Remote Sensing

Article

Optimal Exploitation of the Sentinel-2 Spectral Capabilities for Crop Leaf Area Index Mapping

Katja Richter [1],*, Tobias B. Hank [1], Francesco Vuolo [2], Wolfram Mauser [1] and Guido D'Urso [3]

[1] Department of Geography, Ludwig-Maximilians University Munich, Luisenstr. 37, D-80333 Munich, Germany; E-Mails: tobias.hank@lmu.de (T.H.); w.mauser@lmu.de (W.M.)

[2] Institute of Surveying, Remote Sensing and Land Information (IVFL), University of Natural Resources and Life Sciences (BOKU), Peter Jordan Str. 82, A-1190 Vienna, Austria; E-Mail: francesco.vuolo@boku.ac.at

[3] Department of Agricultural Engineering and Agronomy, University of Naples Federico II, Via Università 100, I-80055 Portici, Italy; E-Mail: durso@unina.it

* Author to whom correspondence should be addressed; E-Mail: k.richter@iggf.geo.uni-muenchen.de; Tel.:+49-89-2180-6684.

Abstract: The continuously increasing demand of accurate quantitative high quality information on land surface properties will be faced by a new generation of environmental Earth observation (EO) missions. One current example, associated with a high potential to contribute to those demands, is the multi-spectral ESA Sentinel-2 (S2) system. The present study focuses on the evaluation of spectral information content needed for crop leaf area index (LAI) mapping in view of the future sensors. Data from a field campaign were used to determine the optimal spectral sampling from available S2 bands applying inversion of a radiative transfer model (PROSAIL) with look-up table (LUT) and artificial neural network (ANN) approaches. Overall LAI estimation performance of the proposed LUT approach (LUT_{N50}) was comparable in terms of retrieval performances with a tested and approved ANN method. Employing seven- and eight-band combinations, the LUT_{N50} approach obtained LAI RMSE of 0.53 and normalized LAI RMSE of 0.12, which was comparable to the results of the ANN. However, the LUT_{N50} method showed a higher robustness and insensitivity to different band settings. Most frequently selected wavebands were located in near infrared and red edge spectral regions. In conclusion, our results emphasize the potential benefits of the Sentinel-2 mission for agricultural applications.

1. Introduction

The increasing demand of accurate quantitative information on land surface properties continues to drive the design and launch of innovative Earth observation (EO) missions. High-quality data delivered from new sensors offer the unique opportunity to continue and improve the derivation of biophysical variables, such as leaf area index (LAI), vegetation cover fraction (fCover) or fraction of photosynthetically-active radiation (fAPAR), being drivers of several important physiological key processes, including evapotranspiration and photosynthesis. These variables describe the spatial distribution of vegetation state and dynamics and therefore provide essential input for a wide range of ecological models in numerous fields of application and research [1,2]. This includes, for instance, the monitoring of forest dynamics [3], climate modeling or the assessment of the carbon and nutrient cycle from regional to global scales [4]. One of the main application fields is the agricultural sector, where biophysical variables, such as LAI, are needed, among others, to support the development of precision farming techniques. This anticipates their implementation into different models for the simulation of crop growth and variability, nutrient demand or irrigation water requirements [5].

The present study was conducted against the background of the future ESA Sentinel-2 (S2) multispectral satellite mission [6], which is embedded in the framework of the Global Monitoring for Environment and Security (GMES). The Sentinel mission's objective is to provide continuity to services depending on current multi-spectral high-resolution optical observations over global terrestrial surfaces, such as the adequate quantification of geo-biophysical variables and the mapping of land-cover/land-change detection [7].

The first from two S2 sensors is scheduled to be launched in the year 2013. Spectral sampling is inherited from sensors that have been used for vegetation monitoring in the last decades, such as SPOT, Landsat and MODIS. The pair of Sentinel-2 satellites will be equipped with visible, near infrared and shortwave infrared sensors. In total, the sensors will have 13 spectral bands (with central wavebands described in the method section), listed as Band 1–8, 8a and 9–12 in [7]. With a spatial resolution of 10 m (4 bands), 20 m (6 bands) and 60 m (3 bands), the new sensors will address high to medium resolution applications. The short revisit time of five days (at the equator under cloud-free conditions) or even 2–3 days (at mid-latitudes) will allow the effective monitoring of vegetation status and dynamics. More detailed information and technical characterization of the mission can be found in the mission requirements document [7] and on the ESA Sentinel-2 website.

Various methods have been proposed, applied and improved in recent decades for the retrieval of biophysical products. Estimation approaches can be pooled into two groups [1,8]. The first group comprises statistical models, principally relying on the learning of the relation between the sought variable and the reflectance. Possibly the simplest statistical approach is represented by vegetation indices (VI), being among the oldest methods to gain information about vegetation characteristics from remote sensing data [9,10]. More complex chemometric and statistical techniques have been developed

to overcome the widely-discussed drawbacks of VIs [11–13]. These include, among others, partial least square regression (PLSR) [14], stepwise multiple linear regression (SMLR) [15], red-edge inflection point (REIP) [16], spectral unmixing (SUM) [17], artificial neural networks (ANN) [18] or kernel-based methods, such as support vector regression (SVR) [19].

The second group involves physically-based approaches, *i.e.*, radiative transfer models (RTM) in combination with different inversion strategies [20–22]. RTMs of different complexities have been developed, describing the interaction of radiation with vegetation, assuming the canopy as a simple one-dimensional (1-D) turbid medium up to more realistic three-dimensional (3-D) architectures [23]. The choice of the model depends on the kind of application, vegetation types monitored and the level of accuracy required.

Regarding the inversion of RTMs, different methodologies have been employed. The most traditional and classical approaches are iterative optimization techniques, e.g., [21,24]. Look-up tables (LUT) are also widely applied inversion strategies [24–26]. Other methodologies involve genetic algorithms [27], SVR [28] or Bayesian methods such as Monte-Carlo Markov Chains (MCMC) [29]. Some of the latter techniques must be regarded as hybrid approaches, since they combine statistical principles with the radiative transfer theory. Prominent representatives of these hybrid methods are artificial neural networks (ANNs) [30]. One of the main concerns for the retrieval of variables with these techniques is that model inversion usually does not correspond to the Hadamard's postulate of well-posedness [24], meaning that more than one unique solution to the problem is possible. Different methodologies have been studied to overcome this problem (for an overview see [1]), among others, the use of *a priori* information of the estimates [24] or increasing of data dimensionality in spectral, spatial or temporal terms. A very promising strategy, for instance, is the use of neighborhood information for regularization, as demonstrated by [31] using data based on Sentinel-2 sensors configuration.

An extensive discussion about different inversion techniques, the ill-posed nature of model inversion and related issues can be found in [8] or [32].

The definition of the S2 sensors spectral sampling was based on careful analyses of spectral regions of interest for geo-/biophysical vegetation variables, but also a range of other applications, constituting the best compromise in terms of user requirements and mission performance [6]. Several studies evaluated spectral regions or band settings of the S2 sensors for vegetation biophysical variable estimation applying empirical (e.g., [33]), or physically-based retrieval methods (e.g., [26]). Though these studies found good-to-optimal results, differences in retrieval accuracy for individual variables may be achieved when employing different retrieval methods with diverse band settings. For instance, some retrieval methods, such as vegetation indices, are limited to only a few spectral bands, which face serious shortcomings as discussed in [34]. On the other hand, physically-based retrieval methods, such as RTM inversion, could obtain better accuracies when using defined spectral sampling compared to the exploitation of the full spectral information available from a sensor [35,36]. This is explained by some studies [35,37] with (excessively) high noise levels in some bands of the spectral measurements or by limitations of the model to properly simulate those bands. The reasons for this effect may be even more multifaceted, depending on the used sensor, field measurement protocol, applied model and model configurations. In fact, the study from [35] showed that the use of spectral subsets does not improve the retrieval performance compared to full spectral resolution when the (full) measured

spectra are well presented by the model simulations. Nevertheless, several studies have revealed that few bands from multi- or hyperspectral data sources are sufficient for the characterization of vegetation and the discrimination between different surface properties, such as soil, water or snow [38,39]. Employing most possible spectral information would only increase the computation time without further improvement of the results. Therefore, to obtain the most accurate results of vegetation biophysical variables possible, together with reasonable computational demands, a defined spectral sampling should be preferred in many cases over the use of the full spectral information.

