Centre for Geo-Information

Thesis Report GIRS-2007-03

RETRIEVING OPEN CANOPY VEGETATION PARAMETERS USING VEGETATION INDICES AND MODEL INVERSION

Detection of iron chlorosis in olive orchards using CASI data and radiative transfer modelling

J.A.E. van Gijsel

February 2007

WAGENINGEN UNIVERSITY

WAGENINGEN UR

To my grandparents

Retrieving open canopy vegetation parameters using vegetation indices and model inversion

Detection of iron chlorosis in olive orchards using CASI data and radiative transfer modelling

J.A.E. van Gijsel

Registration number 80 03 21 290 080

<u>Supervisors</u>:

Dr. P.J. Zarco Tejada
Instituto de Agricultura Sostenible, Consejo Superior de Investigaciones Científicas

Dr. J.G.P.W. Clevers
Wageningen University, Centre for Geo-information

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

February 2007
Wageningen, The Netherlands

Thesis code number: K075-704
Thesis Report: GIRS-2007-03
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Acknowledgements

First of all I would like to deeply thank Dr. Clevers for reviewing this thesis as fast he could in order to make it possible for me to try to graduate in March. The document has greatly improved quality from his comments, questions and suggested changes.

Secondly, my Spanish supervisor Dr. Zarco Tejada is sincerely thanked for the provision of the data, the discussion in Córdoba, the many answers to my questions by e-mail and the reviewing of this work – also at the very last moment. I truly appreciate that you have taken the time to answer my questions and discuss my doubts regardless of your busy schedule.

I would also like to thank Dr. Malenovský for providing the code of his ANMB index, discussing radiative transfer models, arranging access to the linux computer (that I helped to make crash) to allow me to compile and run the FLIGHT model, and for bringing Czech cookies. You know that you, the cookies and the fishes are always welcome whenever you wish to pass by ☺.

My family is warmly thanked as well as without them I would have never been here and without their support I would never have been able to do all the things I wanted to do and learn all the things that I know today. Special thanks go to my mother for sending me "positive energy".

Finally very many thanks go to Raúl for everything that you have done and especially for asking: "Why ..?".

Summary

This thesis deals with the retrieval of canopy variables from remotely sensed data with the main objective to detect iron chlorosis in olive orchards. Iron chlorosis leads to a reduced uptake and distribution of iron in plants, which in turn hampers the creation of chlorophyll and thus the photosynthesis and tree's production. Iron chlorosis can therefore be identified through monitoring for trees with a decreased chlorophyll a + b content (C_{AB}). As changes in C_{AB} have a pronounced effect on the leaf reflectance, we can monitor this C_{AB} by means of remote sensing. CASI images were available for the three study sites at two resolutions: $CASI_{SPATIAL}$ having 8 bands with 1 meter spatial resolution and $CASI_{SPECTRAL}$ having 72 bands with a spatial resolution of 4 meter.

Two retrieval strategies using the CASI observations have been studied: the inversion of a coupled radiative transfer model (the leaf model PROSPECT + the canopy reflectance model FLIGHT) by means of neural networks and the application of statistical relations between canopy reflectance and the chlorophyll content.

The leaf reflectance and transmittance were modelled with PROSPECT and resampled to match the CASI bands. These resampled leaf reflectance and transmittance were used together with variables describing the canopy (leaf area index (LAI), tree dimensions, leaf angle distribution (LAD) etc.), scene conditions (f_{COVER}, background/soil reflectance etc.) and viewing geometry as an input for the FLIGHT model to obtain the canopy reflectance for that scene. Variation was introduced in the inputs for PROSPECT (N, C_M and C_{AB}) and for FLIGHT (LAI, f_{COVER}). For the $CASI_{SPECTRAL}$ simulations, we have also created three classes of soil brightness to be able to study the effect of different background signals on the retrieval methods. The simulations where N and C_M have a constant value (with all other variables being non-constant) are referred to as PF-set 1 and the simulations with variation in N and C_M will be referred to as PF-set 2.

Two different approaches using neural networks (NN) were tested for the retrieval of canopy variables (CV) from $CASI_{SPATIAL}$. The first approach consisted of the 'classical' inversion where the NN were estimating the CV from the canopy reflectance. In the second approach, the NN were doing the reverse: the CV were obtained by inverting this NN estimating the canopy reflectance from a first guess for the CV. The estimated CV were updated using an optimalisation algorithm minimising the difference between the estimated reflectance and the observed reflectance. It was found that the classical NN has a better performance, as the second method had greater uncertainties due to the combination of the imperfect training and inversion of the NN. For the classical NN it was furthermore observed that the use of a-priori information improved the estimation of C_{AB} from PF-set 2, especially knowledge of N and C_M proved to be important. Fixing these two parameters (PF-set 1) resulted in an RMSE of 1.98 $\mu g/cm^2$. Reducing the number of input bands from 7 (band 6 was not used as it corresponds to an oxygen absorption band) to 2 greatly increased the error and indicates that indices using only these 2 bands may have strong limitations. However, the estimation of f_{COVER} and LAI in addition to C_{AB} from PF-set 1 resulted in an RMSE for C_{AB} of 2.57 $\mu g/cm^2$, for f_{COVER} it was 3.91% and for LAI 0.52 m^2/m^2. We concluded that theoretically it is well possible to retrieve multiple variables simultaneously using an inversion of a leaf+canopy model, provided that important crop characteristics (here N and C_M) are known or can be considered constant.

In parallel to the NN approach, we have developed relations between vegetation indices and the chlorophyll content. It was found to be most important to study the effect of other non-constant factors, such as LAI or the soil brightness on the behaviour of an index, as changes in these other factors may induce a trend that could be confused with a change in the main variable under study (the chlorophyll content in this study). The best performing indices in terms of RMSE were the approximated MTCI and MERIS red

edge indices and the developed CASI red edge index. These indices were found to have a fairly constant performance over all tested spatial resolutions (1, 4, 32 and 300 m).

Subsequently we have described the performance of the trained NN and the derived relations between VI and C_{AB} when these were applied to the real CASI imagery. The results were very poor for two of the studied plots and somewhat better for the third plot. The cause for these results was identified in a mismatch between the simulated reflectances and the measured reflectances similar to (Tan et al. 2005). Possible sources of these errors are incorrect parameterisation of the PROSPECT+FLIGHT models such as approximations made on the dimensions of the olive trees or an unrepresentative soil spectrum, limitations of the FLIGHT model at high spatial resolutions, or calibration artefacts in the CASI images. It has shown that careful validation of the results should be done after application of model based relations.

Finally we have concluded that iron chlorosis, observable as reduced chlorophyll contents, can in theory be identified using both vegetation indices and model inversion even at medium spatial resolutions. Non-constant factors influencing the retrieval (like the soil reflectance) should be studied before application of model-based retrieval methods. We recommend testing both retrieval methods with real imagery in an up-scaling study to see the limitations of these methods when confronted with changing influencing factors with changing spatial resolution.

Index

List of figures

List of tables

1 Introduction

The habitat of the olive trees (*Olea europea* L.) is concentrated between the 30° and 45° latitudes, in the climate zone specified as Mediterranean, having very warm and dry summers. The plant is belonging to the botanical family Oleaceae and it is the only comestible species of the 35 species belonging to the genus Olea.
Olives are estimated to have been cultivated for already 6000 years with its origin in the Middle East. The largest cultivation is nowadays taking place in Spain, with 60% of the production in the autonomous community of Andalucía (Barranco et al. 1999). In 2004, about 1.5 million hectares were used for the production of olives in this region (Junta de Andalucía - Consejería de agricultura y pesca 2005). The total acreage under agricultural usage in this year was slightly over 4 million hectares (Junta de Andalucía - Consejería de agricultura y pesca 2005), which indicates the importance of this crop for Andalucía.
Olive production can be hampered by stresses experienced by the trees during various or all development phases. These plant stresses can be defined as the ensemble of environmental (a-)biotic factors negatively influencing the performance of the plant's physiological processes, such as photosynthesis (Lambers et al. 1998). A plant may for instance suffer from stress because of limited water or nutrient availability, predation by pests, or competition for light with other vegetation.
Remote sensing (RS) techniques can be applied to monitor vegetation and to provide a fast way to detect stress in the vegetation that could lead to a production loss. These stresses are visualised by a change in pigment content of the leaves, resulting in a different spectral signature, especially in the green peak and along the red edge (Baranoski and Rokne 2005; Schlemmer et al. 2005; Stone et al. 2001; Zhao et al. 2005). Furthermore, the leaf area index (LAI) is strongly related to the photosynthetic assimilation and as a consequence, the crop yield. LAI is defined as the total surface area covered by green leaves over one unit area of ground surface. Frequent acquisition of LAI data will help to increase the accuracy of crop growth models' forecasts.
A third variable of importance besides the pigment content and LAI is f_{COVER}: the fraction of green vegetation covering a unit area of a horizontal soil, ranging from 0 (bare soil) to 1 (full coverage). It corresponds to the gap fraction in the nadir direction and it can be used to decouple vegetation and soil contribution in energy balance processes and especially evapotranspiration (Weiss et al. 2000).

Different sensors are available to acquire remotely sensed information. We can make a first distinction between airborne and satellite sensors. The airborne sensors usually have a higher spatial resolution, but the data acquisition is not taking place regularly. Secondly, the collection of airborne data is expensive. Satellite sensors normally provide data for one region with a fixed time interval. This does not allow alternative timing, which is a large disadvantage for areas with frequent cloud coverage.
At this moment, the design of a satellite sensor is limited to either having bands with a high spatial resolution or bands having a high spectral resolution. Airborne sensors can overcome this problem by flying at a lower altitude and thereby increasing the ground spatial resolution. Through the linking of airborne data with satellite data, we can attempt to get information with a high temporal, spectral and spatial resolution.

1.1 Problem definition

1.1.1 Estimation of biophysical parameters from remote sensing

For the estimation of important biophysical variables from remote sensing data, we can use empirical relationships such as spectral vegetation indices and a more physically-based approach which involves model inversion. Such an inversion can be achieved using look-up tables (LUT), neural networks (NN), expert systems and genetic algorithms (GA) (Myneni et al. 1995), simulated annealing (Kirkpatrick et al. 1983), tabu search (Glover 1986) and other optimisation techniques.

Spectral vegetation indices (VI) can be seen as optimised regressions based on an empirical relation between reflectance at specific wavelengths and the variable under study. This method includes very simple (single ratio) to complex (combined indices) approximations. Some indices – such as SAVI (Huete 1988), MSAVI (Qi et al. 1994) and TSAVI (Baret et al. 1989) – attempt to correct for the influence of the soil, whilst other indices (e.g. ARVI (Kaufman and Tanre 1992)) try to reduce atmospheric noise. The use of the combined index MCARI/OSAVI (Haboudane et al. 2002) has proved to be successful for the estimation of chlorophyll in olive orchards using 1-m ROSIS data (Zarco-Tejada et al. 2004b).
However, due to the different sensor configurations and the resulting differences in spectral resolutions, a vegetation index that was designed for one sensor cannot always be directly applied to another sensor. Additionally, the relations that were found are often strongly related to the atmospheric condition at the time of measuring, the crop variety under study and the characteristics of the area. This greatly reduces the overall applicability of such indices and this is why physically-based methods may be preferable in some situations, even though they require the acquisition of many input variables and they are far more computationally demanding.

A physically-based method involves inversion of a radiative transfer model (RTM) by adjusting input canopy variables to match the simulated hemispherical-directional reflectance factor (HDRF) with the measured reflectance, which equals the HCRF (hemispherical-conical reflectance factor) for most satellite and airborne sensors (Schaepman-Strub 2004). This difference (HDRF versus HCRF) introduces some small uncertainties besides the model approximation.
We can further find that a calculated HDRF does not have to give a unique solution to the inversion, but it may be the result of different combinations of input canopy variables. This ill-posedness (see definition by (Garabedian 1964) in (Combal et al. 2002)) is also a consequence of measurement and model inaccuracies. The use of prior information, in the form of ancillary measurements, information on canopy architecture or knowledge of the distribution of input canopy variables, can be used to restrict the number of solutions and to overcome the ill-posedness problem (Combal et al. 2002).

Many RTM are available at different scales. The 1-D PROSPECT leaf model (Jacquemoud and Baret 1990) and the SAIL(H) model (Verhoef 1984) working at canopy level have been frequently applied for many vegetation types (Goel and Thompson 1984b; Jacquemoud 1993; Weiss et al. 2000; Zarco-Tejada et al. 2004a; Zhang et al. 2005). Due to their simplicity, they have the advantage of being quick and easily applicable.

Nevertheless, most of the available techniques have been developed for relatively homogeneous and spatially continuous crops. Olive trees are grown in regular patterns with six to twelve meters in between of tree trunks and with an average LAI of 0.5 to 1.5 (Zarco-Tejada et al. 2004b). This results in an open canopy with a large influence of the soil on the obtained reflectance, especially at a low resolution when the pixels contain mixed information. Additionally, the system is complicated further by shadow effects.

Therefore, VI and the two models mentioned before will only produce reliable results at a high resolution where we can clearly separate the green vegetative parts from the (shadowed) soil. When the information inside one pixel is a combination of the three elements (soil, shadows and vegetation), we will need to make use of 3-D models which can include the structure of the vegetation. The simulated reflectance at different wavelengths and pixel sizes can be used to study the effect of up-scaling on the spectral vegetation indices (Zarco-Tejada et al. 2001).

Unfortunately, 3-D models require a lot of computations to solve the radiative transfer and its inversion. To reduce the amount of computations, we can make use of genetic algorithms (GA) and neural networks (NN) to perform the inversion.
A genetic algorithm is an optimisation technique that has originated from cellular automata. A random set of possible solutions is created. From this set, a subset is chosen based on a *fitness* criterion. The solutions in this subset can evolve as a clone to the next generation (elitist selection), cross-over (recombination of two solutions) or mutate. Subsequently the new generation is checked for having solved the defined problem upon which the process stops, or it will continue to evolve until a solution has been found (Wikipedia 2006b).
Artificial neural networks can be seen as a simplified representation of the biological nerve system, both in its architecture and in the way information is presented to the system (Pal and Mitra 1999). Information is sorted and passed on through connected nodes and the weights of the connections between nodes are continuously updated as the problem is being described with more details.

1.1.2 Plant stresses - iron chlorosis

Olive trees often suffer from a lack of iron, which is required by the plant to create chlorophyll molecules. An iron deficiency will therefore lead to a reduction of the chlorophyll content (C_{AB}) in new tissue. The first visible signs of this phenomenon called (iron) chlorosis can be found near the leaf veins and young leaves as the tissue is lighter green or even white in full absence of chlorophyll.
The main cause of an iron deficiency is a high soil pH, as the solubility of iron (hydr)oxids decreases with increasing pH (Janssen and Beusichem 2000). This is called lime induced chlorosis.
Under some circumstances, the chlorosis can be overcome by adding iron chelates to the soil or directly to the plant as a spray. This gives us a perfect opportunity to study the spectral behaviour of chlorophyll to relate it afterwards to the natural conditions and the associated plant stresses.

1.2 Research objectives

In this study we will investigate biochemical and biophysical properties (e.g. C_{AB}, LAI and f_{COVER}) of olive orchards at different spatial resolutions (between 1 and 300 meters).
Our first research objective is to determine the capabilities of hyperspectral indices and RTM for the detection of iron chlorosis, considering different scales.
Our second research objective is to investigate whether it is possible to correctly estimate C_{AB}, LAI and f_{COVER} in open canopies through inverse modelling (at different scales, for different sensors and whether it is possible to retrieve these biophysical properties simultaneously).
As open canopies such as olive orchards present a large spatial heterogeneity and the background (soil) plays an important role, the use of a three-dimensional radiative transfer model would be strongly recommended for the simulation of such a complex scene (Gastellu-Etchegorry et al. 1996; North 1996). Therefore, in this study we will

work with the PROSPECT leaf model coupled with the FLIGHT (Forest LIGHT interaction model) canopy reflectance model.

1.3 Research questions

The research objectives lead us to the following research questions
- Which spectral/spatial prerequisites need to be fulfilled to detect iron chlorosis?
- Is it possible to quantify the uncertainties in the estimated biophysical parameters with respect to the spatial resolution?
- Can vegetation indices be successfully used on reduced spatial resolution data to retrieve C_{AB}?
- Can we estimate chlorophyll accurately if we use 30 m pixel size imagery from olive orchards despite their heterogeneous architecture?
- Can we correctly retrieve multiple variables simultaneously considering the ill-posedness of the radiative transfer?

1.4 Setup of this report

This report has been organised as follows. Chapter two gives an introduction into the modelling of leaf and canopy reflectance. The structure of chapters three till six is visualised in Figure 1-1.

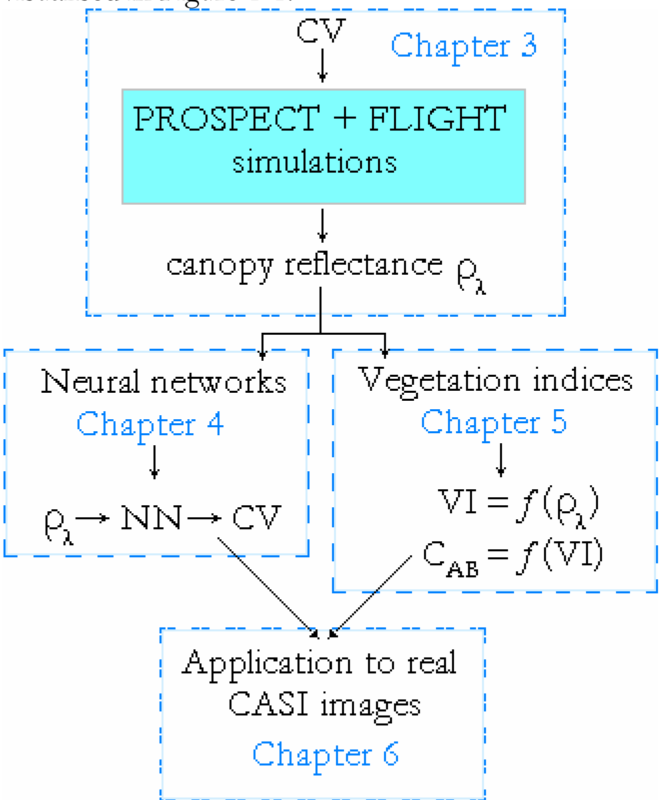

Figure 1-1: Overview of thesis structure (chapter 3 till 6)

As we want to retrieve biophysical and biochemical variables of the trees in the olive orchards, we will follow two approaches: (i) by means of modelling the radiative transfer or (ii) by statistically relating these properties to vegetation indices. In chapter three we describe the study area, the used datasets and the simulations done with the radiative transfer models PROSPECT and FLIGHT. Chapter four covers the application of neural networks for the retrieval of canopy biophysical/biochemical variables (CV) from canopy reflectance data based on these PROSPECT+FLIGHT simulations. Chapter five then presents the application of relations between vegetation indices and the modelled canopy reflectance to retrieve the chlorophyll content specifically. In chapter six, both trained retrieval methods (neural networks and vegetation indices) will be applied to real canopy reflectance data. Conclusions and recommendations are presented in the final chapter (7).

2 Modelling of leaf and canopy reflectance

The reflectance of a leaf or canopy can be modelled using various approaches. Commonly, a leaf (optical) model, describing the leaf's structural and biochemical components, is linked with a canopy model describing the position and inclination of the leaf, the viewing geometry and the background. When dealing with 3D models, the vertical and horizontal distributions of vegetated and other parts have to be incorporated as well. With some models it is also possible to incorporate woody elements such as the trunk and branches of different orders.

The canopy HDRF has been modelled in various ways. We can divide these approaches into three categories: empirical (mathematical relations that have no physical base), semi-empirical (having mathematical approximations for physical relations) and physical models (based on physical relations and a 'constructed' modelled canopy) (Boyd and Danson 2005).

We can also distinguish the different types of models: there are for instance geometric, optical, turbid medium, plate, ray tracing and hybrid models. A number of 3D models have been compared in the RAMI I and II exercises (Pinty et al. 2001; Pinty et al. 2004).

In section 2.1 we will describe the facts that are important for the individual leaf reflectance. Section 2.2 describes the relevant canopy, surface and geometric properties leading to the observed (total) reflectance signal. Spectral vegetation indices are briefly introduced in section 2.3. Section 2.4 deals with model inversion and the use of a-priori knowledge. The application of vegetation indices and model derived relations at lower spatial resolutions is discussed in the section Up-scaling (2.5). The chapter concludes with a summary of the sources of error in remote sensing and the error propagation when using this data to derive canopy properties (2.6).

2.1 Leaf optical properties

We will illustrate the effect of leaf characteristics on the leaf reflectance and transmittance using the PROSPECT leaf model in the following sections (2.1.1 through 2.1.4).

The PROSPECT leaf optical properties model is a one-dimensional radiative transfer model working at the leaf scale. It requires the input of the leaf structural parameter N, the chlorophyll a+b content (C_{AB}), the water content C_W and in later versions the dry matter content C_M as well.

The first version of the PROSPECT leaf model was based on measurements of plant tissue. The spectral samples were taken every 4 nm in the range 400 to 800 nm with a bandwidth of 1 nm. Additional samples were obtained for the range 800 to 2500 nm with steps of 17 nm and a bandwidth of 2 nm (Jacquemoud and Baret 1990). During the LOPEX93 experiment, complementary samples were taken with steps of 1 nm that were later averaged to blocks of 5 nm (Jacquemoud et al. 1996). The model was proven to be numerically invertible in 1993 (Jacquemoud 1993).

As an output, the model will give the leaf's hemispherical reflectance and transmittance for the spectral range of 400 to 2500 nm with steps of 5 nm and a bandwidth of 1 nm (Jacquemoud 2006) matching the specified leaf characteristics as described by the parameters N, C_{AB}, C_W and C_M. The effects of changing these characteristics will be described in the sections below.

2.1.1 The leaf mesophyll structure (structural parameter N)

As PROSPECT is based on Allen's plate model, it has incorporated the same plate (or elementary layer) structure to represent a leaf's internal structure. The thickness (number of layers) is affecting the radiative transfer inside the medium through scattering and absorption. A value of N close to 1 indicates a monocotyledon leaf type, whereas dicotyledon leaves are represented by values between 1.5 and 2.5. Values over 2.5 point at senescent leaves, as their internal structure is rather chaotic (Jacquemoud and Baret 1990).

As we can see in Figure 2-1 top left, an increase of the structural parameter will lead to an increased reflectance and decreased transmittance over the full spectrum. The effect is most pronounced in the NIR plateau.

2.1.2 The chlorophyll content (C_{AB})

Chlorophyll pigments are essential for plant photosynthesis. In most plants, two types of chlorophyll are active: chlorophyll a ($C_{55}H_{72}O_5N_4Mg$) and b ($C_{55}H_{70}O_6N_4Mg$) (Wikipedia 2006a).

Although other pigments such as carotenoids are present in the leaf, chlorophyll has the strongest influence on the radiative transfer as it occurs in higher concentrations. The two chlorophyll types can be distinguished by their absorption spectra, but are usually summed together as C_{AB}.

An increase in C_{AB} reduces both the reflectance and the transmittance in the spectral region between 400 and 750 nm (Figure 2-1 upper right). Beyond this wavelength, a change in C_{AB} does not influence the leaf reflectance or the leaf transmittance.

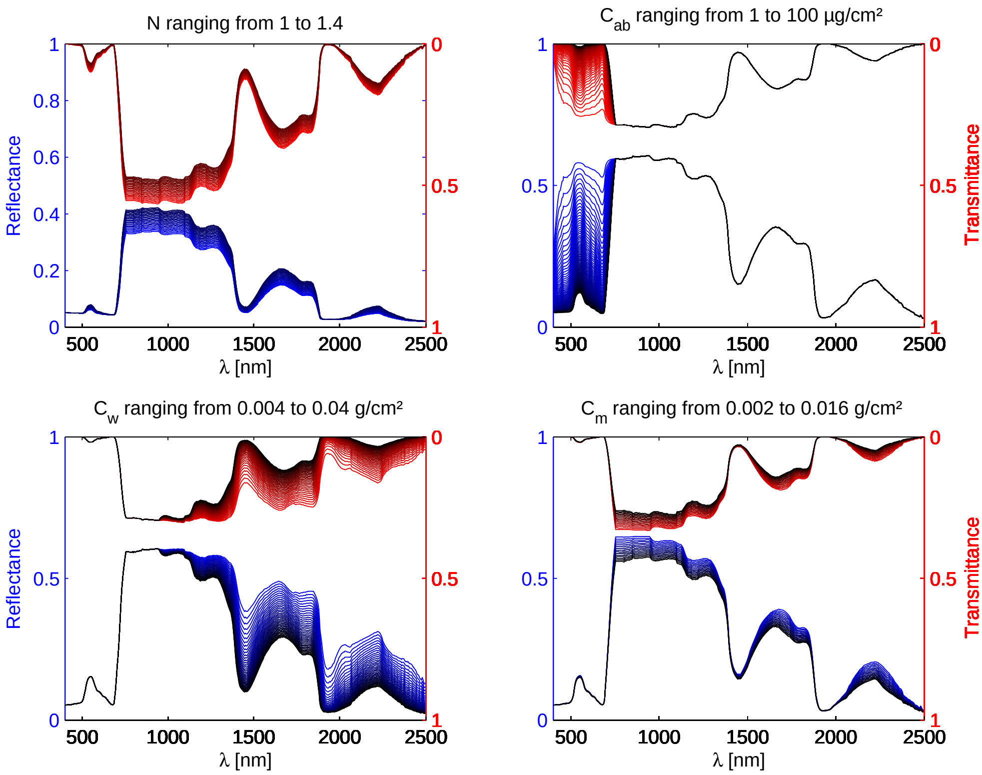

Figure 2-1: Sensitivity of the PROSPECT model to variations in the structural parameter N (upper left), chlorophyll content (upper right), water content (bottom left) and the dry matter content (bottom right). Blue (red) values refer to the reflectance (transmittance) matching the lower threshold specified in the title with increasing obscurity leading to the maximum threshold. If variables were fixed, the following values were used: N = 3, C_{AB} = 75 µg/cm², C_W = 0.025 g/cm² and C_M =0.01 g/cm². This figure has been created using simulations with PROSPECT.

2.1.3 *The water content (C_W)*

The water content is specified as the amount of water per square centimetre leaf [µg/cm^2]. An increase of the leaf water content induces a reduced transmittance and reflectance in the range 900-2500 nm (Figure 2-1 lower left).

2.1.4 *The dry matter content (C_M)*

With increasing dry matter content, we experience a relatively small reduction of the reflectance and transmittance. This effect is very small in the visible range, but most pronounced on the NIR plateau (Figure 2-1 lower right).

2.1.5 *Other leaf biochemicals*

Fourty et al. (Fourty et al. 1996) have studied the retrieval of biochemical components in dry leaves from measurements in the range 1300 to 2400 nm. Sugar, cellulose and hemi-cellulose contributed more to absorption in this range than lignin and protein contents. The first four compounds have similar absorption features as starch and are difficult to separate. Nitrogen in the proteins shows a strong absorption peak around 1900-2000 nm, but the proteins could not be successfully retrieved with the PROSPECT version used in that study. The highest retrieval accuracy was found for grouping all the components together, which has been implemented in later versions of the PROSPECT model.

2.1.6 *Leaf trichomes*

Leaf reflectance is influenced by the leaf's biochemical composition, the surface smoothness and the internal leaf structure. In addition, the leaf reflectance – and, as a consequence, vegetation indices – is also influenced by the presence and density of trichomes (leaf hairs). Levizou et al. (Levizou et al. 2005) found that most indices are strongly affected even at low trichome densities.

The removal of the trichomes of olive leaves led to a reduced reflectance in the VIS, especially in the abaxial (lower side) surface (up to 50% reduction), and an increased reflectance in the NIR (up to 4 %). For the adaxial (top) surface, which is mainly of influence for the reflectance measured by remote sensors, the differences caused by the trichomes were a maximal reduction of 16 percent in the VIS and almost no difference in the NIR.

Although relations found in literature are assumed to be applicable to leaves of the same species, some of the leaf structural components such as trichomes are altering with the leaf's growth stage, nutrient stress of the plant, leaf damaging by pests, prevalent light conditions plus the range in humidity and temperature of the plant's environment (Levizou et al. 2005).

The final effects of the occurring leaf structural differences within a tree may be difficult to assess. For instance, shadowing of a leaf leads to thinner leaves (Lambers et al. 1998; Larcher 2003) and may also lead to lower chlorophyll contents (Demarez et al. 1999), reducing an index such as the NDVI. However, this shadowing also causes the leaf to have a lower trichome density, resulting in an increase of the NDVI. The best performance for the VI under study was given by the red edge index, even at very high leaf hair densities (Levizou et al. 2005).

2.2 *Modelling of canopy reflectance*

It has been shown that the (hemispherical) reflectance of leaves is not sufficient to describe the canopy reflectance, as other factors such as background signals, wooden materials and shadows strongly influence the latter (Colwell 1974).

