

Department für Ökologie und Ökosystemmanagement
Fachbereich für Limnologie

Retrieving canopy variables by radiative transfer model inversion

A regional approach for imaging spectrometer data

Wouter Dorigo

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Wilfried Huber

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Arnulf Melzer
2. Univ.-Prof. Dr. Richard Bamler
3. Prof. Dr. Michael E. Schaepman, Wageningen Univ./
Niederlande (schriftliche Beurteilung)

Die Dissertation wurde am 27.09.2007 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 12.12.2007 angenommen.



GERMAN AEROSPACE CENTER - GERMAN REMOTE SENSING DATA CENTER



Technische Universität München

Copyright © 2007 by Wouter Dorigo

Abstract

High resolution multi- and hyperspectral remote sensing data form the basis of important spatially distributed biogeophysical products in support of crop management at a local to regional scale. Due to the infrequent and irregular data availability from single high resolution sensors, multi-sensor approaches are required to fill the gaps in data sequences and to guarantee data takes at critical time steps during the seasonal life-time of a crop. Prerequisite for such combined use is the consistency of single sensor products and the compatibility between the products of different origin. Developing robust algorithms for the retrieval of vegetation variables from remote sensing data is often hindered by the underdetermined nature of the problem, which is caused by the inferior number of independent earth observation dimensions compared to the large number of canopy elements causing variations in the radiometric signal. The main objective of this thesis is to address this issue in the context of developing an automated, image-based approach which is applicable to a wide variety of airborne and spaceborne high resolution sensors.

The proposed approach, entitled CRASH, provides a concurrent retrieval of the key agricultural variables leaf area index (LAI), leaf chlorophyll content (Cab), leaf water content (Cw), and leaf dry matter content (Cdm), based on the combined inversion of the leaf optical model PROSPECT (Jacquemoud and Baret, 1990; Fourty et al., 1996) and the 1-D turbid medium canopy structure model SAILh (Verhoef, 1984; Verhoef, 1985). The inversion of these radiative transfer models allows for the exploitation of the entire information content contained in the data and facilitates optimizing for illumination/observation geometry and adapting to site specific phenology, background reflectance, and atmospheric conditions. In CRASH, model inversion is based on a lookup table (LUT) approach, which, apart from being computationally fast, offers maximum flexibility to changing input parametrization.

The presented approach entirely relies on the spectral image content and the information products provided by automated preprocessing of the data. Under this assumption, the lack of a priori knowledge on land cover causes the inversion process to be strongly underdetermined and ill-posed, indicating that the system has multiple solutions. Two novel regularization techniques are proposed in this respect: the incorporation of an automated spectral land cover classifier, and the assimilation of preliminary, *a priori* estimates of the output vegetation variables, calculated in place using predictive regression equations. Moreover, a new method, based on the local spatial neighborhood of the pixel under inversion, is suggested for dealing with systematic small scale attribution inconsistencies resulting from the land cover classification.

The suggested land cover classification facilitates a more explicit characterization of spectral uncertainties and appeared very effective in reducing the uncertainties related to the individual inversion results, since the LUTs used for model inversion could be optimized for each land cover type. Moreover, the LUTs that are calculated separately for every land cover class and illumination/observation geometry allow for the generation of semi-empirical predictive equations optimized for each specific situation. These equations are based on the regression between spectral vegetation indices (VIs) calculated from the reflectance spectra contained in the LUT, and the variables used to simulate this canopy reflectance in the forward mode. The predictive equations obtained in this way are subsequently used to calculate a priori estimates of each variable and for every vegetation pixel in the image. The prior estimates, and the assessment of covariance between the several variables based on these estimates, play a decisive role in the stabilization of the model inversion and the reduction of ambiguous results between variables invoking complementary spectral behavior.

The performance and stability of the new approach was extensively tested at three spatial levels, including ground measurements performed with a field spectrometer, data from the airborne full-range imaging spectrometer HyMap, and satellite observations from the multi-directional CHRIS/PROBA sensor. Whereas the field spectrometer measurements allowed for a direct comparison between spectral signal and canopy characteristics while excluding spatial and atmospheric uncertainties, the airborne and satellite measurements offered a deeper insight into the effect of changing observation and illumination properties. The performance of CRASH when applied to the field spectrometer and HyMap data was validated on several intensively used temperate grasslands in southern Germany. Validation measurements included C_w , C_{dm} , and LAI . Ground sampling of C_{ab} and LAI , measured on irrigated cotton fields in Uzbekistan, was used to validate estimates based on the CHRIS data.

Based on resampled field spectra, the influence of different band configurations (HyMap, CHRIS, Landsat ETM+, SPOT HRG, Quickbird) in retrieval performance of C_w , C_{dm} , and LAI was tested. Although the approach is applicable to both multispectral and hyperspectral data, it provided most consistent results for the full-range hyperspectral HyMap configuration, followed by the multispectral Landsat ETM+ configuration. The latter indicates that configurations that show an even distribution of nearly uncorrelated bands across the entire solar-reflective domain contributes more to a robust inversion than a high absolute number of bands in strongly correlating waveband regions, such as provided by CHRIS. The inclusion of SWIR bands clearly lead to regularization of the leaf water retrievals and hence to stabilization of the complete inversion process. The average accuracy obtained for resampled HyMap configuration was around 70% for the 3 considered variables and ranged from 62 - 89% if the grasslands were split up according to phenological classes. The minimum and maximum relative RMS errors (rRMSE) obtained at this spatial level were 14 and 39%, respectively.

At the level of the airborne HyMap data, average estimation accuracy was significantly lower and varied from 37 to 70%, whereas rRMSE laid between 36 - 62%, depending on the variable and on the sun/view constellation. The unsatisfying results obtained at this level were ascribed to the large discrepancy between the small validation plots of $1 \times 1 m^2$ and the average HyMap pixel size of $5 \times 5 m^2$, which especially in the highly heterogeneous pasture played a crucial role. Moreover, consistency of the retrievals among different images of the same area was hindered by anomalies in land cover classification between the scenes, induced by the diverging view/sun constellations of the data takes.

The results obtained at satellite level, based on CHRIS/PROBA multi-angular observations, were not directly comparable with the previous results, since both crop type, spectral sensor configuration, and considered variables differed. Nevertheless, the multi-angular data offered important insight on model performance for a row-structured crop at different view zenith angles. Moreover, it showed the additional gain in stability and accuracy when multiple view angles were simultaneously inverted. For the single view angles, highest accuracy was obtained for the -55° backscatter nominal view angle, while the combination of all view angles except for the extreme forward scattering angle provided best overall performance with an average accuracy of 74% for both C_{ab} and LAI , and a rRMSE of 22 and 33% for C_{ab} and LAI , respectively. The fully automatically generated results were only slightly inferior to those obtained after inclusion of well characterized field measured prior information on the canopy variables. The improvements obtained by using the multi-angular data set underscores the promise residing in this additional information source, even in situations where the radiative transfer model shows considerable discrepancy with the observed canopy.

The results obtained in this thesis for automatically generated biogeophysical products indicate that, based on single sensor observations, full range imaging spectrometers offer most robust radiative transfer model inversion retrievals in cases where little a priori information is available. Even if the obtained results were not accurate enough for multi-sensor applications, for homogeneous vegetation canopies still significant consolidation of the results can be obtained by employing more sophisticated land cover classification schemes. Multi-angular hyperspectral observations in combination with 3-D radiative transfer models seem to bear vast potential for a complete and robust automated characterization of row crops.

Zusammenfassung

Hochaufgelöste multi- und hyperspektrale Fernerkundungsdaten bilden eine wichtige Datengrundlage für die Erstellung räumlich verteilter biogeophysikalischen Produkte zur Unterstützung landwirtschaftlicher Entscheidungsprozessen auf lokaler bis regionaler Ebene. Die eingeschränkte und unregelmäßige Datenverfügbarkeit von einzelnen hochauflösenden Sensoren erfordern multi-sensor Ansätze um die Lücken in den Aufnahmesequenzen zu vervollständigen und Aufnahmen an allen wichtigen Zeitpunkten des saisonalen Wachstumszyklus garantieren zu können. Vorbedingung für solch eine kombinierte Nutzung ist die Konsistenz der einzelnen Sensor-Produkten und die Kompatibilität zwischen den Produkten unterschiedlicher Herkunft. Die Entwicklung von robusten Algorithmen zur Ableitung von Vegetationsvariablen aus Fernerkundungsdaten wird oft erschwert durch den unterbestimmten Charakter des Problems, welcher seinen Ursprung hat in der begrenzten Anzahl der unabhängigen Erdbeobachtungsdimensionen im Vergleich zu den vielen Vegetationsvariablen die für die radiometrischen Schwankungen verantwortlich sind. Die hier vorgestellte Arbeit hat als Hauptziel dieses Problem in den Griff zu bekommen im Rahmen eines neuen vollautomatischen und Bild-basierten Ansatzes der für eine Vielzahl von flugzeug- und satellitenge-tragenen Sensoren entwickelt wurde.

Der Ansatz, namens CRASH, liefert die gleichzeitige Abschätzung vier wichtiger landwirtschaftlicher Variablen: dem Blattflächenindex (LAI), dem Blattchlorophyllgehalt (Cab), dem spezifischen Wassergehalt (Cw), und der spezifischen Trockenmasse (Cdm). Das physikalisch basierte Verfahren beruht auf der kombinierten Inversion des optischen Blattmodells PROSPECT (Jacquemoud und Baret, 1990; Fourty u.A., 1996) und des Vegetations-Strukturmodell SAILh (Verhoef, 1984; 1985). Die Inversion von physikalischen Strahlungstransfer-Modellen ermöglicht es, den gesamten Informationsgehalt der Fernerkundungsdaten auszunutzen und das Modell an die fallspezifische Beleuchtungs- und Aufnahmegeometrie, und der lokalen Phenologie, Hintergrundstrahlung und den atmosphärischen Bedingungen anzupassen. Die Modellinversion erfolgt über Nachschlagetabellen (LUT), welche nicht nur wenig rechenintensiv sind, sondern sich auch sehr flexibel an wechselnde Eingangsparametrisierung anpassen lassen.

CRASH beruht völlig auf dem spektralen Bildinhalt und auf den während der Vorprozessierung entstandenen Zusatzinformationsprodukten. Unter dieser Voraussetzung führt die Abwesenheit von Vorwissen zur Bodenbedeckung dazu dass das Inversionsproblem sehr stark unterdefiniert und schlechtkonditioniert ist. Aus diesem Grund werden zwei neue Regelungsmassnahmen vorgeschlagen: die Einbindung eines automatischen spektralen Bodenbedeckungsklassifikators und die Assimilation einer vorläufigen oder *a priori* Abschätzung des Ergebnisses, die nach einem neuen Verfahren berechnet werden. Des Weiteren wird eine neue Methode vorgestellt die, basierend auf der lokalen Umgebung des zu invertierenden Pixels, kleinräumige Anomalien berücksichtigt die aus der Landnutzungsklassifizierung resultieren.

Die vorgeschlagene Bodenbedeckungsklassifizierung ermöglicht eine explizitere Charakterisierung der radiometrischen Unsicherheiten und zeigt sich sehr effektiv im Reduzieren der Unsicherheiten der einzelnen Inversionsergebnisse. Außerdem bieten die LUTs, die für jede einzelne Bodenbedeckungsklasse und Blick- / Sonnengeometrie erstellt werden, die Möglichkeit semi-empirische Regressionsgleichungen zu generieren, die für jede Situation optimiert sind. Diese Gleichungen basieren auf der Regression zwischen spektralen Vegetationsindizes (VI), berechnet aus den Reflexionsdaten in der LUT, und den Variablen die verwendet wurden um die Bestandesreflexionswerte zu simulieren. Anschließend werden die Regressionsgleichungen verwendet, um für jeden Pixel in der Szene eine vorläufige Abschätzung jeder Variable

zu berechnen. Die a priori Abschätzungen und die Kovarianz, die basierend auf diesen Schätzungen berechnet wird, spielen eine entscheidende Rolle in der Stabilisierung der Modellinversion und in der Reduzierung der Mehrdeutigkeit zwischen Variablen mit sich ergänzender spektraler Auswirkung.

Die Leistung und Stabilität des neuen Verfahrens wurden getestet auf drei räumlichen Skalenniveaus. Dazu gehören Bodenmessungen durchgeführt mit einem Feldspektrometer, Daten aufgenommen mit dem flugzeuggetragenen HyMap-Sensor und multi-direktionale CHRIS/PROBA Satellitenaufnahmen. Während die Feldspektrometermessungen den direkten Vergleich zwischen spektraler Signatur und Bestandesmessung ermöglichten, indem geometrische und atmosphärische Unsicherheiten ausgeschlossen wurden, boten die flugzeug- und satellitengestützten Aufnahmen einen tieferen Einblick in die Effekte verursacht durch wechselnde Aufnahme- und Sonnengeometrie. Für die Feldspektrometermessungen und HyMap-Aufnahmen wurde die Validierung des Verfahrens auf mehreren intensiv bewirtschafteten Grünlandflächen in Südostbayern durchgeführt, wobei C_w , C_{dm} und LAI gemessen wurden. Bodenmessungen von C_{ab} und LAI , erhoben für Baumwolle in Uzbekistan, wurden verwendet um die Abschätzungen basierend auf den CHRIS-Daten zu validieren.

Mittels resampten gemessenen Feldspektren wurde getestet, welche Auswirkung unterschiedliche hyper- und multi-spektrale Bandkonfigurationen (HyMap, CHRIS, Landsat ETM+, SPOT HRG, Quickbird) auf die Ableitungsgenauigkeit von C_w , C_{dm} , and LAI hatten. Die HyMap spektrale Konfiguration, mit gleichmäßig verteilten schmalen Bändern über das gesamte reflektive Spektrum, schnitt am besten ab, gefolgt von der multi-spektralen Landsat ETM+ Konfiguration. Letzteres Ergebnis zeigt, dass Konfigurationen mit gleichmäßig über das Gesamtspektrum verteilten unkorrelierten Kanälen mehr zu einer stabilen Inversion beitragen als eine hohe Zahl von korrelierten Bändern in einem eingeschränkten spektralen Bereich, ähnlich wie beim CHRIS-Sensor. Die Angliederung von SWIR Kanälen führt zu einer deutlichen Stabilisierung der Wassergehaltsabschätzung und damit zu einer Stabilisierung der gesamten Invertierung. Die mittlere Genauigkeit, die für die HyMap Konfiguration erreicht wurde, lag bei 70% für die drei Variablen und variierte zwischen 62-89%, wenn die Grünlandflächen nach phenologischen Klassen aufgeteilt wurden. Die minimale und maximale relative mittlere quadratische Abweichung (rRMSE), die auf diesem räumlichen Niveau erreicht wurden, lagen bei 14% beziehungsweise 39%.

Für die Daten des flugzeuggetragenen HyMap Sensors lag die mittlere Schätzungsgenauigkeit mit 37 - 70% bedeutend niedriger, während die rRMSE abhängig von Variable und Sonnen- / Aufnahmekonstellation zwischen 36 - 63% schwankte. Diese unbefriedigenden Ergebnisse sind der großen Diskrepanz zwischen den kleinen Validierungsplots ($1 \times 1 m^2$) und dem mittleren HyMap Pixelgröße von $5 \times 5 m^2$ zuzuschreiben, welche vor allem in den heterogenen Weiden eine ausschlaggebende Rolle spielten. Hinzu kommt, dass die Konsistenz der Schätzungsgenauigkeit zwischen unterschiedlichen Aufnahmen von einem und dem selben Gebiet durch Anomalien in der Bodenbedeckungsklassifikation beeinflusst wurde, verursacht durch die unterschiedliche Beobachtungs- und Sonnengeometrie der Datenaufnahmen.

Die Ergebnisse, die auf Satellitenebene aus den multi-direktionalen CHRIS/PROBA-Daten erzielt wurden, sind nicht direkt mit den oben erwähnten Ergebnissen vergleichbar, da sowohl der Vegetationstyp als Sensor-Konfiguration und auch berücksichtigte Variablen unterschiedlich waren. Dennoch ermöglichten die multi-direktionalen Daten einen wichtigen Einblick in die Leistung des Schätzungsverfahrens für in Reihen angeordnete Anbaupflanzen unter unterschiedlichen Beobachtungswinkeln. Außerdem wurde gezeigt, dass die gleichzeitige Inversion von mehreren Beobachtungswinkeln zu einer zusätzlichen Stabilität und Genauigkeit des Schätzungsverfahrens beiträgt. Was die einzelnen Beobachtungswinkel angeht, wurde die höchste Genauigkeit für den 55° nominalen Rückstreuungs- Blickwinkel erreicht, während die Kombination von allen Blickwinkeln außer dem maximalen vorwärts gerichteten Winkel die gesamtbeste Genauigkeit erzielte. Im letzten Fall betrug die mittlere Genauigkeit 74%, sowohl für C_{ab} als LAI , und die rRMSE bei 22 und 33% für C_{ab} beziehungsweise LAI . Die völlig automatisch erzeugten Ergebnisse waren nur unwesentlich schlechter als die Ergebnisse unter der Verwendung der Bodenmessungen als a priori Wissen. Die Verbesserungen, die durch Einbeziehung von mehreren Blickwinkeln erreicht wurden, unterstreicht den zusätzlichen nutzbaren Informationsgehalt von solchen Daten sogar in einer Situation, wo das Strahlungstransfermodell eine große Diskrepanz mit der Wirklichkeit vorzeigt.

Die in dieser Dissertation automatisch erzeugten biogeophysikalischen Produkte zeigen dass basierend

auf einzelnen Aufnahmen, abbildende Spektrometer, die den gesamten reflektiven Spektralbereich von 400 - 2500 nm abdecken, die robustesten Ergebnisse liefern falls kein a priori Wissen vorhanden ist. Obwohl die erzielten Genauigkeiten noch nicht ausreichen für multi-sensor Ansätze, kann eine erhebliche zusätzliche Konsolidierung der Ergebnisse erreicht werden durch weiter verfeinerte Bodenbedeckungs- und Landnutzungsklassifikatoren. Multi-direktionale hyperspektrale Aufnahmen zeigen großes Potential für eine komplette und robuste automatisierte Charakterisierung von in Reihen angeordneten Anbaupflanzen.

Acknowledgments

This study was carried out within the framework of the 'High-Tech Offensive Zukunft Bayern' project No. 290 ('Pilotprojekt Waging-Tachinger See' of the Limnological Station of the TU München) financed by the Bavarian State Ministry of Science, Research, and the Arts. Propagation of the work was ascertained by the scientific exchange program of the German Aerospace Center. However, the completion of a thesis is never driven by financial incentives alone. It would never have been possible without the support of the following persons.

First of all, I would like to thank my "Doktorvater" Prof. Dr. Arnulf Melzer for his confidence and support, the pleasant evenings at the Limnological Station in Iffeldorf, and for baking the best pizza that I ever had outside of Italy. Prof. Dr. Richard Bamler and Prof. Dr. Michael Schaepman of Wageningen University, thank you for examining this thesis and for the fruitful comments that you gave me during the writing process. Many thanks to Dr. Uta Raeder, Stefan Zimmermann, and Dr. Thomas Schneider who were the ones that offered me the possibility to start this PhD research in the first place. Tomi, I really enjoyed your enthusiasm during the various field campaigns and the passion you show for spectrodirectional remote sensing.

I would like to thank Dr. Frédéric Baret of INRA in Avignon, France, who introduced me to radiative transfer model inversion. During, and after my repeated stays in his department he was a big source of inspiration and always open to share his ideas and knowledge. Thanks to all the others of the Batiment Sol-Climat for your valuable advices and the pleasant time during and after work. Prof. Dr. Wout Verhoef of NLR/ITC, thank you for your good advice, the intercomparison of SAILh simulations, and for providing me with so much interesting literature.

Andreas Müller, "el jefe" of the Imaging Spectroscopy (AS) Group of DFD, thank you for your fruitful feedback during all these years and your very "unique" way of motivating people, which was the driving force behind the conclusion of this work. Not to forget the frequent eating and drinking occasions that we enjoyed during and after work. A very special thanks goes out to Dr. Rolf Richter, my unofficial supervisor: you were always willing to share parts of your source code with me, giving me very practical feedback on methodology and programming, and proof reading the chapters of this thesis. Martin Bachmann, thank you very much for your viable scientific advices, proof reading parts of the thesis, and the nice chats in and outside of DLR. Steffi Holzwarth, sunshine of the AS team, thank you for always sharing your sparkling good mood with the rest of the team. Martin Habermeyer, Lena Lieckfeld, Wieke Heldens, Hannes Taubenböck, Klaas Scholte, and all colleagues, guest scientists, Diplomanden, and Praktis in the AS group and the rest of DFD-US, thanks for the many chats, sometimes with, mostly without meaningful content, at the coffee machine, in Schlemmerland, or after work.

I wish to thank colleagues of the DLR flight department and IMF: without them I would never have had data to base this thesis on. I am very grateful to Gerd Ruecker and Joerg Grillenberger for the enormous effort they made in organizing and implementing the field campaign

in Uzbekistan and sharing the data with me. Gerd, thank you too for proof reading parts of this thesis. Thanks as well to all the Uzbek students that I don't know personally but who I know they worked hard to collect the field data. The CHRIS/PROBA team of ESA for providing the CHRIS/PROBA data and Luis Alonso of the University of Valencia for the preprocessing tools for CHRIS. Thanks to all hyperspectral remote sensors from RSL, ITC, Wageningen University, Utrecht University, Warsaw University, Tel Aviv University, University of Valencia, and all other institutes I forgot to mention, for the pleasant cooperations, the interesting discussions on imaging spectrometry, and the amusing nights in the pub during conferences.

Prof. Dr. Wolfgang Wagner and Dr. Markus Hollaus, thank you for giving me the opportunity to continue in science as a postdoctoral researcher at the Institute of Photogrammetry and Remote Sensing of the Vienna University of Technology, even without being a PhD yet. I greatly appreciate your support and your generosity allowing me to work on my thesis during working hours when time was getting short.

Bringing a PhD thesis to an end also requires switching off from science every now and then. Arko, David, Raymond, Armin, and Stevie, thank you for being true friends for so many years, sharing ups and downs, and showing me how beautiful life can be. Arko, David and Armin, you know like no one else how tough it sometimes can be to combine a passion for mountains with a career in science. Many thanks to Markus, Johannes, Stefan, Ulf, Wiebe, Robbert, Arne, Julia and all other climbing friends, for the great moments we had together in the Wilder Kaiser, Ortler, Kackar, and all other mountain ranges we explored together on skis, with two ice axes, or just with a chalk bag and a pair of climbing shoes.

I would especially like to thank my mother Clémentine, my father Valentino, and my sister Marlies. They were always there for me and with them I could always be myself, no matter in which mood I was. I am really happy with Pien en Vik as well, who came as a bonus together with my girlfriend Jasminka. Jasminka, I have no words to describe how special you are. Allowing me to move to Munich and Vienna for my career and supporting me whenever I needed it, no matter how evil I could sometimes be to you in times of stress, are just a few examples of the tremendous support and love you have given me during all these years. Thank you.

Contents

Abstract	v
Zusammenfassung	vii
Acknowledgments	xi
List of Figures	xix
List of Tables	xxiii
List of Abbreviations	xxv
List of symbols	xxvii
1 Introduction	1
1.1 Earth observation in a changing environment	1
1.2 User requirements on remote sensing biogeophysical products	2
1.3 The challenge of providing robust environmental indicators	3
1.4 Rationale of the thesis	5
1.5 Research objectives	6
1.6 Structure of the thesis	9
2 Characterizing vegetation canopies by solar-reflective remote sensing	11
2.1 Principles of solar-reflective remote sensing	11
2.1.1 Electromagnetic radiation	11
2.1.2 The solar-reflective domain	13
2.1.3 Atmospheric interaction	13
2.1.4 System requirements	16
2.1.5 Observation properties	16
2.2 Canopy reflectance	17
2.2.1 Leaf optical properties	17
2.2.2 Canopy structure	19
2.2.3 Background reflectance	21
2.2.4 Anisotropy in canopy reflectance	21
2.3 Retrieving canopy variables by statistical approaches	23
2.4 Spectral unmixing for cover fraction estimates	25
2.5 Physical approaches for retrieving vegetation variables	27

2.5.1	Leaf optical models	27
2.5.2	Canopy radiative transfer models	29
2.5.3	Radiative transfer model inversion	31
2.5.3.1	Iterative optimisation techniques	31
2.5.3.2	Lookup table approach	32
2.5.3.3	Artificial neural networks	33
2.5.4	Under-determination and ill-posedness in radiative transfer model inversion	33
2.5.5	Improving retrieval performances	34
2.5.5.1	Increasing the observation dimensionality	34
2.5.5.2	Constraints on input variables	35
2.6	Conclusions	37
3	The automated CRASH approach: theoretical concept and implementation	39
3.1	Introduction	39
3.2	Choosing an appropriate radiative transfer model	40
3.2.1	The leaf optical model PROSPECT	41
3.2.1.1	N compact layers	41
3.2.1.2	Biochemical components	41
3.2.1.3	Simplifications and limitations	45
3.2.2	The canopy reflectance model SAILh	45
3.2.2.1	4-stream approximation	45
3.2.2.2	Canopy characterization	46
3.2.2.3	The hot spot effect	48
3.2.2.4	Assumptions and caveats	49
3.2.3	The soil background model	50
3.2.4	Interaction of variables	51
3.3	The inversion approach	51
3.3.1	Justifying a LUT based inversion scheme	51
3.3.2	Land cover classification for improved retrieval performance	53
3.3.3	Considerations underlying lookup table parametrization	55
3.3.3.1	Variable ranges	56
3.3.3.2	Variable distribution functions	57
3.3.3.3	Sampling strategy	57
3.3.3.4	Sampling distance	58
3.3.3.5	Incorporating view and sun geometry	58
3.3.3.6	Calculating diffuse radiance	59
3.3.3.7	Background reflectance parametrization	59
3.4	The optimization algorithm	60
3.4.1	Exploiting radiometric information	62
3.4.1.1	Adding local spectral variance to overcome classification anomalies	64
3.4.1.2	Covariance matrix inversion	65
3.4.2	Exploiting a priori information on canopy variables	66
3.4.2.1	Defining prior estimates using vegetation indices	66
3.4.2.2	Assigning uncertainty to the prior estimates	70
3.4.2.3	Defining the final solution	71
3.4.3	Accuracy description	72

3.4.4	Accounting for angular anisotropy	72
3.5	Conclusions and preface to Chapter 4 and 5	74
4	Validating CRASH at ground and airborne level: grassland characterization using field spectrometer and HyMap data	77
4.1	Introduction	77
4.1.1	Preceding grassland studies using imaging spectroscopy	78
4.1.2	Objectives	79
4.2	Study site and data	79
4.2.1	Study site Waging-Taching	79
4.2.2	Ground validation measurements	81
4.2.2.1	Configuration of ground sampling locations	81
4.2.2.2	Leaf area index	81
4.2.2.3	Leaf dry matter and water content	84
4.2.2.4	Comparing results obtained with direct and indirect LAI sampling	85
4.2.2.5	Field spectrometer measurements	86
4.2.3	HyMap imaging spectrometer measurements	88
4.2.3.1	Sensor characteristics	88
4.2.3.2	Flight configuration	89
4.2.3.3	Geometric correction	89
4.2.3.4	Calibration and atmospheric correction	91
4.3	Exploring algorithm potential and constraints using field spectrometer data	92
4.3.1	Spectral field characteristics	93
4.3.2	Comparing modeled with measured reflectance	94
4.3.3	Correlation between canopy variables and wavebands	95
4.3.4	Stepwise integration of algorithm components	97
4.3.4.1	Influence of land cover classification	99
4.3.4.2	Quantifying spectral covariance and the influence of sensor configuration	100
4.3.4.3	Introducing prior information on variables	104
4.3.4.4	Introducing covariance between the variables	105
4.3.4.5	Integrating a priori estimates based on predictive regression functions	106
4.3.4.6	Comparing predictions based on regression functions with final RTM estimates	108
4.3.5	Exploring additional regularization	109
4.3.5.1	Estimating biochemicals at canopy level	110
4.3.5.2	Coupling C_w with C_{dm}	111
4.3.6	Model sensitivity to LUT parametrization	112
4.3.6.1	Reproducibility of estimates	112
4.3.6.2	Dependence on variable ranges	114
4.3.7	Discussion	117
4.4	RTM inversion applied to HyMap flight lines	121
4.4.1	Accounting for spectral anisotropy	122
4.4.1.1	Quantifying spectral anisotropy	122
4.4.1.2	Incorporating view angle information in model inversion	125

4.4.2	Influence of land cover classification on retrieval performance	126
4.4.3	Inconsistencies and benefits introduced by the predictive regression functions	129
4.4.4	Validation at test sites	131
4.4.4.1	Observed patterns	131
4.4.4.2	Comparison with ground validation measurements	135
4.5	Conclusions	136
5	Validating CRASH at satellite level: estimating cotton leaf chlorophyll content from multi-angular CHRIS/PROBA observations	141
5.1	Introduction	141
5.1.1	Multi-angular remote sensing	141
5.1.2	Land and water use restructuring in the Uzbek Khorezm region	142
5.1.3	Imaging spectroscopy in cotton studies	142
5.1.4	Objectives	144
5.2	Study site and data	144
5.2.1	Study site Khorezm	144
5.2.2	Biometric sampling	146
5.2.2.1	Sampling scheme	146
5.2.2.2	Chlorophyll	146
5.2.2.3	Leaf area index	147
5.2.2.4	Additional measurements in support of RTM inversion	148
5.2.3	CHRIS multi-angular satellite observations	150
5.2.3.1	Sensor and view characteristics	150
5.2.3.2	Preprocessing	151
5.3	Quantifying spectral anisotropy	152
5.3.1	Observed directional signatures	152
5.3.2	Reconstructing spectral anisotropy with PROSPECT and SAILh	154
5.4	RTM inversion using the fully automated CRASH approach	157
5.4.1	RTM inversion of individual observation angles	157
5.4.2	RTM inversion of combined observation angles	160
5.4.2.1	Adapting the cost function	160
5.4.2.2	Iteratively adding view angles	160
5.5	RTM inversion including measured a priori information on cotton	163
5.5.1	Parametrization of the LUTs	163
5.5.2	Results and discussion	164
5.6	Conclusions	166
6	Synthesis	169
6.1	Introduction	169
6.2	Resuming the performance of the automated approach	170
6.3	Conclusions	175
6.4	Outlook	177
A	SPECL classification rules	179
B	Variable sampling plan for LUTs	181
B.1	SPECL class 2: Dark vegetation	181

B.2	SPECL class 3: Average vegetation	182
B.3	SPECL class 4: Bright vegetation	183
B.4	SPECL class 5: Yellow vegetation	184
B.5	SPECL class 6: Mixed vegetation / soil	186
B.6	SPECL class 12: Dry vegetation / soil	187
B.7	SPECL class 13: Sparse vegetation / soil	188
C Overview of used vegetation indices		191
D Spectral configuration HyEurope 2003		195
E Canopy variables measured at Waging test site		197
F Spectral configuration CHRIS Mode 5 in 2006		201
Bibliography		203
Publications of the author		227
Curriculum Vitae		229

List of Figures

1.1	Processing chain for hyperspectral data operated by DLR-DFD	7
2.1	Atmospheric transmittance in the optical domain (a), schematic sketch of radiance components measured at a detector element of a high altitude sensor (b)	15
2.2	Hyperspectral versus multispectral observation of a vegetation canopy in relation to spectral absorption features of leaf biochemicals	17
2.3	Measurement geometry in solar-reflective remote sensing	18
2.4	Schematic cross section of a dicotyledon leaf	18
2.5	Hemispherical leaf reflectance and transmittance	20
2.6	Spectral similarity of soil and non-photosynthetic vegetation	22
2.7	Effect of volumetric soil water content on soil reflectance	22
2.8	Example of reflectance anisotropy in the red wavelength domain for a maize crop	23
2.9	Example of hot spot effect in HyMap airborne imagery	24
2.10	Exponential empirical relationship between WDVI and LAI	26
2.11	Concept of radiative transfer model inversion	28
3.1	Overview of CRASH inversion approach	40
3.2	PROSPECT leaf optical model	42
3.3	Standardized PROSPECT absorption coefficients for leaf chlorophyll, leaf dry matter, leaf water, and leaf brown pigment content	43
3.4	Effects of N , Cab , Cw , and Cdm on leaf reflectance and transmittance	44
3.5	4-stream approximation employed in SAILh	47
3.6	Graphical versus ellipsoidal formulation of LIDF	48
3.7	Ambiguity between the canopy structure variables LAI and ALA	52
3.8	Effect of posing lower and upper limits on number and sampling density of variables in LUT	54
3.9	Template nadir vegetation reflectance spectra employed by the SPECL code . . .	55
3.10	Effect of classification anomalies on variable retrieval pattern	56
3.11	Sampling in a 2-D variable space resulting from global and stratified random variable selection	58
3.12	Diffuse fraction of incident radiance calculated for a standard atmosphere	60
3.13	Scene based soil characterization employed in CRASH	61
3.14	Evidence for the soil brightness parameter	61
3.15	Range of reflectances simulated for the classes bright and sparse vegetation . . .	62
3.16	Optimization algorithm proposed in this study	63
3.17	Variance-covariance matrices for two SPECL land cover classes	64

3.18	Example of attributing atmospheric, sensor, and model uncertainties to a simulated spectrum	68
3.19	Examples of predictive regression functions between best performing vegetation indices and C_w , LAI, and BS, based on RTM simulations	70
4.1	Location of Waging-Taching test site	80
4.2	Location of selected test fields and plots used for canopy characterization	82
4.3	The three different types of grassland sampled during first HyMap data campaign	82
4.4	Procedure followed for destructive sampling of LAI	84
4.5	Comparison between LAI measured by LAI-2000 and destructive sampling	87
4.6	C_{dm} and C_w based on directly (empirically) and indirectly (LAI-2000) measured LAI	87
4.7	Georectified false color composite of three flightlines recorded at June 30, 2003	90
4.8	Comparison of 3 georectified and image matched subsets of the flightlines recorded at June 30, 2003	91
4.9	Comparison of average field spectrometer and atmospherically corrected HyMap reflectance data	92
4.10	Average spectral signatures of MEA1, MEA2, and PAS for the campaigns HYM1 and HYM2	94
4.11	Comparing modeled and measured reflectance for two sampling locations	96
4.12	Correlation between canopy reflectance and measured and modeled C_w , C_{dm} , and LAI for vigorous green grassland	98
4.13	Correlation between canopy reflectance and measured and modeled C_w , C_{dm} , and LAI for dry, cut meadows	98
4.14	RTM inversion results for C_w , C_{dm} , and LAI applying a global LUT and without using prior information	101
4.15	RTM inversion results for C_w , C_{dm} , and LAI applying LUTs adjusted to SPECL classes, and without using prior information	101
4.16	Effect of accounting for spectral covariance on estimates of C_{ab} , LAI and ALA	104
4.17	RTM inversion results for C_w , C_{dm} , and LAI using LUTs adjusted to SPECL classes and prior information based on field measurements	105
4.18	RTM inversion results for C_w , C_{dm} , and LAI using LUTs adjusted to SPECL classes and including prior information and covariance between the variables based on field measurements	106
4.19	RTM inversion results for C_w , C_{dm} , and LAI using LUTs adjusted to SPECL classes and prior information based on predictive equations	108
4.20	Influence of introducing a prior estimates based on predictive regression equations on C_{ab} retrieval	108
4.21	Reducing ambiguity by accounting for covariance between the variables	109
4.22	Comparing final RTM inversion results of C_w , C_{dm} , and LAI with prior estimates obtained by predictive regression equations	110
4.23	RTM inversion results for canopy water and dry matter content, using the synthetic variables $C_w \times LAI$ and $C_{dm} \times LAI$	111
4.24	Mean and maximum absolute deviation between single estimated variables and the average result over 12 model runs	114

4.25	Sensitivity of RTM inversion results to variable ranges used to construct the LUT for SPECL class 4	116
4.26	Sensitivity of RTM inversion results to variable ranges used to construct the LUT for SPECL class 6	118
4.27	Spectral anisotropy factors for the HyMap scenes WAGING07, WAGING13, and WAGING14	123
4.28	Relative spectral differences in red and NIR reflectance between three overlapping subsets	124
4.29	Effect of spectral anisotropy on NDVI-based <i>LAI</i> estimates	125
4.30	Intercomparison of <i>LAI</i> estimated from different observation/sun constellations .	126
4.31	Comparison of leaf variable estimates from 3 HyMap scenes with varying view/sun geometry	127
4.32	Comparison of canopy variable estimates from 3 HyMap scenes with varying view/sun geometry	128
4.33	Results of automatic SPECL classification of nadir-normalized HyMap data of June 30, 2003	129
4.34	Correlation between biased classification results and deviations in <i>LAI</i> and <i>C_w</i> retrievals	130
4.35	Example of effect of land cover classification and inclusion of local variance on prior estimates on <i>LAI</i> retrieval	131
4.36	Prior estimates of <i>C_{ab}</i> , <i>C_w</i> , and <i>LAI</i> for flightline WAGING13	131
4.37	Relationship between estimates of <i>LAI</i> and <i>ALA</i> for prior estimates and final RTM inversion results	132
4.38	Spatial distribution at MEA1, MEA2, and PAS of leaf variables estimated from flightline WAGING07	133
4.39	Spatial distribution at MEA1, MEA2, and PAS of canopy variables estimated from flightline WAGING07	134
4.40	Estimated versus measured <i>C_w</i> , <i>C_{dm}</i> , and <i>LAI</i> , based on HyMap scenes with differing view/sun constellations	136
4.41	Variables estimated from ASD field spectrometer data versus estimates based on HyMap scene WAGING07	137
5.1	Location of the Khorezm test site	145
5.2	Examples of cotton canopies encountered in the study area at the time of satellite overpass	146
5.3	Regression fit between SPAD and laboratory measurements of chlorophyll a+b .	147
5.4	Spatial distribution of measured <i>C_{ab}</i> and <i>LAI</i>	149
5.5	View/sun constellation of CHRIS observations and overlay of georeferenced multi-angular CHRIS images	151
5.6	Green, red, and NIR HDRF for various fields at 5 different view angles	153
5.7	Anisotropy index (ANIX) and average nadir reflectance for six different fields . .	154
5.8	Influence of changing <i>LAI</i> on spectral anisotropy, simulated with PROSPECT and SAILh for green, red, and NIR reflectance	156
5.9	Influence of changing <i>ALA</i> on spectral anisotropy, simulated with PROSPECT and SAILh for green, red, and NIR reflectance	156

5.10	Influence of changing hot spot parameter on spectral anisotropy, simulated with PROSPECT and SAILh for green, red, and NIR reflectance	157
5.11	Comparison of estimated <i>LAI</i> and <i>Cab</i> with measured <i>LAI</i> , SPAD, and <i>Cab</i> . .	159
5.12	Evolution of average rRMSE and AA in relation to view angle	159
5.13	Evolution of average rRMSE and AA in relation to different combinations of view angles	161
5.14	Comparing the spatial distribution of relative standard errors for mono- and multi-directional model inversion results of <i>Cab</i>	161
5.15	Comparing the spatial distribution of relative standard errors for mono- and multi-directional model inversion results of <i>LAI</i>	162
5.16	Spatial distribution of <i>Cab</i> and <i>LAI</i> for multi-directional model inversion results	163
5.17	Comparing measured with estimated <i>Cab</i> and <i>LAI</i> , obtained from the concurrent inversion of -55, -36, 0, and +36° nominal view angles	163
5.18	Frequency distribution of field measured <i>Cab</i> and <i>LAI</i>	165
5.19	Comparing measured with estimated <i>Cab</i> and <i>LAI</i> , obtained from the concurrent inversion of -55, -36, 0, and +36° nominal view angles using field measurement values for the parametrization of the LUT	165

List of Tables

1.1	Variables indicative for canopy condition and functioning	3
2.1	Categorization of wavelength ranges	13
2.2	Approaches used to model leaf and canopy radiative transfer	29
2.3	Differences between radiative transfer and radiosity approach	30
3.1	Structure variable N and leaf biochemicals for a selection of agricultural crops .	42
3.2	Input parameters used in SAILh	50
3.3	Quality flags used to indicate confidence of model inversion	72
4.1	Field level results of grassland canopy measurements	86
4.2	Technical specifications of the HyMap sensor	88
4.3	Observation and illumination details of the 3 HyMap flight lines used in this study and recorded at June 30, 2003	89
4.4	HyMap bands that were excluded from thematic processing	93
4.5	Distribution of the input variables used for the simulation of the canopy spectra in Figure 4.11	95
4.6	Distribution of the input variables and sampling plan used to construct a global LUT to match all land cover types.	100
4.7	Spectral configuration of HyMap 2003, CHRIS Mode5, Landsat ETM, Quickbird, and SPOT HRG	103
4.8	Retrieval performance for different sensor configurations using two different ra- diometric cost functions	103
4.9	Best predictive equations based on simulated spectra in different SPECL classes	107
4.10	Comparing retrieval accuracy for C_w , C_{dm} , and LAI based on predictive equa- tions with those obtained with complete algorithm	110
4.11	RTM inversion results for C_w , C_{dm} , and LAI when using a constrained rapport between C_{dm} and C_w	113
4.12	Retrieval performance for different sensor configurations when introducing prior information on the variables	119
4.13	Retrieval performance separated according to three phenological classes	120
4.14	Accuracy of estimated C_w , C_{dm} , and LAI for 3 different observation/illumination conditions	137
5.1	Crop management specifications of validation fields	145
5.2	Descriptive statistics of cotton canopy measurements	148
5.3	Variables measured in support of RTM inversion	150

5.4	Observation and illumination details of the CHRIS imagery used in this study . .	150
5.5	Goodness of fit between measured SPAD, <i>Cab</i> , and <i>LAI</i> and the values obtained from 5 separate CHRIS view angles	159
5.6	Distribution of the input variables and the corresponding number of classes of the orthogonal sampling plan used to construct the LUT for cotton	164
A.1	Decision rules used in SPECL classification	180
C.1	Vegetation indices used for predictive equations.	193
D.1	Spectral configuration of HyEurope 2003, arranged according to the 4 detectors .	196
E.1	Results of canopy measurements for grassland measured during HYM1 campaign	198
E.2	Results of canopy measurements for grassland measured during HYM2 campaign	199
F.1	Spectral configuration of CHRIS Mode 5 in 2006	202

List of Abbreviations

AA	Average Accuracy
ALA	Average Leaf Angle
ARES	Airborne Reflective Emissive Spectrometer
ASD	Analytical Spectral Devices
ATCOR	ATmospheric CORrection program (R. Richter)
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
BCRF	BiConical Reflectance Factor
BHRF	BiHemispherical Reflectance Factor
BRDF	Bidirectional Reflectance Distribution Function
BRF	Bidirectional Reflectance Factor
CASI	Compact Airborne Spectrographic Imager
CHRIS	Compact High Resolution Imaging Spectrometer onboard PROBA (ESA)
CRASH	Canopy variable Retrieval Algorithm based on prospect and Sailh
DEM	Digital Elevation Model
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center)
EnMAP	Environmental Mapping and Analysis Program
ENVI	ENvironment for Visualizing Images (ITT-VIS, Inc.)
ESA	European Space Agency
ETM+	Enhanced Thematic Mapper plus onboard Landsat 7
EWT	Equivalent Water Thickness
f APAR	Fraction of Absorbed Photosynthetic Active Radiation
f Cover	Fractional cover
FOV	Field of View
FWHM	Full Width Half Maximum
GPS	Global Positioning System
HCRF	Hemispherical Conical Reflectance Factor
HDRF	Hemispherical Directional Reflectance Factor
HyMap	Hyperspectral Mapper (Integrated Spectronics)
IDL	Interactive Data Language (ITT-VIS, Inc.)
IFOV	Instantaneous Field Of View
LAD	Leaf Angle Distribution
LAI	Leaf Area Index

LIDF	Leaf Inclination Distribution Function
LUT	Look-Up Table
MERIS	MEDium Resolution Imaging Spectrometer (ESA)
MODIS	MODerate resolution Imaging Spectrometer (NASA)
MISR	Multiangle Imaging SpectroRadiometer
MIR	Mid InfraRed part of electromagnetic spectrum
NIR	Near InfraRed part of electromagnetic spectrum
OA	Overall Accuracy
OP	Orthogonal Plane
POLDER	POLarization and Directionality of the Earths Reflectances
PP	Principle Plane
PROBA	PRoject for On-Board Autonomy platform
PROSPECT	leaf optical PROPERTIES SPECTra model
RAMI	RAdiation transfer Model Intercomparison
RMSE	Root Mean Square Error
rRMSE	relative Root Mean Square Error
RTM	Radiative Transfer Model
SAILh	Scattering from Arbitrarily Inclined Leaves, including hot spot parameter h
SNR	Signal to Noise Ratio
SPECL	SPEctral CLassification program (R. Richter, DLR)
SWIR	ShortWave InfraRed part of electromagnetic Spectrum
TM	Thematic Mapper
TOC	Top-Of-Canopy
VI	Vegetation Index
VIS	VISible part of electromagnetic spectrum
VNIR	Visible and Near InfraRed part of the electromagnetic spectrum

List of symbols

λ	Wavelength	[nm, μm]
θ_s, ϑ_s	Solar zenith angle	[$^\circ$],[rad]
θ_v, ϑ_v	View zenith angle	[$^\circ$],[rad]
ψ_s	Solar azimuth angle	[$^\circ$],[rad]
ψ_v	View azimuth angle	[$^\circ$],[rad]
ψ_{rel}	Relative azimuth angle	[$^\circ$],[rad]
Ω	Observation/illumination geometry	
ρ, R	Reflectance	[dimensionless]
τ	Transmittance	[dimensionless]
χ^2	Merit function	
Cab	Leaf chlorophyll content	[$\mu\text{g}/\text{cm}^2$]
Cw	Leaf water content	[mg/cm^2]
Cdm	Leaf dry matter content	[mg/cm^2]
Cbp	Leaf brown pigments	[dimensionless]
E	Irradiance, radiance flux density	[$\text{W}/\text{m}^2\mu\text{m}$]
E^-, E^+	Downward and upward diffuse spectral irradiance vectors	[$\text{W}/\text{m}^2\mu\text{m}$]
E_s	Solar direct spectral irradiance	[$\text{W}/\text{m}^2\mu\text{m}$]
E_o	Direct radiance flux density in direction of observer	[$\text{W}/\text{m}^2\mu\text{m}$]
N	Leaf mesophyll structure parameter	[dimensionless]
LAI	Leaf Area Index	[m^2/m^2]
ALA	Average Leaf Angle	[$^\circ$],[rad]
HS	Hot spot parameter	[dimensionless]
BS	Soil brightness coefficient	[dimensionless]
$SKYL$	Fraction of diffuse irradiation compared to global irradiation	[dimensionless]
μ_i	Average of variable i	
σ_i	Standard deviation of variable i	
R^2	Pearson's correlation coefficient	
COV_{rad}	Covariance matrix of radiometric information (bands, direction)	
COV_{var}	Covariance matrix of canopy variables	
V_{est}	Estimated variable	

Chapter 1

Introduction

1.1 Earth observation in a changing environment

In February 2007, the Intergovernmental Panel on Climate Change reported that if human induced emission atmospheric greenhouse gases is not tempered, an ongoing increase in global warming is inevitable, having dramatic consequences for our planet (IPCC, 2007). In this light, increased effort is put in studying the role of the biosphere in terrestrial carbon balance (GCP, 2003), and in monitoring and modeling the effects of changing climatological conditions on ecosystems and biodiversity (DIVERSITAS, 2002). Such research is often conducted in support of policy at national and international level, issued to meet the criteria posed by political treaties such as The Kyoto Protocol of the United Nations Framework Convention on Climate Change, the Convention on Biological Diversity signed at the United Nations Conference on Environment and Development, or the UN Millennium Development Goals.

Earth observation plays a crucial role in such assessments, being able to provide information on the state of the vegetation cover over extended geographical areas, at a high temporal frequency, and at relatively low cost. This potential has also been recognized by various intergovernmental institutions, as evidenced by the large number of programs adopting remote sensing as one of the key elements in global environmental monitoring. In 2005, the Group on Earth Observations adopted a 10 year implementation plan of an integrated Global Earth Observation System of Systems (GEOSS), established to organize the efforts of different nations on the international level for providing spatially and temporally consistent observations (GEO, 2005). The Global Monitoring for Environment and Security (GMES) initiative, is the European participation in this plan, and is a joint effort of European Union and the European Space Agency (ESA) to provide the data and operational services required for global environmental monitoring and for a wide range of security related issues (Brachet, 2004).

Environmental and security issues adopted in GMES and similar non-European initiatives also involve food related risks, such as the effects of drought and unsustainable land use on crop yield and food security. Several programs concentrate on developing remote sensing based systems for monitoring and forecasting crop prospects in different parts of the world (e.g., MARS-FOOD¹ and GMFS² of the EU and CADRE of the Production Estimates and Crop Assessment Division of the USDA Foreign Agricultural Service³). Whereas these programs mainly intend

to support political decision making at regional to global scales, at a more local scale remote sensing products can assist farmers in adapting their crop management strategies. By applying manure, pesticides, and water supplies according to the needs of the individual plants, yields can be optimized, costs reduced, and impacts on the fragile and ever scarcer available natural resources brought down to a minimum (Delécolle et al., 1992; Inoue, 2003). Such management decisions are often taken at field or intra-field scale thus requiring information products with high spatial resolution.

The vegetation properties provided by earth observation systems rarely directly support decision making and mostly have to be interpreted in a broader environmental context. Sound decision making can only be guaranteed when a complete overview of the environmental processes responsible for ecosystem behavior is provided. In this context, remote sensing products are often employed in conjunction with information on soil properties, plant phenology and uptake behavior, and meteorologic and environmental conditions, to feed or calibrate process models providing the environmental variables that can be interpreted by users and policy makers and help them in their decision-making processes (Bach and Mauser, 2003; Delécolle et al., 1992; Houser et al., 1998; Makowski et al., 2003; Moulin et al., 1998; Oliso et al., 1999; Schaepman, 2007; Verhoef and Bach, 2003a).

1.2 User requirements on remote sensing biogeophysical products

The remote sensing platforms supplying earth observation data have strongly differing properties, including active and passive systems, with spatial resolutions ranging from less than a meter to several kilometers, and visit frequencies varying from one single data take to several measurements every hour. The choice of system, its spatial and spectral resolution, and the temporal frequency of the observations is driven by the required product and envisaged application. For example, ecosystem and terrestrial carbon studies require data with a continental to global coverage at moderate spatial and temporal resolutions, while long term continuity, needed to detect gradual changes from the generally accepted steady-state conditions, should be guaranteed (Van Leeuwen et al., 2003). On the other hand, local scale agricultural applications require data with high spatial resolutions at several critical time steps during the annual growth cycle of a crop (Delécolle et al., 1992; Launay and Guerif, 2005).

Data consistency plays a critical role when long term ecological effects are studied, given the usually relatively limited life time of most sensor missions (Van Leeuwen et al., 2003). But also for intra-annual applications such as precision farming, several independent providers might contemporaneously disseminate basically the same product derived from different sensors and based on different algorithms, making it difficult for a user to decide whether the products are interchangeable and if they can be used to fill gaps in their observation sequences (Verhoef, 2007). Therefore, if thematic remote sensing based products of the biosphere are to be employed by a wide range of users and decision makers, the remote sensing community should put increased effort in developing consistent algorithms and data products.

The success of remote sensing based vegetation products in agricultural applications (Moran

¹URL: <http://agrifish.jrc.it/marsfood/Default.htm>

²URL: <http://www.gmfs.info/>

³URL: <http://www.pecad.fas.usda.gov/cropexplorer/index.cfm>

Table 1.1: Variables indicative for canopy condition and functioning. The third and fourth columns give some examples of studies where these variables have been successfully retrieved from different types of remote sensing data.

<i>Biophysical variable</i>	<i>Main indicator</i>	<i>Sensor type</i>	<i>Remote sensing application</i>
Fraction of absorbed photosynthetically active radiation (fAPAR)	Photosynthesis, plant stress	Reflective	Clevers (1997); Gobron et al. (2006)
Leaf Area Index (LAI)	Plant development, plant stress	Reflective	Doraiswamy et al. (2004); Mo et al. (2005); Moulin et al.,(2003)
Fractional cover (fCover)	Plant development, land degradation	Reflective	Bouman (1995); Bachmann et al. (2004)
Chlorophyll and other pigments	Photosynthesis, plant stress	Reflective	Haboudane et al. (2002); Zhao et al. (2004)
Mineral content (K, P, Ca, Mg)	Crop quality, soil degradation	Reflective	Mutanga et al. (2004)
Plant water content	Drought stress	Reflective, Radar	Moran et al. (1994); Zarco-Tejada et al. (2003)
Above ground biomass / Net Primary Production	Carbon storage; (crop) yield, timber volume	Reflective, Radar, LIDAR	Tucker et al. (1983)
Evapotranspiration	Drought stress	Emissive	Bastiaanssen and Ali,(2003); Hurtado et al. (1994)
Vegetation height, stem density	Plant development, plant stress	Radar, LIDAR, high resolution optical stereo	Hollaus et al. (2006); Richardson et al. (1982)

et al., 1997) and security related issues, such as wild fire propagation (Koltunov and Ustin, 2007; Leblon, 2005) are additionally governed by the need of nearly real-time data delivery, requiring automated preprocessing facilities and thematic algorithms that provide on demand self-consistent products without intervenience from producer or user side (Brazile et al., 2005; De Vries et al., 2007; Habermeyer et al., 2005; Knyazikhin et al., 1999b; Richter et al., 2006b).

1.3 The challenge of providing robust environmental indicators

The way vegetation reacts to changing environmental and climatological conditions is often reflected by a direct or gradual change of its biophysical and biochemical properties. Table 1.1 lists the state variables that are most illustrative for the condition of the canopy, together with the phenological, environmental, or climatological processes which they are indicative for. However, different types of stress (e.g., drought, insects, soil contamination) often manifest themselves in a very similar way in plant physiological response, and not seldom the reaction consists in a concurrent modification of more than one state variable (Fridgen and Varco, 2004; Vyn and Hooker, 2002).

Remote sensing observations do not directly sample the canopy variables indicative for vege-

tation state but provide spatially distributed radiation fluxes reflected or emitted by the surface. The measured radiation fluxes are driven by radiative transfer processes, such as scattering, absorption and emission, intrinsically related to the properties of the observed surface (Koetz et al., 2004). However, the variables controlling radiative transfer and, hence, the signal captured by the earth observation system, are often not directly related to the surface properties of ultimate interest (Verstraete et al., 1996). Therefore, for establishing consistent relationships between earth observation signal and variable of interest, other variables controlling the signal have to be taken into account as well.

Classical empirical approaches calibrated over experimental data sets (Huete, 1988; Richardson and Wiegand, 1977) implicitly assume a direct relationship between measured signal and variable of interest and thus neglect spectral variations induced by changing canopy architecture, background, or observation geometry. As a consequence, these methods lack robust portability and relationships have to be calibrated for every new situation (Baret and Guyot, 1991). Instead, inversion of canopy radiative transfer models appears a promising alternative to the empirical approaches (Bacour et al., 2002a; Bacour et al., 2006; Houborg et al., 2007; Jacquemoud et al., 1995a; Knyazikhin et al., 1999b; Koetz et al., 2004; Weiss et al., 2002). Radiative transfer models summarize our knowledge on the physical processes involved in the photon transport within vegetation canopies, and simulate the radiation field reflected or emitted by the surface for given observational configuration, once the vegetation, the background, and possibly the atmosphere are specified (Baret and Buis, 2007; Goel and Thompson, 2000; Verhoef and Bach, 2007).

Comprehensive modeling of radiation propagation in canopies with a radiative transfer model requires a considerable number of inputs, since absorption and scattering processes at various levels in the canopy are defined by a large number of biochemical (e.g., leaf chlorophyll content, leaf and stem water content) and structural (e.g., leaf density, leaf inclination, leaf size) vegetation properties (Gobron et al., 1997; Kuusk, 1995a; Verhoef, 1984). This number will even increase when reflectance of structurally complex canopies such as forest stands, orchards, or transition communities have to be mimicked (Govaerts and Verstraete, 1998; Verhoef and Bach, 2007). Given the limited number of independent data dimensions in most remote sensing systems (Fourty and Baret, 1997; Gemmell, 2000; Verhoef, 2007; Zhang et al., 2002a; Zhang et al., 2002b), estimating vegetation properties from earth observation data by radiative transfer model inversion or by any other retrieval method, is an under-determined problem (Combal et al., 2002b; Kimes et al., 2000). In addition, several variable combinations show complementary behavior and may lead to very similar spectral signatures, a phenomenon that is reinforced by radiometric, georeferencing, and model errors. Radiative transfer model inversion therefore has a large number of multiple solutions, a phenomenon also known as ill-posedness (Atzberger, 2004; Combal et al., 2002b; Tarantola, 2005).

A robust retrieval of vegetation variables can therefore only be guaranteed if additional assumptions, constraints, or additional independent observations are introduced (Bacour et al., 2006; Combal et al., 2002b; Vohland and Jarmer, 2007; Weiss et al., 2001). Even if assumptions and constraints reduce the dimensionality of the inverse problem, they also limit the transferability of the approach, since they are generally only applicable for a specific situation (Koetz, 2006). For this reason, the science community increasingly focuses on expanding the number of independent observations of a target, either by increased spectral resolution of single sensors (Schlerf et al., 2005; Verhoef, 2007), by assimilation of additional independent data sources of different nature, such as multi-directional and multi-temporal observations (Chopping

et al., 2003; Knyazikhin et al., 1999c; Pinty et al., 2002; Bacour et al., 2002b; Koetz et al., 2005a), or by a combination of both (Koetz et al., 2007; Menenti et al., 2005). The degree to which additional observation dimensions contribute to improved retrieval performance strongly depends on the nature of the complementary information and its collinearity with the information already present. Whereas increased spectral resolution primarily leads to improved determination of foliage biochemistry and species composition (Ustin et al., 2004; Verhoef, 2007), multi-angular observations, laser scanning, and microwave remote sensing principally contribute to enhanced estimation of canopy structure elements (Hyypä et al., 2000; Koetz et al., 2007; Widlowski et al., 2004).

1.4 Rationale of the thesis

Recently, scientists have demonstrated increased activity in developing operational algorithms for the retrieval of biogeochemical variables from solar-reflective remote sensing data based on radiative transfer model inversion. Examples are the leaf area (*LAI*) and vegetation cover (*fCover*) products from POLDER-2 onboard ADEOS-2 (Lacaze, 2005), *LAI* and the fraction of absorbed photosynthetically active radiation (*fAPAR*) from MODIS/MISR on the Terra and Aqua platforms, *LAI*, *fAPAR*, *fCover*, and canopy chlorophyll content from MERIS on ENVISAT (Bacour et al., 2006), and the *LAI*, *fAPAR*, and *fCover* products from VEGETATION observations (Baret et al., 2007; Weiss et al., 2007). These algorithms have all been optimized for the respective sensor while the spatial resolution of the products ranges from 300 meter for MERIS over 1 kilometer for MODIS and VEGETATION up to 6×7 kilometers for POLDER. On the contrary, agricultural applications, such as precision farming, require information products at field or intra-field levels, typically in the range of 1 - 30 meters.

Several commercial companies⁴ already offer precision farming products at high to very high resolutions based on multispectral sensor systems and automated processing chains. But, due to the limited number of bands inherent to such systems, the product palettes offered by these companies are based on underdetermined model problems and therefore mainly include biophysical products that have a more or less straightforward relationship with canopy reflectance, such as yield or biomass. On the other hand, crop stress and nutrition deficiency is often expressed in changes of biochemical constituents, such as leaf chlorophyll and leaf water content. Designing consistent automated algorithms for detecting such changes is often hindered by the low spectral sensitivity of most leaf biochemical variables and by their high correlation with *LAI* and other structural variables (Bacour et al., 2006; Vohland and Jarmer, 2007). Moreover, the spectral changes induced by these variables are not captured by every multispectral system. Whereas the presented medium resolution sensors can fall back on more frequent coverage, thus facilitating land cover classifications in support of retrievals optimized for specific biomes (Houborg et al., 2007; Lotsch et al., 2003), a similar option is not offered to the high resolution systems. A stable inversion thus requires several assumptions and constraints which will obviously limit the transferability of the developed inversion schemes to a wide range of different land cover types. Additional data dimensions are therefore required to guarantee a stable model inversion for a wide range of crops based on high resolution data.

At the advent of upcoming European initiatives providing superspectral and hyperspectral data at high resolutions, such as Sentinel-2 (Gascon and Berger, 2007), EnMAP (Müller

et al., 2004) and VENUS⁵, a solid understanding of the influence of increased spectral information content on radiative transfer model inversion is of growing interest. Theoretically, due to the increased number of independent data dimensions, such sensors should provide increased retrieval performance compared to the common mono-directional multispectral sensors (Verhoef, 2007). This would allow for enhanced accuracy of automatically generated products as well. Until launch into orbit of the spaceborne sensors mentioned above, several high resolution technology demonstrators and operational precursor missions serve the research community with hyperspectral data (Schaepman, 2007). Examples of available sensors are the airborne instruments AVIRIS (Green et al., 1998), HyMap (Cocks et al., 1998), and the currently developed APEX (Schaepman et al., 2004) and ARES (Müller et al., 2005; Richter et al., 2005), and the spaceborne missions of CHRIS/PROBA (Barnsley and Settle, 2004) and Hyperion/EO-1 (Pearlman et al., 2003). These instruments allow us to explore the potential of hyperspectral data content for the characterization of selected targets and at user-defined spatial and observational configurations (Beisl, 2001).

Even if hyperspectral imagery would facilitate stable radiative transfer model inversion based on single mono-directional data takes, the temporal restriction of most data campaigns (airborne sensors) and the cloud cover frequently encountered at mid-latitudes, still lead to intermittent data series. Complementing data products from different sensors are therefore often necessary to obtain data series that span the entire growth cycle.

In prospect of the forthcoming EnMap mission, the German Remote Sensing Center of the German Aerospace Center (DLR-DFD) established an automated processing chain for the calibration, georectification, and atmospheric correction of imaging spectrometer data (Bachmann et al., 2007; Habermeyer et al., 2005) (Figure 1.1). Based on the level-2 surface reflectance cube, a series of value added level-3 products has been developed, such as a spectral land cover classifier (Richter, 2007a) and an algorithm for fractional cover estimates of photosynthetic and non-photosynthetic vegetation (Bachmann et al., 2004). In line with these level-3 products, this thesis explores the potential of hyperspectral data for an automated retrieval of canopy biophysical and biochemical variables for agricultural purposes using radiative transfer model inversion. The approach is driven by the prerequisite that it should be completely image based, using only the information available from preceding processing steps and is not bounded to a single imaging spectrometer.

1.5 Research objectives

There is a clear gap between the fully operative processing chains providing biogeochemical products at medium to low resolution, and the exigence from side of the farmers requiring consistent biochemical and biophysical products in support of their farming management strategies at a local to regional scale. Moreover, the irregular data availability from single sensors calls for the combination of data products originating from different sensors in order to obtain data sequences covering the entire growth season. This requires flexible algorithms that provide consistent results among a variety of sensors. This thesis aims to bridge this gap by completing the

⁴e.g., Spot Image (http://www.spotimage.fr/web/en/640-precision_farming.php), Basfood (<http://www.basfood.nl/>)

⁵<http://smc.cnes.fr/VENUS/>; visited July 2007

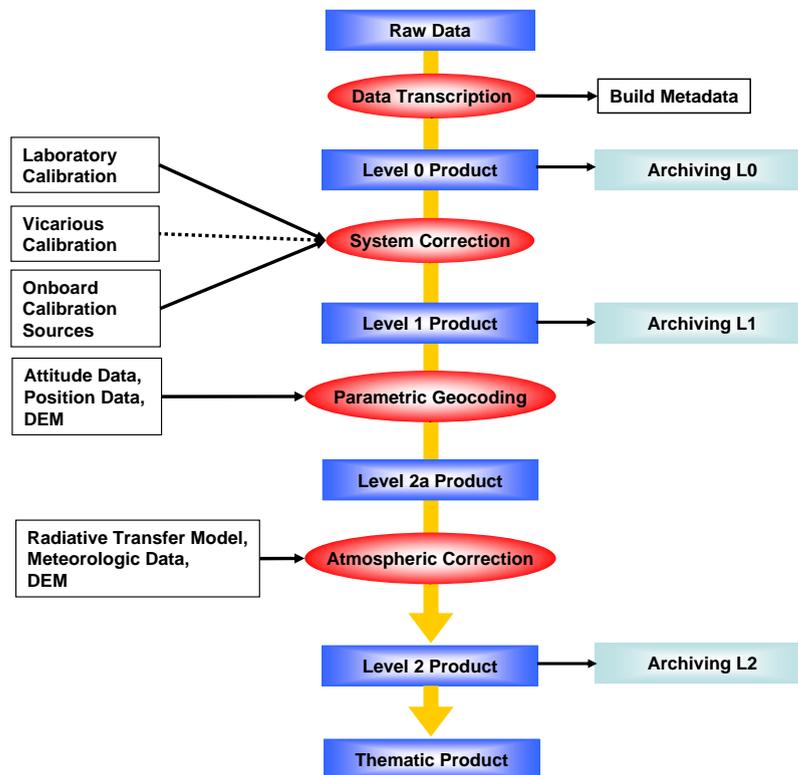


Figure 1.1: Processing chain for hyperspectral data operated by DLR-DFD (Adapted from Bachmann et al. (2007))

following main objective:

- To develop a consistent, completely automated, and image based radiative transfer model inversion approach for the concurrent retrieval of the key vegetation variables **leaf area index**, **leaf chlorophyll content**, **leaf water content**, and **leaf dry matter content** from mono-temporal data recorded with any high resolution airborne or spaceborne imaging spectrometer.

In order to be fully automated, the inversion approach should satisfy the following criteria:

- It completely relies on the top-of-canopy reflectance data and the intermediate products provided by the automated procession chain for imaging spectrometer data available at DLR-DFD.
- It is assumed that no information on land cover is a priori available.
- It should be applicable to a wide variety of field crops and grasslands in different phenological stages.

- The algorithm should be computationally fast, since the complete chain - from data acquisition over (pre-)processing up to data product delivery - should not take more than a couple of days in order to guarantee timely decision support to the user.

In fact, the ultimate goal is not the development of the automated approach as such, but rather the challenge of providing consistent results over a wide variety of sensors, observation/illumination conditions, and agricultural crops. In order to fulfill these requirements, pertinent answers have to be found to the following research questions:

- Is the radiative transfer model underlying the approach well able to mimic top-of-canopy reflectance for a wide range of crops and view/sun geometries?
- Is canopy reflectance sensitive enough to the variables of ultimate interest?
- To what degree are the results dependent on waveband configuration of the sensor?
- Are the results reproducible for a single observation, i.e., does the algorithm provide similar results in different model runs?
- How sensitive is the approach to changing observation and/or illumination conditions and do certain constellations perform better than others?
- How well does the automated approach regularize the ill-posed model inversion, and how good is this performance compared to situations in which there is a significantly higher amount of a priori information?
- Can additional stability and accuracy of the automated approach be provided by inclusion of multi-angular observations?

In the framework of this study, there has been an explicit choice for imaging spectrometers. This choice is motivated by the strongly underdetermined nature of radiative transfer model inversion based on multi-spectral data, especially in absence of thorough a priori knowledge on land cover. The hypothesis underlying this study is therefore, that the increased information content of hyperspectral observations compared to multispectral observations, especially on biochemical leaf constituents (Fourty and Baret, 1997; Verhoef, 2007), has a positive effect on the stability and accuracy of radiative transfer model inversion.

In support of the previously mentioned research activities, a prototype software package called CRASH (Canopy variable Retrieval Algorithm based on PROSPECT and SAILh) has been developed. The software tool was programmed in the Interactive Data Language (IDL; ITT-VIS, Inc.) as a stand-alone widget module in ENVI. This compatibility with the most widely used hyperspectral image processing software and data formats in the remote sensing community should ensure a continued lifetime and further development after conclusion of this thesis.

1.6 Structure of the thesis

Chapter 2 gives a comprehensive state-of-the art overview of the methods used to provide quantitative characterizations of vegetative surfaces. It discusses the benefits and limitations of the different techniques in relation to the underlying physical background of radiation propagation in leaves and canopies. The main part of the chapter will focus on radiative transfer modeling, the problems encountered in radiative transfer model inversion, and the solutions proposed by other authors in order to reduce ill-posed solutions.

Chapter 3 provides the theoretical considerations and the implementational aspects of the automated CRASh approach. These include the choice of the radiative transfer model, an automated land cover classification, and the inversion approach itself. The inversion approach incorporates a new concept for additional regularization based on the spectral content of the image.

In Chapters 4 and 5, the automated approach is validated at three spatial levels: at ground level based on field spectrometer measurements, at airborne level using data from a HyMap sensor, and at satellite level using data from the CHRIS/PROBA multi-angular imaging spectrometer. Chapter 4 tests the performance of the retrieval algorithm at the first two levels against the background of estimating leaf area index, leaf water content, and leaf dry matter content for mid-latitude agricultural grasslands. Based on the field spectrometer measurements, the impact of the single algorithm components is tested, as well as the reproducibility of the results and the sensitivity of the approach to model parametrization. The HyMap data allows for a close inspection of model sensitivity to changing illumination/observation geometry and to the automated land cover classification. Chapter 5 validates the automated approach in the light of leaf chlorophyll and *LAI* estimations over irrigated cotton. First, the performance of the automated algorithm is tested on the five single view angle data sets of the multi-angular CHRIS/PROBA sensor. Subsequently, it is tested whether the concurrent inversion of multiple view angles can further improve estimation accuracy.

Chapter 6 concludes this thesis, synthesizing the findings of the preceding chapters and drawing conclusions in the context of the objective and research questions posed in the first chapter. Finally, an outlook is given and possible focal points of future research in the field of the presented thesis are spelled out.

Chapter 2

Characterizing vegetation canopies by solar-reflective remote sensing*

Remote sensing of vegetation is based on the measurement of spatially distributed electromagnetic radiation fluxes reflected or emitted by the canopy. The observed signal is the combination of scattering, absorption, and emission processes that take place in the atmosphere and on the surface materials found within the sampling unit of the sensor and is therefore rarely directly related to the surface properties of ultimate interest (Verstraete and Pinty, 1996). The challenge is to establish consistent relationships between measured signal and variable of interest while minimizing the influence of other factors interacting with radiation. To be able to establish such relationships, a thorough knowledge of the mechanisms determining the variations in the observed signal is indispensable. Therefore, this chapter first starts with a brief overview of solar-reflective remote sensing, and the different surface properties interacting with the radiance signal. It is followed by an overview of the different methods available for retrieving the vegetation variables of interest, with a special focus on physical approaches based on radiative transfer model inversion.

2.1 Principles of solar-reflective remote sensing

2.1.1 Electromagnetic radiation

The physical quantity that is measured by a remote sensing system is electromagnetic radiation. Electromagnetic radiation is a dynamic form of energy that is capable of propagating through a vacuum and becoming apparent only by its interaction with matter (Suits, 1983). In quantum theory, electromagnetic radiation is described as a stream of discrete particles (photons) carrying fixed amounts of energy (Hunt, 1983). Such packages of energy are emitted when a molecule

*This chapter is partly based on:

Dorigo, W. A., Zurita-Milla, R., de Wit, A. J. W., Brazile, J., Singh, R. and Schaepman, M. E. (2007). A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling, *International Journal of Applied Earth Observation and Geoinformation* **9**: 165–193.

or atom falls from an excited energy state to a lower one. Similarly, electromagnetic radiation is absorbed if the provided energy matches the amount that is required to promote an atom or molecule from one energy state to a higher one.

The energy E carried by a photon is associated to the wave frequency ν by:

$$E = h \cdot \nu \quad (2.1)$$

where h is Planck's constant ($6.626 \cdot 10^{-34} J \cdot s$) and ν is the frequency, which in turn is inversely related to wavelength λ by:

$$\nu = c/\lambda \quad (2.2)$$

with c being the speed of light ($3 \cdot 10^8 m \cdot s^{-1}$). Combining Equation 2.1 and 2.2 reveals that the energy content of electromagnetic radiance is inversely proportional to wavelength. Both wavelength (optical remote sensing) and frequency (radar) are used to categorize radiation type in remote sensing.

If a body (e.g. Sun or Earth) is assumed a blackbody radiator (i.e., it emits radiance at the maximum efficiency possible for a body at its effective temperature T (Schowengerdt, 1997)), the wavelength λ of maximum emission is given by Wien's displacement's law:

$$\lambda_{max} = 2898/T \quad (2.3)$$

where λ is expressed in μm and T in degrees Kelvin. Using this law, the wavelength of maximum radiant exitance lies at $0.49 \mu m$ for the Sun ($T = 5900$ K) and around $9.66 \mu m$ for the Earth ($T = 300$ K). The wavelength at which solar irradiation measured at the Earth is matched by self-emitted thermal energy ranges between 2.5 and $6 \mu m$ (Slater, 1980).

The total amount of radiation E incident on a surface is called *irradiance* and expressed as the flux density (Ψ) per unit area (A):

$$E = d\Psi/dA, Wm^{-2} \quad (2.4)$$

Typically used irradiance quantities in remote sensing are E_{sun} to denote direct solar irradiance and E_{sky} for diffuse hemispheric irradiance. In the context of spectral measurements, irradiance is integrated over a specific wavelength interval and the SI-unit is extended with $[\mu m^{-1}]$. Introducing the direction of the originating irradiance component δ (e.g., incident on a vegetation surface or on a detector element) gives the *radiance* L which is the flux density per unit projected area and per unit solid angle Ω :

$$L = E/\pi = \frac{d^2\Psi}{dA \cos\delta d\Omega}, Wm^{-2}sr^{-1} \quad (2.5)$$

In solar-reflective remote sensing usually the quantity *reflectance* ρ_λ is used to describe the spectral properties of a surface, which is given by the ratio between reflected $L \uparrow$ and incident radiance $L \downarrow$ measured at the bottom of the atmosphere:

$$\rho_\lambda = \pi \cdot L/E \quad (2.6)$$

Measured reflectance strongly depends on the ensemble of illumination and observation direction and on the properties of the considered surface (Section 2.1.5 and Section 2.2, respectively).

2.1.2 The solar-reflective domain

Passive remote sensing systems operate in the optical domain which stretches from around 0.4 to 14 μm . This wavelength range can be roughly divided in a reflective and an emissive part. In the part of the optical domain with wavelengths greater than 5 μm self-emitted radiation from the Earth surface generally dominates. This part is also called the thermal infrared (TIR) domain. In the range up to 2500 nm the signal observed by the sensor is dominated by reflected solar radiance. This (solar-)reflective domain can be subdivided in the visible (VIS; 0.4-0.7 μm), near-infrared (NIR; 0.7-1.3 μm) and shortwave infrared (SWIR; 1.3-2.5 μm , sometimes also referred to as mid-infrared (MIR)). In the VIS, a subdivision can be made between blue, green, and red light (Table 2.1) whereas the SWIR is sometimes separated in two regions, delineated by the major atmospheric water absorption band around 1.9 μm (Section 2.1.3).

In the transition zone between SWIR and TIR, i.e. between 2.5 and 5 μm , the observed signal is constituted of a mixture of reflected and emitted radiance having similar magnitudes. Since the respective contributions are difficult to differentiate, this part of the spectrum is rarely used in vegetation studies, except for detecting (forest) fires (Giglio et al., 2003; Justice et al., 2002). In the following paragraphs the discussion is confined to the solar-reflective domain.

Table 2.1: Categorization of wavelength ranges in the solar-reflective domain.

<i>Spectral domain</i>	<i>Wavelength range [μm]</i>
Visible (VIS)	0.4 - 0.7
- Blue	0.4 - 0.5
- Green	0.5 - 0.6
- Red	0.6 - 0.7
Near Infrared (NIR)	0.7 - 1.3
Shortwave Infrared (SWIR)	1.3 - 2.5
- SWIR 1	1.3 - 1.9
- SWIR 2	1.9 - 2.5

2.1.3 Atmospheric interaction

Incoming or reflected radiance is subject to absorption by atmospheric gases (Richter and Schlaepfer, 2002; Berk et al., 2005). Figure 2.1(a) shows that different elements are respon-

sible for this absorption and that their influence is dependent on wavelength. The principle contributors to atmospheric absorption are i) *ozone* (O_3), absorbing part of the radiance in the red and the ultraviolet, ii) *oxygen* (O_2), responsible for the increased absorption around 760 nm, and iii) carbon dioxide (CO_2) and water vapor (H_2O), who are the principle contributors to absorption in the reflective domain and show major features around 1.4, 1.9 and 2.5 μm and some less pronounced ones at 0.9 and 1.1 μm . Figure 2.1(a) shows that absorption is highly variable across the reflective part of the spectrum. The spectral regions with relatively high transmittance are called “atmospheric windows”. It is usually in these windows that the spectral bands are positioned for the study of land surface properties and vegetation characteristics in particular.

Apart from absorption, the solar radiance flux incident on the Earth surface and finally reaching the sensor is also influenced by atmospheric scattering. Molecular or *Rayleigh* scattering is caused by particles whose size is far less than the wavelength of the radiation component. The amount of this type of scattering is inverse proportional to the fourth power of the wavelength. As a result, shorter wavelengths, such as the blue, are affected more than others. Rayleigh scattering acts symmetrically in the forward and the backward direction and to a lesser extent in perpendicular directions (Liang, 2004).

Mie scattering is caused by particulates with a size similar to the wavelength of the incoming radiation (e.g. smoke, smog, fog, dust, and haze (Liang, 2004)). For large particles Mie scattering is wavelength independent, for smaller ones it is inverse proportional to wavelength, thus increasingly affecting the shorter wavelengths. It acts mainly in the forward specular direction. Mie scattering evolves particularly in the lower atmospheric layers where the concentration of the larger, heavier particles is highest.

A third type of scattering, called non-selective scattering, is induced by particles that are significantly larger than the wavelength (e.g. rain drops, snow, large dust particles). Scattering of this type affects all wavelengths equally and shows isotrope behavior. The total amount of scattered, or diffuse, radiance for a standard atmosphere is around 30% of total radiance in the blue wavelengths and asymptotically reaches a level of 1-2% in the SWIR (Berk et al., 2005).

The combined action of absorption and scattering is related to the distance the solar radiation has to travel through the atmosphere before reaching the Earth surface or sensor. It therefore depends on solar elevation, land surface altitude, and view direction of the sensor. A term often used in this context is the optical thickness which is the product of the atmospheric extinction coefficient and the path length (Richter, 2007a; Berk et al., 2005). It actually represents the total of optical thicknesses of the individual contributors (molecular scattering, aerosol scattering, and molecular absorption).

Figure 2.1(b) gives an overview of the combined action of atmospheric components on the upwelling radiance measured at a detector element of a high-altitude or satellite sensor. Four major fluxes can be distinguished:

1. Path radiance, i.e., photons scattered in the air volume between ground and sensor and reaching the earth observation sensor without prior ground contact.
2. Direct solar radiation incident on a pixel and reflected and transmitted into the instantaneous field of view (IFOV)¹ of the sensor. In terms of reflectance, this contribution is

¹IFOV is the angle subtended by a single detector element on the axis of the optical system

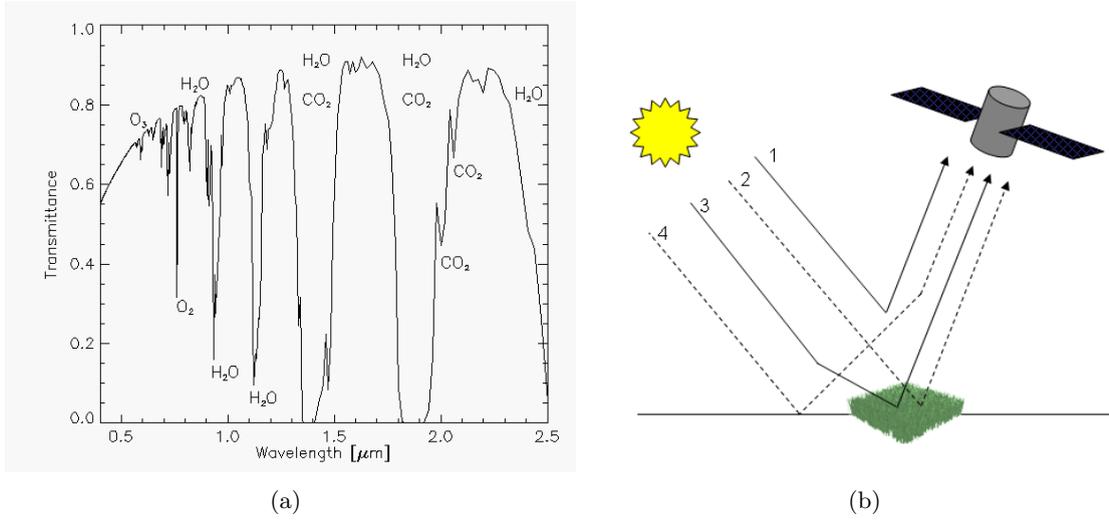


Figure 2.1: (a) Atmospheric transmittance in the optical domain (Richter, 2007a), (b) schematic sketch of radiance components measured at a detector element of a high altitude sensor (b) (Richter et al., 2006a).

denoted as the bidirectional reflectance factor (BRF). Strictly speaking, since the directional contribution assumes the size of the radiance source and sensor infinitely small, it would be more correct to use the term biconical reflectance factor instead (Martonchik et al., 2000; Nicodemus et al., 1977; Schaepman-Strub et al., 2006). However, in most literature the term BRF is still commonly used.

3. Diffuse solar radiation incident on a pixel and reflected into the IFOV of the sensor. The sum of direct and diffuse flux incident on the ground surface is called global flux. Translated into reflectance terminology, the term hemispherical directional reflectance factor (HDRF) is used for component 3. Like in the case of BRF, it would be more appropriate to use the term hemispherical conical reflectance factor (HCRF) instead. In common conceptual terminology, HDRF is used for the combination of directional and hemispherical incoming radiation reflected in the direction of the sensor and therefore equal to the combination of reflectance components 2 and 3 (Martonchik et al., 2000). Schaepman-Strub et al. (2006) already stretched this confounding terminology and the importance of standardizing reflectance nomenclature in order to exclude uncertainties resulting from inappropriate interpretation of the spectral quantities.
4. Adjacency radiance: reflected radiation from the adjacent areas scattered by the air volume into the IFOV of the sensor.

In mathematical form, and expressing the combination of radiance components 2 and 3 as $L_{reflected}$, the total radiance signal measured at the sensor can be written as:

$$L_{sensor} = L_{path} + L_{reflected} + L_{adjacency} \quad (2.7)$$

Only the second radiation component in Equation 2.7 contains information from the currently viewed pixel. Removal of components 1 and 3, and the retrieval of the ground reflectance component 2 can be obtained by atmospheric correction (Richter, 2007a).

2.1.4 System requirements

The properties of spectral observations in the reflective domain mainly depends on the type of sensor that is used. The choice for a particular sensor depends on the application and is always a trade-off between the required spectral resolution, the spatial resolution, and the revisit frequency. Precision agriculture applications require sensors with high spatial resolutions and, if possible, repeated visits throughout the vegetation growth cycle. Besides, spectral coverage should be fine enough to capture all changes in spectral shape that are indicative for the status of the canopy elements the user is interested in (Figure 2.2). Fourty et al. (1996) report that the information content in the NIR and SWIR part of canopy reflectance can be explained by at least 10 selected spectral bands. Including a band in the blue, green, and red, this would lead to a minimum of 13 spectral bands explaining all spectral variation.

Bearing in mind these requirements, hyperspectral remote sensing systems with a high spatial resolution and several evenly distributed revisits throughout the growth season would be ideal (Delécolle et al., 1992; Launay and Guerif, 2005). Hyperspectral sensors, or imaging spectrometers, are systems that may add up to several hundreds of narrow contiguous bands throughout the spectrum thus enabling to detect typical absorption features that would be impossible to detect with many of the commonly used multi-spectral sensors (e.g. Landsat Thematic Mapper or SPOT; Figure 2.2). The only operative full range satellite imaging spectrometer (HYPERION) does not meet the criteria posed by the spatial and temporal domain. However, airborne sensors such as AVIRIS, ARES, and HyMap can partly bridge this gap by providing data with high spatial resolution. Nevertheless, for airborne campaigns the temporal coverage is often not guaranteed.

2.1.5 Observation properties

Before continuing the discourse on canopy reflectance, a few common definitions concerning the view/sun geometry are introduced which are essential for a correct understanding of the work presented in this thesis (Figure 2.3):

- The *zenith angle* of sun (θ_s) and observer (θ_v) are counted relative to nadir², i.e., a zenith angle of 0° is identical to nadir, while the horizon has a $\theta_{s/v}$ of 90° ,
- The *relative azimuth* (ψ_{rel}) is the absolute difference between the solar (ψ_s) and the observation (ψ_v) azimuth and ranges from 0 - 180° . In this work, a ψ_{rel} of 0° means that the observer has the sun in his back, whereas at a ψ_{rel} of 180° the observer faces the sun,
- The *principal plane* is the azimuthal plane parallel to the incoming radiation. The azimuthal plane perpendicular to the direction of the incoming radiation, is called the *orthogonal plane*.

²Nadir is considered relative to a horizontal plane

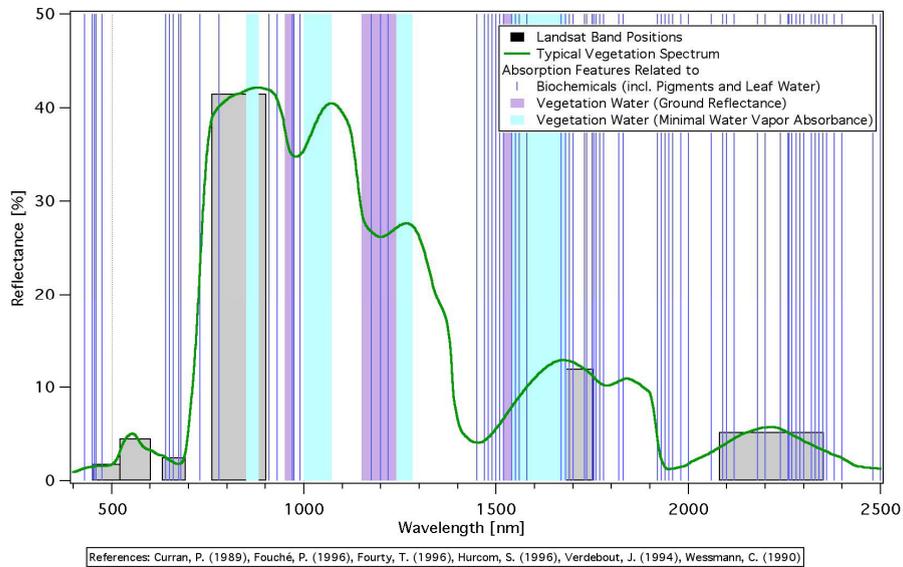


Figure 2.2: Hyperspectral (ASD fieldspec PRO FR) versus multispectral (Landsat 7 ETM+) observation of a vegetation canopy in relation to spectral absorption features of leaf biochemicals identified from literature (Schaeppman, 2007).