ESA Sentinel-2 sensors will be highly requested for operational agricultural applications. Therefore, the objective of the present study was to define the optimal use (number of bands and spectral regions) of available Sentinel-2 bands for crop LAI retrieval. For this purpose, data from a field campaign were exploited by means of RTM-based model inversions using look-up table and neural network algorithms.

2. Material and Methods

2.1. Test Site and Data Acquisition

The data set investigated in the present study was obtained in the framework of the ESA SPectra bARrax Campaigns (SPARC). These campaigns were carried out in July 2003 (therefore called here: SPARC'03) in Barrax, which is an agricultural test area situated within La Mancha region in Southern Spain (approx. 39°3′N, 2°6′W). Briefly, the SPARC'03 campaign aimed at supporting calibration and validation activities of existing algorithms and the development of new ones, such as for geo-/biophysical variable retrievals. Details can be found in [40].

Airborne hyperspectral data of the area were acquired by the HyMap imaging system [41] on 13 July 2003 around 11:20 UTC, with flight lines parallel to the principal plane (towards the sun). The sensor recorded spectral reflectance in 126 spectral channels with a ground sampling distance (GSD) of 5 m. Atmospheric correction and radiometric calibration were carried out by the Laboratory for Earth Observation of the University of Valencia using a modified MODTRAN4 code [42]. Ground LAI measurements were collected non-destructively by means of the LICOR LAI-2000 Plant Canopy Analyzer instrument [43]. In total, 70 LAI measurements of alfalfa, maize, sugar beet, garlic and onion, collected concurrently to the HyMap sensor overpass, were analyzed for the present study. A stratified random sampling strategy was applied with a minimum of 12 measurements per Elementary Sampling Unit (ESU). The mean value of these measurements represented the final LAI value for each ESU. Dimensions of the ESUs corresponded to 20 m × 20 m, being a compromise for the different spatial resolutions of the various remote sensing acquisitions during the SPARC'03 campaign. Detailed description of the measurements can be found in [44,45].

Clumping of the leaves was only partially regarded by the instruments and corresponding software. Moreover, no corrections were applied to account for the influence of non-green plant components, such as stems or senescent leaves. Thus, the term LAI used here for ground measurements corresponds to the effective plant area index (PAI_{eff}) [46,47]. The error arising from the ground LAI measurements can be up to 10% depending on the degree of crop heterogeneity. Moreover, other potential sources of uncertainties may originate from the (optical) instrument, such as illumination conditions, saturation effects, or instrument simplifications [48]. However, the vegetation surface apparent to a space- or

airborne remote sensing instrument corresponds rather to an "effective green area index," since leaf overlapping can lead to saturation of reflectance, in particular for higher LAI values [1]. Therefore, a correction of the clumping effect may not be explicitly necessary when comparing deviates of the LAI-2000 instrument and a remote sensor [49]. Measurement differences between the two instruments might be largest in the presence of non-green plant components, for instance during flowering or later crop growth stages. For a homogenous coverage—such as in a middle growth stage with mainly green plant components—differences between the two approaches may be rather marginal.

2.2. Relative Transfer (RT) Model and Inversion Procedures

The widespread PROSPECT-5 [50] and SAIL models [51] coupled in "PROSAIL" were chosen for the study. Comprehensive descriptions of the models already have been published [52]. Thus, only their main characteristics are briefly sketched here: the SAIL model simulates the bi-directional reflectance of homogeneous canopies as a function of soil reflectance, illumination and viewing geometries, several structural and biophysical variables, such as LAI, average leaf angle (ALA) and a hot spot parameter, implemented by [53]. Leaf optical properties (reflectance and transmittance) are simulated by the PROSPECT-5 model as a function of a structure parameter N, leaf chlorophyll content (C_{ab}), dry matter content (C_m), carotenoids (C_{ar}) and leaf water content (C_w).

For the estimation of the variables from the PROSAIL model, an adequate inversion procedure has to be defined. For this purpose, a look-up table was chosen as the main approach. LUTs belong to the most simple inversion strategies. However, they provide accurate results if an appropriate sampling of the canopy characteristics is realized [1,54]. By means of the LUT method, a global search of the best solution is performed, thus avoiding being trapped into local minima, as can occur with iterative optimization methods [35]. For comparison, an approved artificial neural network inversion approach was selected [37]. ANNs combine two advantages: first, they are computationally very fast, and second, they have the ability to approximate any (non-linear) relationship between different variables.

However, unexpected behavior may occur if the training data base does not well represent the spectral characteristics of the analyzed canopies [55].

For the setup of both inversion strategies, a synthetic data base with a size of 49,152 variable combinations was generated using the PROSAIL model. PROSAIL was configured for the simulation of the future ESA S2 sensors spectral band configurations according to the sensor spectral response functions. Variables and model parameters were randomly sampled within bounds (see Table 1) and applying truncated Gaussian distribution laws representative for different world vegetation types as proposed by [20]. Soil background was approximated by extracting and averaging bare soil signatures from different fields of the HyMap imagery. A simple multiplicative soil reflectance factor (α_{soil}, Table 1) is assumed to mimic variations of reflectance due to changes in superficial soil water content [56].

A stratified sampling scheme was used to ensure that values from each class (N class, Table 1) were combined with values from each other variable class. Illumination and viewing conditions (sun and sensor viewing angle, azimuth between sun and sensor) corresponded to those during the image acquisition.

Table 1. Variables, number of classes (N class), mean values, bounds (min/max) and standard deviation (SD) as input for the PROSAIL model for the generation of the training data base for LUT and ANN approaches.

Variable	N class	Mean	Min	Max	SD
Leaf Model: PROSPECT-5					
C_{ab}[1] [μg/cm^2]	4	45	20	90	30
C_w[2] [cm]	4	0.01	0.001	0.1	0.01
C_m[1,2] [g/cm^2]	4	0.0075	0.002	0.02	0.0075
C_{ar}[3] [μg/cm^2]	2	10	0	25	15
N[1,2]	4	1.5	1.0	2.5	1.0
Canopy Model: SAIL					
LAI[1,3] [m^2/m^2]	6	2	0	6	2
ALA[1,3] [°]	4	60	30	80	20
Hot[1,3]	1	0.1	0.001	1	0.3
α_{soil}[5]	4	1	0.6	1.4	0.5

Ranges of variables were applied according to: [1] [37], [2] [21], [3] [20], [4] [50], [5] [56].