In the following subsections of paragraph 2.2, we will describe the factors that are influencing the canopy reflectance.

2.2.1 *Leaf area index (LAI)*

The leaf area index is usually defined as half of the total surface area of leaves per unit ground area. As it describes the present biomass, knowledge of this variable is essential for the modelling of plant/crop growth, calculation of the net primary production (Gower et al. 1999), to monitor the carbon balance (Aragão et al. 2005) amongst other processes.

The LAI can be measured directly through area harvest, application of allometric equations to stand diameter data and the collection of fallen leaves in autumn (Gower et al. 1999). Alternatively, it can be measured indirectly using non-destructive instruments measuring light transmittance in the field. Finally, it can also be estimated by inferring a relation with remotely sensed data (Berterretche et al. 2005; Lee et al. 2004).

Applying general allometric equations may lead to large errors, because the coefficients of the equations should be adjusted for biotic and abiotic factors concerning the area under study (Gower et al. 1999).

2.2.2 *Leaf angle distribution (LAD)*

The leaf angle distribution describes the orientation of the leaf surface with respect to the normal. A number of distributions have been described in literature and can be approximated as functions as shown in Figure 2-2. Notice that different approximations exist depending on the author (shown as full versus dotted lines).

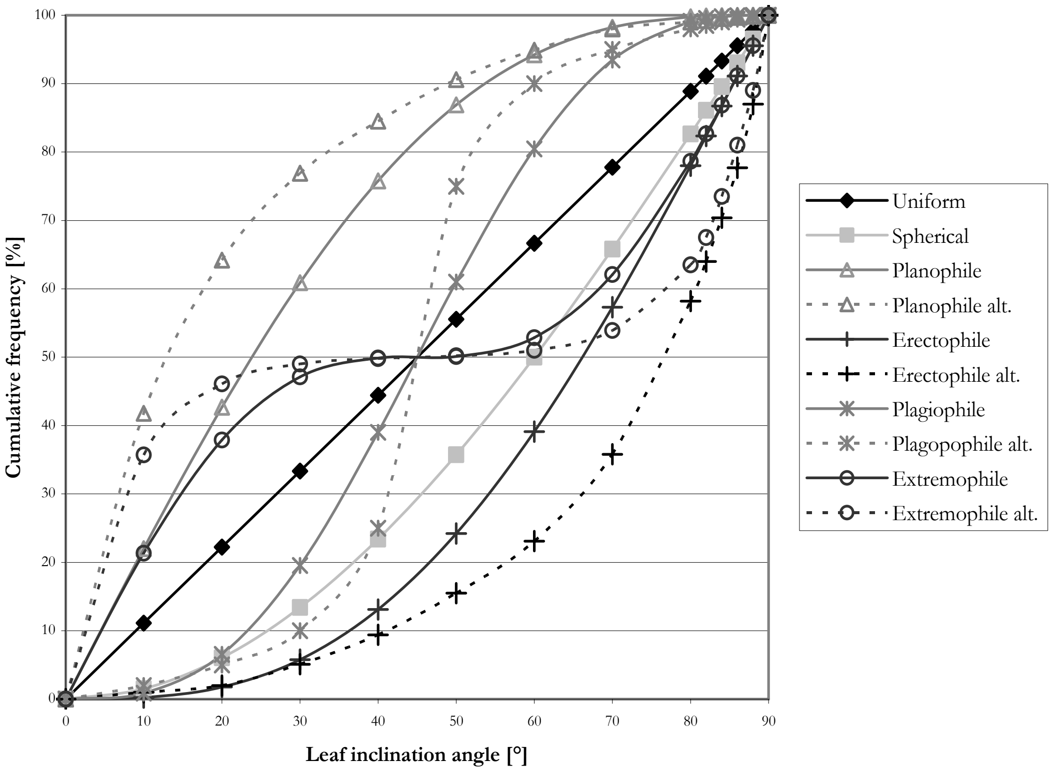

Figure 2-2: Leaf angle distribution functions after (King 1999) (full lines) and (Smith et al. 1977) (dotted lines; noted as alternative: 'alt').

Goel and Thompson (Goel and Thompson 1984c) showed that the average leaf inclination angle (ALA) could be derived from crop reflectance data through inversion of the SAIL model for a soybean canopy (f_{COVER} = 100%), given that the leaf reflectance and transmittance, the soil reflectance and the fraction of diffuse light are known.

2.2.3 *Canopy structure*

A canopy can be split into photosynthetically active (green) and non-photosynthetically active parts. Amongst the second class we can find flowers, fruits and woody elements (e.g. stem, branches and twigs). It is important to consider that the composition will change throughout a growing season; deciduous vegetation will have no photosynthetically active elements in winter periods and flowering as well as fruiting is limited to specific periods. In addition, care should be taken when using different viewing geometries (see section 2.2.5). Finally, the composition or structure of vertically heterogeneous vegetations should be modelled with sufficient detail to ensure that the HDRF corresponds to the RS observations.

2.2.3.1 *Canopy hotspot*

The canopy hotspot is occurring when the phase angle (see section 2.2.5) is close to zero (Camacho-De Coca et al. 2004). In other words, this is when the sun – or in general the light source – and observer are aligned so that no shadow is seen. Information regarding the canopy structure can be derived by studying the changes of this phenomenon resulting from changes in viewing and illumination directions (Jupp and Strahler 1991).
Through simulations was found that canopies with large leaves might show an increased HDRF of 20 to 40% over a wide range of directions due to the hot-spot (Andrieu et al. 1997).
In many radiative transfer models (for example SAIL) a 'hotspot parameter' [m/m] has been implemented to correct for the increase of the canopy reflectance in the backward direction. A frequently used estimation of this hotspot parameter is the ratio between the average leaf length and the canopy height (Jacquemoud et al. 2000; Kuusk 1995).

2.2.3.2 *Effects of shadowing*

The arrangement of leaves inside a canopy is optimised to capture incoming radiation as efficiently as possible. As a plant grows more leaves, the top leaf layer will intercept most of the light and older leaves, now in less illuminated positions, have to adapt to the changed light conditions. This is done by altering the size and number of the chloroplasts inside the leaf and the thickness of the shaded leaf itself (Lambers et al. 1998).
The chlorophyll per dry mass is higher for shaded leaves compared to sunlit leaves (Larcher 2003). The ratio between chlorophyll a and b is normally lowered with decreasing sunlight (Lambers et al. 1998). Shaded leaves tend to be thinner due to a reduced palisade parenchyma (Lambers et al. 1998). As shown in (Gausman 1984), the reflectances of a sunlit leaf and a shaded leaf are therefore distinct.
The changes of chlorophyll content *per area* for different illumination conditions are not stable considering the available literature. According to (Lambers et al. 1998; Larcher 2003), no difference should occur for most plants, although shaded leaves of some species do show an increased chlorophyll content per area (Lambers et al. 1998). This was also described by (Anthony et al. 2002; Barták et al. 1999), whereas the opposite was found by (Demarez et al. 1999; Levizou et al. 2005). Most likely, this variation is caused by a species-specific behaviour and because of the difficulty of converting the total chlorophyll content per unit mass to a chlorophyll content per unit area.
In addition to the leaf biochemical changes, shadowing may also result in changed background reflectances, as the biological activity and the soil moisture content change. Besides the leaf angle distribution it is important to consider the shading effects of the branching architecture of a canopy as well (Read et al. 2006).

2.2.4 *Fractional coverage (f_{COVER})*

The fractional coverage was defined as the fraction of vegetation over a unit area $\left(\frac{A_{VEGETATED}}{A_{TOTAL}}\right)$. When given as a percentage, the fraction should be multiplied by 100%. An f_{COVER} of 1 or 100% indicates that the soil is not directly visible, whereas an f_{COVER} of 0 or 0% is representative of a bare soil or other type of surface.

2.2.4.1 *Vegetation spatial distribution*

Besides the f_{COVER}, we must also consider the spatial distribution of the vegetation inside the scene: we can distinguish regular distributions and random distributions. Regular distributions (vegetation planted in rows or on a regular grid) may show a strong variation in the reflectance signal with changing viewing geometry due to shadowing effects and the (in)visibility of the soil areas intersecting the rows of aligned plants (Boyd and Danson 2005; Goel and Grier 1987). Many models have been adjusted to work with crops planted in rows, such as the Walthall, SAIL and Kuusk models.

2.2.5 *Sun-object-sensor geometry*

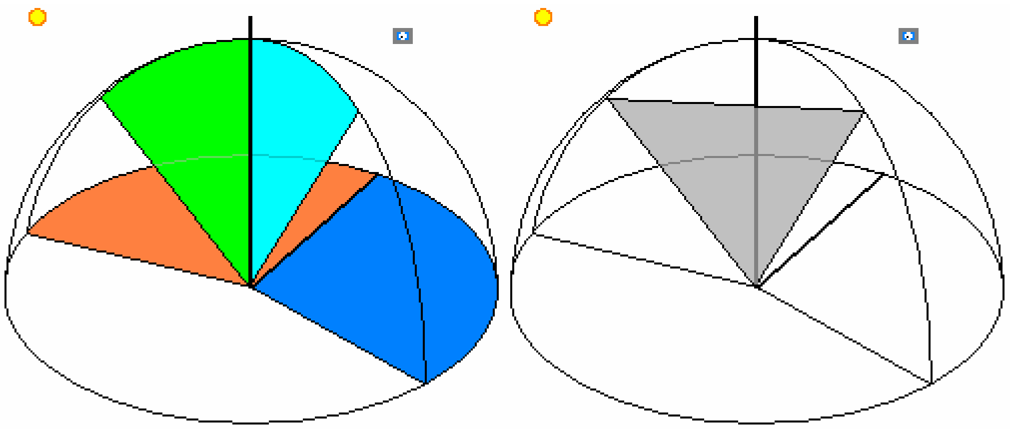

Figure 2-3: Sun-object-sensor geometry with zenith and azimuth angles (left) and with phase angle (right). The thick black line in the horizontal surface indicates the position of the North.

The *phase angle* is the angle between the illuminator-object and object-observer (see Figure 2-3, right part). When it is equal or close to 0, we will observe the hot spot effect (see section 2.2.3.1). The other angles relevant for the viewing geometry are described below.

2.2.5.1 *Solar zenith θ_S*

The solar zenith angle is defined as the angle between the normal and the position of the sun seen in a 2-D vertical slice (see Figure 2-4 and green surface in Figure 2-3, left part). A solar zenith angle of 0° corresponds with the nadir position.

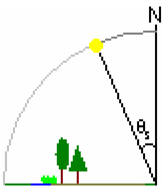

Figure 2-4: Solar zenith angle with respect to the normal N

2.2.5.2 *Solar azimuth angle Φ_S*

The solar azimuth angle is defined as the angle between the North (0°) and the position of the sun seen in a 2-D horizontal slice. A solar azimuth angle of 90° corresponds with the sun being positioned in the East. See coral surface in the left part of Figure 2-3.

2.2.5.3 *View zenith θ_V*

In a similar fashion as the solar zenith, the view zenith is defined as the angle between the normal and the position of the observer. See the light blue surface in Figure 2-3 (left).

2.2.5.4 *View azimuth Φ_V*

The view azimuth corresponds to the angle between the observer and the north on a horizontal plane, as displayed in blue in Figure 2-3 (left part).

2.2.6 *Aerosol optical depth*

The aerosol optical depth (AOD) is defined as the integrated extinction coefficient over a column of air reaching from the surface to the top of the atmosphere (Veefkind 1999) or: $AOD = -\ln\left(\frac{I}{I_0}\right)$, where I is the light intensity and I_0 is the initial intensity. The higher the AOD, the greater quantity of light is reduced over a given distance, hence, the less direct and more diffuse light. The AOD is dependent on the measured wavelength, as the spectral properties of the aerosols are changing with particle size and shape. Note that the particle size may increase due to adhesion of water with increasing relative humidity.

2.2.7 *Diffuse fraction*

The diffuse fraction is derived from the AOD, solar zenith angle and the wavelength.

$$\tau_{RAYLEIGH} = \frac{\frac{pressure}{p_0} \times \left(c_1 \times \lambda^{-4} \times \left(1 + c_2 \times \lambda^{-2} + c_3 \times \lambda^{-4}\right)\right)}{\cos(\theta_s)}$$

, where $pressure/p_0$ was set to 1, c_1 = 0.008569, c_2 = 0.0113 and c_3 = 0.00013 (North 1996).

$$\tau_{AEROSOL} = \frac{f_{DIFF}^{EFF} \times \frac{AOD}{c_4^K} \times \lambda^K}{\cos(\theta_s)}$$

, where f_{DIFF}^{EFF} equals the effective diffuse fraction (set to 0.55), K = the approximated angstrom coefficient for continental aerosols (-1.25) and c_4 = 0.55.

$$dir = e^{-(\tau_{RAYLEIGH} + \tau_{AEROSOL})}$$

$$dif = f_{RAYLEIGH} \times \left(1 - e^{-\tau_{RAYLEIGH}}\right) + f_{AEROSOL} \times \left(1 - e^{-\tau_{AEROSOL}}\right)$$

, where $f_{RAYLEIGH}$ *and* $f_{AEROSOL}$ equal the fractions of Rayleigh and aerosol scattered light to reach the ground surface (0.5 and 0.75 respectively).

The diffuse fraction is then calculated as $f_{DIF} = \frac{dif}{dir + dif}$.

Naturally, the diffuse fraction increases with an increasing AOD. In Figure 2-5 we can observe that the fraction of diffuse radiation f_{DIF} is decreasing with increasing wavelength and decreasing solar zenith. The total reflectance ρ_{TOT} can be derived from the diffuse (ρ_{DIF}) and direct (ρ_{DIR}) reflectance calculated by: $\rho_{TOT} = f_{DIF} \times \rho_{DIF} + (1 - f_{DIF}) \times \rho_{DIR}$.

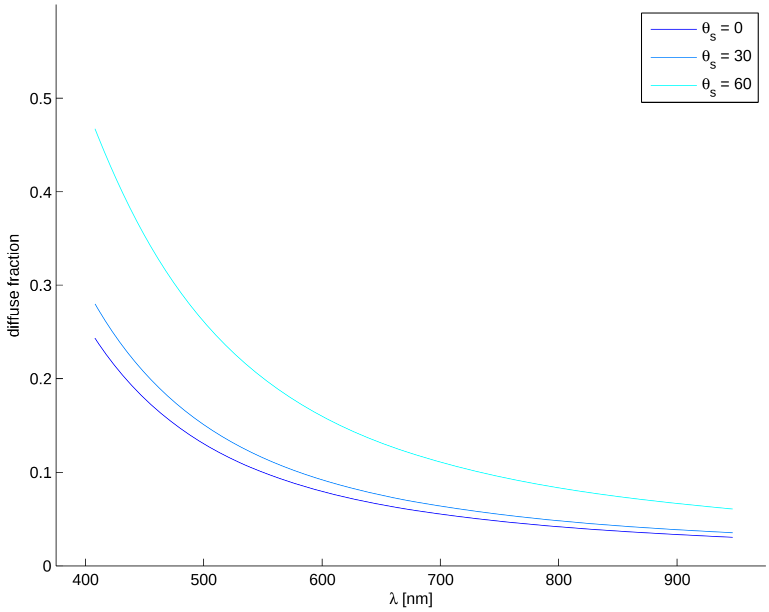

Figure 2-5: Change in diffuse fraction with wavelength λ and solar zenith θ_S. Values calculated with AOD equal to 0.12. This figure was made with results from the adjusted version of the Flight model (see Appendix I).

2.2.8 Soil background

In addition to the vegetation signal, the measured reflectance is also influenced by the soil composition of the area under study. The reflectance signal of a given soil type is determined by the mineral composition of the surface of this soil, as well as the organic matter and soil moisture contents.

The relevance of this signal depends on the f_{COVER}, the LAI, the presence of leaf litter, the terrain surface and the soil's roughness and brightness. Unless the area under study is very small and level, the soil signal cannot be considered constant due to the spatial variation of the soil surface properties.

The influence of the soil on the measured reflectance has been described by Rondeaux et al. in relation to the canopy leaf area index as: $1\text{-}e^{-K*LAI}$ (Rondeaux et al. 1996), where the factor K is an extinction coefficient related to the LAD (section 2.2.2).

2.2.8.1 Soil roughness index

The soil roughness index is used to indicate the roughness of the soil surface. A value of 0 (smooth) indicates that the surface can be represented by a Lambertian surface, whereas a value of 1 (rough, with mean slope of 60°) leads an approximated HDRF based upon a look-up table (North 1996).

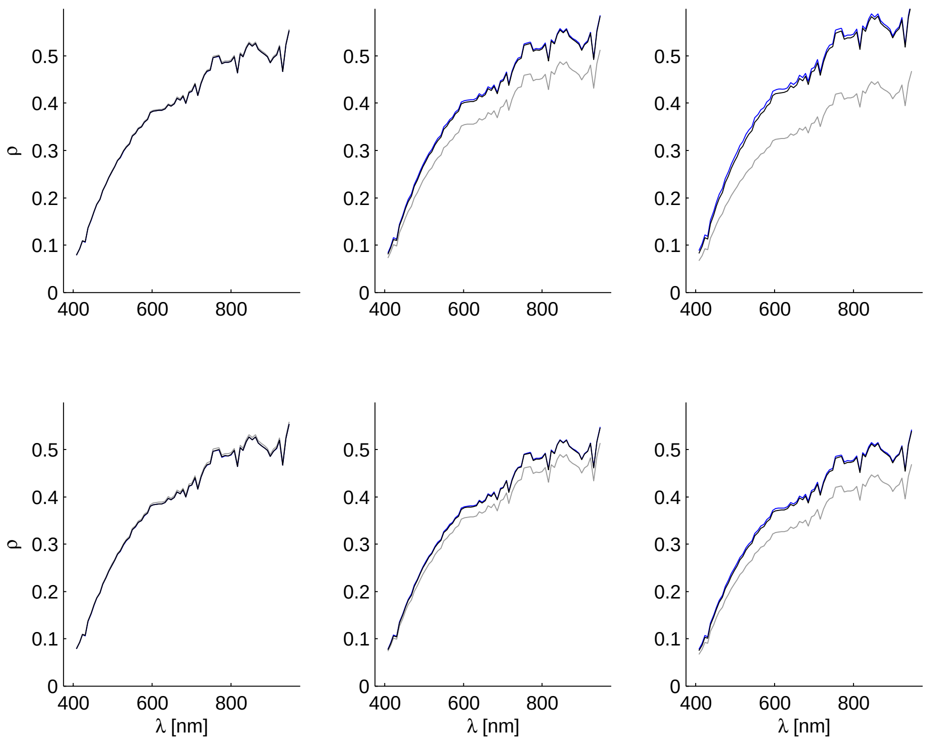

Figure 2-6: Reflectance of a bright soil at nadir for diffuse radiation (grey lines), direct radiation (blue lines) and total radiation (black lines) for simulations at different solar zenith angles (columns: top: $\theta_S = 0°$, down: $\theta_S = 30°$) and soil roughness index values (left to right: 0 to 0.8 with steps of 0.4). This figure was made with results from the adjusted version of the FLIGHT model (see Appendix I). Note that the oxygen absorption feature (762 nm) has been omitted in the figure (band replaced with interpolated point).

Figure 2-6 shows that the differences between the diffuse and direct radiation increase with increasing wavelength. However, the magnitude of the role that the diffuse radiation is playing decreases with the wavelength (see Figure 2-5) and therefore the total reflectance (Figure 2-6 in black) seems to follow the direct reflectance trend (in blue).
Secondly, the differences are enlarged by increasing soil roughness. Furthermore, for the tested configurations, the diffuse radiation always decreased with increasing soil roughness, whereas the direct radiation showed an increase for the simulations with solar zenith equal to 0 changing into a decrease for large zenith angles (not shown). Our third observation is that a lower sun position (increase of solar zenith) leads to a decreased direct reflectance and a minor increased diffuse reflectance, which is also clear from Figure 2-5.

2.2.8.2 Soil brightness

Huete et al. (Huete et al. 1985) studied a number of indices related to canopy greenness and found that the soil brightness strongly influenced these indices. They suggest that the canopy and soil spectra interact in a non-linear way that is partly correlated.
The greenness vegetation index (GVI) has been shown to be sensitive to the soil type and the soil moisture status – even if there was one unique soil type. The perpendicular vegetation index (PVI) is also showing a response to variations in the soil moisture content. Ratio indices such as NIR/RED or the NDVI were shown to have higher greenness values for darker backgrounds when f_{COVER} was kept constant. In contrast, orthogonal indices showed the opposite response to soil brightness.

Reflectance in the red region remained constant (dark soil) or decreased (lighter soils) with increasing vegetation cover. In the near-infrared region, reflectance increased practically linearly with f_{COVER} with steeper slopes for darker soils and a steeper slope for the last 10% to full coverage. They conclude that the soil background cannot be normalised to a constant ratio, because this only removes the soil spectral influence, not the soil brightness influence.

2.3 Spectral (vegetation) indices

Many attempts to relate remotely sensed data to ground observed variables can be found in literature. As most of these statistical relations have been designed for vegetation, they are referred to as vegetation indices (VI). Commonly, these VI are described through a regression between the leaf or canopy reflectance in one or more bands/wavelengths and the variable under study. The VI can be grouped according to the spectral regions used (VIS, VIS+NIR, VIS+SWIR, red-edge, broad bands versus narrow bands, etc.), according to their function (accounting for soil background and/or atmospheric effects, sensitivity to chlorophyll or LAI etc.) or according to their mathematical structure (single ratio, derivatives, etc.).

In (Grossman et al. 1996) it was shown that the band selection in a study relating leaf biochemical components to leaf reflectance through stepwise multivariate regression could not be linked to the absorption features of these components. They found that the band selection was also very sensitive to the samples included in the dataset, the unit of the expressed variable and dataset itself (i.e. selection may differ for different dates and/or locations). Therefore, statistical relations found between biochemical properties and corresponding reflectance spectra may not always be applicable to other sites and/or years (Grossman et al. 1996; Zarco-Tejada et al. 2001). This is why the usage of, or the combination with, models is preferred when looking for general relations.

2.4 Model inversion

The reflectance of a given canopy is considered to be a result of this canopy's biochemical, biophysical and structural properties. These relations have been modelled in numerous canopy reflectance models. Many of these – for instance SAIL (Scattering by arbitrarily inclined leaves) (Goel and Thompson 1984b), PROSPECT (Jacquemoud 1993), Suits' model (Goel and Strebel 1983; Goel et al. 1984; Goel and Thompson 1984a) and INFORM (Invertible Forest Reflectance Model) (Schlerf and Atzberger 2006) – have been proven to be mathematically invertible to obtain various canopy variables using the canopy reflectance in a number of wavelengths as an input.

The model inversion is often 'approximated' by creating a so-called look-up table of reflectances corresponding to combinations of biophysical, biochemical and structural parameters for given illumination and background conditions. Rather than to mathematically invert the model or to minimise the error through model iterations (sequential runs), the measured reflectances are looked up in the table and the matching (originally input) parameters are retrieved. This method can be very time efficient, but depending on the data density and number of requested parameters, no (e.g. because of low sampling of the input variable space) or multiple combinations may be found. The second problem is called ill-posedness, which can be overcome by introducing ancillary data (see section 2.4.1). The first problem (no spectrum found in the look-up table exactly matching the measured spectrum) can for instance be worked around by the application of neural networks trained with the same look-up table. The NN tries to learn the general rules relating canopy variables with the corresponding spectra and will be capable of 'interpolating' in between of the known data.

Many other techniques exist (see section 1.1.1), but in general an optimisation technique is used to minimise a merit function, which usually consists of the difference (for instance as RMSE) between measured and modelled reflectance and/or transmittance.

2.4.1 Ancillary data

Ancillary – or a-priori – data can help to reduce the number of possible solutions caused by the ill-posedness in the inversion process. It can consist of a restricted first guess (Combal et al. 2002) or the use of an additional variable, such as f_{COVER}. This additional information can be determined in various ways. For instance for f_{COVER}, some available methods are fieldwork (creation of a vegetation map that will be intersected with the remotely sensed grid), spectral unmixing (into vegetation/non vegetated signals), pattern recognition software and neural networks.

2.5 Up-scaling

Up-scaling can be defined as the process which ensures that the values of a given biophysical/biochemical property at a given spatial resolution are consistent (equal to the arithmetic average) with the values of that same property independently derived from a higher spatial resolution (Tian et al. 2003).

Four different up-scaling methods have been identified to derive biophysical or biochemical properties from measured canopy reflectance (Zarco-Tejada et al. 2001):

- The creation and application of statistical relations (vegetation indices) linking these properties directly to the canopy reflectance.
- The application of vegetation indices relating the leaf reflectance to the biophysical or biochemical property at the canopy level.
- The scaling-up of vegetation indices derived at the leaf-level through radiative transfer modelling.
- The inversion of a linked canopy+leaf RTM to derive the canopy property from the measured canopy reflectance.

The first two methods are only applicable to a given study site, for the crop under study and viewing geometry.

The optical indices can also be scaled up by means of canopy reflectance models. These scaled-up leaf-level relations can be directly applied to the measured canopy reflectance when the canopy structure and viewing geometry have been accounted for in the modelling (Zarco-Tejada et al. 2001).

With the fourth methodology, there is no need to relate the biophysical or biochemical properties with the reflectance at the leaf level. However, only properties that are used in the leaf model can be derived.

Notice that some leaf properties cannot be up-scaled directly to canopy-aggregate properties and require a transition via other leaf or canopy properties (Read et al. 2006).

2.6 Measurement error in remote sensing and derived products

Remote sensing data has been frequently used to describe ground conditions through statistical relations. In many cases, this is in the form of a regression function. Most of these regressions are of the type $y = ax + b + \varepsilon$, where *a* and *b* are estimated coefficients, *x* is the remotely sensed data, *y* is the variable to be estimated and ε is the residual error. Although in statistics it is normal that *y* is the dependent variable and *x* the independent, in remote sensing studies *x* is in fact the dependent variable as the measured reflectance is (partially) a *consequence* of the ground conditions.

Knowing that sensors do not provide perfect measurements, the *x* presented to the regression model contains measurement errors. This leads to a biased (underestimated) factor *a* ($\hat{a} = a/(1+(\sigma_e^2/\sigma_x^2))$), where σ_e^2 = variance from measurement errors in *x* and σ_x^2

= variance originating from real variation in *x*). When *a* is a positive factor, this leads to overestimates for values greater than the mean and to underestimates for values below the mean.
Errors in ground measurements originate from two sources. The first consists of errors in the measuring device that are normally constant and can be accounted or corrected for. The second error is related to the sampling density and surface, which adds more uncertainty when pixel sizes increase (e.g. how to obtain LAI measurements over a region of 30 by 30m? The error will depend on the number of samples taken inside that area).
Errors in remotely sensed variables have seven sources (Curran and Hay 1986):

- irradiance variation (low and high frequency noise)
- sensor calibration error (primary standard, calibration of and from lab. integrating sphere)
- sensor radiometric resolution (signal and noise, spectral sensitivity)
- sensor drift (signal distortion with time)
- signal digitization (conversion of the signal to digital numbers)
- atmospheric attenuation (very variable error)
- atmospheric path radiance (very variable error)

To this, we can add the uncertainty added because of the point spread function (PSF). Imperfections of the imaging sensor and diffraction of the light cause a signal coming from a point to be somewhat blurred out. This response can be mathematically described by the PSF. It has been estimated that only about half of the radiance value for each pixel can be attributed to the ground area that it is representing (Curran and Hay 1986).
Errors in the correlation of RS variables with ground-based variables can also be attributed to errors in geo-location causing mis-registration and differences caused by time differences between RS image acquisition and ground-based measurements. The latter error will depend on the variation with time of the studied variable (leaf moisture content will fluctuate more than leaf area index, thus requiring a more restricted time frame for both observation moments).

3 Data and models

This chapter describes the study area and datasets (section 3.1), followed by the data pre-processing in section 3.2, the description of the FLIGHT model in section 3.3 and the preparation of the FLIGHT simulations in section 3.4.

3.1 Study area and datasets

Four study sites with olive trees have been established in the Spanish province of Seville and one in Jaén to study leaf biochemistry in open canopy crops. In this study we have worked with the data of three sites, namely Cañaveralejo (referred to as CLO1), Tobalico (referred to as CLO3), and Las Aguilillas (CLO4), as one of the study sites was missed by the plane during the date of image acquisition. The following varieties are planted in the selected test sites:
1) Hojiblanca intensivo (used for the production of olive oil) at CLO1
2) Manzanilla (used for table olives) at site CLO3
3) Arbequina (used for the production of a sweet olive oil) at CLO4.
CLO1 and CLO4 were irrigated whereas CLO3 was rain-fed.

In order to create different levels of chlorophyll content, the study sites have received four different treatments with iron chelates:

a) Control
b) Received one application of 2.3% iron chelate (Q Fe 2.3 ortho-ortho).
c) Received one application of 4.8% iron chelate (Q Fe 4.8).
d) Received one application of Q Fe 4.8 with a 33% higher dose than of group c.
e) Received an application of Q Fe 4.8 with a double dose compared to group c.

