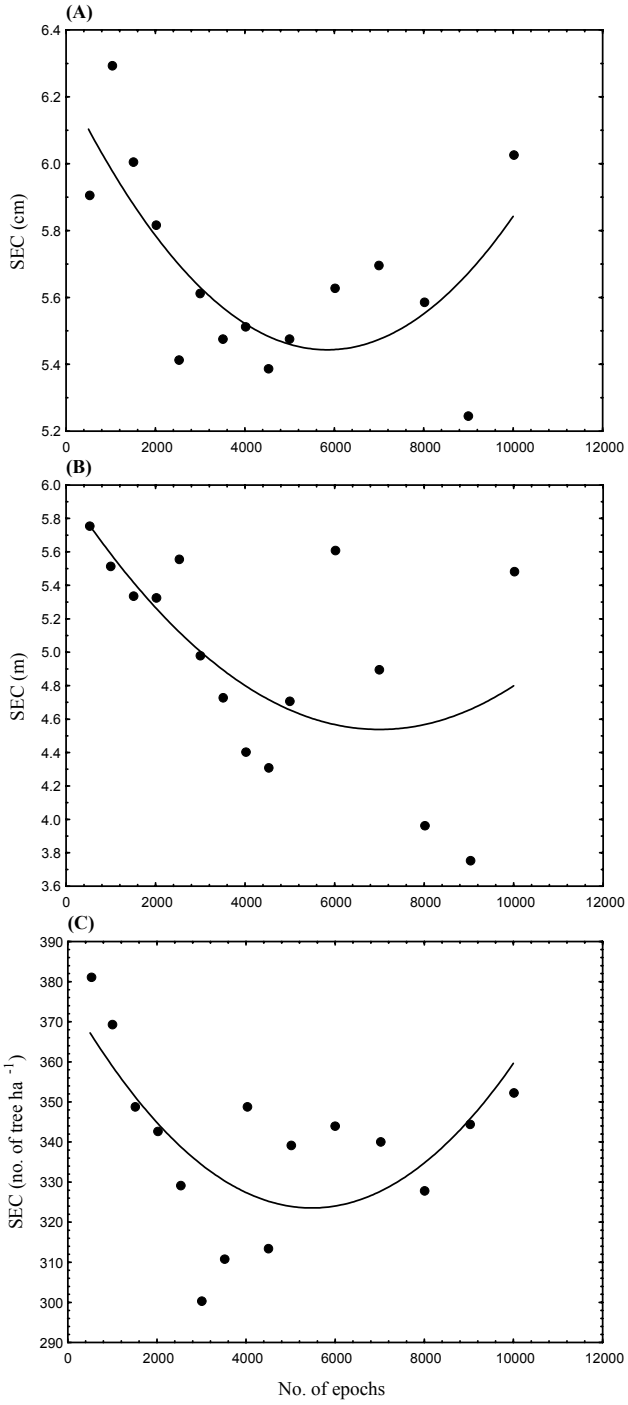


Fig. 7.7. Relationship between standard error of calibration and number of epochs for estimating (A) mean tree diameter-at-breast-height (DBH), (B) mean height and (C) tree density using artificial neural networks.

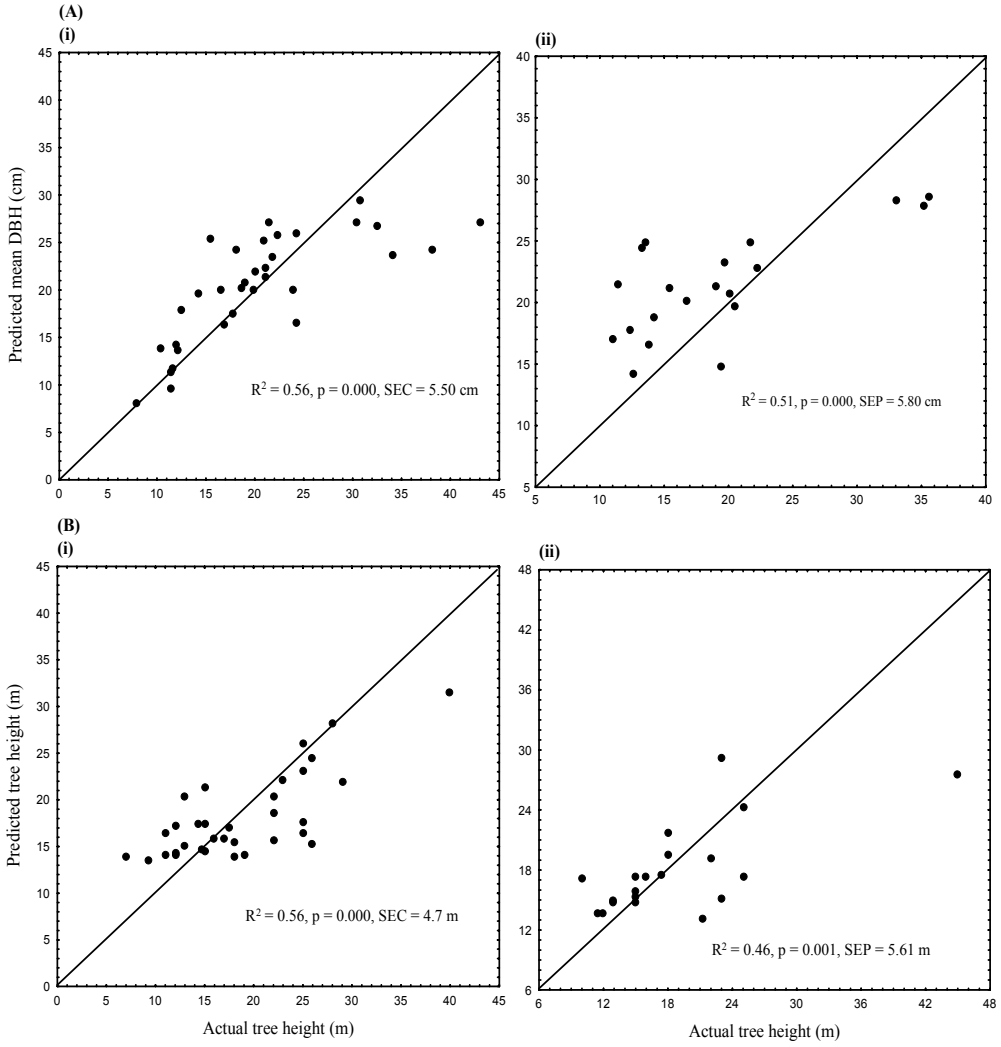


Fig. 7.8 Relationship between predicted and actual values for (A) mean tree diameter-at-breast-height (DBH) and (B) mean tree height using artificial neural networks. (i) calibration and (ii) validation analyses. SEC = standard error of calibration, SEP = standard error of prediction.

7.3.6 Mapping forest structure

The Majella beech forest structure was mapped using the best-predicted parameter, i.e., mean DBH. Mean DBH maps were produced using SARVI and PLS calibration models. Before the maps were produced, beech forest areas were subset from the HyMap image strips using SARVI threshold values, thus eliminating areas occupied by other land-

cover types (mainly grasslands and housing areas). The following calibration equation was used for mapping mean DBH with SARVI:

$$\text{Mean DBH} = 469.83 - 499.4*\text{SARVI} \quad (7.4)$$

The regression coefficients for the PLS model are presented in Table 8. The reflectance values of the predictor bands for each image pixel were mean centred, i.e., by subtracting the mean value in the calibration data set from the pixel value. The mean-centred DBH for each pixel was subsequently calculated from Eq. 7.5.

$$y = x_1b_1+x_2b_2+\ldots+x_nb_n \quad (7.5)$$

where y = mean-centred DBH, x_1 to x_n = predictor bands and b_1 to b_n = regression coefficients. The final mean DBH value for each image pixel was calculated from Eq.7.6.

$$\text{Mean DBH} = \text{predicted mean-centred DBH} + \text{mean DBH of the calibration data} \quad (7.6)$$

Table 7.8 Regression coefficients of partial least squares regression used in predicting beach forest diameter at breast height (DBH).

Predictor bands (nm)	Regression coefficients
466	0.717
695	0.860
725	0.207
740	-0.975
786	-1.795
846	-1.104
895	-0.473
1113	0.870
1215	2.013
1661	1.829
2173	1.094
2359	0.765

The predicted maps of mean DBH using SARVI and PLS models are presented in Fig. 7.9 and their corresponding histograms in Fig. 7.10. The mean DBH for all the image pixels predicted by PLS regression (18.74 cm) was closer to the actual sample mean (18.70 cm, n = 53) compared with that of SARVI (15.80 cm) (Table 7.9). When the 95% confidence interval is taken into account, the predicted DBH values lie between 6.12 cm and 31.36 cm for mean DBH predicted by PLS regression compared

to 2.38 cm to 29.22 cm for SARVI. Apart from the above statistical results, we could also affirm that the spatial distribution of DBH as predicted by PLS regression is more similar to what we observed on the ground during the field campaign than that predicted by SARVI. The map of DBH shows a high heterogeneity of DBH within the various forest patches. There is no clear effect of altitude on the forest structure.

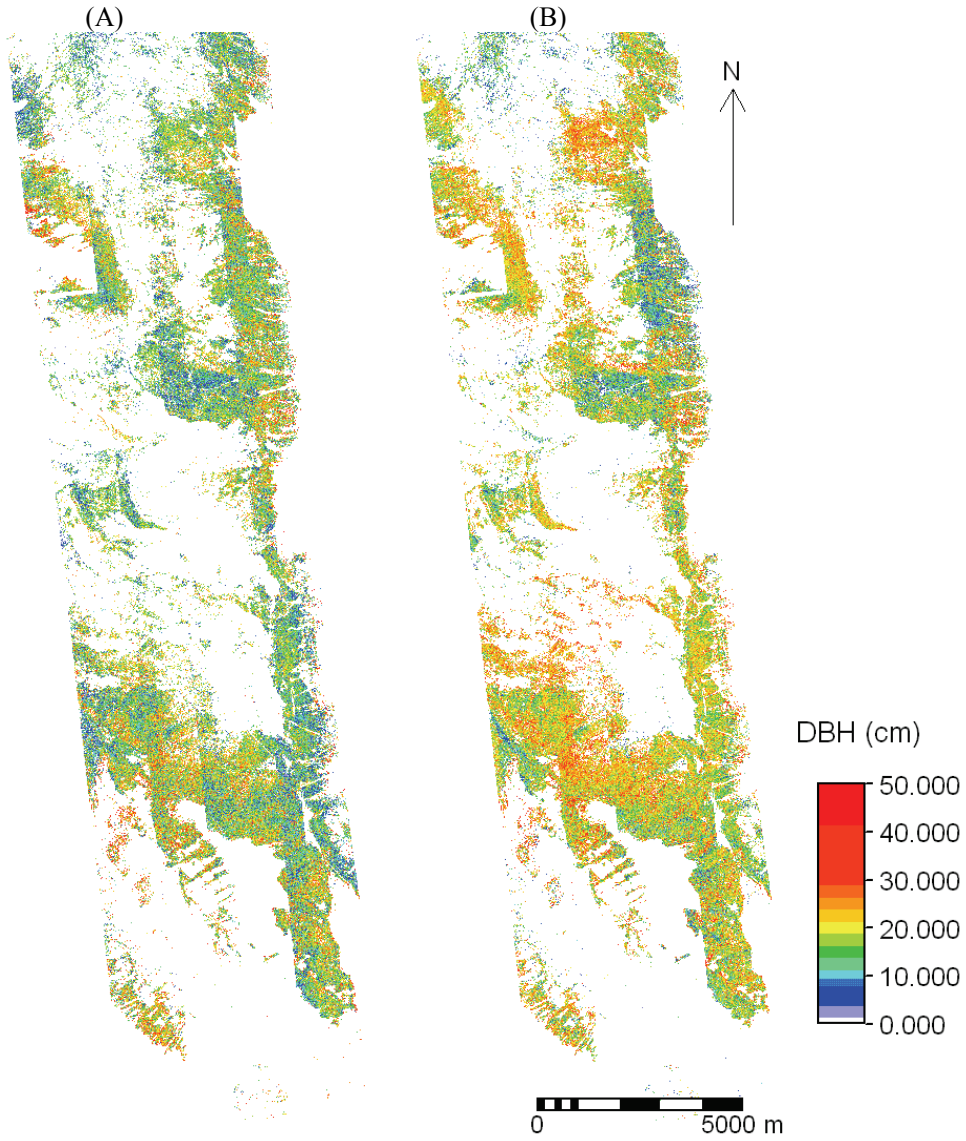


Fig. 7.9 Predicted maps of beech forest tree diameter at breast height in the Majella National Park, Italy using: (A) soil adjusted and atmospherically resistant vegetation index (SARVI) and (B) partial least squares regression.

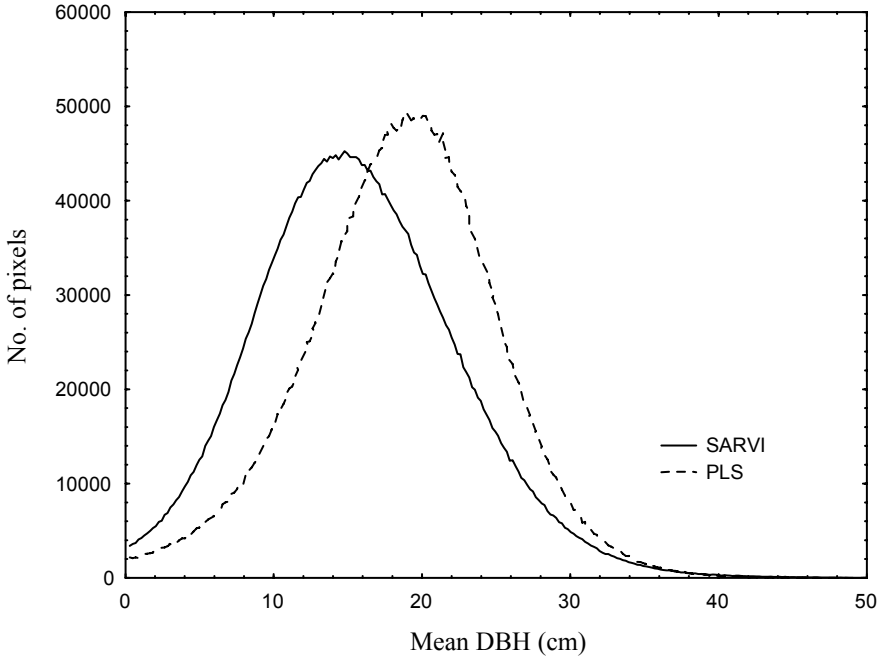


Fig. 7.10 Histograms of predicted diameter at breast height using soil adjusted and atmospherically resistant vegetation index and partial least squares regression.