Atmospheric correction and instrumental noise can result in multiplicative and additive uncertainties. Radiometric calibration might be inaccurate, which leads to systematic errors. Moreover, the used RTM may contain errors depending on its (simplified) description of the radiation regime in a vegetation canopy. Thus, to at least partly account for these uncertainties, the inclusion of noise in the simulations was decided. Fifty random initializations were generated adding and multiplying Gaussian white noise (absolute: 0.01 and relative: 4 %) to the simulated reflectance. This was done band-dependently and band-independently according to [37], who demonstrated that the combination of all these error-terms performed best for variable retrievals:

$$R(\lambda) = R_{sim}(\lambda) \cdot [1 + \varepsilon(0, \sigma_{rel}(\lambda)) + \varepsilon(0, \sigma_{rel}(all))] + \varepsilon(0, \sigma_{abs}(\lambda)) + \varepsilon(0, \sigma_{abs}(all)) \tag{1}$$

$R(\lambda)$ corresponds to the final and $R_{sim}(\lambda)$ to the simulated reflectance by the RTM, $\varepsilon(0,\sigma)$ representing a normal distribution, with $\sigma_{rel}(\lambda)$ and $\sigma_{rel}(all)$ representing the relative uncertainty applied to band λ and to all bands respectively, and with $\sigma_{abs}(\lambda)$ and $\sigma_{abs}(all)$ characterizing the absolute uncertainty added to band λ and to all bands, respectively.

For the LUT, a simple cost function composed of the root mean square error (RMSE) was employed [24,26]. Hereby, the spectra of the closest (radiometric) match with the measured signal were selected. The selection of the final solution was composed of two steps. As a first step, all variables (*i.e.*, LAI) were averaged that corresponded to the spectra within less than 20% of the lowest RMSE value. This value has been chosen according to our own tests and trials. Moreover, the 20%-threshold [22,57], or generally the application of multiple solutions [35], was also proposed by similar studies. As second step and final solution, the mean LAI was computed over fifty random initializations (found as sufficient by [37] for ANNs) with additive/multiplicative noise (Equation (1)). The results of this procedure are abbreviated with "LUT$_{N50}$". For comparison purposes, the LUT retrieval was performed without step two, *i.e.*, the selection procedure was performed only once, comparing measured and simulated spectra without adding/multiplying noise, abbreviated with "LUT$_{N1}$".

For the ANN, a three-layer feed forward, back propagation neural network was designed using the neural network toolbox in MatLAB®. Tan-sigmoid transfer functions were implemented in the hidden layer and linear transfer functions in the output layer.

Performing a sensitivity analysis, we found a number of five neurons for the hidden layer as optimal (not shown). This setting has also been found by [20,37]. The number of input neurons depended on the number of bands used for the training, while the output layer was composed of only a single neuron for the prediction of LAI. The use of a single neuron in the output layer has been suggested by [20,37] and moreover was found as optimal by our own tests.

The synthetic data base was split into three subsets. The first was used for updating the weights and biases of the network (50%-training). The second data set (25%) was employed to check the progress of the training algorithm, thus to prevent over-fitting. This implies that these data were not completely independent but an essential part of the training process to select the right model. The third (25%) subset was then used for independent model evaluation and therefore to obtain confidence of the final model. Whereas the second data subset was a part of the reiteration process, the (third) validation data set was used only once. The final solution of LAI was then calculated as average of all 50 networks.

2.3. Band Sensitivity Analysis

In order to identify the optimal spectral sampling, *i.e.*, how many and which of the available S2 bands would be required for best LAI retrieval performance, a band sensitivity analysis was carried out. For this purpose, the approaches described above were applied to all possible combinations of bands. The synthetic S2-bands were grouped into arrangements, which are defined by different numbers of potentially used bands. Thereby, between two, and up to ten, spectral bands may be included in one arrangement. For each possible arrangement, the maximum number of possible band combinations was calculated (see Table 2). Since the HyMap sensor covers the spectral information of future S2 sensors, all bands of interest for our study (*i.e.*, 10 out of 13) could be included in the analyses. Hence, the wavebands of the HyMap sensor most adjacent to the following central S2 wavebands were incorporated: 490 nm, 560 nm, 665 nm, 705 nm, 740 nm, 783 nm, 842 nm and 865 nm, 1,610 nm and 2,190 nm. This may include uncertainty. However, HyMap does not provide the required high spectral resolution to apply the S2 sensors' spectral response functions.

In this way, only channels providing a GSD of 10 m or 20 m were considered. This decision was taken, because in the context of GMES land monitoring applications, the purpose of these bands will mainly be the mapping of geo-biophysical vegetation variables, land use and land cover, whereas the remaining three bands with a GSD of 60 m (443 nm, 945 nm and 1,375 nm) are foreseen as being used for atmospheric correction [7].

Spatial aspects in view of the three different GSDs of the S2 sensors were not considered in the current study. However, to provide comparability of the remotely-sensed estimates from the HyMap sensor with the in situ LAI measurements, reflectance mean values of 4 × 4 pixels were extracted. For this purpose, the central coordinates of the LAI ESUs were taken. Comparison of measurements and simulations are therefore based on a 20 m ground sampling distance.

Table 2. Number of possible band combinations (N) and retrieved minimum RMSE ($RMSE_{min}$) between measured and estimated LAI of each arrangement for the three approaches: LUT_{N50}, LUT_{N1} and ANN. SPARC'03 campaigns data analyses (best results of each approach in bold).

Arrangement (No. of Included Spectral Bands)	N (No. of Band Combinations)	$RMSE_{min}$		
		LUT_{N50}	LUT_{N1}	ANN
2	45	0.56	0.93	0.91
3	120	0.54	0.77	0.73
4	210	0.54	0.73	**0.58**
5	252	0.54	**0.70**	**0.58**
6	210	0.54	**0.70**	**0.58**
7	120	**0.53**	0.73	0.59
8	45	**0.53**	0.77	0.64
9	10	0.56	0.78	0.67
10	1	0.70	0.95	0.84

3. Results and Discussion

In this section, outcomes of the spectral band analyses are presented and discussed. Three aspects are considered: first, the distribution of RMSE values between measured and estimated LAI ("LAI RMSE") for all possible band arrangements is analyzed. Second, the importance of the different spectral regions for LAI estimation is addressed. Finally, crop specific differences are elaborated.
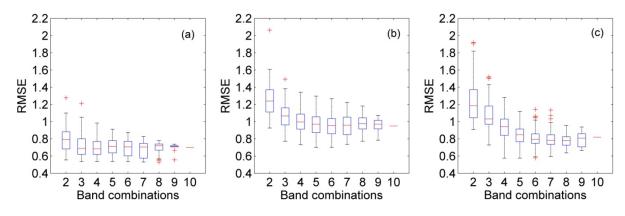
3.1. Optimal Number of Bands

The distribution and variation of the resulting LAI RMSE values for each band arrangement can be well illustrated through box plot diagrams (Figure 1). Overall best accuracy (LAI $RMSE_{min}$ = 0.53, Table 2) was obtained by the LUT_{N50} with the seven/eight-band arrangements (Figure 1(a)). Worst results instead are achieved through LUT_{N1} by a combination within a two-band arrangement (LAI $RMSE_{min}$ = 2.1, Table 2, Figure 1(b)). However, the $RMSE_{min}$ differences between ANN and LUT_{N50} approaches cannot be regarded as significant. These results rather show that the proposed LUT_{N50} approach with implemented noise levels is comparable in terms of retrieval performances with a tested and approved ANN method.

For all three approaches, the size of the boxes (which correspond to 50 % of the data) tends to decrease with increasing number of bands included in the arrangement, meaning that the variability (dispersion) of the results also diminished when higher numbers of bands were used. This is also expressed by the smaller distance of the whiskers for the band arrangements with higher numbers of included bands. For the ANN approach, the decrease only is constant until eight bands are included, Figure 1(c).

