Estimating leaf mass per area and equivalent water thickness based on leaf optical properties: potential and limitations of physical modeling and machine learning.

Jean-Baptiste Féret, G. Le Maire, S. Jay, D. Berveiller, Ryad Bendoula, G. Hmimina, Anice Cheraiet, J.C. Oliveira, F.J. J Ponzoni, T. Solanki, et al.

Estimating leaf mass per area and equivalent water thickness based on leaf optical properties: potential and limitations of physical modeling and machine learning.

J.-B. Féret[1], G. le Maire[2,3,4], S. Jay[5], D. Berveiller[6], R. Bendoula[7], G. Hmimina[6], A. Cheraiet[6], J.C. Oliveira[8], F.J. Ponzoni[9], T. Solanki[10], F. de Boissieu[1], J. Chave[11], Y. Nouvellon[2,3,12], A. Porcar-Castell[10], C. Proisy[13,14], K. Soudani[6], J.-P. Gastellu-Etchegorry[15], M.-J. Lefèvre-Fonollosa[16]

[1]TETIS, Irstea, AgroParisTech, CIRAD, CNRS, Université Montpellier, Montpellier, France

[2]CIRAD, UMR ECO&SOLS, Montpellier, France

[3]Eco&Sols, Univ Montpellier, CIRAD, INRA, IRD, Montpellier SupAgro, Montpellier, France.

[4] Interdisciplinary Center of Energy Planning (NIPE), UNICAMP, 13083-896 Campinas, Brazil

[5]Aix Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, F-13013 Marseille, France

[6]Ecologie Systematique Evolution, University of Paris-Sud, CNRS, AgroParisTech, Université Paris Saclay, F-91400 Orsay, France

[7]ITAP, Irstea, Montpellier SupAgro, Université Montpellier, Montpellier, France

[8]School of Agricultural Engineering - FEAGRI, University of Campinas, São Paulo, Brazil

[9]Instituto Nacional de Pesquisas Espaciais, Sao Jose dos Campos 12227-010, Brazil

[10]Optics of Photosynthesis Laboratory, INAR/Forests, Faculty of Agriculture and Forestry, 00014 University of Helsinki, Finland

[11]Laboratoire Evolution et Diversité Biologique UMR 5174, CNRS, Université Paul Sabatier, Toulouse, France

[12]University of Sao Paulo, ESALQ/USP, Piracicaba 13418-900, Brazil

[13]AMAP, IRD, CIRAD, CNRS, INRA, Univ. Montpellier, Montpellier, France

[14]GEOSMIT, French Institute of Pondicherry, Pondicherry, India

[15]Centre d'Etudes Spatiales de la Biosphère, Toulouse 31400, France

[16]CNES

Abstract

Leaf mass per area (LMA) and leaf equivalent water thickness (EWT) are key leaf functional traits providing information for many applications including ecosystem functioning modeling and fire risk management. In this paper, we investigate two common conclusions generally made for LMA and EWT estimation based on leaf optical properties in the near-infrared (NIR) and shortwave infrared (SWIR) domains: (1) physically-based approaches estimate EWT accurately and LMA poorly, while (2) statistically-based and machine learning (ML) methods provide accurate estimates of both LMA and EWT.

Using six experimental datasets including broadleaf species samples of more than 150 species collected over tropical, temperate and boreal ecosystems, we compared the performances of a physically-based method (PROSPECT model inversion) and a ML algorithm (support vector machine regressions, SVM) to infer EWT and LMA based on leaf reflectance and transmittance. We assessed several merit functions to invert PROSPECT based on iterative optimization and investigated the spectral domain to be used for optimal estimation of LMA and EWT. We also tested several strategies to select the training samples used by the SVM, in order to investigate the generalization ability of the derived regression models.

We evidenced that using spectral information from 1700 to 2400 nm leads to strong improvement in the estimation of EWT and LMA when performing a PROSPECT inversion, decreasing the LMA and EWT estimation errors by 55% and 33%, respectively.

The comparison of various sampling strategies for the training set used with SVM suggests that regression models show limited generalization ability, particularly when the regression model is applied on data fully independent from the training set. Finally, our results demonstrate that, when using an appropriate spectral domain, the PROSPECT inversion outperforms SVM trained with experimental data for the estimation of EWT and LMA. Thus we recommend that estimation of

LMA and EWT based on leaf optical properties should be physically-based using inversion of reflectance and transmittance measurements on the 1700 to 2400 nm spectral range.