2.2 Canopy reflectance

Canopy reflectance can be considered the combined action of scattering and absorption at phytoelements and soil background and depends on the optical properties of the single elements and the arrangement of these elements in a spatial context. In this section, optical and structural properties of vegetation and background optical properties will be discussed in consecutive order.

2.2.1 Leaf optical properties

The spectral properties of vegetation elements such as leaves, needles, twigs, stems, and fruits, can be considered the major determinant of canopy reflectance and influence the shape of the overall spectrum (Jacquemoud and Baret, 1990). Schematically, a leaf can be considered a pile of several distinctive cell layers (Figure 2.4). Based on this mesophyll structure, often a distinction is made between monocotyledon leaves, characterized by a homogeneous parenchyma tissue with few intercellular airspaces, and dicotyledon leaves, where the mesophyll is characteristically differentiated into a densely packed palisade and a loosely packed spongy tissue (Govaerts et al., 1996; Baranoski, 1998).

Radiation incident on a leaf is either scattered (including both reflection and transmission) or absorbed by leaf constituents. According to the dominant processes, three distinctive spectral domains can be differentiated (Figure 2.5).

- The **visible** (VIS) domain is characterized by strong photon absorption caused by leaf pigments. These pigments (e.g., chlorophyll a and b, carotenoids, and brown pigments) are found within the chloroplasts in the mesophyll layers. In green vegetation, absorbance

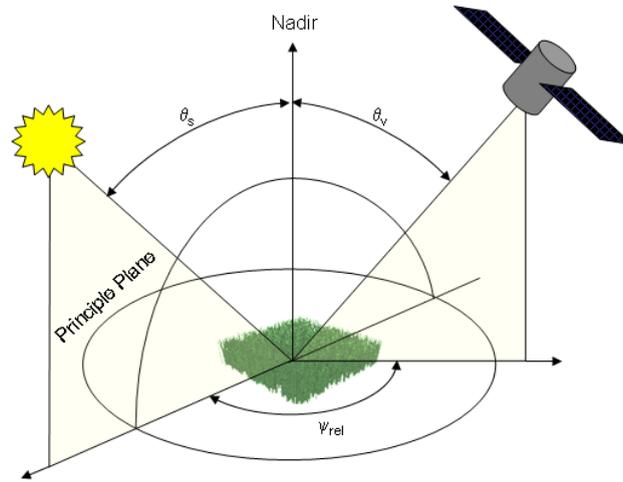


Figure 2.3: Measurement geometry in solar-reflective remote sensing. θ_s = sun zenith angle, θ_v = view zenith angle, ψ_{rel} = relative azimuth angle

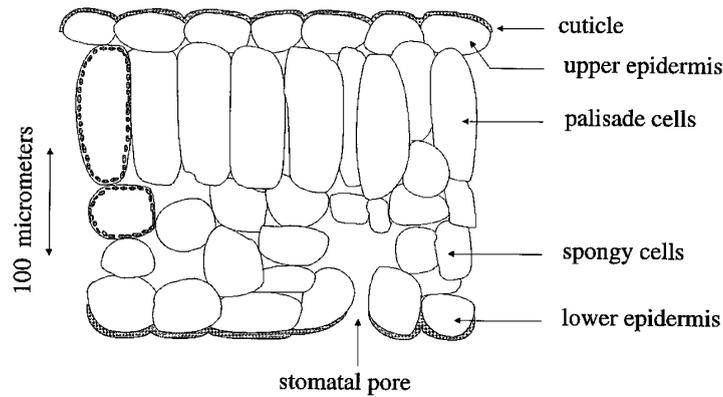


Figure 2.4: Schematic cross section of a dicotyledon leaf. Chloroplasts are drawn in only one cell of palisade and spongy tissues (Govaerts et al., 1996).

in this range is controlled by chlorophyll a and b which use the absorbed energy for photosynthesis. Maximum chlorophyll absorption lies around 450 nm (blue light) and 670 nm (red light) while absorption in the green (560 nm) is minimum. In healthy vegetation, chlorophyll concentration is around ten times higher than that of other pigments such as carotenoids, hence masking out the specific absorption features of these (Jacquemoud and Baret, 1990). Brown pigment concentration (tannis) increases with a decay of the plant (Baranoski, 1998).

- **Near infrared (NIR)** radiance is subject to leaf surface reflectance and intra-leaf scattering. Responsible for this scattering are the structural organization of the leaf tissues and

the arrangement of single cells and air spaces within the mesophylls. Absorption by pigments and other leaf chemicals in the NIR range is low, leading to high overall reflectance and transmittance rates. The strong increase in reflectance around 720 nm is also called the red-edge (Baret et al., 1992), while the range with maximum reflectance is called the NIR plateau.

- Leaf water and various carbon based chemicals composing the frame of the leaf tissues (e.g. lignin, protein, cellulose) dominate spectral absorption in the **shortwave infrared** (SWIR). Water absorption plays a dominant role in the entire SWIR with maximum absorption concentrated around 1200, 1450, 1940 and 2500 nm. In vigorous green vegetation, leaf water content masks the absorption features caused by the other biochemicals (Fourty and Baret, 1997).

Most energy that is absorbed by the leaf is either used for photosynthesis or stored as heat. However, a small portion of this radiance is first absorbed by chlorophyll and then remitted in the form of fluorescence (Camenen et al., 1986; Zarco-Tejada et al., 2002). Fluorescence in the visible domain is active in the region from 660 to 780 nm with peaks around 690 and 740 nm thus increasing reflectance in the red and the shoulder of near infrared plateau (Gamon et al., 1990). Under certain conditions, fluorescence may amount even up to 50% of the reflectance in the red domain (Maier, 2000). Fluorescence intensity depends on photosynthetic activity and is therefore highly variable in the course of a day. As chlorophyll fluorescence is a good indicator of photosynthetic activity, reduced fluorescence is an early sign of stress and often precedes the decay of chlorophyll (Zarco-Tejada et al., 2002).

Leaf optical properties are not at all a static phenomenon but continuously change during periods of growth, maturity, senescence, decay, or stress (e.g. water and nutrient deficiencies, parasites). For example, senescence is marked by an increase of reflectance over the whole spectrum: a rapid decrease of chlorophylls increases VIS reflectance and, at the same time, exposes the absorption characteristics of other pigments such as carotenoids (yellowing of the leaf) and xanthophylls (reddening of the leaf). With an ongoing stress, leaf structures decompose, giving room to extra intra-leaf scattering and hence to an increased NIR signal. At the same time, concentrations of brown pigments, which absorb radiance in the VIS and at the onset of the NIR, may increase leading to a flattening of the red-edge. Absorption in the SWIR decreases due to reduced leaf water content. With a decay of the leaf material, the absorption features characteristic for vegetation gradually disappear and the spectral shape approximates that of the soil background (Nagler et al., 2000).

2.2.2 Canopy structure

Canopy structure represents the spatial configuration of the scattering and absorbing vegetation elements and mainly accounts for the magnitude and the directional variation of the reflected signal (Knyazikhin et al., 1999c). Several factors play a role in this respect: the quantity, the angular distribution, and the form of the scattering and absorbing elements, the spatial distribution of vegetated and non-vegetated areas, both in the horizontal and vertical dimension, and the uniformity of this distribution (Kuusk, 1995a; Verhoef, 1984). Following variables are commonly used to describe canopy architecture:

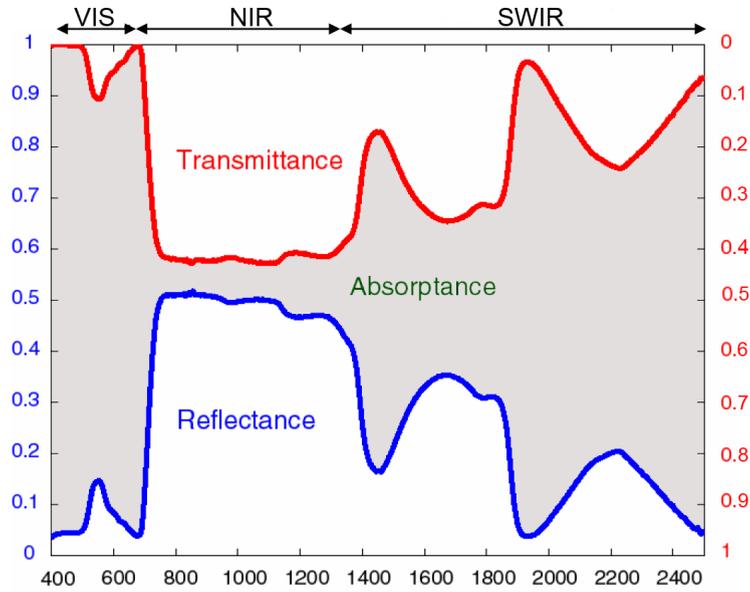


Figure 2.5: Example of hemispherical reflectance and transmittance spectra of a dicotyledon leaf. In gray the fraction of absorbed radiance. Hemispherical reflectance and transmittance spectra are symmetrical in shape but different in magnitude.

- The quantity of leaves and needles is usually expressed by the **leaf area index** or *LAI*. It can be defined as the total one-sided leaf area in a canopy volume divided by the unit horizontal ground surface area (Watson, 1947). Although this definition is clear for flat broad leaves, it may cause problems for needles and non-flat leaves. For this reason, it is also denoted as half the total developed area of leaves per unit ground horizontal surface area (Pinty et al., 2004a). *LAI* characterizes the surface available for exchange of energy and mass between the canopy and the atmosphere (Baret et al., 2005b) and is therefore a key variable when modeling surface evapotranspiration and biomass production as well as yield and yield loss (Broge and Mortensen, 2002). *LAI* typically varies from 0 for a bare soil to 5-6 for agricultural crops and up to 15 for a tropical forest. As multiple scattering within the canopy increases with an increasing *LAI*, canopy reflectance changes accordingly: reflectance in the VIS decreases (more absorption by pigments) while reflectance in the NIR increases. However, at *LAI* values greater than 3 (for the VIS) or 5-6 (NIR), radiation does not penetrate to the lower leaf layers leading to a saturated signal (Haboudane et al., 2004).
- The **leaf angle distribution** indicates the distribution of leaf inclination (zenith) and orientation (azimuth). It varies among vegetation species but also depends on phenological stage, stress (wilking), and sometimes even on the hour of the day (Thanisawanyangkura et al., 1997). A general distinction can be made between planophile (flat) and erectophile (upright) leaf angle inclinations (Verhoef, 1997).
- **Canopy height** and **leaf size** condition the probabilities of photon transmission and interception in the canopy, especially close the direction of the incident radiation. This bidirectional gap-probability (Kuusk, 1985; Kuusk, 1995a) gives rise to strong anisotropy

in canopy reflectance and to the foliage hot spot, which will be discussed more in detail in Section 2.2.4.

- The **spatial distribution** of vegetation elements in a canopy strongly influences the scattering and shading behavior. The definition of spatial heterogeneity depends on observation scale and can be observed at leaf level (*clumping* of leaves and needles (Kuusk, 1995a)), plant level, or canopy level. One can distinguish between horizontal and vertical heterogeneity. Examples of vertical heterogeneity are the layered structures present in a tropical forest or a cereal crop canopy having a top layer dominated by ears, an intermediate layer dominated by leaves, and a bottom layer mainly consisting of senescent material. Strong horizontal heterogeneity is found in open forest canopies, shrub lands, and row crops.

2.2.3 Background reflectance

Background reflectance plays a significant role for sparsely vegetated areas, and has to be well characterized. Background reflectance is usually dominated by soil optical properties although also crop residues, litter, organic crusts, and salt deposits may contribute to it. Soil optical properties differ greatly from those of green vegetation but approximate those of senescent plant material and litter, especially when it contains a considerable amount of organic matter (Figure 2.6). Soil moisture is one of the major determinants of spectral absorption in agricultural soils. A liquid water film at the surface of soil particles causes an almost uniform decrease of reflectance across the solar-reflective domain, whereas additional absorption can be observed around the common water absorption bands already referred to in Section 2.1.3 for atmospheric water (Baumgardner et al., 1985; Verhoef and Bach, 2007) (Figure 2.7).

Soil reflectance is highly non-Lambertian, a property depending mainly on surface roughness (Cierniewski and Verbrugge, 1997). Specular effects can be observed for smooth soil surfaces whereas hot spot effects are observed for soils with higher surface roughness. Moreover, soil improvement measures such as tillage and ploughing may influence the directional heterogeneity of soil reflectance.

The influence of the soil background over the total signal recorded by the sensor is wavelength dependent and largest in the near infrared region. Nevertheless, this influence can be neglected for canopies with LAI greater than 3 in homogeneous canopies (Atzberger et al., 2003) or values greater than 4-6 for row crops when the solar azimuth coincides with the row direction (Schneider, 1994).

2.2.4 Anisotropy in canopy reflectance

In the previous sections, the anisotropic behavior of atmospheric, leaf, canopy, and background scattering has already been pointed out. In the vegetation canopy itself, anisotropy is mainly induced by surface roughness (of leaves and soil particles), volume scattering, which is typical for any 3-D porous medium, shading caused by geometrical objects, and the contrast between leaf and background reflectance in combination with the structural effects of the canopy (Lucht et al., 2000; Roujean and Lacaze, 2002). Anisotropy effects are most pronounced in structurally vertical canopies such as grasses and conifer stands since the fraction of understory that is

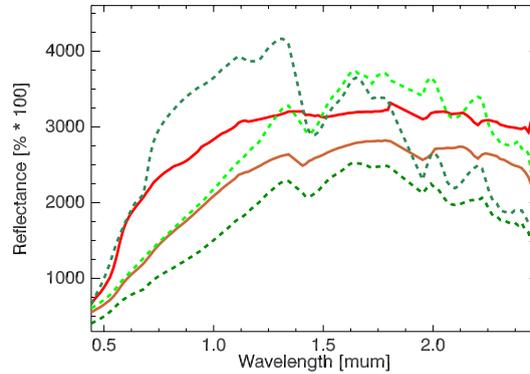


Figure 2.6: Spectral similarity of soils (solid curves) and non-photosynthetic vegetation (NPV, dashed curves) measured in southern Spain. By courtesy of Martin Bachmann.

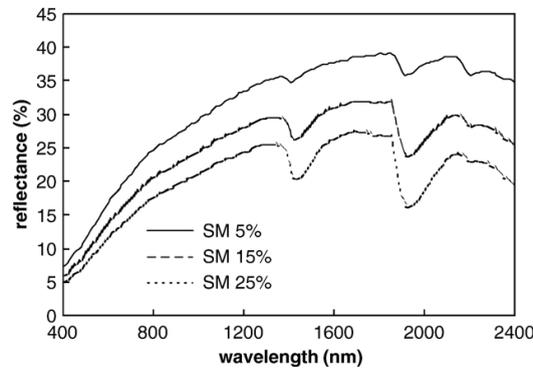


Figure 2.7: Effect of volumetric soil water content on soil reflectance (Verhoef and Bach, 2007).

viewed by the sensor significantly changes with changing view angle (Sandmeier et al., 1998; Sandmeier and Deering, 1999).

Anisotropy effects are particularly strong in spectral ranges of high vegetation absorbance, such as the visible bands, since the influence of background scattering in these ranges is more prominent. In highly reflective bands such as the NIR, multiple scattering processes effectively reduce the contrast between shadowed and illuminated canopy components, resulting in low anisotropy effects (Sandmeier et al., 1998). In the VIS, anisotropy in the principle plane typically has a bell shape for canopies with a vertical structure. This can be explained by the higher fraction of soil background (having relatively high reflectance compared to vegetation) that is observed by a sensor when it is in nadir position. Canopies having a predominant horizontal structure exhibit a more or less pronounced bowl shape anisotropy (Widlowski et al., 2004; Widlowski et al., 2005; Koetz et al., 2005b) (Figure 2.8).

Due to the internal structure of canopy, there is a higher probability that radiance that has been scattered between the elements in the canopy, leaves the canopy again in a direction opposite to the incident radiation (Figure 2.9). This peak in backscatter reflectance is called the hot spot effect (Kuusk, 1985; Verhoef, 1997). Actually, the observed hot spot is a combination of the hot spot resulting from soil background, volume scattering (the bi-directional gap probability), and

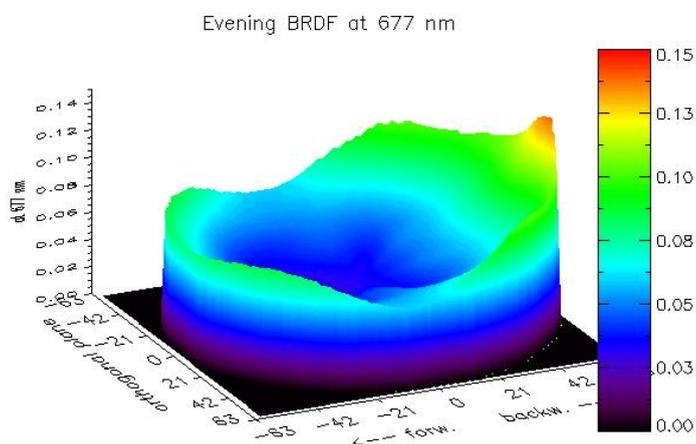


Figure 2.8: Reflectance anisotropy for a maize crop in the red wavelength domain and for view zenith angles between 60° in forward and backward direction. The HDRF was measured with a field goniometer system at low solar elevation (Schneider et al., 2006).

geometrical shading (Verhoef and Bach, 2007). The contribution of the latter can be explained by the fact that the canopy elements - being large compared to the wavelength - cast shadows that are well visible when the angle between incident radiation and observer (phase angle) is large, but hidden by the elements themselves when the phase angle is close to zero (Camacho-de Coca et al., 2004; Roujean, 2000; Sandmeier et al., 1998).

The reflectance anisotropy of ground surfaces can be described by the bidirectional reflectance distribution function (BRDF) which for a given surface depends on wavelength and view/sun constellation. It describes the bidirectional reflectance factor (BRF) as a function of illumination geometry (azimuth and zenith) and the position of the observer in the hemisphere. Under natural conditions, i.e. for all measurements outdoors, the assumption of a single direction of the incident beam is unrealistic since atmospheric scattering will lead to a considerable amount of diffuse hemispherical incoming radiation, especially in the VNIR domain (Schaepman-Strub et al., 2006). In such a situation the BRDF is based on HDRFs instead of BRFs and thus, for an accurate thematic evaluation based on reflectance of BRDF shapes measured outdoors, the diffuse fraction either must first be corrected for or has to be implicitly taken into account in subsequent processing.

2.3 Retrieving canopy variables by statistical approaches

Statistical approaches search for a consistent relationship between the spectral signature of an object, in general the leaf or canopy reflectance, and the biophysical or biochemical variable of interest. For establishing such a relationship, the spectral information is rarely directly used. Many data manipulations have been proposed to enhance subtle spectral features and to reduce undesired effects caused by variations in soil reflectance, sun and view geometry, atmospheric composition, and other leaf or canopy properties. Standard manipulations, often used in imag-



Figure 2.9: Hot spot effect in HyMap airborne imagery. False color image of Barrax, Spain (acquired on June 3, 1999 at solar noon)

ing spectroscopy, involve normalization (Chappelle et al., 1992), logarithmic transformation (Jacquemoud et al., 1995b; Yoder and Pettigrew-Crosby, 1995), continuum removal (Kokaly and Clark, 1999; Mutanga et al., 2004), and the calculation of first or second derivatives (Huang et al., 2004).

However, the most widespread method used to reduce background effects and enhance spectral features is to express spectral reflectance in a combination of a limited number of (transformed) spectral bands to create what is known as a vegetation index (VI). Most VIs concentrate on the red-edge region, which is the region between 680 and 800 nm that is characterized by a sharp decrease of chlorophyll absorption from maximum absorption around 680 nm to almost zero absorption at 800 nm. This makes this wavelength range very well suited to study vegetation characteristics (Baret et al., 1992).

Probably the most widely used VI is the normalized different vegetation index (NDVI), an index based on the early broadband earth observation systems (Rouse et al., 1973). After this early VI, a wide range of VIs have been developed through the years, each trying to reach maximum sensitivity for a specific leaf or canopy variable while minimizing effects of background, atmosphere, and other canopy properties. The coming up of hyperspectral spectrometers enabled the calculation of first and second derivatives, the establishment of the so-called red edge inflection point (REIP) (Baret et al., 1992; Guyot et al., 1988; Dawson and Curran, 1998) and the development of indices based on absorption features caused by biochemical components other than chlorophyll. The reader is referred to (Broge and Mortensen, 2002; Dorigo et al., 2007; Haboudane et al., 2004) or to Appendix C for more extensive overviews of existing VIs.

The traditional way of linking spectral information to the variable of interest is by simple or multiple regression techniques (Clevers, 1989; Jacquemoud et al., 1995b) (Figure 2.10). Recently, more sophisticated statistical approaches, such as partial least square regression and artificial neural networks (ANN), have been introduced (Atzberger, 2004; Huang et al., 2004; Combal et al., 2002b). Whereas most regression techniques assume a linear or exponential relationship between the biophysical variable of interest and the measured spectral signature, the artificial neural network is a sophisticated technique capable of modeling extremely complex nonlinear functions. ANNs can be considered as an interpolation technique that establishes a response surface (M) between the reflectance values R (inputs) and the variables of interest V (outputs), i.e.:

$$V = M(R) + \epsilon$$

where ϵ accounts for the model error. Calibration of the model M is performed on a learning data set. In the training stage, inputs (spectral information) are linked to a selection of output variables, thus optimizing the network for the variables of interest (Combal et al., 2002b). This is because a nonparametric statistical model can nonlinearly project data so that some factors are enhanced while others are suppressed. ANNs are frequently used in combination with radiative transfer model simulations (Section 2.5.3.2), although encouraging results have been obtained as well for measured variables (Huang et al., 2004).

A limiting factor in the use of the statistical approaches, in particular for VIs or other spectral transformations, is the fact that they are never sensitive to the variable(s) of interest only. As a consequence, for different plant species, phenological development stages, soil background, and so on, often a new relationship has to be established between VI and variable of interest. It may be clear that this is a cost and time intensive matter and - even if such is done properly - it is nearly impossible to cover all possible realistic scenarios, which in turn might influence the validity and portability of the relationships (Clevers, 1989; Jacquemoud et al., 1995b). For this reason, several authors proposed to use radiative transfer model simulations for establishing and calibrating the empirical relationships (Haboudane et al., 2002; Baret and Guyot, 1991; Huete et al., 1997; Rondeaux et al., 1996; Verstraete and Pinty, 1996). Radiative transfer simulations have formed the basis of developing a number of operationally used predictive empirical algorithms for medium resolution sensors, such as the MERIS Global Vegetation Index (MGVI) (Gobron et al., 2000), the MODIS back-up algorithm based on NDVI (Knyazikhin et al., 1999a), and the POLDER algorithm based on the difference vegetation index (DVI) computed from bidirectional reflectance factor measurements normalized to a standard geometrical configuration (Roujean and Lacaze, 2002).

Although often effective, VIs are intrinsically limited by the empiricism of their design and the small number of bands concurrently used (generally 2 to 3). This might not be a major problem for $fAPAR$ and $fCover$ which have a nearly linear relationship to canopy reflectance, but would hamper the retrieval of variables such as LAI or chlorophyll content which show higher level of non linearity with reflectance measurements (Baret and Buis, 2007; Weiss et al., 2000). Partly due to this phenomenon increasing interest goes out in the inversion of radiative transfer models for the retrieval of biophysical and biochemical variables.

2.4 Spectral unmixing for cover fraction estimates

Although not directly linking to biophysical or biochemical information content like the statistical methods discussed in the previous section, spectral unmixing is often applied for estimating cover fractions of partly vegetated surfaces and for the differentiation of plant communities (Garcia-Haro et al., 2005). Spectral *mixing* models rely on the assumption that the irradiance measured by a detector element equals the integrated sum of fractional spectral contributions of the different surface elements found within the IFOV (Adams and Smith, 1997). When a spectral mixing model is *linear*, then the spectral contribution of each element, or *endmember*, approximately coincides with its fractional coverage (Adams and Smith, 1997). For the measured

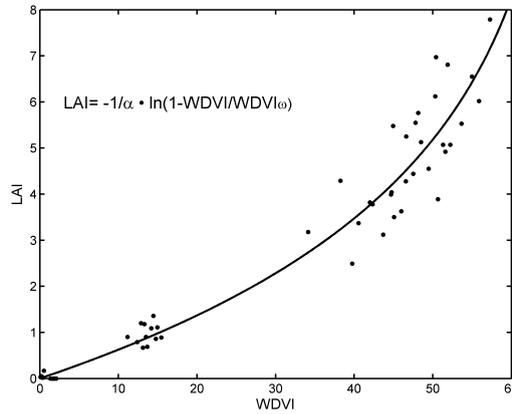


Figure 2.10: Exponential empirical relationship between WDVl and LAI (adapted from (Clevers, 1989)).

reflectance in a single waveband (R_{meas}) this is expressed as:

$$R_{meas} = \sum_{i=1}^n f_i \cdot R_i + \epsilon \quad (2.8)$$

where f_i is the fractional cover and R_i the spectral reflectance of endmember i , while ϵ is an error term describing spectral variation unexplained by the spectral mixing model as well as uncertainties in pre-processing, sensor noise and so on.

Spectral *unmixing* consists in finding the solution of Equation 2.8, i.e., by determining the single contributions of the n endmembers given the measured input reflectance R_{meas} . Endmember reflectance properties R_n are either based on laboratory or field spectrometer measurements or on the signatures of pixels found at the extremes of the spectral feature space covered by the image. Conventional spectral unmixing approaches are based on a predefined set of endmembers for the entire scene. To cover the complete spectral variation of the single endmembers (e.g. due to different biochemical concentrations or variable LAI) a large number of endmembers is required. However, the collinearity between the several endmembers and the large number of endmembers compared to the effective dimensionality of the data can lead to unstable solutions and large errors in the retrieved abundances (Garcia-Haro et al., 2005; Winter et al., 2003). To overcome such complications, multiple endmember spectral mixture analysis (MESMA) approaches have been developed (Asner and Lobell, 2004; Bachmann, 2007; Garcia-Haro et al., 2005; Roberts et al., 1998) which allow for a pixel based optimization of the mixture model based on only a limited number of endmembers per pixel (e.g. soil, non photosynthetic, and photosynthetic active vegetation). For a vast overview of different unmixing strategies the reader is referred to (Bachmann, 2007).

The highly non-linear character of vegetation reflectance, especially in denser canopies, complicates the use of spectral unmixing for open-structured agricultural crops where there is a considerable radiation exchange between soil and vegetation. For this reason, spectral unmixing is typically applied in semi-arid regions where vegetation shows a more patchy distribution

(Bachmann, 2007).

2.5 Physical approaches for retrieving vegetation variables

Physical approaches are based on the inversion of radiative transfer models which simulate the interactions between solar radiation and the elements constituting the canopy, using physical laws (Figure 2.11). Before discussing the inversion of radiative transfer models for the estimation of biophysical variables, a more detailed look at the functioning of these models in the forward mode is indispensable.

The most established way of modeling radiative transfer in canopies is to combine a leaf optical model with a canopy structure and a soil reflectance model to calculate the top-of-canopy reflectance. For remote sensing applications, this modeled reflectance should be in agreement with measured reflectance data corrected for atmospheric influences (i.e., the HDRF). A different approach consists of computing the radiance as it would have been measured by the sensor and therefore encompasses a model for the calculation of radiance propagation in the atmosphere (Baret et al., 2005a; Verhoef and Bach, 2003a; Verhoef and Bach, 2007). In the following sections the discussion will be confined to reflectance at canopy level, and therefore does not consider any atmospheric model. Soil reflectance is an important element in radiative transfer modeling, being the lower boundary condition and having its own spectral properties (e.g., absorption features and directional anisotropy). Knowing soil reflectance properties is fundamental if sparse or low vegetated canopies are to be simulated. This model input is typically measured in the field, taken from the image itself, or can be simulated using soil reflectance models (Hapke, 1981; Jacquemoud et al., 1992; Pinty et al., 1989). Variations in soil reflectance representing differences in soil moisture content, surface roughness, soil organic matter, and inorganic carbon content are often parameterized using simple empirical formulas and scaling factors (Atzberger, 2004; Atzberger et al., 2003; Baret et al., 2005b). Soil reflectance models are beyond the scope of this thesis and will therefore not be discussed in further detail here.

2.5.1 Leaf optical models

Our understanding of leaf microstructures and the distribution of biochemical components in leaves is still very limited. Even if the same is true for the anisotropic scattering of leaves (Jacquemoud and Ustin, 2001), various approaches have been proposed that successfully describe leaf scattering and absorption in a more or less simplified way (Table 2.2).

N-flux models (Fukshansky et al., 1991; Richter and Fukshansky, 1996) are based on the Kubelka-Munk theory and consider the leaf as being a slab of diffusing and absorbing material. The strength of this approximation is its simplicity. However, it is very difficult to link the overall absorption coefficient of the leaf with the specific absorption coefficients and the concentrations of the plant chemicals. A retrieval of concentrations of biochemicals by model inversion is therefore rather complex, if not even impossible (Fukshansky et al., 1991).

Plate models consider the leaf as a pile of N plates separated by $N-1$ air spaces, where N does not have to be a discrete value. The structure parameter N is comparable to the scattering parameter in the Kubelka-Munk theory. It is an uncomplicated model which assumes Lamber-

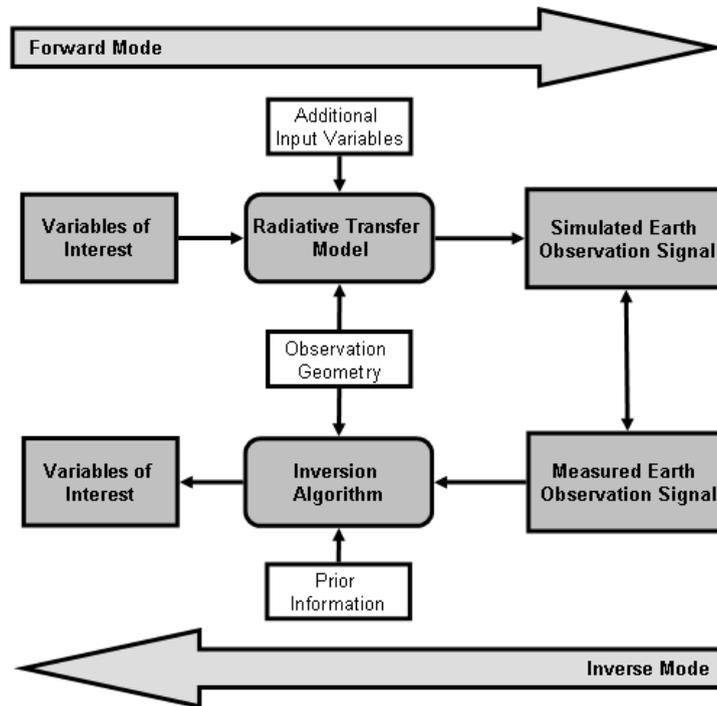


Figure 2.11: Concept of radiative transfer modeling: forward mode (above) and inverse mode (below)

tian scattering and absorption elements (biochemicals) distributed homogeneously throughout the leaf. This assumption makes it suitable for the retrieval of leaf biochemicals by model inversion. However, biochemicals with only relatively small absorption features cannot be accurately retrieved (Fourty et al., 1996). The best known model in this category is the PROSPECT model developed by Jacquemoud and Baret (1990).

Ray tracing models, based on Monte Carlo simulations, are the only type of model that accounts for the complexity of internal leaf structure and are therefore the most realistic (Govaerts et al., 1996). A ray tracing model simulates the propagation of photons within a leaf foliage. This potency is at the same time its weakness: it requires a very detailed description of individual cells, their optical constants, and their arrangement within the foliage. This complex description of light propagation makes the model computationally very intensive and model inversions numerically difficult to implement.

Radiosity models describe the total amount of energy leaving a surface per unit time per unit area (Liang, 2004). The leaf is divided in discrete reflecting and transmitting elements with distinct shapes, positions and orientations. The major advantage of this method is that once a solution has been found for radiative transport, leaf and canopy reflectance can be simulated at any view angle for any wavelength. The major limitation of the method is the initial computational load in forming the view factor matrix and solving for radiative transport. The Algorithmic BDF Model of Baranoski and Rokne (Baranoski and Rokne, 1997) is the best known leaf radiosity model.

Stochastic models are based on Markov chains where transition probabilities from one

Table 2.2: Examples of various approaches used to model leaf and canopy radiative transfer

<i>Medium</i>	<i>Type</i>	<i>Leaf model</i>	<i>Canopy Model</i>
Homogeneous	1D radiative transfer	(Fukshansky et al., 1991)	SAIL (Verhoef, 1984), KUUSK (Kuusk, 1995b)
	Plate model	PROSPECT (Jacquemoud and Baret, 1990)	
Heterogeneous	3D radiative transfer	-	DISORD (Myneni et al., 1992)
	Geometric	-	SGM (Chopping et al., 2003), (Chen and Leblanc, 1997)
	Hybrid	-	DART (Gastellu-Etchegorry et al., 1996), GeoSAIL (Huemmrich, 2001), TRIM (Goel and Grier, 1988), 4SAIL2 (Verhoef and Bach, 2007)
	Ray tracing	RAYTRAN (Govaerts et al., 1996)	RAYTRAN (Govaerts and Verstraete, 1998), SPRINT (Goel and Thompson, 2000)
	Radiosity	ABM (Baranoski and Rokne, 1997)	PARCINOPY (Chelle and Andrieu, 1998)
	Stochastic	SLOP (Maier et al., 1999)	(Shabanov et al., 2000)

radiation state to another, for instance from reflected to absorbed, are described. Unfortunately these models are more computation-intensive than plate and N-flux models and are not suited for direct inversions. An example of such a model is SLOP (Maier et al., 1999).

Although it is very well known that leaf reflectance and transmittance are strongly anisotropic (non-Lambertian), depending on angle of incidence, up to date these phenomena rarely have been captured in the leaf optical models (Bousquet et al., 2005; Jacquemoud and Ustin, 2001).

2.5.2 Canopy radiative transfer models

A wide range of canopy structure models have been developed, for which the difference principally lies in the grade of detail and the canopy type the model intends to describe (Table 2.2). Traditionally, canopy radiative transfer models assume that the canopy is a **turbid medium** where the canopy elements (leaves) are treated as infinitely small, randomly distributed absorbing and scattering elements with no physical size. A one-dimensional approximation (Verhoef, 1984) assumes the canopy to be horizontally homogeneous and infinite but vertically variable and finite. These assumptions, together with the fact that leaf area is explicitly taken into account, make this type of model well suited for describing radiance propagation in denser canopies where the single vegetation elements are smaller than the canopy height, which is the case for most agricultural crops. Most models of this type are based on the Kubelka-Munk theory and consider four different radiative fluxes describing both specular and diffuse downwelling and outgoing radiation. A hot spot function describes the gap probability in the principal plane but only accounts for volume scattering and not for mutual shading. Several models of this type have seen the

Table 2.3: Differences between radiative transfer and radiosity approach (after Gerstl and Borel (1992)).

<i>Radiative Transfer</i>	<i>Radiosity</i>
Volume scattering	Surface reflection and transmission
Continuous medium	Discrete and oriented surfaces
Averaged scattering phase function	Explicit scattering characteristics
No physically based spatial correlations of leaves	Spatial correlations retained
No holes or clumps in canopy	Holes and clumps describable
Multiple Scattering	Multiple scattering
Integrodifferential equation	System of coupled linear equations

light of day during the last two decades (Kuusk, 1995b; Verhoef, 1984; Gobron et al., 1997).

For situations in which the assumption of a horizontally homogeneous and infinite canopy does not apply (e.g. heterogeneous canopies like row crops and orchards with isolated tree crowns), different approaches based on the principle of radiative transfer have been proposed. **Geometrical models** have been formulated to describe radiation propagation in sparse canopies where multiple scattering can be ignored and mutual shading is negligible due to low zenith angles (Chen and Leblanc, 1997). In this approach the canopy is described as a ground surface with opaque geometrical objects with known spatial distribution and optical properties. Often geometrical and turbid medium approaches have been combined to form so-called **hybrid models**, in which the canopy is made up of translucent geometrical objects (plants) to which the turbid medium radiative transfer equation is applied. Therefore, this type of model is particularly suited for representing forests or sparse canopies. The Three Dimensional Radiation Interaction Model (TRIM (Goel and Grier, 1988)), GeoSAIL (Huemmrich, 2001), and the Discrete Anisotropic Radiative Transfer (DART) model (Gastellu-Etchegorry et al., 1996) are examples of such 3D hybrid radiative transfer models. Various other approaches based on radiative transfer have been suggested to account for both vertical and horizontal heterogeneities (Kuusk, 1995a; Myneni et al., 1992; Shabanov et al., 2000).

Monte-Carlo ray tracing models (Goel and Thompson, 2000; Govaerts and Verstraete, 1998) and *radiosity models* (Borel et al., 1991; Chelle and Andrieu, 1998; Gerstl and Borel, 1992) give a more realistic representation of the radiation transfer in the canopy and offer the possibility to study the statistical nature of a radiation behavior. The high level of detail used to describe radiation propagation makes this type of model computationally very intensive. Table 2.3 summarizes the main differences between radiosity and 1-D turbid medium canopy modeling.

The choice of a canopy structure model is not only driven by the canopy type under consideration but also depends on the spatial resolution of the observation (Pinty et al., 2004a; Widlowski et al., 2006a). A vast intercomparison of the performance of several existing canopy radiative transfer models at several spatial resolutions and varying canopy complexity is given in (Pinty et al., 2001; Pinty et al., 2004b; Widlowski et al., 2006b). For a detailed overview of the various leaf and canopy radiative transfer models described in this section, and the equations used to simulate radiation propagation, the reader is referred to the original publications and to (Liang, 2004).

2.5.3 Radiative transfer model inversion

Inverting a radiative transfer model (RTM) consists in finding the set of input variables that leads to the best match between the HDRF (or top-of-atmosphere radiance) simulated with the model and the reflectance (or radiance) measured by the sensor (Combal et al., 2002b). Considering that this quantity is a function of the canopy variables V and the measurement geometry Ω , the reflectance R can be written as:

$$R = f(V, \Omega) + \epsilon \quad (2.9)$$

where ϵ represents the error, both for measurement (sensor noise, data pre-processing, etc) and model uncertainties (model simplifications, assumptions, etc.). Assuming that illumination and observation properties are a priori known, the inverse problem consists in estimating the set of variables V that leads to the observed spectrum R . Since ϵ is not precisely known, usually rather a best matching spectrum is sought, rather than a unique solution (Section 2.5.4). Different methods have been developed to find this best fit between modeled and measured reflectance.

2.5.3.1 Iterative optimisation techniques

Iterative optimisation techniques search for the best fit between the simulated and the measured reflectance by iteratively running the radiative transfer model with different sets of input variables. The minimization of a cost function that traditionally accounts for differences between the simulated and the measured reflectance is used as a stopping criterium for this optimisation problem. The cost function to minimize may be theoretically derived from the maximum likelihood estimator indicating that measurement uncertainty is introduced leading to a solution that is not exactly determined but constituted by the maximum probability (Bacour, 2001; Tarantola, 2005). In the case where only radiometric information is considered and where uncertainties for each spectral band i are assumed independent and gaussian, the cost function (χ^2) to be minimized is expressed by:

$$\chi^2 = \sum_{i=1}^{n_{meas}} \frac{(R_{meas}^i - R_{sim}^i)^2}{\sigma_i^2} \quad (2.10)$$

where the distance is minimized between measured reflectance values (R_{meas}^i) and those simulated with the radiative transfer model (R_{sim}^i), weighted by the variance associated to both reflectance measurements and non-systematic model uncertainties (σ_i^2). However, because of the difficulty of providing an estimate of the variance, in particular that of model uncertainties, several approximations of Equation 2.10 have been proposed, as summarized in Bacour (2001) and Baret and Buis (2007).

A wide range of mathematical and statistical approaches is used to converge to the minimum, such as Quasi-Newton and Simplex algorithms (Bacour et al., 2002b; Vohland and Jarmer, 2007), simulated annealing, genetic algorithms (De Wit, 1999), and Bayesian probability statistics (Lavergne et al., 2007; Zhang et al., 2005; Verhoef, 2007). The latter category has been recently

discovered as a vigorous alternative to the classical numerical approaches and are promising both in terms of increased model stability and computation efficiency.

The classical iterative approaches present a few drawbacks: (i) they require an initial guess of the solution. The correctness of this initial guess can be critical to get a successful inversion. If an inadequate inversion algorithm is used the solution may get trapped in a local minimum (Qiu et al., 1998). However, this can be partly avoided by confining the range of variation of the variables to be estimated (Baret and Buis, 2007). (ii) The approach is computationally too intensive to be suited for operational use or for the inversion of complex three dimensional radiative transfer models (Liang, 2004), although there are ways to speed up the process by limiting the number of model runs for each iteration by using the adjoint model that provides an analytical expression of the gradient of the cost function (Lavergne et al., 2007). (iii) Apart from the Bayesian approaches, iterative algorithms do not consider an *a posteriori* probability distribution of the retrieved variables.

Despite the several drawbacks, various authors have successfully applied this technique for various vegetation types (Bacour et al., 2002b; De Wit, 1999; Fang et al., 2003; Jacquemoud et al., 1995a; Qiu et al., 1998). Its strength mainly resides in its flexibility, allowing to adapt to changing observational configurations. It is even possible to invert radiative transfer models concurrently over several configurations (Baret and Buis, 2007), opening great potential for the exploitation of additional temporal, angular, or spatial constraints (Section 2.5.5).

2.5.3.2 Lookup table approach

In the lookup table (LUT) approach, the radiative transfer model is used in a direct way to simulate a large number of possible spectra, depending on different combinations of input parameter values. In a next step, a cost function (e.g. one similar to Equation 2.10) searches in the LUT for the spectrum that has the highest similarity to the measured one. The set of variables used to simulate this spectrum is considered the final solution of the inversion. Usually, not a single spectrum but the average or median of a limited number of best fitting spectra close to absolute minimum is proposed as the solution (Weiss et al., 2000). This allows one to derive a posteriori variable distribution indicating the "confidence level" around the final estimate (Koetz et al., 2004). The LUT approach has the advantage of being considerably faster than iterative optimisation techniques, although computation time can significantly increase if large lookup tables, representing a wide range of canopy realizations, are used. The technique can also deal with more complex radiative transfer models, such as ray tracing models, since the model is used only in the forward mode.

A major drawback of this method is that a LUT has to be generated for each new situation, and specific sun and view geometry, although Gastellu-Etchegorry et al. (2003) found a satisfactory way to interpolate between view, sun, and azimuth angles. A point of concern which may strongly influence the final result of the inversion is the composition of the LUT. A poor choice of the distribution of the different variables (e.g. uniform or gaussian) or the step size between successive values of the variables in the LUT may lead to an over- or under representation of certain variable ranges and thus to a non representative result (Combal et al., 2002b; Weiss et al., 2000). The composition of the LUT is especially critical for models having many input arguments. For these models, the size of the LUT dramatically expands if input variables are sampled at small intervals (Combal et al., 2002b; Weiss et al., 2000). Another point of criticism

may be the fact that the solution is not always based on a solid theoretical background but rather depends on an arbitrarily chosen threshold (Baret and Buis, 2007). The cases selected as possible solutions are either defined as a fraction of the initial population of cases (Combal et al., 2002b; Weiss et al., 2000), or can be defined by a threshold corresponding to measurement and model uncertainties as in Knyazikhin et al. (1999c). LUT inversion is being effectively applied for the retrieval of LAI and fAPAR on a global scale from MISR/MODIS data (Knyazikhin et al., 1999a).

2.5.3.3 Artificial neural networks

Inversion based on an Artificial Neural Network (ANN) can be seen as the most prominent member of a collection of hybrid approaches that combine physical and statistical models (Liang, 2004). Like in the LUT approach, a radiative transfer model is used in the direct mode to build a large synthetic data set that represents a wide variety of canopy realizations. While the LUT approach searches for the simulated spectrum closest to the measured one, the ANN minimizes the distance between canopy biophysical variables. ANNs are fast after the training stage and according to Combal et al. (2002b) less sensitive to model uncertainties than the iterative optimization and LUT approach. ANNs can effectively adapt for multispectral and multi-angular data, making it an adequate technique for estimating variables from sensors with large swath angles (Liang, 2004). Bacour et al. (2006) found satisfying results estimating LAI, fCover, and fAPAR for the globally operating medium resolution MERIS sensor. As the radiative transfer model is used only in the forward mode, ANNs are suited for inverting more complex models such as 3-D or ray tracing models. ANNs are often criticised of having a black-box character, which makes it difficult to study cause-effect relationships.

2.5.4 Under-determination and ill-posedness in radiative transfer model inversion

Estimating canopy variables by radiative transfer model inversion is mathematically often an under-determined problem, indicating that the number of unknowns is generally larger than the number of independent radiometric information sources. Baret and Buis (2007) state that even a simplified representation of canopy directional reflectance, based on a combination of simple soil, leaf, and canopy reflectance models, already requires at least 13 input variables. In this case, theoretically, at least the same number of observational configurations (e.g. wave bands, view directions, or a combination of both) would be required to retrieve the variables from the remote sensing data by model inversion. Although several satellite systems (e.g. MODIS and MERIS both with 15 bands, Hyperion with 220, and CHRIS with 18 to 62 bands in the solar-reflective domain) would fulfill this requirement, the actual dimensionality of remote sensing measurements is much smaller than the number of available configurations, given the high level of redundancy between bands and view directions (Gemmell, 2000; Verhoef, 2007; Zhang et al., 2002a; Zhang et al., 2002b).

A major drawback of the inversion of physically based radiative transfer models is the fact that the inversion does not fulfill the Hadamard's postulates of well-posedness (Hadamard, 1902). Hadamard stated that a problem is well posed if and only if its solution exists, this solution is unique, and depends continuously on the data. The inversion of radiative transfer models

is ill-posed for two reasons: (i) the solution is not necessarily unique because of the compensation between several variables that affect canopy reflectance in a similar way (Combal et al., 2002b; Fourty and Baret, 1997). For example, in the combined leaf and canopy radiative transfer model PROSPECT (Jacquemoud and Baret, 1990) and SAIL (Verhoef, 1984), the spectral reflectance of a sparse canopy (low LAI) with planophile leaf orientation is very similar to one of a dense, erectophile canopy (Baret and Guyot, 1991; Jacquemoud et al., 1995a). (ii) Model uncertainties and assumptions and approximations used for radiative transfer modeling may induce large variations in the solution of the inverse problem (Combal et al., 2002a; Combal et al., 2002b; Privette et al., 1996). This is particularly true for well developed canopies, where a small variation in the measured reflectance can translate into large variation of variables such as LAI , for which reflectance 'saturates', i.e., is very little sensitive to variation of LAI (Baret and Buis, 2007). The problem of ill-posedness augments with an increasing number of input variables that are left free during inversion (Combal et al., 2002b).

2.5.5 Improving retrieval performances

As seen in the previous section, radiative transfer model inversion in general has an under-determined and ill-posed nature. Several regularization techniques can be proposed to reduce the ambiguities involved and get a more stable and reliable solution, either by increasing the dimensionality of the observation or by introducing constraints on the variables (Tarantola, 2005). Although these methods are discussed in the context of radiative transfer model inversion, they apply to all inverse remote sensing problems, so just as well to the statistical methods discussed in Section 2.3.

2.5.5.1 Increasing the observation dimensionality

Hyperspectral remote sensing Increasing the number of spectral bands leads to an increase of the spectral dimension. However, increased spectral dimensionality does not necessarily imply an increased information content. Additional bands should be positioned at wavelengths where canopy reflectance is significantly influenced by changes in one or more variables. Moreover, only little collinearity should exist between the additional bands and the channels already available. Verhoef (2007) reports that a well calibrated single-view nadir looking hyperspectral sensor contains up to 12 independent data dimensions, while this dimensionality significantly reduces with increasing noise levels. The additional dimensions offered by imaging spectrometers compared to multispectral sensors, appear to contribute mainly to improved estimates of leaf biochemical variables (Verhoef, pers. comm.).

The large amount of redundancy in the data expresses the need of carefully weighting the separate bands. If the collinearity is explicitly accounted for in the minimization procedure, the assumptions made in Equation 2.10 are no longer valid, which leads to the following cost function:

$$\chi^2 = (R_{meas}^i - R_{sim}^i)^T \cdot COV_{rad}^{-1} \cdot (R_{meas}^i - R_{sim}^i) \quad (2.11)$$

where COV_{rad} is the matrix containing the covariance between the spectral bands and should account for both model and measurement uncertainties. The latter depend both on canopy type and observation/illumination properties and are therefore strongly variable and difficult to estimate. For this reason covariance description is often omitted. The diagonal element of COV_{rad} is the variance of the single bands and equals the σ^2 of Equation 2.10.

Multi-angular observations As seen in Section 2.2.2, each vegetation canopy exhibits a characteristic anisotropic behavior due to differences in canopy architecture. Multi-angular remote sensing is a system able to capture such behavior and, if a sufficient number of directions is available, can even provide albedo estimates (Knyazikhin et al., 1999c). The use of multi-angular observations increases the dimensionality of the inverse problem thus enabling a more stable inversion (Lavergne et al., 2007; Verhoef, 2007). The number of configurations of a target is significantly increased when directional information is added to the spectral information (e.g. the amount of configurations of MISR adds up to 36 when all view angles are included (9 view angles \times 4 bands), for POLDER it totals 84 (14 directions \times 6 spectral bands)), and for CHRIS it may even add up to 310 (5 view angles \times 62 bands).

Even if there is a large redundancy of information, similar as in hyperspectral observations, especially the retrieval of structural canopy variables, such as *LAI*, shrub density, and vegetation height, benefits from the increased dimensionality (Widlowski et al., 2004; Chopping et al., 2003; Gobron et al., 2006) and the information content of directional information alone permits the retrieval of 2-3 structural variables (Gemmell, 2000; Barnsley et al., 1997). The increased dimensionality offered by the angular anisotropy is therefore complementary to those offered by hyperspectral observations (Verhoef, 2007). Based on field goniometer measurements, Schneider et al. (2006) not only found improved estimates of structural variables, but also obtained improved leaf chlorophyll retrievals when multi-angular instead of mono-directional observations were used. Moreover, improved stability of the estimates with respect to changing sun zenith angles was observed. Also land cover classifications appear to become more accurate when directional anisotropy is included (Sandmeier and Deering, 1999; Brown de Colstoun and Walthall, 2006). As seen in Section 2.2.4, the highest information content on structural parameters is found in the principle plane and in particular around the hot spot (Jacquemoud et al., 2000). However, due to their polar space orbits, in practice, satellite observations around the hot spot are difficult to realize.

In the case of including multi-angular observations in the cost function, COV_{rad} in Equation 2.11 should also contain the covariance between the different view angles.

2.5.5.2 Constraints on input variables

Variable fixation The most straightforward way of reducing the under-determinedness of the inverse problem is to set one or more variables to a fixed value. Usually, those variables are fixed that cause little variation in the modeled reflectance and show little variation for the observed canopy. Fixation of variables requires thorough a priori knowledge of the canopy under observation since the fixed value should lie closely to the value actually present in the canopy. Even if in many situations the inversion process becomes more stable, it does not necessarily lead to improved estimates of the retrieved state variables, as usually poorer fitting occurs between

measured and simulated reflectance (Widlowski et al., 2005).

A special category of variable fixation is constituted by fixing the ratio between two variables. This is often done for leaf water and leaf dry matter content since relative water content is usually close to 80% for green leaves and around 20% for senescent leaves (Bacour et al., 2006; Baret et al., 2005b; Baret et al., 2005a). Other authors proposed to use a ratio of 1:3.2 (Vohland and Jarmer, 2007) or to allow for some variation within the given ratio of 1:4 (Vohland and Mader, 2007).

Synthetic variables Some variables in the radiative transfer model always appear in interaction. Such interaction is called ambiguity or equi-finality (Baret and Buis, 2007) and often occurs between the content of an absorbing (leaf) material and the *LAI*. Although not appearing formally in the radiative transfer model, the product of these interacting variables would be a physically meaningful variable corresponding to the actual optical thickness of the medium (Weiss et al., 2000). Several authors achieved improvements when retrieving such synthetic variables instead of the single variables separately (Combal et al., 2002b; Fourty and Baret, 1997; Jacquemoud et al., 1995a).

Introducing a priori information A priori information is information on the distribution of the variables already available and might be collected from literature, field measurements, other sensors, or previous experiments (Baret and Buis, 2007). Prior knowledge on the distribution of variables includes information about upper and lower limits of the variable ranges and probability distributions of the expected result. In fact, also the variable fixation discussed in the previous paragraph is a form of using prior information. Various authors report considerable improvements in the retrieval performance when a priori information is included (Combal et al., 2002a; Combal et al., 2002b; Knyazikhin et al., 1999b; Lavergne et al., 2007; Verhoef, 2007).

When prior information is included in the cost function searching for the maximum of a probability density function then Equation 2.11 would turn into (Combal et al., 2002b; Rodgers, 2000; Tarantola, 2005):

$$\chi^2 = \underbrace{(R_{meas}^i - R_{sim}^i)^T \cdot COV_{rad}^{-1} \cdot (R_{meas}^i - R_{sim}^i)}_{\text{Radiometric information}} + \underbrace{(V_{prior}^j - V_{sim}^j)^T \cdot COV_{var}^{-1} \cdot (V_{prior}^j - V_{sim}^j)}_{\text{prior information}} \quad (2.12)$$

where V_{prior}^j is the vector of prior estimates of the variables j that is left free in the inversion. V_{sim}^j is the set of free variables that was used in the RTM to simulate R_{sim}^i . The covariance matrix of the variables (COV_{var}) accounts for collinearity between the variables and weights the distance between the estimated variables and the prior information. There is no agreement on the weights the radiometric and a priori parts in the cost function should receive. While Combal et al. (2002a) used separate weights for each variable left free during inversion, others exploited radiometric and a priori information in two successive steps (Koetz et al., 2005a).

A priori information is also introduced in LUT and ANN inversions when the upper and lower boundaries and the sampling scheme of the variables used to generate the LUT and the training database are chosen.

In general, well-defined and possibly restricted variable ranges will enhance variable retrieval, whereas broad ranges lead to a vaguer prior estimate and increased uncertainties as expressed by the covariance matrix (Baret and Buis, 2007). Therefore, various authors use classification schemes to split the problem into sub-domains for which prior information is attributed separately (Knyazikhin et al., 1999a; Chen et al., 2002). Nevertheless, this approach may introduce problems due to misclassification and attribution errors and lead to artefacts at the limit between classes translating into more chaotic spatial or temporal variation of the solution (Lotsch et al., 2003). Moreover, many of the biomes found at the Earth surface constitute a mixture of vegetation species, for which rather a gradual transition than a crisp classification applies.

Including spatial information So far, the discussion was limited to the inversion for single spectra. Atzberger (2004) proposed to include the neighborhood radiometric information of a pixel, to get what he called an “object signature”. The neighborhood data can either be extracted from moving windows or land cover maps. Particularly for LAI, results were significantly better, which he attributed to a reduction of the confounding effects between LAI and leaf angle distribution typical for inversion approaches based on single spectra (Atzberger, 2004; Jacquemoud et al., 1995a).

The method proposed by Atzberger (2004) only includes statistical distributions within a given object. Methods including additional geo-statistical constraints could bear additional information, particularly for imagery with higher spatial resolutions (Garrigues et al., 2006a).

Temporal constraints Dynamics in vegetation canopies, and hence reflectance, occur under the influence of changing meteorologic, climatologic, and environmental conditions and take place at different time scales ranging from diurnal and seasonal cycles to long-term gradual impacts. As such changes are rarely of abrupt nature (with the exception of accidental events such as fire, flooding, harvesting, or lodging), the smooth character of the dynamics of canopy variables may be exploited as additional constraint in the retrieval process (Baret and Buis, 2007).

Temporal constraints can be introduced either by directly using the series of multi-temporal observations (CROMA, 2002) or by using additional models which mimic the evolution of one or more variables over time. Last category ranges from simple semi-empirical approaches (Koetz et al., 2005a) to advanced soil-vegetation-atmosphere (SVAT) models where the evolution of the canopy status is modeled in detail and remote sensing and SVAT model predictions are intercalibrated (Verhoef and Bach, 2003b; Launay and Guerif, 2005; Guérif and Duke, 2000). When a model is employed that introduces additional free variables, the undetermined nature of the problem will only reduce if the number of multi-temporal remote sensing observations exceeds the number of newly introduced variables of the SVAT model and are well-spread across the observation period (Delécolle et al., 1992; Launay and Guerif, 2005).

2.6 Conclusions

This chapter presented an overview of the methods currently in use for the retrieval of canopy biochemical and biophysical variables from solar-reflective remote sensing data. As the observed spectral reflectance is the result of many scattering and absorbing processes in the canopy and

boundary layers, a clear characterization of these processes is the premise for an adequate and stable retrieval of the variables of interest. This can only be done against a strong physical background such as offered by radiative transfer models.

The performance of radiative transfer model inversion greatly relies on the ability of the chosen model to represent the observed canopy. Although our ability of modeling radiance propagation in complex canopies has considerably evolved during the last decades, using such models in the inverse mode does not necessarily lead to improved variable estimations. A more realistic representation of canopy architecture requires additional input variables, which, in cases where the ranges of these variables are not adequately known, will even increase the under-determination of the inverse problem. A compromise should therefore be found between the model's capability in describing radiation propagation and the degree of freedom allowed by the remote sensing observation(s) and ancillary data sources.

Regularization techniques, aiming at reducing the ill-posedness and under-determined nature of model inversion, intrinsically change the dimensionality of the problem, either by increasing the number of independent information sources, or by restricting the number of free variables or the limits in which the variables are allowed to vary. Although imaging spectroscopy is generally believed to contain additional information content compared to systems with a limited number of spectral bands, especially on biochemical composition of the leaves, only few studies explicitly quantify this profit (Bacour et al., 2002a; Verhoef, 2007).

During the last few years, the scientific community repeatedly pointed out the potential of the combined use of reflective information and signatures originating from the angular (Widlowski et al., 2004; Knyazikhin et al., 1999b; Chopping et al., 2003), the spatial (Atzberger, 2004), or temporal (Koetz et al., 2005a) dimension in improving variable retrievals. Also the synergy between spectral measurements and independent observations of different nature, such as LIDAR (Koetz et al., 2007) and RADAR (Treuhaft et al., 2002; Treuhaft et al., 2004) holds a promise.

Even if additional information dimensions are very appealing for improved variable estimation, the actual information content has to be clearly addressed, especially in cases where high correlation exists between the different information sources. Addressing the high dimensionality and regularizing ill-posed model inversions will be an even more challenging task in the context of an automated approach, like the one presented in the next chapter, when very little additional information on the observed surface is a priori known.

Chapter 3

The automated CRASh approach: theoretical concept and implementation

3.1 Introduction

An automated, image based retrieval approach based on radiative transfer modeling should find a careful balance between the lacking knowledge on land cover and phenological conditions on one hand, and accurate retrieval performances for the various agricultural land cover types encountered in the scene, at the other. The poorly available a priori information on surface conditions hampers a strict definition of canopy variables input to the radiative transfer model and will accordingly exacerbate the ill-posedness of the inversion. The challenge of the automated approach is therefore to exploit the full information content contained in the image itself by considering not only the spectral dimension, but also by taking advantage of the information offered by the other pixels in the scene.

In the previous chapter it was already pointed out that a successful radiative transfer model inversion depends on three major decisions: i) the choice of the radiative transfer models used to characterize leaf and canopy reflectance, ii) the choice of a suited inversion approach, and iii) the definition of the final solution. While Section 3.2 addresses the choice of an appropriate canopy reflectance model, the inversion approach is discussed in Section 3.3. The definition of the retrieval procedure itself is tackled in Section 3.4.

The ensemble of automated processing steps that is introduced in this thesis was entitled CRASh (Canopy variable Retrieval Algorithm based on PROSPECT and SAILh). The basic processing flow of the automated CRASh procedure is schematically depicted in Figure 3.1. Input is the atmospherically corrected top-of-canopy (TOC) reflectance image which is first submitted to a spectral classification based on the SPECL module (Richter, 2007a)(Section 3.3.2). The land cover classification forms the backbone of the radiative transfer model (RTM) inversion as the classes of spectral similarity allow a description of uncertainty associated to the measurements and, to some degree, a refinement of the variable ranges considered during model inversion (Section 3.3.3).

Although land cover classification reduces the number of multiple solutions, it only partly overcomes the ill-posedness of the inversion, and may even give rise to additional drawbacks

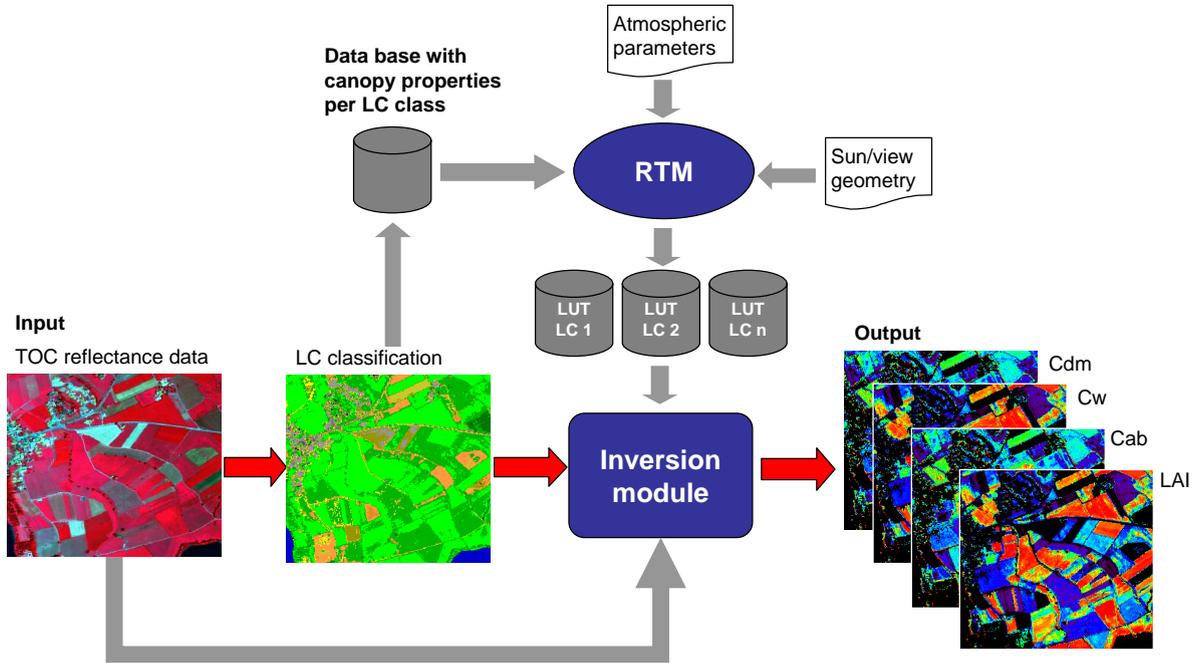


Figure 3.1: Overview of CRASH inversion approach. TOC = Top Of Canopy, LC = Land Cover, RTM = radiative transfer model, Cdm = leaf dry matter content, Cw = leaf water content, Cab = leaf dry matter content, LAI = leaf area index, θ_v = view zenith angle, θ_s = sun zenith angle, ψ_{rel} = relative azimuth. Atmospheric parameters include those that were configured during atmospheric correction. See text for a detailed explanation of single components.

associated to classification errors. For this reason, in Section 3.4.2 a new method for an image based retrieval of a priori information on the variables is introduced.

In addition to the use of prior information, two additional information sources are exploited for their potential in regularizing the inverse problem, i.e., the spatial neighborhood and, if available, the angular anisotropy of the observation (Section 3.4.4). In the following, the above mentioned components of the inversion approach will be introduced step by step, while summarizing existing knowledge and proposing new solutions to common problems.

3.2 Choosing an appropriate radiative transfer model

An appropriate canopy reflectance model has to fulfill the following requirements:

1. In the inverse mode, it must be able to deliver the products the user is interested in, being either primary variables (i.e., variables that are input to the model) or secondary variables (variables that are no direct model inputs, but can be calculated based on these).
2. It must be able to accurately describe the reflectance behavior of the canopy of interest. That is, the parametrization of the model must be able to describe the optical properties

of the various vegetation elements, and the complexity and spatial heterogeneity of the canopy.

3. The variables left free during inversion should have significant influence on canopy reflectance and should possibly be restricted in number (Pinty et al., 2004a; Verstraete et al., 1996). Since in an operational environment only little a priori knowledge is present on the exact distribution of the input variables, increasing the number of free variables will lead to a substantial rise in multiple solutions.

In the framework of this study, the choice of the model was based on the capability of retrieving the variables leaf area index, leaf chlorophyll content, leaf water content, and organic dry matter. The complete canopy reflectance model used in this study is composed of three individual sub-models, accounting for leaf optical properties, canopy architecture, and background reflectance, respectively.

3.2.1 The leaf optical model PROSPECT

In order to be able to estimate leaf chlorophyll, leaf water, and leaf dry matter content, these variables should be part of the leaf optical model or should at least have an unambiguous relationship with its constituents. The commonly used PROSPECT model (Jacquemoud and Baret, 1990; Fourty et al., 1996) explicitly describes the absorption of these three components and therefore appears to be a good candidate for modeling the optical leaf properties.

3.2.1.1 N compact layers

PROSPECT describes the leaf as a pile of N compact layers with specific absorbing and refractive properties (Figure 3.2). The intercellular air spaces which are responsible for multiple scattering within a leaf and are thus the driving force of high NIR reflectance and transmittance (Figure 3.3b), are represented by the $N-1$ empty spaces between the layers (Bacour, 2001).

The leaf mesophyll structure parameter N is not necessarily an integer but can be any real value within a realistic range. It varies between 1.0 for most compact leaves (such as young cereal leaves) up to 4.5 for thick leaves with well developed spongy mesophyll, or for senescent leaves having disorganized mesophyll with a large amount of air spaces (Baret et al., 2005b; Jacquemoud et al., 1996) (Table 3.1).

3.2.1.2 Biochemical components

The overall absorption of a leaf is the sum of the absorption caused by the single leaf biochemicals contained in PROSPECT plus a rest term accounting for absorption by chemicals that are not included in the PROSPECT model. The absorbing elements are assumed to be uniformly distributed across the leaf and their contribution to overall absorption is calculated by multiplying their concentration per unit area with an absorption coefficient. These absorption coefficients are specific for every biochemical and vary over wavelength (Figure 3.3a). The current version of PROSPECT considers 4 specific absorbing materials:

3.2. Choosing an appropriate radiative transfer model

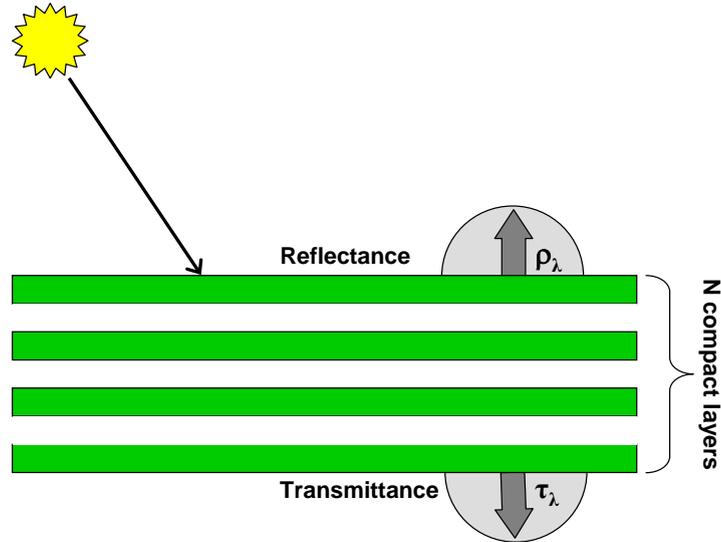


Figure 3.2: Schematic representation of PROSPECT N-layer radiative transfer model for leaves (Adapted from: Jacquemoud and Ustin (2001))

Table 3.1: Structure variable N , leaf chlorophyll content C_{ab} , dry matter content C_{dm} , and leaf water content C_w for a selection of agricultural crops. N and C_w are given for fresh (left) and dry (right) leaves (Jacquemoud et al., 1996; Bacour, 2001)

<i>Species</i>	N		$C_{ab}[\mu g \cdot cm^{-2}]$	$C_{dm}[g \cdot cm^{-2}]$	$C_w[g \cdot cm^{-2}]$	
<i>Monocotyledons</i>						
Sorghum (grass)	1.49	-	39.03	0.0038	0.0125	-
Maize	1.42	2.41	58.52	0.0045	0.0122	0.0002
Rice	2.03	3.6	30.25	0.0068	0.0089	0.0005
<i>Dicotyledons</i>						
Clover	1.87	-	46.71	0.0030	0.0100	-
Salad	1.92	2.11	35.20	0.0021	0.0231	0.0002
Alfalfa	1.54	1.66	40.52	0.0028	0.0107	0.0001
Potato	1.78	1.98	43.37	0.0026	0.0170	0.0001
Sun flower	1.72	2.63	53.23	0.0038	0.0191	0.0002
Tomato	1.66	1.62	20.21	0.0026	0.0184	0.0001
Vine	1.53	2.00	16.92	0.0056	0.0121	0.0003

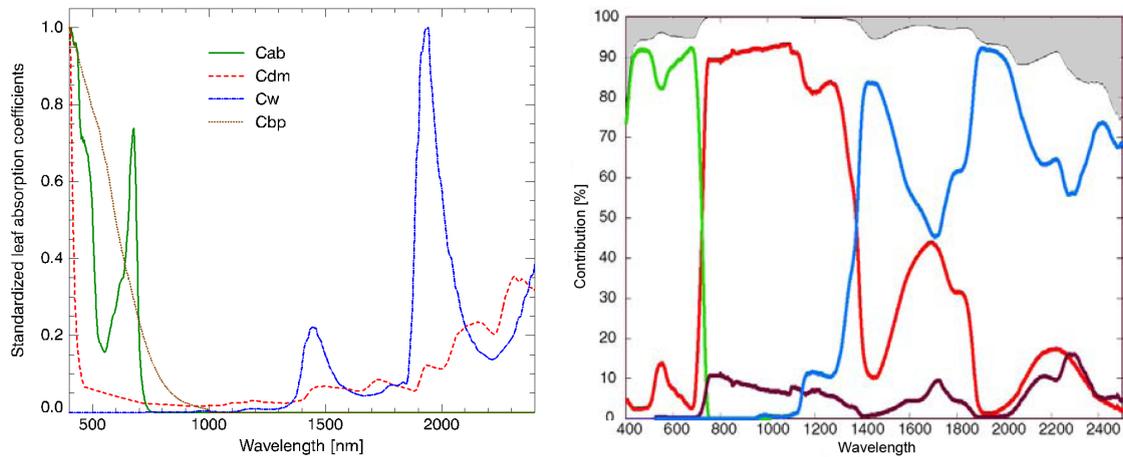


Figure 3.3: Standardized PROSPECT absorption coefficients for leaf chlorophyll, leaf dry matter, leaf water, and leaf brown pigment content. For each variable the contribution relative to the maximum absorption coefficient is given (left); Relative contribution of different leaf absorbing materials to leaf transmittance modeled with PROSPECT. Green corresponds to chlorophyll concentration, red to structure parameter N , blue stands for water content, and brown for dry matter content. The gray curve indicates the sum of contributions, the shaded area the part remaining unexplained by PROSPECT (Pavan, unpublished; op.cit. Ustin et al. (2005)) (right)

- **Leaf chlorophyll content (Cab)** [$\mu g \cdot cm^{-2}$] actually corresponds to the content of chlorophyll a, chlorophyll b, and carotenoids, as the strong correlation between these components makes it very difficult to distinguish them based on remote sensing measurements (Sims and Gamon, 2002). Cab normally varies between 0 and $100 \mu g \cdot cm^{-2}$ although green leaves typically have values greater than $15 \mu g \cdot cm^{-2}$ (Baret et al., 2005b).
- **Dry matter content (Cdm)** [$g \cdot cm^{-2}$], also called specific leaf weight, is the inverse of the specific leaf area (SLA), a unit often used by ecophysiologicals, and is directly related to above ground (foliage) biomass when multiplied with the leaf area index. Cdm is actually the sum of the carbon based absorbing materials cellulose, hemi-cellulose, lignin, starch, and protein which were combined into a single variable when the distinct absorption features appeared to be too insignificant to allow a retrieval of the individual components from fresh leaves (Baret and Fourty, 1997a; Fourty and Baret, 1997). It typically varies from 0.002 up to $0.02 g \cdot cm^{-2}$ and absorbs over the whole reflective domain, while its relative importance compared to other absorbing materials is highest in the NIR and SWIR-2. In vigorous leaves, the specific absorption features of Cdm are usually masked by those of leaf water.
- **Leaf water content (Cw)** [$g \cdot cm^{-2}$], or equivalent water thickness (Danson et al., 1992), dominates absorption in the SWIR and generally ranges from 0.01 to $0.03 g \cdot cm^{-2}$. Multiplying it with leaf area gives the total foliage water, a quantity that can be easily measured. Several studies proposed to approximate relative water content to a value close to 80% for green leaves, and 20% for senescent leaves, allowing to establish a fixed rapport of 4:1 between Cw and Cdm (Baret et al., 2005b) and thus to reduce the number of free variables in RTM inversion.

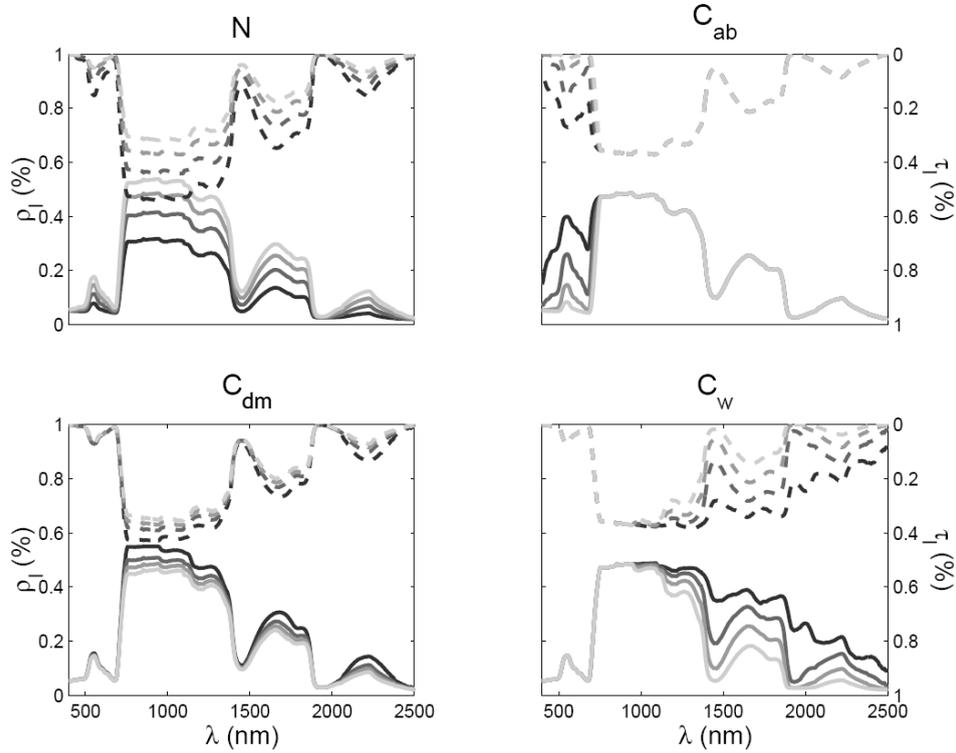


Figure 3.4: Effects of N , C_{ab} , C_w , C_{dm} , and C_{bp} on leaf reflectance (continuous lines) and transmittance (dashed lines), modeled with PROSPECT, based on the average set of parameters: $N = 1.7$; $C_{ab} = 50 \mu\text{g} \cdot \text{cm}^{-2}$; $C_{dm} = 0.015 \text{ g} \cdot \text{cm}^{-2}$; $C_w = 0.025 \text{ g} \cdot \text{cm}^{-2}$; $C_{bp} = 0.001$. The variables used in the simulation for the different leaf variables are respectively $N = [1 ; 1.5 ; 2 ; 2.5]$, $C_{ab} = [5 ; 20 ; 50 ; 100]$, $C_{dm} = [0.002 ; 0.010 ; 0.015 ; 0.02]$, and $C_w = [0.001 ; 0.010 ; 0.025 ; 0.05]$. Dark colors correspond to the lowest concentrations, bright colors to the highest (Bacour, 2001).

- **Leaf brown pigment content (C_{bp})** [relative units] increases at senescence of a leaf. C_{bp} and C_{ab} are spatially exclusive which means that they never appear at the same part of the leaf (Baret et al., 2002). C_{bp} typically varies from 0 for green leaves up to 3.5 for senescent dark brown leaves.

Figure 3.3b shows for an average leaf the wavelength dependent influence of each PROSPECT variable to leaf transmittance (Pavan, unpublished; op.cit. Ustin et al. (2005)). Figure 3.4 demonstrates the effect of the natural variation of each variable on leaf reflectance. Some clear trends can be distinguished from these two figures: the VIS is dominated by C_{ab} absorption, whereas N dominates scattering in the NIR plateau. The gradient of the red-edge, the sharp rise in reflectance between VIS and NIR, can be affected by the content of brown pigments (not shown). In the SWIR, leaf optical properties are dominated by C_w and to a lesser extent by N . In fresh leaves, like the ones simulated in the lower figure, the effects of C_{dm} are largely masked out by the previous two components, insinuating that a retrieval of this parameter by model inversion may be difficult for fresh vegetation.

3.2.1.3 Simplifications and limitations

Intrinsic to every model is the simplified description of processes compared to reality. The PROSPECT model is no exception to this. First of all, it starts from the assumption that optical properties of both leaf faces are equal, an assumption that in many situations obviously violates reality. In section 2.2.1 it was indeed shown that chlorophyll absorption mainly takes place in the parenchyma close to the leaf surface, intending that a leaf cannot be considered symmetrical. Secondly, the PROSPECT model considers the leaf as a Lambertian surface, an approximation that is valid in all view directions, except for the specular direction (Sanz et al., 1997). However, mainly due to the simplified parametrization and because of the limited number of input arguments, the model appears a good candidate for model inversion when little is known on the distribution of biochemicals. Several studies performed at leaf level (Fourty and Baret, 1997; Jacquemoud et al., 1996) and at canopy level in combination with a canopy structure model (Jacquemoud et al., 1995a; Jacquemoud et al., 2000; Koetz et al., 2005a; Weiss et al., 2000; Zarco-Tejada et al., 2005b) demonstrated the ability of the model in accurately modeling leaf scattering and absorption.

3.2.2 The canopy reflectance model SAILh

In Chapter 2 it was already pointed out that the use of complex canopy radiative transfer models, though accurately describing radiation behavior in the canopy, may enhance the ill-posed and under-determined nature of the inverse problem when input variables cannot be accurately characterized. In this light it would be appropriate to prefer a model that requires relatively few input variables while still being able to accurately describe the canopy radiation budget. For homogeneous agricultural crops, one-dimensional turbid medium models broadly fulfill this requirement. Widlowski et al. (2005) even state that when both a 1-D and complex 3-D models account for all features of the measured radiance field, then - in absence of further information regarding the nature and structure of the target - the use of the latter may even lead to an over-interpolation of the available data and therefore appear inappropriate. An additional benefit of using a 1-D model is the relatively low computation time involved.

Several 1-D turbid medium models based on the radiative transfer equation have been proposed (Kuusk, 1995b; Kuusk, 1995a; Verhoef, 1984; Gobron et al., 1997). The choice fell on SAILh (Verhoef, 1984; Verhoef, 1985), based on intercomparison results (Jacquemoud et al., 2000; Bacour et al., 2002a) and the satisfying results obtained by the remote sensing community in a wide range of applications. Moreover, model results for heterogeneous scenes lie close to those obtained with complex 3-D models for standard situations, although for observations in the principle plane reflectance values are generally higher than those simulated with 3-D models (Pinty et al., 2004b).

3.2.2.1 4-stream approximation

SAILh (Scattering by Arbitrarily Inclined Leaves; h stands for the incorporation of the hot spot function which will be discussed later) is based on a four stream approximation (Verhoef, 1985; Verhoef, 1997) in which the TOC radiance regime is separated into four major fluxes (Figure

3.5): i) a direct incoming flux from the source of illumination E_s , ii) a diffuse hemispherical downward flux E^- , iii) a diffuse hemispherical upward flux E^+ , and iv) the radiance flux in the direction of the observer E_o . All fluxes depend on wavelength.

For calculating the different fluxes at each wavelength, a set of 4 differential equation has to be solved in relation to the vertical position z in the canopy:

$$dE_s/dz = kE_s \quad (3.1)$$

$$dE^-/dz = -sE_s + aE^- - \sigma E^+ \quad (3.2)$$

$$dE^+/dz = s'E_s + \sigma E^- aE^+ \quad (3.3)$$

$$dE_o/dz = w'E_s + v'E^- + vE^+ - KE_o \quad (3.4)$$

where k is the extinction coefficient for the directional incoming flux, and s (s') the backward (forward) scatter coefficient for the directional flux contributing to E^- (E^+); a (σ) is the forward (backward) scattering coefficient for the downward (upward) hemispherical flux. K is the extinction coefficient in the viewing direction, w the bidirectional scattering coefficient, and v (v') the backward (forward) scattering coefficient for E^- (E^+). The definition of the coefficients is completely regulated by the structural properties of the vegetation canopy, the optical properties of leaves (reflectance, transmittance, and absorption) and soil, and the constellation of illumination source and observation. For a detailed description of the different fluxes the reader is referred to (Verhoef, 1997).

In the case of atmospherically corrected data, the flux that is finally observed by the sensor E_o is the sum of the fractions of direct and diffuse incoming radiation that is reflected by canopy and soil in the direction of the sensor (=HDRF). Therefore, in order to be able to compare modeled with measured HDRF, the diffuse incoming radiance component has to be characterized. In SAILh, this so-called *skylight* component (*SKYL*) is expressed as the wavelength dependent fraction of diffuse and global irradiance and depends on atmospheric composition, solar elevation, and altitude.

3.2.2.2 Canopy characterization

The original SAIL model (Verhoef, 1984) is a pure turbid medium model, which considers the canopy a horizontal infinite layer which in the vertical direction is bounded by the atmosphere (top) and the soil (bottom). Leaves are the only scattering and absorbing elements, having an infinitely small physical size and exhibiting a horizontally and vertically uniform random distribution. However, with the incorporation of the hot spot parameter, which is related to canopy height and leaf width, leaves obtained a physical size. Therefore, strictly spoken, SAILh (Verhoef, 1997) is a hybrid model, even if it is mostly still denoted as turbid medium (Verhoef and Bach, 2007). Average spectral properties of the leaves are described by the hemispherical transmittance and reflectance modeled with PROSPECT model. The amount of radiance reflected in a certain direction is the combined action of extinction, absorption, and scattering within the canopy, depending on *LAI*, leaf angle distribution, and the hot spot parameter.

The leaf angle distribution (LAD) describes the average distribution of leaf orientation and inclination. In SAILh *leaf orientation* (azimuth) is considered uniform. In contrast, *leaf inclina-*

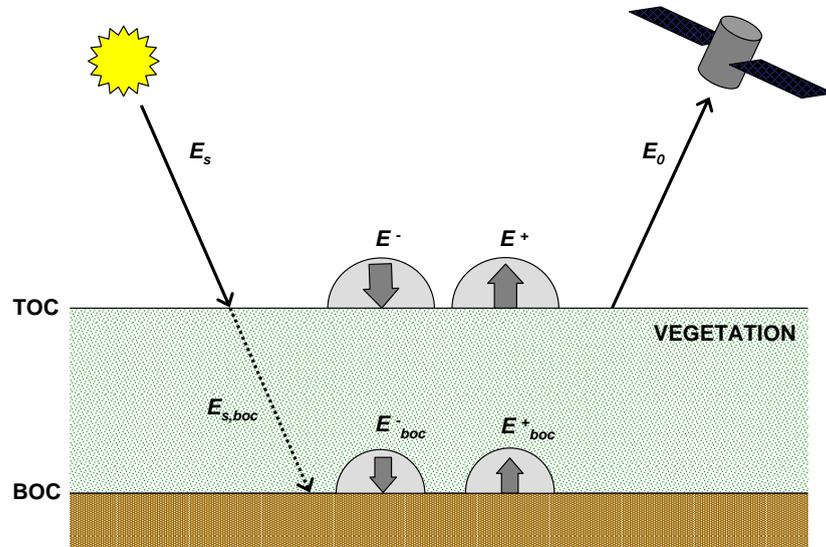


Figure 3.5: 4-stream radiation fluxes used to approximate the top of canopy (TOC) and bottom of canopy (BOC) radiance regimes. See text for explanation. The suffix $_{boc}$ is used to denote the radiance fluxes at the bottom of the canopy. Note that soil reflectance is assumed Lambertian (Adapted from Verhoef and Bach (2003b))

tion is highly variable. Leaf inclination distributions are commonly categorized into six typical distributions (Bunnik, 1978; De Wit, 1965; Verhoef, 1997):

- planophile : majority of leaves has a horizontal orientation,
- erectophile : majority of leaves has a vertical orientation,
- plagiophile : leaf inclination is centered around 45° ,
- extremophile : leaf inclination is concentrated around the two extremes, 0° and 90° ,
- uniform : probability of finding a particular leaf angle is equal for all inclinations,
- spherical : occurrence probability is as if the leaves were placed on sphere, with a higher probability found at extremere inclination angles.