An overview of the obtained $RMSE_{min}$ from the combinations of each investigated band arrangement is presented in Table 2: whereas the ANN and LUT_{N1} approaches reached absolute minima with 4 to 6 (ANN) and 5/6 (LUT_{N1}) band combinations, the LUT_{N50} approach obtained the $RMSE_{min}$ with 7 and 8 bands. Similar—though not identical—results have been found in previous studies: Verger *et al*. [37] found seven out of 62 bands as optimal for LAI estimation using Compact

High Resolution Imaging Spectrometer (CHRIS)/Proba sensor data and applying an ANN approach. Weiss *et al*. [58] selected six from nine synthetic bands (simulated with a RTM), obtaining $RMSE_{min}$ between measured and simulated LAI using a look-up table approach. Fourty *et al*. [59] found five to eight wavebands for estimating accurately different canopy biophysical variables, using multiple linear regression on simulated data with PROSAIL. In another study, exploiting the PROSAIL model with hyperspectral airborne DAIS data, 22 from 30 bands performed best for LAI retrieval. However, the strongest reduction of RMSE was found from using up to eight bands [36].

> **Figure 1.** Distribution and variation of the resulting LAI RMSE values for each band arrangement for the LUT based approaches (**a**): LUT_{N50} and (**b**): LUT_{N1} as well as the ANN approach (**c**). The bottom and the top of the boxes describe the 25th and the 75th percentiles; the red line within the boxes corresponds to the 50th percentile (or median); lower and upper ends of the whiskers show lowest and highest values within the 1.5 interquartile range (IQR). Outliers are marked with a red cross.

Whereas the absolute lowest RMSE values are on a similar level for ANN and LUT_{N50}, the latter provides lower variability of RMSE differences (*i.e.*, highest $RMSE_{min}$ − lowest $RMSE_{min}$): 0.17 of LUT_{N50} compared to 0.25 of the LUT_{N1} and 0.33 of the ANN methods. This indicates higher robustness and a lower sensitivity of the LUT_{N50} method regarding the optimal number of spectral bands.

As demonstrated, the results of such analyses may depend on the algorithms employed. However, generally it can be said that the optimal number of bands is around six to eight for the estimation of LAI. Whereas the use of only a few bands (two to four) enhances the ill-posed inverse problem and therefore the retrieval uncertainty, the employment of too many bands (more than eight/nine) may again lead to decreasing accuracy of the estimates. This can be caused by redundant spectral information, noise in the measured reflectance or the inability of the RTM to appropriately simulate certain spectral regions [37].

3.2. Optimal Spectral Sampling

In order to identify the spectral bands most often used by the approaches, all band combinations of each arrangement between the $RMSE_{min}$ and the 0.05 quantile of $RMSE_{min}$ were selected from LUT_{N50} and ANN approaches. All cases within the 0.05 quantile, instead of simply selecting the best band combination, were included in order to account for some uncertainty. The 0.05 quantile included between one (for nine-band arrangement) and 13 (for five-band arrangement) cases. In Figure 2, the

frequency of the selected bands is presented. Whereas Figure 2(a) shows the frequency of selection for each single band, Figure 2(b) indicates the most frequent selection grouped per spectral region: visible (VIS, 490 nm, 560 nm and 665 nm), red edge (705 nm, 740 nm and 783 nm), near infrared (NIR, 842 nm, 865 nm) and short wave infrared (SWIR, 1,610 nm and 2,190 nm).
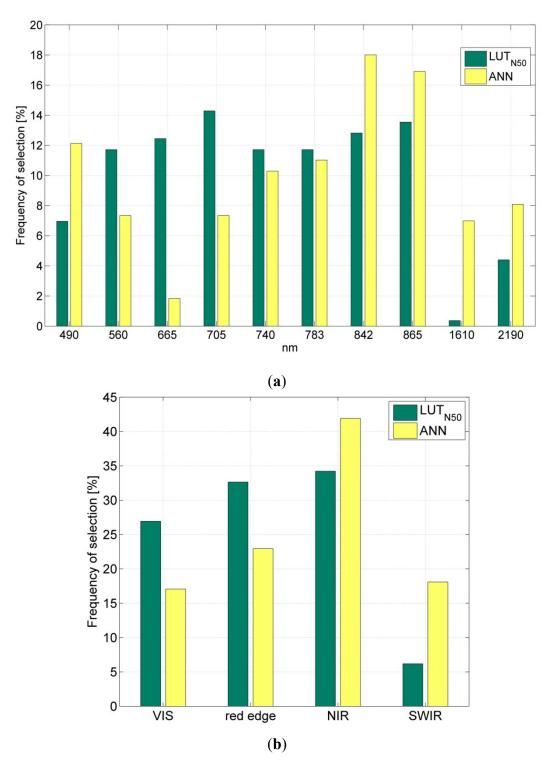
In some aspects, both approaches (*i.e.*, LUT_{N50} and ANN) show the same tendency: no band and thus no spectral region were completely excluded by the algorithms. However, there are some strong differences, visible in Figure 2(a): the LUT_{N50} approach selected most often the red edge band (705 nm), closely followed by the two NIR bands and the red and green visible (665 nm and 560 nm). Less often selected bands were located in the blue visible (490 nm), but in particular in the SWIR region. Instead, the ANN approach prioritized the two NIR bands, followed by the blue visible, then red edge, green visible and the two SWIR bands. The red visible was the less often selected spectral band. Looking at the spectral groups (Figure 2(b)), these differences diminish to a more similar pattern: most often selected bands were located in the NIR, followed by the red edge and VIS domains, or by the SWIR respectively, in case of the ANN. To some extent, the same tendency as for the optimal number of bands was found: the LUT_{N50} methods exhibit lower sensitivity to the selection of bands than the ANN approach, at least from visible to NIR domains.

The frequent selection of NIR bands was expected: multiple scattering between the spongy mesophyll cells is very pronounced in this spectral region (e.g., [60]). Therefore, it is a well-known fact that reflectance increases with increasing leaf material—thus LAI. The dominance of NIR reflectance in this context has been also found by other studies, for instance [36], where the absolute minimum RMSE for LAI retrieval was reached after selecting the majority of available NIR bands (15) and four bands in the visible region. In the study of [37] the NIR was also found to be the spectral region of most interest for LAI retrieval, selecting five bands from the NIR and two from the red visible domains.

There are diverging opinions in the literature concerning the importance of the red edge spectral region for LAI estimation (e.g., [33,61]). Our results suggested that the red edge has more influence on LAI retrieval than visible and SWIR bands. In the study of [58], two of the six selected bands were located in the red edge domain, three in the visible and one band in NIR region.

Spectral bands located in the SWIR range were less frequently chosen than most others, as also found by [36] who selected only two from 21 in the SWIR domain for an optimal LAI retrieval. Nevertheless, the SWIR bands also contributed to the (relatively) high retrieval accuracies in our study as well as in others, for instance [59]. Even though only wavelengths from 880 to 2,380 nm were considered, five of the six selected bands for optimal LAI estimation were located in the SWIR [59]. Due to the limited data availability from sensors operating in the SWIR, previous studies often could employ only visible and NIR bands. Since some studies demonstrated that the inclusion of SWIR improved the retrieval accuracy of LAI [61,62], data delivery from Sentinel-2 in this spectral region will certainly be valuable. Moreover, the SWIR bands may play an important role for discriminating the spectral signal for different soil and vegetation variables, such as dry matter or water content [63]. However, further research is still required in this regard.

Figure 2. Frequency of selected bands from RMSE$_{min}$ to the 0.05 quantile, LUT$_{N50}$ and ANN approaches. (**a**) per band and (**b**) per spectral group.