Table 7.9 Statistics of predicted tree diameter at breast height using soil adjusted and atmospherically resistant vegetation index and partial least squares regression for 3749094 pixel elements.

Method	Minimum	Maximum	Mean	Standard deviation	- 95% CI	+95 CI
SARVI	0.10	49.98	15.80	6.71	2.38	29.22
Partial least squares	0.10	49.99	18.74	6.31	6.12	31.36

CI = confidence interval

7.4 Discussion

The main objective was to investigate whether multivariate techniques perform better than univariate regression methods based on vegetation indices in estimating and mapping forest structural parameters.

7.4.1 Predictive performance of various methods

Overall, mean DBH and mean height were better predicted using the various statistical methods compared with tree density. The prediction

errors for mean DBH and mean height ranged between 28 to 38% of the mean compared with tree density (65 to 75% of mean). PLS regression and ANN, the multivariate statistical methods adopted in this study, respectively, showed a slight improvement in the prediction of mean DBH (SEP = 5.50 cm, 28 % of mean) and mean height (5.61 cm, 30% of mean) compared with univariate regression models based on vegetation indices. PLS regression has rarely been applied for estimating forest attributes from remotely sensed data. However, in two related studies Hansen and Schjoerring (2003) and Cho et al. (In press) showed that PLS regression performs better than univariate techniques involving vegetation indices in predicting grass biomass. PLS regression has some advantages over ANN: (i) PLS is simple to implement and (ii) does not produce variable estimation accuracy in contrast to ANN whereby, the estimation accuracy varies because of the stochastic nature of the model (Skidmore et al, 1997). In this study, PLS regression models based on a few selected bands and on all HyMap bands produced similar calibration and validation accuracies. The spectral information content required for estimating forest structural parameters might be contained in a few narrowbands. Therefore, an optimum band selection procedure would enhance the PLS model parsimony.

SARVI was the best vegetation index for the prediction of mean DBH (6.03 cm, 30% of mean) and mean height (6.25 cm, 33% of mean). SARVI might have performed better than NDVI because of its ability to mitigate atmospheric effects. The prediction of mean DBH by multiple stepwise regression was affected by collinearity amongst the predictor bands, thus supporting the assertions by Curran (1989) and De Jong et al. (2003).

The prediction of the forest structural attributes in this study reveals the phenomenon of ‘local bias’. Local bias occurs when high values of the response variable are predicted low and the low values predicted high (Geladi et al, 1999). Geladi et al. (1999) argue that some of the deviations from the diagonal representing the 1:1 relationship between the predicted and actual values may be attributed to random noise. However, when the bias becomes systematic, as was the case in our study, it may be attributed to non-linearity in the true physical relationship (Geladi et al. 1998). Ingram et al. (2006) also observed local bias in the prediction of basal area of tropical forest in southeastern Madagascar. They attributed it to changes in basal area that might have

occurred between time of field data collection and the acquisition of the satellite imagery used in the study. But we argue that the 'local bias' observed in our study may be largely attributed to non-linearity in relationship between forest structural parameters and spectral data. The saturation of the spectral signal in dense and multi-layered canopy cover is a well-known phenomenon (Sellers, 1985; Gao et al., 2000). Although several authors argue that PLS and ANN can be used to model nonlinear relationships (Miller et al., 1995; Paola and Schowengerdt, 1995; Mas et al., 2004), albeit mildly for PLS (Geladi et al., 1999), PLS and ANN modelling of beech forest parameters as illustrated in this study still produced non-linear relationships. Nevertheless, they provided an improvement over traditional vegetation indices for estimating mean DBH and mean tree height. However, a possible drawback of empirical models in remote sensing of vegetation parameters as observed in this study and suggested by Curran (1994) and Gobron et al. (1997) is that they might be site, sensor or season specific.

The red-edge indices, namely, Carter index and Vogelman index, and REP showed poor predictive capabilities for forest structural parameters. These indices have shown high sensitivity to leaf chlorophyll and grass biomass (Vogelmann et al., 1993; Carter, 1994; Mutanga and Skidmore, 2004; Cho et al., 2006). Their performance for predicting structural parameters could be limited by foliage chlorophyll content because the beech forest parameters were weakly correlated with the chlorophyll (visible) spectrum (see Fig. 7.2).

7.4.2 Predictive map of DBH and implications for beech forest management

The maps show a high heterogeneity of mean DBH within the various forest patches. This pattern could be attributed to the forest management practice in the park. A combination of thinning and the occurrence of avalanches in the Majella National Park, have given rise to a compound coppice, which is a mixture of coppice and high beech forest. Thus, the spatial information of mean DBH revealed by the maps could be used to identify areas for specific management measures, e.g. thinning of low mean DBH coppice to promote diameter growth, to assess the effects of management on the beech forest or to detect changes in the forest structure caused by external factors such as avalanches.

7.5 Conclusions

The results of this study highlight the potential utility of multivariate techniques such as ANNs and PLS regression based on several narrowbands for providing improved estimates of beech forest structural attributes compared with univariate methods based on vegetation indices. PLS regression has rarely been applied in this sense and thus offers new possibilities for mapping and monitoring beech forest structure.

Acknowledgments

The International Institute for Geo-Information Science and Earth Observation (ITC) provided financial support for this study. We extend our gratitude to Fabio Corsi, Roshanak Darvishzadeh and Jane Bemigisha all colleagues of the Natural Resources Department of ITC for their assistance during the field campaign. We also appreciate the generous help provided by many people at the Majella National Park, Italy and particularly by Dr Theodoro Andrisano.

Chapter 8

Synthesis

Estimating biochemical and biophysical parameters with hyperspectral remote sensing

The derivative red-edge "double peak feature", a nuisance or an opportunity?

8.1 Introduction

Improved quantification and monitoring of biochemical and biophysical attributes is required to acquire deeper understanding of the carbon cycle and predict the response of ecosystems to climate change (Scurlock and Prince, 1993). Remote sensing, using current or anticipated technology, is widely viewed as a time- and cost-efficient way to proceed with large-scale monitoring of vegetation parameters. However, the information that can be extracted from remotely sensed data depends on the sophistication of the sensors, our understanding of how radiation interacts with a vegetation canopy and on spectral information extraction techniques (Curran, 1989).

Most studies have focused on developing empirical relationships between ground-measured vegetation parameters and spectral indices commonly known as vegetation indices. The most frequently used remote sensing product is the normalised difference vegetation index (NDVI) (Rouse et al. 1974) derived from broadband or multispectral sensors e.g. Landsat TM/ETM+ with seven spectral bands (Richardson et al., 1983; Wiegand et al., 1991; Anderson et al., 1993; Duchemin, 1999; Van Wagtendonk and Root, 2003). Multispectral sensors integrate radiance data over wide bands of the electromagnetic spectrum, resulting in loss of critical information available in specific narrowbands (Wessman et al., 1989; Gong et al., 2003). Several studies show that broadband NDVI can be unstable, varying with soil colour, atmospheric conditions and illumination/view angles (Huete and Jackson, 1988, Middleton, 1991, Kaufman and Tanré, 1992, Qi et al. 1995, Todd et al. 1998). Furthermore, NDVI also asymptotically saturates at higher biomass or LAI (Sellers, 1985, Gao et al. 2000).

The advent of hyperspectral remote sensing or imaging spectrometry provides additional bands within the visible, near-infrared (NIR) and shortwave infrared (SWIR). Most hyperspectral sensors (e.g. Airborne visible/infrared Imaging Spectrometer (AVIRIS) and Hyperion) acquire radiance information in less than 10 nm bandwidths (Asner, 1998). Hyperspectral data contain more information on subtle spectral features, which could otherwise be masked by the broad nature of multispectral imagery (Wessman et al., 1989). For example, the spectral shift of the red-edge (670-780 nm) slope and its inflection point termed the red-edge position (REP) associated with leaf chlorophyll content, phenological

state and vegetation stress, is not accessible with broadband sensors (Collins, et. al., 1977; Horler, et al., 1983). Furthermore, leaf chlorophyll concentration can be used to provide more accurate estimates of productivity (Chappelle et al., 1992) and to improve the discrimination of species by monitoring the phenological dynamics (Blackburn, 1998). The REP could be considered as a spectral index of interest because of its low sensitivity to disturbing factors such as atmospheric conditions and soil brightness (Clevers et al., 2001).

The REP is commonly defined by the maximum first derivative of the red-edge spectrum. However, the limitation of this approach is that the maximum first derivatives of contiguous spectra have been shown to occur within two principal spectral regions (around 700 and 725 nm) causing a bimodal distribution of REP data and a discontinuity in the REP/chlorophyll relationship (Horler et al. 1983). Experimental studies show that low leaf chlorophyll concentration is associated with REP values near 700 nm, while high chlorophyll concentration in combination with leaf internal scattering influence REP values near 725 nm (Horler et al., 1980; Boochs et al., 1990; Lamb et al., 2002). Zarco-Tejada et al. (2003) demonstrated in an experimental study that the existence of the double peak is due to chlorophyll fluorescence emission at about 690 nm and 730 nm. Earlier studies had shown that the chlorophyll fluorescence ratio (F690/F730) decreases with increasing chlorophyll content of developing leaves (Hák et al., 1990; Babani et al., 1996) and increases during autumnal chlorophyll breakdown of various tree leaves (D'Ambrosio et al., 1992). Thus, could more accurate estimates of leaf chlorophyll or leaf nitrogen concentration be obtained with REP derived by tracking changes that occur at the dominant peaks?

The spectral response of vegetation is directly determined by leaf and stem area, leaf and stem orientation, foliage clumping and leaf biochemical contents (Asner, 1998; Jacquemoud et al., 1995; Lefsky et al., 1999) and indirectly by canopy structural attributes such as biomass, tree diameter at breast height (DBH) and tree height (Lefsky et al., 1999), as well as non-plant attributes such as soil colour, shadows and dead material (Asner, 1998). Thus, the estimation of biomass, DBH and tree height may depend on fully exploiting the spectral information content of leaf and stem biophysical and biochemical spectral features that directly influence vegetation reflectance. However, two-band vegetation indices utilise a limited amount of the total spectral information available in a

high spectral resolution image (Lee et al. 2004). The question is whether multivariate statistical techniques based on more than two bands can improve estimation of grass biomass and forest structural parameters when compared to vegetation indices?

Thus, the objectives of this study were:

(i) to develop a technique for locating the REP that mitigates the destabilising effect of the double-peak feature on the REP data and tracks changes around the dominant chlorophyll-sensitive peaks (700 and 725 nm)
(ii) to test the performance of the new method vis-à-vis other REP techniques for estimating foliar chlorophyll and nitrogen concentrations, discriminating species and estimating green grass biomass under a wide range of canopy and environmental conditions.
(iii) to investigate the performance of other empirical techniques based on multiple hyperspectral bands for estimating grass biomass and forest structural attributes.

8.2 Towards red-edge positions less sensitive to canopy structure for chlorophyll/nitrogen estimation

We developed a new technique for estimating the REP that maximally estimates leaf chlorophyll or nitrogen content with minimal effects from LAI, leaf thickness, leaf mass and leaf developmental stage. The new technique is designed to mitigate the destabilising effect of the derivative double-peak feature on the correlation between nitrogen and REP (determined as the maximum first derivative) and to track variations near the low and high chlorophyll sensitive wavebands (700 nm and 725 nm) as illustrated in Fig. 8.1.

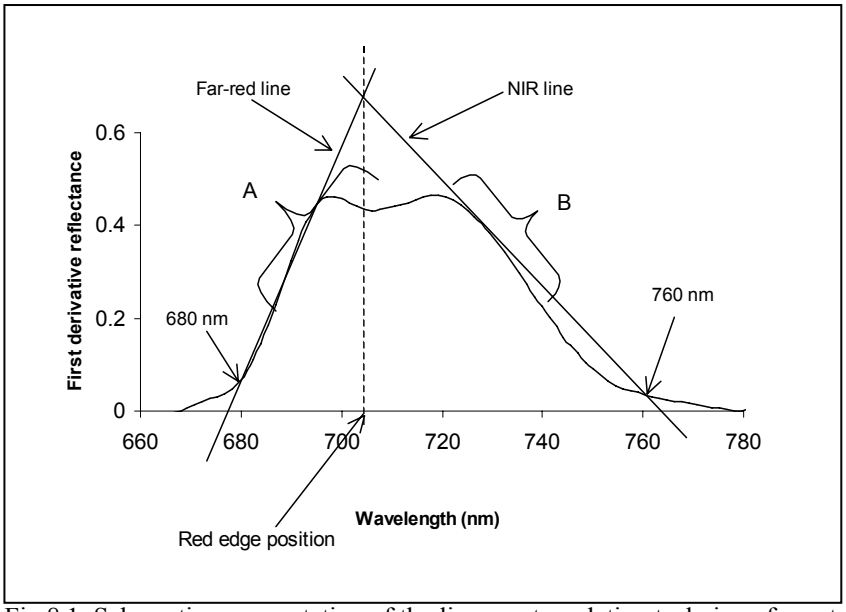


Fig.8.1. Schematic representation of the linear extrapolation technique for extracting the red edge position (REP) – wavelength of the meeting point between two straight lines extrapolated on the far-red and NIR flanks of the first derivative spectrum.

The new technique is based on linear extrapolation of two straight lines (Eqs. 8.1 and 8.2) through two points on the far-red (680 nm to 700 nm) and two points on the NIR (725 nm to 760 nm) flanks of the first derivative reflectance spectrum of the red edge region as illustrated in Fig. 8.1. The REP is then defined by the wavelength value at the intersection of the straight lines (Eq. 8.3).