Several mathematical formulations have been proposed to describe leaf inclination, including trigonometrical (Bunnik, 1978), beta (Goel and Strebel, 1984), elliptical (Kuusk, 1995a), and ellipsoidal functions (Campbell, 1986; Campbell, 1990). Verhoef (1984; 1997) proposed a number of 10-12 typical leaf inclination distribution functions (LIDFs) which are based on a graphical method and described by a combination of two coefficients (Figure 3.6, upper part). For the current study it was decided to replace these discrete distributions with the smooth ellipsoidal function proposed by Campbell (1990)(Figure 3.6, bottom). The advantage of this formulation is that it requires only one input parameter, the average leaf angle (ALA), and thus reduces the number of free input variables in the inversion. Nevertheless, not all distributions as proposed by Verhoef can be exactly reconstructed using the ALA (e.g., the plagiophile distribution in Figure 3.6). The elipsoidal formulation of leaf inclination distribution in SAILh has however been successfully implemented by several other authors (Baret et al., 2005b; Jacquemoud et al.,

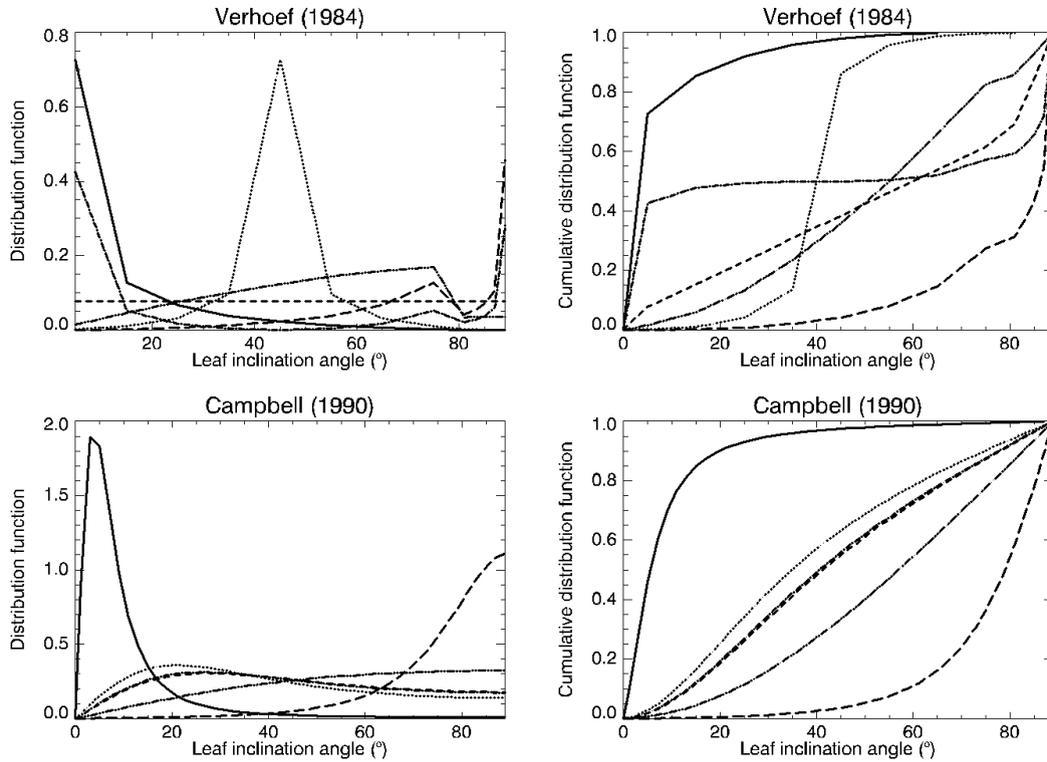


Figure 3.6: Graphical (Verhoef, 1984; Verhoef, 1997) versus ellipsoidal (Campbell, 1986; Campbell, 1990) formulation of LIDF, for absolute (left) and cumulative (right) density distribution functions, respectively. The following distribution functions are shown: planophile (solid line), plagiophile (dotted), uniform (dashed), extremophile (dash dot), spherical (dash dot dot), and erectophile (long dashes)

2000; Koetz et al., 2005a) and forward RTM simulations show satisfying agreement compared to the parametrization proposed by Verhoef, except for very extreme leaf inclination or view angles (Verhoef, personal communication).

3.2.2.3 The hot spot effect

In pure turbid media, in absence of geometrical shading and gaps, the hot spot effect is exclusively caused by volume scattering (Qin and Xiang, 1994; Widlowski et al., 2005). In SAILh where leaves have a physical size but no geometric shading is accounted for, the hot spot effect is approximated by the bi-directional gap probability based on the theory of Kuusk (1985). Its angular width is directly related to the typical sizes of the individual scatterers and can be regulated by the hot spot size parameter (HS) which is equal to the ratio of the correlation length of leaf projection in the horizontal plane and the canopy height. In order to reduce the number of variables during model inversion Verhoef and Bach (2003a) proposed to slave the hot spot parameter to canopy LAI by the equation $hs = 0.5/LAI$, hence describing crops that tend to grow mainly in vertical direction by addition of new leaves rather than by growth of existing leaves. In the automated mode however, HS is left free since no preferred canopy

growth direction can be assumed in advance.

3.2.2.4 Assumptions and caveats

The choice of SAILh was mainly based on its relatively simple parametrization and its proven capability of simulating radiance propagation in homogeneous canopy types. However, despite the more or less justified assumption of a turbid medium, the use of SAILh for modeling agricultural crops brings about a few concerns:

- *Horizontal homogeneity*: some crops, like maize or cotton, exhibit a row-like distribution of the plants, which is present especially during early growth. For these situations a hybrid model that physically describes these rows would be more appropriate. Nevertheless, some studies obtained satisfying results with SAILh even for row crops (Koetz et al., 2005a). A row extension to the SAIL model (RowSAIL) was developed in the context of the CROMA project (CROMA, 2002), but this version did not account for the hot spot effect and up to date has not been made public. Horizontal homogeneity also has to be seen within the context of spatial resolution of the data. With increasing resolution, horizontal discontinuity plays a major role since the principle of SAILh is based on the average properties of a pixel and does not account for horizontal fluxes. In very high resolution data of heterogeneous canopies the pixels can thus not be longer considered radiatively independent units (Pinty et al., 2004a; Widlowski et al., 2006a).
- *Vertical homogeneity*: at plant level, leaves often show a non-homogeneous distribution. Clumping of leaves of shoots may occur and the fraction of senescent and vital leaves depends on the vertical position within the canopy. The youngest members of the SAIL family do account for non-homogeneous vertical distributions of leaves (leaf color gradients) and clumping effects (Verhoef and Bach, 2003b; Verhoef and Bach, 2007), but up-to-date these versions have not been made public.
- *Scattering and absorbing elements*: in SAILh, leaves constitute the only scattering and absorbing canopy elements. In certain phenological stages however, flowers, stems, branches, twigs, fruits, and ears may be even more prominent than leaves. Distinguishing between leaves and other elements is appropriate since they possess quite different absorption properties. For this reason, in certain development stages, especially at flowering and ripening when flowers, fruits or ears appear at the top of the canopy, the use of SAILh might be inappropriate.

Although in some situations the use of the original SAILh version may lead to some oversimplification, it was nevertheless preferred for the automated approach. Based on the fact that row structures (directions and height), crop type, and phenological development stage are not a priori known and in most situations cannot be recognized from the image itself, this would lead to increased uncertainty in model parametrization. The variables finally entering the SAILh model are summarized in Table 3.2.

Table 3.2: Input parameters used in SAILh. * Soil reflectance and the soil brightness parameter bs are discussed in Section 3.2.3

LAI	Leaf Area Index	$[m^2/m^2]$
ALA	Average Leaf Angle	$[^\circ]$
ρ_{leaf}	Leaf reflectance calculated with PROSPECT	-
τ_{leaf}	Leaf transmittance calculated with PROSPECT	-
hs	Hot spot parameter	-
ρ_{soil}	Soil reflectance *	-
bs	Soil brightness parameter *	-
$SKYL$	Fraction of diffuse irradiance	-
θ_s	Solar zenith angle	$[^\circ, \text{rad}]$
θ_v	Observation zenith angle	$[^\circ, \text{rad}]$
ψ_{rel}	Relative azimuth angle	$[^\circ, \text{rad}]$

3.2.3 The soil background model

The lower boundary layer of the SAILh model is constituted by the optical properties of the soil background. Actually, the background may also comprise non-green materials found in the bottom layer of the canopy, including litter and non-photosynthetic vegetation and which are spectrally very similar to soil (Nagler et al., 2000). In contrast, green (and hence absorbing) understory is considered part of the green vegetation canopy. The background reflectance signal depends on illumination and view geometry, surface roughness, mineral composition, the amount of dead organic matter, and soil moisture content.

The soil reflectance parametrization used in this study is based on the soil line concept, which allows to synthesize the effect of geometrical conditions, roughness and moisture within a single parameter which is assumed independent of wavelength (Baret et al., 2005b; Weiss et al., 2002; Bacour et al., 2002a). Using this concept, the background reflectance $R_b(\lambda, \Omega_i, H_j, z_k)$ for any wavelength λ , observation geometrical configuration Ω_i , moisture H_j and roughness z_k is assumed proportional to the reflectance background for the same wavelength λ but with different observation geometrical configuration Ω_l , moisture H_m and roughness z_n (Baret et al., 2005b):

$$R_b(\lambda, \Omega_i, H_j, z_k) = BS \cdot \rho_b(\lambda, \Omega_l, H_m, z_n) \quad (3.5)$$

where BS is the brightness parameter that does not depend on wavelength λ , but on all other factors (Ω, H, z) . Representing background spectral variation by a multiplication coefficient implies Lambert background scattering, an assumption that might be criticized, especially in situations with low vegetation cover (Verhoef and Bach, 2007). However, in the case of the automated approach presented in this thesis, average background reflectance and variations are calculated from the image itself and, except for the multi-angular case, represent a fixed view/sun geometry. BS has thus only to account for variations in surface roughness and soil moist.

An additional point of concern may be the fact that soil moist does not induce linear spectral shifts that are similar for all wavebands. However, most earth observation systems do not cover the spectral regions covered by the strongest water absorption bands. Moreover, in the VIS, the domain with highest spectral contrast between soil and vegetation, no specific water absorption features are present (Verhoef and Bach, 2007). Therefore, in this region soil moist variations induce a quasi linear spectral shift that is equal for all wavebands (Bacour et al., 2006; Baret et al., 2005a). The general validity of the brightness concept in standard situations (dry soil, medium roughness, no hot spot configuration) has been confirmed in several preceding studies (Bacour et al., 2002a; Koetz et al., 2005a; Weiss et al., 2002). A similar soil brightness concept has also been implemented for the grass and crop biomes of the MODIS algorithm (Knyazikhin et al., 1999c).

3.2.4 Interaction of variables

Like already emphasized, several variable combinations can induce very similar spectral effects thus giving rise to ill-posed nature of RTM inversion. Probably the strongest bias present in the concurrent inversion of PROSPECT and SAILh is introduced by the opposite effect on canopy reflectance of LAI and ALA (Figure 3.7). Various authors report the spectral similarity of a sparse canopy with planophile leaf orientation and a dense, erectophile canopy (Atzberger, 2004; Baret and Guyot, 1991; Jacquemoud, 1993; Jacquemoud et al., 1995a). Regularization of this interaction is difficult due to the generally poor a priori knowledge on the distribution of both variables. Up to date, only very few studies are known that properly address leaf angle distribution¹.

Partial compensation takes place between LAI and Cab in the VIS domain (Bacour et al., 2002a; Baret and Buis, 2007; Jacquemoud et al., 1995a), which particularly for sensors covering only the VNIR could lead to bias in the estimates. Many other variable interactions, though less pronounced, should be accounted for, such as the opposite effect of N and average leaf angle (Jacquemoud, 1993), the counteracting of LAI and Cw in the SWIR (Jacquemoud et al., 1995a), and the ambiguity between LAI , Cdm , and N . Regularization of these interactions can be induced by a sharp definition of the expected estimates or by combining two strongly interacting variables into a synthetic variable. Last mentioned is often done by combining one of the absorbing leaf materials (Cab , Cdm , Cab) with LAI (Combal et al., 2002b; Fourty and Baret, 1997; Jacquemoud et al., 1995a; Koetz et al., 2004; Vohland and Jarmer, 2007).

3.3 The inversion approach

3.3.1 Justifying a LUT based inversion scheme

Different inversion approaches have been proposed for the retrieval of canopy variables by radiative transfer model inversion, all having their specific advantages and disadvantages (Section

¹A good example is given by the FIFE experiment conducted by the The Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC): <http://daacsti.ornl.gov/FIFE/FIFE.Home.html>

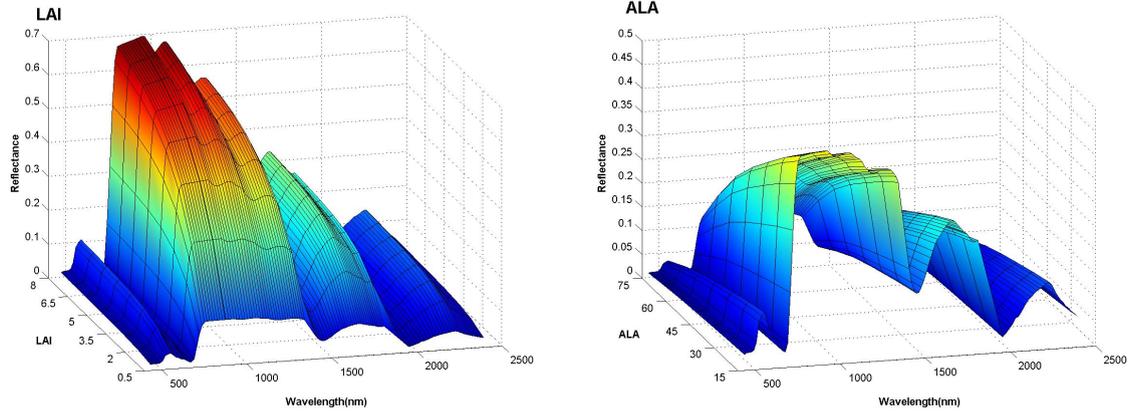


Figure 3.7: Ambiguity between the canopy structure variables LAI (left) and ALA (right). Changing their values shows an opposite effect on reflectance. The spectra have been modeled with PROSPECT and SAILh, based on the average set of parameters: $N = 1.7$; $Cab = 50 \mu g \cdot cm^{-2}$; $Cdm = 0.010 g \cdot cm^{-2}$; $Cw = 0.025 g \cdot cm^{-2}$; $Cbp = 0.01$; $ALA = 50^\circ$, $LAI = 3$, $BS = 0$, $HS = 0.15$, $\theta_s = 30^\circ$, $\theta_v = 0^\circ$, $\psi_{rel} = 0^\circ$

2.5.3). In the context of a regional, image based inversion scheme, such as presented in this study, the choice of the inversion approach is driven by the following prerequisites:

1. The inversion algorithm should be fast,
2. optimized for, or easily adaptable to, scene specific view and sun geometry,
3. easily adaptable to site-specific crop types and phenology and open to a priori information available from the user site,
4. and applicable to a wide variety of hyper- and multispectral, being able to provide consistent data products that depend as little as possible on the specifications of the used sensor.

An inversion scheme based on LUTs appears the best compromise for this combination of requirements. Even if for iterative optimization approaches it is slightly less problematic to incorporate changing view/sun geometry, band characteristics, and phenology, their iterative nature makes them computationally too demanding for large data sets. On the other hand, artificial neural networks which - once trained - are very fast, are optimized for specific view/sun constellations, canopy types, and sensor specifications, and would have to be trained again if best possible results want to be obtained for new conditions. Moreover, the large number of database entries in the LUT allows for the definition of preliminary or priori estimates, as will be seen in Section 3.4.2.1.

As stated in Chapter 2, the use of LUTs also brings along some drawbacks, such as the large number of LUT entries required to cover all possible canopy realizations, the increased processing time coinciding with large synthetic data sets and hyperspectral information content, and the dependence of the solution on the configuration (number and distribution of samples) of the LUT and the cost function used to select the samples from the LUT forming the end result.

For a successful retrieval, these issues have to be implicitly accounted for. In the rest of this chapter these issues will be addressed one-by-one.

3.3.2 Land cover classification for improved retrieval performance

Possessing knowledge on vegetation type or the phenological stage of a certain crop allows one to confine the range of possible solutions to the actual conditions, thus reducing the ill-posedness of the inverse problem. For this reason, the current study employs an inversion scheme based on land cover classification, where the LUT-inversion is performed for each land cover class separately.

For the LUT approach, a land cover based inversion implies some important advantages compared to a global scheme: apart from reducing the number of ill-posed solutions, limiting the range of variation of the variables either leads to i) a reduced number of LUT entries, and hence reduced processing requirements, when the sampling *distance* is kept constant, or ii) an intensified sampling density in cases where the *number* of samples per variable is kept constant (Figure 3.8). Moreover, the use of land cover classes facilitates a less vague description of the covariance characterizing the uncertainties and collinearity in radiometric and a priori information (Section 3.4). Finally, a land cover classification permits the user to retrieve canopy variables for the land cover types of interest only, thus reducing calculation time.

In ideal situations, a land cover or crop map is already available, either generated by supervised classification or field mapping. However, in most situations, up-to-date information on land cover and land type is not at hand, and an automated classification should be provided by the inversion model itself. The CRASh program offers both options, whereas in the following the discussion will be confined to the automated approach.

The complication underlying an automatic spectral classification based on a single image, is the fact that crop type and phenologic conditions actually present in a scene are not known in advance and no information on the seasonal spectral course of a crop, which could make classification more reliable (Houborg et al., 2007; Houborg and Boegh, In press; Strahler et al., 1999), is present. On the other hand, class definitions that are too restrictive or specialized are more sensitive to misclassifications (e.g., induced by variations in spectral brightness or errors originating from preprocessing), which in turn may lead to significant accuracy loss in the retrieval process, or to artefacts at class transitions which translate into more chaotic spatial variations in the solution (Lotsch et al., 2003). For this reason, it was decided to apply a less definite land cover classification based on some general reflectance characteristics of the pixel under observation and on partly overlapping classes.

Land cover classification is based on the SPECL module which is integrated in ATCOR (Richter, 2007a). The SPECL code performs a spectral classification of the reflectance cube based upon template spectra for the Landsat Thematic Mapper reference wavelengths (i.e., 0.48, 0.56, 0.66, 0.83, 1.6, 2.2 μm) and returns an image with class indices. The template spectra consist of typical vegetation covers, soil, sand, infrastructure, and water. If at the reference wavelengths the spectral reflectance signature agrees within a 10% margin with one of the class template spectra it is put into this class, otherwise it belongs to the class undefined. The decision rules used by SPECL can be found in Appendix A. The classes containing a green vegetation fraction are (Figure 3.9):

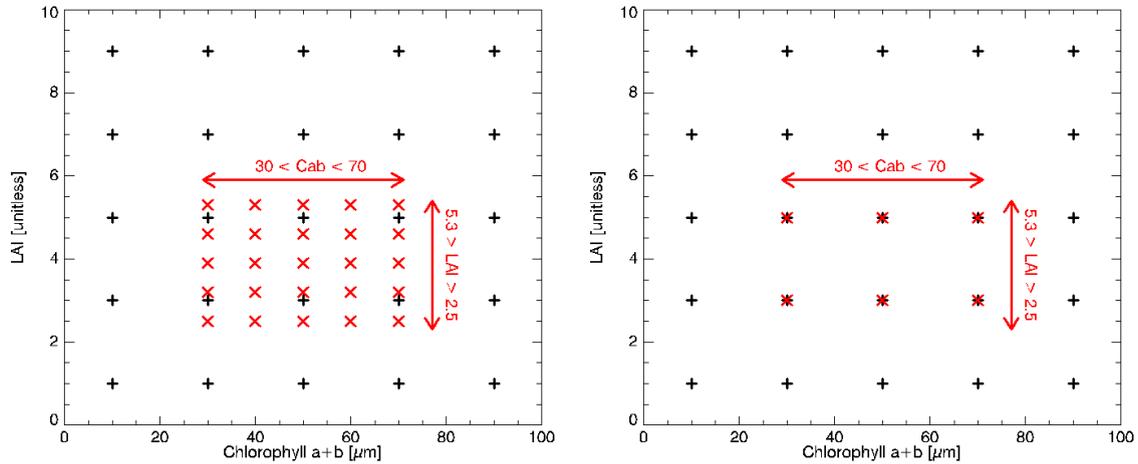


Figure 3.8: Effect of posing lower and upper limits (e.g., by a land cover classification) on number and sampling density of variables in LUT: the black crosses represent a regular sampling over the full range of Cab and LAI that is commonly found within a scene. Confining the range of variation for a certain land cover type either leads to an intensified sampling density if the number of samples is kept constant (left), or a reduction of the total number of sample points (right)

- *dark vegetation*, containing green vegetation with low NIR reflectance, usually induced by a large shading component due to vertically oriented canopy structures. Mainly coniferous tree canopies fall into this category, but also agricultural land use with heterogeneous medium scale canopy or erectophile structures such as vine yards and horticultures. This class has a strong spectral similarity with canopies composed of a mix of soil and vegetation: for this reason the class should also partly describe canopy types of this kind,
- *average vegetation*, principally containing green crops and grasses in intermediate phenological stages and having medium to high NIR reflectance,
- *bright vegetation* mainly including green crops and grasses in full development (high NIR reflectance),
- *yellow vegetation*, accounting for healthy green vegetation (high NIR reflectance) containing flourishing yellow flowers (low green and red absorption), such as rapeseed or sunflower. This is a problematic class since optical properties of flowers are not included in the PROSPECT model,
- *sparse vegetation/soil*, being a class where the soil signature is dominant but which still contains a fraction of photosynthetic active vegetation or green crop residues,
- *mix of vegetation and soil*, representing a land cover type where green vegetation does not yet fully cover the underlying soil background, so primarily crops in early development stages, but also cut meadows or harvested green crops,
- *dry vegetation/soil*, encompassing vegetation with a high content of senescent material, such as cut meadows or mature cereals prior to harvest. This class has very high spectral similarity to soil signatures (Figure 3.9).

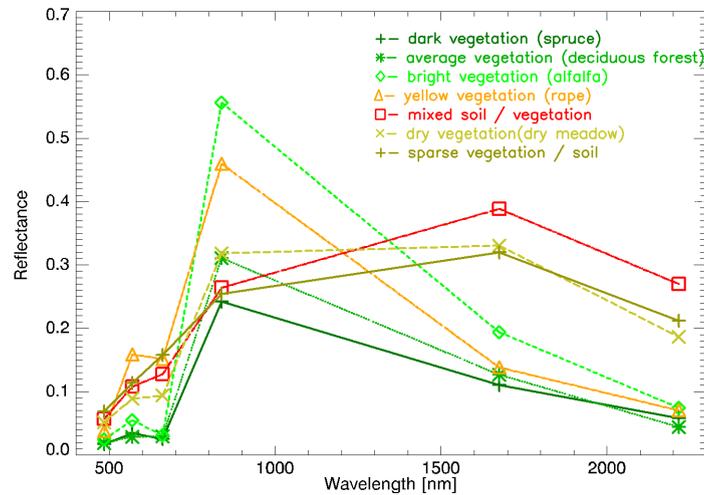


Figure 3.9: Template nadir vegetation reflectance spectra employed by the SPECL code (Richter, 2007a)

Beisl (2001) observed that, even if the spectral criteria used in SPECL may be susceptible to changes in reflectance intensity (e.g., due to angular effects), it still outperformed the commonly used spectral angle mapper (Kruse et al., 1993) based on the same classes.

It was already emphasized that splitting the whole domain into a set of sub-domains may introduce problems due to misclassification and attribution errors. Even if a jump between classes and, hence, a sharp change in predicted parameters is justified for two adjacent fields, such classification steps may not occur at field level. This problem is schematically illustrated in Figure 3.10, where a hyperspectral image (left) is first submitted to a SPECL classification (center) and successively to a land cover based inversion scheme (right). The spectral similarity of the pixels within the central field is depicted by the spectra drawn in the plot below the image. Despite their spectral similarity, several pixels (some examples are found in the in red circles) were attributed to classes different from the prevailing *average vegetation* class. Since the ranges of variables for each of these classes are defined in a different way (See next Section), this could have an impact on the estimated values if these misclassifications are not properly accounted for (e.g. *LAI*). In this study, a new method is presented to reduce the influence of classification errors on the variable retrieval, and which is based on the incorporation of local neighborhood information in the definition of the covariance matrices describing the uncertainty in measurements and a priori information. This will be introduced in Section 3.4.2.

3.3.3 Considerations underlying lookup table parametrization

In the previous section we have seen that a land cover based inversion scheme breaks the inverse problem down from a global inversion schedule into a set of sub problems, each optimized for a specific class. Even if the number of LUT entries can be significantly reduced when the LUT is optimized for the canopy properties found in the specific land cover class, the risk of getting a LUT that is either too large, or that undersamples specific parts of the variable space, continues

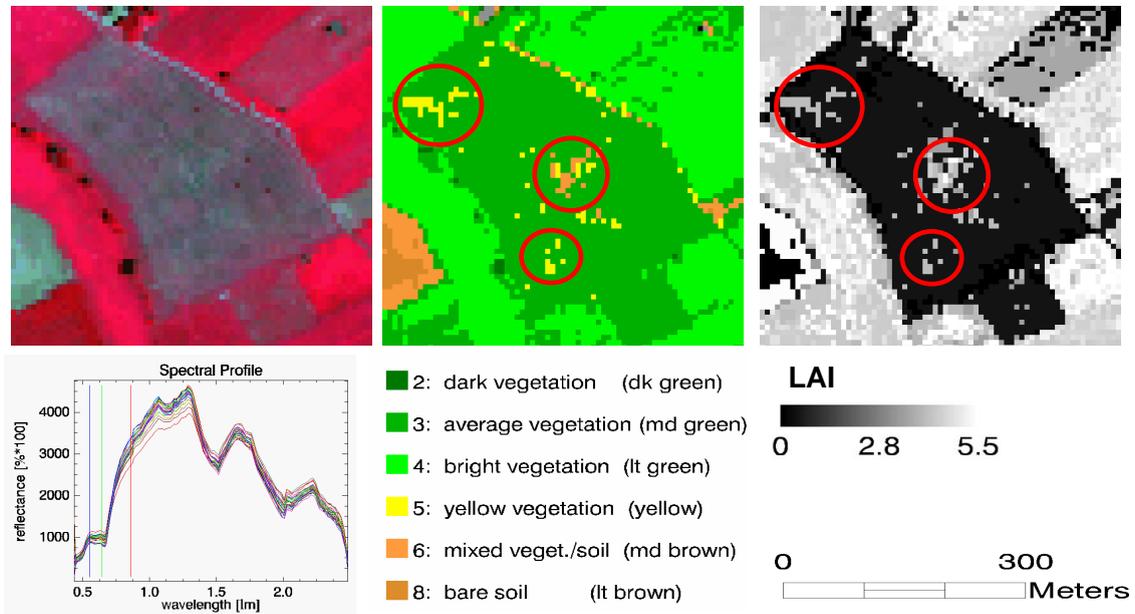


Figure 3.10: Effect of classification anomalies on variable retrieval pattern. A hyperspectral image (left) is first submitted to a SPECL classification (center) and successively to a land cover based inversion scheme (right). Although pixels have a very close spectral similarity (left below), retrieved LAI in the misclassified pixels strongly deviates from their neighborhood pixels

to exist. Last concern is unfortunately too often disregarded. Considering the composition of the LUT, four aspects have to be taken into account: i) the upper and lower boundaries of the variable ranges, ii) the distribution of the variables, iii) the number of samples/drawings per variable, and, directly related to this, iv) the sampling interval. Moreover, atmospheric conditions and soil characterization should be representative for the image under consideration.

3.3.3.1 Variable ranges

The variable ranges should be set in such way that all possible canopy realizations within the land cover class are covered but tight enough to allow for a significant reduction of the range compared to the global range of solutions. This may be less trivial than it seems, especially in the case of an automated procedure such as the one presented here, since only little might be known on the land cover type or the phenological stage actually presented by the different pixels attributed to a specific class. For example, the class *bright vegetation* can contain different canopy types, ranging from cereals and grasses to maize and sugar beets, each having their particular biochemical and structural characteristics. Besides, a certain overlap between different classes must exist, so that in case of a misclassification - which is probable to occur to spectra at the class extremes - the pixel is still covered by the class-specific LUT. The variable ranges for the vegetation classes used in this study are listed in Appendix B, together with the argumentation used to define these ranges. The ranges have been selected based on expert knowledge and a range of publications (Atzberger, 2004; Combal et al., 2002a; Hosgood et al., 1995; Koetz

et al., 2005a; Weiss et al., 2000; Weiss and Baret, 1999).

3.3.3.2 Variable distribution functions

Once the variable ranges are set, one has to decide according to which distribution (e.g. uniform, gaussian) the variables are sampled between the lower and the upper limit. In literature no agreement exists about the distribution rule that should be employed. Although some studies emphasize that the space of canopy realizations has to be sampled in such a way that surface response is represented, i.e., with denser sampling in those ranges where sensitivity of reflectance to specific canopy variables is higher (Combal et al., 2002b; Weiss et al., 2000; Baret and Buis, 2007), most authors practice probability distributions that mimic the actual distribution (Atzberger, 2004; Koetz et al., 2005a). Notice that the choice of a particular distribution function already contains a certain amount of a priori knowledge (Section 3.4.2).

Based on literature review and expert knowledge, it was decided to use different sampling distributions for different variables. Cw , and Cdm were sampled according to a uniform distribution between lower and upper limit, since very little is known about their actual distributions. Cbp , N , the hot spot HS , and soil brightness parameter BS were sampled assuming a Gaussian distribution with class mean μ and standard deviation σ that vary among the classes. The Gaussian distribution allows for denser sampling in ranges that are more likely to occur. Finally, Cab and LAI were sampled using the distribution functions proposed by (Combal et al., 2002b) which emphasize those domains where canopy reflectance is more sensitive to variations in the considered variable (Equation 3.6 and 3.7). This is a justified decision if one wants to cover the complete spectral variation.

$$Cab_{trans} = e^{Cab/100} \quad (3.6)$$

$$LAI_{trans} = e^{-LAI/2} \quad (3.7)$$

3.3.3.3 Sampling strategy

The most common way of choosing a variable set for a forward model run, is the combination of variables that are randomly selected according to a specific distribution function for each variable. Such an approach may lead to undersampling of certain ranges, while oversampling others. This phenomenon is illustrated by Figure 3.11a where the sampling of 9 free variables (which is the common number of variables used for the combination of PROSPECT and SAILh) is assumed. Only 2 dimensions (Cab and LAI) are shown. If we distribute the selections equally through the variable space and for each variable we consider 5 different values (black crosses), a structural combination of all possibilities will lead to a LUT of $5^9 \approx 2$ million canopy realizations. If we have a similar situation (e.g. 2 million canopy realizations) based on random drawings from a uniform distribution, the sampling density in a 2-D space may look like the pattern demarcated by the red crosses in the same figure, in which clear zones of under- and oversampling are visible.

In order to avoid areas of under- and oversampling, a sampling method based on a full orthogonal scheme is proposed (Bacour et al., 2002a; Bacour et al., 2006). Using such a stratified

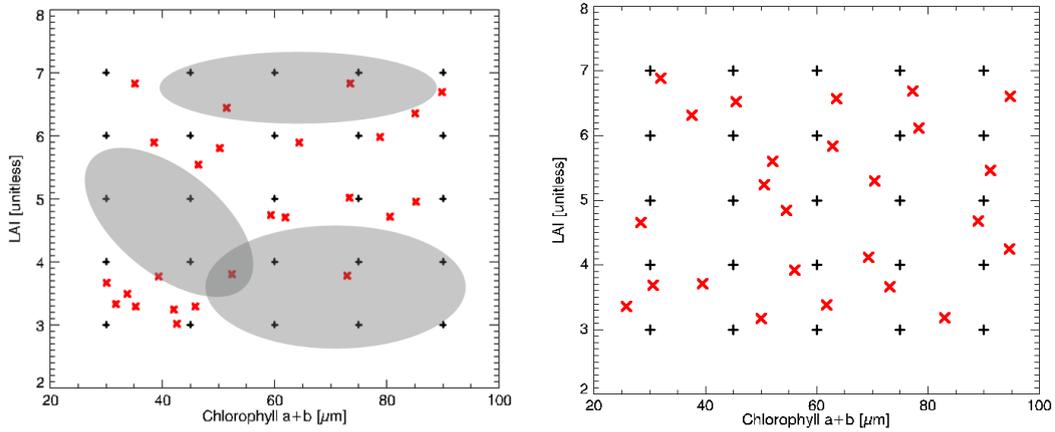


Figure 3.11: Sampling in a 2-D variable space resulting from global (left) and stratified (right) random variable selection: $25 \leq Cab \leq 95$, $2 \leq LAI < 7$. The scenario represents a case where a total of 2 million canopy realizations is generated for 9 free variables. The shaded ellipses in (a) roughly indicate the zones of undersampling when a global random sampling scheme is used.

scheme, the range of variation of each variable is split into a given number of intervals from which samples are drawn on a random basis. All combinations of intervals are sampled once. This procedure allows accounting for all the interactions between variables while preserving their natural random variation (Figure 3.11b).

3.3.3.4 Sampling distance

Based on the previous, still the number of intervals for the stratified scheme has to be defined. In order to keep the processing time within reasonable limits, we assume as a rule of thumb that the total size of the LUT should not exceed 300,000 entries. With 9 free variables this would lead to a total number of $\sqrt[9]{300000} \approx 4$ intervals (data points) per variable. Although for variables with little sensitivity (i.e. BS for a dense canopy) this may be appropriate, for variables exhibiting large influence on the spectrum (e.g. LAI) this will not be sufficient and lead to undersampling. Therefore, a land cover class specific attribution scheme is applied, in which the sampling intervals depend on the sensitivity and range of the parameters in a specific class. The considerations that were taken into account for the definition of the classes were based on expert knowledge and literature reviews and can be found in Appendix B.

3.3.3.5 Incorporating view and sun geometry

The LUTs have to be adapted to the view/sun constellation present in the remote sensing imagery. The solar geometry (ψ_s and θ_s) depends on the geographical location of the scene (longitude, latitude) and the day of the year and can be adopted from the preceding atmospheric correction, where these parameters have to be specified as well.

The inclusion of view geometry evolves in a different way for satellite and airborne sensors. Common high resolution satellite sensors such as Landsat TM, have a field of view (FOV) of

only a few degrees, and view zenith and azimuth vary only very little across the scene. In this case, the use of a single LUT for the whole scene is justified. In the case of multi-angular satellite observations (e.g. CHRIS-PROBA) a single LUT is created for every view angle (Chapter 5).

For airborne high resolution imagery however, FOV may exceed 60° , giving rise to considerable anisotropy within a single image, specifically for observations in the principle plane and at high solar elevations (Cfr. Figure 2.9). To account for such anisotropy within a single image, the HDRF is sampled at an angular view interval of 3° . For view/sun configurations with ψ_{rel} within $\pm 30^\circ$ to the principle plane, this is done at a sampling interval of 1° . The sampling is performed for both azimuth directions (which equals the flight heading $\pm 90^\circ$), separately.

For large FOV data usually a parametric georectification is employed, providing view zenith and azimuth angle of every single pixel in the form of a scan angle file (Section 4.2.3.3). This file can be employed in the RTM approach for the coupling of the generated LUTs to the relevant pixels in the scene. In cases where such information is lacking, the scan angle file is simulated, based on average flight heading, the FOV of the sensor, and the number of pixels per line.

3.3.3.6 Calculating diffuse radiance

In this work, radiative transfer model (RTM) inversion is based on the spectral similarity of measured and modeled top-of-canopy HDRF in the direction of the sensor (Section 2.2.4). To calculate HDRF in the forward simulations, the fraction of diffuse hemispherical incoming radiation (*SKYL*) has to be known.

The fraction of diffuse radiance strongly depends on atmospheric composition and solar elevation. As far as it concerns the automatic approach, an atmospheric correction has been carried out in advance with ATCOR (Cfr. Section 4.2.3.4). The atmospheric parameters that were appointed during this operation (solar zenith, average ground elevation, visibility, aerosol type, and water vapor column) are used for the calculation of the fraction of diffuse radiation based on the MODTRAN4 code implemented in ATCOR (Berk et al., 2003; Richter, 2007a). In situations where atmospheric parameters are unknown, a standard atmosphere (visibility = 35 km, ground elevation = 0.5 km, water vapor column = 4 cm, rural aerosol type) is assumed and a standard diffuse fraction for the complete reflective domain is calculated (Figure 3.12). The figure shows that particularly in the VIS the skylight fraction is high. Not properly accounting for the skylight fraction may therefore considerably affect C_{ab} retrieval (Schaepman-Strub et al., 2006).

3.3.3.7 Background reflectance parametrization

The regionally based approach employed in this study allows retrieving background reflectance and its spectral variation directly from the scene itself by means of the pixels that are allocated to the various soil classes during the automatic land cover classification (Section 3.3.2). This allows optimizing background reflectance for local soil conditions and composition, and for the prevailing sun/observation geometry. After a filter procedure for the removal of noisy and falsely classified pixels, scene average background reflectance is calculated from all pixels that were assigned to one of the soil classes in the image (Figure 3.13). As the reflectance of litter and crop residues are very similar to that of soil (Nagler et al., 2000), this is automatically contained in

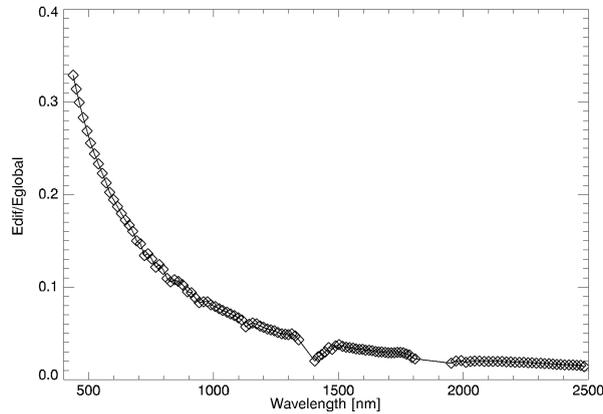


Figure 3.12: Diffuse fraction of incident radiance calculated for a solar zenith of 30° and a standard atmosphere: visibility = 35 km, ground elevation = 0.5 km, water vapor column = 4 cm, rural aerosol type.

the SPECL soil classes, and aggregated in the background reflectance characterization. Some of the soil spectra might also contain a small fraction of green vegetation (e.g., at crop germination or early leaf stages).

The validity of using a single brightness parameter to describe spectral variation as a function of average background reflectance is shown in Figure 3.14. The left plot shows 5 different HyMap spectra (Chapter 4) that all have been attributed to one of the soil classes. The red spectrum represents the average scene reflectance, calculated from all soil pixels in the image. The right plot shows the spectra of the left plot divided by the average spectrum. In the ideal case this should provide a straight line for each spectrum, with the y-value representing the brightness value. For most spectra this assumption holds reasonably well, with the exception of spectrum "soil 4", which, as observed in the shape of the original spectrum, features some fraction of photosynthetic active vegetation. In general however, the use of single brightness coefficient is reasonable well able to describe variation in soil reflectance. For the parametrization of the LUTs, a different range of brightness variation is used for the single land cover classes, since it is assumed that soil moist and surface roughness vary as well among the different land cover types (Appendix B).

Figure 3.15 gives an indication of the spectral variation found within and between two different vegetation classes. Parametrization of the simulations were based on one of the HyMap scenes discussed in Chapter 4.

3.4 The optimization algorithm

This section deals with the inversion algorithm itself, i.e., with the way in which, based on an input spectrum, entries are selected from the LUT in order to calculate the final result. The solution of the RTM inversion has to fulfill two conditions: first of all, the modeled HDRF has to lie close to the one that was measured by the earth observation system. As already pointed out in previous sections, diverging combinations of vegetation variables may render a

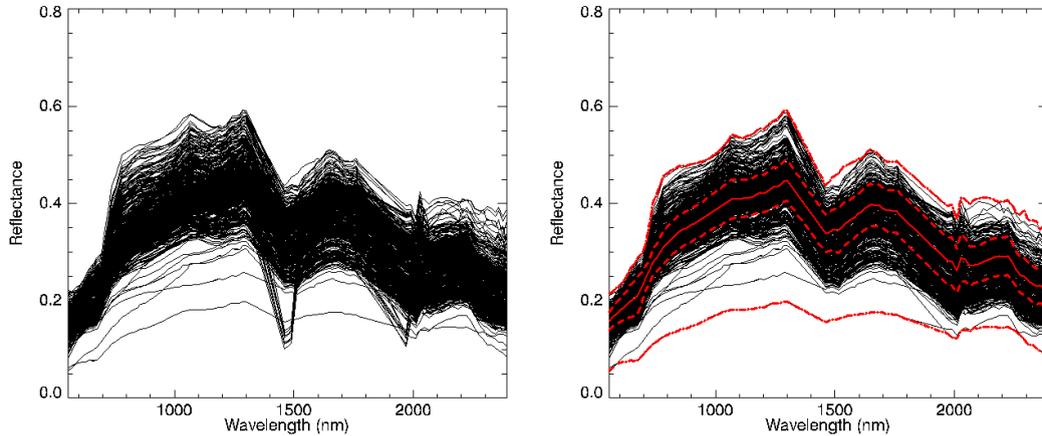


Figure 3.13: Scene based soil characterization: soil spectra are collected from the scene, based on a preceding land cover classification (left). Noisy and falsely classified spectra are removed and average background characteristics are expressed by: average background reflectance (solid), standard deviation (dashed), and min/max reflectance (dotted) (right)

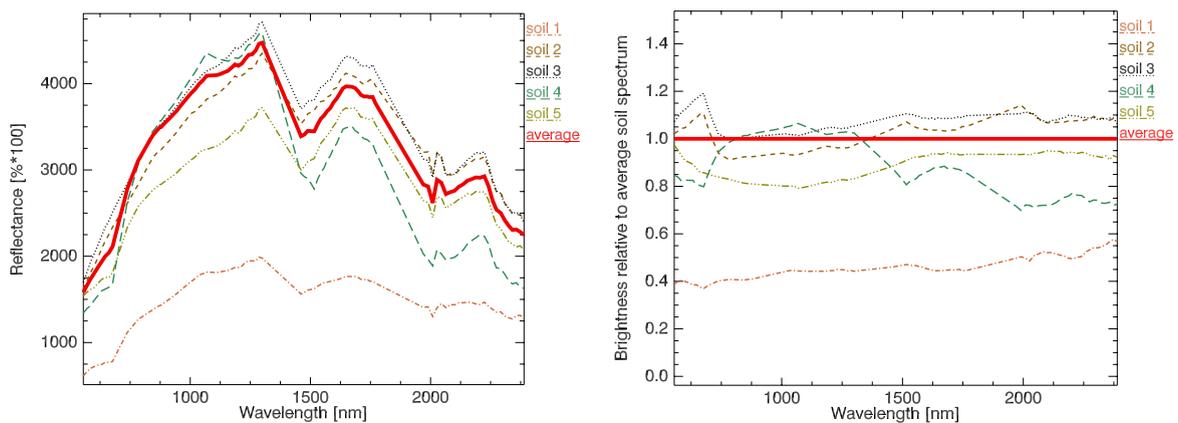


Figure 3.14: Evidence for the soil brightness parameter. The left graph shows 5 selected HyMap spectra encountered in the imagery used in Chapter 4. The red curve shows average soil reflectance as calculated from all soil pixels in the image. Dividing the individual spectra by the scene average spectrum results in the plot depicted at the right. The average y-value of the spectra denotes their approximate brightness with respect to the average soil spectrum

very similar signal, thus giving rise to the ill-posedness of the inverse problem. For this reason, a second condition should be met, i.e., the retrieved variables should fall within plausible margins around the expected, or *a priori*, solution (second term of the right-hand side of Equation 2.12). Last condition is not so straightforward as it may sound, since in an automated approach little information is available about the canopy under observation and therefore the solution can only be roughly approximated. For this reason, many authors usually ignore the second condition, or at most pose upper and lower boundaries to the expected range of variation, and reduce their inversion to a radiometric minimization problem.

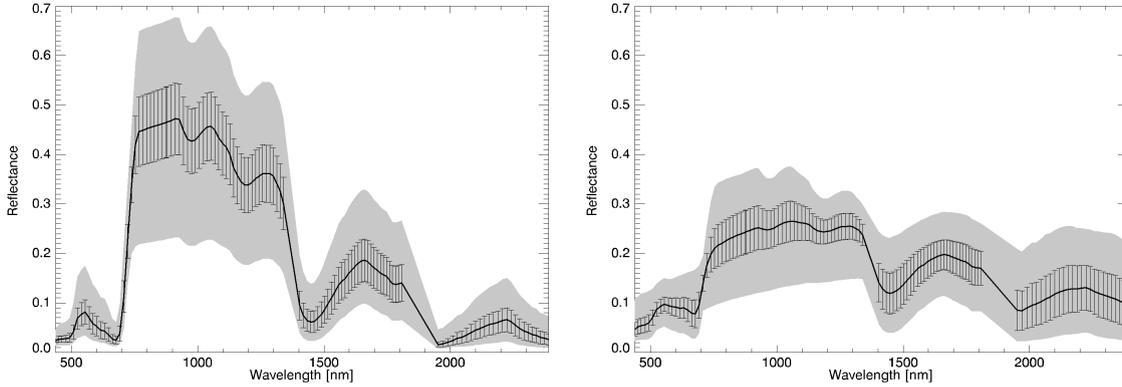


Figure 3.15: Range of reflectances simulated for the classes *bright* (a) and *sparse vegetation* (b). Simulations were based on the sampling schemes found in Appendix B, a solar zenith of 45° and a nadir viewing sensor position. The solid lines indicate average class reflectance, error bars the standard deviation of the simulations, and the shaded area the range of maximum variability.

In this section, a new method for the definition of the second term of the right hand side of Equation 2.12 is proposed. As we will see in Section 3.4.2, the prior estimates partly rely on the data itself, the reason for which it would be incorrect to speak of *a priori* knowledge in a strict sense. Since it rather involves an additional regularization step, it was decided to abandon common theoretical perception where both parts of the cost function (Equation 2.12) are simultaneously explored and weighted according to the respective covariance matrices (Lavergne et al., 2007; Rodgers, 2000; Tarantola, 2005), and instead to explore radiometric information and a priori information in two successive steps (Figure 3.16). The successive exploration of radiometric and a priori information has been successfully implemented by other authors as well (Combal et al., 2002a; Koetz et al., 2005a). Local spatial variance of radiometric values and a priori estimates were introduced in both steps to cope with classification anomalies, as explained later.

3.4.1 Exploiting radiometric information

To accelerate the inversion algorithm, only those LUT entries are considered that for each band have a reflectance value within $\pm 20\%$ of the input spectrum. A range of 40% is considered broad enough to include all measurement and model errors. However, in order to prevent any ambiguity, e.g. in bands with a low signal-to-noise ratio or bands affected by strong atmospheric scattering, only the Landsat TM bands are considered. In addition, for wavebands with a reflectance lower than 0.1 (10%) an absolute range of ± 0.02 (2%) is employed:

$$\widehat{LUT} \in \{LUT_c | 0.8 \cdot R_{meas} \leq R_{LUT} \leq 1.2 \cdot R_{meas}\} \quad (3.8)$$

where LUT_c is the initial LUT generated for class c , R_{meas} the measured reflectance, and R_{LUT} the reflectance spectra contained in LUT_c . If less than 30 LUT-spectra meet above criterion,

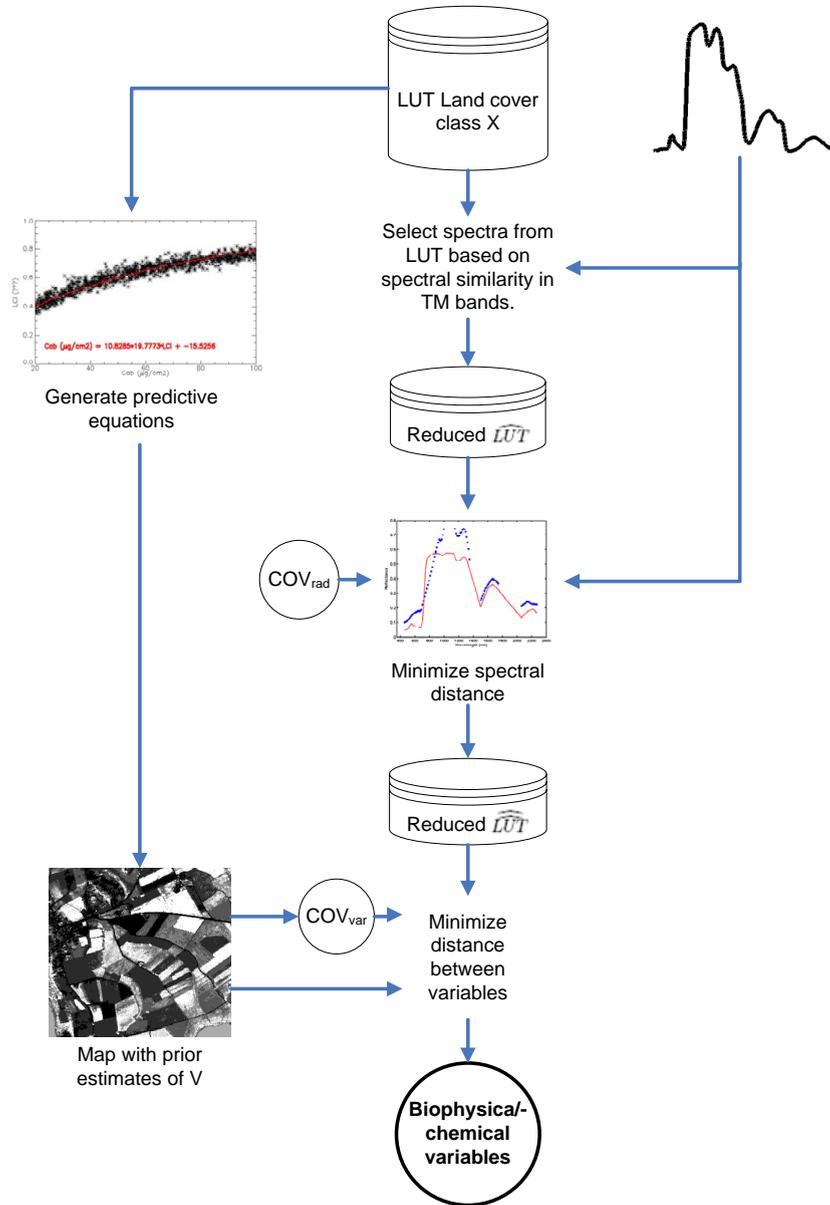


Figure 3.16: Optimization algorithm proposed in this study

the range is automatically extended to $\pm 50\%$, and, if again this range appears too restrictive, eventually all LUT spectra are included. This usually only occurs to misclassified pixels or to spectra that are constituted of mixed signals, and if such happens, this is marked by quality flag (Section 3.4.3).

Based on the original wavebands, the reduced LUT is sorted according to the cost function χ_{rad}^2 corresponding to the weighted spectral distance between the measured reflectance R_{meas} and the simulated reflectance R_{LUT} found in \widehat{LUT} :

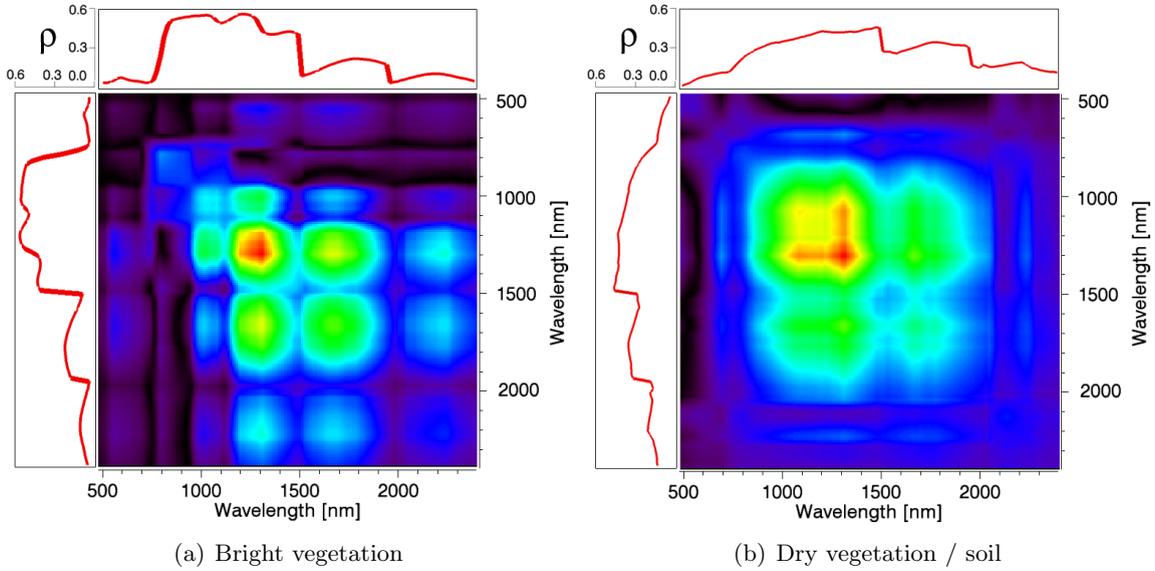


Figure 3.17: Examples showing variance-covariance matrices for two SPECL land cover classes. Dark colors indicate low (co-)variance, red colors high values. The calculation is based on HyMap data of the Waging-Taching area (Chapter 4). On the top/left of the figure a characteristic spectrum of each class is shown.

$$\chi_{rad}^2 = (R_{meas} - R_{\widehat{LUT}})^T \cdot COV_{rad}^{-1} \cdot (R_{meas} - R_{\widehat{LUT}}) \quad (3.9)$$

COV_{rad} , the matrix describing the covariance between the different spectral bands, should account for non-systematic uncertainties associated to the measurements (noise, geocoding, atmospheric correction, classification) and to the radiative transfer model. The covariance between wavebands is not known in advance since it changes with target properties, illumination conditions, and so on. But, using the land cover classification based on spectral similarity measures, we can approximate COV_{rad} by calculating the matrix from the spectra contained within each land cover class. This results in a different covariance matrix for each land cover class (Figure 3.17). It should be noticed however, that in this way only non-systematic radiometric uncertainties are accounted for, thus disobeying possible systematic errors in model assumptions and parameterizations.

Of the sorted LUT, the possible solutions considered were those that were within 20% of the best absolute radiometric match, leading to a reduced LUT denoted as $\widehat{\widehat{LUT}}$. The threshold of 20% is consistent with what other authors proposed in earlier studies (Combal et al., 2002b; Koetz et al., 2005a).

3.4.1.1 Adding local spectral variance to overcome classification anomalies

Intermediary model test results revealed that in the proposed inversion approach systematic errors (e.g. by RTM assumptions and parametrization) are in many cases larger than the normally distributed random uncertainties of a certain object as captured by the covariance

matrix. In cases of large systematic errors, a too strict definition of radiometric uncertainties may therefore lead to poor retrieval results. In Figure 3.10 it was already shown that a land cover classification that splits the whole spectral domain into a set of sub-domains may introduce problems due to attribution errors which in turn are translated into crisp jumps in the retrieval results. The covariance matrices that are constructed for the separate classes are unable to compensate for such larger local systematic uncertainties. To overcome such local systematic deviations originating from attribution anomalies, additional variance is added to the covariance matrix. This radiometric variance is calculated from the spectral values in a 5×5 pixels window around the currently inverted pixel. This local spectral variance (VAR_{loc}) is added to the class specific covariance matrix COV_{rad}^c to constitute the covariance matrix that is finally implemented in the radiometric minimization function (Equation 3.9):

$$COV_{rad} = COV_{rad}^c + VAR_{loc} \quad (3.10)$$

or in matrix notation:

$$COV_{rad} = \begin{pmatrix} var_{cla}^0 & \cdots & cov_{cla}^{0n} \\ \vdots & \ddots & \vdots \\ cov_{cla}^{n0} & \cdots & var_{cla}^n \end{pmatrix} + \begin{pmatrix} var_{loc}^0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & var_{loc}^n \end{pmatrix} \quad (3.11)$$

var_{cla}^n denotes the class variance, var_{loc}^n the local variance of waveband n . cov_{cla}^{nm} is the class specific covariance between band n and m . If the local neighborhood of the pixel shows strong heterogeneity, the variance contained on the diagonal of the covariance matrix is increased, whereas, in case of a homogeneous local neighborhood, there will be barely any changes compared to the class specific covariance matrix COV_{rad}^c . In cases where pixels lack a spatial context, which occurs when the neighboring pixels belong to a class that is excluded from processing, the second part of the right hand term in Equation 3.9 is omitted and the class specific covariance term COV_{rad}^c is employed for COV_{rad} instead.

3.4.1.2 Covariance matrix inversion

One of the major drawbacks of using imaging spectrometer data in the proposed way is of numerical nature. The high degree of linear dependency between bands, especially in cases where the spectra are smoothed, e.g. during atmospheric correction, is reflected in COV_{rad} , provoking numerical problems when this matrix has to be inverted. A matrix is not invertible or *singular* when its determinant is zero. To find out if COV_{rad} is invertible, the determinant is calculated using LU decomposition, a procedure for decomposing an $N \times N$ matrix into a product of a lower triangular matrix L and an upper triangular matrix U. This numerical operation is performed using the Linear Algebra Package (LAPACK) library implemented in IDL (ITT-VIS Inc, Boulder, CO). If the determinant has a non-zero value, the matrix is inverted, again using LU decomposition.

In cases where the covariance matrix is singular, the number of bands is reduced. Although more sophisticated methods based on singular value decomposition have been proposed to reduce the number of dimensions (Rodgers, 2000; Wang et al., In press), it was decided to employ a more

pragmatic approach and to halve the number of bands by taking every second channel. Based on this reduced number of bands, COV_{rad} is recalculated and tested again for its singularity. If singularity continues to exist, bands are reduced again by taking a third and a fourth of the initial set of bands, respectively. Eventually, if singularity still persists, the bands proposed by Fourty and Baret (1997) complemented with a band in the blue, green, and red, are used. In practice however, when using spectral data that are not strongly polished, matrix inversion appears usually directly possible or at least after the first reduction of bands. Reducing the number of bands by structured selection at regular intervals is justified by the equal spacing of bands in imaging spectrometer systems. Doing so, the original distribution of bands within the system is preserved and no displacement of the weights attributed to particular wavelength ranges occurs, which also favors a direct comparison between different land cover classes.

3.4.2 Exploiting a priori information on canopy variables

Even though the use of land cover classes considerably reduces the range of possible canopy realizations and offers an opportunity to quantify the uncertainty contained in the measurements, it only partly overcomes the ambiguity induced by the ill-posed nature of the inverse problem, thus still leading to multiple and instable solutions. Additional regularization, in the form of prior information, is required.

Although up to this point some prior information has already been included in constructing the LUTs by restraining the ranges of variation of the variables and defining their statistical distributions, test results showed that this is not sufficient for accurate regularization, so that other ways have to be explored. Since in an automated, image based, approach a priori estimates based on expert knowledge, predictions from additional information sources, or the temporal evolution of one or more parameters (Cfr. Section 2.5.5.2) cannot be presumed available, additional regularization should be retrieved from the spectral information content of the image itself.

3.4.2.1 Defining prior estimates using vegetation indices

A novel approach, based on the spectral information content contained in the image and the lookup tables, is presented for obtaining a first estimate of the solution and defining the level of confidence around this estimate. First estimates of all variables are based on a predictive regression equation between a selected spectral vegetation index (VI) and the variable of interest which is applied to the image reflectance data (See Section 2.3). Vegetation indices have the advantage that they are computation efficient and based on only a limited number of independent wavebands. Strictly speaking, first estimates defined in this way are not pure a priori information, since the estimates partly rely on the data itself. Even if it might be more correct to speak of an additional regularization, the term *a priori* will be retained due to the similarity with "real" a priori information.

Since ground reference measurements of the canopy variables are absent, the regression functions are created by using radiative transfer modeling. This allows one to generate a virtually endless number of canopy reflectance spectra for which the input variables are known and illumination and view conditions can be controlled (Ceccato et al., 2001; Haboudane et al., 2004; Zarco-Tejada et al., 2005a; Zarco-Tejada et al., 2001). The advantage of such an RTM based approach

compared to an approach based on empirical measurements, is the fact that every regression function can be optimized for the view/sun geometry and soil and atmospheric conditions encountered in the image. Moreover, they can be calculated for each specific land cover class and for every canopy variable separately.

Simulating reflectance At this point, no additional radiative transfer model simulations have to be carried out, since a large number of canopy realizations and accompanying variables for each land cover class are already contained in the LUT. In order to make the regression function more robust, approximations for sensor, measurement, and model uncertainties are attributed to the simulated spectra in the LUT to form $R_{LUT}^*(\lambda)$. For the individual uncertainty sources, a random gaussian distribution rn of average noise contributions is assumed:

$$R_{LUT}^*(\lambda) = rn \cdot N_{model} \cdot (R_{LUT}(\lambda) + R_{LUT}(\lambda) \cdot rn \cdot (N_{sens}(\lambda) + N_{atm}(\lambda))) \quad (3.12)$$

Sensor noise N_{sens} , is expressed by the Noise Equivalent Delta Reflectance ($NE\Delta\rho$) which roughly corresponds to the reflectance divided by the signal-to-noise ratio ($\simeq \frac{\rho}{SNR}$). According to the average sensor specifications of various sensors (e.g. HyMap, MODIS, Landsat TM), the noise level is set to 0.01 (1%) for wavelengths severely affected by water absorption and to 0.003 (0.3%) for the other wavelengths (Cocks et al., 1998)(Figure 3.18a - blue line). Measurement and atmospheric uncertainties (N_{atm} ; Figure 3.18a - red line), mainly resulting from errors in atmospheric correction, are set to 0.005 (0.5%) for the blue range, and 0.001 (0.1%) for the rest of the spectrum (Berk et al., 2003). Finally, modeling errors (N_{model} ; Figure 3.18b) have been approximated by the relative standard deviation (i.e. the standard deviation divided by the average of measurements) of spectral measurements of a garden cress performed under laboratory conditions (Richter et al., 2006a). Such a laboratory set-up excludes any disturbing influences of atmosphere, positioning, or view constellation and is therefore representative for variations in the canopy alone. As shown in Figure 3.18(c), simulated uncertainty is representative for an average quality sensor under good illumination conditions.

Choosing appropriate VIs The predictive performance of a certain vegetation index depends on the variable to be estimated, and the canopy, atmospheric, soil, and illumination/view conditions at the time of data recording. Therefore, it was decided not to select one specific vegetation index, but to look in every new situation and for each variable which VI shows highest performance. For this purpose, a wide range of well-known VIs, developed for the analysis of green vegetation, is calculated and plotted versus the different canopy variables. In order to find the best predictive equation, both linear and exponential fits are tested for each combination of variable and VI. Based on the combination of the largest coefficient of determination (R^2) and the smallest root mean square error (RMSE), the best performing regression function is selected for each variable in each class. The final regression functions are written to a text file, including the accompanying R^2 and RMSE values.

The single VIs and their formulas are listed in Appendix C. For a detailed discussion of their unique characteristics the reader is referred to the original publications or to several publications describing and comparing various VIs (Broge and Mortensen, 2002; Haboudane et al., 2004;

3.4. The optimization algorithm

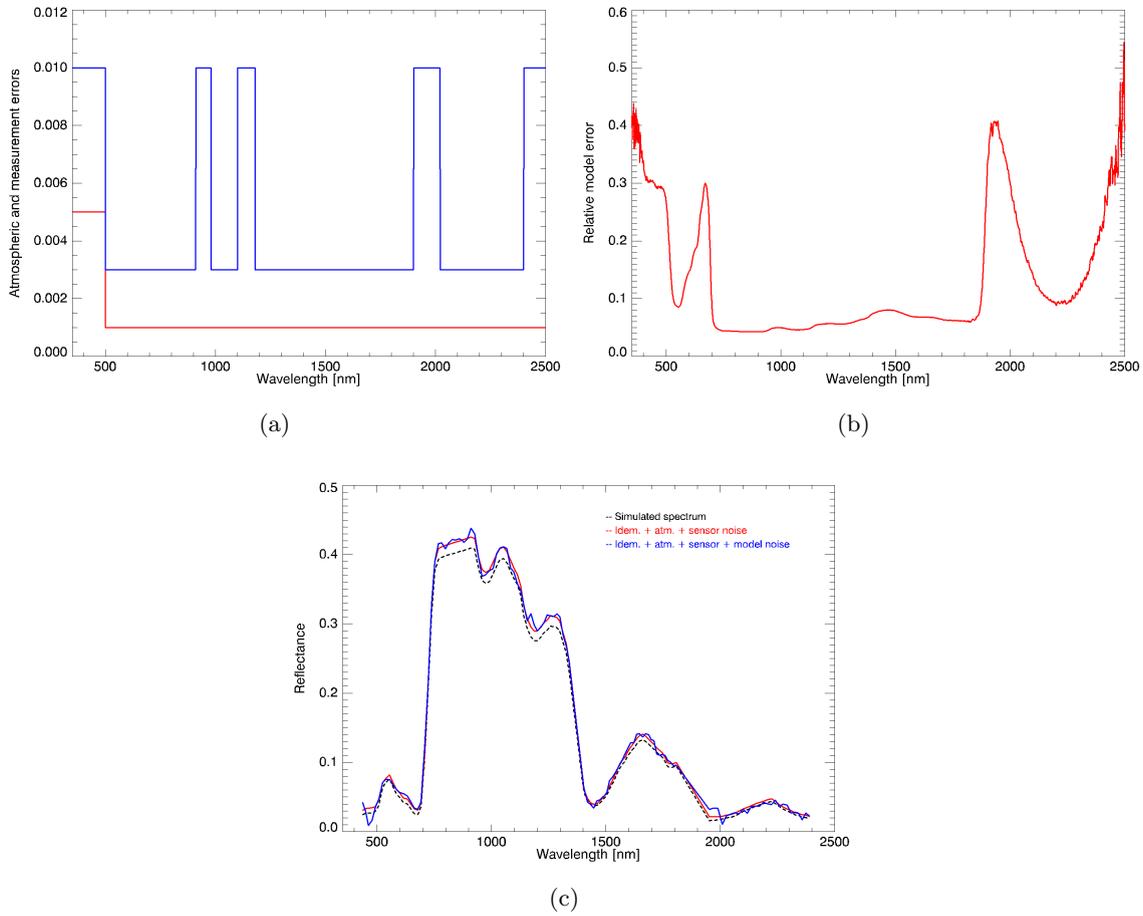


Figure 3.18: Example of attributing atmospheric, sensor, and model uncertainties to a simulated spectrum. (a) shows the $(NE\Delta\rho)$ of a typical sensor (blue line) and the uncertainties involved in atmospheric correction (red). (b) shows the relative modeling errors per wavelength (see text for explanation). (c) shows the result of adding sensor and atmospheric errors (red) and model errors (blue) in successive steps to a modeled HyMap spectrum (black; $\theta_s = 45^\circ$, $\phi_{rel} = 0^\circ$; $C_{ab} = 50 \mu g \cdot cm^{-2}$, $C_w = 0.03 g \cdot cm^{-2}$, $C_{dm} = 0.01 g \cdot cm^{-2}$, $N = 1.7$, $LAI = 3$, $ALA = 57^\circ$, $HS = 0.015$). Uncertainties are defined according to Equation 3.12 by applying a random gaussian error with a wavelength dependent standard deviation as given in plot a and b.

Ruecker et al., 2006; Zarco-Tejada et al., 2005a; Zarco-Tejada et al., 2005b). The selected VIs can be roughly subdivided into 4 different categories:

- *Broadband vegetation indices.* VIs of this type were originally designed to estimate canopy structural variables such as LAI and $fCover$ and all include at least a waveband in the red and in the NIR. Some of them include an additional band in the blue to make them more resistant to disturbing atmospheric influences (e.g. SARVI, EVI) or a green band to render them more susceptible to changes in chlorophyll concentration (e.g., TVI, MTVI1).
- *Chlorophyll indices.* VIs in this category were designed to estimate leaf chlorophyll content and contain 2 or more narrow bands around the red edge. Some members of this group of

the aim at characterizing the red-edge inflection point (REIP), the wavelength of the red edge where the 1st derivative of the spectrum is maximal (or the 2nd derivative is zero).

- *Water indices.* Designed to estimate plant water content or detect water stress. Indices of this type have 1 or more bands in spectral regions sensitive to water absorption, which at the same time makes them susceptible to atmospheric water vapor. Therefore, their performance and portability strongly depend on the quality of the atmospheric correction. Water band indices that in simulation studies were found too sensitive (e.g. Normalized Difference Water Index (NDWI) (Gao, 2002), or Normalized Difference Water Index - Mid Infrared (NDWI-MIR) (Chen et al., 2003), were not used.
- *Dry matter indices.* The indices contained in this category have in common that they are sensitive to one or more of chemicals composing C_{dm} , and include indices sensitive to lignin absorption (NDNI, NDLI) and cellulose absorption (CAI). Notice that absorption features due to C_{dm} in green vegetation are usually masked by water absorption. Therefore, indices of this category only really make sense in estimating C_{dm} in the class *dry vegetation*.

Although the different indices were designed for detecting specific canopy characteristics, they are all tested for every variable. Finally, for all vegetation pixels a prior estimate is calculated, leading to an image with nine bands, one for each free variable input to PROSPECT and SAILh. Even if in this way a map with preliminary results is provided, the use of vegetation indices alone cannot substitute the actual radiative transfer model, as only a few spectral bands are used and the collinearity between variables is not accounted for. The importance of the latter will be exemplified in the next chapter.

Sensitivity of VIs The predictive power of the regression equation between a selected VI and a canopy variable strongly depends on land cover type, illumination/observation geometry, and the variable in question. Figure 3.19 shows some examples derived from the sun/view constellation of the HyMap imagery in the next Chapter. The examples show the large difference in sensitivity among the different variables and classes. Spectral response to changing C_w appears very large for a dense canopy, whereas for a sparse canopy reflectance shows no response at all. For LAI , a clear saturation is observed for high values, whereas for a sparse canopy, sensitivity is high. In neither of the considered cases spectral reflectance is sensitive to soil brightness. Even if these plots are only a few selected examples, they very well illustrate the high variability among the predictive equations.

Two of the VIs that were selected for the 3 more sensitive cases, were actually designed for the considered canopy variable, i.e., LWVI2 for the estimation of leaf water content and RDVI for the estimation of canopy structural variables such as LAI . Nevertheless, the best performing VI for the prediction of LAI in sparse vegetation is also the LWVI2, which can be ascribed to the similar spectral response in the SWIR of leaf water and LAI . It should be noticed that several of the best performing VIs in the examples are based on narrow hyperspectral bands in the SWIR. Thus, in the case of a broadband multispectral sensor such as Landsat Thematic Mapper, or imaging spectrometers covering only a limited part of the reflective domain, not all VIs can be calculated and, as a consequence, the predictive equation would be less accurate.

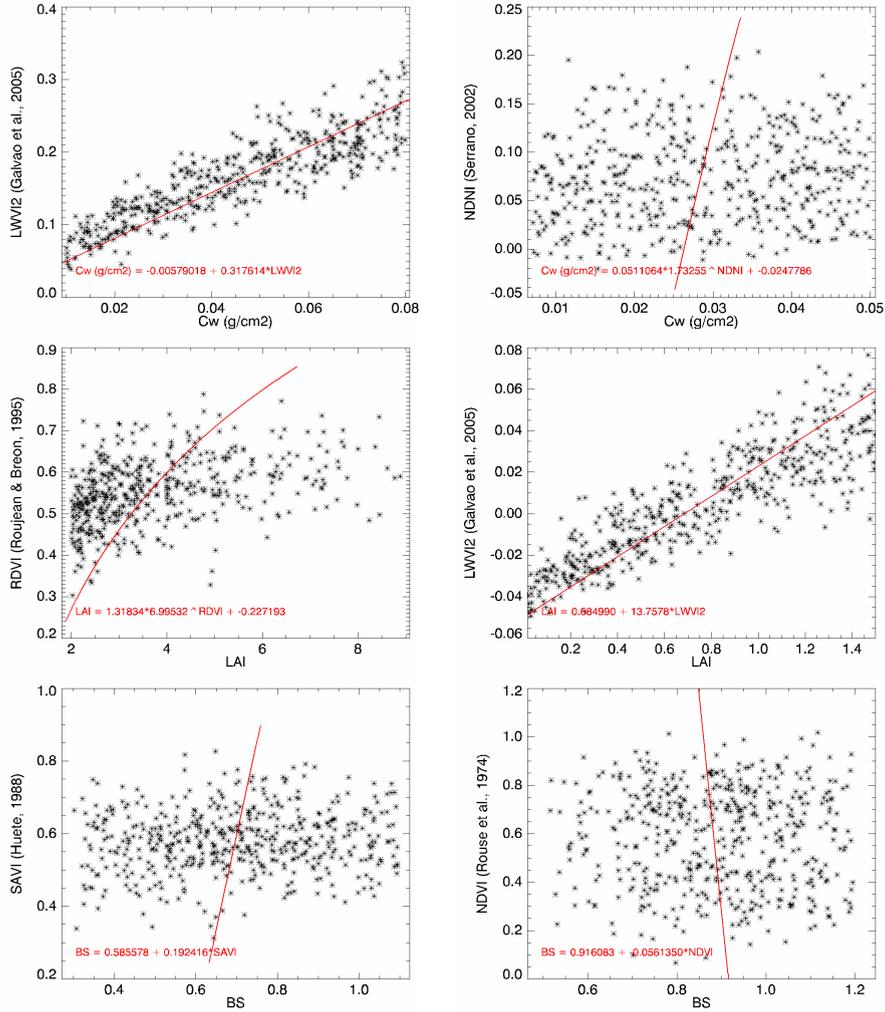


Figure 3.19: Examples of predictive regression functions between best performing vegetation indices and Cw, LAI, and BS, based on RTM simulations. The left column shows regression equations for the class *bright vegetation*, the right row *mix soil/vegetation* (view zenith = 20°; sun zenith = 30°; relative azimuth = 28°).

3.4.2.2 Assigning uncertainty to the prior estimates

Minimizing the distance χ_{var}^2 between the a priori estimates and the samples used to generate reflectances in the reduced LUT (i.e. $\widehat{\widehat{LUT}}$) is performed according to:

$$\chi_{var}^2 = (V_{prior} - V_{\widehat{\widehat{LUT}}})^T \cdot \alpha_v \cdot COV_{var}^{-1} * (V_{prior} - V_{\widehat{\widehat{LUT}}}) \quad (3.13)$$

where V_{prior} is the vector with the prior estimates of the j variables that are left free in the inversion, $V_{\widehat{\widehat{LUT}}}$ is the set of free variables that was used to simulate $R_{\widehat{\widehat{LUT}}}$ in Equation 3.9, and COV_{var} is the matrix describing the covariance between the different variables. α_v is a diagonal

matrix containing the weights that different variables receive in the cost function. These weights correspond to the coefficients of determination that were assigned during the establishment of the predictive regression equations, with high R^2 values corresponding to high weights. The α coefficients are indicative for the quality of the regression model (Cfr. Figure 3.19). They represent the systematic errors and are therefore complementary to COV_{var} . Thus, prior estimates that are based on regression models with a high accuracy are emphasized while those based on poor predictive equations are suppressed.

The covariance matrix COV_{var} in Equation 3.4.2.3 accounts for the non-systematic model errors related to the predictive equations themselves and for the random errors in the predicted values resulting from uncertainties in reflectance (noise, co-registration, and so on). For land cover class c the covariance between the variables is calculated from all pixels of the given class contained within the image, and is denoted COV_{var}^c .

Analogous to Equations 3.10 and 3.11 used to compensate for systematic radiometric uncertainties induced by classification anomalies, local variance of the prior estimates is calculated from a kernel of 5×5 pixels around the pixel under consideration, and is added to the class specific covariance matrix of prior estimates:

$$COV_{var} = COV_{var}^c + VAR_{loc}^c = \begin{pmatrix} var_{cla}^0 & \cdots & cov_{cla}^{0n} \\ \vdots & \ddots & \vdots \\ cov_{cla}^{n0} & \cdots & var_{cla}^n \end{pmatrix} + \begin{pmatrix} var_{loc}^0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & var_{loc}^n \end{pmatrix} \quad (3.14)$$

with var_{cla}^n denoting the class variance, and VAR_{loc}^n the local variance of variable n . cov_{cla}^{nm} is the class specific covariance between variable n and m . If the neighborhood of the pixel shows strongly heterogeneous a priori estimates, the diagonal of the covariance matrix is affected likewise, on the other hand, when the local neighborhood is homogeneous, the covariance matrix COV_{var} will be barely altered with respect to COV_{var}^c . Notice, that more than one land cover class can be found in the kernel. In cases where pixels lack a spatial context, which occurs when the neighboring pixels belong to a class that is excluded from processing, COV_{var} in Equation 3.4.2.3 is substituted by the class specific covariance term COV_{var}^c .

3.4.2.3 Defining the final solution

The smallest 20% of sorted χ_{var}^2 values is used to select those entries from $\widehat{\widehat{LUT}}$ that are used for calculating the final solution of variable V^j . This is done in such a way that the set of variables leading to the smallest χ_{var}^2 receives largest weight:

$$V_{est}^j = \sum_{k=1}^n (W_k \cdot V_k^j) \quad (3.15)$$

where n is the number of entries selected from the reduced $\widehat{\widehat{LUT}}$ and V_k^j the value of variable V for LUT-entry k . The weight W_k , attributed to a each particular entry selected from the LUT, is defined by:

Table 3.3: Quality flags used to indicate the confidence of the inversion.

<i>Code</i>	<i>Error</i>	<i>Consequence and action</i>
0	No error/limitations encountered	
2	Less than 30 LUT spectra fall within range of $\pm 20\%$ around input spectrum (Eq. 3.8)	Parametrization LUT is not representative for input spectrum \rightarrow range of spectral similarity is enlarged to $\pm 50\%$ from input spectrum
3	Less than 30 LUT spectra fall within $\pm 50\%$ of input spectrum (Eq. 3.8)	Class is probably not representative for input spectrum \rightarrow all spectra contained within LUT are used for minimization.
7	Too few neighborhood pixels to calculate local variance terms in COV_{rad} and COV_{var}	Class specific covariance matrices are used instead \rightarrow local variations are not accounted for
8	COV_{rad} cannot be inverted	Spectral covariance is not accounted for \rightarrow cost function is based on maximum likelihood estimator (class and local variance) COV_{var}

$$W_k = \frac{1/\chi_k^2}{\sum_{k=1}^n 1/\chi_k^2} \quad (3.16)$$

where χ_k^2 is the result of the cost function (Equation) for LUT entry k .

3.4.3 Accuracy description

It must be reemphasized that the solution is not a single value but based on the concept of maximum a posteriori probability (MAP), indicating that the solution is the *most probable* one within a certain distribution of solutions. Therefore, the uncertainty around the final solution can be described by the standard deviation, which is equal to $\sqrt{variance}$ and assumes a normal distribution around the mean. The standard deviation is calculated from the same selected entries of the reduced \widehat{LUT} that were used for the calculation of V_{est}^j (Equation 3.15).

In addition to the normal distribution around the estimated mean, each pixel is documented with a *quality flag* indicating important decisions that were taken during processing and concern the quality of the solution (Table 3.3). The final quality flag is the sum of the values in the table (e.g., a value of 9 (=2+7) indicates that the solution was based on all LUT spectra within a range of $\pm 50\%$ around the input spectrum and that a class based COV_{rad} and COV_{var} without local variance was used.

3.4.4 Accounting for angular anisotropy

So far, the discussion was confined to the processing of mono-directional datasets while large field-of-view sensors were regarded as a kind of "collection" of mono-directional sensors, each of them representing a unique combination of view angle and relative azimuth. The local moving

windows, used to describe variance in the direct neighborhood of a pixel, partly overcome contingent inaccuracies in view angle attributions and the possible crisp jumps in prior estimates that may occur at the transition from one view angle interval to the next. Although several view/sun constellations are present in one scene, pixels are only viewed from one direction, and therefore are still regarded as mono-directional data.

The retrieval procedure for multi-angular data evolves in a way similar to the one for mono-directional data, with a few modifications regarding the cost function and the description of covariance. The procedure starts with the generation of a distinct LUT for each available observation constellation. For each land cover class this is done according to the range and number of variables and intervals presented in Appendix B while only the view/sun constellation (view zenith and/or relative azimuth) varies.

Exploring radiometric information Analogous to Equation 3.8, only the LUT entries that in the Landsat TM bands deviate not more than 20% from the input spectrum are considered, while extending the range to 50% and eventually to all LUT entries when the criterium is not met. Notice, that this criterium must be met for *all* view constellations simultaneously, and thus not only the reflectance in LUT^i should approximate the spectrum in view constellation i , but the complete anisotropy should match. This is expressed in:

$$\widehat{LUT}_c^i \in \{LUT_c^i | 0.8 \cdot R_{meas}^i \leq R_{LUT}^i \leq 1.2 \cdot R_{meas}^i\} \quad (3.17)$$

where LUT_c^i is used to denote the set of LUTs generated for the i different available view/sun constellations in class c , R_{meas}^i the measured reflectance anisotropy, and R_{LUT}^i the simulated reflectance anisotropy.

The cost function used to characterize the closeness between modeled and measured reflectances is based on Equation 3.9 and adapted in order to incorporate the different view constellations:

$$\chi_{rad}^2 = \sum_{i=1}^{n_{dir}} \left((R_{meas}^i - R_{\widehat{LUT}}^i)^T \cdot COV_{rad}^{-1} \cdot (R_{meas}^i - R_{\widehat{LUT}}^i) \right) \quad (3.18)$$

where n_{dir} is the number of view configurations. COV_{rad} describes the radiometric covariance between the wavebands in every separate view angle. Notice that in this way no covariance/weighting between the different view constellations is included. This is conform to the definitions proposed by Knyazikhin et al. (1999b) and Knyazikhin et al. (1999c) for MISR and is justified by the lack of a priori information on the sensitivity and model accuracy of the different view angles.

Exploring a priori information on variables A priori estimates on the variables are calculated for each view configuration separately using the entries in the respective LUTs (Section 3.4.2.1). Having i different view constellations, for each pixel and for each variable, the number of i a priori estimates and accompanying ranges of uncertainty (σ) are averaged to form the final a priori estimate and uncertainty. The same is done for the determination coefficients providing

the weights α_v in Equation 3.4.2.3.

Covariance between the variables is calculated following Equation 3.14. Alternatively, in cases where neighborhood information on variables is missing, COV_{var} is based on the class specific covariance matrix like specified in Section 3.4.2.2.

3.5 Conclusions and preface to Chapter 4 and 5

In this chapter the background of the image based CRASh inversion approach and the implementation of it in a practical automated environment was discussed. The approach, which is based on existing radiative transfer models, focuses on improving retrieval performance by reducing the number of multiple ill-posed solutions. This is done by considering the pixel (and its spectral values) not as an independent measurement, but by also taking into account its spectral, spatial and, if available, angular context. The simultaneous exploitation of the various information sources is extremely important in cases where only very little knowledge is available on land cover type and phenology and where additional information sources such as multi-temporal observations or data from other sensors is lacking.

The novelty of the presented approach mainly resides in the ensemble of processing steps and the way regularization is obtained, namely by the real-time calculation of a priori variable estimates based on vegetation indices and predictive regression equations optimized for land cover type, and view/illumination geometry. Although the idea of basing RTM inversion on land cover classes is not new (Houborg et al., 2007; Knyazikhin et al., 1999b; Knyazikhin et al., 1999c), up to date no automatic approach is known that utilizes the land cover classes as a basis for the characterization of spectral covariance.

Even if in this chapter the advantages and some possible limitations (e.g., the limited capability of the RTM in mimicking radiation transfer in complex, heterogeneous canopies) of the proposed approach have already been pointed out, the approach should first be extensively tested on a wide range of data and canopy types before any general conclusions can be drawn. Although exhaustive testing in terms of sensor configurations and canopy types would go beyond the scope of this thesis, it was decided to work out two test cases dealing with observations taken by imaging spectrometers with strongly differing sensor specifications and observation properties. In addition, it was chosen to test the approach at three different spatial levels, using field spectrometer measurements with a footprint of less than a square meter, data from an airborne sensor with a pixel size of $5 \times 5 \text{ m}^2$ and satellite data with a spatial resolution of $17 \times 17 \text{ m}^2$, respectively.

The first validation chapter (Chapter 4) discusses the potential of the model for the characterization of temperate grasslands. Based on field spectrometer measurements, the contribution of the different algorithm components and information sources (e.g., land cover classification, description of covariance between wavebands, introduction of a priori information) to the end product will be tested in successive order. Moreover, the influence of the band configuration is tested, even as the sensitivity of the model to class parametrization and spectral uncertainties. In the second part of Chapter 4, the approach is tested on data from the airborne HyMap sensor. This part focuses on the sensitivity of the model to changing observation and illumination conditions and on the influence of the proposed spectral land cover classification on the end product.

The second validation chapter (Chapter 5) emphasizes on the estimation of cotton chlorophyll content from the hyperspectral CHRIS/PROBA satellite. Apart from having a different band configuration compared to the HyMap sensor, this sensor is also able to observe a target from up to five different view directions. In this way, an additional information source, independent from the spectral and spatial one, is included in model inversion. The focus of this chapter will therefore be the concurrent exploitation of the spatial-spectrodirectional information.

Although a validation of the model which is based on data simulated with a radiative transfer model would certainly provide a better control on the input variables and hence a more comprehensive data set, simulated data would not account for possible mismatches between the radiative transfer model and reality, while radiometric and spatial uncertainties could only be roughly approximated. For this reason, validation was exclusively based on real remote sensing data.

Chapter 4

Validating CRASh at ground and airborne level: grassland characterization using field spectrometer and HyMap data

4.1 Introduction

In summer 2003, the Waging-Taching experiment, carried out in southeast Germany, was set up to validate the automatic retrieval algorithm proposed in the previous chapter. The study involved the instantaneous sampling of agricultural grasslands by HyMap imaging spectrometer data and ground validation measurements. However, merely validating the results obtained by RTM inversion with the ground based biophysical and -chemical variables would not reveal the added value of the proposed approach compared to other statistical and RTM-based approaches. Validation should therefore also involve a quantification of the merits supplied by the single algorithm components. Moreover, comprehensive analysis of model sensitivity to land cover class parametrization, data quality, and observation/illumination geometry is required to guarantee the portability of model performance to future data sets and other sensors. The current chapter will address these aspects based on field spectrometer and airborne hyperspectral data sets. Whilst the field spectrometer data exclude uncertainties attributable to atmospheric and geometric correction and thus facilitate a direct comparison between spectral signal and the canopy properties of interest, the HyMap imaging spectrometer data allows for studying the spatial distribution of retrieved variables, even as the influence of changing observation properties on retrieval performance. The study concentrated on temperate grasslands, which are one of the predominant land cover types in central Europe and therefore of high economic value.