(**a**)

(**b**)

3.3. Crop Specific Differences

The RT model (PROSAIL) applied here is based on a (1-D) turbid medium assumption and thus has a limited capability to simulate complex canopy architectures [57]. Mainly for this reason, retrieval accuracies may vary between the different crops exhibiting different canopy structures and growth stages, as found by several studies (e.g., [26,31,57]). Therefore, crop-specific accuracies were calculated

and depicted in Figure 3 in the form of scatter plots for the LUT_{N50} and the ANN approaches. LAI estimation was performed, employing the band combinations providing the minimum RMSE ($RMSE_{min}$), *i.e.*, an eight-band combination from LUT_{N50}, including the first eight bands without the SWIR and a five-band combination from ANN, including the green visible, the two NIR and the two SWIR bands.
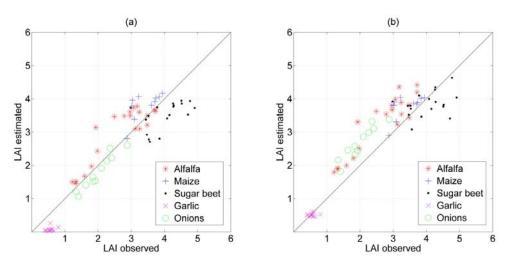
The use of a single statistical measure (such as RMSE) only provides limited information of the retrieval performance: differences in absolute number, magnitude, range or spatial patterns of the measured/simulated values can influence the indicators. Thus, to give a valid overview of model performance, a statistical indicator set proposed by [64] was calculated (see Table 3): coefficient of determination (R^2), RMSE, normalized RMSE (NRMSE)—which is the RMSE, divided by the range of the reference measurements—and Nash-Sutcliffe efficiency index (NSE, [65], Equation (2)). The NSE index, which can range between $-\infty$ and 1, gives a good indication of a model's prediction capability. A value of NSE below 0 indicates that the estimated variable values obtain lower accuracies than simply the mean of the observed (measured) variables. Therefore, model reliability is only provided for NSE > 0. The index is calculated according to the following equation (Equation (2)):

$$NSE = 1 - \frac{\sum_{i=1}^{n} (V_{obs}^i - V_{est}^i)^2}{\sum_{i=1}^{n} (V_{obs}^i - \overline{V}_{obs})^2} \quad (2)$$

V_{obs}^i is the observed (measured) variable i and V_{est}^i the corresponding estimated value. The mean value of all observed variables is indicated with $\overline{V}_{obs}$.

In fact, it is notable that the retrieval performance depends on the crop type (Figure 3, Table 3).

Figure 3. Scatter plots of observed *versus* estimated LAI from SPARC'03 campaigns data, best results ($RMSE_{min}$) of (**a**) LUT_{N50} approach, (**b**) ANN approach.

Whereas the highest accuracy was found by all statistical indicators for onions and alfalfa using the LUT_{N50} approach, the ANN method obtained better results for maize, sugar beet and garlic. For these three crops, however, the spatial patterns (indicated by R^2) were better reproduced by the LUT_{N50} approach, although the correlations were still low (from $R^2 \sim 0.1$ to 0.4). Moreover, all NSE are < 0,

implying that the mean value of the observed LAI would obtain a higher accuracy than the estimated LAI values.
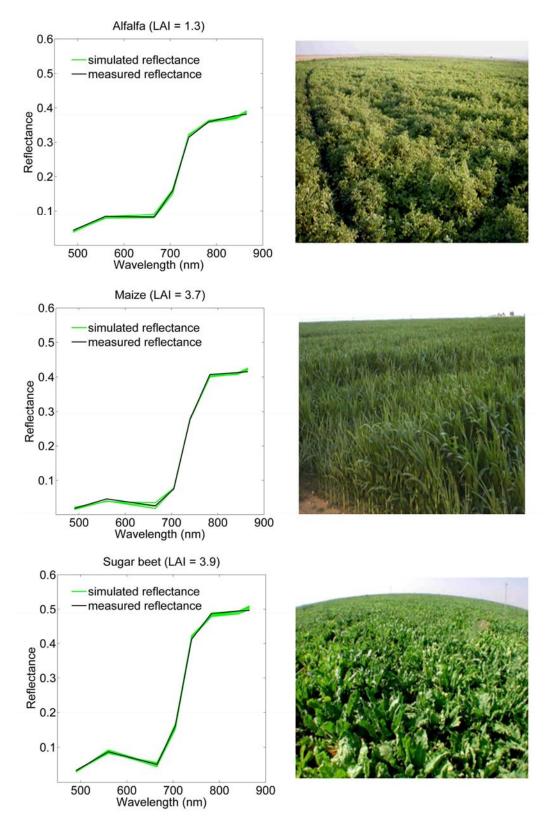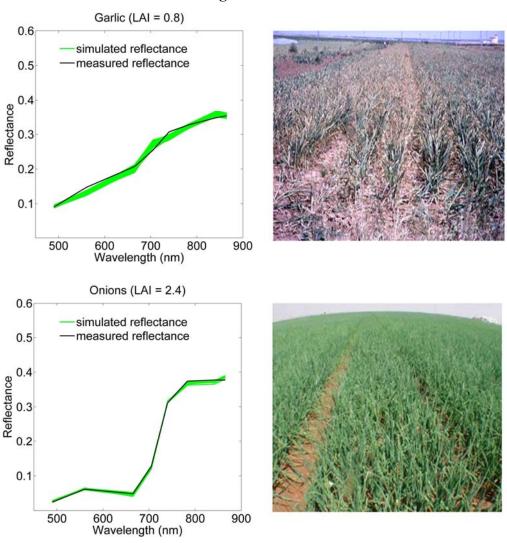
Table 3. Goodness-of-fit statistics (R^2, RMSE, NRMSE and NSE) between observed and estimated LAI from SPARC'03 data. Higher performance values between the two approaches are emphasized in bold.

Crop Type	N	Range (Observed)	LUT$_{N50}$: 'Best Band' Combination (8 Bands)				ANN: 'Best Band' Combination (5 Bands)			
			R^2	RMSE	NRMSE	NSE	R^2	RMSE	NRMSE	NSE
alfalfa	20	1.22–3.72	0.80	**0.50**	**0.20**	**0.63**	**0.84**	0.71	0.28	0.26
maize	10	2.87–3.94	**0.44**	0.50	0.47	−0.75	0.39	**0.47**	**0.44**	**−0.56**
sugar beet	18	2.98–4.91	**0.3**	0.71	0.37	−0.79	0.13	**0.55**	**0.29**	**−0.08**
garlic	11	0.43–0.80	**0.11**	0.49	1.23	−21.35	0.01	**0.12**	**0.30**	**−0.34**
onions	11	1.34–2.88	**0.92**	**0.25**	**0.16**	**0.68**	0.88	0.70	0.45	−1.51
all crops	70	0.43–4.91	**0.86**	**0.53**	**0.12**	**0.83**	0.85	0.58	0.13	0.80

Thus, the models prediction capabilities are doubtful. This may be due to the above-mentioned limitations of the used RT model: the erectophile canopy of garlic, for instance, may lead to a strong influence of the soil, being a very critical factor for model inversion [31,37]. Looking at Figure 4, the garlic spectra resemble bare soil signatures (both measured and simulated) due to absent chlorophyll absorption even though the field measurements indicated a LAI of 0.8. According to the picture taken during the campaign, garlic exhibits already a senescent growth stage. The measurement from the LAI-2000 instrument (PAI$_{eff}$) is therefore strongly influenced by non-green plant components. This suggests that the LAI estimated by the LUT$_{N50}$ approach is closer to the green LAI value than estimated by the ANN approach (see also Figure 3) or measured in the field. Despite the fact that the measured and ANN estimated LAI-values of garlic are very close, the LUT$_{N50}$ approach seems to give the most accurate interpretation of the spectral signature.