Far-red line: $FDR = m_1\lambda + c_1$ (8.1)

NIR line: $FDR = m_2\lambda + c_2$ (8.2)

where m and c represent the slope and intercept of the straight lines, and FDR = first derivative reflectance. At the intersection, the two lines have equal λ (wavelength) and FDR values. Therefore, the REP, which is the λ at the intersection, is given by:

$$REP = \frac{-(c_1 - c_2)}{(m_1 - m_2)} \quad (8.3)$$

In summary, the technique is simple to implement, as is the case with the linear four-point interpolation method developed by Guyot and Baret (1988). Only four coordinate points (or wavebands) are required to calculate the REP by the linear extrapolation method; for instance, two bands near 680 nm and near 700 nm to calculate m_1 and c_1 for the far-red line and two bands near 725 nm and near 760 nm to calculate m_2 and c_2 for the NIR line.

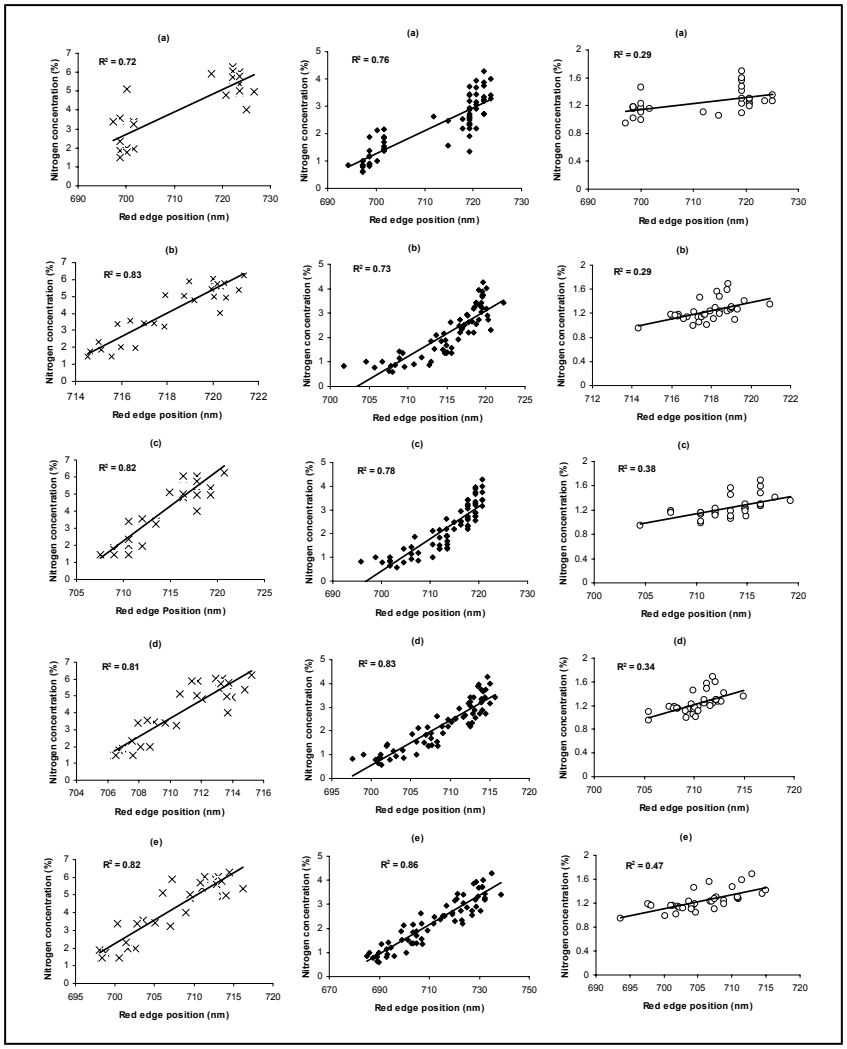


Fig. 8.2. Straight-line model fits for the regression between foliar nitrogen concentration and red edge position (REP) extracted using (a) maximum first derivative, (b) linear interpolation, (c) polynomial fitting, (d) inverted Gaussian modelling and (e) linear extrapolation (new approach) techniques. x = rye canopy, ♦ = maize leaves and ○ = mixed grass/herb leaf stack spectral data sets.

Far-red wavebands at 680 nm and 694 nm in combination with NIR wavebands at 724 nm and 760 nm or at 732 nm and 760 nm were identified as the optimal combinations for calculating nitrogen-sensitive REPs for three spectral data sets (rye canopy, maize leaves at different stages of development and mixed grass/herb leaf stacks). REPs extracted by the linear extrapolation method showed high correlations with a wide range of foliar nitrogen concentrations for both narrow and wider bandwidth spectra, being comparable with results obtained using the traditional linear four-point interpolation (Guyot and Baret, 1988), polynomial and inverted Gaussian fitting techniques (Bonham-Carter, 1988). However, the linear extrapolation method performed better than various REP alternatives in the case of maize leaves at different stages of development and mixed grass/herb leaf stacks (Fig. 8.2). This suggests that the relationship between linear extrapolation REP and leaf nitrogen concentration might be minimally confounded by differences in leaf phenology, leaf density or species.

The performance of the linear extrapolation method for chlorophyll content was further investigated under extreme leaf and canopy characteristics including variable biophysical parameters, soil brightness, solar zenith angle and sensor noise. Synthetic data were created using well-known leaf and canopy radiative transfer models (PROSPECT-SAILH). The linear extrapolation REPs involving 680, 694, 724 and 760 nm produced the highest correlation (R^2 = 0.75) with leaf chlorophyll content with minimal effects of leaf area index and leaf dry matter content compared with various REP alternatives (R^2 = 0.49 to 0.62). In addition, the linear extrapolation REPs were insensitive to changes in solar zenith angle. However, the advantage of using the linear extrapolation method compared to the various alternative methods diminished with increasing sensor noise and decreasing spectral resolution.

In conclusion, the linear extrapolation method demonstrates that REPs based on tracking changes that occur near the dominant red-edge peaks are less sensitive to the effect of differences in leaf LAI, developmental stage and species. These results have some important implications for ecological studies and precision agriculture. Sims and Gamon (2002) argue that for spectral indices to be most useful in ecological studies e.g. in assessing vegetation health, their relationships with chlorophyll content should be generalisable across species and leaf developmental

stages. In addition, early estimates of crop chlorophyll or nitrogen concentration with REP could be useful in yield forecasting or to manipulate nutrient application in order to improve crop yield in precision agriculture (Haboudane et al., 2002; Goel et al., 2003), provided they are insensitive to variable LAI, leaf thickness and leaf biomass (Pinar and Curran, 1996).

8.2 Application of the linear extrapolation method for discriminating species at leaf and canopy scales

Experimental studies demonstrate poor signal propagation from the leaf to canopy scale because of the complexities introduced by the canopy (Verhoef, 1984; Kuusk, 1991; Jacquemoud et al., 1995; Yoder and Pettigrew-Crosby, 1995; Asner, 1998). However, the low sensitivity of the linear extrapolation method to leaf and canopy structural characteristics might be advantageous for up-scaling leaf level spectral information to the canopy scale. This conclusion was made based on an investigation involving six species (3 shrubs and 3 trees). Among several spectral indices including NDVI, Carter index (Carter, 1994), Vogelman Index (Vogelman et al., 1993), Gitelson and Merzylak index (Gitelson and Merzylak, 1997), Photochemical reflectance index (Gamon et al., 1992), Carotenoid reflectance index (Gitelson et al., 2002), and REPs derived by various methods, the linear extrapolation REPs were least sensitive to the change in measurement scales and showed the highest potential to discriminate the same pairs of species at both scales. Linear extrapolation REPs could significantly discriminate 10 out of a total of 15 pairs of species at both scales compared with none for NDVI involving 800 and 670 nm. The photochemical reflectance index was the best discriminator at the canopy scale (Fig. 8.3). But the significant finding in this study is that canopy indices have a superior discriminating power than leaf level indices as illustrated in Fig. 8.3. This is essential for remote sensing of species at the ecosystem level.

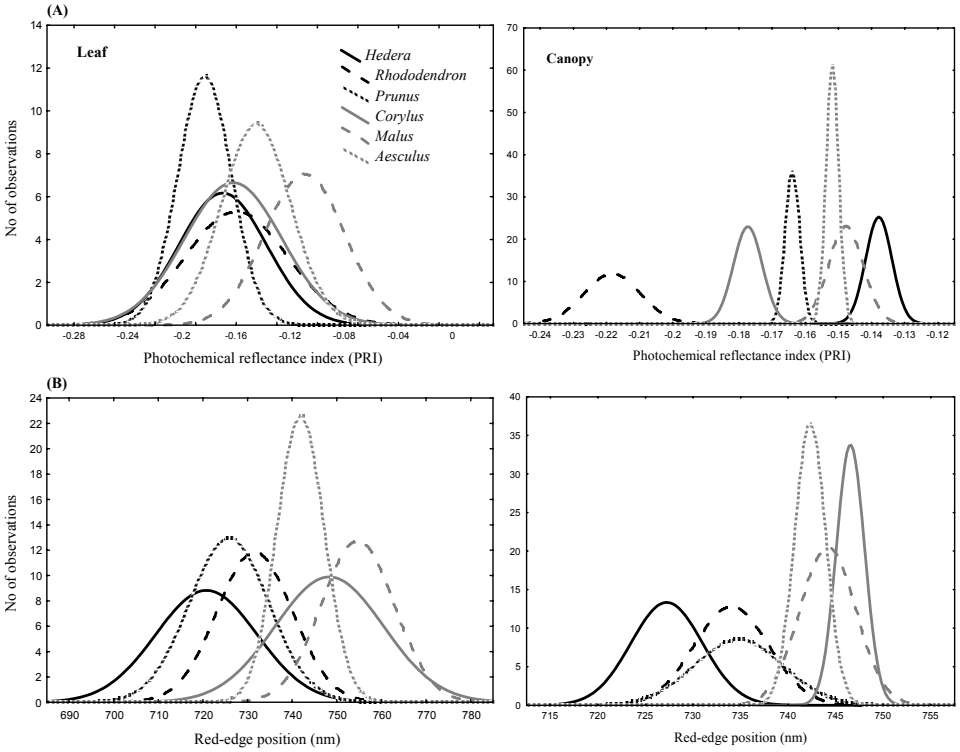


Fig.8.3. Histograms of leaf and canopy indices, namely (A) Photochemical reflectance index (PRI) and (B) linear extrapolation REP for six species of plants. The histograms illustrate the ability of the indices to differentiate species at the leaf and canopy scales.

8.3 Comparing univariate and multivariate statistical techniques for estimating vegetation structural parameters with hyperspectral data

It has been argued that two-band vegetation indices utilise a limited amount of the total spectral information available in a high spectral resolution image (Lee et al. 2004). We therefore, investigated the utility of multivariate statistical techniques based on more than two bands in general and partial least squares (PLS) regression in particular for estimating grass biomass and forest structural parameters. PLS regression model based on continuum-removed bands produced the lowest prediction error for estimating green grass/herb biomass in the Majella National Park (standard error of prediction, SEP = 149 g m^{-2}) compared to NDVI and REP models (SEP = 261 to 295 g m^{-2}). Furthermore, more accurate estimates of beech (*Fagus sylvatica* L.) forest mean diameter-at-

breast height (DBH) (Fig. 8.4) were obtained with PLS regression model than with models based on vegetation indices. PLS regression has an advantage over traditional multiple regression technique in that it can deal with the problem of a large number of collinear predictor bands inherent in hyperspectral data (Geladi and Kowalski, 1986).

The critical question at this juncture is whether the use of two-band vegetation indices such as NDVI is still relevant in the context of hyperspectral remote sensing? NDVI may be simple to implement but could be lacking in terms of exploiting the information content inherent in several narrowbands. The use of multivariate statistical techniques is recommended for hyperspectral remote sensing of vegetation structural parameters.

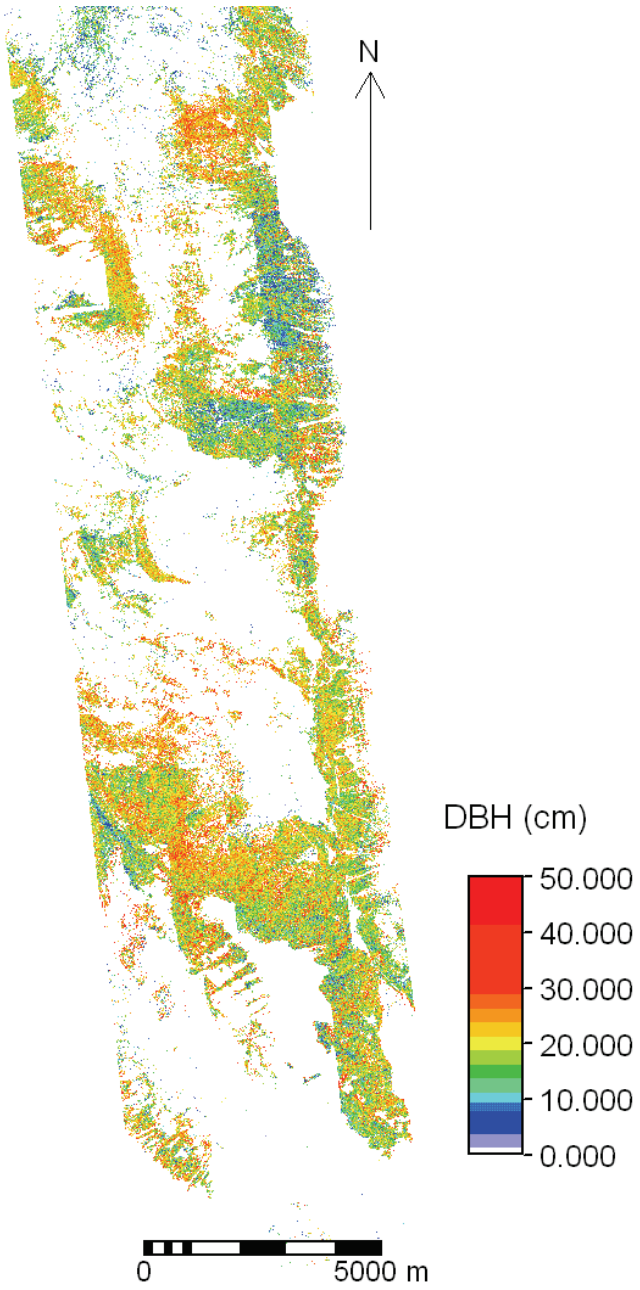


Fig.8.4. Spatial distribution of beech forest tree diameter at breast height (DBH) in the Majella National Park, Italy using partial least squares regression.