4.1.1 Preceding grassland studies using imaging spectroscopy

Grasslands occupy around 25% of the Earth's land surface and are therewith the largest food suppliers to wildlife and dairy worldwide (Brown, 1998). In cultivated, intensively used dairy husbandry systems, the way in which grasslands are managed determines to a large degree the productivity and profitability of a farm (Rougoor et al., 1999). In this context, accurate assessment of biochemical (e.g., Nitrogen (N) and chlorophyll content) and biophysical (leaf area index (*LAI*), dry matter yield) status of the fields provides valuable information on the necessary management strategies to be followed, the finetuning of nutrient and water supplies, harvest planning, and the intercomparison of fields and farms (Schut, 2003). The required accuracy of these information sources should be at least 10% in order to support proper decision making (Harmony et al., 1997).

Detailed knowledge on the nutrient balance of a farm does not only bring commercial benefits but also helps to reduce excessive use of manure and fertilizer and, hence, emissions from farming systems into the environment (Smit et al., 2003). Washout of manure, the predominant nutrient supplier for grasslands under temperate climatic conditions, influences the quality of fresh water resources and hence their ecological state (Melzer, 1999; Pinnel, 2007). Allocating the actual status of the agricultural grasslands in support of optimized dairy system management would therefore also help to reduce ecological impacts.

As regular field inventories with sufficient coverage are too cost-intensive and time-consuming, radiometric measurements provided by multi- or hyperspectral sensors are increasingly explored for their suitability of deriving key grassland parameters (Clevers et al., 2005; Psomas et al., 2007; Schut, 2003; Vohland and Jarmer, 2007). At laboratory level, the use of imaging spectroscopy for the characterization of grass swards was intensively studied by Schut (2003). He explored the potential for growth monitoring, detection of nitrogen and drought stress, dry matter yield, clover content, nutrient content, feeding value, sward heterogeneity and production capacity, using a close range spectroscopy system applicable in the laboratory or for mini experiments.

Based on the work Schut, Clevers et al. (2005) showed that the small-scale experiments could be extrapolated to field level using a multi-spectral non-destructive close sensing system. Using multiple linear and partial least square regression they obtained good results for biomass ($R^2=0.77$), dry matter yield ($R^2=0.68$) and reasonable accuracy for dry matter content ($R^2=0.58$). Nutrient contents like N were only poorly predicted.

The upscaling of approaches developed at ground level to airborne or satellite based observations is hampered by the heterogeneity of the grassland canopy in space and time. Unlike most agricultural crops which build rather homogeneous canopies and exhibit a more or less predictable annual life cycle, grasslands are composed of different plant species that all have their unique morphologic and chemical features and whose distribution is strongly dependent on periods of growth and regrowth (Vohland and Jarmer, 2007). Moreover, composition and architecture are strongly influenced by management strategy and use (e.g., number of cuts, meadow vs. pasture). This implies that developed statistical approaches for the retrieval of grassland variables from remote sensing data will have to be calibrated for land use, phenological stage, and species composition.

But even physical approaches based on radiative transfer modeling, often proposed as a robust alternative to statistical approaches, face similar limitations since canopy parametrization is often given by an average set of variables which mostly does not account for single species

within the canopy. This is confirmed by Vohland and Jarmer (2007) who were able to accurately retrieve *LAI* from field spectrometer measurements while only moderate results were obtained for leaf water and leaf dry matter content. Even if the application of additional constraints to canopy parametrization (i.e., by assuming a nearly fixed rapport between leaf water and leaf dry matter) in their case brought additional improvement, the use of such constraints in a full operational approach would only be of limited value since very little is a priori known about land cover and the expected variable ranges. Moreover, upscaling RTM inversion to airborne or satellite level would involve additional uncertainties ensuing from atmospheric and geometric preprocessing.

4.1.2 Objectives

The focus of this chapter is to find out to what precision the automated CRASh inversion approach proposed in the previous chapter is able to estimate important grassland canopy state variables from data of the airborne HyMap imaging spectrometer. HyMap overflights took place at two time steps (June 30 and August 4) during the summer of 2003, a summer that was characterized by extreme drought and excessively high temperatures (Deutscher Wetterdienst¹). The ultimate goal was to see if possible advancing drought stress, expressed by an altering water content, could be detected by remote sensing. Due to technical problems, the second HyMap data set could not be used for further purposes. Nevertheless, the canopy sampling and field spectrometer data collected during this second campaign could still be used for validation purposes at ground level.

Based on the spectral measurements and on the accompanying ground validation measurements for grasslands of leaf area index, leaf dry matter content, and leaf water content, the following issues are addressed:

- Contribution of the single algorithm elements to the final estimates.
- Absolute accuracy of the automatic approach, both for the field spectrometer and HyMap data.
- Sensitivity of the approach to land cover classification, LUT parametrization, and radiometric uncertainties.
- Stability of the algorithm in terms of changing flight and solar constellation.
- Effect of incorporating local spatial relationships.

4.2 Study site and data

4.2.1 Study site Waging-Taching

The study was performed in the catchment of Lake Waging-Taching which is situated in the foreland of the Bavarian Alps, close to Salzburg (Figure 4.1). The lake catchment has a size of

¹URL: <http://www.dwd.de>

4.2. Study site and data

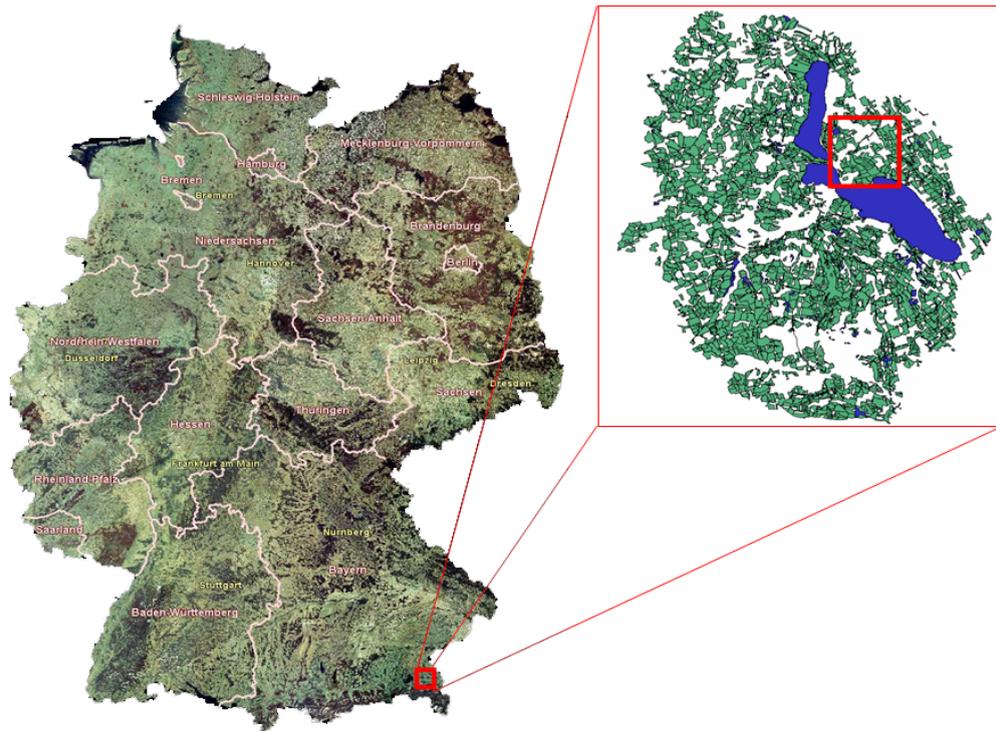


Figure 4.1: Location of Waging-Taching test site. The image at the right shows the agricultural parcels in the area. The red square indicates the actual study area

approximately 120 km² and its elevation ranges from about 450 to 700 meters above sea level.

The substrate in the area is merely constituted by glacially transformed impermeable clay deposits which form a perfect base for the development of peat. Both upland moors (on hill tops) and low moors (near the lake) are present. The soil types are heterogeneously distributed, corresponding to the small scale variations in geomorphology and elevation. Climate is moderate central European with relatively mild winters, temperate summers, and precipitation the whole year round.

Land use within the catchment is dominated by agriculture, though also forestry and to a lesser extent recreational use play an important role. Two third of the agricultural area is constituted by grassland, both meadows and pastures, while first mentioned prevail. Grassland use is predominantly intensive, which is represented by up to seven cuts a year and additional manuring for meadows, and frequently grazing (and the resulting constant dung input) for pastures. The parcels of different use and management intensity are evenly scattered through the catchment and range in area from the size of an average backyard to over 25 hectares.

While remote sensing data were acquired over the entire catchment, biophysical validation and field spectrometer measurements were concentrated on 3 fields situated near the lake shore around the center coordinates 47°57'N, 12°46'E (Figure 4.1 and 4.2).

4.2.2 Ground validation measurements

4.2.2.1 Configuration of ground sampling locations

Ground validation measurements were performed close to the time of the HyMap overflights to test the absolute retrieval performance of the inversion algorithm. For a comprehensive testing, dry and wet biomass and leaf area index were sampled at various locations in the area and, in order to facilitate a direct comparison between the measured canopy characteristics and their spectral signatures, field spectral measurements were performed exactly at the same locations. The measurements were carried out at 3 fields that during the first flight campaign varied either in land use or phenological stage, and were coded MEA1, MEA2, and PAS (Figure 4.2).

MEA1 and MEA2 are both intensively managed meadows, with the difference that at the time of the first HyMap campaign (from now referred to as HYM1; for the second campaign the abbreviation HYM2 will be used) MEA1 had a dense vegetation cover whereas MEA2 had been recently cut (Figure 4.3). A look at the pictures reveals that species composition is dominated by perennial ryegrass (*Lolium perenne* L.) whereas also different white clover types (*Trifolium repens*, cv. *Blanca* L.) are present. While MEA1 merely consists of green, vital vegetation, MEA2 is characterized by a certain amount of dry harvest residues, stubbles, and a small fraction of bare soil.

The epasture PAS is characterized by an inhomogeneous vegetation composition where at short distance grazed patches with low vegetation height alternate with patches mainly containing stinging nettle (*Urtica dioica* L.). As shown in Figure 4.3, the fraction of species other than grasses is relatively high compared to MEA1 and MEA2.

Each field was sampled on 5 to 7 locations in order to get the spatial distribution of the biophysical and biochemical characteristics within the field (Figure 4.2). In MEA2 and PAS the plots were positioned close to the corners and in the center of the field. As spectral properties in MEA1 showed a clear trend, from high NIR reflectance values in the southwest part of the field toward lower values in the northeast corner, a transect-like sampling strategy was applied to this field, allowing to capture possible trends in canopy variables too.

The location of each plot was obtained by a duplicated measurement with a non-differential handheld Garmin eTrex Vista GPS (Garmin International Inc., Olathe, KS). The accuracy of a single measurement, expressed in RMSE, is according to the instrument specifications less than 15 m in the horizontal plane. Although duplicating the measurements increases accuracy, it should be noticed that location errors, still persist and in several cases span more than one pixel.

On each plot location, an area of 1x1 m² was subject to a detailed examination, including measurements of leaf area index, wet and dry biomass, and canopy height. Canopy height was measured at 25 locations within the plot area, taking the height of the highest shoot at each point.

4.2.2.2 Leaf area index

Indirect sampling Indirect sampling was performed with the LAI-2000 Plant Canopy Analyzer instrument (LICOR, 2000). This instrument determines the gap fraction based on the ratio of

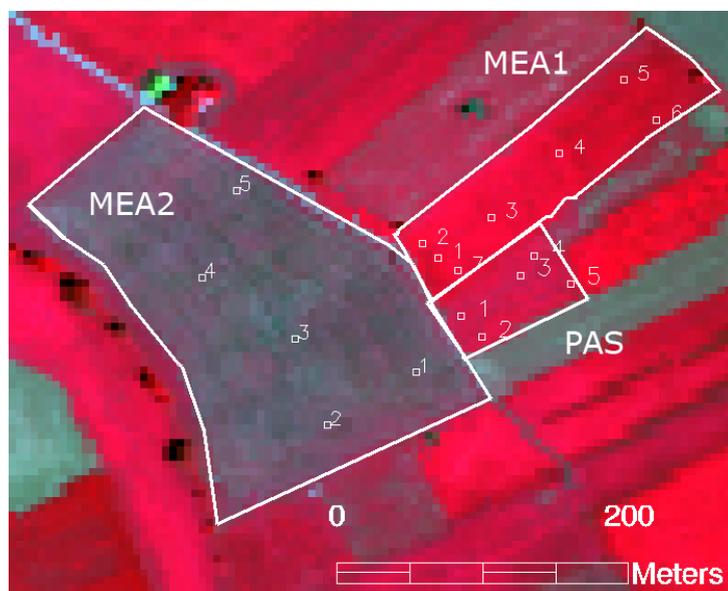


Figure 4.2: Location of selected test fields and plots used for canopy characterization. The background shows a subset of the WAGING07 HyMap flightline recorded at June 30, 2003 (Red = band 29 (849 nm); green = band 15 (646 nm); blue = band 9 (555 nm)).

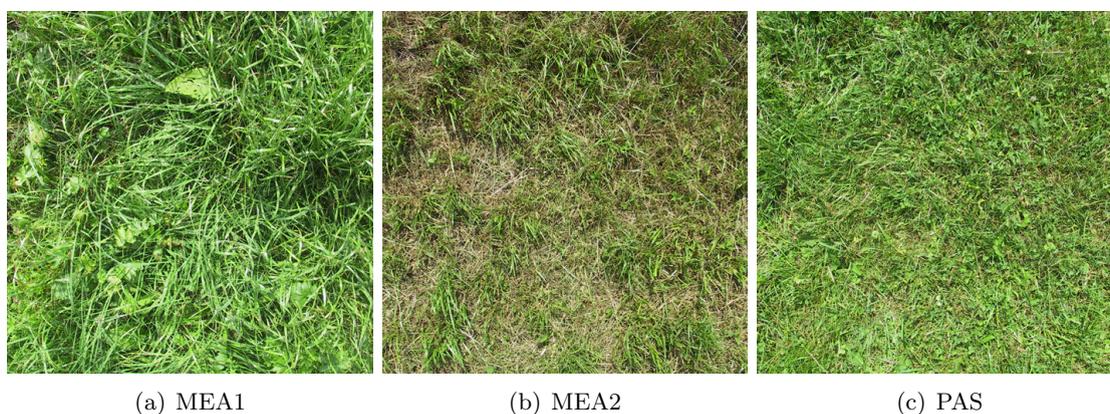


Figure 4.3: The three different types of grassland sampled during the first HyMap data campaign [HYM1]. MEA1 is an intensively used meadow in full development, characterized by a high vegetation height (a), MEA2 a recently cut intensively used meadow, characterized by low vegetation height, a relatively high content of senescent material and a small amount of soil fraction (b), and PAS, an intensively used pasture, characterized by a low, dense vegetation cover and a high fraction of vegetation species other than grasses (c).

diffuse irradiance at 490 nm measured above and below the canopy (Weiss et al., 2004). Radiance is captured by an optical fish-eye lens with a zenith cutoff angle of 74 degrees and is divided into 5 concentric detector rings each of them representing a different view zenith range. Since measurements in every ring are integrated over all azimuth directions, no information on the spatial distribution or the clumpiness of the leaves is available (Baret et al., submitted). The measurement principle, requiring upwards looking measurements from the bottom of the canopy,

in combination with the thickness of the measurement unit of approximately 5 cm, makes it difficult to characterize very low vegetation canopies.

Strictly speaking, not leaf area index but plant area index, or crop area index, is measured with the LAI-2000, since the instrument is not able to separate between radiance absorbed by leaves or by other phyto-elements (Broge and Mortensen, 2002). This may lead to overestimated *LAI* values in canopies with high amounts of stems, ears, and senescent leaves. Moreover, the model used for the calculation of the gap fraction considers the leaves as complete absorbers, allowing no transmittance or reflectance. Therefore, measurements should be taken at conditions of diffuse irradiance, i.e. at homogeneously covered skies or during sunrise/sunset, in order to reduce specular leaf reflectance and the stronger model violations thus involved. Nevertheless, despite some generalizations, the LAI-2000 usually provides results that for green crops are in good agreement with direct field observations (White et al., 2000).

To avoid man-induced disturbances within the field-of-view, a view cap of 180 azimuth was used to cover the part of the hemisphere oriented toward the person performing the measurement. One measurement consisted of 1 reading above the canopy layer followed by 5 readings at the bottom. This was repeated 25 times for one plot, following a regular grid sampling of 5 samples in both X and Y directions. Average *LAI* and standard deviation were retrieved from the ensemble of 25 measurements and based on all concentric rings. The use of all rings was preferred to a calculation based on ring 1-4 or on ring 4 alone, like recommended by (Weiss et al., 2004), as the latter provided values exceeding 8.5, a value that according to literature is somewhat unrealistic for grasslands.

Direct sampling Direct measurements of leaf area were performed in order to proof the consistency of the LAI-2000 measurements in situations where the instrument could not be placed well below the canopy, i.e., in the case of recently cut meadows and grazed pastures.

The leaf area measurements were performed on exactly the same plot locations where LAI-2000 measurements had been carried out before. First of all, the complete $1 \times 1 \text{ m}^2$ was cut at ground level and directly weighted with the aim of establishing the fresh wet above ground biomass (Section 4.2.2.3).

For practical reasons it was impossible to scan the area of the complete harvested square meter. Therefore, it was decided to determine the leaf area of some subsamples and to extrapolate the values thus obtained by using an empirical regression between fresh biomass and leaf area. This was done in the following way (Figure 4.4): i) 3 to 5 random subsamples were taken from the total available amount of biomass per plot and their fresh weights established. ii) Each subsample was spread out on a A3-sized paper in such a way that no overlap existed between adjacent leaves, after which the sheet was photographed from nadir with a digital camera (Figure 4.4a). iii) After correcting for image distortion, the digital image was subdivided into a plant material and a background class using the ENVI (ITT-VIS, Boulder, Colorado) unsupervised K-means classifier (Figure 4.4b). The number of pixels falling in the plant material class were related to the total area of the A3 sheet in order to get plant cover fraction. The procedure was repeated for the plots of MEA1 during the HYM1, and all the plots of the HYM2 campaign. This amounted to a total of 53 data pairs of wet biomass and leaf surface which were plotted in order to obtain a linear regression function (Figure 4.4c):

$$\text{Leaf area [cm}^2\text{]} = 34.6 \cdot \text{wet biomass[g]} + 103.4 \quad (4.1)$$

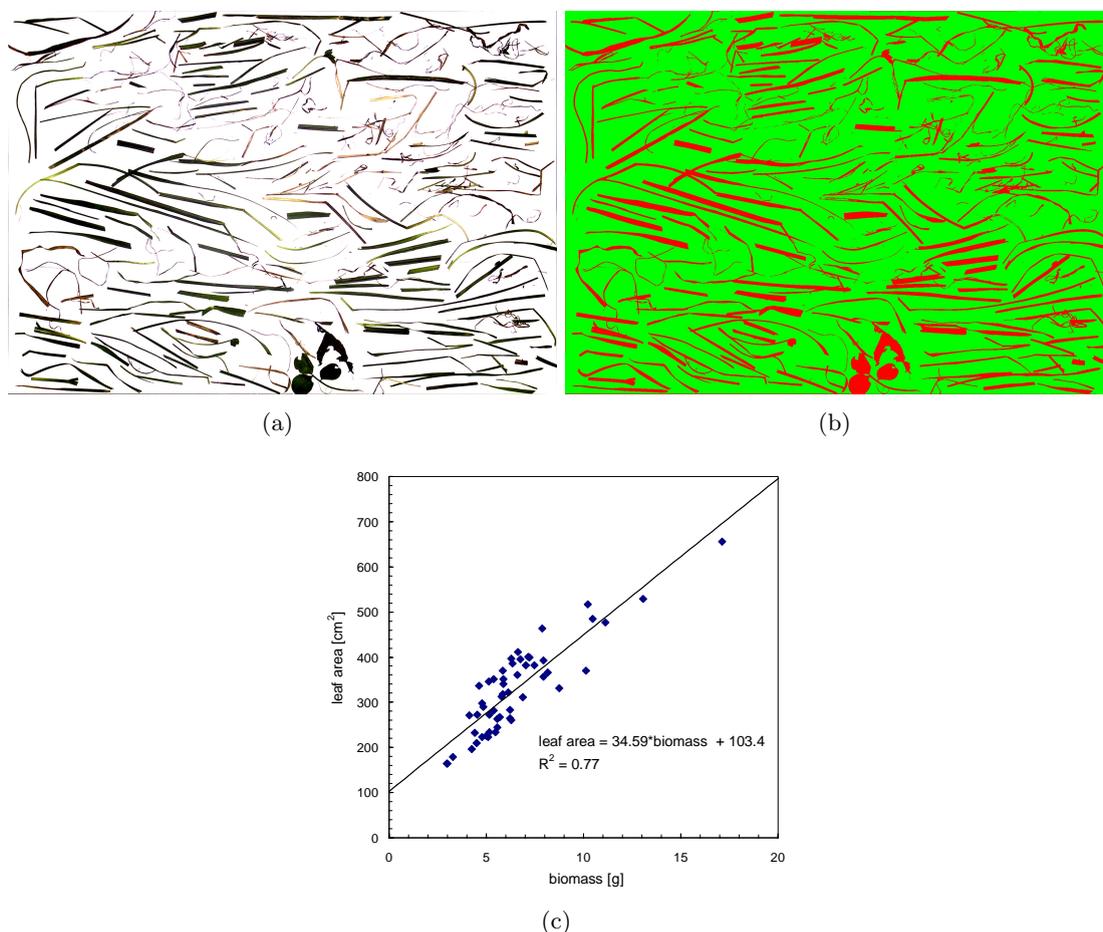


Figure 4.4: Procedure followed for destructive sampling of LAI : a subsample is taken from the total collected amount of biomass, laid out on an A3 paper sheet, photographed, and corrected for distortions (a). The image is classified in order to establish the leaf surface of the subsample (red fraction in b). The paired combinations of wet biomass and leaf surface are plotted to create a predictive regression function between the two variables (c) ($n=53$)

The plotted regression function, has a R^2 of 0.77 and a standard deviation of 170.9 cm^2 , which for an average sample weight of 10 g gives a relative standard error 38%. The established regression equation was used to calculate LAI from the fresh above ground biomass collected at each sampling plot of $1 \times 1 \text{ m}^2$. Notice, that for the composition of this equation no distinction is made between grasses and other vegetation species, or between green and senescent plant material, an assumption that is however in line with the $LAI-2000$ measurements.

4.2.2.3 Leaf dry matter and water content

The difference between fresh and dry biomass weight allows for the calculation of leaf water content (C_w) expressed as the equivalent water thickness (EWT ; $g \cdot \text{cm}^{-2}$ or cm). EWT or C_w

corresponds to a hypothetical thickness of a single layer of water averaged over the whole leaf area and is calculated by (Ceccato et al., 2001; Danson et al., 1992):

$$Cw = \frac{FW - DW}{A}, [g \cdot cm^{-2} \text{ or } cm] \quad (4.2)$$

where FW [g] stands for leaf fresh weight, DW [g] for leaf dry weight, and A [cm^2] for the leaf surface. In a similar way, the leaf dry matter content (Cdm) or specific leaf weight (SLW) is calculated:

$$Cdm = \frac{DW}{A}, [g \cdot cm^{-2}] \quad (4.3)$$

Leaf fresh weight (FW) was determined by harvesting the total above ground biomass of the square meter used for the characterization of spectral properties and leaf area. Leaf dry weight (DW) was determined after oven-drying the samples at 70° C for 36 hours. Cw and Cdm were calculated by dividing the total amount of water (= $FW-DW$) and dry weight by the measured LAI . This was done for both the indirect and direct estimations of LAI .

4.2.2.4 Comparing results obtained with direct and indirect LAI sampling

The results of the canopy variables that were measured for each validation plot during both campaigns can be found in Appendix E while the characteristics per field are summarized in Table 4.1. The latter shows that there is significant variation in LAI and canopy height within and between the different fields. Within-field variation is highest for the pasture, thus properly reflecting the pattern of alternating grazed and ungrazed patches. Intrafield variability is lowest for MEA2, like one would also expect for a recently cut meadow.

Figure 4.5 plots the LAI values measured in the direct and indirect way. Although values generally agree, some significant deviations can be observed, particularly for PAS (left). An interesting trend appears when the difference between the two measurements is divided by the $LAI-2000$ measurement (right): the relative difference shown in this way reveals that divergence between direct and indirect measurement is particularly large at low LAI -values. These values mainly correspond to the recently cut meadows and hence to low canopy heights.

The discrepancy between directly and indirectly sampled LAI at lower vegetation heights is also reflected in the calculated Cdm and Cw values (Figure 4.6). Cdm and Cw coincide quite well for fields where the LAI values calculated in both ways generally agree. In cases where LAI did not agree well (i.e., MEA2 and, to a lesser extent, PAS) Cdm and Cw values based on $LAI-2000$ measurements significantly deviate from average values and in some cases even exceed the ranges one would expect based on the values known from literature (Cfr. Table 3.1) with Cdm (Cw) values up to 0.0168 (0.0764) [$g \cdot cm^{-2}$] observed at MEA1 during the first campaign. This suggests that $LAI-2000$ measurements underestimate LAI of low canopies. For the sake of comparison, the average vegetation height of each field was added to the charts, showing a strong negative correlation between the vegetation height and the discrepancy of both methods used to calculate Cdm and Cw .

Another interesting aspect that can be distinguished from Figure 4.6 is that in general Cw

Table 4.1: Field level results of grassland canopy measurements. In brackets the standard errors.

	<i>Unit</i>	<i>MEA1</i>	<i>MEA2</i>	<i>PAS</i>
<i>HYM1</i>				
LAI-2000	m^2/m^2	5.04 (1.26)	0.76 (0.17)	4.10 (1.90)
LAI empirical	m^2/m^2	4.23 (1.50)	1.92 (0.22)	5.71 (1.57)
Cdm LAI-2000	$g \cdot cm^{-2}$	0.0065 (0.0008)	0.0141 (0.0026)	0.0103 (0.0029)
Cdm empirical	$g \cdot cm^{-2}$	0.0079 (0.0012)	0.0055 (0.0008)	0.0066 (0.0012)
Cw LAI-2000	$g \cdot cm^{-2}$	0.0208 (0.0024)	0.0606 (0.0117)	0.0438 (0.0161)
Cw empirical	$g \cdot cm^{-2}$	0.0252 (0.0015)	0.0233 (0.0007)	0.0273 (0.0012)
Canopy height	<i>cm</i>	33.8 (13.9)	8.3 (1.4)	12.6 (4.4)
<i>HYM2</i>				
LAI-2000	m^2/m^2	2.37 (0.44)	1.01 (0.69)	4.90 (2.69)
LAI empirical	m^2/m^2	1.33 (0.44)	1.20 (0.39)	3.28 (1.38)
Cdm LAI-2000	$g \cdot cm^{-2}$	0.0053 (0.0012)	0.0166 (0.0078)	0.0076 (0.0054)
Cdm empirical	$g \cdot cm^{-2}$	0.0097 (0.0022)	0.0115 (0.0011)	0.0080 (0.0019)
Cw LAI-2000	$g \cdot cm^{-2}$	0.0130 (0.0046)	0.0295 (0.0130)	0.0204 (0.0091)
Cw empirical	$g \cdot cm^{-2}$	0.0225 (0.0023)	0.0207 (0.0012)	0.0244 (0.0019)
Canopy height	<i>cm</i>	15.5 (3.2)	6.5 (1.1)	18.3 (11.3)

(*Cdm*) is lower (higher) during the second campaign. This reflects the lower relative water content measured during the second campaign (Appendix E) and matches well to the general drought stress encountered during this excessively hot summer.

The canopy observations revealed some clear differences concerning the variation within and between the fields. But also differences between the applied measurement techniques were observed. Since both techniques underly some assumptions and uncertainties, it can not be concluded which one is more accurate, although we do have some indications that LAI-2000 underestimates foliage area for lower canopies. However, no distinction was made between different species or between leaves and other plant materials such as leaves and senescent material. Especially last component has a relatively high presence in MEA2 (Cfr. Figure 4.3), which may have lead to an overestimation of green *LAI*. Therefore, initially, both measurements will be included in the validation of the radiative transfer model inversions.

4.2.2.5 Field spectrometer measurements

Spectral properties of each single plot were measured exactly on the location where subsequently the measurements for canopy characterization would take place. This enabled a direct comparison between the structural and chemical composition of the plots and their spectral properties.

Spectro-radiometric measurements were taken using a portable Fieldspec PRO FR spectrom-

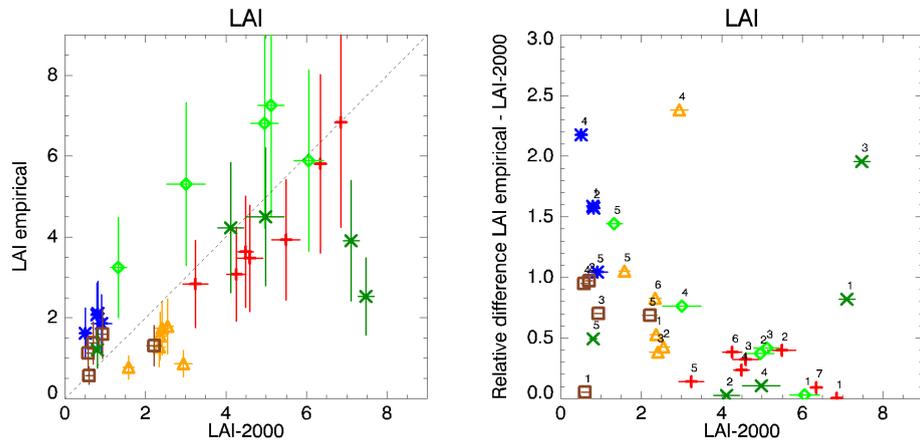


Figure 4.5: Comparison of LAI obtained by LAI-2000 and destructive sampling during HYM1 and HYM2 campaigns. At the left the absolute values are shown, at the right the relative difference between the two measurements, which is the difference divided by the LAI-2000 measurement. Red pluses = HYM1_MEA1, blue asterisks = HYM1_MEA2, green diamonds = HYM1_PAS, yellow triangles = HYM2_MEA1, brown squares = HYM2_MEA2, dark green crosses = HYM2_PAS. Error bars indicate the standard deviation associated to the measurements, the numbers indicate the plot numbers.

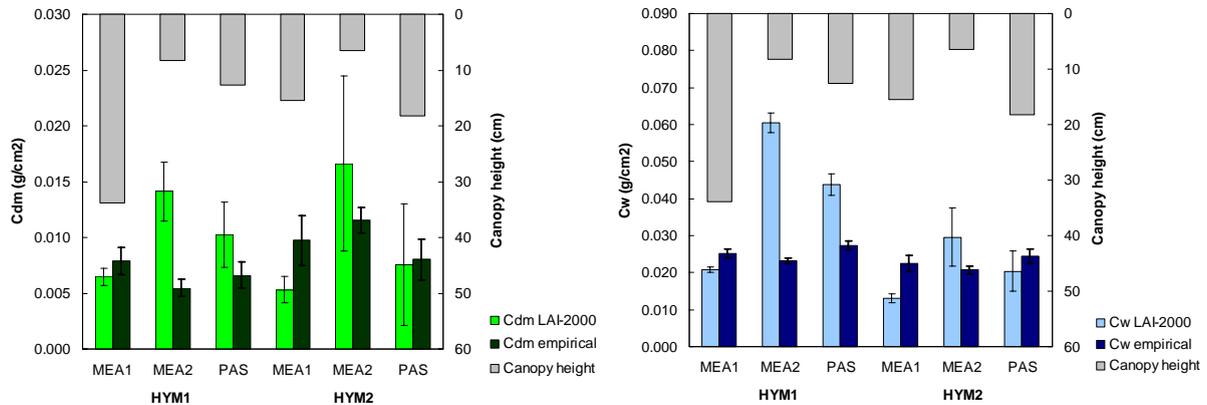


Figure 4.6: C_{dm} and C_w based on directly (empirically) and indirectly (LAI-2000) measured LAI

eter (Analytical Spectral Devices, Inc.). This instrument measures radiance in the solar-reflective domain at a sampling interval between 1,4 nm and 2 nm, and with a spectral resolution between 3 and 10 nm. In total 10 measurements were taken of each plot. This was done without using any fore optic, resulting in a FOV of 25° . The radiance measurements were directly converted into reflectance by taking a SpectralonTM panel as a white reference. The single spectra were corrected for the spectral properties of the applied Spectralon panel, deviations of the white reference off the 100 % reflectance line, and the spectral jump between the VNIR and the SWIR1 detector. The latter was done using an additive correction and taking the SWIR1 detector as a reference (Dorigo et al., 2006). Subsequently, average reflectance, standard deviation, and minimum/maximum reflectance were calculated. To enable a direct comparison of RTM results with the HyMap imaging spectrometer data, the field spectra were resampled to match the spectral

Table 4.2: Technical specifications of the HyMap sensor (Cocks et al., 1998)

Scanner principle	Whiskbroom
Field of view (FOV)	61.3°
Instantaneous field of view (IFOV)	2.5 mrad along track 2.0 mrad across track
Pixels per scan line	512
Ground resolution	3 - 10 m
Number of bands	126 - 128
Spectral coverage	400 - 2500 nm
Full Width Half Maximum (FWHM)	15 - 20 nm
Spectral sampling interval	13 - 17 nm
Radiometric Resolution	16 bit

configuration of this sensor during the HyEurope 2003 campaign, including band positions and spectral response curves (Appendix D). This was done likewise for the characteristics of the sensors discussed in Section 4.3.4.2 (Table 4.7).

Due to technical spectrometer problems, not all measurements of the first HyMap campaign (HYM1) could be used for further evaluation. Eventually, all spectral measurements of MEA1 could be used, plot 5 of MEA2, and plot 2 and 3 of PAS. However, to get a more comprehensive dataset, all measurements performed at ground level during the second campaign that year (HYM2) were included as well, leading to a total set of 27 sampling plots for which both spectral field measurements and canopy biophysical/-chemical properties were available.

4.2.3 HyMap imaging spectrometer measurements

4.2.3.1 Sensor characteristics

The HyMap (“Hyperspectral Mapper”) sensor is an airborne imaging spectrometer that provides almost contiguous spectral coverage throughout the solar-reflective domain (Cocks et al., 1998). Its technical details are summarized in Table 4.2. In total, it provides 128 wavebands over the range from 400 - 2480 nm with a spectral resolution of 13 - 17 nm. Signal to noise ratio is quoted at 500:1 outside the water vapor bands and its band configuration is within 1/10th of a pixel. The instrument measures reflected radiance at a field of view of 61.3° divided over 512 pixels per scan line. For the recommended spatial resolution of 5 m (Cocks et al., 1998), this leads to a swath width of approximately 2.3 km. The radiance that is captured by a rotating mirror is divided over 4 different detectors, each of them covering 32 bands. During the HyEurope 2003 campaign, the sensor had a slightly different setup compared to the original one with a reduced number of 126 bands. The exact setup of that year is found in Appendix D and from now on will be entitled HyMap 2003.

Table 4.3: Observation and illumination details of the 3 HyMap flight lines used in this study and recorded at June 30, 2003. Relative azimuth is the heading $\pm 90^\circ$

<i>Flight line</i>	<i>Time (UTC)</i>	<i>Heading [°]</i>	<i>Solar azimuth [°]</i>	<i>Relative azimuth [°]</i>	<i>Solar zenith [°]</i>
WAGING07	08:52	180.04	118.96	28.92/151.08	37.43
WAGING13	09:44	89.98	136.2	43.78/136.22	30.57
WAGING14	09:52	269.86	139.79	40.07/139.93	29.56

4.2.3.2 Flight configuration

HyMap data with a spatial resolution of 5 m at nadir were acquired by the German Aerospace Center (DLR) on June 30, 2003. Even though the complete Waging-Taching catchment was covered, only 3 flight lines were selected for further analysis (Table 4.3). These three flight lines were positioned in such a way that the field reference sites were viewed from 3 different view directions, thus enabling to study the effect of varying view/sun geometries on retrieval performances (Figure 4.7).

4.2.3.3 Geometric correction

Remote sensing data from airborne sensors are subject to large distortions due to variations in the flightpath (3-dimensional position) and in attitude (roll, pitch, yaw) of the aircraft. Since georectification techniques using polynomial transformations based on ground control points (GCPs) are not able to match sufficiently well these complex movements, a parametric approach is required that performs a pixel-by-pixel transformation and takes into account the exact position and attitude of the aircraft. The latter are usually available from Global Positioning Systems (GPS) and Inertial Measurement Units (IMU) aboard the plane. In addition, the HyMap scanner, which is mounted on a stabilized platform, contains an integrated IMU which, independently from the aircraft IMU system, records the exact sensor attitude during the flight.

The software package PARGE (PARAmetric GEocoding) (Schlaepfer and Richter, 2002) was used to orthorectify the data. Apart from navigation and sensor movements, PARGE is also able to correct for distortions due to topographic effects such as inclination, exposition, and distance of the pixel to the aircraft. Such correction is based on a digital elevation model (DEM). For the current study, a DEM with a resolution of 5 m was used, based on the digitized contour lines of the German Land Register Map (Deutsche Flurkarte) 1:5000 which were made available through the Bavarian Topographic Service² and have a height accuracy of approximately 1 m. Even if the DEM does not have the highest possible accuracy (which can be obtained for example by laser altimetry), this does not constitute a big problem, given the nearly flat topography at the location of the study site.

Geometric correction implies a resampling of the original data in order to match the new position. It was decided to employ the nearest neighbor resampling algorithm implemented in PARGE, in order to preserve the originally measured radiance values.

²Landesverband für Vermessung und Geoinformation Bayern

4.2. Study site and data

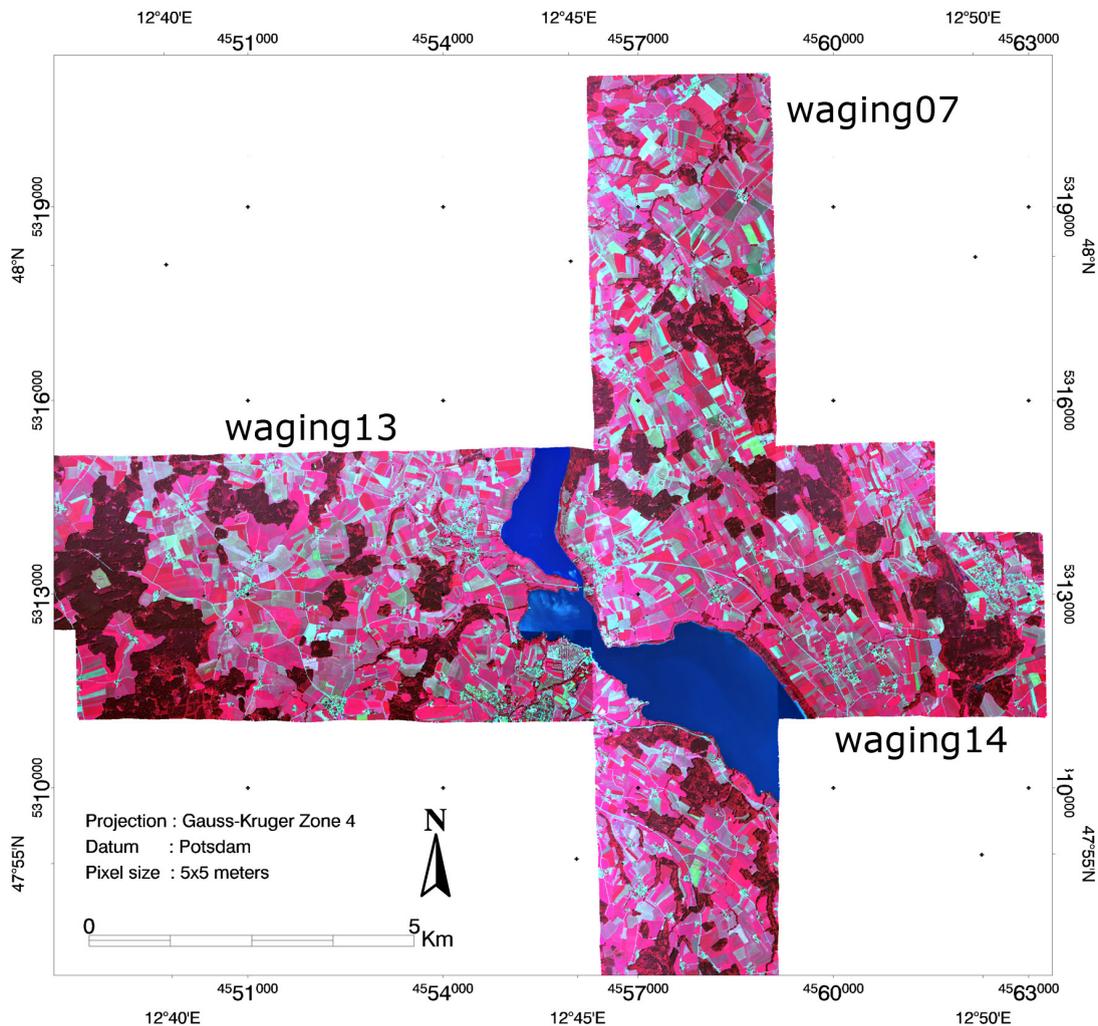


Figure 4.7: Georectified false color composite (red = band 29 (859 nm), green = band 17 (677 nm), blue = band 9 (555 nm)) of the flightlines WAGING07, WAGING13, and WAGING14, recorded at June 30, 2003.

The accuracy of the geometric correction was tested and optimized by including an average number of 20-30 GCPs per image taken from the German Topographical Map 1:25000 (TK25¹). GCPs were iteratively attributed and omitted until for each pixel a RMSE of less than 1 pixel in both x- and y- direction was obtained. Finally, for the small subset of intersecting flight lines at the study site, a manual image-to-image matching was performed. The image-to-image accuracy is demonstrated by the examples in Figure 4.8 and is about 1 pixel (5 meters) in x- and y-direction.

In addition to the georectified data, PARGE provides a scan angle file which includes information on view zenith and azimuth angle of every pixel in the scene. This important information is required for accurately assessing view/sun geometry during the RTM inversion process.

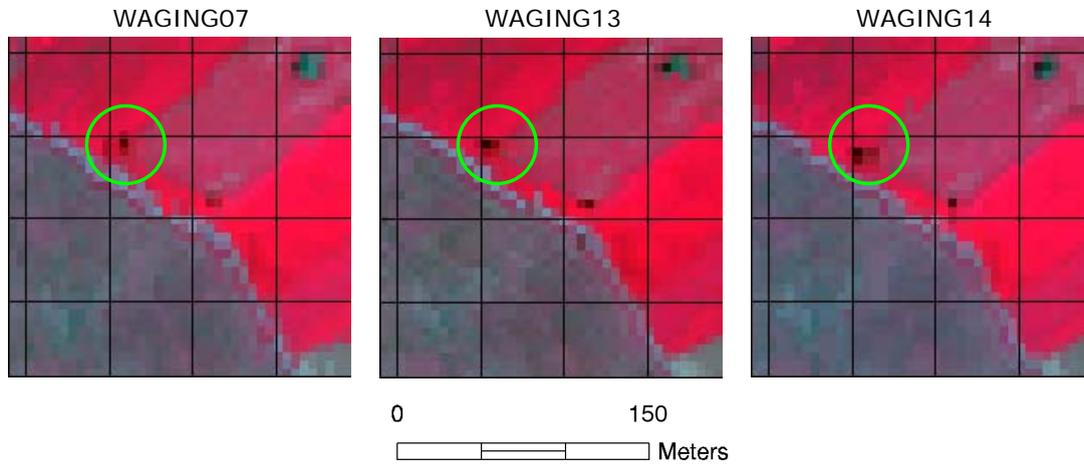


Figure 4.8: Comparing three georectified and image matched subsets of the flightlines recorded at June 30, 2003: (red = band 29 (849 nm), green = band 15 (646 nm), blue = band 9 (555 nm)). The shelter contained within the green circle shows a minor displacement from one image to another which, apart from co-registration errors, also results from changing view zenith and azimuth angles.

4.2.3.4 Calibration and atmospheric correction

The data, which is recorded in Digital Numbers (DN), was first transformed into radiance data by Integrated Spectronics Pty Ltd, the tenderer of the HyMap sensor, who corrected for dark current and electronic offsets. The required calibration coefficients were acquired at pre-flight laboratory measurements for sensor characterization and by onboard dark current measurements. A planned vicarious calibration using selected bright homogeneous targets within the scene, which is commonly performed for airborne sensors in order to quantify radiometric shifts between laboratory and operational conditions, could not be performed due to a malfunctioning field spectrometer (Section 4.2.2.5).

The atmosphere has a large influence on the solar radiation that is received by the sensor (Section 2.1.3) Atmospheric correction has to be performed in order to separate the signal originating from the actual pixel from that of neighboring pixels and the atmosphere and to obtain the hemispherical directional reflectance factor (HDRF) at the top-of-canopy that is required for the RTM inversion.

The data was corrected for atmospheric influences using the parametric model ATCOR4 (Richter and Schlaepfer, 2002; Richter, 2007a). This software package is based on the MODTRAN4 atmospheric radiative transfer code (Berk et al., 2003) and concurrently corrects for radiance differences caused by topographic effects. For the latter purpose, the same DEM was adopted as for the geometric correction (Section 4.2.3.3). Atmospheric correction was performed using a rural aerosol type while the water vapor column was automatically estimated from the bands in the water vapor absorption regions around 940 and 1130 nm, and identified 2 cm. The visibility was calculated by comparing dark water pixels of Lake Waging-Taching with reference spectra in the ATCOR spectral data base and finally for all flight lines fixed at 23 km.

Figure 4.9 compares average reflectance of MEA1, MEA2 and PAS for the atmospherically corrected flightline WAGING07 with the average spectra of all ASD field spectrometer measurements in the respective field. Despite the absence of a vicarious calibration, field spec-

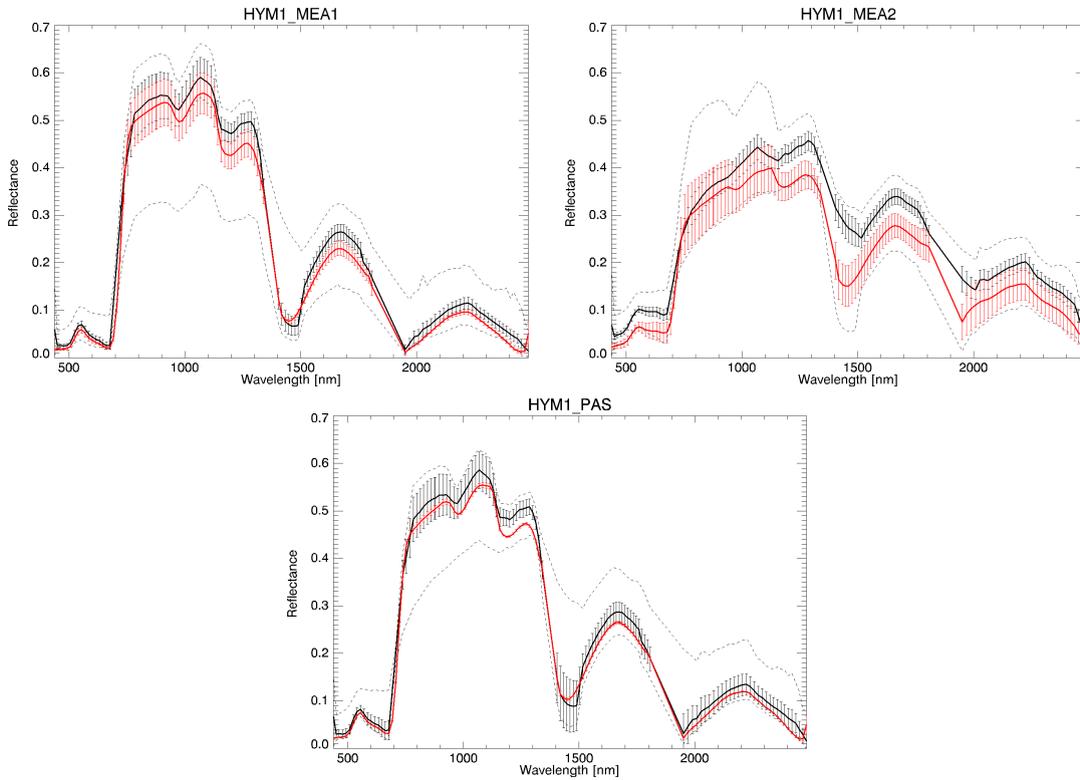


Figure 4.9: Comparison of average field spectrometer (red curves) and atmospherically corrected HyMap (black curves) reflectance data for the 3 test fields. The HyMap spectra reflect the average field reflectance taken from flightline WAGING07. Error bars indicate one standard deviation, the dashed black curves the minimum and maximum HyMap reflectance measured within the respective field.

trometer and HyMap data agree very well for MEA1 and PAS, while there is some discrepancy for MEA2. In general, HyMap reflectance is somewhat higher than reflectance measured with the field spectrometer. Nevertheless, all field spectrometer measurements fall within the range between minimum and maximum reflectance identified by the HyMap sensor for each field. It has to be kept in mind, however, that the spectral characterization using the ASD data is based on 3 to 7 point measurements, whereas the spectral characterization using the HyMap data is based on several hundreds of pixels for each field.

From the initial set of 126 bands, the ones falling within the known regions of enhanced atmospheric (water) absorption were excluded from further thematic processing, leading to final set of 95 sound bands (Table 4.4).

4.3 Exploring algorithm potential and constraints using field spectrometer data

Even if the theoretical framework of the approach proposed in the previous chapter seems justified, a thorough insight in the model and its components is required to draw any solid conclusions and to recognize potential restraints and sources of uncertainty. In this respect, not only insight

Table 4.4: HyMap bands that were excluded from thematic processing. In brackets the center wavelength in nanometers

<i>Spectral range [nm]</i>	<i>Affected bands (nm)</i>
< 470	1 (438), 2 (450), 3 (462)
1300-1530	60 (1313), 61 (1327), 62 (1340), 63 (1404), 64 (1419), 65 (1433), 65 (1447), 66 (1461), 67 (1475), 68 (1489), 69 (1503), 70 (1516)
1780-2100	92 (1783), 93 (1795), 94 (1807), 95 (1951), 96 (1970), 97 (1990), 98 (2009), 99 (2027), 100 (2045), 101 (2064), 102 (2082)
> 2400	121 (2405), 122 (2421), 123 (2437), 124 (2453), 125 (2468), 126 (2483)

into the contribution of each individual algorithm component is important, but also the ability of the underlying PROSPECT+SAILh combination of reconstructing the measured field spectra, and the distinctiveness of the spectral contribution of the individual measured canopy variables.

It seemed more appropriate to test the model performance on the field spectrometer data than on the HyMap data, as the field spectral measurements were taken exactly at the locations of biophysical and biochemical sampling. Thus, errors due to georeferencing or atmospheric interaction with radiance could be excluded and a direct relationship could be established between spectral signal and canopy variable of interest. In the next section, estimations will be based on the HyMap flight lines, for which uncertainties, mainly resulting from scaling, orthorectification, atmospheric correction, and the canopy variation found within a specific pixel, are considerably higher compared to the previous case (Cfr. Figure 4.1). In contrast, the latter will allow to study the influence of the spatial context of the pixel while the large number of available spectra in each land cover class enables a more accurate description of the covariance between wavebands and variables.

4.3.1 Spectral field characteristics

To get an idea of the spectral variation of the 27 field spectra used for this study (Section 4.2.2.5), the average spectrum of each field was plotted in Figure 4.10. Although the spectral variation within each field is not visible, some evident characteristics can be distinguished: the fields that are characterized by relatively high *LAI* (i.e., HYM1_MEA1, HYM1_PAS, and HYM2_PAS, Table 4.1) show high NIR reflectance and relatively low VIS and SWIR reflectance, the latter due to high chlorophyll and water absorption, respectively. In contrast, fields characterized by relatively low *LAI* (i.e., HYM1_MEA2, HYM2_MEA1, and HYM2_MEA2), show lower NIR reflectance, while HYM2_MEA2 also clearly shows reduced chlorophyll (VIS) and water absorption (SWIR), resulting either from a higher soil signal or an increased amount of senescent plant material and stubbles. It is noteworthy that for the latter 3 spectra, decreasing NIR reflectance is consistent with the decreasing *LAI* values measured by destructive sampling (Table 4.1), a trend that is consistent with expectations. This parallel between spectral reflectance and measured *LAI* cannot be observed for the LAI-2000 measurements. For this reason, and because of the excessively high leaf water content previously noticed, evaluations in this Chapter are

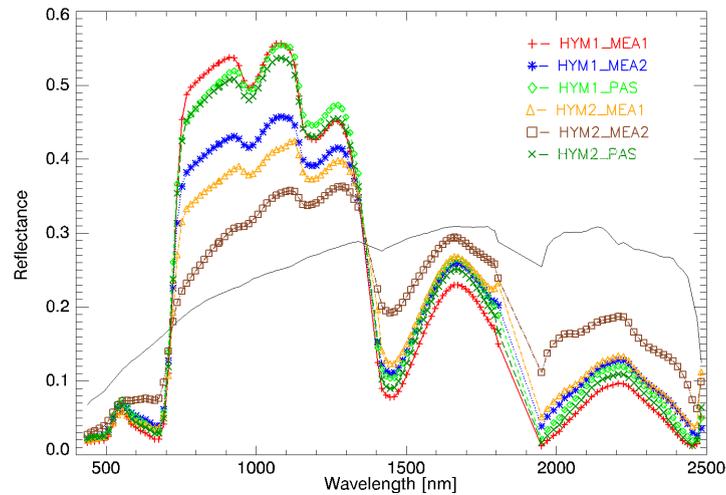


Figure 4.10: Average spectra of MEA1, MEA2, and PAS, measured during campaigns HYM1 and HYM2 with an ASD Fieldspec PRO portable field spectrometer and resampled to the spectral configuration of HyMap in 2003. The black continuous line represents an average soil spectrum measured in the immediate vicinity of the test fields

performed with canopy variables based on the destructively measured LAI . If done otherwise, this is explicitly mentioned.

4.3.2 Comparing modeled with measured reflectance

Prerequisite for a successful radiative transfer model inversion is the ability of the RTM of correctly mimicking measured reflectance in the forward mode. Introducing the variables measured at the validation fields in the RTM should provide canopy spectra that are within a plausible range from the ones measured with the field spectrometer. If this is not the case, potential difficulties in the model inversion can be discovered in advance.

Modeled and measured reflectance were compared for two of the most deviating spectra observed during the two HyMap campaigns: plot 1 in MEA1 measured during the HYM1 campaign (HYM1_MEA1 1) and plot 5 in MEA2 measured during the HYM2 campaign (HYM2_MEA2 5). In order to compare modeled and measured reflectance, the measured values of C_w , C_{dm} , and LAI (average and standard deviation; Appendix E) were used as input to forward PROSPECT and SAILh modeling. The characterization of the variables not verified by ground validation measurements were based on the SPECL classes to which the spectra were assigned by the automatic land cover classification (See Section 4.3.4.5 and Table 4.9): *bright vegetation* for HYM1_MEA1 1 and *dark vegetation* for HYM2_MEA2 5. The minimum and maximum values and the sampling plan of each variable were similar to the ones used for the characterization of the SPECL-based LUTs (Table 4.5).

Figure 4.11 shows that for the two previously mentioned spectra, modeled and measured reflectance are in good agreement and always fall within one standard deviation. Slightly larger relative deviations can be observed in the VIS for HYM1_MEA1 1 and in SWIR1 for HYM2_MEA2 5. A possible explanation for the former deviation could be the fact that average chlorophyll

Table 4.5: Distribution of the input variables used for the simulation of the canopy spectra in Figure 4.11. Average and standard deviation of Cw , Cdm , and LAI are based on field validation measurements (Appendix), parametrization of the other variables is based on the sampling schemes used for the SPECL class parametrization (Appendix B)

Variable	Distribution	HYM1_MEA1 plot 1				HYM2_MEA2 plot 5			
		Min.	Max.	μ	σ	Min.	Max.	μ	σ
N [-]	Gaussian	1	2.5	1.63	1.	1	3.5	2.0	1.0
Cab [$\mu g \cdot cm^{-2}$]	Combal	20	100	-	-	20	75	-	-
Cw [$g \cdot cm^{-2}$]	Gaussian	0.0100	0.0800	0.0227	0.0086	0.0150	0.060	0.0209	0.0080
Cdm [$g \cdot cm^{-2}$]	Gaussian	0.0050	0.0250	0.0061	0.0023	0.0035	0.0150	0.0113	0.0040
Cbp [-]	Gaussian	0	0	0	0	0	1.5	0	0.6
LAI [m^2/m^2]	Gaussian	2	9	6.83	2.59	0.1	6.0	1.31	0.50
ALA [$^\circ$]	Gaussian	30	70	57	20	25	70	57	20
HOT [-]	Gaussian	0.001	0.3	0.2	0.2	0.05	0.2	0.2	0.2
BS [-]	Gaussian	0.3	1.1	0.7	0.3	0.3	1.1	0.7	0.3

content of the simulated spectra ($= 60 \mu g \cdot cm^{-2}$) is lower than the actual one, although this value is already higher than the values that Vohland and Jarmer (2007) reported for grassland. A plausible explanation for the latter deviation could involve an inaccurate sampling of leaf structure parameter N with values that are too low with respect to the substantial fraction of dead harvest remainders in the observed canopy (Figure 4.3). Nevertheless, the presented modeled reflectance represents the average of several thousands of spectra which, as demonstrated by the ranges of variation, on a whole agree very well with measured reflectance. The results presented in Figure 4.11 therefore confirm the capability of the PROSPECT+SAILh combination and the proposed SPECL class parametrization of appropriately modeling canopy reflectance for the grasslands found in the area.

4.3.3 Correlation between canopy variables and wavebands

The degree to which a single variable can be estimated by radiative transfer model inversion mainly depends on its influence on the overall canopy spectral signature. For this reason, a correlation analysis was performed between the structural and biochemical parameters of interest and the canopy reflectances measured in the field for each wavelength. This was done separately for the green, healthy vegetation (represented by the spectra and variables measured in HYM1_MEA1, HYM1_PAS, and HYM2_PAS) and the partly stressed, cut meadows (HYM1_MEA2, HYM2_MEA1, and HYM2_MEA2). Figure 4.12 shows the results for the first case. The plot at the upper left reveals a high collinearity between Cw and LAI , an indication that separating the spectral sensitivity to both variables could be problematic. The observed collinearity originates from the almost constant water content in combination with the highly variable LAI (Cfr. Appendix E). The latter is confirmed by the unnatural behavior of correla-

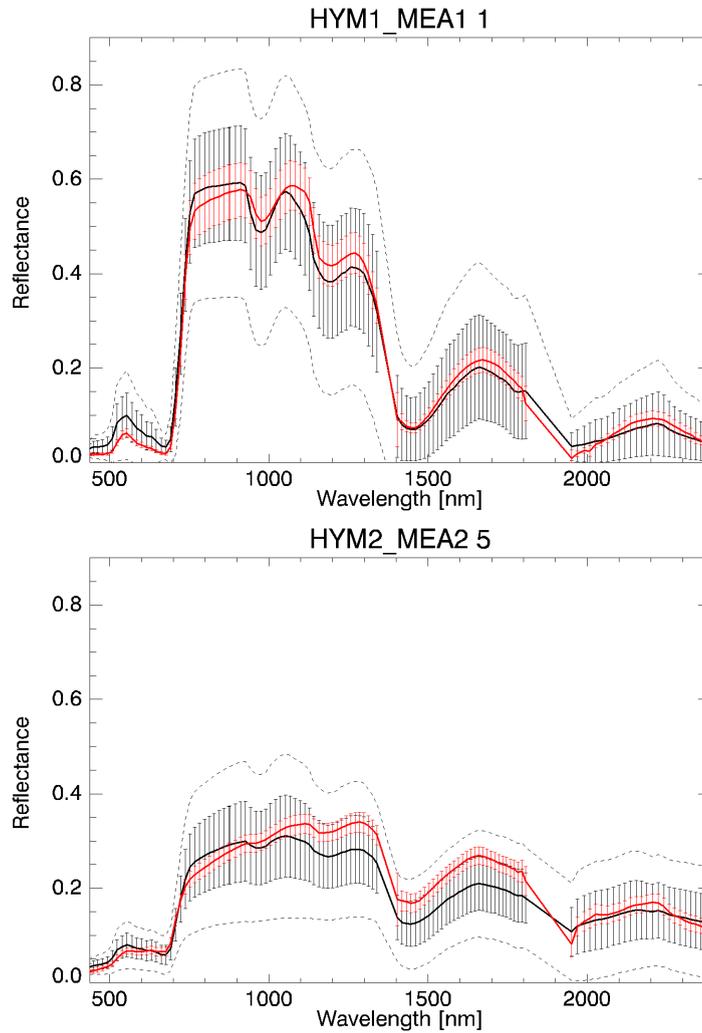


Figure 4.11: Comparing modeled (black) with measured (red) average reflectance for two sampling locations (HYM1_MEA1 plot 1 and HYM2_MEA2 plot 5). Error bars indicate 1 standard deviation of both measured and modeled reflectance and the dashed black curves a range of 2 standard deviations from average modeled reflectance. For the used input variables see Table 4.5.

tion between Cw and reflectance at wavelengths where water absorption is supposed zero (i.e., in the VIS). A second trend that can be observed from the figure is the complementary behavior of Cdm to Cw and LAI . This behavior is inherent to the applied measurement technique where an increase of leaf water content necessarily leads to an decrease of dry matter content and vice versa.

The upper right graph in Figure 4.12 shows the correlation between wavelength and Cw , Cdm , and LAI , based on RTM simulations with PROSPECT and SAILh. The correlation analysis is based on a total of 43,000 simulations for a green canopy and therefore represents a much larger range of variation in the variables (ranges and parametrization are equal to the ones used for the land cover class 4: *bright vegetation*). Although the large differences compared to the correlation coefficients based on measured spectra are evident, some analogy to the latter can

still be observed, like the complementary behavior of C_{dm} to LAI (and, to a smaller degree, to C_w) and the moderate collinearity between C_w and LAI in the NIR. However, responses for the simulated data seem more natural than for the measured ones. The zero correlation between C_w and the VIS is a good example of this, even as the correlation coefficients close to zero between C_{dm} and the wavelengths with high water absorption around 1400, 1900, and 2400 nm. The correlation between leaf chlorophyll content (C_{ab}) and modeled reflectance has been added for reasons of comparison (gray curve). It can be seen that for the modeled spectra, spectral variations in the VIS can be almost exclusively explained by variations of C_{ab} .

The graphs at the bottom of Figure 4.12 show the correlation coefficients when C_{dm} and C_w are calculated at canopy level ($C_{dm} \times LAI$ and $C_w \times LAI$, respectively). For the measured data, the very high collinearity between the two leaf constituents and LAI is confirmed, whereas for the modeled spectra, despite more pronounced collinearity, a significant degree of independent variation continues to exist. C_{ab} at canopy level persists to be primarily responsible for spectral variations in the VIS.

The correlation diagrams presented in Figure 4.12 provide important insight into the setbacks often involved in radiative transfer model inversion of canopy spectra, and the potential gap between model and praxis. Based on the modeled results, all 3 canopy variables of interest show distinctive spectral responses and could therefore be estimated from reflectance spectra simulated with PROSPECT and SAILh. However, this distinctive character can only be partly confirmed for the variables and reflectances measured in this study. Given the high collinearity between the measured leaf constituents and LAI , a robust retrieval of the biochemicals from the field spectra, either at leaf or canopy level, seems unfeasible for green, healthy vegetation, regardless of the applied inversion algorithm or technique. LAI and C_{ab} appear to be the key variables at canopy level and should therefore be well retrievable from measured spectral reflectance.

At leaf level, the results for the drier, cut meadows vastly agree with the results found for the green grasslands (Figure 4.12) whereas at the canopy level, collinearity between dry matter and spectral reflectance is absent. However, despite the more distinctive character, retrieval of C_{dm} will still remain problematic, since correlation between canopy reflectance and C_{dm} is low and significantly less than between canopy reflectance and LAI .

4.3.4 Stepwise integration of algorithm components

Although the previous subsection stretched the theoretical constraints of RTM inversion for the estimation of the leaf constituents C_{dm} and C_w , the inversion approach itself can still play a decisive role in finding the optimal set of variables within the limits dictated by the spectral response. Moreover, estimates of variables that do have significant influence on the measured canopy reflectance, such as C_{ab} or LAI , may very well depend on the chosen retrieval algorithm.

This subsection will study in detail the influence of the algorithm components introduced in the previous Chapter. In consecutive order, the influence of land cover classification, spectral covariance description, the introduction of priori estimates (with and without covariance description), and the use of the predictive regression functions will be discussed. Starting point of the new situation will always be the preceding stage. Until introducing spectral covariance, the spectral cost function will be composed of a maximum likelihood estimator (Equation 2.10), characterized by the average and standard deviation of the spectral measurements performed

4.3. Exploring algorithm potential and constraints using field spectrometer data

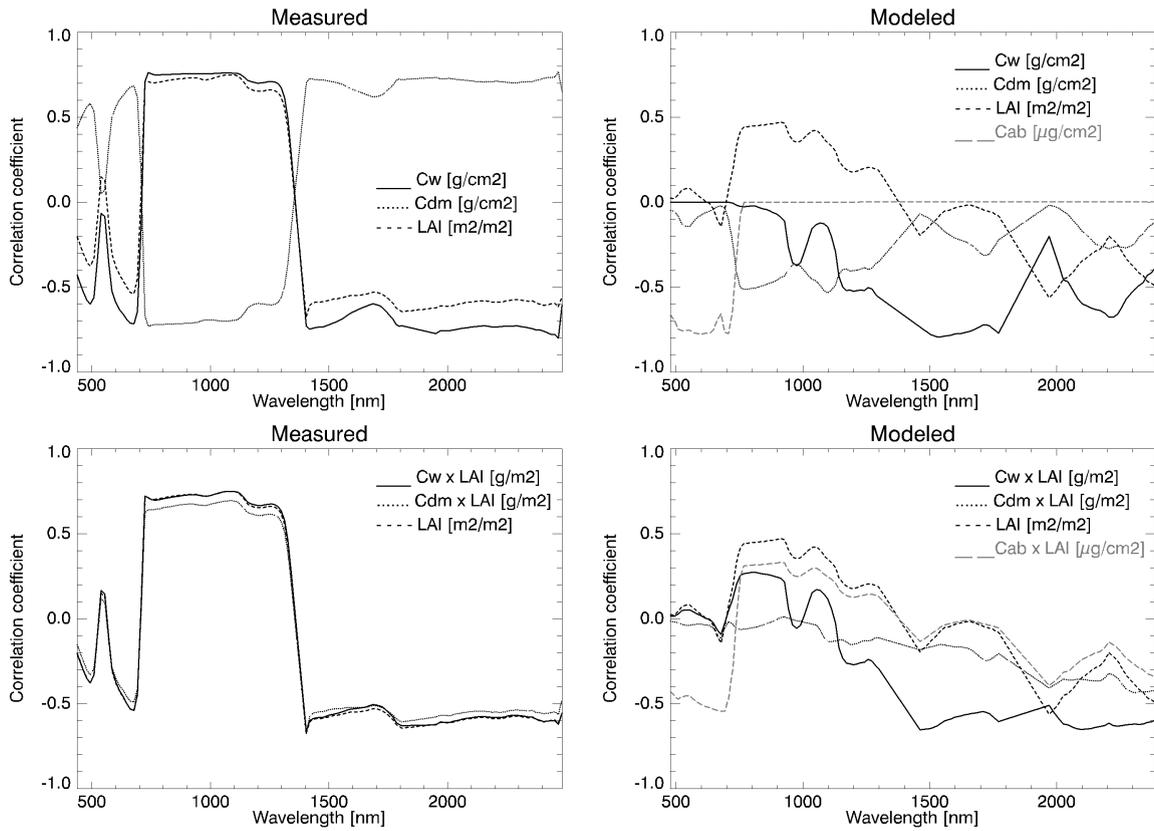


Figure 4.12: Coefficient of determination (R^2) between reflectance and measured (left) and modeled (right) C_w , C_{dm} , and LAI for vigorous green grassland, displayed both at leaf level (top) and at canopy level (bottom). For measured reflectance, $n=15$, for modeled reflectance, $n=43200$

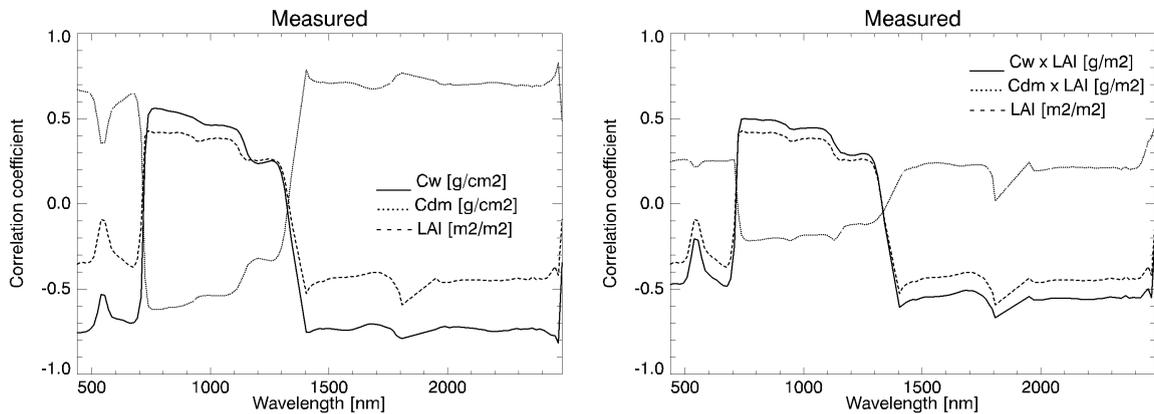


Figure 4.13: Coefficient of determination (R^2) between canopy reflectance and measured C_w , C_{dm} , and LAI for dry, cut meadows, expressed at leaf level (left) and at canopy level (right; $n=12$)

at the specific location. For all information levels, soil reflectance, skylight fraction and observation/illumination properties are identical. Input soil reflectance is an average soil spectrum measured close to the test site and was already presented in Figure 4.10. Atmospheric parameters used to calculate the skylight fraction are conform the ones parametrized during atmospheric correction of the HyMap scenes (Section 4.2.3.4), solar zenith (θ_s) is set to 30° , and view zenith (θ_v) is set to nadir (0°).

Two commonly used descriptive measures of similarity are used to describe the goodness of fit between measured and estimated variables. The relative root mean square error of the estimates ($rRMSE_{\bar{V}}$), expressed as a percentage, is the root mean square error ($RMSE_{\bar{V}}$) divided by the average of the measured variables $\mu_{V,m}$ and is given by:

$$rRMSE_{\bar{V}} = \frac{RMSE_{\bar{V}}}{\mu_{V,m}} \times 100\% \quad (4.4)$$

Average accuracy as a percentage ($AA\%$) forms the complement of the average absolute deviation and is expressed by:

$$AA\% = \frac{1}{n} \frac{\sum_{i=1}^n (V_{i,m} - |V_{i,m} - \bar{V}_i|)}{\sum_{i=1}^n \bar{V}_i} \times 100\% \quad (4.5)$$

where n is the number of observations, $V_{i,m}$ the measured, and \bar{V}_i the estimated value of variable V for observation i . Overall accuracy (OA) is in this study used to indicate the mean of the average accuracies for Cw , Cdm , and LAI . To visualize observed trends in the estimates, most results are also graphically presented. To facilitate visual comparison of the results throughout the various steps, the plotted ranges in all steps are kept constant, which in some cases might lead to high plot densities for which individual data points or error bars cannot be properly distinguished.

4.3.4.1 Influence of land cover classification

The influence of land cover classification on retrieval performance is twofold: first of all, it helps to restrain the range of possible canopy realizations, second, it allows for the construction of predictive regression functions for the estimation of a priori values, optimized for each land cover class. The first argument is discussed at this point, while the generation of class specific predictive equations will be discussed in Section 4.3.4.5.

To study the effect of incorporating a land cover classification, RTM inversion of the measured spectra is first performed with a single "global" LUT which is able to describe every potential canopy type found within the scene. The variable ranges used to construct this LUT are spanned by the extreme values found in the single LUTs in Appendix B and have been summarized in Table 4.6. The LUT is conform with the variable ranges and distributions that Bacour et al. (2006) used for the description of global canopy variation. The retrieval algorithm is conform the approach described in Section 3.4 while minimization between variables is excluded and the cost function used to minimize for radiometric distance is identical to the simple maximum likelihood estimator given in Equation 2.10.

4.3. Exploring algorithm potential and constraints using field spectrometer data

Table 4.6: Distribution of the input variables and sampling plan used to construct a global LUT to match all land cover types.

	Variable	Unit	Distribution	Minimum	Maximum	Mean	σ	# intervals
Leaf	N	Unitless	Gaussian	1	4.5	1.5	1	3
	Cab	$\mu g/cm^2$	Combal	1	100	(50)	(50)	6
	Cw	g/cm^2	Uniform	0.0050	0.0800	(0.0250)	(0.0300)	4
	Cdm	g/cm^2	Uniform	0.0020	0.020	(0.0075)	(0.0075)	4
	Cbp	Unitless	Gaussian	0	1.5	0.001	0.6	3
Canopy	LAI	m^2/m^2	Combal	0	9	(2.5)	(3.5)	6
	ALA	$^\circ$	Gaussian	20	85	57	20	5
	HOT	Unitless	Gaussian	0.001	1	0.1	0.3	5
Soil	BS	Unitless	Gaussian	0.3	1.3	0.8	0.3	3
							Total #	388,800

Figure 4.14 plots the estimated variables for the inversion based on the global LUT described above. Average RTM inversion results for Cw are in good agreement with the measured canopy variables based on destructive LAI sampling, with an RMSE of $0.0050 g \cdot cm^{-2}$ and an average retrieval accuracy of 83.8%. Dry matter content is estimated less accurately, with an RMSE of $0.0036 g \cdot cm^{-2}$ and an average accuracy of 59.9%. For both variables, the uncertainties attributed to the result (indicated by the error bars representing 1 standard deviation) are very large, while for several observations relative standard errors ($= \sigma/\mu$) even amount to more than 60%. Estimation of LAI in general shows poor performance, with an RMSE of 1.545 and average accuracy of only 31.5%. LAI retrieval shows good agreement for the low values measured at HYM2_MEA2, followed by a general overestimation up to values of around 4. Beyond 4.5, i.e., for the high values of HYM1_MEA1 and HYM2_PAS2, LAI is underestimated.

In the case where the spectra are first submitted to a SPECL classification (for classification results see Table 4.9) and subsequently a class specific LUT (Appendix B) is used for inversion, a moderate improvement can be observed for LAI and Cdm , while Cw estimations slightly deteriorate (Figure 4.15). However, a considerable decrease of retrieval uncertainties can be observed in particular for Cw and Cdm , which can be attributed to the reduced range of canopy realizations covered by each LUT. The better defined solution is also an indicator of improved stability of the retrieval. Nonetheless, estimation accuracy of both LAI (AA = 59.1%) and Cw (AA = 70.1%) are still insufficient, calling for additional regularization.

4.3.4.2 Quantifying spectral covariance and the influence of sensor configuration

In the previous examples, only the variance σ_i^2 of the spectral measurements was used to account for radiometric uncertainties. The variance matrix is a special case of the covariance matrix COV_{rad} in Equation 3.9, where the diagonal elements describe the variance of each waveband while the off-diagonal elements are all set to 0. Therefore, using variance does not compensate

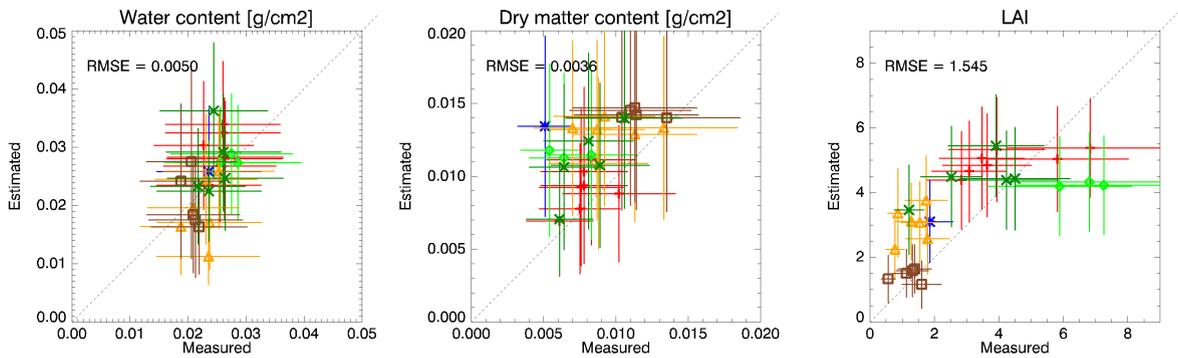


Figure 4.14: RTM inversion results for C_w , C_{dm} , and LAI applying the global LUT described in Table 4.6 and without using prior information. Colors and symbols correspond to the ones used in Figure 4.10: red pluses = HYM1.MEA1, blue asterisks = HYM1.MEA2, green diamonds = HYM1.PAS, yellow triangles = HYM2.MEA1, brown squares = HYM2.MEA2, dark green crosses = HYM2.PAS

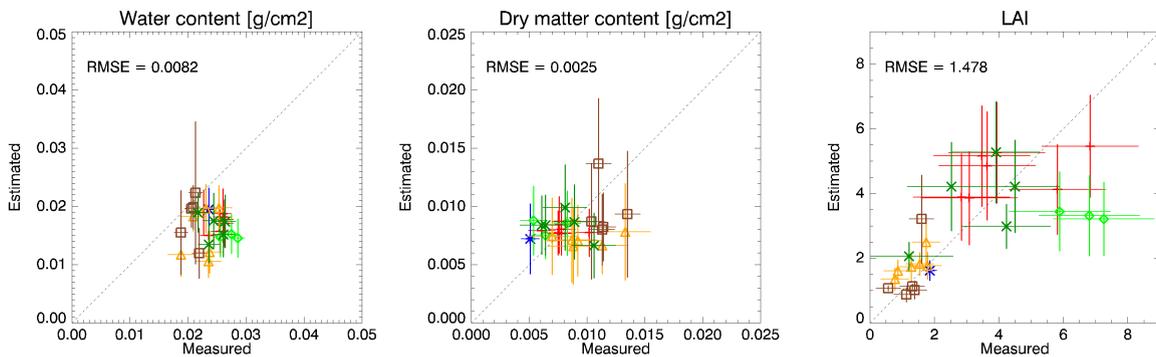


Figure 4.15: RTM inversion results for C_w , C_{dm} , and LAI applying LUTs adjusted to SPECL classes, and without using prior information. Red pluses = HYM1.MEA1, blue asterisks = HYM1.MEA2, green diamonds = HYM1.PAS, yellow triangles = HYM2.MEA1, brown squares = HYM2.MEA2, dark green crosses = HYM2.PAS

for the correlation between wavebands or the distribution and number of bands in a specific spectral region. Theoretically, introducing spectral covariance description would help to emphasize spectral regions of low reflectance while suppressing others. In practice, for vigorous green vegetation and using a full range sensor, this would imply an enhancement of spectral changes taking place in the VIS (for example induced by changing chlorophyll content) while subduing spectral variations in the highly reflective NIR.

To better study the effect of quantifying covariance in relation to the distribution and number of bands, RTM inversion was applied to different waveband combinations. For this purpose, the original Fieldspec reflectance data was resampled to the spectral properties of five commonly used, high spatial resolution sensors, differing either in the number of wavebands, or in the spectral domain they cover (Table 4.7). Apart from the HyMap sensor already introduced before, the sensor configurations included CHRIS/PROBA (hyperspectral, covering VNIR), Landsat ETM+ (multispectral, VNIR-SWIR), SPOT5 HRG (multispectral, VNIR-SWIR), and Quickbird (multispectral, VNIR). RTM inversion is applied to the different spectral configurations based on the

SPECL land cover classification and for two different scenarios. In the first scenario, radiometric uncertainty is expressed by the standard deviation of the measurements determined at each single plot, and minimization is performed by the maximum likelihood estimator of Equation 2.10. This case is similar to the approach described in the previous paragraph. In the second scenario, spectral distance is calculated according to Equation 3.18 while calculation of the spectral covariance matrix (COV_{rad}) for each land cover class is based on the different spectra attributed to that particular class (Cfr. Table 4.9).

Table 4.8 lists the results of both minimization functions. Concerning the sensor configuration, differences in retrieval performance between the various sensors are largest for Cw . As expected, Quickbird, covering only the VNIR, performs very poorly, which can be directly ascribed to the absence of bands in the SWIR, the spectral range most affected by leaf water absorption (Cfr. Figure 3.4). The CHRIS sensor, having a few wavebands at the onset of leaf water absorption and a band in the prominent water absorption feature around 970 nm, performs better, followed by the sensors having one or more bands in the highly affected SWIR. In the case where only variance is considered, the superior performance of SPOT, having only one band in the SWIR, in the retrieval of Cw may seem a little bit surprising. This is however compensated by a low retrieval accuracy of LAI . As seen before, both variables show highly collinear spectral behavior, which is expressed as complementing behavior in the retrieval of both variables. Nevertheless, the presented results suggest that, notwithstanding the very limited distinctive spectral response of Cw demonstrated in the previous section, estimations of Cw are still highly determined by adequate placement of wavebands in the spectral regions affected by leaf water absorption. Concerning Cdm and LAI , differences between sensor configurations are not as obvious as for leaf water content, although Quickbird performs somewhat less than the other sensors in estimating LAI .

At first sight, the introduction of covariance between the spectral wavebands seems to induce some odd effects on retrieval performances. On average, accuracy reduces when spectral covariance is introduced. This is what is expected for sensors having only a few wavebands in different characteristic spectral regions, since correlation between these bands is only little, and introducing covariance based on a class with only a few members may even increase measurement errors and hence the inaccuracy of the end product. Such a decrease of accuracy can be observed for LAI estimations of Landsat, SPOT, and Quickbird. On the other hand, introducing class covariance to the hyperspectral observations increases to some degree retrieval accuracy of LAI , while predictions of Cdm slightly worsen.

The explanation for the somewhat reduced retrieval accuracy when covariance description is introduced, has to be sought in the ensemble of variables that is accounted for. This is illustrated by Figure 4.16 which shows the estimations of Cab , LAI , and ALA , with, and without accounting for covariance between wavebands. The figure reveals that a shift in the estimated variables also takes place for the variables that are not validated in this study. Introducing the covariance description sort of redistributes the weights of the single wavebands in the cost function. While some variables take advantage of this, for others retrieval accuracy is reduced. The latter is exemplified by the loss of accuracy for LAI estimates from multi-spectral band configurations, probably at the benefit of Cab estimates which spectrally dominate the VIS. A quite unnatural behavior is observed for many of the estimates of Cab , which seem to be "trapped" around the upper boundary of variation permitted by the LUT, and which can be explained by the mismatch between measured and modeled reflectance in the VIS (Figure 4.11). However, it has to be beared in mind that in this example quantification of covariance has been based only on a very

Table 4.7: Spectral configuration of HyMap 2003, CHRIS Mode5, Landsat ETM, Quickbird, and SPOT HRG. For further details regarding CHRIS/PROBA see Chapter 5 and Appendix F

<i>Sensor</i>	<i>Spectral domain</i>	<i>Number of bands</i>	<i>Band position (Band width [nm])</i>
HyMap 2003	VNIR-SWIR	126	Contiguous at 13 - 17 nm distance (11-22)
CHRIS Mode 5	VNIR	37	Contiguous at 6 - 30 nm distance (6-47)
Landsat 7 ETM	VNIR-SWIR	6	478 (71), 570 (80), 662 (61), 874 (126), 1648 (200), 2224 (280)
Quickbird	VNIR	4	485 (70), 560 (80), 660 (60), 830 (140)
SPOT HRG	VNIR-SWIR	4	545 (90), 645 (70), 835 (110), 1665 (170)

Table 4.8: Retrieval performance for different sensor configurations using two different radiometric cost functions. SE stands for the case where the standard error of the plot measurements is used to describe radiometric uncertainty, COV for the case where covariance between different class members is used. The last column shows the overall average accuracy for all variables together. In bold the best result per column.

Sensor		<i>C_w</i>		<i>C_{dm}</i>		<i>LAI</i>		<i>OA [%]</i>
		<i>RMSE [g·cm⁻²]</i>	<i>AA [%]</i>	<i>RMSE [g·cm⁻²]</i>	<i>AA [%]</i>	<i>RMSE [m²/m²]</i>	<i>AA [%]</i>	
HyMap 2003	SE	0.0082	70.4	0.0025	77.1	1.478	59.1	68.9
	COV	0.0104	62.1	0.0032	64.0	1.427	63.0	63.0
CHRIS Mode 5	SE	0.0121	51.5	0.0036	58.2	1.335	54.8	56.2
	COV	0.0120	55.4	0.0047	43.8	1.405	59.3	52.8
Landsat 7 ETM	SE	0.0072	74.7	0.0027	70.9	1.529	57.6	67.7
	COV	0.0074	73.6	0.0029	74.1	1.440	45.3	64.3
SPOT 5 HRG	SE	0.0060	78.3	0.0032	62.9	1.391	49.6	63.6
	COV	0.0100	62.0	0.0022	76.0	1.399	39.9	59.3
Quickbird	SE	0.0183	23.9	0.0043	50.2	1.545	57.3	43.8
	COV	0.0181	24.5	0.0026	71.2	1.678	44.7	46.8

limited number of spectra and that results may deviate when a more extensive quantification takes place. Moreover, systematic errors between radiative transfer model and measurement may have occurred which are not accounted for by the covariance matrices. Nevertheless, the still unsatisfying accuracy of the estimations show the need for additional regularization, which could be provided by a priori information on the variables.