Moreover, the presence of row structures, which are not accounted for by the 1-D RT model, may lead to inaccuracies, for instance as is often the case in maize [26]. However, the maize already reached LAI values between 3 and 4 and thus almost approached a homogenous coverage (see also picture Figure 4). In fact, seven from ten values are located on the 1:1 line (for both approaches). Moreover, with RMSE values of 0.5 (LUT$_{N50}$) and 0.47 (ANN), the accuracy is higher than the average. The influence of leaves not randomly distributed as assumed by the model but clumped, can result in an underestimation of high LAI values [20], as it occurred in the actual growth stage of sugar beet for both approaches. However, as discussed in Section 2.1., the "effective LAI" is measured, rather than the true LAI, also by the optical ground-based instrument. A proper interpretation of the overestimation is therefore difficult in this case. It could be speculated that the higher LAI values obtained with the LAI-2000 were an effect of beginning leaf senescence, apparently leading to higher LAI values. However, as shown in Figure 4, this is not as strong as in case of garlic.

Figure 4. Selected simulated (PROSAIL generated database for LUT and ANN approaches) and HyMap spectra, one for each crop type; band setting of best results (RMSE$_{min}$) of LUT$_{N50}$ approach, *i.e.*, first eight S2 bands used in the study (*i.e.*, from 490 nm to 865 nm). All simulated spectra with radiometric RMSE value within 20% of the lowest RMSE are presented. Measured LAI values are indicated in the titles. On the right the corresponding pictures of each crop field, taken during the campaign, are presented.

Figure 4. *Cont.*

Even when looking at the crop-specific results, it cannot be concluded that one method outperforms the other. Both approaches, LUT_{N50} and ANN, reveal similar performances with reasonable results. However, problems are still present depending on the architecture and actual growth stages of the crops.

In order to obtain an idea of the model's ability to reproduce the HyMap reflectance data, some exemplary spectra are presented in Figure 4. For this purpose, the simulated spectra obtaining the best radiometric match within the LUT_{N50} approach were chosen (eight bands). It is clearly visible in Figure 4 that in all cases the HyMap reflectance is appropriately reproduced by the model.

3.4. Limitations of the Study

In the retrieval of biophysical vegetation variables (such as LAI), various components can influence the estimation accuracy. These include, for instance, the type of remote sensor with its spectral and radiometric characteristics, crop type monitored as well as the applied model and retrieval (inversion) methods. Moreover, the validation of the final estimates is influenced by the instruments and strategy used for acquisition of the *in situ* reference data. The contribution to the overall uncertainty of each of these components may vary from case to case. Still further research efforts are required to reduce or at least mitigate these uncertainties and errors and to guarantee high retrieval qualities within the context of both

current and future satellite missions. Therefore, the question of the ideal number and position of spectral bands for LAI retrieval cannot, of course, be entirely answered solely with the results of our study.

Our study was conducted using a part of the extensive database generated during one of the largest agricultural field campaigns in the last years, the SPARC campaign in the Barrax area. Nevertheless, the application of the method to other environmental sites and sensors and thus confirmation of the algorithms and outcomes would be desirable.

As in our study from 2009 [26], we can draw the conclusion that the inversion strategies have only a minor influence on the LAI retrieval accuracy when using well-established approaches such as ANN or LUT including noise (*i.e.*, LUT_{N50}). The inversion strategy is, however, of secondary importance, since it will not compensate for problems related to the choice of the appropriate radiative transfer model. We have again selected the PROSAIL model for our study because it has been widely applied and tested and has been proven to be a feasible compromise between accuracy, variable number input and computation time. However, due to the model's turbid medium assumption it is also well known (and found in our own studies) that PROSAIL has limitations, especially for crops in particular growth stages, where clumping occurs or the underlying soil and row structures affect the spectral signal.

Moreover, as also shown by our results, optimum spectral sampling can depend on the retrieval method. The employment of other band selection/elimination methods, for instance from SVR [66], may again lead to diverging results. The band setting, however, has a major influence on retrieval accuracy in contrast to the findings of [26]. This was demonstrated in the present study with a more dedicated band selection process: with the box plots, a reliable indicator of the required quantity of spectral information for biophysical variable (LAI) retrieval is presented. By means of this tool, the optimal number of spectral bands can be detected. On the one hand, this avoids the use of limited spectral information (e.g., with VIs), diminishing the ill-posed inverse problem, while on the other, the employment of too many bands, for instance in RTM inversion schemes, can be prevented, reducing inaccuracies and computation time due to redundant spectral information.

Therefore, the application of the presented analyses to other sites as well as further tests of the applied method would enhance the validity of our results. This enhancement is required to ensure high quality biophysical data products from Sentinel-2 sensors.

4. Conclusions

In this study, spectral issues for the retrieval of leaf area index, one of the major biophysical variables required for agricultural applications, were addressed. We focused on the question of the optimal spectral sampling for LAI retrieval from future Sentinel-2 sensor data using two Look-up tables and an approved artificial neural network approach. In summary, results from our analyses lead to the following conclusions: though LAI can be roughly estimated using only a few bands (*i.e.*, two or three, as with VI approaches), a high retrieval uncertainty must be taken into account when using this minimum spectral information. Box plots in Figure 1 demonstrate that this uncertainty can be diminished by including a higher number of spectral bands (six to eight), thus adding important information until spectral redundancy may outweigh the information gain. Regarding the band positions, NIR and red edge spectral regions provide the most relevant information for LAI, confirming previous literature findings. Moreover, the proper inclusion (*i.e.*, 50 times) of additive and

multiplicative noise accounting for uncertainties from atmospheric correction, instrument and RT model into the LUT$_{N50}$ method significantly improved the retrieval results of this approach.

The best result from the band selection process (RMSE of 0.53 from LUT$_{N50}$) corresponds to a normalized RMSE of 12 % from the LUT$_{N50}$ approach (13 % from ANN, respectively). Results of both approaches are therefore comparable. More importantly, the LUT$_{N50}$ provided a higher robustness and lower sensitivity to band selection, as indicated by the low variability between the RMSE$_{min}$ values and the more equal selection of bands between visible and NIR domains.

Looking at crop specific results, NRMSE values were below 16%. A range of 15%–20% is regarded by [67] as the currently achievable accuracy for LAI from remote sensing observations. Hence, the spectral channels planned for the Sentinel-2 sensors offer a valid information basis for LAI retrieval. However, a retrieval accuracy of 10 % for LAI is targeted for the mission [7]. Thus, improvements would be desirable. With Sentinel-2 sensors data, such improvements could be achieved by employing spatial [31] and/or temporal information [22,68].

The improved retrieval of biophysical variables may further encourage the development of advanced strategies for the use of Earth observation EO data. By these means, the assimilation of remote sensing data into land surface process models [69] may largely contribute to an enhanced application of Sentinel-2 products.

