Remote Sensing of Environment 112 (2008) 3846–3864

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass

Guerric le Maire [a,b,*], Christophe François [a], Kamel Soudani [a], Daniel Berveiller [a], Jean-Yves Pontailler [a], Nathalie Bréda [c], Hélène Genet [c], Hendrik Davi [d], Eric Dufrêne [a]

[a] Laboratoire Ecologie Systématique et Evolution, UMR CNRS - Univ Paris Sud - AgroParisTech, 91405 Orsay, France
[b] CIRAD, UR 80 Fonctionnement et pilotage des écosystèmes de plantations, Montpellier, F-34398, France
[c] Ecologie et Ecophysiologie Forestières, UMR INRA-UHP 54820 Champenoux, France
[d] Ecologie des Forêts Méditerranéennes, UR 629, INRA, 84914 Avignon, France

ARTICLE INFO

Article history:
Received 12 October 2007
Received in revised form 6 June 2008
Accepted 6 June 2008

Keywords:
Chlorophyll
LMA
SLA
Leaf biomass
EO1 Hyperion
ASD Fieldspec
LAI
PROSPECT
SAIL
PROSAIL

ABSTRACT

This article aims at finding efficient hyperspectral indices for the estimation of forest sun leaf chlorophyll content (CHL, µg cm^{-2}_{leaf}), sun leaf mass per area (LMA, $g_{dry\ matter}$ m^{-2}_{leaf}), canopy leaf area index (LAI, m^2_{leaf} m^{-2}_{soil}) and leaf canopy biomass (B_{leaf}, $g_{dry\ matter}$ m^{-2}_{soil}). These parameters are useful inputs for forest ecosystem simulations at landscape scale. The method is based on the determination of the best vegetation indices (index form and wavelengths) using the radiative transfer model PROSAIL (formed by the newly-calibrated leaf reflectance model PROSPECT coupled with the multi-layer version of the canopy radiative transfer model SAIL). The results are tested on experimental measurements at both leaf and canopy scales. At the leaf scale, it is possible to estimate CHL with high precision using a two wavelength vegetation index after a simulation based calibration. At the leaf scale, the LMA is more difficult to estimate with indices. At the canopy scale, efficient indices were determined on a generic simulated database to estimate CHL, LMA, LAI and Bleaf in a general way. These indices were then applied to two Hyperion images (50 plots) on the Fontainebleau and Fougères forests and portable spectroradiometer measurements. They showed good results with an RMSE of 8.2 µg cm^{-2} for CHL, 9.1 g m^{-2} for LMA, 1.7 m^2 m^{-2} for LAI and 50.6 g m^{-2} for Bleaf. However, at the canopy scale, even if the wavelengths of the calibrated indices were accurately determined with the simulated database, the regressions between the indices and the biophysical characteristics still had to be calibrated on measurements. At the canopy scale, the best indices were: for leaf chlorophyll content: $ND_{chl}=(\rho_{925}-\rho_{710})/(\rho_{925}+\rho_{710})$, for leaf mass per area: $ND_{LMA}=(\rho_{2260}-\rho_{1490})/(\rho_{2260}+\rho_{1490})$, for leaf area index: $D_{LAI}=\rho_{1725}-\rho_{970}$, and for canopy leaf biomass: $ND_{Bleaf}=(\rho_{2160}-\rho_{1540})/(\rho_{2160}+\rho_{1540})$.

1. Introduction

Forest ecosystems are well-studied at the stand scale. However, in order to better understand their functioning and response to environmental changes, it is necessary to up-scale this knowledge to the scale of the entire forest or small region (Landsberg, 2003; Makela et al., 2000). One way to reach this objective is to use ecosystem models that are validated with local-scale observations and applied to larger areas. For a large scale simulation, a selection of spatially-parameterized input parameters is necessary. The selection of the main spatial parameters should meet the following criteria (le Maire et al., 2005):

(i) to be a parameter to whish the model is sensitive,
(ii) to be spatially variable at the scale of interest (for instance between stands), and to have a larger variability at this scale than at finer scale (e.g., inter-stand vs. intra-stand variability),
(ii) to have a non-linear model response: this strengthens the need for spatialization of the parameter if the simulation results are averaged.

A study with a particular forest process-based ecosystem model has shown that a number of parameters are sensitive in this model (Dufrêne et al., 2005). Many of these parameters are spatially variable between stands, some of them having a non-linear response (Davi

* Corresponding author. Laboratoire Ecologie Systématique et Evolution, UMR CNRS - Univ Paris Sud - AgroParisTech, 91405 Orsay, France.
E-mail address: guerric.le_maire@cirad.fr (G. le Maire).

0034-4257/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2008.06.005

et al., 2006). Among these parameters, in addition to soil parameters driving the soil water budget, the following vegetation parameters were identified:

- leaf nitrogen content (N_{leaf}), which is directly involved in the photosynthesis calculation. Experimental measurements have shown that this parameter is highly correlated with leaf chlorophyll content (CHL, µg cm^{-2}) for sun leaves (see Section 4.7),
- annual maximum leaf mass per area (LMA) of sun leaves (leaves of the canopy that are not shadowed by other leaves), a parameter that enables conversion of leaf area to leaf biomass (B_{leaf}), which is used in many processes in the model,
- annual maximum leaf area index (LAI), which drives many processes like radiation interception, canopy photosynthesis and litter amount.

The objective of the present study is to assess the possibility of estimating essential parameters of forest ecosystem models (LMA, LAI, CHL and B_{leaf}) using hyperspectral satellite images on large areas, and to estimate the obtained accuracy.

