Coupled canopy-atmosphere modelling for radiance-based estimation of vegetation properties

Valérie Laurent

Coupled canopy-atmosphere modelling for radiance-based estimation of vegetation properties

Valérie Laurent

Thesis committee

Thesis supervisor
Prof. Dr. sc. nat. M.E. Schaepman
Professor of Remote Sensing, University of Zurich, Switzerland
Professor of Geo-information Science and Remote Sensing,
Wageningen University, the Netherlands

Thesis co-supervisors
Prof. Dr. Ing. W. Verhoef
Professor of Advanced Earth Observation for Water Resources Applications and
Professor of Quantitative Remote Sensing in Natural Resources
University of Twente, the Netherlands

Dr. Ir. J.G.P.W. Clevers
Associate Professor at the Laboratory of Geo-information Science and Remote Sensing,
Wageningen University, The Netherlands

Other members
Prof. Dr. M.C. Krol, Wageningen University, the Netherlands
Prof. Dr. A. Bégué, CIRAD, France
Prof. Dr. F. Baret, INRA, France
Dr. R. Zurita-Milla, University of Twente, the Netherlands

This research was conducted under the auspices of the C.T. de Wit Graduate School of
Production Ecology & Resource Conservation (PE&RC)

Coupled canopy-atmosphere modelling for radiance-based estimation of vegetation properties

Valérie Laurent

Table of contents

Chapter 1

Introduction

1.1 Optical remote sensing for vegetation monitoring

During the last few centuries, human activities have risen exponentially and so have their impacts on planet Earth (Vitousek et al. 1997). The use of fossil fuel combined with land cover and land use changes have caused a dramatic augmentation of the greenhouse gases concentration in the atmosphere, affecting the Earth's climate (IPCC 2007). Vegetation plays an important role in the carbon exchange of terrestrial ecosystems (Schimel et al. 2001) and vegetation variables such as leaf area index (LAI) and leaf chlorophyll content (Cab) are therefore major inputs for dynamic global vegetation models (DGVM) (Sitch et al. 2008; Smith et al. 2001). When exploring possible future scenarios with DGVMs, it is important to have an accurate initial situation. Accurate and spatially explicit data about vegetation variables are therefore essential.

Historically, these data were acquired by means of costly and time-consuming field campaigns, resulting into very sparse spatial and temporal sampling schemes (Ramankutty and Foley 1999). Inaccessible and dangerous areas were not sampled at all. Furthermore, these data were usually acquired nationally or regionally using expert knowledge, and the data harmonization for use at global scale is often problematic (FAO 2010). In this context, satellite remote sensing (RS) provides a unique opportunity to collect spatially continuous data at the global scale (Hansen et al. 2002; Myneni et al. 2002). Spaceborne sensors offer standardized and replicable data at regular temporal intervals, if calibrated and validated well. Field campaigns are still important for validating the RS products, although this validation is not trivial because of the scale mismatch between footprints of the field and satellite measurements. This scale gap may be bridged using RS sensors having a pixel footprint intermediate between the field and satellite data footprints (Tian et al. 2002), such as airborne sensors (Morisette et al. 2002).

Optical sensors operate in the solar reflective domain of the electromagnetic spectrum, between 400 and 2400 nm. They measure the radiance reflected by the Earth in a number of spectral bands, each of them carrying different types of information about the vegetation canopy. Optical data contain information related to both the bio-chemical composition (Kokaly et al. 2009; Ustin et al. 2009) and the physical structure (Asner 1998; Bacour et al. 2006a; Baret et al. 2007) of the vegetation canopy.

Traditionally, the measured radiance data are pre-processed into top-of-canopy (TOC) reflectances, before being used for estimating vegetation variables. The pre-processing is still a cumbersome and error-prone task. The TOC approach has, however, proven to be successful in the last decades, enabling the production of high level data products (Garrigues et al. 2008; Myneni et al. 2002; Yang et al. 2006a), which are now being used as inputs for DGVMs.

Two main categories of approaches can be used for estimating the vegetation variables. Empirical approaches rely on statistical relationships between the RS data and the vegetation variables. The obtained relationships are specific to vegetation type and to acquisition conditions (acquisition geometry, sensor spectral characteristics, geographical area). Vegetation indices are the most prominent example of empirical approaches (Huete et al. 2002).

Physically-based approaches rely on canopy reflectance models based on the radiative transfer (RT) theory and have been used for almost 30 years (Sellers 1985). They are generic and therefore easily adaptable to different vegetation types and acquisition conditions, which makes them more widely applicable than empirical approaches (Darvishzadeh et al. 2011), and therefore more suited for global vegetation monitoring (Myneni et al. 2002). To estimate vegetation variables, however, canopy RT models have to be inverted, and the inversion is often under-determined and ill-posed (Combal et al. 2002). Several regularization methods have been proposed to obtain more accurate estimates of the vegetation variables (Baret and Buis 2008).

The following sections present physically-based approaches (section 1.2), and associated regularization methods (section 1.3), before discussing the limitations related to the extensive pre-processing required in the TOC approach (section 1.4). A new top-of-atmosphere (TOA) approach, which does not require pre-processing, presenting a number of advantages over the TOC approach is then introduced (section 1.5). Finally the objectives (section 1.6) and outline (section 1.7) of this thesis are presented.

1.2 Physically-based approaches

1.2.1 Canopy radiative transfer models

Canopy RT models describe the interaction of the electromagnetic radiation with and within the canopy based on the RT theory. They simulate the bi-directional reflectance factor (BRF) of the canopy based on physical parameters. The input parameters include: canopy structural parameters, leaf and soil optical properties, solar and observation geometry. There are four main types of canopy RT models (Goel 1988):

1) turbid medium models simulate the canopy as a homogeneous medium containing randomly distributed infinitely small canopy elements,

2) geometric-optical models assume solid basic geometric shapes such as ellipsoids and cones, e.g. for tree crowns, casting shadows on the soil background,

3) ray-tracing models simulate the trajectory and interactions of many photons with a detailed 3-dimensional representation of the canopy and then record the proportion of photons exiting the scene in each direction, and

4) hybrid models combine elements of turbid medium and geometric-optical models, treating each foliage clump (crown) as a turbid medium.

Hybrid models are therefore better suited to model heterogeneous vegetation canopies than turbid and geometric-optical models. They are, however, not as realistic as ray-tracing models but are more computationally efficient and therefore represent a good trade-off between realism and need of simplification for practical applications (Gemmell et al. 2002; Pinty and Verstraete 1992). One such model was therefore used throughout this thesis.

1.2.2 Estimating the vegetation variables: model inversion

The parameters of interest (e.g. LAI and Cab) are hereafter referred to as variables. Canopy RT modelling enables solving the forward problem: from the vegetation variables to the reflectance properties. To estimate the vegetation variables from the RS reflectance data, however, the canopy RT models have to be inverted (Kimes et al. 1998). Because of the complexity of these models, there is no analytical solution to the inverse problem, and therefore many numerical methods have been developed in order to estimate the vegetation variables. Model inversion methods fall in two main categories (Baret and Buis 2008): canopy variables-driven and radiometric data-driven approaches.

Canopy variables-driven approaches require calibrating a parametric inverse model over a training dataset. They are generally tuned for a limited set of observation configurations, and extending this set implies a severe increase of the size of the training dataset.

Radiometric data-driven approaches are based on finding the best match between simulated and measured radiometric data. These methods are much more flexible than canopy variables-driven approaches. They can easily accommodate data from various sensors and observation conditions and were therefore used in this thesis, despite their higher computational load.

Several algorithms have been developed for implementing radiometric data-driven approaches: classical optimization (Combal et al. 2002; Verhoef 2007), look-up tables (LUT) (Darvishzadeh et al. 2008; Gemmell et al. 2002), artificial neural networks (Kimes et al. 1998), and, more recently, genetic algorithms (Fang et al. 2003) and support vector machines (Durbha et al. 2007).

The problem of inverting the canopy RT model is under-determined because of the limited information content (effective dimensionality) of the radiometric signal compared to the high number of unknowns (both variables and parameters) influencing the canopy reflectance (Jacquemoud et al. 2009). The inversion problem is also ill-posed because several sets of inputs can yield very similar spectra (Combal et al. 2002). A limited model accuracy and measurement uncertainties are additional obstacles to the inversion of the canopy RT model. Therefore, regularization methods

have been developed in order to find more stable and reliable solutions (Baret and Buis 2008; Durbha et al. 2007).

1.3 Regularization methods

Regularization methods act on three main aspects of the inverse problems: 1) the number of unknowns, 2) the data information content, and 3) the variables space (number of variables and their variation range). Decreasing the number of unknowns and/or increasing the data information content allow reducing the under-determination of the inverse problem. Reducing the variables space allows reducing the ill-posedness.

1.3.1 Model coupling

The spectral properties of the leaf and soil that are required as inputs in the canopy RT model represent a large number of inputs (leaf reflectance and transmittance, and soil reflectance for each wavelength). Replacing these unknown spectral properties by leaf and soil RT models having only a few input parameters allows reducing the number of inputs in the canopy RT model. The number of unknowns in the variables in the inversion is then smaller and closer to the dimensionality of the data. Consequently, coupling leaf and soil RT models to the canopy RT model efficiently reduces the degree of under-determination and ill-posedness of the inverse problem (Baret and Buis 2008).

1.3.2 Multi-angular data

Multi-angular data are a set of (quasi-)simultaneous observations acquired under varying solar and/or observation geometry. Because the reflectance anisotropy of the canopy is related to its internal structure, multi-angular data contain additional information compared to mono-angular data. This additional information content translates into an increased dimensionality of the data, thereby reducing the under-determination of the inversion. Consequently, multi-angular data have more potential for estimating vegetation variables (Liang et al. 2000; Schaepman 2007). In addition, multi-angular data may be used for constraining and decreasing the ill-posedness of the inversion problem (Asner et al. 1998; Diner et al. 1999). The use of multi-angular data for inverting physical canopy RT models, however, is still relatively scarce (Schaepman et al. 2009).

1.3.3 A priori information

A priori information consists of the knowledge that is available about the unknowns before the inversion. It can be collected from field data, expert knowledge, or agricultural databases, at various spatial scales such as agricultural fields or land

cover types. A priori information allows avoiding unlikely variable combinations by reducing the variable space to a smaller subspace, thus facilitating and reducing the ill-posedness of the inversion (Combal et al. 2002; Li et al. 2001). A priori information has been widely and successfully used to constrain different inversion algorithms: optimization (Combal et al. 2002; Verhoef 2007), look-up tables (Darvishzadeh et al. 2008; Gemmell et al. 2002) and neural networks (Kimes et al. 1998). When including the a priori information in the cost function, one obtains the Bayesian approach (Tarantola 2005). In the Bayesian approach, the distinction between knowns and unknowns becomes fuzzy because nothing is perfectly known or totally unknown. The Bayesian cost function can be used in combination with all inversion methods, but it has been most frequently used with optimization (Lavergne et al. 2007; Meroni et al. 2004; Pinty et al. 2007; Verhoef 2007; Yao et al. 2008).

1.3.4 Spatial constraints

The proximity principle in the spatial domain states that closer pixels are more similar than pixels that are further away. It is therefore possible to define a homogeneous object around the target pixel. This object can then be used either to increase the data information content, or to constrain the inversion. The data information content can be increased by using object statistics in addition to the pixel spectral signature (Atzberger 2004), or by simultaneously using the signatures of all the pixels in the object (Lauvernet et al. 2008) for inverting the canopy RT model. The object can be used for constraining the inversion by enforcing spatial smoothness of the variables in the object (Wang et al. 2008). In the strictest case, the variables can be considered constant for all pixels in the object (Houborg et al. 2009). The object size may be varied depending on the variable. For example, 3 x 3 windows may be used for soil brightness and canopy hotspot parameter (Atzberger and Richter 2012), agricultural fields may be used for the leaf inclination distribution function (Atzberger and Richter 2012; Houborg et al. 2009), while the entire image may be suitable for the atmospheric parameters (Lauvernet et al. 2008). It is also possible to implement an object size hierarchy to fine-tune the implementation of the spatial constraints (Atzberger and Richter 2012).

1.3.5 Temporal constraints

Similarly to the spatial domain, when a time series of acquisitions is available, the proximity principle can be used in the temporal domain. The temporal neighbourhood of a pixel observation then includes previous and later observations of the same pixel. This neighbourhood can then be used to simultaneously invert the canopy RT model for all observations in the neighbourhood (Lauvernet et al. 2008), thus increasing the data information content. The smoothness principle (Quaife and Lewis 2010; Wang et al. 2008) also can be used to apply temporal constraints. The variables

can be considered fixed (Houborg et al. 2007), or they can be constrained by their expected variation through the time period (e.g. using a plant growth model) (CROMA 2000; Koetz et al. 2007).

1.3.6 Combining the regularization methods

In order to obtain the most stable and reliable estimates of the vegetation variables, the regularization methods can be combined (Baret and Buis 2008). Although it has not been attempted yet, it would be possible to estimate the vegetation variables using a coupled soil-leaf-canopy RT model from a time series of multi-angular RS images, using a priori information and implementing spatial and temporal constraints. The most comprehensive example of combining the regularization methods is the study of Lauvernet et al. (2008), which was performed on a time series of three 5x5-pixel synthetic images. They estimated vegetation variables by simultaneously inverting a coupled RT model over the 75 (3x5x5) available synthetic observations, thus increasing the data information content in both spatial and temporal domains. They also used a Bayesian cost function to include a priori information in the inversion.

1.4 Limitations of the top-of-canopy approach

The performance of the TOC approach for estimating vegetation variables is limited by the sequential correction of atmospheric, adjacency, directional and topographic effects in the pre-processing chain. Furthermore, the mismatch between the simulated and measured reflectance physical quantities, which are compared during the inversion of the canopy RT model, also limits the accuracy of the estimates. This will be explained in the following subsections.

1.4.1 Correction of the atmospheric effect

Atmospheric correction is the main step of the pre-processing chain. It aims at removing the atmospheric effects from the measured radiances, in order to obtain TOC reflectances. The atmospheric correction, however, relies on the inversion of a physically-based atmosphere RT model (Rahman and Dedieu 1994; Richter 2009). Similarly to the inversion of a canopy RT model, this inversion is an ill-posed problem, which causes inherent limitations in the atmospheric correction. Atmospheric RT models simulate the atmospheric optical properties based on surface reflectance signatures, atmospheric variables and vertical concentration profiles of the main atmospheric constituents. They are usually very complex and have long running times (Berk et al. 2003; Vermote et al. 1997b). Therefore, operational atmospheric correction algorithms rely on pre-computed LUTs for inverting the atmospheric RT model. These LUTs store the model outputs for a number of

combinations of the main atmospheric variables (visibility, aerosol type, water vapour concentration), and of viewing and solar geometry. The combination of precise values for one given observation is usually not in the LUT and it is therefore necessary to interpolate between the closest LUT entries (Gao et al. 2009; Guanter et al. 2009; Richter 2009). Furthermore, in order to obtain an analytical formula to obtain the TOC reflectance from the observed radiance, most operational atmospheric correction algorithms require the assumption that the surface is Lambertian, which is generally not true, especially for vegetation (Bicheron and Leroy 2000). In addition, the atmospheric correction often also ignores adjacency and topography effects.

1.4.2 Correction of the adjacency effect

Because of the multiple scattering in the atmosphere, photons from surrounding areas can be scattered into the sensor's instantaneous field of view (IFOV), thus contaminating the signal. The contamination coming from the surrounding pixels is described by the atmospheric point spread function (PSF). The atmospheric PSF is modelled by a two-dimensional function having a sharp peak centred on the target pixel, the number of photons coming from the target and close-by pixels being much higher than the number of photons coming from pixels further away (Verhoef and Bach 2003b). The adjacency effect causes a partial blurring of the RS image and is especially strong for images having a small pixel size (Vermote et al. 1997a). It is usually corrected for by spatially filtering the image (Semenov et al. 2011), but the adjacency effect is already included in the image, so it is like filtering the image twice. In addition, the adjacency effect is influenced by the atmospheric conditions, by surface anisotropy, and by the terrain topography (Tanré et al. 1987). These effects, however, are usually ignored by the correction filter.

1.4.3 Correction of the directional effect

The surface directional properties are described by the bi-directional reflectance function (BRDF). This anisotropy is often regarded as an undesired effect which needs to be corrected before using the image (Bacour et al. 2006b; Schaaf et al. 2002), especially for images with large field of view (FOV), for image mosaicking, or multi-temporal studies. The BRDF correction consists in normalizing all pixels of the image to nadir viewing and to a given sun geometry (Duchemin et al. 2002; Schaaf et al. 2002). It is generally applied to the TOC reflectance image, but correcting for surface directional effects is not consistent with the Lambertian surface assumed in the atmospheric correction step. Most BRDF correction methods are semi-empirical and ignore the impact of the topography on the surface BRDF. Furthermore, land cover information is required to correctly adjust the algorithm parameters to the directional properties of each land cover (Schaaf et al. 2002; Weyermann et al. Accepted).

1.4.4 Correction of the topographic effect

The terrain topography affects the surface reflectance. For a given land cover, a surface on a slope facing towards the sun appears brighter and has higher reflectance than a surface on a slope facing away from the sun. Therefore, when the terrain topography is rough, it may be necessary to correct for the topographic effects. Topographic correction requires a digital elevation model (DEM) to calculate the slope, aspect and sky view factor of each pixel (Gu and Gillespie 1998; Richter and Schläpfer 2002). The topography effect also depends on the lower boundary condition of the atmosphere, and on the surface land cover and associated BRDF (Li et al. 1998; Proy et al. 1989). Various correction approaches are available, some assuming a Lambertian surface, while others use empirical methods to account for the surface anisotropy (Riaño et al. 2003). It is controversial whether the topographic correction should be integrated with the atmospheric correction, or performed subsequently.

1.4.5 Mismatch in physical quantities in the model inversion

At the end of the pre-processing, one obtains hemispherical-conical reflectance factor (HCRF) data (Schaepman-Strub et al. 2006), which are then used as the reference data for inverting the canopy RT model. Most canopy RT models, however, simulate the BRF, not the HCRF (Dorigo et al. 2007). Assuming that the directional effects within the sensor's IFOV are negligible, the HCRF can be used as an approximation of the HDRF. The HDRF includes the BRF, but also the diffuse sky irradiance. This mismatch between the two quantities compared in the inversion is another inherent limitation of the TOC approach.

1.4.6 Problem definition

Physically, atmospheric, adjacency, directional, and topographic effects are inter-related and influence each other (Gao et al. 2009; Tanré et al. 1987). In the pre-processing chain, however, the required corrections are applied sequentially and independently, the correction of each effect requiring assumptions about the others. These assumptions introduce errors that propagate from one correction to the next, finally accumulating in the TOC reflectance data. Consequently, the estimation of vegetation variables by the TOC approach is subject to errors both in the simulations, because of modelling errors and parameter uncertainties, and in the reference TOC reflectance data, because of the error propagation in the pre-processing chain (Rahman 2001). Furthermore, the two reflectance quantities compared during the inversion of the canopy RT model are physically different: the HCRF data obtained from the pre-processing chain are compared to the BRF simulated by the canopy RT model, which is one more limitation of the TOC approach.

1.5 The top-of-atmosphere approach

All the limitations of the TOC approach presented in the previous section arise from the need to correct the measured TOA radiance data to TOC reflectance. A solution to avoid this correction is to estimate the vegetation variables directly from the measured TOA radiance data. This requires coupling the canopy and atmosphere RT models in order to simulate the TOA radiance. The coupled canopy-atmosphere model can then be inverted against the measured TOA radiance.

1.5.1 Implementation

The TOA approach relies on the inversion of a coupled canopy-atmosphere model for estimating the vegetation variables. The same soil, leaf, canopy and atmosphere RT models as in the TOC approach may be involved, but they are all coupled in order to simulate the TOA radiance. Coupled canopy-atmosphere RT models have been developed using either hard or soft coupling. In hard coupling, both the canopy and atmosphere RT models have to be run for each new TOA radiance simulation (Fourty and Baret 1997; Gastellu-Etchegorry et al. 2004; Lauvernet et al. 2008). In soft coupling, the atmospheric RT model is run in advance, and atmospheric variables are calculated and stored for later use. Each new TOA radiance simulation requires running only the canopy RT model and combining its output with the stored atmospheric variables (Börner et al. 2001; Rahman et al. 1993; Verhoef and Bach 2003b). Soft coupling is much more computer efficient than hard coupling and therefore better suited for estimating and monitoring vegetation variables at global scale. Only a few studies (Baret et al. 2006; Lauvernet et al. 2008; Rahman et al. 1993; 2007) have started to investigate the potential of the inversion of the coupled canopy-atmosphere model for estimating vegetation variables.

1.5.2 Towards more accurate estimation of vegetation variables

1.5.2.1. Avoiding the limitations of the TOC approach

In the TOA approach, there is no need to correct the measured TOA radiance data for the atmospheric effects. Instead, the latter are included in the forward coupled canopy-atmosphere model. In the TOA approach, the vegetation variables are estimated by inverting the coupled canopy-atmosphere model, inside of which both canopy and atmosphere models are used in forward mode. Therefore, in theory, the inversion of the coupled model is more accurate than successive inversion of the atmosphere and canopy models. Furthermore, topographic, directional and adjacency effects, which had to be corrected for separately in the TOC approach, can be included in the forward coupled canopy-atmosphere model (Börner et al. 2001; Verhoef and Bach 2003b).

The TOA approach also avoids the approximations needed in the TOC approach for comparing the model simulation to the measurement. Indeed, the coupled model simulates the TOA radiance, which is the physical quantity measured by the sensor. Sensor properties such as spectral response functions, spatial resolution, and modulation transfer functions can be included in the coupled modelling, and the simulation can be directly compared to the measurements (Verhoef and Bach 2003b).

Finally, in the TOA approach, the errors related to both canopy and atmosphere models and their parameter values are contained in the simulation. This makes it easier to carry out error propagation studies to evaluate their impact on the estimates of the vegetation variables. A comparison of pre-processing and model simulation between TOC and TOA approaches is presented in Table 1.1.

1.5.2.2. Improving the regularization

Similarly to the TOC approach, the inversion of the coupled canopy-atmosphere model is ill-posed. The same regularization methods as in the TOC approach (section 1.3) can be used to obtain more stable and reliable solutions.

The coupled model includes all the components that affect the radiation from its entrance in the atmosphere to the sensor's measurement instrument. It is therefore the maximum implementation of the model coupling regularization method.

Table 1.1 Comparison of pre-processing and modelling efforts required in the TOC and TOA approaches, from (Laurent et al. 2011b).

		TOC approach	TOA approach
Observations (Pre-processing)	Radiometric	Radiometric calibration	Radiometric
	Geometric	Geometric correction	Geometric correction
	Terrain	Ortho-rectification	Ortho-rectification
	Correction to TOC reflectance	Atmospheric correction: inversion of atmosphere model using a LUT and assuming Lambertian surface Filter for adjacency effect Correct topography effects Correct BRDF effects	None
	Physical quantity	HCRF approximates HDRF	Radiance
	Errors	Measurement + modelling + assumptions errors + error	Measurement errors
Simulations	Coupled RT models	Soil, Leaf, Canopy	Soil, Leaf, Canopy, Atmosphere
	Effects	Anisotropy Topography Adjacency	Anisotropy Topography Adjacency Atmospheric
	Model output	TOC BRF	TOA radiance
	Errors	Modelling errors	Modelling errors
Comparison	Simulation vs observation	Approximation: BRF ~ HDRF ~ HCRF	Direct

Both the canopy and the atmosphere have an anisotropic behaviour (Schaepman-Strub et al. 2006). The coupled model is therefore better suited to accurately invert multi-angular data than the TOC approach, especially if adjacency, directional and topographic effects would be included in the forward coupled model (Pandya et al. 2007).

As can be seen from Table 1.1, the pre-processing for the TOA approach is reduced to only radiometric calibration, geometric correction and ortho-rectification. Because of this minimal pre-processing, the TOA approach could also facilitate the use of multi-temporal data for regularizing the inversion of the coupled model, whether the time series is acquired by one or several sensors.

Regularization using spatial constraints and a priori information can be applied using the same methods as in the TOC approach. A priori values for the atmospheric variables can be acquired in the same way as is being done for performing the atmospheric correction in the TOC approach (Richter et al. 2006).

1.6 Objectives and research questions

This thesis contributes to improving the accuracy of estimates of bio-chemical and bio-physical vegetation variables from optical spaceborne remote sensing data, focussing on the most general approach, namely the radiometric data-driven inversion of physically-based RT models. The core hypothesis is that the TOA approach allows estimating the vegetation variables more accurately than the TOC approach.

The main objective of this thesis is to explore the potential of the TOA approach for estimating biophysical and biochemical vegetation variables. In order to obtain the most accurate estimates from mono-temporal radiance data, multi-angular information, a priori information (Bayesian approach), and spatial constraints (objects) regularization methods can be integrated with the TOA approach. Therefore, this thesis investigates the following research questions:

A. How accurately can the TOA approach estimate vegetation variables directly from the radiance data measured by a spaceborne sensor?
B. What is the added value of multi-angular radiance data for estimating vegetation variables using the TOA approach?
C. What is the potential of a Bayesian object-based approach for estimating LAI and Cab using a coupled canopy-atmosphere model?
D. What is the applicability of the Bayesian object-based approach for estimating LAI and Cab from a TOA image for which less a priori information is available?

1.7 Outline

This thesis consists of four main chapters, each addressing one of the research questions presented in section 1.6.

Chapter 2 compares the vegetation variables estimated using both TOA and TOC approaches from near-nadir RS data for three Norway spruce stands in the Czech Republic (research question A).

Chapter 3 investigates the additional information content of multi-angular TOA data for estimating vegetation variables by increasing the number of observation directions used for inverting the coupled canopy-atmosphere model, in the same three Czech Norway spruce stands as used in Chapter 2 (research question B).

Chapter 4 introduces a Bayesian object-based approach and investigates its potential for estimating LAI and Cab from at-sensor imaging spectrometer data with high spatial resolution for 246 agricultural fields of eight different crop types in Switzerland (research question C).

Chapter 5 explores the applicability of the Bayesian object-based approach to a TOA image of the Zurich area (Switzerland) where no field data were available (research question D).

This thesis is concluded by Chapter 6, where the findings for each research question are presented and discussed in regard of the core hypothesis. Chapter 6 ends with an outlook and suggestions for further research.

Chapter 2

Estimating forest variables from top-of-atmosphere radiance satellite measurements using coupled radiative transfer models

Abstract

Traditionally, it is necessary to pre-process remote sensing data to obtain top of canopy (TOC) reflectances before applying physically-based model inversion techniques to estimate forest variables. Corrections for atmospheric, adjacency, topography, and surface directional effects are applied sequentially and independently, accumulating errors into the TOC reflectance data, which are then further used in the inversion process. This paper presents a proof of concept for demonstrating the direct use of measured top of atmosphere (TOA) radiance data to estimate forest biophysical and biochemical variables, by using a coupled canopy-atmosphere radiative transfer model. Advantages of this approach are that no atmospheric correction is needed and that atmospheric, adjacency, topography, and surface directional effects can be directly and more accurately included in the forward modelling.

In the case study, we applied both TOC and TOA approaches to three Norway spruce stands in Eastern Czech Republic. We used the SLC soil-leaf-canopy model and the MODTRAN4 atmosphere model. For the TOA approach, the physical coupling between canopy and atmosphere was performed using a generic method based on the 4-stream radiative transfer theory which enables full use of the directional reflectance components provided by SLC. The method uses three runs of the atmosphere model for Lambertian surfaces, and thus avoids running the atmosphere model for each new simulation. We used local sensitivity analysis and singular value decomposition to determine which variables could be estimated, namely: canopy cover, fraction of bark, needle chlorophyll, and dry matter content. TOC and TOA approaches resulted in different sets of estimates, but had comparable performance. The TOC approach, however, was at its best potential because of the flatness and homogeneity of the area. On the contrary, the capacities of the TOA approach would be better exploited in heterogeneous rugged areas. We conclude that, having similar performances, the TOA approach should be preferred in situations where minimizing the pre-processing is important, such as in data assimilation and multi-sensor studies.

16

2.1 Introduction

Forests are important ecosystems on Earth: they cover about 30% of the land surface (FAO 2006), provide a wide range of services and have a major influence on the global climate. Dynamic global vegetation models (DGVM) increasingly require forest biophysical and biochemical variables, such as leaf area index (LAI), fractional vegetation cover (fCover), and chlorophyll content, as inputs. These variables can be estimated and monitored using remotely sensed data (Bacour et al. 2006a; Baret et al. 2007; Myneni et al. 2002).

Traditionally, remote sensing data are pre-processed to obtain top-of-canopy (TOC) reflectance data, which are then used to estimate forest variables. We refer to this approach as the TOC approach.

Several categories of methods can be used to estimate forest variables: (semi-) empirical and physically-based are amongst the most frequently used. Empirical methods rely on statistical correlations between image information and forest variables. They depend on many local characteristics, such as vegetation type, background reflectance, sun-target-sensor geometry, and sensor spectral bands, and are therefore site and time specific (Dorigo et al. 2007; Ustin et al. 2009).

Physically-based methods are more general because they rely on physical relationships (Gemmell et al. 2002; Malenovský et al. 2008). Most of these methods use canopy reflectance models based on the radiative transfer (RT) theory. To estimate forest variables, the canopy RT model has to be inverted. The inversion, however, is not straightforward because it is an ill-posed and ill-conditioned problem due to the limited information content of the radiometric signal (Jacquemoud et al. 2009) and to measurement and model uncertainties (Combal et al. 2002).

The ill-posedness of the inversion can be reduced by using regularization methods: prior information, spatial (Atzberger 2004) and temporal constraints (Lauvernet et al. 2008) allow reducing the variable search space, and model coupling allows decreasing the number of free variables (Baret and Buis 2008; Jacquemoud et al. 2009).

The TOC approach has proven to be successful in the last decades, enabling the production of high level data products (Garrigues et al. 2008; Yang et al. 2006a), which are now being used as inputs for DGVMs. Pre-processing of remote sensing data to TOC reflectances, however, is still a cumbersome task, and each pre-processing step has a number of limitations.

Atmospheric correction often relies on the inversion of an atmospheric RT model, which is a limitation in itself because this inversion is ill-posed. In practice, most operational methods use look up tables (LUT) to perform the inversion, leading to frequent interpolation between LUT entries (Guanter et al. 2009; Richter 2009). In addition, it is often necessary to assume that the surface is Lambertian, which is

generally not true, especially for forests (Bicheron and Leroy 2000). In heterogeneous scenes, adjacency effects are usually dealt with by spatially filtering the image, which is not completely accurate because the measured pixel data are affected by the adjacency effects through the atmospheric point-spread function. Topographic effects are influenced by factors such as sky view factor, surface anisotropy, and lower boundary conditions of the atmosphere, which makes their correction complex (Richter and Schläpfer 2002), especially for forest environments (Soenen et al. 2005). Finally, for images with a large field of view (Schlerf et al. 2005), for image mosaicking (Schaaf et al. 2002) or multi-temporal studies (Bacour et al. 2006b), it may be necessary to correct for surface directional effects, which are described by the bi-directional reflectance distribution function (BRDF). Correcting for BRDF effects is not consistent with the Lambertian surface assumption used in other pre-processing steps.

Usually, corrections for atmospheric, topographic, adjacency, and directional effects are applied sequentially and independently. In addition to error propagation issues, sequential processing does not reflect the physical interactions between these effects (Gao et al. 2009). Because all these effects are inter-related, it is not possible to correct for each effect without making simplifying assumptions about the others or using an explicit physically-based approach.

Following a proper pre-processing sequence, hemispherical-conical reflectance factor (HCRF) (Schaepman-Strub et al. 2006) data can be derived. Based on the assumption that there are no directional effects within the very small instantaneous field of view of satellite sensors, HCRF is commonly used as an approximation of the hemispherical-directional reflectance factor (HDRF). Most studies, however, use the bi-directional reflectance factor (BRF) output of the canopy RT model in the model inversion (Dorigo et al. 2007; Malenovský et al. 2008; Verhoef and Bach 2007; Verrelst et al. 2010c). This mismatch between the two quantities compared, is another inherent limitation of the TOC approach.

Finally, the estimation of forest variables is subject to errors both in the simulations, because of modelling errors and parameter uncertainties, and in the reference TOC reflectance data, because of the error propagation in the pre-processing chain (Rahman 2001).

Because most limitations of the TOC approach arise from the need to correct the remote sensing data from top of atmosphere (TOA) to TOC level, a solution is to use the TOA data directly (Verhoef and Bach 2003a). This requires coupling the canopy and atmosphere models to assess the canopy variables from TOA level. This coupling also enables the inclusion of surface BRDF, adjacency and topography effects in the model. Thus, the forward modelling of the coupled surface-atmosphere system is more accurate than applying a series of corrections as in the TOC approach. In addition, the coupled model simulates the TOA radiance, which is the physical quantity measured by the sensor. Therefore, by including sensor properties such as

spectral bands, spatial and spectral resolution, and modulation transfer functions in the modelling, the simulations can be directly compared with the measurements (Verhoef and Bach 2003b), and the coupled model can be inverted directly against the measured radiance data to estimate the forest variables. Finally, in the TOA approach, all modelling errors and parameter uncertainties are contained in the simulation, which makes it easier to study their impact on the estimates. Table 1.1 presents a comparison of the pre-processing and modelling steps between TOC and TOA approaches.

Coupled canopy-atmosphere models already exist (Börner et al. 2001; Fourty and Baret 1997; Gastellu-Etchegorry et al. 2004; Rahman et al. 1993; Verhoef and Bach 2007). Their ability to simulate sensor-like data has been used for sensor feasibility studies (Schläpfer and Schaepman 2002), for preparing processing chains to test potential product performance of instruments (Hüni 2009), and for investigating the impact of atmospheric parameter uncertainties on the estimation of biophysical and biochemical variables (Gobron et al. 2008; Rahman 2001). Because of its minimal pre-processing, the TOA approach could facilitate data assimilation, multi-temporal and multi-sensor studies, which already proved their high potential to estimate surface variables (Dorigo et al. 2007; Kötz et al. 2005; Lauvernet et al. 2008; Olioso et al. 2005; Verhoef and Bach 2003a). In addition, the estimation of surface variables involving the downwelling radiance such as fraction of absorbed photosynthetically active radiation (fAPAR) and albedo could benefit from the atmosphere coupling implemented in the TOA approach (Pinty et al. 2005; Verhoef and Bach 2007). Verhoef and Bach (2003b) demonstrated the feasibility of simulating TOA radiance images for data assimilation and Verhoef (2007) successfully estimated surface variables from simulated hyperspectral-multidirectional TOA radiance data. Finally, an operational algorithm has recently been developed for producing fCover, LAI and fAPAR products from TOA MERIS reflectance images (Baret et al. 2006), but it assumes a Lambertian surface and ignores adjacency effects.

The present study provides a proof of concept for the TOA approach by testing it in forward and inverse modes, and comparing the results with the TOC approach. The generic canopy-atmosphere coupling set-up allows making full use of the canopy directional reflectances in the framework of the 4-stream RT theory. Seven atmospheric parameters are calculated, avoiding re-running the atmospheric model for each new simulation. The case study focussed on three Norway spruce stands in Eastern Czech Republic for which field and near-nadir CHRIS (Compact High Resolution Imaging Spectrometer) (Cutter et al. 2000) data were available. The simulations were performed using the SLC (soil-leaf-canopy) model (Verhoef and Bach 2007) and the MODTRAN4 atmospheric model (Berk et al. 2003).

2.2 Materials and methods

2.2.1 TOA radiance simulation

The 4-stream theory provides a simple but reasonably accurate framework for physically based RT modelling. The four fluxes considered are: 1) direct solar flux, 2) downward diffuse flux, 3) upward diffuse flux, and 4) direct observed flux. Following the terminology of Verhoef and Bach (2003b), we use ρ for the reflectance of a layer through volume scattering, τ for the transmittance through a layer, and r for the reflectance of a surface. Subscripts are used to indicate the direction of the radiation: o for the direction of observation, s for the direction of the sunrays, and d for diffuse hemispherical radiation. The first subscript is for the incident direction and the second for the exiting direction. For clarity, the spectral dependence of the variables and parameters is omitted in the equations. With these notations, the TOA radiance in the observing direction L_o can be calculated as (Verhoef and Bach 2003b):

$$L_o = \frac{E_s}{\pi} \cos\theta_s \left(\rho_{so} + \frac{\tau_{ss}\overline{r_{sd}} + \tau_{sd}\overline{r_{dd}}}{1 - \overline{r_{dd}}\rho_{dd}} \tau_{do} + \frac{\tau_{sd} + \tau_{ss}\overline{r_{sd}}\rho_{dd}}{1 - \overline{r_{dd}}\rho_{dd}} r_{do}\tau_{oo} + \tau_{ss}r_{so}\tau_{oo} \right) \tag{2.1},$$

where E_s is the extraterrestrial solar irradiance on a plane perpendicular to the sunrays, θ_s is the local solar zenith angle, and the bars indicate low-pass spatial filtering over the environment of the target pixel using the point spread function to account for the adjacency effect. The quantity ρ_{dd} is the bi-hemispherical reflectance (BHR) of the bottom of the atmosphere for upwelling diffuse radiation. It is sometimes called 'spherical albedo' of the atmosphere, but, unlike the term 'albedo' might suggest, it is a spectral quantity, not a spectrally integrated one. Equation 2.1 can be simplified using the atmospheric path radiance L_{atm} and the atmospheric gain factors G as defined below (Bach and Verhoef 2009):

$$L_o = \underbrace{L_{atm} + \frac{G_{ssdo}\overline{r_{sd}} + G_{sddo}\overline{r_{dd}}}{1 - \overline{r_{dd}}\rho_{dd}}}_{\text{PATH}} + \underbrace{\frac{G_{sdoo} + G_{mult}\overline{r_{sd}}}{1 - \overline{r_{dd}}\rho_{dd}} r_{do} + \underbrace{G_{ssoo}r_{so}}_{\text{GSUN}}}_{\text{GRFL}} \tag{2.2},$$

with $\quad L_{atm} = \dfrac{E_s \cos\theta_s}{\pi} \rho_{so}$, $\qquad\qquad$ $G_{ssoo} = \dfrac{E_s \cos\theta_s}{\pi} \tau_{ss}\tau_{oo}$,

$$G_{ssdo} = \frac{E_s^o \cos\theta_s}{\pi} \tau_{ss}\tau_{do}, \qquad\qquad G_{sddo} = \frac{E_s \cos\theta_s}{\pi} \tau_{sd}\tau_{do},$$

$$G_{sdoo} = \frac{E_s \cos\theta_s}{\pi} \tau_{sd}\tau_{oo}, \text{ and} \qquad G_{mult} = \frac{E_s \cos\theta_s}{\pi} \tau_{ss}\rho_{dd}\tau_{oo}.$$

Equation 2.2 shows that surface-atmosphere coupling can be achieved using four surface reflectance components and seven atmospheric parameters. The surface reflectance components can be provided by any 4-stream surface RT model, and the atmospheric parameters can be calculated using any atmospheric model that provides the total path radiance (PATH), the sunlight ground-reflected radiance (GSUN) and the total ground-reflected radiance (GRFL) (see annotations on Equation 2.2). For a given atmosphere, three model runs for Lambertian surfaces of different albedo are needed. The seven atmospheric parameters can then be calculated from the nine equations obtained for PATH, GSUN, and GRFL for the three runs, for example for albedo values 0, 0.5, and 1. The gain factors contain products of spectrally variable parameters, which co-vary strongly over the spectral interval and should therefore be calculated by averaging the parameter products and not by calculating the products of the averaged parameters. The obtained atmospheric parameters can be used to physically couple any surface, including non-Lambertian, to the atmosphere. Indeed, the directional effects of the surface, as described by the four reflectance components, are fully used in Equation 2.2. This method is efficient because the atmospheric parameters can be stored on hard disk, avoiding re-running the atmosphere model for every new surface. This approach based on gain factors was developed to solve problems with averaging over finite spectral intervals (Bach and Verhoef 2009).

2.2.2 Atmospheric correction

The aim of the atmospheric correction is to calculate the TOC reflectance of the target, given the TOA radiance L_o and the atmospheric parameters. There are four unknown reflectance components in Equation 2.2, which causes the atmospheric correction to be an underdetermined inverse problem. To decrease the number of unknowns, it is necessary to assume that both the target and its environment are Lambertian, with reflectances r_t and r_e. Equation 2.2 then becomes:

$$L_o = L_{atm} + \frac{G_t r_t + G_e r_e}{1 - r_e \rho_{dd}} \qquad\qquad (2.3),$$

with $G_t = G_{ssoo} + G_{sdoo}$ and $G_e = G_{ssdo} + G_{sddo}$. A second equation is obtained by exploiting the low-pass filtered radiance $\overline{L_o}$, and r_t can be calculated as:

$$r_t = \frac{L_o - L_{atm} + \frac{G_e}{G_t}\left(L_o - \overline{L_o}\right)}{G_t + G_e + \rho_{dd}\left(\overline{L_o} - L_{atm}\right)} \qquad (2.4).$$

2.2.3 Cost function

To allow comparing the results obtained from the TOC and TOA approaches, which have different units, the relative results $r_{rel} = r / r_{ref}$ were used, where r is either r_{so} at TOC level, or L_o at TOA level, and the subscript *ref* refers to the reference signature. In addition, the use of relative results avoids giving more importance to the bands where the reflectance of radiance values are higher (Weiss et al. 2000). The cost function χ, based on the root mean square difference (RMSD) between the simulation and the reference signature, was thus defined as:

$$\chi = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\Delta r_{rel}(\lambda_i)\right]^2} \qquad (2.5),$$

where λ is the wavelength, and n is the number of spectral bands. A χ value of 0 corresponds to a perfect match ($r = r_{ref}$), and a χ value of 1 corresponds to an error of 100% (e.g. $r = 2r_{ref}$, or $r = 0$).