1. INTRODUCTION

Global climate change and biodiversity loss strongly impact species and ecosystem functions, which directly influences processes at landscape and regional scales, and disrupts global biogeochemical cycles (Chapin, 2003). These ecosystem functions are tightly connected with species composition and can be partly described and explained using plant traits (Diaz and Cabido, 2001; Eviner and Chapin, 2003). By definition, plant traits correspond to morphological, physiological or phenological features measurable at the individual level, and functional traits are defined as these features impacting individual fitness via their effects on growth, reproduction and/or survival, the three components of individual performance (Violle et al., 2007). Therefore, our understanding of the interactions between climate, human activity and ecosystem functioning strongly depends on our capacity to monitor critical functional traits across space and time (Asner and Martin, 2016).

Leaf mass per area (LMA) is defined as the ratio of leaf dry mass (DW) to leaf area (A):

$$LMA = \frac{DW}{A}\ (mg.cm^{-2}) \qquad \text{Eq. 1}$$

It is a plant functional trait widely used as an indicator of plant functioning and ecosystem processes. In the leaf economic spectrum theory, the biophysical constraints explain the high coordination between organs properties and available resources: for instance, plants that have high trunk water conductivity generally have high stomatal conductance, low LMA and high photosynthetic capacities, developed root system and nutrient uptake, high turnover rate of resource acquisition organs, high growth rates. LMA is therefore a very significant trait because it correlates with key plant functional properties (de la Riva et al., 2016; Oren et al., 1986; Reich et al., 1997), therefore capturing a great proportion of the functional variation in the ecosystem.

LMA is important for the description of plant strategies and photosynthetic capacity over various vegetation types and climates (Asner et al., 2011; Gratani and Varone, 2006; Osnas et al., 2013; Puglielli et al., 2015; Reich et al., 1997, 1998; Weng et al., 2017). It is also a predictor of relative growth rate (Antúnez et al., 2001; Rees et al., 2010) and is usually correlated with mass-based maximum photosynthetic rate (Wright et al. 2004). At broader scales, it is also identified as a critical plant trait for the global monitoring of functional diversity, and for the determination of species fitness in their environment, affecting various ecosystem processes (Poorter et al., 2009; Schimel et al., 2015). Measurement of LMA is also relevant for many other applications, such as fire risk assessment (Cornelissen et al., 2017). Finally, LMA allows the conversion of traits expressed on an area basis into mass basis and vice versa. This is important since physical models usually express leaf constituent content per surface unit, whereas ecologists and plant physiologists may use constituent content per surface unit or per mass unit (Osnas et al., 2013; Wright et al., 2004).

The second important functional trait discussed in this study is the equivalent water thickness (EWT), defined as:

$$EWT = \frac{FW - DW}{A}\ (mg.cm^{-2}) \quad \text{Eq. 2}$$

with FW the leaf fresh mass. EWT is the area-weighted moisture content. It is related to a range of physiological and ecosystem processes, including leaf-level tolerance to dehydration, and ecological strategy. Indeed, species with large EWT tend to have lower construction costs, and are predominantly fast-growing and pioneer species (Wright et al., 2004).

The ability to accurately estimate both EWT and LMA is also critical for applications such as fire danger assessment: fuel moisture content (FMC, Chuvieco et al., 2002), also referred to as gravimetric water content (GWC, Datt, 1999), is a critical variable affecting fire interactions with fuel (Yebra et al., 2013). The accurate estimation of FMC is usually limited by the uncertainty associated to the estimation of LMA (Riano et al., 2005). Destructive measurements of LMA and EWT are time-consuming and logistically complex in remote environments. Alternative methods based on leaf spectroscopy have showed good performances for the estimation of various constituents (Asner et al., 2011, 2009; Ceccato et al., 2001; Colombo et al., 2008; Feilhauer et al., 2015; Féret et al., 2017; Fourty and Baret, 1998). Two main types of methods have been developed for the estimation of vegetation properties from their optical properties (including leaf chemistry but also canopy biophysical properties): *physically-based* methods and *data-driven* methods, also referred to as “radiometric data-driven approaches” and “biophysical variable driven approaches” respectively, by Baret and Buis (2008). In this study, we will only use the terms physically-based methods and data-driven methods in order to avoid confusion.