8.5 Utility of Empirical methods

Empirical models derived from vegetation indices have been criticised as highly site, time and sensor specific (Curran, 1994; Gobron et al. 1997). The results of this study do support the above assertion for both univariate models based on vegetation indices and multivariate statistical models involving more than two hyperspectral bands. For example, models relating vegetation spectra and beech forest structural characteristics were significant for 2004 and not 2005. Differences in phenological and atmospheric conditions might have affected the relationships.

However, the results of this study show that predictive models derived from REPs computed by the Lagrangian interpolation (Dawson and Curran, 1998) and linear extrapolation methods for 2004 HyMap image were able to produce more accurate estimates of the fresh grass/herb biomass for 2005 and vice versa compared with normalised difference vegetation indices, in the Majella National Park, Italy (e.g. Fig. 8.5). Narrowband NDVI involving 786 and 725 nm showed high correlations with grass/herb biomass for both years (R^2 >0.50), but the regression model for 2004 produced a low prediction accuracy for 2005 and vice versa. One reason that could have accounted for the stability of the REP/biomass relationship for different years is that the REPs are less sensitive to varying soil and atmospheric conditions, sensor view angle and illumination angle (Curran et al. 1995, Blackburn and Pitman, 1999, Clevers et al. 2001; Cho et al., In review). This may apply particularly to the Lagrangian and linear extrapolation REPs that are computed from derivative spectra. Derivative analysis enhances absorption features and suppresses contributions of non-vegetative reflectance components (Boochs et al., 1990; Curran et al., 1991). Given the limited nature of these results in terms of the spatial/temporal coverage and number of sensors involved, it could be premature to draw a definite conclusion about the stability of the Lagrangian and linear extrapolation REPs for modelling and monitoring grass biomass over large areas on yearly basis. Further research is needed.

The lack of robustness and portability of empirical models is therefore of critical concern. The validity of empirical models is limited to the environmental conditions where it has been developed (Asner et al., 2003). The retrieval of biophysical variables using canopy reflectance or

radiative transfer models is viewed as an alternative (e.g. Atzberger, 2004; Schlerf and Atzberger, 2006). Radiative transfer models are physically based or process driven. This class of models describe the interactions between the electromagnetic radiation and the vegetation canopy, so that many different types of environmental conditions can be handled (Iaquinta et al. 1997). However, empirical models that are less sensitive to environmental conditions such as models based on the linear extrapolation REP could be used to support the development of physically based models, particularly to estimate the value of the model parameters, or to refine the underlying concepts on which the model is constructed (Skidmore, 2002). The future of hyperspectral remote sensing could hinge on enhancing the link between empirical and physically based approaches.

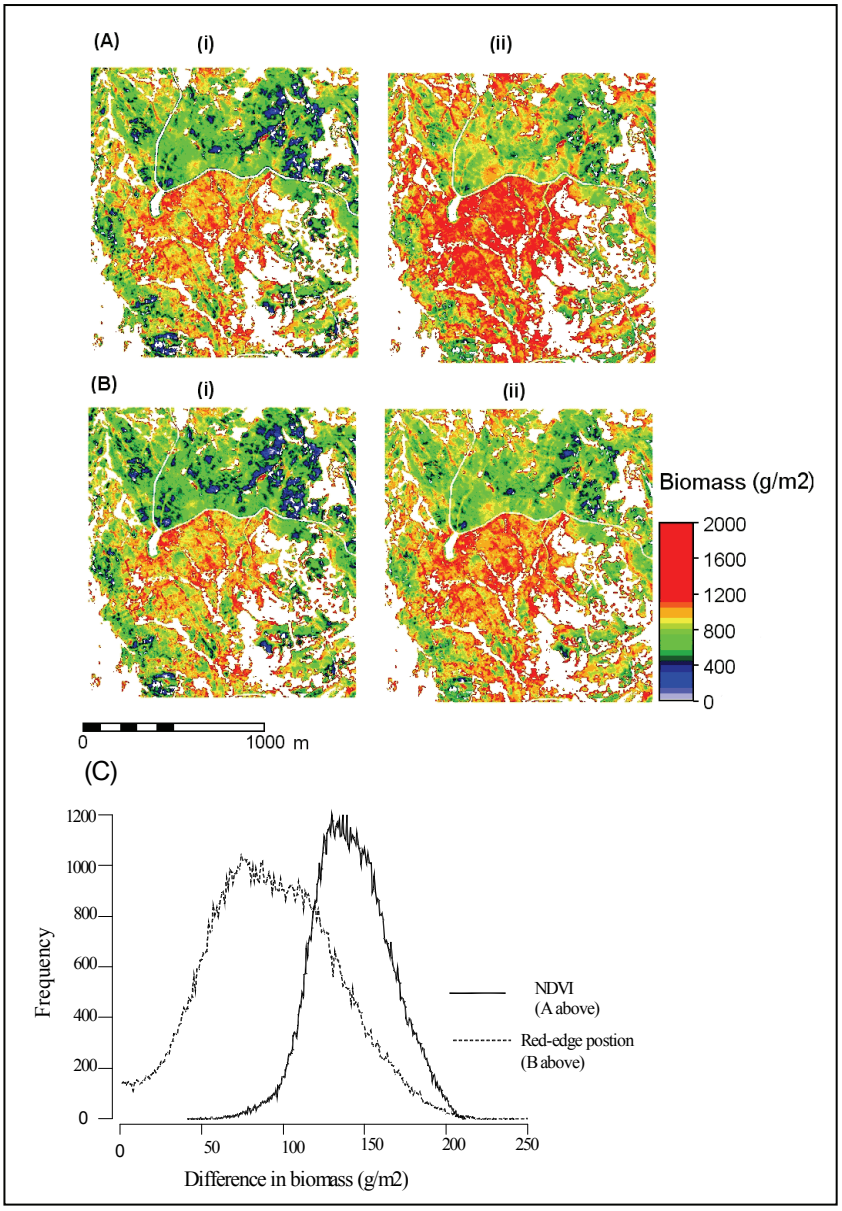


Fig.8.5. Predicted green grass/herb biomass for a subset area of the 2005 HyMap image based on (i) 2005 and (ii) 2004 regression models for (A) NDVI (786 & 725 nm) and (B) red-edge position extracted by the linear extrapolation method. (C) Histograms showing the differences between (i) and (ii), i.e. number of pixels against difference in biomass.

8.4. Conclusions

A new technique for estimating the chlorophyll-sensitive REP (Horler, 1983) was developed by considering changes that occur at the dominant red-edge inflection points (double-peak feature). The method is termed the linear extrapolation technique (Cho and Skidmore, 2006). The principle underlining the linear extrapolation method contradicts earlier curve fitting approaches (Bonham-Carter, 1988), which tend to avoid the double-peak feature. The major contribution of this study is that REPs extracted by the linear extrapolation provide optimal estimates of leaf chlorophyll or nitrogen content with minimal influence of leaf/canopy structural and background confounders. By minimising the effects of canopy structure, the linear extrapolation may be useful for detecting early physiological stresses associated with changes in leaf chlorophyll/nitrogen levels.

The linear extrapolation method also shows high potential for discriminating tree and shrub species at both the leaf and canopy scales. Lastly, it could be used as a more stable predictor for monitoring green grass biomass in the Majella National park, Italy compared with two band vegetation indices. The method is simple to implement but is sensitive to spectral noise. Spectral smoothing is recommended when noise is a problem.

The study also highlights the utility of PLS regression based on several HyMap bands for estimating grass biomass and beech (*Fagus sylvatica* L.) forest mean DBH. PLS regression produced more accurate estimates grass biomass and beech forest mean DBH than univariate regression involving vegetation indices. The use of multivariate statistical techniques is recommended for hyperspectral remote sensing of vegetation parameters.

In line with many recent studies (e.g. Dawson and Curran, 1998; Kokaly and Clark, 1999), this study makes a contribution in the domain of information extraction from hyperspectral data for estimating vegetation parameters. Considerable advances have been achieved in harnessing the predictive capability of spectral data for specific biochemical and biophysical parameters by minimising the effects of environmental conditions. These include, derivation of new vegetation indices, development of new methods to locate the REP, and the use of spectral

transformations such as continuum-removal and derivative analysis. Other studies are focused on developing physically based methods given the lack of robustness and portability of empirical models for varying environmental conditions. However, empirical models that are less sensitive to environmental conditions such as models based on the linear extrapolation REP could be used to support the development of physically based models, particularly to estimate the value of the model parameters, or to refine the underlying concepts on which the model is constructed. The future of hyperspectral remote sensing could hinge on enhancing the link between empirical and physically based approaches.

References

Achard, F. and Estreguil, C., 1995. Forest classification of South East Asia using NOAA AVHRR data. Remote Sensing of Environment, 54: 198-208.

Anderson, G.L., Hanson, J.D. and Haas, R.H., 1993, Evaluating Landsat thematic Mapper derived vegetation indices for estimating above-ground biomass on semiarid rangelands. Remote Sensing of Environment, 45(2): 165-175.

Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sensing of Environment, 64(234-253).

Asner, G.P., Hicke, J.A. and Lobell, D.B., 2003. Per-pixel analysis of forest structure: Vegetation indices, spectral mixture analysis and canopy reflectance modeling. In: M.A. Wulder and FranklinS.E. (Editors), Remote sensing of forest environments. Concepts and case studies. Kluwer Academic Publishers, Boston, pp. 209-254.

Atkinson, P.M. and Tatnall, A.R.L., 1997. Neural networks in remote sensing. International Journal of Remote Sensing, 18(4): 699-709.

Atzberger, C., Jarmer, T., Schlerf, M., Kötz, B. and Werner, W., 2003, Retrieval of wheat bio-physical attributes from hyperspectral data and SAILH+PROSPECT radiative transfer mode. In: M. Habermeyer, A. Müller and S. Holzwarth (Editors), 3rd EARSeL workshop on imaging spectroscopy, 13-16 May, 2003 Germany, Herrchnig, pp. 473-482.

Atzberger, C., 2004. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. Remote Sensing of Environment, 93(1-2): 53-67.

Babani, F., Lichtenthaler, H.K. and Richter, P., 1996. Changes of chlorophyll fluorescence signatures during greening of etoilated barley seedlings as measured with CCD-OMA fluorometer. Journal of Plant Physiology, 148: 471-477.

Bacour, C., Jacquemoud, S., Tourbier, Y., Dechambre, M. and Frangi, J.-P., 2002, Design and analysis of numerical experiments to compare four canopy reflectance models. Remote Sensing of Environment, 79(1): 72-83.

Baldwin Jr., V.C. et al., 2000. The effects of spacing and thinning on stand and tree characteristics of 38-year-old Loblolly Pine. Forest Ecology and Management, 137(1-3): 91-102.

Baranoski, G.V.G. and Rokne, J.G., 2005. A practical approach for estimating the red edge position of plant leaf reflectance. International Journal of Remote Sensing, 26(3): 503-521.

Barton, C.V.M. and North, P.R.J., 2001. Remote sensing of canopy light use efficiency using the photochemical reflectance index: Model and sensitivity analysis. Remote Sensing of Environment, 78(3): 264-273.

Bausch, W.C. and Duke, H.R., 1996. Remote sensing of plant nitrogen status in corn. Transactions - American Society of Agricultural Engineers, 39(5): 1869-1875.

Blackburn, G.A., 1998. Quantifying chlorophylls and caroteniods at leaf and canopy scales: An evaluation of some hyperspectral approaches. Remote Sensing of Environment, 66: 273-285.

Blackburn, G.A. and Pitman, J.I., 1999. Biophysical controls on the directional spectral reflectance properties of bracken (*Pteridium aquilinum*) canopies: results of a field experiment. International Journal of Remote Sensing, 20(11): 2265-2282.

Boegh, E., Soegaard, H., Broge, N., Hasager, C.B., Jensen, N.O., Schelde, K.and Thomsen, A. 2002. Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. Remote Sensing of Environment, 81(2-3): 179-193.

Bonan, G.B., 1993. Importance of leaf area index and forest type when estimating photosynthesis in boreal forests. Remote Sensing of Environment, 43(3): 303-314.

Boneau, C.A., 1960. The effects of violations of assumptions underlying the t test. Psychol. Bull, 57: 49-64.

Bonham-Carter, G.F., 1988. Numerical procedures and computer program for fitting an inverted gaussian model to vegetation reflectance data. Computer & Geosciences, 14(3): 339-356.

Boochs, F., Kupfer, G., Dockter, K. and Kuhbauch, W., 1990. Shape of the red-edge as vitality indicator for plants. International Journal of Remote Sensing, 11(10): 1741-1753.

Brauner, N. and Shacham, M., 1998. Identifying and removing sources of imprecision in polynomial regression. Mathematics and computers simulation, 48: 75-91.

Broge, N.H. and Leblanc, E., 2000. Comparing prediction power and stability of broadband and Hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Remote Sensing of Environment, 76: 156-172.

Bruzzone, L., Conese, C., Maselli, F. and Roli, F., 1997. Multisource classification of complex rural areas by statistical and neural-network approaches. Photogrammetric Engineering and Remote Sensing, 63(5): 523-533.

Buschmann, C. and Nagel, E., 1993. In vivo spectroscopy and internal optics of leaves as basis for remote sensing of vegetation. International Journal of Remote Sensing, 14(4): 711-722.

Carter, G.A., 1994. Ratios of leaf reflectance in narrow wavebands as indicator of plant stress. International Journal of remote sensing, 15: 697-704.

Carter, G.A. and Knapp, A.K., 2001. Leaf optical properties of higher plants: Linking spectral characteristics to stress and chlorophyll concentration. American Journal of Botany, 88(4): 677-684.

Chang, S.H. and Collins, W., 1983. Confirmation of the airborne biogeophysical mineral exploration technique using laboratory methods. Economic Geology and the Bulletin of the society of Economic Geologists, 78: 723-736.

Chappelle, E.W., Kim, M.S. and McMurtrey III, J.E., 1992, Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll A, chlorophyll B, and caroteniods in soybean leaves. Remote Sensing of Environment, 39(3): 239-247.

Chen, J.M., 1996. Evaluation of vegetation indices and a modified simple ratio for boreal applications. Canadian Journal of Remote Sensing, 22: 229-242.