2.2.4 Local sensitivity analysis and dimensionality

The local sensitivity analysis (LSA) relies on the Jacobian matrix **J**. **J** is the matrix of the partial derivatives of the relative model output r_{rel} with respect to each input parameter p, normalized assuming a uniform distribution over its potential variation range:

$$\mathbf{J} = \lfloor j_{i,k} \rfloor_{1 \le i \le n,\, 1 \le k \le m}, \text{ with } j_{i,k} = \frac{\partial r_{rel}(\lambda_i)}{\partial p_k} \qquad (2.6),$$

where m is the number of parameters. Each parameter was varied by 1% of its potential variation range, and the main atmospheric parameters were included in the TOA analyses. To evaluate the influence of parameter k on the cost function, the indicator α_k was defined using the same structure as χ:

$$\alpha_k = \sqrt{\frac{1}{n} \sum_{i=1}^{n} j_{i,k}^2} \qquad (2.7),$$

To allow comparing the parameter influences between TOC and TOA level, the α values were normalized (α_{norm}). Only the most influential parameters having high α_{norm} values can be estimated.

The maximum number of parameters that can be estimated based on the model and data information content is the dimensionality of the estimation problem. The dimensionality can be evaluated using the singular value decomposition (SVD) of $\mathbf{J}$ (Verhoef 2007):

$$\mathbf{J} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} \qquad (2.8),$$

where $\mathbf{S}$ is a diagonal matrix containing the singular values, the superscript T is the transpose operator, and $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices. In particular, $\mathbf{U}$ and $\mathbf{V}$ are orthogonal, so: $\mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}$ and $\mathbf{V}\mathbf{V}^{\mathrm{T}} = \mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$. Because $\mathbf{J}$ relates the vector of normalized parameter variation $\Delta\mathbf{p}$ to the vector of relative model output difference $\Delta\mathbf{r}_{rel}$ as $\Delta\mathbf{r}_{rel} = \mathbf{J}\Delta\mathbf{p}$, one can infer:

$$\mathbf{U}^{\mathrm{T}}\Delta\mathbf{r}_{rel} = \mathbf{S}\mathbf{V}^{\mathrm{T}}\Delta\mathbf{p} \qquad (2.9).$$

Because $\mathbf{S}$ is diagonal, Equation 2.9 shows that there is a one-to-one relationship between the transformed model output differences $\mathbf{U}^{\mathrm{T}}\Delta\mathbf{r}_{rel}$ and the transformed parameter variations $\mathbf{V}^{\mathrm{T}}\Delta\mathbf{p}$. Null singular values indicate linear combinations of parameter changes which have no effect on the spectral output. Therefore, the rank of $\mathbf{S}$ gives the dimensionality of the estimation problem.

2.2.5 Radiative transfer models

The soil-leaf-canopy (SLC) (Verhoef and Bach 2007) model was chosen because it is a hybrid model and thus offers a good trade-off between realistic representation of the canopy and computational effort (Gemmell et al. 2002; Pinty and Verstraete 1992). It simulates the four reflectance components of the canopy. A description of the main features of SLC is given here. For more details, the reader is referred to Verhoef and Bach (2007). SLC couples three models for the soil, leaf and canopy.

The soil BRF model is called 4SOIL. It is an extended version of the Hapke model (Hapke 1981) which includes hotspot and spectral soil moisture effect. It was not used in this study because two of the three stands had vegetated understory and the third stand had a very high crown cover.

The leaf reflectance model is a robust version of PROSPECT (Jacquemoud and Baret 1990), modified to include brown pigments (Verhoef and Bach 2003b). The

specific absorption coefficients of chlorophyll, dry matter and water and the refractive index of leaf material at 1 nm resolution were taken from PROSPECT-4 (Feret et al. 2008).

The canopy reflectance model is 4SAIL2 (Verhoef and Bach 2007), an advanced 2-layer version of the SAIL model (Verhoef 1984) which includes the crown clumping effect thanks to the introduction of two additional inputs: crown cover (Cv) and tree shape factor (Zeta). Zeta is defined as the crown diameter divided by the height of the crown centre above ground. The 2-layer feature allows mixing green and brown elements in the canopy using the fraction of brown plant area (fB) and the dissociation factor D. For D = 0, the brown and green elements are homogeneously mixed in the canopy, and for D = 1, all the green elements are in the top layer and all the brown elements are in the bottom layer. The brown elements can be used to simulate the bark and other woody material. The crown clumping and the 2-layer set-up with brown elements enable a better representation of forest scenes.

MODTRAN4 (Berk et al. 2003) describes the radiative transfer in the atmosphere, and simulates, among other quantities, PATH, GSUN and GRFL. It is a state-of-the-art model and is commonly used in conjunction with atmospheric correction software.

2.2.6 Variable estimation

The LUT method was chosen for estimating the variables because of its capacity of finding the global minimum of the cost function χ. The results of the LSA were used to select only the most influential variables to keep free in the LUTs (Table 2.2). The models were run for each combination of variables values. The r_{so} signatures were stored in the TOC LUT and the L_o signatures in the TOA LUT.

2.3 Case study

2.3.1 Study area

The study area is located in Eastern Czech Republic, at the Bily Kriz experimental research site in the Moravian-Silesian Beskydy Mountains, (18.54°E, 49.50°N; altitude 936 m above sea level). A detailed description of the environmental conditions can be found in (Kratochvilová et al. 1989). The forest area is dominated by montane Norway spruce (*Picea abies* (L.) Karst.). Three stands of different ages and structures were selected (Table 2.1): YOUNG, OLD1 and OLD2. The data were collected in the first half of September 2006.

Table 2.1 Stand characteristics and model inputs.

Stand		YOUNG	OLD1	OLD2
Age (years)		29	100	75
Density (trees/ha)		1450	160	420
DBH (cm)		14	53	37
# CHRIS pixels		4	13	10
Canopy	PAI (-)	8.88	5.73	7.35
	fB (-)	0.13	0.23	0.4
	D (-)	0.4	0.4	0.1
	Hot (-)	0.01	0.01	0.01
	LIDF (-)	Spherical	Spherical	Spherical
	Cv (-)	0.9	0.55	0.7
	Zeta (-)	0.34	0.24	0.26
Needle	Cab (μg/cm^2)	55	60	65
	Cw (cm)	0.02	0.02	0.02
	Cdm (g/cm^2)	0.04	0.04	0.04
	Cs (-)	0	0	0
	N (-)	1.8	1.7	1.7
Bark	Cab (μg/cm^2)		10	
	Cw (cm)		0	
	Cdm (g/cm^2)		0.5	
	Cs (-)		15	
	N (-)		10	

2.3.2 Remote sensing data

The remote sensing data was a near-nadir image acquired on September 12th, 2006, by CHRIS/PROBA. The image was acquired in chlorophyll mode (mode 4), resulting in 18 spectral bands in the range 485-802 nm at a spatial resolution of 17 m. The viewing zenith and azimuth angles were respectively 16° and 138.1° and the solar zenith and azimuth angles were 46.6° and 162.7°. The CHRIS image was radiometrically calibrated by the data provider and was further de-striped, geo-corrected and ortho-rectified using nearest neighbour interpolation. Band 15, centred at 761 nm, was not used because it sampled one of the oxygen absorption features and was noisy.

An AISA (Airborne Imaging Spectro-radiometer for Applications) Eagle image was acquired on September 14th, 2006. It had 40 spectral bands in the range 450-830 nm and a spatial resolution of 40 cm. The AISA image was radiometrically calibrated, geo-corrected, ortho-rectified, and atmospherically corrected using the ATCOR4 software.

2.3.3 Field data

The plant area index (PAI), defined as half of the total plant area (needles and non photosynthetic plant material) per unit of ground surface area (Chen 1996), was estimated in each stand by three methods: LAI-2000 plant canopy analyzer, hemispherical photograph, and TRAC (Tracing Radiation and Architecture of

Canopies) (2007). The obtained values were averaged to obtain one PAI value for each stand. The vertical crown cover (Cv) was estimated by classifying the AISA image (Lukeš 2009).

Ten sample trees in the YOUNG stand and 20 in the OLD1 stand were selected for canopy and needle measurements. Only canopy measurements were made in the OLD2 stand. Canopy structure measurements included tree height, crown radius, and length of live and dead crown, all measured with a laser rangefinder Impulse 200.

The spectral properties of the main soil components, understory species and bark were measured in the field at 1 nm resolution using an ASD FieldSpec Pro spectroradiometer.

2.3.4 Models parameterization

The background signatures were calculated as the weighted average of the main soil and understory components. The background of the YOUNG stand was a mixture of litter, humus, and soil, whereas the background of the OLD1 and OLD2 stands consisted of a majority of blueberry, blackberry, and grass. The backgrounds were considered Lambertian because of the lack of directional data and of the high canopy cover.

Because the default PROSPECT model is not applicable to needles (Malenovský et al. 2006), the input parameters were manually tuned to fit the measured optical properties rather than using chemistry measurements. For the bark, the PROSPECT parameters were optimized to match the measured signature. For the canopy, the leaf inclination distribution function (LIDF) was chosen spherical (LIDFa = -0.35, and LIDFb = -0.15), and the dissociation factor D was calibrated manually. The values used as inputs in SLC for the three stands are presented in Table 2.1.

In MODTRAN4, the urban aerosol type was chosen because there is an urban-industrial zone located 20 km north of the study area and north wind is dominant in the area. In addition, high air concentrations of SO_2 were recorded at the study area on the day of AISA acquisition (Marek et al. 2007). The visibility (vis) was chosen as the smallest value allowing all radiances in the CHRIS image to be higher than the simulated atmospheric path radiance L_{atm}. That value was 80 km. The default ozone (O_3 = 7.05 g/m^2) and water vapour (H_2O = 1.835834 g/cm^2) columns were used. Following Guanter et al. (2009), it was chosen to use the DISORT algorithm with 8 'streams', the correlated-k option with 17 values and the 5 cm^{-1} MODTRAN4 band model.

The adjacency effect was neglected because the scene was mostly covered and surrounded by forests, thus simplifying Equation 2.2 to:

$$L_o = L_{atm} + \frac{G_{ssdo}r_{sd} + G_{sddo}r_{dd}}{1 - r_{dd}\rho_{dd}} + \frac{G_{sdoo} + G_{mult}r_{sd}}{1 - r_{dd}\rho_{dd}}r_{do} + G_{ssoo}r_{so} \quad (2.10),$$

and Equation 2.4 to:

$$r_t = \frac{L_o - L_{atm}}{G_t + G_e + \rho_{dd}(L_o - L_{atm})}$$ (2.11).

2.3.5 Adjustments for CHRIS data

The four canopy reflectance components and the seven atmospheric parameters were resampled to the CHRIS bands using Gaussian approximations of the sensor response functions before feeding them into Equations 10 and 11. Because the CHRIS spectral bands are not equidistant, weights were used in the cost function:

$$\begin{cases} w_1 = (\lambda_2 - \lambda_1) \\ w_i = (\lambda_{i+1} - \lambda_{i-1})/2, \text{ for } 2 \le i \le n-1 \\ w_n = (\lambda_n - \lambda_{n-1}) \end{cases}$$ (2.12),

so the cost function became:

$$\chi = \sqrt{\frac{\sum_{i=1}^{n} w_i (\Delta r_{rel}(\lambda_i))^2}{\sum_{i=1}^{n} w_i}}$$ (2.13),

and the influence factor:

$$\alpha_k = \sqrt{\frac{\sum_{i=1}^{n} w_i j_{i,k}^2}{\sum_{i=1}^{n} w_i}}$$ (2.14).

The near-nadir image used in this study was acquired in the backward direction, not far from the principal plane (relative azimuth angle = 25°), causing the hotspot parameter (hot) to be considerably influential. Since the value of 0.01 is well known for forests, it was decided to exclude it from the sensitivity analysis to obtain more general results.
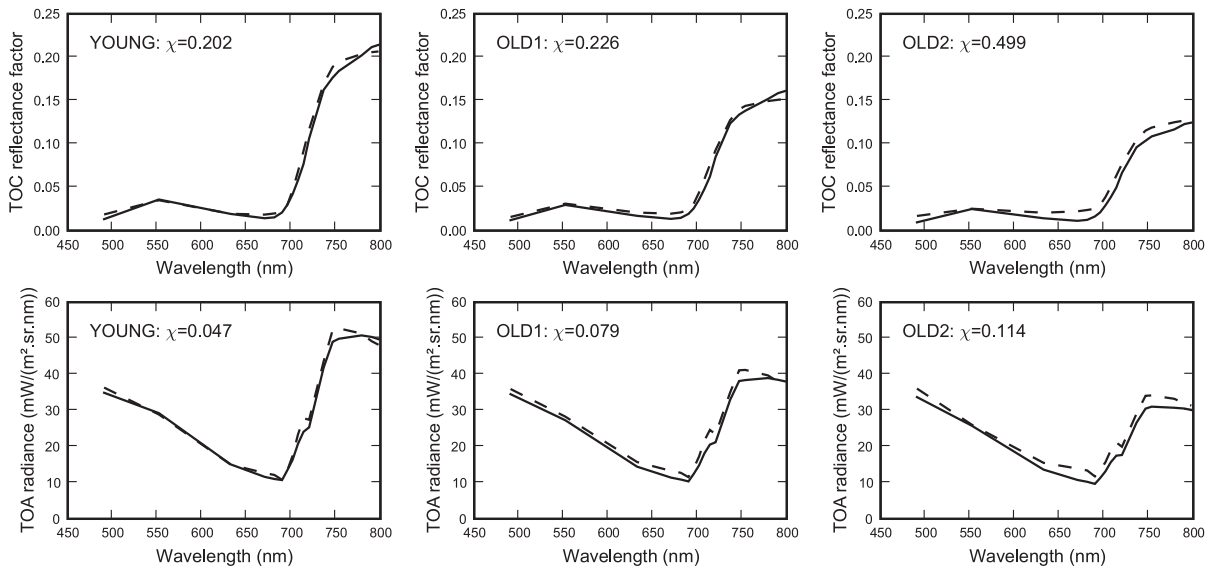
2.4 Results

2.4.1 Simulations

A good match was obtained between the PROSPECT simulations and the measured needle signatures for the YOUNG and OLD1 stands (χ(YOUNG) = 0.07 and χ(OLD1) = 0.08). No field measurements were available for the OLD2 stand, but the simulated reflectance and transmittance were similar to those of the OLD1 stand. To obtain low simulations similar to the measured optical properties, it was necessary to have high Cab and Cdm values (Table 2.1).

The optical properties of the bark could also be well simulated by PROSPECT (χ = 0.11). As expected, the best agreement was obtained for high values of N, Cdm, and Cs, low value of Cab, and Cw = 0 (Table 2.1). The simulated transmittance was null, as required for bark material. A small discrepancy persisted in the blue band.

The TOC and TOA simulations, obtained using the values in Table 2.1, are presented in Figure 2.1. Good fits were obtained, but small discrepancies between simulated and measured data appeared in the blue, red and NIR shoulder, with a tendency to overestimation. The χ values were smaller at TOA than at TOC level, indicating a better fit at TOA level. At both TOC and TOA level, the χ values decreased from the YOUNG to the OLD1 to the OLD2 stand.
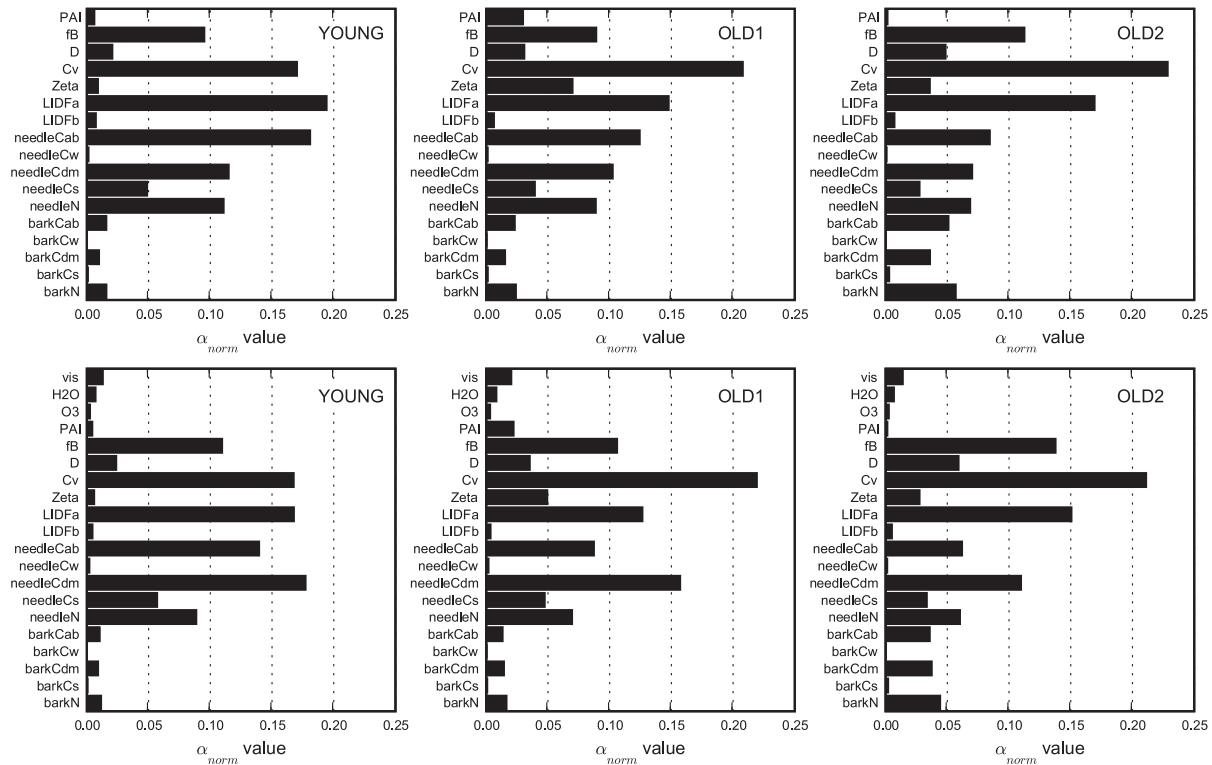

Figure 2.1 Comparison of simulation outputs (dashed lines) with CHRIS data (plain lines). Top row: SLC r_{so} simulation and CHRIS atmospherically corrected TOC reflectance. Bottom row: coupled model L_o simulation and CHRIS TOA radiance.
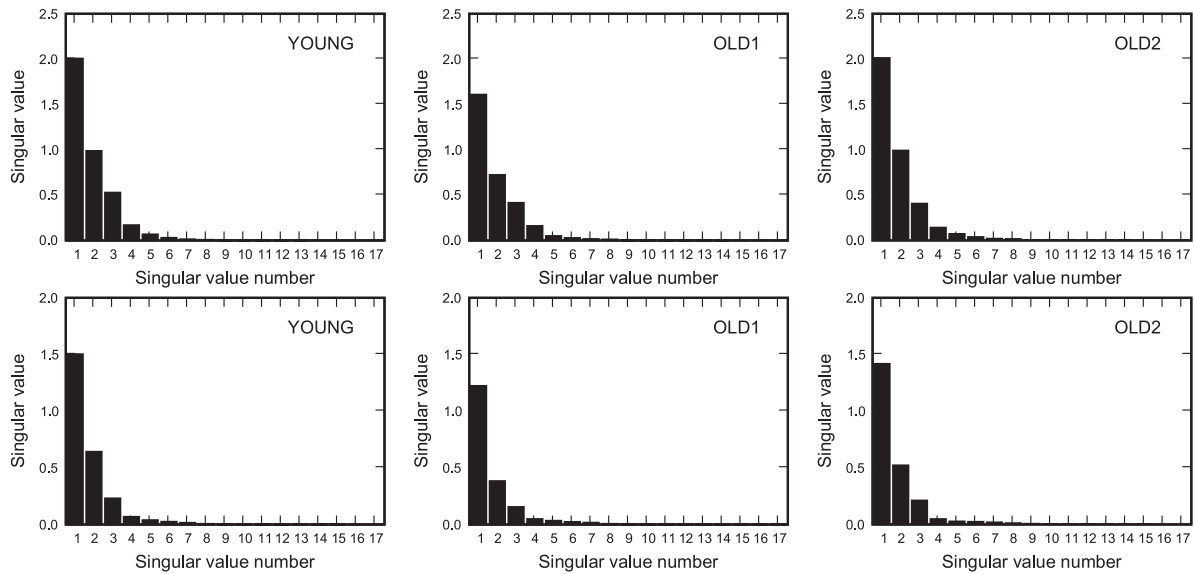
2.4.2 Local sensitivity analyses and dimensionality

The α$_{norm}$ values, which indicate the parameter influences, are presented in Figure 2.2. At TOC level, the most influential parameters for the YOUNG stand were LIDFa, Cv, needleCab, needleCdm, needleN and fB; they were well isolated with α$_{norm}$ values larger than 0.1. For the OLD2 stand, Cv was the most influential (α$_{norm}$ = 0.36) and many other parameters had secondary influence with α$_{norm}$ values in the range 0.05 to 0.2, including LIDFa, fB, needleCab, needleCdm, needleN, similarly to YOUNG, and additionally D, Zeta, barkCab, barkCdm, and barkN. The OLD1 stand had an intermediary situation, with Cv as the most influential parameter and only Zeta as additional secondary parameter.

The least influential parameters were LAI, needleCw, bark, and atmospheric parameters. Although fB was very influential, the PROSPECT bark parameters, and thus the bark signature, were not important. Grouping the parameters in categories, the canopy parameters were more influential than the needle parameters, which were more influential than the bark parameters, than the atmospheric parameters.

These tendencies were similar at TOA level. Another difference was that while needleCab was more influential than needleCdm at TOC level, it was the other way around at TOA level. Overall, there were six most influential parameters for the three

Figure 2.2 Parameter influences (α$_{norm}$ values). Top row: TOC level. Bottom row: TOA level.

Figure 2.3 Singular values of the Jacobian matrix. Top row: TOC level. Bottom row: TOA level.

stands for both TOC and TOA level: Cv, fB, LIDFa, needleCab, needleCdm, and needleN.

The singular values are plotted in Figure 2.3. Seventeen singular values were obtained, as the number of bands was limiting (17 bands versus 17 parameters at TOC level and 20 at TOA level). The singular values were slightly smaller at TOA than at TOC. In all cases, the rank of S was considered to be 3, giving a dimensionality of the estimation problem of 3.

2.4.3 Variable estimations

Based on the results of the local sensitivity analyses, a maximum of three variables could theoretically be estimated whereas there were six influential variables. Since it was not possible to discriminate three most influential variables among Cv, fB, LIDFa, needleCab, needleCdm, and needleN, only the variables most relevant for applications (forest health, fuel moisture, carbon stock…) were kept free. These were: Cv, fB, needleCab and needleCdm. Each of these four variables was sampled from a uniform distribution using regular steps (Table 2.2). The LUTs were built by running the SLC and the SLC-MODTRAN4 model according to the CHRIS spectral bands for all 27,951 combinations of the four variables, keeping the other parameters fixed at the values used in forward modelling (Table 2.1).

Table 2.2 Parameter ranges and intervals used for the generation of the LUT.

Parameter	Min	Max	Step	#values
fB (-)	0	1	0.1	11
Cv (-)	0	1	0.1	11
NeedleCab (μg/cm^2)	0	100	5	21
NeedleCdm (g/cm^2)	0	0.05	0.005	11

Table 2.3 Estimates of Cv, fB, needleCab, and needleCdm obtained for the 3 stands at TOC and TOA level.

Stand	YOUNG		OLD1		OLD2	
Level	TOC	TOA	TOC	TOA	TOC	TOA
Cv (-)	0.8	0.8	0.6	0.7	0.7	0.7
fB (-)	0	0	0	0.2	0	0.3
NeedleCab (μg/cm^2)	55	55	70	70	85	100
NeedleCdm (g/cm^2)	0.04	0.04	0.05	0.05	0.05	0.05
χ (-)	0.143	0.041	0.139	0.043	0.333	0.103

The estimates were selected from the LUTs by finding the variable combination having the minimum χ value, calculated using all CHRIS bands, except band 15. The estimates and their associated χ values are presented in Table 2.3. Similarly to the simulations, the obtained χ values were smaller at TOA than at TOC level. The two approaches gave the same values for the estimates of all four variables for the YOUNG stand. Comparing the estimates with the values used in the forward modelling (Table 2.1), both approaches were able to estimate Cv reasonably well, considering its very high values in the stands. The TOC approach gave better results than the TOA approach for the Cv in the OLD1 stand and the needleCab in the OLD2 stands. The needleCab and needleCdm were overestimated by both approaches, even reaching the upper boundary for the needleCdm estimates for the OLD1 and OLD2 stands, whereas fB was better estimated by the TOA approach in the OLD1 and OLD2 stands. Overall, neither approach outperformed the other for all variable estimations.

2.5 Discussion

2.5.1 Models and simulations

Although PROSPECT was developed to simulate the optical properties of broad leaves (Jacquemoud and Baret 1990), it could be used to simulate needles with adequate accuracy, but at the expense of the physical meaning of the parameter values. This may be avoided by recalibrating PROSPECT (Malenovský et al. 2006) or by using the LIBERTY model (Dawson et al. 1998) which was designed for needles. The presence of brown pigments in this PROSPECT version allowed obtaining a realistic simulation of the bark material. The small discrepancy in the blue band may be due to missing chemical components in PROSPECT. For the purpose of this proof of concept, however, whether the parameter values reflected the truth was of secondary importance. The goal was to generate realistic representations of the canopy elements and then to check how well the estimates would match the original values.

Coniferous forests are one of the most complex medium from a radiative transfer point of view because of the strong hierarchy in the arrangement of the canopy components (Cudlín et al. 2001). The clumping of the needles into shoots, branches and crowns influences the multiple scattering in the canopy by trapping photons, thus decreasing the reflectance of the canopy (Stenberg 2007). Although 4SAIL2 is a simple 2-layer model which describes the canopy as a turbid medium clumped only into crowns, it provided good simulations. The slight tendency to oversimulating the reflectance might be due to the absence of branch and shoot level clumping.

For comparison with other studies, the RMSD values at TOC were 0.007 for the YOUNG and OLD1 stands, and 0.009 for the OLD2 stand. These values are comparable with values obtained for forests in other studies using various models, including DART simulations in the blue, green, red and one NIR band for the same YOUNG stand (Malenovský et al. 2008).

Most studies that included simulations of TOA radiance did not compare the simulations with actual remote sensing data, so it was not possible to compare our TOA results with other studies. MODTRAN4, however, is a state-of-the-art atmosphere model, so it was not surprising that the good simulations at TOC level translated into good simulations at TOA level. The visibility had to be very high (80 km) so that all radiances in the CHRIS data would be higher than the atmospheric path radiance. The very low CHRIS radiances may also be a factor explaining the oversimulation trend at TOA and TOC level.

The smaller χ values obtained at TOA level than at TOC level are due to the atmospheric path radiance. L_{atm} is a constant which is added to the contribution of the surface, lifting up the signal and thus reducing the relative difference between signatures. An illustration can be found in Gerstl and Zardecki (1985). At TOC, the reflectance values for forests are very low, so that a small absolute difference in the simulation translates into a large relative difference. This is why a χ value as high as 0.5 was obtained for the OLD2 stand.

In addition, the higher χ values obtained for the OLD2 stand than for the other stands at both TOC and TOA level may be explained by the lower quality of the field data, as no needle measurements were available for this stand. The fact that the results were better for the YOUNG than for the OLD1 stand might be due to the very high Cv and tree density in the YOUNG stand, as this situation is closest to the turbid medium assumption used in 4SAIL2.

2.5.2 Parameter influences and dimensionality

The local sensitivity analyses provided useful insights about the parameter influences. The most striking result was that the three stands presented similar influence profiles at TOC and TOA level. The influential parameters at TOC level were also influential at TOA level. The small influence of the visibility may be due to its very high value of 80 km. In such clear atmosphere small changes in visibility

do not affect the atmosphere transmittance much. The small influence of the atmospheric parameters means that small inaccuracies in the atmospheric data do not affect the simulations and do not hamper the estimation of the forest variables. This small atmospheric influence, however, might only hold for very clear atmospheric conditions and near-nadir viewing.

The PAI was also found to have little influence. This could be expected, considering that the PAI consists mostly of needles (proportion 1 - fB) and is thus closely related to the LAI, which is well-known for saturating the reflectance for high values (Baret and Guyot 1991; Clevers and Verhoef 1993).

The large influence of the needle parameters is in accordance with the fact that needles represent most of the canopy material. The only exception was needleCw which was not influential because the water absorption features start progressively around 750 nm and therefore weakly affected only a small part of the simulated signature.

The large influence of fB is in agreement with other studies (Malenovský et al. 2008; Verrelst et al. 2010c). In addition, the local sensitivity analyses revealed that the bark signature is far less important than fB.

Cv was an influential parameter in all cases, regardless of background type. Its α_{norm} values at TOC and TOA level, however, were higher in the OLD1 and OLD2 stands, where the backgrounds were vegetated, than in the YOUNG stand, where the background was not vegetated.

The LIDFa parameter controls the average leaf slope. Although its influence was high, for a coniferous forest, one can hardly assume anything else but a random orientation of the needles, which corresponds to a spherical leaf angle distribution and therefore to a LIDFa of -0.35.

The SVD showed that the dimensionality of the estimation problem was 3 for both TOC and TOA level, meaning that both approaches have similar potential for estimating biophysical and biochemical forest variables, despite the smaller relative differences between simulation and reference data at TOA level. The dimensionality of 3 is small compared to the number of model parameters, but corresponds to a common dimensionality usually observed from remote sensing data over coniferous forests. The small dimensionality was due to correlation and parameter compensation effects in the simulations. These cause inversion instabilities and thus ill-posedness (Jacquemoud et al. 2009).

2.5.3 Variable estimations

The results obtained for the three stands of this study agree with the conclusion of Verrelst et al. (2010c) that the higher the Cv and the lower the fB, the better the needleCab estimates. This could also be seen in the LSA where the α_{norm} value for needleCab decreased from the YOUNG to the OLD1 to the OLD2 stand. This study

also extends the results of Verhoef (2007), who had demonstrated the possibility to estimate variables from simulated TOA radiance data, to measured data.

One can note that the quality of the estimates was not directly related to their influence. For example, needleCdm had higher α_{norm} values at TOA than at TOC, so one would expect to obtain better estimates from TOA level, but the two approaches gave the same value for each stand. This might be due to the limited number of entries in the LUTs.

2.5.4 Performance comparison of the TOC and TOA approaches

Overall, the performances of the TOC and TOA approaches were comparable. Similar performance is already sufficient to prefer the TOA approach over the TOC approach for applications, such as data assimilation and multi-sensor studies, where minimal pre-processing is advantageous.

The stands used in this study, like all coniferous forest environments, strongly depart from Lambertian surfaces (Deering et al. 1994). However, near nadir viewing combined with a rather flat and homogeneous study area where adjacency and topography effects could be ignored were good conditions for the TOC approach. Indeed, under these conditions one has least limitations in the atmospheric correction and can obtain good TOC reference data. The potential of the TOA approach would be better exploited in rugged terrain with heterogeneous land cover at off-nadir angles, where the TOC approach is most limited. In addition, the TOC approach benefited from the atmospheric correction method used in this study which used the exact geometry in the MODTRAN4 runs. The TOC reference data was therefore more accurate than that provided by atmospheric correction softwares, which would have interpolated the atmospheric parameters from their pre-computed LUT.

The MODTRAN4 runs are computing-intensive, but, when using the method of three Lambertian runs, the seven atmospheric parameters can be stored and MODTRAN4 does not need to be re-run for every new simulation. One could even build a LUT of the atmospheric parameters in advance and would still benefit from the fully integrated forward modelling with full inclusion of the surface directionality effects and possibility to include adjacency and topography effects in the forward set-up, although it would introduce a need for interpolation (Bach and Verhoef 2009).

The TOA approach reduces the pre-processing errors and improves the modelling of the surface-atmosphere radiative interactions. Regarding the ill-posedness of the variable estimation problem, however, one can use the same regularization methods as in the TOC approach: prior information and spatio-temporal constraints.

2.6 Conclusion

The SLC model was successfully applied to simulate the TOC reflectance, and the SLC-MODTRAN4 coupled model to simulate the TOA radiance of three Norway

spruce stands as measured by CHRIS in near-nadir direction. Based on the local sensitivity analyses and SVDs, four variables were estimated: Cv, fB, needleCab, and needleCdm. The TOC and TOA approaches showed similar performances for estimating forest biophysical and biochemical variables. The TOC approach was at its best potential because of the near-nadir viewing and the flat and homogeneous area, which are good conditions for the atmospheric correction. The surface-atmosphere coupling implemented in the TOA approach already allows full use of the simulated surface directional properties, but it could be further improved by including adjacency and topography effects in the forward modelling. It would be interesting to test the full capacities of the TOA approach in a heterogeneous rugged area viewed from an off-nadir direction. The minimum pre-processing requirements of the TOA approach would benefit data assimilation and multi-sensor applications. Further research will extend the testing of the TOA approach to the multi-angular dataset acquired by CHRIS.

Acknowledgements

Chapter 3

Inversion of a coupled canopy-atmosphere model using multi-angular top-of-atmosphere radiance data: a forest case study

Abstract

Since the launch of sensors with angular observation capabilities, such as CHRIS and MISR, the additional potential of multi-angular observations for vegetation structural and biochemical variables has been widely recognized. Various methods have been successfully implemented to estimate forest biochemical and biophysical variables from atmospherically-corrected multi-angular data, but the use of physically based radiative transfer (RT) models is still limited. Because both canopy and atmosphere have an anisotropic behaviour, it is important to understand the multi-angular signal measured by the sensor at the top of the atmosphere (TOA). Coupled canopy-atmosphere RT models allow linking surface variables directly to the TOA radiance measured by the sensor and are therefore very interesting tools to use for estimating forest variables from multi-angular data.

We investigated the potential of TOA multi-angular radiance data for estimating forest variables by inverting a coupled canopy-atmosphere physical RT model. The case study focussed on three Norway spruce stands located at the Bily Kriz experimental site (Czech Republic), for which multi-angular CHRIS and field data were acquired in September 2006. The soil-leaf-canopy RT model SLC and the atmospheric model MODTRAN4 were coupled using a method allowing to make full use of the four canopy angular reflectance components provided by SLC. The TOA radiance simulations were in good agreement with the spectral and angular signatures measured by CHRIS. Singular value decompositions of the Jacobian matrices showed that the dimensionality of the variable estimation problem increased from 3 to 6 when increasing the number of observation angles from 1 to 4. The model inversion was conducted for two cases: 4 and 7 variables. The most influential parameters were chosen as free variables in the look-up tables, namely: vertical crown cover (Cv), fraction of bark material (fB), needle chlorophyll content (needleCab), needle dry matter content (needleCdm) for the 4-variable case, and additionally, tree shape factor (Zeta), dissociation factor (D), and needle brown pigments content (needleCs) in the 7-variable case. All angular combinations were tested, and the best estimates were obtained with combinations using two or three angles, depending on the number of variables and on the stand used. Overall, this case study showed that, although making use of its full potential is still a challenge, TOA multi-angular radiance data do have a higher potential for variable estimation than mono-angular data.

3.1 Introduction

Forest environments cover about 30% of the Earth surface (FAO 2006) and play an important role in the carbon and water cycles. Projection scenarios based on dynamic global vegetation models can therefore benefit from accurate information about forest variables such as leaf area index (LAI), chlorophyll content and canopy cover. These variables can be efficiently monitored using satellite data, which provide regular and spatially continuous coverage.

The reflectance of most land surfaces, including forests, depends on the acquisition (illumination-target-observation) geometry. This anisotropy is often regarded as an undesired effect which needs to be corrected before using the image (Bacour et al. 2006b; Schaaf et al. 2002). Another view is that the reflectance anisotropy contains additional information about the target, and that consequently multi-angular data has more potential for estimating surface variables than mono-angular data (Asner et al. 1998; Diner et al. 1999; Liang et al. 2000; Schaepman 2007). In the last decades, several space-borne instruments have been launched to sample the radiation field of land surfaces in both spectral and angular dimensions, in a near-simultaneous way (e.g. CHRIS, MISR, POLDER). For forest applications, multi-angular data has been used for monitoring foliage condition (Hilker et al. 2009; Hilker et al. 2011), separating the understory from the overstory contribution (Rautiainen et al. 2008), and estimating canopy biochemical (Huber et al. 2010; Kneubühler et al. 2008) and structural variables (Chen et al. 2003; Chopping et al. 2008; Heiskanen 2006; Verrelst et al. 2010a; Verrelst et al. 2010b; Widlowski et al. 2004). In addition, algorithms have been developed to jointly retrieve structure (LAI) and biophysical variables (fraction of absorbed photosynthetically absorbed radiation) from multi-angular data (Gobron et al. 2002; Knyazikhin et al. 1998).

Many methods have been tested to make full use of the information contained in the multi-angular data to estimate surface variables. The angular dependency of the traditional vegetation indices has been evaluated (Verrelst et al. 2008) and exploited (Hall et al. 2008), and new angular indices have been proposed (Chen et al. 2003). Other statistical methods, such as multiple linear regression (Huber et al. 2010), artificial neural networks (Heiskanen 2006), and measure theory (Knyazikhin et al. 1999; Knyazikhin et al. 1998), have been shown to perform well on multi-angular data, and retrieval algorithms based on parametric radiative transfer (RT) models have been developed (Gobron et al. 2002; Lavergne et al. 2007). The use of physical RT models, however, is still limited (Schaepman et al. 2009), although multi-angular data may help constraining the inversion problem (Asner et al. 1998; Diner et al. 1999). So far, the use of physical RT models in a multi-angular context has mostly been limited to explanatory and exploratory studies. For example, Verrelst et al. (2008) generalized their results concerning the angular sensitivity of spectral vegetation indices using the coupled PROSPECT-FLIGHT model and Rautiainen et al. (2008) evaluated the contribution of the forest background in the canopy

reflectance using the FRT model (Kuusk and Nilson 2000). Simulations of multi-angular data have been conducted to compare the performance of various RT models (Schlerf et al. 2007) and to explore a new concept for spectro-angular sampling (Simic and Chen 2008). Physical models have also been used for investigating the feasibility of estimating surface variables based on simulated data (Timmermans et al. 2009; Weiss et al. 2000).

Because both the canopy and the atmosphere have an anisotropic behaviour (Schaepman-Strub et al. 2006), a complete understanding of the angular properties of the coupled canopy-atmosphere system is required for successful implementation of algorithms making full use of multi-angular data (Pandya et al. 2007). Coupled canopy-atmosphere physical RT models allow directly relating the surface variables to the radiance measured at the top of the atmosphere (TOA) (Verhoef and Bach 2007). Using such a coupled model allows directly inverting the TOA radiance data to estimate the forest variables and also gives a more accurate inversion framework while minimizing data pre-processing efforts (Laurent et al. 2011b). Heiskanen (2006) showed that the estimates of tree cover and tree height obtained from MISR multi-angular TOA data were more accurate than those obtained from mono-angular data, and also than the estimates obtained from the atmospherically corrected data, because of a quilted pattern caused by the tiles used for atmospheric correction. Physically-based studies that estimate surface variables from TOA data remain scarce, but Laurent et al. (2011b) found that the direct inversion of TOA near-nadir radiance data performed as well as the traditional approach based on atmospherically-corrected data for estimating biophysical and biochemical variables in Norway spruce stands.

The objective of the present study was to investigate the potential of multi-angular TOA radiance data for estimating forest variables, using a coupled canopy-atmosphere physical RT model. The case study focused on three Norway spruce stands in the Czech Republic for which multi-angular CHRIS data and field data were available. The canopy-atmosphere system was simulated by coupling the SLC soil-leaf-canopy (Verhoef and Bach 2007) and the MODTRAN4 atmosphere (Berk et al. 2003) models, based on the 4-stream RT theory, which allows fully using the canopy angular reflectance factors provided by SLC in the simulations. The model inversion was performed by the look-up (LUT) approach.

3.2 Materials

The study area was located at the Bily Kriz experimental research site in the east of the Czech Republic, in the Moravian-Silesian Beskydy Mountains (18.54°E, 49.50°N; 936 m above sea level). A detailed description of the environmental conditions can be found in Kratochvilová et al. (1989). Three stands of montane

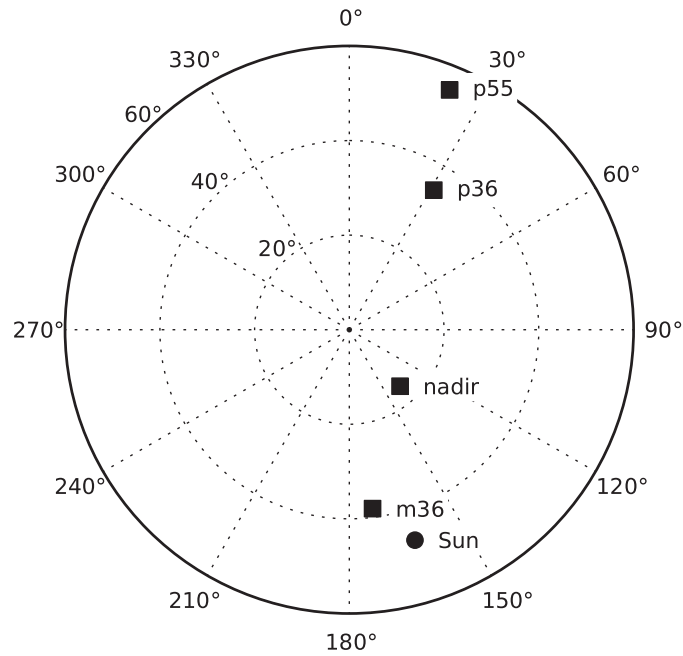**Table 3.1** Stand characteristics and model input.