Acknowledgements

References

1. Baret, F.; Buis, S. Estimating canopy characteristics from remote sensing observations: Review of methods and associated problems. In *Advances in Land Remote Sensing: System, Modeling, Inversion and Application*; Liang, S., Ed.; Springer: Dordrecht, The Netherlands, 2008; pp. 173-201.
2. Widlowski, J.L.; Pinty, B.; Gobron, N.; Verstraete, M.M.; Diner, D.J.; Davis, A.B. Canopy structure parameters derived from multi-angular remote sensing data for terrestrial carbon studies. *Clim. Change* **2004**, *67*, 403-415.
3. Wolter, P.T.; Townsend, P.A.; Sturtevant, B.R. Estimation of forest structural parameters using 5 and 10 meter SPOT-5 satellite data. *Remote Sens. Environ.* **2009**, *113*, 2019-2036.
4. Potter, C.S.; Klooster, S.; Brooks, V. Interannual variability in terrestrial net primary production: Exploration of trends and controls on regional to global scales. *Ecosystems* **1999**, *2*, 36-48.
5. D'Urso, G.; Richter, K.; Calera, A.; Osann, M.A.; Escadafal, R.; Garatuza-Pajan, J.; Hanich, L.; Perdigao, A.; Tapia, J.B.; Vuolo, F. Earth observation products for operational irrigation management in the context of the pleiades project. *Agric. Water Manag.* **2010**, *98*, 271-282.

6. Martimort, P.; Berger, M.; Carnicero, B.; Del Bello, U.; Fernandez, V.; Gascon, F.; Silvestrin, P.; Spoto, F.; Sy, O.; Arino, O.; *et al.* Sentinel-2: The optical high-resolution mission for GMES operational services. In *ESA Bulletin*; 2007; Volume 131, pp. 18-23.

7. Drusch, M.; Gascon, F.; Berger, M. *GMES Sentinel-2 Mission Requirements Document*; European Space Agency: 2010; p. 42. http://esamultimedia.esa.int/docs/GMES/Sentinel-2_MRD.pdf (accessed date 02 February 2012).

8. Dorigo, W.A.; Zurita-Milla, R.; de Wit, A.J.W.; Brazile, J.; Singh, R.; Schaepman, M.E. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. *Int. J. Appl. Earth Obs. Geoinf.* **2007**, *9*, 165-193.

9. Baret, F.; Guyot, G. Potentials and limits of vegetation indices for LAI and apar assessment. *Remote Sens. Environ.* **1991**, *35*, 161-173.

10. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127-150.

11. Glenn, E.; Huete, A.; Nagler, P.; Nelson, S. Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors* **2008**, *8*, 2136-2160.

12. Govaerts, Y.M.; Verstraete, M.M.; Pinty, B.; Gobron, N. Designing optimal spectral indices: A feasibility and proof of concept study. *Int. J. Remote Sens.* **1999**, *20*, 1853-1873.

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

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

15. Atzberger, C.; Guerif, M.; Baret, F.; Werner, W. Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Comput. Electron. Agric.* **2010**, *73*, 165-173.

16. Cho, M.A.; Skidmore, A.K.; Atzberger, C. Towards red-edge positions less sensitive to canopy biophysical parameters for leaf chlorophyll estimation using properties optique spectrales des feuilles (prospect) and scattering by arbitrarily inclined leaves (Sailh) simulated data. *Int. J. Remote Sens.* **2008**, *29*, 2241-2255.

17. Byambakhuu, I.; Sugita, M.; Matsushima, D. Spectral unmixing model to assess land cover fractions in mongolian steppe regions. *Remote Sens. Environ.* **2010**, *114*, 2361-2372.

18. Atkinson, P.M.; Tatnall, A.R.L. Introduction neural networks in remote sensing. *Int. J. Remote Sens.* **1997**, *18*, 699-709.

19. Camps Valls, G.; Bruzzone, L.; Rojo Álvarez, J.L.; Melgani, F. Robust support vector regression for biophysical variable estimation from remotely sensed images. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 339-343.

20. Baret, F.; Hagolle, O.; Geiger, B.; Bicheron, P.; Miras, B.; Huc, M.; Berthelot, B.; Nino, F.; Weiss, M.; Samain, O.; *et al.* LAI, FaPAR and fcover cyclopes global products derived from vegetation—Part 1: Principles of the algorithm. *Remote Sens. Environ.* **2007**, *110*, 275-286.

21. Jacquemoud, S.; Baret, F.; Andrieu, B.; Danson, F.M.; Jaggard, K. Extraction of vegetation biophysical parameters by inversion of the prospect plus sail models on sugar-beet canopy reflectance data - application to tm and aviris sensors. *Remote Sens. Environ.* **1995**, *52*, 163-172.

22. Koetz, B.; Baret, F.; Poilve, H.; Hill, J. Use of coupled canopy structure dynamic and radiative transfer models to estimate biophysical canopy characteristics. *Remote Sens. Environ.* **2005**, *95*, 115-124.

23. Goel, N.S. Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data. *Remote Sens. Rev.* **1988**, *4*, 1-212.

24. Combal, B.; Baret, F.; Weiss, M.; Trubuil, A.; Mace, D.; Pragnere, A.; Myneni, R.; Knyazikhin, Y.; Wang, L. Retrieval of canopy biophysical variables from bidirectional reflectance—Using prior information to solve the ill-posed inverse problem. *Remote Sens. Environ.* **2003**, *84*, 1-15.

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

26. Richter, K.; Atzberger, C.; Vuolo, F.; Weihs, P.; D'Urso, G. Experimental assessment of the Sentinel-2 band setting for rtm-based lai retrieval of sugar beet and maize. *Can. J. Remote Sens.* **2009**, *35*, 230-247.

27. Fang, H.L.; Liang, S.L.; Kuusk, A. Retrieving leaf area index using a genetic algorithm with a canopy radiative transfer model. *Remote Sens. Environ.* **2003**, *85*, 257-270.

28. Durbha, S.S.; King, R.L.; Younan, N.H. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sens. Environ.* **2007**, *107*, 348-361.

29. Zhang, Q.Y.; Xiao, X.M.; Braswell, B.; Linder, E.; Baret, F.; Moore, B. Estimating light absorption by chlorophyll, leaf and canopy in a deciduous broadleaf forest using MODIS data and a radiative transfer model. *Remote Sens. Environ.* **2005**, *99*, 357-371.

30. Walthall, C.; Dulaney, W.; Anderson, M.; Norman, J.; Fang, H.L.; Liang, S.L. A comparison of empirical and neural network approaches for estimating corn and soybean leaf area index from Landsat ETM+ imagery. *Remote Sens. Environ.* **2004**, *92*, 465-474.

31. Atzberger, C.; Richter, K. Spatially constrained inversion of radiative transfer models for improved lai mapping from future Sentinel-2 imagery. *Remote Sens. Environ.* **2011**, accepted.

32. Liang, S. Recent developments in estimating land surface biogeophysical variables from optical remote sensing. *Progr. Phys. Geogr.* **2007**, *31*, 501-516.

33. Delegido, J.; Verrelst, J.; Alonso, L.; Moreno, J. Evaluation of Sentinel-2 red-edge bands for empirical estimation of green lai and chlorophyll content. *Sensors* **2011**, *11*, 7063-7081.

34. Atzberger, C.; Richter, K.; Vuolo, F.; Darvishzadeh, R.; Schlerf, M. Why confining to vegetation indices? Exploiting the potential of improved spectral observations using radiative transfer models. *Proc. SPIE* **2011**, *8174*, 81740Q.

35. Darvishzadeh, R.; Skidmore, A.; Schlerf, M.; Atzberger, C. Inversion of a radiative transfer model for estimating vegetation lai and chlorophyll in a heterogeneous grassland. *Remote Sens. Environ.* **2008**, *112*, 2592-2604.

36. Meroni, M.; Colombo, R.; Panigada, C. Inversion of a radiative transfer model with hyperspectral observations for lai mapping in poplar plantations. *Remote Sens. Environ.* **2004**, *92*, 195-206.