Cho, M.A. and Skidmore, A.K., 2006, A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. Remote Sensing of Environment, 101(2): 181-193.

Cho, M.A., Skidmore, A.K. and Atzberger, C., 2006. Towards red edge positions less sensitive to canopy biophysical parameters using PROSPECT-SAILH simulated data, ISPRS conference Proceedings - Remote Sensing: From Pixels to Processes, Enschede, the Netherlands.

Cho, M.A., Skidmore, A.K., Corsi, F., van Wieren, S.E. and Sobhan, I., in press. Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. International Journal of Applied Earth Observation and Geoinformation.

Cho, M.A. and Skidmore, A.K., In review. Hyperspectral predictors for monitoring biomass production in Mediterranean mountain grasslands: Majella National Park, Italy. International Journal of Remote Sensing.

Ciancio O., Corona, P., Lamonaca, A., Portoghesi, L. and Travaglini, D., 2006. Conversion of clearcut beech coppices into high forests with continuous cover: A case study in central Italy. Forest Ecology and Management, 224: 235-240.

Clevers, J.G.P.W., 1997. A simplified approach for yield prediction of sugar beet based on optical remote sensing data. Remote Sensing of Environment, 61(2): 221-228.

Clevers, J.G.P.W., de Jong, S.M., Epema, G.F., Van der Meer, F., Bakker W.H., Skidmore, A.K., Scholte, K.H., 2002. Derivation of the red edge index using MERIS standard band setting. International Journal of Remote Sensing, 23(16): 3169-3184.

Clevers, J.G.P.W., de Jong, S.M., Epema, G.F., Van der Meer, F., Bakker W.H., Skidmore, A.K., Addink, E.A., 2001. MERIS and the red-edge position. JAG, 3(4): 313-319.

Clevers, J.G.P.W., Kooistra, L. and Salas, E.A.L., 2004. Study of heavy metal contamination in river floodplains using the red-edge position in spectroscopic data. International Journal of Remote Sensing, 25(19): 3883-3895.

Cochran, W.G., 1947. Some consequences when the assumptions for analysis of variance are not satisfied. Biometrics, 3: 22-38.

Cohen, W.B., Maiersperger, T.K., Gower, S.T. and Turner, D.P., 2003. An improved strategy for regression of biophysical variables and Landsat ETM+ data. Remote Sensing of Environment, 84(4): 561-571.

Collins, W., Raines, G.L. and Canney, F.C., 1977. Airborne spectroradiometer discrimination of vegetation anomalies over sulphide mineralisation - a remote sensing technique. Abstract with Programmes, Geological Society of America, Seattle, Washington, 7-9 November 1977, pp. 932-933.

Collins, W., 1978. Remote sensing of crop type and maturity. Photogrammetric Engineering and Remote Sensing, 44: 43-55.

Collins, W., Chang, S.H., Raines, G., Canney, F. and Ashley, R., 1983, Airborne biogeophysical mapping of hidden mineral deposits. Economic Geology and the bulletin of the society of Economic Geologists, 78: 737-749.

Colombo, R., Bellingeri, D., Fasolini, D. and Marino, C.M., 2003. Retrieval of leaf area index in different vegetation types using high resolution satellite data. Remote Sensing of Environment, 86(1): 120-131.

Combal, B. Baret, F., Weiss, M, Trubuil, A., Mace´, D., Pragne`re, A., Myneni, R, Knyazikhin, Y., Wang, L. 2002, Retrieval of canopy biophysical variables from bidirectional reflectance using prior information to solve the ill-posed inverse problem. Remote Sensing of Environment, 84, 1-15.

Conti, F., 1998, Flora D'Abruzzo: An annotated checklist of the flora of the Abruzzo. Herbarium Mediterraneum Panormitanum, Palermo, Italy.

Curran, P.J., 1989. Remote sensing of foliar chemistry. Remote Sensing of Environment, 30(3): 271-278.

Curran, P.J., Dungan, J.L., Macler, B.A. and Plummer, S.E., 1991. The effect of a red leaf pigment on the relationship between red edge and chlorophyll concentration. Remote Sensing of Environment, 35: 69-76.

Curran, P.J., 1994, Imaging spectrometry. Progress in Physical Geography, 18(2): 247– 266.

Curran, P.J., Windham, W.R. and Gholz, H.L., 1995. Exploring the relationship between reflectance red edge and chlorophyll concentration in slash pine leaves. Tree Physiology, 15: 203-206.

Curran, P.J., Dungan, J.L. and Peterson, D.L., 2001. Estimating the foliar biochemical concentration of leaves with reflectance spectrometry. Testing the Kokaly and Clark methodologies. Remote Sensing of Environment, 76: 349-359.

D'Ambrosio, N., Szabo, K. and Lichtenthaler, H.K., 1992. Increase of chlorophyll fluorescence ratio F690/F735 during autumnal chlorophyll breakdown. Radiat. Environ. Biophys., 31: 51-62.

Danson, F.M. and Plummer, S.E., 1995, Red-edge response to forest leaf area index. International Journal of Remote Sensing, 16(1): 183-188.

Daughtry, C.S.T., Gallo, K.P., Goward, S.N., Prince, S.D. and Kustas, W.P., 1992. Spectral estimates of absorbed radiation and phytomass production in corn and soybean canopies. Remote Sensing of Environment, 39(2): 141-152.

Daughtry, C.S.T., Walthall, C.L., Kim, M.S., de Colstoun, E.B. and McMurtreyIII, J.E., 2000, Estimating Corn Leaf Chlorophyll

Concentration from Leaf and Canopy Reflectance. Remote Sensing of Environment, 74(2): 229-239.

Davies, A.M.C., 2001. Uncertainty testing in PLS regression. Spectroscopy Europe, 13: 16–19.

Dawson, T.P. and Curran, P.J., 1998. A new technique for interpolating red edge position. International Journal of Remote Sensing, 19(11): 2133-2139.

De Jong, S.M., Pebesma, E.J. and Lacaze, B., 2003. Above-ground biomass assessment of Mediterranean forests using airborne imaging spectrometry: the DAIS Peyne experiment. International Journal of Remote Sensing, 24(7): 1505-1520.

Duchemin, B., 1999. NOAA/AVHRR Bidirectional Reflectance: Modeling and Application for the Monitoring of a Temperate Forest. Remote Sensing of Environment, 67(1): 51-67.

Everitt, J.H., Richardson, A.J. and Gausman, H.W., 1985. Leaf reflectance-chlorophyll relations in buffelgrass. Photogrammetric engineering and remote sensing, 51: 463-466.

Everitt, J.H., Escobar, D.E. and Richardson, A.J., 1989, Estimating grassland phytomass production with near-infrared and mid-infrared spectral variables. Remote Sensing of Environment, 30(3): 257-261.

Fang, H., Liang, S. and Kuusk, A., 2003. Retrieving leaf area index using a genetic algorithm with a canopy radiative transfer model. Remote Sensing of Environment, 85: 257-270.

Feudale, R.N. and Brown, S.D., 2005. An inverse model for target detection. Chemometrics and Intelligent Laboratory Systems, 77(1-2): 75-84.

Filella, I. and Peñuelas, J., 1994. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. International Journal of Remote Sensing, 15(7): 1459-1470.

Foody, G.M., Boyd, D.S. and Cutler, M.E.J., 2003. Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. Remote Sensing of Environment, 85(4): 463-474.

Fourty, T., Baret, F., Jacquemoud, S., Schmuck, G. and Verdebout, J., 1996, Leaf optical properties with explicit description of its biochemical composition: Direct and inverse problems. Remote Sensing of Environment, 56(2): 104-117.

Franco-Lopez, H., Ek, A.R. and Bauer, M.E., 2001. Estimation and mapping of forest stand density, volume, and cover type using the

k-nearest neighbors method. Remote Sensing of Environment, 77(3): 251-274.

Fuhr, M., Nasi, R. and Delegue, M.-A., 2001. Vegetation structure, floristic composition and growth characteristics of Aucoumea klaineana Pierre stands as influenced by stand age and thinning. Forest Ecology and Management, 140(2-3): 117-132.

Gallego Fernández, J.B., García Mora, M.R. and García Novo, F., 2004, Vegetation dynamics of Mediterranean shrublands in former cultural landscape at Grazalema Mountains, South Spain. Plant Ecology, 172(1): 83-94.

Gamon, J.A., Peñuelas, J. and Field, C.B., 1992, A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. Remote Sensing of Environment, 41: 35-44.

Gamon, J.A. and Surfus, J.S., 1999. Assessing leaf pigment content and activity with a reflectometer. New Phytol, 143: 105-117.

Gao, B., 1996, NDWI - A normalized difference water index for remote sensing of vegetation water from space. Remote Sensing of Environment, 58(3): 257-266.

Gao, X., Huete, A.R., Ni, W. and Miura, T., 2000, Optical-biophysical relationships of vegetation spectra without background contamination. Remote Sensing of Environment, 74: 609-620.

Gates, D.M., Keegan, H.J., Schleter, J.C. and Weidner, V.R., 1965. Spectral properties of plants. Applied Optics, 4(1): 11-20.

Gausman, H.W., 1977. Reflectance of leaf components. Remote Sensing of Environment, 6: 1-9.

Geladi, P. and Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. Analytica Chimica Acta, 185: 1-17.

Geladi, P., Hadjiiski, L. and Hopke, P., 1999. Multiple regression for environmental data: nonlinearities and prediction bias. Chemometrics and Intelligent Laboratory Systems, 47(2): 165-173.

Gilabert, M.A., Gandia, S. and Melia, J., 1996. Analyses of spectral-biophysical relationships for a corn canopy. Remote Sensing of Environment, 55(1): 11-20.

Gitelson, A.A. and Merzlyak, M.N., 1997, Remote estimation of chlorophyll content in higher plant leaves. International Journal of Remote Sensing, 18(12): 2691-2697.

Gitelson, A.A., Zur, Y., Chivkunova, O.B. and Merzlyak, M.N., 2002. Assessing Carotenoid Content in Plant Leaves with Reflectance Spectroscopy. Photochemistry and Photobiology, 75(3): 272-281.

Gobron, N., Pinty, B. and Verstraete, M.M., 1997, Theoretical limits to the estimation of the leaf area index on the basis of visible and near-infrared remote sensing data. IEEE Transactions on Geoscience and Remote Sensing, 35(6): 1438– 1445.

Goel, P.K., Prasher, S.O., Landry, S.A., Patel, R.M., Bonnel, R.B., Viau, A.A. and Miller, J.R., 2003. Potenttial of airborne hyperspectral remote sensing to detect nitrogen deficiency and weed infestation in corn. Computers and Electronics in Agriculture, 38(2):99-124.

Gong, P., Pu, R. and Heald, R.C., 2002. Analysis of in situ hyperspectral data for nutrient estimation of giant sequoia. International Journal of Remote Sensing, 23(9): 1827-1850.

Gong, P., Pu, R., Biging, G.S. and Larrieu, M.R., 2003. Estimation of forest leaf area index using vegetation indices derived from Hyperion hyperspectral Data. IEEE Transactions on Geoscience and Remote Sensing, 41: 1355–1362.

Goward, S.N. and Huemmrich, K.F., 1992, Vegetation canopy PAR absorptance and the normalized difference vegetation index: An assessment using the SAIL model. Remote Sensing of Environment, 39(2): 119-140.

Gower, S.T., Kucharik, C.J. and Norman, J.M., 1999. Direct and Indirect Estimation of Leaf Area Index, fAPAR, and Net Primary Production of Terrestrial Ecosystems. Remote Sensing of Environment, 70(1): 29-51.

Gupta, R.K., Vijayan, D. and Prasad, T.S., 2003, Comparative analysis of red-edge hyperspectral indices. Advance Space Research, 32(11): 2217-2222.

Guyot, G. and Baret, F., 1988. Utilisation de la haute résolution spectrale pour suivre l'état des couverts végétaux, Proceedings of the 4th International colloquium on spectral signatures of objects in remote sensing. ESA SP-287, Assois, France, pp. 279-286.

Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J. and Strachan, I.B., 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modelling and validation in the context of precision agriculture. Remote Sensing of Environment, 90: 337-352.

Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J. and Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sensing of Environment, 81: 416-426.

Hák, R., Lichtenthaler, H.K. and Rinderle, U., 1990. Decrease of the chlorophyll fluorescence ratio (F690/F730) during greening and development of leaves. Radiat. Environ. Biophys., 29: 329-336.

Hansen, P.M. and Schjoerring, J.K., 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sensing of Environment, 86: 542-553.

Hare, E.W., Miller, J.R. and Edward, G.R., 1984. Studies of vegetation red reflectance edge in geobotanical remote sensing in eastern Canada., Proceedings of the 9th Canadian Symposium on Remote Sensing, held at St, John's, Newfoundland, 13-17 August 1984, (Ottawa, Canadian Aeronautics and Space Institute), pp. 433-440.

Horler, D.N.H., Barber, J. and Barringer, A.R., 1980. Effects of heavy metals on the absorbance and reflectance spectra of plants. International Journal of Remote Sensing, 1: 121.

Horler, D.N.H., Dockray, M. and Barber, J., 1983. The red edge of plant leaf reflectance. International Journal of Remote Sensing, 4(2): 273-288.

Huang, Z., Turner, B.J., Dury, S.J., Wallis, I.R. and Foley, W.J., 2004. Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. Remote Sensing of Environment, 93(1-2): 18-29.

Huete, A.R., 1988, A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment, 25(3): 295-309.

Huete, A.R. and Jackson, R.D., 1988. Soil and atmosphere influences on the spectra of partial canopies. Remote Sensing of Environment, 25(1): 89-105.

Huete, A.R., Hua, G., Qi, J., Chehbouni, A. and van Leeuwen, W.J.D., 1992, Normalization of multidirectional red and NIR reflectances with the SAVI. Remote Sensing of Environment, 41(2-3): 143-154.

Huete, A., Justice, C. and Liu, H., 1994. Development of vegetation and soil indices for MODIS-EOS. Remote Sensing of Environment, 49(3): 224-234.

Iaquinta, J., Pinty, B., Privette, J.L., 1997. Inversion of a physically based bidirectional reflectance model of vegetation. IEEE Transactions on Geoscience and Remote sensing, 35: 687-698.