Physically-based methods are based on radiative transfer models (RTM) providing a mechanistic link between leaf traits and their optical properties. They aim at minimizing the residuals between measured and modeled radiometric data (hence the term “radiometric data-driven approach” by Baret and Buis, 2008). The PROSPECT model (Jacquemoud and Baret, 1990; Féret et al., 2017) is the most widespread model, due to its relative simplicity and computational efficiency combined with excellent modeling performances for a broad range of leaf types. Several retrieval algorithms have been developed to estimate leaf chemistry from their optical properties, taking advantage of physical modeling. These include look-up-table (LUT) methods (Ali et al., 2016) and iterative optimization based on minimization algorithms (Jacquemoud et al., 1996). Physically-based methods do not require calibration data, but they are computationally demanding.

Data-driven methods use a calibration dataset of measured leaf optical properties and traits in order to adjust regression models for the estimation of leaf chemistry (Verrelst et al., 2016). These include regression models derived from spectral indices, one of the most classic approaches (Gitelson et al., 2006; Main et al., 2011). More complex multivariate methods such as partial least square regression (Asner et al., 2011), and machine learning algorithms (ML) are also extensively used in the domain of remote sensing. These include support vector machine (SVM, Cortes and Vapnik, 1995; Drucker et al., 1996), random forest (Breiman, 2001), and artificial neural networks (Hornik et al., 1989). ML algorithms have been extensively used for remote sensing applications during the past decades, most of them at the canopy level when it comes to the estimation of biochemical constituents (Brown et al., 2000; Gualtieri, 2009; Lardeux et al., 2009; le Maire et al., 2011; Schmitter et al., 2017; Stumpf and Kerle, 2011; Zhang et al., 2017) , and a limited number of studies focusing on the leaf/needle scale (Conejo et al., 2015; Dawson et al., 1998; le Maire et al., 2004). ML algorithms usually show good performances in terms of prediction ability and high computational efficiency. The capacity of data-driven approaches to accurately predict leaf chemistry from their optical properties is inherently dependent on the dataset used to train the algorithm and regression model. The experiments performed in this study aim at quantifying this assertion over an extensive experimental dataset. This implies that correct implementation of data-driven methods using experimental data for training requires substantial efforts for the measurement of leaf optical properties and chemical constituents with destructive methods, whereas physical modeling only requires leaf optical properties.

Note that a third type of approach, namely, *hybrid* methods, could also be mentioned here (Verrelst et al., 2015). Such methods use data-driven algorithms trained with spectral properties simulated with physical models. These methods are particularly developed at the canopy scale, and combine the advantages of physically-based and data-driven methods: they do not require destructive measurements to build an experimental training dataset, and they are computationally efficient.

LMA and EWT both influence leaf optical properties in the near-infrared (NIR) and shortwave infrared (SWIR) domains (Bowyer and Danson, 2004). However, physically-based methods have often been reported to perform poorly for the estimation of LMA (Colombo et al., 2008; le Maire et al., 2008; Riano et al., 2005; Wang et al., 2011). Several reasons have been mentioned in the literature, including suboptimal modeling (Qiu et al., 2018), optical data collection (Merzlyak et al., 2004) or inversion (Colombo et al., 2008; Qiu et al., 2018; Riano et al., 2005; Sun et al., 2018; Wang et al., 2011, 2015).

A first reason related to modeling is that the influence of LMA on the optical properties modeled by PROSPECT is defined by a single specific absorption coefficient (SAC), although various non-pigment organic materials (cellulose, hemicellulose, lignin, proteins, starch) influence leaf optics individually (Jacquemoud et al., 1996). Therefore, this single SAC assumes that the relative proportion of each of these single constituents is constant among leaves, which may not be the case. Another reason may be due to an imperfect modeling of light propagation within the leaf. From that perspective, Qiu et al. (2018) proposed a refined version of PROSPECT (named PROSPECT-g) including an anisotropic-scattering factor in order to improve the estimation of LMA, and developed an iterative inversion procedure specifically dedicated to this model.

Experimental uncertainty should also be considered when discrepancies between measurements and simulations are observed. Indeed, accurately measuring leaf optical properties remains challenging despite the high performances of field and lab spectroradiometers, leading to possible experimental bias which is usually unaccounted for. As an example, Merzlyak et al. (2004) reported the difficulty to accurately measure leaf optical properties in the NIR domain due to incomplete collection of the light leaving the highly scattering tissue. They proposed a correcting factor for transmittance based on the hypothesis that leaf absorption in the NIR domain is negligible. For these reasons, the relevance of systematically using the full spectral domain (especially the NIR domain) can be questioned.