Stand		YOUNG	OLD1	OLD2
Age (years)		29	100	75
Density (trees/ha)		1450	160	420
DBH (cm)		14	53	37
Atm.	Aerosols		Urban	
	VIS (km)		100	
	H_2O		Default	
	O_3		Default	
Canopy	PAI (-)	8.88	5.73	7.35
	fB (-)	0.13	0.23	0.4
	D (-)	0	0	0.1
	Hot (-)	0.01	0.01	0.01
	LIDF	Spherical	Spherical	Spherical
	Cv (-)	0.9	0.55	0.7
	Zeta (-)	0.34	0.24	0.26
Needle	Cab ($\mu g/cm^2$)	55	60	65
	Cw (cm)	0.02	0.02	0.02
	Cdm (g/cm^2)	0.04	0.04	0.04
	Cs (-)	0	0	0
	N (-)	2.7	2.5	2.3
Bark	Cab ($\mu g/cm^2$)		10	
	Cw (cm)		0	
	Cdm (g/cm^2)		0.5	
	Cs (-)		15	
	N (-)		10	

Norway spruce (*Picea abies* (L.) Karst.) having different ages and structures were selected for the study (Table 3.1): YOUNG, OLD1 and OLD2. The data campaign took place in the first half of September 2006.

3.2.1 Remote sensing data

A set of multi-angular CHRIS data (Cutter et al. 2000) was acquired on September 12th, 2006. Four images covered the study area: m36, nadir, p36, p55 (m stands for minus, meaning looking in the backward scattering direction, and p stands for plus, meaning forward scattering direction). The acquisition geometry is shown in Figure 3.1, where it can be seen that the near-nadir image was acquired in the backward direction, with a zenith angle of -16°. The four images were acquired within a time frame of three minutes, in CHRIS chlorophyll mode (mode 4), resulting in 18 spectral bands in the range 485-802 nm at a spatial resolution of 17 m. The images were radiometrically calibrated by the data provider and were further de-striped, geo-corrected and ortho-rectified using nearest neighbour interpolation. Band 15, centred at 761 nm, was not used in this study because it sampled one of the oxygen atmospheric absorption features and it was noisy. Using RT model inversion, there is no incentive reducing the number of spectral bands any further (Schaepman

et al. 2009). The centre wavelength and full-width-half-maximum of the 17 bands used in the study are presented in Table 3.2.

Figure 3.1 Acquisition geometry of the multi-angular CHRIS/PROBA data.

Table 3.2 Centre wavelength (λ_{centre}) and full-width-half-maximum (FWHM) of the CHRIS bands used in the study.

Band number	λ_{centre} (nm)	FWHM (nm)
1	490.7	11.6
2	552.0	12.9
3	632.4	14.1
4	670.3	10.8
5	681.3	11.2
6	689.9	5.8
7	695.7	5.9
8	701.7	6.0
9	707.7	6.1
10	713.9	6.3
11	720.2	6.4
12	736.5	13.3
13	746.6	6.9
14	753.5	7.0
16	778.8	14.9
17	790.2	7.6
18	797.9	7.8

3.2.2 Field data

In the above mentioned test site, plant area index (PAI), defined as half of the total plant area (needles and non-photosynthetic plant material) per unit of ground surface area (Chen 1996), was estimated using three methods: a LAI-2000 plant canopy analyser, hemispherical photographs, and a TRAC instrument (Tracing Radiation and Architecture of Canopies) (Homolová et al. 2007). For each stand, the PAI was taken as the average of the three obtained values. The woody-to-total area ratio used by Homolová et al. (2007) to obtain LAI from PAI was taken as the fraction of brown material in the PAI (fB). The crown cover (Cv) was estimated from a classification of existing airborne imaging spectrometer data (Lukeš 2009).

Canopy and needle measurements were performed on 10 sample trees in the YOUNG stand and 20 in the OLD1 stand. Canopy measurements alone were performed in the OLD2 stand. Canopy structure measurements included tree height, crown radius, and crown length. Spectral properties of the main background components (soil, humus, litter, understory species) and bark were measured in the field using an ASD field spectrometer.

3.3 Methods

3.3.1 TOA radiance simulation

The 4-stream theory provides a simple but powerful framework for radiative transfer modelling. When ignoring the adjacency effect, the TOA radiance in the observation direction L_o can be calculated as (Laurent et al. 2011b):

$$L_o = L_{atm} + \underbrace{\frac{G_{ssdo}r_{sd} + G_{sddo}r_{dd}}{1 - r_{dd}\rho_{dd}}}_{\text{PATH}} + \underbrace{\frac{G_{sdoo} + G_{mult}r_{sd}}{1 - r_{dd}\rho_{dd}}r_{do} + \underbrace{G_{ssoo}r_{so}}_{\text{GSUN}}}_{\text{GRFL}} \quad (3.1),$$

where L_{atm} is the atmospheric path radiance, ρ_{dd} is the spherical albedo of the atmosphere, the *r* terms are the directional reflectance factors of the canopy, the *G* terms are directional atmospheric gain factors for the double pass in the atmosphere, and the subscripts indicate the direction of the radiation: *s* for the sun direction, *o* for the observer direction and *d* for diffuse hemispherical radiation. The wavelength dependency of the variables is omitted for clarity. L_{atm}, ρ_{dd} and the *G* factors can be calculated using any atmospheric RT model simulating the total path radiance (PATH), the sunlight ground-reflected radiance (GSUN), and the total ground-reflected (GRFL) radiance (see annotations on Equation 3.1). Laurent et al. (2011b) showed that three runs of the atmospheric model for Lambertian surfaces are sufficient to calculate L_{atm}, ρ_{dd} and the *G* factors.

3.3.2 Radiative transfer models and parameterization

The Soil-Leaf-Canopy (SLC) model simulates the four reflectance factors of the canopy. It couples the 4SOIL soil reflectance model, the PROSPECT leaf optical properties model (Jacquemoud and Baret 1990), and the 4SAIL2 canopy RT model (Verhoef and Bach 2007).

4SAIL2 is the latest version of the SAIL model. It includes the crown clumping effect through the crown cover (Cv) and the tree shape factor (Zeta), which is defined as the crown diameter divided by the height of the crown centre above ground. The leaf inclination distribution function (LIDF) was set to spherical (LIDFa = -0.35 and LIDFb = -0.15), as this is the best description of randomly oriented needles. 4SAIL2 also allows mixing green and brown elements in the canopy by using fB and the dissociation factor (D), which describes the distribution of the brown material between the two canopy layers. The green elements were used for the needles and the brown elements for the bark.

The optical properties of needle and bark material were simulated using PROSPECT. The inputs are the concentration of chlorophyll (Cab), dry matter (Cdm), and water (Cw), and the leaf structure parameter (N). The model applied was a modified version which includes also the concentration in senescent material (Cs) (Verhoef and Bach 2003b). The specific absorption coefficients and refractive index of the leaf material at 1 nm resolution were taken from PROSPECT-4 (Feret et al. 2008). The parameters for the bark were optimized to match the measured bark signature, and the parameters for the needles were tuned using the four angular measurements at TOA level, together with the D parameter (Table 3.1).

The soil model was not used in this study because of high canopy cover and vegetated background. Instead, the background signature for each stand was calculated as the average of the measured signatures of the background components, weighted by their fractional area.

The MODTRAN4 model was used for the atmosphere. The following options were selected: DISORT algorithm with 8 streams, medium speed correlated-k option with 17 values, and 5 cm^{-1} spectral database (Guanter et al. 2009). The atmosphere was assumed to be constant for the four images, while the path length differences due to changing observational geometry are accounted for within MODTRAN4. Urban aerosols were used because of the dominant north wind blowing from an industrial zone and high air concentration of SO_2. The visibility was chosen as the smallest value (100 km) for which the simulated atmospheric path radiance in each observation direction was smaller than all radiances in the corresponding CHRIS image.

The *r* factors provided by SLC and the *G* factors calculated from the MODTRAN4 outputs were resampled to the CHRIS spectral bands using Gaussian approximations of the sensor response functions before feeding them into Equation 3.1.

3.3.3 Parameter influences

For each observation direction o, the Jacobian matrix $\mathbf{J}_o$ is defined as the matrix of the partial derivatives of the model output L_o with respect to the normalized input parameters p_{norm}:

$$\mathbf{J}_o = \lfloor j_{o,i,k} \rfloor_{1 \leq i \leq n_b,\ 1 \leq k \leq n_p}, \text{ with } j_{o,i,k} = \frac{\partial L_o(\lambda_i)}{\partial p_{norm,k}} \tag{3.2},$$

where λ_i is the central wavelength of the ith band, n_b is the number of bands and n_p is the number of parameters. Each parameter p was normalized assuming a uniform distribution over its potential variation range:

$$p_{norm} = \frac{p}{p_{max} - p_{min}} \tag{3.3},$$

where p_{min} and p_{max} are the potential minimum and maximum values of p. Canopy and atmospheric parameters were varied by 1% of their potential variation range, except the canopy hotspot parameter (hot) which was varied by 0.005 because of its very small value. For the multi-angular analysis, the Jacobian matrix $\mathbf{J}$ was calculated over the ensemble Θ of observation directions by vertically stacking the $\mathbf{J}_o$ matrices for o in Θ. For each parameter p_k, the influence indicator α_k was defined as:

$$\alpha_k = \sqrt{\frac{\sum\limits_{o \in \Theta} \sum\limits_{i=1}^{n_b} w_i j_{o,i,k}^2}{\sum\limits_{o \in \Theta} \sum\limits_{i=1}^{n} w_i}} \tag{3.4},$$

where the w terms are weights that account for the irregular spectral distance between the CHRIS bands:

$$\begin{cases} w_1 = (\lambda_2 - \lambda_1) \\ w_i = (\lambda_{i+1} - \lambda_{i-1})/2, \ 2 \leq i \leq n-1 \\ w_n = (\lambda_n - \lambda_{n-1}) \end{cases} \tag{3.5}.$$

To allow easier comparison between stands, the α values were normalized so that their sum was equal to one (α_{norm}). Only the most influential parameters having high α_{norm} values can be estimated. The number of parameters that can be estimated depends on the dimensionality (see following section).

3.3.4 Dimensionality

The dimensionality of the estimation problem is the maximum number of parameters that can theoretically be estimated. The dimensionality was assessed by applying a Singular Value Decomposition (SVD) to $\mathbf{J}$ (Laurent et al. 2011b; Verhoef 2007), yielding: $\mathbf{J} = \mathbf{USV}^{\mathrm{T}}$, where $\mathbf{S}$ is a diagonal matrix containing the singular values, and $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices so that $\mathbf{UU}^{\mathrm{T}} = \mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}$ and $\mathbf{VV}^{\mathrm{T}} = \mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$. Because $\mathbf{J}$ relates the vector of normalized parameter change $\Delta\mathbf{p}$ to the stacked multi-angular vector of radiance difference $\Delta\mathbf{L}$ as: $\Delta\mathbf{L} = \mathbf{J}\Delta\mathbf{p}$, one can infer:

$$\mathbf{U}^{\mathrm{T}}\Delta\mathbf{L} = \mathbf{SV}^{\mathrm{T}}\Delta\mathbf{p} \qquad (3.6).$$

Equation 3.6 shows that, because $\mathbf{S}$ is diagonal, there is a one-to-one relationship between the successive elements of the transformed vector of parameter variations $\mathbf{V}^{\mathrm{T}}\Delta\mathbf{p}$ and the transformed vector of output differences $\mathbf{U}^{\mathrm{T}}\Delta\mathbf{L}$. Therefore, the rank of $\mathbf{S}$ gives the dimensionality of the estimation problem in the transformed spaces. Because $\mathbf{V}^{\mathrm{T}}$ and $\mathbf{U}^{\mathrm{T}}$ are orthonormal, the dimensionality can easily be translated back to the original parameter and output spaces, since in that case the inverses of $\mathbf{V}^{\mathrm{T}}$ and $\mathbf{U}^{\mathrm{T}}$ are simply equal to $\mathbf{V}$ and $\mathbf{U}$, respectively. After ordering the singular values in decreasing order, the dimensionality was taken as the number of singular values needed to reach 95% of the sum of all singular values.

3.3.5 Variables estimation

In this paper, the term "variable" refers to the parameters of interest, the ones that were estimated. The number of variables n_v was chosen as the dimensionality plus one, in order to have one adjustment variable, and the n_v most influential parameters were selected as variables. The LUT inversion method was chosen because of its ability to find the global minimum of the cost function. For each stand, a LUT was built by sampling each variable from a uniform distribution using regular steps, between minimum and maximum values defined using prior knowledge about the structure of the three stands (Table 3.3). For each combination of the variable values,

Table 3.3 Variable sampling scheme used for building the LUTs.

Parameter	Min	Max	Step	# values
fB (-)	0	0.6	0.1	7
Cv (-)	0.4	1	0.1	7
D (-)	0	1	0.2	6
Zeta (-)	0.2	0.4	0.04	6
NeedleCab (μg/cm²)	40	80	10	5
NeedleCdm (g/cm²)	0.01	0.07	0.01	7
NeedleCs (-)	0	0.03	0.01	4

the coupled model was run, keeping the values of the parameters to the values used for the simulations (Table 3.1), and the L_o simulations in the CHRIS spectral bands for each of the four observation directions were stored in the LUT. The combinations having a non-influential variable were summarized into one entry having a value of -99 for that variable (e.g. Zeta when Cv = 1, and D when fB = 0). The variable estimation was performed using the multi-angular cost function χ, defined using the same structure as the α indicator:

$$\chi = \sqrt{\frac{\sum\limits_{o \in \Theta}\sum\limits_{i=1}^{n_b} w_i \left(L_o(\lambda_i) - L_{o,ref}(\lambda_i)\right)^2}{\sum\limits_{o \in \Theta}\sum\limits_{i=1}^{n} w_i}} \tag{3.7},$$
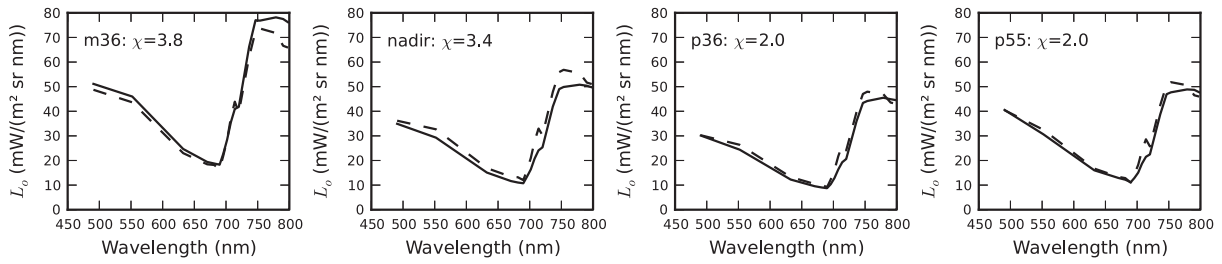
where $L_{o,ref}$ is the measured CHRIS radiance in the observing direction *o*. For each angular combination Θ, the entries having the same set of variable values were used to calculate the χ value. For a given stand and angular combination, the solution space was defined as the set of LUT entries whose χ value were smaller than a threshold value χ_{thresh} (see section 3.4.4). For each variable, the estimate was chosen as the median of the values in the solution space, and the estimation uncertainty as the standard deviation. If there were some -99 values, they were excluded before calculating the median and standard deviation. Because of the large number of variables and angular combinations, an indicator of the quality of the estimates was built. Because of the different units, the variable values were normalized in a similar fashion as for the Jacobian matrix (Equation 3.3). The quality indicator δ was defined as the average absolute difference between the reference value v_{ref} (Table 3.1) and the estimate for each variable *v*, ignoring the variables for which the estimate was non-influential:

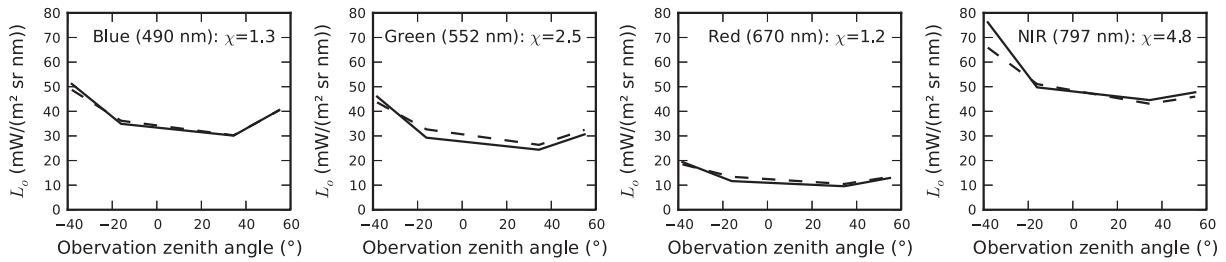$$\delta = \frac{1}{n_v}\sum_{i=1}^{n_v}\left|v_i - v_{i,ref}\right| \tag{3.8}.$$

3.4 Results

3.4.1 Simulations

The simulation results obtained for the YOUNG stand using the field data (Table 3.1) are presented in Figures 3.2 and 3.3. The simulated spectral signatures (Figure 3.2) matched the CHRIS data well in all observation directions, but the χ values were

Figure 3.2 Simulated (dashed lines) and measured (solid lines) spectral TOA radiance signatures for each CHRIS observation direction for the YOUNG stand.

Figure 3.3 Simulated (dashed lines) and measured (solid lines) angular TOA radiance signatures for four selected CHRIS spectral bands for the YOUNG stand.
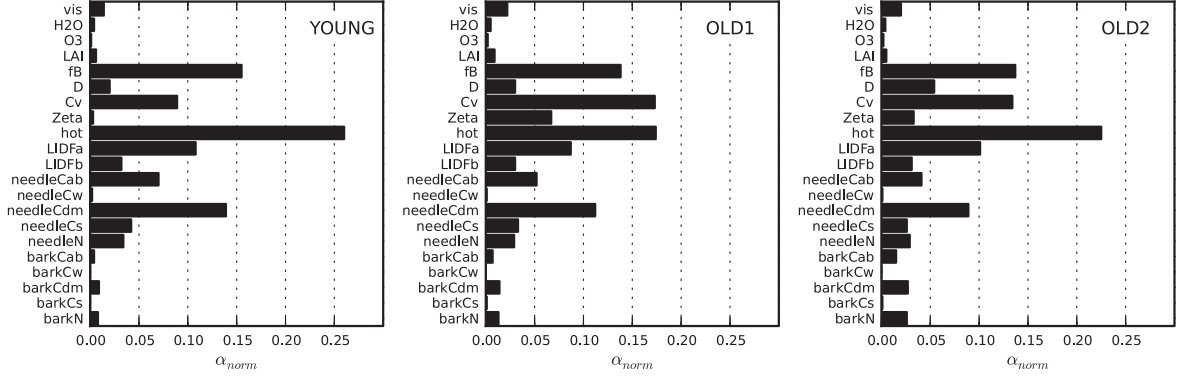
smaller in the forward than in the backward direction. The signatures were overestimated in the nadir, p36, and p55 directions and underestimated in the m36 direction. The angular signatures (Figure 3.3) matched the bowl shape reflectance anisotropy pattern present in the CHRIS data. This bowl shape is typical for dense coniferous forests (Verrelst et al. 2010a). The χ values were smaller in the visible domain, especially in the blue and red band, and larger in the NIR band. The χ value for the YOUNG stand for Θ = {m36, nadir, p36, p55} was 2.9 mW/(m^2 sr nm).

Similar trends were observed for the OLD1 and OLD2 stands, but with slightly higher χ values: the χ values for Θ = {m36, nadir, p36, p55} were 4.8 mW/(m^2 sr nm) for the OLD1 stand and 3.9 mW/(m^2 sr nm) for the OLD2 stand.

3.4.2 Parameter influences

The local sensitivity analyses were conducted for the three stands, for all possible angular combinations. The three stands present similar profiles of parameter influences for Θ = {m36, nadir, p36, p55} (Figure 3.4): needle and canopy parameters were more influential than the bark and atmosphere parameters. This hierarchy was also present for the other angular combinations (in the YOUNG stand for example, Table 3.4). LIDFa, LIDFb, fB, and needleCdm were more influential in the forward direction (p36 and p55), whereas the hotspot parameter was very influential in the backward direction and Cv was most influential in the nadir and p36 directions which were closest to nadir viewing. Zeta was not influential because of

Figure 3.4 Parameter influences (α_{norm}) for the three stands for Θ = {m36, nadir, p36, p55}.

the very high Cv value. Similar trends were observed for the OLD1 and OLD2 stands, except that the Zeta parameter was influential in the backward direction (α_{norm} of about 0.05). PAI was not influential in all cases. Regarding the atmosphere, VIS was the most influential parameter, with a α_{norm} value between 0.01 and 0.05 for all stands, the maximum value was observed in the p55 direction. When considering Θ = {m36, nadir, p36, p55}, however, the α_{norm} value of VIS was lower than 0.02 for all stands (Figure 3.4).

3.4.3 Dimensionality

The local dimensionalities obtained using all possible angular combinations are presented in Table 3.5. For each combination, the three stands had similar values, with the YOUNG stand having slightly smaller dimensionality. The dimensionality increased from 3 to 6 when increasing the number of observation directions in Θ from one to four, thus revealing an increasing potential to estimate more variables when increasing the angular sampling.

3.4.4 Variable estimations

The two extreme cases of minimal and maximal dimensionality were investigated. For each stand, two LUTs were built: one with four free variables, and one with seven variables. The free variables were chosen based on the parameter influence results. LIDFa, LIDFb and hot were influential, but a value of 0.01 for hot is common for forests, and a spherical LIDF i the most appropriate description of the random orientation of the needles in coniferous forest, so these parameters were not chosen as variables to estimate. For each stand, the 4-variable LUT was built using fB, Cv, needleCdm, and needleCab, while needleCs, D and Zeta were added in the 7-variable LUTs. The same variable s sampling scheme was used for both LUTs (Table 3.3). The 4-variable LUTs had1,715 entries, and the 7-variable LUT had 191,660 entries after removing the insensitive combinations as described in section 3.3.5. After building the LUTs, for each stand, the variables were estimated for each angular combination (see section 3.3.5). The threshold value used to delimit the solution

Table 3.4 Parameter influences (α_{norm}) for the YOUNG stand for all possible angular combinations.

Θ	barkN	barkCs	barkCdm	barkCw	barkCab	needleN	needleCs	needleCdm	needleCw	needleCab	LIDFb	LIDFa	hot	Zeta	Cv	D	fB	PAI	O3	H2O	VIS
{m36}	0.008	0.000	0.009	0.000	0.004	0.041	0.038	0.125	0.001	0.071	0.004	0.063	0.390	0.005	0.060	0.018	0.138	0.009	0.001	0.004	0.010
{nadir}	0.009	0.000	0.010	0.000	0.005	0.041	0.044	0.146	0.002	0.077	0.004	0.140	0.171	0.004	0.148	0.021	0.161	0.004	0.001	0.005	0.008
{p36}	0.009	0.000	0.010	0.000	0.004	0.027	0.051	0.174	0.002	0.079	0.059	0.180	0.053	0.001	0.111	0.025	0.193	0.002	0.001	0.005	0.013
{p55}	0.010	0.000	0.011	0.000	0.004	0.016	0.064	0.219	0.002	0.093	0.073	0.131	0.008	0.000	0.047	0.032	0.244	0.003	0.002	0.007	0.034
{nadir, m36}	0.008	0.000	0.009	0.000	0.004	0.040	0.039	0.128	0.001	0.071	0.004	0.093	0.321	0.004	0.096	0.018	0.141	0.008	0.001	0.004	0.009
{nadir, p36}	0.009	0.000	0.010	0.000	0.004	0.035	0.046	0.153	0.002	0.076	0.036	0.152	0.133	0.003	0.131	0.022	0.170	0.003	0.001	0.004	0.010
{nadir, p55}	0.009	0.000	0.010	0.000	0.004	0.033	0.048	0.163	0.002	0.078	0.039	0.129	0.132	0.003	0.118	0.024	0.181	0.003	0.001	0.005	0.019
{m36, p36}	0.008	0.000	0.008	0.000	0.004	0.035	0.039	0.130	0.001	0.068	0.028	0.099	0.318	0.004	0.071	0.019	0.144	0.008	0.001	0.004	0.010
{m36, p55}	0.008	0.000	0.009	0.000	0.004	0.035	0.041	0.138	0.002	0.071	0.030	0.076	0.326	0.004	0.053	0.020	0.153	0.008	0.001	0.005	0.016
{p36, p55}	0.009	0.000	0.010	0.000	0.004	0.022	0.056	0.190	0.002	0.083	0.064	0.158	0.040	0.000	0.088	0.028	0.212	0.003	0.002	0.006	0.024
{nadir, m36, p36}	0.008	0.000	0.009	0.000	0.004	0.036	0.040	0.132	0.001	0.070	0.024	0.108	0.283	0.004	0.095	0.019	0.146	0.007	0.001	0.004	0.009
{nadir, m36, p55}	0.008	0.000	0.009	0.000	0.004	0.036	0.041	0.137	0.002	0.071	0.025	0.094	0.286	0.004	0.087	0.020	0.151	0.007	0.001	0.004	0.014
{nadir, p36, p55}	0.009	0.000	0.010	0.000	0.004	0.031	0.049	0.164	0.002	0.077	0.045	0.144	0.115	0.002	0.115	0.024	0.183	0.003	0.001	0.005	0.017
{m36, p36, p55}	0.008	0.000	0.009	0.000	0.004	0.032	0.042	0.140	0.002	0.070	0.036	0.100	0.284	0.003	0.066	0.020	0.156	0.007	0.001	0.004	0.015
{nadir, m36, p36, p55}	0.008	0.000	0.009	0.000	0.004	0.034	0.042	0.139	0.002	0.070	0.032	0.108	0.260	0.003	0.089	0.020	0.155	0.006	0.001	0.004	0.014

Table 3.5 Dimensionality based on the singular value decomposition for all possible angular combinations.

Θ	YOUNG	OLD1	OLD2
{m36}	3	3	3
{nadir}	3	3	3
{p36}	3	3	3
{p55}	4	4	4
{nadir, m36}	4	4	4
{nadir, p36}	4	5	5
{nadir, p55}	5	5	5
{m36, p36}	4	4	4
{m36, p55}	4	5	5
{p36, p55}	5	5	5
{nadir, m36, p36}	5	5	5
{nadir, m36, p55}	5	6	6
{nadir, p36, p55}	5	6	6
{m36, p36, p55}	5	6	6
{nadir, m36, p36, p55}	5	6	6

space was defined in the form $\chi_{thresh} = k\chi_{min}$, where χ_{min} is the minimum χ value found in the LUT. Increasing values of k were tried, until some of the solution spaces based on the 7-variable LUT became larger than 0.5% of the total number of entries in the LUT, the proportion after which Weiss et al. (2000) found that the estimation error started to increase. In the end, the value $\chi_{thresh} = 1.05\chi_{min}$ was chosen and used for all LUTs.

The estimation results of the 4-variable case for the YOUNG stand are presented in Table 3.6. The solution spaces consisted of less than 8 entries. The best radiance match (smallest χ_{thresh} value) was provided by Θ = {nadir, p36, p55}, whereas the best estimate match (smallest δ value) was provided by Θ = {m36, p36, p55}. In the 7-variable case (Table 3.7), the solution spaces were larger (30 to 250 entries), making the estimate uncertainty more statistically relevant. Similarly to the 4-variable case, the best radiance match was provided by Θ = {nadir, p36, p55}. The best estimates were provided by two combinations having the same δ value: Θ = {nadir, p36, p55} and Θ = {m36, p36, p55}.

The solution spaces for the OLD1 and OLD2 stands were bigger than for the YOUNG stand: up to about 100 entries for the 4-variable case, and up to about 1,000 entries for the 7-variable case. Regarding the angular combinations, the results for the OLD2 stand were very similar to the YOUNG stand: best radiance match obtained with Θ = {nadir, p36, p55} in both cases, and best estimates with Θ = {nadir, m36, p55} for the 7-variable case, with Θ = {nadir, p55} for the4-variable case. The results for the OLD1 stand are slightly different, as the best radiance match was obtained for Θ = {nadir} in both cases. The best estimates were

Table 3.6 Estimates obtained from the 4-variable LUT for the YOUNG stand for all possible angular combinations. The values in brackets represent the uncertainties of the estimates.

Θ	Cv (-)	fB (-)	needleCab (µg/cm²)	needleCdm (g/cm²)	δ (-)	χ$_{thresh}$ (-)	# entries
{m36}	0.9 (0.00)	0.3 (0.00)	50.0 (0.00)	0.01 (0.00)	0.450	1.815	1
{nadir}	0.8 (0.05)	0.0 (0.05)	80.0 (4.71)	0.04 (0.00)	0.416	1.118	3
{p36}	0.7 (0.00)	0.05 (0.05)	70.0 (0.00)	0.035 (0.01)	0.416	1.019	2
{p55}	0.85 (0.07)	0.3 (0.07)	70.0 (4.71)	0.02 (0.01)	0.494	1.059	6
{nadir, m36}	0.6 (0.00)	0.1 (0.08)	50.0 (0.00)	0.02 (0.01)	0.502	1.974	3
{nadir, p36}	0.7 (0.00)	0.0 (0.00)	70.0 (0.00)	0.04 (0.00)	0.416	0.761	1
{nadir, p55}	0.8 (0.00)	0.15 (0.08)	80.0 (4.33)	0.035 (0.01)	0.364	0.813	4
{m36, p36}	0.8 (0.00)	0.3 (0.00)	50.0 (0.00)	0.01 (0.00)	0.537	1.828	1
{m36, p55}	1.0 (0.00)	0.2 (0.07)	55.0 (5.00)	0.03 (0.01)	0.234	1.947	4
{p36, p55}	0.85 (0.05)	0.25 (0.08)	70.0 (4.33)	0.025 (0.01)	0.407	0.787	4
{nadir, m36, p36}	0.6 (0.05)	0.1 (0.07)	55.0 (5.00)	0.03 (0.01)	0.373	1.534	8
{nadir, m36, p55}	0.7 (0.07)	0.2 (0.09)	60.0 (4.52)	0.03 (0.01)	0.364	1.759	7
{nadir, p36, p55}	0.8 (0.00)	0.1 (0.07)	75.0 (5.00)	0.04 (0.01)	0.286	0.674	4
{m36, p36, p55}	0.9 (0.05)	0.2 (0.07)	60.0 (0.00)	0.03 (0.01)	0.191	1.562	5
{nadir, m36, p36, p55}	0.7 (0.00)	0.1 (0.08)	60.0 (0.00)	0.03 (0.01)	0.329	1.361	3

provided by Θ = {nadir, m36} and Θ = {nadir, m36, p36} in the 4-variable case, and by Θ = {m36, p55} and Θ = {nadir, m36, p55} in the 7-variable case.

The best δ values in the 4-variable case were: 0.191 for the YOUNG stand, 0.277 for the OLD1 stand and 0.130 for the OLD2 stand. In the 7-variable case, they were 0.426 for the YOUNG stand, 0.396 for the OLD1 stand, and 0.292 for the OLD2 stand.

3.5 Discussion

3.5.1 Simulations

The atmospheric visibility had to be chosen very high so that all radiances in each CHRIS image would be higher than the corresponding simulated L_{atm}. L_{atm} may also explain the different χ values obtained for the angular signatures in the visible and NIR domain: in the visible most of the radiance comes from L_{atm}, whereas in the NIR band L_{atm} is small and most of the radiance comes from the canopy. Therefore, inaccuracies in the canopy reflectance were less important in the visible than in the NIR, leading to smaller χ values, whereas they had more impact on the NIR angular signature, leading to larger χ values.

Despite its relatively simple description of the forest canopy as a turbid medium clumped into crowns and having two layers of different compositions, the SLC model performed well for simulating the multi-angular CHRIS data. The trend to underestimation in the m36 direction and overestimation in the nadir, p36, and p55 directions might be due to the LIDF. In addition, SLC includes only crown level

Table 3.7 Estimates obtained from the 7-variable LUT for the YOUNG stand for all possible angular combinations. The values in brackets represent the uncertainties of the estimates.

Θ	Cv (-)	fB (-)	needleCab (μg/cm²)	needleCdm (g/cm²)	D (-)	Zeta (-)	needleCs (-)	δ (-)	χ_{thresh} (-)	# entries
{m36}	0.9 (0.04)	0.4 (0.09)	50.0 (0.00)	0.01 (0.00)	0.4 (0.18)	0.28 (0.07)	0.02 (0.01)	0.544	1.699	49
{nadir}	0.8 (0.04)	0.2 (0.14)	70.0 (5.00)	0.05 (0.01)	1.0 (0.31)	0.28 (0.07)	0.02 (0.01)	0.742	1.088	169
{p36}	0.8 (0.08)	0.3 (0.19)	70.0 (4.01)	0.04 (0.01)	0.6 (0.25)	0.28 (0.07)	0.02 (0.01)	0.544	0.981	204
{p55}	0.8 (0.09)	0.4 (0.14)	70.0 (0.00)	0.02 (0.01)	0.4 (0.22)	0.28 (0.07)	0.02 (0.01)	0.594	1.018	211
{nadir, m36}	0.6 (0.03)	0.3 (0.20)	50.0 (3.81)	0.02 (0.00)	0.8 (0.24)	0.32 (0.05)	0.02 (0.01)	0.772	1.756	102
{nadir, p36}	0.7 (0.05)	0.2 (0.13)	70.0 (3.87)	0.04 (0.00)	0.6 (0.30)	0.24 (0.06)	0.01 (0.01)	0.559	0.752	142
{nadir, p55}	0.8 (0.00)	0.2 (0.17)	80.0 (4.50)	0.05 (0.01)	0.8 (0.45)	0.32 (0.06)	0.02 (0.01)	0.673	0.793	167
{m36, p36}	0.7 (0.03)	0.3 (0.18)	50.0 (3.51)	0.02 (0.01)	0.6 (0.23)	0.4 (0.02)	0.02 (0.01)	0.643	1.686	97
{m36, p55}	1.0 (0.00)	0.3 (0.16)	50.0 (4.96)	0.03 (0.01)	0.4 (0.25)	Non-influent	0.02 (0.01)	0.485	1.891	30
{p36, p55}	0.9 (0.02)	0.3 (0.16)	80.0 (4.50)	0.035 (0.01)	0.4 (0.27)	0.3 (0.07)	0.02 (0.01)	0.460	0.747	192
{nadir, m36, p36}	0.7 (0.05)	0.2 (0.18)	60.0 (0.00)	0.02 (0.01)	0.6 (0.24)	0.4 (0.02)	0.01 (0.01)	0.589	1.379	101
{nadir, m36, p55}	0.8 (0.05)	0.3 (0.16)	60.0 (1.41)	0.03 (0.01)	0.6 (0.29)	0.4 (0.03)	0.01 (0.01)	0.539	1.673	247
{nadir, p36, p55}	0.8 (0.00)	0.1 (0.13)	80.0 (5.00)	0.04 (0.01)	0.4 (0.39)	0.28 (0.07)	0.02 (0.01)	0.426	0.656	171
{m36, p36, p55}	0.9 (0.06)	0.4 (0.16)	60.0 (3.35)	0.03 (0.01)	0.4 (0.25)	0.36 (0.05)	0.02 (0.01)	0.426	1.537	240
{nadir, m36, p36, p55}	0.8 (0.05)	0.3 (0.16)	60.0 (0.67)	0.03 (0.01)	0.4 (0.30)	0.4 (0.03)	0.01 (0.01)	0.440	1.302	222

clumping, whereas branch and shoot level clumping play an important role in coniferous stands by increasing the trapping of photons in the canopy (Stenberg 2007). Poorer simulation results at the top of the canopy for the oblique directions are not uncommon (Gascon et al. 2007; Schlerf et al. 2007). The assumption of constant atmospheric parameters for the four angles might have limited the quality of the TOA simulations. However, the time difference between the first and last image recorded was smaller than three minutes, and the water vapour absorption in the range 700 to 800 nm is very weak and only affects the CHRIS bands with the longest wavelengths, so that the spatial variation of the water vapour was considered negligible.

It was possible to obtain a good simulation of the needle and bark material using PROSPECT, but at the expense of the physical meaning of the parameters. For this study, however, the goal was to have a realistic representation of the canopy elements, and to see how well the obtained estimates would match the original values. This might be avoided by recalibrating PROSPECT (Malenovský et al. 2006), or by using the LIBERTY model (Dawson et al. 1998) which was designed specifically for needles. In addition, the values for needleN and D were calibrated for $\Theta = \{m36, nadir, p36, p55\}$, and therefore differ from the values which were used in Laurent et al. (2011b): D was decreased to 0, which gives a homogeneous distribution of bark material over the canopy. This is in better agreement with the fact that more bark material was visible in the canopy from oblique directions than from nadir (2008). Schlerf (2007) also used a high needleN value of 3 for spruce stands in Germany. The better results obtained for the YOUNG stand may be due to its high Cv and tree density, as this type of canopy is closest to the turbid medium assumption used in 4SAIL2.

The good performance of SLC and the high quality of the state-of-the-art MODTRAN4 model explain the good simulations of the spectral and angular signatures measured by CHRIS. Another factor explaining these good results is the fact that the four angular reflectance components provided by SLC are fully used in the canopy-atmosphere coupling (Equation 3.1). The results of the present study extend to multi-angular data the results of Laurent et al. (2011b), who had successfully simulated the TOA radiance of the same three stands in the near-nadir direction.

3.5.2 Parameter influence

The analysis of the α_{norm} values showed that the canopy and needle parameters were most influential, meaning that the canopy biophysical and biochemical variables could be directly estimated from the TOA radiance data. In addition, the small influence of the atmospheric parameters found in this case study is an advantage, because inaccuracies in atmospheric parameter values will not influence the simulations and therefore not the estimations either. Because this was true in all

four observation directions, one expects to obtain similar results as in studies based on atmospherically-corrected data, that is, higher potential of multi-angular data for estimating canopy variables. The small atmospheric influence, however, might be limited to very clear atmospheres only.

The PAI was not influential because of its very high value. Most of the PAI consists of needles (in proportion 1-fB), and therefore a high PAI saturates the signal in the same way as a high LAI does. fB was one of the most influential parameters, but the bark parameters were not influential, probably due to saturation effects caused by the very high input values used in PROSPECT. The very high influence of the hotspot parameter in the backward direction was due to the wide hotspot effect caused by the high PAI of the three stands. The m36 direction was close to the hotspot, and the nadir image which was acquired in the backward direction was also close to the principal plane (relative azimuth = 25°); these two directions were therefore the most sensitive to the hotspot parameter.

3.5.3 Dimensionality

A dimensionality of 3 is commonly observed for mono-angular spectral data. It is much lower than the total number of spectral bands, because of spectrally contiguous information. This redundancy also appeared in the angular domain, where the dimensionality would theoretically be 2 x 3 = 6, but was effectively only 4. The data redundancy in the angular domain was also found by Simic and Chen (2008) for black spruce and aspen stands in Canada. Settle (2004) found that the dimensionality of hyperspectral multi-angular data was much lower than the number of observations, by performing a SVD on extensive lab measurements of vegetation canopies. In our study, however, the dimensionality assessment was based on the model sensitivity (Jacobian matrix) and did not take into account measurement or calibration errors in the CHRIS data.

3.5.4 Variable estimations

The same variable sampling scheme was used for both 4-variable and 7-variable LUT, leading to very disproportionate sizes of the LUTs. Four to seven sample points were used for all variables, which would be a sufficient sampling density in combination with linear interpolation between LUT entries according to Barnsley et al. (2000). Because prior knowledge was used to reduce the variable space, the interpolation was not considered necessary here.

In relation with the dimensionality results, the expectations regarding the variable estimations were that all angular combinations would be able to provide good estimates in the 4-variable case, and that the more directions used in the combination, the better the estimates in the 7-variable case. The results, however, showed that, in both cases, the best radiance match and the best estimates were obtained using two or

three angles only, and not with the full angular combination. No clear trend could be observed about the size of the solution space.

Over all LUTs, two angular combinations gave the best radiance match: Θ = {nadir} for the OLD1 stand, and Θ = {nadir, m36, p55} for the YOUNG and OLD2 stands. However, these combinations were able to give the best estimates only once: for the YOUNG stand in the 7-variable case. This finding supports Gascon et al. (2007) who pointed out the difficulty to fully exploit the information contained in multi-angular data, as they obtained worse estimation results when using multi-angular data for inverting the PROSAIL model. To improve the estimation performance, one might introduce angular weights to account for poorer simulations at very oblique directions, or use a more complex canopy model, but at the expense of higher LUT computational load. In addition, the CHRIS measurements were assumed error-free in this study. Including the measurement uncertainty in the inversion process, for example through Bayesian techniques (Lauvernet et al. 2008; Lavergne et al. 2007; Pinty et al. 2007; Timmermans et al. 2009), might help to improve the accuracy of the estimates. Another cause of low estimation performance might be compensation effects in the cost function. For example, a set of variable values having good radiance match in the forward direction and bad in the backward direction might give a similar χ value as another set of variable values having opposite angular radiance match characteristics. Finally, the quality of the reference data to validate the estimation results is often problematic. In our case, Cv was derived by classifying an airborne dataset and no in-situ measurements were available for D.

The pixels included in the plot of each stand were selected by hand from each of the four images, thus avoiding co-registration issues. The quality of the spatial pre-processing of the images, however, would be critical for any image-based multi-angular study (Ma et al. 2010).