Ingram, J.C., Dawson, T.P. and Whittaker, R.J., 2005. Mapping tropical forest structure in southeastern Madagascar using remote sensing

and artificial neural networks. Remote Sensing of Environment, 94(4): 491-507.

Jacquemoud, S. and Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. Remote Sensing of Environment, 34(2): 75-91.

Jacquemoud, S., Baret, F., Andrieu, B., Danson, F.M. and Jaggard, K., 1995, Extraction of vegetation biophysical parameters by inversion of the PROSPECT + SAIL models on sugar beet canopy reflectance data. Application to TM and AVIRIS sensors. Remote Sensing of Environment, 52(3): 163-172.

Jacquemoud, S., Bacour, C., Poilve, H. and Frangi, J.-P., 2000, Comparison of Four Radiative Transfer Models to Simulate Plant Canopies Reflectance: Direct and Inverse Mode. Remote Sensing of Environment, 74(3): 471-481.

Jago, R.A., Cutler, M.E.J. and Curran, P.J., 1999. Estimating Canopy Chlorophyll Concentration from Field and Airborne Spectra. Remote Sensing of Environment, 68: 217-224.

Jensen, J.R., 2000, Remote sensing of the environment: an earth resource perspective. Prentice Hall series in Geographic Information Science. Prentice Hall, New Jersey, 544 pp.

Johnson, L.F., Hlavka, C.A. and Peterson, D.L., 1994. Multivariate analysis of AVIRIS data for canopy biochemical estimation along Oregon transect. Remote Sensing of Environment, 47: 216-230.

Jongschaap, R.E.E. and Booij, R., 2004. Spectral measurements at different spatial scales in potato: relating leaf, plant and canopy nitrogen status. International Journal of Applied Earth Observation and Geoinformation, 5: 204-218.

Kasischke, E.S., Melack, J.M. and Dobson, M.C., 1997. The use of imaging radar for ecological applications – a review. Remote Sensing of Environment, 59:141-156.

Kaufman, Y.J. and Tanré, D., 1992. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. IEEE Transactions on Geoscience and Remote Sensing, 30(2): 261-270.

Knapp, A.K. and Carter, G.A., 1998. Variability in Leaf Optical Properties Among 26 Species from a Broad Range of Habitats. American Journal of Botany, 85(7): 940-946.

Knipling, E.B., 1970. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. Remote Sensing of Environment, 1(3): 155-159.

Kokaly, R.F. and Clark, R.N., 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and

stepwise multiple linear regression. Remote Sensing of Environment, 67: 267-287.

Kooistra, L., Salas, A.L., Clevers, J.G.P.W., Wehrens, R., Leuven, R.S.E.W., Nienhius, P.H., and Buydens, L.M.C., 2004. Exploring field vegetation reflectance as an indicator of soil contamination in river floodplains. Environmental Pollution, 127: 281–290.

Kubinyi, H., 1996. Evolutionary variable selection in regression and PLS analyses. Journal of Chemometrics, 10: 119– 133.

Kumar, L., Schmidt, K.S. and Skidmore, A.K., 2001. Imaging spectroscopy and vegetation science. In: S.M.D. Jong (Editor), Imaging spectrometry. Academic Kluwer Press, Dordrecht, The Netherlands, pp. 111-155.

Kuusk, A., 1991. The angular distribution of reflectance and vegetation indices in barley and clover canopies. Remote Sensing of Environment, 37(2): 143-151.

Lamb, D.W., Steyn-Ross M., Schaare P., Hanna, M.M., Silvester, W., Steyn-Ross, A., 2002. Estimating leaf nitrogen concentration in ryegrass (*Lolium* spp.) pasture using the chlorophyll red-edge: theoretical modelling and experimental observations. International Journal of Remote Sensing, 23(18): 3619-3648.

Lee, K.-S., Cohen, W.B., Kennedy, R.E., Maiersperger, T.K. and Gower, S.T., 2004. Hyperspectral versus multispectral data for estimating leaf area index in four different biomes. Remote Sensing of Environment, 91(3-4): 508-520.

Lefsky, M.A. Cohen, W.B., Acker, S.A., Spies, T.A., Parker, G.G. and Harding, D., 1999. Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir Western Hemlock Forests. Remote Sensing of Environment, 70(3): 339-361.

Lefsky, M.A. Cohen, W.B. and Spies, T.A., 2001. An Evaluation of alternate remote sensing products for forest inventory, monitoring and mapping of Douglas-fir forests in western Oregon. Canadian Journal of Forest Research, 31:78-87.

Lichtenthaler, H.K., Gitelson, A.A. and Lang, M.J., 1996, Non-destructive determination of chlorophyll content of leaves of a green and aurea mutant tobacco by reflectance measurements. Journal of Plant Physiology, 148: 483-493.

Martens, H. and Martens, M., 2000. Modified jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). Food Quality and Preference, 11: 5 – 16.

Mas, J.F., Puig, H., Palacio, J.L. and Sosa-Lopez, A., 2004. Modelling deforestation using GIS and artificial neural networks. Environmental Modelling & Software, 19(5): 461-471.

McGarigal, K., Cushman, S. and Stafford, S., 2000. Multivariate statistics for wildlife and ecology research. Springer, New York.

Merzlyak, M.N., Gitelson, A.A., Chivkunova, O.B., Solovchenko, A.E. and Pogosyan, S.I. (2003). Application of reflectance spectroscopy for analysis of higher plant pigments. Russian Journal of Plant Physiology, 50(5): 704-710.

Messina, M.G., 1992. Response of Eucalyptus regnans F. Muell. to thinning and urea fertilization in New Zealand. Forest Ecology and Management, 51(4): 269-283.

Middleton, E.M., 1991. Solar zenith angle effects on vegetation indices in tallgrass prairie. Remote Sensing of Environment, 38(1): 45-62.

Miller, D.M., Kaminsky, E.J. and Rana, S., 1995. Neural network classification of remote-sensing data. Computers & Geosciences, 21(3): 377-386.

Miller, J.R., Hare, E.W. and Wu, J., 1990. Quantitative characterization of the red edge reflectance. An inverted-Gaussian reflectance model. International Journal of Remote Sensing, 11(10): 1755-1773.

Mooney, H.A. (Editor), 1986. Photosynthesis. Plant ecology. Blackwell scientific publication, Oxford, United Kingdom, 345-373 pp.

Mutanga, O., Skidmore, A. and van Wieren, S., 2003. Discriminating tropical grass (*Cenchrus ciliaris*) canopies grown under different nitrogen treatments using spectroradiometry. ISPRS Journal of Photogrammetry & Remote Sensing, 57: 263-272.

Mutanga, O. and Skidmore, A.K., 2004. Narrow band vegetation indices overcome the saturation problem in biomass estimation. International Journal of Remote Sensing, 25: 1-16.

Mutanga, O., Skidmore, A.K. and Prins, H.H.T., 2004. Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. Remote Sensing of Environment, 89: 393-408.

Nagendra, H., 2001. Using remote sensing to assess biodiversity. International Journal of Remote Sensing, 22: 2377-2400.

Nisson, M. 1996. Estimation of tree height and stand volume using an airborne lidar system. Remote Sensing of Environment, 56:1-7.

Nguyen, H.T. and Lee, B.-W., 2006. Assessment of rice leaf growth and nitrogen status by hyperspectral canopy reflectance and partial least square regression. European Journal of Agronomy, 24(4): 349-356.

Novozamsky, I., Houba, V.J.G., van Eck, R. and van Vark, W., 1983. A novel digestion technique for multi-element plant analysis. Communications in soil science and plant analysis, 14: 239-249.

Paola, J.D. and Schowengerdt, R.A., 1995. A review and analysis of back-propagation neural networks for classification of remotely sensed multispectral imagery. International Journal of Remote Sensing, 16(16): 3033-3058.

Peñuelas, J., Baret, F. and Filella, I., 1995, Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. Photosynthetica, 31: 221-230.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J-M., Tucker, C.J. and Stenseth, N.C., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. TRENDS in Ecology and Evolution, 20(9): 503-510.

Pierce, L.L., Running, S.W. and Walker, J., 1994. Regional-scale relationships of leaf-area index to specific leaf-area and leaf nitrogen-content. Ecological Applications, 4: 313-321.

Pinar, A. and Curran, P.J., 1996. Grass chlorophyll and the reflectance red-edge. International Journal of Remote Sensing, 17 (2), 135-357.

Posten, H.O., Yeh, H.C. and Owen, D.B., 1982. Robustness of the two-sample t test under violations of the homogeneity of variance assumptions. Commnic. statist. - Theor.Meth., 11: 109-126.

Price, J.C., 1992. Estimating vegetation amount from visible and near infrared reflectances. Remote Sensing of Environment, 41(1): 29-34.

Pu, R., Gong, P., Biging, G.S. and Larrieu, M.R., 2003. Extraction of red edge optical parameters from Hyperion data for estimation of forest leaf area index. IEEE Transactions on Geoscience and Remote Sensing, 41(4): 916-921.

Qi, J., Moran, M.S., Cabot, F. and Dedieu, G., 1995, Normalization of sun/view angle effects using spectral albedo-based vegetation indices. Remote Sensing of Environment, 52(3): 207-217.

Qi, J., Cabot, F., Moran, M.S. and Dedieu, G., 1995, Biophysical parameter estimations using multidirectional spectral measurements. Remote Sensing of Environment, 54(1): 71-83.

Richardson, A.J., Wiegand, C.L., Arkin, G.F., Nixon, P.R. and Gerbermann, A.H., 1982, Remotely-sensed spectral indicators of sorghum development and their use in growth modeling. Agricultural Meteorology, 26(1): 11-23.

Richardson, A.J., Everitt, J.H. and Gausman, H.W., 1983. Radiometric estimation of biomass and nitrogen content of Alicia grass. Remote Sensing of Environment, 13: 179-184.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W. and Harlan, J.C., 1974. Monitoring the vernal advancement and retrogradation of natural vegetation, NASA/GSFC, Type III Final Report, M.D. Greenbelt, pp. 371.

Saney, G.B. and Elliott, R.L., 2002. Capacity of AVHRR data in discriminating rangeland cover mixtures. International Journal of Remote Sensing, 23: 299-312.

Savitzky, A. and Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least-squares procedures. Analytica Chemistry., 36(8): 1627–1639.

Schlerf, M. and Atzberger, C., 2006. Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data. Remote Sensing of Environment, 100(3): 281-294.

Schmidt, K.S. and Skidmore, A.K., 2001. Exploring spectral discrimination of grass species in African rangelands. International Journal of Remote Sensing, 22(17): 3421-3434.

Schmidtlein, S. and Sassin, J., 2004. Mapping of continuous floristic gradients in grasslands using hyperspectral imagery. Remote Sensing of Environment, 92(1): 126-138.

Scurlock, J.M.O. and Prince, S.D., 1993. Remote sensing of biomass and productivity. In: D.O. Hall, J.M.O. Scurlock and H.R. Bolhar-Nordenkampf (Editors), Photosynthesis and production in a changing environment: A field and laboratory manual. Chapman & Hall, London, pp. 23-46.

Sellers, P.J., 1985, Canopy reflectance, photosynthesis and transpiration. International Journal of Remote Sensing, 6(8): 1335-1372.

Senft, R.L., Rittenhouse, L.R. and Woodmansee, R.B., 1985. Factors influencing patterns of cattle grazing behavior on shortgrass steppe. Journal of Range Management, 38(82–87).

Skidmore, A.K., Brinkhof, T.W. and Knowles, E., 1997. Performance of a neural network: Mapping forest using GIS and remotely sensed

data. Photogrammetric Engineering and remote sensing, 63(5): 501-514.

Skidmore, A.K., 2002. Taxonomy of environmental models in the spatial sciences. In: A.K. Skidmore (Editor), Environmental modelling. Taylor & Francis, London and New York, pp. 8-25.

Smith, K.L., Steven, M.D. and Colls, J.J. 2004. Use of hyperspectral derivative ratios in the red edge region to identify plant stress responses to gas leak. Remote Sensing of Environment, 92: 207-217.

Sullivan, D.G, Shaw, J.N., Mask, P.L., Rickman, D., Luvall, J., Wersinger, J.M. (2004). Evaluating corn nitrogen variability via remote-sensed data. Communications in Soil Science and Plant Analysis, 35(17-18): 2465-2483.

Thenkabail, P.S., Smith, R.B. and De Pauw, E., 2000, Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. Remote Sensing of Environment, 71: 158-182.

Thenkabail, P.S., Enclona, E.A., Ashton, M.S., Legg, C. and De Dieu, M.J., 2004. Hyperion, IKONOS, ALI, and ETM+ sensors in the study of African rainforests. Remote Sensing of Environment, 90(1): 23-43.

Thenkabail, P.S., Enclona, E.A., Ashton, M.S. and Van Der Meer, B., 2004. Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. Remote Sensing of Environment, 91(3-4): 354-376.

Thomas, J.R. and Oerther, G.F., 1972. Estimating nitrogen content of sweet pepper leaves by reflectance measurement. Agronomy Journal, 64: 11-13.

Todd, S.W., Hoffer, R.M. and Milchunas, D.G., 1998, Biomass estimation on grazed and ungrazed rangelands using spectral indices. International Journal of Remote Sensing, 19(3): 427-438.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment, 8: 127–150.

Uno, Y. et al., 2005. Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data. Computers and Electronics in Agriculture, 47(2): 149-161.

Vaiphasa, C., Ongsomwang, S., Vaiphasa, T. and Skidmore, A.K., 2005. Tropical mangrove species discrimination using hyperspectral data: A laboratory study. Estuarine, Coastal and Shelf Science, 65: 371-379.

van den Broek, W.H.A.M., Derks, E.P.P.A., van de Ven, E.W., Geladi, W.P. and Buydens, M.C., 1996. Plastic identification by remote sensing spectroscopic NIR imaging using kernel partial least squares (KPLS). Chemometrics and Intelligent Laboratory Systems, 35(2): 187-197.

Van Wagtendonk, J.W. and Root, R.R., 2003. The use of multitemporal normalised difference vegetation index (NDVI) data for mapping fuel models in Yosemite National Park, USA. International Journal of Remote Sensing, 24: 1639-1651.

Verhoef, W., 1984, Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. Remote Sensing of Environment, 16(2): 125-141.

Verstraete, M.M., Pinty, B. and Myneni, R.B., 1996. Potential and limitations of information extraction on the terrestrial biosphere from satellite remote sensing. Remote Sensing of Environment, 58(2): 201–214.