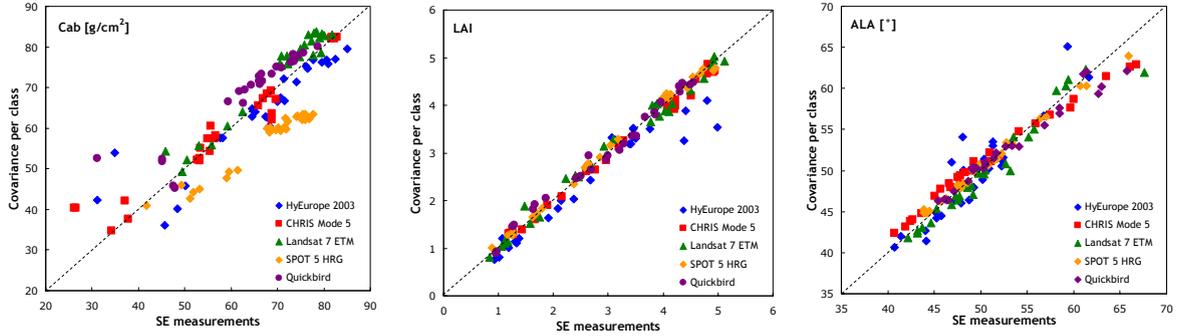


Figure 4.16: Effect of accounting for spectral covariance on estimates of C_{ab} , LAI and ALA . The effect is shown for 5 different sensor configurations. The x-axis represents the estimations when only measurement variance is used, the y-axis when covariance based on spectra within the land cover classes is introduced.

4.3.4.3 Introducing prior information on variables

The effect of integrating a priori information on the variables is studied only for the HyMap 2003 configuration, while radiometric minimization includes a SPECL classification and the spectral covariance description presented in the previous section. Analogous to the process described in Section 3.4, a priori information is explored in a successive step. Initially, the cost function χ^2 , applied for the minimization between variables in the LUT (V_{LUT}^j) and a priori estimates (V_{prior}^j), is a maximum likelihood estimator:

$$\chi = \sum_{j=1}^{n_{var}} \frac{(V_{prior}^j - V_{LUT}^j)^2}{\sigma_{V,j}^2} \quad (4.6)$$

with $\sigma_{V,j}^2$ being the uncertainty associated to the prior estimate of variable j .

In the following example, a priori information and its range of uncertainty is composed by the in situ canopy measurements based on destructive LAI sampling, and therefore very well-defined and close to the actual values. The importance of including prior information is reflected by the results for C_w and C_{dm} in Figure 4.17 (top), which have a significantly higher accuracy compared to the case where only radiometric information is used, with RMS errors being approximately only half of the ones retrieved in the pure radiometric case (rRMSE = 14.8 and 15.1%) and an average accuracy of 88.1 and 88.3%, respectively. In contrast, overall LAI retrieval does not improve (AA = 70.8% rRMSE = 51.0%), which can be explained by the high level of measurement uncertainty $\sigma_{V,LAI}^2$ associated to samples having high leaf area index values and which therefore receive relatively little weight in the cost function. This is better illustrated by the plots at the bottom of Figure 4.17 where the LAI -values measured with LAI-2000 are used as a basis for the prior estimates. In this case, the high uncertainties associated to the measurements of C_w and C_{dm} and the low range of uncertainty attributed to the LAI measurements push the solution in favor of the latter.

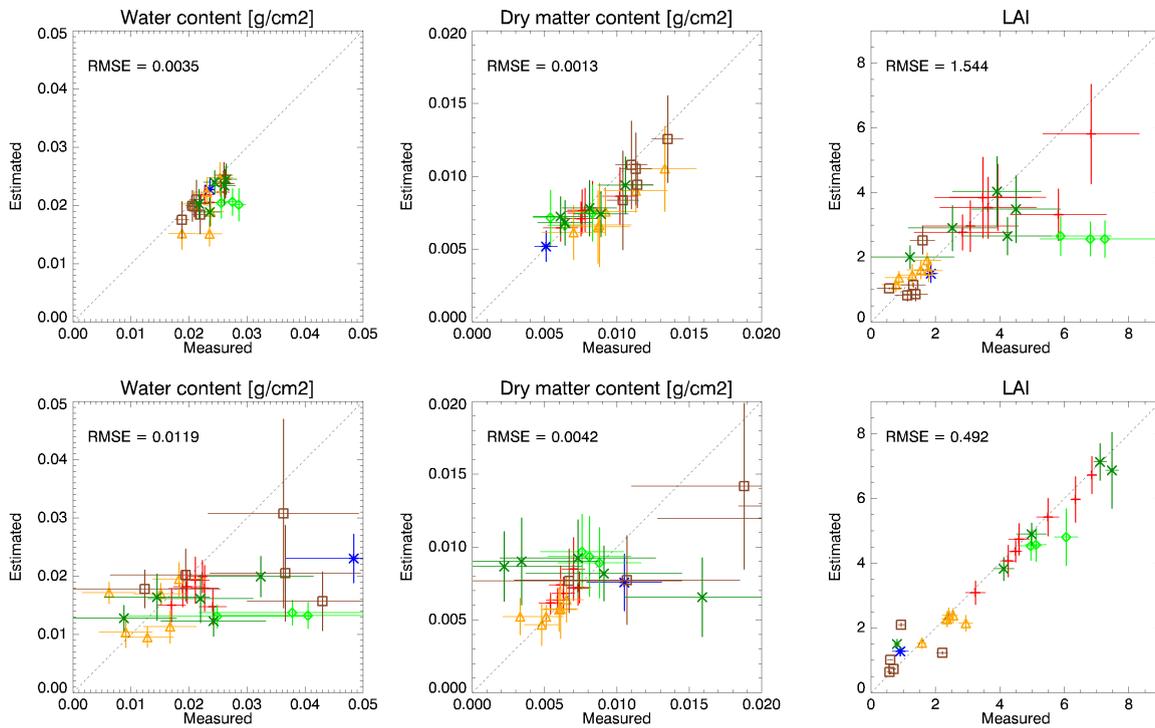


Figure 4.17: RTM inversion results for C_w , C_{dm} , and LAI using LUTs adjusted to SPECL classes and prior information based on field measurements. The top row shows the results when canopy variables based on direct LAI measurements are used, the bottom row shows the results when variables based LAI-2000 measurements are used. Red pluses = HYM1_MEA1, blue asterisks = HYM1_MEA2, green diamonds = HYM1_PAS, yellow triangles = HYM2_MEA1, brown squares = HYM2_MEA2, dark green crosses = HYM2_PAS

4.3.4.4 Introducing covariance between the variables

The demonstrated strong dependency of overall retrieval results on the prior estimates of single variables reveal the importance of well defined prior estimates of all variables, and properly weighting between them. Weighting between the variables can be obtained by introducing their covariance (Equation 3.4.2.3). A sufficient number of data points from similar canopy realizations is required for an explicit characterization of the covariance matrix. In the image based inversion presented in the next section, this information can be retrieved from the prior estimates of pixels falling within a specific land cover class, whereas in the spectrum based approach presented here, it can be approximated by using the canopy measurements found within a single SPECL class (Table 4.9). As can be seen in the table, only class 4 really contains sufficient observations for a reliable description of the covariance. Figure 4.18 shows the results when covariance description between the variables is retrieved from the canopy measurements based on destructive LAI sampling. While results for C_w and C_{dm} become less accurate, there is a prominent improvement in LAI retrievals, being most evident for plots with high LAI . Introducing the covariance between the variables clearly induces a kind of "balancing" between the estimates of the different variables, a phenomenon that is also reflected by the similar levels of average accuracy (and rRMSE) for C_w , C_{dm} , and LAI , amounting to 76.3, 78.2, and 81.7% (29.6, 26.0, 22.0%), respectively. The example reveals the importance of taking into consideration

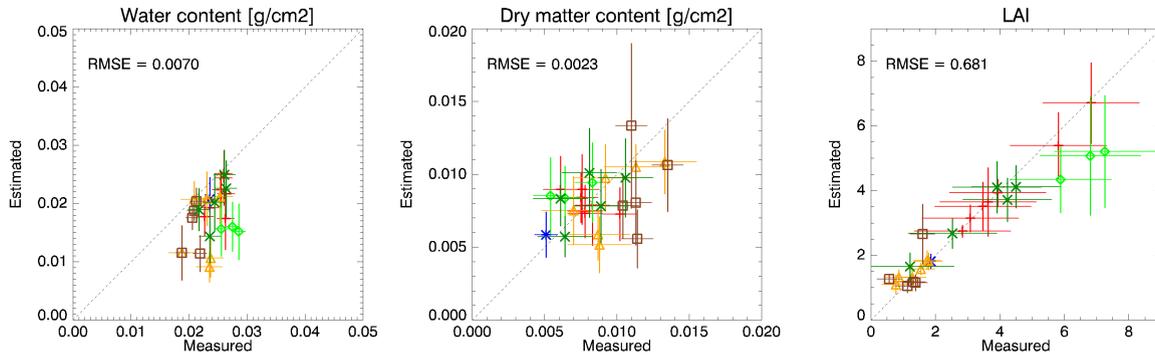


Figure 4.18: RTM inversion results for C_w , C_{dm} , and LAI using LUTs adjusted to SPECL classes and including prior information and covariance between the variables based on field measurements. Red pluses = HYM1_MEA1, blue asterisks = HYM1_MEA2, green diamonds = HYM1_PAS, yellow triangles = HYM2_MEA1, brown squares = HYM2_MEA2, dark green crosses = HYM2_PAS.

the covariance between variables when a concurrent retrieval of several variables is envisaged.

4.3.4.5 Integrating a priori estimates based on predictive regression functions

In practice, a priori information with the accuracy of the field measurements used in the previous example is usually not available, and even if such information would be available, the example showed that the eventual retrieval accuracy of each canopy variable is highly dependent on the accuracy of the single sources of prior information. The challenge is therefore to define prior information in such a way that each canopy variable in the RTM is represented and that estimates are as accurate as possible. The calculation of prior estimates based on predictive regression equations, like proposed in the previous chapter, is at least able to provide an estimate for every variable, although its degree of uncertainty strongly depends on the ensemble of canopy, observation, and illumination characteristics, which, in turn, determines the radiometric effect of each variable. This is illustrated in Table 4.9 where the predictive equations of the variables of interest are given for the current situation. For example, the predictive power of VIs in estimating LAI (expressed by R^2 and RMSE) decreases with increasing "greenness" and density of the canopy. A reverse trend can be observed for leaf water content and, to a smaller degree, for leaf dry matter. The trend observed for C_{dm} is somewhat remarkable since the spectral response of this leaf constituent is usually masked by leaf water in healthy vegetation. The trends in predictability are partly reflected in Figure 4.19, showing more accurate LAI retrievals for plots with low LAI values and slightly better estimates of C_w and C_{dm} for plots with high LAI . The average retrieval accuracy (71.7, 73.2, 66.3% for C_w , C_{dm} , and LAI , respectively) is only little less than in the case where prior information resulting from the field measurements is directly included, especially in the case of C_w and C_{dm} (See previous section).

Introducing prior estimates based on regression functions brings only little improvement compared to the case where only radiometric information from the reduced SPECL LUTs is used (Cfr. Figure 4.15). However, although not validated, chlorophyll estimates seem to give far more realistic values in the case where prior estimates are introduced for all variables, especially for recently cut meadows with a relatively high amount of dead material and stubbles (Figure

Table 4.9: Best predictive equations based on simulated spectra in different SPECL classes. For the acronyms of the selected vegetation indices see: Appendix C. The last column shows the results of the SPECL classification.

<i>SPECL class</i>	<i>Variable</i>	<i>Regression function</i>	R^2	<i>RMSE</i>	<i>Spectra allocated to class</i>
2 dark vegetation	Cw [g/cm^2]	$0.028 + 0.425 * LWVI1$	0.66	0.0097	HYM2.MEA2: 1,3,5
	Cdm [g/cm^2]	$0.011 - 0.082 * LWVI1$	0.50	0.0029	
	LAI [m^2/m^2]	$0.843 * 14.8327^{RDVI} - 1.076$	0.76	0.807	
	Cab [$\mu g/cm^2$]	$16.961 * 127.750^{REIP-Guyot} - 498.107$	0.78	12.6	
3 average vegetation	Cw [g/cm^2]	$0.016 + 0.202 * LWVI2$	0.67	0.0118	HYM1.MEA2: 5
	Cdm [g/cm^2]	$0.014 - 0.113 * LWVI1$	0.57	0.0039	HYM2.MEA1: 1,3,4
	LAI [m^2/m^2]	$0.024 * 13.069^{RDVI} - 1.322$	0.76	0.935	HYM2.PAS: 2,5
	Cab [$\mu g/cm^2$]	$17.843 * 140.726^{REIP-Guyot} - 564.621$	0.86	11.7	
4 bright vegetation	Cw [g/cm^2]	$-0.006 + 0.323 * LWVI2$	0.87	0.0101	HYM1.MEA1: 1-7;
	Cdm [g/cm^2]	$0.016 - 0.111 * LWVI1$	0.58	0.0047	HYM1.PAS: 1-3
	LAI [m^2/m^2]	$1.279 * 8.922^{RDVI} - 0.628$	0.48	1.369	HYM2.PAS: 1,3,4
	Cab [$\mu g/cm^2$]	$-42.007 + 178.939 * LCI$	0.88	10.8	
6 mix soil / vegetation	Cw [g/cm^2]	$0.025 * 39.335^{LWVI2} - 0.001$	0.43	0.0112	HYM2.MEA1: 2,5,6
	Cdm [g/cm^2]	$0.016 - 0.150 * LWVI1$	0.47	0.0058	HYM2.MEA2: 2,4
	LAI [m^2/m^2]	$24.616 * 0.312^{MSI} - 0.748$	0.85	0.411	
	Cab [$\mu g/cm^2$]	$10.0346 + 96.9097 * LCI$	0.59	16.3	

4.20). This can be explained by the well defined a priori estimates of chlorophyll, which are expressed by high R^2 (up to 0.88) and low RMSE ($< 16 \mu g \cdot cm^{-2}$) values. So, even if in this case the introduction of prior estimates derived from predictive regression equations does not significantly influence the estimates of ultimate interest, it certainly seems to improve the ensemble of estimated variables, including chlorophyll, one of the major determinants of canopy reflectance and a key parameter in the assessment of grassland vitality.

Probably the most important contribution of incorporating a priori information is not the improvement of absolute retrieval performance, but the strong reduction of ambiguity, a term used for the high interaction between variable pairs (Baret and Buis, 2007). This is illustrated in Figure 4.21 where ambiguity between the estimated variables LAI , ALA , and Cab is shown for three different sensor configurations and for the case where prior information based on predictive regression functions is introduced. In the case where covariance between the variables is not accounted for (top row), high correlation, expressed by the R^2 values, exists between the several data pairs, especially between LAI and ALA . Ambiguity decreases with an increasing number of spectral bands. Accounting for the covariance between variables shows a dramatic decrease of ambiguity for all variables and sensor configurations (bottom row). Only the correlation between LAI and Cab for the HyMap 2003 configuration slightly increases. A moderate correlation between LAI and Cab is however expected, since the meadows with low LAI were in most cases recently cut and contained a higher fraction of senescent material and stubbles. This example

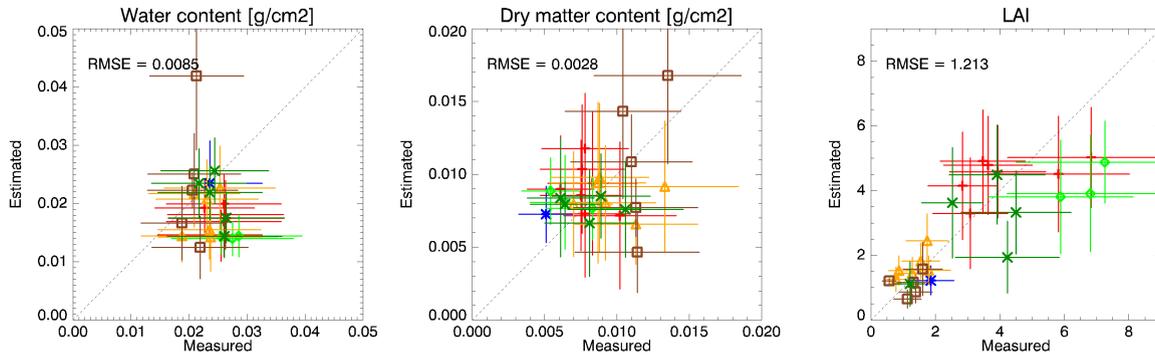


Figure 4.19: RTM inversion results for C_w , C_{dm} , and LAI using LUTs adjusted to SPECL classes and prior information based on predictive equations. Red pluses = HYM1_MEA1, blue asterisks = HYM1_MEA2, green diamonds = HYM1_PAS, yellow triangles = HYM2_MEA1, brown squares = HYM2_MEA2, dark green crosses = HYM2_PAS

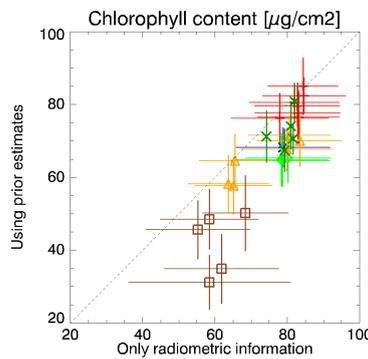


Figure 4.20: Influence of introducing a priori estimates based on predictive regression equations on Cab retrieval. Red pluses = HYM1_MEA1, blue asterisks = HYM1_MEA2, green diamonds = HYM1_PAS, yellow triangles = HYM2_MEA1, brown squares = HYM2_MEA2, dark green crosses = HYM2_PAS

illustrates that, even if for some variables prior estimates are somewhat loosely defined, they still provide considerable stabilization of the algorithm when their interaction is accounted for.

4.3.4.6 Comparing predictions based on regression functions with final RTM estimates

The approach presented in this study actually includes a dual estimation of canopy variables based on radiative transfer modeling: i) using the predictive regression equations based on the RTM simulations in the LUT, and ii) applying RTM inversion directly to the spectra, while prior estimates based on the predictive equations are included as prior information. Theoretically, and as shown in the previous section, the second run is required to account for the interaction between the different variables. However, it might be questioned whether model results are really improved compared to the first estimates based on predictive equations alone, when RTM inversion is additionally implemented.

Table 4.10 compares model results obtained with the predictive equations (i.e., the prior

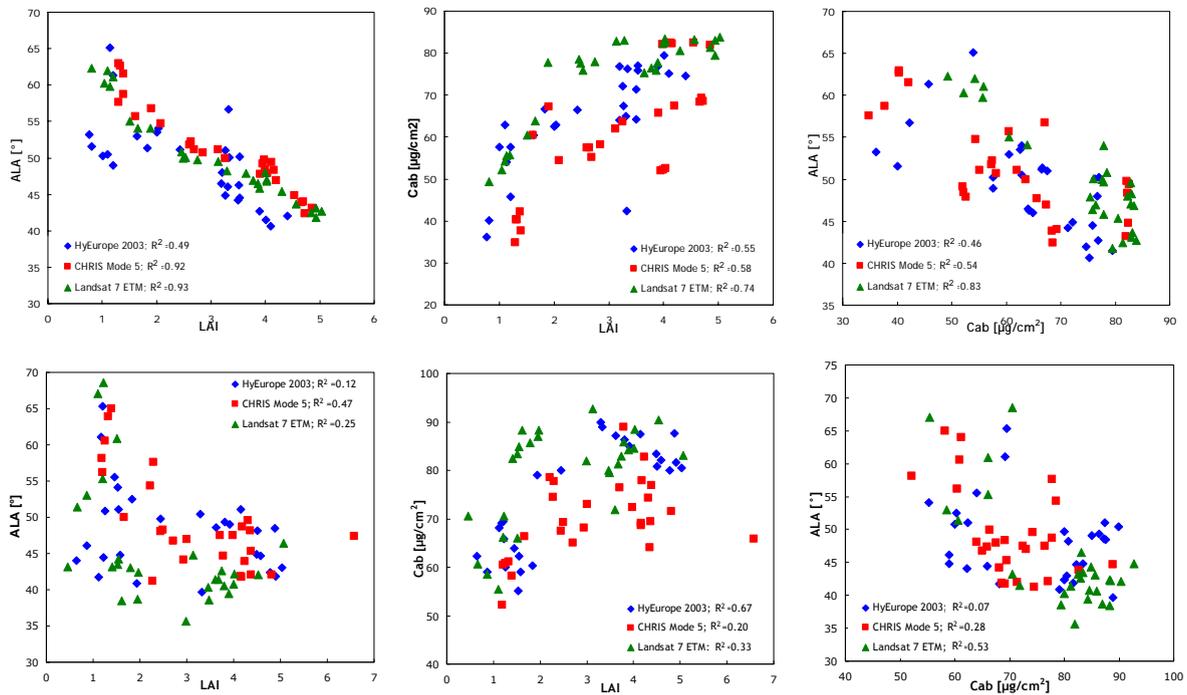


Figure 4.21: Reducing ambiguity by accounting for covariance between the variables. The top row shows the results when minimization in the variable space is based on variance alone, the bottom row the results when also covariance is accounted for.

estimates) with those obtained by implementing the complete algorithm, including RTM inversion. It is noticed that the prior estimates are significantly less accurate than the final results. The differences are graphically presented in Figure 4.22, where it becomes evident that some variables and plots are more affected than others. Whereas estimates of LAI on the whole show good agreement, divergence for the other 2 variables is larger, in particular for $HYM2.MEA2$. Moreover, prior estimates of Cw seem to be significantly overestimated compared to the field measurements and to the values known from literature. As shown in the presented case, the prior estimates do not necessarily dictate the final solution, but rather enable describing the covariance between the several variables, hence leading to stabilization of the retrieval procedure. The results clearly demonstrate the added value of a complete inversion with respect to a sole regression equation, not only in the case of the inversion algorithm proposed in this study, but also compared to empirical approaches based on VIs in general.

4.3.5 Exploring additional regularization

The results presented in the previous section provided some remarkable results concerning the estimates of Cdm and Cw . Despite their very limited unique spectral response (Cfr. Figure 4.12 and Figure 4.13) for the given data set, they were estimated with a reasonable accuracy. However, accuracy of none of the variables is good enough (i.e. maximum error of 10%) to fulfill the requirement posed by many precision farming applications (Harmony et al., 1997).

Table 4.10: Comparing retrieval accuracy for Cw , Cdm , and LAI based on predictive equations (prior) with those obtained by implementing the complete optimization algorithm (RTM)

RWC	Cw			Cdm			LAI			
	RMSE [g/cm^{-2}]	rRMSE [%]	AA [%]	RMSE [g/cm^{-2}]	rRMSE [%]	AA [%]	RMSE [m^2/m^2]	rRMSE [%]	AA [%]	OA [%]
prior	0.0137	57.4	55.1	0.0033	38.2	62.7	1.408	46.5	48.1	55.3
RTM	0.0081	35.6	71.7	0.0028	31.8	73.2	1.213	40.1	66.3	70.4

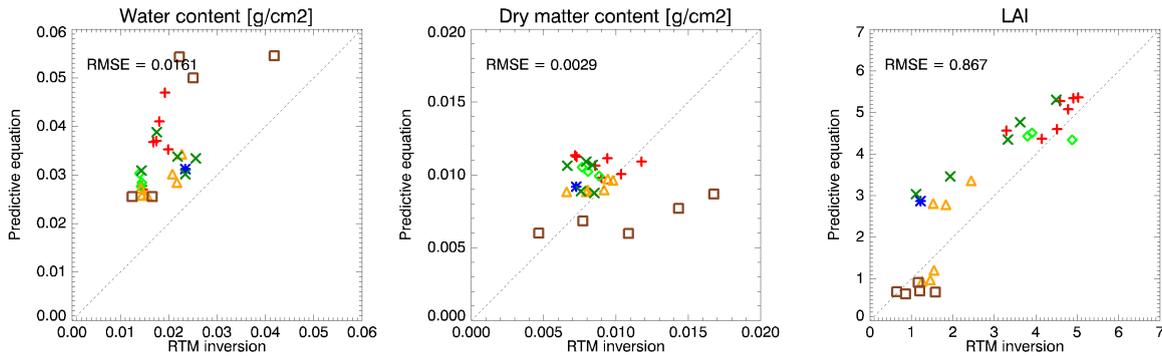


Figure 4.22: Comparing RTM inversion results of Cw , Cdm , and LAI with prior estimates obtained by predictive regression equations. Red pluses = HYM1_MEA1, blue asterisks = HYM1_MEA2, green diamonds = HYM1_PAS, yellow triangles = HYM2_MEA1, brown squares = HYM2_MEA2, dark green crosses = HYM2_PAS

Therefore, in this section two additional regularization techniques are explored for their potential of improving retrieval accuracy.

4.3.5.1 Estimating biochemicals at canopy level

Various authors observed increased accuracy and robustness when synthesized variables were used directly during inversion (Cfr. Section 3.2.4). In particular, the product between the concentration of leaf absorbing materials and leaf area index appears to be a suited candidate, having a physical meaning corresponding to the actual optical thickness of the medium (Weiss et al., 2000). Furthermore, in our case, one of the variables of ultimate interest is the above ground biomass, which, for a square meter, is the product between Cdm and LAI , and can be directly validated with the measured above ground biomass.

Figure 4.23 compares measured with estimated above ground biomass [$g \cdot m^{-2}$], canopy water [$g \cdot m^{-2}$], and LAI for the case in which covariance between variables is calculated from the prior estimates based on predictive equations. At the top row, the results are shown when Cw and Cdm estimated at leaf level are a posteriori multiplied with LAI , at the bottom the result when the products of both Cw and Cdm with LAI ($Cw \times LAI$ and $Cdm \times LAI$) are directly introduced in the cost function. However, for none of the variables, direct inclusion of the product in RTM

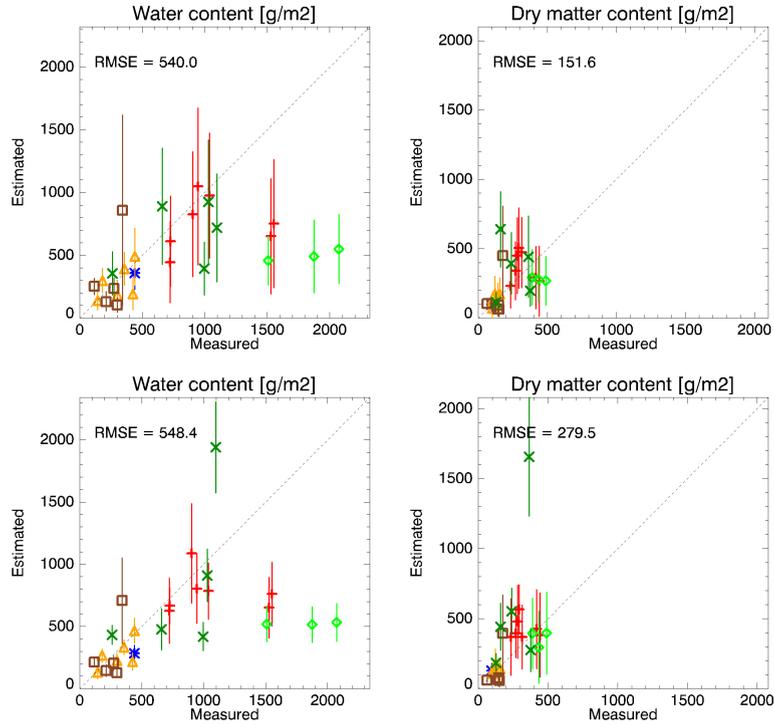


Figure 4.23: RTM inversion results for canopy water and dry matter content using the synthetic variables $C_w \times LAI$ and $C_{dm} \times LAI$. The top row shows the results when the product is calculated a posteriori to the retrieval, bottom row shows the results when the products are included in inversions. Red pluses = HYM1.MEA1, blue asterisks = HYM1.MEA2, green diamonds = HYM1.PAS, yellow triangles = HYM2.MEA1, brown squares = HYM2.MEA2, dark green crosses = HYM2.PAS

inversion leads to significant improvement which can be explained by the very low ambiguity between the single leaf constituents and LAI in the original case ($R^2 = 0.08$ and 0.02 for C_w and C_{dm} , respectively). In addition, it was tested if introducing the product between C_{ab} and LAI ($C_{ab} \times LAI$), which shows moderate correlation with LAI (Cfr. Figure 4.21) brought additional improvement. However, neither in the case where all three products were introduced, nor in the case where only $C_{ab} \times LAI$ was used, significant improvements could be observed for any of the variables (in the latter case: $RMSE=0.0082 g \cdot cm^{-2}$, $0.0037 g \cdot cm^{-2}$, 1.357 for C_w , C_{dm} and LAI , respectively; $AA = 73.1$, 61.7 , and 69.6%). This is probably due to the fact that the covariance matrix between the variables already accounts for proper weighting of the variables in the cost function, thus stabilizing the result. This conclusion is supported by the fact that no significant reduction in ambiguity between C_{ab} and LAI could be observed (not shown).

4.3.5.2 Coupling C_w with C_{dm}

Several authors proposed to tie the contents of leaf dry matter and leaf water at a fixed ratio, usually 1:4 for fresh vegetation (Bacour et al., 2006; Baret et al., 2005b; Vohland and Jarmer, 2007). The advantage of doing such is that the intrinsic dimensionality of the under-determined inversion problem is reduced, which may lead to a more stable and accurate retrieval of all

parameters. It was tested if this fixed relationship would bring improvement for the current situation as well. However, given the significant variation in relative water content encountered in the field measurements (Appendix E), a fixed rapport of $C_{dm}:C_w$ of 1:4 would be unrealistic. Therefore it was decided to repeat the estimation while allowing some degree of randomness in the ratio $C_{dm}:C_w$, based on the natural variation in relative water content (58-84%) measured at the test site. Extending the range to both sides to account for uncertainties in the measurements (i.e., to the range 55-85%), C_w was tied to C_{dm} using the following formula:

$$C_w = (FMC_{min} + \alpha \cdot (FMC_{max} - FMC_{min})) \times C_{dm}, [g \cdot cm^{-2}] \quad (4.7)$$

where α is a randomly selected number from a uniform distribution between 0 and 1, FMC_{min} the minimum foliage moisture content and FMC_{max} the maximum foliage moisture content FMC. FMC is expressed as $C_w \times C_{dm}^{-1}$. In our case $FMC_{min}=1.2$ (55/45) and $FMC_{max}=5.7$ (85/15). This range reflects the actual values found for the canopy under observation but is very large compared to the range of 3.8-4.2 proposed by Vohland and Jarmer (2007), a difference that can be ascribed to the considerable amount of senescent material and different plant species contained within in the plots considered in this study. However, the added value of introducing such a limited degree of freedom as proposed by these authors (reflecting a relative water content between 79.2-80.8%), compared to a complete fixation of the tie, might be questioned in our case.

Table 4.11 summarizes the results of tying C_w to C_{dm} . Both for the completely fixed rapport and for the case where C_w is allowed to vary within the ranges given by the field measurements, overall retrieval accuracy deteriorates compared to the original case where all variables were left free. For the first case the cause is quite clear, since the fixed rapport does not allow for the varying moisture content actually measured in the field. The reason for the reduced accuracy in the second case remains unclear.

Based on the observation that the two commonly applied regularization techniques proposed in this subsection did not yield any improvements compared to the original algorithm where no additional constraints were implemented, it is suggested that introducing covariance description between variables accounts for the maximum attainable reduction of ambiguity between variables. For this reason it was decided to continue the further evaluations with the original set of 9 input variables and the variable sampling plans proposed in Chapter 3.

4.3.6 Model sensitivity to LUT parametrization

4.3.6.1 Reproducibility of estimates

A prerequisite of a high-quality model is the consistency in reproducing similar results for alternating model runs. This is not as trivial as it might seem, since the generation of the LUTs is based on a semi-random selection of the variables. Even if the stratified sampling plan forces the variable selection to take place over the complete domain, within the several intervals there still remains considerable room for variation. The variable parametrization of the LUT does not only influence the optimization between LUT entries and measured reflectance, it also influences the generation of regression fits used for the prediction of the prior estimates. To test the consistency

Table 4.11: RTM inversion results for C_w , C_{dm} , and LAI when using a constrained rapport between C_{dm} and C_w for the generation of the LUTs. Results are shown for fixed relative water content (RWC) of 80% and for the case where RWC is allowed to vary in the range defined by the field validation measurements (55-85%). For the latter, the link between C_{dm} and C_w is established according to Equation 4.7. The 'free' case shows the results when C_w and C_{dm} are untied and was already presented in Table 4.12.

RWC	C_w			C_{dm}			LAI			
	RMSE [$g \cdot cm^{-2}$]	rRMSE [%]	AA [%]	RMSE [$g \cdot cm^{-2}$]	rRMSE [%]	AA [%]	RMSE [m^2/m^2]	rRMSE [%]	AA [%]	OA [%]
80%	0.0076	31.8	78.4	0.0047	54.2	52.3	1.730	57.1	60.1	63.6
50-85%	0.0095	40.1	68.8	0.0034	38.7	61.1	1.371	45.3	65.2	65.0
'Free'	0.0081	35.6	71.7	0.0028	31.8	73.2	1.213	40.1	66.3	70.4

of the variable retrievals, the model was run 12 times on the complete set of 27 field spectra. For each variable output from the model and for each model run, the absolute deviation between the estimated variable (\bar{V}_i) and the mean of all estimations for one plot $\bar{V}_{i,\mu}$ was calculated and in turn averaged over the number of spectra, to form the mean absolute deviation from the average of all estimates, expressed as a percentage ($MD\%$):

$$MD\% = \frac{1}{n} \sum_{i=1}^n \frac{|\bar{V}_{i,\mu} - \bar{V}_i|}{\bar{V}_{i,\mu}} \times 100\% \quad (4.8)$$

where n is the number of plots.

The bar diagram presented in Figure 4.24 shows for every variable output from the model the mean and maximum absolute deviation. The results have been categorized according to the SPECL classification result in order to be able to distinguish trends between the classes. Mean absolute deviation for all parameters is low, usually $< 5\%$, except for most variables in class 2 and LAI in class 6. The maximum absolute deviations show a similar trend, although percentages are significantly higher. While for most variables and classes maximum absolute deviation remains below 20%, some large outliers can be observed, with up to 55% for C_{bp} in class 2. The large variability in class 2 can be explained by the LUT parametrization of this class, which originally had been set up for canopy types with a high shading component, such as row crops. The parametrization is thus not representative for the canopy types that were assigned to this class and which, based on morphological criteria, would better fit into class 6. The relatively high deviations for LAI in class 6 result from the high percentual deviations that small absolute differences can have on low average estimates which in this class are around $1 m^2/m^2$. This can be an indication that the proposed sampling distances are not short enough in variable ranges where reflectance is very sensitive to changes of the variable.

Regarding the individual variables, except for the problematic cases mentioned above, variation between the model runs is low to moderate. On average, the hot spot parameter HS is the least steadily predicted variable, which can be attributed to its limited influence on canopy reflectance in the studied view/observation geometry. The minor spectral variations caused by this variable will therefore serve for filling the gaps between modeled and measured reflectance.

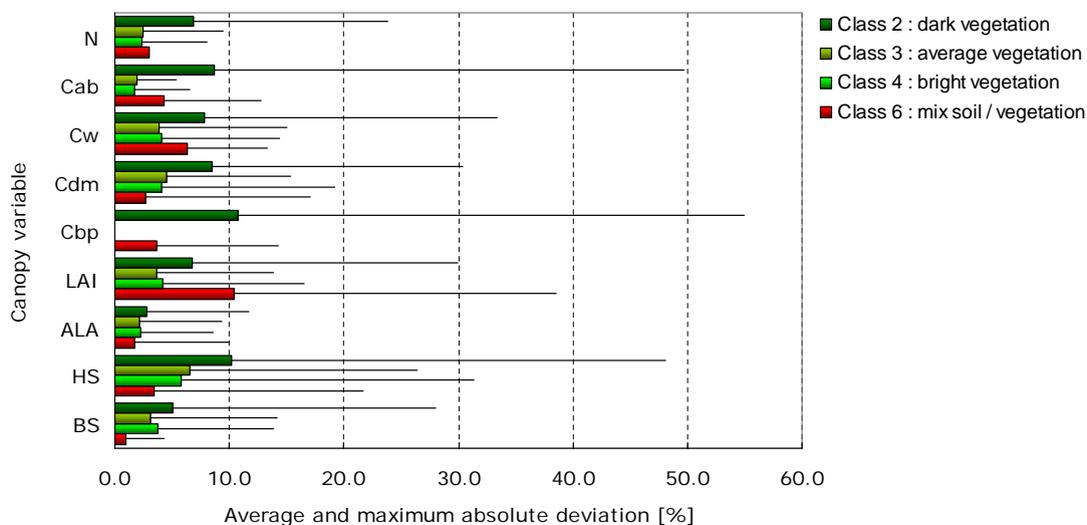


Figure 4.24: Mean (bars) and maximum (indicated with error bars) absolute deviation between single estimated variables and the average result over 12 model runs. For class 2: n=3; class 3: n=6; class 4: n=13; class 6: n=5. For the single plots belonging to each class, see Table 4.9

Most stable predicted is *ALA*, which is also related to the Gaussian parametrization of this variable in the LUT.

The results presented in this paragraph show that, although most variables are predicted in a consistent way, additional stabilization and, hence, improvement can be obtained by a repetitive model run. This aspect is however time consuming and therefore difficult to realize in an operational chain where large quantities of data have to be processed in a short time.

4.3.6.2 Dependence on variable ranges

In Chapter 3, the parametrization of LUTs (i.e., variable ranges and distributions) for the various classes has been based on expert knowledge and existing literature values. It should however be tested if the model is sensitive to these settings and, if this would be the case, whether the distributions from which the variables are selected, are characterized in a way that allows for maximum accuracy of the estimations. Thus, it was analyzed to what degree the estimates depended on the class characterization. This was done for the two most dissimilar classes the measured spectra had been assigned to: class 4 *bright vegetation*, and class 6 *mix of soil and vegetation*. These were at the same time the classes containing most members, with n=14 and n=5 for class 4 and 6, respectively (Table 4.9).

For the variables of primary interest in this study (i.e., *Cw*, *Cdm*, and *LAI*) and for *Cab*, causing the major spectral variation in the VIS, it was tested to what degree the results depended on the range of variation used to construct the LUT. This was done for one variable at the time while the ranges of the other variables were kept constant. Of the original class range of each variable, first the lower boundary (LB) was gradually changed, while keeping the upper boundary (UB) fixed at the predefined value of the respective SPECL class. In a next step, the upper boundary (UL) was gradually changed while LB remained unaltered. The ranges that were

applied can be found in the plots presenting the results (Figure 4.25 and 4.26). Results look different for class 4 and 6.

Class 4: bright vegetation For none of the 4 variables, reducing the LB does really affect the retrieval performance for the measured variables. In contrast, restraining the range of variation (i.e., increasing LB) causes the estimates of the variable subject to variation to become less accurate. This is especially the case when LB approaches the values of the measured variables (Cfr. Table 4.1 and Appendix E). Except for C_{dm} in the case of changing C_w , accuracy of the other variables is only little affected when one of the variables is inadequately parametrized. Remarkable are the results for changing the LB of C_{ab} : highest maximum overall accuracy is obtained when C_{ab} is allowed to vary only between 80 and 100 $\mu\text{g} \cdot \text{cm}^{-2}$, a range that does not seem realistic for grassland species (Vohland and Jarmer, 2007). Also this phenomenon can be attributed to the mismatch between the modeled and measured VIS reflectance previously presented in Figure 4.11. The nearly constant accuracy of LAI retrievals in this case confirms the capability of the algorithm in separating spectral influence of C_{ab} and LAI in the VIS.

Presupposed that the solution is within the LUT range spanned by lower and upper boundary, the position of UB does hardly influence the result. This might be due to the saturating character of spectral reflectance in this class. In contrast to the results found for the LB, the UB of C_{ab} does matter for overall accuracy and significantly influences estimation accuracy of the measured variables, especially of C_{dm} .

Class 6: mix soil/vegetation More than for class 4, the results for class 6 are characterized by some evident optimal ranges, where overall accuracy (rRMSE of the single variables) reaches a maximum (minimum). Extending the range below (above) the LB (UB) decreases accuracy again. This phenomenon is particularly apparent for C_w and C_{dm} and could be ascribed to the increased sensitivity to spectral reflectance compared to class 4. It is noteworthy that a reduced estimation accuracy for LAI also influences the results of C_w while for the opposite case this does not hold true. Actually, estimation accuracy of LAI almost exclusively depends on parametrization of LAI itself, underscoring the dominating role of this variable in spectral reflectance.

In all considered cases, and for both SPECL classes, maximum achievable overall accuracy is around 80%. Using the originally predefined LB and UB values provides an overall accuracy (76.2% for class 4 and 78.0% for class 6) that in almost every case is at, or close to this maximum value. Only at the level of single variables, the variable ranges used to construct the LUTs can still be significantly optimized. It has to be emphasized, however, that the results presented in both figures are far from exhaustive since prediction accuracy depends on the parametrization of all variables simultaneously and not, like in this case from the range of a single variable. Moreover, the analysis was based only on a small number of spectra and canopy measurements representing a limited amount of canopy variation. A comprehensive sensitivity analysis of the model to LUT parametrization should account for all possible interactions and typify all possible land cover realizations. Additionally, this should be done for a wide range of observation/illumination geometries, which would obviously go beyond the scope of this study. Nevertheless, in the next Section, sensitivity of the model to radiometric and angular uncertainties is illustrated with several examples taken from the HyMap data and by comparing the retrieval

4.3. Exploring algorithm potential and constraints using field spectrometer data

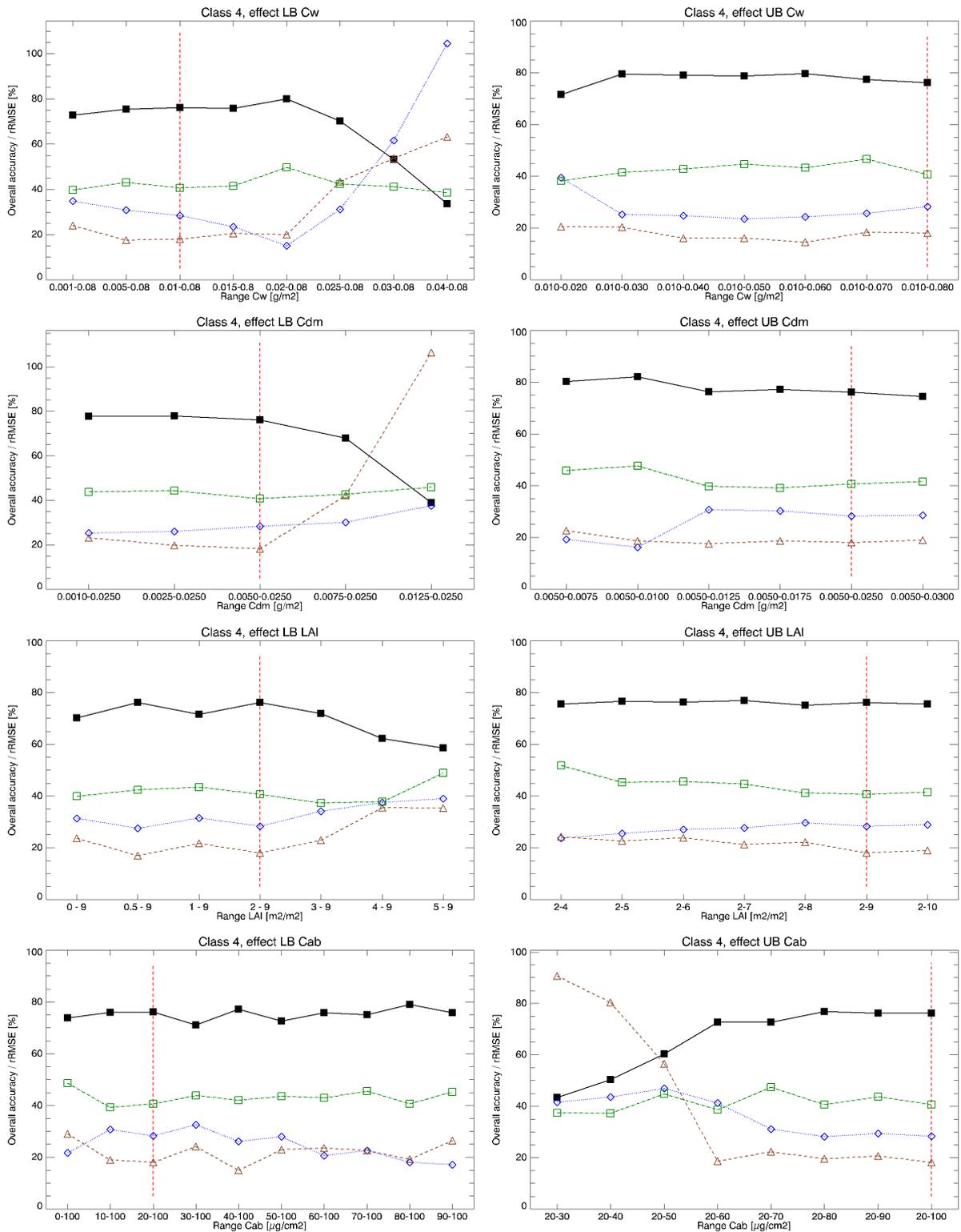


Figure 4.25: Sensitivity of RTM inversion results to changing lower and upper boundary of LUT variable ranges for SPECL class 4. Shown is the rRMSE in % for C_w (blue diamonds), C_{dm} (brown triangles), and LAI (green squares), and the overall accuracy (black squares), which is the sum of the average accuracy for the 3 variables divided by 3. The dotted red line indicates the range as predefined in Appendix B (n=14).

results based on the field spectrometer data with the ones that were retrieved from the HyMap measurements.

4.3.7 Discussion

The original approach, like proposed in the previous chapter, contains two levels of regularization: a restriction of the range of variables contained in the LUT, and the introduction of prior estimates on the variables, obtained from predictive regression equations. Restriction of the LUT ranges in particular lead to better defined estimates, with significantly reduced a posteriori levels of uncertainty. The resampled field spectrometer measurements showed that the benefit of introducing a priori estimates is related to the spectral configuration of the simulated sensor (Table 4.12): while sensors having an even distribution of bands throughout the complete VNIR-SWIR (i.e., HyMap 2003 and Landsat 7 ETM+) perform significantly better compared to the case where only spectral information is used, the accuracy for other sensor configurations is considerably reduced. The latter has several causes. Partly, it is induced by the stabilization of the ambiguity between variables, which is demonstrated by the increased estimation accuracy when covariance description between variables is omitted (Table 4.12, shown in brackets). Probably, also the relatively strong weighting in the cost function of C_{ab} , which in most cases has a high predictability, plays an important role. Another, and maybe the most important reason, is the fact that the limited number of bands of sensors other than HyMap, do not allow for calculating all spectral indices, and therefore lead to less well-defined a priori estimates. This is confirmed by the finding that for the estimation of C_w and C_{dm} based on the HyMap 2003 configuration, for all classes, LWVI1 or LWVI2 was the best performing VI. They both require a band around 1100, and at 1000 or 1200 nm, bands that are not provided by the other sensors. The same applies for the VIs selected for estimating C_{ab} (LCI and REIP_Guyot), which are both based on narrow spectral bands in the red edge. Nevertheless, in cases where sufficient wavebands are available, introducing covariance between the a priori estimates brings additional stabilization of the retrieval process and strongly reduces the ambiguity between the results. The power of this methodology is reinforced by the findings that two commonly used types of additional regularization (i.e., introducing leaf constituents at canopy level, and tying C_w to C_{dm}) did not bring any improvement to the final estimates, neither a reduction of the ambiguity.

The study presented in this section revealed the importance of a well chosen set of wavebands if an adequate, concurrent estimation of several important canopy variables is envisaged. Sensors covering the complete solar-reflective domain give the best overall estimations whereas sensors covering only the VNIR are very weak in predicting C_w , leading to lower accuracy for other variables as well. The similar performance of the HyMap 2003 and the Landsat ETM+ configuration suggests that the 6 bands of the latter are positioned in a way that allows for maximum uncorrelated information content. Nevertheless, the estimates based on the hyperspectral configuration clearly show less ambiguity, an indication that hyperspectral data lead to more robust estimates.

The benefit of accounting for covariance between wavebands is not straightforward. Although in theory one should account for spectral collinearity in order to be able to obtain the real dimensionality of spectral differences between modeled and measured reflectance, the impact this has on estimation accuracy is ambiguous and depends on canopy type and the envisaged variable. The results presented in this section showed that, especially in the case of multispectral sensors,

4.3. Exploring algorithm potential and constraints using field spectrometer data

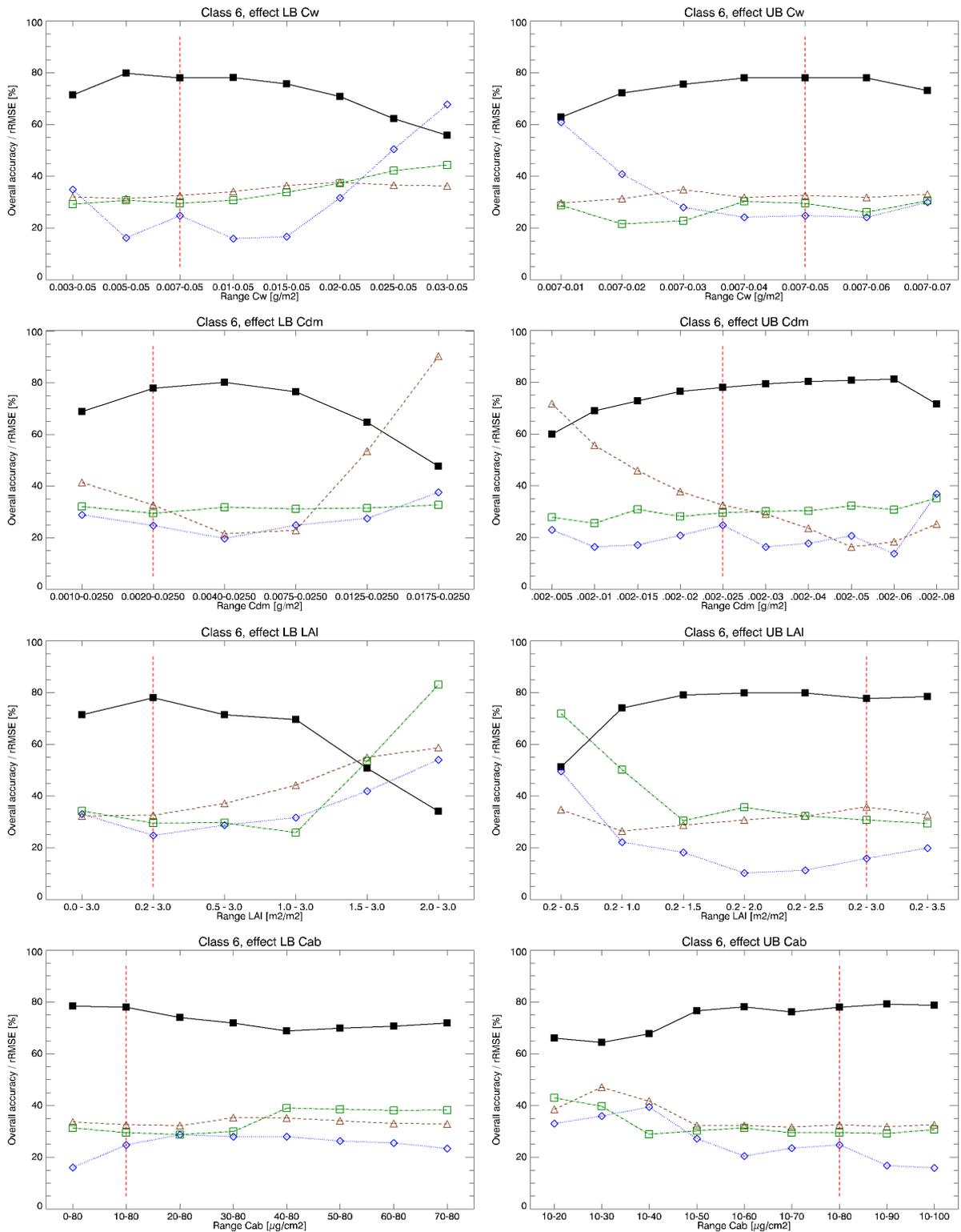


Figure 4.26: Sensitivity of RTM inversion results to changing lower and upper boundary of LUT variable ranges for SPECL class 6. Shown is the rRMSE in % for C_w (blue diamonds), C_{dm} (brown triangles), and LAI (green squares), and the overall accuracy (black squares), which is the sum of the average accuracy for the 3 variables divided by 3. The dotted red line indicates the range as predefined in Appendix B ($n=5$).

Table 4.12: Retrieval performance for different simulated sensor configurations when introducing prior information on the variables, estimated with predictive regression equations. The values without brackets represent the cases when covariance between the variables is introduced, in brackets the case where only the variance per variable is used.

Sensor	C_w		C_{dm}		LAI		Overall acc. [%]
	RMSE [$g \cdot cm^{-2}$]	Average acc. [%]	RMSE [$g \cdot cm^{-2}$]	Average acc. [%]	RMSE [m^2/m^2]	Average acc. [%]	
HyMap 2003	0.0081 (0.0052)	71.7 (82.5)	0.0028 (0.0025)	73.2 (77.3)	1.213 (1.496)	66.3 (61.9)	70.4 (73.9)
CHRIS Mode 5	0.0211 (0.0180)	18.8 (23.8)	0.0042 (0.0035)	52.6 (59.2)	1.505 (1.393)	47.3 (49.0)	39.6 (44.0)
Landsat 7 ETM+	0.0059 (0.0061)	78.0 (78.2)	0.0027 (0.0025)	71.5 (71.9)	1.486 (1.378)	65.3 (52.8)	71.6 (67.7)
SPOT 5 HRG	0.0114 (0.0081)	64.8 (68.4)	0.0040 (0.0041)	58.8 (53.7)	1.732 (1.390)	29.4 (43.1)	51.0 (55.1)
Quickbird	0.0185 (0.0188)	23.9 (22.2)	0.0041 (0.0038)	52.7 (54.8)	1.672 (1.512)	43.7 (49.5)	43.7 (42.2)

introducing spectral covariance deteriorates estimation accuracy, in particular for LAI , probably for the benefit of variables that play a dominant role in the VIS (C_{ab}). In this study, the HyMap 2003 configuration appeared to be the best overall performer, both in terms of accuracy as well as in terms of minimum amounts of ambiguity. Second best was Landsat 7 ETM+. Quickbird, having only 4 wavebands in the VNIR performed worst with lowest overall accuracy and highest RMS errors.

For a better understanding of the estimation accuracy based on the HyMap 2003 configuration, the 27 samples were grouped into the three prevailing phenological types: the samples of HYM1_MEA2, HYM2_MEA1, and HYM2_MEA2 were combined into a class representing recently cut meadows (*meadow-cut*), all PAS samples of HYM1 and HYM2 were grouped to represent the class *pasture*, and all samples of MEA1 represented the class containing mature meadows (*meadow-long*). Based on this subdivision, descriptive statistics were calculated for each group separately. The retrieval accuracy for each class is summarized in Table 4.13, in brackets the results are shown when estimates are compared to the canopy variables based on LAI-2000 measurements.

Meadow-cut According to their spectral properties, the plots within this group were categorized in 3 different SPECL classes (Table 4.9): *dark vegetation*, *average vegetation*, and a *mixture of soil and vegetation*. This reveals the limitation posed by such a crisp class definition, where spectra from plots with similar phenologic states are assigned to classes whose description is based on strongly differing assumptions on vegetative conditions. However, the partly overlap of variable range definitions between SPECL classes should avoid results that show unnatural jumps due to classification differences. For *meadow-cut*, C_w and LAI are estimated with moderate precision (rRMSE = 37.1, and 38.8%, respectively), whereas C_w is estimated with a very high accuracy (rRMSE = 14.1%). The relatively low accuracy of LAI estimates is probably

Table 4.13: Retrieval performance separated according to three phenological classes. In brackets the results for the measurements based on LAI-2000.

Land cover	<i>Cw</i>		<i>Cdm</i>		<i>LAI</i>		Overall acc. [%]
	<i>RMSE</i> [$g \cdot cm^{-2}$]	<i>Average acc.</i> [%]	<i>RMSE</i> [$g \cdot cm^{-2}$]	<i>Average acc.</i> [%]	<i>RMSE</i>	<i>Average acc.</i> [%]	
meadow-cut	0.0031 (0.0144)	88.8 (40.3)	0.0037 (0.0050)	69.1 (50.0)	0.513 (0.719)	61.6 (62.0)	70.1 (50.8)
meadow-long	0.0041 (0.0044)	86.6 (83.1)	0.0010 (0.0031)	88.2 (57.7)	1.519 (1.163)	80.7 (80.5)	85.2 (73.8)
Pasture	0.0075 (0.0137)	77.1 (61.4)	0.0022 (0.0040)	74.9 (36.8)	2.280 (2.108)	60.0 (65.6)	70.7 (54.6)

attributed to the ambiguity in classification. Validation of measurements based on the LAI-2000 performs poor for *Cw* and *Cdm*, and reaches moderate accuracy for *LAI*.

Meadow-long Without any exception, the spectra contained within this group were attributed to the SPECL class *bright vegetation*, a classification that seems to fit well to plots in this development stage. Results for all 3 variables are in good agreement with the field measurements and estimates all fall within one standard deviation from the measured values. Validation of measurements based on LAI-2000 performs well in this class, especially for *Cw* and *LAI*.

Pasture The spectra belonging to the *pasture* group are classified either as *average* or *bright vegetation*, which both seem plausible. *Cdm* is estimated with a very good accuracy, whereas RMS errors for *Cw* and *LAI* are quite high. Deviations of *LAI* increase with increasing value, underscoring the difficulty of estimating high *LAI* values due to saturation of the reflectance signal. In addition, and as already pointed out in Section 4.2, this class suffers from considerable uncertainty in the field measurements. Estimations based on LAI-2000 measurements perform quite weakly.

The categorization of the ensemble of plots into 3 different phenological classes exposes some important information on the proposed inversion approach. First of all, estimation performance appears to be quite stable for *Cw* and *Cdm*, both in terms of RMSE and average estimation accuracy. This is quite surprising as the correlation analysis performed in the previous section pointed out the high collinearity between these leaf constituents and *LAI*. For *LAI* the case looks somewhat different: although average estimation accuracy remains more or less stable, RMSE increases with increasing density of the vegetation cover. From values of 5/6 upwards, the algorithm suffers from saturation effects of the reflectance signal which cannot be solved by the incorporation of prior estimates based on predictive regressions either, since these suffer from saturation effects too. This is a common problem in solar-reflective remote sensing and can only be solved when a priori information from completely independent sources or signals is integrated. Nevertheless, with overall estimation accuracies between 70 and 85%, depending on the canopy type, the results are still quite satisfying, especially considering the completely

automated nature of the approach where no a priori information on variables or land use is known in advance. Besides, obtained overall accuracies are only 5-10% less accurate than in the case where measured variables are directly incorporated as prior information.

It should be kept in mind that the results presented in this section were based on a total of 27 grassland data sets (spectral reflectance and canopy variables), some of them of uncertain quality, especially in the case of the pasture samples. The data set is thus far from representative for all agricultural canopy types. Moreover, the limited number of spectra contained within the single classes may have lead to an unrealistic description of spectral collinearity and covariance between the variables. At least the latter can be addressed in the next Section where CRASH will be employed to HyMaP image data of the area.

4.4 RTM inversion applied to HyMap flight lines

The previous section stretched the potential and constraints of CRASH in estimating canopy variables from spectral data measured at ground level. It was shown that on average accuracy lies between 70-80%. Measuring spectral reflectance at ground level was ideal in terms of geometric and atmospheric uncertainties, as these could be practically excluded. Moreover, no spatial discontinuities of the canopy were involved, since the measured quantities were exactly those responsible for the observed radiometric signal. Transferring the approach to airborne or satellite level brings along the challenge of addressing these radiometric and spatial uncertainties, even as the directional uncertainties involved when the model is deployed to imagery from tilting sensors or sensors with a large field-of-view.

Despite the several drawbacks that have to be overcome, a regional inversion scheme also offers the opportunity for a better spectral and spatial characterization of vegetation and soil. This aspect is of major interest in the radiative transfer model approach presented in this study, since it allows for a more accurate description of radiometric (co-)variance for each land cover class. Besides, local neighborhood information may help to overcome unnatural crisp transitions in estimated values in cases of small-scale classification anomalies (Cfr. Figure 3.10).

The model simulations presented in this section are all based on the complete algorithm like presented in the previous chapter, including the SPECL land cover classification, the use of spectral covariance, and introduction of prior estimates based on predictive regression equations. In contrast to the previous section, characterization of covariance between the prior estimates is in this section based on the class specific covariance plus a local variance term introduced to compensate for potential large systematic errors due to land cover classification (Equation 3.14). Although the generation of the predictive regression equations and the different LUTs were based on distinctive view angle intervals, characterization of the covariance matrices for the different classes was based on the pixels found across the entire scene, and thus on all view angles found within a land cover class. This solution was preferred to circumvent deficient covariance description that could occur when too few pixels of a class are found within a specific view angle interval. Especially for small images, like the subsets presented later on, this chance would be relatively high.

4.4.1 Accounting for spectral anisotropy

4.4.1.1 Quantifying spectral anisotropy

One of the main advantages of using radiative transfer model approaches for the estimation of canopy variables is the possibility of accounting for changing observation and illumination properties. Several studies showed that anisotropic spectral behavior of canopies depends on their structural properties (Bacour and Bréon, 2005; Sandmeier et al., 1998; Sandmeier and Deering, 1999). If this were also the case for the observation/sun geometries encountered in this study, it would make sense to perform model inversion for each class separately, while optimizing the radiative transfer model for prevailing structural class properties. It should first be questioned, however, if adaptation of the model is required for view/sun geometries far from the principle plane while sensor observation angle does not exceed 30° , such as the case for the HyMap flightlines covered in this study.

Spectral anisotropy in reflectance data can be expressed by the anisotropy factor (ANIF) which is simply a normalization with nadir reflectance (Sandmeier et al., 1998; Sandmeier and Deering, 1999):

$$ANIF(\lambda, \theta_s, \psi_s, \theta_v, \psi_v) = \frac{\rho(\lambda, \theta_s, \psi_s, \theta_v, \psi_v)}{\rho_0(\lambda, \theta_s, \psi_s)} \quad [dimensionless] \quad (4.9)$$

where ρ is the bidirectional reflectance factor, ρ_0 is the nadir reflectance factor, λ is wavelength, θ_s (ψ_s) is illumination zenith (azimuth) angle, and θ_v (ψ_v) is viewing zenith (azimuth) angle.

According to the physical mechanisms in vegetation canopies, spectral anisotropy is particularly strong in ranges of high vegetation absorbance such as the visible blue and red chlorophyll absorbance bands. In highly reflective near-infrared bands, multiple scattering processes effectively reduce contrast between shadowed and illuminated canopy components resulting in lower anisotropy effects. In addition, BRDF effects are most pronounced in erectophile canopies with strongly contrasting soil/background reflectance, and are reduced in planophile surfaces (Sandmeier et al., 1998; Sandmeier and Deering, 1999). Both these theoretic assumptions are confirmed by Figure 4.27, which for three different wavelengths, and for the three dominant vegetation classes, shows the anisotropy factors averaged over the entire length of the flightline. Although some fluctuations are observed due to the fact that not every view angle interval contains a representative number of pixels for every class, general trends are very well displayed. For all 3 cover types, BRDF effects show highest dynamics for the red (dotted) and green (continuous) bands, while effects are least pronounced for near-infrared reflectance (dashed). Moreover, for all three sun/observation constellations, anisotropy is significantly larger for the class *dark vegetation* which is known to be constituted mainly of coniferous forest canopies and row-structured crops exhibiting a large amount of mutual shading. Dense, homogeneous vegetation covers (*bright vegetation*) in general show lowest anisotropy. Against expectancy, spectral anisotropy in the forward direction is sometimes larger than in the backward direction. A possible explanation could be the fact that observations so far off the principle plane do not have pronounced backscattering peaks.

Figure 4.28 shows the relative spectral differences³ between the 3 overlapping subsets in the

³ $\delta R_{rel} = \frac{|R_{\lambda WAGINGx} - R_{\lambda WAGINGy}|}{R_{\lambda WAGINGx}} \times 100\%$; R_{λ} is the reflectance at wavelength λ and x/y is the flight line

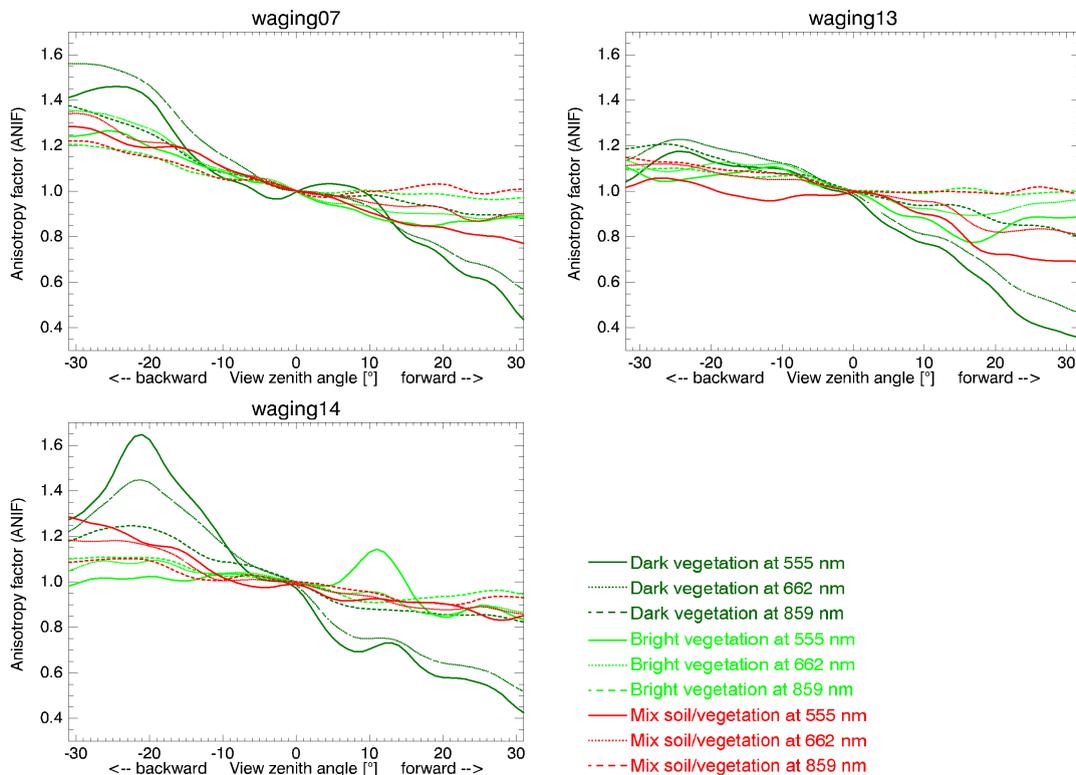


Figure 4.27: Spectral anisotropy factors (ANIFs) for the HyMap scenes WAGING07, WAGING13, and WAGING14. Results are shown for the 3 dominant vegetation classes for green (band 9, 555 nm), red (band 16, 662 nm), and NIR (band 29, 859 nm) reflectance

red and NIR domain. It can be observed that spectral anomalies between the different images are largest at extreme backscatter view angles and are most pronounced for the combination of WAGING07 and one of the two other flight lines. Differences between WAGING13 and WAGING14, which have an opposed view geometry, are smaller and relatively equally distributed across the scene. From this figure it becomes clear that relative deviations are much stronger in the red wavelength. Small inconsistencies in the radiative transfer model itself in this domain, in parametrization of variables active at these wavelengths, or of the employed inversion approach, may therefore lead to considerable inaccuracies in the estimated canopy variables. This is especially true for estimation of leaf chlorophyll content, which is the dominant player in this spectral domain, but would also be transferred to the other variables.

Figure 4.29 shows the influence of angular anisotropy encountered in class *dark vegetation* of WAGING07 on the values of a vegetation index, in this case the Normalized Difference Vegetation Index (NDVI; Rouse et al. (1973)), and *LAI* estimates based on this VI. For a simple empirical equation between *LAI* and NDVI, such as the one proposed by Gardner and Blad (1986) (Equation 4.10) this would already impose a variation up to 15% only due to angular effects, and this for a scene with a view/sun constellation that is far from the principle plane. The example underscores the importance of incorporating information on view geometry, not only during model inversion itself, but also for the generation of predictive equations for the

number.

estimation of prior values on the canopy variables.

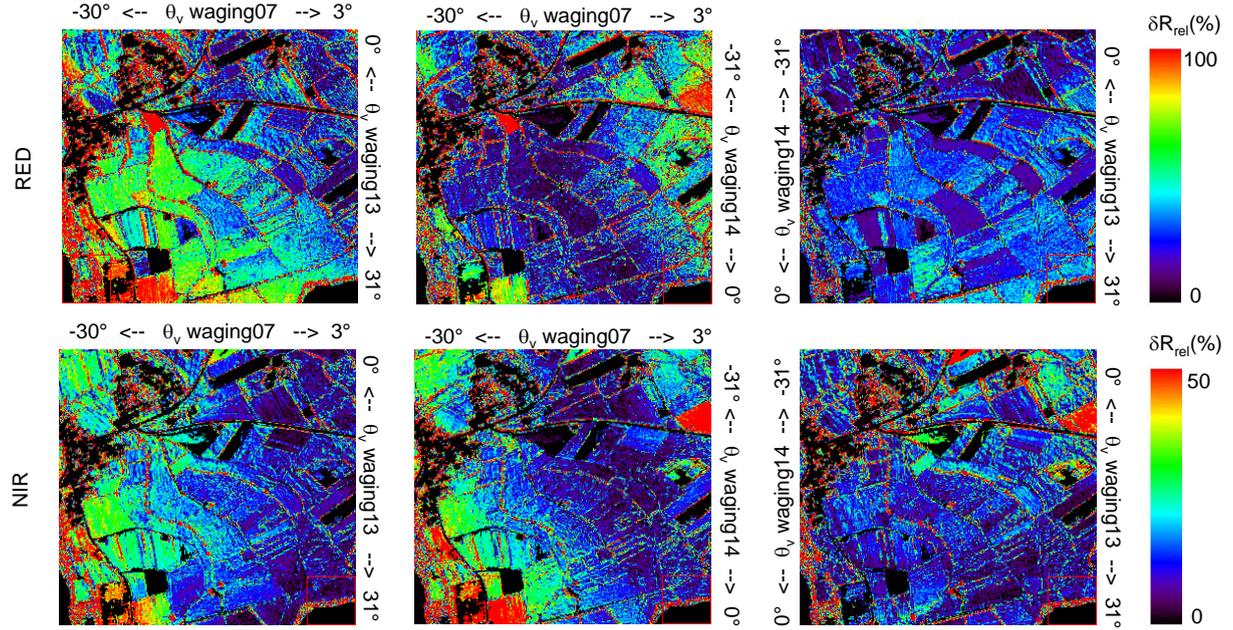


Figure 4.28: Relative spectral differences δR_{rel} in red (band 16; 662 nm) and NIR (band 29; 859 nm) reflectance between the three overlapping subsets. $\delta R_{rel} = \frac{|R_{\lambda WAGINGx} - R_{\lambda WAGINGy}|}{R_{\lambda WAGINGx}} \times 100\%$ where R_{λ} is the reflectance at wavelength λ and x/y is the flight line number. θ_v indicates the view zenith angle of the indicated flightline, where negative values stand for observations in the backscattering direction and positive values for observations in the forward scattering direction. For clarity of comparison, the non-vegetative classes have been masked out.

$$LAI = -1.248 + 5.839 \cdot NDVI \quad (4.10)$$

$$NDVI = \frac{\rho_{859} - \rho_{662}}{\rho_{859} + \rho_{662}} \quad (4.11)$$

Apart from its considerable role in the RTM inversion process, spectral anisotropy also influences the land cover classification, and therewith the inversion results, since the decision rules employed in SPECL rely on spectral ratios and thresholds based on nadir reflectance. Therefore, classification of the image has to be based on nadir-normalized reflectance data. Normalization of the data was performed by dividing the original reflectance cube through the ANIF factors (three of which were shown in Figure 4.27), and is identical to the algorithm offered by ATCOR4 (Richter, 2007a). Radiative transfer model inversion was explicitly applied to the original reflectance data, since nadir normalization would introduce additional radiometric uncertainties, and the spectral integration over distinctive land cover types could level out the unique angular spectral features typical for several vegetation covers.

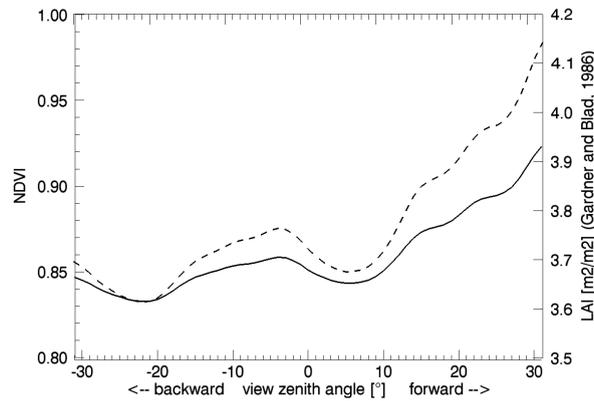


Figure 4.29: Effect of spectral anisotropy on NDVI-based LAI estimates for scene WAGING07. Continuous curve shows the NDVI values, the dashed curve LAI estimates based on NDVI using the regression function proposed by Gardner and Blad (1986) (Equation 4.10).

4.4.1.2 Incorporating view angle information in model inversion

The anisotropy plots shown in Figure 4.27 clearly demonstrate the need to accurately account for angular anisotropy in the HyMap data, even when observations are relatively far from the principle plane. Moreover, they show the need to distinguish between different land cover types, since each of them exhibits characteristic anisotropy behavior. The question still remains whether the radiative transfer model is able to accurately mimic the angular behavior of the canopy and if so, whether the land cover classes are parametrized correctly thus being able to describe such behavior. In other words, if both radiative transfer model and parametrization are correct, and the inversion algorithm is robust, inversion of any selected pixel in the intersecting area of flightlines WAGING07, WAGING13 and WAGING14 should provide very similar results for all three view/sun constellations. This was evaluated by applying RTM inversion independently to the three subsets while view angle information was included according to procedure described in Section 3.3.3.5: a separate LUT was generated for every 3° view angle interval using the scene specific solar geometry. Estimated canopy variables were compared for 27 fields throughout the entire scene by calculating the average of a 5×5 pixels subset to account for co-registration errors (Figure 4.30).

Results of the intercomparison are shown in Figure 4.31 for leaf variables, and in Figure 4.32 for canopy variables. Very good agreement is obtained for Cab , LAI , and ALA , whereas Cw shows moderate to good concurrence between the various sun/observation geometries. Estimates of Cdm , one of the primary variables of interest, show very little congruence, even as the hot spot parameter HS . This can be explained by the fact that influence of these variables on overall canopy reflectance is relatively low and often masked by other variables (e.g., Cw). Nevertheless, the inconsistency of Cdm retrievals is higher than those obtained at ground level for this variable. In contrast, the high variability of HS estimates had already been diagnosed at ground level (Cfr. Figure 4.24). The soil brightness parameter BS coincides well for high values and poorly for low values, a trend that is expected, given the sensitivity of this parameter to exposed (and therefore brighter) soils at low vegetation covers. Positively surprising is the leaf mesophyll structure variable N for which, except for a few outliers, coincidence is generally

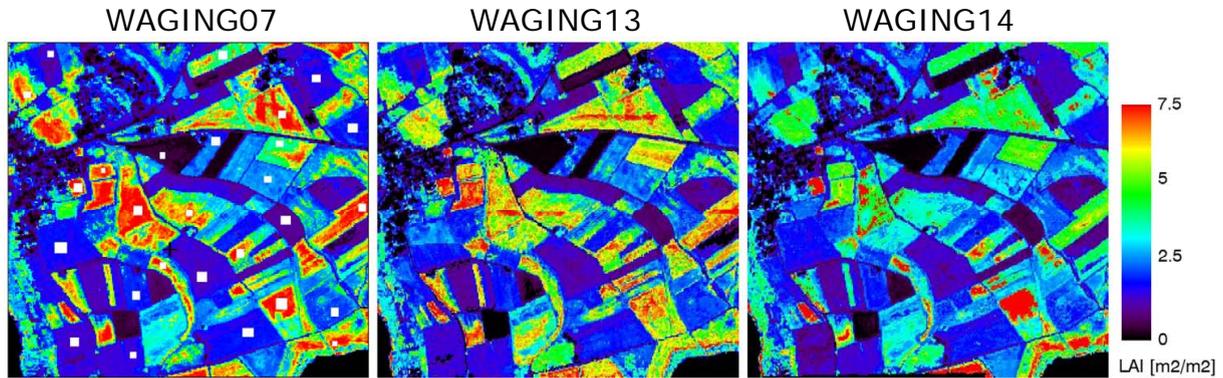


Figure 4.30: Intercomparison of LAI estimated from different observation/sun constellations. for observation/illumination details see Table 4.3. The white squares in the WAGING07 scene indicate the locations used for the intercomparison plots in Figure 4.31 and 4.32.

good.

The stability of the estimates generally follows the trend already observed for the predictive regression functions (Table 4.9), with stable retrieval performance for variables that have dominating influence on one or more domains of the spectrum, and poor stability for variables having only minor influence. For the variables of prime interest this affects especially the reliability of C_{dm} estimates.

It was also studied if there was any relationship between anomalies in estimated canopy variables and spectral deviations resulting from differing view/sun geometries (Figure 4.28). However, for none of the variables, any consistent trend could be observed between spectral dissimilarity and retrieval anomaly (not shown). This indicates that the canopy radiative transfer model (SAILh) is capable of correctly mimicking the angular anisotropy observed in the subsets and that the explanation for the divergence has to be sought elsewhere. Apart from the possible causes already suggested in the previous paragraphs, it would be valuable to study the relationship between retrieval bias and the classification results of the single subsets.

4.4.2 Influence of land cover classification on retrieval performance

The different sun/observation constellations of the HyMap flightlines allow us to study more in detail the influence of radiometric discontinuity on the SPECL classification and on retrieval performance. Figure 4.33 shows the SPECL classification results of the three intersecting subsets. It can be seen that, despite the nadir normalization of the reflectance data prior to classification, still significant anomalies exist between the different observation geometries. There appears to be some confusion especially between average and bright vegetation, and for canopies where the background has a moderate influence. Classification of surface types belonging to the latter category vary from *mix soil/vegetation* to *dark* and *average vegetation* which, for example, is visible for example in the southwestern and northeastern part of the subset. Similar classification bias dependent on changing view/sun geometry was observed by (Beisl, 2001), who found a decreasing accuracy with increasing solar zenith angle.

The effect of classification anomaly on estimation accuracy of LAI and C_{ab} is shown in Figure 4.34 which shows the relative deviation as a percentage (rD) between WAGING07 and

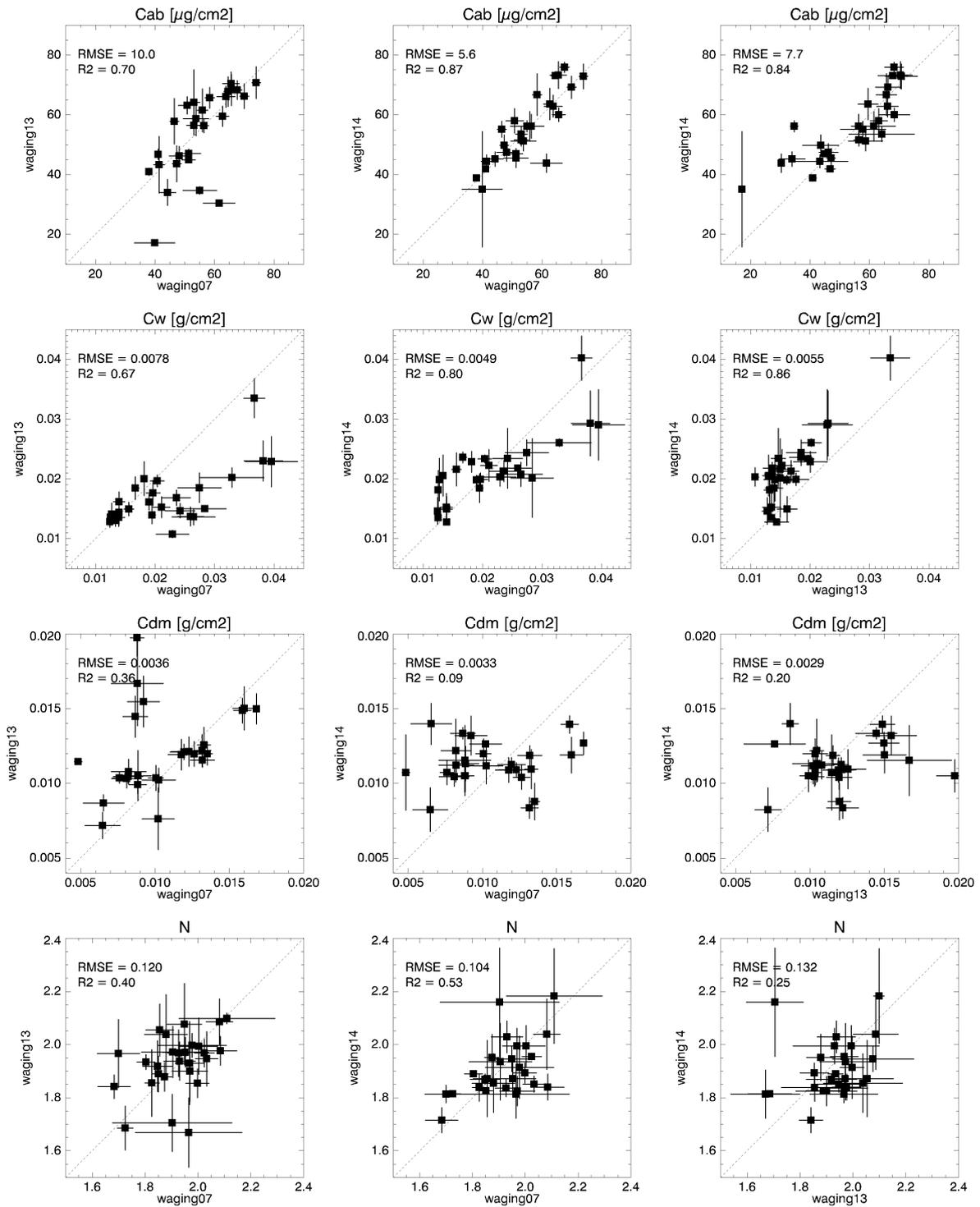


Figure 4.31: Comparison of leaf variable estimates from from 3 HyMap scenes with varying view/sun geometry; *Cab* = leaf chlorophyll a+b content, *Cw* = leaf water content, *Cdm* = leaf dry matter content, *N* = leaf mesophyll structure parameter. Error bars indicate the variation of the variable within the 5×5 pixels window.

4.4. RTM inversion applied to HyMap flight lines

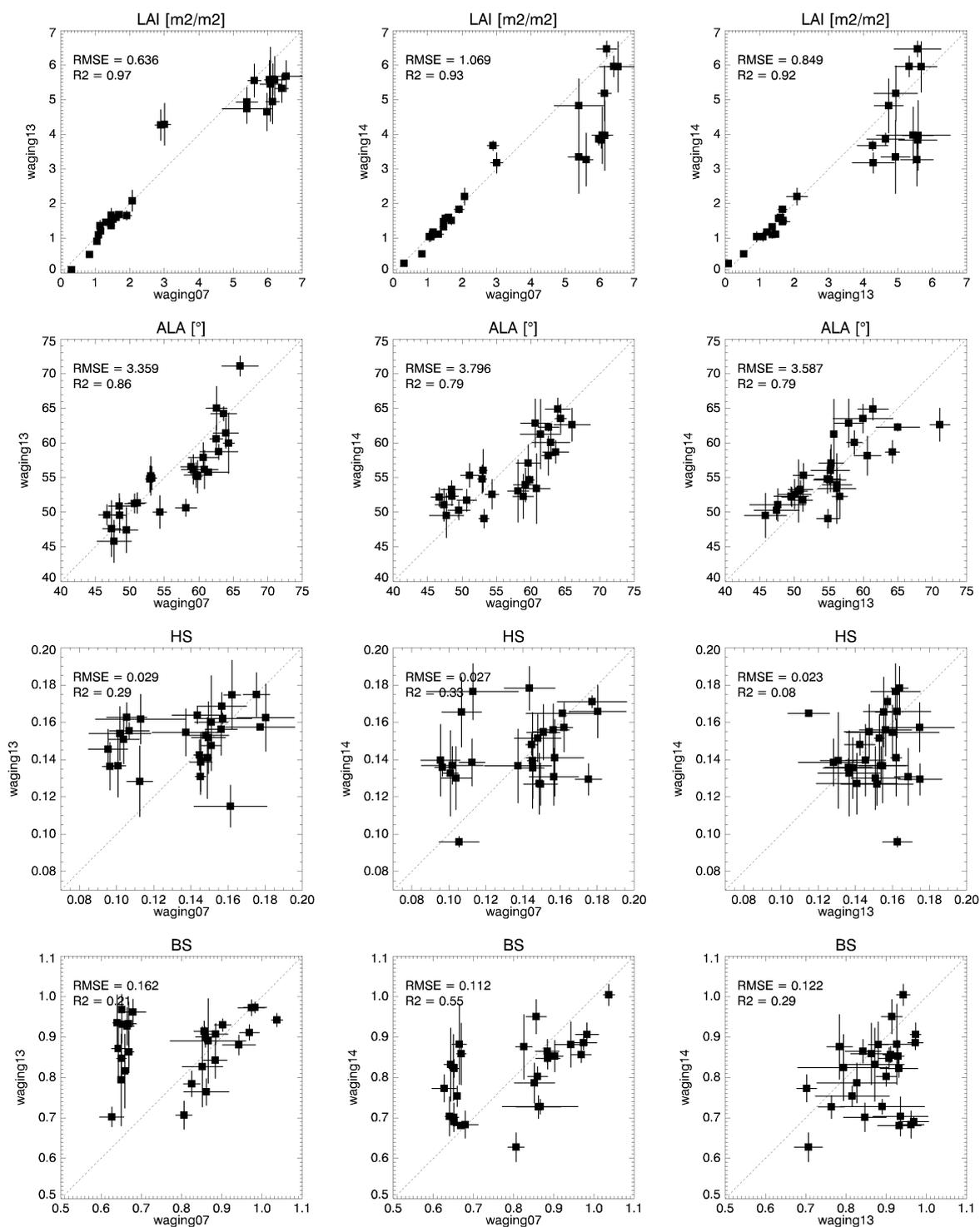


Figure 4.32: Comparison of canopy variable estimates from 3 HyMap scenes with varying view/sun geometry; *LAI* = leaf area index, *ALA* = average leaf angle, *HS* = hot spot parameter, *BS* = soil brightness parameter. Error bars indicate the variation of the variable within the 5×5 pixels window.