3.6 Conclusion

This paper reported on the first case study of forest variable estimation conducted from measured multi-angular TOA radiance data by inverting a physically-based canopy-atmosphere RT model. The SVD of the Jacobian matrix showed that the dimensionality of the estimation problem increased from 3 to 6 when increasing the number of observation directions from one to four, thus revealing the higher potential of multi-angular data for estimating forest biochemical and biophysical variables. The Jacobian matrix was also used to identify the most influential model parameters, and only the most influential parameters were included as free variables in the LUTs. The analysis of the estimates obtained from all possible angular CHRIS combinations, however, pointed out the difficulty of fully exploiting the additional information contained in the multi-angular data.

The coupled SLC-MODTRAN model provided good simulations of the TOA radiance spectral and angular signatures of the three Norway spruce forest stands, as measured by CHRIS, thanks to the brown material and crown clumping features included in 4SAIL2. The canopy-atmosphere coupling based on the 4-stream RT theory allows making full use of the four canopy angular reflectance components provided by SLC, which benefits the simulation accuracy and therefore the estimation performance.

Acknowledgements

Chapter 4

A Bayesian object-based approach for estimating vegetation biophysical and biochemical variables from at-sensor APEX data

Abstract

Vegetation variables such as leaf area index (LAI) and leaf chlorophyll content (Cab) are important inputs for dynamic global vegetation models. LAI and Cab can be estimated from remote sensing data using either empirical or physically-based approaches. The latter are more generally applicable because they can easily be adapted to different sensors, acquisition geometries, and vegetation types. They estimate vegetation variables through inversion of radiative transfer models. Such inversions are ill-posed but can be regularized by coupling models, by using a priori information, and spatial and/or temporal constraints. Striving to improve the accuracy of LAI and Cab estimates from single remote sensing images, this contribution proposes a Bayesian object-based approach to invert at-sensor radiance data, combining the strengths of regularization by model coupling, as well as using a priori data and object-level spatial constraints.

The approach was applied to a study area consisting of homogeneous agricultural fields, which were used as objects for applying the spatial constraints. LAI and Cab were estimated from at-sensor radiance data of the Airborne Prism EXperiment (APEX) imaging spectrometer by inverting the coupled SLC-MODTRAN4 canopy-atmosphere model. The estimation was implemented in two steps. In the first step, up to six variables were estimated for each object using a Bayesian optimization algorithm. In the second step, a look-up-table (LUT) was built for each object with only LAI and Cab as free variables, constraining the values of all other variables to the values obtained in the first step. The results indicated that the Bayesian object-based approach estimated LAI more accurately ($R^2 = 0.45$) than a LUT with a Bayesian cost function (LUT-BCF) approach ($R^2 = 0.22$), and Cab with a smaller absolute bias.

The results of this study are an important contribution to further improve the regularization of ill-posed RT model inversions. The proposed approach allows reducing uncertainties of estimated vegetation variables, which is essential to support various environmental applications. The definition of objects and a priori data in cases where less extensive ground data is available, as well as the definition of the observation covariance matrix, are critical issues which require further research.

Keywords

4.1 Introduction

Vegetation is an important component of the Earth's biosphere and therefore plays an important role in the climate and carbon cycles (Foley et al. 2000). Vegetation variables such as leaf area index (LAI) and leaf chlorophyll content (Cab) are therefore essential inputs in dynamic global vegetation models. Remote sensing (RS) allows obtaining spatially continuous maps of vegetation variables at regular time intervals, including LAI (Baret et al. 2007; Myneni et al. 2002) and canopy chlorophyll content (Dash et al. 2010).

There are two main approaches to estimate vegetation variables from RS data. Empirical approaches are based on statistical relationships between the variables of interest and the RS data. Such statistical relationships require extensive field data, are specific to the vegetation type, development stage, study site and are only valid for a given sensor spectral configuration and acquisition geometry (Dorigo et al. 2007; Ustin et al. 2009).

Physically-based approaches rely on canopy radiative transfer (RT) models, which simulate the optical properties of a vegetation canopy using structural and biochemical variables (Goel 1988). They can easily be adjusted for various sensor types and acquisition geometries, and are therefore more general than empirical approaches (Darvishzadeh et al. 2011; Gemmell et al. 2002). To estimate vegetation variables, however, the canopy RT models have to be inverted. Look-up tables (LUT) are a common inversion method (Atzberger and Richter 2012; Darvishzadeh et al. 2008; Kimes et al. 2000; Soenen et al. 2009; Weiss et al. 2000) because of their simplicity, but also optimization (Combal et al. 2002; Lauvernet et al. 2008; Lavergne et al. 2007; Pinty et al. 2007), neural nets (Atzberger 2004), support vector machines (Durbha et al. 2007) and genetic algorithms (Fang et al. 2003) can be used. The inversion of canopy RT models is an under-determined and ill-posed problem because of i) the limited information content of the measured RS data (Jacquemoud et al. 2009), ii) uncertainties of models and measurements (Baret and Buis 2008; Combal et al. 2002) and iii) a similar response of radiometric signals to different variable combinations (Atzberger 2004). Several regularization methods can be implemented to reduce the ill-posedness of the inverse problem: model coupling, a priori information, and temporal and spatial constraints.

Model coupling allows reducing the number of input parameters in the forward model and therefore reduces the number of free variables during model inversion (Baret and Buis 2008). This efficiently reduces the risk of under-determined inversion problems because the number of retrieved variables is closer to the dimensionality of the data, which usually varies between 3 for mono-angular data and 6 for multi-angular data (Laurent et al. 2011c). Traditional model couplings involve soil, leaf and canopy RT models, but also atmospheric RT models can be

added to the coupling set-up. This allows estimating leaf and canopy variables directly from the top of atmosphere radiance data (Baret et al. 2006; Laurent et al. 2011b; Lauvernet et al. 2008; Verhoef and Bach 2003a), and also reduces the image pre-processing to only radiometric calibration, geometric correction, and possibly ortho-rectification. The complex atmospheric correction step, which has a number of limitations, is not needed and working directly with at-sensor radiance data using a coupled canopy-atmosphere model is therefore theoretically more accurate (Laurent et al. 2011b).

The use of a priori information allows avoiding unlikely variable combinations by reducing the variable space to a smaller subspace, thus facilitating the inversion (Combal et al. 2002; Li et al. 2001). A priori information can be collected from field data, expert knowledge, or agricultural databases, at various spatial scales such as agricultural fields or land cover types. It includes the knowledge of which variables can be kept fixed and which are important for a given level, and can be as extensive as the full prior probability distribution (prior) for each variable, or simply consist of the main statistics (e.g., mean, minimum, maximum, variance). When including the a priori information in the cost function, one obtains the Bayesian approach (Tarantola 2005). The Bayesian cost function can be used in combination with all inversion methods, but it has been most frequently used with optimization (Lavergne et al. 2007; Meroni et al. 2004; Pinty et al. 2007; Verhoef 2007; Yao et al. 2008). In addition to its use as regularization method, the Bayesian approach also provides an interesting framework for data assimilation (Lewis et al. 2012; Verhoef 2007).

Temporal constraints can be applied when several RS images from a time series are available. One approach is to use all the images concurrently in the inversion, constraining some variables to be constant over the acquisition time period (Houborg et al. 2007; Lauvernet et al. 2008). Another approach is to use the images sequentially in the inversion process, constraining the variation of some variables (usually LAI, because of its slower change rate) over the acquisition time period, either by the expected growth of the plants (CROMA 2000; Koetz et al. 2007), or by imposing temporal smoothness (Quaife and Lewis 2010).

Spatial constraints were introduced by Atzberger (2004) who demonstrated, using synthetic Landsat data, an increase of the inversion performance when using object statistics in addition to the pixel spectral signature. The appropriate object size depends on the variable of interest. For example, 3 x 3 windows may be used for soil brightness and the canopy hot spot parameter (Atzberger and Richter 2012), agricultural fields may be used for the leaf inclination distribution function (LIDF) (Atzberger and Richter 2012; Houborg et al. 2009), while the entire image may be suitable for atmospheric parameters (Lauvernet et al. 2008). It is also possible to implement an object size hierarchy to fine-tune the implementation of the spatial constraints (Atzberger and Richter 2012). Finally, similarly to the temporal case, spatial smoothness constraints can also be applied (Wang et al. 2008).

Spatial constraints are a recent development in the regularization of RT model inversion. All studies available in literature used either synthetic data (Atzberger 2004; Lauvernet et al. 2008), or satellite sensors with few spectral bands (Atzberger and Richter 2012; Houborg et al. 2009; Wang et al. 2008). Despite appropriate pixel sizes for measuring LAI and leaf chlorophyll content at their typical process length scales, and very high spectral resolutions, airborne imaging spectrometer data have not been used yet for investigating the potential of using spatial constraints.

This contribution proposes the use of a Bayesian object-based approach to invert a coupled canopy-atmosphere RT model to estimate vegetation variables from airborne imaging spectrometer at-sensor radiance data. The approach combines the strengths of three inversion regularization strategies, namely i) the coupling of models, ii) the incorporation of a priori data, and iii) the application of object-level spatial constraints. The approach was applied to an agricultural study area in Switzerland, using homogeneous fields serving as objects to apply the spatial constraints. LAI and leaf chlorophyll content (Cab) were estimated from at-sensor radiances of the Airborne Prism Experiment (APEX) imaging spectrometer (Jehle et al. 2010), using the coupled SLC-MODTRAN canopy-atmosphere RT model (Laurent et al. 2011b). The specific objectives of this study were to: 1) introduce the Bayesian object-based approach, 2) evaluate its performance for estimating LAI and Cab, and 3) highlight the added value of spatial constrains for model inversion by comparing the new approach against a LUT with Bayesian cost function (LUT-BCF) approach. The results of this study are an important contribution to further improve the regularization of ill-posed RT model inversions. The proposed approach allows reducing uncertainties in the estimates of vegetation variables, which is essential to support various environmental applications.

4.2 Materials and methods

4.2.1 Study area

The study area was located south of Basel, Switzerland, near the village of Oensingen, at +47° 16' 44" N, +7° 43' 53" E, and at 523 m above sea level (a.s.l.). The Oensingen test site is an agricultural area with flat topography and consists of 246 homogeneous and well-managed agricultural fields of eight crop types: winter wheat, corn, sugar beet, bean, grass, pea, rapeseed, and clover. The climate is temperate continental, with an annual rainfall of 1100 mm and an average annual temperature of 9°C.

4.2.2 APEX airborne imaging spectrometer data

APEX data were acquired on June 26th, 2010, around 11:30 a.m. local time, from nadir viewing, with a solar zenith angle of 27.1° and a solar azimuth angle of 146.4°.

The flight pattern consisted of a single flight line heading north, at 5000 m a.s.l. The resulting ground pixel size was approximately 2.2 m. APEX is a dispersive pushbroom imaging spectrometer covering the spectral region between 380 nm and 2500 nm in 313 contiguous spectral bands. The sampling interval varies between 0.55 nm and 10 nm, and the spectral resolution (full width half-maximum) varies between 0.6 nm and 11 nm, depending on wavelength (D'Odorico et al. 2011; Hüni 2009; Jehle et al. 2010). The data were radiometrically calibrated to obtain at-sensor radiances. The geometric correction was performed using PARGE (Schläpfer et al. 1998). Using 15 ground control points, the root mean square error (RMSE) was evaluated to be 2.4 m, which represents an accuracy of about one ground pixel. The standard deviation was evaluated to be 1.1 m, indicating a good spatial consistency of the accuracy. The at-sensor radiance image in raw sensor geometry was used for processing, in order to use the highest quality radiance data. The final outputs were geometrically corrected to be compared with the ground measurements.

4.2.3 Ground measurements

The ground measurements were acquired within 30 minutes of the APEX overflight. Extensive measurements were carried out in five fields: winter wheat (id 6), corn (id 36), sugar beet (id 37), pea (id 50), and clover (id 58) (Damm et al. 2010). In each field, four points, representative of an area of 2 m^2, were chosen along a transect following the longest direction of the field to measure canopy optical, biochemical, biophysical and structural properties. The largest field had an extent of 70 x 220 m, so the four points were considered to be representative of a field. Cab was non-destructively measured with a SPAD-502 chlorophyll-meter, using specific transfer functions for post-processing from previous campaigns (CEFLES2, FLUXPATH). During the CEFLES2 campaign (Rascher et al. 2009) an inter-comparison of two SPAD devices was performed on several crop types. The device-related uncertainty was evaluated to be at most 4 % (not shown), which was considered negligible. In homogeneous canopies (i.e., winter wheat, pea, sugar beet, clover), ten leaves were sampled at each point and five measurements were taken in the middle of each leaf. In heterogeneous canopies (i.e., corn), to account for vertical gradients (Ciganda et al. 2008; Winterhalter et al. 2012), five leaves were sampled in each of four canopy layers (total 20 leaves) for each point. Three measurements were then taken in each of the base, middle and top part (total 9 measurements) of each sampled leaf. The leaf dry matter (Cdm) and water (Cw) content were determined from the fresh and dry weight (oven drying at 75°C until constant weight) measurements of the sampled leaves in the laboratory. Effective LAI measurements were performed with a LI-COR LAI 2000 instrument under diffuse illumination conditions. Soil reflectance was measured using an ASD FieldSpec 3 spectrometer in a bean field, which had very low canopy cover (Cv = 0.1), between the rows of seedlings.

4.2.4 Coupled canopy-atmosphere RT model and parameterization

The soil-leaf-canopy model SLC (Verhoef and Bach 2007) was chosen to model crop canopies. It is a hybrid model with reasonable accuracy and short computation time, making it suitable for use in the optimization algorithm. SLC combines the 4SOIL soil reflectance model (Verhoef and Bach 2007) with the PROSPECT leaf optical properties model and the 4SAIL2 canopy reflectance model. The PROSPECT version used in this study is a robust version of PROSPECT-3 (Jacquemoud and Baret 1990), which uses the 1-nm specific absorption coefficients of PROSPECT-4 (Feret et al. 2008) and has been extended to include brown pigments (Cs) (Verhoef and Bach 2003b). The refraction index of the leaf tissue and the angle of incidence of incoming radiation of PROSPECT-3 were retained. 4SAIL2 is based on the SAIL model (Verhoef 1984), and also assumes arbitrarily inclined leaves, as described by two LIDF parameters: LIDFa, which controls the average leaf slope, and LIDFb, which controls the bimodality of the distribution. A fraction of the total LAI can be assigned to brown leaves (fB) (Verhoef and Bach 2003b). The vertical distribution of the brown leaves in the two canopy layers is controlled by the dissociation parameter (D): if D = 0, the brown leaves are homogeneously distributed, if D = 1, all the brown leaves are at the bottom of the canopy. The crown-level clumping is modelled using porous spheres and is controlled by the canopy cover (Cv), and the crown shape parameter (Zeta), calculated at the ratio of the crown diameter to the height of the crown centre above ground. This is an interesting feature to model young crops having discontinuous cover and where the individual plants are distinguishable. Finally, the semi-empirical hot spot parameter (Hot) controls the angular width of the hot spot reflectance peak in the angular domain.

The atmosphere was simulated using the MODTRAN4 RT model (Berk et al. 2003). It is a state-of-the-art model with rather high computational requirements, depending on spectral resolution and range. The canopy-atmosphere coupling approach of Verhoef and Bach (2003b), as adapted by Laurent et al. (2011b), however, requires only three runs for surface albedo values of 0, 0.5, and 1. These can be computed in advance. This approach avoided repeating time-consuming MODTRAN4 runs for each new radiance simulation, and allowed the coupled SLC-MODTRAN4 model to be used in the optimization. The canopy-atmosphere coupling relies on the 4-stream theory, making full use of the four directional components of the canopy reflectance as provided by SLC. SLC and MODTRAN4 simulations were performed at 1 nm spectral resolution and the final radiance output signature was convolved to the APEX bands using Gaussian response functions (D'Odorico et al. 2013).

The soil spectrum measured in the field was used instead of invoking the 4SOIL model. For all crop types, the brown leaf parameters were set as follows: Cab = 5 µg/cm^2, Cw = 0.001 cm, Cdm = 0.005 g/cm^2, Cs = 1, and the mesophyll parameter (N) was set to 2. MODTRAN4 was run between ground and APEX flight altitudes,

using the "mid-latitude summer" model atmosphere with "rural" aerosols and the 1 cm^{-1} band model. The visibility was set to 20 km, which is the value usually observed in summer in Switzerland. The water vapour was estimated per pixel using the algorithm of Richter and Schläpfer (2002). The average water vapour column over the study area was 2.070 g/cm^2. The default surface pressure (908.9 hPa) and the ozone column (7.07 g/m^2) of the "mid-latitude summer" model atmosphere were used. The surface temperature was set to 20°C. Soil spectrum, brown leaf parameters and atmospheric parameters were considered constant over the study area.

4.2.5 Bayesian object-based estimation of LAI and Cab

4.2.5.1 Objects, classification and a priori information

The study area is intensively managed and the agricultural fields were therefore assumed homogenous and used as objects to apply the spatial constraints. The 246 fields were digitized manually, avoiding mixed pixels at the edges. The image pixels were classified into the eight crop types present in the area using a maximum likelihood classifier. The classification was run on the first ten bands of a minimum noise fraction (MNF) transformed image to minimize the effects of image along-track striping. The crop type with the highest number of pixels in the object was taken as the object crop type.

Three levels of information were used to set the canopy and green leaf variables: crop types, crop field objects, and pixels. For each crop type, Cs, D, LIDFa, LIDFb, Zeta and Hot were fixed to the values indicated in Table 4.1 according to expert knowledge. The other variables were free in the Bayesian object-based optimization at the object level (step 1). Their a priori information was defined for each crop, using extensive expert knowledge (Table 4.2) accounting for the individual growth stage of the crops. For example, winter wheat and rapeseed were well-developed, with full canopy cover and fruits in the canopy, whereas beans were in an early stage, with only a few leaves per plant and very low canopy cover. Cv was fixed to 1 for well-developed crops, and kept free for younger crops, which had discontinuous cover and where the individual plants were distinguishable, because the porous spheres are the best approximation of discontinuous canopies available in the SLC

Table 4.1 Fixed parameter values for the SLC model for the eight crop types (W: winter wheat, Co: corn, S: sugar beet, B: bean, G: grass, P: pea, R: rapeseed, Cl: clover).

Parameter (unit)	W	Co	S	B	G	P	R	Cl
Cs (-)	0.05	0.2	0.05	0.05	0.05	0.05	0.05	0
D (-)	0	0	0	0	0	0	0	0
LIDFa (-)	-0.7	-0.35	0	0.8	-0.35	-0.35	-0.35	0.8
LIDFb (-)	0	-0.15	-0.8	0	-0.15	-0.15	-0.15	0
Zeta (-)	1	1	1	0.5	1	1	1	1
Hot (-)	0.05	0.05	0.2	0.05	0.05	0.05	0.05	0.05

Table 4.2 A priori values (v_a), standard deviations (σ_v) and ranges for each variable [v_{min}, v_{max}] in the Bayesian optimization, for the eight crop types (n/a: not applicable). For explanation of crop types see Table 4.1.

Variable (unit)	W	Co	S	B	G	P	R	Cl
LAI (-)	**2.5** (2) [0, 8]	**1.3** (1) [0, 8]	**2.3** (2) [0, 8]	**0.5** (1) [0, 8]	**3** (2) [0, 8]	**1.5** (2) [0, 8]	**2.5** (2) [0, 8]	**4.5** (2) [0, 8]
Cv (-)	**1** (n/a) [n/a, n/a]	**0.8** (0.2) [0.5, 1]	**0.8** (0.2) [0.5, 1]	**0.3** (0.2) [0.1, 0.5]	**1** (n/a) [n/a, n/a]	**0.8** (0.2) [0.5, 1]	**1** (n/a) [n/a, n/a]	**1** (n/a) [n/a, n/a]
Cab (μg/cm^2)	**45** (20) [0, 100]	**30** (20) [10, 100]	**40** (20) [10, 100]	**30** (20) [10, 100]	**30** (20) [10, 100]	**50** (20) [10, 100]	**45** (20) [10, 100]	**30** (20) [10, 100]
Cw (cm)	**0.04** (0.01) [0.005, 0.05]	**0.03** (0.01) [0.005, 0.05]	**0.035** (0.01) [0.005, 0.05]	**0.02** (0.01) [0.005, 0.05]	**0.02** (0.01) [0.005, 0.05]	**0.03** (0.01) [0.005, 0.05]	**0.04** (0.01) [0.005, 0.05]	**0.02** (0.01) [0.005, 0.05]
Cdm (g/cm^2)	**0.005** (0.001) [0.001, 0.01]	**0.003** (0.001) [0.001, 0.01]	**0.005** (0.001) [0.001, 0.01]	**0.004** (0.001) [0.001, 0.01]	**0.004** (0.001) [0.001, 0.01]	**0.003** (0.001) [0.001, 0.01]	**0.005** (0.001) [0.001, 0.01]	**0.004** (0.001) [0.001, 0.01]
N (-)	**1.3** (0.5) [1, 1.5]	**1.5** (0.5) [1, 1.5]	**2** (0.5) [1.5, 2.5]	**1.8** (0.5) [1, 2]	**1.5** (0.5) [1, 2]	**1.8** (0.5) [1, 2]	**1.3** (0.5) [1, 1.5]	**1.8** (0.5) [1, 2]
fB (-)	**0.2** (0.1) [0, 0.5]	**0** (n/a) [n/a]	**0** (n/a) [n/a]	**0** (n/a) [n/a]	**0.1** (0.1) [0, 0.3]	**0** (n/a) [n/a]	**0.2** (0.1) [0, 0.3]	**0** (n/a) [n/a]

model. fB was assumed smaller than 0.3 for grass and rapeseed, and smaller than 0.5 for winter wheat. Based on visual assessment on the field, the green and brown leaves were assumed vertically homogeneously distributed in the canopy, so the dissociation parameter D was set to zero. For the other crops, fB was zero, and D was therefore non-influential and set to zero. The a priori values for N were chosen according to expert knowledge and simulations with trial and error. Contrary to other studies using broad ranges for N in order to include various plant species or development stages (Atzberger 2004; Combal et al. 2002; Féret et al. 2011), N was assumed to vary little within a crop type, and smaller variation ranges were therefore used (Darvishzadeh et al. 2011). At the pixel level, only LAI and Cab were kept free, and all other variables were fixed to their object optimized value obtained in step 1 or to their value in Table 4.1 (step 2).

4.2.5.2 Bayesian optimization at the object level (step 1)

To find the maximum likelihood object estimates of LAI, Cab, Cw, Cdm, N, and eventually Cv or fB according to Table 4.2, the Bayesian cost function χ^2 was calculated as (Tarantola 2005):

$$\chi^2 = \tfrac{1}{2}(\mathbf{L_o} - \mathbf{L})^{\mathrm{T}} \mathbf{C_o}^{-1}(\mathbf{L_o} - \mathbf{L}) + \tfrac{1}{2}(\mathbf{v_a} - \mathbf{v})^{\mathrm{T}} \mathbf{C_a}^{-1}(\mathbf{v_a} - \mathbf{v}) \qquad (4.1),$$

where $\mathbf{L_o}$ is the vector of average observed APEX radiance of the object, $\mathbf{L}$ is the vector of simulated radiance, $\mathbf{C_o}$ is the covariance matrix containing the observation and model uncertainties in each spectral band, $\mathbf{v}$ is the vector of variable values, $\mathbf{v_a}$ is the vector of a priori variable values, and $\mathbf{C_a}$ is the covariance matrix of the a priori variables. The first term of the cost function is related to the difference between the model simulation and the observation (radiometric cost), whereas the second term is related to the difference between the variable values used in the model simulations and the a priori variable values (a priori cost). Each term is weighted by the inverse of its covariance matrix, which represents the Bayesian degree of belief. The measurement uncertainties for the a priori data and APEX observations were assumed uncorrelated, so the two covariance matrices were diagonal. The diagonal of the $\mathbf{C_a}$ matrix was filled with the a priori variance of the variables, calculated from Table 4.2. The noise of the APEX data was assumed to be equal to 1 $[\mathrm{mW}/(\mathrm{m}^2 \ \mathrm{sr} \ \mathrm{nm})]^2$ for all spectral bands, and the modelling errors were assumed negligible, so the $\mathbf{C_o}$ matrix was equal to the identity matrix. A detailed discussion on the $\mathbf{C_o}$ matrix can be found in section 4.4.4.2.

From the cost function, one can derive the following formula to calculate the update vector $\Delta\mathbf{v}$ to be applied to $\mathbf{v}$ in each iteration of the optimization algorithm:

$$\Delta\mathbf{v} = [\mathbf{J}^{\mathrm{T}}\mathbf{C_o}^{-1}\mathbf{J} + \mathbf{C_a}^{-1}]^{-1}[\mathbf{J}^{\mathrm{T}}\mathbf{C_o}^{-1}(\mathbf{L_o} - \mathbf{L}) + \mathbf{C_a}^{-1}(\mathbf{v_a} - \mathbf{v})] \qquad (4.2),$$

where **J** is the Jacobian matrix. In each column, **J** contains the partial derivatives of the model, obtained by varying each variable by 1 % of its allowed range (Laurent et al. 2011b). Leaving aside the terms related to the a priori data, Equation 4.2 boils down to the Gauss-Newton algorithm, which relies on the Jacobians to calculate the update vector and therefore only works well if **v** + **Δv** is within the neighbourhood of **v** where the model can be considered linear. A common problem when using the Gauss-Newton algorithm is that some iteration steps may lead to an increase in the cost function. A solution to this problem, as implemented in the Levenberg-Marquardt algorithm, is to introduce an extra term which allows rotating the update vector towards the direction of the steepest descent. In our case with a priori information, the formula to calculate the update vector becomes:

$$\mathbf{\Delta v} = [\mathbf{J}^T \mathbf{C_0}^{-1} \mathbf{J} + \mathbf{C_a}^{-1} + \mu \mathbf{J}^T \mathbf{J}]^{-1} [\mathbf{J}^T \mathbf{C_0}^{-1} (\mathbf{L_0} - \mathbf{L}) + \mathbf{C_a}^{-1} (\mathbf{v_a} - \mathbf{v})] \tag{4.3},$$

where μ is a damping factor controlling the weight of the steepest descent part of the formula. At the beginning of each optimization iteration, μ is set to an initial value of zero. If the calculated update vector leads to an increase in cost, μ is set to one. If the new value of μ again leads to an increase in cost, μ is multiplied by 10, and so on, until a decrease in cost is obtained, or until μ reaches 10^{10} (set as threshold).

The a priori knowledge about the variation range of each variable v was included in the optimization algorithm by constraining the variables between their minimum (v_{min}) and maximum (v_{max}) values throughout the iterations. This constraint was implemented by replacing each variable v by a transformed variable u:

$$u = \arcsin\left(2\frac{v - v_{\min}}{v_{\max} - v_{\min}} - 1\right) \tag{4.4},$$

which varies within the interval $[-\pi/2; \pi/2]$. The iteration step is then performed using the vector **u** of transformed variables, and **C$_a$** and **J** have to be expressed in terms of transformed variables. Equation 4.3 then allows calculating the update vector **Δu**. If some values in **u** + **Δu** are outside of the interval $[-\pi/2; \pi/2]$, the corresponding values in the original variable space are reshuffled within the interval $[v_{min}, v_{max}]$, because the inverse transformation involves the sine function which is periodical:

$$v = v_{\min} + \frac{1}{2}(v_{\max} - v_{\min})(\sin(u) + 1) \tag{4.5}.$$

The initial values of all variables were set to the centre of their a priori variation range, except LAI whose initial value was set to 2, where it is more influential. The optimization was terminated when the update values (Equation 4.3) for all the

variables became smaller than 1% of their a priori variation range, or when 10 iterations had been performed.

4.2.5.3 LUT inversion at the pixel level (step 2)

LAI and Cab were estimated at pixel level using LUTs and applying spatial constraints at object level. For each object, Cv, fB, Cw, Cdm, and N were fixed to the values obtained for the object in step 1, and the LAI and Cab sampling was tailored as described below.

First, the LAI sampling was determined because LAI has influence over the whole spectral domain, and has consequently more influence on the cost function than Cab, which only has influence in the range from 400 to 800 nm. LAI has the greatest influence (highest absolute value of the Jacobians in the local sensitivity analysis, not shown) in the range from 800 to 880 nm. Therefore, LAI_{min} and LAI_{max} were determined for each object using the pixels having lowest and highest mean radiance in the range from 800 to 880 nm. For each of these two pixels, LAI was estimated using the Bayesian optimization algorithm of step 1, starting from the object-optimized values (step 1), with a maximum of five iterations, and using only the range from 800 to 880 nm. LAI was then sampled between LAI_{min} and LAI_{max} using a regular sampling scheme. To ensure that the step size would have comparable magnitude for fields having different LAI ranges, the number of steps was increased from three steps when the LAI range was smaller than 0.5, up to eight steps when the LAI range was larger than 6. This sampling scheme ensures that the step size is smaller than 1 LAI unit, representing the usual LAI measurement accuracy.

Second, for each LAI value, the Cab range was determined according to the LAI-Cab scatterplot of all objects in the study area (step 1). The minimum Cab was set to 10 μg/cm^2 for LAI < 4, and to 15 μg/cm^2 for LAI $\geq$4. The maximum Cab increased from 60 μg/cm^2 for LAI < 1 to 100 μg/cm^2 for LAI $\geq$4. The Cab range was then sampled using a step size of 5 μg/cm^2.

The Bayesian cost values were calculated for each LUT entry (Equation 4.1), using the pixel signature as $\mathbf{L_o}$ and the LAI and Cab values obtained in step 1 as a priori information. The entry with the lowest cost was chosen as estimate of LAI and Cab.

4.2.6 LUT with Bayesian cost function (LUT-BCF) approach

To assess the added value of the object spatial constraints, the performance of the Bayesian object-based approach was compared to that of a pixel-based approach. Applying the Bayesian optimization for each pixel would be very computationally demanding, so a LUT-BCF approach was chosen. One LUT was built for each crop, using the crop-specific values given in Tables 4.1 and 4.2 to parameterize the coupled SLC-MODTRAN4 model. The variables LAI, Cv, Cab, and Cw were kept free within their a priori ranges and 25,000 combinations were randomly sampled

assuming uniform distributions between their a priori minimum and maximum values. The coupled model was run for each variable combination, filling the LUT with simulated at-sensor radiance signatures. For each pixel in the study area, the appropriate LUT was selected using its crop type. The Bayesian cost values were calculated for each LUT entry (Equation 4.1) using the pixel signature as L_o and the crop data of Table 4.2 as a priori information. The entry with the lowest cost was chosen as estimate of LAI, Cv, Cab, and Cw.
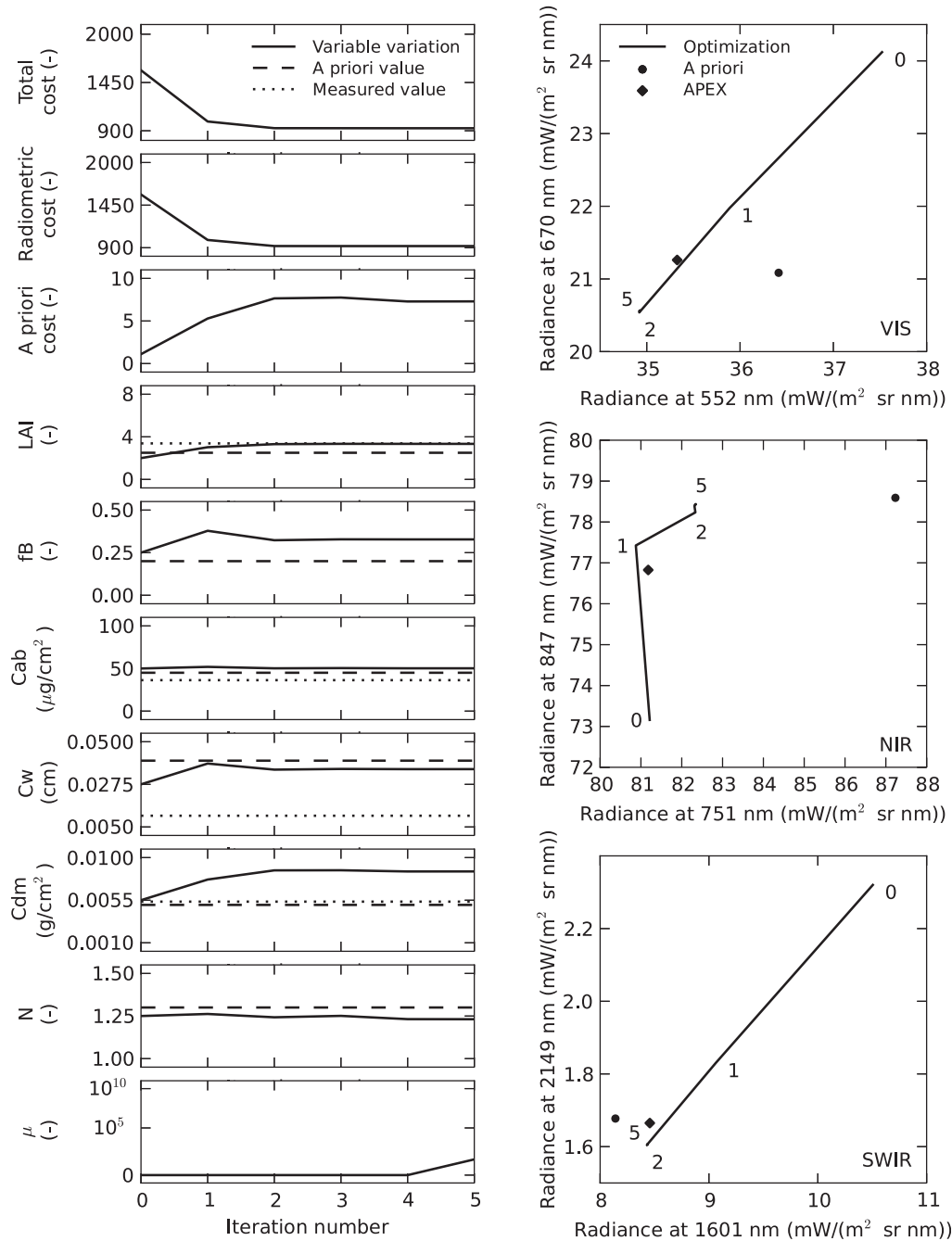
4.3 Results

4.3.1 Bayesian object-based approach

4.3.1.1 Bayesian optimization of a winter wheat object

The optimization procedure is described in detail for the example of a winter wheat field (object id = 6) for which ground measurements were available. The left side of Figure 4.1 shows the variation of the total, radiometric and a priori costs, the variation of the six free variables and the variation of the damping factor through the iterations of the optimization algorithm. Iteration 0 represents the starting point of the optimization, and the algorithm stopped at iteration 5 because the changes for all variables became smaller than 1% of their a priori variation range. The values at iteration 5 therefore are the estimates. The comparison with reference data (see Table 4.3) shows that good estimates were obtained for LAI and Cab, whereas Cw and Cdm were overestimated. Through the iterations, the total cost decreased from about 1600 to close to 900. The total cost consisted almost entirely of the radiometric cost, because the a priori cost was very low (< 8). This situation of very high radiometric cost and very small a priori cost was similar for all objects in the study area. The effect of constraining the variables in their a priori range cannot be seen in this example. After iteration 2, the changes in the variable values became small, until the damping factor (μ) increased between iterations 4 and 5, leading to the termination of the optimization.
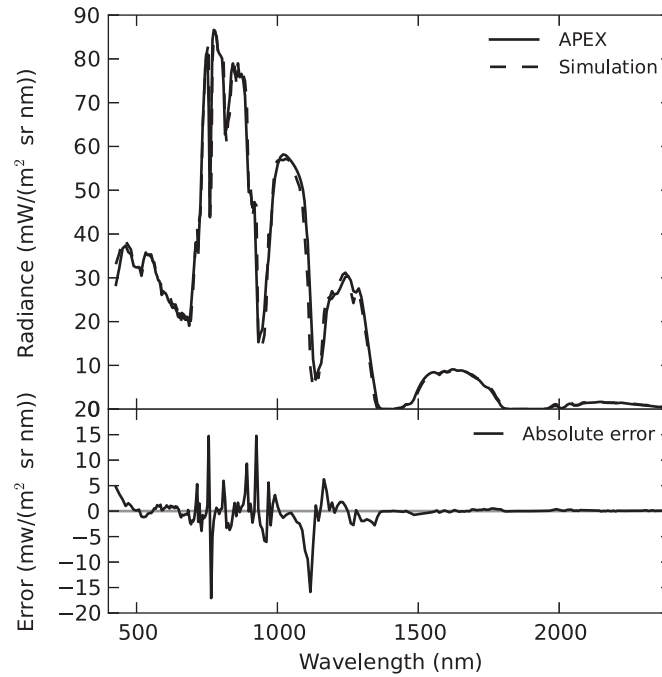
The right side of Figure 4.1 shows the variations of the simulated radiance through the optimization in visible (VIS), near-infrared (NIR) and shortwave infrared (SWIR) feature spaces. Iteration 1 brought the simulation closer to the APEX measurements in the three spectral domains. Iteration 2 provided an improvement in the SWIR, but not in the VIS and NIR. The small changes in the parameter values in iteration 3 to 5 caused small changes in the simulated radiance. No changes were visible in the VIS and SWIR, but the simulated radiance at 847 nm drifted slightly away from the APEX measurement. The variable changes in iterations 3 to 5, however, decreased the cost value, so that they must have provided improvements in spectral bands not shown in Figure 4.1.

The final match between APEX measurement and model simulation at the end of the optimization is presented in Figure 4.2. The overall agreement was very good, with slightly higher simulated radiances in the VIS and SWIR spectral domains. The

Figure 4.1 Variation of the costs, variables, damping factor μ (left) and radiance values (right) for a winter wheat object (id = 6) through the optimization iterations. The allowed variation range for the variables and damping factors were used as limits for the y axes of the corresponding plots. In the right part of the figure, the numbers refer to iterations 0, 1, 2 and 5, and VIS, NIR and SWIR refer to the visible, near-infrared and shortwave infrared domains.

Figure 4.2 Radiometric match and absolute radiometric error at the end of the Bayesian optimization of a winter wheat object (id = 6).

absolute error was less than 1 mW/(m^2 sr nm) in the SWIR, and less than 5 mW/(m^2 sr nm) in the VIS. The obvious spikes in the error curve in the NIR domain have no apparent trend to over- or under-simulation.
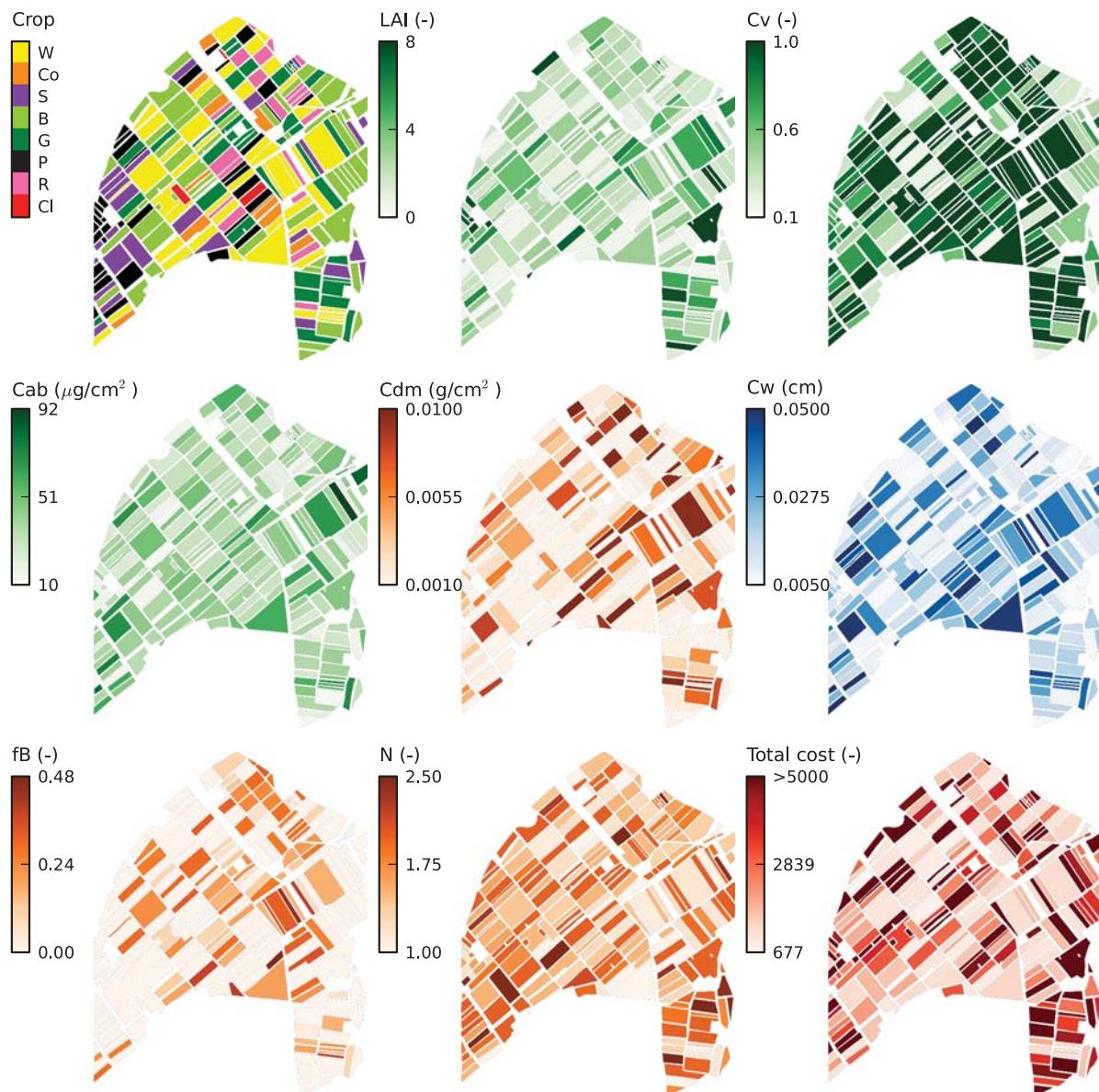
4.3.1.2 Object-level results

The results obtained from the Bayesian optimization at object level are presented in Figure 4.3. The object-level crop classification had an overall accuracy of 79 %. The misclassifications concerned mostly the fields having low canopy cover, such as short grass, bean and pea. For all variables, the estimated values covered the entire a priori ranges, except for very high values of Cab (Cab$_{max}$ was 100 μg/cm^2) and fB (fB$_{max}$ was 0.5). A visual check with the true colour APEX image (not shown) revealed that the pattern of fields having low estimated Cv is realistic. Further, very high LAI values are mostly found in winter wheat objects, but some are found in objects having low Cv values and may therefore be erroneous. The leaf variables show a lot of variation from object to object, but it is not possible to visually evaluate the plausibility of the results. The cost values are a good proxy to judge the estimation performance. Low costs indicate a good radiometric match and a reliable estimate and vice versa. More than half of the study area had costs lower than 2839 and only a few objects had very high costs indicating unrealistic estimates. These were mainly fields having low canopy cover (short grass, bean and pea).