The chlorophyll content is expected to increase with application of iron chelates as this is reducing the chlorosis. In CLO1 and CLO4, there are 100 treated trees of which 50 have been selected for evaluation (10 per group). In study site CLO3, we have 40 trees (8 per group).

3.1.1 Measurement of chlorophyll content

Chlorophyll concentration data has been collected for the 3 study sites in 2003. 12 leaves per tree were sampled around the crown, placed in bags and stored at -23 °C prior to analysis. Two 2.3 cm circle samples were cut out of each leaf. One circle was ground into liquid N_2, weighed, and placed in a 15 ml centrifuge tube. The second circle was weighed, oven-dried at 80 °C for 24 h, and re-weighed. Ten milliliters of N,N-dimethylformamide (Spectralanalyzed grade, Fisher) was added to the tube, and 3 ml of supernatant was placed in a cuvette and the absorbance measured at 663.8, 646.8 and 480 nm with a Cary 1 spectrophotometer. Chlorophyll a, chlorophyll b, and total carotenoid concentrations were calculated using the extinction coefficients derived by (Wellburn 1994). Subsequently, the retrieved chlorophyll levels were averaged per tree. The results are shown in Table 3-1. Note that some of the trees were not sampled.

Table 3-1: Average chlorophyll a + b concentration [µg/cm²] for the different iron chelate treatments and the three study sites

Group / *Site*	*a*	*b*	*c*	*d*	*e*	*Total #*
CLO1	59.82	69.31	72.57	74.71	79.40	48
CLO3	51.69	60.64	54.49	57.70	57.08	40
CLO4	74.55	76.08	78.21	80.05	82.85	39

We can observe that on site CLO3 the iron chelate application has not (yet) resulted in the anticipated behaviour of an increased C_{AB} content with an increasing application dose. However, in this study we only require the availability of a range of C_{AB} and therefore this is not a problem.

3.1.2 Airborne imagery

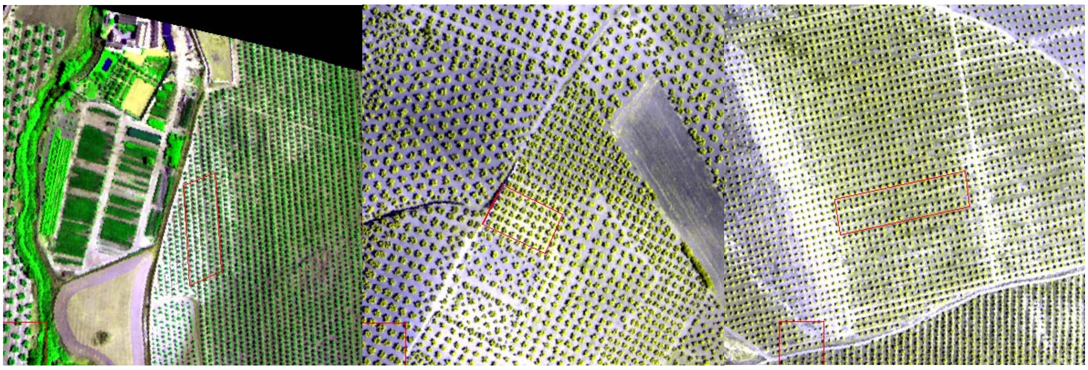

Figure 3-1: CASI imagery of Cab-study sites (sites indicated with a red polygon). C_{AB} site 1 (left), C_{AB} site 3 (middle) and C_{AB} site 4 (right)

On the 20th of July 2003, images of the study sites were collected at 1 m and 4 m pixel size with the compact airborne spectrographic imager (CASI) over the 3 sites where C_{AB} had been measured (Figure 3-1).
CASI was the first commercially available airborne hyperspectral sensor that was programmable (Lillesand and Kiefer 1999). It has a maximum of 288 bands between 0.4 and 0.9 μm. In this study we have used 8 bands in the spatial mode (1 m) and 72 bands in the spectral mode (4 m), see Table 3-2 and Table 3-3. The 12-bit radiometric resolution data collected by CASI were processed to at-sensor radiance using calibration coefficients derived in the laboratory by the Centre for Research in Earth and Space Technology (CRESTech). The aerosol optical depth that was determined in the field during image acquisition at 550 nm was used to subsequently process the image data to ground reflectance with the CAM5S atmospheric correction model (O'Neill et al. 1997; Zarco-Tejada et al. 2001).

Table 3-2 $CASI_{SPATIAL}$ band positions

Band	Band center [nm]	FWHW [nm]
1	490.744	12.3317
2	550.623	12.3805
3	670.355	10.5794
4	700.709	6.7989
5	750.191	6.8137
6	762.591	4.9087
7	775.004	6.8209
8	799.864	10.6563

Table 3-3 $CASI_{SPECTRAL}$ band positions

Band	Band center [nm]	FWHW [nm]	Band	Band center [nm]	FWHW [nm]
1	408.924	4.2789	37	678.884	4.3442
2	416.337	4.2808	38	686.471	4.346
3	423.755	4.2827	39	694.062	4.3477
4	431.178	4.2846	40	701.658	4.3494
5	438.606	4.2865	41	709.259	4.3511
6	446.039	4.2884	42	716.864	4.3528

7	453.477	4.2902	43	724.474	4.3545
8	460.921	4.2921	44	732.088	4.3562
9	468.369	4.294	45	739.707	4.3579
10	475.822	4.2958	46	747.331	4.3596
11	483.28	4.2977	47	754.959	4.3613
12	490.743	4.2995	48	762.591	4.363
13	498.21	4.3014	49	770.228	4.3647
14	505.683	4.3032	50	777.869	4.3663
15	513.161	4.305	51	785.515	4.368
16	520.643	4.3069	52	793.165	4.3697
17	528.131	4.3087	53	800.82	4.3713
18	535.623	4.3105	54	808.479	4.373
19	543.12	4.3123	55	816.143	4.3746
20	550.622	4.3141	56	823.811	4.3763
21	558.129	4.3159	57	831.483	4.3779
22	565.641	4.3178	58	839.159	4.3796
23	573.157	4.3195	59	846.84	4.3812
24	580.678	4.3214	60	854.526	4.3828
25	588.204	4.3231	61	862.216	4.3845
26	595.735	4.3249	62	869.91	4.3861
27	603.27	4.3267	63	877.608	4.3877
28	610.811	4.3285	64	885.31	4.3893
29	618.356	4.3302	65	893.017	4.3909
30	625.905	4.332	66	900.729	4.3925
31	633.46	4.3338	67	908.444	4.3941
32	641.019	4.3355	68	916.164	4.3957
33	648.582	4.3373	69	923.888	4.3973
34	656.151	4.339	70	931.616	4.3988
35	663.724	4.3408	71	939.348	4.4004
36	671.301	4.3425	72	947.085	4.402

3.2 Data pre-processing

3.2.1 Co-registration

The CASI images have been recorded without differential GPS. This resulted in a shift between images covering the same site (see Figure 3-2). To correct for this, we have co-registered the $CASI_{SPECTRAL}$ with $CASI_{SPATIAL}$ re-sampled to 4 meter, using the $CASI_{SPATIAL}$ at 1 meter as a reference through a geographic link with the 4 meter $CASI_{SPATIAL}$. (This is because ENVI does not correctly co-register images of different pixel sizes.)

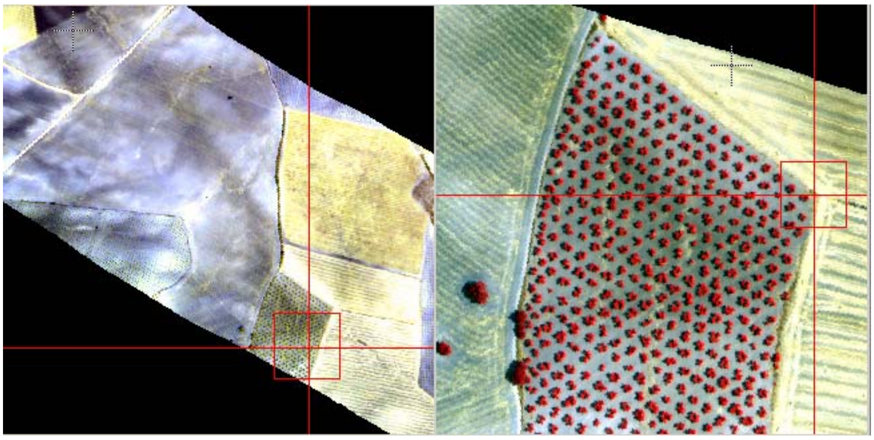

Figure 3-2: Shift between $\mathrm{CASI_{SPECTRAL}}$ (left) and $\mathrm{CASI_{SPATIAL}}$ (right) before co-registration

For the images containing field CLO4, we have collected 20 ground control points. Using a 1st degree polynomial warp, the RMSE is 0.407. The procedure was repeated for CLO3, giving an RMSE of 0.456 for 20 ground control points. Similarly, for CLO1, we obtained an RMSE of 0.498 for 30 ground control points.

3.2.2 Conversion to reflectance

The data values had been transformed from double to long integer to save disk space. Therefore, they had to be converted back to reflectance by applying the following formula according to: $\frac{x}{50000} - 0.1$.

3.2.3 Data subsetting and soil classification

In order to reduce storage and computational requirements, the $\mathrm{CASI_{SPECTRAL}}$ images were spatially subset to regions of 100 by 100 pixels.

3.2.4 Masking of the vegetated areas

In order to identify the vegetation and individual crowns within the CASI images, four different vegetation masks have been created and tested.
These masks consisted of:
1) a manual selection of the crowns,
2) tree surroundings where NDVI > 0.3,
3) tree surroundings where NDVI > 0.4, and
4) sunlit, vegetated areas of the tree where NDVI > 0.4.
The last mask was created by applying a maximum likelihood classification over the CASI image trained by a selection of prototype spectra for sunlit or shaded vegetation/background. All pixels other than sunlit vegetation were subsequently masked out.

3.3 The FLIGHT model

Heterogeneous canopies such as forests and open canopies present complex radiative transfer systems, both on a detailed level (interactions within the canopy) as well as on a scene level (influence of the background signal, mutual shadowing, etc.). The FLIGHT model (Forest LIGHT interaction model) (North 1996) was designed to deal with this kind of canopies and their associated multiple scattering events.
The 3-D model is based on Monte Carlo simulation of photon transport and simulates the bidirectional reflectance of a scene with a forest canopy represented by geometrical elements. A single tree can be modelled as a sphere or cone of leaves (the crown) with

the possibility of adding senescent leaves and woody elements (through a fraction of bark and/or the presence of a stem). The dimensions of this crown need to be specified as well as the position of the tree with respect to the scene centre. Furthermore, the model requires as an input the optical properties of the green leaves (reflectance and transmittance) together with its spatial distribution (LAD), the leaf size and the leaf area index, as well as the optical properties of the senescent leaves and bark, if applicable. The atmospheric conditions are estimated from the aerosol optical depth (see section 2.2.6 for the calculation of the fraction of diffuse and direct radiation) and the soil background reflectance signal has to be presented in combination with an estimate of the soil roughness index (section 2.2.8.1). These inputs are completed with the viewing geometry, the extent of the scene (choice of pixel size) and the setting of the output hemi-spherical grid (in terms of Φ_V & θ_V) and the number of photons (# photons) to be used in the simulation. The latter affects the accuracy of the estimated HDRF, which can be estimated as (North 1996): the mean error [%] = $100 \times \sqrt{\frac{\theta grid \times \phi grid}{\# \, photons}}$, where θ_{GRID} and Φ_{GRID} are specifying the number of angle bins sampling the grid in terms of θ_V and Φ_V (see Figure 3-3; each polygon (enclosed by the lines) corresponds to one angle bin).

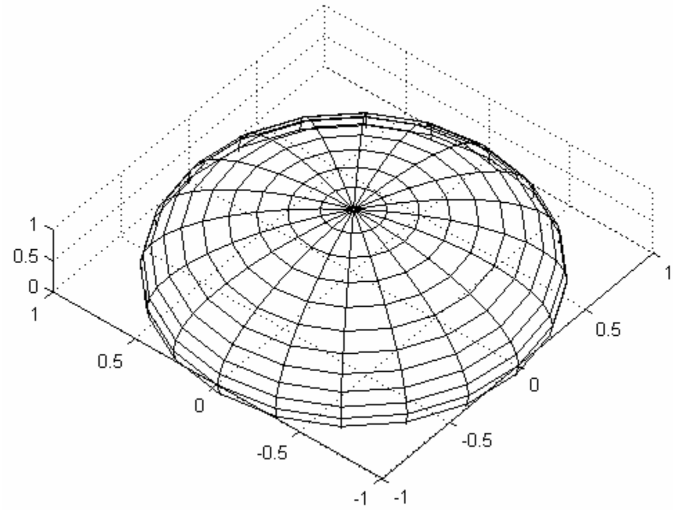

Figure 3-3: Example of an angle bin grid as used by the FLIGHT model

In this study we want to retrieve biophysical and biochemical properties (canopy variables (CV)) of the olive trees in the CLO fields. To do so, we can follow 2 approaches, namely by means of modelling the radiative transfer for these fields or by statistically relating vegetation indices with these properties. The first approach is described in chapter 4 and requires simulations with the PROSPECT and the FLIGHT model. The preparation of these simulations will be described in following parts of this chapter.

The second approach to retrieve biochemical/biophysical parameters using vegetation indices is described in chapter 5, where we focus on the retrieval of C_{AB}.

We have made some adjustments to version 5.0 of the FLIGHT model. These adjustments can be found in Appendix I.

3.4 Preparation of the Flight scenarios

This section describes the inputs and runs of the PROSPECT+FLIGHT model to obtain the canopy reflectances for different canopy and scene conditions. Figure 3-4 describes the inputs for and results of the PROSPECT and FLIGHT models.

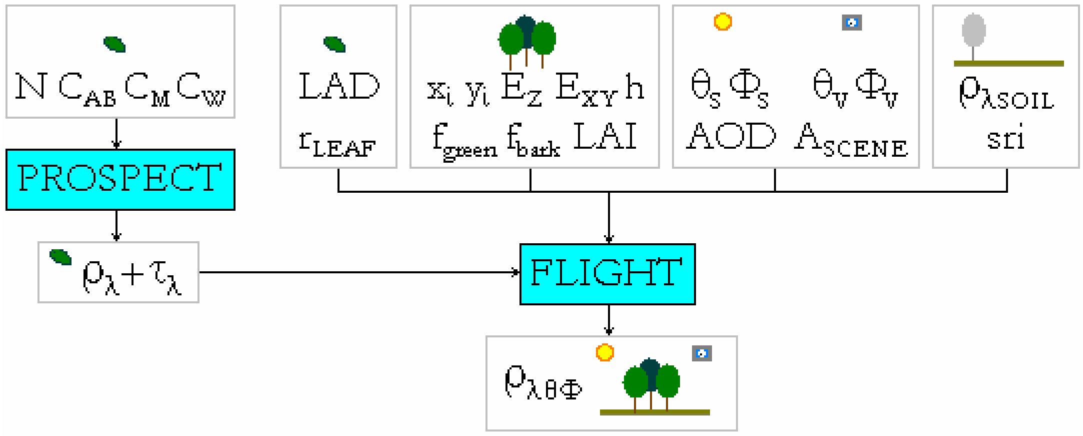

Figure 3-4: Flowchart of inputs into PROSPECT and FLIGHT with their outputs

3.4.1 FLIGHT inputs derived from the PROSPECT model

The FLIGHT model requires the input of leaf transmittance and reflectance data. To obtain these, we have used the PROSPECT leaf model. The model uses C_{AB}, C_M, C_W and N as an input. The factor C_W does not influence the reflectance or transmittance over the wavelengths used in our study and can therefore be set to a constant value. As C_M and N are normally considered to be constant for a specific species under the same conditions (i.e. in the same field) (Zarco-Tejada et al. 2005), an estimate of these values for olive leaves has been made. The reflectance (ϱ) and transmittance (τ) of 30 olive leaves were measured in a laboratory. These spectra have been used to invert the PROSPECT model (section 3.4.1.1). To test whether genetic algorithms (GA) could outperform a least squares minimisation function, first a set of simulated spectra was created. Both methods used the same evaluation function that consisted of the RMSE between the measured and estimated leaf ϱ & τ. It was observed that the GA performed very slowly and that the minimisation function reached a lower RMSE, especially if noise (±5%, see Figure 3-5) was added to the spectra. Therefore it was chosen to apply the minimisation function to invert the measured spectra to obtain the CV.

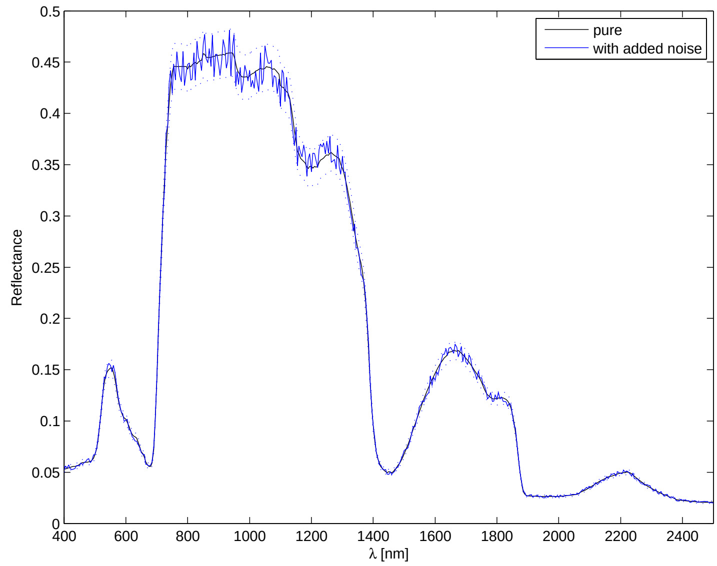

Figure 3-5: Example of adding normally distributed noise (between plus/minus 5%) to the leaf reflectance. Black line indicates the original spectrum (N=1.8, C_{AB} =42 µg/cm², C_M = 0.0158 g/cm² and C_W = 0.049 g/cm²), blue line shows the noisy signal and the dotted blue line indicates the plus or minus 5% border.

3.4.1.1 Inversion of PROSPECT to retrieve N and C_M

To obtain a first guess for the structural parameter N and the dry matter content C_M, we have used the reflectance and transmittance spectra of 30 olive leaves that were measured in a laboratory (see Figure 3-6). The leaves had been measured between 352 nm and 1001 nm, but for the inversion we have only used the range between 500 and 800 nm, as outside of this range the signal was very noisy – especially before 400 nm and after 900 nm. To further reduce the noise, the leaf reflectance and transmittance were spectrally resampled by averaging to an interval of 5 nm as in (Jacquemoud et al. 1996).

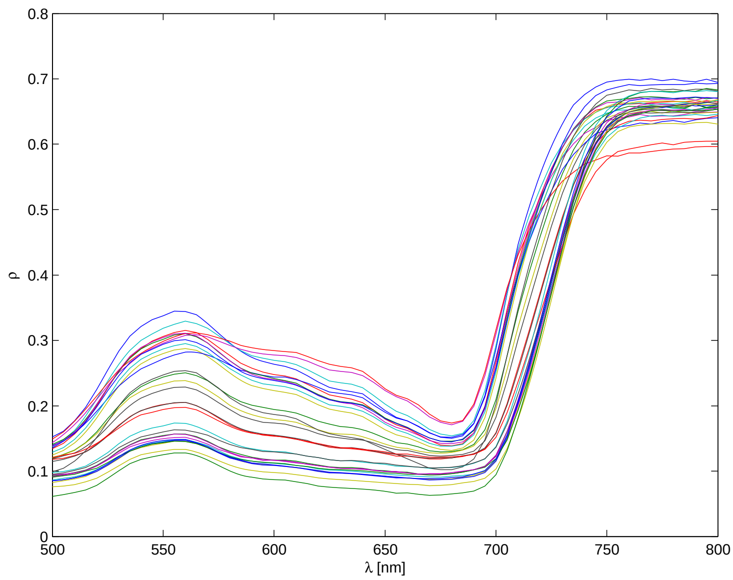

Figure 3-6: Reflectance over 500 to 800 nm from 30 olive leaf measurements averaged over 5 nm intervals

These resampled reflectances and transmittances were subsequently inverted by adjusting the input parameters for Prospect according to a minimisation of the RMSE between the simulated ϱ & τ between 500 and 800 nm and the measured ϱ & τ.
The chosen initial values were N = 2, C_{AB} = 20 µg/cm^2 and C_M = 0.01 g/cm^2. As stated in section 1.3, the water content does not influence the spectra in the chosen range and can therefore be chosen freely (chosen was C_W = 0.02 g/cm^2).
The mean retrieved structural parameter was 3.8 with a standard deviation of 0.29 (7.5%). The dry matter content was estimated to be 0.009 g/cm^2 on average with a standard deviation of 0.004. These retrieved values differ from those used in (Zarco-Tejada et al. 2004b) (N = 2.8; C_M = 0.025 g/cm^2), but can be considered normal for olive trees (P. Zarco-Tejada, personal communication).
With different first guesses, the means for the parameters do not change substantially (0-6%).

3.4.1.2 Spectral resampling of the Prospect leaf model to match the CASI sensor characteristics

The Prospect model was designed to return leaf reflectance and transmittance between 400 and 2500 nm with steps of 5 nm (see section 2.1). These radiometric estimates will be used as input for the Flight model. However, the model does not have the concept of bandwidths implemented and therefore we have to adjust the Prospect output to the observed bands and corresponding band widths.
In order to do this, each 5 nm interval of the Prospect model is linearly interpolated. Subsequently, we average this continuous dataset per band over the observed wavelengths. These new reflectance and transmittance data are then used as inputs for

the Flight model and correspond exactly with the wavelengths used by the CASI sensor. See the figure below for two examples.

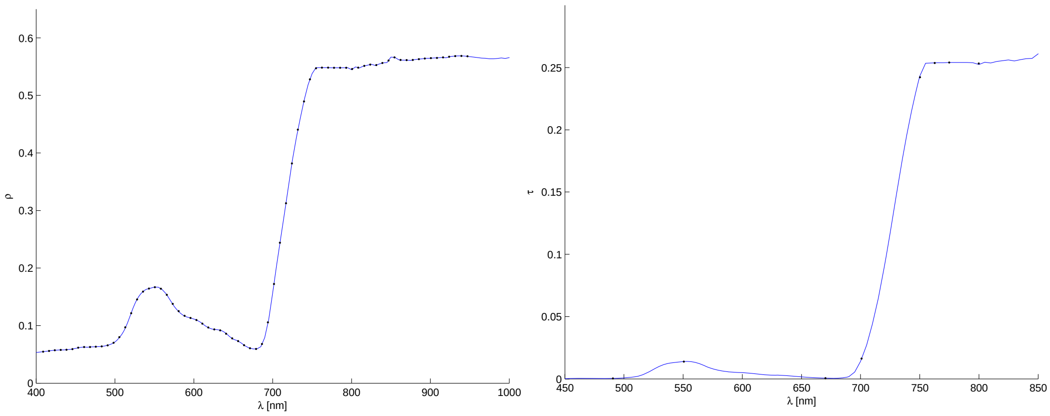

Figure 3-7: Adjusted leaf reflectance (left) and transmittance (right) for the $CASI_{SPECTRAL}$ (left) and the $CASI_{SPATIAL}$ (right) sensors. Original Prospect simulations are in blue, resampled simulations are plotted as black dots. Used parameters for this example: N = 3.5, C_{AB} = 75 µg/cm², C_W = 0.025 g/cm² and C_M = 0.01 g/cm².

3.4.2 Other FLIGHT constant input parameters

This section will describe the other parameters that have been used for the FLIGHT simulations.

3.4.2.1 Tree dimensions (E_Z, E_{XY} and h)

The horizontal dimensions of the trees were estimated from the CASI imagery at 1 meter. The 50 trees in plot CLO4 were first identified. Subsequently, the image was classified using a maximum likelihood classification into sunlit + shadowed vegetation and background. The perimeters of the trees were then intersected with the areas classified as vegetation (both sunlit and shadowed). As each pixel covers an area of 1 m^2, we can estimate the vegetated area per tree by summing the total number of vegetated pixels. As we need to approximate the tree dimensions with an ellipsoid, the radius of the ellipsoid in the horizontal plane can be estimated by taking the square root of the area divided by pi. In summary: $r = \sqrt{\frac{\sum_{n_{vegetated}} pixels}{\pi}}$.

We find that in CLO4 the mean tree radius (E_{XY}) equals 1.91 meter, with a standard deviation of 22 cm. This was rounded to 1.9 meter.
The tree height (h) is mainly of interest to correctly estimate the length of the shadows. With a solar azimuth of approximately 45°, the height of the tree is smaller than, or at maximum equals the length of a tree's shadow. Considering the extent of the tree's shadow and the shape of the crown, we can get a rough estimation of the tree height. For plot CLO4, it was found to range between 2.5 and 3.5 meters. This is consistent with the value used in (Zarco-Tejada et al. 2004b). We will work with a crown centre located at 2.25 above the ground, reaching one meter up and one meter down (minimum height = 1.25 m and maximum height = 3.25 m, E_Z = 2 m). For the retrieval of biophysical parameters, the uncertainty related to this estimation is not of relevance, as long as we obtain at least one simulated spectrum of a shadowed background for different f_{COVER} and LAI.

Applying the same procedure for the two other fields, we find a mean tree radius of 3.04 m with a standard deviation of 25 cm for CLO3 and a mean tree radius of 1.88 m with a standard deviation of 24 cm for CLO1. The latter is very similar to the values found for CLO4 and therefore we can use the same rounded dimensions. We have to simulate a different set of tree dimensions for CLO3, which will be a crown radius of 3.0 meters, a tree height of 5 m and a crown vertical radius of 2 m.

3.4.2.2 AOD and viewing geometry

The AOD used in this study was determined during image acquisitions and was found to be 0.12 at 550 nm. This value is considered to reflect a nearly clean (upper limit) atmosphere over a continental surface (Veefkind 1999).

The solar zenith (θ_S) and azimuth (Φ_S) angles have been set corresponding to the sun position during image acquisition (Table 3-4). FLIGHT uses the following coordinate convention: the azimuth is counted counter clockwise with 0° being located in the East. Therefore the Φ_S values were transposed to match a compass orientation (0° = North, 90° = East),

The model calculates the HDRF for different viewing zenith (θ_V) and azimuth (Φ_V) angles based on θ_{GRID} and Φ_{GRID}. 10 bins were selected for the zenith plane (θ_{GRID}) and 36 bins were selected for the azimuthal plane (Φ_{GRID}), resulting in angle bins of 10°, except for zenith angles between 0 - 5° and 85 - 90° where the bin size equals 5°.

Table 3-4: Solar zenith and azimuth angles matching the solar position during image acquisition

	θ_S [°]	Φ_S [°]
CLO1	39.0620	103.4390
CLO3	43.0197	99.4169
CLO4	45.4886	97.0938

3.4.2.3 Background signal

The soil spectra have been taken from the CASI imagery (CLO4). For the simulations with $CASI_{SPATIAL}$ only a single soil spectrum has been used as an input (Figure 3-8 left). We have included 3 soils with different grades of brightness for the simulations in spectral mode (Figure 3-8 right).

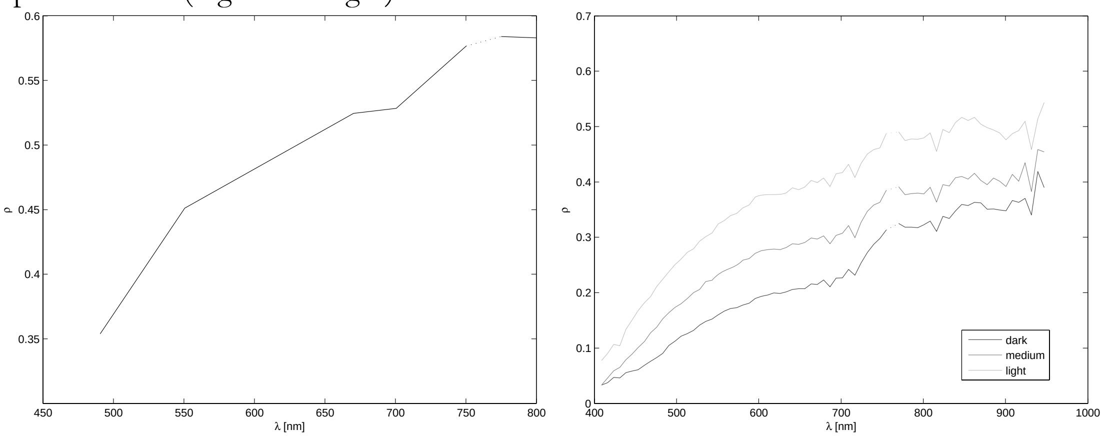

Figure 3-8: Soil spectra used in Flight simulations. Left: $CASI_{SPATIAL}$ soil spectrum, right: $CASI_{SPECTRAL}$ soil spectra. The dip at 760 nm (oxygen absorption feature) has been omitted in the figure (dotted lines).