Viscarra Rossel, R.A., 2005. ParLeS., Pre-processing of data, Principal Component Analysis and Partial Least Squares Regression with Leave-one-out Cross Validation. http://www.usyd.edu.au/su/agric/acpa/people/rvrossel/soft01.htm.

Vogelmann, J.E., Rock, B.N. and Moss, D.M., 1993. Red-edge spectral measurements from sugar maple leaves. International Journal of Remote Sensing, 14: 1563-1575.

Webster, R., 2000. Statistics to support soil research and their presentation. European Journal of Soil Science, 52: 331-340.

Wenjiang, H., Jihua, W., Zhijie, W., Jiang, Z., Liangyun, L., Jindi, W., 2004. Inversion of foliar biochemical parameters at various physiological stages and grain quality indicators of winter wheat with canopy reflectance. International Journal of Remote Sensing, 25(12): 2409-2419.

Wessman, C.A., Aber, J.D. and Peterson, D.L., 1989. An evaluation of imaging spectrometry for estimating forest canopy chemistry. International Journal of Remote Sensing, 10(8): 1293-1316.

Wiegand, C.L., Richardson, A.J., Escobar, A.J. and Gerbermann, A.H., 1991. Vegetation indices in crop assessment. Remote Sensing of Environment, 35: 105-119.

Woodcock, C.E., Collins, J., Jakabhazy, V.D., Li, X., Macomber, S.A., and Wu, Y., 1997. Inversion of the Li-Strahler canopy reflectance model for mapping forest structure. IEEE Transactions on Geoscience and Remote Sensing, 35(2): 405-414.

Wylie, B.K., Meyer, D.J., Tieszen, L.L. and Mannel, S., 2002. Satellite mapping of surface biophysical parameters at the biome scale over the North American grasslands: A case study. Remote Sensing of Environment, 79(2-3): 266-278.

Yoder, B.J. and Pettigrew-Crosby, R.E., 1995. Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2500 nm) at leaf and canopy scales. Remote Sensing of Environment, 53: 199-211.

Zar, J.H., 1996. Biostatistics Analysis. Prentice-Hall International INC., pp. 235-276.

Zarco-Tejada, P.J., Pushnik, J.C., Dobrowski, S. and Ustin, S.L., 2003. Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. Remote Sensing of Environment, 84(2): 283-294.

Author's Biography

Curriculum vitae

Moses Azong Cho was born on the 7th of September 1967 in Mankon, Cameroon.

He completed his Ordinary level and Advanced level education at Presbyterian Secondary School (PSS) Mankon (1980-1985) and Cameroon Collage of Arts Science and Technology (CCAST) Bambili (1985-1987), respectively. He proceeded to the University of Yaoundé in 1987 where he obtained a Bachelor's degree in the Natural Sciences (Specialisation – Botany) in 1991. He worked as a Secondary and High school science teacher (Biology and Mathematics) between 1992 and 1998 and between 1998-2000, he worked as a science teacher trainer and nature conservation research assistant with the German Development Service In-service Training programme in Cameroon and Northwest Development Authority, respectively. In 2000, he joined the International Master programme at the University of Greenwich, the UK/Institute of Spatial planning and Environmental Science at Saxion Deventer, the Netherlands, where he obtained an MSc in Nature Conservation and Biodiversity Management with a distinction in September 2001. He conducted his MSc research during a six-month internship in wetland conservation with the Institute for Inland Water Management and Waste Water Treatment (RIZA), Lelystad, the Netherlands. He went back to Cameroon and worked as a wetland researcher between 2001-2003 with the Center for Biodiversity Conservation (CBC), which he founded in 2001. CBC is currently implementing a huge wetland project in the Douala-Edea estuary in partnership with the Watershed Task Group (WTG). The project is funded by IUCN, The Netherlands. He was awarded a scholarship in June 2003 by the International Institute Geoinformation Science and Earth Observation (ITC) to pursue a PhD degree at ITC and Wageningen University in The Netherlands.

Scientific Publications

Cho, M.A. and Skidmore, A.K., **2006**. A new technique for extracting the red edge position from Hyperspectral data: The linear extrapolation method. *Remote Sensing of Environment*, 101:181-193.

Cho, M.A., Skidmore, A.K., Corsi, F., van Wieren, S.E. and Sobhan, I. (**In press**). Estimating green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *International Journal of Applied Earth Observation and Geoinformation*.

Cho, M.A.**,** Skidmore, A.K. and Atzberger, C.G., **2006**. Towards red - edge positions less sensitive to canopy biophysical parameters using PROSPECT-SAILH simulated data. In: ISPRS 2006: ISPRS mid-term symposium 2006 remote sensing: from pixels to processes, 8-11 May 2006, Enschede, the Netherlands. Enschede: ITC, 2006. 6 p. (**in review after revision** – *International Journal of Remote Sensing*)

Cho, M. A, Sobhan, I. Skidmore A. K., **2006**. Estimating fresh grass/herb biomass from HYMAP data using the red edge position. Proc. SPIE Vol. 6298, 629805 (Sep. 27, 2006).

Cho, M.A. and Skidmore, A.K. (**In review**). Hyperspectral predictors for monitoring biomass production in Mediterranean grasslands using HyMap data. *International Journal of Remote sensing*.

Cho, M.A., Skidmore, A.K. and Sobhan, I. (**In review**), Mapping beech (*Fagus sylvatica* L.) forest structure using partial least squares regression on airborne hyperspectral imagery. *Forest Ecology and management*.

Cho, M.A., Sobhan, I., Skidmore, A.K. and de Leeuw, J. (**In Prep**). Discriminating species using hyperspectral indices at leaf and canopy scales.

Cho, M.A., **2003**. A draft management plan for the Waza Logone floodplains, Cameroon. In P.Q. Tersptra (Editor). Wetland management planning: Five years of case studies from West Africa. Wetlands International Publication, Dakar Senegal. p. 135-152.

Cho, M.A. 2002 (MSc. Thesis). Mitigation of wetland degradation in the Lake Chad Basin. University of Greenwich, UK.

ITC DISSERTATION LIST

1. **Akinyede** (1990), Highway cost modelling and route selection using a geotechnical information system
2. **Pan He Ping** (1990), 90-9003-757-8, Spatial structure theory in machine vision and applications to structural and textural analysis of remotely sensed images
3. **Bocco Verdinelli, G.** (1990), Gully erosion analysis using remote sensing and geographic information systems: a case study in Central Mexico
4. **Sharif, M**. (1991), Composite sampling optimization for DTM in the context of GIS
5. **Drummond, J.** (1991), Determining and processing quality parameters in geographic information systems
6. **Groten, S.** (1991), Satellite monitoring of agro-ecosystems in the Sahel
7. **Sharifi, A.** (1991), 90-6164-074-1, Development of an appropriate resource information system to support agricultural management at farm enterprise level
8. **Zee, D. van der** (1991), 90-6164-075-X, Recreation studied from above: Air photo interpretation as input into land evaluation for recreation
9. **Mannaerts, C.** (1991), 90-6164-085-7, Assessment of the transferability of laboratory rainfall-runoff and rainfall - soil loss relationships to field and catchment scales: a study in the Cape Verde Islands
10. **Ze Shen Wang** (1991), 90-393-0333-9, An expert system for cartographic symbol design
11. **Zhou Yunxian** (1991), 90-6164-081-4, Application of Radon transforms to the processing of airborne geophysical data
12. **Zuviria, M. de** (1992), 90-6164-077-6, Mapping agro-topoclimates by integrating topographic, meteorological and land ecological data in a geographic information system: a case study of the Lom Sak area, North Central Thailand
13. **Westen, C. van** (1993), 90-6164-078-4, Application of Geographic Information Systems to landslide hazard zonation
14. **Shi Wenzhong** (1994), 90-6164-099-7, Modelling positional and thematic uncertainties in integration of remote sensing and geographic information systems

15. **Javelosa, R.** (1994), 90-6164-086-5, Active Quaternary environments in the Philippine mobile belt
16. **Lo King-Chang** (1994), 90-9006526-1, High Quality Automatic DEM, Digital Elevation Model Generation from Multiple Imagery
17. **Wokabi, S.** (1994), 90-6164-102-0, Quantified land evaluation for maize yield gap analysis at three sites on the eastern slope of Mt. Kenya
18. **Rodriguez, O.** (1995), Land Use conflicts and planning strategies in urban fringes: a case study of Western Caracas, Venezuela
19. **Meer, F. van der** (1995), 90-5485-385-9, Imaging spectrometry & the Ronda peridotites
20. **Kufoniyi, O.** (1995), 90-6164-105-5, Spatial coincidence: automated database updating and data consistency in vector GIS
21. **Zambezi, P.** (1995), Geochemistry of the Nkombwa Hill carbonatite complex of Isoka District, north-east Zambia, with special emphasis on economic minerals
22. **Woldai, T.** (1995), The application of remote sensing to the study of the geology and structure of the Carboniferous in the Calañas area, pyrite belt, SW Spain
23. **Verweij, P.** (1995), 90-6164-109-8, Spatial and temporal modelling of vegetation patterns: burning and grazing in the Paramo of Los Nevados National Park, Colombia
24. **Pohl, C.** (1996), 90-6164-121-7, Geometric Aspects of Multisensor Image Fusion for Topographic Map Updating in the Humid Tropics
25. **Jiang Bin** (1996), 90-6266-128-9, Fuzzy overlay analysis and visualization in GIS
26. **Metternicht, G.** (1996), 90-6164-118-7, Detecting and monitoring land degradation features and processes in the Cochabamba Valleys, Bolivia. A synergistic approach
27. **Hoanh Chu Thai** (1996), 90-6164-120-9, Development of a Computerized Aid to Integrated Land Use Planning (CAILUP) at regional level in irrigated areas: a case study for the Quan Lo Phung Hiep region in the Mekong Delta, Vietnam
28. **Roshannejad, A.** (1996), 90-9009-284-6, The management of spatio-temporal data in a national geographic information system
29. **Terlien, M.** (1996), 90-6164-115-2, Modelling Spatial and Temporal Variations in Rainfall-Triggered Landslides: the integration of hydrologic models, slope stability models and GIS

for the hazard zonation of rainfall-triggered landslides with examples from Manizales, Colombia

30. **Mahavir, J.** (1996), 90-6164-117-9, Modelling settlement patterns for metropolitan regions: inputs from remote sensing
31. **Al-Amir, S.** (1996), 90-6164-116-0, Modern spatial planning practice as supported by the multi-applicable tools of remote sensing and GIS: the Syrian case
32. **Pilouk, M.** (1996), 90-6164-122-5, Integrated modelling for 3D GIS
33. **Duan Zengshan** (1996), 90-6164-123-3, Optimization modelling of a river-aquifer system with technical interventions: a case study for the Huangshui river and the coastal aquifer, Shandong, China
34. **Man, W.H. de** (1996), 90-9009-775-9, Surveys: informatie als norm: een verkenning van de institutionalisering van dorp - surveys in Thailand en op de Filippijnen
35. **Vekerdy, Z.** (1996), 90-6164-119-5, GIS-based hydrological modelling of alluvial regions: using the example of the Kisaföld, Hungary
36. **Pereira, Luisa** (1996), 90-407-1385-5, A Robust and Adaptive Matching Procedure for Automatic Modelling of Terrain Relief
37. **Fandino Lozano, M.** (1996), 90-6164-129-2, A Framework of Ecological Evaluation oriented at the Establishment and Management of Protected Areas: a case study of the Santuario de Iguaque, Colombia
38. **Toxopeus, B.** (1996), 90-6164-126-8, ISM: an Interactive Spatial and temporal Modelling system as a tool in ecosystem management: with two case studies: Cibodas biosphere reserve, West Java Indonesia: Amboseli biosphere reserve, Kajiado district, Central Southern Kenya
39. **Wang Yiman** (1997), 90-6164-131-4, Satellite SAR imagery for topographic mapping of tidal flat areas in the Dutch Wadden Sea
40. **Saldana-Lopez, Asunción** (1997), 90-6164-133-0, Complexity of soils and Soilscape patterns on the southern slopes of the Ayllon Range, central Spain: a GIS assisted modelling approach
41. **Ceccarelli, T.** (1997), 90-6164-135-7, Towards a planning support system for communal areas in the Zambezi valley, Zimbabwe; a multi-criteria evaluation linking farm household analysis, land evaluation and geographic information systems
42. **Peng Wanning** (1997), 90-6164-134-9, Automated generalization in GIS

43. **Lawas, C.** (1997), 90-6164-137-3, The Resource Users' Knowledge, the neglected input in Land resource management: the case of the Kankanaey farmers in Benguet, Philippines
44. **Bijker, W.** (1997), 90-6164-139-X, Radar for rain forest: A monitoring system for land cover Change in the Colombian Amazon
45. **Farshad, A.** (1997), 90-6164-142-X, Analysis of integrated land and water management practices within different agricultural systems under semi-arid conditions of Iran and evaluation of their sustainability
46. **Orlic, B.** (1997), 90-6164-140-3, Predicting subsurface conditions for geotechnical modelling
47. **Bishr, Y.** (1997), 90-6164-141-1, Semantic Aspects of Interoperable GIS
48. **Zhang Xiangmin** (1998), 90-6164-144-6, Coal fires in Northwest China: detection, monitoring and prediction using remote sensing data
49. **Gens, R.** (1998), 90-6164-155-1, Quality assessment of SAR interferometric data
50. **Turkstra, J.** (1998), 90-6164-147-0, Urban development and geographical information: spatial and temporal patterns of urban development and land values using integrated geo-data, Villaviciencia, Colombia
51. **Cassells, C.** (1998), 90-6164-234-5, Thermal modelling of underground coal fires in northern China
52. **Naseri, M.** (1998), 90-6164-195-0, Characterization of Salt-affected Soils for Modelling Sustainable Land Management in Semi-arid Environment: a case study in the Gorgan Region, Northeast, Iran
53. **Gorte B.G.H.** (1998), 90-6164-157-8, Probabilistic Segmentation of Remotely Sensed Images
54. **Tegaye, Tenalem Ayenew** (1998), 90-6164-158-6, The hydrological system of the lake district basin, central main Ethiopian rift
55. **Wang Donggen** (1998), 90-6864-551-7, Conjoint approaches to developing activity-based models
56. **Bastidas de Calderon, M.** (1998), 90-6164-193-4, Environmental fragility and vulnerability of Amazonian landscapes and ecosystems in the middle Orinoco river basin, Venezuela