The object estimates of LAI, Cab, Cdm and Cw were compared with the average of the four point measurements taken in the corresponding field. Table 4.3 shows that

LAI was estimated most accurately, followed by Cw, and Cdm. Cab was poorly estimated. A positive bias was observed for LAI and Cab.

4.3.1.3 Pixel-level results: LAI and Cab estimates

The LAI, Cab and cost maps obtained at pixel level from the object-based LUTs (Figure 4.4) present similar spatial patterns to those observed at the object level (Figure 4.3), but intra-object variability appears. This intra-object variability was
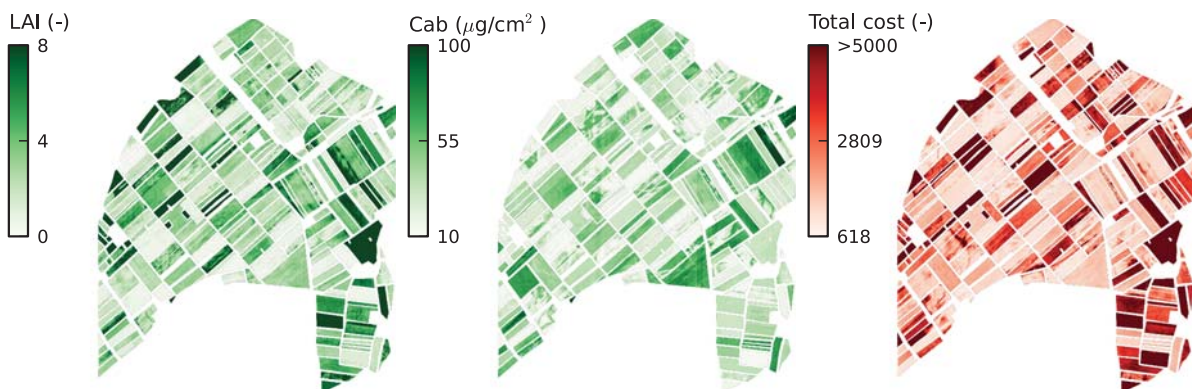

Figure 4.3 Object crop map, maps of the estimates of the seven variables at the object level, and map of the total cost, obtained from step 1 of the Bayesian object-based approach (raw image geometry, north towards the top).

Table 4.3 Comparison of the object-level estimates (step 1 of the Bayesian object-based approach) with the ground measurements (averages of the four point measurements taken in each field). For explanation of crop types see Table 4.1.
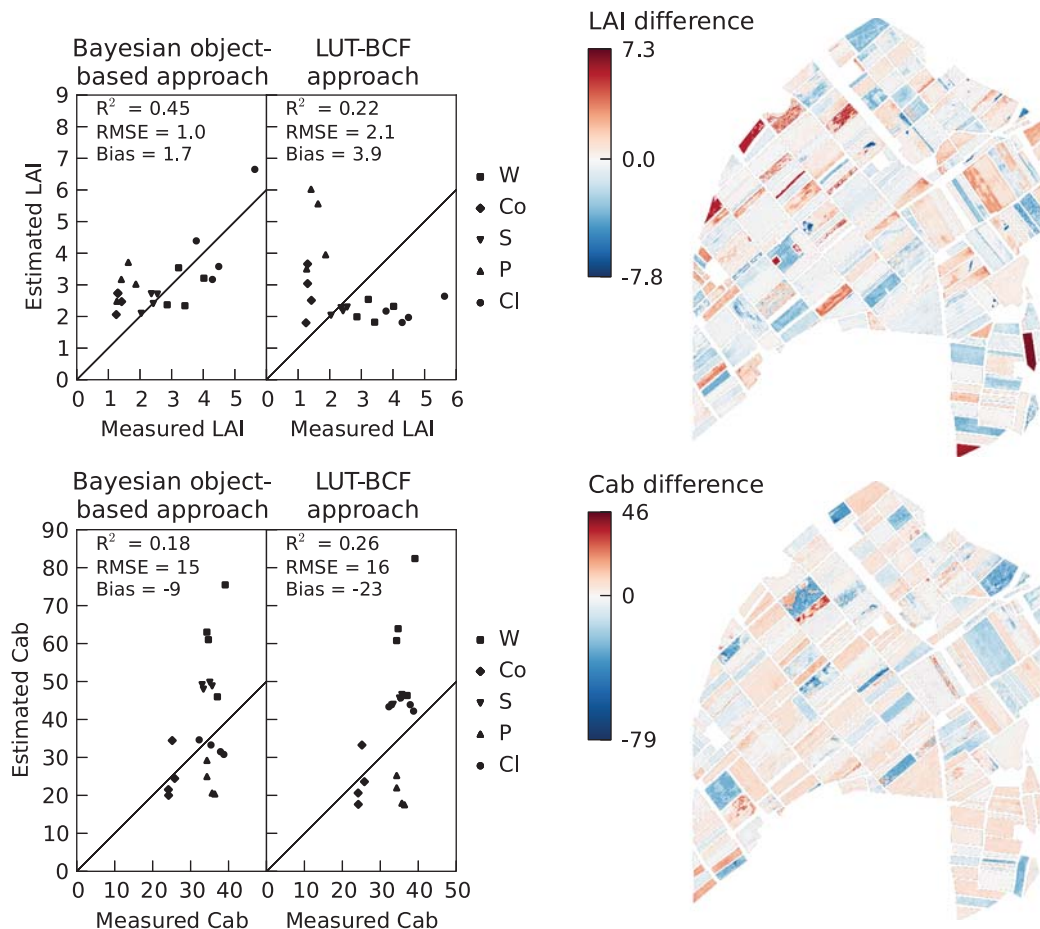
Crop	Object id	LAI (-)		Cab (μg/cm^2)		Cdm (g/cm^2)		Cw (cm)	
		Measured	**Estimated**	**Measured**	**Estimated**	**Measured**	**Estimated**	**Measured**	**Estimated**
W	6	3.4	3.3	36	50	0.005	0.009	0.01	0.004
Co	36	1.3	1.9	25	32	0.003	0.003	0.01	0.02
S	37	2.3	2.1	34	48	0.005	0.004	0.03	0.03
P	50	1.5	2	35	39	0.003	0.004	0.02	0.02
Cl	58	4.5	3.8	36	27	0.004	0.007	0.01	0.01
R^2		0.95		0.11		0.36		0.68	
RMSE		1.1		23		0.005		0.012	
Bias		1.0		16		-0.0006		0.002	

expected and is related to small scale variations due to management practices of the farmers or due to local variation in the soil conditions. Another spatial pattern appearing on the pixel-level Cab map was a very narrow vertical striping. This effect is typical for pushbroom sensors like APEX and can be explained by small deviations of the response of detector elements across track. In the objects, the cost values are higher in areas with less vegetation, indicating a higher estimation uncertainty. A closer look at this effect reveals that objects presenting a wide range of Cab values are typically characterized by a low LAI, whereas objects with higher LAI have a lower variability in Cab. This reflects the smaller sensitivity to Cab in canopies having a smaller LAI.

The estimates were compared with the 20 available ground measurements using 3 by 3 averaging windows. The comparison (left plots in Figure 4.5) shows that the quality of the estimates was rather poor and that LAI ($R^2 = 0.45$) was estimated more accurately than Cab ($R^2 = 0.18$).

Figure 4.4 Maps of the LAI and Cab estimates and of the total cost obtained from step 2 of the Bayesian object-based approach (raw image geometry, north towards the top).
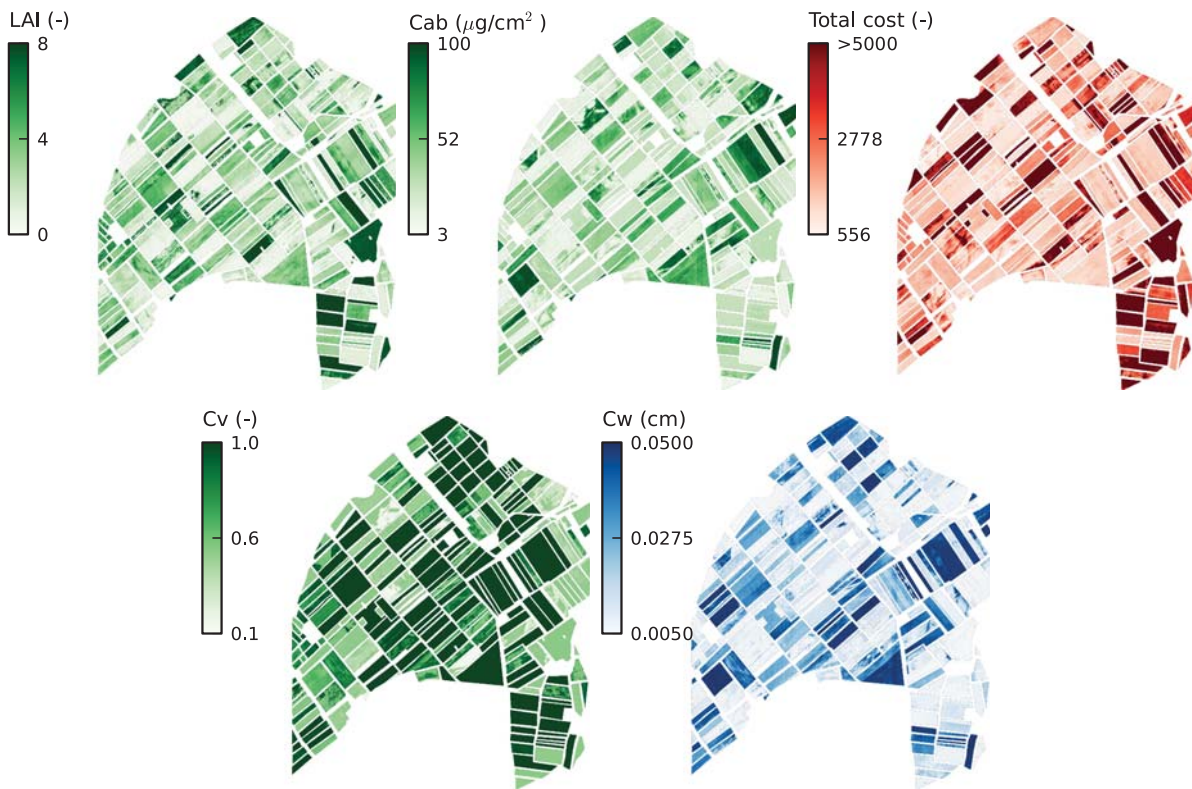
Figure 4.5 Comparison of the LAI and Cab estimates obtained from the two approaches. The plots on the left side present the comparison of the estimates with the ground measurements (the line is the 1:1 line), and the maps on the right side present the absolute difference between the estimates obtained from the two approaches, red indicates that the estimate from the Bayesian object-based approach had a higher value than that of the LUT-BCF approach (raw image geometry, north towards the top). The LAI values are unitless, and the Cab values are in μg/cm^2.

4.3.2 LAI and Cab from the LUT-BCF approach

The output maps of the LUT-BCF approach (Figure 4.6) show inter- and intra-object variability. The overall patterns of LAI and Cv correspond quite well with the vegetation patterns seen on the true colour APEX image, with some exception of objects having high LAI and very low Cv. The comparison of the estimates to the ground measurements (right plots of Figure 4.5) shows that the quality of the estimates was poor and that Cab ($R^2 = 0.26$) was estimated more accurately than LAI ($R^2 = 0.22$).

Figure 4.6 Maps of the estimates of LAI, Cv, Cab and Cw and of the total cost, obtained from the pixel-based LUT-BCF approach (raw image geometry, north towards the top).

4.3.3 Comparison of the two approaches

A visual comparison of the output maps of the LUT-BCF approach (Figure 4.6) with the output maps of the Bayesian object-based approach (Figure 4.4 for LAI, Cab and total cost, and Figure 4.3 for Cv and Cw) shows that the obtained spatial patterns are similar. The R^2 between the two estimated LAI maps is 0.78, and reaches 0.92 between the two Cab maps. This similarity is supported by the difference maps calculated for LAI and Cab (Figure 4.5). Most of the study area has differences close to zero, especially for Cab. Two objects in the south east of the study area, however, show high differences in LAI, mostly caused by misclassifications.

The LAI plots obtained from the two approaches show different trends. The LUT-BCF approach strongly overestimated low LAI values and strongly underestimated high LAI values. By contrast, the Bayesian object-based approach provided more realistic estimates closer to the 1:1 line. For this reason, the R^2 of the Bayesian object-based approach is higher than the R^2 of the LUT-BCF approach, indicating a better performance and higher accuracy of the Bayesian object-based approach. In addition, the RMSE and bias of the Bayesian object-based approach are smaller than those of the LUT-BCF approach. The Cab plots obtained from the two approaches are very similar, which is consistent with the very high R^2 observed between the two

Cab maps. Despite its smaller R^2 value, the Bayesian object-based approach had a bias closer to zero (- 9 µg/cm^2) than the LUT-BCF approach (- 23 µg/cm^2).

4.4 Discussion

We structure the discussion starting with the most general findings first, and then focusing on aspects requiring more detailed descriptions.

4.4.1 Bayesian object-based approach versus LUT-BCF approach

The performance of the two approaches in terms of R^2 or RMSE values was poorer than reported for grasslands (Darvishzadeh et al. 2011; Darvishzadeh et al. 2008) and maize (Yao et al. 2008). The improvement of the performance of the estimation when including spatial constraints, however, supports the results of previous studies on synthetic (Atzberger 2004; Lauvernet et al. 2008) and actual data (Atzberger and Richter 2012). The better performance of the Bayesian object-based approach over the LUT-BCF approach can be explained by two main factors. First, the higher number of free variables in step 1 allows improving the values for more variables, thus reducing the sensitivity to crop misclassifications and use of erroneous a priori values. Second, the tailoring of the LUTs of step 2 to each object effectively avoids unrealistic values of LAI and Cab and variable compensation effects that occur in the LUT-BCF approach, as for example between LAI and Cv for two objects in the south-east of the study area which had very high LAI differences (Figure 4.5). Finally, one can note that the different ranges used for the variables in the LUTs of the two approaches did not affect the cost values directly, because only the mean and variance of the a priori variables are used in the cost function.

4.4.2 Objects and spatial constraints

The spatial constraints were implemented using the concept of variable smoothness (Wang et al. 2008) strictly, imposing that all parameters and variables, except LAI and Cab, have the same values in all pixels of a given object. This approach was also used in other studies, but for LIDF (Atzberger and Richter 2012), and additionally N and a clumping parameter (Houborg et al. 2009). The object definition depends on the study area, on the variables of interest and on their associated assumptions (constraints). In other studies where less strict spatial constraints were desired, smaller objects (e.g., 3 x 3 pixel windows) were added to allow smooth spatial variation of some variables inside the fields (Atzberger and Richter 2012). Larger objects (e.g. a few km^2) were also used for atmospheric parameters (Lauvernet et al. 2008). For operational implementation, an automatic segmentation algorithm may be used to obtain the objects (Blaschke 2010).

4.4.3 Assigning the a priori data

Assigning the correct crop type and associated a priori data to each object is crucial for a good estimation performance (Houborg and Boegh 2008; Yang et al. 2006b). The good accuracy of 79 % of the object crop classification, however, is likely to decrease towards the west and east edges of the image because of lack of ground truth in those areas, and because of view angle effects in the APEX data. Misclassifications could partly explain the very high cost values observed in Figures 4.3, 4.4, and 4.5. Very detailed a priori data could be used because of the availability of extensive ground data and of the homogeneity of the study area in terms of sowing date. In general, however, the availability of a priori knowledge may be problematic and labour-intensive, because of local spatial variations of crop development stages, and because of the dependency of the a priori values on the time of year (Ciganda et al. 2008). In other situations, where the crop types may not be known in advance, broader classes such as several levels of vegetation greenness may be used (Dorigo et al. 2009).

4.4.4 Bayesian optimization algorithm

4.4.4.1 Inversion algorithm

Inversion of the PROSPECT + SAIL model by downhill simplex numerical optimization was found to be more accurate than inversion by artificial neural networks or LUTs (Vohland et al. 2010). This may be thanks to the full use of the sensitivity of the model variables contained in the Jacobian matrix, as illustrated in Yao et al. (2008). The Jacobian matrix was also fully exploited in the Bayesian optimization algorithm.

Contrary to other studies, which relied on a LUT inversion at the object-level (Atzberger and Richter 2012; Houborg et al. 2009), LIDFa and LIDFb were kept fixed because they caused confusions with other variables in the optimization. These parameters have a strong influence on the radiance signal and they were therefore carefully set. It may be possible to include them as free variables in the optimization if fixing other variables (e.g., N, fB or Cv) or if multi-angular observations are available (Migdall et al. 2009).

4.4.4.2 Observation covariance matrix

In theory, the C_o matrix contains both observation and model uncertainties (Kimes et al. 2000; Li et al. 2001). Model uncertainties are very difficult to assess because they depend on the parameter values and because of the lack of perfect reference observations (Widlowski et al. 2007), and they are therefore usually ignored. Observation uncertainties arise from multiple factors (Widlowski et al. 2001; Yao et al. 2008) related to the instrument (e.g. sensitivity, spectral shifts, stability in time) and to the pre-processing of the data (e.g. radiometric calibration,

geometric correction), and are therefore complex to evaluate. Most studies avoid these problems by summarizing $\mathbf{C_o}$ and $\mathbf{C_a}$ matrices into a single weight vector containing the ratios of information contained in the observations to the information contained in the a priori data for each variable (Combal et al. 2002; Li et al. 2001; Meroni et al. 2004). These ratios are based on expert knowledge and vary between 0.25 and 1, giving equal or more weight to the a priori data. Other approaches used a diagonal $\mathbf{C_o}$ matrix, assuming uncorrelated uncertainties with a constant noise percentage level for the observations (Lauvernet et al. 2008; Lavergne et al. 2007; Pinty et al. 2007). Finally, a recent approach used the covariance matrix of all pixels belonging to each land cover type separately (Dorigo et al. 2009). Such a matrix, however, also includes the noise due to the spatial variation of the environment.

The approach using a weight vector to balance radiometric and a priori information is subjective, and, to our opinion, gives too much weight to the a priori data. On the other hand, approximating $\mathbf{C_o}$ by the variance matrix of the vegetation pixels of the APEX image (not shown) gave very high weight to the observations (in the order of 10^6), with emphasis on the SWIR bands because they have a smaller variance. Consequently, the optimized radiance matched the APEX observation very well in the SWIR, but the match in the visible and NIR ranges was poor, leading to erroneous estimates of LAI and Cab. This is why a compromise between these two approaches was chosen.

4.4.4.3 Effectiveness of the use of the a priori information

The much lower magnitude of the a priori cost, as compared to the radiometric cost, means that the a priori data has virtually no influence on the cost function. Consequently, the optimization is almost entirely driven by the radiometric match between model simulation and APEX data. In practice, however, the a priori information prevents the possibility of a singular matrix in the inverted term of Equation 4.2, and allows strictly constraining the variables in their ranges. Contrary to the suggestion of Li and Strahler (1996), the a priori values were not used for initializing the optimization in order to have an objective initialization of the algorithm. No assumption was made on the prior distribution of the variables, making the Bayesian optimization more flexible and objective. This, however, prevents obtaining the posterior distribution of the variables and the posterior confidence intervals for the estimated values.

4.4.5 Model and data related issues

4.4.5.1 Model choice

Yao et al. (2008) found that the ROW model, a RT model specifically designed for row crops, performed better than the SAIL model for estimating vegetation variables in row-planted corn. The results of our study, however, showed that the

coupled SLC-MODTRAN4 model was able to provide adequate radiance simulations for all crop types and good estimates for the reference objects where ground measurements were available. This may be because of the clumping feature included in SLC to extend the capabilities of SAIL to discontinuous canopies. SLC, like SAIL, models the vegetation canopy as a turbid medium, which may explain why the crown shape parameter and hot spot parameter were important, despite the nadir viewing of APEX. In addition, the high estimated fB values (up to 0.48 for a few objects) compensate for non-leaf elements such as stems and fruits, which have a smaller transmittance than leaves and which are not included in the SAIL model.

The use of a single soil spectral signature and constant atmospheric parameters might have limited the performance of the coupled SLC-MODTRAN model. These assumptions were, however, considered reasonable for the small extent of the study area. The soil signature, which was measured in a bean field having very low Cv and LAI, was slightly contaminated by vegetation and showed a very faint red-edge. This may explain the higher cost values in the areas of some objects which had smaller LAI (Figure 4.4).

4.4.5.2 Imaging spectrometer APEX data

Mono-angular multispectral data usually have a dimensionality of 3, and the imaging spectrometer APEX data used here had a dimensionality of 4 (the first 4 principal components of the image, masked to use only the 246 investigated fields, accounted for 99 % of the variance). The Bayesian optimization, however, was able to estimate up to six variables. This high performance is related to the use of full radiance signature vectors in Equation 4.2, which allowed taking full advantage of the very high spectral resolution of APEX. This, in turn, enabled a fine distinction of the variables thanks to their detailed sensitivity profiles as contained in the Jacobian matrix.

The APEX data suffer from slight along-track striping and were not smile-corrected. Small spectral shifts may therefore explain the spikiness of the absolute error curve observed on Figure 4.2.

4.4.5.3 Ground measurements

Four measurement points per field were considered to be sufficient to represent within-field inhomogeneities and to compare the performance of the two approaches. To thoroughly evaluate the added value of the spatial constraints, however, it would be desirable to have a dense and regular spatial sampling scheme in fields of varying heterogeneity. Furthermore, LAI indirect measurements can suffer from uncertainties up to about 1 LAI unit (Verger et al. 2009). The Cab measurements suffered from errors introduced by the use of specific transfer functions from other field campaigns. In addition, in the cut leaf samples, Cab decreases over time, from the field to the lab, leading to underestimation of in vivo Cab. Finally, despite all precautions taken,

the leaves may also have dried during transport, leading to inaccurate, smaller Cw measurements. Therefore, LAI values should be considered with a confidence interval of 1 LAI unit, and in vivo Cab and Cw values are expected to be higher than the measured values, which may partly explain the overestimation of Cab and Cw.

4.5 Conclusions

Striving to improve the accuracy of LAI and Cab estimates from single RS images, this contribution proposed a new approach combining the strengths of model coupling, inclusion of a priori data, and object-level spatial constraints in order to regularize the inverse problem as much as possible. Inverting the coupled SLC-MODTRAN4 canopy-atmosphere model against APEX at-sensor radiance data, this novel Bayesian object-based approach improved the LAI estimates ($R^2 = 0.45$) as compared to using a pixel-based LUT-BCF approach ($R^2 = 0.22$), and estimated Cab with a smaller absolute bias.

The results of this study are an important contribution to further improve the regularization of the ill-posed RT model inversion. The proposed approach allows reducing uncertainties of estimated vegetation variables, thus increasing the quality of RS vegetation products, which is essential to support various environmental applications.

Two main issues were identified as critical and require further research. The first is to improve the strategy for defining objects and assigning them a land cover class with associated a priori data. This is essential to be able to apply the proposed method in areas with limited ground data. The second issue concerns a better implementation of the observation covariance matrix, which is critical to properly balance the radiometric information versus the a priori information.

Acknowledgements

Chapter 5

Bayesian object-based estimation of LAI and chlorophyll from a simulated Sentinel-2 top-of-atmosphere radiance scene

Abstract

Leaf area index (LAI) and chlorophyll content (Cab) are important vegetation variables which can be monitored using remote sensing (RS). Physically-based approaches have higher transferability and are therefore better suited than empirically-based approaches for estimating LAI and Cab at global scales. These approaches, however, require the inversion of radiative transfer (RT) models, which is an ill-posed and underdetermined problem. Four regularization methods have been proposed, allowing finding stable solutions: 1) model coupling, 2) using a priori information (e.g. Bayesian approaches), 3) spatial constraints (e.g. using objects), and 4) temporal constraints. For mono-temporal data, only the first three methods can be applied.

In an earlier study, we presented a Bayesian object-based algorithm for inverting the SLC-MODTRAN4 coupled canopy-atmosphere RT model, and compared it with a Bayesian LUT inversion. The results showed that the object-based approach provided more accurate LAI estimates. This study, however, heavily relied on expert knowledge about the objects and vegetation classes. Therefore, in this new contribution, we investigated the applicability of the Bayesian object-based inversion of the SLC-MODTRAN4 model to a situation where no such knowledge was available.

The case study used a 16 x 22 km^2 simulated top-of-atmosphere scene of the upcoming Sentinel-2 sensor, covering the area near the city of Zurich, Switzerland. Seven APEX radiance images were nadir-normalized using the parametric Li-Ross model, spectrally and spatially resampled to Sentinel-2 specifications, geometrically corrected, and mosaicked. The atmospheric effects between APEX flight height and top-of-atmosphere level were added using the MODTRAN4 model. The vegetation objects were automatically identified and extracted using a segmentation algorithm, and classified in four levels of brightness in the visible domain. The LAI and Cab maps obtained from the Bayesian object-based inversion of the coupled SLC-MODTRAN4 model presented realistic spatial patterns. The impact of the parametric Li-Ross nadir-normalization was evaluated. The differences between the LAI and Cab maps obtained from the data with or without nadir normalization increased from the centre towards the edges of the across-track direction. The results of this study contribute to preparing the RS community for the arrival of Sentinel-2 data in the near future, and generalize the applicability of the Bayesian object-based approach for estimating vegetation variables to cases where no field data are available.

5.1 Introduction

Global climate and carbon cycles are strongly influenced by the Earth's biosphere, and in particular by its vegetation component. Vegetation variables, such as leaf area index (LAI) and leaf chlorophyll content (Cab), are therefore important inputs in dynamic global vegetation models (DGVM) (Foley et al. 2000). These vegetation inputs can be provided in a spatially continuous way and at global scale by satellite remote sensing (Bacour et al. 2006a; Baret et al. 2007; Myneni et al. 2002).

Usually, remote sensing data are first atmospherically corrected to top-of-canopy (TOC) reflectance data before they are used for estimating the vegetation variables. The variables can be estimated by using two main approaches. Empirical approaches rely on statistical relationships between the vegetation variables and the TOC reflectance data. The statistical relationships, however, require extensive field data collection and are only valid for the specific conditions for which they were developed, including sensor, acquisition geometry, and vegetation type (Dorigo et al. 2007; Ustin et al. 2009). Physically based approaches rely on vegetation canopy reflectance models, which are mostly based on radiative transfer (RT) theory, and are therefore more general because they can be adapted for different sensors, acquisition geometry and be parameterized for various vegetation types (Gemmell et al. 2002; Malenovský et al. 2008).

5.1.1 Physically-based estimation of vegetation variables

To estimate the vegetation variables from the TOC reflectance data, however, the canopy RT model has to be inverted. This inverse problem is ill-posed (Combal et al. 2002; Jacquemoud et al. 2009), and four types of regularization methods have been proposed (Baret and Buis 2008): 1) coupling models, 2) using a priori data, 3) using spatial constraints and 4) using temporal constraints, or combinations of these.

Model coupling allows reducing the number of input parameters, thereby reducing the under-determined nature of the inversion. The maximum model coupling set-up involves soil, leaf, canopy, and atmosphere RT models (Baret and Buis 2008). Using such a coupled model allows working directly with the top-of-atmosphere (TOA) radiance data, skipping the atmospheric correction step (Laurent et al. 2011b). The atmospheric correction requires inverting the atmospheric RT model, whereas, when working at TOA level, the atmospheric RT model is used in forward mode, which is more accurate and allows for better inclusion of canopy directional effects (Laurent et al. 2011c), topography and adjacency effects in the coupled canopy-atmosphere model.

A priori information allows restricting the variable space to a smaller subspace, thus facilitating the inversion (Combal et al. 2002). Bayesian approaches use the a priori data directly in the cost function, and have been widely used for estimating

vegetation variables (Lavergne et al. 2007; Li et al. 2001; Pinty et al. 2007). Spatial constraints allow using the information contained in the neighbouring pixels in the inversion (Atzberger 2004; Atzberger and Richter 2012; Houborg et al. 2009), while temporal constraints allow using the information contained in a time series of remote sensing observations (Kötz et al. 2005; Lauvernet et al. 2008).

For a single RS image, the maximum regularization set-up involves a coupled canopy-atmosphere RT model, a priori information, and spatial constraints (Laurent et al. In review-b).

5.1.2 Sentinel-2

Sentinel-2 is a scheduled multispectral and high spatial resolution mission which is part of the Global Monitoring for Environment and Security (GMES) program (Berger et al. 2012; Drusch et al. 2012; Malenovský et al. 2012). The spectral and spatial characteristics for the Sentinel-2 mission have been specified so as to provide enhanced continuity for SPOT and Landsat missions in the visible (VIS), near infrared (NIR) and short-wave infrared (SWIR) spectral domains. The Multi Spectral Instrument (MSI) on-board Sentinel-2 will have 13 spectral bands in the range from 400 to 2400 nm, with pixel sizes of 10, 20, or 60 m, depending on the spectral band (Drusch et al. 2012; Sentinel-2 PDGS Project Team 2011). The first of two satellites is planned to be launched in 2014. Until Sentinel-2 data are available, several studies have investigated the potential of Sentinel-2 for vegetation applications.

Most of the Sentinel-2 exploratory studies focussed on the spectral dimension, selecting appropriate bands from surrogate sensors such as CHRIS (Atzberger and Richter 2012; Delegido et al. 2011) and HyMap (Richter et al. 2012), or convolving the bands of hyperspectral sensors such as CASI (Richter et al. 2011a; Richter et al. 2009) or field spectrometers (Clevers and Gitelson 2012; Herrmann et al. 2011) to the Sentinel-2 bands. Limited by the spectral range of the surrogate sensor used, most of these studies were not able to simulate the blue and the SWIR Sentinel-2 bands in full.

Further, few studies included the varying pixel size in their simulated Sentinel-2 data (Richter et al. 2011b), and even fewer investigated the potential of the spatial characteristics of the Sentinel-2 data (Hedley et al. 2012; Verrelst et al. 2012). Only two studies made use of top-of-atmosphere simulated Sentinel-2 data, and they focussed on cloud detection and correction (Hagolle et al. 2010; Richter et al. 2011b), but not on vegetation.

Therefore, despite its potential use for supporting the development of (pre)processing algorithms in advance, a full TOA simulated scene compliant with all spectral and spatial characteristics of the Sentinel-2 mission so far was still missing.

5.1.3 Objectives

The two main objectives of this study were to: 1) build a realistic TOA Sentinel-2 scene with full spectral and spatial characteristics as specified in the Sentinel-2 documentation, and 2) estimate LAI and Cab from the Sentinel-2 scene by inverting of a coupled canopy-atmosphere RT model.

The Sentinel-2 scene was built using seven flight lines of the APEX imaging spectrometer covering the area around the city of Zurich, Switzerland, which includes a wide range of land cover types (e.g., agriculture, forest, lakes, an airport and urban areas). The APEX images were normalized to nadir viewing before being spatially and spectrally resampled and mosaicked to obtain a simulated Sentinel-2 scene covering 16 x 22 km^2. In order to obtain the most accurate LAI and Cab estimates as possible from this single scene, the Bayesian object-based approach of (Laurent et al. In review-b) was chosen, because it combines the strengths of model coupling, a priori data and spatial constraints regularization methods. This latter study, however, relied on manual digitization of the objects used to apply the spatial constraints and was based on extensive field data on vegetation classes and their associated a priori data. Therefore, two specific objectives were added to the main objectives: 1) evaluate the effect of the normalization to nadir viewing of the APEX images, and 2) propose an automatic image-based approach for extracting objects, and a general vegetation classification associated with a priori data which does not require field data.

The results of this study contribute to preparing the RS community for the arrival of Sentinel-2 data in the near future, and generalize the applicability of the Bayesian object-based approach of Laurent et al. (In review-b) for estimating vegetation variables to cases where no field data are available, as is generally the case for studies in less accessible regions as well as global studies.

5.2 Materials and methods

5.2.1 Study area and APEX data

The study area covered the city of Zurich, Switzerland (+ 47° 22' 2" N, + 8° 32' 40" E) and its surrounding rural areas. A wide range of land covers and land uses are present in the area: buildings, urban parks, an airport, lakes, rivers, forests, and crop fields at various phenological stages. The lake is framed by two hill slopes, with altitudes varying between 392 and 871 m a.s.l. The average altitude of the area was 485 m a.s.l.

The remote sensing data were acquired on June 26th, 2011, under cloud-free conditions, with the APEX instrument. APEX is a dispersive push broom imaging spectrometer which records 1000 pixels per scan line with a field of view (FOV) of 28°. It covers the spectral region between 380 nm and 2500 nm in 313 spectral bands

with a sampling interval varying between 0.4 and 10 nm (Hüni 2009; Jehle et al. 2010). The flight pattern consisted of seven flight lines with headings alternating between north (34.2°) and south (214.2°). The data were acquired at an altitude of about 6400 m a.s.l., resulting in a raw across-track ground sampling distance of approximately 3.0 m. The solar zenith angle varied between 34.7 and 43.2°, with an average of 38.6°, and the solar azimuth angle varied between 107.6 and 122.7°, with an average of 115.3°. The seven resulting APEX images were radiometrically calibrated to obtain at-sensor radiance images.

5.2.2 Simulated Sentinel-2 scene

Each APEX radiance image was pre-processed individually before mosaicking. The pre-processing included: nadir-normalization, spectral resampling, geometric correction, and spatial resampling.

First the images were normalized to nadir viewing by applying a Li-Ross BRDF (bidirectional reflectance distribution function) correction based on a land cover classification with five classes (two vegetation structural classes, soil, water and urban surfaces) (Weyermann et al. Accepted). The land cover maps were obtained by classifying the atmospherically-corrected images using spectral angle mapping, with reference signatures extracted from the images. The Li-Ross BRDF correction was then applied individually to each land cover. For each land cover, the kernels for geometric (Li-dense or Li-sparse) and volumetric (Ross-thick or Ross-thin) scattering were selected based on a RMSE analysis.

The nadir-normalized data were spectrally convolved to the Sentinel-2 sensor response functions. They were then geo-corrected and ortho-rectified to the Swiss National Grid (CH1903) with a ground sampling distance of 10 m, using bi-linear interpolation. The seven APEX radiance images were then mosaicked. The Sentinel-2 bands of the mosaic having resolutions of 20 and 60 m were then further spatially resampled.

Finally, the Sentinel-2 TOA radiances were calculated from the radiances in the mosaic that correspond to measurements at the height of the APEX flight (L_{APEX}) using the following equation (Green et al. 2003):

$$L = L_{path} + T_{atm} L_{APEX} \qquad (5.1),$$

where L_{path} is the atmospheric path radiance, and T_{atm} is the total atmospheric upward transmittance between the APEX flight height and the top of the atmosphere. These two variables were calculated using MODTRAN4 as described in (Guanter et al. 2009) for surface albedo values of 0 and 0.5. The MODTRAN4 runs were performed with a "ground altitude" of 6 km, the maximum allowed in MODTRAN4, instead of the APEX flight height of 6.4 km, and the maximum water column of 0.36308 g/cm^2, which was the allowed value closest to the water column estimated in the

atmospheric correction. The final image spatially and spectrally corresponds to a Sentinel-2 level 1c product.

5.2.3 Objects segmentation and classification

The simulated Sentinel-2 image was segmented in spectrally homogeneous objects using the multi-resolution segmentation algorithm as implemented in eCognition (version 8.0.1). This is a region-growing algorithm which minimizes the internal spectral and spatial heterogeneity of the objects. The segmentation is performed using three user-defined parameters: the scale parameter (*scale*) which controls the internal homogeneity and therefore the size of the objects, the shape parameter (*shape*) which weights the importance of the spatial information versus the spectral information, and the compactness parameter (*compact*) which weights the importance of the compactness versus the smoothness of the object within the shape weight (Benz et al. 2004).

The segmentation was performed on the Sentinel-2 bands having the highest spatial resolution of 10 m: band 2 (blue, B), band 3 (green, G), band 4 (red, R), and band 8 (near infra-red, NIR). The three segmentation parameters were chosen by trial and error, paying attention to keeping agricultural fields and forest stands as individual objects. They were set as follows: *scale* = 8, *shape* = 0.2 and *compact* = 0.5.

The objects were then classified into five land cover classes: no vegetation, and four vegetation classes (see Table 5.1). The normalized difference vegetation index (NDVI) was used to discriminate vegetated and non-vegetated objects. The vegetated objects were then classified in four classes having similar radiative transfer regime, as related to their structural complexity. The structural complexity is inversely related to the brightness in the visible domain and the four vegetation classes (VB1 to VB4) were therefore defined using the sum of the R, G, and B bands (R+G+B).

Table 5.1 Decision rules used for classifying the objects.

NDVI (-)	R+G+B (mW/(m² sr nm))	Class
NDVI > 0.3	R+G+B ≤ 86	VB1
	86 < R+G+B ≤ 100	VB2
	100 < R+G+B < 130	VB3
	R+G+B ≥ 130	VB4
NDVI ≤ 0.3		No vegetation

5.2.4 Bayesian object-based estimation of LAI and Cab

5.2.4.1 Estimation algorithm

LAI and Cab were estimated directly from the TOA radiance data using a Bayesian object-based algorithm (Laurent et al. In review-b). This algorithm consists of two steps which are briefly described below.

In the first step, six to seven variables, including LAI and Cab, are estimated for each object, using a Bayesian optimization algorithm. This algorithm seeks to find the maximum likelihood estimate by minimizing the cost function χ^2 (Tarantola 2005):

$$\chi^2 = \tfrac{1}{2}(\mathbf{L_o} - \mathbf{L})^{\mathrm{T}}\mathbf{C_o}^{-1}(\mathbf{L_o} - \mathbf{L}) + \tfrac{1}{2}(\mathbf{v_a} - \mathbf{v})^{\mathrm{T}}\mathbf{C_a}^{-1}(\mathbf{v_a} - \mathbf{v}) \qquad (5.2),$$

where $\mathbf{L_o}$ is the vector of average observed radiance for the object, $\mathbf{L}$ is the vector of simulated radiance, $\mathbf{C_o}$ is the covariance matrix containing the observation and model uncertainties in each spectral band, $\mathbf{v_a}$ is the vector of a priori variable values, $\mathbf{v}$ is the vector of variable values, and $\mathbf{C_a}$ is the covariance matrix of the a priori variables. The first term of Equation 5.2 is the radiometric cost, and the second term is the a priori cost.

The update vector $\mathbf{\Delta v}$ to apply to $\mathbf{v}$ in each iteration of the optimization is given by:

$$\mathbf{\Delta v} = [\mathbf{J}^{\mathrm{T}}\mathbf{C_o}^{-1}\mathbf{J} + \mathbf{C_a}^{-1} + \mu\mathbf{J}^{\mathrm{T}}\mathbf{J}]^{-1}[\mathbf{J}^{\mathrm{T}}\mathbf{C_0}^{-1}(\mathbf{L_o} - \mathbf{L}) + \mathbf{C_a}^{-1}(\mathbf{v_a} - \mathbf{v})] \qquad (5.3),$$

where $\mathbf{J}$ is the Jacobian matrix, and μ is the damping factor. $\mathbf{J}$ contains the partial derivatives of the model, obtained by varying each variable, by 1 % of its allowed range (Laurent et al. 2011b). The damping factor was introduced to avoid the occasional cost increase which can happen with the Gauss-Newton-like formula by rotating $\mathbf{\Delta v}$ towards the direction of steepest descent, similarly to the Levenberg-Marquardt algorithm. It is set to zero at the beginning of each iteration. In case of increasing cost, μ is set to one, and subsequently multiplied by 10, until a decrease in cost is obtained, or until μ reaches 10^{10}.

Each variable v is constrained between its a priori minimum (v_{min}) and maximum (v_{max}) values by a transformation into a variable u:

$$u = \arcsin\left(2\frac{v - v_{\mathrm{min}}}{v_{\mathrm{max}} - v_{\mathrm{min}}} - 1\right) \qquad (5.4).$$

The optimization is performed in the transformed variable space, so the $\mathbf{C_a}$ and $\mathbf{J}$ matrices are also expressed in the transformed variables. At the end of the

optimization, the optimized variable values are transformed back into the original variable space.

In the second step, LAI and Cab are estimated for each pixel in the object using a LUT, constrained by the outputs of the Bayesian optimisation at the object level. All variables, except LAI and Cab are fixed to their optimized values. The LAI range is chosen by optimizing LAI for the two pixels having the minimum and maximum radiance in the NIR bands (bands 8 and 9), using the Bayesian optimization algorithm of the first step. LAI is then uniformly sampled within the obtained range, increasing the number of samples with increasing LAI range. For each sampled LAI, the range for Cab was chosen according to the LAI value. Cab was then sampled in its range using steps of 10 μg/cm^2. Finally, the coupled model was run for all LAI-Cab combinations to populate the LUT.