The soil roughness index (sri) was set to 0, which corresponds to a smooth surface.

3.4.2.4 Leaf angle distribution

The olive leaves have been considered plagiophile with a distribution as specified in Table 3-5.

Table 3-5: Fraction of leaves in 10 degree zenith angles (LAD5 = 0-10°) for a plagiophile distribution, estimated after (Smith et al. 1977)

LAD	*Fraction*
5	0.02
15	0.03
25	0.05
35	0.15
45	0.5
55	0.15
65	0.05
75	0.03
85	0.02

3.4.2.5 Canopy composition

As during fieldwork no measurements of the bark of the olive trees had been done, both the spectral signal and the fraction of visible bark (in branches, twigs and stems) were unknown. For this reason it was assumed that the trees as seen from nadir position only consist of green vegetation. The Flight model then only requires the input of leaf reflectance and transmittance, which has been modelled using the Prospect leaf model. No trunk was modelled.

3.5 Simulations with the PROSPECT+FLIGHT models

Using the values described in the previous section, the PROSPECT + FLIGHT models have been run for CLO3 at a spatial resolution of 1 m and the $CASI_{SPATIAL}$ spectral resolution using different values for C_{AB}, LAI, f_{COVER} and the position of the observed part of the vegetation with respect to the crown centre (resulting in differences in light intensity/shadowing). These simulations will be referred to as PF-set 1. Considering that there may be an uncertainty in the estimation of N and C_M as they have not been validated, more simulations have been done where N and C_M were not constant. This set will be referred to as PF-set 2.

In addition to these simulations for CLO3 at 1 m with $CASI_{SPATIAL}$ spectral properties, we have also simulated CLO1 and CLO4 at this resolution (see Table 3-6) and together with CLO3 at 4 m with $CASI_{SPECTRAL}$ properties. Furthermore, we have also simulated CLO4 at additional resolutions (see Table 3-7): 32 m to assess the retrieval methods corresponding to spatial resolutions of sensors like Landsat and 300 m to work at medium resolution resolutions such as used for MODIS and MERIS.

Table 3-6: Available variation for CASI simulations in spatial mode. × indicates that the variable was having a range of values.

field	*pixel size [m]*	*x**	*y*	f_{COVER}	C_M	*N*	*LAI*	C_{AB}	*Total # sim*
CLO1	1	×	×	×	×		×	×	1792
CLO3	1	×	×	×	×	×	×	×	5764
CLO4	1	×	×	×		×	×	×	3567
Range of variation	Min	-3	-3	0	0.01	2.5	0.5	25	-
	Max	3	3	100	0.025	3.5	5.5	100	

Table 3-7: Available variation for CASI simulations in spectral mode

field	*pixel size [m]*	*x*	*y*	f_{COVER}	C_M	*N*	*LAI*	C_{AB}	*soil*	*Total # sim*
CLO1	4	×	×	×			×	×	×	5800
CLO3	4	×	×	×	×		×	×	×	5880
CLO4	4	×	×	×			×	×	×	7203
CLO4	32					×	×	×	×	1170
CLO4	300**			×	×	×	×	×	×	368

* At high spatial resolutions, different illumination conditions and f_{COVER} have been introduced by varying the position of the tree with respect to the observation area (see Figure 3-9).

** Note that for this simulation the tree dimensions have been varied in order to generate 2 values for f_{COVER}.

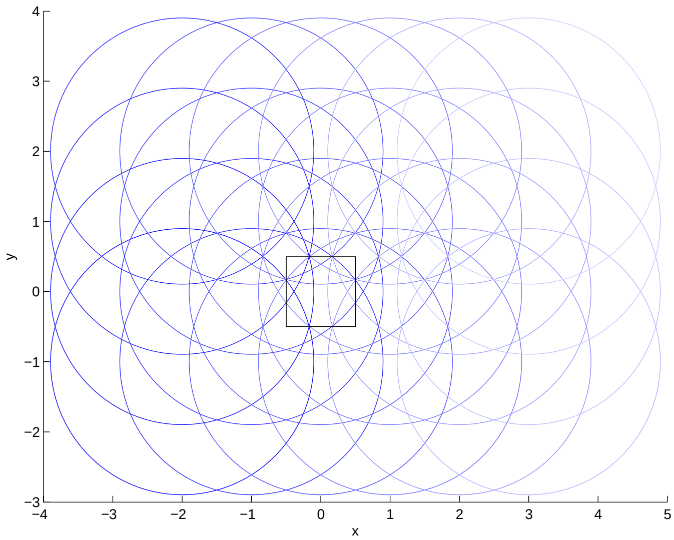

Figure 3-9: Location of observed crown (blue/grey surroundings) with respect to the observed scene (black square). Position of the crown is moved to simulate different f_{COVER} and different degrees of shadowing.

4 Application of neural networks to retrieve biophysical parameters

4.1 *Neural networks*

Environmental modelling is nowadays no longer restricted by technical complexity, yet by the lack of input data at the right scale. Remote sensing can provide many of these data, but we need easily applicable methods to extract the information from the remotely sensed images. Neural networks (NN) can provide such a method, as they have been found to be quite capable of dealing with datasets having a great internal diversity, especially when the relationships between inputs and output are not (fully) understood (Pal and Mitra 1999; Schultz and Wieland 1997). Different applications of neural networks in remote sensing can be found in literature (see (Egmont-Petersen et al. 2002) for an extensive review on the usage of NN for image processing), amongst which detection of clouds (Jang et al. 2006), change detection (Nemmour and Chibani 2006) and the retrieval of biophysical canopy variables (Atzberger 2004; Bacour et al. 2006; Gong et al. 1999; Weiss et al. 2000).

Neural networks (NN) are composed of simple elements operating in parallel, inspired by the human nervous system. As in the human system, the network function depends on the connections/nodes between the elements (neurons). When the input is passing through a connection (see Figure 4-1), the values are multiplied by the weight of the connection and if biases are included in the model, those are added after the multiplication.

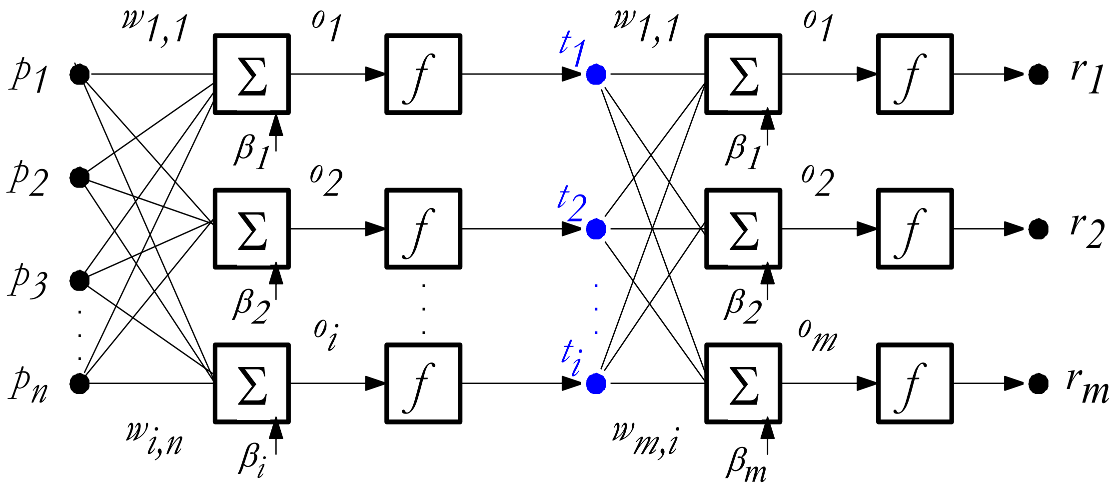

Figure 4-1: Flowchart of a two-layer NN. *P* is the input data, *w* are the weights, *β* are the biases, *o* are the inputs for the transfer functions *f*, *t* are intermediate results and *r* are the output results. Figure adapted after (Hagan et al. 1996).

The obtained value is used as input for a transfer function which will return an output. NN can be trained to simulate a function by repetitively feeding it an input and a matching target output. During this supervised learning (we inform the system about the a-priori knowledge of the output), the connections (weights and biases) between the elements will be adjusted until a proper weighting between them leads to (a close approximation of) the desired output. Different configurations of neurons can be chosen (# in layer, # of layers), as well as a set of training functions and transfer functions.

The major advantage of using NN is that we can solve a problem when we do not know what exactly the function is that converts the input into a given output. However, we should be careful not to overfit a network to the input dataset and therefore a proper validation is required. Over-fitting means that the network has been adapted to (almost) perfectly match the given training input to the given targets. Three main techniques are available to attempt to prevent this. The first technique is called *early stopping*. To ensure that the network is not overfitting to the training input data, we can add a validation set to the training. If the error for the validation set has increased when the error for the training set is still decreasing, the function has overfit and the process is stopped. The weights and biases obtained for the smallest error over the validation set will be used. A smoother response can be retrieved with the second technique: *regularisation*. Regularisation adds a parameter to the system that is being solved. The regularisation parameter gives information about the size or complexity of the model. The system will attempt to reduce the error (reach the goal), whilst keeping the complexity of the model as low as possible. A third solution to increase generalisation is to *prune* the network after training it. This is done by removing a number of weights. In this study we have selected the early stopping technique because contrarily to the two other techniques, it can be easily implemented and does not require objective decisions.

One of the disadvantages of using neural networks is that the optimal number of layers and neurons cannot be predetermined (de Vos and Rientjes 2005) and that for each application the network should be tested with different designs. Also the initialisation parameters might play a role (see section 4.3.1). The best NN can be chosen from this series by minimising the RMSE of the output provided by each network.

Amongst the several types of NN that are available, feed forward backpropagation NN are commonly used for model inversion (Bacour et al. 2006; Gong et al. 1999; Kimes et al. 2002). These networks are also known as multi-layer perceptrons (MLP). MLP can give reasonable estimates when they are presented with input that they have never 'seen', provided that the network is properly trained. In this study we will apply MLP to retrieve canopy variables (CV) from CASI imagery. The construction of the NN and training with simulated data will be described in the following paragraphs of this chapter.

4.2 Retrieval of canopy variables using neural networks

In this section we describe the retrieval of CV using neural networks. First we shall explain two methodologies used: the classical approach – training a network with the observed reflectance values as an input to estimate the CV as the target – followed by the opposite case: the inversion of a neural network trained with the CV as input to estimate the corresponding canopy reflectance as a target. This is in principle describing the problem more naturally: the reaction – changes in canopy reflectance – is the result of a cause (changes in canopy variables), whereas the first approach deals with the reverse as was described in section 2.6. We will illustrate both cases for CLO3 at 1 m spatial resolution and the $CASI_{SPATIAL}$ band configuration.

4.2.1 Inversion of FLIGHT using neural networks

We have trained MLP with canopy reflectance as modelled by the FLIGHT model for CLO3 in $CASI_{SPATIAL}$ configuration. These simulations have been described in section 3.5. The networks were taught to estimate CV based on 7 $CASI_{SPATIAL}$ bands (band 6 was excluded because it is measuring at an oxygen absorption peak). The general procedure is illustrated in Figure 4-2.

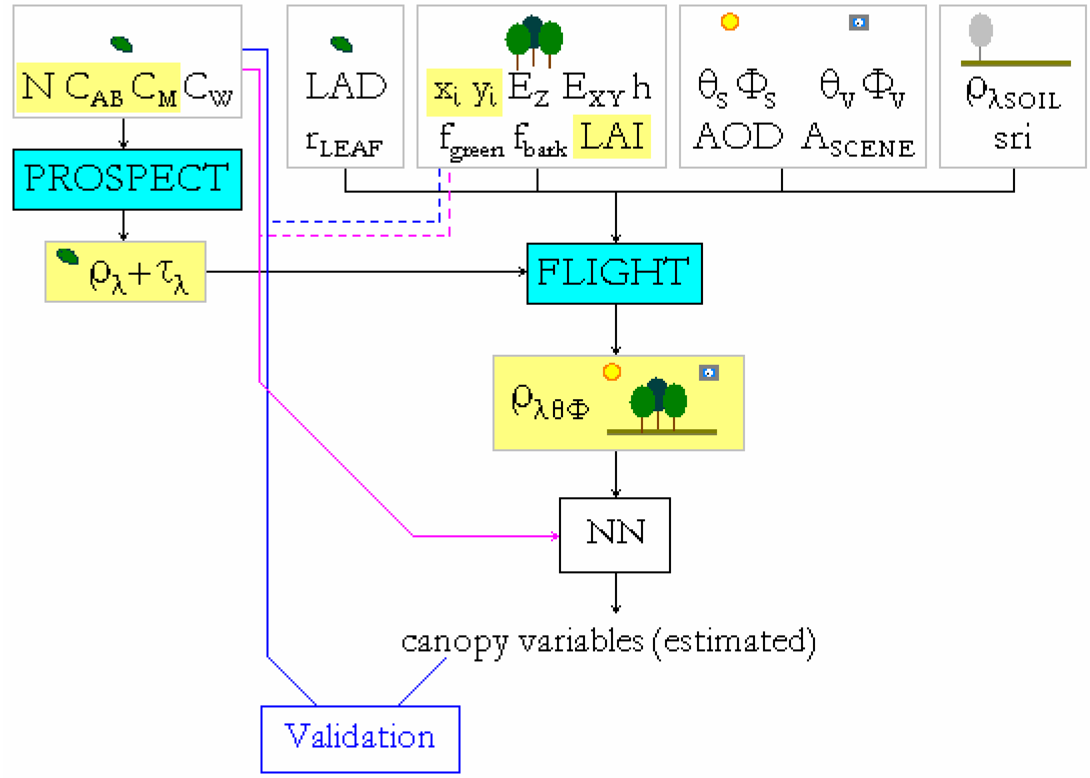

Figure 4-2: Flowchart of the methodology to retrieve biophysical canopy variables using a neural network (NN) trained with output from PROSPECT+FLIGHT to estimate the corresponding canopy variables (target path in pink) used to run the RTM. The path in blue shows the validation during the training phase and compares all estimated canopy variables with the 'known' values. Items shown in light yellow are non constant. Different fractional covers are created through changes in x and y (see Figure 3-9).

First, the leaf reflectance and transmittance are calculated using PROSPECT with N, C_{AB}, C_M and C_W. This information is passed on to the FLIGHT model together with a number of constant input parameters and the non-constant LAI, x and y position of the tree centre and consequently f_{COVER}. The canopy reflectance at given viewing conditions as calculated by the FLIGHT model is then presented to the NN, which is trained to estimate the values of the desired canopy variables (C_{AB} or C_{AB} + LAI + f_{COVER}) through learning which changes in the CV lead to particular features in the modelled canopy reflectance. The trained NN is subsequently validated by letting the NN estimate CV from a set of canopy reflectance values and comparing these estimates with the CV corresponding to the modelled canopy reflectances.

In this section we will test the response of the network to the complexity of the simulated reflectances. This is done by using the two sets (PF-sets 1 and 2) that have been described in section 3.5. First we will train the NN for the complex case in which the variability is large. Secondly, the NN are trained for the simplified case in which N and C_M are constant (PF-set 1). Both sets have been split into input data for training (90%) and validation (10%). The training data was furthermore split into 8/9 for the actual training and 1/9 to prevent overfitting using the early stopping technique. The validation results have been used to select the best NN and to assess the quality of the inversion in terms of RMSE and R^2.

To reduce the ill-posedness of the system, we can provide the NN with a-priori information in addition to the modelled reflectance (Combal et al. 2002). We have

examined the influence of knowing LAI, C_M, N and f_{COVER} on the performance of the retrieval method for C_{AB} using PF-set 2. Even though these variables do not influence the chlorophyll content, they do influence the canopy reflectance, which the NN should be able to detect.
As it is common in remote sensing that a high spatial resolution is achieved at the cost of the spectral resolution, we have also investigated the restriction of the canopy reflectance to two bands. One band in the red (band 3) and one band in the NIR (band 8) have been selected as many vegetation indices efficiently make use of these wavelengths to estimate CV (Baret and Guyot 1991; Broge and Leblanc 2000; Daughtry et al. 2000; Stenberg et al. 2004).
Finally, for environmental modelling it is of interest to simultaneously retrieve other variables than C_{AB} from remotely sensed images. Therefore, we have tested the possibility of retrieving multiple variables (C_{AB}, f_{COVER} and LAI) simultaneously. This also allows us to assess the reduction in the accuracy of the retrieval of C_{AB} by comparing with the retrieval of C_{AB} as the single CV.

As mentioned in the introduction of this chapter (section 4.1), the configuration of the NN has to be optimised for each specific problem. In this study we have carried out a number of tests to find out the most suitable NN configuration (# of layers, # of neurons, type of neurons, type of training and learning functions).

4.2.2 Inversion of forward neural networks

It is also possible to train the NN to mimic PROSPECT+FLIGHT for a given set of input parameters (scene composition, viewing geometry, etc.) and to subsequently invert the NN. If the NN has learned the "rules" of these models, although it is an approximation, the model should be able to interpolate between the known combinations used for training. The inversion of this NN can be done much faster than by directly inverting PROSPECT+FLIGHT. The NN is taught to estimate the canopy reflectance as was modelled by PROSPECT+FLIGHT by training it with the same variable inputs (CV: f_{COVER}, LAI, C_{AB}, etc.) that were used as input for PROSPECT+FLIGHT (see Figure 4-3). The NN should thus learn how each CV affects the canopy reflectance in a given wavelength. This is why we call this type of NN 'forward' NN.
After the training is completed, the NN will be inverted for each 'pixel' (corresponds here to a single simulation with PROSPECT+FLIGHT). This inversion is done using a constrained minimisation algorithm that is first calculating the RMSE between the estimated canopy reflectance from a first guess for the CV and the modelled reflectance. The values of the estimated CV ($\hat{CV}$) are slightly altered and a new value for the canopy reflectance is calculated. The algorithm will iteratively update the $\hat{CV}$ whilst attempting to minimise the RMSE, until a minimum has been found (successful inversion) or a lower/upper boundary for CV has been reached (unsuccessful inversion). The final estimates for CV are then compared with the true CV in the validation process.
As before, we test the sensitivity of the NN to the complexity of the simulations by comparing NN trained with PF-set 1 and 2. The modelled simulations have been split into 80% for training, 5% to ensure that the neural network does not overfit, 5% to evaluate the training and 10% for inversion of the network.

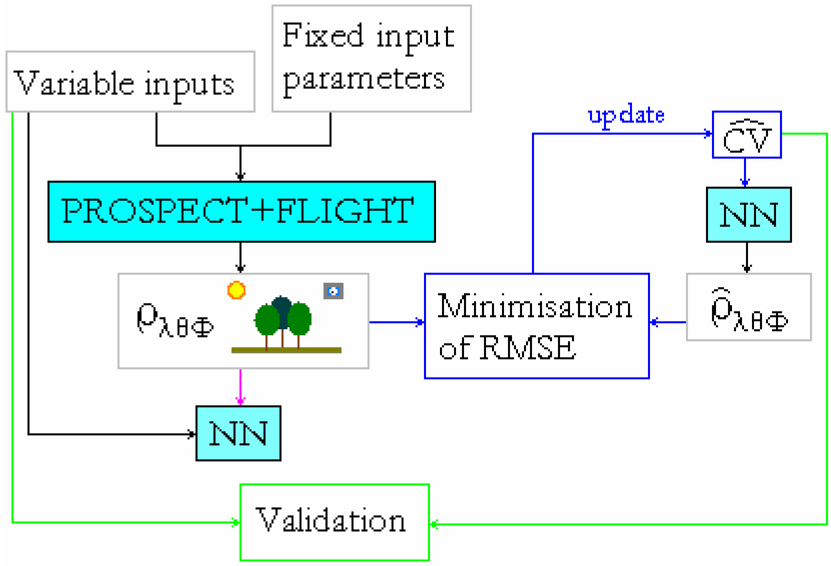

Figure 4-3: Flowchart of the methodology for the retrieval of biophysical variables by means of an inversion of a neural network (NN). NN is trained with the variable inputs to PROSPECT+FLIGHT and with the output of the linked models as the target (target path indicated in pink). The trained NN is subsequently inverted using a minimisation algorithm that minimises the RMSE between the estimated and modelled reflectance. The algorithm iteratively updates the input canopy variables for the NN (blue path), until it reaches minimal RMSE or an upper/lower boundary for the CV. The final estimated CV are compared with the true CV (denoted as Variable inputs) during the validation (green path).

4.2.2.1 *Individual band estimation followed by inversion of all bands simultaneously*

For each wavelength, a single NN has been created to estimate the reflectance at nadir in that band using the position of the tree inside the scene, the fractional cover, the dry matter content, the chlorophyll content, the leaf parameter N and the leaf area index, considering all other factors such as the viewing geometry constant. The used simulations correspond to PF-set 2.

The trained networks have been inverted using a minimisation algorithm, evaluating the root mean square error between all of the estimated and observed reflectances given by the 7 NN together by optimising the prior input parameters (position (x and y), f_{COVER}, C_{AB}, C_M, N and LAI).

4.2.2.2 *All bands simulated by a single NN*

In this step, one forward NN has been trained to estimate all 7 bands simultaneously from the 7 input variables that have been described before. Again the used simulations for training and the inversion correspond to the complex case (PF-set 2).

The trained networks have been inverted using a minimisation algorithm, evaluating the root mean square error between the estimated and observed reflectances by optimising the prior input parameters (position, f_{COVER}, C_{AB}, C_M, N and LAI).

4.2.2.3 *All bands simulated by a single NN, estimation of C_{AB}, LAI and f_{COVER}*

Subsequently, one forward NN has been trained to estimate all 7 bands simultaneously from only C_{AB}, LAI and f_{COVER}. In this less complex situation, the dry matter content and the parameter N are kept constant (i.e. PF-set 1 was used). The variation in shadowing due to the position of the observed canopy part with respect to the illumination source is still present in the data, but is now no longer used as an input for the network. In the inversion, C_{AB}, LAI and f_{COVER} are retrieved by minimising the RMSE between the modelled reflectance in all seven bands and the estimated reflectances by the NN through optimisation of these three parameters.

4.3 *Results*

This section will first describe the chosen NN configuration followed by the results of the classical inversion making use of a NN and concluding with the results of the inversion of the forward NN.

4.3.1 *NN configuration*

After some initial tests, we have decided to work with a two-layer network, see Figure 4-1. Such a network consists of one hidden layer and an output layer. In the hidden layer, we have chosen to have double the number of input variables of neurons. This is consistent with (Kanellopoulos and Wilkinson 1997) where it is stated that for a single hidden layer, the number of neurons should be at least equal to the number of inputs or outputs (whichever is greater). The layer is of the type tangent-sigmoid neurons which can deal well with inputs at different scales and input values ranging from -∞ to ∞ are rescaled to values between -1 and 1 using $O = \frac{e^{i} - e^{-i}}{e^{i} + e^{-i}}$, where O is the output and *i* is the input (Hagan et al. 1996).

The output layer is filled with linear (purelin) neurons, where the number is depending on the number of output vectors (1 vector requires 1 neuron). The output layer will take the values ranging between -1 and 1 coming from the hidden layers and convert them linearly to any output range. The training algorithm giving the best results was *trainlm*, which uses the Levenberg-Marquardt optimisation to update the weights and biases. The default learning function was used.

We note here that based on initialisation tests where it was found that a random combination of initialisation values might result in a dysfunctional network, it was decided to re-create each network in a given configuration for at least 20 initialisation tests. The best model was then chosen from these NN based on the lowest RMSE of the validation set. The effect of this initialisation randomness was larger than the influence of the number of used neurons and therefore preference was given to work around the initialisation problem over further optimising the number of neurons for the individual study cases.

4.3.2 *Results for the inversion of PROSPECT+FLIGHT by means of NN*

4.3.2.1 *Use of a-priori knowledge*

Table 4-1 shows the comparison of having a-priori knowledge and training a NN only with reflectance data. The mean C_{AB} of the validation set is approximately 60 μg/cm^2, therefore the error is between 4 and 8%.

Table 4-1: Estimation of C_{AB} from $CASI_{SPATIAL}$ bands using a neural network inversion

Input variables	*RMSE of C_{AB} [μg/cm^2]*	R^2
All bands + N	2.38	0.9896
All bands + C_M	2.53	0.9882
All bands + LAI	3.08	0.9826
All bands + f_{COVER}	3.51	0.9773
All bands without a-priori	4.66	0.9601

As expected, the RMSE without a-priori knowledge is greater than when the NN have additional information. Even though the modelled variation in N and C_M is smaller than that of LAI and f_{COVER}, the neural network benefited more from knowing N or C_M.

4.3.2.2 *Effect of constant (known) N and C_M*

Rather than specifying N and C_M as a-priori variables, for a species under specific conditions we can consider that N and C_M are constant. Using PF-set 1, we then further reduce the uncertainties in the estimation: the RMSE for the NN trained with only the reflectance where N and C_M were fixed equalled 1.98 µg/cm^2 with the coefficient of determination equal to 0.9926, which indicates that the use of a known N and C_M indeed improves the estimation of the chlorophyll content.

4.3.2.3 *Band reduction to single red + NIR bands*

To see the applicability of this type of NN for broad-band sensors, we can restrict the input bands to a band in the red domain + a NIR band ($CASI_{SPATIAL}$ bands 3+8). The found RMSE for C_{AB} was 17.21 µg/cm^2 with R^2 = 0.4408. These results are worse than for the estimation from all bands with variable N and C_M. We must hence conclude that the information from the supplementary bands adds more value than knowledge of the values for N and C_M.

4.3.2.4 *Simultaneous estimation of multiple CV*

As stated in the methodology, for environmental modelling it is of interest to be able to simultaneously retrieve information on multiple variables from the same source. We have studied the retrieval of C_{AB} together with LAI and f_{COVER} from the reflectance in seven bands with known C_M and N (PF-set 1).
Although the C_{AB} estimate is somewhat worse than estimating only C_{AB} from reflectance, the accuracy of the f_{COVER} and LAI estimates are quite fair (Table 4-2).

Table 4-2: Estimation of biophysical parameters from $CASI_{SPATIAL}$ bands using a neural network inversion. N and C_M were set to a fixed value.

Input variables	*Output variable*	*RMSE*	*R^2*
All bands	C_{AB}	2.57 µg/cm^2	0.9874
	f_{COVER}	3.91%	0.9905
	LAI	0.5201 m^2/m^2	0.7528

4.3.3 *Results inversion forward NN*

4.3.3.1 *Single NN per band, simultaneous inversion of all NN*

We have created one NN for each of the seven bands where the position, C_{AB}, C_M, LAI, f_{COVER} and N were used as input for the network to estimate the canopy reflectance for the given wavelength corresponding to the used input variables.
The training results were very good: the errors were less than 1.8% (RMSE/mean(ρ_λ)), see Table 4-3.

Table 4-3: RMSE [-]and R^2 between the modelled and estimated reflectances of the $CASI_{SPATIAL}$ bands

	B1	B2	B3	B4	B5	B7	B8
RMSE	0.0019	0.0029	0.0025	0.0032	0.0054	0.0048	0.0049
R^2	0.9997	0.9995	0.9998	0.9996	0.9972	0.9978	0.9977

The inversion was applied for 200 samples using the same minimisation algorithm calculating the RMSE between the 7 estimated and observed reflectance values. It was observed that this minimisation did not manage to solve the system of equations properly: the algorithm constantly reached one or more of the upper or lower limits.
Theoretically these NN could give a good performance, as they are highly specialised, but unfortunately the results were not very promising. This can be attributed to the ill-posedness of the system (multiple combinations of canopy/scene properties may lead to the same reflectance) and the fact that the uncertainties of 7 networks are combined.

4.3.3.2 *One NN to estimate ρ in 7 bands, followed by inversion to retrieve 7 CV*

To reduce the uncertainties due to the linking of 7 networks, we have created a single NN that was taught to estimate the reflectance in all 7 bands from the 7 input variables.
This network was inverted as in the previous section and gave more consistent results: the RMSE for C_{AB} was 16.79 µg/cm^2 and for f_{COVER} it was 0.1679. Considering that 7 variables had to be estimated, these results are relatively good.

4.3.3.3 *One NN to estimate ρ in 7 bands, followed by inversion to retrieve C_{AB}, f_{COVER} and LAI*

To reduce the ill-posedness of the inversion procedure, we have restricted the number of CV to be inverted to C_{AB}, LAI and f_{COVER}. C_M and N were set to fixed quantities (PF-set 1).
The RMSE for C_{AB} equalled 8.15 µg/cm^2, for f_{COVER} this was 0.0456 and for LAI 1.20 m^2/m^2, which is worse than an estimate equal to the mean of the dataset would be (considering a standard deviation of 1.04 m^2/m^2).