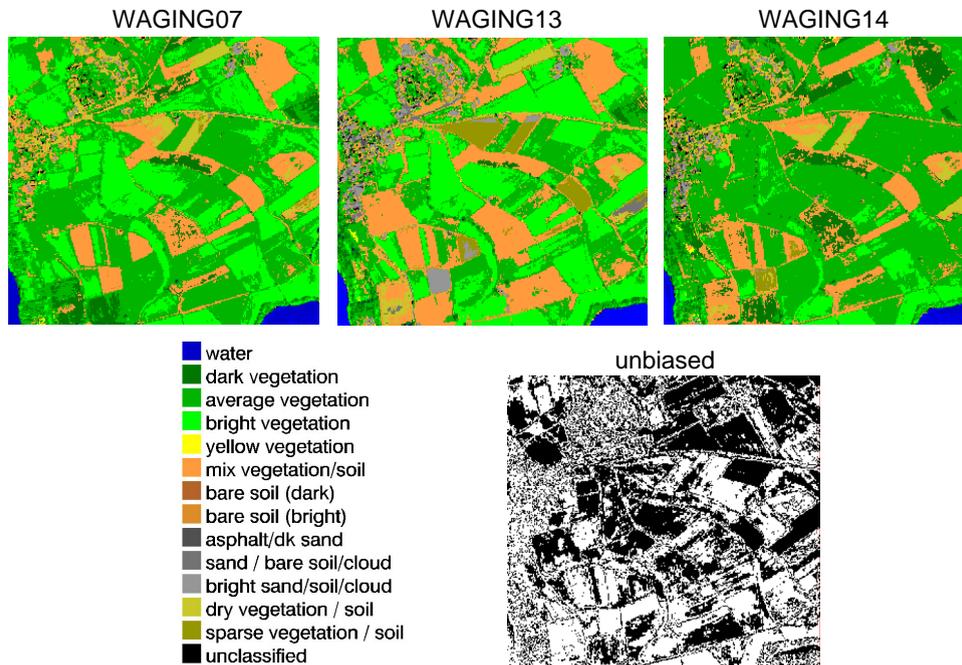


Figure 4.33: Results automatic SPECL classification of nadir-normalized HyMap data of June 30, 2003. The image at the lower right shows the areas (white) that are categorized into the same class for all 3 subsets.

WAGING14⁴. It appears that deviating results are strongly correlated to classification ambiguities between *average vegetation* and *bright vegetation*. This is somewhat surprising as these classes show a broad overlap in parametrization of the LUTs. The answer has to be sought in the low spectral response of the dense canopies belonging to the latter class, leading to less stable *LAI* predictions and, as a consequence, to an altered covariance description. As can be observed in the figure, classification confusion between classes with a soil fraction (e.g., the two fields in upper right and the lower left corners of Figure 4.33) is almost completely resolved during processing and does not show higher deviations than unambiguously classified areas. This contrasts with the results found at ground level for which estimates were least consistent for the classes *dark vegetation* and *mix soil/vegetation*.

Figure 4.35 shows the complication of the SPECL based inversion scheme in the case of smooth gradual reflectance shifts observed within a single field. In this case, one and the same field is classified into 3 different classes. The crisp classification poses a large unnatural jump in *LAI*-values at the transition from one class to the next. Although the characterization of local variance presented in Equation 3.14 is able to reduce some smaller attribution errors, larger systematic artefacts, such as the ones posed by gradual transitions, are only partly compensated.

4.4.3 Inconsistencies and benefits introduced by the predictive regression functions

In Figure 4.30, displaying the spatial distribution of *LAI* estimates for flightline WAGING13, some horizontally structured artefacts could be observed for fields with high *LAI* values (center

⁴ $rD = \frac{|WAGING07 - WAGING14|}{WAGING07} \times 100\%$

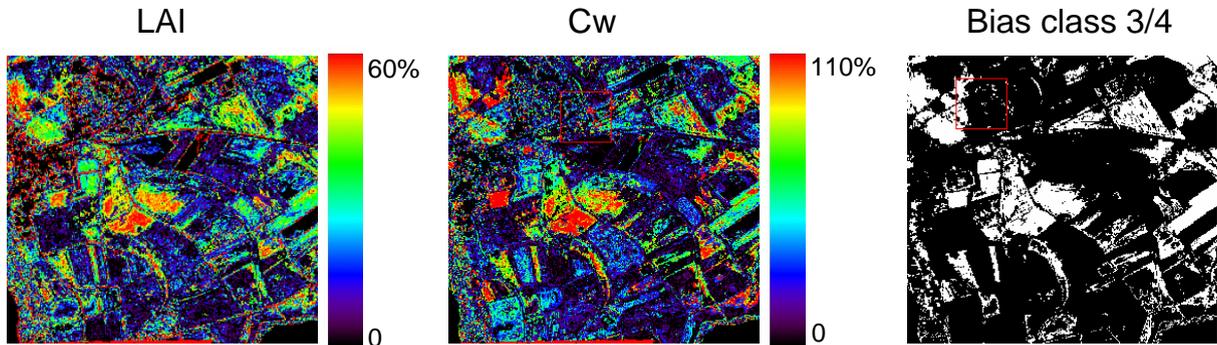


Figure 4.34: Correlation between biased classification results and deviations in LAI and Cw retrievals. The images show the relative absolute deviations between WAGING07 and WAGING13. The white areas in the right image indicate the parts where classification shows confusion between class 3 (*average vegetation*) and 4 (*bright vegetation*).

of image). It seems that the origin of these artefacts lies in the jump of prior estimates from one view interval to the next, and occur specifically at locations with high a priori estimates of Cab and LAI . The reason why they happen in particular for these combinations, has to be sought in the saturating signal occurring at high variable values. Accordingly, the predictive regression functions are highly susceptible to small changes in reflectance. Nevertheless, inconsistencies between adjacent view intervals amount only a few percent and do not occur for all transitions and variables (Figure 4.36). Surprisingly, some of the anomalies in the final results are larger than for the prior estimates themselves, and do not occur *between* one interval and the next, but lie *at* the transition itself. Apparently, this effect is induced by the increased variance within the predefined local window in the cost function. While some transitions are emphasized, most anomalies completely disappear so that jumps between view angle intervals that were present in the prior estimates, are no longer visible.

Even if the inclusion of prior estimations sometimes poses some hurdles, integration of such information, and especially the covariance description based on this, remains important, particularly for the reduction of the ambiguity between different variables in the solution. This had already been observed at ground level and is confirmed by Figure 4.37 which shows the relationship between LAI and ALA for all vegetation pixels in subset WAGING07 for the prior estimates based on predictive regression equations (left) and for the final estimates based on RTM inversion which includes the covariance between the model variables (right). Aside from reducing the ambiguity between the estimates, the final estimates seem to decouple the estimates from classification results, a trend that is still very well visible for the prior estimates. A similar reduction of ambiguity, although not as obvious, was observed for the combinations $LAI - Cab$ and $Cab - ALA$ (not shown).

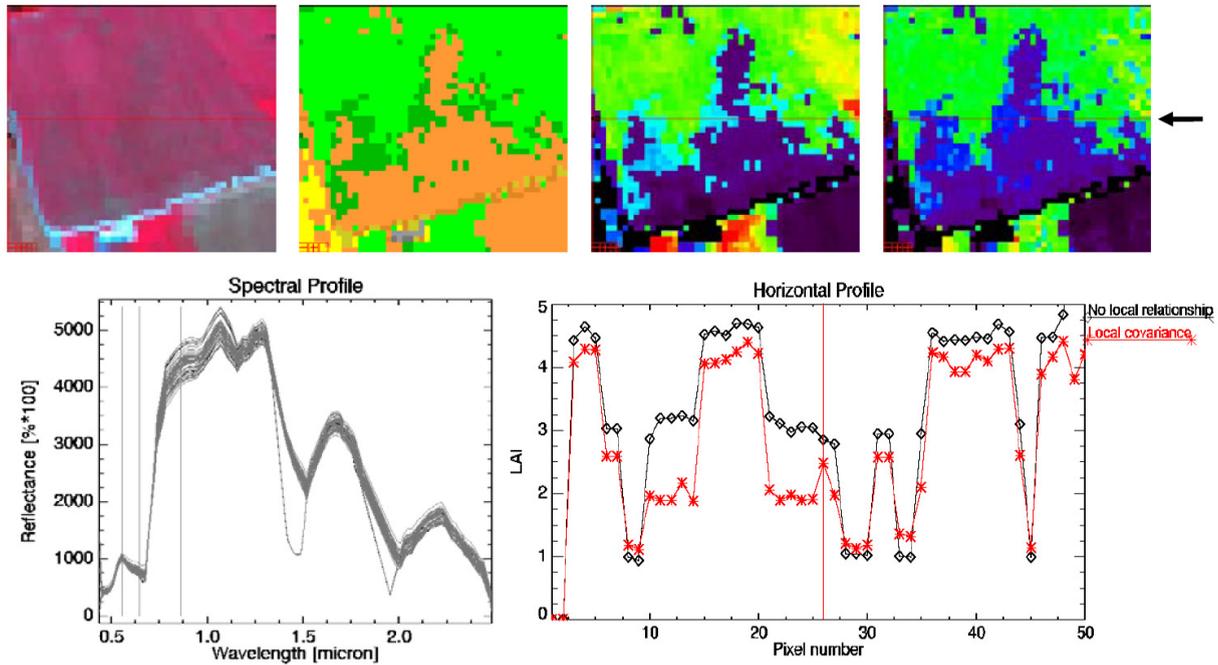


Figure 4.35: Example of effect of land cover classification and inclusion of local variance on variable retrieval. The CIR image shown at the upper left is characterized by gradual spectral changes (lower left). The crisp SPECL classification (2nd image on top row; for legend see Figure 4.33) leads to unnatural transitions in retrieved LAI (3rd image top row). Introducing local variance in the covariance matrices only partly polishes these artefacts (4th image top row). This is elucidated by the LAI profile taken along the line indicated by the arrow (lower right).

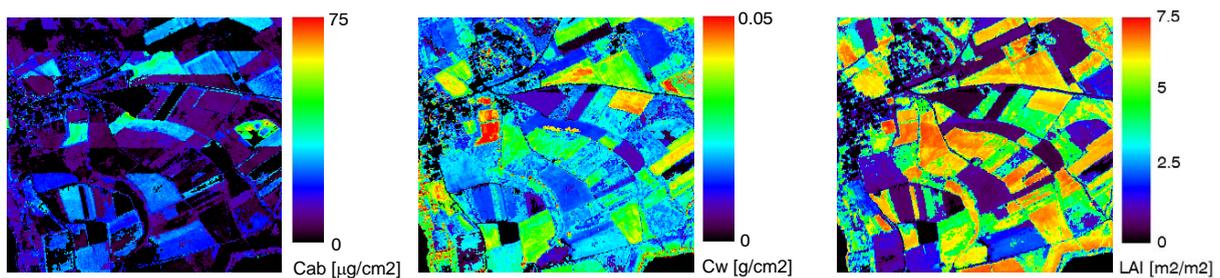


Figure 4.36: Prior estimates of Cab , Cw , and LAI for flightline WAGING13

4.4.4 Validation at test sites

4.4.4.1 Observed patterns

In the previous subsections, several general tendencies concerning the variable estimates were highlighted. Major concern were the anomalies in estimates between different view/sun constellations in cases where bias occurred between land cover class 3 (*average vegetation*) and 4 (*bright vegetation*). This bias is also present for MEA1, one of the validation fields considered in this study. Figures 4.38 and 4.39 show for flightline WAGING07 the results that were obtained at

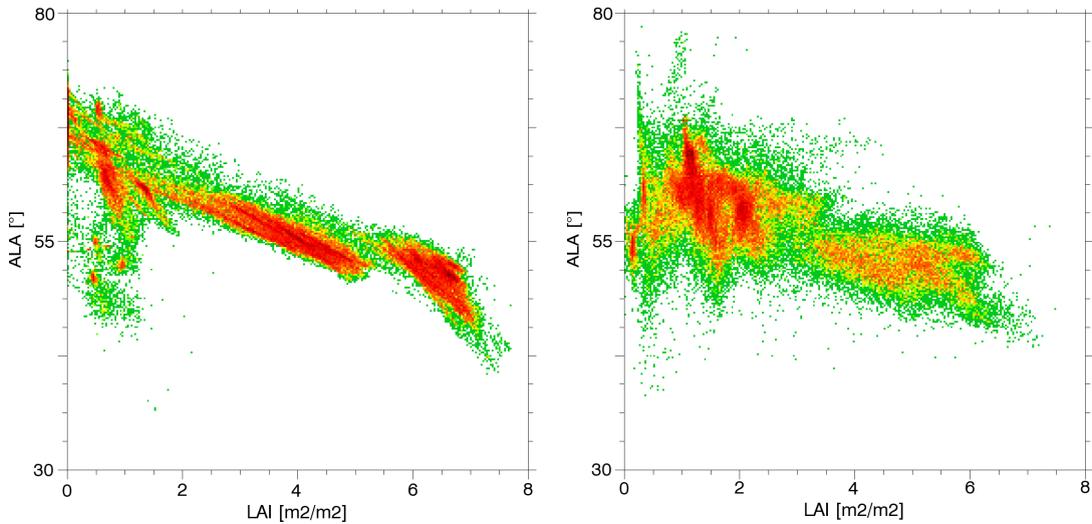


Figure 4.37: Relationship between estimates of LAI and ALA for prior estimates (left) and final RTM inversion results (right) of subset WAGING07. The density of the samples ranges from single points (green) to dense (dark red) ($n=85311$).

the 3 test fields indicated in Figure 4.3 for leaf and canopy variables, respectively.

Leaf variables Cab estimates are highest for MEA1, a result that is expected given the high vitality of this field. On average, values for MEA2 seem too high, knowing that this field contains a considerable fraction of harvest remnants and stubbles. Uncertainties in predicted Cab are relatively low, a result that is in accordance with the high canopy spectral response to this variable and with the stability of the estimates over several model runs (Figure 4.24). The tendency in Cw estimates is opposite to what one would expect, with higher values at MEA2 than at the fields with vital, green vegetation. While for the first field estimates are in line with the measured values, MEA1 and PAS seem to suffer from underestimates. For most pixels, Cdm seems in good agreement with the field measurements, whereas for part of MEA1 they seem quite high. Little is known about the accuracy of Cbp estimations, although uncertainties are very high. The black areas coincide with the classes for which this variable was fixed at zero.

Canopy variables Patterns observed for LAI match quite well expectancy, although in the length direction of MEA1 they seem to drop off to values that are too low for this field. Uncertainties, are on average between 20 and 30% with a very irregular distribution for the pasture, which probably reflects the patchy vegetation structure of this land cover type. ALA is highest for MEA2. Although this coincides with the erectophile structure of the cut meadow, there also seems to be a moderate inverse correlation between ALA with LAI , a tendency that already had been observed in Figure 4.37. Uncertainty for ALA estimates is low. The uniform distribution of the HS parameter concurs with the homogeneity of the vegetation species, although values seem somewhat high. However, too little is known about the actual values of this variable to draw any stakeholding conclusions. Trends for the soil brightness parameter BS seem to be in

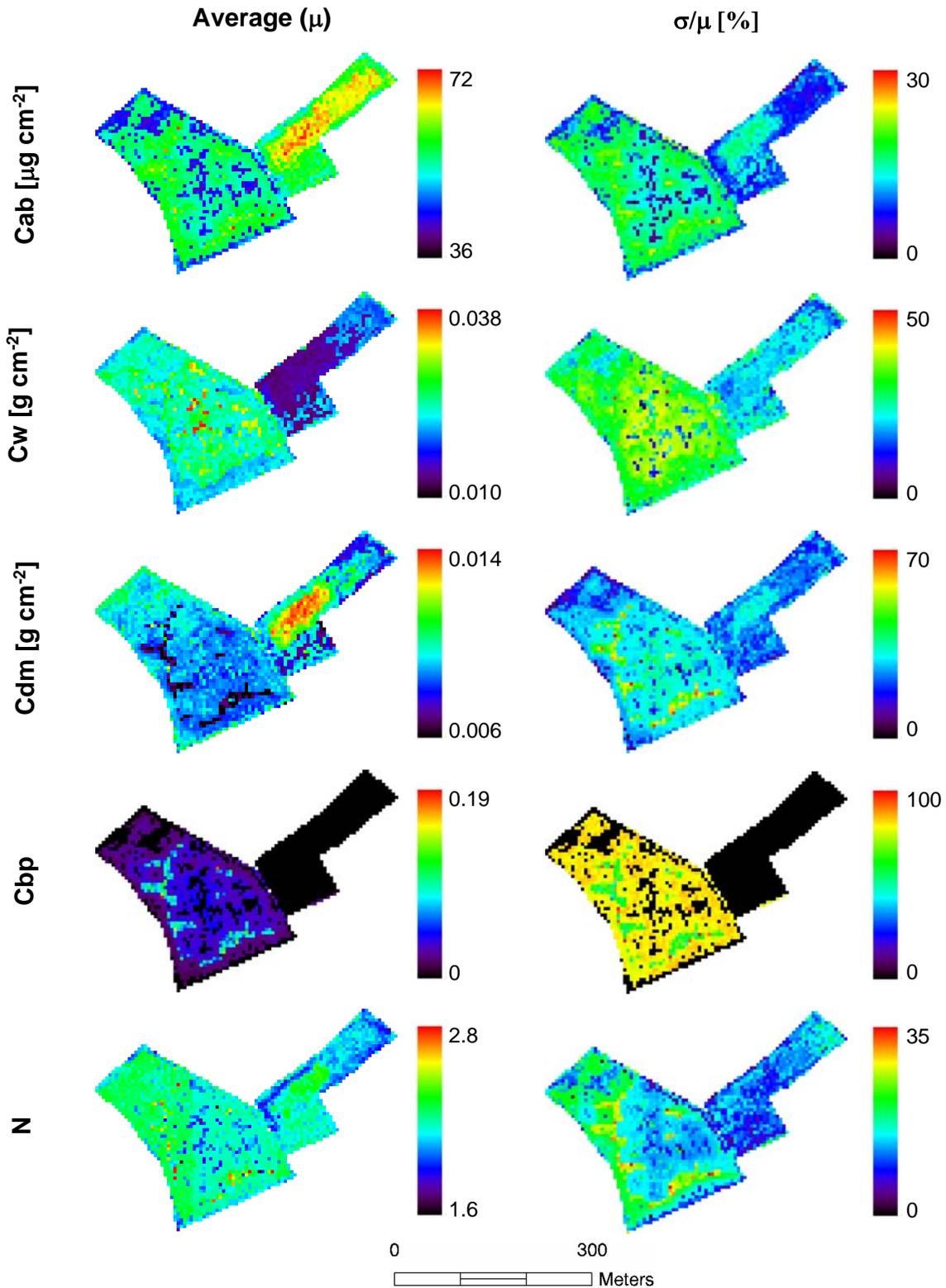


Figure 4.38: Spatial distribution at MEA1, MEA2, and PAS of leaf variables estimated from flightline WAG-ING07. The left column shows the average estimated values μ , the right column shows the standard error σ divided by μ and multiplied with 100%. For the exact delineation of the fields, see Figure 4.2.

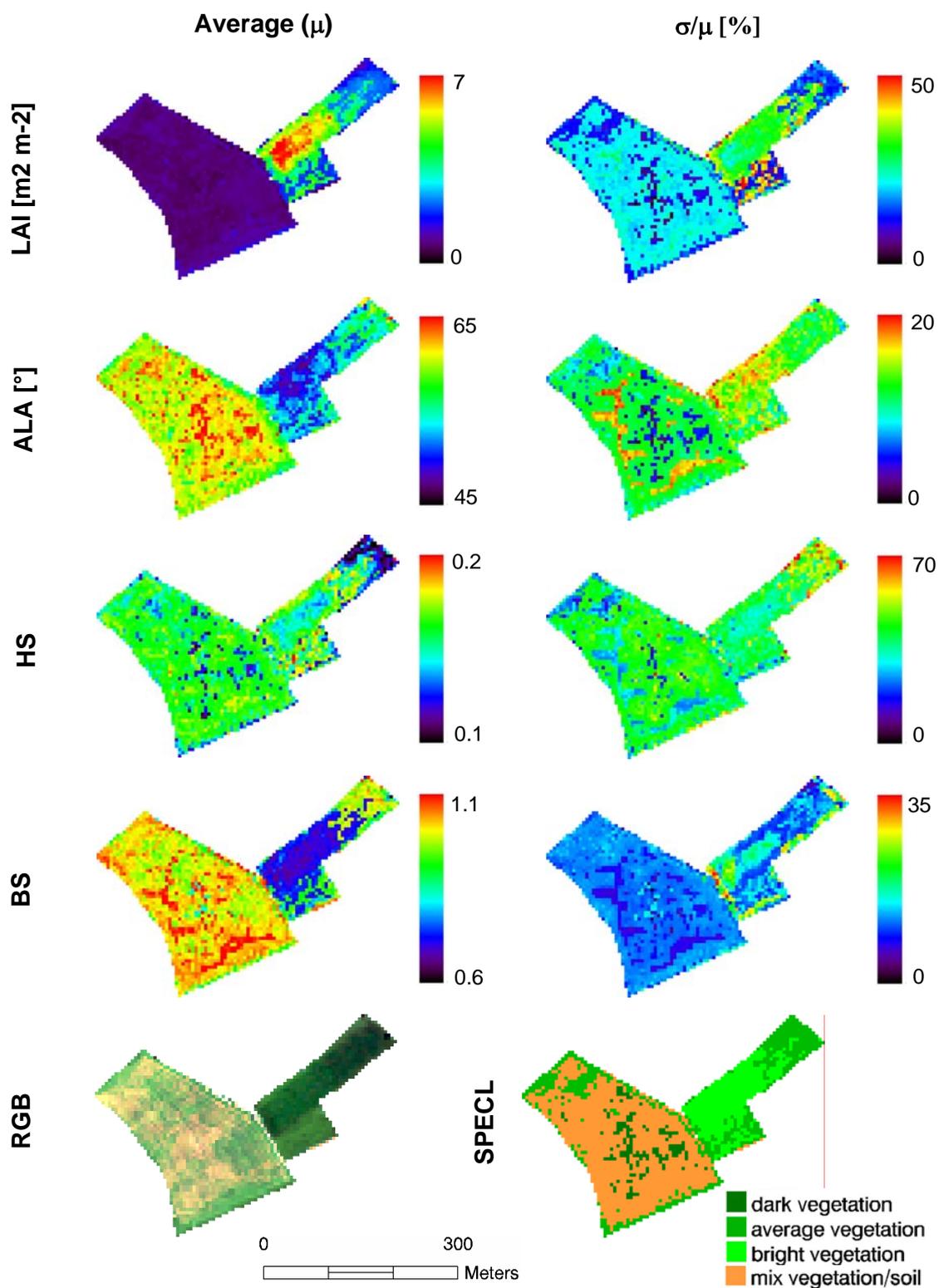


Figure 4.39: Spatial distribution at MEA1, MEA2, and PAS of canopy variables estimated from flightline WAGING07. The left column shows the average estimated values μ , the right column shows the standard error σ divided by μ and multiplied with 100%. At the lower left a true color composite of the HyMap data is shown, at the lower right the SPECL classification results. For the exact delineation of the fields, see Fig. 4.2.

line with expectations, given the low visibility of the soil in dense canopies.

Spatial incongruities Part of the trends in the patterns observed above reflect the anomalies in classification results (Figure 4.39, lower left). Whereas discontinuities in MEA2 (class 2, 3, and 6) are reflected in artificial jumps, especially in Cab and Cbp estimates, incongruity between class 3 and 4, like observed in MEA1 and PAS, are reflected in aberrant transitions for LAI , ALA , and BS . The higher degree of uncertainty at such transitions, as offered by the increased variance of the prior estimates within the local window of 5×5 pixels, appears not to be able to level out such transitions accurately enough. Nevertheless, the polishing character of the moving window, in this case unsolicited, can be very well observed at the field boundaries of MEA1, in particular for Cab , Cdm , and LAI .

4.4.4.2 Comparison with ground validation measurements

Absolute retrieval performance was tested by comparing the field validation measurements of Cw , Cdm , and LAI based on destructive sampling (Section 4.2.2) with the average values of a 2×2 pixels window around the measured GPS coordinate of each plot. The latter was done in order to compensate for allocation uncertainties. Results of this procedure are presented for each of the three observation/sun constellations (Figure 4.40 and Table 4.14). Error bars have been omitted to allow for a clearer comparison.

The trends observed in the previous paragraphs already heralded the limited absolute accuracy of the results: for all three different illumination/observation constellations and for all variables, estimates are significantly less accurate than the ones obtained for the field spectrometer based inversions (Cfr. Table 4.12). Part of the explanation for this reduced accuracy is the mismatch between the field sample size of 1×1 m² and the average HyMap pixel size of 5×5 m². This, in combination with the fact that leaf area characterization was subject to high uncertainties, is responsible for the disparity between estimated and measured LAI , especially in the case of the highly heterogeneous pasture. But the consistent underestimation of Cw and overestimation of Cdm indicate that the algorithm is subject to systematic deviations. It is suggested that different sensor calibrations and offsets due to atmospheric correction cause this shift. This is reinforced by the low spectral response induced by these variables, the reason for which they receive little weight in the cost function.

The influence of radiometric shifts on retrieval performance of the single variables is demonstrated by Figure 4.41 which for every variable (except for Cbp) compares the divergence between the estimates based on HyMap data (flightline WAGING07) and the ASD spectra discussed in the previous section. Comparison of the estimates clearly distinguishes between variables that show consistent results among both sensor types (N , LAI , ALA , and HS) and variables that show large fluctuation (Cab , Cw , Cdm , and BS). Especially for Cab , the HyMap data seem to provide far more realistic results, whereas the ASD-based inversion results appear to be trapped at the upper boundary of the range of variation defined in the LUT. The improved accuracy of Cab estimates is confirmed by the findings of Vohland and Jarmer (2007) who found values for grassland to range between $30\text{-}40 \mu\text{g}\cdot\text{cm}^{-2}$. Given the very limited spectral deviations in the VIS previously shown in Figure 4.11, the high sensitivity of Cab estimates seems remarkable, but can be well explained by the high sensitivity of this variable at high contents and the large weight

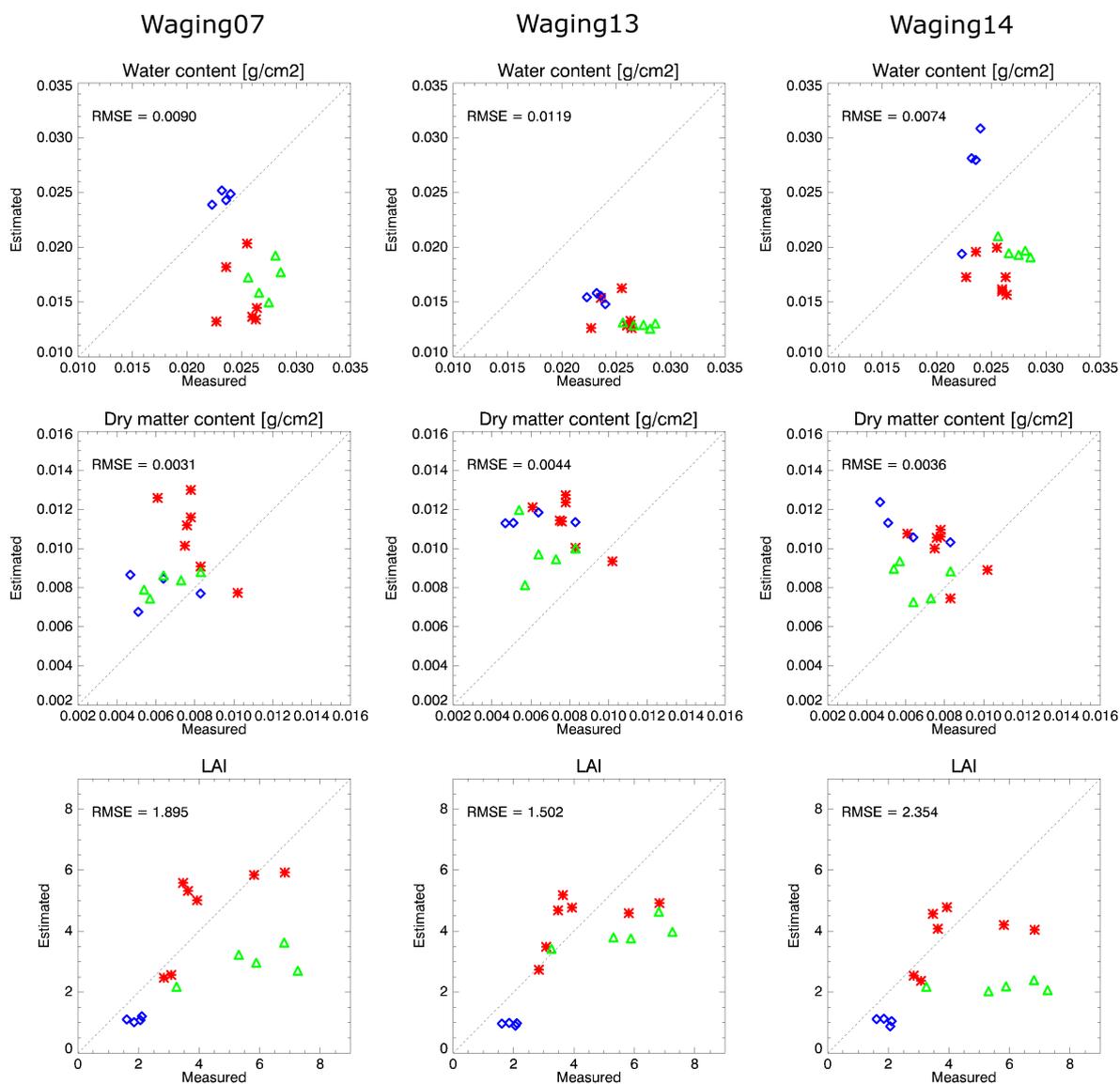


Figure 4.40: Estimated versus measured C_w , C_{dm} , and LAI , based on HyMap scenes with differing view/sun constellations: WAGING07, WAGING13, and WAGING14. Red asterisks = MEA1; blue squares = MEA2; Green triangles = PAS. Error bars have been omitted to allow for a clearer comparison.

the visible domain receives in the cost function due to the incorporation of spectral covariance.

4.5 Conclusions

Extensive testing of the proposed automated CRASH approach provided a good insight into the potential and constraints inherent to the approach itself, but also to radiative transfer model inversion in general.

Table 4.14: Accuracy of estimated C_w , C_{dm} , and LAI for 3 different observation/illumination conditions. rRMSE = relative RMSE; AA = average accuracy; OA (overall accuracy) = the average of AA over the 3 variables.

Flightline	C_w		C_{dm}		LAI		
	rRMSE [%]	AA [%]	rRMSE [%]	AA [%]	rRMSE [%]	AA [%]	OA [%]
WAGING07	35.6	69.7	43.3	60.8	46.1	63.9	64.8
WAGING13	46.7	55.3	61.8	37.3	36.5	68.2	53.6
WAGING14	47.5	63.2	50.6	51.5	57.3	59.9	58.2

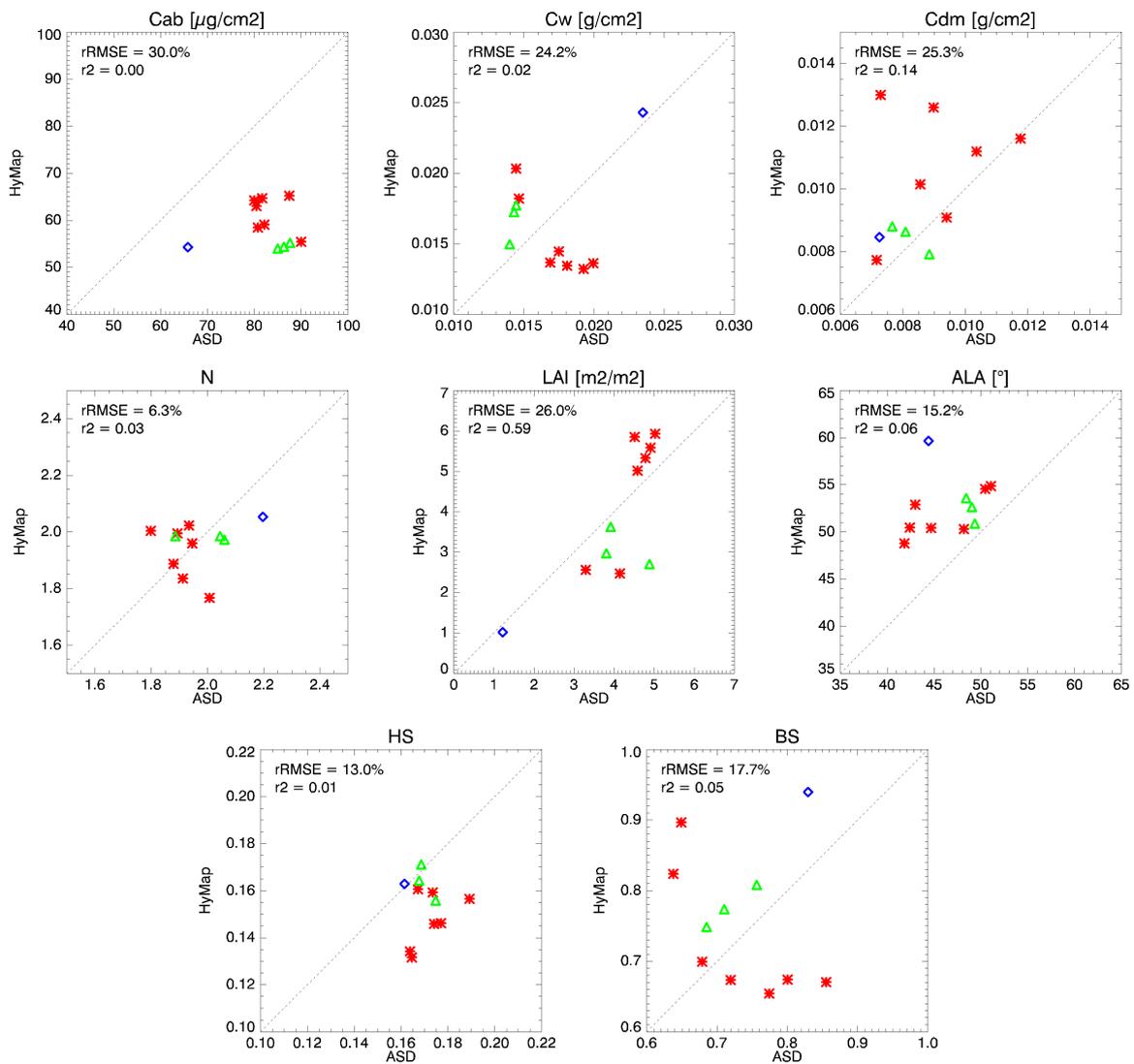


Figure 4.41: Variables estimated from field spectrometer data versus estimates based on HyMap scene WAGING07. Red asterisks = MEA1; blue square = MEA2; Green triangles = PAS.

Generally, radiative transfer model inversion suffers from the limited spectral influence of several variables in particular situations, and the saturating spectral signal occurring at high biochemical concentrations and high *LAI*. In vigorous vegetation, spectral variations are dominated by changes in chlorophyll concentration in the VIS and by *LAI* in the NIR-SWIR. Relative water content in such canopy types fluctuate around 80% and deviations are only limited. The influence of dry matter is broadly masked by those of canopy water. Nevertheless, and despite the demonstrated very high collinearity of the leaf chemicals with *LAI*, at ground level both *Cw* and *Cdm* were estimated with satisfying accuracy, especially for plots for which ground validation measurements were accurate.

In contrast, the results obtained at image level were unsatisfying. The explanation for this loss in accuracy in the first place has to be sought in scaling inconsistencies between the small sample plots and the pixel size of 5×5 meters. A more intensive field sampling campaign is therefore required to extrapolate the results of the presented test case to the overall performance of the approach. In addition, there was an increase of geometric uncertainties, of radiometric uncertainties due to increased atmospheric path length, and of directional uncertainties resulting from changing illumination/observation geometries and the attribution of different view angle intervals. These sources of uncertainty, in turn, influenced the different components of the algorithm such as the land cover classification, the predicted prior estimates, and the characterization of spectral covariance. The characterization of spectral covariance appeared to play a decisive role in redistributing the weights of the single wavebands in the radiometric cost function. The emphasized weight received by the VIS makes the approach susceptible to radiometric uncertainties in this domain. This is particularly true for green vegetation with high VIS absorption for which small spectral changes can lead to large deviations in predicted chlorophyll concentrations. Through the covariance terms, these discrepancies are propagated to the other variables. This assumption is confirmed by the viable results obtained for the cut meadow, a canopy type that is not subject to saturation of spectral sensitivity. It is therefore recommended not to use spectral covariance in cases where no explicit information on the quality of the radiometric calibration and atmospheric correction is available.

The land cover based inversion approach appeared to be a powerful tool in optimizing estimations for the observed canopy type and in reducing uncertainties related to the estimated values. Nonetheless, the quality of the classification plays a decisive role for the accuracy that is finally obtained. Classifications based on a single image face the difficulty of lacking information about the phenological course of a canopy which hampers a subdivision based on crop types. As a consequence the class specific LUT parametrization had to be kept general, thus impeding an accurate regularization of the inverse problem. In this respect, spectral classification should be combined with other classification techniques (e.g., texture based approaches and edge detection algorithms) to improve demarcation of field boundaries and to eliminate unnatural transitions at intra-field level.

The ultimate goal of this chapter, namely detecting progressive changes in water content, was overshadowed by the instability of the approach, as presented above. Moreover, spectral variations in the characterized meadows were mainly induced by morphological changes induced by cutting and regrowth.

Nevertheless, the results presented in this chapter emphasize the importance of testing a model for several view/observation geometries and different qualities of radiometric data. Even if also at image level improved accuracy would have been obtained if the model had been fed with better-defined a priori information measured during the field campaign, this was not the

purpose of this chapter, as this thesis emphasizes on achieving maximum accuracy based on a fully automated approach. Therefore, additional regularization should be explored, provided by completely independent information sources. The next chapter will focus on including an additional independent information source, provided by the diverging behavior of angular anisotropy of different land cover types.

Chapter 5

Validating CRASh at satellite level: estimating cotton leaf chlorophyll content from multi-angular CHRIS/PROBA observations

5.1 Introduction

Radiative transfer model inversion bears the potential of concurrently accounting for the multiple elements responsible for canopy reflectance. However, the ill-posed nature of radiative transfer model (RTM) inversion requires considerable regularization of the system in order to bring multiple solutions down to a minimum. Especially in situations where only little a priori information on the canopy under observation is available, like in the case of an automated inversion, this might constitute a severe limitation. The previous chapter showed that significant consolidation of the results could be obtained by introducing a land cover classification and prior information on the variables. Nevertheless, neither the high dimensionality of hyperspectral data, nor the use of vegetation indices provided sufficient regularization in all situations, mainly due to the low radiometric sensitivity to several variables and the resulting complicated interaction between the various variables that were left free during inversion. Therefore, additional regularization is required in order to obtain a data quality that is accurate enough to be reliably incorporated in spatially distributed process models. Multi-directional observations of a single target can provide such supplemental regularization.

5.1.1 Multi-angular remote sensing

Several studies showed that multi-angular observations of reflectance anisotropy provide an independent and complementary information source to the spectral signature of a canopy (Bacour et al., 2002b; Beisl, 2001; Gemmell, 2000). They have proved to be diagnostic for structural

surface properties and are helpful to complement the spectral measurements for a complete and robust characterization of a vegetation canopy (Bacour and Bréon, 2005; Koetz et al., 2006). Spectrodirectional anisotropy is known to be particularly large for canopies having a complex 3-D structure and intermediate density such as open coniferous forest stands (Koetz et al., 2006; Widłowski et al., 2004). Also a row-structured crop like cotton could give rise to increased anisotropy since the fraction of the spectrally contrasting soil background that is viewed by the sensor strongly varies with view zenith angle. Due to the complementary effect between the various input variables in RTM inversions, improved characterization of structural elements should enhance estimations of other variables, such as foliage biochemical components, as well.

The aim of the current chapter is to explore the benefit of introducing the directional data dimension in RTM inversion in addition to the hyperspectral dimension already addressed in the previous chapter. For this purpose, data originating from the CHRIS (Compact High Resolution Imaging Spectrometer) sensor was exploited for its potential of assessing chlorophyll concentration in cotton leaves. To date, the CHRIS onboard the PROBA (Project for On-board Autonomy) platform is the only spaceborne initiative providing hyperspectral and multi-angular observations of the earth surface at high spatial resolutions (Barnsley and Settle, 2004). In summer 2006, the CHRIS sensor collected multiple viewing angle data of an area covered mainly with cotton and rice fields in the Khorezm region of Uzbekistan. In this region, a long long-term, interdisciplinary ZEF/UNESCO pilot research project for the economic and ecological restructuring of land- and water use is carried out¹.

5.1.2 Land and water use restructuring in the Uzbek Khorezm region

The Khorezm region in the Uzbek part of the Aral Sea Basin is characterized by irrigation-based agriculture. Throughout the years, the region has increasingly suffered from the consequences of unsustainable use of natural resources, pesticides, and fertilizers. This has led to a variety of severe ecological and economical problems, such as salinization, over-fertilization, soil and water pollution, and as a consequence, higher expenses and lower incomes for the farmers. Crop produce is often far from optimal, partly due to inefficient use of fertilizers. This is also the case for cotton (*Gossypium hirsutum L.*), the dominant crop in the region. The application recommendations that were established during Soviet times are often not adapted to the prevailing, site-specific environmental conditions and recent plant varieties.

Fertilizer trials carried out in the region showed that Nitrogen (N) is the most limiting nutrient in Khorezm soils (Kienzler, In prep.). Therefore, maps showing the spatial distribution of N status in cotton plants are an important information basis for the local farmers which, combined with knowledge on soil condition and plant-nutrient uptake behavior, would support adjusting fertilizer inputs according to the actual and site-specific N requirements and farmers' individual crop production strategies (Ruecker et al., 2006).

5.1.3 Imaging spectroscopy in cotton studies

Fertilizer trials and crop simulation models in support of detecting nutrient deficits have been carried out at several test plots in the region (Kienzler, In prep.). In contrast, satellite-based

¹<http://www.khorezm.uni-bonn.de/>

approaches have the potential to rapidly assess the intra-field crop nutrient status over a larger region. Nitrogen is known to strongly correlate with chlorophyll content, a quantity that can be well detected using various remote sensing approaches (Baret and Fourty, 1997b; Ruecker et al., 2006). Nevertheless, nutrient stress anomalies not only manifest themselves in premature senescence, as evidenced by yellowing or chlorosis of older leaves, but also affect crop structural variables such as *LAI* and biomass development (Fridgen and Varco, 2004; Zarco-Tejada et al., 2005b). Concurrent estimation of canopy structure variables and biochemical constituents may therefore serve as a more robust indicator of hampered vegetation development.

Numerous studies addressed the potential of multi- and hyperspectral remote sensing approaches in assessing cotton quality and yield estimates, either based on stepwise multiple regression (Thenkabail et al., 2000), broadband and narrow band vegetation indices (Thenkabail et al., 2000; Zarco-Tejada et al., 2005b; Zhao et al., 2005) or partial least regression (Fridgen and Varco, 2004; Read et al., 2002). Within-field variability of yield and growth development can be well detected using high spatial resolution hyperspectral and multispectral remote sensing data or false color infrared aerial photography (Plant et al., 2000; Yang et al., 2001; Yang et al., 2003; Zarco-Tejada et al., 2005b). Identifying regional anomalies of specific nutrient stresses appears more difficult due to the combined impact of nitrogen stress on both biochemical and structural canopy composition. In addition, observed anomalies, for example in chlorophyll content, are not always a direct result of N deficiency but can be just as well induced by other stress factors (Fridgen and Varco, 2004).

Several studies, using statistical approaches and conducted at leaf level, achieved high correlations between vegetation indices and leaf chlorophyll or N content (Fridgen and Varco, 2004; Tarpley et al., 2000). At plant level, reported results were usually less accurate than at leaf scale (Zhao et al., 2005). Based on field spectrometer measurements conducted at fertilizer trials, Read et al. (2002) found a good correlation between the R415/R695 ratio and chlorophyll a+b content in cotton plants. However, practical use of this ratio is hindered by the absence of the 415 nm band in most imaging spectrometer systems and the susceptibility of this channel to atmospheric influences. Up to date, only few attempts have been reported on the applicability of statistically based methods for the spatially distributed detection of nutrient deficits in cotton plants under real crop production conditions. Such farmer-managed fields usually show less intra-field variation than fertilizer plots due to more uniform management, making it more difficult to establish clear correlations (Grillenberger, 2007; Ruecker et al., 2006).

Thenkabail et al. (2000) and Zhao et al. (2005) reported the importance of center wavelength position and band width on the performance of the proposed statistical relationships between reflectance data and biochemical compositions. In this regard, there is general agreement on the added value of hyperspectral data for the detection of biochemical concentrations in cotton leaves (Zarco-Tejada et al., 2005b). Nevertheless, the limitation of statistical approaches in addressing nitrogen and chlorophyll status, is the implicit assumption that spectral variations in regions sensitive to leaf pigments (i.e., the blue and red domains) are solely induced by changes in the concentrations of these foliage biochemicals. However, the strong absorption in these regions is also governed by other foliage and structural properties (Cfr. Chapter 2). Simultaneous assessment of the other variables is therefore necessary for an accurate retrieval of leaf chlorophyll content and N at plant, field, and regional level. Such a simultaneous assessment can be offered by radiative transfer model inversion.

5.1.4 Objectives

The main objective of this chapter is to discover to which degree the automated CRASh approach proposed in Chapter 3 is capable of estimating chlorophyll content in cotton leaves from hyper-spectral multi-angular CHRIS observations. Although featured spatial resolution and spectral coverage are different from the HyMap sensor presented in the previous section, the anisotropy contained in the CHRIS data is said to contain a similar amount of uncorrelated information as the full resolution HyMap (Verhoef, 2007). RTM inversion based on multi-angular data should therefore be able to provide accuracies that are comparable with the ones obtained in the previous section. To gain a more comprehensive insight into the potential offered by the multi-directional data, the following specific objectives were defined:

- Quantifying the spectral anisotropy present in the data and testing the potential of the radiative transfer model (PROSPECT+SAILh) in reproducing this anisotropy.
- Assessing the accuracy of chlorophyll a+b and *LAI* estimations obtained by inversion of the single view directions.
- Identifying the combination of view angles that provide maximum accuracy of the solution.
- Evaluating the degree to which retrieval accuracy can be improved when, based on the current RTM, a priori knowledge from the field experiments is included in constructing the LUTs.

5.2 Study site and data

5.2.1 Study site Khorezm

The study area is located south of the Aral Sea basin, at the lower Amu Darya River, in the Khorezm region of Uzbekistan (Figure 5.1). This region is situated at an altitude of ca. 96 m above sea level, and characterized by an extremely continental arid to semi-arid climate with large daily and seasonal temperature differences. The monthly averaged temperatures are 35-36°C in July and dominated by frost with temperatures down to -20°C in January. Precipitation rates are very low with an average of 92 mm per year which, in combination with the high summer temperatures, permits only irrigated agriculture. The inappropriate use of water has led to a high salinisation of water and soils, a problem often encountered in irrigation-based agriculture in arid and semiarid landscapes (Bastiaanssen and Ali, 2003). Cotton is the most important crop in the region, covering in recent years ca. 50% of the irrigated land, while rice, wheat, and other crops have a markedly smaller area share (Djanibekov, 2007).

Validation of the CHRIS-based estimates of chlorophyll and *LAI* was performed on the Amir Temur Shirkat collective farm on four selected fields with the Khorezm-127 cotton variety, which were intensively sampled between July 11-18, 2006 (Figure 5.1). During the growth season, the fields received similar amounts of N fertilizer and water while irrigation was always applied within a few days after fertilization (Table 5.1). Due to varying environmental conditions, as parametrized by different soil fertility, or bonitet¹ levels, large inter- and intra-field variabilities

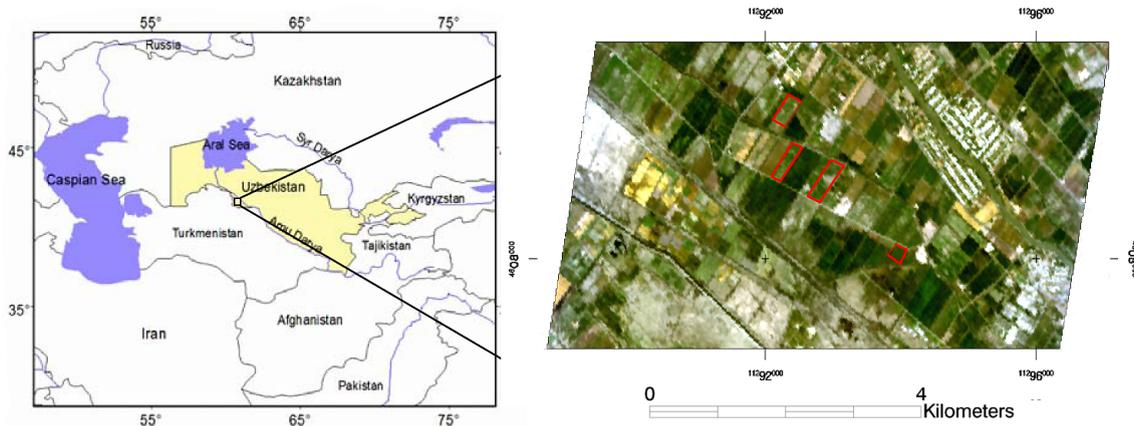


Figure 5.1: Location of the Khorezm region in Uzbekistan (left) and site overview as observed by PROBA-1/CHRIS mode 5 imagery on July 16, 2006 (bands 8, 4, 2). The red polygons indicate the validation fields used in this study which from the northwest to the southeast are numbered in ascending order C1-C4 (right). By courtesy of Gerd Ruecker.

Table 5.1: Crop management specifications of validation fields.

Field	Planting date	Fertilisation date	Fertilizer amount	irrigation date	Bonitet level	Row azimuth [°]
C1	11.04.06	04.05.06	400 kg N/ha	09.05.06	8	31
		14.06.06	200 kg N/ha	16.06.06		
		18.07.06	200 kg N/ha	20.07.06		
C2	28.04.06	10.07.06	400 kg N/ha	15.07.06	5-8	34
		-	-	20.07.06		
C3	14.04.06	10.05.06	200 kg N/ha	15.05.06	5-8	34
		02.06.06	200 kg N/ha	04.06.06		
		05.07.06	200 kg N/ha	10.07.06		
		01.08.06	200 kg N/ha	04.08.06		
C4	20.04.06	09.05.06	400 kg N/ha	12.05.06	5-7	32
		07.06.06	200 kg N/ha	08.06.06		
		07.07.06	200 kg N/ha	10.07.06		

in plant development were observed, which determined to a large extent the fraction of exposed bare soil between the rows (Figure 5.2). The planting direction of the single cotton rows was always orthogonal to the main roads and had an azimuth direction between 31-34°.

¹The bonitet parameter defines the potential soil fertility class and is based on soil texture and depth (Kuziev, 1989). Bonitet is scaled from 1-100, with 100 representing maximum fertility.



Figure 5.2: Examples of cotton canopies encountered in the study area at the time of satellite overpass. The left photo was taken in field C1, the right photo in C4. Photos by Jörg Grillenberger

5.2.2 Biometric sampling

5.2.2.1 Sampling scheme

The validation fields in Amir Temur Shirkat were sampled for leaf chlorophyll content and LAI within two days before or after the satellite image acquisition. An approximate X-shaped sampling pattern was followed with sampling points ca. every 15 m (Figure 5.4). It was ascertained that measurements were taken at least 20 m from the field boundaries in order to avoid that measurements would fall into pixels constituting a mixed signal of adjacent fields. Each measurement point was marked using a hand-held non-differential GARMIN GPS 12 receiver. A total number of 46 point measurements was taken on each field.

5.2.2.2 Chlorophyll

Rapid non-destructive sampling of leaf chlorophyll was obtained with the SPAD-502 chlorophyll meter (hereafter simply called SPAD; Minolta, Inc.). The instrument calculates an estimate of the relative chlorophyll content based on the absorbed fraction of red and near-infrared light emitted by the instrument. To be able to study the contribution of each leaf layer to the remote sensing signal, SPAD readings were taken at the different shoot levels of the plant, counting from the top downwards. For each level, the final SPAD values were constituted by the average of 3 measurements.

The SPAD instrument was calibrated for plant variety and local growth conditions in order to obtain a chlorophyll a+b content that was comparable with the ones retrieved from remote sensing observations (Cab ; $\mu g \cdot cm^{-2}$). Calibration was based on 100 randomly sampled leaves from 24 cotton plants of the Khorezm-127 variety that were cultivated on a fertilizer trial plot near Urgench University (Kienzler, In prep.). The leaves were selected in such a way that a wide SPAD range (3.3 - 70.2) was covered. Five SPAD readings were taken from each leaf before being harvested. The collected leaves were placed in an iced, air-sealed container and transported to the laboratory where chlorophyll and carotenoid concentrations were determined after (Lichtenthaler, 1987). The pigment concentrations were converted to $\mu g \cdot cm^{-2}$ by relating

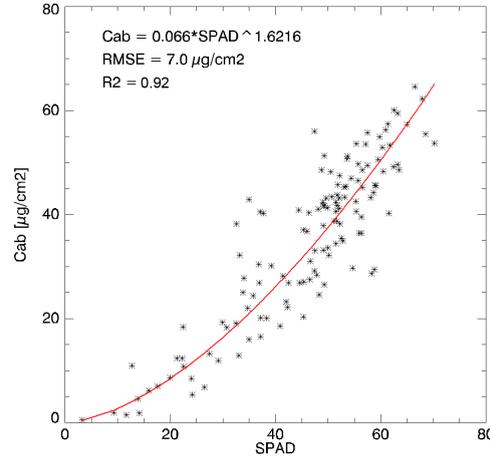


Figure 5.3: Regression fit between SPAD and laboratory measurements of chlorophyll a+b (Cab).

the concentration to the leaf area of the sample. A relationship between SPAD and laboratory measured Cab was established using an exponential regression fit with correlation coefficient (R^2) of 0.92 and a root mean square error of $7.0 \mu\text{g} \cdot \text{cm}^{-2}$ (Figure 5.3):

$$\text{Chlorophyll } a + b [\mu\text{g} \cdot \text{cm}^{-2}] = 0.066 \cdot \text{SPAD}^{1.6216} \quad (5.1)$$

Equation 5.1 was used to convert the SPAD values measured at the validation fields into Cab . Errors of the above equation were propagated to the Cab estimates using a relative RMSE^2 of 15.8%.

5.2.2.3 Leaf area index

LAI was measured at each sampling location using a LAI-2000 plant canopy analyzer (LICOR, 2000). The LAI-2000 provides an effective plant area index including green foliage and stems rather than just the green leaf area per unit ground surface area. For a more detailed description of the instrument and the measuring principle, the reader is referred to Section 4.2.2.2.

LAI was measured in four directions around the stem of the same plants that had been used for the characterization of chlorophyll content. To upscale these measurements to the level of the entire canopy, and to account for the intra-row spaces without vegetation cover, digital nadir-taken photographs of each sampling location were classified into a green-vegetation and a background fraction representing the uncovered soil between the rows. The fraction of plant material obtained in this way was in turn used to calculate LAI at canopy level (LAI_c) assuming a linear relationship between vegetation cover and LAI :

$$LAI_c = LAI_p \cdot fCover \quad (5.2)$$

²rRMSE = RMSE / mean of predictions

Table 5.2: Descriptive statistics of cotton canopy measurements, summarized per validation field. Shown is the average for value and, in brackets, the coefficient of variation measured in each field. Cab and SPAD values are shown for the top layer leaves (TL), the bottom leaves (BL), and for the average of all leaf layers (AL).

<i>Variable</i>	<i>Unit</i>	<i>C1</i>	<i>C2</i>	<i>C3</i>	<i>C4</i>
SPAD (TL)	unitless	36.6 (19.4%)	34.1 (16.7%)	38.3 (12.4%)	34.9 (14.8%)
SPAD (BL)	unitless	46.2 (19.8%)	38.6 (17.5%)	42.8 (13.2%)	41.4 (12.6%)
SPAD (AL)	unitless	42.5 (15.9%)	35.7 (14.7%)	41.3 (7.9%)	37.8 (9.6%)
Cab (TL)	$\mu g \cdot cm^{-2}$	23.0 (31.4%)	20.4 (26.9%)	24.6 (20.7%)	21.2 (23.2%)
Cab (BL)	$\mu g \cdot cm^{-2}$	34.8 (14.6%)	25.0 (27.6%)	29.4 (21.3%)	27.9 (20.7%)
Cab (AL)	$\mu g \cdot cm^{-2}$	29.2 (11.3%)	22.0 (11.6%)	27.6 (16.0%)	24.0 (15.0%)
LAI	m^2/m^2	1.35 (39.9%)	1.01 (55.0%)	1.61 (34.5%)	1.79 (30.2%)
<i>n</i> samples		44	43	46	46

with LAI_p being the LAI measured below the plant with LAI-2000, and $fCover$ the fractional cover of green foliage estimated from the digital photographs. This procedure was repeated for all point measurements. The eventual uncertainty of the LAI measurements was calculated by multiplying the standard error of the average LAI-2000 measurements by the fraction of uncovered soil.

Table 5.2 shows some descriptive statistics of the canopy measurements. SPAD and Cab are shown for the top level leaves (TL), the leaves found at the bottom level (BL) and the average of all leaf layers (AL). Measured SPAD and Cab are significantly higher for the bottom leaves than for the top level leaves. Although this trend is consistent with earlier results obtained by Ruecker et al. (2006), it does not coincide with the general assumption that upper leaf layers generally show higher chlorophyll contents per leaf area (Vohland and Jarmer, 2007). Cotton plants appear to develop according to a complicated growth pattern with allocations of high chlorophyll concentrations possible in every layer (Kienzler, pers. comm.). Within-field variability is significantly higher for LAI than for Cab , while the latter also shows reduced inter-field variability. Spatial distributions of measured Cab (averaged over all shoot levels) and LAI are displayed in Figure 5.4. The figure reveals that the relatively low values of Cab and LAI measured on cotton field C2 coincide well with the less intense green (for Cab) and red (for LAI) colors of the background image, indicating lower absorption of red light and reduced NIR reflectance, respectively. Similarly, the intenser background colors observed for C3 and C4 are consistent with the higher measured levels of both variables, whereas C1 constitutes intermediate levels.

5.2.2.4 Additional measurements in support of RTM inversion

Several other biochemical and biophysical variables were measured in order to be able to employ realistic ranges of variation in the radiative transfer model simulations (Section 5.5). These

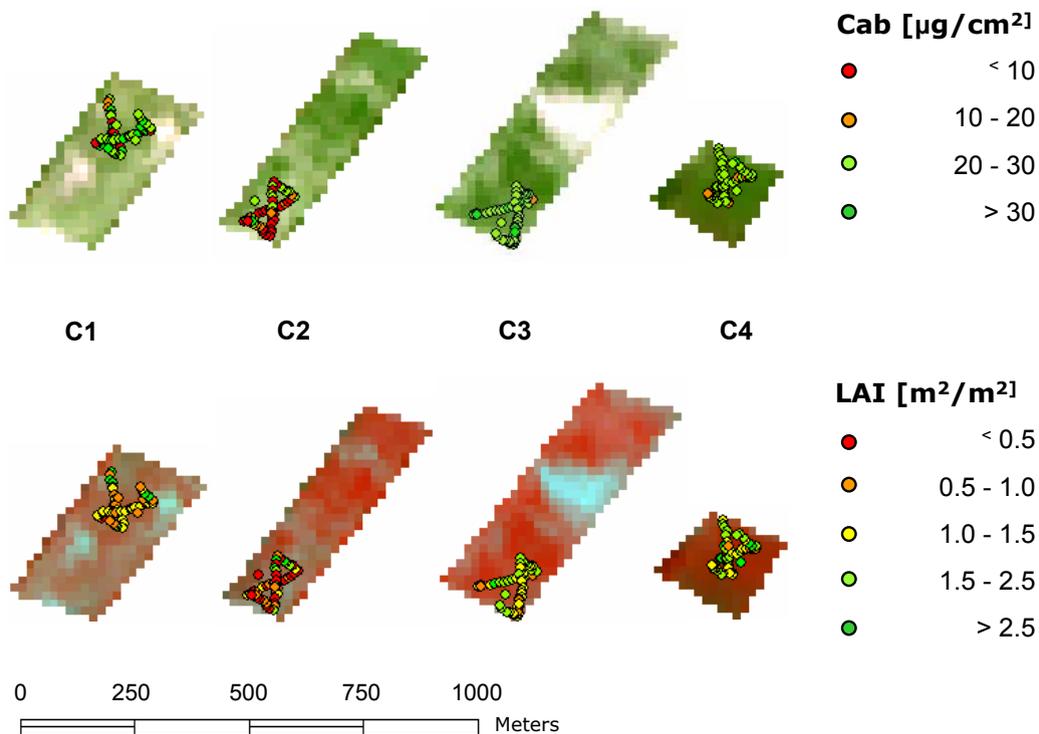


Figure 5.4: Spatial distribution of the sampling points and the measured Cab (above) and LAI (below) values. The sampling points are projected on a true color (red = band 8, green = band 4, blue = band 2) and a false color composite (red = band 25, green = band 8, blue = band 4) of the CHRIS nadir observation of July 16, 2006, for Cab and LAI respectively.

variables were measured on fertilizer trial plots near Urgench University, which were representative for the wide range of vitality rates of the Khorezm-127 variety that could be potentially encountered in the area.

Leaf dry matter (Cdm ; $g \cdot \text{cm}^{-2}$) and leaf water content (Cw ; $g \cdot \text{cm}^{-2}$) were sampled for a selected number of leaves ($n=33$). First, the area of the single leaves was determined using a Li-3000 planimeter (Li-Cor. Inc., Lincoln, Nebraska). After determining their fresh weight, the scanned leaves were oven-dried at 70°C for 36 hours. Differences between fresh and dry weight in combination with the leaf surface facilitated the calculation of Cw and Cdm according to Equations 4.2 and 4.3 introduced in Chapter 4.

Canopy height and leaf width were obtained to allow for the calculation of the hot spot parameter HS , which is equal to the ratio of the correlation length of leaf projections in the horizontal plane and the canopy height (Verhoef and Bach, 2007). A summary of the additionally measured variables is given in Table 5.3.

Table 5.3: Variables measured in support of RTM inversion.

<i>Variable</i>	<i>Unit</i>	<i>mean</i>	σ	<i>min</i>	<i>max</i>
<i>Cw</i>	$g \cdot cm^{-2}$	0.0265	0.0153	0.0024	0.0852
<i>Cdm</i>	$g \cdot cm^{-2}$	0.0093	0.0052	0.0011	0.0284
Canopy height	<i>cm</i>	80		51	108
Leaf width	<i>cm</i>	13		10	15

Table 5.4: Observation and illumination details of the CHRIS imagery used in this study.

<i>Date</i>	<i>Nominal θ_v [°]</i>	<i>Actual θ_v [°]</i>	ψ_v [°]	θ_s [°]	ψ_s [°]	ψ_{rel} [°]	<i>Time (UTC)</i>
16.07.06	+55	48.31	6.77	23.4	145.6	138.8	07:06:32
	+36	27.44	359.54			146.1	07:07:21
	0	8.01	224.29			78.7	07:08:10
	-36	32.96	203.20			57.6	07:08:59
	-55	51.9	199.01			53.4	07:09:48

5.2.3 CHRIS multi-angular satellite observations

5.2.3.1 Sensor and view characteristics

On July 16, 2006, spectrodirectional data of the test site was collected with the CHRIS sensor onboard the PROBA platform (Barnsley and Settle, 2004). The CHRIS instrument is able to acquire observations of one and the same target at 5 different view angles, two in the forward looking direction, one at nadir, and two in backward viewing direction. Due to the necessity of sideways pointing, nominal and actual observation angles usually deviate a few degrees (Table 5.4). Although the angular resolution of CHRIS is small compared to other spaceborne multi-angular sensors such as POLDER (14 view directions) or MISR (9 view directions), the sensor distinguishes itself from the others by its high spatial and spectral resolution (up to 17 m, and up to 63 bands in the VNIR, respectively). Another unique feature of the instrument is its capability of observing the surface in different spectral and spatial modes according to the requirements of the user.

The scenes recorded for this project were collected using the *half swath mode* (Mode 5), imaging an area of 7×7 km² (744 \times 748 pixels) with the for CHRIS highest possible spatial resolution of 17 m at nadir. For the two extreme view angles ($\pm 55^\circ$), resolution deteriorates by approximately a factor of two. Mode 5 covers the spectral range between 442 to 1025 nm with a total of 37 spectral bands having a full-width-half-maximum (FWHM) of 6-30 nm (Appendix F). The observation details of the data set used in this study is summarized in Table 5.4. Figure 5.5(a) shows the position (view zenith and azimuth) of the observations relative to the test fields.

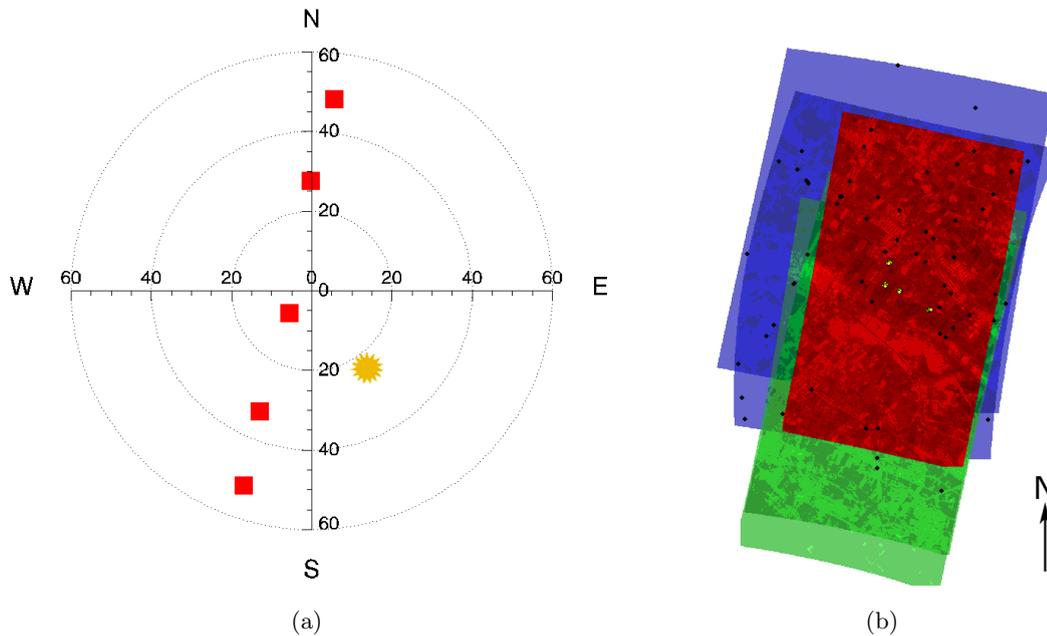


Figure 5.5: (a) Polar plot of CHRIS image acquisition and illumination geometry as of July 16, 2006 (red squares). The red squares indicate the position of the observation (in polar coordinates) relative to the sensed area in the origin. In the north (N) $\psi_v=0$, in the East (E) $\psi_v = 90^\circ$, and so on. (b) Overlay of georeferenced multi-angular CHRIS images. Blue scenes: $\theta_{v,nominal} = +/- 55^\circ$; green scenes: $\theta_{v,nominal} = +/- 36^\circ$; red: $\theta_{v,nominal} = +/- 55^\circ$. The yellow dots indicate the position of the field sampling locations, the black dots, the positions of the total of GCPs available for georeferencing. By courtesy of Richard Fuchs.

5.2.3.2 Preprocessing

Image enhancement and georeferencing All data were first corrected for bad lines and striping using the ESA HDFclean V2 algorithm (Cutter, 2006). Subsequently, the images were georeferenced using a non-parametric approach based on ground control points (GCPs). For this purpose, the coordinates of several prominent infrastructures, such as road crossings and bridges over wider irrigation channels, within the expected coverage of the images were marked using a non-differential hand-held Garmin GPS 12 receiver while the direct vicinity of the validation fields was more densely sampled.

The entire scenes were projected into the local Gauss Kruger zone 11 (Pulkova 1942) coordinate system using a second degree polynomial transformation and nearest neighbor interpolation (Figure 5.5(b)). GCPs were iteratively added until a minimum RMSE of the projected GCPs was obtained. The total number of used GCPs depended on view geometry and ranged from 16-28 (Grillenberger, 2007). Due to the relatively small size of the fields, a good image to image registration was required to allow for a reliable intercomparison between the different view angles. Thus, image to image registration was performed for the area around the test sites taking the nadir observation as a reference and using a 1st degree polynomial fit based on 10-15 image based GCPs. The original data taken at extreme view angles were resampled to a pixel size of $17 \times 17 \text{ m}^2$ in order to match the observations taken at nadir. The final uncertainty (total

RMSE) of the georectification process ranged from 0.33-0.70 pixels in X- and Y-direction.

Vicarious calibration and atmospheric correction Since standard calibration coefficients accompanying the data did not yield satisfying accuracy, vicarious calibration was performed simultaneously to atmospheric correction using the inflight calibration option in ATCOR (Richter, 2007b). For this purpose, spectral properties of a homogeneous bare surface of approximately $500 \times 200 \text{ m}^2$ were measured with a portable Fieldspec PRO FR spectrometer (Analytical Spectral Devices, Inc.). In total, 100 bare fiber ($\text{FOV} = 25^\circ$) radiance measurements were taken and directly converted into reflectance by taking a SpectralonTM panel as a white reference. The single spectra were first corrected for the spectral properties of the applied Spectralon panel, deviations of the white reference from the 100 % reflectance line, and the spectral jump between the VNIR and the SWIR1 detector, using an additive correction and taking the SWIR1 detector as a reference (Dorigo et al., 2006). Subsequently, average reflectance, standard deviation, and minimum/maximum reflectance were defined and resampled to match the sensor characteristics of CHRIS.

Atmospheric correction was performed using a rural aerosol type, a water vapor column of 2 cm, and an average visibility of 41 km. Flat terrain was assumed and ground elevation was fixed at 100 m above sea level. Of the initial set of 37 bands, 6 bands subject to increased calibration uncertainty were excluded from further processing, including band 1 (442.5 nm), 2 (490.4 nm), 3 (530.2 nm), 19 (759.7 nm), 20 (766.8 nm), and 30 (930.4 nm).

5.3 Quantifying spectral anisotropy

5.3.1 Observed directional signatures

According to the physical mechanisms of photon transport in vegetation canopies, spectral anisotropy should be particularly strong in ranges of high vegetation absorbance such as the visible blue and red chlorophyll absorption bands where spectral contrast between canopy constituents and background is most pronounced. In the highly reflective near-infrared bands, multiple scattering processes effectively reduce contrast between shadowed and illuminated canopy components resulting in lower anisotropy effects. In addition, BRDF effects are most pronounced in erectophile canopies and are reduced in planophile surfaces (Sandmeier et al., 1998; Sandmeier and Deering, 1999). According to these principles, one would expect increased anisotropy effects in the visible channels for the open, row-structured cotton canopies where the bright, saline soil background should lead to a bell shape pattern with maximum reflectance at nadir (Widlowski et al., 2004).

Figure 5.6 shows the average HDRF measured at each test field for the green, red, and NIR channels. Despite the fact that the observation and solar azimuthal plane exhibit an angle between 33.9 and 78.7° , significant anisotropy can be observed. In the visible wavebands, the expected bell shape, however, cannot be observed. In contrast, angular signatures in the visible wavelengths show a quasi linear shape with maximum HDRF in the view direction closest to the hot spot geometry (i.e., $\theta_v = -33.0^\circ$), whereas NIR reflectance shows a slightly bowl shaped pattern. The observed shapes are typical for homogeneous vegetation surfaces and are conform

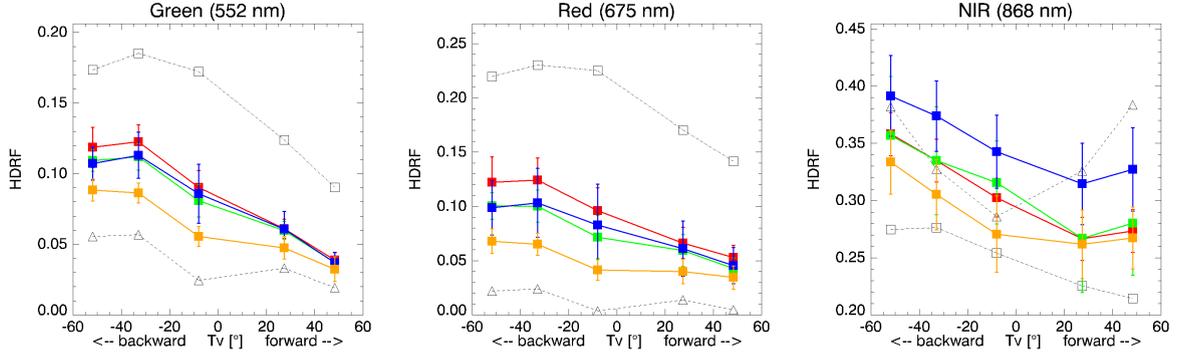


Figure 5.6: Green, red, and NIR HDRF for various fields at 5 different view angles. Red = test field C1, green = C2, blue = C3, orange = C4. The error bars indicate the range of variation encountered in the respective field. For reasons of comparison also the HDRF for a neighboring rice field (black triangles / dashed curve) and bare soil (black squares / dashed-dotted curve) are shown.

to what others found for either measured (Widlowski et al., 2004; Verhoef and Bach, 2007) or modeled anisotropy in structurally homogeneous canopies with rather planophile leaf angle distributions (Pinty et al., 2004b; Verhoef and Bach, 2007). This finding suggests that, despite the row structure of cotton, in the considered case it is appropriate to use the 1-D SAILh model.

To give an indication of directional anisotropy of the cotton fields compared to other land cover types in the area, the spectral variation of a bare soil and a rice paddy in the immediate vicinity of the cotton fields are shown too in Figure 5.6. Soil anisotropy shows a moderate bowl shape in all wavelengths, with maximum reflectance at the view constellation closest to the hot spot. This is exactly in line with the observations made by Verhoef and Bach (2007) and is explained by the fact that on a smooth surface, backward scattering is stronger than forward scattering. The observed directional soil anisotropy is an important recognition with respect to the SAILh parametrization used in this study, which assumes soil reflectance to be anisotropic, and may have a considerable effect in the visible part of the spectrum at low vegetation covers where spectral contrast between vegetation and soil is high. The rice paddy shows a directional behavior that for the visible wavelengths is very similar to the ones observed for the cotton fields, with the only exception being the reduced absolute reflectance, which is ascribed to the increased vegetation density as reported by field observations. In contrast, directional behavior observed in the NIR is very distinctive from that of the cotton fields, showing a pronounced bowl shape. The bowl shape in the NIR is characteristic for vertically oriented canopy structures, a result that coincides very well with the observed erectophile structure of rice (Pinty et al., 2001; Pinty et al., 2004b; Verhoef and Bach, 2007; Widlowski et al., 2006b).

To visualize directional anisotropy for every waveband, the anisotropy index (ANIX) was calculated, which is defined as the ratio between the maximum and minimum reflectance values in the principal plane (or defined azimuth plane) per spectral band (Sandmeier et al., 1998; Sandmeier and Deering, 1999):

$$ANIX(\lambda, \Omega) = \frac{R_{max}(\lambda)}{R_{min}(\lambda)} \quad [dimensionless] \quad (5.3)$$

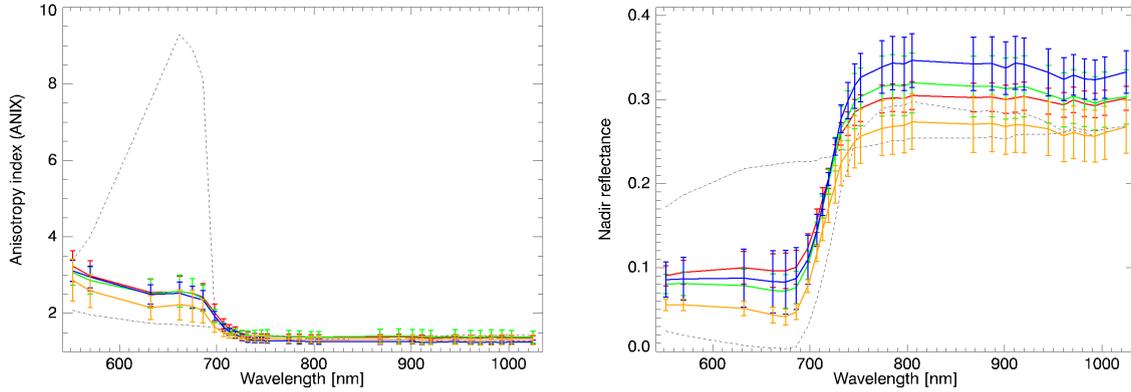


Figure 5.7: Anisotropy index (ANIX; left) and average nadir reflectance (right) for six different fields. Red = test field 1, green = test field 2, blue = test field 3, orange = test field 4. The error bars indicate the range of variation encountered in the respective field. For reasons of comparison also the ANIX and reflectance of a neighboring rice field (dashed curve) and bare soil (dashed-dotted curve) are shown.

R_{max} is the maximum bidirectional reflectance factor and R_{min} the minimum bidirectional reflectance factor observed for wavelength λ at observation/illumination constellation Ω . In Figure 5.7(left) it can be seen that ANIX for all considered land cover types is largest in the visible domain and presents a relatively sharp decrease to lower variation at the onset of the red edge. The high ANIX peak observed for rice in the red wavebands is a result of dividing maximum reflectance by a very low minimum reflectance at nadir (Figure 5.6) and suggests that radiometric calibration at these wavebands is inaccurate, providing reflectances that are too low. On the other hand, at the time of observation, the rice paddy was filled with water, which due to its strong absorption could have induced the nadir reflectance to be close to zero in the red waveband. For the rest, it is noteworthy that, although very different in shape, the amplitude of anisotropy in the NIR is very similar for the cotton and rice fields. Soil generally shows less variation.

The additional discriminative power of angular anisotropy is shown when the left plot in Figure 5.7 is compared the right one, in which the average nadir spectrum of each of the 4 cotton fields, the rice paddy, and the bare soil is plotted. Whereas soil and vegetation can be clearly distinguished based on spectral properties alone, discrimination between cotton and rice is subject to much more ambiguity, especially when spectral variation is included as well. The examples in Figure 5.7 clarify that canopy variable retrievals, but also land cover classifications, can be significantly improved when both the spectral and angular dimensions are concurrently explored.

5.3.2 Reconstructing spectral anisotropy with PROSPECT and SAILh

The examples shown in the previous paragraphs and the cited literature elucidate the dependence of directional anisotropy on structural vegetation elements and illumination/observation geometry. In this context, it is important to gather insight into the degree to which the various structural canopy variables are responsible for the observed shifts in anisotropy. But even more

important it is to see whether the PROSPECT+SAILh combination is able to reproduce the directional behavior observed for the cotton canopy, since SAILh assumes a horizontally homogeneous canopy and does not account for vegetation clumping in the form of row structures. Anisotropy is therefore a function of volume scattering while a term describing mutual shading is absent. To study the ability of SAILh in representing the current situation, spectral reflectance (HDRF) of the cotton canopy was simulated for the 5 different observation geometries as delineated by the CHRIS observations. For the canopy characterization, average measured Cab ($= 25 \mu g \cdot cm^{-2}$) and LAI ($= 1.5 m^2/m^2$) were used. The other variables were either measured at the calibration plot ($Cw = 0.0265 g \cdot cm^{-2}$, $Cdm = 0.0093 g \cdot cm^{-2}$, $ALA = 30^\circ$, $HS=0.15$) or taken from literature ($Cbp = 0.001$, $N = 2$, $BS = 1.0$). Atmospheric conditions were similar to those employed for atmospheric correction, while the nadir soil spectrum plotted in Figure 5.7(left) was used as background reflectance.

Figures 5.8, 5.9, and 5.10 show the respective anisotropy behavior when one of the structural variables LAI , ALA , and HS is gradually changed. Only for green reflectance, changing LAI significantly alters the anisotropy signature. However, also for this wavelength changes in absolute reflectance remain rather low. With LAI approaching 0, anisotropy vanishes, which coincides with the assumed Lambert background reflectance. For the red and NIR domain, changing LAI mainly induces a shift in absolute reflectance. The effects of varying ALA are apparent. The erectophile canopies allow enhanced view of the soil background at nadir sensor position which, due to the large contrast in soil and vegetation reflectance in the red leads to the well known bell shape in this domain (see also Figure 5.7(left)). In the NIR the opposite is true: the erectophile canopy structure reduces multiple scattering in the nadir direction, leading to the observed bowl shape. In the green waveband, significant changes in absolute reflectance can be observed, whereas alterations in anisotropy at this wavelength are little but still show a minor bowl shape for erectophile canopies. Varying HS induces similar effects in the green and NIR domain, where increasing HS (i.e., increasing leaf size) changes the anisotropy shape from a slight concave form into a moderate bell form with maximum reflectance at the constellation closest to the direction of maximum backscattering. Also for red reflectance an enhanced bell shape is observed when HS is increased. Nevertheless, changes in the simulated constellations are not as large as one would expect close to the hot spot where spectral reflectance is much more sensitive to this parameter.

Even if the simulations show the spectral sensitivity of SAILh at the considered view/sun constellation, some differences compared to the measured CHRIS reflectance still occur. Although the decline of reflectance in the forward scattering direction, generally observed for all wavebands in the CHRIS data, is partly reproduced by the simulations, the extent of the decrease is much greater in the measured data. Especially for the green waveband, reflectance measured for the extreme forward scattering angle drops to a value that is far less than any of the values reproduced by the simulations. This could be explained by the fact that SAILh lacks a module describing shading caused by geometrical objects (rows): in the forward scattering direction the sensor views the shaded sides of the cotton rows, leading to decrease in observed reflectance with respect to a turbid medium surface. A second anomaly between observed and modeled anisotropy, is that all simulations in the red domain show a more or less bell shaped anisotropy, while this pattern cannot be recognized in the measured data.

Although some of the differences between observed and modeled anisotropy can be ascribed to the generalization of the input variables (e.g., averaging of the measured variables and assuming standard values for the other variables) this would not explain all anomalies between the

5.3. Quantifying spectral anisotropy

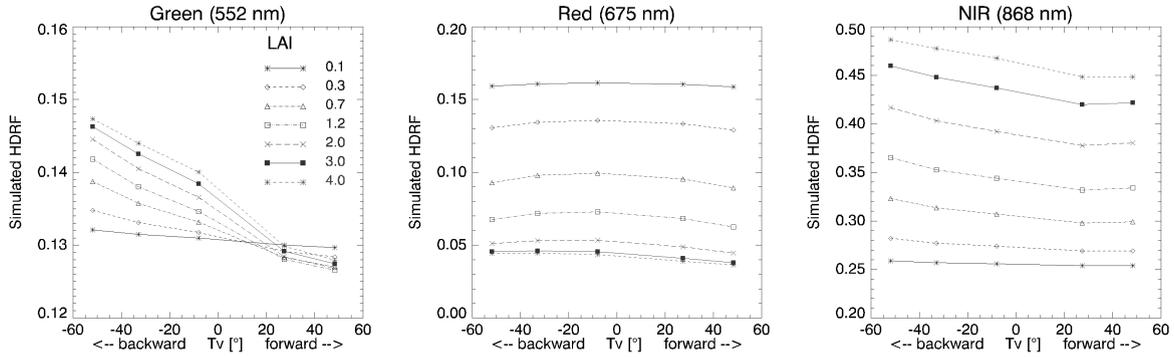


Figure 5.8: Influence of changing LAI on spectral anisotropy, simulated with PROSPECT and SAILh for green, red, and NIR reflectance. Tv [°] = view direction with negative values in the backscatter direction and positive values in the forward scatter direction. Simulations were based on the following input variables: $Cab = 25 \mu g \cdot cm^{-2}$, $Cw = 0.030 g \cdot cm^{-2}$, $Cdm = 0.010 g \cdot cm^{-2}$, $Cbp = 0.001$, $N = 2$, $LAI = [0.1, 0.3, 0.7, 1.2, 2.0, 3.0, 4.0] m^2/m^2$, $ALA = 50^\circ$, $HS = 0.015$, $BS = 1.0$.

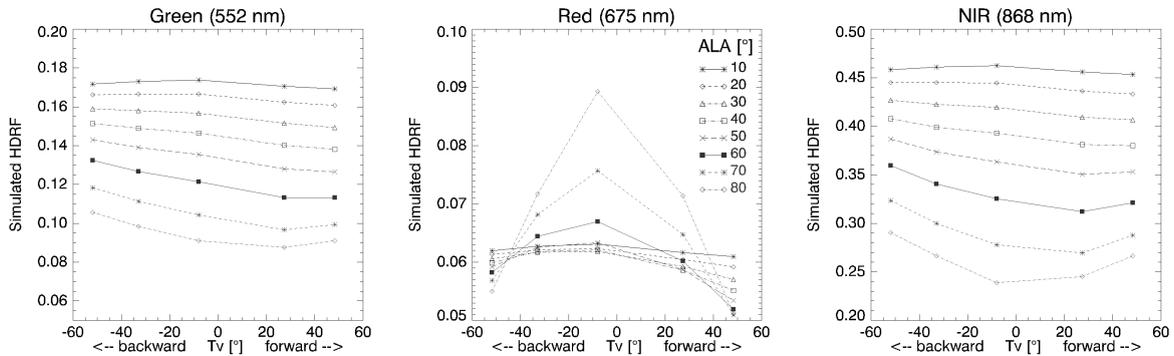


Figure 5.9: Influence of changing ALA on spectral anisotropy, simulated with PROSPECT+SAILh for green, red, and NIR reflectance. Tv =view direction with negative values in the backscatter direction and positive values in the forward scatter direction. Simulations were based on the following input variables: $Cab = 25 \mu g \cdot cm^{-2}$, $Cw = 0.030 g \cdot cm^{-2}$, $Cdm = 0.010 g \cdot cm^{-2}$, $Cbp = 0.001$, $N = 2$, $LAI = 1.5 m^2/m^2$, $ALA = [10, 20, 30, 40, 50, 60, 70, 80]^\circ$, $HS = 0.015$, $BS = 1.0$.

anisotropy characterizations. Undoubtedly, part of the deviations can be ascribed to the deficiency of SAILh in describing gap-driven anisotropy caused by the row structures of the canopy. This assumption is confirmed by the observed ability of the model in accurately reproducing the typical bowl shaped NIR anisotropy for the dense rice canopy, a surface type for which the heterogeneity assumption is legitimate (result not shown).