5.2.4.2 Coupled canopy-atmosphere model

The vegetation canopy was modelled with the soil-leaf-canopy model SLC (Verhoef and Bach 2007). It was chosen because it is a hybrid RT model with good accuracy and fast running time for use in the optimization algorithm. SLC couples the soil model 4SOIL with the leaf model PROSPECT and the canopy model 4SAIL2.

PROSPECT simulates the leaf optical properties based on the leaf concentrations in chlorophyll (Cab), dry matter (Cdm), water (Cw), and on the leaf mesophyll parameter (N) (Jacquemoud and Baret 1990). The refractive index of leaf material and the specific absorption coefficients of chlorophyll, dry matter and water at 1 nm resolution were taken from Feret et al. (2008). In addition, the model was extended to include the concentration in brown pigments (Cs) (Verhoef and Bach 2003b).

4SAIL2 is based on the SAIL model (Verhoef 1984), which assumes arbitrarily inclined leaves, as described by the leaf inclination distribution (LIDF). 4SAIL2 allows for two horizontal canopy layers, which can have different proportions of brown and green leaves in their total leaf area (fB). The vertical distribution of the brown and green leaves is controlled by the dissociation parameter (D): if D = 0, the brown and green leaves are homogeneously distributed, if D = 1, the brown leaves are all in the bottom layer of the canopy, and the green leaves in the top layer. The crown-level clumping is modelled assuming spherical crown shapes and is controlled by the canopy cover (Cv), and the crown shape parameter (Zeta), defined as the ratio of the crown diameter to the height of the crown centre above ground (Verhoef and Bach 2007). The crown clumping is an interesting feature to model open canopies. Finally, the hot spot parameter (Hot) controls the width of the hot spot reflectance peak in the angular domain.

The atmosphere was simulated using the MODTRAN4 RT model (Berk et al. 2003). It is a state-of-the-art model with rather long running time. The canopy-atmosphere coupling approach of Laurent et al. (2011b; 2011c), however, requires

only three runs for surface albedo values of 0, 0.5, and 1 (Verhoef and Bach 2003b) which can be performed in advance. Because the atmospheric parameters were considered constant, this allowed using the coupled SLC-MODTRAN4 model in the optimization. The canopy-atmosphere coupling relies on the 4-stream theory, and allows making full use of the four directional components of the canopy reflectance provided by SLC.

5.2.4.3 Parameterization of the model and estimation algorithm

A soil spectrum measured with an ASD spectroradiometer in a neighbouring test site with similar soil type (Oensingen, Switzerland, +47° 16' 44" N, +7° 43' 53" E) was used, instead of using the 4SOIL model, and it was considered constant in the study area. The SLC parameters were fixed independently for each vegetation class, using expert knowledge (Table 5.2). The three MODTRAN4 runs were conducted for a visibility of 60 km, a water vapour column of 1.3 g/cm^2, between the average height of the study area and TOA level, using the 1 cm^{-1} band model. The visibility of 60 km and water vapour column of 1.3 g/cm^2 were obtained by averaging the values calculated for the atmospheric correction of the individual flight lines. The SLC and MODTRAN4 simulations were performed at 1 nm spectral resolution, and the final radiance output signature was convolved to the Sentinel-2 bands using the spectral response functions as defined in Drusch et al. (2012).

LAI and Cab were estimated using the Bayesian object-based estimation algorithm (section 5.2.4.1) for all vegetated objects whose size was larger than 1000 m^2 (ten pixels of 10 m). The free variables in the object-based Bayesian optimization were: LAI, Cv, Cab, Cw, Cdm, N (of the green leaves), and additionally fB for the VB1 and VB2 objects, which correspond to forest or mature crop fields. The a priori data for these six or seven variables for each vegetation class were chosen using expert knowledge (Table 5.3). The initial values of all the variables were set to the middle of their a priori variation range, except LAI whose initial value was 2. The optimization was terminated when the update changes (Equation 5.3) for all the

Table 5.2 Fixed parameter values for the SLC model for the four vegetation classes.

		Vegetation class			
	Parameter (unit)	VB1	VB2	VB3	VB4
Canopy	LIDF (-)	Spherical	Spherical	Spherical	Spherical
	D (-)	0	0	0	0
	Zeta (-)	0.5	0.5	1	1
	Hot (-)	0.01	0.05	0.05	0.05
Brown leaves	Cab (µg/cm^2)	10	5	5	5
	Cw (cm)	0	0.001	0.001	0.001
	Cdm (g/cm^2)	0.5	0.005	0.005	0.005
	Cs (-)	15	1	1	1
	N (-)	10	2	2	2
Green leaves	Cs (-)	0.05	0.05	0.05	0.05

variables were smaller than 1 % of their a priori variable range, or when ten iterations had been performed. The measurement uncertainties for the a priori data and the Sentinel-2 observations were assumed to be uncorrelated, so the two covariance matrices were both diagonal. The diagonal of the $\mathbf{C_a}$ matrix was filled with the a priori variance of the variables, calculated from Table 5.3. Following Laurent et al. (In review-b), the noise of the Sentinel-2 data was assumed uncorrelated and equal to 1 $[mW/(m^2 \, sr \, nm)]^2$ for all spectral bands, and the model errors were ignored, so the $\mathbf{C_o}$ matrix was equal to the identity matrix.

5.2.5 Evaluation of the effects of the parametric nadir-normalization

There is a conceptual issue when using a parametric RT model to normalize the radiance data to nadir viewing, and then estimating the variables from the nadir-normalized data using a physically-based RT model. Indeed, the radiance values are corrected to nadir-viewing according to the angular behaviour of the parametric model, which may differ from that of the physically-based model, and therefore reduce the performance of the inversion of the physically-based model. First, the discrepancy between the Li-Ross and SLC-MODTRAN4 angular signatures was investigated, and second the impact of the Li-Ross nadir-normalization on the LAI and Cab estimates was evaluated. These two analyses were performed at APEX flight height, before applying Equation 5.1. The results are easily transferred to TOA level, because the TOA radiance is a linear transformation of the radiance at APEX flight height (Equation 5.1). The westernmost image (flight line 9, FL9) was selected because it had the most vegetation. The solar zenith angle for FL9 was 34.7° and the solar azimuth angle was 122.7°. MODTRAN4 was run in advance, between the

Table 5.3 A priori mean, standard deviation, minimum and maximum values of the canopy and green leaves variables for the four vegetation classes.

Variable	Vegetation class			
(unit)	VB1	VB2	VB3	VB4
LAI	**4** (2)	**3** (2)	**1** (2)	**0.5** (2)
(-)	[2, 8]	[1, 8]	[0, 4]	[0, 2]
Cv	**1** (0.2)	**1** (0.2)	**0.8** (0.2)	**0.3** (0.2)
(-)	[0.7, 1]	[0.5, 1]	[0.3, 1]	[0, 0.5]
Cab	**40** (20)	**40** (20)	**40** (20)	**40** (20)
(μg/cm^2)	[10, 100]	[10, 100]	[10, 100]	[10, 100]
Cw	**0.02** (0.01)	**0.02** (0.01)	**0.02** (0.01)	**0.02** (0.01)
(cm)	[0.005, 0.05]	[0.005, 0.05]	[0.005, 0.05]	[0.005, 0.05]
Cdm	**0.005** (0.001)	**0.005** (0.001)	**0.005** (0.001)	**0.005** (0.001)
(g/cm^2)	[0.001, 0.01]	[0.001, 0.01]	[0.001, 0.01]	[0.001, 0.01]
N	**1.8** (0.5)	**1.8** (0.5)	**1.8** (0.5)	**1.8** (0.5)
(-)	[1, 2.5]	[1, 2.5]	[1, 2.5]	[1, 2.5]
fB	**0.2** (0.1)	**0.2** (0.1)	**0** (n/a)	**0** (n/a)
(-)	[0, 0.5]	[0, 0.5]	[n/a]	[n/a]

average altitude of the study area and the APEX flight height, for viewing zenith angles (VZA) between -14° and +14° (APEX FOV) in steps of 2°. The outputs were stored in an atmospheric database. Positive VZA values indicate viewing directions opposite to the sun (forward direction).

5.2.5.1 Angular signatures

The angular signatures simulated by the SLC-MODTRAN4 and Li-Ross models were compared for the four vegetation classes for the four 10 m bands of Sentinel-2. For each class, the SLC-MODTRAN4 was run using the default parameter values (Table 5.2) and the a priori values (Table 5.3), for each of the 1000 across-track APEX pixels, interpolating the MODTRAN4 outputs from the atmospheric database. The obtained angular signatures were then fed to the Li-Ross nadir-normalization algorithm (see section 5.2.2). The kernels and associated weights chosen by the algorithm were then used to simulate the Li-Ross angular signatures.

5.2.5.2 LAI and Cab estimates

The impact of the parametric Li-Ross nadir-normalization on the LAI and Cab estimates was evaluated by comparing the LAI and Cab maps obtained by the Bayesian object-based inversion of the SLC-MODTRAN4 model (section 5.2.4) from:

1. the nadir-normalized image, using nadir viewing geometry in the simulations (NN approach)
2. the original image, using the original viewing geometry in the simulations (OG approach)

The objects and associated classes needed to apply the Bayesian object-based estimation algorithm to FL9 were obtained by taking a spatial subset of the objects and associated vegetation classes obtained for the Sentinel-2 scene (section 5.2.3). For the OG approach, an image containing the original viewing geometry (zenith and azimuth angles) for each pixel was created in raw APEX geometry, and then geometrically corrected in the same way as FL9. The average viewing geometry for each object were then calculated and used in the SLC-MODTRAN4 simulations.

5.3 Results

5.3.1 Simulated Sentinel-2 scene

The simulated Sentinel-2 scene has an extent of 16 x 22 km^2. It has full spectral and spatial characteristics of the future Sentinel-2 satellite. True and false colour composites of the 10 m resolution bands are presented in Figure 5.1. Despite the nadir-normalization applied to each of the seven APEX images, a brighter vertical

Figure 5.1 True (R: band 4, G: band 3, B: band 2) and false (R: band 8, G: band 4, B: band 3) colour composites of the simulated Sentinel-2 scene with 10 m resolution. The blue rectangle indicates the area used for Figures 5.2 and 5.4, and the yellow rectangle indicates the area of FL9 used for Figure 5.6.
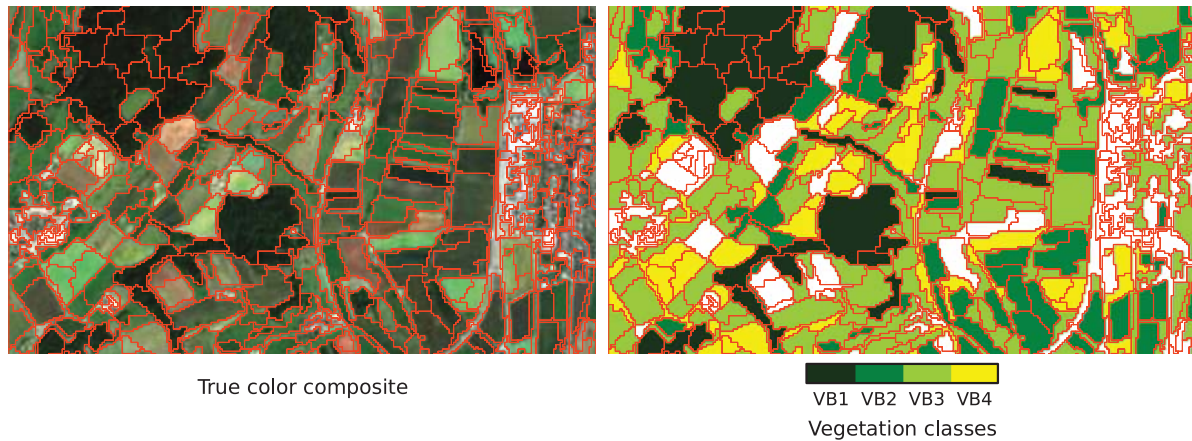
stripe appears on the west side of the scene, and there is a general trend of increasing brightness from east to west. This is discussed in section 5.4.3.

5.3.2 Objects identification and classification

The objects and classes for a small area in the south west of the scene are presented in Figure 5.2. The size of the objects is related to the heterogeneity of the area. Small objects are predominantly found in urban areas where many types of surfaces can be found within a few pixels, while larger objects are found in forested areas, which possess lower spectral and spatial internal variability. Agricultural fields are spectrally and spatially more homogenous than forest patches, but objects in agricultural areas are smaller because of the sharp transition from one crop to the next at the edges of the small Swiss fields.

The classification criteria (Table 5.1) efficiently discriminated vegetated and non vegetated objects such as urban areas and bare fields. As expected, most forest objects were classified as VB1, and brown field objects having low vegetation cover were classified as VB4. Vegetated objects located within urban areas and larger than 1000 m^2 were included in the LAI and Cab estimation.

True color composite

VB1 VB2 VB3 VB4
Vegetation classes

Figure 5.2 Objects (red polygons) with the true colour composite (left) and with the vegetation classes (right) for a zoomed area (blue rectangle on Figure 5.1). The white areas were classified as no vegetation.
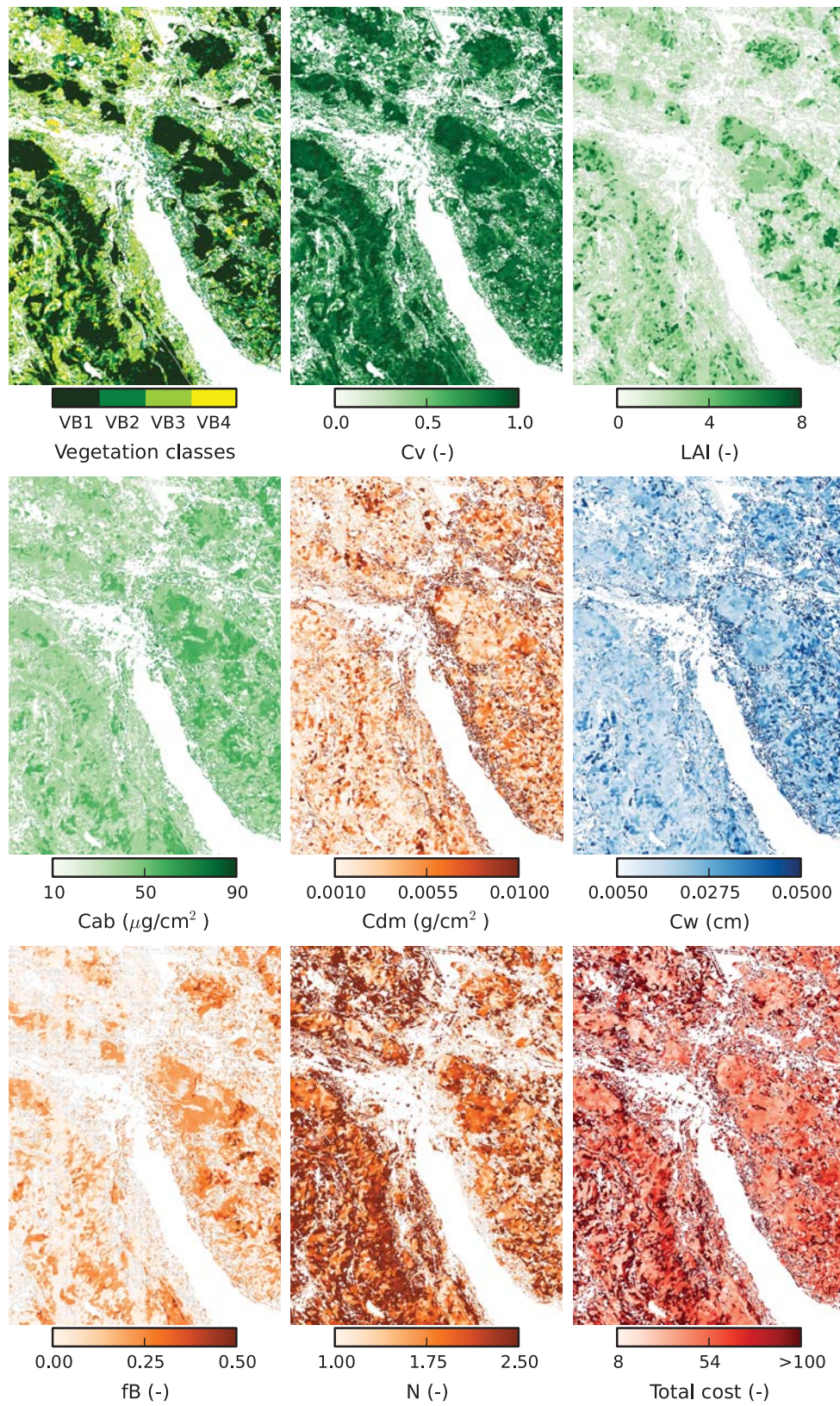
5.3.3 Bayesian object-based estimation

The object-level maps of the estimates of the seven free variables in the Bayesian optimization algorithm are presented in Figure 5.3, as well as the full vegetation classification map, and the cost map. Most objects had high Cv values, except for VB4 objects and a few VB3 objects, accordingly to the Cv ranges used in the a priori data (Table 5.3). The highest LAI values were found in VB1 objects. These objects also had higher Cab and higher fB than the objects classified as VB2 to VB4. The differences between object classes were less marked for Cdm, Cw, and N. The cost map was represented using the value range from 8 to 100. Because the a priori costs were much smaller than the radiometric costs, this corresponds to an average radiance mismatch of 1.1 to 3.9 mW/(m^2 sr nm) in each band, assuming an equal radiance mismatch in all 13 bands. This can be related to the magnitude of the radiance values, which is usually in the range of 20 to 50 mW/(m^2 sr nm) in the visible, and in the range of 40 to 200 mW/(m^2 sr nm) in the NIR for vegetation. Most objects had cost values smaller than 100, but there was a trend to higher cost values towards the west of the scene. The brighter vertical stripe on the west side of the scene translated into visible stripes on the object-level maps of LAI, Cdm, fB, and cost.

The pixel-level maps of LAI and Cab (not shown) show the same spatial patterns of higher values in VB1 objects, the values were higher than those obtained at the object level. Most pixels still had cost values smaller than 100, but much higher cost values than at object level were reached locally, especially in very high VB1 objects.

The LAI and Cab estimates at object and pixel levels are presented in more detail for a small area in the south west of the scene (Figure 5.4). The object and pixel maps of each variable have similar spatial patterns, with more details appearing in the pixel maps. The added variability at the pixel level concerns mostly the VB1 and

Figure 5.3 Object-based maps of vegetation classes, LAI, Cv, fB, Cab, Cdm, Cw, N, and cost values obtained from the simulated Sentinel-2 image.

Figure 5.4 Object- and pixel-level maps of LAI and Cab for a zoomed area (blue rectangle on Figure 5.1), with objects boundaries (red polygons).

VB2 objects (see vegetation classes on Figure 5.2). From object to pixel level, the maximum values increased from 7.6 to 8 for LAI, and from 57 to 100 for Cab.

5.3.4 Effects of the parametric nadir normalization

5.3.4.1 Angular signatures

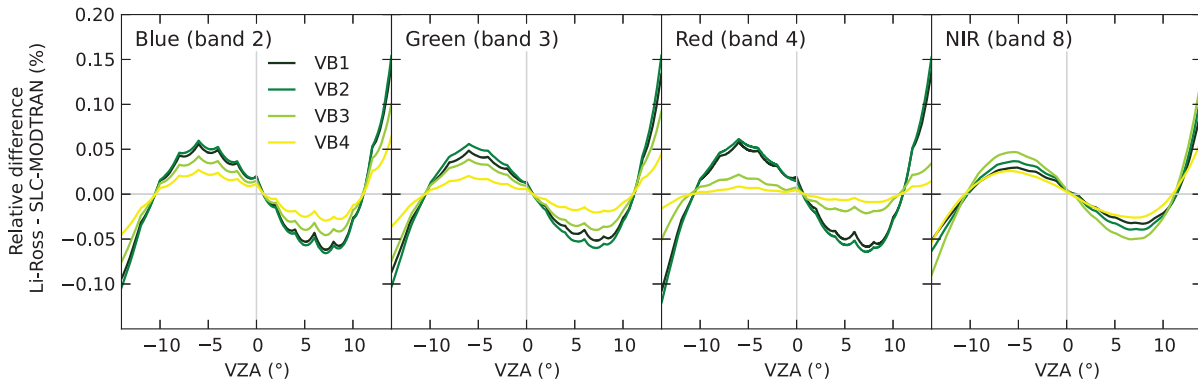For the Li-Ross simulations, the nadir normalization algorithm selected the same kernels for all four vegetation classes and for all four 10 m resolution bands: Ross-thin for the volumetric scattering and Li-dense for the geometric scattering (see section 5.2.2). For all bands and vegetation classes, the spectral profiles of the kernel weights were similar (not shown). The weights of the isotropic scattering kernel were the highest, while the weights of the geometric kernel were close to zero, and the weights for the volumetric kernel were negative.

The relative differences between the Li-Ross and SLC-MODTRAN4 angular signatures for the four 10 m resolution bands are presented in Figure 5.5. The difference curves for all four vegetation classes present the same sinusoidal shape for the four bands, revealing systematic differences between the two models. The Li-Ross simulations were lower than the SLC-MODTRAN4 simulations for $VZA < -11°$ and for $VZA \in [1°, 12°]$, and higher than the SLC-MODTRAN4

Figure 5.5 Relative difference between the Li-Ross and SLC-MODTRAN4 angular signatures for the four 10 m bands and for the four vegetation classes.

simulations for VZA ∈ [- 11°, 1°] and VZA >12°. The Li-Ross simulations therefore differed from the SLC-MODTRAN4 simulations for nadir viewing. For the VB1 and VB2 vegetation classes, the relative differences were lower in the NIR than in the visible bands, whereas for the VB3 and VB4 vegetation classes, the relative differences were lowest in the red bands. All absolute difference values were smaller than 0.15 mW/(m^2 sr nm) in absolute value.

5.3.4.2 LAI and Cab estimates

The LAI, Cab and cost maps obtained from the NN and OG approaches (not shown) had R^2 values of 0.96 for LAI, 0.98 for Cab and 0.93 for the cost, showing a good consistency between the two approaches. The difference maps (Figure 5.6) show the spatial patterns of the differences between the two approaches. The objects are much more visible on the LAI and cost difference maps than on the Cab difference map, which appears more scattered. For all three maps, as expected, the differences are smallest in the across-track centre of the image, and increase towards the edges. The highest LAI differences are found in very high scattering objects, where LAI values are higher. The spatial pattern of very high cost differences roughly follows the pattern of the VB1 and VB2 objects.

5.4 Discussion

5.4.1 Objects and spatial constraints

The multi-resolution segmentation algorithm was chosen to demonstrate the feasibility of object identification within a Sentinel-2 simulated scene. The area of the city of Zurich and its surroundings presented discrete landscape units with clear boundaries, favouring object-based approaches. The multi-resolution segmentation algorithm implemented in eCognition is the most frequently used algorithm in object-

Figure 5.6 Maps of the difference between the LAI, Cab, and cost obtained from the NN and OG approaches for the south part of FL9 (area in the yellow rectangle of Figure 5.1). Blue colours indicate areas where the NN approach gave lower values than the OG approach.

based studies (Blaschke 2010) and has been successfully used for identification and classification of operational vegetation units (Johansen et al. 2007; Lamonaca et al. 2008; Massada et al. 2012; Wang et al. 2004). This region-growing segmentation algorithm enables automatic identification and delineation of objects of varying sizes, because it evaluates the spectral and spatial heterogeneity at each 'growing' step of the object segmentation process, until the user-defined maximum internal heterogeneity (the *scale* parameter). If this threshold is reached early, at a 'high resolution', the object will be small, e.g. in urban areas (Figure 5.2). On the contrary, if the threshold is reached at a 'low resolution' the object will be bigger, e.g. in forested areas (Figure 5.2). For this reason, careful tuning of the *scale* parameter is important in order to obtain objects homogenous enough to apply the spatial constraints and to obtain accurate estimates of LAI and Cab at the pixel level.

Indeed, in this study, the smoothness principle (Wang et al. 2008) was applied strictly, fixing the object-level variables to their optimized values for all pixels in the object. If, for example, an object consists of pixels belonging to two vegetation classes, the object-level variables will be optimized to match the area-weighted average signature of the two classes. The obtained values will not be suitable for either of the two classes, and will lead to inaccurate LAI and Cab estimates at the pixel level. Multi-level image segmentation yields a hierarchical object network (Benz et al. 2004), which allows obtaining meaningful objects (Blaschke 2010) such as agricultural fields and urban tree lines. For implementation of the spatial constraints, however, the homogeneity of the objects is primordial, and single-level segmentation was more adapted, as it was preferred to have several objects in a single field rather than two fields in the same object.

The 10 m resolution of four of the Sentinel-2 bands was appropriate for deriving objects using multi-resolution segmentation. Contrary to the study of (Laurent et al. In review-b), the mixed pixels at the edges of the objects were not discarded, because of the spatial resolution of Sentinel-2 was not as fine.

5.4.2 Vegetation classes and a priori data

A general, straightforward to apply, vegetation classification was adopted because, contrary to the study of Laurent et al. (In review-b), no extensive land cover data was available. The four levels of brightness in the visible domain were related to vegetation vertical structure, and general expert knowledge could then be used to define the a priori data (Table 5.3). Because of the broadness of the classes, the confidence in the a priori values was low. This was reflected by large values for the a priori standard deviations. This classification is therefore well applicable without the need of labour- (or cost-) intensive generation of a priori data.

5.4.3 Li-Ross modelling and nadir normalization

In the parametric Li-Ross model each wavelength is treated independently, which may result in different kernels and kernel weights for each band, whereas the physically-based SLC-MODTRAN4 model uses the same values of the canopy parameters for all bands. In addition, the volume scattering component of the Li-Ross model is based on the assumption of a spherical LIDF and equal leaf reflectance and transmittance. The latter assumption is better satisfied in the NIR than in the visible domain, so one would expect smaller angular deviations between Li-Ross and SLC-MODTRAN4 in the NIR than in the visible bands. This expectation was only fulfilled for the VB1 and VB2 vegetation classes. The deviations between the SLC-MODTRAN4 and Li-Ross model were expected to be small because the spherical LIDF assumption used in the Li-Ross volumetric scattering component matches the spherical LIDF used for defining the vegetation classes (Table 5.2). Although the

deviations between the two models were small in radiance units, the they were systematic, which might be due to the much smaller number of degrees of freedom of the Li-Ross model.

These systematic deviations may have impacted the quality of the nadir-normalization of the APEX images. The main limitation, however, was that the changes in solar geometry over the two-hour acquisition period were not accounted for in the Li-Ross BRDF correction algorithm. This explains the trend of increasing brightness from the east to the west (Figure 5.1) of the Sentinel-2 mosaic. Another important issue is that the nadir-normalization algorithm relies on a 5-classes land cover classification, only two of which concerning vegetation. This implies that the quality of the results depend on the spatial distribution of pixels of each class over the sensor's FOV. If observations of a scattering type are rather sparse and not well distributed over the FOV, the Li-Ross kernel weights can be less stable and correction may become less reliable. Another interfering effect is caused by topography, which is not accounted for in the algorithm. The brighter vertical stripe in the western part may be due to a combination of the described uncertainties. The Li-Ross BRDF correction cannot distinguish between the anisotropy of the surface and that of the atmosphere, and therefore corrected them both simultaneously. This is consistent with the nadir viewing convention used for creating the simulated Sentinel-2 scene. Enhancements on the BRDF correction approach are subject to current work. A more robust correction could be achieved by inverting the model on a set of observations generated from several flight lines if not from all. Applying different error metrics than least-RMSE for the inversion process, e.g. the least median of error squares (Susaki et al. 2004) or least variance of white-sky albedo (Gao et al. 2001) could further enhance the quality of the mosaic.

5.4.4 Simulated Sentinel-2 TOA scene

5.4.4.1 Suitability of APEX data

APEX data has unprecedented spectral, spatial and radiometric resolution, allowing to simulate each Sentinel-2 band with at least three APEX bands (D'Odorico et al. 2013) and each Sentinel-2 pixel with at least nine APEX pixels. The APEX signal-to-noise ratio (SNR) is also well above the expected SNR of Sentinel-2. The only pitfall simulating the Sentinel-2 scene is the 2-hour duration of the APEX acquisition, which resulted in variation of the solar geometry (zenith: 34.7 to 43.2°, azimuth: 107.6 to 122.7°). The effects on this changing solar geometry on the nadir normalization were discussed in section 5.4.3.

5.4.4.2 Structural limitations: aspects which were not included in the simulated scene

The simulated Sentinel-2 scene has an extent of 16 x 22 km^2, which is smaller than the 290 km swath and than the 100 x 100 km^2 tiles planned for the Sentinel-2

mission. Despite its restricted extent, however, the Zurich simulated scene covers a wide range of land cover types including agricultural fields, pastures, forests, water bodies, urban areas and even an airport. This variety makes this scene an interesting dataset for preparing and/or testing operational algorithms for Sentinel-2.

The simulated scene consists of a mosaic of nadir-normalized APEX images and therefore contains nadir-normalized radiance data. If the scene would be at the centre of the Sentinel-2 swath, the VZA would vary between - 0.6° and + 0.6°, which is small enough to assume that all pixels are viewed from nadir. The geometry of the Sentinel-2 MSI is relatively complex because it involves 12 detector modules, each recording the radiance in 13 spectral bands. Each detector module and spectral band has a slightly different viewing angle, which causes a significant parallax effect. This parallax effect, however, was ignored because it will only be visible for targets which are not at the ground altitude such as clouds and will therefore not be ortho-rectified properly.

5.4.4.3 Other limitations / sources of error

Each step of the processing described in section 5.2.2 introduced inaccuracies, all of which propagated in the final simulated Sentinel-2 TOA scene. First are the inaccuracies associated with the nadir-normalization of the APEX images discussed in section 5.4.3. The fact that the nadir-normalization ignored the atmospheric and topographic effects means that these effects were kept in the final simulated Sentinel-2 scene. This is good because those effects would also be included in the real Sentinel-2 imagery. The very high spectral resolution of the APEX data allowed limiting the inaccuracies related to the spectral convolution, and its high spatial resolution allowed for precise geometric correction and spatial resampling. The mosaicking process averaged the radiance values from up to three nadir-normalized APEX images, but which still included directional effects related to the solar geometry. The effects of the part of the atmosphere above APEX flight height was then added using Equation 5.1, which ignored the adjacency effects above APEX flight height. In addition, the values for the atmospheric parameters at APEX flight height were not known, so the surface values were used when possible, or replaced by the maximum value allowed in MODTRAN4. Considering the limitations of the processing to obtain the simulated Sentinel-2 scene, and despite the much higher SNR of APEX compared to that of Sentinel-2, no noise was added to the simulated Sentinel-2 scene.

5.4.5 LAI and Cab estimates

Unfortunately, no field data was available for evaluating the accuracy of the LAI and Cab estimates. It would be possible to perform an indirect validation using independent remote sensing products such as MODIS or MERIS LAI and Cab products. These products, however, have very different pixel size and are also subject

to uncertainties. In the context of this proof of concept study for applying the Bayesian object-based approach in areas without field data, it was not considered necessary to implement such an indirect validation.

The accuracy of the LAI and Cab estimates is limited by the accuracy of the simulated Sentinel-2 data, as discussed above, and by the performance of the estimation method, which includes the SLC-MODTRAN4 model and the Bayesian object-based approach. The Bayesian object-based approach was shown to be more accurate than a LUT with Bayesian cost function approach by a previous study where the results were extensively discussed (Laurent et al. In review-b). An indication on the performance of the estimation method is given by the cost values: higher cost values correspond to less reliable estimates. At the object level, most cost values were under 100 (Figure 5.3). At the pixel-level, however, the cost values increased, especially in VB1 objects. This may be related to the 10 m pixel size of the simulated Sentinel-2 data, which is too small for SLC to model forest environments: a 10 m pixel does not cover enough trees, violating the assumption made in SLC of internal homogeneity within the pixel. Therefore, the object-level LAI and Cab estimates for VB1 vegetation are more reliable than those at pixel level. Indeed, the lower LAI values obtained at object level are more realistic because the forests around Zurich are regularly thinned. Finally, the adjacency and topography effects were ignored in the coupled canopy-atmosphere RT modelling.

5.5 Conclusions

The Bayesian object-based approach allows implementing the maximum regulation of the RT model inversion to obtain the most accurate LAI and Cab estimates: coupling canopy and atmospheric RT models, using a priori data in the Bayesian optimization at the object level, and using spatial constraints for building the object-specific LUTs. Thanks to an automatic multi-resolution segmentation to identify vegetation objects and to a general vegetation classification based on four levels of brightness in the visible domain, the Bayesian object-based approach was successfully applied to estimate LAI and Cab from a simulated Sentinel-2 scene for which no field data was available. The results of this study therefore extend the applicability of the Bayesian object-based approach to cases where no field data is available, which is essential for large-scale environmental studies relying on RS-based LAI and Cab maps. Another important contribution of this study is the demonstration of systematic, although small, differences between the angular simulations of the physically-based SLC-MODTRAN4 model and the parametric Li-Ross model. Therefore, nadir-normalization of the data should be avoided, and it should be preferred to use original data, including the original geometry for each object, as was demonstrated here.

Acknowledgements

Chapter 6

Synthesis

6.1 Main results

Recognizing the need of the scientific community for better vegetation-related inputs for dynamic global vegetation models (DGVM), this thesis contributes to improving the accuracy of vegetation variables estimated from mono-temporal satellite optical remote sensing data. A globally applicable approach is required, and this thesis thus focussed on the most general and adaptable estimation approach: the radiometric data-driven inversion of physically-based radiative transfer (RT) models (section 1.2). The traditional top-of-canopy (TOC) approach, however, suffers from limitations related to the pre-processing of the measured data to TOC reflectances (section 1.4), and this thesis therefore proposes to work directly with the radiance data measured at top-of-atmosphere (TOA) level. Such a TOA-based approach allows avoiding the limitations of the TOC approach (section 1.5.2.1) and improving the regularization of the model inversion (section 1.5.2.2). The core hypothesis of this work was therefore that the TOA approach allows estimating the vegetation variables more accurately than the TOC approach. Based on this hypothesis, four research questions were raised (section 1.6). Each of them is answered below.

6.1.1 Research question A: Proof of concept of the TOA approach

A. How accurately can the TOA approach estimate vegetation variables directly from the radiance data measured by the spaceborne sensor?

This research question was addressed in chapter 2. The aim of this chapter was to introduce the TOA approach and to compare it with the traditional TOC approach in a case study using near-nadir CHRIS data of three Norway spruce stands in Czech Republic. In order to provide a convincing proof of concept for the TOA approach, question A was investigated in three steps.

First, the accuracy of the coupled canopy-atmosphere RT model was evaluated (Laurent et al. 2009). This is essential because the vegetation variables can be estimated in the TOA approach, by inverting the coupled model, only if the coupled model is accurate enough. The coupling set-up, based on the 4-stream theory, allowed making full use of the four directional reflectance components simulated by the SLC canopy model in the TOA radiance simulations. For the proof of concept, the coupled model should be at least as accurate as the canopy model. To compare the accuracy of the two models, which simulate different physical quantities, the relative errors were used to calculate the cost function. For the three stands, the costs were smaller at TOA level (between 0.05 and 0.14) than at TOC level (between 0.2 and 0.5), but this was due to the atmospheric path radiance which increased the radiance signal. Therefore, the accuracy of the coupled canopy-atmosphere model was similar to that of the canopy model for this case study.

Second, the dimensionality and the sensitivity to the inputs of the two models were investigated. This is important because the number of variables that can be estimated is limited by the dimensionality, and only the variables having a significant influence on the model simulations can be reliably estimated. An influence indicator was calculated for each model input based on the Jacobian matrix. The local sensitivity analysis showed that, for the three stands, and at both TOA and TOC levels, the canopy structure parameters were most influential, followed by the leaf parameters and the brown material parameters. At TOA level, the atmospheric parameters had the smallest influence, and therefore they did not limit the estimation of vegetation variables from the TOA approach. The singular value decomposition (SVD) of the Jacobian matrix evaluated the dimensionality to be 3 for the three stands for both TOA and TOC level. Therefore, four of the most influential model inputs were chosen as variables to estimate: canopy crown cover (Cv), fraction of brown material (fB), and leaf chlorophyll (Cab) and dry matter (Cdm) contents.

Third, the performance of the TOA approach for estimating these four variables was compared to that of the TOC approach. This constitutes the core of research question A. For each stand, two look-up tables (LUT) with Cv, fB, Cab and Cdm as free variables were built: one with TOA radiances for the TOA approach and the other with TOC reflectances for the TOC approach. The results showed that TOA and TOC approaches had similar estimation performance for the three stands (Laurent et al. 2010a).

Finally, this case study demonstrated that, having similar forward modelling accuracy, similar dimensionality and similar parameter influence profile as the TOC approach, the TOA approach was able to estimate the vegetation variables with similar accuracy as the TOC approach (Laurent et al. 2011b). Similar performance, however, is already sufficient to prefer the TOA approach over the TOC approach for applications where minimal pre-processing of the data is advantageous, such as the estimation of vegetation variables from multi-angular data where each dataset has to be pre-processed individually in the TOC approach.

6.1.2 Research question B: TOA approach and multi-angular data

B. What is the added value of multi-angular radiance data for estimating vegetation variables using the TOA approach?

This research question was addressed in chapter 3, using the same three Norway spruce stands as in research question A, but this time making use of the three available multi-angular CHRIS datasets in addition to using (near-)nadir data only. Because the TOA approach is new, no study estimating vegetation variables from measured TOA multi-angular data was found in the literature. It was therefore considered necessary to proceed with the same three steps as for research question A,

but comparing mono- and multi-angular approaches instead of TOC and TOA approaches.

First, the accuracy of the coupled model for simulating the off-nadir measurements should be comparable to its accuracy for the nadir simulations. The cost values were based on the root mean square error between simulated and measured spectral signatures. For the three stands, the off-nadir costs had the same magnitude as the nadir cost. The cost values, extended to include all four available observation directions, varied between 2.9 and 4.8 mW/(m^2 sr nm) for the three stands. The magnitude of the forest radiance usually varies between 20 and 50 mW/(m^2 sr nm) in the visible, and between 40 to 150 mW/(m^2 sr nm) in the near-infrared. Therefore, the coupled model was able to provide realistic simulations of the measured multi-angular data.

Second, the dimensionality and the sensitivity to the inputs of the coupled model for the mono- and multi-angular cases were compared. The dimensionality and influence of the model inputs were evaluated for all possible combinations of the four observation directions, by vertically stacking the Jacobian matrices obtained for each of the observation directions used in the combination. For all angular combinations, the canopy structure parameters were most influential, followed, in decreasing order, by the leaf parameters, the brown material parameters, and the atmospheric parameters. For the three stands, the SVD of the Jacobian matrix indicated that the dimensionality increased from 3 to 6 when increasing the number of observation directions from 1 to 4 (Laurent et al. 2010b; Laurent et al. 2011a). Therefore, the extreme dimensionality cases were tested by estimating either four or seven variables. fB, Cv, Cdm and Cab were chosen for the 4-variable case, and the leaf concentration in brown pigments (Cs), the dissociation factor (D), and the tree shape factor (Zeta) were added for the 7-variable case.

Third, the performances of all angular combinations for estimating these four and seven variables using the TOA approach were compared. For each stand, a 4-variable LUT and a 7-variable LUT were built. The variables were estimated from each LUT, using all possible angular combinations. For the three stands and for both cases, the angular combinations which provided the best estimates had two or three observation directions. These results did not match the expectation that the more observation directions used, the more accurate the estimates, thus pointing out the difficulty of fully exploiting multi-angular data.

Finally, this case study showed that the dimensionality increased when increasing the number of observation directions, therefore demonstrating a higher potential of multi-angular data over mono-angular data for estimating vegetation variables. This added value of multi-angular data was already known for TOC reflectance data, but this study showed that this is also true at TOA level (Laurent et al. 2011c).

6.1.3 Research question C: Bayesian object-based approach

C. What is the potential of a Bayesian object-based approach for estimating LAI and Cab using a coupled canopy-atmosphere model?

This research question was addressed in chapter 4. The aim of this chapter was to introduce a Bayesian object-based approach implementing three regularization methods: model coupling, a priori information and spatial constraints. The approach relies on spectrally homogeneous objects which are treated independently. For each object, the approach consists of two steps: 1) Bayesian optimization of up to six variables (LAI, Cab, Cdm, leaf water content Cw, and leaf mesophyll parameter N, and either Cv or fB), and 2) LUT estimation of LAI and Cab for each pixel, fixing all other variables to their optimized values. The potential of this new approach for estimating LAI and Cab was tested on APEX airborne imaging spectrometer data over an agricultural test site in Oensingen, Switzerland. Three steps were required to answer research question C.

First, the performance of the Bayesian optimization for estimating the variables at object level was evaluated. This is important because accurate object-level estimates are crucial for building appropriate object-specific LUTs and therefore to the accuracy of the pixel-level LAI and Cab estimates. Although the dimensionality of the APEX data was only 4, the Bayesian optimization was able to estimate up to six variables. This was possible thanks to the use of the Jacobian matrix in the updating rule, which allowed to distinguish the effects of the variables by their full spectral sensitivity profiles (Laurent et al. 2012). The a priori information allowed strictly constraining the variables in their ranges, and the a priori covariance matrix efficiently avoided obtaining a singular matrix in the update rule. The a priori values themselves, however, were virtually not used, because the observation covariance matrix gave very high weight to the observations. Consequently, the optimization was almost entirely driven by the radiometric match between the model simulation and the observations. Visually, the patterns of low and high Cv and LAI corresponded well with the brown and green fields on the true colour image. LAI was most accurately estimated (R^2 = 0.95), followed by Cw (R^2 = 0.68) and Cdm (R^2 = 0.36). Cab was poorly estimated (R^2 = 0.11).