4.4 *Conclusions and discussion*

In conclusion, the classical NN were successful in retrieving the CV. The use of a-priori knowledge led to an increased accuracy of the estimates. Simultaneous retrieval of LAI, C_{AB} and f_{COVER} had very good results for the last two variables ($R^2 = 0.99$, RSME for C_{AB} was 2.57 µg/cm^2 and the RMSE for f_{COVER} was < 4%) and a reasonable accuracy for LAI (RMSE= 0.52 m^2/m^2).
Although it appears more natural to model the radiance as a result of a set of biophysical parameters (giving a known and constant viewing geometry etc.), training a neural network to run in forward mode did not give very promising results at high spatial resolutions. As shown in this chapter, the uncertainties due to imperfect training and ill-posed inversion led to a significantly reduced performance of these networks with respect to the classical inversion.
One source of uncertainties when applying NN trained with these simulated datasets to real imagery is shadowing. As the FLIGHT model does not consider vegetated parts that are outside of the scene, but that would influence the reflectance of that scene by casting shadows upon it, the modelled photon path is therefore not corresponding to the reality. Shadowed bare soils could for instance not be modelled and it is likely that this will cause misinterpretations when the retrieval algorithm is presented with such an input spectrum.

5 Application of vegetation indices to retrieve the chlorophyll content from canopy reflectance data

A vegetation index (VI) is defining a relationship between the reflectance of a leaf or canopy measured by a sensor and a property of that leaf or canopy which we shall call CV. Mathematically, the VI approximates how the reflectance ϱ is changing with an observed change of CV: VI = $f(\frac{\partial \rho}{\partial CV})$. Ideally, it should be sensitive only to that specific property and not influenced by other factors. Although vegetation indices have been found to be applicable only under a number of restrictions, they have the advantage that they can be computed much faster than canopy reflectance models can be inverted (Houborg et al. 2007).

In this chapter we present the retrieval of chlorophyll content from canopy reflectance data using vegetation indices. All relations and results are based on simulations with PROSPECT+FLIGHT. Section 5.1 describes the used indices and the relations used for the retrieval of C_{AB} using vegetation indices. The results of these relations applied to the validation data are presented in section 5.2 and a comparison with the NN results can be found in section 5.3 with the concluding remarks in paragraph 5.4.

5.1 Retrieval of C_{AB} from vegetation indices

In this study, we have CASI imagery at 1 and 4 m with different spectral resolutions. In this chapter we will first present the relations with C_{AB} based on VI derived from modelled reflectance at CASI 1 m in spatial configuration for the three plots (5.1.1). In section 5.1.2 we present the relations between C_{AB} and $CASI_{SPECTRAL}$ simulated reflectance, where we also have introduced a variability in the soil brightness. The performance of these indices has been evaluated at different spatial resolutions (4, 32 and 300 m) in section 5.1.3.

5.1.1 Application of indices at 1 m spatial resolution and with $CASI_{SPATIAL}$ bands

A number of vegetation indices have been tested in this study for the retrieval of biophysical properties and the sensitivity to other non-constant variables. Table 5-1 describes those VI used for the CASI images in spatial configuration. We will refer to the indices used with the CASI sensor in spatial configuration as $VI_{SPATIAL}$. The majority of the enlisted indices have proven to be suitable for the estimation of C_{AB}; MTVI1 and RDVI, as well as the NDVI have been associated with LAI (Haboudane et al. 2004a; Stenberg et al. 2004) and VARI was found to be suitable for the estimation of f_{COVER} (Gitelson et al. 2002). OSAVI – sensitive to LAI and f_{COVER} but not to the soil background – has been included because of its usage in the ratio of indices TCARI/OSAVI and MCARI/OSAVI.

Table 5-1: Vegetation indices used in $CASI_{SPATIAL}$ configuration

Index name	*Original wavelengths*	*Used bands*	*Source*
GM_1	R_{750}/R_{700}	R_{750}: B5 R_{700}: B4	(Gitelson and Merzlyak 1996)
GM_2	R_{750}/R_{550}	R_{750}: B5 R_{550}: B2	(Gitelson and Merzlyak 1996)
G	R_{554}/R_{677}	R_{677}: B3 R_{554}: B2	(Zarco-Tejada et al. 2005)
NDVI = PSNDa	$(R_{NIR} - R_{RED})/(R_{NIR} + R_{RED})$ $(R_{800} - R_{675})/ (R_{800} + R_{675})$	R_{NIR}: B8 R_{RED}: B3	(Blackburn 1998)

RDVI	$(R_{800} - R_{670})/(R_{800} - R_{670})^{0.5}$	R_{800}: B8 R_{670}: B3	(Haboudane et al. 2004a)
RVI = NIR/R	R_{NIR}/R_{RED} or R_{800}/R_{675}	R_{NIR}: B8 R_{RED}: B3	(Jordan 1969)
VARI	$(R_{GREEN} - R_{RED})/(R_{GREEN} + R_{RED} - R_{BLUE})$	R_{RED}: B3 R_{GREEN}: B2 R_{BLUE}: B1	(Gitelson et al. 2002)
TVI	$0.5 \times (120 \times (R_{750} - R_{550}) - 200 \times (R_{670} - R_{550})$	R_{750}: B5 R_{670}: B3 R_{550}: B2	(Broge and Leblanc 2000)
MTVI1	$1.2 \times (1.2 \times (R_{800} - R_{550}) - 2.5 \times (R_{670} - R_{550}))$	R_{800}: B8 R_{670}: B3 R_{550}: B2	(Haboudane et al. 2004a)
MCARI	$((R_{700}-R_{670}) - 0.2 \times (R_{700}-R_{550})) \times (R_{700}/R_{670})$	R_{700}: B4 R_{670}: B3 R_{550}: B2	(Daughtry et al. 2000)
TCARI	$3 \times ((R_{700}-R_{670}) - (0.2 \times (R_{700}-R_{550}) \times (R_{700}/R_{670})))$	R_{700}: B4 R_{670}: B3 R_{550}: B2	(Baret et al. 1989)
OSAVI	$(1 + 0.16) \times (R_{800}-R_{670})/(R_{800} + R_{670} + 0.16)$	R_{800}: B8 R_{670}: B3	(Rondeaux et al. 1996)
MCARI/OSAVI		"	(Daughtry et al. 2000)
TCARI/OSAVI		"	(Haboudane et al. 2002)
MTCI[a]	$(R_{753.75} – R_{708.75})/(R_{708.75} - R_{681.25})$	$R_{753.75}$: B5 $R_{708.75}$: B4 $R_{681.25}$: B3	(Dash and Curran 2004)
MERIS red-edge[a]	$705 + (48.75 \times (((R_{665} + R_{775})/2- R_{705})/(R_{753.75}-R_{705}))$	R_{775}: B7 $R_{753.75}$: B5 R_{705}: B4 R_{665}: B3	(Clevers et al. 2002)
CASI red-edge[b]	$700.7088 + (49.4822 \times (((R_{670} + R_{775})/2- R_{701})/(R_{750}-R_{701}))$	R_{775}: B7 R_{750}: B5 R_{701}: B4 R_{670}: B3	(Guyot and Baret 1988)

[a] Note that the actual MERIS band configuration may differ (for instance, metadata of a MERIS image taken in 2003 indicated that MERIS band 9 was located at 708.42 nm).
[b] The coefficients used in the CASI red-edge indices have been calculated according to (Guyot and Baret 1988).

For the three study sites, the simulated HDRF at 1 m have been used to calculate each of the indices. The corresponding C_{AB} content was then related to the VI by describing the relation through a trend line. Eight functions have been tested to fit such a trend line, which are given in Table 5-2.
Three parameters (α, β and δ) were iteratively optimised using an unconstrained non-linear minimisation of the RMSE between the estimated C_{AB} from the fitted equation and the observed C_{AB}. The best fitting function was chosen for each VI based upon the highest R^2 of the fitted equations. This model was then validated using a validation dataset and the quality of this best fit assessed in terms of R^2 and RMSE.

Table 5-2: Equations of the tested fitting equations

Model	*Equation*
1	$C_{AB} = \alpha \times \log((VI + \delta) \times \beta)$
2	$C_{AB} = \alpha \times (VI + \delta)^{\beta}$
3	$C_{AB} = \alpha \times e^{(VI+\delta)\times\beta}$
4	$C_{AB} = 1 - (\alpha \times \log((VI + \delta) \times \beta))$
5	$C_{AB} = 100 - \left(\alpha \times (VI + \delta)^{\beta}\right)$
6	$C_{AB} = 1 - \alpha \times e^{(VI+\delta)\times\beta}$
7	$C_{AB} = \alpha \times \log\left(1 - \frac{VI + \delta}{\beta}\right)$
8	$C_{AB} = \alpha \times VI + \beta + \delta$ *

* Note that the term δ is redundant in this equation.

To have an indication of how sensitive the VI are to f_{COVER} and/or LAI, two subsets of the simulations have been created. The first set is limited to values for f_{COVER} greater than 0.5 and the second set was limited to f_{COVER} greater than 0.2 together with LAI greater than 2. These partially overlapping subsets with a stronger vegetation signal will be compared with the full dataset.
As we have seen that the NN showed an increased estimation capability for cases when C_M and N were known (see section 4.4), we will also compare the performance of the indices for this situation.
The input data have been split into a calibration dataset consisting of 70% of the data and a validation set containing the remaining 30% of the data. All fitted indices have been evaluated using the coefficient of determination R^2 and the (normalised) RMSE between the estimated C_{AB} values on the trend line and the modelled C_{AB} concentration of the validation sets as this allows for comparison with the results of the NN.

5.1.2 Performance of indices for increased spectral resolution to retrieve C_{AB}

In this section we describe the estimation of C_{AB} using vegetation indices derived at spectral resolutions matching the CASI sensor in spectral mode.
Some absorption features can only be observed over very narrow wavebands. By increasing the spectral resolution and adding more bands, we want to see if we can improve the estimation of C_{AB}. These new bands offer the usage of additional VI. All used indices and band combinations are listed in Table 5-3. They will be referred to as $VI_{SPECTRAL}$.

Table 5-3: Vegetation indices used in $CASI_{SPECTRAL}$ configuration

Index name	*Original wavelengths*	*Used bands*	*Source*
GM_1	R_{750}/R_{700}	R_{750}: B46 R_{700}: B40	(Gitelson and Merzlyak 1996)
GM_2	R_{750}/R_{550}	R_{750}: B46 R_{550}: B20	(Gitelson and Merzlyak 1996)
MCARI	$((R_{700}-R_{670}) - 0.2 \times (R_{700}-R_{550})) \times (R_{700}/R_{670})$	R_{700}: B40 R_{670}: B36 R_{550}: B20	(Daughtry et al. 2000)
TCARI	$3 \times ((R_{700}-R_{670}) - (0.2 \times (R_{700}-R_{550}) \times (R_{700}/R_{670})))$	R_{700}: B40 R_{670}: B36 R_{550}: B20	(Baret et al. 1989)

OSAVI	$(1 + 0.16) \times (R_{800}-R_{670})/(R_{800} + R_{670} + 0.16)$	R_{800}: B53 R_{670}: B36	(Rondeaux et al. 1996)
MCARI/OSAVI		"	(Daughtry et al. 2000)
TCARI/OSAVI		"	(Haboudane et al. 2002)
NDVI = PSNDa	$(R_{NIR} - R_{RED})/(R_{NIR} + R_{RED})$ $(R_{800} - R_{675})/(R_{800} + R_{675})$	R_{NIR}: B53 R_{RED}: B36	(Blackburn 1998)
TVI	$0.5 \times(120 \times (R_{750} - R_{550}) - 200 \times (R_{670} - R_{550})$	R_{750}: B46 R_{670}: B36 R_{550}: B20	(Broge and Leblanc 2000)
G	R_{554}/R_{677}	R_{677}: B37 R_{554}: B20	(Zarco-Tejada et al. 2005)
RDVI	$(R_{800} - R_{670})/(R_{800} - R_{670})^{0.5}$	R_{800}: B53 R_{670}: B36	(Haboudane et al. 2004a)
MTVI1	$1.2 \times (1.2 \times(R_{800} - R_{550}) - 2.5 \times (R_{670} - R_{550}))$	R_{800}: B53 R_{670}: B36 R_{550}: B20	(Haboudane et al. 2004a)
MTCI[a]	$(R_{753.75} – R_{708.75})/(R_{708.75} - R_{681.25})$	$R_{753.75}$: B47 $R_{708.75}$: B41 $R_{681.25}$: B37	(Dash and Curran 2004)
MERIS red-edge[a]	$705 + (48.75 \times(((R_{665} + R_{775})/2-R_{705})/(R_{753.75}-R_{705}))$	R_{775}: B50 $R_{753.75}$: B46 R_{705}: B40 R_{665}: B36	(Clevers et al. 2002)
VARI	$(R_{GREEN} - R_{RED})/(R_{GREEN} + R_{RED} - R_{BLUE})$	R_{RED}: B36 R_{GREEN}: B20 R_{BLUE}: B12	(Gitelson et al. 2002)
NPCI	$(R_{680} - R_{430})/(R_{680} + R_{430})$	R_{680}: B37 R_{430}: B4	(Peñuelas et al. 1994)
Red/green	R_{RED}/R_{GREEN}	R_{RED}: B36 R_{GREEN}: B20	(Gamon and Surfus 1999)
Sipi	$(R_{800} - R_{445})/(R_{800} + R_{680})$	R_{800}: B53 R_{680}: B37 R_{445}: B6	(Peñuelas et al. 1995)
PSSRa = RVI	$R_{800}/R_{675} = R_{NIR}/R_{RED}$	R_{800}: B53 R_{675}: B36	(Blackburn 1998)
PSSRb	R_{800}/R_{650}	R_{800}: B53 R_{650}: B33	(Blackburn 1998)
Yellowness index	$(R(\lambda_{-1}) + 2 \times R(\lambda_0) + R(\lambda_{+1}))/\Delta\lambda^2$ Implemented as: (B24 – 2×B30 + B36) /(((671.3019-625.9056)/1000)^2)	B36 B30 B24	(Adams et al. 1999)
CASI red-edge[b]	$701.6588 + (38.049 \times(((R_{664} + R_{778})/2-R_{702})/(R_{740}-R_{702}))$	R_{778}: B51 R_{740}: B45 R_{702}: B40 R_{664}: B36	(Guyot and Baret 1988)

ANMB	$\dfrac{\frac{1}{2}\sum_{j=1}^{j=n-1}(\lambda_{j+1}-\lambda_j)\times(R_{j+1}+R_j)}{\max\ band\ depth}$, where *n* equals the total number of bands for the wavelengths between 650 and 725 nm with *j=1* equal to the first band within this range, etc.	B33 until B43	(Malenovský et al. 2006)

[a] The actual MERIS band configuration may differ (for instance, metadata of a MERIS image taken in 2003 indicated that MERIS band 9 was located at 708.42 nm). MTCI2 uses B46 instead of B47 and B40 instead of B41.
[b] The coefficients used in the CASI red-edge indices have been calculated according to (Guyot and Baret 1988).

Similar to section 5.1.1, we will estimate the C_{AB} content by relating its values with simulated VI.

5.1.3 *Performance of vegetation indices with increasing spatial resolution*

In this section we first explore the response of one chosen VI – the ANMB index – to demonstrate the influence of other factors than C_{AB} on the performance of VI-C_{AB} relations. This index was developed for the estimation of C_{AB} under conditions of relatively high LAI (2 to 9 m^2/m^2) and for a very high spatial resolution (Malenovský et al. 2006). First, the size of the observed scene/pixel (A_{SCENE}) was set to 1 m^2 and f_{COVER} was kept constant (100%) and a total of 11 bands was used with a dark soil background. The simulated chlorophyll contents ranged from 20 microgram/cm^2 to 80 microgram/cm^2 with steps of 5. LAI was simulated from 0.5 to 4.5 (step size 1 m^2/m^2). The performance of the index was compared with simulations done with A_{SCENE} = 32 ×32 m^2, where f_{COVER} was corresponding to field conditions. As this involved a large proportion of bare soil, the behaviour of the index was studied for different soil brightness conditions.
Subsequently, the performance of all $VI_{SPECTRAL}$ at 4, 32 and 300 m is compared.

5.2 *Results*

Note that all presented results in this section refer to the (n)RMSE and R^2 of the validation datasets.

5.2.1 *Application of vegetation indices using $CASI_{SPATIAL}$ bands at 1m*

In Figure 5-1 we have presented the results of the fitting procedure for all study sites at 1 m using the full input datasets. We can see that most indices give a normalised RMSE (nRMSE = RMSE/$\overline{C_{AB}}$) of around 0.4. The two red-edge indices and the MTCI have the best performance (see Table 5-4), followed by MCARI/OSAVI and TCARI/OSAVI for fields CLO1 and CLO4. All of these 5 indices had a worse performance for CLO3 compared to CLO1 and CLO4. Considering that the trees in this plot have greater dimensions (E_{XY}, E_Z and h), this might be explained by the trees having a greater structural complexity and more intense differences between sunlit and shaded parts of the canopy. On the other hand, the f_{COVER} was on average greater than for the CLO1 and CLO4 datasets and therefore the soil background signal was weaker.

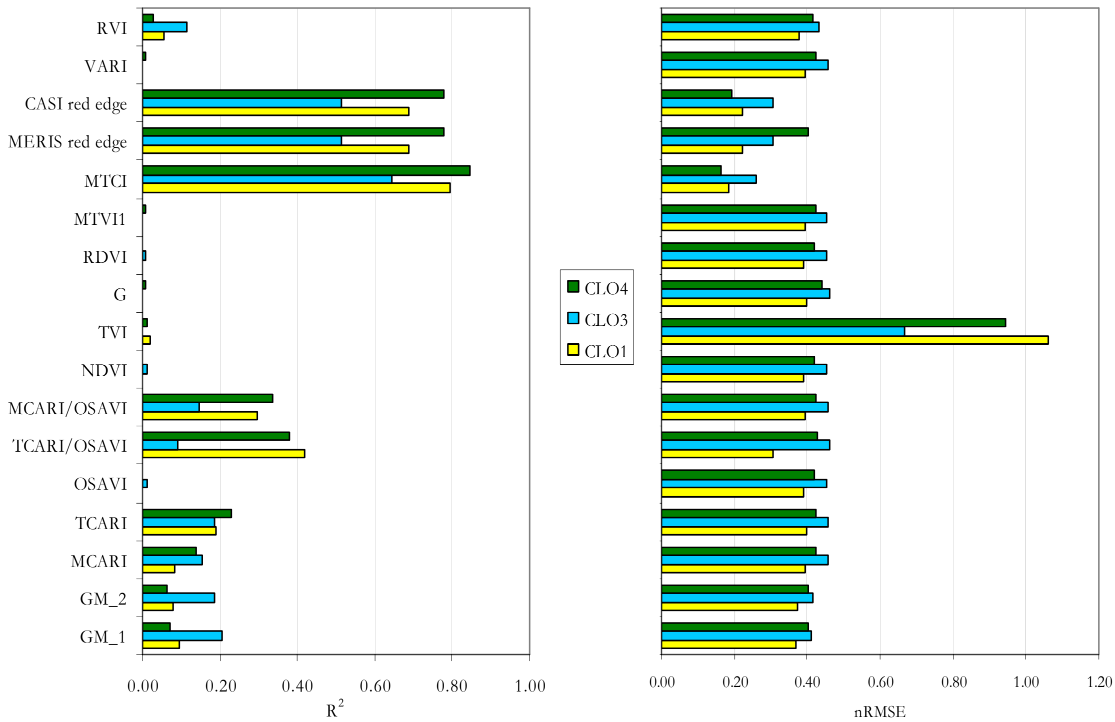

Figure 5-1: Obtained R^2 and normalised RMSE values for the fitted equations relating C_{AB} content with $VI_{SPATIAL}$ at 1 meter applied to the validation set.

Table 5-4: Fitted equations for the 3 best performing indices for the three study sites

Field	*Index*	*Equation*	*RMSE*	*R^2*
CLO1	MTCI	$C_{AB} = 100 - \left(43.35 \times (VI - 28.37)^{-0.05}\right)$	9.21	0.79
	$MERIS_{RE}$	$C_{AB}=7.08 \times VI - 5095$	11.30	0.69
	$CASI_{RE}$	$C_{AB}=6.98 \times VI - 4992$	11.30	0.69
CLO3	MTCI	$C_{AB} = 59.32 \times \log\left(1 - \frac{VI - 0.99}{-3.70}\right)$	13.17	0.65
	$MERIS_{RE}$	$C_{AB}=6.77 \times VI - 4870$	15.37	0.51
	$CASI_{RE}$	$C_{AB}=6.67 \times VI - 4770$	15.37	0.51
CLO4	MTCI	$C_{AB} = 73.79 \times \log\left(1 - \frac{VI - 1.79}{-4.93}\right)$	8.28	0.85
	$CASI_{RE}$	$C_{AB}=10.15 \times VI - 7293$	9.65	0.78
	$MERIS_{RE}$	$C_{AB} = 1.32 \times 10^9 \times \log\left(1 - \frac{VI - 618}{-2.57 \times 10^9}\right)$	20.38	0.78

RMSE is given in µg/cm² and can be converted to nRMSE by dividing by the mean C_{AB} of each dataset (50.39 for CLO1, 60.01 for CLO3 and 56.67 for CLO4). Note that for CLO4 the error is relatively high when estimating with the MERIS red-edge index. GM_1 and GM_2 had a lower RMSE, yet their R^2 was very low.

5.2.1.1 *Influence of other CV on relation $VI_{SPATIAL}$-C_{AB}*

In the following part we highlight the influence of other variables on the $VI_{SPATIAL}$-C_{AB} relations by means of the behaviour of four chosen indices as examples.

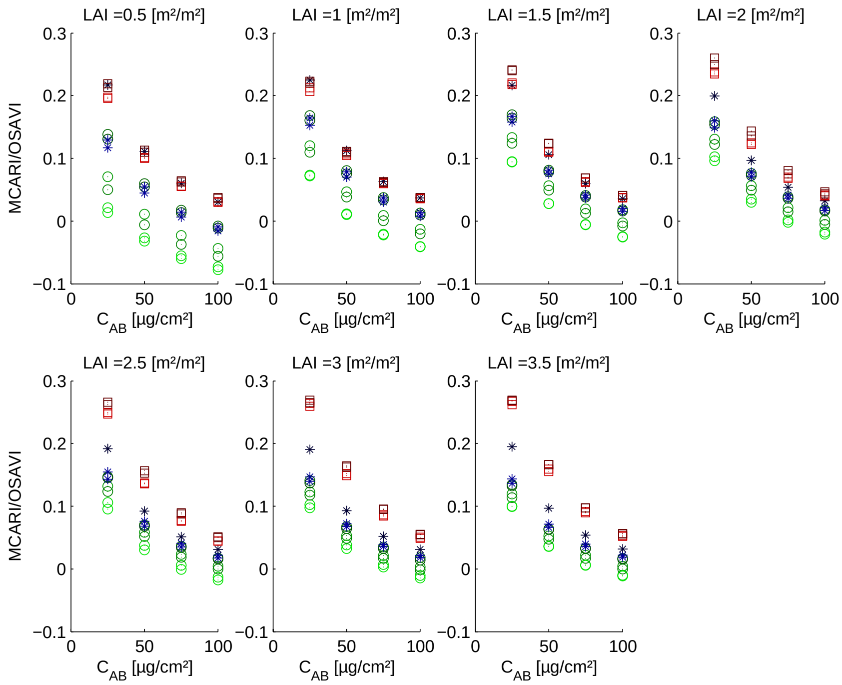

Figure 5-2: MCARI/OSAVI under different C_{AB}, LAI and f_{COVER} conditions. Green circles: f_{COVER}=13; blue stars: f_{COVER}=38; red squares: f_{COVER}=95; Darkness indicates relative shadowing of the observed scene.

In Figure 5-2 the MCARI/OSAVI index is presented for different conditions of C_{AB}, LAI and f_{COVER}. We can see that small variations in MCARI/OSAVI exist for the same combination of LAI and C_{AB}. These are mainly a result of (self) shadowing by the vegetation (sun position with respect to the observed canopy part) and partially because of small differences in the observed tree volume. Furthermore, we can see that this index is sensitive to changes in LAI for LAI<2: the index has increasing values for increasing LAI. In these cases, the index cannot be used correctly to retrieve C_{AB} unless LAI is known.

Figure 5-3 shows the TCARI/OSAVI index for different f_{COVER}. For low f_{COVER}, sunlit scenes have a lower value for TCARI/OSAVI than shadowed scenes. The differences between sunlit and shadowed scenes get smaller with increasing LAI.

Under f_{COVER} =38% sunlit scenes still have a lower value than shadowed pixels, unless LAI > 2 and C_{AB} <= 50 µg/cm². Scenes with a high fractional coverage (95/100%) that are sunlit have a higher value than shadowed areas for LAI > 0.5. The index becomes less sensitive to C_{AB} at high values of C_{AB}. We can conclude that the response of the index is a complex combination of LAI, f_{COVer}, C_{AB} and the light conditions.

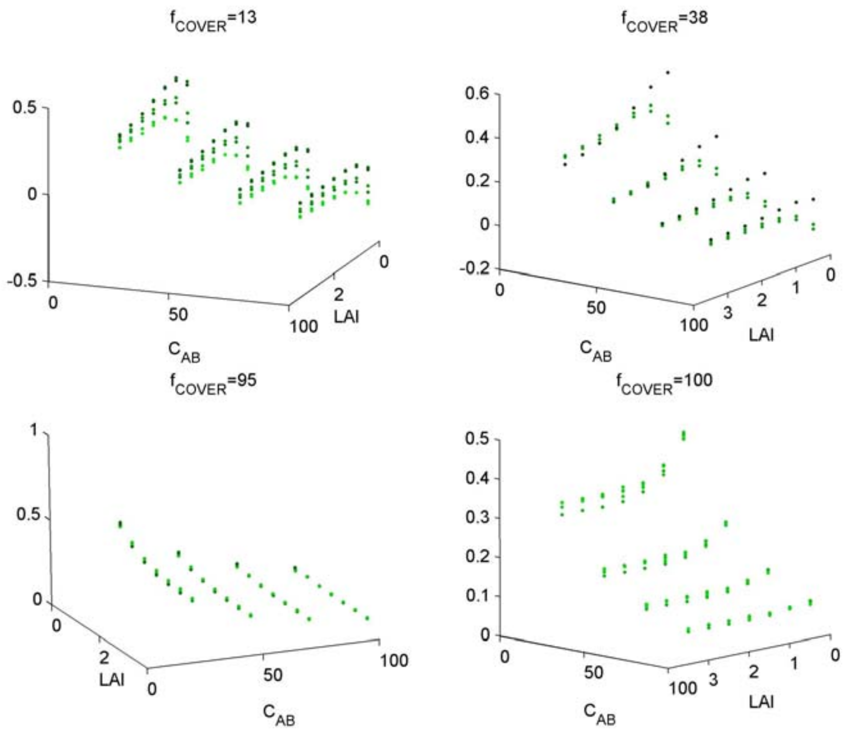

Figure 5-3: TCARI/OSAVI for different fractional coverages, LAI and chlorophyll contents. Darkness indicates relative shadowing of the observed scene.

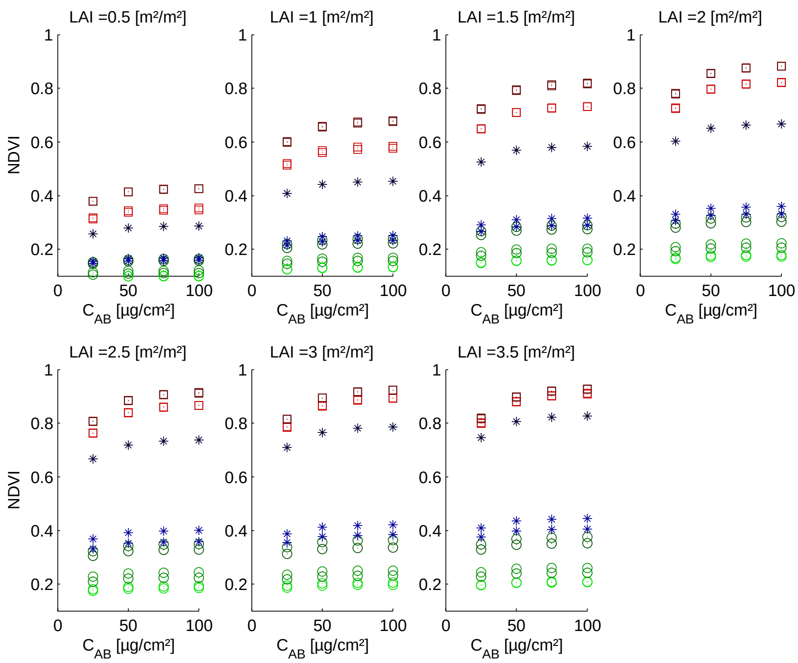

Figure 5-4: NDVI trends with C_{AB} for different LAI and fractional coverages (green circles: f_{COVER}=13; blue stars: f_{COVER}=38; red squares: f_{COVER}=95). Obscurity represents shadowing.