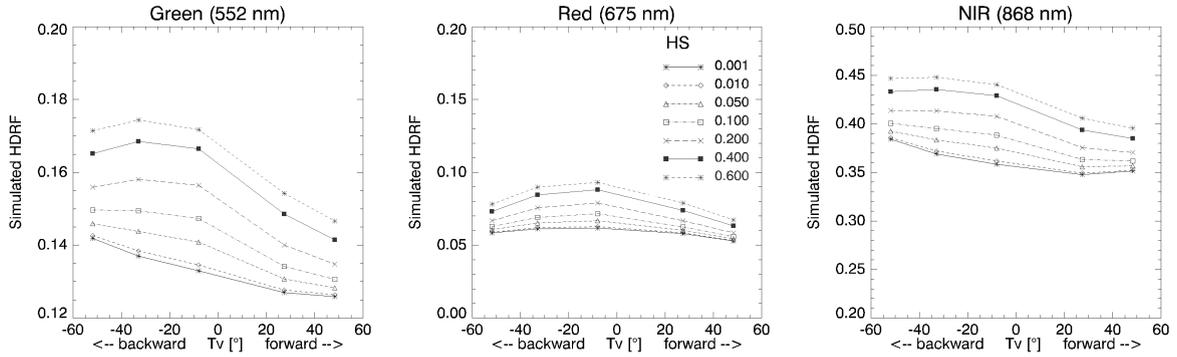


Figure 5.10: Influence of changing hot spot parameter HS on spectral anisotropy, simulated with PROSPECT and SAILh for green, red, and NIR reflectance. Tv = view direction with negative values in the backscatter direction and positive values in the forward scatter direction. Simulations were based on the following input variables: $Cab = 25 \mu\text{g}/\text{cm}^2$, $Cw = 0.030 \text{ g}/\text{cm}^2$, $Cdm = 0.010 \text{ g}/\text{cm}^2$, $Cbp = 0.001$, $N = 2$, $LAI = 1.5 \text{ m}^2/\text{m}^2$, $ALA = 50^\circ$, $HS = [0.001, 0.01, 0.05, 0.1, 0.2, 0.4, 0.6]$, $BS = 1.0$

5.4 RTM inversion using the fully automated CRASH approach

The previous section demonstrated that, even if the SAILh model is able to reproduce angular anisotropy for the fields under observation, the magnitude of the variations is not completely mimicked. If deviations are large, the introduction of multiple view angles might even introduce additional uncertainty to the result instead of restraining the set of viable solutions. For this reason, first the retrieval performance for the single view angles was tested before feeding the model with multi-directional data.

5.4.1 RTM inversion of individual observation angles

For the single observation angles, the complete approach, like introduced in Section 3 and validated in Section 4, was employed, including the SPECL land cover classification, the generation of a separate LUT for each view/sun constellation, and incorporation of a priori estimates. However, the examples shown in Chapter 4 demonstrated that in cases where there is a mismatch between modeled and measured reflectance, the introduction of covariance between spectral bands gives excessively high weight to the visible wavebands, thus leading to overestimations of leaf chlorophyll content. Since there was some indefiniteness on the nature of the very low red reflectance observed at nadir for the rice paddy (being either due to water absorption or to inaccurate calibration), it was decided to use spectral variance (i.e., class variance plus local variance) instead of covariance. The spectral cost function was therefore similar to the maximum likelihood estimator presented in Equation 2.10.

Following the same argumentation line as for the radiometric cost function, also for minimization in the variable space the use of variance instead of covariance was preferred. In a later phase, this decision was supported by the notice that no significant ambiguity between the estimates of Cab , LAI , and ALA would be observed when variance instead of covariance was used for the prior estimates (Cfr. Section 4.3.4.5). This is probably related to the less satu-

rated reflectance signals compared to the HyMap observations discussed in the previous section. When the function to minimize in the variable space χ_{var}^2 does no longer account for covariance between the variables, it is expressed as:

$$\chi_{var}^2 = \sum_{k=1}^{n_{var}} \frac{(V_{prior}^k - V_{LUT}^k)^2}{\sigma_k^2} \quad (5.4)$$

V_{prior}^k is the prior estimate of variable k left free during inversion and constitutes the average of the prior estimates predicted for the separate view angles. V_{LUT}^k is the LUT entry for variable k . Similarly to the spectral variance, σ_k^2 involves both class variance and the variance encountered in the direct neighborhood of the pixel. It has to be memorized that SPECL land cover classification, which has been designed for classifying nadir reflectance signatures, was applied to all view angles independently, which may lead to unforeseen classification results at high view zenith angles.

The validation of model inversion results was performed by comparing the average estimate of a 2×2 pixels window with the average of all field measurements found within this window. This was done in order to compensate for co-registration errors and small scale spatial variations. Table 5.5 summarizes the results obtained for all view angles. Although not being a biochemical quantity, for reasons of comparison also the goodness of fit between chlorophyll estimates and measured SPAD values is presented. For SPAD and chlorophyll, the results are given both for the top layer (TL) and for the average of all layers (AL). The best and worst performance for each variable have been marked in bold and italics, respectively.

The results in the table reveal that, except for the $+36^\circ$ nominal view angle, *LAI* is moderately well predicted, with relative RMS errors (Equation 4.4) around 40% and average accuracies (Equation 4.5) between 62.6 and 74.5%. Figure 5.11 (left) shows *LAI* estimates for the -55° nominal view angle. It can be observed that *LAI* is predicted evenly well for all fields, whereas part of the values measured at field 1 and 4 are underestimated. Remarkably, the estimated values of *Cab* coincide better with measured SPAD values than with measured *Cab*. The latter is in most cases overestimated, (Figure 5.11 (right)). Due to the nearly linear relationship between SPAD and measured *Cab* (Cfr. Figure 5.3), the patterns in the relationship between measured and predicted values do not significantly change, but are rather reflected by a shift (Figure 5.11 (middle-right)).

A very interesting aspect is revealed by Figure 5.12 which shows the column means of the values given in Table 5.5, i.e., average rRMSE and average AA, in relation to the observation angle. Although the accuracy descriptions are based on only 2 independent variables, a clear trend can be distinguished, showing decreasing retrieval accuracy with increasing forward scattering direction. This phenomenon agrees very well with the findings made in the previous section, which divulged the deficiency of SAILh in reproducing the decrease in reflectance in the forward scattering direction. The fact that on average, estimations are best in the directions of maximum backward scattering coincide well with the results found by Bacour (2001). In addition, at increased view angles the sensor captures less gaps within the vegetation canopy and, as a consequence, the observed surface again approaches a turbid medium leading to a reduced discrepancy between SAILh simulations and observations. The better estimates for the extreme backscatter direction occurred despite the large difference in land cover classification encountered for this view angle with respect to the nadir direction (result not shown).

Table 5.5: Goodness of fit between measured SPAD, *Cab*, and *LAI* and the values obtained from 5 separate CHRIS view angles (given is the nominal view angle). rRMSE = relative RMSE (%); AA = average accuracy (%); TL = top layer; AL = average of all layers. The best result for each variable is represented in bold, whereas the least accurate result is marked in italics.

Variable	unit	-55°		-36°		0°		$+36^\circ$		$+55^\circ$	
		rRMSE	AA	rRMSE	AA	rRMSE	AA	rRMSE	AA	rRMSE	AA
SPAD (TL)	unitless	30.1	75.2	19.8	83.3	26.8	77.2	33.2	69.7	<i>41.0</i>	<i>65.1</i>
SPAD (AL)	unitless	<i>38.0</i>	72.6	21.8	83.4	22.6	82.3	37.4	<i>69.1</i>	34.9	72.9
<i>Cab</i> (TL)	$\mu\text{g} \cdot \text{cm}^{-2}$	30.0	59.4	46.2	28.6	57.9	9.3	41.5	50.6	<i>75.4</i>	<i>-17.8</i>
<i>Cab</i> (AL)	$\mu\text{g} \cdot \text{cm}^{-2}$	28.9	62.6	40.0	46.2	50.3	27.0	38.3	59.0	<i>68.0</i>	<i>-0.4</i>
<i>LAI</i>	m^2/m^2	37.1	74.5	44.5	68.9	36.9	74.3	55.0	<i>37.3</i>	36.4	62.6

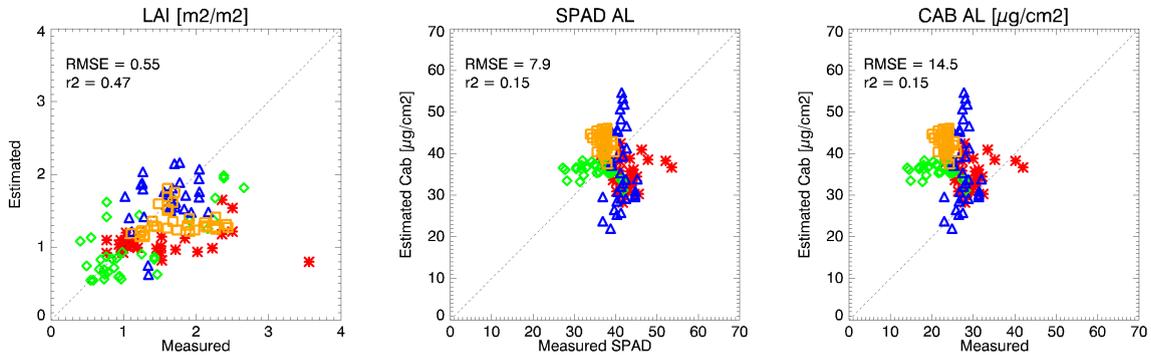


Figure 5.11: Comparison of estimated *LAI* and *Cab* with measured *LAI*, SPAD, and *Cab*. *LAI* estimates are based on the -55° nominal view angle, *Cab* estimates on the -36° nominal view angle. SPAD and *Cab* are the average values of all canopy layers (AL). Red asterisks, green triangles, blue diamonds, and orange squares indicate measurements taken at field C1, C2, C3, and C4, respectively.

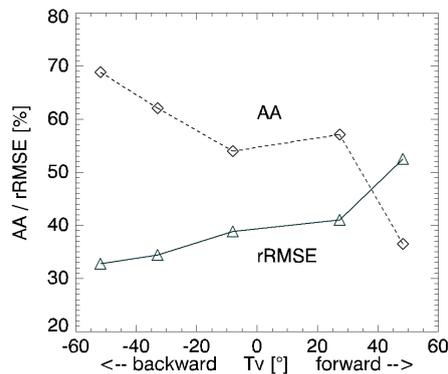


Figure 5.12: Evolution of average rRMSE and AA in relation to view angle. The values represent the column means of the values reported in Table 5.5.

5.4.2 RTM inversion of combined observation angles

5.4.2.1 Adapting the cost function

The results presented in the previous section revealed a clear trend toward increased accuracy for view angles with enhanced backscattering. Based on this notion, the -55° view angle was used to test whether it was a right decision not to include spectral covariance. Several cost functions were tested (with and without using prior information, using variance/covariance for radiometric information and prior estimates) and it appeared that including spectral covariance in this case indeed deteriorated retrieval accuracy. The cost function used to minimize for the multi-angular observations χ_{rad} is thus expressed as:

$$\chi_{rad}^2 = \sum_{i=1}^{n_{dir}} \sum_{j=1}^{n_{bands}} \frac{(R_{meas}(\Omega_i; \lambda_j) - R_{LUT}(\Omega_i; \lambda_j))^2}{\sigma_{i,j}^2} \quad (5.5)$$

where n_{dir} is the number of view geometries Ω_i , and n_{bands} the number of spectral CHRIS bands, in this case 31. R_{meas} and R_{LUT} denote the measured and simulated reflectance in the LUT, respectively. $\sigma_{i,j}^2$ is the uncertainty associated to waveband j in view direction i and which for each pixel is the sum of the variance of the land cover class the pixel belongs to, and the local variance encountered in a 5×5 pixels moving window (Section 3.4.1). It was decided to apply the same weight to all observation geometries in the cost function since no a priori information was available on the sensitivity and accuracy of the different view angles. Even if the backscatter direction is generally more sensitive to spectral changes and should therefore receive increased weight (Bacour et al., 2002a), this is compensated by the increased model uncertainties at extremer view angles (Verhoef and Bach, 2007).

5.4.2.2 Iteratively adding view angles

Starting with the best performing single view angle (-55°) and using the cost functions in Equation 5.5 and 5.4, view angles were iteratively added. The results were validated only for the variables of ultimate interest, namely the LAI and the Cab averaged over all shoot levels. The latter was preferred over using only the top level measurements since the greater number of measurements made the measurement more robust. Figure 5.13 visualizes the behavior of the rRMSE and AA when the view angles are iteratively added.

Introducing additional view angles gradually improves estimation results until all nominal view angles between -55° and $+36^\circ$ are included. Additionally incorporating the $+55^\circ$ nominal view angle induces a sharp loss in accuracy. The total absolute gain amounts to approximately 6% both for rRMSE and average accuracy. Although the gain compared to the mono-directional observation at -55° is not very high, the uncertainty of the estimations is significantly reduced when multi-angular observations are introduced, especially for the areas with less dense vegetation cover (Figure 5.14 and 5.15). Despite the fact that the view angles that have been added perform less than the -55° view angle in the mono-directional inversion scheme, the more restrictive spectral constraints posed by the multi-angular inversion, which requires the modeled

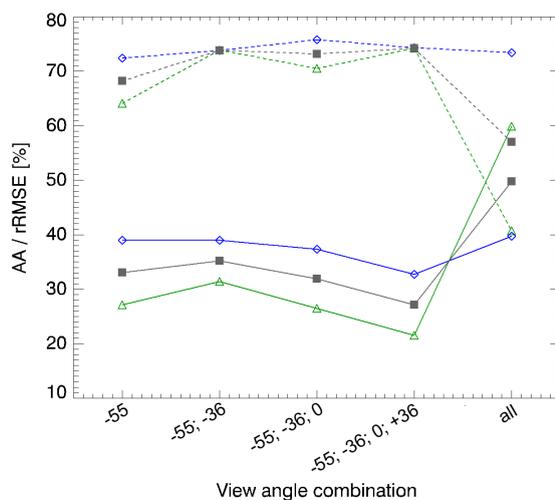


Figure 5.13: Evolution of average rRMSE (solid) and AA (dashed) in relation to different combinations of view angles. Green triangles represent *Cab*, blue squares *LAI*, and the gray squares the average of both.

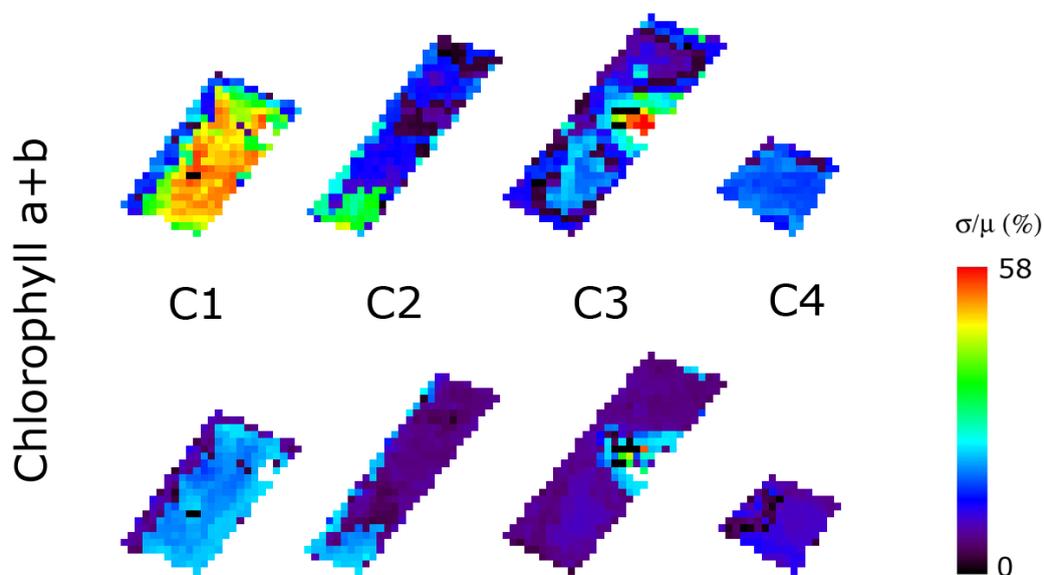


Figure 5.14: Comparing the spatial distribution of relative standard errors for mono- and multi-directional model inversion results of *Cab*. The upper row shows the errors associated to the inversion of the -55° nominal view angle, the bottom row the errors associated to the concurrent inversion of the nominal view angles -55° , -36° , 0° , and $+36^\circ$. The relative standard error is obtained by dividing the uncertainty of the estimate by the estimate itself and multiplying it with 100%.

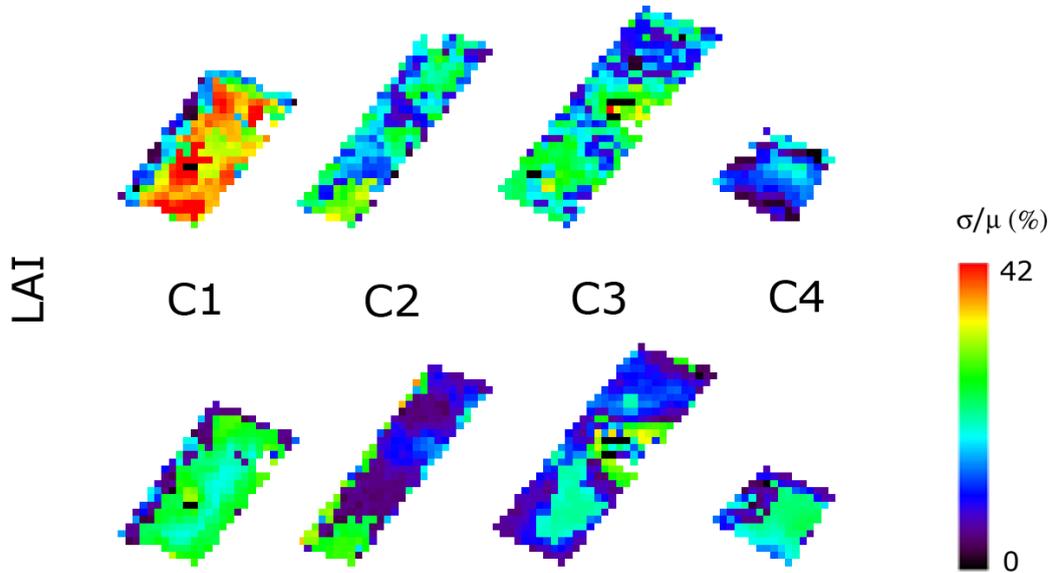


Figure 5.15: Comparing the spatial distribution of relative standard errors for mono- and multi-directional model inversion results of LAI . The upper row shows the errors associated to the inversion of the -55° nominal view angle, the bottom row the errors associated to the concurrent inversion of the nominal view angles -55 , -36 , 0 , and $+36^\circ$. The relative standard error is obtained by dividing the uncertainty of the estimate by the estimate itself and multiplying it with 100%

anisotropy to be in line with the measured one, puts additional limitations on the range of possible canopy structure variables resulting in more accurate estimations. Additionally, the prior information input to model inversion is based on the average of the prior estimates predicted with four different regression equations (one for each view direction), and therefore more robust.

The accuracy and stability of the estimated values significantly increased when the inversion is based on multiple view angles, leading to acceptable results in terms of RMSE and average accuracy. Figure 5.16 shows the spatial distribution of estimated Cab and LAI . Altering zones of high and low Cab and LAI within the fields are well visible and generally coincide with the patterns observed for the field measurements in Figure 5.4, suggesting that the obtained maps can be used for more targeted fertilizer applications at intra-field level. However, Cab estimates observed for field C3 and C4, show a considerably higher variation compared to the measured values, including some remarkable overestimations (Figure 5.17). The reason for this is yet unclear. In the next section it is investigated if these inaccuracies can be further reduced when the range of variation of the variables in the LUTs are limited to the ones actually measured in the field, thus optimizing the inversion for the cotton Khorezm-127 variety in the current situation.

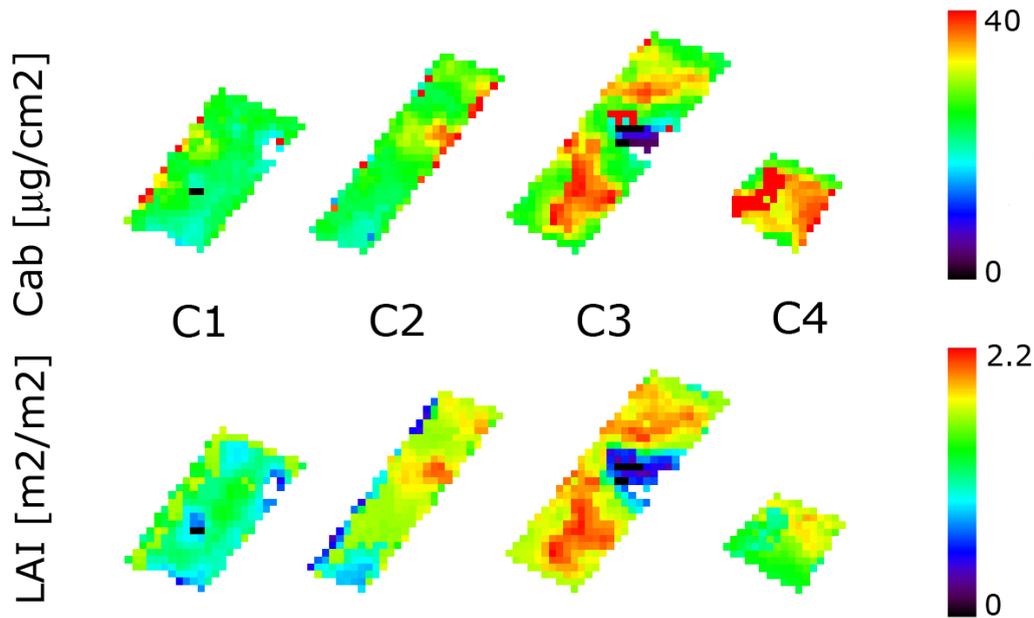


Figure 5.16: Spatial distribution of Cab and LAI for multi-directional model inversion based on the nominal view angles -55° , -36° , 0° , and $+36^\circ$.

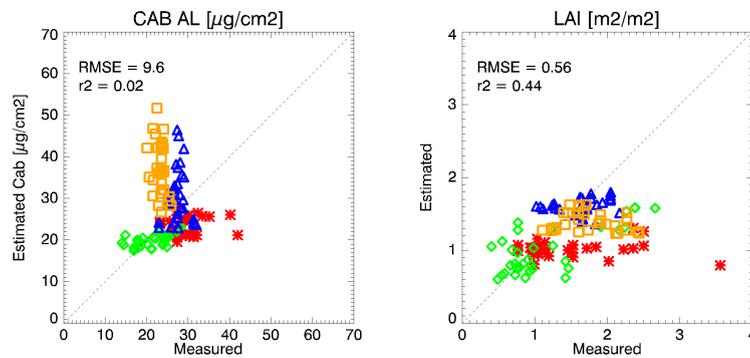


Figure 5.17: Comparing measured with estimated Cab and LAI , obtained from the concurrent inversion of -55° , -36° , 0° , and $+36^\circ$ nominal view angles. Cab values are based on the average of all canopy layers (AL). Red asterisks, green triangles, blue diamonds, and orange squares indicate measurements taken at field C1, C2, C3, and C4, respectively.

5.5 RTM inversion including measured a priori information on cotton

5.5.1 Parametrization of the LUTs

Up to this point, the obtained results were all based on the complete automatic approach proposed in Chapter 3. Parametrization of the LUTs was based on the SPECL classification and, accordingly, the range and distribution of the variables from which the solution of the inverse

Table 5.6: Distribution of the input variables and the corresponding number of classes of the orthogonal sampling plan used to construct the LUT for cotton.

	Variable	Unit	Distribution	Minimum	Maximum	Mean	σ	# intervals
Leaf	N	Unitless	Gaussian	1.5	2.5	2	1	3
	Cab	$\mu g \cdot cm^{-2}$	Gaussian	10	50	25	6	5
	Cw	$g \cdot cm^{-2}$	Gaussian	0.0024	0.0852	0.0265	0.0153	3
	Cdm	$g \cdot cm^{-2}S$	Gaussian	0.0011	0.028	0.0093	0.0052	5
	Cbp	Unitless	Gaussian	0	0	-	-	1
Canopy	LAI	m^2/m^2	Gaussian	0.1	4.7	1.5	0.9	5
	ALA	$^\circ$	Gaussian	5	60	25	20	5
	HOT	Unitless	Gaussian	0.05	0.3	0.12	0.1	3
Soil	BS	Unitless	Gaussian	0.7	1.3	1.0	0.2	3
							Total #	50625

problem was selected could have been far from optimal for the selected cotton species. To test whether estimations of Cab and LAI could be further improved when LUT parametrization is optimized for the Khorezm-127 variety in the considered phenological stage, distribution and ranges of the variables like they were measured in the field experiments were used as input for the construction of the LUTs (Table 5.6). The distributions used for sampling Cab and LAI were based on the probability density distributions composed of all point measurements (Figure 5.18). Distributions based on average measured values per pixel would probably have been more centered around the mean values since the inclusion of several measurements in one pixel will level out the observed extremes. Except for the leaf structure parameter N and the average leaf angle, all variables were measured during the field campaign. The distribution of N was based on literature review, whereas from field photographs a planophile leaf angle distribution could be established for ALA .

The SPECL land cover classification was abandoned, and instead, all 4 cotton fields were grouped into one "cotton" class for which one and the same LUT was deployed. Also computation of the class based variance terms was thus based on all pixels in the 4 cotton fields, irrespective of the observed vegetation density.

5.5.2 Results and discussion

RTM inversion was performed for the best performing combination of view angles of the previous section, i.e., for the combination of the -55° , -36° , 0° , and $+36^\circ$ nominal view angles. Figure 5.19 shows that in this case good agreement is achieved between measured and estimated Cab , which is confirmed by the low rRMSE (18.2%) and high average accuracy (75.9%). Nevertheless, retrieval accuracy is only a few percent better than in the fully automated case.

For LAI , rRMSE (41.8%) is slightly worse, and average accuracy (76.5%) is similar compared to the values obtained with the fully automated inversion scheme based on SPECL land cover

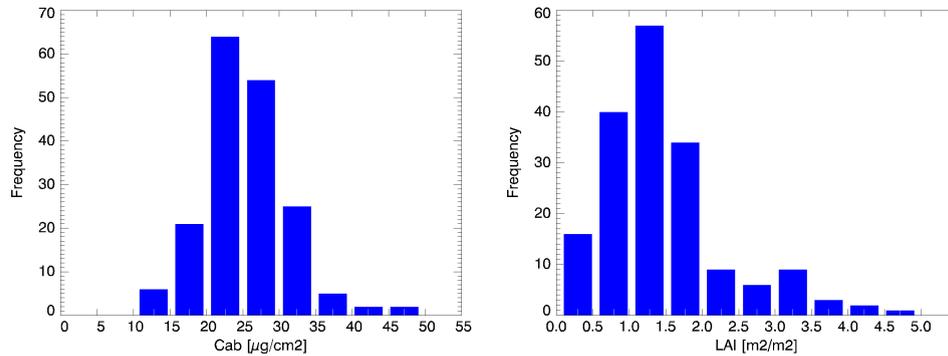


Figure 5.18: Frequency distribution of field measured Cab and LAI.

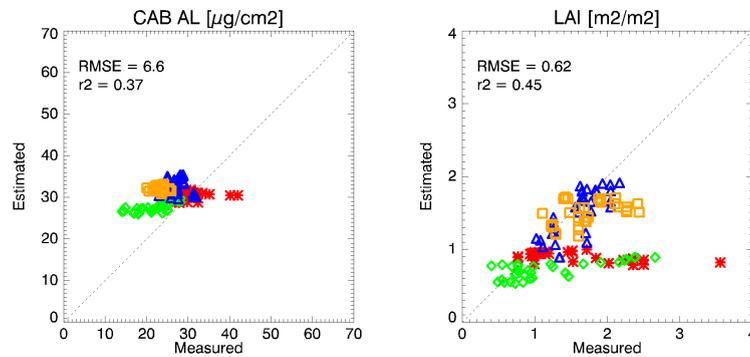


Figure 5.19: Comparing measured with estimated Cab and LAI, obtained from the concurrent inversion of -55 , -36 , 0 , and $+36^\circ$ nominal view angles using field measurement values for the parametrization of the LUT. Measured Cab is based on the average of all canopy layers (AL). Red asterisks, green triangles, blue diamonds, and orange squares indicate measurements taken at field C1, C2, C3, and C4, respectively.

characterization. LAI estimates for the more densely vegetated fields C3 and C4 (Cfr. Figure 5.4) clearly correlate with the measured values. For fields C1 and C2 LAI estimates seem to be trapped around 1. Similar behavior of estimates for C1 and C2, though less pronounced, could already be observed in the automated approach (Cfr. Figure 5.17). Probably, the inability of SAILh in completely mimicking anisotropy in sparse, row-structured canopies, is compensated by an unnatural variation in several canopy variables (e.g., dry matter, leaf water, and ALA). When in turn these variables are forced to adopt a value that is close to the actual one, the dissimilarity between measured reflectance anisotropy and model simulation is counterbalanced by adjusting LAI which relative to its spectral sensitivity has a large degree of freedom. This assumption is underscored by the finding that estimations for the denser canopies correlate well with measured values. Similar findings were obtained in a theoretical study performed by Widłowski et al. (2005), who observed underestimation of LAI in the cases where a 1-D turbid medium RTM was used to describe a complex 3-D canopy. Further, also scaling issues play a role when the limited number of point measurements is transferred to an area of 2×2 pixels. This scaling issue will have to be addressed in a future study.

5.6 Conclusions

This chapter explored the capability of the automated CRASH approach presented in Chapter 3 in retrieving cotton leaf chlorophyll content from multi-angular CHRIS/PROBA data. It was shown that the 1-D turbid medium radiative transfer model SAILh, which forms the basis of the approach, was not able to completely reproduce the spectrodirectional signatures of the cotton fields as observed by the sensor. Bias between modeled and measured reflectance was smallest in the extreme backscatter directions, which can be attributed to the reduced effects of mutual shading induced by the planting rows, a phenomenon that is not accounted for by SAILh. Thus, in this view direction the turbid medium assumption is most valid. Due to the larger shading component common to the forward scattering region, large bias between modeled and measured reflectance was observed for the $+55^\circ$ nominal view angle. The increased discrepancy between model and observation toward forward scattering directions was also reflected in the decreased accuracy in the estimates of both Cab and LAI based on RTM inversion of the mono-directional observations. In particular the $+55^\circ$ nominal view angle provided results that were significantly less accurate than the ones obtained for the other view angles. The best performing mono-directional data set (i.e., -55° nominal view angle) provided an rRMSE of 28.9 and 37.1%, and an average accuracy of 62.6 and 74.5% for Cab averaged over all leaf layers and LAI , respectively.

Inclusion of multi-angular data sets in RTM based inversions does not necessarily lead to improved retrieval performance. The benefit of introducing additional view angles is strongly determined by the ability of the RTM in mimicking the complete anisotropy of the canopy present in the imagery. When discrepancy between modeled and measured reflectance is larger than the range of variation allowed by the inversion procedure, too restrictive spectral constraints bear the potential of forcing the solution into unnatural variable combinations. In this study, progressively adding less performing observation angles to the best performing monodirectional data set (i.e., -55° nominal view angle) gradually improved retrieval performance until all viewing angles except for the extreme forward scattering angle were included. Model inversion of this 4-angle combination provided an rRMSE of 21.5 and 32.7%, and an average accuracy of 74.2 and 74.3% for Cab and LAI , respectively. Correlation coefficients were generally poor (0.02 and 0.43 for Cab and LAI , respectively) due to the scattered nature of the estimates around the 1:1 line. Additionally including the $+55^\circ$ viewing angle dramatically reduced retrieval performance.

In the considered case, the most remarkable effect of including angular anisotropy in model inversion is the reduction of the uncertainty around the final estimates, especially for sparsely vegetated areas. In some cases, errors could even be halved in comparison with the results obtained for mono-directional data. This underscores the potential of multi-angular observations not only in improving variable estimates but also in stabilizing the result and reducing the uncertainties associated to the retrievals.

The fact that discrepancy between measured and modeled anisotropy is mainly responsible for deviating model estimates was confirmed by a multi-angular RTM inversion where the variables measured for cotton during the field campaign were used to construct the LUT. While the inclusion of this well-defined a priori information led to significantly better Cab retrievals for all fields, only for the denser canopies an improvement in LAI could be observed. LAI estimates of sparser canopies were estimated less accurately than in the case where the variables were allowed to move more freely.

The work presented in this chapter demonstrated the potential of automated RTM inversion

of multi-angular hyperspectral remote sensing data for more accurate and stable retrieval of chlorophyll content in cotton leaves. In conjunction with *LAI* estimates, a more comprehensive picture of the spatial distribution of the current crop status was provided which could be a useful guideline to the farmers in supporting their crop management and fertilizer strategies. Due to the 1-D turbid medium assumption of the underlying RTM, application of the model should be limited to later growth stages and fertilizing dates when canopies are fully developed. In early crop development stages with sparse vegetation cover, SAILh will probably be less capable of accurately mimicking the anisotropy induced by leaf clumping and the directional behavior of soil reflectance. Model sensitivity to plant development will be more intensively addressed in 2008, when an extensive field validation campaign covering the entire growth season is planned. Moreover, the incorporation of an RTM accounting for row structures and bidirectional soil reflectance is envisaged. It is expected that the use of such an RTM will lead to a significant additional increase in accuracy.

Chapter 6

Synthesis

6.1 Introduction

Proper crop management decision making can be supported by spatially distributed biogeophysical products based on high resolution remote sensing systems. Due to the infrequent and irregular data availability from single high resolution sensors, multi-sensor approaches are required to fill the gaps in data sequences and to guarantee data takes at critical time steps during the seasonal life-time of a crop. Even if several contractors provide high and very high resolution precision farming commodities on a commercial basis, the degree of interchangeability between these products is often not clear. Since most of these products are based on mono-temporal, multispectral data sets, they underlie an underdetermined model inversion problem, induced by the large number of canopy variables potentially causing the observed spectral variations compared to the limited number of available independent information dimensions. Therefore, the generation of information products based on such sensors require several assumptions and restrictions.

To bridge the gap between farmers' needs and the commercial products already available on the market, this thesis presented a completely automated, image based, radiative transfer model inversion approach for the concurrent retrieval of the key agricultural variables leaf area index, leaf chlorophyll content, leaf water content, and leaf dry matter content from mono-temporal data recorded with any high resolution airborne or spaceborne imaging spectrometer. The choice to concentrate on imaging spectrometers was motivated by the demonstrated increased information content of the data provided by such sensors, especially on plant biochemistry (Fourty and Baret, 1997; Verhoef, 2007). The use of hyperspectral data is therefore assumed to reduce the underdetermined nature of algorithms focusing on the retrieval of canopy variables from remote sensing data.

In contrast to statistical-empirical approaches which usually concentrate on a limited number of spectral bands, the inversion of radiative transfer models offers the potential of concurrently exploiting the entire information content contained in the data. In addition, they facilitate optimizing for illumination / observation geometry and adapting to site specific phenology, background reflectance, and atmospheric conditions. To guarantee maximum flexibility of the approach, model inversion was based on lookup tables (LUTs) which, apart from being compu-

tationally fast, have an open structure which makes them easily adaptable to changing sensor, observation, illumination, and phenological properties.

The presented approach, named CRASH, entirely relies on the spectral content contained within the image and the information products provided by the automated preprocessing (i.e., calibration, georectification, and atmospheric correction) of the data. The challenge of such a scene-based, automated approach is to overcome the lack of a priori knowledge on land cover, causing the inversion process to be a strongly underdetermined and ill-posed problem. To bring underdetermination down to minimum, CRASH was based on the relatively simple leaf optical model PROSPECT (Jacquemoud and Baret, 1990; Fourty et al., 1996) and the 1-D turbid medium canopy structure model SAILh (Verhoef, 1984; Verhoef, 1985) which taken together result in 9 variables that are not a priori known and are left free during inversion.

Two novel regularization techniques were proposed to address the ill-posedness. First, an automated spectral land cover classifier (SPECL (Richter, 2007a)) was introduced. The land cover classification facilitated a more explicit characterization of spectral uncertainty (Chapter 3) and enabled the model inversion to take place in more explicitly defined and restricted variable spaces. Although the concept of performing RTM inversion on a land cover basis is not new (Knyazikhin et al., 1999b; Lotsch et al., 2003; Houborg et al., 2007), no studies are known where the classification is directly implemented in the retrieval approach itself and is based on the same mono-directional data set. Moreover, in combination with the land cover based inversion scheme, this study proposed a new method, based on the local neighborhood of the pixel under inversion, for dealing with small scale attribution inconsistencies.

The land cover classification also played an important role in the second proposed regularization technique. The LUTs that were calculated separately for every land cover class and illumination / observation geometry allowed for the generation of semi-empirical predictive equations optimized for each specific situation. These equations were based on the regression between spectral vegetation indices (VIs) calculated from the reflectance spectra contained in the LUT, and the variables used to simulate these LUT reflectances, using the RTM in the forward mode. The predictive equations obtained in this way were subsequently used to calculate a priori estimates for each variable and for every vegetation pixel in the image (Chapter 3). In turn, the solution of the model inversion was forced to lie within plausible margins around the prior estimates as defined by their level of uncertainty.

6.2 Resuming the performance of the automated approach

The ultimate goal of this study was not the development of the automated approach as such, but rather the challenge of providing consistent results for a wide range of canopy types, sun/view constellations, and sensor configurations. For this purpose, performance and stability of the new approach were extensively tested at three spatial levels for different sensor configurations, including ground field spectrometer measurements, data from the airborne full-range imaging spectrometer HyMap, and satellite observations from the multi-directional CHRIS/PROBA sensor. Whereas the field spectrometer measurements allowed for a direct comparison between spectral signal and canopy characteristics while excluding spatial uncertainties and uncertainties in atmospheric correction, the airborne and satellite measurements offered a deeper insight into the effect of changing observation and illumination properties and spatial scaling issues. Validation

of model performance for field spectrometer and HyMap data was performed for several intensively used temperate grasslands in southern Germany (Chapter 4), and included the biometric sampling of leaf water content (Cw), leaf dry matter content (Cdm), and leaf area index (LAI). Ground sampling of leaf chlorophyll content (Cab) and LAI measured on irrigated cotton fields in Uzbekistan was used to test the performance of the approach for the CHRIS data (Chapter 5).

To provide insight into the different aspects of the approach, and to expose potential sources of uncertainty, several research questions have been put forward in Chapter 1. In the following, these issues will be recapitulated one-by-one against the background of the results obtained in this thesis.

Choice of the radiative transfer model Prerequisite for a successful retrieval of canopy variables by radiative transfer model (RTM) inversion is the choice of a RTM that accurately mimics the spectral behavior of the considered canopy. The combination of the leaf optical model PROSPECT (Jacquemoud and Baret, 1990; Fourty et al., 1996) and the 1-D turbid medium canopy structure model SAILh (Verhoef, 1984; Verhoef, 1985) was selected as a trade-off between model accuracy for different types of canopies on one hand and the limited number of input variables on the other. The latter may become extremely important in situations where little a priori knowledge is present on the distribution of the input variables (Pinty et al., 2004a).

For the various grasslands studied in Section 4, the combination of PROSPECT and SAILh appeared well capable in mimicking top-of-canopy reflectance within the the ranges of uncertainty posed by the field measurements. This was true for all observation geometries provided by the HyMap sensor, confirming the results found by others for homogeneous canopies (Baret et al., 2005b; Pinty et al., 2004b). In contrast, reflectance anisotropy of the cotton canopies in Chapter 5 could not be properly reproduced. The open row structure of these moderately vegetated fields gives rise to anisotropy caused by shadow-casting and non-Lambert soil reflectance, phenomena that are both not accounted for by SAILh. It was observed that SAILh performed better in the backscatter directions, where geometrical shadowing played a less significant role and the turbid medium assumption was most valid. Incorporation of geometrical functions describing row geometry would therefore be necessary for improved formulations of the radiative transfer equations in heterogeneous row crops.

Sensitivity of canopy reflectance to variables of interest The success of retrieving a certain canopy variable from remote sensing data by RTM inversion strongly hinges on the response of overall canopy reflectance to the variable in question (Verstraete and Pinty, 1996). This sensitivity differs from situation to situation and depends on canopy composition and illumination/observation geometry. The spectral correlation analysis performed in Chapter 4 clearly demonstrated the distinctive spectral sensitivity to Cw , Cdm , and LAI . In effect, for vigorous meadows, LAI could be considered the only variable responsible for spectral variations while Cw and Cdm showed highly collinear behavior with LAI . This is in good agreement with the results found by Vohland and Jarmer ((2007)). For the recently cut meadows, the situation was somewhat different, with Cdm showing a sensitivity that was clearly discernible from LAI and Cw . Nevertheless, the driving canopy variable in most situations is obviously LAI .

In the considered grassland cases, measured and modeled reflectance showed distinct sensitivity to changing variable values, with the latter showing only low to moderate collinearity

between the leaf absorbing materials and *LAI*. This discrepancy has its implications on the transferability of the predictive equations, based on the radiative transfer simulations, to the actual image data. Whereas, based on the predictive regression equations, *C_{dm}* and *C_w* should correlate quite well with spectral reflectance, with coefficients of determination (R^2) ranging from 0.43 to 0.87, this was not confirmed by the ground validation measurements. Nevertheless, RMS errors of predictive equations and final model inversion results agree fairly well.

Influence of sensor configuration The influence of system configuration on retrieval performance is one of the issues that is most difficult to get a grip on. It strongly depends on the number of bands in a specific wavelength region and the cost function used to minimize between modeled and measured reflectance. The definition of the cost function mainly governs the weights the single wavebands and, hence, the single variables, receive. In this thesis, a novel method was proposed to account for spectral covariance based on the spectral variation found within a specific land cover class. Introducing such covariance terms led to increased weight in the cost function of spectral regions with low reflectance, mainly the VIS. The influence this covariance term has on the accuracy of the final results mainly depends on the variable in question, on sensor calibration, atmospheric correction, and on the ability of the RTM to describe the canopy under consideration. Whereas the variance-covariance matrix should explicitly account for all these uncertainty terms, the way it was calculated in this study assumed that all uncertainty was enclosed in the spectral information of the respective land cover class and the local neighborhood of the considered pixel. This obviously violates reality in cases where systematic errors are particularly large, and is probably the reason that incorporation of the covariance terms degraded retrieval accuracy for the CHRIS observations (Chapter 5). In the case of the well-calibrated ASD field spectrometer measurements and atmospherically corrected HyMap observations of Chapter 4, it principally led to a redistribution of the retrieved variables. The decision whether to incorporate spectral covariance should therefore be guided by the quality of the spectral data, the suitability of RTM model, and the envisaged canopy variable. Such a decision could be supported by quality measures obtained during preprocessing (Bachmann et al., 2007).

Regardless of the use of covariance measures or the inclusion of a priori information, Chapter 4 revealed some clear relationships between spectral band configuration and the estimation accuracy of *C_w*, *C_{dm}*, and *LAI*. Based on resampled field spectra, the performance of different band configurations (HyMap, CHRIS, Landsat ETM+, SPOT HRG, Quickbird) was tested. Although the approach is applicable to any multispectral or hyperspectral band configuration, most consistent results were obtained for the HyMap configuration spanning the entire VNIR-SWIR domain, followed by the Landsat ETM+ configuration. This indicates that an even distribution of nearly uncorrelated bands across the entire solar-reflective domain is more important for a stable inversion than a high absolute number of bands in strongly correlating waveband regions. Obviously, the 6 bands of ETM+ provide significantly more uncorrelated information than the 37 bands of CHRIS which are all concentrated in the VNIR. The inclusion of SWIR bands clearly leads to regularization of the leaf water retrievals and hence to stabilization of the complete inversion. This is confirmed by the poor *C_w* predictions of the Quickbird, CHRIS, and to some degree, SPOT HRG configurations. Part of the deviations can also be explained by the sometimes better defined a priori estimates for hyperspectral data, which are a consequence of the enhanced number of sensitive vegetation indices to choose from for the generation of the

predictive equations.

A third aspect where band configuration plays a role, and which is directly related to the previously mentioned inversion stability, is the reduction of ambiguity between the estimated variables. Ambiguity is a direct result of the ill-posed nature of RTM inversion and indicates the complementary behavior of several variables in model inversion results. In Chapter 4 it could be seen that, in the case where covariance between the variables was neglected, this phenomenon is very apparent for the multispectral configuration, whereas for the hyperspectral configurations of CHRIS, and especially HyMap, this ambiguity is significantly lower. Concluding, it can be stated that a high number of spectral bands, equally spread across the solar-reflective domain, provides the best precondition for an accurate and stable RTM inversion of mono-directional data. This notice provides evidence for the increased information content contained in hyperspectral data as found by Verhoef (2007).

Reproducibility of results and sensitivity to LUT parametrization Due to the semi-random selection of LUT-variables, different model runs can provide diverging results (Chapter 4). The maximum degree of variation in the obtained results is inherent to the sensitivity of a specific variable in the considered situation but in most cases stayed within 20% of the average of several repeated model runs. Uncertainty is largest in situations of low spectral response to the variables, such as the case for most leaf biochemical and canopy structural variables of a sparsely vegetated area, or for the hot spot parameter in observation directions far from the principle plane. The repetition of 12 model runs firmly consolidated the results and for most situations reduced the uncertainty to less than 5%. Therefore, if computer resources allow so, such a repeated model run would be highly recommendable.

In Chapter 4 it was tested to which degree the inversion was dependent on the upper and lower boundaries posed to variable ranges used to construct the LUTs. It was shown that for the retrieval of C_{dm} , C_w , and LAI from the considered grassland spectra the default variable ranges as specified by CRASH were all close to the optimum. Nevertheless, only the dependence of the results on the change of one variable at a time was studied. Future work should therefore focus on multivariate approaches to study the effects of simultaneously changing input LUT ranges. A further optimization of the ranges could help to reduce the number sampling intervals per variable or to intensify sampling density in variable spaces where spectral response is high. Additionally, the sensitivity of retrieval results to the sampling distribution functions should be considered which however is a precarious task, since spectral sensitivity strongly depends on canopy composition and observation/illumination geometry.

Sensitivity of the approach to changing view/sun geometry The HyMap imagery presented in Chapter 4 offered a unique opportunity of studying the effects of changing view/sun geometry on estimation performance. It was shown that retrievals of variables with significant effect on canopy reflectance such as C_{ab} and LAI , but also average leaf angle (ALA), showed high consistency among the different constellations, whereas variables with low effect such as C_{dm} and the hot spot parameter HS showed very low consistency. Estimates of C_w , the leaf mesophyll parameter N , and the soil brightness coefficient BS showed intermediate congruity. These results clearly correlate with the trends revealed in the previous paragraph.

The bias in the obtained results between scenes with different view/sun geometry showed no direct parallel to the spectral anomalies between the scenes. This indicates that the PROSPECT

and SAILh combination was well able to mimic reflectance anisotropy resulting from the different view/sun constellations of the HyMap scenes. In contrast, discrepancies between different constellations were primarily induced by anomalies in land cover classification results which occurred even after nadir-normalization. Whereas irregularities in land cover classification had hardly any effect on sparse to medium-dense canopies, for dense canopies their influence was very large. It is suggested that due to the saturating spectral response of such dense canopies, small anomalies in the predictive equations for LAI and Cab have a large impact on the prior estimates of these variables, and hence on the final results of all estimates.

The study case presented in Chapter 5 provided interesting insight into the relationship between retrieval performance and view geometry for cases in which the assumptions made in the RTM are inadequate for describing radiation propagation in the canopy under consideration. It disclosed the superior retrieval accuracy for oblique backscatter view directions compared to nadir and forward looking directions. Geometrical shading, a component not accounted for by SAILh, is negligible in the backscatter directions, leading to a better approximation to the turbid medium assumption and, hence, to reduced discrepancy between modeled and observed canopy reflectance.

Performance of CRASH compared to better determined inversions The automated CRASH approach presented in this thesis had to deal with the lack of a priori knowledge on land cover and phenology, which severely hampered the definition of well-constrained variable ranges input to the generation of the LUTs. An inbuilt automated land cover classification and the inclusion of a priori estimates based on semi-empirical predictive regression equations were proposed to overcome this deficiency. The integration of both components led to improved accuracy of the solutions but above all contributed to better defined and more stable solutions. While the land cover classification primarily reduced the uncertainty of the solution, the integration of covariance between the generated a priori estimates significantly reduced collinearity and ambiguity between the estimates.

Several knowledge based regularization techniques were tested for their potency of providing additional improvement of retrieval accuracy, such as fixing the ratio between C_{dm} and C_w and a synthesis of leaf biochemicals and LAI . None of these regularization techniques provided better variable estimates in the considered grassland cases. Apparently, the incorporation of the covariance between variables in the cost function already led to the maximum possible reduction of ambiguity in the results. Even the introduction of the ground validation measurements of C_w , C_{dm} , and LAI as prior information in the inversion process, improved estimation accuracy (relative RMSE and overall accuracy) only about 10%, which indicates the fairly good performance of CRASH in cases where no a priori information on land cover is available. This additional gain in accuracy could be obtained for cases in which land cover and phenology is previously known so that RTM parameterization can be optimized for the considered situation.

For the row-structured cotton canopies, introducing a high amount of a priori information even led to decreased estimation accuracy for LAI . This observation provides important insight into the behavior of model inversion in the case of an ill-posed and under-determined problem. Very tight definitions of a priori information only leads to significantly better performance when the employed RTM is capable of accounting for all dominating radiation transfer processes and of providing reflectance factors that agree very well with the measured ones. Since there is always some discrepancy between measured and modeled reflectance, e.g., due to model assumptions

and radiometric and atmospheric uncertainties, always some small degree of variation should be consented to the several variables influencing canopy reflectance, even when their effect is not so straightforward. This would namely allow for better retrievals of variables with significant influence on the spectral signal at the cost of the prediction accuracy of variables having little influence on canopy reflectance.

Potential of additional regularization based on multi-angular observations Chapter 5 demonstrated the high potential residing in the simultaneous inversion of several view angles, even in the case where RTM and observed canopy are not completely compatible. The multi-angular observations provide a highly complementary information source to the hyperspectral dimension and reveal additional information on canopy structure. The enhanced regularization of canopy structure variables stabilized the entire retrieval procedure and at the same time led to improved estimations of leaf chlorophyll content as well. This potential has been recognized for many years by the research community (Bacour et al., 2002b; Knyazikhin et al., 1999c; Menenti et al., 2005; Pinty et al., 2002; Rast, 2004; Verhoef, 2007; Widlowski et al., 2004).

Absolute accuracy For the grassland study, the average accuracy obtained at ground level for the HyMap configuration was around 70% for all 3 considered variables and ranged from 61.6 - 88.8% if the grasslands were split up according to phenological classes. The minimum and maximum relative RMS errors (rRMSE) obtained at this level were 14.1 and 38.8%, respectively. At the level of the airborne HyMap data, average estimation accuracy was significantly lower and varied from 37.3 to 69.7%, depending on variable and sun/view constellation, while rRMSE laid between 35.6 - 61.8 % (Chapter 4). The unsatisfying results obtained at image level are ascribed to the large discrepancy between the small validation plots of $1 \times 1 m^2$ and the average HyMap pixel size of $5 \times 5 m^2$. Especially in the highly heterogeneous pasture this played a decisive role.

The results obtained CHRIS/PROBA observations are significantly better than the ones obtained at image level from HyMap data, although they are not directly comparable, since both canopy type, spectral sensor configuration, and considered variables differed. For the single view angles, highest accuracy was obtained for the 55° backscatter nominal view angle, providing rRMSE of 28.9 and 37.1%, and an average accuracy of 62.6 and 74.5% for *Cab* averaged over all leaf layers and *LAI*, respectively. Regarding the simultaneous inversion of several view angles, best overall performance when all view angles but the extreme forward scattering angle ($+55^\circ$ nominal view angle) were combined. This provided a rRMSE of 21.5 and 32.7% and an average accuracy of 74.2 and 74.3% for *Cab* and *LAI*, respectively, marking a significant improvement compared to the best performing single view angle.

6.3 Conclusions

Given the fairly accurate results obtained from the field spectrometer data and the encouraging results obtained for CHRIS, the approach presented in this thesis provides a substantial contribution to the development of automated algorithms for the generation of consistent biophysical and biochemical products based on mono-temporal high resolution earth observation

data. The large discrepancy between the results obtained at ground and airborne level indicate that additional effort should be put in validating the model with ground campaigns that are representative for the spatial resolution of the sensor (Baret et al., submitted). Only in this way the full potential of the approach can be explored and optimal use can be made of the spectral and spatial information dimensions provided by the image content. Several data campaigns supporting such validation and calibration activities are in progress or planned for the coming years and should provide a more exhaustive overview of model performance over varying land cover types, phenological stages, view/sun geometries, and sensor configurations. With regard to the latter, validation should be concentrated on full range superspectral or hyperspectral sensors such as the airborne sensors AVIRIS (Green et al., 1998), HyMap (Cocks et al., 1998), APEX (Schaepman et al., 2004), and ARES (Müller et al., 2005; Richter et al., 2005), and the planned Sentinel-2 and EnMap satellite missions (Gascon and Berger, 2007; Müller et al., 2004). This study demonstrated that, even if data redundancy is high for such systems, full range imaging spectrometers provide significantly more uncorrelated information than most multi-spectral sensor systems, or than imaging spectrometers covering only the VNIR. Inclusion of the SWIR region is indispensable for robust allocations of leaf biochemics such as leaf water content, whereas a high number of bands significantly reduces the ambiguity in the estimates of variables inducing complementary spectral behavior. With this notice, the study confirms the theoretical work performed by Verhoef (2007).

The high correlation between classification anomalies and inconsistencies in the estimated variables suggests that considerable inversion improvement can be obtained by applying more sophisticated land cover classification schemes. Such schemes should be less dependent on pure thresholds and could for example combine the spectral decision rules with segmentation algorithms for object recognition and fuzzy classifiers (Lucieer and Stein, 2002). Unnatural within-field transitions can thus be prevented and spectral characterization can be based on coherent and well defined single fields. It is believed that optimizing for land cover classification in the proposed way, would provide consistent estimations for homogeneous canopies of spectrally sensitive variables such as *LAI* and leaf chlorophyll content, but that consistent results could also be obtained for less sensitive variables such as leaf water and leaf dry matter content.

Even if there is still some room for optimization of the algorithm, especially for homogeneous surfaces, portability of the approach to more complex surfaces, such as row crops, is hampered by the limited intrinsic dimensionality of the data. Verhoef (2007) reports that, depending on the noise level, the number of uncorrelated dimensions in single-view, full-range, hyperspectral data varies between 6 and 12. Whereas for a homogeneous canopy theoretically all 9 variables could be retrieved from the TOC reflectance data (given that the variables are uncorrelated), the number of variables required for a simple representation of a row canopy (i.e., the 9 previously mentioned variables plus 4 parameters describing row height, distance, width, and direction (Kuusk, 1995b; Kuusk, 1995a; CROMA, 2002)), already exceeds the dimensionality of the data. Thus, a stable model inversion for such canopies can only be obtained by incorporating additional independent information dimensions, such as the temporal evolution of the vegetation (Koetz et al., 2005a), LIDAR data (Koetz et al., 2005c; Koetz et al., 2007), RADAR backscatter coefficients (Treuhaft et al., 2002; Treuhaft et al., 2004), or multi-angular observations (Menenti et al., 2005; Rast, 2004; Verhoef, 2007; Widlowski et al., 2004).

The automated CRASH retrieval approach presented in this thesis was based on the presumption that all the information should be obtainable from the (hyperspectral) remote sensing data source itself. In practice however, this would be a rather experimental constraint, as in precision

farming applications information on land cover, crop management, phenology, soil condition and so on is usually previously available. Therefore, CRASh has been programmed in such a way that, if complementary information on canopy variables or land cover is available, this can be integrated into the model, which in such cases will lead to more stable and better defined solutions. Nevertheless, the purpose of this study was to demonstrate which accuracy could be obtained by a fully automated and completely image based approach, starting from the theoretical baseline of zero a priori information, and to discover to which degree individual information sources and regularization techniques can contribute to improved retrieval performance.

6.4 Outlook

As demonstrated in Chapter 5, and as evidenced by many other authors, great potential lies in the synergetic use of imaging spectrometry and multi-angular observations for enhanced characterization of complex vegetated surfaces. The angular dimension principally provides information on canopy structure and is therewith highly complementary to the hyperspectral dimension being diagnostic for biochemistry. Although after the dismissal of the SPECTRA mission (Rast, 2004) in the near future the experimental CHRIS sensor will remain the only sensor providing spectrodirectional observations with high spatial resolutions, potential is also seen for the upcoming German EnMAP full-range imaging spectrometer (Müller et al., 2004). Through its tilting capacity up to 30° this sensor will be able to provide multi-angular composites of images collected during successive satellite overpasses. Nevertheless, precision farming products based on spectrodirectional observations of both CHRIS and EnMap will probably have a rather experimental character.

Instead of further exploiting the joint hyperspectral and angular dimensions of high resolution earth observation data, in the coming decade international political programs and commercial efforts mainly seem to concentrate on increasing the temporal coverage of multi- and super-spectral systems (Schreier and Dech, 2005). The system requirements of the ESA Sentinel-2 satellite, which should provide data continuity to Landsat and SPOT-like sensors, are mainly driven by the GMES Service Elements, requiring earth observation data for natural hazards management, food security, global change issues, humanitarian aid, and forest monitoring (Gascon and Berger, 2007). Very promising, in the context of improved product continuity and consistency for agricultural areas, will be the privately owned multi-satellite system RapidEye¹ which is to be launched in 2008. Although the spectral configuration contains only 5 spectral bands, the five-satellite constellation should guarantee near daily revisit capability. This would clearly contribute to improved land cover identification (Lotsch et al., 2003; Houborg et al., 2007) and to additional regularization of the inverse problem offered by the information on phenological development (Koetz et al., 2005a).

Quantitative characterization of the biosphere is moving toward integrated approaches, fusing earth observation data of different sensors, spatial resolutions, radiometric quantities, and revisit frequencies (Schaepman, 2007). The remote sensing signal will thus no longer be considered as an independent information source but interpreted in the context of the different data dimensions available prior to and during the time of observation. However, assimilation strategies will not merely rely on complementary remote sensing data sources, but will increasingly incorporate in situ measurements, canopy state variables generated by physical earth surface process

models, and ground based GIS information layers (INSPIRE, 2007; Schaepman, 2007; Schreier and Dech, 2005). The challenge of future remote sensing research will be to develop algorithms that carefully weight these multiple information sources while taking into account their uncertainties (Dorigo et al., 2007; Launay and Guerif, 2005). In this respect, great potential has been recognized in the coupling of soil-vegetation-atmosphere (SVAT) models with radiative transfer model approaches, and in other methods for incorporating a priori information with varying uncertainties in radiative transfer model inversion schemes (Lavergne et al., 2007; Verhoef and Bach, 2003a; Verhoef, 2007). In addition, to be able to correctly combine and interpret historic data records from different sensors, and to extend these records to future missions, further effort should be put in improving retrieval consistency for mono-temporal data sets and in developing reliable cross-sensor translation functions (Van Leeuwen et al., 2003).

¹URL: <http://www.rapideye.de>

Appendix A

SPECL classification rules

<i>Class name</i>	<i>Decision rules</i>									
snow	$b4/b3 \leq 1.3$	AND	$b3 \geq 0.2$	AND	$b5 \leq 0.12$	AND	$b5 \geq 0.2$			
cloud	$b4 \geq 0.25$	AND	$0.85 \leq b1/b4 \leq 1.15$	AND	$b4/b5 \geq 0.9$	AND				
bare soil	$b4 \geq 0.15$	AND	$1.3 \leq b4/b3 \leq 3.0$	AND						
bare soil (dark)	$b4 \geq 0.15$	AND	$1.3 \leq b4/b3 \leq 3.0$	AND	$b2 \leq 0.10$	AND	$b3 \leq 0.055$			
average vegetation	$b4/b3 \geq 3.0$	AND	$b1/b3 \geq 0.8$ OR $b3 \leq 0.15$	AND	$0.28 \leq b4 \leq 0.4$	AND				
bright vegetation	$b4/b3 \geq 3.0$	AND	$b1/b3 \geq 0.8$ OR $b3 \leq 0.15$	AND	$b4 \geq 0.4$	AND	$b4 \leq 0.28$			
dark vegetation	$b4/b3 \geq 3.0$	AND	$b1/b3 \geq 0.8$ OR $b3 \leq 0.15$	AND	$b3 \leq 0.08$	AND	$b4 \leq 0.28$			
yellow vegetation	$b4/b3 \geq 2.0$	AND	$b2 \geq b3$	AND	$b3 \geq 8.0$	AND	$b4/b5 \geq 1.5^a$			
mixed veget./soil	$2.0 \leq b4/b3 \leq 3.0$	AND	$5.0 \leq b3 \leq 15.0$	AND	$b4 \geq 15.0$	AND				
asphalt/dark sand	$b4/b3 \leq 1.6$	AND	$5.0 \leq b3 \leq 20.0$	AND	$5.0 \leq b4 \leq 20.0^a$	AND				...
	$5.0 \leq b5 \leq 25.0$	AND	$b5/b4 \geq 0.7^a$	AND						
sand/bare soil/cloud	$b4/b3 \leq 2.0$	AND	$b4 \geq 0.15$	AND	$b5 \geq 15.0^a$	AND				
bright sand/soil/cloud	$b4/b3 \leq 2.0$	AND	$b4 \geq 0.15$	AND	$(b4 \geq 0.25^b$	OR	$b5 \geq 0.30^b)$			
dry veget. / soil	$1.7 \leq b4/b3 \leq 2.0$	AND	$b4 \geq 0.25^c$	OR	$(1.4 \leq b4/b3 \leq 2.0$	AND	$b7/b5 \leq 0.83^c)$			
sparse veg. / soil	$1.4 \leq b4/b3 \leq 1.7$	AND	$b4 \geq 0.25^c$	OR	$(1.4 \leq b4/b3 \leq 2.0$	AND	...			
	$b7/b5 \leq 0.83$	AND	$b5/b4 \geq 1.2^c)$	AND						
turbid water	$b4 \leq 0.11$	AND	$b5 \leq 0.05^a$	AND						
clear water	$b4 \leq 0.02$	AND	$b5 \leq 0.02^a$	AND						
clear water over sand	$b3 \geq 0.02$	AND	$b3 \geq b4 + 0.005$	AND	$b5 \leq 0.02^a$	AND				

Table A.1: Decision rules used in SPECL classification, based on reflectance measured at Landsat TM central wave bands: b1 is located at 0.48 μm , b2 at 0.56 μm , b3 at 0.66 μm , b4 at 0.83 μm , b5 at 1.6 μm , b7 at 2.2 μm .

^aThese expressions are optional and only used if b5 is present.

^bDecision rule depends on presence of b5.

^cDecision rule depends on presence of b7

Appendix B

Variable sampling plan for LUTs

Distributions, variable ranges, and sampling intervals used generating LUTs for each SPECL land cover class in the automatic mode. Below the considerations are listed that were taken into account for assigning the parameters in the table.

B.1 SPECL class 2: Dark vegetation

	Variable	Distribution	Minimum	Maximum	Mean	σ	# intervals
Leaf	N	Gaussian	1.	3.5	2.0	1.0	3
	Cab	Comb	20	90	(50)	(50)	8
	Cw	Uniform	0.010	0.060	(0.015)	(0.005)	5
	Cdm	Uniform	0.0035	0.015	(0.01)	(0.003)	5
	Cbp	Gaussian	0.	1.5	0.	0.6	1
Canopy	LAI	Comb	0.5	6	(4)	(4.)	8
	ALA	Gaussian	25	70	57	20	3
	HOT	Gaussian	0.001	0.2	0.02	0.1	3
Soil	BS	Gaussian	0.3	1.1	0.7	0.3	1
Total # samples							43,200

- *Principle land cover characteristics:* green vegetation with low NIR reflectance, induced by a large shading component caused by vertical canopy structures. Mainly tree canopies fall into this category, but also agricultural land use with an inhomogeneous large scale canopy structures such as vine yards and horticultures. Strong spectral similarity with canopies composed of a mix of soil and vegetation: for this reason the class should also partly describe canopy types of this kind.
- N : should cover both fresh (green row crops) and partly senescent leaves in the case of

classifying mix of soil and vegetation, therefore average of 2.0 and sigma of 1.0 (Table 3.1. Relatively little influence on spectrum, therefore only 3 classes/steps.

- *Cab*: considered medium-high. Large influence on VIS. Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. One of end products → finer sampling steps.
- *Cw*: Little known about distribution → uniform distribution. Around 4 times higher than Cdm (In healthy vegetation relative water content is around 80%)
- *Cdm*: Little known about distribution → uniform distribution. Around 0.25 times Cw (In healthy vegetation relative water content is around 80%)
- *Cbp*: For the possibility of receiving misclassified pixels of mix soil/vegetation or senescent material it was decided to include Cbp with $\mu = 0$ $\sigma=0.6$.
- *LAI*: can cover a wide range of canopy types including a mix soil/vegetation and denser canopies. Therefore range is set from 0.5 - 6.0 (Forest canopies, which can have higher values are not considered in this case). Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. One of end products → finer sampling steps.
- *ALA*: Little is known about ALA. As the class mainly consists of shrub-like, a slightly planophile ALA (45) is assumed: Erectophile- and planophile distributions are less likely, so a Gaussian distribution with a σ of 20 assumed.
- *HS*: Vertically structured canopies are assumed: therefore leaf size is assumed small compared to canopy height and a Gaussian distribution with a small mean is adopted.
- *Bs*: Influence of soil reflectance is assumed little: Bs is reduced to 1 class with Gaussian variation around average. Because of high shading component (and soil moist), average Bs is set to 0.7 and σ to 0.2

B.2 SPECL class 3: Average vegetation

- *Principle land cover characteristics*: green crops and grasses in intermediate phenological stages, medium to high NIR reflectance
- *N*: fresh leaves, therefore average of 1.63 and sig. of 0.5 (Table 3.1. Relatively little influence on spectrum, therefore only 3 classes/steps.
- *Cab*: considered medium-high. Large influence on VIS. Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. One of end products → finer sampling steps.
- *Cw*: Little known about distribution → uniform distribution. Around 4 times higher than Cdm (In healthy vegetation relative water content is around 80%)

	Variable	Distribution	Minimum	Maximum	Mean	σ	# intervals
Leaf	N	Gaussian	1.	2.5	1.63	0.5	3
	Cab	Combal	20	100	(50)	(50)	8
	Cw	Uniform	0.0100	0.0700	(0.025)	(0.010)	5
	Cdm	Uniform	0.0035	0.0250	(0.007)	(0.003)	5
	Cbp	Fixed	0.	-	-	-	1
Canopy	LAI	Combal	1.	7	(4)	(4.)	8
	ALA	Gaussian	30	70	57	20	3
	HOT	Gaussian	0.001	0.3	0.05	0.2	3
Soil	BS	Gaussian	0.3	1.1	0.7	0.3	1
Total # samples							43,200

- Cdm : Little known about distribution \rightarrow uniform distribution. Around 0.25 times Cw (In healthy vegetation relative water content is around 80%)
- Cbp : Vegetation canopy itself is considered free of brown leaves, so Cbp is set to fixed value of 0. Incidental brown leaves are assumed to be included in soil signature.
- LAI : considered medium-high, with values up to 7 possible for dense crop canopies. Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. One of end products \rightarrow finer sampling steps.
- ALA : since very little is known about ALA , an average value of 57 (being a spherical distribution) is assumed. Erectophile- and planophile distributions are less likely, so a Gaussian distribution with a σ of 20 assumed.
- HOT : leaf size assumed intermediate compared to canopy height \rightarrow gaussian distr. with small mean.
- Bs : Influence of soil reflectance is assumed little: Bs is reduced to 1 class with gaussian variation around average. Because of high shading component and soil moist, average Bs is set to 0.7 and σ to 0.2
-

B.3 SPECL class 4: Bright vegetation

- *Principle land cover characteristics*: green crops and grasses in full development (high NIR reflectance).
- N : fresh leaves, therefore average of 1.63 and sig. of 0.5 (Table 3.1. Relatively little influence on spectrum, therefore only 3 classes/steps.

	Variable	Distribution	Minimum	Maximum	Mean	σ	# intervals
Leaf	N	Gaussian	1.	2.5	1.63	0.5	3
	Cab	Combal	20	100	(50)	(50)	8
	Cw	Uniform	0.01	0.08	(0.015)	(0.005)	5
	Cdm	Uniform	0.0050	0.0250	(0.01)	(0.003)	5
	Cbp	Fixed	0.	-	-	-	1
Canopy	LAI	Combal	2	9	(4)	(4.)	8
	ALA	Gaussian	30	70	57	20	3
	HOT	Gaussian	0.001	0.3	0.2	0.2	3
Soil	BS	Gaussian	0.3	1.1	0.7	0.3	1
Total # samples							43,200

- Cab : considered medium-high. Large influence on VIS. Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. One of end products \rightarrow finer sampling steps.
- Cw : Little known about distribution \rightarrow uniform distribution. Around 4 times higher than Cdm (In healthy vegetation relative water content is around 80%)
- Cdm : Little known about distribution \rightarrow uniform distribution. Around 0.25 times Cw (In healthy vegetation relative water content is around 80%)
- Cbp : Vegetation canopy itself is considered free of brown leaves, so Cbp is set to fixed value of 0. Incidental brown leaves are assumed to be included in soil signature.
- LAI : considered medium-high, with values up to 7 possible for dense crop canopies. Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. One of end products \rightarrow finer sampling steps.
- ALA : since very little is known about ALA , an average value of 57 (being a spherical distribution) is assumed. Erectophile- and planophile distributions are less likely, so a gaussian distribution with a σ of 20 assumed.
- HOT : leaf size assumed small compared to canopy height \rightarrow gaussian distr. with small mean.
- Bs : Influence of soil reflectance is assumed little: Bs is reduced to 1 class with gaussian variation around average. Because of high shading component and soil moist, average Bs is set to 0.7 and σ to 0.2

B.4 SPECL class 5: Yellow vegetation

- Principle land cover characteristics: healthy green vegetation (high NIR reflectance) containing flourishing yellow flowers (low green and red absorption), such as rapeseed or

	Variable	Distribution	Minimum	Maximum	Mean	σ	# intervals
Leaf	N	Gaussian	1.	2.5	1.63	1.	3
	Cab	Combal	20	100	(50)	(50)	8
	Cw	Uniform	0.01	0.08	(0.015)	(0.005)	5
	Cdm	Uniform	0.005	0.025	(0.01)	(0.003)	5
	Cbp	Fixed	0.	-	-	-	1
Canopy	LAI	Combal	1.5	7	(4)	(4.)	8
	ALA	Gaussian	30	70	57	20	3
	HOT	Gaussian	0.001	0.3	0.05	0.2	3
Soil	BS	Gaussian	0.3	1.1	0.7	0.3	1
Total # samples							43,200

sunflower.

- N : fresh leaves, therefore average of 1.63 and sig. of 0.5 (Table 3.1. Relatively little influence on spectrum, therefore only 3 classes/steps.
- Cab : considered medium-high. Large influence on VIS. Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. One of end products \rightarrow finer sampling steps.
- Cw : Little known about distribution \rightarrow uniform distribution. Around 4 times higher than Cdm (In healthy vegetation relative water content is around 80%)
- Cdm : Little known about distribution \rightarrow uniform distribution. Around 0.25 times Cw (In healthy vegetation relative water content is around 80%)
- Cbp : Vegetation canopy itself is considered free of brown leaves, so Cbp is set to fixed value of 0. Incidental brown leaves are assumed to be included in soil signature.
- LAI : considered medium-high, with values up to 7 possible for dense crop canopies. Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. One of end products \rightarrow finer sampling steps.
- ALA : since very little is known about ALA , an average value of 57 (being a spherical distribution) is assumed. Erectophile- and planophile distributions are less likely, so a gaussian distribution with a σ of 20 assumed.
- HOT : leaf size assumed small compared to canopy height \rightarrow gaussian distr. with small mean.
- Bs : Influence of soil reflectance is assumed little: Bs is reduced to 1 class with gaussian variation around average. Because of high shading component and soil moist, average Bs is set to 0.7 and σ to 0.2

B.5 SPECL class 6: Mixed vegetation / soil

	Variable	Distribution	Minimum	Maximum	Mean	σ	# intervals
Leaf	<i>N</i>	Gaussian	1	3.5	1.7	1.	3
	<i>Cab</i>	Combaf	10	80	(30)	(50)	7
	<i>Cw</i>	Uniform	0.007	0.05	(0.015)	(0.005)	5
	<i>Cdm</i>	Uniform	0.002	0.025	(0.01)	(0.003)	5
	<i>Cbp</i>	Gaussian	0.	0.5	0.	0.5	2
Canopy	<i>LAI</i>	Combaf	0.2	3.	(1.5)	(4.)	5
	<i>ALA</i>	Gaussian	30	80	57	20	3
	<i>HOT</i>	Gaussian	0.01	0.3	0.2	0.3	3
Soil	<i>BS</i>	Gaussian	0.5	1.2	0.9	0.2	3
Total # samples							141,750

- *Principle land cover characteristics*: green vegetation does not fully cover the underlying soil background, so primarily crops in early development stages, but also cut meadows or harvested green crops.
- *N*: Class includes both young and cut/harvested green vegetation → N maximum is extended to 3.5. Gaussian distribution with average of 1.7 and sigma of 1. Relatively little influence on spectrum, therefore only 3 classes/steps.
- *Cab*: is considered low-medium. Distribution according to (Combaf et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. Range slightly reduced because of reduced range.
- *Cw*: Water content ranges from low to high. Little known about distribution → uniform distribution. LB is lowered as also dry vegetation might be included. Average water content of $\pm 80\%$ is no longer valid for this class.
- *Cdm*: Little known about distribution → uniform distribution.
- *Cbp*: Small fraction of senescent leaves might be available → gaussian distr. with mean=0 and $\sigma = 0.5$.
- *LAI*: considered low-medium, with values up to 3.5. Number of classes is reduced because of reduced range. Distribution according to (Combaf et al., 2002b) in order to allow denser sampling in ranges of high sensitivity.
- *ALA*: since very little is known about ALA, an average value of 57 (bspherical distribution) is assumed. Erectophile- and planophile distributions are less likely, so a gaussian distribution with a σ of 20 is assumed. Max is increased a little since cut meadows will have an erectophile distribution.

- *HOT*: canopy height is assumed low compared to leaf width → Gaussian distr. with mean 0.2.
- *Bs*: Influence of soil reflectance is assumed medium-large → Number of classes is increased to 3. Because of larger fraction of unshaded bare soil, and reduced soil moist, average Bs is increased to 0.9 with a σ to 0.2

B.6 SPECL class 12: Dry vegetation / soil

	Variable	Distribution	Minimum	Maximum	Mean	σ	# intervals
Leaf	<i>N</i>	Gaussian	1.5	4.0	2.2	1.	3
	<i>Cab</i>	Combal	0	20	(10)	(50)	3
	<i>Cw</i>	Uniform	0.001	0.01	(0.003)	(0.004)	5
	<i>Cdm</i>	Uniform	0.002	0.015	(0.01)	(0.003)	5
	<i>Cbp</i>	Gaussian	0.	1.5	0	0.6	3
Canopy	<i>LAI</i>	Combal	0	1.5	(0.5)	(4.)	5
	<i>ALA</i>	Gaussian	30	70	57	20	3
	<i>HOT</i>	Gaussian	0.01	0.8	0.2	0.2	1
Soil	<i>BS</i>	Gaussian	0.7	1.3	1.0	0.2	3
Total # samples							30,375

- *Principle land cover characteristics*: vegetation with a high content of senescent material, such as cut meadows or mature cereals prior to harvest. Low cover.
- *N*: Class mainly includes senescent/harvested vegetation → increased N. min is increased to 1.5, maximum is extended to 4. A gauss. distr. with average of 2 and σ of 1 is used. Relatively little influence on spectrum, therefore only 3 classes/steps.
- *Cab*: is considered low. Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. Reduced number of intervals because of reduced range.
- *Cw*: Little known about distribution → uniform dist. LB is lowered as also dry vegetation might be included.
- *Cdm*: Little known about distribution → uniform
- *Cbp*: Fraction of senescent leaves is likely → gaussian distr. with mean=0 and $\sigma = 0.5$. 3 classes
- *LAI*: considered low, with values up to 1.5 Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. Nr. of classes is kept at 5 since spectrum is very sensitive LAI in this range.

- *ALA*: since very little is known about ALA, an average value of 57 (being a spherical distribution) is assumed. Erectophile- and planophile distributions are less likely, so a gaussian distribution with a σ of 20 is assumed.
- *HOT*: canopy height is assumed low compared to leaf width \rightarrow gaussian distr. with mean 0.2. Since canopy is relatively low and sparse, Hot spot has little effect \rightarrow Nr. classes is reduced to 1.
- *Bs*: Soil reflectance is dominant \rightarrow Number of classes is increased to 3. Because of larger fraction of unshaded bare soil, and reduced soil moist, average Bs is increased to 1.0. As soil conditions are very close to those found under bare soil conditions, minimum, maximum, and standard deviation are the ones that were deducted from scene variation 3.2.3

B.7 SPECL class 13: Sparse vegetation / soil

	Variable	Distribution	Minimum	Maximum	Mean	σ	# intervals
Leaf	<i>N</i>	Combal	1.	4.0	1.7	1.	3
	<i>Cab</i>	Uniform	0	40.	(20)	(50)	5
	<i>Cw</i>	Uniform	0.005	0.03	(0.015)	(0.005)	5
	<i>Cdm</i>	Uniform	0.002	0.020	(0.01)	(0.003)	5
	<i>Cbp</i>	Gaussian	0.	0.5	0.	0.5	2
Canopy	<i>LAI</i>	Combal	0.01	1.5	(0.5)	(1.)	5
	<i>ALA</i>	Gaussian	30	70	57	20	3
	<i>HOT</i>	Gaussian	0.01	0.8	0.2	0.2	1
Soil	<i>BS</i>	Gaussian	0.7	1.3	1.0	0.2	3
Total # samples							33,750

- *Principle land cover characteristics*: soil signature is dominant but still containing a fraction of photosynthetic active vegetation or green crop residues.
- *N*: Class includes very young and harvested green vegetation \rightarrow Relatively little influence on spectrum, therefore only 3 classes/steps.
- *Cab*: is considered low-medium. Distribution according to (Combal et al., 2002b) in order to allow denser sampling in ranges of high sensitivity. Finer steps are chosen because of large influence on lower concentrations.
- *Cw*: Little known about distribution \rightarrow Uniform dist. LB is lowered as also dry vegetation might be included.
- *Cdm*: Little known about distribution \rightarrow uniform
- *Cbp*: Small fraction of senescent leaves might be available \rightarrow gaussian distr. with mean=0 and $\sigma = 0.5$.