Second, the performance of the Bayesian object-based approach for estimating LAI and Cab at the pixel level was evaluated by comparing the estimates with the ground data. The obtained spatial patterns were similar to those obtained at object-level, but intra-object variability appeared. LAI (R^2 = 0.45) was estimated more accurately than Cab (R^2 = 0.18).

Third, in order to evaluate the impact of the spatial constraints, the performance of the Bayesian object-based approach was compared with that of a traditional LUT approach, but using a Bayesian cost function (LUT-BCF). The LUT-BCF approach estimated Cab (R^2 = 0.26) more accurately than LAI (R^2 = 0.22). Both approaches

suffered from misclassification issues, leading to the use of erroneous a priori data. The Bayesian object-based approach, however, has a higher degree of freedom in the Bayesian optimization and the spatial constraints allow avoiding variable compensation effects in the LUT. These two features allow the Bayesian object-based approach to provide more accurate LAI estimates than the LUT-BCF approach.

Finally, this case study showed that the Bayesian object-based approach effectively combined the strengths of model coupling, a priori information and spatial constraints regularization methods, and that the spatial constraints allow increasing the accuracy of the LAI estimates (Laurent et al. In review-b). In order to better balance the observation and the a priori information, more research on the implementation of the observation covariance matrix is needed. Furthermore, for this proof of concept study, the objects were manually digitized, and extensive a priori information was available. Therefore, the applicability of the Bayesian object-based approach in cases where less extensive knowledge about objects and land cover is available still needs to be investigated.

6.1.4 Research question D: Extending the Bayesian object-based approach

D. What is the applicability of the Bayesian object-based approach for estimating LAI and Cab from a TOA image for which less a priori information is available?

This research question was addressed in chapter 5. The aim of this chapter was to propose a method for applying the Bayesian object-based approach to an image for which there is no available knowledge about the objects or the land cover. The proposed method uses automatic image segmentation and four general vegetation classes having similar radiative transfer regime, based on increasing levels of brightness in the visible domain (classes VB1 to VB4). It was tested on a simulated Sentinel-2 scene of the Zurich area, which was mosaicked from seven APEX images. The scene covered an extent of 16 x 22 km^2 and was spatially and spectrally equivalent to the Sentinel-2 level 1c TOA radiance product. The main source of error in the simulated data was the nadir-normalization of the APEX images, which relied on a parametric model. The differences between this parametric model and the physically-based model used for estimating LAI and Cab, as well as the impact of the nadir-normalization on the estimates, were discussed in Chapter 5. Three steps were required to answer research question C.

First, the ability of multi-resolution segmentation to automatically derive objects from the simulated Sentinel-2 scene was tested. Bands 2, 3, 4, and 8 had 10 m pixel size, which is appropriate to detect vegetation objects in both urban and rural areas. Based on those four bands, the segmentation was able to detect and identify objects automatically using only three user-defined parameters: *scale*, *shape*, and *compactness*. Careful tuning of the *scale* parameter is critical for obtaining spectrally

homogeneous objects to apply spatial constraints. In case of doubt, smaller more homogeneous objects should be preferred.

Second, the usability of the four vegetation classes in the Bayesian optimization was investigated. Because of the broadness of the classes, general expert knowledge could be used for defining the a priori information: the VB1 class (low visual brightness) roughly corresponds to forest environments with high LAI and Cv, the VB4 class (high visual brightness) corresponds to almost-bare fields with very low LAI and Cv, the other two classes falling in between. High a priori variances were used to reflect the low confidence in the a priori values.

Third, the performance of the Bayesian object-based approach for estimating LAI and Cab based on the automatically segmented objects and on the four broad vegetation visible brightness classes was evaluated. No field data was available to validate the LAI and Cab maps, but the patterns of LAI, and Cv obtained at object-level were realistic, and the cost values were mostly low. The cost values were higher at pixel level, indicating a lesser performance. This increase, however, was mostly limited to VB1 objects, where the assumption of pixel internal homogeneity of the SLC canopy RT model was violated. For VB1 areas, the object-level estimates were therefore more reliable than the pixel-level ones.

Finally, automatic multi-resolution segmentation and the vegetation classification using four levels of visible brightness allowed successful implementation of the Bayesian object-based approach for the simulated Sentinel-2 scene, which had a much larger extent and smaller quantity of a priori information than was available in the APEX scene used for research question C (Laurent et al. In review-a). Because of the limitations related to the nadir-normalization, however, it should be preferred to work directly with the original data, using the original pixel viewing geometry. The accuracy of the LAI and Cab estimates would also be enhanced if topography and adjacency effects, important in the rugged terrain of the Zurich area, were also included in the coupled canopy-atmosphere model.

6.2 General conclusions

Having proposed to work directly with the measured radiance data, the main objective of this thesis was to explore the potential of the TOA approach for estimating biophysical and biochemical vegetation variables. The main conclusions from the work presented in this thesis are:

- The TOA approach had similar forward modelling accuracy, similar parameter influence profiles and the same dimensionality of 3 as the TOC approach for a case study with near-nadir CHRIS data. The small influence of the atmospheric parameters allowed the TOA approach to estimate Cv, fB, Cab and Cdm for the three Norway spruce stands as

accurately as the TOC approach. This similar performance is already sufficient to prefer the TOA approach over the TOC approach for applications where minimal data pre-processing is advantageous, such as multi-angular studies.

- The SVD of the Jacobian matrices for all possible angular combinations of the four available CHRIS observation directions for the same three Norway spruce stands showed that the dimensionality increased from 3 to 6 when increasing the number of directions from 1 to 4. Therefore, the higher information content of multi-angular data compared to that of mono-angular data, already recognized at TOC level, is also true at TOA level. The estimation results, however, showed that, similar to TOC level, the exploitation of this higher information content is still challenging.

- The Bayesian object-based approach allowed a high number of free variables in the Bayesian optimization (up to six) and efficiently implemented spatial constraints at the object level. These two features enabled the Bayesian object-based approach to provide more accurate LAI estimates than the LUT-BCF approach for the Oensingen agricultural area.

- Automatic multi-resolution image segmentation and general vegetation classification based on four levels of visible brightness allowed successful implementation of the Bayesian object-based approach for estimating LAI and Cab from the simulated Sentinel-2 scene of the Zurich area where no knowledge about the objects and a priori data was available.

6.3 Reflection

This thesis contributes to improving the accuracy of vegetation variables estimated from optical remote sensing data, by directly using TOA radiance measurements to invert a coupled canopy-atmosphere model. Intrinsic advantages of the TOA approach are the inclusion of the atmospheric effects in the forward modelling, and the comparison of the same physical quantities in the inversion. In addition, the 4-stream coupling set-up used in the four case studies of this thesis enabled to include the canopy directional effects in the modelling. Adjacency and topography effects, which were ignored in this thesis, could be included in the forward modelling to further improve the estimation accuracy.

The Jacobian matrix was used throughout this thesis for two purposes. In the first two case studies, it was used to investigate the parameter sensitivities and to diagnose the dimensionality of the estimation problem. This then allowed choosing the most influential variables and determining how many of them can be estimated. In the last two case studies, the Jacobian matrix was used to calculate the variable

update vector in the Bayesian optimization. This allowed making full use of the sensitivity spectral profile of the variables, thus enabling the estimation of up to six variables when the data dimensionality was only 4 in chapter 4.

Another recurring theme throughout this thesis was the issue of the weighting of the radiometric data. Spectral weights were used in chapter 2, the possible use of angular weights was discussed in chapter 3, while chapter 4 raised the issue of the implementation of the observation covariance matrix, which acts as weight for the radiometric data. The common underlying issue to this weighting problem is the evaluation of the information content of each radiance measurement, as related to its uncertainties (noise) and to its concurrent spectral and or/angular measurements. Furthermore, the observation covariance matrix should, in theory, also contain the model errors. In practice, however, these uncertainties are still unknown. Improved forward modelling and more appropriate weights might allow better exploiting the higher information content of multi-angular data.

The observation covariance matrix also contributes, together with the a priori covariance matrix, to the balance between the radiometric data and the a priori information. This is the core of the Bayesian approach, because the covariance matrices represent the degrees of belief. This balance issue is also related to the cost of the collection of the a priori data, which consists of both land cover information and a priori values, range and standard deviations. Chapter 5, however, showed that realistic estimates could still be obtained by the Bayesian object-based approach when using four broad vegetation classes and imprecise a priori data.

Exploring the potential of the TOA approach for estimating vegetation variables from mono-temporal data, this thesis successfully implemented four regularization methods in the TOA approach: model coupling, using multi-angular data, using a priori information, and applying spatial constraints. Working directly with the TOA radiance measurements facilitated the case studies because data pre-processing was reduced to a minimum. This is especially true for chapter 3, where four angular datasets would have had to be pre-processed independently for use in the traditional TOC approach.

Finally, the atmospheric parameters were kept fixed throughout this thesis. The coupled canopy-atmosphere modelling involved in the TOA approach, however, would allow including the atmospheric parameter as free variables to be estimated simultaneously with the vegetation variables. Atmospheric variables could then be spatially constrained on areas larger than vegetation objects.

6.4 Outlook

As was seen in this thesis, the accurate estimation of vegetation variables using physically-based approaches relies on three main pillars: 1) accurate forward modelling, 2) regularization methods, and 3) high quality radiometric data and

inclusion of their uncertainties in the model inversion. Further research should therefore contribute to these three pillars. Following this thesis, the following specific research directions have been identified for each pillar:

1. Include adjacency and topography effects in the coupled canopy-atmosphere model to increase the accuracy of the forward model
2. Implement temporal constraints to further regularize the inversion of the coupled model
3. Investigate observation and model uncertainties in order to build a sensible and comprehensive observation covariance matrix which will allow for better balancing of a priori and radiometric costs in the Bayesian approach

The minimal data pre-processing required in the TOA approach will facilitate research on the simultaneous use of several RS images in the model inversion, such as for multi-angular, multi-sensor and multi-temporal studies. For example, time series could be used for estimating vegetation variables in a data assimilation framework. The Bayesian approach used in chapters 4 and 5 is especially suited for this purpose, because the estimates obtained on a given date could be used as a priori data for the estimation on the next date. Such a set-up would also allow decreasing the amount of required a priori data and its related cost. This line of research will be supported by the upcoming generation of satellites having high revisit frequency, such as the Sentinel-2 mission(s) of the European Space Agency.

Direct data assimilation through the TOA approach will allow enhanced continuous monitoring of vegetation variables for DGVMs. More research will, however, be needed to meet the user requirements of the DGVM community regarding the estimation accuracy, as well as temporal and spatial resolutions of the remote sensing vegetation variable products. Furthermore, more accurate vegetation products will benefit many vegetation-related applications, such as forest management, precision agriculture, and carbon stock estimation in natural biomes.

References

Asner, G.P. (1998). Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment, 64* (3), 234-253

Asner, G.P., Braswell, B.H., Schimel, D.S., & Wessman, C.A. (1998). Ecological research needs from multiangle remote sensing data. *Remote Sensing of Environment, 63* (2), 155-165

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

Atzberger, C., & Richter, K. (2012). Spatially constrained inversion of radiative transfer models for improved LAI mapping from future Sentinel-2 imagery. *Remote Sensing of Environment, 120*, 208-218

Bach, H., & Verhoef, W. (2009). Sentinel-3 Land scene generator for Optical System Performance Simulator. Wessling, Germany: VISTA Geowissenschaftliche Fernerkundung GmbH, 59 pp

Bacour, C., Baret, F., Béal, D., Weiss, M., & Pavageau, K. (2006a). Neural network estimation of LAI, fAPAR, fCover and LAI×Cab, from top of canopy MERIS reflectance data: Principles and validation. *Remote Sensing of Environment, 105* (4), 313-325

Bacour, C., Bréon, F.M., & Maignan, F. (2006b). Normalization of the directional effects in NOAA-AVHRR reflectance measurements for an improved monitoring of vegetation cycles. *Remote Sensing of Environment, 102* (3-4), 402-413

Baret, F., & Buis, S. (2008). 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* (pp. 173-201): Springer

Baret, F., & Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment, 35* (2-3), 161-173

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J.L., & Leroy, M. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. Part 1: Principles of the algorithm. *Remote Sensing of Environment, 110* (3), 275-286

Baret, F., Pavageau, K., Béal, D., Weiss, M., Berthelot, B., & Regner, P. (2006). Algorithm Theoretical Basis Document for MERIS Top of Atmosphere Land Products (TOA_VEG). Avignon: INRA-CSE, 37 pp

Barnsley, M.J., Lewis, P., O'Dwyer, S., Disney, M.I., Hobson, P., Cutter, M., & Lobb, D. (2000). On the potential of CHRIS/PROBA for estimating vegetation canopy properties from space. *Remote Sensing Reviews, 19* (1-4), 171-189

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing, 58* (3-4), 239-258

Berger, M., Moreno, J., Johannessen, J.A., Levelt, P.F., & Hanssen, R.F. (2012). ESA's sentinel missions in support of Earth system science. *Remote Sensing of Environment, 120*, 84-90

Berk, A., Anderson, G.P., Acharya, P.K., Hoke, M.L., Chetwynd, J.H., Bernstein, L.S., Shettle, E.P., Matthew, M.W., & Adler-Golden, S.M. (2003). MODTRAN4 Version 3 Revision 1 User's manual. Airforce Research Laboratory, Hanscom, MA, USA, 97 pp

Bicheron, P., & Leroy, M. (2000). Bidirectional reflectance distribution function signatures of major biomes observed from space. *Journal of Geophysical Research D: Atmospheres, 105* (D21), 26669-26681

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 65* (1), 2-16

Börner, A., Wiest, L., Keller, P., Reulke, R., Richter, R., Schaepman, M., & Schläpfer, D. (2001). SENSOR: A tool for the simulation of hyperspectral remote sensing systems. *ISPRS Journal of Photogrammetry and Remote Sensing, 55* (5-6), 299-312

Chen, J.M. (1996). Optically-based methods for measuring seasonal variation of leaf area index in boreal conifer stands. *Agricultural and Forest Meteorology, 80* (2-4), 135-163

Chen, J.M., Liu, J., Leblanc, S.G., Lacaze, R., & Roujean, J.L. (2003). Multi-angular optical remote sensing for assessing vegetation structure and carbon absorption. *Remote Sensing of Environment, 84* (4), 516-525

Chopping, M., Moisen, G.G., Su, L., Laliberte, A., Rango, A., Martonchik, J.V., & Peters, D.P.C. (2008). Large area mapping of southwestern forest

crown cover, canopy height, and biomass using the NASA Multiangle Imaging Spectro-Radiometer. *Remote Sensing of Environment, 112* (5), 2051-2063

Ciganda, V., Gitelson, A., & Schepers, J. (2008). Vertical profile and temporal variation of chlorophyll in maize canopy: Quantitative "crop vigor" indicator by means of reflectance-based techniques. *Agronomy Journal, 100* (5), 1409-1417

Clevers, J.G.P.W., & Gitelson, A.A. (2012). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. *International Journal of Applied Earth Observation and Geoinformation, In press*

Clevers, J.G.P.W., & Verhoef, W. (1993). LAI estimation by means of the WDVI: A sensitivity analysis with a combined PROSPECT-SAIL model. *Remote Sensing Reviews, 7*, 43-64

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

CROMA (2000). Crop Reflectance Operational Models for Agriculture. Description of work., Energy, Environment and Sustainable Development work programme: EF5/PhD/0035.00, 44 pp

Cudlín, P., Novotný, R., Moravec, I., & Chmelíková, E. (2001). Retrospective evaluation of the response of montane forest ecosystems to multiple stress. *Ekologia Bratislava, 20* (1), 108-124

Cutter, M.A., Lobb, D.R., & Cockshott, R.A. (2000). Compact High Resolution Imaging Spectrometer (CHRIS). *Acta Astronautica, 46* (2), 263-268

D'Odorico, P., Gonsamo, A., Damm, A., & Schaepman, M.E. (2013). Experimental Evaluation of Sentinel-2 Spectral Response Functions for NDVI Time-Series Continuity. *Geoscience and Remote Sensing, IEEE Transactions on, PP* (99), 1-13

D'Odorico, P., Guanter, L., Schaepman, M.E., & Schläpfer, D. (2011). Performance assessment of onboard and scene-based methods for Airborne Prism Experiment spectral characterization. *Applied Optics, 50* (24), 4755-4764

Damm, A., Borgeaud, M., Brunner, D., Buchmann, B., Hüsler, F., Feigenwinter, C., Kneubühler, M., Parlow, E., Popp, C., Psomas, A., Schaepman, M., Stamenkovic, J., Thiran, J.-P., Wunderle, S., & Zimmermann, N. (2010). Data Acquisition Report 2010. Zurich: Remote Sensing Laboratories, 53 pp

Darvishzadeh, R., Atzberger, C., Skidmore, A., & Schlerf, M. (2011). Mapping grassland leaf area index with airborne hyperspectral imagery: A comparison study of statistical approaches and inversion of radiative transfer models. *ISPRS Journal of Photogrammetry and Remote Sensing, 66* (6), 894-906

Darvishzadeh, R., Skidmore, A., Schlerf, M., & Atzberger, C. (2008). Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. *Remote Sensing of Environment, 112* (5), 2592-2604

Dash, J., Jeganathan, C., & Atkinson, P.M. (2010). The use of MERIS Terrestrial Chlorophyll Index to study spatio-temporal variation in vegetation phenology over India. *Remote Sensing of Environment, 114* (7), 1388-1402

Dawson, T.P., Curran, P.J., & Plummer, S.E. (1998). LIBERTY - Modeling the effects of Leaf Biochemical Concentration on Reflectance Spectra. *Remote Sensing of Environment, 65* (1), 50-60

Deering, D.W., Middleton, E.M., & Eck, T.F. (1994). Reflectance anisotropy for a spruce-hemlock forest canopy. *Remote Sensing of Environment, 47* (2), 242-260

Delegido, J., Verrelst, J., Alonso, L., & Moreno, J. (2011). Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. *Sensors, 11* (7), 7063-7081

Diner, D.J., Asner, G.P., Davies, R., Knyazikhin, Y., Muller, J.P., Nolin, A.W., Pinty, B., Schaaf, C.B., & Stroeve, J. (1999). New Directions in Earth Observing: Scientific Applications of Multiangle Remote Sensing. *Bulletin of the American Meteorological Society, 80* (11), 2209-2228

Dorigo, W., Richter, R., Baret, F., Bamler, R., & Wagner, W. (2009). Enhanced automated canopy characterization from hyperspectral data by a novel two step radiative transfer model inversion approach. *Remote Sensing, 1* (4), 1139-1170

Dorigo, W.A., Zurita-Milla, R., de Wit, A.J.W., Brazile, J., Singh, R., & 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* (2), 165-193

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment, 120*, 25-36

Duchemin, B., Berthelot, B., Dedieu, G., Leroy, M., & Maisongrande, P. (2002). Normalisation of directional effects in 10-day global syntheses derived from VEGETATION/SPOT: II.

Validation of an operational method on actual data sets. *Remote Sensing of Environment, 81* (1), 101-113

Durbha, S.S., King, R.L., & Younan, N.H. (2007). Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment, 107* (1-2), 348-361

Fang, H., Liang, S., & Kuusk, A. (2003). Retrieving leaf area index using a genetic algorithm with a canopy radiative transfer model. *Remote Sensing of Environment, 85* (3), 257-270

FAO (2006). Chapter 2: Extent of forest resources. *Global forest resources assessment 2005: Progress towards sustainable forest management* (pp. 11-36)

FAO (2010). *Global Forest Resources Assessment 2010: Main Report*. Rome: Food and Agriculture Organization of the United Nations, 378 pp

Feret, J.B., François, C., Asner, G.P., Gitelson, A.A., Martin, R.E., Bidel, L.P.R., Ustin, S.L., Le Maire, G., & Jacquemoud, S. (2008). PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment, 112* (6), 3030-3043

Féret, J.B., François, C., Gitelson, A., Asner, G.P., Barry, K.M., Panigada, C., Richardson, A.D., & Jacquemoud, S. (2011). Optimizing spectral indices and chemometric analysis of leaf chemical properties using radiative transfer modeling. *Remote Sensing of Environment, 115* (10), 2742-2750

Foley, J.A., Levis, S., Costa, M.H., Cramer, W., & Pollard, D. (2000). Incorporating dynamic vegetation cover within global climate models. *Ecological Applications, 10* (6), 1620-1632

Fourty, T., & 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* (1), 34-45

Gao, B.C., Montes, M.J., Davis, C.O., & Goetz, A.F.H. (2009). Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. *Remote Sensing of Environment, 113* (SUPPL. 1), S17-S24

Gao, F., Schaaf, C.B., Strahler, A.H., & Lucht, W. (2001). Using a multikernel least-variance approach to retrieve and evaluate albedo from limited bidirectional measurements. *Remote Sensing of Environment, 76* (1), 57-66

Garrigues, S., Lacaze, R., Baret, F., Morisette, J.T., Weiss, M., Nickeson, J.E., Fernandes, R., Plummer, S., Shabanov, N.V., Myneni, R.B., Knyazikhin, Y., & Yang, W. (2008). Validation and intercomparison of global Leaf Area Index products derived from remote sensing data.

Journal of Geophysical Research G: Biogeosciences, 113 (2)

Gascon, F., Gastellu-Etchegorry, J.P., & Leroy, M. (2007). Using multi-directional high-resolution imagery from POLDER sensor to retrieve leaf area index. *International Journal of Remote Sensing, 28* (1), 167-181

Gastellu-Etchegorry, J.P., Martin, E., & Gascon, F. (2004). DART: A 3D model for simulating satellite images and studying surface radiation budget. *International Journal of Remote Sensing, 25* (1), 73-96

Gemmell, F., Varjo, J., Strandstrom, M., & Kuusk, A. (2002). Comparison of measured boreal forest characteristics with estimates from TM data and limited ancillary information using reflectance model inversion. *Remote Sensing of Environment, 81* (2-3), 365-377

Gerstl, S.A.W., & Zardecki, A. (1985). Coupled atmosphere/canopy model for remote sensing of plant reflectance features. *Applied Optics, 24* (1), 94-103

Gobron, N., Pinty, B., Aussedat, O., Taberner, M., Faber, O., Mélin, F., Lavergne, T., Robustelli, M., & Snoeij, P. (2008). Uncertainty estimates for the FAPAR operational products derived from MERIS - Impact of top-of-atmosphere radiance uncertainties and validation with field data. *Remote Sensing of Environment, 112* (4), 1871-1883

Gobron, N., Pinty, B., Verstraete, M.M., Widlowski, J.L., & Diner, D.J. (2002). Uniqueness of multiangular measurements - Part II: Joint retrieval of vegetation structure and photosynthetic activity from MISR. *IEEE Transactions on Geoscience and Remote Sensing, 40* (7), 1574-1592

Goel, N.S. (1988). Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data. *Remote Sensing Reviews, 4* (1), 1-221

Green, R.O., Pavri, B.E., & Chrien, T.G. (2003). On-orbit radiometric and spectral calibration characteristics of EO-1 hyperion derived with an underflight of AVIRIS and In situ measurements at Salar de Arizaro, Argentina. *IEEE Transactions on Geoscience and Remote Sensing, 41* (6 PART I), 1194-1203

Gu, D., & Gillespie, A. (1998). Topographic normalization of Landsat TM images of forest based on subpixel Sun-canopy-sensor geometry. *Remote Sensing of Environment, 64* (2), 166-175

Guanter, L., Richter, R., & Kaufmann, H. (2009). On the application of the MODTRAN4 atmospheric radiative transfer code to optical remote sensing. *International Journal of Remote Sensing, 30* (6), 1407-1424

Hagolle, O., Huc, M., Pascual, D.V., & Dedieu, G. (2010). A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENµS, LANDSAT and SENTINEL-2 images. *Remote Sensing of Environment, 114* (8), 1747-1755

Hall, F.G., Hilker, T., Coops, N.C., Lyapustin, A., Huemmrich, K.F., Middleton, E., Margolis, H., Drolet, G., & Black, T.A. (2008). Multi-angle remote sensing of forest light use efficiency by observing PRI variation with canopy shadow fraction. *Remote Sensing of Environment, 112* (7), 3201-3211

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., Dimiceli, C., & Carroll, M. (2002). Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment, 83* (1–2), 303-319

Hapke, B.W. (1981). Bi-directional reflectance spectroscopy: 1. Theory. *Journal of Geophysical Research, 86*, 3039−3054

Hedley, J., Roelfsema, C., Koetz, B., & Phinn, S. (2012). Capability of the Sentinel 2 mission for tropical coral reef mapping and coral bleaching detection. *Remote Sensing of Environment, 120*, 145-155

Heiskanen, J. (2006). Tree cover and height estimation in the Fennoscandian tundra-taiga transition zone using multiangular MISR data. *Remote Sensing of Environment, 103* (1), 97-114

Herrmann, I., Pimstein, A., Karnieli, A., Cohen, Y., Alchanatis, V., & Bonfil, D.J. (2011). LAI assessment of wheat and potato crops by VENµS and Sentinel-2 bands. *Remote Sensing of Environment, 115* (8), 2141-2151

Hilker, T., Coops, N.C., Coggins, S.B., Wulder, M.A., Brownc, M., Black, T.A., Nesic, Z., & Lessard, D. (2009). Detection of foliage conditions and disturbance from multi-angular high spectral resolution remote sensing. *Remote Sensing of Environment, 113* (2), 421-434

Hilker, T., Gitelson, A., Coops, N.C., Hall, F.G., & Black, T.A. (2011). Tracking plant physiological properties from multi-angular tower-based remote sensing. *Oecologia, 165* (4), 865-876

Homolová, L., Malenovský, Z., Hanuš, J., Tomášková, I., Dvořáková, M., & Pokorný, R. (2007). Comparison of different ground techniques to map leaf area index of Norway spruce forest canopy. 10th ISPMSRS, Davos, Switzerland, 6 pp

Houborg, R., Anderson, M., & Daughtry, C. (2009). Utility of an image-based canopy reflectance modeling tool for remote estimation of LAI and leaf chlorophyll content at the field scale. *Remote Sensing of Environment, 113* (1), 259-274

Houborg, R., & Boegh, E. (2008). Mapping leaf chlorophyll and leaf area index using inverse and forward canopy reflectance modeling and SPOT reflectance data. *Remote Sensing of Environment, 112* (1), 186-202

Houborg, R., Soegaard, H., & 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* (1), 39-58

Huber, S., Koetz, B., Psomas, A., Kneubühler, M., Schopfer, J., Itten, K., & Zimmermann, N.E. (2010). Impact of multiangular information on empirical models to estimate canopy nitrogen concentration in mixed forest. *Journal of Applied Remote Sensing, 4* (043530)

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., & Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment, 83* (1-2), 195-213

Hüni, A., Bieseman, J., Dell'Endice, F., Alberti, E., Meuleman, K., & Schaepman, M. (2009). The structure of the APEX (airborne prism experiment) Processing and Archiving Facility. First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS '09), p 1-4

IPCC (2007). *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press, 996 pp

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

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., & Ustin, S.L. (2009). PROSPECT + SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment, 113* (SUPPL. 1), S56–S66

Jehle, M., Hueni, A., Damm, A., D'Odorico, P., Weyermann, J., Kneubühler, M., Schläpfer, D., Schaepman, M.E., & Meuleman, K. (2010). APEX - Current status, performance and validation concept. IEEE Sensors 2010 Conference, Waikoloa, HI, USA, p 533-537

Johansen, K., Coops, N.C., Gergel, S.E., & Stange, Y. (2007). Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. *Remote Sensing of Environment, 110* (1), 29-44

Kimes, D.S., Knyazikhin, Y., Privette, J.L., Abuelgasim, A.A., & Gao, F. (2000). Inversion methods for physically-based models. *Remote Sensing Reviews, 18* (2), 381-439

Kimes, D.S., Nelson, R.F., Manry, M.T., & Fung, A.K. (1998). Attributes of neural networks for extracting continuous vegetation variables from optical and radar measurements. *International Journal of Remote Sensing, 19* (14), 2639-2662

Kneubühler, M., Koetz, B., Huber, S., Schaepman, M.E., & Zimmermann, N.E. (2008). Space-based spectrodirectional measurements for the improved estimation of ecosystem variables. *Canadian Journal of Remote Sensing, 34* (3), 192-205

Knyazikhin, Y., Glassy, J., Privette, J.L., Tian, Y., Lotsch, A., Zhang, Y., Wang, Y., Morisette, J.T., Votava, P., Myneni, R.B., Nemani, R.R., & Running, S.W. (1999). 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.V., Diner, D.J., Myneni, R.B., Verstraete, M., Pinty, B., & Gobron, N. (1998). Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from atmosphere-corrected MISR data. *Journal of Geophysical Research D: Atmospheres, 103* (D24), 32239-32256

Koetz, B., Kneubuehler, M., Huber, S., Schopfer, J., & Baret, F. (2007). LAI estimation based on multi-temporal chris/proba data and radiative transfer modeling. European Space Agency, (Special Publication) ESA SP-636, Montreux, p 1-6

Kokaly, R.F., Asner, G.P., Ollinger, S.V., Martin, M.E., & Wessman, C.A. (2009). Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment, 113* (SUPPL. 1), S78-S91

Kötz, B., Baret, F., Poilve, H., & Hill, J. (2005). Use of coupled canopy structure dynamic and radiative transfer models to estimate biophysical canopy characteristics. *Remote Sensing of Environment, 95* (1), 115-124

Kratochvilová, I., Janouš, D., Marek, M., Barták, M., & Řiha, L. (1989). Production activity of mountain cultivated Norway spruce stands under the impact of air pollution. I. General description of problems. *Ekológia, 8* (4), 407-419

Kuusk, A., & Nilson, T. (2000). A directional multispectral forest reflectance model. *Remote Sensing of Environment, 72* (2), 244-252

Lamonaca, A., Corona, P., & Barbati, A. (2008). Exploring forest structural complexity by multi-scale segmentation of VHR imagery. *Remote Sensing of Environment, 112* (6), 2839-2849

Laurent, V.C.E., Clevers, J.G.P.W., Schaepman, M., & Verhoef, W. (2009). Simulation of forest radiance at top-of-atmosphere level using coupled radiative transfer models. 6th EARSeL SIG IS Workshop, Tel Aviv, Israel, 6 pp

Laurent, V.C.E., Schaepman, M.E., Verhoef, W., Weyermann, J., & Chávez, R.O. (In review-a). Bayesian object-based estimation of LAI and chlorophyll from a simulated Sentinel-2 top-of-atmosphere radiance scene. *Remote Sensing of Environment*

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W., & Schaepman, M. (2010a). Estimating forest parameters from top-of-atmosphere radiance measurements using coupled radiative transfer models. Hyperspectral Workshop 2010, Frascati, Italy, 8 pp

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W., & Schaepman, M. (2010b). Estimating forest parameters from top of atmosphere multi-angular radiance data using coupled radiative transfer models. Recent Advances in Quantitative Remote Sensing, Valencia, Spain, 6 pp

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W., & Schaepman, M.E. (2011a). Added value of multiangular measurements for estimating forest variables from the top of the atmosphere using coupled radiative transfer models. EARSeL 7th SIG-Imaging Spectroscopy Workshop, Edinburgh, 6 pp

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W., & Schaepman, M.E. (2011b). Estimating forest variables from top-of-atmosphere radiance satellite measurements using coupled radiative transfer models. *Remote Sensing of Environment, 115* (4), 1043-1052

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W., & Schaepman, M.E. (2011c). Inversion of a coupled canopy-atmosphere model using multi-angular top-of-atmosphere radiance data: A forest case study. *Remote Sensing of Environment, 115* (10), 2603-2612

Laurent, V.C.E., Verhoef, W., Damm, A., Schaepman, M.E., & Clevers, J.G.P.W. (In review-b). A Bayesian object-based approach for estimating vegetation biophysical and biochemical variables from at-sensor APEX data. *Remote Sensing of Environment*

Laurent, V.C.E., Verhoef, W., Schaepman, M.E., Damm, A., & Clevers, J.G.P.W. (2012). Mapping LAI and chlorophyll content from at-sensor APEX data using a Bayesian optimization of a coupled canopy-atmosphere model. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, 4 pp

Lauvernet, C., Baret, F., Hascoët, L., Buis, S., & Le Dimet, F.X. (2008). Multitemporal-patch ensemble inversion of coupled surface-atmosphere radiative transfer models for land surface

References

characterization. *Remote Sensing of Environment, 112* (3), 851-861

Lavergne, T., Kaminski, T., Pinty, B., Taberner, M., Gobron, N., Verstraete, M.M., Voβbeck, M., Widlowski, J.L., & Giering, R. (2007). Application to MISR land products of an RPV model inversion package using adjoint and Hessian codes. *Remote Sensing of Environment, 107* (1-2), 362-375

Lewis, P., Gómez-Dans, J., Kaminski, T., Settle, J., Quaife, T., Gobron, N., Styles, J., & Berger, M. (2012). An Earth Observation Land Data Assimilation System (EO-LDAS). *Remote Sensing of Environment, 120*, 219-235

Li, W.H., Weeks, R., & Gillespie, A.R. (1998). Multiple scattering in the remote sensing of natural surfaces. *International Journal of Remote Sensing, 19* (9), 1725-1740

Li, X., Gao, F., Wang, J., & Strahler, A. (2001). A priori knowledge accumulation and its application to linear BRDF model inversion. *Journal of Geophysical Research D: Atmospheres, 106* (D11), 11925-11935

Li, X., & Strahler, A. (1996). Knowledge-based inversion of physical BRDF model and three examples. International Geoscience and Remote Sensing Symposium (IGARSS), Lincoln, NE, USA, p 2173-2176

Liang, S., Strahler, A.H., Barnsley, M.J., Borel, C.C., Gerstl, S.A.W., Diner, D.J., Prata, A.J., & Walthall, C.L. (2000). Multiangle remote sensing: Past, present and future. *Remote Sensing Reviews, 18* (2), 83-102

Lukeš, P. (2009). Retrieval of canopy cover of the Norway spruce stands in the Bily Kriz area (CZ) from classification of AISA Eagle data. Personal communication on 22 January 2009

Ma, J., Chan, J.C.W., & Canters, F. (2010). Fully automatic subpixel image registration of multiangle CHRIS/proba data. *IEEE Transactions on Geoscience and Remote Sensing, 48* (7), 2829-2839

Malenovský, Z., Albrechtová, J., Lhotáková, Z., Zurita-Milla, R., Clevers, J.G.P.W., Schaepman, M.E., & Cudlín, P. (2006). Applicability of the PROSPECT model for Norway spruce needles. *International Journal of Remote Sensing, 27* (24), 5315-5340

Malenovský, Z., Martin, E., Homolová, L., Gastellu-Etchegorry, J.P., Zurita-Milla, R., Schaepman, M.E., Pokorný, R., Clevers, J.G.P.W., & Cudlín, P. (2008). Influence of woody elements of a Norway spruce canopy on nadir reflectance simulated by the DART model at very high spatial resolution. *Remote Sensing of Environment, 112* (1), 1-18

Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M.E., García-Santos, G., Fernandes, R., & Berger, M. (2012). Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sensing of Environment, 120* (0), 91-101

Marek, M.V., Malenovský, Z., Homolová, L., Hanuš, J., Kaplan, V., & Lukeš, P. (2007). Spectral-spatial scaling from leaf to canopy level using spectro-directional approaches in support of the GMES Sentinel 2: 'Superspectral' mission - 3rd ESA/PECS project report for period of January 1st 2007 – December 31st 2007. Institute of Systems Biology and Ecology, Academy of Sciences of the Czech Republic (ISBE ASCR), 27 pp

Massada, A.B., Kent, R., Blank, L., Perevolotsky, A., Hadar, L., & Carmel, Y. (2012). Automated segmentation of vegetation structure units in a mediterranean landscape. *International Journal of Remote Sensing, 33* (2), 346-364

Meroni, M., Colombo, R., & Panigada, C. (2004). Inversion of a radiative transfer model with hyperspectral observations for LAI mapping in poplar plantations. *Remote Sensing of Environment, 92* (2), 195-206

Migdall, S., Bach, H., Bobert, J., Wehrhan, M., & Mauser, W. (2009). Inversion of a canopy reflectance model using hyperspectral imagery for monitoring wheat growth and estimating yield. *Precision Agriculture*, 1-17

Morisette, J.T., Privette, J.L., & Justice, C.O. (2002). A framework for the validation of MODIS Land products. *Remote Sensing of Environment, 83* (1-2), 77-96

Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G.R., Lotsch, A., Friedl, M., Morisette, J.T., Votava, P., Nemani, R.R., & Running, S.W. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment, 83* (1-2), 214-231

Olioso, A., Inoue, Y., Ortega-Farias, S., Demarty, J., Wigneron, J.P., Braud, I., Jacob, F., Lecharpentier, P., Ottlé, C., Calvet, J.C., & Brisson, N. (2005). Future directions for advanced evapotranspiration modeling: Assimilation of remote sensing data into crop simulation models and SVAT models. *Irrigation and Drainage Systems, 19* (3-4), 377-412

Pandya, M.R., Singh, R.P., & Panigrahy, S. (2007). Directional reflectance of vegetation targets: Simulation of its space measurements by coupling atmospheric and biophysical radiative transfer models. *Indian Journal of Radio & Space Physics, 36* (3), 219-228

Pinty, B., Lattanzio, A., Martonchik, J.V., Verstraete, M.M., Gobron, N., Taberner, M., Widlowski, J.L., Dickinson, R.E., & Govaerts, Y. (2005). Coupling diffuse sky radiation and surface albedo. *Journal of the Atmospheric Sciences, 62* (7 II), 2580-2591

Pinty, B., Lavergne, T., Vossbeck, M., Kaminski, T., Aussedat, O., Giering, R., Gobron, N., Taberner, M., Verstraete, M.M., & Widlowski, J.L. (2007). Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS)-Multiangle Imaging Spectroradiometer (MISR) albedo products. *Journal of Geophysical Research D: Atmospheres, 112* (10)

Pinty, B., & Verstraete, M.M. (1992). On the design and validation of surface bidirectional reflectance and albedo models. *Remote Sensing of Environment, 41* (2-3), 155-167

Proy, C., Tanré, D., & Deschamps, P.Y. (1989). Evaluation of topographic effects in remotely sensed data. *Remote Sensing of Environment, 30* (1), 21-32

Quaife, T., & Lewis, P. (2010). Temporal constraints on linear BRDF model parameters. *IEEE Transactions on Geoscience and Remote Sensing, 48* (5), 2445-2450

Rahman, H. (2001). Influence of atmospheric correction on the estimation of biophysical parameters of crop canopy using satellite remote sensing. *International Journal of Remote Sensing, 22* (7), 1245-1268

Rahman, H., & Dedieu, G. (1994). SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. *International Journal of Remote Sensing, 15* (1), 123-143

Rahman, H., Verstraete, M.M., & Pinty, B. (1993). Coupled surface-atmosphere reflectance (CSAR) model 1. Model description and inversion on synthetic data. *Journal of Geophysical Research, 98* (D11), 20779-20789

Ramankutty, N., & Foley, J.A. (1999). Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles, 13* (4), 997-1027

Rascher, U., Agati, G., Alonso, L., Cecchi, G., Champagne, S., Colombo, R., Damm, A., Daumard, F., De Miguel, E., Fernandez, G., Franch, B., Franke, J., Gerbig, C., Gioli, B., Gómez, J.A., Goulas, Y., Guanter, L., Gutiérrez-De-La-Cámara, É., Hamdi, K., Hostert, P., Jiménez, M., Kosvancova, M., Lognoli, D., Meroni, M., Miglietta, F., Moersch, A., Moreno, J., Moya, I., Neininger, B., Okujeni, A., Ounis, A., Palombi, L., Raimondi, V., Schickling, A., Sobrino, J.A., Stellmes, M., Toci, G., Toscano, P., Udelhoven, T., Van Der Linden, S., & Zaldei, A.

(2009). CEFLES2: The remote sensing component to quantify photosynthetic efficiency from the leaf to the region by measuring sun-induced fluorescence in the oxygen absorption bands. *Biogeosciences, 6* (7), 1181-1198

Rautiainen, M., Lang, M., Mõttus, M., Kuusk, A., Nilson, T., Kuusk, J., & Lükk, T. (2008). Multi-angular reflectance properties of a hemiboreal forest: An analysis using CHRIS PROBA data. *Remote Sensing of Environment, 112* (5), 2627-2642

Riaño, D., Chuvieco, E., Salas, J., & Aguado, I. (2003). Assessment of different topographic corrections in landsat-TM data for mapping vegetation types (2003). *IEEE Transactions on Geoscience and Remote Sensing, 41* (5 PART 1), 1056-1061

Richter, K., Atzberger, C., Vuolo, F., & D'Urso, G. (2011a). Evaluation of Sentinel-2 spectral sampling for radiative transfer model based LAI estimation of wheat, sugar beet, and maize. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 4* (2), 458-464

Richter, K., Atzberger, C., Vuolo, F., Weihs, P., & D'Urso, G. (2009). Experimental assessment of the Sentinel-2 band setting for RTM-based LAI retrieval of sugar beet and maize. *Canadian Journal of Remote Sensing, 35* (3), 230-247

Richter, K., Hank, T.B., Vuolo, F., Mauser, W., & D'Urso, G. (2012). Optimal exploitation of the sentinel-2 spectral capabilities for crop leaf area index mapping. *Remote Sensing, 4* (3), 561-582

Richter, R. (2009). Atmospheric / Topographic Correction for Satellite Imagery (ATCOR-2/3 User Guide, version 7.0, January 2009). Wessling, Germany: DLR - German Aerospace Center - Remote Sensing Data Center, 152 pp

Richter, R., & Schläpfer, D. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: Atmospheric/topographic correction. *International Journal of Remote Sensing, 23* (13), 2631-2649

Richter, R., Schläpfer, D., & Müller, A. (2006). An automatic atmospheric correction algorithm for visible/NIR imagery. *International Journal of Remote Sensing, 27* (10), 2077-2085

Richter, R., Wang, X., Bachmann, M., & Schläpfer, D. (2011b). Correction of cirrus effects in Sentinel-2 type of imagery. *International Journal of Remote Sensing, 32* (10), 2931-2941

Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C., Zhang, X., Jin, Y., Muller, J.P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., D'Entremont, R.P., Hu, B., Liang, S., Privette, J.L., & Roy, D. (2002). First operational BRDF, albedo nadir reflectance products from MODIS.