For the NDVI (Figure 5-4) it is clear that sunlit areas have a lower value than shadowed vegetation. This can lead to confusion with sunlit areas with a higher fractional cover (Figure 5-4). The NDVI response shows a decreasing sensitivity for increasing chlorophyll contents as well as a saturation for LAI > 2.5. These effects have been extensively described in literature (see for instance (Gitelson et al. 1996; Haboudane et al. 2004b; Stenberg et al. 2004; Thenkabail et al. 2000)) .

We conclude that the index may be of use to retrieve C_{AB} for areas with a high fractional coverage and high LAI, so that a small change in f_{COVER} or LAI is not influencing the index and as a result the retrieval of C_{AB}. The index is also sensitive to changes in the structural parameter N (not shown), which should therefore be known or constant. Vegetation with low LAI and/or low fractional coverage can show large confusion with the background spectrum, especially for soils with a high brightness. Under those circumstances, the index should not be applied to retrieve C_{AB}.

5.2.1.2 *Reduction of the non-vegetation signal*

We have compared the use of all simulated spectra with the use of spectra that have a minimal proportion of vegetation inside the observed pixel. Two of such thresholds were tested, namely setting the f_{COVER} to a minimum of 0.5 and secondly, restricting f_{COVER} to a minimum of 0.2 and the LAI to values over 2.

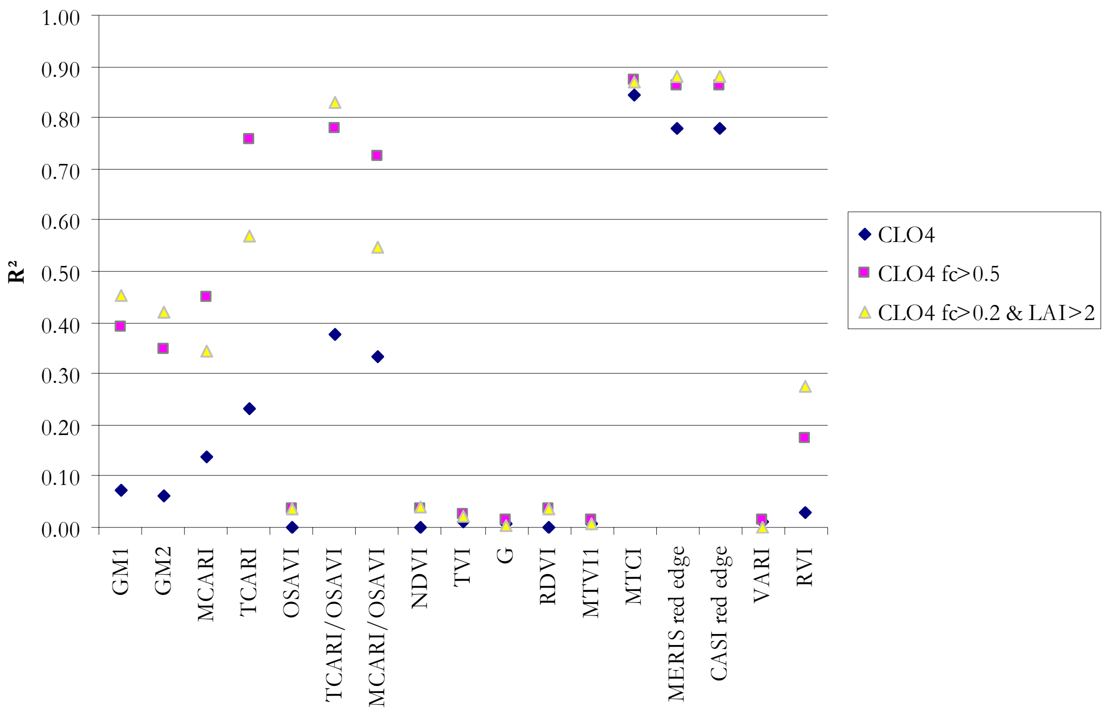

Figure 5-5: R^2 between C_{AB} estimated from VI-C_{AB} relation and modelled C_{AB} for CLO4

Figure 5-5 presents the R^2 between the estimated and modelled C_{AB} contents for CLO4. We see that all indices computed with a greater average vegetation proportion have resulted in a better fit, especially the GM and *CARI* indices (MCARI, TCARI, MCARI/OSAVI and TCARI/OSAVI). The extent of the improvement depends on the sensitivity of each vegetation index to background effects (low f_{COVER} and/or low LAI values). Some indices are therefore not very suitable for C_{AB} estimation in open canopies whereas the red edge indices and the MTCI appeared to be well applicable for both open and more closed canopies.

5.2.1.3 *A-priori knowledge of N and C_M*

Table 5-5 presents the results for the comparison of $VI_{SPATIAL}$ created with PF-sets 1 (N and C_M are constant) and 2 (N and C_M were not constant in simulations) for CLO3 at 1m spatial resolution. We see that all $VI_{SPATIAL}$ except the MCARI, TCARI and MCARI/OSAVI indices have an improved coefficient of determination when N and C_M are constant. For all indices, the nRMSE has improved. This corresponds with generally good fits of vegetation for site- and species specific conditions that are often found (Baret and Guyot 1991; Fang et al. 2003). We must note here that the fitting procedure does not necessarily reproduce the same fit when run with the same settings.

Table 5-5: Comparison of fitting results for PF-sets 1 and 2

	R^2		nRMSE	
	PFset 1	PFset 2	PFset 1	PFset 2
GM1	0.23	0.21	0.33	0.35
GM2	0.21	0.19	0.33	0.35
MCARI	0.13	0.15	0.38	0.38
TCARI	0.19	0.19	0.38	0.38
OSAVI	0.01	0.01	0.37	0.38
TCARI/OSAVI	0.11	0.09	0.38	0.39
MCARI/OSAVI	0.14	0.15	0.38	0.38
NDVI	0.01	0.01	0.37	0.38
TVI	0.00	0.00	0.37	0.56
G	0.00	0.00	0.37	0.39
RDVI	0.01	0.01	0.37	0.38
MTVI1	0.00	0.00	0.37	0.38
MTCI	0.79	0.65	0.18	0.22
MERIS red edge	0.67	0.51	0.23	0.26
CASI red edge	0.67	0.51	0.23	0.26
VARI	0.00	0.00	0.37	0.38
RVI	0.15	0.12	0.34	0.36

5.2.2 *Application of vegetation indices for the retrieval of C_{AB} using $CASI_{SPECTRAL}$ bands*

Figure 5-6 shows the performance of the $VI_{SPECTRAL}$ for the estimation of C_{AB} for the three plots. Although CLO3 presents a higher degree of complexity as C_M is not constant as with CLO1 and CLO4, the average f_{COVER} is higher due to the larger tree dimensions (see Table 3-7). The latter however also increases the differences in shadowing.

We can observe that this causes different responses by the indices for the three plots. In general, the nRMSE is lowest for CLO3, with the exception of the CASI red edge index. The performance in terms of the coefficient of determination decreases on CLO3 for the indices with an $R^2 > 0.2$ except YI, MTCI2 and TCARI/OSAVI. In this case we do see that the R^2 improves for MCARI and TCARI if they are taken as a ratio with OSAVI to correct for the influence of the soil. Furthermore, we see that the nRMSE is ranging between 0.20 and 0.40, which is similar to the values found for the $VI_{SPATIAL}$ in section 5.2.1. This points out that the variability in soil brightness gives a similar effect in terms of error as having variability in N and/or C_M.

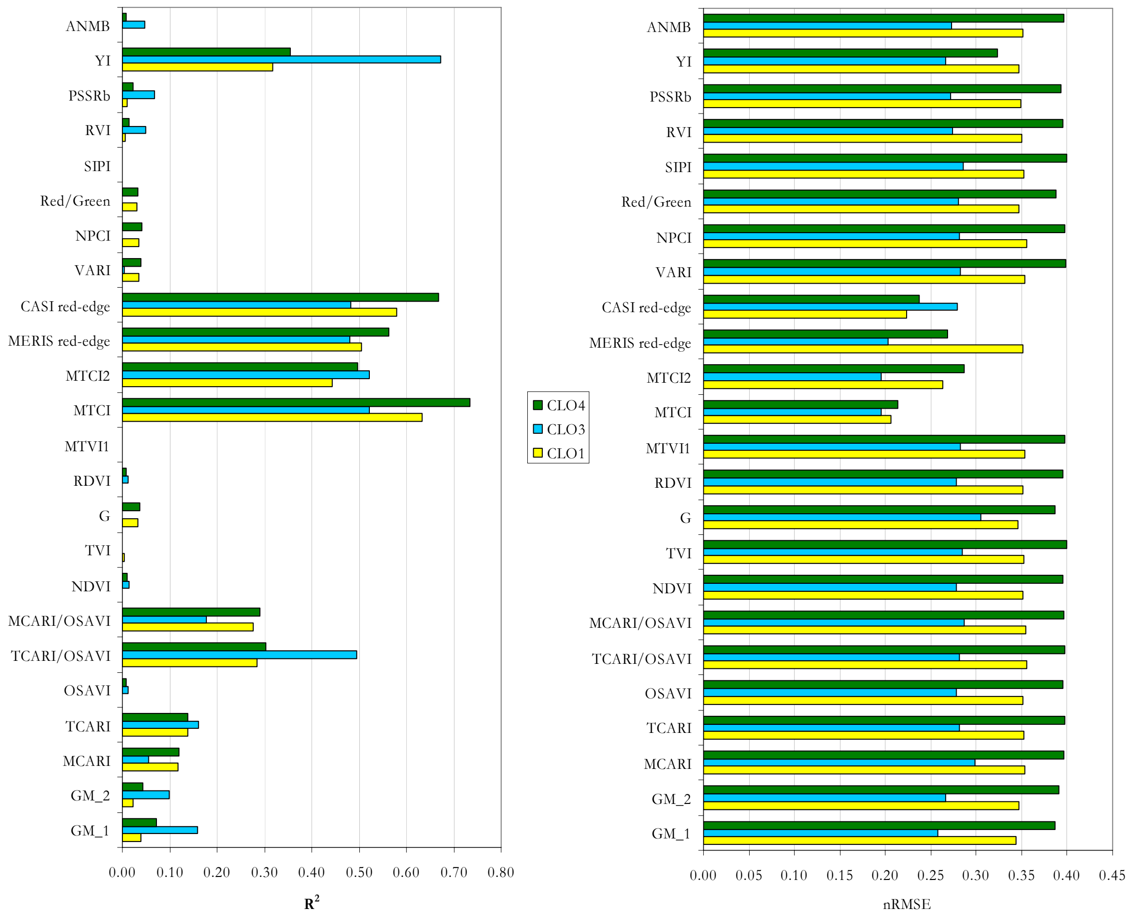

Figure 5-6: R^2 (left) and normalised RMSE (right) for the C_{AB} estimation using $VI_{SPECTRAL}$ at 4 meter.

5.2.3 Application of vegetation indices at different spatial resolutions

We have chosen one $VI_{SPECTRAL}$ – the ANMB index – to illustrate the sensitivity to other factors than C_{AB} and the dependence of these factors on the spatial resolution.

Under constant f_{COVER} (100%) and with a dark soil background, simulations have been carried out with the PROSPECT+FLIGHT model. The simulated chlorophyll contents ranged from 20 microgram/cm^2 to 80 microgram/cm^2 with steps of 5 µg. LAI was simulated from 0.5 to 4.5 with steps of 1 m^2/m^2.

We have made the following observations:

- The index increases with increasing LAI, but shows a saturation for higher values (largest response to increases at low values).
- The lower the chlorophyll content, the larger the response to LAI.
- Increase of the structural parameter N leads to a decrease in ANMB.
- Increase of LAI leads to a higher ANMB.
- Increase of C_{AB} leads to a higher ANMB.

For greater pixel sizes, we will have a lower fractional coverage, as the olive trees do not have overlapping crowns and thus some background signal will be recorded. We have investigated the use of this index at a pixel size of 32 m where we have a fractional coverage of 23%. As this involves a large proportion of bare soil, we have also simulated three different soil brightness classes.

The role of the background signal was found to significantly influence the usage of the index (see Figure 5-7). On the darkest soil, the index was showing the desired response of an increase with increasing chlorophyll contents. However, on the light soil, after an initial increase, the index was remaining equal (for low LAI) or started to decrease (LAI > 1.5) for chlorophyll contents over 50-60 µg/cm^2.

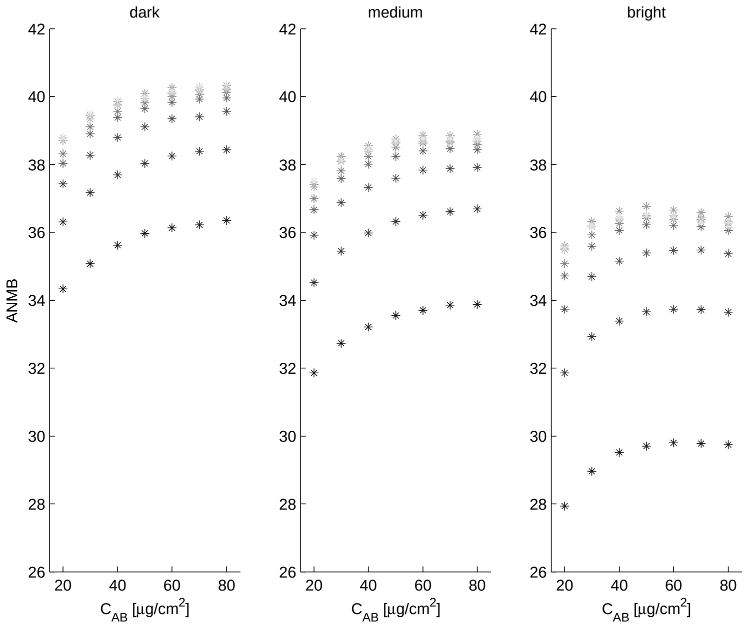

Figure 5-7: Response of ANMB to changes in background, LAI (shown as brightness with higher values having a higher light intensity) and C_{AB}. f_{COVER} = 23%, N = 4.

For other indices, a similar checking should be carried out to identify possible error sources, as the responses of indices cannot be generalised. When considering the same variation in other variables, the influence of, for example, the soil brightness on the trend of an index with changes in C_{AB} has been found to be a complex relation. For instance, the position of the MERIS red-edge shifted to lower wavelengths for increased soil brightness (not shown). This effect was however reduced when the chlorophyll content increased. The MTCI2 had a large variation in values for the index with a dark soil background while the MTCI with only slightly different wavelengths was more clearly concentrated (not shown). This indicates that care should be taken to investigate the sensitivity of an index to CV other than the single variable which will be related to by that index, when heterogeneity of these other CV is known to exist in a given study area.

At high spatial resolutions, the positions of each tree with respect to the plot are not of relevance as we observe parts of a tree up to the full tree. At low spatial resolutions we cannot distinguish the individual tree crowns and rather see the forest or plot. The spacing between trees becomes important with respect to f_{COVER} and mutual shadowing, as well as the soil background signal. At an intermediate resolution, the reflectance is both influenced by tree and inter-tree characteristics (Colwell 1974).

In Figure 5-8 we compare the performance of all $VI_{SPECTRAL}$ at different spatial resolutions for CLO4. We see that several indices such as Red/Green, TCARI and MCARI and the Greenness index (G) have suddenly an increased performance at 32 m. This can be partially attributed to the fact that the f_{COVER} and C_M were constant for simulations at this spatial resolution.

When we compare the results for MTCI and MTCI2, we also see that the chosen wavelengths, even though very close, do matter. Surprisingly, the MCARI/OSAVI and TCARI/OSAVI have a better performance at 300 m than at 4 m in this study, although the indices were developed for high spatial resolutions (Zarco-Tejada et al. 2004b). They

were also specifically designed to work with open canopy crops, such as olive trees, but the performance was much lower than that of the red edge indices and of MTCI.
Finally, Appendix II summarises a short study on the retrieval capabilities of VI at the leaf level.

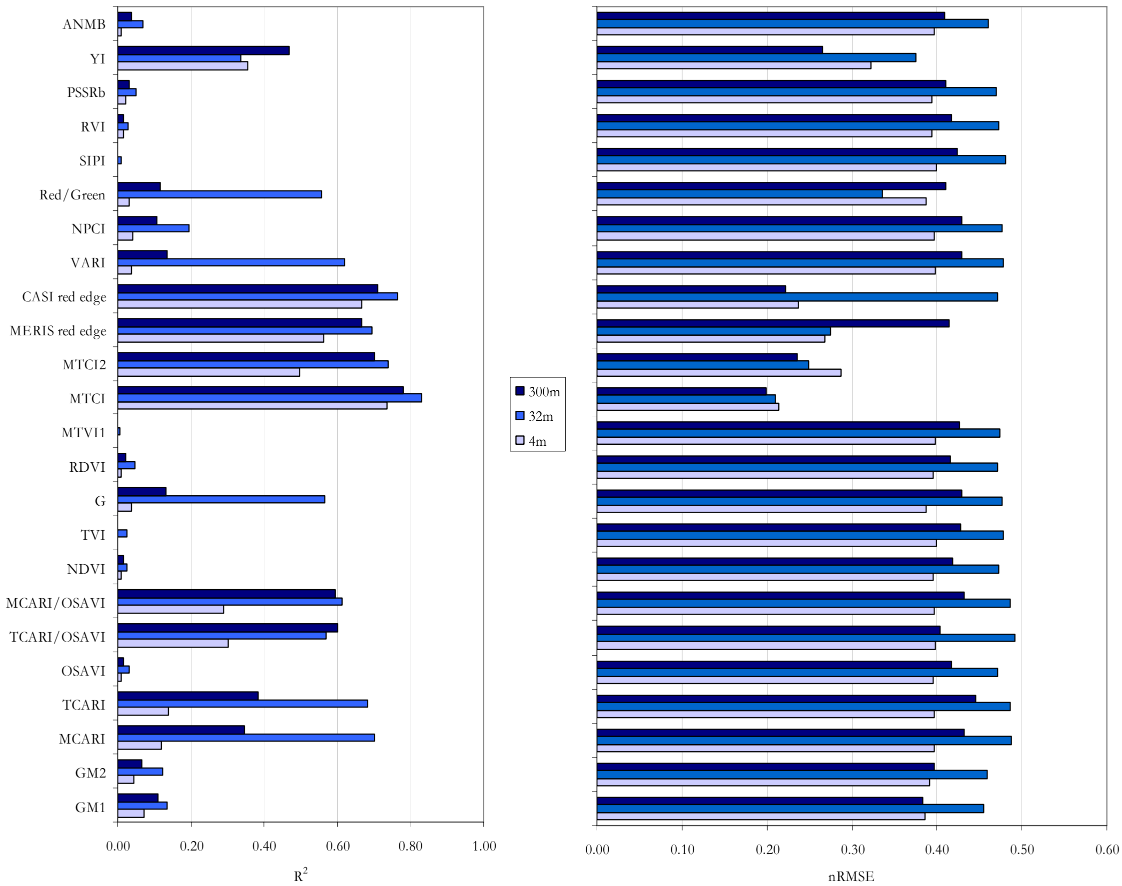

Figure 5-8: Comparison of the performance of vegetation indices at 4 m, 32 m and 300 m. Left: R^2, right: normalised RMSE.

5.3 Comparison with the results obtained from model inversion by neural networks

The results given by the indices can be compared with those obtained by the NN approaches. For this comparison, we have taken the simulations carried out with the $CASI_{SPATIAL}$ simulations for CLO3 with PF-sets 1 and 2. The three indices with the best performance have been shown in Table 5-6 together with the results for the classical NN approach.

Table 5-6: Fitting results for the MTCI, MERIS and CASI red edge indices. nRMSE was calculated by dividing the RMSE by the mean C_{AB} of the datasets

	R^2		nRMSE	
	PF-set 1	PF-set 2	PF-set 1	PF-set 2
MTCI	0.79	0.65	0.18	0.22
MERIS red edge	0.67	0.51	0.23	0.26
CASI red edge	0.67	0.51	0.23	0.26
NN only C_{AB}	0.99	0.96	0.03	0.08
NN C_{AB} + f_{COVER} + LAI	-	0.99	-	0.04

5.4 Conclusions and discussion

Indices should be evaluated for their response to non-constant variables other than the one we seek to find a relation with. In literature it has been stated that the goodness of fit should be presented in the form of a derivative (Baret and Guyot 1991; Huete et al. 1994; Ji and Peters 2007) rather than as a single value. Such an evaluation would for instance allow easy identification of saturation of the index. However, the performance of the NN cannot be evaluated in such a way as they do not describe a single relation.

Medium spatial resolution sensors such as MODIS and MERIS provide coverage of the Earth's surface with a relatively high temporal frequency. Although the spatial resolution does not allow for a detailed analysis of individual trees, it gives us an opportunity to monitor the field status throughout the growing season. As MERIS contains 15 bands in the VIS+NIR region with a resolution of 300 meter, a number of spectral indices have been developed for this sensor to derive canopy characteristics. The first index that was evaluated in this study is the MERIS terrestrial chlorophyll index (MTCI (Dash and Curran 2004)). The second index that we have evaluated is the MERIS red edge index (λ_{RE_MERIS} (Clevers et al. 2002)). Together with the developed CASI red edge index, these indices have proven to work most efficiently at both a high spatial resolution and lower spatial resolutions and to remain stable under conditions of low vegetation proportions.

However, we have seen that the neural networks were much better able to retrieve the chlorophyll content from the same datasets. Even when the inversion was complicated by adding two more CV to be retrieved, the results are still better. The only classical network that had a similar accuracy to that of the best indices was the one trained with only one band in the red and a single band in the NIR (R^2 was 0.44, nRMSE = 0.29). Considering that the MTCI has used three bands and the red edge indices have used four bands, we can conclude that some of the lower performance can be attributed to the fewer amount of available spectral information, yet the NN appear to be more flexible in relating C_{AB} to the reflectance. In addition, indices have been shown to have a larger sensitivity to other CV and may show saturation effects. Further research is needed to identify which wavelengths are essential to not only derive chlorophyll but to compensate for influences by other CV.

6 Chapter 6: Application of retrieval algorithms to CASI imagery

6.1 Retrieval algorithms developed with simulated data applied to real CASI data

In this section we show the results of the retrieval algorithms that were trained in chapter four and five. The CASI images were masked to identify each individual tree. The four masks (k) that have been described in section 3.2.4 were applied iteratively. For the VI, the mean spectrum of each masked crown was used to calculate all $VI_{SPATIAL}$ and the relations derived in chapter five were applied (see Figure 6-1). The estimated chlorophyll content was compared with the measured chlorophyll content for each tree.
A similar strategy was followed for the NN: the trees were identified, masked and the mean reflectance spectrum of each crown was used as input for the NN. The estimated C_{AB} values were compared with those measured for each of the trees.

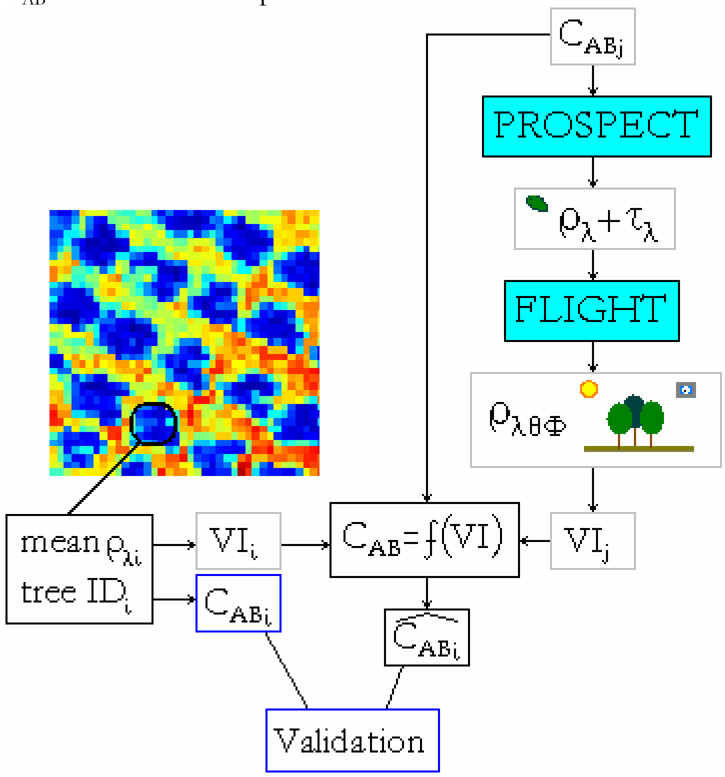

Figure 6-1: Application of the relations derived between VI and C_{AB} from PROSPECT+FLIGHT

The relations that were developed between the VI and the C_{AB} from the modelled simulations gave bad results for most $VI_{SPATIAL}$ when applied to the real data.
On CLO1, most $VI_{SPATIAL}$ underestimated C_{AB} except for the red edge indices. The restriction to more 'vegetation signal' increased the RMSE. The lowest RMSE were obtained with the red-edge indices (RMSE = 9.30 µg/cm^2 compared to a standard deviation in the observations of 9.65 µg/cm^2), but the R^2 only reached 0.11.
On CLO4, all relations based on the indices let to an underestimated C_{AB}. The RMSE was always greater than the standard deviation σ of the observations. The best estimates were given by the relation based on the MTCI trained with $f_{COVER} > 0.5$: RMSE =13.861 µg/cm^2 (compared to σ=4.979 µg/cm^2) with a low R^2 of 0.19.
For CLO3, the C_{AB} estimates produced by the $VI_{SPATIAL}$ were most often overestimating the chlorophyll content, except for GM1, GM2, MTCI and RVI that always produced an underestimation (Table 6-1). When restricting the training set to a higher degree of vegetation in the spectrum (minimum f_{COVER} and/or LAI threshold, see Table 6-1), we

see that more indices (NDVI, VARI, OSAVI, RDVI and MTVI1) produce an underestimation of C_{AB} for some k.

Table 6-1: Best estimates for C_{AB} compared to measured C_{AB} in CLO3 for different masks

Set	*k*	*Index*	*Mean est.*	*RMSE*	*R^2*
f_{COVER}>0 LAI>0	1	MCARI/OSAVI	50.80	6.845	0.26
	2	MTCI	49.47	8.550	0.35
	3	MTCI	50.33	7.916	0.33
	4	MTCI	48.64	9.355	0.30
f_{COVER}>0.2 LAI > 2	1	MCARI/OSAVI	58.66	6.383	0.26
	2	MTCI	52.76	6.162	0.34
	3	MTCI	53.74	5.768	0.32
	4	MTCI	51.79	7.023	0.30
f_{COVER}>0.5 LAI > 0	1	MCARI/OSAVI	60.80	7.389	0.26
	2	MTCI	55.39	5.037	0.34
	3	MTCI	56.41	5.104	0.32
	4	MTCI	54.38	5.664	0.30
Observations			μ=56.62	σ=6.07	n/a

μ=mean, σ=standard deviation

Inversion of the forward NN resulted in the minimisation algorithm continuously reaching the upper boundary for chlorophyll and was therefore unsuccessful.
In Table 6-2 we see that most classical NN also overestimated the chlorophyll content. Only the NN that were trained to estimate C_{AB}, LAI + f_{COVER} from all 7 $CASI_{SPATIAL}$ bands underestimated C_{AB} for k 2, 3 and 4. The NN producing the lowest RMSE had a very low R^2, indicating that all estimates were close to the mean observed C_{AB}, but uncorrelated with the observations.

Table 6-2: Estimation results for application for CLO3 of PF-trained NN

# bands	*# CV*	*PF-set*	*Mean est.*	*RMSE*	*R^2*	*k*
7	3	1	60.58	10.05	0.19	1
			44.34	14.71	0.25	2
			43.63	15.50	0.22	3
			41.54	17.02	0.16	4
7	3	2	205.52	166.52	0	1
			-213.45	300.63	0.02	2
			-256.89	330.24	0.05	3
			-107.55	234.80	0.03	4
7	1	1	213.80	157.86	0.04	1
			228.61	173.26	0.11	2
			215.71	161.69	0.05	3
			127.90	78.86	0	4
7	1	2	142.70	153.78	0	1
			1066	1028.6	0	2
			1092	1064.9	0	3
			694	684.3	0	4
2	1	1	59.30	6.61	0.02	1
			58.73	6.36	0.01	2
			57.83	6.20	0	3
			58.55	6.27	0.03	4

2	3	1	57.12	6.39	0.04	1
			69.05	13.89	0.02	2
			69.12	14.05	0	3
			71.09	15.70	0.10	4
		Observations	μ=56.62	σ=6.07	n/a	n/a

μ=mean, σ=standard deviation

The best estimation (shown in Figure 6-2) was done by the NN trained with PF-set 1 to estimate C_{AB}, LAI and f_{COVER}. The smallest RMSE equalled 10.04 (using mask 1, R^2 = 0.188) and the largest R^2 was found with mask 2 (0.254 with RMSE = 14.71).