37. Verger, A.; Baret, F.; Camacho, F. Optimal modalities for radiative transfer-neural network estimation of canopy biophysical characteristics: Evaluation over an agricultural area with chris/proba observations. *Remote Sens. Environ.* **2011**, *115*, 415-426.

38. Price, J. An approach for analysis of reflectance spectra. *Remote Sens. Environ.* **1998**, *64*, 316-330.

39. Thenkabail, P.S.; Enclona, E.A.; Ashton, M.S.; Van Der Meer, B. Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. *Remote Sens. Environ.* **2004**, *91*, 354-376.

40. Moreno, J.; Alonso, L.; Fernández, G.; Fortea, J.C.; Gandía, S.; Guanter, L. The Spectra Barrax Campaign (Sparc): Overview and First Results from CHRIS Data. In *Proceedings of 2nd CHRIS/PROBA Workshop*, Frascati, Italy, 28–30 April 2004.

41. Cocks, T.; Jenssen, R.; Stewart, A.; Wilson, I.; Shields, T. The Hymap Airborne Hyperspectral Sensor: The System, Calibration and Performance. In *Proceedings of 1rd EARSeL Workshop on Imaging Spectrometry*, Zurich, Switzerland, 6–8 October 1998; pp. 37-42.

42. Guanter, L.; Richter, R.; Moreno, J. Spectral calibration of hyperspectral imagery using atmospheric absorption features. *Appl. Opt.* **2006**, *45*, 2360-2370.

43. Welles, J.M.; Norman, J.M. Instrument for indirect measurement of canopy architecture. *Agron. J.* **1991**, *83*, 818-825.

44. Martinez, B.; Baret, F.; Camacho-de Coca, F.; Garcia-Haro, F.J.; Verger, A.; Melia, J. Validation of MSG vegetation products: Part I. Field retrieval of LAI and FVC from hemispherical photographs. *Proc. SPIE* **2004**, *5568*, 57-68.

45. Martinez, B.; Cassiraga, E.; Camacho, F.; Garcia-Haro, J. Geostatistics for mapping leaf area index over a cropland landscape: Efficiency sampling assessment. *Remote Sens.* **2010**, *2*, 2584-2606.

46. Leblanc, S.G.; Chen, J.M.; Fernandes, R.; Deering, D.W.; Conley, A. Methodology comparison for canopy structure parameters extraction from digital hemispherical photography in boreal forests. *Agric. For. Meteorol.* **2005**, *129*, 187-207.

47. Ryu, Y.; Nilson, T.; Kobayashi, H.; Sonnentag, O.; Law, B.E.; Baldocchi, D.D. On the correct estimation of effective leaf area index: Does it reveal information on clumping effects? *Agric. For. Meteorol.* **2010**, *150*, 463-472.

48. Garrigues, S.; Shabanov, N.V.; Swanson, K.; Morisette, J.T.; Baret, F.; Myneni, R.B. Intercomparison and sensitivity analysis of leaf area index retrievals from LAI-2000, accupar, and digital hemispherical photography over croplands. *Agric. For. Meteorol.* **2008**, *148*, 1193-1209.

49. Soudani, K.; Francois, C.; le Maire, G.; Le Dantec, V.; Dufrene, E. Comparative analysis of IKONOS, SPOT, and ETM+ data for leaf area index estimation in temperate coniferous and deciduous forest stands. *Remote Sens. Environ.* **2006**, *102*, 161-175.

50. Feret, J.-B.; Francois, C.; Asner, G.P.; Gitelson, A.A.; Martin, R.E.; Bidel, L.P.R.; Ustin, S.L.; le Maire, G.; Jacquemoud, S. Prospect-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sens. Environ.* **2008**, *112*, 3030-3043.

51. Verhoef, W. Light scattering by leaf layers with application to canopy reflectance modeling: The scattering by arbitrarily inclined leaves (sail) model. *Remote Sens. Environ.* **1984**, *16*, 125-141.

52. Jacquemoud, S.; Verhoef, W.; Baret, F.; Bacour, C.; Zarco-Tejada, P.J.; Asner, G.P.; Francois, C.; Ustin, S.L. Prospect plus sail models: A review of use for vegetation characterization. *Remote Sens. Environ.* **2009**, *113*, S56-S66.

53. Kuusk, A. The hot spot effect in plant canopy reflectance. In *Photon and Vegetation Interactions*; Myneni, R.B.; Ross, J., Eds.; Springer-Verlag: Berlin, Germany, 1991; pp. 140-159.

54. Jones, H.G.; Vaughan, R.A. *Remote Sensing of Vegetation: Principles, Techniques and Applications*; Oxford University Press: Oxford, UK, 2010.

55. Schlerf, M.; Atzberger, C. Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data. *Remote Sens. Environ.* **2006**, *100*, 281-294.

56. Richter, K.; Vuolo, F.; D'Urso, G.; Palladino, M. Evaluation of near-surface soil water status through the inversion of soil-canopy radiative transfer models in the reflective optical domain. *Int. J. Remote Sens.* **2012**, in press.

57. Dorigo, W.A. Improving the robustness of cotton status characterisation by radiative transfer model inversion of multi-angular chris/proba data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2011**, *PP*, 1-12.

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

59. Fourty, T.; Baret, F. Vegetation water and dry matter contents estimated from top-of-the-atmosphere reflectance data: A simulation study. *Remote Sens. Environ.* **1997**, *61*, 34-45.

60. Gausman, H.W.; Allen, W.A.; Cardenas, R. Reflectance of cotton leaves and their structure *Remote Sens. Environ.* **1969**, *1*, 19-22.

61. Darvishzadeh, R.; Atzberger, C.; Skidmore, A.K.; Abkar, A.A. Leaf area index derivation from hyperspectral vegetation indices and the red edge position. *Int. J. Remote Sens.* **2009**, *30*, 6199-6218.

62. Brown, L.; Chen, J.M.; Leblanc, S.G.; Cihlar, J. A shortwave infrared modification to the simple ratio for lai retrieval in boreal forests: An image and model analysis. *Remote Sens. Environ.* **2000**, *71*, 16-25.

63. Khanna, S.; Palacios-Orueta, A.; Whiting, M.L.; Ustin, S.L.; Riaño, D.; Litago, J. Development of angle indexes for soil moisture estimation, dry matter detection and land-cover discrimination. *Remote Sens. Environ.* **2007**, *109*, 154-165.

64. Richter, K.; Hank, T.B.; Atzberger, C.; Mauser, W. Goodness-of-fit measures: What do they tell about vegetation variable retrieval performance from earth observation data. *Proc. SPIE* **2011**, *8174*, 81740R.

65. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models Part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282-290.

66. Archibald, R.; Fann, G. Feature selection and classification of hyperspectral images with support vector machines. *IEEE Geosci. Remote Sens. Lett.* **2007**, *4*, 674-677.

67. Baret, F. Biophysical Vegetation Variables Retrieval from Remote Sensing Observations. In *Proceedings of Remote Sensing for Agriculture, Ecosystems, and Hydrology XII*, Toulouse, France, 20–22 September 2010; Neale, C.M.U.; Maltese, A., Eds.; SPIE: Bellingham, WA, USA, 2010; Volume 7824, p. xvii-xix, front matter.

68. Lauvernet, C.; Baret, F.; Hascoet, L.; Buis, S.; Le Dimet, F.-X. Multitemporal-patch ensemble inversion of coupled surface-atmosphere radiative transfer models for land surface characterization. *Remote Sens. Environ.* **2008**, *112*, 851-861.

69. Bach, H.; Mauser, W. Methods and examples for remote sensing data assimilation in land surface process modeling. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1629-1637