Remote Sensing of Environment, 83 (1-2), 135-148

Schaepman-Strub, G., Schaepman, M.E., Painter, T.H., Dangel, S., & Martonchik, J.V. (2006). Reflectance quantities in optical remote sensing-definitions and case studies. *Remote Sensing of Environment, 103* (1), 27-42

Schaepman, M.E. (2007). Spectrodirectional remote sensing: From pixels to processes. *International Journal of Applied Earth Observation and Geoinformation, 9* (2), 204-223

Schaepman, M.E., Ustin, S.L., Plaza, A.J., Painter, T.H., Verrelst, J., & Liang, S. (2009). Earth system science related imaging spectroscopy-An assessment. *Remote Sensing of Environment, 113* (SUPPL. 1), S123–S137

Schimel, D.S., House, J.I., Hibbard, K.A., Bousquet, P., Ciais, P., Peylin, P., Braswell, B.H., Apps, M.J., Baker, D., Bondeau, A., Canadell, J., Churkina, G., Cramer, W., Denning, A.S., Field, C.B., Friedlingstein, P., Goodale, C., Heimann, M., Houghton, R.A., Melillo, J.M., Moore Iii, B., Murdiyarso, D., Noble, I., Pacala, S.W., Prentice, I.C., Raupach, M.R., Rayner, P.J., Scholes, R.J., Steffen, W.L., & Wirth, C. (2001). Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature, 414* (6860), 169-172

Schläpfer, D., & Schaepman, M. (2002). Modeling the noise equivalent radiance requirements of imaging spectrometers based on scientific applications. *Applied Optics, 41* (27), 5691-5701

Schläpfer, D., Schaepman, M.E., & Itten, K.I. (1998). PARGE: Parametric geocoding based on GCP-calibrated auxiliary data. Imaging Spectrometry IV, San Diego, p 334-344

Schlerf, M., Atzberger, C., & Hill, J. (2005). Remote sensing of forest biophysical variables using HyMap imaging spectrometer data. *Remote Sensing of Environment, 95* (2), 177-194

Schlerf, M., Verhoef, W., Buddenbaum, H., Hill, J., Atzberger, C., & Skidmore, A. (2007). Comparing three canopy reflectance models with hyperspectral multi-angular satellite data. 10th International Symposium on Physical Measurements and Signatures in Remote Sensing (ISPMSRS'07), Davos, Switzerland, p 404-407

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

Semenov, A.A., Moshkov, A.V., Pozhidayev, V.N., Barducci, A., Marcoionni, P., & Pippi, I. (2011). Estimation of Normalized Atmospheric Point Spread Function and Restoration of Remotely Sensed Images. *IEEE Transactions on Geoscience and Remote Sensing, 49* (7), 2623-2634

Sentinel-2 PDGS Project Team (2011). GMES space component, Sentinel-2 payload data ground segment (PDGS), Products definition document. GMES-GSEG-EOPG-TN-09-0029, issue 2, revision 1. ESA, 93 pp

Settle, J. (2004). On the dimensionality of multi-view hyperspectral measurements of vegetation. *Remote Sensing of Environment, 90* (2), 235-242

Simic, A., & Chen, J.M. (2008). Refining a hyperspectral and multiangle measurement concept for vegetation structure assessment. *Canadian Journal of Remote Sensing, 34* (3), 174-191

Sitch, S., Huntingford, C., Gedney, N., Levy, P.E., Lomas, M., Piao, S.L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C.D., Prentice, I.C., & Woodward, F.I. (2008). Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Global Change Biology, 14* (9), 2015-2039

Smith, B., Prentice, I.C., & Sykes, M.T. (2001). Representation of vegetation dynamics in the modelling of terrestrial ecosystems: Comparing two contrasting approaches within European climate space. *Global Ecology and Biogeography, 10* (6), 621-637

Soenen, S.A., Peddle, D.R., & Coburn, C.A. (2005). SCS+C: A modified sun-canopy-sensor topographic correction in forested terrain. *IEEE Transactions on Geoscience and Remote Sensing, 43* (9), 2148-2159

Soenen, S.A., Peddle, D.R., Coburn, C.A., Hall, R.J., & Hall, F.G. (2009). Canopy reflectance model inversion in multiple forward mode: Forest structural information retrieval from solution set distributions. *Photogrammetric Engineering and Remote Sensing, 75* (4), 361-374

Stenberg, P. (2007). Simple analytical formula for calculating average photon recollision probability in vegetation canopies. *Remote Sensing of Environment, 109* (2), 221-224

Susaki, J., Hara, K., Kajiwara, K., & Honda, Y. (2004). Robust estimation of BRDF model parameters. *Remote Sensing of Environment, 89* (1), 63-71

Tanré, D., Deschamps, P.Y., Duhaut, P., & Herman, M. (1987). Adjacency effect produced by the atmospheric scattering in Thematic Mapper data. *Journal of Geophysical Research, 92* (D10), 12000-12006

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Philadelphia, USA: Society for Industrial and Applied Mathematics, 342 pp

Tian, Y., Woodcock, C.E., Wang, Y., Privette, J.L., Shabanov, N.V., Zhou, L., Zhang, Y., Buermann, W., Dong, J., Veikkanen, B., Häme, T., Andersson, K., Ozdogan, M., Knyazikhin, Y., &

Myneni, R.B. (2002). Multiscale analysis and validation of the MODIS LAI product I. Uncertainty assessment. *Remote Sensing of Environment, 83* (3), 414-430

Timmermans, J., Verhoef, W., Van der Tol, C., & Su, Z. (2009). Retrieval of canopy component temperatures through Bayesian inversion of directional thermal measurements. *Hydrology and Earth System Sciences, 13* (7), 1249-1260

Ustin, S.L., Gitelson, A.A., Jacquemoud, S., Schaepman, M., Asner, G.P., Gamon, J.A., & Zarco-Tejada, P. (2009). Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. *Remote Sensing of Environment, 113* (SUPPL. 1), S67–S77

Verger, A., Martínez, B., Coca, F.C.D., & García-Haro, F.J. (2009). Accuracy assessment of fraction of vegetation cover and leaf area index estimates from pragmatic methods in a cropland area. *International Journal of Remote Sensing, 30* (10), 2685-2704

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

Verhoef, W. (2007). A Bayesian optimisation approach for model inversion of hyperspectral-multidirectional observations: The balance with a priori information. 10th ISPMSRS, Davos, Switzerland, 6 pp

Verhoef, W., & Bach, H. (2003a). Remote sensing data assimilation using coupled radiative transfer models. *Physics and Chemistry of the Earth, 28* (1-3), 3-13

Verhoef, W., & Bach, H. (2003b). Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models. *Remote Sensing of Environment, 87* (1), 23-41

Verhoef, W., & 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* (2), 166-182

Vermote, E.F., El Saleous, N., Justice, C.O., Kaufman, Y.J., Privette, J.L., Remer, L., Roger, J.C., & Tanré, D. (1997a). Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: Background, operational algorithm and validation. *Journal of Geophysical Research D: Atmospheres, 102* (14), 17131-17141

Vermote, E.F., Tanré, D., Deuzé, J.L., Herman, M., & Morcrette, J.J. (1997b). Second simulation of the satellite signal in the solar spectrum, 6s: an overview. *IEEE Transactions on Geoscience and Remote Sensing, 35* (3), 675-686

Verrelst, J., Clevers, J.G.P.W., & Schaepman, M.E. (2010a). Merging the Minnaert-k parameter with spectral unmixing to map forest heterogeneity with CHRIS/PROBA data. *IEEE Transactions on Geoscience and Remote Sensing, 48* (11), 4014-4022

Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera, J.P., Camps-Valls, G., & Moreno, J. (2012). Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sensing of Environment, 118*, 127-139

Verrelst, J., Schaepman, M.E., & Clevers, J.G.P.W. (2010b). Spectrodirectional Minnaert-k retrieval using CHRIS/PROBA data. *Canadian Journal of Remote Sensing, 36* (6), 631-644

Verrelst, J., Schaepman, M.E., Koetz, B., & Kneubülhler, M. (2008). Angular sensitivity analysis of vegetation indices derived from CHRIS/PROBA data. *Remote Sensing of Environment, 112* (5), 2341-2353

Verrelst, J., Schaepman, M.E., Malenovský, Z., & Clevers, J.G.P.W. (2010c). Effects of woody elements on simulated canopy reflectance: Implications for forest chlorophyll content retrieval. *Remote Sensing of Environment, 114* (3), 647-656

Vitousek, P.M., Mooney, H.A., Lubchenco, J., & Melillo, J.M. (1997). Human domination of Earth's ecosystems. *Science, 277* (5325), 494-499

Vohland, M., Mader, S., & Dorigo, W. (2010). Applying different inversion techniques to retrieve stand variables of summer barley with PROSPECT + SAIL. *International Journal of Applied Earth Observation and Geoinformation, 12* (2), 71-80

Wang, L., Sousa, W.P., & Gong, P. (2004). Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. *International Journal of Remote Sensing, 25* (24), 5655-5668

Wang, Y.F., Yang, C., & Li, X. (2008). Regularizing kernel-based BRDF model inversion method for ill-posed land surface parameter retrieval using smoothness constraint. *Journal of Geophysical Research D: Atmospheres, 113* (13)

Weiss, M., Baret, F., Myneni, R.B., Pragnère, A., & Knyazikhin, Y. (2000). Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. *Agronomie, 20* (1), 3-22

Weyermann, J., Damm, A., Kneubühler, M., & Schaepman, M.E. (Accepted). Correction of reflectance anisotropy effects of vegetation on airborne spectroscopy data and derived products. *IEEE Transactions on Geoscience and Remote Sensing*

References

Widlowski, J.L., Pinty, B., Gobron, N., & Verstraete, M.M. (2001). Detection and characterization of boreal coniferous forests from remote sensing data. *Journal of Geophysical Research D: Atmospheres, 106* (D24), 33405-33419

Widlowski, J.L., Pinty, B., Gobron, N., Verstraete, M.M., Diner, D.J., & Davis, A.B. (2004). Canopy structure parameters derived from multi-angular remote sensing data for terrestrial carbon studies. *Climatic Change, 67* (2-3), 403-415

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.E., Martin, E., Mõttus, M., North, P.R.J., Qin, W., Robustelli, M., Rochdi, N., Ruiloba, R., Soler, C., Thompson, R., Verhoef, W., Verstraete, M.M., & Xie, D. (2007). Third Radiation Transfer Model Intercomparison (RAMI) exercise: Documenting progress in canopy reflectance models. *Journal of Geophysical Research D: Atmospheres, 112* (9)

Winterhalter, L., Mistele, B., & Schmidhalter, U. (2012). Assessing the vertical footprint of reflectance measurements to characterize nitrogen uptake and biomass distribution in maize canopies. *Field Crops Research, 129* (0), 14-20

Yang, W., Shabanov, N.V., Huang, D., Wang, W., Dickinson, R.E., Nemani, R.R., Knyazikhin, Y., & Myneni, R.B. (2006a). Analysis of leaf area index products from combination of MODIS Terra and Aqua data. *Remote Sensing of Environment, 104* (3), 297-312

Yang, W., Tan, B., Huang, D., Rautiainen, M., Shabanov, N.V., Wang, Y., Privette, J.L., Huemmrich, K.F., Fensholt, R., Sandholt, I., Weiss, M., Ahl, D.E., Gower, S.T., Nemani, R.R., Knyazikhin, Y., & Myneni, R.B. (2006b). MODIS leaf area index products: From validation to algorithm improvement. *IEEE Transactions on Geoscience and Remote Sensing, 44* (7), 1885-1896

Yao, Y., Liu, Q., & Li, X. (2008). LAI retrieval and uncertainty evaluations for typical row-planted crops at different growth stages. *Remote Sensing of Environment, 112* (1), 94-106

Summary

Vegetation is an important component of the Earth's biosphere and therefore plays a crucial role in the carbon exchange of terrestrial ecosystems. Vegetation variables, such as leaf area index (LAI) and leaf chlorophyll content (Cab), can be monitored at global scale using remote sensing (RS). There are two main categories of approaches for estimating the vegetation variables from RS data: empirical and physically-based approaches. Physically-based approaches are more widely applicable because they rely on radiative transfer (RT) models, which can be adapted to the observation conditions and to the observed vegetation. For estimating the vegetation variables, however, the RT model has to be inverted, and this inversion is usually an ill-posed and under-determined problem. Several regularization methods have been proposed to allow finding stable and unique solutions: model coupling, using multi-angular data, using a priori information, as well as applying spatial or temporal constraints.

Traditionally, radiance data measured at top-of the atmosphere (TOA) are pre-processed to top-of-canopy (TOC) reflectances. Corrections for atmospheric effects, and, if needed, for adjacency, directional, or topographic effects are usually applied sequentially and independently. Physically, however, these effects are inter-related, and each correction introduces errors. These errors propagate to the TOC reflectance data, which are used to invert the canopy RT model. The performance of the TOC approach is therefore limited by the errors introduced in the data during the pre-processing steps.

This thesis proposes to minimize these errors by directly using measured TOA radiance data. In such a TOA approach, the atmospheric RT model, which is normally inverted to perform the atmospheric correction, is coupled to the canopy RT model. The coupled canopy-atmosphere model is inverted directly using the measured radiance data. Adjacency, directional and topographic effects can then be included in the coupled RT model. The same regularization methods as used for TOC approaches can be applied to obtain stable and unique estimates. The TOA approach was tested using four case studies based on mono-temporal data.

A) The performance of the TOA approach was compared to a TOC approach for three Norway spruce stands in the Czech Republic, using near-nadir Compact High Resolution Imaging Spectrometer (CHRIS) data. The coupled model included canopy directional effects and simulated the CHRIS radiance data with similar accuracy as the canopy model simulated the atmospherically-corrected CHRIS data. Local sensitivity analyses showed that the atmospheric parameters had much less influence on the simulations than the vegetation parameters, and that the sensitivity profiles of the latter were very similar for both TOC and TOA approaches. The dimensionality

of the estimation problem was evaluated to be 3 for both approaches. Canopy cover (Cv), fraction of bark material (fB), Cab, and leaf dry matter content (Cdm) were estimated using look-up tables (LUT) with similar accuracy with both approaches.

B) Regularization using multi-angular data was tested for the TOA approach, using four angular CHRIS datasets, for the same three stands as used in A). The coupled model provided good simulations for all angles. The dimensionality increased from 3 to 6 when using all four angles. Two LUTs were built for each stand: a 4-variable LUT with fB, Cv, Cdm, and Cab, and a 7-variable LUT where leaf brown pigment concentration (Cs), dissociation factor (D), and tree shape factor (Zeta) were added. The results did not fully match the expectation that the more angles used, the more accurate the estimates become. Although their exploitation remains challenging, multi-angular data have higher potential than mono-angular data at TOA level.

C) A Bayesian object-based approach was developed and tested on at-sensor Airborne Prism Experiment (APEX) radiance data for an agricultural area in Switzerland. This approach consists of two steps. First, up to six variables were estimated for each crop field object using a Bayesian optimization algorithm, using a priori information. Second, a LUT was built for each object with only LAI and Cab as free variables, thus spatially constraining the values of all other variables to the values obtained in the first step. The Bayesian object-based approach estimated LAI more accurately than a LUT with a Bayesian cost function approach. This case study relied on extensive field data allowing defining the objects and a priori data.

D) The Bayesian object-based approach proposed in C) was applied to a simulated TOA Sentinel-2 scene, covering the area around Zurich, Switzerland. The simulated scene was mosaicked using seven APEX flight lines, which allowed including all spatial and spectral characteristics of Sentinel-2. Automatic multi-resolution segmentation and classification of the vegetated objects in four levels of brightness in the visible domain enabled defining the objects and a priori data without field data, allowing successful implementation of the Bayesian object-based approach.

The research conducted in this thesis contributes to the improvement of the use of regularization methods in ill-posed RT model inversions. Three major areas were identified for further research: 1) inclusion of adjacency and topography effects in the coupled model, 2) addition of temporal constraints in the inversion, and 3) better inclusion of observation and model uncertainties in the cost function. The TOA approach proposed here will facilitate the exploitation of multi-angular, multi-temporal and multi-sensor data, leading to more accurate RS vegetation products. These higher quality products will support many vegetation-related applications.

Samenvatting

Vegetatie is een belangrijk onderdeel van de biosfeer van de aarde en speelt daarom een cruciale rol in de koolstofuitwisseling van terrestrische ecosystemen. Vegetatievariabelen, zoals de bladoppervlakte-index (LAI) en het chlorofylgehalte (Cab) van het blad, kunnen op mondiale schaal gemonitord worden met behulp van remote sensing (RS). Er zijn twee hoofdcategorieën van methoden voor het schatten van de vegetatievariabelen met RS gegevens: empirische en fysische benaderingen. Fysische benaderingen zijn meer algemeen toepasbaar omdat zij gebaseerd zijn op stralingsinteractie (RT) modellen, die aangepast kunnen worden aan de waarnemingsomstandigheden en de waargenomen vegetatie. Voor het schatten van de vegetatievariabelen moeten de RT modellen echter geïnverteerd worden en deze inversie is gewoonlijk slecht gesteld (bijna singulier) en onderbepaald. Verschillende regularisatiemethodes zijn voorgesteld om stabiele en unieke oplossingen te vinden: modelkoppeling, gebruik multi-kijkhoek gegevens, gebruik *a priori* informatie, en de toepassing van ruimtelijke of temporele restricties.

Vanouds wordt de gemeten radiantie aan de top van de atmosfeer (TOA) voorbewerkt tot top van de vegetatie (TOC) reflecties. Correcties voor atmosferische effecten en, indien nodig, voor omgevingseffecten, directionele of topografische effecten worden meestal sequentieel en onafhankelijk toegepast. Fysisch zijn deze effecten echter met elkaar verbonden, en elke correctie introduceert fouten. Deze fouten planten zich voort naar de TOC reflectiegegevens, die gebruikt worden om het vegetatie RT model te inverteren. De prestatie van de TOC benadering wordt derhalve beperkt door de fouten geïntroduceerd in de gegevens gedurende de voorbewerkingsstappen.

Dit proefschrift stelt voor om deze fouten te minimaliseren door direct gemeten TOA radianties te gebruiken. In een dergelijke TOA benadering wordt het atmosferische RT model, dat normaliter geïnverteerd wordt om de atmosferische correctie uit te voeren, gekoppeld aan het vegetatie RT model. Het gekoppelde vegetatie-atmosfeer model wordt rechtstreeks geïnverteerd met behulp van de gemeten radiantie gegevens. Omgevings-, directionele en topografische effecten kunnen vervolgens opgenomen worden in het gekoppelde RT model. Dezelfde regularisatiemethodes als voor TOC benaderingen kunnen gebruikt worden om stabiele en unieke schattingen te verkrijgen. De TOA benadering is getest met behulp van vier case studies op basis van monotemporele gegevens.

A) De prestatie van de TOA aanpak is vergeleken met een TOC benadering voor drie fijnspar opstanden in de Tsjechische Republiek, met behulp van bijna-nadir 'Compact High Resolution Imaging Spectrometer' (CHRIS) gegevens. Het gekoppelde model bevatte directionele effecten van de vegetatie en simuleerde de

CHRIS radiantie gegevens met een nauwkeurigheid vergelijkbaar met de simulatie van de atmosferisch gecorrigeerde CHRIS gegevens door het vegetatiemodel. Lokale gevoeligheidsanalyses toonden aan dat de atmosferische parameters veel minder invloed hadden op de simulaties dan de vegetatieparameters, en dat de gevoeligheid voor parameters vergelijkbaar was voor zowel TOC en TOA benaderingen. De dimensionaliteit van het schattingsprobleem is geëvalueerd als zijnde 3 voor beide benaderingen. Bedekkingsgraad (Cv), fractie schors materiaal (fB), Cab, en droge stofgehalte van het blad (Cdm) werden geschat met behulp van opzoektabellen (LUT) met een vergelijkbare nauwkeurigheid voor beide benaderingen.

B) Regularisatie met multi-kijkhoek gegevens is getest voor de TOA benadering middels vier CHRIS datasets verkregen onder verschillende kijkhoeken voor dezelfde drie fijnspar opstanden als gebruikt in A). Het gekoppelde model leverde goede simulaties op voor alle hoeken. De dimensionaliteit nam toe van 3 naar 6 als alle vier hoeken gebruikt werden. Twee LUTs werden gebouwd voor elke opstand: een 4-variabele LUT met fB, Cv, Cdm en Cab, en een 7-variabele LUT waar de concentratie aan bruine pigmenten van het blad (Cs), de dissociatiefactor (D), en de boomvormfactor (Zeta) toegevoegd werden. De resultaten voldeden niet volledig aan de verwachting dat hoe meer hoeken gebruikt werden, des te nauwkeuriger de schatting werd. Hoewel het gebruik uitdagend blijft, hebben multi-kijkhoek gegevens meer potentieel op TOA niveau dan gegevens verkregen bij slechts één kijkhoek.

C) Een Bayesiaanse object-gebaseerde aanpak is ontwikkeld en getest op 'Airborne Prism Experiment' (APEX) radiantie gegevens, gemeten op sensorniveau, voor een agrarisch gebied in Zwitserland. Deze aanpak bestaat uit twee stappen. Allereerst werden tot zes variabelen geschat voor elk gewas met behulp van een Bayesiaans optimalisatie-algoritme onder gebruikmaking van *a priori* informatie. Vervolgens werd voor elk object een LUT gebouwd met slechts LAI en Cab als vrije variabelen, waardoor alle andere variabelen ruimtelijk de waarden opgelegd kregen die verkregen waren in de eerste stap. De Bayesiaanse object-gebaseerde benadering schatte LAI nauwkeuriger dan een LUT aanpak met een Bayesiaanse kostenfunctie. Deze case studie was gebaseerd op uitgebreide veldgegevens waardoor het definiëren van de objecten en *a priori* gegevens mogelijk was.

D) De Bayesiaanse object-gebaseerde aanpak voorgesteld in C) is toegepast op een gesimuleerd TOA Sentinel-2 beeld, dat het gebied rond Zürich in Zwitserland bestrijkt. Het gesimuleerde beeld werd samengesteld uit zeven APEX vluchtlijnen, hetgeen bestudering van alle ruimtelijke en spectrale eigenschappen van Sentinel-2 mogelijk maakte. Automatische multi-resolutie segmentatie en classificatie van vegetatie-objecten volgens vier niveaus van helderheid in het zichtbare domein maakte het mogelijk objecten en *a priori* gegevens te definiëren zonder veldgegevens, waardoor een succesvolle implementatie van de Bayesiaanse object-gebaseerde aanpak mogelijk was.

Het onderzoek beschreven in dit proefschrift draagt bij aan de verbetering van het gebruik van regularisatiemethoden in slecht gestelde RT modelinversies. Drie belangrijke gebieden werden geïdentificeerd voor verder onderzoek: 1) opname van effecten van omgeving en topografie in het gekoppelde model, 2) toevoeging van temporele restricties in de inversie, en 3) een betere integratie van observatie- en modelonzekerheden in de kostenfunctie. De hier voorgestelde TOA aanpak zal het gebruik van multi-kijkhoek, multitemporele en multi-sensor data vergemakkelijken, hetgeen leidt tot nauwkeurigere RS vegetatieproducten. Deze producten met hogere kwaliteit zullen vele vegetatie-gerelateerde toepassingen ondersteunen.

Résumé

La végétation est un composant important de la biosphère et joue un rôle crucial dans les échanges de carbone des écosystèmes terrestres. Les variables de végétation, comme l'index de surface foliaire (LAI) et la concentration des feuilles en chlorophylle (Cab), peuvent être suivies à l'échelle mondiale par télédétection. Deux catégories principales d'approches peuvent être utilisées pour estimer les variables de végétation a partir de données de télédétection : les approches empiriques et les approches physiques. Les approches physiques sont plus largement applicable parce qu'elles s'appuient sur des modèles de transfert radiatif (MTR) qui peuvent être adaptés aux conditions d'observation et à la végétation observée. Cependant, pour estimer les variables de végétation, un MTR doit être inversé, et cette inversion est en général un problème mal posé et sous-déterminé. Plusieurs méthodes de régularisation ont été proposées pour permettre de trouver des solutions stables et uniques : le couplage de modèles, l'utilisation de données multi-angulaires, l'utilisation de données a priori, ainsi que l'application de contraintes spatiales ou temporelles.

Traditionnellement, les données de radiance mesurées au sommet de l'atmosphère (SA) sont prétraitées en données de réflectance au sommet de la canopée (SC). Les corrections des effets atmosphériques, et, si besoin, des effets topographiques, directionnels et de proximité, sont effectuées séquentiellement et indépendamment. Cependant, physiquement, ces effets sont interdépendants et chaque correction introduit des erreurs. Ces erreurs se propagent dans les données de réflectance SC, qui sont utilisées pour inverser le MRT. La performance de l'approche SC est donc limitée par les erreurs introduites dans les données pendant le prétraitement.

Cette thèse propose de minimiser ces erreurs en utilisant directement les données de radiance SA. Dans une telle approche SA, le MTR de l'atmosphère, qui est normalement inversé pour accomplir la correction atmosphérique, est couplé au MTR de la canopée. Le MTR couplé canopée-atmosphère est inversé en utilisant directement les données de radiance SA. Les effets topographiques, directionnels et de proximité peuvent alors être inclus dans le MRT couplé. Les méthodes de régularisation utilisées pour l'approche SC peuvent être utilisées pour l'approche SA afin d'obtenir des estimations stables et uniques. L'approche SA a été testée dans quatre études de cas basées sur des données mono-temporelles.

A) La performance de l'approche SA a été comparée à celle de l'approche SC pour trois parcelles d'épicéa commun en République Tchèque, en utilisant des données proches du nadir, issues du « Compact High Resolution Imaging Spectrometer » (CHRIS). Le MRT couplé, qui incluait les effets directionnels de la canopée, a simulé les données de radiances de CHRIS avec une précision similaire à celle avec laquelle le MTR de canopée simulait les données CHRIS corrigées pour

les effets atmosphériques. Les analyses locales de sensitivité ont montré que les paramètres atmosphériques avaient beaucoup moins d'influence sur les simulations que les paramètres de la végétation, et que les profils de sensitivité de ces derniers étaient très similaires pour les approches SC et SA. La dimensionnalité du problème d'estimation a été évaluée à 3 pour les deux approches. Le couvert de la canopée (Cv), la fraction d'écorce (fB), Cab, et la concentration des feuilles en matière sèche (Cdm) ont été estimés, en utilisant des tables de correspondances (TC), avec une précision similaire pour les deux approches.

B) La régularisation par l'utilisation de données multi-angulaires a été testée pour l'approche SA avec quatre jeux de données angulaires CHRIS, pour les trois mêmes parcelles déjà utilisées en A). Le MTR couplé a fourni de bonnes simulations pour tous les angles et la dimensionnalité a augmenté de 3 à 6 en utilisant les quatre angles. Deux TC ont été construites pour chaque parcelle : une avec fB, Cv, Cdm, et Cab (4 variables), et une autre avec la concentration des feuilles en pigments bruns (Cs), le facteur de dissociation (D), et le facteur de forme des arbres (Zeta) en plus (7 variables). Les résultats ne correspondirent pas avec l'attente que la précision des estimations augmenterait avec le nombre d'angles utilisés. Bien que leur exploitation reste difficile, les données multi-angulaires ont plus de potentiel que les données mono-angulaires au niveau SA.

C) Une approche Bayésienne basée sur des objets (ABBO) a été développée et testée avec des données de radiance du capteur « Airborne Prism EXperiment » (APEX), pour une zone agricole en Suisse. L'approche consiste en deux étapes. Premièrement, un algorithme Bayésien d'optimisation a été utilisé pour estimer jusqu'à six variables pour chaque objet (champ agricole) en incluant des informations a priori. Deuxièmement, une TC a été construite pour chaque objet avec seulement LAI et Cab comme variables libres, contraignant ainsi toutes les autres variables à leurs valeurs obtenues dans la première étape. L'ABBO a estimé LAI plus précisément qu'une approche TC avec une fonction de coût Bayésienne. Cette étude de cas reposait sur de nombreuses données de terrain pour la définition des objets et des informations a priori.

D) L'ABBO proposée en C) a été appliquée à une scène Sentinel-2 de radiance SA simulée, couvrant la zone autour de Zurich en Suisse. La scène simulée a été mosaïquée a partir de sept images APEX, ce qui a permis d'inclure toutes les caractéristiques spatiales et spectrales de Sentinel-2. Une segmentation multi-résolution automatique et une classification des objets de végétation en quatre niveaux de brillance dans le domaine visible ont permis de définir les objet et les informations a priori sans données de terrain, ce qui a permis d'appliquer l'ABBO avec succès.

Les recherches menées dans le cadre de cette thèse contribuent à l'amélioration de l'utilisation des méthodes de régularisation dans les inversions mal-posées de MTR. Trois sujets principaux ont été identifiés pour des recherches futures : 1) l'inclusion

des effets topographiques et directionnels dans le MTR couplé, 2) l'addition de contraintes temporelles dans l'inversion, et 3) l'amélioration de l'inclusion des incertitudes des observations et du MTR dans la fonction de coût. L'approche SA proposée ici facilitera l'exploitation des données multi-angulaires, multi-temporelles et multi-capteurs, ce qui entrainera l'amélioration de la précision des produits sur la végétation issus de la télédétection. Ces produits de meilleure qualité amélioreront de nombreuses applications pour la végétation.

Acknowledgements

Looking back and reflecting on the past five years, many people contributed to the successful completion of my PhD thesis and to my PhD life. To all of them, I wish to express my gratitude.

First are of course my supervisors, whose constant support enabled me to complete this thesis. Without Michael's constancy in asking me whether I was interested in doing a PhD every time we met, I might not have seriously considered the question. Later on, I appreciated his scientific guidance and our technical and more general discussions. Jan always found time to discuss my numerous latest figures and options how to proceed. His realism and positivism were very good support throughout my thesis. Wout provided me with patient, precise and in-depth scientific and technical support. He not only helped me to see the forest through the trees, but also to look carefully at each tree. The three of you were complementary and made a very good supervision team.

My other co-authors also provided significant contributions to some thesis chapters. Alex's questions unblocked a number of issues when developing the Bayesian optimisation algorithm. He also provided results and paper writing advice. Joerg introduced me to the world of parametric models and Roberto's expertise on image segmentation was very helpful. Collaborating with you all was a pleasure and extended my knowledge.

I also wish to thank Allard, who set up Python scripts for me to run the SLC and MODTRAN models without learning Fortran. It would probably have taken me months to get the models running, and using freeware made my successive moves much easier. Lucie and Petr also provided me with valuable support for the Bily Kriz data.

Next are all my office-mates who contributed to a nice and enjoyable office life: Jochem and Gerd, followed by Richard, Erika and Michael S. at CGI, then Hans and Loïc in the Hora building in Ede, and finally Petra, Lucie, Lucia, Hossein, Peter, Titia and Parviz at RSL. The other PhD students and colleagues also participated of the nice atmosphere. Among all of these, I would especially like to thank Roberto, Daniela, Daniel, Noriko, Laure, Nandika, Kim, Rogier, Reik, and Irene.

For logistics and technical support, I am grateful to Raul and Anne for accommodation for ITC visits, to Lucie and Lucia for accommodation for RSL visits, and to Petra and Sara for last minute cat accommodation in Zurich.

For moral support, I would like to thank everyone who participated in wadlopen, biking, hiking, skiing, sledding, swimming, shopping, camping, concerts, cooking, jogging, crafting, BBQ, drinks and dinner events, including many of the people already mentioned above, and Marcela, Daniel W., Adrian, Mila, Corina, Marc,

Michael R., Bernardo, and especially Livia, who was my housemate for my last PhD months. A special mention also goes to my paranymphs Titia and Roberto.

From the French side, I am thankful to my friends from forestry, especially Morgan and Vincent, who followed my PhD adventures and reminded me about how life without/after a PhD can be like.

I am also very grateful to all the people who contributed to pass on the taste for studying onto me during my childhood, including my great grand-mother Suzanne, my grand-mothers Jeanine and Eva, as well as my primary school teachers Alain, Luc, and Arlette. This served me well throughout this thesis.

Finally, I wish to thank my family, my parents Bruno and Lucette, who always believed in me and pushed and supported me when needed, my sister Stéphanie and her partner Rémy with whom I always enjoyed time off of the PhD, although they did have to listen to some work-related chatter. Last but not least, I am also thankful to Aurélien, who followed me back to Wageningen when I started my PhD and who believed in me and did his best to support me throughout my research work, including even ArcGIS models.

Valérie
Montpellier, 13-02-2013

List of publications

Peer reviewed journals

Damm, A., Schaepman, M.E., **Laurent, V.C.E.**, Schickling, A., Guanter, L., & Rascher, U. (In review). Retrieval of sun-induced chlorophyll fluorescence from airborne spectroscopy data. *Geophysical Research Letters*

Damm, A., Schaepman, M.E., & **Laurent, V.C.E.** (In preparation). Assessment of sun-induced chlorophyll fluorescence and its functional relationship to plant pigments. *IEEE Geoscience and Remote Sensing Letters*

Laurent, V.C.E., Schaepman, M.E., Verhoef, W., Weyermann, J., & Chávez, R.O. (In review). Bayesian object-based estimation of LAI and chlorophyll from a simulated Sentinel-2 top-of-atmosphere radiance scene. *Remote Sensing of Environment*

Laurent, V.C.E., Verhoef, W., Damm, A., Schaepman, M.E., & Clevers, J.G.P.W. (In review). A Bayesian object-based approach for estimating vegetation biophysical and biochemical variables from at-sensor APEX data. *Remote Sensing of Environment*

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W., & Schaepman, M.E. (2011). Inversion of a coupled canopy-atmosphere model using multi-angular top-of-atmosphere radiance data: A forest case study. *Remote Sensing of Environment*, 115 (10), 2603-2612

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W., & Schaepman, M.E. (2011). Estimating forest variables from top-of-atmosphere radiance satellite measurements using coupled radiative transfer models. *Remote Sensing of Environment*, 115 (4), 1043-1052

Other scientific publications

Laurent, V.C.E., Verhoef, W., Schaepman, M.E., Damm, A., & Clevers, J.G.P.W. (2012). Mapping LAI and chlorophyll content from at-sensor APEX data using a Bayesian optimization of a coupled canopy-atmosphere model. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, 4 pp

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W., & Schaepman, M.E. (2011). Added value of multiangular measurements for estimating forest variables from the top of the atmosphere using coupled radiative transfer models. EARSeL 7th SIG-Imaging Spectroscopy Workshop, Edinburgh, 6 pp

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W., & Schaepman, M. (2010). Estimating forest parameters from top of atmosphere multi-angular radiance data using coupled radiative transfer models. Recent Advances in Quantitative Remote Sensing, Valencia, Spain, 6 pp

Laurent, V.C.E., Verhoef, W., Clevers, J.G.P.W., & Schaepman, M. (2010). Estimating forest parameters from top-of-atmosphere radiance measurements using coupled radiative transfer models. Hyperspectral Workshop 2010, Frascati, Italy, 8 pp

Laurent, V.C.E., Clevers, J.G.P.W., Schaepman, M., & Verhoef, W. (2009). Simulation of forest radiance at top-of-atmosphere level using coupled radiative transfer models. 6th EARSeL SIG IS Workshop, Tel Aviv, Israel, 6 pp

Laurent, V.C.E. (2007). Estimation of forest parameters using multi-angular CHRIS-PROBA data and the Li-Strahler GOMS model. Master thesis. Wageningen, the Netherlands: Laboratory of Geo-Information Science and Remote Sensing, Wageningen University, 89 pp

Short biography

Valérie Laurent was born in Cannes, France, on June 22nd, 1983. She attended the neighbourhood primary school, *École Saint-François*, in Grasse. She loved this school where each of the two classrooms hosted one teacher and about 25 pupils of three different levels. In Valerie's view, the best feature of this set-up was that the pupils were allowed to read books at their desks when they had finished their assignments. She then attended the *Collège Canteperdix* secondary school and the *Lycée Amiral de Grasse* high school in Grasse, where she obtained her *Baccalauréat scientifique* (scientific A-level) in 2000. Throughout her school years, Valérie especially enjoyed the mathematics classes. In her free time, she read a lot, practiced violin, and enjoyed outdoors activities with her local scouting group.

She entered the *classe préparatoire de Biologie, Chimie, Physique, et Sciences de la Terre* of the *Lycée Masséna* in Nice. This consisted in three years of very intensive studying of biology, physics, chemistry and mathematics. During these years, Valérie had to stop reading in order to fit studying and sleeping in her 24 hours days. She still liked mathematics best and also enjoyed vegetal biology. In 2003, she was admitted to the *Ecole Nationale du Génie Rural, des Eaux et des Forêts* (ENGREF) in Nancy, where she studied forestry. In 2005, she went as an Erasmus student to Wageningen University, the Netherlands. There she followed courses about geographic information systems and discovered remote sensing. Valérie then organized a joint Master thesis between ENGREF and Wageningen University on the "Estimation of forest parameters using multi-angular CHRIS/PROBA data and the Li-Strahler geometric-optical mutual shadowing model", which allowed her to obtain both a forest engineer master diploma from ENGREF and an MSc in Geo-Information Sciences from Wageningen University. She missed both graduation ceremonies due to being on the other side of the world for a remote sensing internship at the *Institut de Recherche pour le Developpement* in Nouméa, New Caledonia, France.

In 2008, she got the opportunity to undertake her PhD research at the Laboratory of Geo-information Science and Remote Sensing of Wageningen University. She worked on the coupling of canopy and atmosphere radiative transfer models to estimate vegetation variables directly from radiance data. She first worked with CHRIS multi-angular data, and, in 2012, with APEX imaging spectrometry data at the Remote Sensing Laboratories of the University of Zurich.

Valérie's research interests are related to remote sensing of vegetation, especially forests, with emphasis on physically-based approaches. This includes canopy and atmosphere radiative transfer modelling, scaling issues, multi-sensor studies, data assimilation and plant physiology.

PE&RC PhD Education Certificate

With the educational activities listed below, the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC), which comprises of a minimum total of 32 ECTS (= 22 weeks of activities).

Review of literature (6 ECTS)

- Multi-sensor image simulation using coupled radiative transfer models, validation and model inversion for deriving biophysical and - chemical (2008)

Writing of project proposal (4.5 ECTS)

- Multi-sensor image simulation and validation for driving biophysical and - chemical variables of the Earth's surface (2008)

Post-graduate courses (6.6 ECTS)

- IDL Programming; CGI (2008)
- 2nd Hyper-i-Net summer school; CGI (2008)
- ESA Training course on land remote sensing; ESA (2009)
- Summer course inverse modelling; UvA (2009)
- Bayesian statistics; PE&RC (2009)

Laboratory training and working visits (4.5 ECTS)

- Bayesian optimization algorithm; ITC, Enschede (2011)
- APEX, Sentinel-2, object-based approach; RSL, Zürich (2012)

Invited review of (unpublished) journal manuscript (2 ECTS)

- IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing: land cover classification using multi-angular data (2011)
- Remote Sensing: sentinel-2 spectral capabilities for LAI mapping (2012)
- Computer and Geoscience: 3D radiative transfer modelling of forest on slopes (2012)

Competence strengthening / skills courses (2.5 ECTS)

- PhD Competence assessment; WGS (2008)
- Mobilizing your – scientific – network; WGS (2008)
- Techniques for writing and presenting a scientific paper; CENTA (2008)

PE&RC Annual meetings, seminars and the PE&RC weekend (2.1 ECTS)

- PE&RC Weekend (2008)
- PE&RC Day: scaling from molecules to ecosystems (2008)
- PE&RC Day: accelerate the scientific progress (2009)
- PE&RC Day: selling science (2010)
- PE&RC Day: innovation for sustainability: what are the neighbours doing? (2010)

Discussion groups / local seminars / other scientific meetings (5.1 ECTS)

- Spatial Methods , SPAM (2008-2011)
- Forest and Conservation Ecology (2008-2010)
- Seminar on remote sensing; presentation; Amsterdam (2009)
- Statistics, mathematics and modelling in PE&RC; organizer (2009-2010)
- WIMEK/SENSE Symposium: modelling and observation earth compartments; Wageningen (2011)

International symposia, workshops and conferences (9 ECTS)

- Physense: the second workshop of the Nordic Network on Physiologically based remote sensing of forests; Helsinki (2008)
- Sensing a changing world, Wageningen (2008)
- Remote sensing of the environment, Wageningen (2009)
- EARSeL SIG-IS Workshop; poster; Tel-Aviv (2009)
- Remote sensing seminar; presentation; Amsterdam (2009)
- Hyperspectral workshop; presentation; Frascati (2010)
- Recent Advances in Quantitative Remote Sensing Symposium, Valencia,(2010)
- EARSeL SIG-IS Workshop; presentation; Edinburgh (2011)
- Swiss Geoscience Meeting; presentation; Zürich (2011)

Lecturing / supervision of practical's / tutorials; 7 days (2.1 ECTS)

- HydroLand: retrieval of land surface hydrological parameters; lecture (2010)
- Remote sensing basics; practical supervision (2010)

Supervision of a MSc student (3 ECTS)

- Temporal variability of probability density functions of the leaf area index for boreal and temperate forests