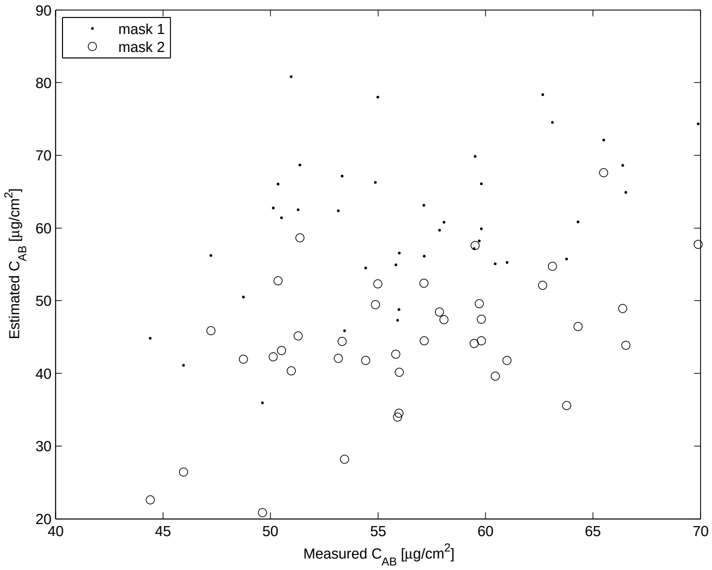

Figure 6-2: Measured C_{AB} content versus estimated C_{AB} by NN. Application of two masks is shown: first mask = manual selection of tree perimeters (includes partial soil signal), second mask = NDVI threshold of 0.3 over the first applied masks.

6.2 Identification of a mismatch between simulated and measured CASI data

In this section we describe the main reason for the bad results of the application of the retrievals with VI and NN. A substantial mismatch was found between the simulated and the observed canopy reflectances (see Figure 6-3), which strongly influenced the estimation of chlorophyll using the relations created based on the simulated data. The NDVI mask was applied as the overall means would not be comparable due to the large proportion of bare soil spectra, whereas with the simulations data few bare soil spectra were present to avoid redundancy.

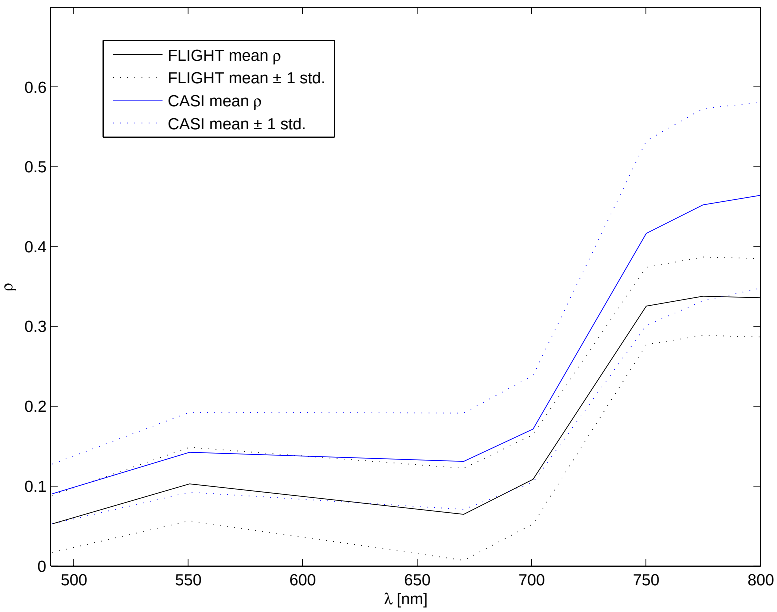

Figure 6-3: Mean reflectance (full lines) ± 1 standard deviation (dashed-dotted lines) of the FLIGHT simulations (in blue) and the $CASI_{SPATIAL}$ image for CLO3 after application of an NDVI filter (NDVI>0.3) on both datasets to obtain vegetation signals only

The overestimation of the chlorophyll content in CLO3 is contrarily to what might be expected from the offset between the CASI simulations and real data, as an increase in C_{AB} causes the leaf reflectance in the VIS to decrease whereas the observations required an increase in the VIS. However, at the canopy level we found for the crown pixels used to generate Figure 6-3 that the measured chlorophyll content was completely uncorrelated with the measured reflectance (absolute coefficient of correlation < 0.1 for bands 1 till 4).

Different reasons for the mismatch between the simulated and real data can be thought of. It could be caused by incorrect parameterisation of the PROSPECT+FLIGHT models, the processing of the CASI images and/or limitations of the PROSPECT+FLIGHT model. For instance, the FLIGHT model was not specifically designed to work at high spatial resolutions. One of the problems that we have encountered is that the illumination and shadowing was inappropriate; the model is only considering the objects (vegetation) that is present in the observed 'pixel' (see red area in Figure 6-4).

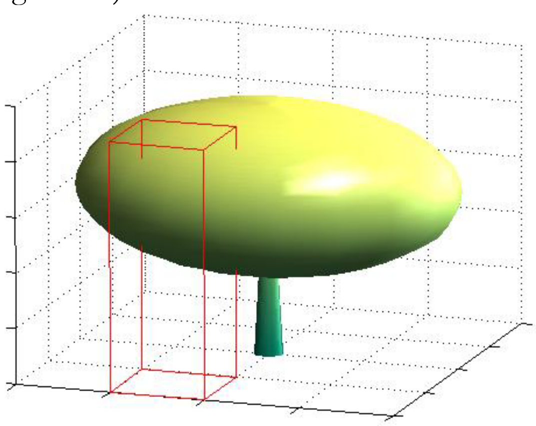

Figure 6-4: Example of tree with observation scene (indicated with the red box)

If we consider that this pixel covers a part of the tree crown on the shaded side, we know that the radiation reaching this part has already passed through other parts of the canopy. This is however not considered by the model causing a disagreement with the true illumination conditions, and as a result of the modelled reflectance. Additionally, under conditions of low LAI, the neglecting of non-vegetation elements (stem, branches, etc.) may have led to further deviations from the observed reflectances. Further work is therefore required to optimise the model for high spatial resolutions and to optimise the parameters to reduce differences between simulated and measured reflectances.

A similar mismatch problem was found by (Tan et al. 2005) in the LAI retrieval algorithm for MODIS. They revised the look-up tables by tuning the albedos for the red and NIR bands to maximise the overlap between the modelled and observed reflectances. (Malenovský 2007) has experienced such a problem as well when modelling a spruce forest stand at high spatial resolutions with the DART model. Simulated reflectances are thought to deviate from the AISA measurements because of the tree shape approximation and model parameterisation.

6.3 Creation of VI-C_{AB} relations from real data

In this section we test the estimation capabilities of VI derived from $CASI_{SPATIAL}$ for C_{AB} to see what minimum error could potentially be achieved.
The first step consisted in identifying the individual trees. Subsequently, iteratively one of the masks was applied over the selected pixels. The obtained spectra were then used to calculate the indices. The relation between the indices and the measured C_{AB} was calculated using the procedure described in section 5.1. The fit was evaluated in terms of RMSE and R^2.
Table 6-3 presents the results of fitting the $VI_{SPATIAL}$ to the C_{AB} for all three fields together (see Appendix III for the individual fields). Note that these numbers refer to all the available measurements that were used for the calibration and that therefore this only gives an indication of the best possible accuracy. We see that amongst the best indices again we find the MTCI. The lowest RMSE is given by MCARI/OSAVI (see also Figure 6-5 and Figure 6-6). Although TCARI/OSAVI has the highest R^2, its RMSE is one of the highest, indicating a systematic bias in the estimates. VARI did not have very good results with the simulated data, but with the real CASI images it is performing very well.

Table 6-3: Fitting results for $VI_{SPATIAL}$ for all three CLO fields together. Mean C_{AB} = 68.75 µg/cm², standard deviation = 11.46 µg/cm²

Index	*Mask*	*Fitting model*	*R^2*	*RMSE*
GM1	3	6	0.088	10.904
GM2	3	6	0.131	10.644
MCARI	2	8	0.338	9.290
TCARI	1	3	0.358	11.504
OSAVI	1	8	0.155	10.499
TCARI/OSAVI	2	6	0.515	11.544
MCARI/OSAVI	2	8	0.497	8.097
NDVI	1	8	0.126	10.676
TVI	1	5	0.280	9.670
G	1	6	0.467	11.567
RDVI	1	8	0.195	10.248
MTVI1	1	3	0.304	11.802
MTCI	4	5	0.497	8.230
MERIS red edge	1	8	0.271	9.572

CASI red edge	1	8	0.271	9.572
VARI	1	8	0.468	8.329
RVI	1	3	0.138	12.081

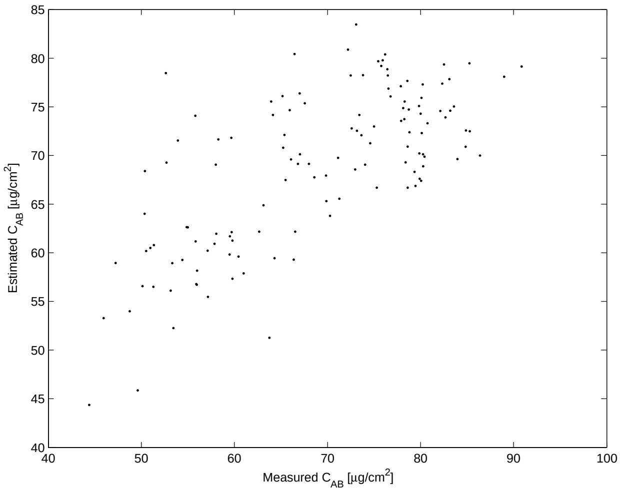

Figure 6-5: Estimates versus measurements for the chlorophyll content of 127 trees in CLO 1, 3 and 4 by applying fitted relation between MCARI/OSAVI and measured C_{AB}

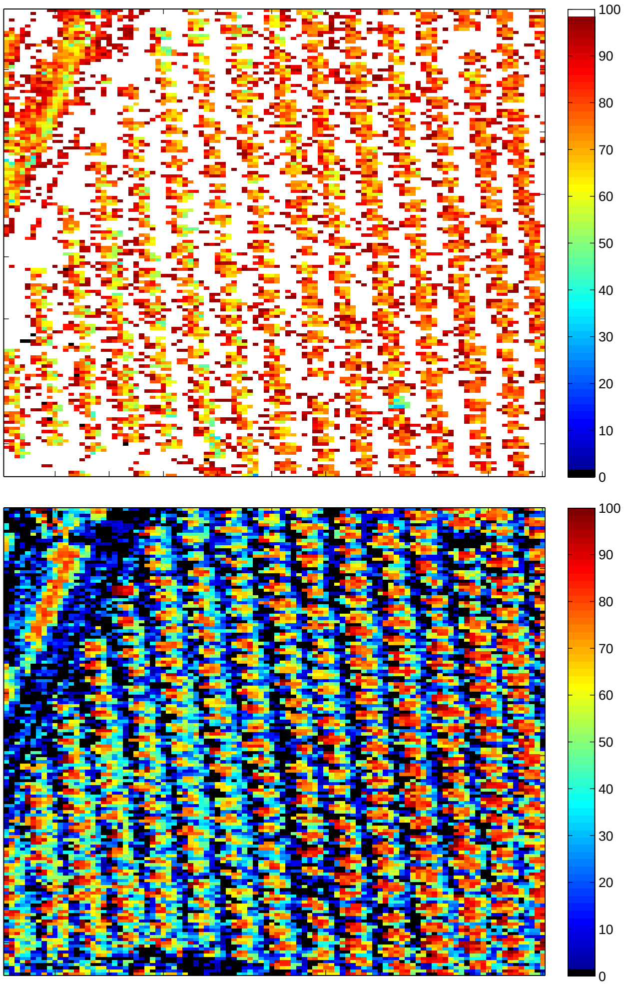

Figure 6-6: Top: Chlorophyll content estimated for CLO1 by applying relation with MCARI/OSAVI. Down: Chlorophyll content estimated for CLO1 by application of relation with MTCI. Black (white) points were having estimated values below 0 (above 100) µg/cm^2

When we compare the relations with MTCI and MCARI/OSAVI applied for CLO1 (Figure 6-6), we see that the chlorophyll contents found for the trees correspond quite well. However, the values for bare soil are in some places very different. Care should therefore be taken to isolate the canopy. A mask could be applied over the soil pixels to avoid confusion. See Appendix IV for the estimates using MCARI/OSAVI over CLO3 and CLO4.

The results for CLO3 (Appendix III Table III-2) showed that the RMSE and the R^2 for the estimation using the measured CASI reflectances and the MTCI-derived relation based on simulations are very close to the best fit-results based on measured reflectance. We can conclude that this band combination compensated quite well for the mismatch between the simulations and measured reflectances of CLO3.

6.4 Conclusions

We conclude that the application of retrieval algorithms based on simulated data should be done with care ensuring that the simulations respond to what is seen in the field. Sources of error should be identified to give an indication of the quality of the retrieved estimates. The mismatch that was found between the simulated and measured ϱ led to a large overestimation of the chlorophyll content.

For the simulated data at 1 m, the best $VI_{SPATIAL}$ to estimate C_{AB} were MTCI and the CASI and MERIS red edge indices. Relations based on simulated data applied to the CASI images showed that C_{AB} was mostly underestimated and badly estimated on CLO1 and CLO4. Best performing indices were the red-edge indices (CLO1) and the MTCI (CLO4). The results were somewhat better for CLO3, although here C_{AB} was often overestimated. The best results were achieved with MCARI/OSAVI (no masking) and MTCI (tree crowns are masked with NDVI threshold). The estimates for CLO3 by the NN trained with simulated data were worse than these indices.

Relations derived from the real data indicated that the chlorophyll content could be best assessed with MCARI/OSAVI, MTCI or VARI. The differences in best performance between simulated and real data may originate from the CASI calibration or discrepancies between simulated and field variability.

7 Conclusions and recommendations

This chapter summarises the conclusions of this thesis and our answers to the research questions.

One of main objectives of this work was to test methods for the retrieval of chlorophyll content (C_{AB}) in open canopies from remotely sensed data to identify trees affected by iron chlorosis. Two main methods were tested: 1) retrieval of canopy variables (CV) through the inversion of the linked PROSPECT+FLIGHT models by means of neural networks, and 2) application of empirical relations between vegetation indices (VI) and C_{AB} that were defined using PROSPECT+FLIGHT simulations. The simulations and the collection of inputs for the models have been described in chapter three. The chlorophyll content was varied in addition to the structural parameter N, the dry matter content C_M, leaf area index (LAI) and f_{COVER}. Different soil brightness classes were used for the simulations in spectral mode. Two sets of simulations were distinguished: PF-set 1 where the dry matter content C_M and the leaf structural parameter N were set to a fixed value and PF-set 2 where these two variables were varied.

Chapter four relates about creating the neural networks (NN) for the retrieval of CV. Two approaches were defined: the "classical" approach and the inversion of a "forward" neural network. In the classical approach, the NN were trained with simulated reflectances (PF-set 2) from PROSPECT+FLIGHT in seven bands matching the $CASI_{SPATIAL}$ wavelengths to estimate CV that were used to generate those reflectances. It was found that the use of a-priori information significantly increased the estimation accuracy of C_{AB} (RMSE decreased from 4.66 μg/cm^2 to 2.38 μg/cm^2 for training with N as additional information besides the reflectance ϱ). N and C_M were found to be providing the most useful information, followed by LAI and finally f_{COVER}. Upon using PF-set 1 (with constant N and C_M) estimation of C_{AB} from the same seven bands improved further to reach an RMSE of 1.98 μg/cm^2. When limiting the training inputs to two bands (one red and one NIR), estimates were worse (RSME =17.21 μg/cm^2) than with variable N and C_M. We therefore concluded that the information from the supplementary bands added more value than knowledge of the values for N and C_M. The simultaneous retrieval of LAI and f_{COVER} in addition to C_{AB} was also tested using the ϱ of the seven input bands and PF-set 1. The RMSE for C_{AB} then equalled 2.57 μg/cm^2, for f_{COVER} it was 3.91% and for LAI 0.52 m^2/m^2. We conclude that theoretically it is well possible to retrieve multiple variables simultaneously using an inversion of a leaf+canopy model, provided that important crop characteristics (here N and C_M) are known or can be considered constant.

In the second approach, the NN were trained with the CV as an input to estimate the canopy ϱ. The trained NN were subsequently inverted to find the CV corresponding to the canopy ϱ. The accuracy of the retrievals was lower than for the classical approach. This was attributed to the combination of the training uncertainties with the uncertainties in the inversion process. It was concluded that the classical approach led to a higher stability of the retrievals.

In chapter five we have developed relations between vegetation indices and the chlorophyll content. It was found to be most important to study the effect of other non-constant factors, such as LAI or the soil brightness on the behaviour of an index, as changes in these other factors may induce a trend that could be confused with a change in the main variable under study (the chlorophyll content in this study). The best performing indices in terms of RMSE were the approximated MTCI and MERIS red edge indices and the developed CASI red edge index. These indices were found to have a fairly constant performance over all tested spatial resolutions (1, 4, 32 and 300 m).

In chapter six we have described the performance of the trained NN and the derived relations between VI and C_{AB} when applied to the real CASI imagery. The results were very poor for CLO1 and CLO4 and not very good for CLO3. We identified a mismatch between the simulated reflectances and the measured reflectances similar to (Tan et al. 2005). Possible sources of error are incorrect parameterisation of the PROSPECT+FLIGHT models such as approximations made on the dimensions of the olive trees or the chosen soil spectrum that may not have been representative, limitations of the FLIGHT model at high spatial resolutions and calibration artefacts in the CASI images. It was shown that careful validation of the results should be done after application of model based relations. Although the simulated spectra can be transformed to correspond better to the measured reflectances (Tan et al. 2005), this could increase uncertainties and therefore requires a thorough analysis of the error propagation.

The first research question "Which spectral/spatial prerequisites need to be fulfilled to detect iron chlorosis?" can be answered by examining the effects of iron chlorosis on a plant, causing a distorted iron uptake and distribution. Iron is required by the plant to create chlorophyll molecules. An iron deficiency will therefore lead to a reduction of the chlorophyll content in new tissue. The first visible signs of this phenomenon called (iron) chlorosis can be found near the leaf veins and young leaves as the tissue is lighter green or even white in full absence of chlorophyll.
The following factors play a role in iron chlorosis (Janssen and Beusichem 2000):

- a high concentration of bi-carbonates + a high pH as a result of a high lime content of the soils
- a high sequence of precipitative events resulting in a reduced gas exchange in the soil followed by an increase of the concentration of bi-carbonates
- a high application dose of alkaline fertilisers
- a high light intensity
- an excess of heavy metals such as cadmium, cupper and nickel
- a high concentration of phosphorus in the soil or inside the plant
- specific viruses.

The extent of the area affected by iron chlorosis depends on the cause of the iron chlorosis. As stated before, if the iron chlorosis has recently been induced, it will practically only be visible in the new leaves. To be able to detect such a stress, a high spatial resolution is required.
However, the main cause of an iron deficiency is a high soil pH, as the solubility of iron (hydr)oxids decreases with increasing pH (Janssen and Beusichem 2000). As the soil pH does not increase suddenly under normal conditions, this means that most of the iron chlorosis will appear over large areas where most to all of the plant has been affected by the chlorosis during its life cycle. As a result, areas with lime induced chlorosis (iron chlorosis caused by a high soil pH) will show a uniform lower chlorophyll level than areas without iron chlorosis. In that case, the remotely sensed imagery does not have to be very spatially detailed to identify the affected areas, as the red edge indices and the MTCI showed to be well applicable for simulated olive orchards at a spatial resolution of 300 m.
With respect to the spectral resolution, we have seen in chapters 4 and 5 that the reflectance in at least three wavelengths was needed for both the best VI and the best NN. There was no significant difference in the obtained accuracy between the best indices applied from $CASI_{SPATIAL}$ or from $CASI_{SPECTRAL}$. We have shown that indices that required many narrow spectral bands such as the ANMB do not necessarily perform better. Nevertheless, no true broad bands have been tested in this study and therefore we

cannot indicate what the maximum band widths are with which accurate retrievals can be done.

We can only partially answer the second research question “Is it possible to quantify the uncertainties in the estimated biophysical parameters with respect to the spatial resolution”. With changing resolution, the factors that play a role in the reflectance observed at the sensor change as well. For instance, shadowing will be an important factor at high spatial resolutions, but at medium to low resolutions, it will be far less relevant than the very strong background (soil) signal present in open canopies.
This change in disturbing factors with upscaling may cause indices to work well at an intermediate resolution, whilst giving worse results at low or high resolutions. The applicability of an index on a certain scale will therefore depend on its sensitivity to the most important factors at that resolution. In this study we have seen that for the simulated and real data, the MTCI was a very good estimator of chlorophyll at all resolutions (leaf level, 1 m to 300 m) despite the different factors influencing the observed reflectance. A series of images with different spatial resolutions should be obtained with a single sensor to be able to validate if the observed behaviour of the indices with simulated data corresponds with the real situation.
We conclude that a general answer to this question does not exist. The uncertainties will depend on the complexity of the observed landscape, the used spectral resolutions and the used methodology to retrieve the biophysical parameters.

The third research question “Can indices be successfully used on reduced spatial resolution data to retrieve C_{AB}?” has been answered positively in the discussion of the previous question, although we must note that at least it holds true for simulated data. Again, we recommend testing with real imagery of different spatial resolutions.

The fourth question “Can we estimate chlorophyll accurately if we use 30 m pixel size imagery from olive orchards despite their heterogeneous architecture?” can be answered positively as well. We believe that the use of deriving relations between the chlorophyll content and indices based on modelling of the olive orchards should be well applicable provided that the simulations match the observed reflectances. Inversion of these models by means of neural networks is expected to give a greater accuracy, especially if a-priori knowledge can be incorporated into the retrieval process.

We have attempted to answer the fifth question “Can we correctly retrieve multiple variables simultaneously considering the ill-posedness of the radiative transfer” only by means of inversion of the PROSPECT+FLIGHT model using neural networks. If “correct” is interpreted as having a certain accuracy in the retrieval, the answer will depend on the purpose of the retrieval. If we seek to identify different groups (ranges), this will definitely be possible. In chapter 4 it was shown that from simulated data, the simultaneous retrieval of C_{AB}, LAI and f_{COVER} could be done with a reasonable accuracy for the three variables (RMSE for C_{AB} = 2.57 μg/cm^2, RMSE for f_{COVER} = 3.91% and for LAI was 0.52 m^2/m^2). However, if a very high accuracy is required, retrieval of each variable separately (by creating a NN for each variable) may give better results as the NN would be highly specialised. Finally, the combination of indices estimating different CV using the same input image may give satisfactory results as well.

As indicated before, we stress that it is very important to test the sensitivity of an index to other factors than the one you are interested in. All possible sources of variation in the study site should be assessed to see which variables can cause conflictive changes in the

index. This is also of most interest when the application of the index at different spatial resolutions is intended, as the sources of variation will change.
We further recommend testing the retrieval by means of neural networks using CASI spectral images. It would be interesting to test which bands are most optimal to have a stable retrieval considering all other factors influencing the canopy reflectance. In addition, identifying the soil brightness by means of a classification of the bare soil visible in the CASI images to present this as a-priori knowledge to the NN may have a considerable positive effect on the accuracy of the estimated CV.
Finally, we must realise that modelling will always be an approximation of reality: "The best model of a cat is another cat, or preferably the same cat" (Wiener and Rosenblueth 1945).

8 References

Adams, M.L., Philpot, W.D., & Norvell, W.A. (1999). "Yellowness index: an application of spectral second derivatives to estimate chlorosis of leaves in stressed vegetation". *International Journal of Remote Sensing, **20**(18)*, 3663-3675

Andrieu, B., Baret, F., Jacquemoud, S., Malthus, T., & Steven, M. (1997). "Evaluation of an improved version of SAIL model for simulating bidirectional reflectance of sugar beet canopies". *Remote Sensing of Environment, **60**(3)*, 247-257

Anthony, P.A., Holtum, J.A.M., & Jackes, B.R. (2002). "Shade acclimation of rainforest leaves to colonization by lichens". *Functional Ecology, **16**(6)*, 808-816

Aragão, L.E.O.C., Shimabukuro, Y.E., Espírito Santo, F.D.B., & Williams, M. (2005). "Landscape pattern and spatial variability of leaf area index in Eastern Amazonia". *Forest Ecology and Management, **211**(3)*, 240-256

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

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

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

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., Guyot, G., & Major, D.J. (1989). "TSAVI: A vegetation index which minimizes soil brightness effects on LAI and APAR estimation". *Digest - International Geoscience and Remote Sensing Symposium (IGARSS), **3***1355-1358

Barranco, D., Fernandez-Escobar, R., & Rallo, L. (1999). *El cultivo del olivo*. Madrid: Junta de Andalucía

Barták, M., Raschi, A., & Tognetti, R. (1999). "Photosynthetic characteristics of sun and shade leaves in the canopy of Arbutus unedo L. trees exposed to in situ long-term elevated CO2". *Photosynthetica, **37**(1)*, 1-16

Berterretche, M., Hudak, A.T., Cohen, W.B., Maiersperger, T.K., Gower, S.T., & Dungan, J. (2005). "Comparison of regression and geostatistical methods for mapping Leaf Area Index (LAI) with Landsat ETM+ data over a boreal forest". *Remote Sensing of Environment, **96**(1)*, 49-61

Blackburn, G.A. (1998). "Spectral indices for estimating photosynthetic pigment concentrations: a test using senescent tree leaves". *International Journal of Remote Sensing, **19**(4)*, 657-675

Boyd, D.S., & Danson, F.M. (2005). "Satellite remote sensing of forest resources: Three decades of research development". *Progress in Physical Geography, **29**(1)*, 1-26

Broge, N.H., & Leblanc, E. (2000). "Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density". *Remote Sensing of Environment, **76**(2)*, 156-172

Camacho-De Coca, F., Bréon, F.M., Leroy, M., & Garcia-Haro, F.J. (2004). "Airborne measurement of hot spot reflectance signatures". *Remote Sensing of Environment, **90**(1)*, 63-75

Clevers, J.G.P.W., De Jong, S.M., Epema, G.F., Van der Meer, F.D., Bakker, W.H., Skidmore, A.K., & Scholte, K.H. (2002). "Derivation of the red edge index using

the MERIS standard band setting". *International Journal of Remote Sensing, **23**(16)*, 3169-3184
Colwell, J.E. (1974). "Vegetation canopy reflectance". *Remote Sensing of Environment, **3**(3)*, 175-183
Combal, B., Baret, F., Weiss, M., Trubuil, A., Macé, 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
Curran, P.J., & Hay, A.M. (1986). "The importance of measurement error for certain procedures in remote sensing at optical wavelengths." *Photogrammetric Engineering & Remote Sensing, **52**(2)*, 229-241
Dash, J., & Curran, P.J. (2004). "The MERIS terrestrial chlorophyll index". *International Journal of Remote Sensing, **25**(23)*, 5403-5413
Daughtry, C.S.T., Walthall, C.L., Kim, M.S., Brown de Colstoun, E., & McMurtrey III, J.E. (2000). "Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance". *Remote Sensing of Environment, **74**(2)*, 229-239
de Vos, N.J., & Rientjes, T.H.M. (2005). "Constraints of artificial neural networks for rainfall-runoff modelling: Trade-offs in hydrological state representation and model evaluation". *Hydrology and Earth System Sciences, **9**(1-2)*, 111-126
Demarez, V., Gastellu-Etchegorry, J.P., Mougin, E., Marty, G., Proisy, C., Dufrêne, E., & Le Dantec, V. (1999). "Seasonal variation of leaf chlorophyll content of a temperate forest. Inversion of the PROSPECT model". *International Journal of Remote Sensing, **20**(5)*, 879-894
Egmont-Petersen, M., de Ridder, D., & Handels, H. (2002). "Image processing with neural networks - a review". *Pattern Recognition, **35**(10)*, 2279-2301
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
Fourty, T., Baret, F., Jacquemoud, S., Schmuck, G., & Verdebout, J. (1996). "Leaf optical properties with explicit description of its biochemical composition: Direct and inverse problems". *Remote Sensing of Environment, **56**(2)*, 104-117
Gamon, J.A., & Surfus, J.S. (1999). "Assessing leaf pigment content and activity with a reflectometer". *New Phytologist, **143**(1)*, 105-117
Garabedian, P. (1964). *Partial differential equations*. New York: John Wiley and Sons
Gastellu-Etchegorry, J.P., Demarez, V., Pinel, V., & Zagolski, F. (1996). "Modeling radiative transfer in heterogeneous 3-D vegetation canopies". *Remote Sensing of Environment, **58**(2)*, 131-156
Gausman, H.W. (1984). "Evaluation of factors causing reflectance differences between sun and shade leaves." *Remote Sensing of Environment, **15**(2)*, 177-181
Gitelson, A.A., Kaufman, Y.J., & Merzlyak, M.N. (1996). "Use of a green channel in remote sensing of global vegetation from EOS- MODIS". *Remote Sensing of Environment, **58**(3)*, 289-298
Gitelson, A.A., Kaufman, Y.J., Stark, R., & Rundquist, D. (2002). "Novel algorithms for remote estimation of vegetation fraction". *Remote Sensing of Environment, **80**(1)*, 76-87
Gitelson, A.A., & Merzlyak, M.N. (1996). "Signature analysis of leaf reflectance spectra: Algorithm development for remote sensing of chlorophyll". *Journal of Plant Physiology, **148**(3-4)*, 494-500
Glover, F. (1986). "Future paths for integer programming and links to artificial intelligence." *Computers and Operations Research, **13**(5)*, 533-549