57. **Moameni, A.** (1999), Soil quality changes under long-term wheat cultivation in the Marvdasht plain, South-Central Iran
58. **Groenigen, J.W. van** (1999), 90-6164-156-X, Constrained optimisation of spatial sampling: a geostatistical approach
59. **Cheng Tao** (1999), 90-6164-164-0, A process-oriented data model for fuzzy spatial objects
60. **Wolski, Piotr** (1999), 90-6164-165-9, Application of reservoir modelling to hydrotopes identified by remote sensing
61. **Acharya, B.** (1999), 90-6164-168-3, Forest biodiversity assessment: A spatial analysis of tree species diversity in Nepal
62. **Akbar Abkar, Ali** (1999), 90-6164-169-1, Likelihood-based segmentation and classification of remotely sensed images
63. **Yanuariadi, T.** (1999), 90-5808-082-X, Sustainable Land Allocation: GIS-based decision support for industrial forest plantation development in Indonesia
64. **Abu Bakr**, **Mohamed** (1999), 90-6164-170-5, An Integrated Agro-Economic and Agro-Ecological Framework for Land Use Planning and Policy Analysis
65. **Eleveld, M.** (1999), 90-6461-166-7, Exploring coastal morphodynamics of Ameland (The Netherlands) with remote sensing monitoring techniques and dynamic modelling in GIS
66. **Yang Hong** (1999), 90-6164-172-1, Imaging Spectrometry for Hydrocarbon Microseepage
67. **Mainam, Félix** (1999), 90-6164-179-9, Modelling soil erodibility in the semiarid zone of Cameroon
68. **Bakr, Mahmoud** (2000), 90-6164-176-4, A Stochastic Inverse-Management Approach to Groundwater Quality
69. **Zlatanova, Z.** (2000), 90-6164-178-0, 3D GIS for Urban Development
70. **Ottichilo, Wilber K.** (2000), 90-5808-197-4, Wildlife Dynamics: An Analysis of Change in the Masai Mara Ecosystem
71. **Kaymakci, Nuri** (2000), 90-6164-181-0, Tectono-stratigraphical Evolution of the Cankori Basin (Central Anatolia, Turkey)
72. **Gonzalez, Rhodora** (2000), 90-5808-246-6, Platforms and Terraces: Bridging participation and GIS in joint-learning for watershed management with the Ifugaos of the Philippines
73. **Schetselaar, Ernst** (2000), 90-6164-180-2, Integrated analyses of granite-gneiss terrain from field and multisource remotely sensed data. A case study from the Canadian Shield

74. **Mesgari, Saadi** (2000), 90-3651-511-4, Topological Cell-Tuple Structure for Three-Dimensional Spatial Data
75. **Bie, Cees A.J.M. de** (2000), 90-5808-253-9, Comparative Performance Analysis of Agro-Ecosystems
76. **Khaemba, Wilson M**. (2000), 90-5808-280-6, Spatial Statistics for Natural Resource Management
77. **Shrestha, Dhruba** (2000), 90-6164-189-6, Aspects of erosion and sedimentation in the Nepalese Himalaya: highland-lowland relations
78. **Asadi Haroni, Hooshang** (2000), 90-6164-185-3, The Zarshuran Gold Deposit Model Applied in a Mineral Exploration GIS in Iran
79. **Raza, Ale** (2001), 90-3651-540-8, Object-oriented Temporal GIS for Urban Applications
80. **Farah, Hussein** (2001), 90-5808-331-4, Estimation of regional evaporation under different weather conditions from satellite and meteorological data. A case study in the Naivasha Basin, Kenya
81. **Zheng**, **Ding** (2001), 90-6164-190-X, A Neural - Fuzzy Approach to Linguistic Knowledge Acquisition and Assessment in Spatial Decision Making
82. **Sahu, B.K.** (2001), Aeromagnetics of continental areas flanking the Indian Ocean; with implications for geological correlation and reassembly of Central Gondwana
83. **Alfestawi, Y.** (2001), 90-6164-198-5, The structural, paleogeographical and hydrocarbon systems analysis of the Ghadamis and Murzuq Basins, West Libya, with emphasis on their relation to the intervening Al Qarqaf Arch
84. **Liu, Xuehua** (2001), 90-5808-496-5, Mapping and Modelling the Habitat of Giant Pandas in Foping Nature Reserve, China
85. **Oindo, Boniface Oluoch** (2001), 90-5808-495-7, Spatial Patterns of Species Diversity in Kenya
86. **Carranza, Emmanuel John** (2002), 90-6164-203-5, Geologically-constrained Mineral Potential Mapping
87. **Rugege, Denis** (2002), 90-5808-584-8, Regional Analysis of Maize-Based Land Use Systems for Early Warning Applications
88. **Liu, Yaolin** (2002), 90-5808-648-8, Categorical Database Generalization in GIS
89. **Ogao, Patrick** (2002), 90-6164-206-X, Exploratory Visualization of Temporal Geospatial Data using Animation

90. **Abadi, Abdulbaset M.** (2002), 90-6164-205-1, Tectonics of the Sirt Basin – Inferences from tectonic subsidence analysis, stress inversion and gravity modelling
91. **Geneletti, Davide** (2002), 90-5383-831-7, Ecological Evaluation for Environmental Impact Assessment
92. **Sedogo, Laurent G.** (2002), 90-5808-751-4, Integration of Participatory Local and Regional Planning for Resources Management using Remote Sensing and GIS
93. **Montoya, Lorena** (2002), 90-6164-208-6, Urban Disaster Management: a case study of earthquake risk assessment in Carthago, Costa Rica
94. **Ahmad, Mobin-ud-Din** (2002), 90-5808-761-1, Estimation of Net Groundwater Use in Irrigated River Basins using Geo-information Techniques: A case study in Rechna Doab, Pakistan
95. **Said, Mohammed Yahya** (2003), 90-5808-794-8, Multiscale perspectives of species richness in East Africa
96. **Schmidt, Karin** (2003), 90-5808-830-8, Hyperspectral Remote Sensing of Vegetation Species Distribution in a Saltmarsh
97. **Lopez Binnquist, Citlalli** (2003), 90-3651-900-4, The Endurance of Mexican Amate Paper: Exploring Additional Dimensions to the Sustainable Development Concept
98. **Huang, Zhengdong** (2003), 90-6164-211-6, Data Integration for Urban Transport Planning
99. **Cheng, Jianquan** (2003), 90-6164-212-4, Modelling Spatial and Temporal Urban Growth
100. **Campos dos Santos, Jose Laurindo** (2003), 90-6164-214-0, A Biodiversity Information System in an Open Data/Metadatabase Architecture
101. **Hengl, Tomislav** (2003), 90-5808-896-0, PEDOMETRIC MAPPING, Bridging the gaps between conventional and pedometric approaches
102. **Barrera Bassols, Narciso (**2003), 90-6164-217-5, Symbolism, Knowledge and management of Soil and Land Resources in Indigenous Communities: Ethnopedology at Global, Regional and Local Scales
103. **Zhan, Qingming** (2003), 90-5808-917-7, A Hierarchical Object-Based Approach for Urban Land-Use Classification from Remote Sensing Data

104. **Daag, Arturo S.** (2003), 90-6164-218-3, Modelling the Erosion of Pyroclastic Flow Deposits and the Occurrences of Lahars at Mt. Pinatubo, Philippines
105. **Bacic, Ivan** (2003), 90-5808-902-9, Demand-driven Land Evaluation with case studies in Santa Catarina, Brazil
106. **Murwira, Amon** (2003), 90-5808-951-7, Scale matters! A new approach to quantify spatial heterogeneity for predicting the distribution of wildlife
107. **Mazvimavi, Dominic** (2003), 90-5808-950-9, Estimation of Flow Characteristics of Ungauged Catchments. A case study in Zimbabwe
108. **Tang, Xinming** (2004), 90-6164-220-5, Spatial Object Modelling in Fuzzy Topological Spaces with Applications to Land Cover Change
109. **Kariuki, Patrick** (2004), 90-6164-221-3, Spectroscopy and Swelling Soils; an integrated approach
110. **Morales, Javier** (2004), 90-6164-222-1, Model Driven Methodology for the Design of Geo-information Services
111. **Mutanga, Onisimo** (2004), 90-5808-981-9, Hyperspectral Remote Sensing of Tropical Grass Quality and Quantity
112. **Šliužas, Ričardas V.** (2004), 90-6164-223-X, Managing Informal Settlements: a study using geo-information in Dar es Salaam, Tanzania
113. **Lucieer, Arko** (2004), 90-6164-225-6, Uncertainties in Segmentation and their Visualisation
114. **Corsi, Fabio** (2004), 90-8504-090-6, Applications of existing biodiversity information: Capacity to support decision-making
115. **Tuladhar, Arbind** (2004), 90-6164-224-8, Parcel-based Geo-information System: Concepts and Guidelines
116. **Elzakker, Corné van** (2004), 90-6809-365-7, The use of maps in the exploration of geographic data
117. **Nidumolu, Uday Bhaskar** (2004), 90-8504-138-4, Integrating Geo-information models with participatory approaches: applications in land use analysis
118. **Koua, Etien L.** (2005), 90-6164-229-9, Computational and Visual Support for Exploratory Geovisualization and Knowledge Construction
119. **Blok, Connie A.** (2005), Dynamic visualization variables in animation to support monitoring of spatial phenomena

120. **Meratnia, Nirvana** (2005), 90-365-2152-1, Towards Database Support for Moving Object Data
121. **Yemefack, Martin** (2005), 90-6164-233-7, Modelling and monitoring Soil and Land Use Dynamics within Shifting Agricultural Landscape Mosaic Systems
122. **Kheirkhah, Masoud** (2005), 90-8504-256-9, Decision support system for floodwater spreading site selection in Iran
123. **Nangendo, Grace** (2005), 90-8504-200-3, Changing forest-woodland-savanna mosaics in Uganda: with implications for conservation
124. **Mohamed, Yasir Abbas** (2005), 04-15-38483-4, The Nile Hydroclimatology: impact of the Sudd wetland (Distinction)
125. **Duker, Alfred, A.** (2005), 90-8504-243-7, Spatial analysis of factors implicated in *mycobacterium ulcerans* infection in Ghana
126. **Ferwerda, Jelle, G.,** (2005), 90-8504-209-7, Charting the Quality of Forage: Measuring and mapping the variation of chemical components in foliage with hyperspectral remote sensing
127. **Martinez, Javier** (2005), 90-6164-235-3, Monitoring intra-urban inequalities with GIS-based indicators. With a case study in Rosario, Argentina
128. **Saavedra, Carlos** (2005), 90-8504-289-5, Estimating spatial patterns of soil erosion and deposition in the Andean region using Geo-information techniques. A case study in Cochabamba, Bolivia
129. **Vaiphasa, Chaichoke** (2006), 90-8504-353-0, Remote Sensing Techniques for Mangrove Mapping
130. **Porwal, Alok** (2006), 90-6164-240-X, Mineral Potential Mapping with Mathematical Geological Models
131. **Werff, Harald van der** (2006), 90-6164-238-8, Knowledge-based remote sensing of complex objects: recognition of spectral and spatial patterns resulting from natural hydrocarbon seepages
132. **Vlag, Daniël van de** (2006), 90-8504-384-0, Modeling and visualizing dynamic landscape objects and their qualities
133. **Joshi, Chudamani** (2006), 90-8504-470-7, Mapping cryptic invaders and invisibility of tropical forest ecosystems: *Chromolaena odorata* in Nepal
134. **Bandara, K.M.P.S.** (2006), 90-8504-406-5, Assessing irrigation performance by using remote sensing
135. **Dilo, Areti** (2006), 90-8504-461-8, Representation of and Reasoning with Vagueness in Spatial Information. A system for handling vague objects

136. **Debba, Pravesh** (2006), 90-8504-462-6, Sampling scheme optimization from hyperspectral data
137. **Huisman, Marco** (2006), 90-6164-246-9, Assessment of rock mass decay in artificial slopes
138. **Lemmens, Rob** (2006), 90-6164-250-7, Semantic interoperability of distributed geo-services
139. **Chacón Moreno, Eulogio** (2007), 90-8504-559-2, Ecological and spatial modeling: Mapping ecosystems, landscape changes, and plant species distribution in Llanos del Orinoco, Venezuela
140. **Amer, Sherif** (2007), 90-6164-253-1, Towards Spatial Justice in Urban Health Services Planning
141. **Obakeng, Obolokile Thothi** (2007), 90-6164-254-X, Soil moisture dynamics and evapotranspiration at the fringe of the Botswana Kalahari, with emphasis on deep rooting

PE&RC PhD Education Statement Form

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 22 credits (= 32 ECTS = 22 weeks of activities)

Review of Literature (5 credits)

- Hyperspectral remote sensing of biomass and net primary productivity in Mediterranean ecosystem (2003)

Writing of Project Proposal (7 credits)

- Hyperspectral remote sensing of biomass and net primary productivity in Mediterranean ecosystem (2003)

Post-Graduate Courses (4.2 credits)

- Multivariate statistics; PE&RC (2004)
- Advanced geostatistical methods; ITC (2005)
- IDL basic programming; ITC (2006)

Deficiency, Refresh, Brush-up and General courses (0.6 credits)

- Spatial statistical analysis; ITC (2003)
- Accuracy of image classification; ITC (2004)

Competence Strengthening / Skills courses (2 credits)

- Scientific writing course; ITC (2004)
- Oral presentation; ITC (2004)

PhD Discussion Groups (7 credits)

- A forth-nightly PhD discussion group; ITC (2003-2007)

PE&RC Annual Meetings, Seminars and Introduction Days (1.8 credits)

- PE&RC day (2005)
- PE&RC day (2006)
- Three natural resources days organized by the Department of Natural Resources; ITC (2005, 2006, 2007)

- PhD Master class with president of ESRI (2005)

International Symposia, Workshops and Conferences (6 credits)

- ISPRS technical commission VII symposium: Remote sensing from pixel to processes; Thematic processing, modelling and analysis of remotely sensed data (2006)
- SPIE symposium on optics and photonics: Remote sensing and modelling of ecosystems for sustainability III (2006)