The use of indices on hyperspectral images has two major advantages. First, the most informative wavelengths of the 400–2500 nm region can be selected. Second, it allows the use of a narrow spectrum feature necessary for assessing vegetation biochemical properties (Broge & Mortensen, 2002). Many studies have shown that hyperspectral measurements can be used to quantify biophysical characteristics of the vegetation at leaf scale (Gitelson et al., 2003; le Maire et al., 2004; Zhao et al., 2005) or at canopy scale using in situ data, airborne sensors like AVIRIS, CASI and HyMap, or spaceborne sensors like Hyperion and CHRIS.

Different methods exist to retrieve canopy characteristics from reflectance measurements (Blackburn, 2007; Kimes et al., 2000; Weiss & Baret, 1999):

(i) *Indices and/or multiple regressions*: the principle is to combine several reflectances measured on narrow or large spectral bands into mathematical combinations and to correlate them to a particular characteristic of the observed surface. These relationships are calibrated based on an experimental or simulated reflectance database (built up on radiative transfer models). These methods are simple, but have some limits: when calibrated to an experimental database, the representativeness of the relationships is limited to the representativeness of the database. Moreover, indices and multiple regressions may be sensitive to more than one single characteristic. They are also sensitive to atmospheric conditions, view geometry, and spatial resolution, and therefore they must usually be calibrated for each image.

(ii) *Model inversions*: this method uses models that simulate reflectance spectra from canopy and soil characteristics. As noted by Bacour et al. (2006), inversion techniques based on pre-computed reflectance database are often preferred to more computationally heavy iterative methods for operational applications. Among the computationally efficient methods often used are Look-Up Tables (e.g. Knyazikhin et al., 1998) and Neural Networks (e.g. Bacour et al., 2006; Baret et al., 2007). Both methods are dependent on the simulated training database. Inversion of such models often gives a large number of different possible solutions. Moreover, uncertainties in measurements and models may result in large variation in results (Combal et al., 2003).

The best way to find efficient indices would be to use a large measurement database, with many images and canopy conditions. Such a large database with hundreds of measurements is feasible at the leaf scale but is not conceivable at the forest scale. Moreover, indices calibrated on a particular forest canopy database could be unsuitable in other forests.

This issue leads us to create a large synthetic database containing reflectance spectra and their corresponding canopy characteristics. Such a database has many advantages: many canopy characteristics are represented (thousands of spectra); the influence of each characteristic can be totally decoupled from that of others; and the effect of a particular characteristic on the spectra is based on physical processes that are modeled at a small scale. Therefore, well established indices obtained on such a large simulated database may potentially be applied to a wide range of spectra. However, the use of a model relies on its capacity to correctly simulate the reflectance of a wide range of canopies. Thus, it is essential to test these indices on experimental measurements. The representativeness of the simulated database is therefore critical.

In this study, we generate two simulated databases, one at leaf scale with the PROSPECT model, and one at canopy scale coupling the PROSPECT leaf model with the SAIL canopy radiative transfer model (PROSAIL). At leaf scale, we continue the study done in le Maire et al. (2004) using an improved and newly calibrated version of the PROSPECT model (Feret et al., 2008), and a larger experimental database. The work at this scale is a first step to interpret the results at the canopy scale and explain possible discrepancies. At canopy scale, we use a multi-layer version of the SAIL model (Weiss et al., 2001), which is able to represent the vertical LMA profile. The study is restricted to canopies with LAI greater than 3 to correspond with the big-leaf representation of SAIL. These simulated databases are used to find best indices of CHL and LMA at leaf scale, and CHL, LMA, LAI and B_{leaf} at canopy scale.

Results are tested against measurements at both scales. At leaf scale we used a large database of 246 spectra and 49 species. At canopy scales, experimental measurements consist of ground measurements on small and mature canopies with a portable spectroradiometer, and hyperspectral images for two distinct forests measured with the Hyperion satellite.

We first describe the PROSPECT and SAIL models, simulated databases and the determination method of best spectral indices. Then, we present the experimental protocols for the measurements (leaf reflectance measurements, in situ measurements and satellite remote sensing data). The results are given at leaf and canopy scale for the determination of CHL, LMA, LAI and leaf biomass.

2. Model description, simulated databases and best indices determination method

2.1. The Leaf reflectance model PROSPECT

An improved (1-nm resolution) and recalibrated version of the leaf reflectance model PROSPECT has been used in this study (Feret et al., 2008). The PROSPECT model (Jacquemoud & Baret, 1990; Jacquemoud et al., 1996) considers the leaf as a succession of absorbing layers. The new version calculates the leaf hemispherical reflectance and transmittance between 400 and 2500 nm with a 1-nm step as a function of leaf structure index (N_{struc}), leaf chlorophyll content (CHL, µg/cm^2), leaf water content (Cw, g/cm^2), and leaf mass area (LMA, g/m^2).

2.2. Multi-layer PROSAIL model description

The SAIL radiative tranfer model is a turbid medium model. It describes the canopy as horizontally homogeneous, where leaves absorb, reflect, and transmit radiation (Verhoef, 1984). This model has been validated by many studies on numerous vegetation types (e.g. Andrieu et al., 1997; Goel & Thomson, 1984; Major et al., 1992). The radiative transfer equation is solved using an *n*-flux approximation. The radiation is considered as four fluxes: diffuse, direct, upward and downward (Kubelka & Munk, 1931; Suits, 1972). The system is described as four differential equations for the four fluxes. The solar