Goel, N.S., & Grier, T. (1987). "Estimation of canopy parameters of row planted vegetation canopies using reflectance data for only four view directions". *Remote Sensing of Environment, **21**(1)*, 37-51

Goel, N.S., & Strebel, D.E. (1983). "Inversion of vegetation canopy reflectance models for estimating agronomic variables. I. Problem definition and initial results using the Suits model." *Remote Sensing of Environment, **13**(6)*, 487-507

Goel, N.S., Strebel, D.E., & Thompson, R.L. (1984). "Inversion of vegetation canopy reflectance models for estimating agronomic variables. II. Use of angle transforms and error analysis as illustrated by Suits' model." *Remote Sensing of Environment, **14**(1-3)*, 77-111

Goel, N.S., & Thompson, R.L. (1984a). "Inversion of vegetation canopy reflectance models for estimating agronomic variables. III. Estimation using only canopy reflectance data as illustrated by the Suits model." *Remote Sensing of Environment, **15**(3)*, 223-236

Goel, N.S., & Thompson, R.L. (1984b). "Inversion of vegetation canopy reflectance models for estimating agronomic variables. IV. Total inversion of the SAIL model." *Remote Sensing of Environment, **15**(3)*, 237-253

Goel, N.S., & Thompson, R.L. (1984c). "Inversion of vegetation canopy reflectance models for estimating agronomic variables. V. Estimation of leaf area index and average leaf angle using measured canopy reflectances." *Remote Sensing of Environment, **16**(1)*, 69-85

Gong, P., Wang, D.X., & Liang, S. (1999). "Inverting a canopy reflectance model using a neural network". *International Journal of Remote Sensing, **20**(1)*, 111-122

Gower, S.T., Kucharik, C.J., & Norman, J.M. (1999). "Direct and indirect estimation of leaf area index, f(APAR), and net primary production of terrestrial ecosystems". *Remote Sensing of Environment, **70**(1)*, 29-51

Grossman, Y.L., Ustin, S.L., Jacquemoud, S., Sanderson, E.W., Schmuck, G., & Verdebout, J. (1996). "Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data". *Remote Sensing of Environment, **56**(3)*, 182-193

Guyot, G., & Baret, F. (1988). Utilisation de la haute resolution spectral pour suivre l'état des couvert vegetaux. In, *4th International Colloquium on Spectral Signatures of Objects in Remote Sensing* (pp. 279-286). Aussois, France: European Space Agency

Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J., & Strachan, I.B. (2004a). "Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture". *Remote Sensing of Environment, **90**(3)*, 337-352

Haboudane, D., Miller, J.R., Tremblay, N., Pattey, E., & Vigneault, P. (2004b). Estimation of leaf area index using ground spectral measurements over agriculture crops: Predication capability assessment of optical indices. In, *XXth ISPRS congress: "Geo-Imagery Bridging Continents", Commission VII, WG VII/1* (p. 6). Istanbul

Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., & Dextraze, L. (2002). "Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture". *Remote Sensing of Environment, **81**(2-3)*, 416-426

Hagan, M.T., Demuth, H.B., & Beale, M. (1996). *Neural network design*. Boston: PWS Publishing company

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
Huete, A., Justice, C., & Liu, H. (1994). "Development of vegetation and soil indices for MODIS-EOS". *Remote Sensing of Environment, **49**(3)*, 224-234
Huete, A.R. (1988). "A soil-adjusted vegetation index (SAVI)". *Remote Sensing of Environment, **25**(3)*, 295-309
Huete, A.R., Jackson, R.D., & Post, D.F. (1985). "Spectral response of a plant canopy with different soil backgrounds." *Remote Sensing of Environment, **17**(1)*, 37-53
Jacquemoud, S. (1993). "Inversion of the PROSPECT+SAIL canopy reflectance model from AVIRIS equivalent spectra: theoretical study". *Remote Sensing of Environment, **44**(2-3)*, 281-292
Jacquemoud, S. (2006). personal communication
Jacquemoud, S., Bacour, C., Poilvé, H., & Frangi, J.-P. (2000). "Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode". *Remote Sensing of Environment, **74**(3)*, 471-481
Jacquemoud, S., & Baret, F. (1990). "PROSPECT: A model of leaf optical properties spectra". *Remote Sensing of Environment, **34**(2)*, 75-91
Jacquemoud, S., Ustin, S.L., Verdebout, J., Schmuck, G., Andreoli, G., & Hosgood, B. (1996). "Estimating leaf biochemistry using the PROSPECT leaf optical properties model". *Remote Sensing of Environment, **56**(3)*, 194-202
Jang, J.-D., Viau, A.A., Anctil, F., & Bartholomé, E. (2006). "Neural network application for cloud detection in SPOT VEGETATION images". *International Journal of Remote Sensing, **27**(4)*, 719-736
Janssen, B.H., & Beusichem, M.L.v. (2000). *Syllabus Bodem-plant relaties*. Wageningen, The Netherlands: Wageningen University
Ji, L., & Peters, A.J. (2007). "Performance evaluation of spectral vegetation indices using a statistical sensitivity function". *Remote Sensing of Environment, **106**(1)*, 59-65
Jordan, C.F. (1969). "Derivation of Leaf-Area Index from Quality of Light on the Forest Floor". *Ecology, **50**(4)*, 663-666
Junta de Andalucía - Consejería de agricultura y pesca (2005). "http://www.juntadeandalucia.es/agriculturaypesca/portal/opencms/portal/DGPAgraria/Estadisticas/estadisticasagrarias?entrada=servicios&servicio=201"
Jupp, D.L.B., & Strahler, A.H. (1991). "A hotspot model for leaf canopies". *Remote Sensing of Environment, **38**(3)*, 193-210
Kanellopoulos, I., & Wilkinson, G.G. (1997). "Strategies and best practice for neural network image classification". *International Journal of Remote Sensing, **18**(4)*, 711-725
Kaufman, Y.J., & Tanre, D. (1992). "Atmospherically resistant vegetation index (ARVI) for EOS-MODIS". *IEEE Transactions on Geoscience and Remote Sensing, **30**(2)*, 261-270
Kimes, D., Gastellu-Etchegorry, J., & Estève, P. (2002). "Recovery of forest canopy characteristics through inversion of a complex 3D model". *Remote Sensing of Environment, **79**(2-3)*, 320-328
King, R.L. (1999). SAIL. In http://www.ece.msstate.edu/~rking/downloads/sail.pdf (Ed.). accessed 24/09/2006
Kirkpatrick, S., Gelatt Jr., C.D., & Vecchi, M.P. (1983). "Optimization by simulated annealing". *Science, **220**(4598)*, 671-680
Kuusk, A. (1995). "A Markov chain model of canopy reflectance". *Agricultural and Forest Meteorology, **76**(3-4)*, 221-236
Lambers, H., Chapin, F.S.I., & Pons, T.L. (1998). *Plant physiological ecology*. New York - Berlin - Heidelberg: Springer-Verlag

Larcher, W. (2003). *Physiological plant ecology: Ecophysiology and stress physiology of functional groups*. Berlin - Heidelberg - New York: Springer-Verlag

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

Levizou, E., Drilias, P., Psaras, G.K., & Manetas, Y. (2005). "Nondestructive assessment of leaf chemistry and physiology through spectral reflectance measurements may be misleading when changes in trichome density co-occur". *New Phytologist, **165**(2)*, 463-472

Lillesand, T.M., & Kiefer, R.W. (1999). *Remote sensing and image interpretation*. New York, U.S.A.: John Wiley & Sons, Inc.

Malenovský, Z. (2007). personal communication

Malenovský, Z., Ufer, C., Lhotáhová, Z., Clevers, J.G.P.W., Schaepman, M.E., Albrechtová, J., & Cudlín, P. (2006). "A new hyperspectral index for chlorophyll estimation of a forest canopy: area under curve normalised to maximal band depth between 650-725 nm". *EARSeL eProceedings, **5**(2)*, 161-172

Myneni, R.B., Maggion, S., Iaquinta, J., Privette, J.L., Gobron, N., Pinty, B., Kimes, D.S., Verstraete, M.M., & Williams, D.L. (1995). "Optical remote sensing of vegetation: Modeling, caveats, and algorithms". *Remote Sensing of Environment, **51**(1)*, 169-188

Nemmour, H., & Chibani, Y. (2006). "Fuzzy neural network architecture for change detection in remotely sensed imagery". *International Journal of Remote Sensing, **27**(4)*, 705-717

North, P.R.J. (1996). "Three-dimensional forest light interaction model using a Monte Carlo method". *IEEE Transactions on Geoscience and Remote Sensing, **34**(4)*, 946-956

O'Neill, N.T., Zagolski, F., Bergeron, M., Royer, A., Miller, J.R., & Freemantle, J. (1997). "Atmospheric correction validation of casi images acquired over the boreas southern study area". *Canadian Journal of Remote Sensing, **23**(2)*, 143-161

Pal, S.K., & Mitra, S. (1999). *Neuro-fuzzy pattern recognition*: John Wiley & Sons

Peñuelas, J., Baret, F., & Filella, I. (1995). "Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance". *Photosynthetica, **31**(2)*, 221-230

Peñuelas, J., Gamon, J.A., Fredeen, A.L., Merino, J., & Field, C.B. (1994). "Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves". *Remote Sensing of Environment, **48**(2)*, 135-146

Pinty, B., Gobron, N., Widlowski, J.-L., Gerstl, S.A.W., Verstraete, M.M., Antunes, M., Bacour, C., Gascon, F., Gastellu, J.-P., Goel, N., Jacquemoud, S., North, P., Qin, W., & Thompson, R. (2001). "Radiation transfer model intercomparison (RAMI) exercise". *Journal of Geophysical Research D: Atmospheres, **106**(D11)*, 11937-11956

Pinty, B., Widlowski, J.-L., Taberner, M., Gobron, N., Verstraete, M.M., Disney, M., Gascon, F., Gastellu, J.-P., Jiang, L., Kuusk, A., Lewis, P., Li, X., Ni-Meister, W., Nilson, T., North, P., Qin, W., Su, L., Tang, S., Thompson, R., Verhoef, W., Wang, H., Wang, J., Yan, G., & Zang, H. (2004). "Radiation Transfer Model Intercomparison (RAMI) exercise: Results from the second phase". *Journal of Geophysical Research D: Atmospheres, **109**(6)*, D06210 06211-06219

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., & Sorooshian, S. (1994). "A modified soil adjusted vegetation index". *Remote Sensing of Environment, **48**(2)*, 119-126

Read, C., Wright, I.J., & Westoby, M. (2006). "Scaling-up from leaf to canopy-aggregate properties in sclerophyll shrub species". *Austral Ecology, **31**(3)*, 310-316

Rondeaux, G., Steven, M., & Baret, F. (1996). "Optimization of soil-adjusted vegetation indices". *Remote Sensing of Environment, **55**(2)*, 95-107

Schaepman-Strub, G. (2004). "Spectrodirectional reflectance analysis and definition for the estimation of vegetation variables". PhD thesis (128 pages), Department of Geography, University of Zurich, Switzerland

Schlemmer, M.R., Francis, D.D., Shanahan, J.F., & Schepers, J.S. (2005). "Remotely measuring chlorophyll content in corn leaves with differing nitrogen levels and relative water content". *Agronomy Journal, **97**(1)*, 106-112

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

Schultz, A., & Wieland, R. (1997). "The use of neural networks in agroecological modelling". *Computers and Electronics in Agriculture, **18**(2-3)*, 73-90

Smith, J.A., Oliver, R.E., & Berry, J.K. (1977). "A comparison of two photographic techniques for estimating foliage angle distribution". *Australian Journal of Botany, **25**(5)*, 545-553

Stenberg, P., Rautiainen, M., Manninen, T., Voipio, P., & Smolander, H. (2004). "Reduced simple ratio better than NDVI for estimating LAI in Finnish pine and spruce stands". *Silva Fennica, **38**(1)*, 3-14

Stone, C., Chisholm, L., & Coops, N. (2001). "Spectral reflectance characteristics of eucalypt foliage damaged by insects". *Australian Journal of Botany, **49**(6)*, 687-698

Tan, B., Hu, J., Huang, D., Yang, W., Zhang, P., Shabanov, N.V., Knyazikhin, Y., Nemani, R.R., & Myneni, R.B. (2005). "Assessment of the broadleaf crops leaf area index product from the Terra MODIS instrument". *Agricultural and Forest Meteorology, **135**(1-4)*, 124-134

Thenkabail, P.S., Smith, R.B., & De Pauw, E. (2000). "Hyperspectral vegetation indices and their relationships with agricultural crop characteristics". *Remote Sensing of Environment, **71**(2)*, 158-182

Tian, Y., Wang, Y., Zhang, Y., Knyazikhin, Y., Bogaert, J., & Myneni, R.B. (2003). "Radiative transfer based scaling of LAI retrievals from reflectance data of different resolutions". *Remote Sensing of Environment, **84**(1)*, 143-159

Veefkind, J.P. (1999). "Aerosol satellite remote sensing". PhD thesis (131 pages pages), Faculty of Physics and Astronomics, University of Utrecht

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

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

Wellburn, A.R. (1994). "The spectral determination of chlorophylls a and b, as well as the total carotenoids using various solvents with spectrophotometers of different resolutions". *Journal of Plant Physiology, **144**(3)*, 307-313

Wiener, N., & Rosenblueth, A. (1945). Philosophy of science

Wikipedia (2006a). Chlorophyll (http://en.wikipedia.org/wiki/Chlorophyll). In

Wikipedia (2006b). Genetic algorithm (http://en.wikipedia.org/wiki/Genetic_algorithm). In

Zarco-Tejada, P.J., Berjón, A., López-Lozano, R., Miller, J.R., Martín, P., Cachorro, V., González, M.R., & De Frutos, A. (2005). "Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy". *Remote Sensing of Environment, **99**(3)*, 271-287

Zarco-Tejada, P.J., Miller, J.R., Harron, J., Hu, B., Noland, T.L., Goel, N., Mohammed, G.H., & Sampson, P. (2004a). "Needle chlorophyll content estimation through model inversion using hyperspectral data from boreal conifer forest canopies". *Remote Sensing of Environment, **89**(2)*, 189-199

Zarco-Tejada, P.J., Miller, J.R., Morales, A., Berjón, A., & Agüera, J. (2004b). "Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops". *Remote Sensing of Environment, **90**(4)*, 463-476

Zarco-Tejada, P.J., Miller, J.R., Noland, T.L., Mohammed, G.H., & Sampson, P.H. (2001). "Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data". *IEEE Transactions on Geoscience and Remote Sensing, **39**(7)*, 1491-1507

Zhang, Q., Xiao, X., Braswell, B., Linder, E., Baret, F., & Moore III, B. (2005). "Estimating light absorption by chlorophyll, leaf and canopy in a deciduous broadleaf forest using MODIS data and a radiative transfer model". *Remote Sensing of Environment, **99**(3)*, 357-371

Zhao, D., Reddy, K.R., Kakani, V.G., & Reddy, V.R. (2005). "Nitrogen deficiency effects on plant growth, leaf photosynthesis, and hyperspectral reflectance properties of sorghum". *European Journal of Agronomy, **22**(4)*, 391-403

9 Acronyms, abbreviations and symbols

θ_S	solar zenith
θ_v	viewing azimuth
Φ_s	solar azimuth
Φ_v	viewing azimuth
3-D	three-dimensional
A_{SCENE}	surface (area) of the scene
B#	band number
C_{AB}	chlorophyll a + b content [$\mu g/cm^2$]
CASI	Compact Airborne Spectrographic Imager
C_M	dry matter content [g/cm^2]
CV	canopy (biophysical/biochemical) variables
C_W	water content [g/cm^2]
DART	discrete anisotropic radiative transfer model
E_{XY}	Radius of the tree (horizontal plane)
E_Z	Radius of the tree crown (vertical plane)
f_{COVER}	fractional coverage (ground cover)
FLIGHT	Forest LIGHT interaction model
g	gram
GA	genetic algorithm(s)
h	height of the tree
HCRF	hemispherical conical reflectance factor
HDRF	hemispherical directional reflectance factor
LAD	leaf angle distribution (function)
LAI	leaf area index
LUT	look-up table(s)
m	meter
μ	micro (10^{-6}) or mean
MERIS	MEdium Resolution Imaging Spectrometer
MODIS	MODerate resolution Imaging Spectroradiometer
n	nano (10^{-9})
N	nitrogen or leaf structural parameter N
NIR	near infrared region of the electromagnetic spectrum (700 to 1300 nm)
NN	neural network(s)
LAI	leaf area index
PSF	point spread function
ϱ	reflectance
R_{xxx}	Reflectance in band or wavelength *xxx*
RS	remote sensing
RTM	radiative transfer model(s)
σ	standard deviation
SWIR	short wave infrared (1300 to 2400 nm)
τ	transmittance
TIR	thermal infrared
VI	vegetation index/indices
VIS	visible region of the electromagnetic spectrum (400 nm to 700 nm)

Appendix I Alterations of the source code of Flight version 5.0

In this appendix we have listed the lines of code that were altered to make the Flight model run correctly.

Original lines 669:673:

```
while ((WLIST[wno][i]>=0) && (NOSTANDS>0) && (res<=0)) {
        res=inside_stand(photon,WLIST[wno][i]);
                i++;
        }
}
```

Replaced line 669 with:

```
while ((WLIST[wno][i]>=0) && (NOSTANDS>0) && (res<=0) && (i<NOSTANDS))
{
```

Justification: i was increasing unlimited and accessing illegal (non-existing) stand numbers, resulting in segmentation faults in line 2264 (access beyond array stands inside the function inside_stand with stands[stand_number]).

Original lines 2431:2438:

```
while ((inflag==0) && (NOSTANDS>0) && (WLIST[wno][i]>=0) ) {
                        if (inside_stand(photon,WLIST[wno][i])>0)
                        {
                                stand_number=WLIST[wno][i];
                                inflag=1;
                        }
                        i++;
}
```

Replaced line 2431 with:

```
while ((inflag==0) && (NOSTANDS>0) && (WLIST[wno][i]>=0) &&
(i<NOSTANDS) ) {
```

Justification: As with the previous error, i was increasing unlimited and accessing illegal (non-existing) stand numbers, resulting in segmentation faults in line 2264 (access beyond array stands inside the function inside_stand with stands[stand_number]).

Original line 2727:

```
for (theta_o=0;theta_o< PYVAL ;theta_o=theta_o++) {
```

Replaced with:

```
for (theta_o=0;theta_o< PYVAL ; theta_o++) {
```

Justification: theta_0= was obsolete and introduced loop problems with some compilers.

Original line 5245:

```
if ((MODE!='s') && (FRAC_COV>0.0) && (TOTAL_LAI>0.0)) {
```

Replaced with:

```
if ((MODE!='s') && (FRAC_COV>0.0) && ((TOTAL_LAI>0.0) || (FIELD_DATA
==1))) {
```

Justification: The hot spot array was not read and the green vegetation was ignored if the LAI was specified independently per tree (TOTAL_LAI<0).

Line inserted at 5601:

```
fflush(fplog);fclose(fplog);
```

Justification: The log files of the direct and diffuse radiation were not written and closed. This caused a segmentation fault upon re-use of fplog and the two log files were empty.

Appendix II Relating VI with reflectance at the leaf level

We have also conducted an analysis of the estimation capabilities of VI for C_{AB}. Simulations were done with the PROSPECT leaf model with the modelled reflectance resampled to match the CASI spectral configurations. In the first step, only C_{AB} was varied. Secondly, 5% noise was added to these simulations. In the third step, C_{AB}, C_M and N were all three varied. Subsequently, 5% noise was added to these reflectances. For the four sets we calculated the VI from the reflectances. These VI were related to the input C_{AB}. The spectra from the 30 leaves that were inverted in section 3.4.1.1 were resampled to be able to calculate the VI. The original leaf spectra were inverted using the PROSPECT model to estimate the C_{AB} level of each leaf by minimising the RMSE between the modelled and measured leaf reflectance and transmittance. Subsequently, the relations that were derived from the simulations were applied to the VI values derived for the leaves. These estimated C_{AB} contents were finally compared with the ones retrieved from the inversion of PROSPECT. The standard deviation equals 24.53 $\mu g/cm^2$ for the full dataset (3 free variables) and 25.25 $\mu g/cm^2$ for the simple dataset (2 fixed variables, only C_{AB} free). For the 30 leaves, the mean estimated C_{AB} content is 62.40 $\mu g/cm^2$ with a standard deviation of 28.17 $\mu g/cm^2$.

At the leaf level, RVI, GM_1, GM_2 and MTCI were the $VI_{SPATIAL}$ having the best relation with the chlorophyll content. Best results were obtained for the set where noise had been added to the spectra simulated with variation in C_{AB} only. In the simple dataset with 'known' C_M and N, we can observe that the TVI and MTVI1 have a bad performance; they do not appear to be sensitive to changes in chlorophyll. If we add variation in C_M and N, we see that G, VARI and MCARI are now loosing their abilities to observe changes in C_{AB}. We also see that even though MCARI/OSAVI and TCARI/OSAVI have a reasonably good R^2, the RMSE is very high. Finally, when we add 5% noise to the input spectra, the overall performance of the indices is lowered as expected, but only moderately.

For $VI_{SPECTRAL}$, the best index was GM_1 that was trained with variation in C_{AB} + 5% noise. The other well-performing indices (MTCI, GM_2 and MTCI2) were trained with variation in C_{AB} only. RMSE was around 2.9-3.4 $\mu g/cm^2$ for the estimation of the 30 leaves. MCARI/OSAVI and TCARI/OSAVI had a good coefficient of correlation for the fit-line, but a high RMSE and have therefore been taken out of the 'top five'.

Appendix III Obtained fit results for the individual fields

Table III-1: Fit results for CLO1.

μ: 71.2262 σ: 9.6525 min: 50.4060 max: 90.8650 n: 48

Index	*Mask*	*Fitting model*	R^2	*RMSE*
GM1	3	1	0.105	9.039
GM2	3	1	0.068	9.222
MCARI	2	8	0.133	8.890
TCARI	2	8	0.136	8.878
OSAVI	3	1	0.060	9.263
TCARI/OSAVI	2	8	0.153	8.790
MCARI/OSAVI	2	8	0.160	8.750
NDVI	3	4	0.043	9.345
TVI	3	3	0.052	9.262
G	2	3	0.053	9.658
RDVI	3	3	0.070	9.222
MTVI1	3	6	0.047	9.234
MTCI	4	4	0.120	8.961
MERIS red edge	2	8	0.107	9.540
CASI red edge	2	8	0.107	9.540
VARI	2	3	0.052	9.632
RVI	3	1	0.041	9.356

Table III-2: Fit results for CLO3.

μ: 56.5154 σ: 6.0689 min: 44.4011 max: 69.9001 n: 40

Index	*Mask*	*Fitting model*	R^2	*RMSE*
GM1	4	1	0.033	5.892
GM2	3	1	0.088	5.722
MCARI	3	8	0.216	5.307
TCARI	2	8	0.2155	5.321
OSAVI	1	5	0.033	6.240
TCARI/OSAVI	2	8	0.255	5.173
MCARI/OSAVI	2	8	0.279	5.090
NDVI	1	5	0.037	6.184
TVI	4	3	0.119	5.625
G	2	8	0.090	5.715
RDVI	4	3	0.077	5.777
MTVI1	4	3	0.115	5.639
MTCI	2	3	0.357	4.806
MERIS red edge	2	7	0.262	5.974
CASI red edge	2	7	0.262	5.795
VARI	2	8	0.091	5.714
RVI	1	6	0.034	5.891

Table III-3: Fit results for CLO4.

μ: 78.2504 σ: 4.9789 min: 65.9451 max: 86.3960 n: 39

Index	*Mask*	*Fitting model*	R^2	*RMSE*
GM1	2	1	0.047	4.799
GM2	2	1	0.083	4.707
MCARI	1	8	0.025	4.913

TCARI	1	8	0.011	4.888
OSAVI	2	5	0.092	4.683
TCARI/OSAVI	1	5	0.036	4.995
MCARI/OSAVI	1	5	0.084	4.981
NDVI	2	1	0.081	4.712
TVI	2	5	0.142	21.656
G	2	3	0.066	4.749
RDVI	2	5	0.097	4.671
MTVI1	4	6	0.093	4.680
MTCI	3	4	0.201	4.396
MERIS red edge	3	8	0.182	4.445
CASI red edge	3	8	0.182	4.445
VARI	2	6	0.068	4.744
RVI	2	4	0.068	4.745

Appendix IV Chlorophyll contents of CLO3 and CLO4 estimated by MCARI/OSAVI

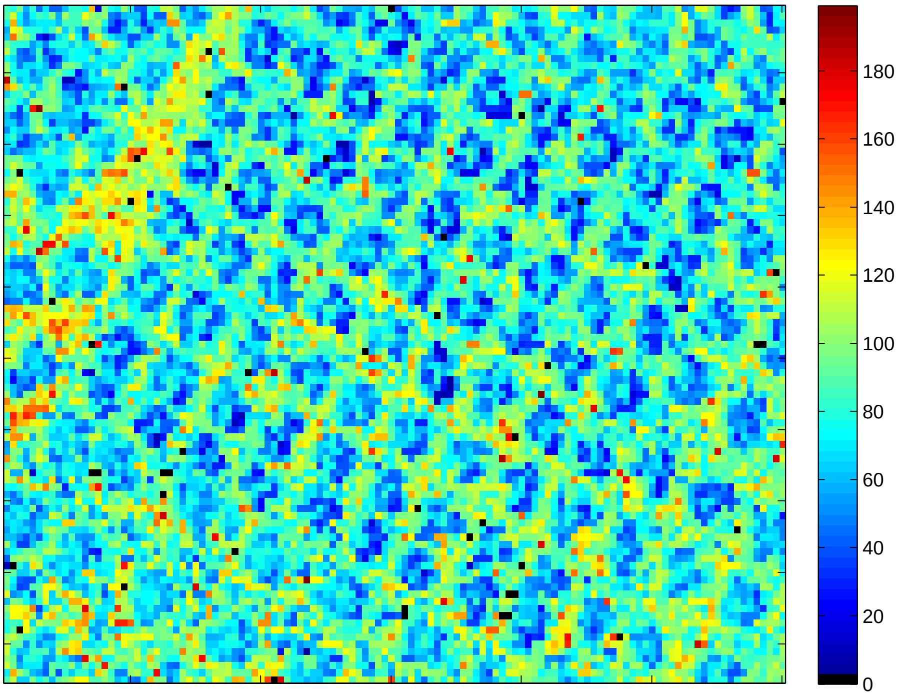

Figure IV-1: C_{AB} estimates for CLO3 by MCARI/OSAVI

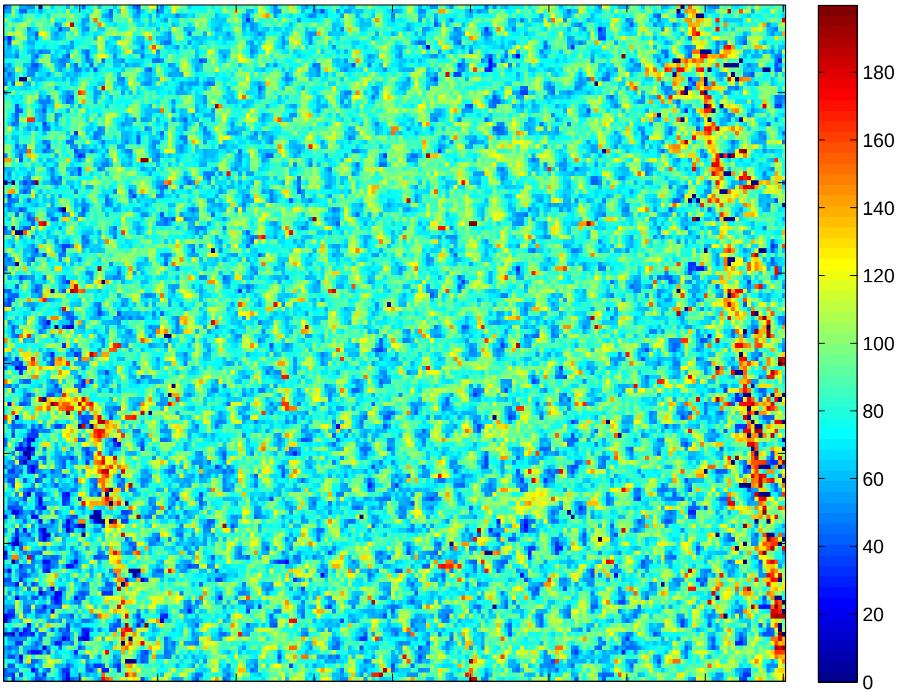

Figure IV-2: C_{AB} estimates for CLO4 by MCARI/OSAVI