Finally, several authors suggested that classical least-squares inversion based on the use of leaf reflectance and transmittance over the full spectral domain was suboptimal for physically-based estimation of LMA, especially due to the lower influence of LMA on leaf optical properties in the SWIR domain as compared to EWT (Colombo et al., 2008; Riano et al., 2005). More elaborated inversion procedures have thus been proposed to improve LMA estimation. Some of them are based on complex iterative procedures consisting in successively estimating different PROSPECT parameters using unweighted merit functions computed over specific spectral domains (Qiu et al., 2018 ; Li and Wang, 2011 ; Wang et al., 2015). When using the full spectral domain from 400 to 2500 nm, Sun et al. (2018) showed that LMA estimation based on PROSPECT inversion and an unweighted merit function was more accurate when using only reflectance or only transmittance instead of reflectance plus transmittance. When using bidirectional reflectance measurements, Li et al. (2018) developed an approach (PROCWT) coupling PROSPECT with continuous wavelet transform in order to suppress surface reflectance effects. PROCWT was shown to perform better than PROSPECT and a simplified version of PROCOSINE (Jay et al., 2016) for the estimation of LMA.

All of these studies demonstrate the complexity of a direct estimation of LMA from leaf optical properties using physically-based methods, and the difficulty to clearly identify the origin of current limitations. In the case of data-driven methods, the estimation of LMA has seldom been investigated comprehensively: training and test data are usually collected following a unique protocol specific to a unique set of equipment and by the same team of operators. This means that possible experimental biases due to protocol, equipment and/or operators may be embedded into the resulting regression model, leading to poor generalization ability when applied to independent datasets collected under different conditions or with different equipment.

The objective of this study is to assess the relative performances of physically-based and data-driven approaches for the estimation of LMA and EWT based on leaf optical properties. Our working questions are (1) what are the limitations of PROSPECT for LMA and EWT estimation, and is there any solution to overcome these limitations, and (2) what is the generalization ability of data-driven approaches when independent datasets are used for training and validation? We gathered six datasets in temperate, tropical and boreal ecosystems, with joint measurements of broadleaf optical properties, LMA and EWT (Section 2). Then, we designed specific protocols to address questions (1) and (2), and to perform an objective comparison of their performances (Section 3). This includes the selection of specific spectral information for PROSPECT inversion, and different strategies for the sampling of the training dataset for ML algorithms. Section 4 presents the results obtained with the different approaches, including a comparison of the validation with the six experimental datasets. Finally, section 5 discusses the potential and current limitations of the approaches and section 6 provides a conclusion.

2. MATERIALS

a. Global description of the datasets

For this study, six datasets were collected over various ecoregions, ranging from tropical forests, to temperate and boreal ecosystems (Table 1). LOPEX and ANGERS are publicly available and used in many publications. HYYTIALA, ITATINGA, NOURAGUES and PARACOU are unpublished datasets.

- The ANGERS[1] dataset was collected in 2003 at INRA (Institut national de la recherche agronomique) in Angers (France). It encompasses physical measurements and biochemical

[1] http://opticleaf.ipgp.fr/index.php?page=database

analyses collected over 43 species and varieties of woody and herbaceous plants. ANGERS was used for the calibration of the SAC for chlorophylls, carotenoids and anthocyanins in the latest versions of PROSPECT (Féret et al., 2017, 2008).

- The Leaf Optical Properties Experiment (LOPEX[1,2]) dataset was collected in 1993 in Italy during a campaign conducted at the Joint Research Centre (Ispra, Italy) (Hosgood et al., 1994). It encompasses physical measurements and biochemical analyses collected over more than 50 species of woody and herbaceous plants, and has been widely used by the remote sensing community (Bowyer and Danson, 2004; Féret et al., 2008; Mobasheri and Fatemi, 2013; Romero et al., 2012). The full LOPEX dataset includes dry and fresh samples and was used for the calibration of the SAC of LMA (Féret et al., 2008), as well as broadleaf and needleleaf samples. However, only broadleaf samples were used in the current study, all fresh leaves except for one set of five dry maize leaf samples.
- The HYYTIALA dataset was collected in July 2017 at the Hyytiälä Forestry Field Station in Southern Finland in the frame of the Fluorescence Across Space and Time (FAST) campaign. This station is located in the boreal belt and is dominated by mixed forest of Scots pine, Norway spruce and silver birch. This dataset encompasses physical measurements and biochemical analyses collected over various native and non-native broadleaf species located in the field station.
- The ITATINGA dataset was collected in October 2015 as part of the IPEF-Eucflux project and HYPERTROPIK project (TOSCA, CNES, France), from experimental *Eucalyptus* stands planted in November 2009 near the University of São Paulo forestry research station at Itatinga Municipality (São Paulo State, southeastern Brazil). ITATINGA includes sixteen genotypes and four species of *Eucalyptus*, eventually with hybrids, provided by different forestry companies in