- *LAI*: considered low-medium, with values up to 1.5. Uniform distribution.
- *ALA*: since very little is known about ALA, an average value of 57 (being a spherical distribution) is assumed. Erectophile- and planophile distributions are less likely, so a gaussian distribution with a σ of 20 is assumed.
- *HOT*: canopy height is assumed low compared to leaf width \rightarrow gaussian distr. with mean 0.2. Hot spot has little effect at low cover \rightarrow Nr. classes is reduced to 1.
- *Bs*: Influence of soil reflectance is assumed large \rightarrow Number of classes is increased to 3. Because of larger fraction of unshaded bare soil, and reduced soil moist, average Bs is increased to 1.0. As soil conditions are very close to those found under bare soil conditions, minimum, maximum, and standard deviation are the ones that were deducted from scene variation

Appendix C

Overview of used vegetation indices

<i>Vegetation Index</i>	<i>Equation</i>	<i>Reference</i>
<i>Broadband vegetation indices</i>		
Normalized Difference Vegetation Index	$NDVI = \frac{\rho_{850} - \rho_{670}}{\rho_{850} + \rho_{670}}$	(Rouse et al., 1973)
Ratio Vegetation Index	$RVI = \frac{\rho_{850}}{\rho_{670}}$	(Pearson and Miller, 1972)
Soil Adjusted Vegetation Index	$SAVI = \frac{(\rho_{850} - \rho_{670}) * (1 + L)}{(\rho_{850} + \rho_{670} + L)}$	(Huete, 1988)
Soil Adjusted Vegetation Index 2	$SAVI_2 = \frac{\rho_{850}}{\rho_{670} + a^b / b^c}$	(Major et al., 1990)
Modified SAVI	$MSAVI = \frac{\rho_{850}}{\sqrt{(\rho_{850} + 0.5)^2 - 2(\rho_{850} - \rho_{670})}} + 0.5$	(Qi et al., 1994)
Optimized SAVI	$OSAVI = (1 + 0.16) * \frac{\rho_{850} - \rho_{670}}{\rho_{850} + \rho_{670} + 0.16}$	(Rondeaux et al., 1996)
Transformed SAVI	$TSAVI = a * \frac{\rho_{850} - a\rho_{670} - b}{a\rho_{850} + \rho_{670} - ab}$	(Baret et al., 1989)
Adjusted Transformed SAVI	$ATSAVI = a * \frac{\rho_{850} - a\rho_{670} - b}{a\rho_{850} + \rho_{670} - ab + X(1 + a^2)}$	(Baret and Guyot, 1991)
Soil and Atmospherically Resistant Index	$SARVI = \frac{(\rho_{850} - \rho_{RB})(1 + L)}{\rho_{850} + \rho_{RB} + L}$ $\rho_{RB} = \rho_{670} - \beta(\rho_{480} - \rho_{670})$	(Kaufman and Tanré, 1992)
Enhanced Vegetation Index or SARVI2	$EVI = \frac{2.5(\rho_{850} - \rho_{670})}{1 + \rho_{850} + 6\rho_{670} - 7.5/\rho_{480}}$	(Huete et al., 1997; Huete et al., 2002)
Ratio Difference Vegetation Index	$RDVI = \frac{\rho_{850} - \rho_{670}}{\sqrt{\rho_{850} + \rho_{670}}} = \sqrt{NDVI * RVI}$	(Roujean and Bréon, 1995)
Triangular Vegetation Index	$TVI = 60(\rho_{750} - \rho_{550}) - 100(\rho_{670} - \rho_{550})$	(Broge and Mortensen, 2002)

(continued on next page)

(continued)

<i>Vegetation Index</i>	<i>Equation</i>	<i>Reference</i>
Modified Triangular Vegetation Index 1	$MTVI1 = 1.2 * [1.2(\rho_{800} - \rho_{550}) - 2.5(\rho_{670} - \rho_{550})]$	(Haboudane et al., 2004)
Modified Triangular Vegetation Index 2	$MTVI2 = \frac{1.5 * [1.2(\rho_{800} - \rho_{550}) - 2.5(\rho_{670} - \rho_{550})]}{\sqrt{(2\rho_{800} + 1)^2 - (6\rho_{800} - 5\sqrt{\rho_{670}}) - 0.5}}$	(Haboudane et al., 2004)
Chlorophyll Absorption Ratio Index	$CARI = \frac{\frac{\rho_{700}}{\rho_{670}} * a\rho_{670} + \rho_{670} + b}{a^2 + 1}$ $a = \frac{\rho_{700} - \rho_{550}}{150}$ $b = \rho_{550} - 550a$	(Kim et al., 1994)

Chlorophyll indices

Transformed CARI	$TCARI = 3 [(\rho_{700} - \rho_{670}) - 0.2(\rho_{700} - \rho_{550})] * (\rho_{700} / \rho_{670})$	(Haboudane et al., 2002)
TCARI/OSAVI		(Haboudane et al., 2002)
Modified CARI	$MCARI = [(\rho_{700} - \rho_{670}) - 0.2(\rho_{700} - \rho_{550})](\rho_{700} / \rho_{670})$	(Daughtry et al., 2000)
Modified CARI 1	$MCARI1 = 1.2[2.5(\rho_{800} - \rho_{670}) - 2.5(\rho_{800} - \rho_{550})]$	(Haboudane et al., 2004)
Modified CARI 2	$MCARI2 = \frac{1.5[2.5(\rho_{800} - \rho_{670}) - 1.3(\rho_{800} - \rho_{550})]}{\sqrt{(2\rho_{800} + 1)^2 - (6\rho_{800} - 5\sqrt{\rho_{670}}) - 0.5}}$	(Haboudane et al., 2004)
MERIS Terrestrial Chlorophyll Index	$MTCI = \frac{\rho_{754} - \rho_{709}}{\rho_{709} - \rho_{681}}$	(Dash and Curran, 2004)
Leaf Chlorophyll Index	$LCI = \frac{\rho_{850} - \rho_{710}}{\rho_{850} + \rho_{710}}$	(?)
Simple ratio at 705 nm	$SR705 = \frac{\rho_{750}}{\rho_{705}}$	(Sims and Gamon, 2002)
Greenness Index	$GI = \frac{\rho_{671}}{\rho_{549}}$	(Zarco-Tejada et al., 2005b)
Photochemical Reflectance Index	$PRI = \frac{\rho_{529} - \rho_{569}}{\rho_{529} + \rho_{569}}$	(Penuelas et al., 1994)
Carter Stress Index 2	$CSI2 = \frac{\rho_{695}}{\rho_{760}}$	(Carter, 1994; Carter et al., 1994)
1st Order Derivative Green vegetation Index	$DGVI1 = \sum_{\lambda_{680}}^{\lambda_{760}} \rho'(\lambda_i) \Delta\lambda_i$	(Elvidge and Chen, 1995)
2nd Order Derivative Green vegetation Index	$DGVI1 = \sum_{\lambda_{680}}^{\lambda_{760}} \rho''(\lambda_i) \Delta\lambda_i$	(Elvidge and Chen, 1995)
Red Edge Inflection Point (index)	$REIP_{guyot}(\lambda) = 700 + \frac{740}{700} * \frac{\rho_i - \rho_{700}}{\rho_{740} - \rho_{700}}$ $\rho_i = 0.5 \frac{\rho_{780}}{\rho_{670}}$	(Guyot et al., 1988)

(continued on next page)

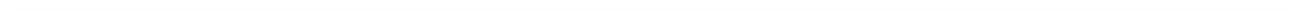
(continued)

<i>Vegetation Index</i>	<i>Equation</i>	<i>Reference</i>
REIP (maximum of 1 st order Savitzky-Golay filter)	$REIP_{sav1}(\lambda) = \max(\frac{\partial \rho}{\partial \lambda})$	(Savitzky and Golay, 1964)
REIP (minimum of 2 nd order Savitzky-Golay filter)	$REIP_{sav2}(\lambda) = \min(\frac{\partial^2 \rho}{\partial^2 \lambda})$	(Savitzky and Golay, 1964)
REIP (Lagrangian model)	$REIP_{Lagr}(\lambda) = \frac{A(\lambda_i + \lambda_{i+1}) + B(\lambda_{i-1} + \lambda_{i+1}) + C(\lambda_{i-1} + \lambda_i)}{2(A+B+C)}$ $A = \frac{D_{\lambda(i-1)}}{(\lambda_{i-1} - \lambda_i)(\lambda_{i-1} - \lambda_{i+1})}$ $B = \frac{D_{\lambda(i)}}{(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})}$ $C = \frac{D_{\lambda(i+1)}}{(\lambda_{i+1} - \lambda_{i-1})(\lambda_{i+1} - \lambda_i)}$	(Dawson and Curran, 1998)
<i>Water indices</i>		
Moisture Stress Index	$MSI = \frac{\rho_{1600}}{\rho_{820}}$	(Hunt and Rock, 1989)
Leaf Water Vegetation Index 1	$LWVI1 = \frac{\rho_{1094} - \rho_{983}}{\rho_{1094} + \rho_{983}}$	(Galvao et al., 2005)
Leaf Water Vegetation Index 2	$LWVI2 = \frac{\rho_{1094} - \rho_{1205}}{\rho_{1094} + \rho_{1205}}$	(Galvao et al., 2005)
Disease Water Stress Index 5	$DWSI5 = \frac{\rho_{800} + \rho_{550}}{\rho_{1660} + \rho_{680}}$	(Apan et al., 2004)
<i>Dry matter indices</i>		
Normalized Difference Nitrogen Index	$NDNI = \frac{LOG(\rho_{1510})^{-1} - LOG(\rho_{1680})^{-1}}{LOG(\rho_{1510})^{-1} + LOG(\rho_{1680})^{-1}}$	(Serrano et al., 2002)
Normalized Difference Lignin Index	$NDLI = \frac{LOG(\rho_{1754})^{-1} - LOG(\rho_{1680})^{-1}}{LOG(\rho_{1754})^{-1} + LOG(\rho_{1680})^{-1}}$	(Serrano et al., 2002)
Cellulose Absorption Index	$CAI = 0.5 * (\rho_{2015} + \rho_{2195}) - \rho_{2106}$	(Nagler et al., 2000)
Shortwave Infrared Green Vegetation Index	$SWIRVI = 37.27 * (\rho_{2210} - \rho_{2090}) + 26.27(\rho_{2280} - \rho_{2090}) - 0.57$	(Lobell et al., 2001)

Table C.1: Vegetation indices used for predictive equations. L = 0.5 (Baret and Guyot, 1991), a = 1.2 (Baret and Guyot, 1991; Broge and Mortensen, 2002), b = 0.04 (Baret and Guyot, 1991), X=0.08 (Baret and Guyot, 1991), $\beta = 1$ (Kaufman and Tanré, 1992)

Appendix D

Spectral configuration HyEurope 2003



<i>VIS</i>			<i>NIR</i>			<i>SWIR1</i>			<i>SWIR2</i>		
<i>Band</i>	λ	<i>FWHM</i>	<i>Band</i>	λ	<i>FWHM</i>	<i>Band</i>	λ	<i>FWHM</i>	<i>Band</i>	λ	<i>FWHM</i>
	[nm]	[nm]		[nm]	[nm]		[nm]	[nm]		[nm]	[nm]
1	438	11	31	878	16	63	1404	16	95	1951	22
2	450	12	32	895	16	64	1419	16	96	1970	22
3	462	16	33	911	16	65	1433	15	97	1990	21
4	478	16	34	927	16	66	1447	15	98	2009	21
5	493	17	35	943	16	67	1461	15	99	2027	21
6	508	16	36	959	16	68	1475	15	100	2045	21
7	524	16	37	975	16	69	1489	14	101	2064	21
8	539	17	38	990	17	70	1503	14	102	2082	21
9	555	16	39	1006	15	71	1516	15	103	2100	21
10	570	15	40	1022	15	72	1530	14	104	2118	21
11	585	15	41	1037	16	73	1543	14	105	2136	20
12	600	16	42	1053	16	74	1557	14	106	2154	20
13	616	16	43	1068	15	75	1570	14	107	2171	20
14	632	15	44	1083	15	76	1583	14	108	2188	20
15	646	15	45	1098	16	77	1596	14	109	2206	22
16	662	16	46	1113	15	78	1609	14	110	2224	20
17	677	16	47	1128	15	79	1622	14	111	2241	20
18	692	16	48	1142	15	80	1635	14	112	2258	20
19	707	16	49	1157	15	81	1648	14	113	2274	19
20	723	16	50	1172	15	82	1661	14	114	2292	20
21	738	16	51	1186	15	83	1673	14	115	2309	19
22	753	16	52	1200	15	84	1686	13	116	2325	20
23	768	16	53	1215	15	85	1698	13	117	2341	19
24	783	17	54	1229	15	86	1710	13	118	2357	19
25	798	17	55	1243	15	87	1723	13	119	2373	18
26	813	17	56	1257	15	88	1735	13	120	2389	19
27	828	17	57	1272	15	89	1747	13	121	2405	19
28	844	18	58	1286	15	90	1759	13	122	2421	19
29	859	17	59	1299	15	91	1771	13	123	2437	18
30	874	17	60	1313	15	92	1783	13	124	2453	18
			61	1327	15	93	1795	12	125	2468	18
			62	1340	14	94	1807	12	126	2483	18

Table D.1: Spectral configuration of HyEurope 2003, arranged according to the 4 detectors (VIS, NIR, SWIR1, and SWIR2). λ indicates the central wavelength, FWHM for the Full Width Half Maximum.

Appendix E

Canopy variables measured at Waging test site

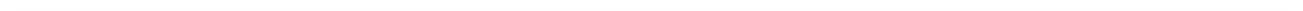


Table E.1: Results of canopy measurements for grassland measured at each plot during HYMI campaign

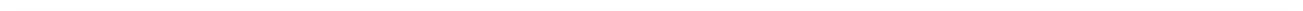
Field	plot	LAI		C _{dm} (g · cm ⁻²)		C _w (g · cm ⁻²)		Rel. water content (%)	Canopy height (cm)
		LAI-2000	emp.	LAI-2000	emp.	LAI-2000	emp.		
MEA1	1	6.85 (0.14)	6.83 (2.59)	0.0061 (0.0001)	0.0061 (0.0023)	0.0227 (0.0005)	0.0227 (0.0086)	0.79	53.2 (10.1)
	2	5.49 (0.35)	3.93 (1.49)	0.0054 (0.0003)	0.0075 (0.0028)	0.0189 (0.0012)	0.0264 (0.0100)	0.78	31.6 (7.6)
	3	4.59 (0.32)	3.47 (1.32)	0.0059 (0.0004)	0.0078 (0.0030)	0.0197 (0.0014)	0.0260 (0.0099)	0.77	25.2 (8.4)
	4	4.49 (0.20)	3.63 (1.38)	0.0063 (0.0003)	0.0078 (0.0030)	0.0211 (0.0009)	0.0260 (0.0099)	0.77	22.8 (5.4)
	5	3.24 (0.31)	2.84 (1.08)	0.0073 (0.0007)	0.0083 (0.0031)	0.0223 (0.0021)	0.0255 (0.0097)	0.75	22.9 (5.7)
	6	4.26 (0.24)	3.08 (1.17)	0.0074 (0.0004)	0.0102 (0.0039)	0.0170 (0.0010)	0.0236 (0.0090)	0.70	26.9 (5.8)
	7	6.35 (0.15)	5.81 (2.21)	0.0070 (0.0002)	0.0076 (0.0029)	0.0241 (0.0006)	0.0263 (0.0100)	0.78	54.1 (7.7)
MEA2	1	0.80 (0.12)	2.07 (0.79)	0.0168 (0.0025)	0.0065 (0.0025)	0.0577 (0.0087)	0.0223 (0.0085)	0.77	8.3 (2.3)
	2	0.82 (0.13)	2.11 (0.80)	0.0143 (0.0023)	0.0056 (0.0021)	0.0597 (0.0095)	0.0232 (0.0088)	0.81	7.9 (3.0)
	3								6.5 (2.3)
PAS	4	0.51 (0.08)	1.62 (0.62)	0.0150 (0.0024)	0.0047 (0.0018)	0.0764 (0.0120)	0.0240 (0.0091)	0.84	8.3 (4.2)
	5	0.91 (0.25)	1.86 (0.71)	0.0105 (0.0029)	0.0051 (0.0019)	0.0484 (0.0133)	0.0236 (0.0090)	0.82	10.5 (4.9)
	1	6.06 (0.37)	5.89 (2.24)	0.0081 (0.0005)	0.0083 (0.0032)	0.0249 (0.0015)	0.0256 (0.0097)	0.75	18.6 (9.5)
	2	4.96 (0.34)	6.81 (2.59)	0.0088 (0.0006)	0.0064 (0.0024)	0.0378 (0.0026)	0.0275 (0.0105)	0.81	14.6 (4.8)
	3	5.12 (0.33)	7.26 (2.76)	0.0076 (0.0005)	0.0054 (0.0020)	0.0405 (0.0026)	0.0286 (0.0109)	0.84	13.6 (4.3)
4	3.01 (0.48)	5.31 (2.02)	0.0128 (0.0020)	0.0073 (0.0028)	0.0469 (0.0075)	0.0266 (0.0101)	0.79	7.7 (3.7)	
5	1.33 (0.20)	3.25 (1.24)	0.0140 (0.0021)	0.0057 (0.0022)	0.0688 (0.0103)	0.0281 (0.0107)	0.83	8.8 (3.6)	

Table E.2: Results of canopy measurements for grassland measured at each plot during HYM1 campaign

Field	plot	LAI		C_{dm} ($g \cdot cm^{-2}$)		C_w ($g \cdot cm^{-2}$)		Rel. water content (%)	Canopy height (cm)
		LAI-2000	emp.	LAI-2000	emp.	LAI-2000	emp.		
MEA1	1	2.37 (0.16)	1.55 (0.59)	0.0060 (0.0004)	0.0092 (0.0035)	0.0151 (0.0010)	0.0231 (0.0088)	0.71	16.4 (4.1)
	2	2.55 (0.15)	1.79 (0.68)	0.0061 (0.0004)	0.0087 (0.0033)	0.0167 (0.0010)	0.0237 (0.0090)	0.73	13.5 (5.8)
	3	2.41 (0.21)	1.74 (0.66)	0.0051 (0.0004)	0.0070 (0.0027)	0.0183 (0.0016)	0.0253 (0.0096)	0.78	19.8 (5.0)
	4	2.94 (0.22)	0.87 (0.33)	0.0033 (0.0002)	0.0113 (0.0043)	0.0062 (0.0005)	0.0209 (0.0079)	0.65	18.4 (5.4)
	5	1.58 (0.14)	0.77 (0.29)	0.0065 (0.0006)	0.0133 (0.0051)	0.0091 (0.0008)	0.0188 (0.0071)	0.59	12.2 (3.3)
	6	2.34 (0.15)	1.28 (0.49)	0.0048 (0.0003)	0.0088 (0.0034)	0.0128 (0.0008)	0.0235 (0.0089)	0.73	12.6 (4.3)
MEA2	1	0.60 (0.09)	0.57 (0.22)	0.0107 (0.0016)	0.0114 (0.0043)	0.0194 (0.0029)	0.0206 (0.0078)	0.64	6.5 (2.7)
	2	0.70 (0.09)	1.38 (0.52)	0.0206 (0.0026)	0.0104 (0.0040)	0.0430 (0.0035)	0.0219 (0.0083)	0.68	6.8 (2.7)
	3	0.94 (0.10)	1.60 (0.61)	0.0188 (0.0020)	0.0110 (0.0042)	0.0363 (0.0039)	0.0213 (0.0081)	0.66	4.9 (2.4)
	4	0.58 (0.08)	1.13 (0.43)	0.0262 (0.0036)	0.0135 (0.0051)	0.0366 (0.0050)	0.0188 (0.0071)	0.58	6.4 (3.1)
	5	2.21 (0.18)	1.31 (0.50)	0.0067 (0.0005)	0.0113 (0.0043)	0.0124 (0.0010)	0.0209 (0.0080)	0.65	8.0 (2.5)
PAS	1	7.11 (0.20)	3.91 (1.49)	0.0034 (0.0001)	0.0061 (0.0023)	0.0145 (0.0004)	0.0264 (0.0100)	0.81	36.2 (8.9)
	2	4.12 (0.32)	4.23 (1.61)	0.0091 (0.0007)	0.0089 (0.0034)	0.0242 (0.0019)	0.0236 (0.0090)	0.73	14.6 (4.8)
	3	7.48 (0.20)	2.53 (0.96)	0.0022 (0.0001)	0.0064 (0.0024)	0.0088 (0.0002)	0.0261 (0.0099)	0.80	15.6 (5.4)
	4	4.98 (0.47)	4.50 (1.71)	0.0073 (0.0007)	0.0081 (0.0031)	0.0220 (0.0021)	0.0244 (0.0093)	0.75	19.5 (10.6)
	5	0.81 (0.09)	1.21 (0.46)	0.0159 (0.0018)	0.0106 (0.0040)	0.0324 (0.0036)	0.0217 (0.0082)	0.67	5.4 (1.9)

Appendix F

Spectral configuration CHRIS Mode 5 in 2006



<i>Band nr.</i>	λ (nm)	<i>FWHM</i> (nm)
1	442.5	10.5
2	490.4	11.6
3	530.2	11.5
4	551.6	12.9
5	570.2	10.7
6	631.8	14.1
7	661.5	15.8
8	675.0	11.0
9	686.2	11.5
10	697.9	11.8
11	706.9	6.1
12	713.1	6.3
13	719.3	6.3
14	725.7	6.5
15	732.3	6.6
16	739.0	6.7
17	745.7	6.8
18	752.6	6.9
19	759.7	7.1
20	766.8	7.2
21	774.1	7.4
22	785.4	15.2
23	796.9	7.8
24	804.7	7.9
25	868.3	18.1
26	886.9	18.8
27	901.1	9.6
28	910.8	9.7
29	920.5	9.9
30	930.4	10.1
31	945.6	20.4
32	961.1	10.3
33	971.5	10.6
34	982.1	10.5
35	992.9	10.8
36	1003.6	10.9
37	1025.4	32.9

Table F.1: Spectral configuration of CHRIS Mode 5 in 2006. λ stands for the center wavelength, FWHM for the Full Width Half Maximum.

Bibliography

- Adams, J. and Smith, M. (1997). Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site., *Journal of Geophysical Research* **91**: 8098–8112.
- Apan, A., Held, A., Phinn, S. and Markley, J. (2004). Detecting sugarcane 'orange rust' disease using EO-1 Hyperion hyperspectral imagery, *International Journal of Remote Sensing* **23**(19): 3939–3958.
- Asner, G. and Lobell, D. (2004). Spectral unmixing of vegetation, soil and dry carbon cover in arid regions: comparing multispectral and hyperspectral Observations, *International Journal of Remote Sensing* **25**: 489–498.
- Atzberger, C. (2004). Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models, *Remote Sensing of Environment* **93**: 53–67.
- Atzberger, C., Jarmer, T., Schlerf, M., Koetz, B. and Werner, W. (2003). Retrieval of wheat biophysical attributes from hyperspectral data and sailh+prospect radiative transfer model, *Proceedings of the 3rd EARSeL Workshop on Imaging Spectroscopy*, EARSeL, Herrsching, Germany.
- Bach, H. and Mauser, W. (2003). Methods and examples for remote sensing data assimilation in land surface process modeling, *IEEE Transactions on Geoscience and Remote Sensing* **41**: 1629–1637.
- Bachmann, M. (2007). *Automatisierte Ableitung von Bodenbedeckungsgraden durch MESMA-Entmischung*, PhD thesis, Bayerischer Julius-Maximilians-Universität, Würzburg, Germany.
- Bachmann, M., Holzwarth, S., Habermeyer, M., Richter, R. and Müller, A. (2007). Including quality measures in an automated processing chain for airborne hyperspectral data, *Proceedings of the 5th EARSeL Workshop on Imaging Spectroscopy*, EARSeL, Bruges, Belgium.
- Bachmann, M., Müller, A., Habermeyer, M. and Dech, S. (2004). An iterative unmixing approach in support of fractional cover estimation in semi-arid environments., in M. Ehlers, F. Posa, H. Kaufmann, U. Michel and G. De Carolis (eds), *Proceedings SPIE - 11th International Symposium on Remote Sensing, Maspalomas, Gran Canaria, 13-16.9.2004*, pp. 205–213.
-

- Bacour, C. (2001). *Contribution à la détermination des paramètres biophysiques des couverts végétaux par inversion de modèles de réflectance: analyse de sensibilité et configurations optimales*, PhD thesis, Université Paris 7 - Denis Diderot, Paris, France. 206 pp.
- Bacour, C., Baret, F., Béal, D., Weiss, M. and Pavageau, K. (2006). Neural network estimation of LAI, fAPAR, fCover and LAIxCab, from top of canopy MERIS reflectance data: principles and validation, *Remote Sensing of Environment* **105**: 313–325.
- Bacour, C., Baret, F. and Jacquemoud, S. (2002a). Information content of HyMap hyperspectral imagery, *Proceedings 1st International Symposium on Recent Advances in Quantitative Remote Sensing*, Valencia, Spain, pp. 503–508.
- Bacour, C. and Bréon, F.-M. (2005). Variability of biome reflectance directional signatures as seen by polder, *Remote Sensing of Environment* **98**: 80–95.
- Bacour, C., Jacquemoud, S., Leroy, M., Hauteœur, O., Weiss, M., Prévot, B. N. and Chauki, H. (2002). Reliability of the estimation of vegetation characteristics by inversion of three canopy reflectance models on polder data, *Agronomie* **22**: 555–565.
- Bacour, C., Jacquemoud, S., Tourbier, Y., Dechambre, M. and Frangi, J. P. (2002a). Design and analysis of numerical experiments to compare four canopy reflectance models, *Remote Sensing of Environment* **79**: 72–83.
- Baranoski, G. V. G. (1998). *Biologically and physically-based rendering of natural scenes*, PhD thesis, Department of Computer Science, Calgary, Alberta.
- Baranoski, G. V. G. and Rokne, J. G. (1997). An algorithmic reflectance and transmittance model for plant tissue, *EUROGRAPHICS'97*, Vol. 16, pp. 141–150.
- Baret, F. and Buis, S. (2007). Estimating canopy characteristics from remote sensing observations. review of methods and associated problems., in S. Liang (ed.), *Advances in Land Remote Sensing: System, Modeling, Inversion and Application*, Springer.
- Baret, F. and Fourty, T. (1997a). Estimation of leaf water content and specific leaf weight from reflectance and transmittance measurements, *Agronomie* **17**: 455–464.
- Baret, F. and Fourty, T. (1997b). Radiometric estimates of nitrogen status in crops, in G. Lemaire (ed.), *Diagnosis of the Nitrogen Status in Crops*, Springer Verlag, New York, pp. 201–227.
- Baret, F. and Guyot, G. (1991). Potentials and limits of vegetation indices for lai and apar assessment, *Remote Sensing of Environment* **35**: 161–173.
- Baret, F., Guyot, G. and Major, D. J. (1989). TSAVI: a vegetation index which minimizes soil brightness effects on lai and APAR estimation, *Quantitative remote sensing: An economic tool for the Nineties; Proceedings of IGARSS '89 and 12th Canadian Symposium on Remote Sensing*, New York, Institute of Electrical and Electronics Engineers, pp. 1355–1358.
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J. L. and Leroy, M. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm, *Remote Sensing of Environment* **110**: 275–286.

-
- Baret, F., Jacquemoud, S., Guyot, G. and Leprieur, C. (1992). Modeled analysis of the biophysical nature of spectral shifts and comparison with information content of broad bands, *Remote Sensing of Environment* **41**: 133–142.
- Baret, F., Koetz, B. and Bruguier, N. (2002). Wp 1400: Characterization of maize structure and optical properties. INRA-CSE, Avignon, France.
- Baret, F., Pavageau, K., Bacour, C., Béal, D., Weiss, M., Bruniquel, V. and Regner, P. (2005a). Algorithm theoretical basis document for MERIS top of atmosphere land products (toa_veg), *Technical report*, INRA & Noveltis.
- Baret, F., Pavageau, K., Bacour, C., Béal, D., Weiss, M., Bruniquel, V. and Regner, P. (2005b). Algorithm theoretical basis document for meris top of canopy land products (toc_veg), *Technical report*, INRA & Noveltis.
- Baret, F., Weiss, M., Allard, D., Garrigue, S., Leroy, M., Jeanjean, H., Fernandes, R., Myneni, R., Privette, J., Morisette, J., Bohbot, H., Bosseno, R., Dedieu, G., Di Bella, C., Duchemin, B., Espana, M., Gond, V., Gu, X. F., Guyon, D., Lelong, C., Maisongrande, P., Mougín, E., Nilson, T., Veroustraete, F. and Vintilla, R. (submitted). VALERI: a network of sites and a methodology for the validation of medium spatial resolution satellite products, *Remote Sensing of Environment*.
- Barnsley, M. J., Allison, D. and Lewis, P. (1997). On the information content of multiple view angle (MVA) images, *International Journal of Remote Sensing* **18**: 1937–1960.
- Barnsley, M. J. and Settle, J. (2004). The PROBA/CHRIS mission: A low-cost smallsat for hyperspectral multiangle observations of the earth surface and atmosphere, *IEEE Transactions on Geoscience and Remote Sensing* **42**: 1512–1520.
- Bastiaanssen, W. G. M. and Ali, S. (2003). A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan, *Agriculture, Ecosystems & Environment* **94**: 321–340.
- Baumgardner, M., Silva, L., Biehl, L. and Stoner, E. (1985). Reflectance properties of soils, *Advances in Agronomy* **38**: 1–44.
- Beisl, U. (2001). *Correction of bidirectional effects in imaging spectrometer data*, PhD thesis, Remote Sensing Laboratories, Department of Geography, University of Zurich, Zurich, Switzerland.
- Berk, A., Anderson, G. P., Acharya, P. K., Bernstein, L. S., Muratov, L., Lee, J., Fox, M., Adler-Golden, S., Chetwynd, J., Hoke, M., Lockwood, R., Gardner, J., Cooley, T., Borel, C. and Lewis, P. (2005). MODTRAN5: a reformulated atmospheric band model with auxiliary species and practical multiple scattering options, in S. S. Shen and P. E. Lewis (eds), *SPIE Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, Vol. 5806, pp. 662–667.
- Berk, A., Anderson, G. P., Acharya, P. K., Hoke, M. L., Chetwynd, J. H., Bernstein, L. S., Shettle, E. P., Matthew, M. W. and Adler-Golden, S. M. (2003). MODTRAN4 Version 3 Revision 1 User’s Manual, *Technical report*, Hanscom, MA.
-

- Borel, C. C., Gerstl, S. A. W. and Powers, B. J. (1991). The radiosity method in optical remote sensing of structured 3-D surfaces, *Remote Sensing of Environment* **36**: 13–44.
- Bouman, B. (1995). Crop modeling and remote-sensing for yield prediction, *Netherlands Journal of Agricultural Sciences* **43**: 143161.
- Bousquet, L., Lacherade, S., Jacquemoud, S. and Moya, I. (2005). Leaf BRDF measurements and model for specular and diffuse components differentiation, *Remote Sensing of Environment* **98**: 201–211.
- Brachet, G. (2004). From initial ideas to a European plan: GMES as an exemplar of European space strategy, *Space Policy* **20**: 7–15.
- Brazile, J., Kaiser, J., Schlaepfer, D., Nieke, J., Schaepman, M. and Itten, K. I. (2005). Parallelization of APEX airborne imaging spectrometer product generation, in B. Zagajewski and M. Sobczak (eds), *Proceedings of the 5th EARSeL Workshop on Imaging Spectroscopy*.
- Broge, N. H. and Mortensen, J. V. (2002). Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data, *Remote Sensing of Environment* **81**: 45–57.
- Brown de Colstoun, E. and Walthall, C. (2006). Improving global scale land cover classifications with multi-directional POLDER data and a decision tree classifier, *Remote Sensing of Environment* **100**: 474–485.
- Brown, J. (1998). Mapping global grassland ecosystems: a comparison of four data sets, *Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS '98.*, Vol. 100.
- Bunnik, N. (1978). *The multispectral reflectance of shortwave radiation in relation with their morphological and optical properties*, PhD thesis, Wageningen Agricultural University.
- Camacho-de Coca, F., Bréon, F. M., Leroy, M. and García-Haro, F. J. (2004). Airborne measurement of hot spot reflectance signatures, *Remote Sensing of Environment* **90**: 63–75.
- Camenen, L., Goulas, Y., Guyot, G., Cerovic, Z., Schmuck, G. and Moya, I. (1986). Estimation of the chlorophyll fluorescence lifetime of plant canopies : validation of a deconvolution method based on the use of a 3-D canopy mockup, *Remote Sensing of Environment* **58**: 157–168.
- Campbell, G. S. (1986). Extinction coefficients for radiation in plant canopies calculated using an ellipsoidal inclination angle distribution, *Agricultural and Forest Meteorology* **36**: 317–321.
- Campbell, G. S. (1990). Derivation of an angle density function for canopies with ellipsoidal leaf angle distributions, *Agricultural and Forest Meteorology* **49**: 173–176.
- Carter, G. A. (1994). Ratios of leaf reflectances in narrow wavebands as indicators of plant stress, *International Journal of Remote Sensing* **15**: 697–704.
- Carter, G. A., Dell, T. and W.G., C. (1994). Spectral characteristics and digital imagery of a pine needle blight in the southeastern united states, *Canadian Journal of Forest Research* **26**: 402–407.

- Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S. and Grégoire, J.-M. (2001). Detecting vegetation leaf water content using reflectance in the optical domain, *Remote Sensing of Environment* **77**: 22–33.
- Chappelle, E. W., Kim, M. S. and McMurtrey, I. J. E. (1992). Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll a, chlorophyll b, and carotenoids in soybean leaves, *Remote Sensing of Environment* **39**: 239–247.
- Chelle, M. and Andrieu, B. (1998). The nested radiosity model for the distribution of light within plant canopies, *Ecological Modelling* **111**: 75–91.
- Chen, D., Jackson, T., Li, F., Cosh, M., Walthall, C. and Anderson, M. (2003). Estimation of vegetation water content for corn and soybeans with a normalized difference water index (NDWI) using Landsat Thematic Mapper data, *Proceedings of 2003 IEEE International Geoscience and Remote Sensing Symposium*, Vol. 4, IGARSS, pp. 2853–2856.
- Chen, J. M. and Leblanc, S. G. (1997). A four-scale bidirectional reflectance model based on canopy architecture, *IEEE Transactions on Geoscience and Remote Sensing* **35**: 1316–1337.
- Chen, J. M., Pavlic, G., Brown, L., Cihlar, J., Leblanc, S. G., White, H. P. and et al. (2002). Derivation and validation of Canada wide coarse resolution leaf area index maps using high resolution satellite imagery and ground measurements, *Remote Sensing of Environment* **80**: 165–184.
- Chopping, M. J., Rango, A., Havstad, K. M., Schiebe, F. R., Ritchie, J. C., Schmutz, T. J., French, A. N., Su, L., McKee, L. and Davis, M. R. (2003). Canopy attributes of desert grassland and transition communities derived from multiangular airborne imagery, *Remote Sensing of Environment* **85**: 339–354.
- Cierniewski, J. and Verbrugge, M. (1997). Influence of soil surface roughness on soil bidirectional reflectance, *International Journal of Remote Sensing* **18**: 1277–1288.
- Clevers, J. G. P. W. (1989). Application of a weighted infrared-red vegetation index for estimating leaf area index by correcting for soil moisture, *Remote Sensing of Environment* **29**: 25–37.
- 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**: 221–228.
- Clevers, J. G. P. W., van der Heijden, G. and Schut, A. (2005). A method using different remote sensing techniques for estimating grassland bio-physical variables, *Proceedings of the 5th EARSeL Workshop on Imaging Spectroscopy*, EARSeL, Warsaw, Poland, pp. 835–849.
- Cocks, T., Jenssen, R., Stewart, A., Wilson, I. and Shields, T. (1998). The HyMap airborne hyperspectral sensor: the system, calibration and performance, *Proceedings of the 1st EARSeL Workshop on Imaging Spectroscopy*, EARSeL, pp. 37–42.
- Combal, B., Baret, F. and Weiss, M. (2002a). Improving canopy variables estimation from remote sensing data by exploiting ancillary information. Case study on sugar beet canopies, *Agronomie* **22**: 205–215.

- Combal, B., Baret, F., Weiss, M., Trubuil, A., Macé, D., Pragnère, A., Myneni, R., Knyazikhin, Y. and Wang, L. (2002b). 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.
- CROMA (2002). Crop reflectance operational models for agriculture. final report. URL: <http://www.astrium-space.com/croma>.
- Cutter, M. (2006). HDFclean V2. URL: <http://earth.esa.int/object/index.cfm?fobjectid=4409>, visited: March 6, 2007.
- Danson, F. M., Steven, M. D., Malthus, T. J. and Clark, J. A. (1992). High spectral resolution data for determining leaf water content, *International Journal of Remote Sensing* **13**: 461–470.
- Dash, J. and Curran, P. J. (2004). The MERIS terrestrial chlorophyll index, *International Journal of Remote Sensing* **25**: 5403–5413.
- Daughtry, C. S. T., Walthall, C. L., Kim, M. S., de Colstoun, E. B. and McMurtrey III, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance, *Remote Sensing of Environment* **74**: 229–239.
- Dawson, T. P. and Curran, P. J. (1998). A new technique for interpolating the reflectance red edge position, *International Journal of Remote Sensing* **19**: 2133–2139.
- De Vries, C., Danaher, T., Denham, R., Scarth, P. and Phinn, S. (2007). An operational radiometric calibration procedure for the landsat sensors based on pseudo-invariant target sites, **107**: 414–429.
- De Wit, A. J. W. (1999). The application of a genetic algorithm for crop model steering using noaa-avhrr data, *Remote sensing for earth science, ocean, and sea ice applications; Proceedings of the Meeting*, Vol. 3868, Society of Photo-Optical Instrumentation Engineers (SPIE).
- De Wit, C. (1965). Photosynthesis of leaf canopies, *Agricultural Research Report 663*, PUDOC, Wageningen.
- Delécolle, R., Maas, S. J., Guérif, M. and Baret, F. (1992). Remote sensing and crop production models - present trends, *ISPRS Journal of Photogrammetry and Remote Sensing* **47**: 145–161.
- DIVERSITAS (2002). DIVERSITAS SCIENCE PLAN; DIVERSITAS Report No. 1. 40 pp., Paris.
- Djanibekov, N. (2007). *A micro-economic analysis of farm restructuring in Khorezm region, Uzbekistan*, PhD thesis, University of Bonn, Bonn, Germany.
- Doraiswamy, P. C., Hatfield, J. L., Jackson, T. J., Akhmedov, B., Prueger, J. and Stern, A. (2004). Crop condition and yield simulations using Landsat and MODIS, *Remote Sensing of Environment* **92**: 548–559.

-
- Dorigo, W. A., Zurita-Milla, R., De Wit, A. J. W., Brazile, J., Singh, R. and Schaepman, M. E. (2007). A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling, *International Journal of Applied Earth Observation and Geoinformation* **9**.
- Dorigo, W., Bachmann, M. and Heldens, W. (2006). AS Toolbox & processing of field spectra – user’s manual, *Technical report*, DLR-DFD, Imaging Spectroscopy Group. URL: http://cocoon.caf.dlr.de/astools_en.html.
- Dorigo, W., Baret, F., Richter, R., Ruecker, G., Schaepman, M. and Mueller, M. (2007). Retrieving canopy variables by radiative transfer model inversion – an automated regional approach for imaging spectrometer data, *Proceedings of the 5th EARSeL Workshop on Imaging Spectroscopy*, EARSeL, Bruges, Belgium.
- Elvidge, C. D. and Chen, Z. (1995). Comparison of broad-band and narrow-band red and near-infrared vegetation indices, *Remote Sensing of Environment* **54**: 38–48.
- 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.
- Fourty, T. and Baret, F. (1997). Vegetation water and dry matter contents estimated from top-of-the-atmosphere reflectance data: A simulation study, *Remote Sensing of Environment* **61**: 34–45.
- 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**: 104–117.
- Fridgen, J. and Varco, J. (2004). Dependency of cotton leaf nitrogen, chlorophyll, and reflectance on nitrogen and potassium availability, *Agronomy Journal* **96**: 63–69.
- Fukshansky, L., Fukshansky-Kazarinova, N. and Martinez v. Remisowsky, A. (1991). Estimation of optical parameters in a living tissue by solving the inverse problem to the multiflux radiative transfer, *Applied Optics* **30**: 3145–3153.
- Galvao, L., Formaggio, A. and Tisot, D. (2005). Discrimination of sugarcane varieties in South-eastern Brazil with EO-1 Hyperion data, *Remote Sensing of Environment* **94**: 523–534.
- Gamon, J., Field, C., Bilger, W., Björkman, O., Fredeen, A. and Penuelas, J. (1990). Remote sensing of the xanthophyll cycle and chlorophyll fluorescence in sunflower leaves and canopies, *Oecologia* **85**: 1–7.
- Gao, B.-C. (2002). NDWI a normalized difference water index for remote sensing of vegetation liquid water from space, *Remote Sensing of Environment* **58**: 257–266.
- Garcia-Haro, F. J., Sommer, S. and Kemper, T. (2005). A new tool for variable multiple end-member spectral mixture analysis (VMESMA), *International Journal of Remote Sensing* **26**: 2135–2162.
- Gardner, B. and Blad, B. (1986). Evaluation of spectral reflectance models to estimate corn leaf area while minimizing the influence of soil background effects, *Remote Sensing of Environment* **20**: 183–193.

- Garrigues, S., Allard, D., Baret, F. and Weiss, M. (2006a). Influence of landscape spatial heterogeneity on the non-linear estimation of leaf area index from moderate spatial resolution remote sensing data, *Remote Sensing of Environment* **105**: 286–298.
- Garrigues, S., Allard, D., Baret, F. and Weiss, M. (2006b). Quantifying spatial heterogeneity at the landscape scale using variogram models, *Remote Sensing of Environment* **103**: 81–96.
- Gascon, F. and Berger, M. (2007). GMES Sentinel-2 Mission Requirements Document, *Technical report*, ESA. issue 2 revision 0 (11/02/07), EOP.SM/1163/MR-dr.
- Gastellu-Etchegorry, J. P., Demarez, V., Pinel, V. and Zagolski, F. (1996). Modeling radiative transfer in heterogeneous 3-D vegetation canopies, *Remote Sensing of Environment* **58**: 131–156.
- Gastellu-Etchegorry, J. P., Gascon, F. and Esteve, P. (2003). An interpolation procedure for generalizing a look-up table inversion method, *Remote Sensing of Environment* **87**: 55–71.
- GCP (2003). The Global Carbon Project, Science Framework and Implementation. Earth System Science Partnership (IGBP, IHDP, WCRP, DIVERSITAS) Report No. 1; Global Carbon Project Report No. 1. 69 pp., Canberra.
- Gemmell, F. (2000). Testing the utility of multi-angle spectral data for reducing the effects of background spectral variations in forest reflectance model inversion, *Remote Sensing of Environment* **72**: 46–63.
- GEO (2005). Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan. URL: <http://www.earthobservations.org/docs/10-Year%20Implementation%20Plan.pdf>.
- Gerstl, S. and Borel, C. (1992). Principles of the radiosity method versus radiative transfer for canopy reflectance modeling, *IEEE Transactions on Geoscience and Remote Sensing* **30**: 271–275.
- Giglio, L., Descloitres, J., Justice, C. O. and Kaufman, Y. (2003). An enhanced contextual fire detection algorithm for MODIS, **87**: 273–282.
- Gobron, N., Pinty, B., Aussedat, O., Chen, J., Cohen, W., Fensholt, R., Gond, V., Huemmrich, K., Lavergne, T., Mélin, F., Privette, J., Sandholt, I., Taberner, M., Turner, D., Verstraete, M. M. and Widlowski, J.-L. (2006). Evaluation of fraction of absorbed photosynthetically active radiation products for different canopy radiation transfer regimes: methodology and results using Joint Research Center products derived from SeaWiFS against ground-based estimations, *Journal of Geophysical Research* **111**. doi:10.1029/2005JD006511.
- Gobron, N., Pinty, B., Verstraete, M. M. and Govaerts, Y. (1997). A semidiscrete model for the scattering of light by vegetation, *Journal of Geophysical Research* **102**: 9431–9446.
- Gobron, N., Pinty, B., Verstraete, M. M. and Widlowski, J.-L. (2000). Advanced vegetation indices optimized for up-coming sensors - design, performance, and applications, *IEEE Transactions on Geoscience and Remote Sensing* **38**: 2489–2505.

- Goel, N. S. and Grier, T. (1988). Estimation of canopy parameters for inhomogeneous vegetation canopies from reflectance data: III. Trim: A model for radiative transfer in heterogeneous three-dimensional canopies, *Remote Sensing of Environment* **25**: 255–293.
- Goel, N. S. and Strebel, D. (1984). Simple beta distribution representation of leaf orientation in vegetation canopies, *Agronomy Journal* **76**: 800–802.
- Goel, N. S. and Thompson, R. L. (2000). A snapshot of canopy reflectance models and a universal model for the radiation regime, *Remote Sensing Reviews* **18**: 197–225.
- Govaerts, Y. M., Jacquemoud, S., Verstraete, M. M. and Ustin, S. L. (1996). Three-dimensional radiation transfer modeling in a dicotyledon leaf, *Applied Optics* **35**: 6585–6598.
- Govaerts, Y. M. and Verstraete, M. M. (1998). Raytran - a monte carlo ray-tracing model to compute light scattering in three-dimensional heterogeneous media, *IEEE Transactions on Geoscience and Remote Sensing* **36**: 493–505.
- Green, R., Eastwood, M., Sarture, C., Chrien, T., Aronsson, M., Chippendale, B., Faust, J., Pavri, B., Chovit, C., Solis, M., Olah, M. and Williams, O. (1998). Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), *Remote Sensing of Environment* **65**: 227–248.
- Grillenberger, J. (2007). Spatial distribution of leaf area index and leaf chlorophyll in cotton fields of Khorezm, Uzbekistan estimated at leaf, plant and regional scale by spectroscopy, remote sensing and GIS modeling. Master thesis, Center for Geoinformatics (Z.GIS), Paris Lodron University, Salzburg.
- Guérif, M. and Duke, C. L. (2000). Adjustment procedures of a crop model to the site specific characteristics of soil and crop using remote sensing data assimilation, *Agriculture Ecosystems & Environment* **81**: 57–69.
- Guyot, G., Baret, F. and Major, D. J. (1988). High spectral resolution: determination of spectral shifts between the red and the near infrared, *International Archives of Photogrammetry and Remote Sensing* **11**: 750–760.
- Habermeyer, M., Mueller, A., Holzwarth, S., Richter, R., Mueller, R., Seitz, K.-H., Seifert, P. and Strobl, P. (2005). Implementation of the automatic processing chain for ARES, in B. Zagajewski and M. Sobczak (eds), *Proceedings of the 4th EARSeL Workshop on Imaging Spectroscopy*, pp. 93 – 101.
- Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J. and Strachan, I. B. (2004). Hyper-spectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling 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.
- Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique, *Princeton University Bulletin* pp. 49–52.

- Hapke, B. (1981). Bidirectional reflectance spectroscopy. I. Theory, *Journal of Geophysical Research* **86**: 3039–3054.
- Harmoney, K., Moore, K., George, J., Brummer, E. and Russell, J. (1997). Determination of pasture biomass using four indirect methods, *Agronomy Journal* **89**: 665–672.
- Hollaus, M., Wagner, W., Eberhöfer, C. and Karel, W. (2006). Accuracy of large-scale canopy heights derived from lidar data under operational constraints in a complex alpine environment, *ISPRS Journal of Photogrammetry & Remote Sensing* **60**: 323–338.
- Hosgood, B., Jacquemoud, S., Andreoli, G., Verdebout, J., Pedrini, A. and Schmuck, G. (1995). The JRC Leaf Optical Properties Experiment (LOPEX'93), *Technical report*.
- Houborg, R. and Boegh, E. (In press). Mapping leaf chlorophyll and leaf area index using inverse and forward canopy reflectance modeling and SPOT reflectance data, *Remote Sensing of Environment*.
- Houborg, R., Soegaard, H. and Boegh, E. (2007). Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using Terra and Aqua MODIS reflectance data, *Remote Sensing of Environment* **106**: 39–58.
- Houser, P., Shuttleworth, W., Famiglietti, J., Gupta, H., Syed, K. and Goodrich, D. (1998). Integration of soil moisture remote sensing and hydrologic modeling using data assimilation., *Water Resources Research* **34**: 3405–3420.
- Huang, Z., Turner, B. J., Dury, S. J., Wallis, I. R. and Foley, W. J. (2004). Estimating foliage nitrogen concentration from hmap data using continuum removal analysis, *Remote Sensing of Environment* **93**: 18–29.
- Huemmrich, K. F. (2001). The GeoSail model: a simple addition to the SAIL model to describe discontinuous canopy reflectance, *Remote Sensing of Environment* **75**: 423–431.
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X. and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices, *Remote Sensing of Environment* **83**: 195–213.
- Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI), *Remote Sensing of Environment* **25**: 295–309.
- Huete, A. R., Liu, H. Q., Batchily, K. and van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS, *Remote Sensing of Environment* **59**: 440–451.
- Hunt, G. (1983). Electromagnetic radiation: the communication link in remote sensing, in B. Siegal and A. Gillespie (eds), *Remote sensing in geology*, John Wiley, New York, pp. 5–45.
- Hunt, J. E. R. and Rock, B. N. (1989). Detection of changes in leaf water content using near- and middle-infrared reflectances, *Remote Sensing of Environment* **30**: 43–54.

-
- Hurtado, E., Artigao, M. and Caselles, V. (1994). Estimating maize (*Zea mays*) evapotranspiration from NOAA-AVHRR thermal data in the Albacete area, Spain, *International Journal of Remote Sensing* **15**: 2023–2037.
- Hyypä, J., Hyypä, H., Inkinen, M., Engdahl, M., Linko, S. and Zhu, Y.-H. (2000). Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes, *Forest Ecology Management* **128**: 109–120.
- Inoue, Y. (2003). Synergy of remote sensing and modeling for estimating ecophysiological processes in plant production, *Plant Production Science* **6**: 3–16.
- INSPIRE (2007). Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007, establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), *Technical report*, the European Parliament and the Council of the European Union. URL: <http://www.ec-gis.org/>.
- IPCC (2007). Intergovernmental panel on climate change - fourth assessment report (AR4), *Technical report*, IPCC, Geneva, Switzerland. URL: <http://www.ipcc.ch>.
- Jacquemoud, S. (1993). Inversion of the PROSPECT + SAIL canopy reflectance model from AVIRIS equivalent spectra: Theoretical study, *Remote Sensing of Environment* **44**: 281–292.
- Jacquemoud, S., Bacour, C., Poilvé, 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**: 471–481.
- Jacquemoud, S. and Baret, F. (1990). PROSPECT: A model of leaf optical properties spectra, *Remote Sensing of Environment* **34**: 75–91.
- Jacquemoud, S., Baret, F., Andrieu, B., Danson, F. M. and Jaggard, K. (1995a). 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**: 163–172.
- Jacquemoud, S., Baret, F. and Hanocq, J. F. (1992). Modeling spectral and bidirectional soil reflectance, *Remote Sensing of Environment* **41**: 123–132.
- Jacquemoud, S. and Ustin, S. L. (2001). Leaf optical properties: A state of the art, *Proceedings of the 8th International Symposium on Physical Measurements & Signatures in Remote Sensing*, CNES.
- Jacquemoud, S., Ustin, S. L., Verdebout, J., Schmuck, G., Andreoli, G. and Hosgood, B. (1996). Estimating leaf biochemistry using the PROSPECT leaf optical properties model, *Remote Sensing of Environment* **56**: 194–202.
- Jacquemoud, S., Verdebout, J., Schmuck, G., Andreoli, G. and Hosgood, B. (1995). Investigation of leaf biochemistry by statistics, *Remote Sensing of Environment* **54**: 180–188.
- Justice, C. O., Giglio, L., Korontzi, S., Owens, J., Morisette, J., Roy, D., Descloitres, J., Alleaume, S., Petitcolin, F. and Kaufman, Y. (2002). The MODIS fire products, **83**: 244–262.

- Kaufman, Y. and Tanré (1992). Atmospherically resistant vegetation index ARVI for EOS-MODIS, *IEEE Transactions on Geoscience and Remote Sensing* **30**: 261–270.
- Kienzler, K. (In prep.). *Potential for increasing nitrogen use efficiency in cotton and winter wheat for irrigated agriculture in Khorezm, Uzbekistan*, PhD thesis, University of Bonn, Bonn, Germany.
- Kim, M. S., Daughtry, C. S. T., Chapelle, E. W. and McMurtrey, J. E. (1994). The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (APAR), *Proceedings of ISPRS '94*, Val d'Isere, France, pp. 299–306.
- Kimes, D. S., Privette, J. L., Abuelgasim, A. A., Knyazikhin, Y. and Gao, F. (2000). Inversion methods for physically-based models (for extracting vegetation variables using bidirectional and spectral data from modern-borne sensors), *Remote Sensing Reviews* **18**: 381–439.
- Knyazikhin, Y., Glassy, J., Privette, J. L., Tian, Y., Lotsch, A., Zhang, Y., Wang, Y., Morisette, J. T., Votava, T., Myneni, R. B., Nemani, R. R. and Running, S. W. (1999a). MODIS leaf area index (LAI) and fraction of photosynthetically active radiation absorbed by vegetation (FPAR) product (MOD15) algorithm theoretical basis document.
- Knyazikhin, Y., Martonchik, J., Diner, D., Myneni, R. B., Verstraete, M., Pinty, B. and Gobron, N. (1999b). Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from atmosphere-corrected MISR data, *Journal of Geophysical Research* **103**: 32239–32256.
- Knyazikhin, Y., Martonchik, J., Myneni, R. B., Diner, D. and Running, S. W. (1999c). Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data, *Journal of Geophysical Research* **103**: 32257–32276.
- Koetz, B. (2006). *Estimating biophysical and biochemical properties over heterogeneous vegetation canopies - radiative transfer modeling in forest canopies based on imaging spectrometry and LIDAR*, PhD thesis, Remote Sensing Laboratories, Department of Geography, University of Zurich, Zurich, Switzerland.
- Koetz, B., Baret, F., Poilvé, H. and Hill, J. (2005a). Use of coupled canopy structure dynamic and radiative transfer models to estimate biophysical canopy characteristics, *Remote Sensing of Environment* **95**: 115–124.
- Koetz, B., Kneubühler, M., Widlowski, J.-L., Morsdorf, F., Schaepman, M. and Itten, K. (2005). Assessment of canopy structure and heterogeneity from multi-angular CHRIS-PROBA data, *Proceedings of the 9th International Symposium on Physical Measurements and Signatures in Remote Sensing (ISPMRS)*, Beijing, China, pp. 73–78.
- Koetz, B., Morsdorf, F., Sun, G., Ranson, K. J., Itten, K. and Allgower, B. (2005). Inversion of a LIDAR waveform model for forest biophysical parameter estimation, *IEEE Geoscience and Remote Sensing Letters* **3**: 49–53.
- Koetz, B., Schaepman, M., Morsdorf, F., Bowyer, P., Itten, K. and Allgower, B. (2004). Radiative transfer modeling within a heterogeneous canopy for estimation of forest fire fuel properties, *Remote Sensing of Environment* **92**: 332–344.

-
- Koetz, B., Sun, G., Ranson, K. J., Kneubühler, M., Itten, K. and Allgöwer, B. (2007). Fusion of imaging spectrometer and LIDAR data over combined radiative transfer models for forest canopy characterization, *Remote Sensing of Environment* **106**: 449–459.
- Koetz, B., Widlowski, J.-L., Morsdorf, F., Verrelst, J., Schaepman, M. and Kneubühler, M. (2006). Suitability of the parametric model RPV to assess canopy structure and heterogeneity from multi-angular CHRIS-PROBA data, *Proceedings of the 4th ESA CHRIS/Proba Workshop*, Frascati, Italy.
- 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.
- Koltunov, A. and Ustin, S. (2007). Early fire detection using non-linear multitemporal prediction of thermal imagery, *Remote Sensing of Environment* **110**: 18–28.
- Kruse, F., Lefkoff, A., Boardman, J., Heidebrecht, K., Shapiro, A., Barloon, P. and Goetz, A. (1993). The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data, *Remote Sensing of Environment* **44**: 145–163.
- Kuusk, A. (1985). The hot spot effect of a uniform vegetation cover, *Soviet Journal of Remote Sensing* **3**: 645–658.
- Kuusk, A. (1995a). A fast, invertible canopy reflectance model, *Remote Sensing of Environment* **51**: 342–350.
- Kuusk, A. (1995b). A markov chain model of canopy reflectance, *Agricultural and Forest Meteorology* **76**: 221–236.
- Kuziev, R. (1989). Methodical manual on valuation of irrigated soils.
- Lacaze, R. (2005). POLDER-2 land surface level-3 products, user manual, algorithm description & product validation, issue 1.40, *Technical report*, MEDIAS-France.
- Launay, M. and Guerif, M. (2005). Assimilating remote sensing data into a crop model to improve predictive performance for spatial applications, *Agriculture, Ecosystems & Environment* **111**: 321–339.
- Lavergne, T., Kaminski, T., Pinty, B., Taberner, M., Gobron, N., Verstraete, M. M., Vossbeck, M., Widlowski, J. L. and Giering, R. (2007). Application to MISR land products of an RPV model inversion package using adjoint and Hessian codes, *Remote Sensing of Environment* **107**: 362–375.
- Leblon, B. (2005). Monitoring forest fire danger with remote sensing, *Natural Hazards* **35**: 343–359.
- Liang, N. S. (2004). *Qualitative Remote Sensing of Land Surfaces*, John Wiley & Sons.
- Lichtenthaler, H. (1987). Chlorophylls and carotenoids: pigments of photosynthetic biomembranes, *Methods of Enzymology* **148**: 350–382.
-

- LICOR (2000). LAI-2000 plant canopy analyzer operating manual, *Technical report*, Lincoln, Nebraska. URL: http://www.licor.com/env/PDF_Files/LAI2000.pdf.
- Lobell, D., Asner, G., Law, B. and Treuhaft, R. (2001). Subpixel canopy cover estimation of coniferous forests in oregon using SWIR imaging spectrometry, *Journal of Geophysical Research* **106**: 5151–5160.
- Lotsch, A., Tian, Y., Friedl, M. and Myneni, R. (2003). Land cover mapping in support of LAI and fPAR retrievals from EOS-MODIS and MISR: classification methods and sensitivities to errors, *International Journal of Remote Sensing* **24**: 1997–2016.
- Lucht, W., Schaaf, C. and Strahler, A. (2000). An algorithm for the retrieval of albedo from space using semiempirical BRDF models, *IEEE Transactions on Geoscience and Remote Sensing* **38**: 977–998.
- Lucieer, A. and Stein, A. (2002). Existential uncertainty of spatial objects segmented from satellite sensor imagery, *IEEE Transactions on Geoscience and Remote Sensing* **40**: 2518–2521.
- Maier, S. W. (2000). *Modeling the radiative transfer in leaves in the 300 nm to 2.5 μ m wavelength region taking into consideration chlorophyll fluorescence - the leaf model SLOP*, PhD thesis, German Aerospace Center (DLR), Wessling, Germany.
- Maier, S. W., Ludeker, W. and Gunther, K. P. (1999). SLOP: A revised version of the stochastic model for leaf optical properties, *Remote Sensing of Environment* **68**: 273–280.
- Major, D. J., Baret, F. and Guyot, G. (1990). A vegetation index adjusted for soil brightness, *International Journal of Remote Sensing* **11**: 727–740.
- Makowski, D., Jeuffroy, M. and Guerif, M. (2003). Bayesian methods for updating crop-model predications, applications for predicting biomass and grain protein content, *Proceedings of the Frontis workshop on Bayesian Statistics and Quality Modelling in the agro-food production chain*, Wageningen, The Netherlands, pp. 57–68.
- Martonchik, J., Bruegge, C. and Strahler, A. (2000). A review of reflectance nomenclature used in remote sensing, *Remote Sensing Reviews* **19**: 9–20.
- Melzer, A. (1999). Aquatic macrophytes as tools for lake management, *Hydrobiologica* **395/396**: 181–190.
- Menenti, M., Rast, M., Bach, H., Baret, F., Hurk, B., Jia, L., Li, Z.-l., Knorr, W., Probeck, M., Mauser, W., Moreno, J., Schaepman, M., Verhoef, W. and Verstraete, M. (2005). Understanding vegetation response to climate variability from space with hyper-spectral, multi-angular observations, *9th International Symposium on Physical Measurements and Signature in Remote Sensing (ISPMSRS2005), 17-19 October 2005, Beijing, China, XXXVI: 7/W20, 72-75.*, pp. 72–75. URL: <http://lsiit.u-strasbg.fr/Publications/2005/MRBBHJKPMMMSV05>.
- Mo, X., Liu, S., Lin, Z., Xu, Y., Xiang, Y. and McVicar, T. (2005). Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain, *Ecological Modelling* **183**: 301–322.

-
- Moran, M., Clarke, T., Inoue, Y. and Vidal, A. (1994). Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index, *Remote Sensing of Environment* **49**: 246–263.
- Moran, M., Inoue, Y. and Barnes, E. (1997). Opportunities and limitations from image-based remote sensing in precision crop management, *Remote Sensing of Environment* **61**: 319–346.
- Moulin, S., Bondeau, A. and Delécolle, R. (1998). Combining agricultural crop models and satellite observations: from field to regional scales, *International Journal of Remote Sensing* **19**: 1021–1036.
- Moulin, S., Zurita-Milla, R., Guérif, M. and Baret, F. (2003). Characterizing the spatial and temporal variability of biophysical variables of a wheat crop using hyper-spectral measurements, *Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 4, pp. 2206–2208.
- Müller, A., Richter, R., Habermeyer, M., Dech, S., Segl, K. and Kaufmann, H. (2005). Spectroradiometric requirements for the reflective module of the airborne spectrometer ARES, *IEEE Geoscience and Remote Sensing Letters* **2**(3): 329–332.
- Müller, A., Richter, R., Kaufmann, C., Stuffer, T., Hofer, S., Mehl, H., Schreier, G., Langemann, M., Haydn, R. and Doll, B. (2004). EnMAP (ENvironmental Monitoring and Analysis Program) - a German hyperspectral initiative, *Proceedings of the 12th Australasian Remote Sensing and Photogrammetry Conference*.
- 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.
- Myneni, R. B., Asrar, G. and Hall, F. G. (1992). A three-dimensional radiative transfer method for optical remote sensing of vegetated land surfaces, *Remote Sensing of Environment* **41**: 105–121.
- Nagler, P. L., Daughtry, C. S. T. and Goward, S. N. (2000). Plant litter and soil reflectance, *Remote Sensing of Environment* **71**: 207–215.
- Nicodemus, F., Richmond, J., Hisia, J., Ginsberg, I. and Limperis, T. (1977). Geometrical considerations and nomenclature for reflectance, *Technical report*, National Bureau of Standards, U.S. Department of Commerce, Washington D.C.
- Olioso, A., Chauki, H., Courault, D. and Wigneron, J. (1999). Estimation of evapotranspiration and photosynthesis by assimilation of remote sensing data into SVAT models., *Remote Sensing of Environment* **68**: 341–356.
- Pearlman, J., Barry, P., Segal, C., Shepanski, J., Beiso, D. and Carman, S. (2003). Hyperion, a space-based imaging spectrometer, *IEEE Transactions on Geoscience and Remote Sensing* **41**: 1160–1173.
- Pearson, R. L. and Miller, L. D. (1972). Remote mapping of standing crop biomass for estimation of the productivity of the short-grass prairie, Pawnee National Grasslands, Colorado,

- Proceedings of the 8th International Symposium on Remote Sensing of Environment, ERIM International*, Ann Arbor, MI, USA, pp. 1357–1381.
- Penuelas, J., Gamon, J. A., Fredeen, A. L., Merino, J. and Field, C. B. (1994). Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves, *Remote Sensing of Environment* **48**: 135–146.
- Pinnel, N. (2007). *A method for mapping submerged macrophytes in lakes using hyperspectral remote sensing*, PhD thesis, Limnological Institute, Technical University of Munich, Munich, Germany. 191 pp.
- Pinty, B., Gobron, N., Widlowski, J. L., Gerstl, S. A. W., Verstraete, M. M., Antunes, M., Bacour, C., Gascon, F., Gastellu, J. P. and Goel, N. (2001). Radiation transfer model intercomparison (RAMI) exercise, *Journal of Geophysical Research* **106**: 11, 937–956.
- Pinty, B., Gobron, N., Widlowski, J. L., Lavergne, T. and Verstraete (2004a). Synergy between 1-D and 3-D radiation transfer models to retrieve vegetation canopy properties from remote sensing data, *Journal of Geophysical Research* **109**: doi:10.1029/2004JD005214.
- Pinty, B., Verstraete, M. and Dickinson, R. (1989). A physical model for predicting bi-directional reflectances over bare soil, *Remote Sensing of Environment* **27**: 273–288.
- Pinty, B., Widlowski, J. L., Gobron, N., Verstraete, M. M. and Diner, D. (2002). Uniqueness of multiangular measurements: Part I. an indicator of surface heterogeneity from MISR, *IEEE Transactions on Geoscience and Remote Sensing* **40**: 1560–1573.
- Pinty, B., Widlowski, J. L., Taberner, M., Gobron, N., Verstraete, M. M., Disney, M., Gascon, F., Gastellu, J. P., Jiang, L. and Kuusk, A. (2004b). Radiation transfer model intercomparison (RAMI) exercise - results from the second phase, *Journal of Geophysical Research. D. Atmospheres* **109**.
- Plant, R., Munk, D., Roberts, B., Vargas, R., Rains, D., Travis, R. and Hutmacher, R. (2000). Relationships between remotely sensed reflectance data and cotton growth and yield, *Transactions of the ASAE* **43**: 535–546.
- Privette, J. L., Morisette, J., Baret, F., Gower, S. and Myneni, R. B. (1996). Summary of the international workshop on LAI product validation, *Earth Observer* **13**: 18–22.
- Psomas, A., Kneubuehler, M., Itten, K. and Zimmermann, N. (2007). Hyperspectral remote sensing for seasonal estimation of above ground biomass in Swiss grasslands, *Proceedings of the 10th International Symposium on Physical Measurements and Signatures in Remote Sensing, ISPMSRS'07*, Davos, Switzerland.
- Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H. and Sorooshian, S. (1994). A modified soil adjusted vegetation index, *Remote Sensing of Environment* **48**: 119–126.
- Qin, W. and Xiang, Y. (1994). On the hotspot effect of leaf canopies: modeling study and influence of leaf shape, *Remote Sensing of Environment* **50**: 95–106.
- Qiu, J., Gao, W. and Lesht, B. M. (1998). Inverting optical reflectance to estimate surface properties of vegetation canopies, *International Journal of Remote Sensing* **19**: 641–656.

- Rast, M. (2004). SPECTRA - Surface Processes and Ecosystem Changes Through Response Analysis, *Technical report*, ESA Publication Division, Noordwijk, The Netherlands.
- Read, J. J., Tarpley, L., McKinion, J. M. and Reddy, K. R. (2002). Narrow-waveband reflectance ratios for remote estimation of nitrogen status in cotton, *Journal of Environmental Quality* **31**: 1442–1452.
- Richardson, A. J. and Wiegand, C. L. (1977). Distinguishing vegetation from soil background information (by gray mapping of Landsat MSS data), *Photogrammetric Engineering and Remote Sensing* **43**: 1541–1552.
- Richardson, A., Wiegand, C., Arkin, G., Nixon, P. and Gerbermann, A. (1982). Remotely-sensed spectral indicators of sorghum development and their use in growth modelling, *Agricultural Meteorology* **26**: 11–23.
- Richter, R. (2007a). Atmospheric / topographic correction for airborne imagery, (ATCOR-4 User Guide, version 4.2, *Technical report*, DLR - German Aerospace Center, Remote Sensing Data Center.
- Richter, R. (2007b). Atmospheric / topographic correction for satellite imagery ATCOR-2/3 user guide, version 6.3, *Technical report*, DLR - German Aerospace Center, Remote Sensing Data Center. DLR IB 565-03/06.
- Richter, R., Bachmann, M., Dorigo, W. and Mueller, A. (2006a). Influence of the adjacency effect on ground reflectance measurements, *IEEE Geoscience and Remote Sensing Letters* **4**(3): 565–569.
- Richter, R., Müller, A., Habermeyer, M., Dech, S., Segl, K. and Kaufmann, H. (2005). Spectral and radiometric requirements for the airborne thermal imaging spectrometer, *International Journal of Remote Sensing* **26**(15): 3149 – 3162.
- Richter, R. and Schlaepfer, D. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: atmospheric/topographic correction, *International Journal of Remote Sensing* **23**: 2631–2649.
- Richter, R., Schlaepfer, D. and Mueller, A. (2006b). An automatic atmospheric correction algorithm for visible/NIR imagery, *International Journal of Remote Sensing* **27**(10): 2077–2085.
- Richter, T. and Fukshansky, L. (1996). Optics of a bifacial leaf: 1. A novel combined procedure for deriving the optical parameters, *Journal of Photochemistry and Photobiology* **63**: 507–516.
- Roberts, D., Gardner, M., Church, R., Ustin, S., Scheer, G. and Green, R. (1998). Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models, *Remote Sensing of Environment* **65**: 267–279.
- Rodgers, C. (2000). *Inverse methods for atmospheric sounding. Theory and praxis*, Vol. 2 of *Series on Atmospheric, Oceanic and Planetary Physics*, World Scientific, Oxford.

- Rondeaux, G., Steven, M. and Baret, F. (1996). Optimization of soil-adjusted vegetation indices, *Remote Sensing of Environment* **55**: 95–107.
- Rougoor, C., Vellinga, T., Huirne, R. and Kuipers, A. (1999). Influence of grassland and feeding management on technical and economic results of dairy farms, *Netherlands Journal of Agricultural Science* **47**: 135–151.
- Roujean, J.-L. (2000). A parametric hot spot model for optical remote sensing applications, *Remote Sensing of Environment* **71**: 197–206.
- Roujean, J.-L. and Bréon, F.-M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements, *Remote Sensing of Environment* **51**: 375–384.
- Roujean, J. M. and Lacaze, R. (2002). Global mapping of vegetation parameters from POLDER multiangular measurements for studies of surface-atmosphere interactions: A pragmatic method and its validation., *Journal of Geophysical Research* **107**: doi: 10.1029/2001JD000751.
- Rouse, J. W., Haas, R. H., Schell, J. A. and Deering, J. A. (1973). Monitoring vegetation systems in the Great Plains with ERTS, *Proceedings of the Third Symposium on Significant Results Obtained with ERTS*, pp. 309–317.
- Ruecker, G., Dorigo, W., Lamers, J., Ibragimov, N., Kienzler, K., Strunz, G. and Vlek, P. (2006). Remote sensing of chlorophyll and nitrogen in cotton fields in Khorezm, Uzbekistan, *Proceedings of the 4th ESA CHRIS/Proba Workshop*, ESA, Frascati, Italy.
- Sandmeier, S. and Deering, D. (1999). Structure analysis and classification of boreal forests using airborne hyperspectral BRDF data from ASAS, *Remote Sensing of Environment* **69**: 281–295.
- Sandmeier, S., Müller, C., Hosgood, B. and Andreoli, G. (1998). Physical mechanisms in hyper-spectral brdf data of grass and watercress, *Remote Sensing of Environment* **66**: 222–233.
- Sanz, C., Espana, M., Baret, F., Weiss, M., Vaillant, L., Hanocq, J., Sarrouy, C., Clastre, P., Bruguier, N., Chelle, M., Andrieu, B. and Zurfluh, O. (1997). Bi-directional characteristics of leaf reflectance and transmittance: measurement and influence on canopy bi-directional reflectance, in G. Guyot and T. Phulpin (eds), *7th International Symposium on physical measurements and signatures in remote sensing*, Courchevel, France, pp. 583–590.
- Savitzky, A. and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures, *Analytical Chemistry* **36**: 1627–1639.
- Schaepman, M. E. (2007). Spectrodirectional remote sensing: From pixels to processes, *International Journal of Applied Earth Observation and Geoinformation* **9**: 204–223.
- Schaepman, M. E., Itten, K., Schläpfer, D., Kaiser, J., Brazile, J., Debruyne, W., Neukom, A., Feusi, H., Adolph, P., Moser, R., Schilliger, T., de Vos, L., Brandt, G., Kohler, P., Meng, M., Piesbergen, J., Strobl, P., Gavira, J., Ulbrich, G. and Meynart, R. (2004). APEX: current status of the airborne dispersive pushbroom imaging spectrometer, *Proceedings of SPIE: Sensors, Systems, and Next-Generation Satellites VII*, vol. 5234, pp. 202–210.

-
- Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S. and Martonchik, J. V. (2006). Reflectance quantities in optical remote sensing—definitions and case studies, *Remote Sensing of Environment* **103**: 27–42.
- Schlaepfer, D. and Richter, R. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 1: parametric orthorectification, *International Journal of Remote Sensing* **23**: 2609 – 2630.
- Schlerf, M., Atzberger, C. and Hill, J. (2005). Remote sensing of forest biophysical variables using hysmap imaging spectrometer data, *Remote Sensing of Environment* **95**: 177–194.
- Schneider, T. (1994). *Möglichkeiten und Grenzen der spektralen Trennbarkeit ackerbaulicher Oberflächentypen - eine Abschätzung anhand spektroskopischer Untersuchungen über die Vegetationsperiode*, PhD thesis, Ludwig-Maximilians-Universität, Munich.
- Schneider, T., Dorigo, W., Huber, K. and Schneider, W. (2006). Goniometer measurements for biophysical parameter retrieval in support of CHRIS data evaluations, *Proceedings of the 4th ESA CHRIS/Proba Workshop*, ESA, Frascati, Italy.
- Schowengerdt, R. A. (1997). *Remote Sensing, Models and Methods for Image Processing*, second edition edn, Academic Press, San Diego.
- Schreier, G. and Dech, S. (2005). High resolution earth observation satellites and services in the next decade - a European perspective, *Acta Astronautica* **57**: 520–533.
- Schut, A. G. T. (2003). *Imaging spectroscopy for characterization of grass swards*, PhD thesis, Wageningen University.
- Serrano, L., Peñuelas, J. and Ustin, S. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals, *Remote Sensing of Environment* **81**: 355–364.
- Shabanov, N. V., Knyazikhin, Y., Baret, F. and Myneni, R. B. (2000). Stochastic modeling of radiation regime in discontinuous vegetation canopies, *Remote Sensing of Environment* **74**: 125–144.
- Sims, D. A. and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages, *Remote Sensing of Environment* **81**: 337–354.
- Slater, P. N. (1980). *Remote Sensing, Optics and Optical Systems*, Addison-Wesley, London.
- Smit, A., Guiking, L., Kater, A., Pronk, F., Ruijter, R., Schreuder, J., Schroder, J., Schut, A., De Haan, M., van Dam, A., van der Schoot, J. and Reuler, H. (2003). Maatregelenpakketten om te voldoen aan de MINAS2003 normen., *Technical report*, Plant Research International, Wageningen, the Netherlands.
- Strahler, A., Muchoney, D., Borak, J., Friedl, M., Gopal, S., Lambin, E. and Moody, A. (1999). MODIS Land Cover Product Algorithm Theoretical Basis Document (ATBD), Version 5.0 MODIS Land Cover and Land-Cover Change, *Technical report*, Center for Remote Sensing, Department of Geography, Boston University, Boston, MA. 72 pp.
-

- Suits, G. (1983). The nature of electromagnetic radiation, in D. Simonett and F. Ullaby (eds), *Manual of remote sensing*, 2 edn, Vol. 1, American Society of Remote Sensing, pp. 37–60.
- Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics. 358 pp.
- Tarpley, L., Reddy, K. and Sassenrath-Cole, G. (2000). Reflectance indices with precision and accuracy in predicting cotton leaf nitrogen concentration, *Crop Science* **40**: 1814–1819.
- Thanisawanyangkura, S., Sinoquet, H., Rivet, P., Cretenet, M. and Jallas, E. (1997). Leaf orientation and sunlit leaf area distribution in cotton, *Agricultural and forest meteorology* **86**: 1–15.
- 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.
- Treuhhaft, R., Asner, G., Law, B. and Van Tuyl, S. (2002). Forest leaf area density profiles from the quantitative fusion of radar and hyperspectral data, *Journal of Geophysical Research* **107**: 148–227.
- Treuhhaft, R., Law, B. and Asner, G. (2004). Forest attributes from radar interferometric structure and its fusion with optical remote sensing, *Bioscience* **54**: 561–571.
- Tucker, C., Vanpraet, C., Boerwinkel, E. and Gaston, A. (1983). Satellite remote sensing of total dry matter production in the Senegalese Sahel, *Remote Sensing of Environment* **13**: 46–474.
- Ustin, S. L., Jacquemoud, S., Palacios-Orueta, A., Li, L. and Whiting, M. (2005). Remote sensing based assessment of biophysical indicators for land degradation and desertification, *Remote sensing and geoinformation processing in the assessment and monitoring of land degradation and desertification*, Trier, Germany, pp. 2–21.
- Ustin, S. L., Roberts, D. A., Gamon, J. A., Asner, G. P. and Green, R. O. (2004). Using imaging spectroscopy to study ecosystem processes and properties, *BioScience* **54**: 523–534.
- Van Leeuwen, W., Orr, B., Marsh, S. and Herrmann, S. (2003). Multi-sensor NDVI data continuity: uncertainties and implications for vegetation monitoring applications, *Remote Sensing of Environment* **100**: 67–81.
- Verhoef, W. (1984). Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model, *Remote Sensing of Environment* **16**: 125–141.
- Verhoef, W. (1985). Earth observation modeling based on layer scattering matrices, *Remote Sensing of Environment* **17**: 165–178.
- Verhoef, W. (1997). *Theory of radiative transfer models applied in optical remote sensing of vegetation canopies*, PhD thesis, Wageningen Agricultural University. ISBN: 90-5485-804-4.

-
- Verhoef, W. (2007). A Bayesian optimisation approach for model inversion of hyperspectral-multidirectional observations: the balance with a priori information, *Proceedings of the 10th International Symposium on Physical Measurements and Signatures in Remote Sensing*, Davos, Switzerland.
- Verhoef, W. and Bach, H. (2003a). Remote sensing data assimilation using coupled radiative transfer models, *Physics and Chemistry of the Earth, Parts A/B/C* **28**: 3–13.
- Verhoef, W. and Bach, H. (2003b). Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models, *Remote Sensing of Environment* **87**: 23–41.
- Verhoef, W. and Bach, H. (2007). Coupled soil-leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data, *Remote Sensing of Environment* **109**: 166–182.
- Verstraete, M. and Pinty, B. (1996). Designing optimal spectral indexes for remote sensing applications, *IEEE Transactions on Geoscience and Remote Sensing* **34**: 1254.
- Verstraete, M., Pinty, B. and Myneni, R. (1996). Potential and limitations of information extraction on the terrestrial biosphere from satellite remote sensing, *Remote Sensing of Environment* **58**: 201–214.
- Vohland, M. and Jarmer, T. (2007). Estimating structural and biochemical parameters for grassland from spectroradiometer data by radiative transfer modelling (PROSPECT + SAIL), *International Journal of Remote Sensing* **In press**.
- Vohland, M. and Mader, S. (2007). Numerical minimisation and artificial neural networks: Two different approaches to retrieve parameters from a canopy reflectance model, *Proceedings of the 6th EARSeL Workshop on Imaging Spectroscopy*, EARSeL, Bruges, Belgium.
- Vyn, T. and Hooker, D. (2002). Assessment of multiple- and single-factor stress impacts on corn, *Field Crops Research* **75**: 123–137.
- Wang, Y., Li, X., Nashed, Z., Zhao, F., Yang, H., Guan, Y. and Zhang, H. (In press). Regularized kernel-based brdf model inversion method for ill-posed land surface parameter retrieval, *Remote Sensing of Environment* .
- Watson, D. (1947). Comparative physiological studies in growth of field crops. I. variation in net assimilation rate and leaf area between species and varieties, and within and between years, *Annals of Botany* **11**: 41–76.
- Weiss, M. and Baret, F. (1999). Evaluation of canopy biophysical variable retrieval performances from the accumulation of large swath satellite data, *Remote Sensing of Environment* **70**: 293–306.
- Weiss, M., Baret, F., Garrigues, S. and Lacaze, R. (2007). LAI and fAPAR CYCLOPES global products derived from VEGETATION. Part 2: validation and comparison with MODIS collection 4 products, *Remote Sensing of Environment* **110**: 317–331.
-

- Weiss, M., Baret, F., Leroy, M., Hauteceœur, O., Bacour, C., Prévot, L., and Bruguier, N. (2002). Validation of neural net techniques to estimate canopy biophysical variables from remote sensing data, *Agronomie* **22**: 547–553.
- Weiss, M., Baret, F., Myneni, R. B., Pragnère, A. and Knyazikhin, Y. (2000). Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data, *Agronomie* **20**: 3–22.
- Weiss, M., Baret, F., Smith, G. J., Jonckheere, I. and Coppin, P. (2004). Review of methods for in situ leaf area index (LAI) determination: Part II. estimation of LAI, errors and sampling, *Agricultural and Forest Meteorology* **121**: 37–53.
- Weiss, M., Troufleur, D., Baret, F., Chauki, H., Prévot, L., Olioso, A., Bruguier, N. and Brisson, N. (2001). Coupling canopy functioning and radiative transfer models for remote sensing data assimilation, *Agricultural and Forest Meteorology* **108**: 113–128.
- White, M. A., Asner, G. P., Nemani, R. R., Privette, J. L. and Running, S. W. (2000). Measuring fractional cover and leaf area index in arid ecosystems: Digital camera, radiation transmittance, and laser altimetry methods, *Remote Sensing of Environment* **74**: 45–57.
- Widlowski, J. L., Pinty, B., Gobron, N., Verstraete, M. M., Diner, D. and Davis, A. (2004). Canopy structure parameters derived from multi-angular remote sensing data for terrestrial carbon studies, *Climatic Change* pp. 403–415.
- Widlowski, J. L., Pinty, B., Lavergne, T., Verstraete, M. and Gobron, N. (2005). Using 1-D models to interpret the reflectance anisotropy of 3-D targets: issues and caveats, *IEEE Transactions on Geoscience and Remote Sensing* **43**: 2008–2017.
- Widlowski, J. L., Pinty, B., Lavergne, T., Verstraete, M. M. and Gobron, N. (2006a). Horizontal radiation transport in 3-D forest canopies at multiple spatial resolutions: Simulated impact on canopy absorption, *Remote Sensing of Environment* **103**: 379–397.
- Widlowski, J.-L., Taberner, M., Pinty, B., Bruniquel-Pinel, V., Disney, M., Fernandes, R., Gastellu-Etchegorry, J.-P., Gobron, N., Kuusk, A., Lavergne, T., Leblanc, S., Lewis, P., Martin, E., Möttus, M., North, P. J. R., Qin, W., Robustelli, M., Rochdi, N., Ruiloba, R., Soler, C., Thompson, R., Verhoef, W., Verstraete, M. M. and Xie, D. (2006b). The third Radiation transfer Model Intercomparison (RAMI) exercise: Documenting progress in canopy reflectance modelling, *Journal of Geophysical Research* **112**.
- Winter, M., Lucey, P. and Steutel, D. (2003). Examining hyperspectral unmixing error reduction due to stepwise unmixing, in S. Shen and P. Lewis (eds), *Proceedings SPIE, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX.*, Vol. 5093, pp. 280–289.
- Yang, C., Bradford, J. M. and Wiegand, C. (2001). Airborne multispectral imagery for mapping variable growing conditions and yields of cotton, grain sorghum, and corn, *Transactions of the ASAE* **44**: 1983–1994.
- Yang, C., Bradford, J. M. and Wiegand, C. (2003). Evaluation of cotton defoliation strategies using airborne multispectral imagery, **46**: 869–876.

- 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.
- Zarco-Tejada, P. J., Berjon, A., Lopez-Lozano, R., Miller, J. R., Martin, P., Cachorro, V., Gonzalez, M. R. and De Frutos, A. (2005a). Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy, *Remote Sensing of Environment* **99**: 271–287.
- Zarco-Tejada, P. J., Miller, J. R., Mohammed, G. H., Noland, T. L. and Sampson, P. H. (2002). Vegetation stress detection through chlorophyll a+b estimation and fluorescence effects on hyperspectral imagery, *Journal of Environmental Quality* **31**: 1433–1441.
- Zarco-Tejada, P. J., Miller, J. R., Noland, T. L., Mohammed, G. H. and Sampson, P. H. (2001). Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data, *IEEE Transactions on Geoscience and Remote Sensing* **39**: 1491–1507.
- Zarco-Tejada, P. J., Rueda, C. A. and Ustin, S. L. (2003). Water content estimation in vegetation with modis reflectance data and model inversion methods, *Remote Sensing of Environment* **85**: 109–124.
- Zarco-Tejada, P. J., Ustin, S. L. and Whiting, M. L. (2005b). Temporal and spatial relationships between within-field yield variability in cotton and high-spatial hyperspectral remote sensing imagery, *Agronomy Journal* **97**: 641–653.
- Zhang, Q., Xiao, X., Braswell, B., Linder, E., Baret, F. and Moore III, B. (2005). Estimating light absorption by chlorophyll, leaf and canopy in a deciduous broadleaf forest using MODIS data and a radiative transfer model, *Remote Sensing of Environment* **99**: 357–371.
- Zhang, Y., Shabanov, N., Knyazikhin, Y. and Myneni, R. (2002a). Assessing the information content of multiangle satellite data for mapping biomes II. Theory., *Remote Sensing of Environment* **80**: 435–446.
- Zhang, Y., Tian, Y., Myneni, R., Knyazikhin, Y. and Woodcock, C. (2002b). Assessing the information content of multiangle satellite data for mapping biomes I. Statistical analysis., *Remote Sensing of Environment* **80**: 418–434.
- Zhao, C., Liu, L., Wang, J., Huang, W., Song, X., Li, C. and Wang, Z. (2004). Methods and application of remote sensing to forecast wheat grain quality, *International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 6, pp. 4008–4010.
- Zhao, D. H., Li, J. L. and Qi, J. G. (2005). Identification of red and NIR spectral regions and vegetative indices for discrimination of cotton nitrogen stress and growth stage, *Computers and Electronics in Agriculture* **48**: 155–169.

Publications of the author

Peer reviewed journal publications

- Dorigo, W. A., Zurita-Milla, R., de Wit, A. J. W., Brazile, J., Singh, R. and Schaepman, M. E. (2007). A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling, *International Journal of Applied Earth Observation and Geoinformation* **9**: 165–193.
- Richter, R., Bachmann, M., Dorigo, W. and Mueller, A. (2006). Influence of the adjacency effect on ground reflectance measurements, *IEEE Geoscience and Remote Sensing Letters* **4**(3): 565–569.

Conference Proceedings

- Dorigo, W., Baret, F., Richter, R., Ruecker, G., Schaepman, M. and Mueller, M. (2007). Retrieving canopy variables by radiative transfer model inversion an automated regional approach for imaging spectrometer data, *Proceedings of the 5th EARSeL Workshop on Imaging Spectroscopy*, EARSeL, Brughes, Belgium.
- Schopfer, J., Huber, S., Schneider, T., Dorigo, W., Oppelt, N., Odermatt, D., Koetz, B., Kneubuehler, M. and Itten, K. (2007). Towards a comparison of spaceborne and ground based spectrodirectional reflectance data, *Proceedings of the ENVISAT Symposium*, ESA, Montreux, Switzerland.
- Schneider, T., Schopfer, J., Oppelt, N., Dorigo, W., Vreeling, W. and Gege, P. (2007). - gonioexp06 a field goniometer intercomparison campaign, in support of physical model inversion and upscaling methods for hyperspectral, multidirectional rs data, *Proceedings of the ENVISAT Symposium*, ESA, Montreux, Switzerland.
- Schneider, T., Dorigo, W., Huber, K. and Schneider, W. (2006). Goniometer measurements for biophysical parameter retrieval in support of CHRIS data evaluations, *Proceedings of the 4th ESA CHRIS/Proba Workshop*, ESA, Frascati, Italy.
- Ruecker, G., Dorigo, W., Lamers, J., Ibragimov, N., Kienzler, K., Strunz, G. and Vlek, P. (2006). Remote sensing of chlorophyll and nitrogen in cotton fields in Khorezm, Uzbekistan., *Proceedings of the 4th ESA CHRIS/Proba Workshop*, ESA.
- Huber, K., Dorigo, W. A., Bauer, T., Eitzinger, S., Haumann, J., Kaiser, G., Linke, R., Postl, W., Rischbeck, P., Schneider, W., Suppan, F. and Weihs, P. (2006). Changes in spectral reflectance of crop canopies due to drought stress, *AIP Conference Proceedings 852, Earth Observation for Vegetation Monitoring and Water Management*, University of Naples, Naples, Italy.
- Huber, K., Dorigo, W. A., Bauer, T., Eitzinger, S., Haumann, J., Kaiser, G., Linke, R., Postl, W.,
-

-
- Rischbeck, P., Schneider, W., Suppan, F. and Weihs, P. (2005). Changes in spectral reflectance of crop canopies due to drought stress, *Proceedings of the 11th SPIE International Symposium on Remote Sensing*, Bruges, Belgium.
- Dorigo, W., Richter, R. and Mueller, A. (2005). A lookup table approach for biophysical parameter retrieval by radiative transfer model inversion applied to wide field of view data, *Proceedings of the 4th EARSeL Workshop on Imaging Spectroscopy*, Warsaw, Poland.
- Feingersh, T., Dorigo, W., Richter, R. and Ben-Dor, E. (2005). A new model-driven correction factor for brdf effects in hyperspectral remote sensing data, *Proceedings of the 4th EARSeL Workshop on Imaging Spectroscopy*, Warsaw, Poland.
- Mohammed, A., Dorigo, W., Habermeyer, M., Mueller, A. and Schneider, T. (2005). Characterization of landcover types by classification and biophysical parameter retrieval using HyMap imagery, *Proceedings of the 4th EARSeL Workshop on Imaging Spectroscopy*, Warsaw, Poland.
- Nielsen, A., Mueller, A. and Dorigo, W. (2004). Hyperspectral data, change detection and the MAD transformation, *Proceedings of the 12th Australasian Remote Sensing & Photogrammetry Association Conference*, Fremantle, Australia. CD-ROM.
- Dorigo, W. A. (2003a). Estimating nutrient uptake of grasslands by assimilating multi- and hyperspectral remote sensing data in growth models, a first approach, *Tagungsband des 20. DFD - Nutzerseminar*, DLR-DFD, Neustrelitz, Germany. CD-ROM.
- Dorigo, W. A. (2003b). A first approach towards estimating nutrient uptake of grasslands by assimilating multi- and hyperspectral remote sensing data in growth models, *Proceedings of the 3rd EARSeL Workshop on Imaging Spectroscopy*, EARSeL.

Technical reports

- Dorigo, W., Bachmann, M. and Heldens, W. (2006). AS Toolbox & processing of field spectra – user manual, *Technical report*, DLR-DFD, Imaging Spectroscopy Group. URL: http://cocoon.caf.dlr.de/astools_en.html.

Master Thesis

- Dorigo, W. and Van Groenendaal, Y. (2000). *Modeling interception loss in a semi-arid environment: a study using remotely sensed data performed in southeastern Spain*, Master's thesis, Department of Physical Geography, Utrecht University, the Netherlands.

Curriculum Vitae



Wouter Dorigo

Date and place of birth July 18, 1975, Arnhem, the Netherlands

Professional experience

- 06/2007 - present Research associate at the Institute of Photogrammetry and Remote Sensing (I.P.F.), Vienna University of Technology (TU Wien), Vienna, Austria
- 11/2002 - 10/2006 Research associate at the Imaging Spectroscopy team of the German Remote Sensing Data Center, German Aerospace Center (DLR), Oberpfaffenhofen, Germany
- 3-10/2002 Project manager Actual Height model of the Netherlands (AHN), Ministry of Transport, Public Works, and Water Management, Survey Department, Delft, the Netherlands
- 10/2001 -2/2002 GIS specialist at CSO environmental consultancy, Bunnik, the Netherlands

Educational background

- 11/2002 - 10/2005 PhD candidate at the Limnological Station, Technical University of Munich, Germany
- 9/1994 - 6/2001 MA Italian Language and Literature at the Utrecht University, the Netherlands
- 9/1995 - 1/2000 MSc Physical Geography at Utrecht University, the Netherlands
Thesis title: Modeling interception loss in a semi-arid environment, a study using remotely sensed data performed in south-eastern Spain
- 10/1993 - 6/1994 Università per Stranieri, Perugia, Italy
- 9/1987 - 6/1993 Gymnasium, Arnhem, the Netherlands