[2] http://teledetection.ipgp.jussieu.fr/opticleaf/lopex.htm

different regions of Brazil. For each genotype, leaves corresponding to various developmental stages were collected, from juvenile to mature to senescent, and various locations within the crown (shaded leaves from the lower part of the crown, leaves from mid crown and sunlit leaves from the upper part of the crown). This dataset is the only genus-specific dataset. Hence, in spite of the large variability in terms of developmental stages, the ranges of LMA and EWT show significantly lower variability than those observed for the other datasets (Table 1). See Oliveira et al. (2017) for more details.

- The NOURAGUES dataset was collected at the CNRS Nouragues experimental research station, French Guiana, in September 2015, in the frame of the HYPERTROPIK project. This site is a lowland Amazonian forest, protected since 1996 by a Natural Reserve status. This dataset includes four to ten leaf samples from 38 emerging tropical tree species, collected from both shaded and sunlit parts of the crown. The Nouragues station is also a pilot site for remote sensing studies of tropical ecosystems (Réjou-Méchain et al., 2015).
- The PARACOU dataset was collected at the CIRAD-INRA Paracou experimental research station, French Guiana, in September 2015 (HYPERTROPIK project). This dataset includes four to ten leaf samples from 28 emerging tropical tree species, collected from both shaded and sunlit parts of the crown. Paracou is located in coastal lowland Amazonian forest. Various experiments are ongoing, including disturbance experiments, CO_2 flux experiments, fertilization and long-term studies in forest dynamics and biodiversity.

b. Measurements of leaf optical properties

For all the samples, directional-hemispherical reflectance and transmittance (Schaepman-Strub et al., 2006) of the upper surface of the leaves were measured with a spectroradiometer and an integrating sphere in the visible (VIS), NIR and SWIR domains between 400 and 2500 nm. Here, we used the infrared domain ranging from 900 to 2400 nm, due to the low influence of LMA and EWT on leaf optical properties below 900 nm, and to the low signal-to-noise ratio (SNR) beyond 2400 nm. All datasets shared the same protocol for the measurement of leaf optical properties, and included spectral calibration for stray light in order to correct the imperfect collimation of the lamp beam as well as compensation for the optical properties of the coating of the integrating sphere when measuring leaf reflectance and transmittance (Asner et al., 2009; Carter and Knapp, 2001). The datasets were collected by different operators, and using different devices. Despite efforts to share a unique protocol for the acquisition of leaf optical properties, this diversity of operators, equipment and conditions of acquisition, is a possible source of bias that we discuss here.

c. Measurements of LMA and EWT

The measurement of EWT and LMA shared the same protocol among experimental datasets. Leaf samples were collected in the field, stored in a cooler and measured in an experimental facility equipped with a precision scale and a drying oven. Minutes after measuring the leaf optical properties, disks of fresh leaf material were sampled using a cork borer, and immediately weighted using the precision scale to obtain FW (Eq. 2). The disks were then placed in a drying oven at 85°C for at least 48 hours until constant mass was attained, and immediately weighted when out of the oven in order to determine DW (Eq. 1 and Eq. 2) (Cornelissen et al., 2003; Pérez-Harguindeguy et al., 2013). EWT and LMA were then computed based on Eq. 1 and Eq. 2.

Table 1 summarizes basic statistics and information for each dataset. LMA and EWT were systematically measured for each sample in each dataset, except for the PARACOU dataset which only includes LMA measurements. Similarly to optical properties, various sources of uncertainty may have affected EWT and LMA measurements, including errors in the area sampled on leaf material due to imperfect circular sampling disks, loss in water content between leaf optics measurements and weighting of fresh mass, or rehydration between drying and weighting of dry mass. However, care was paid to standardize data collection, so as to minimize the influence of these possible biases. EWT and LMA show no correlation for ITATINGA, weak correlation for LOPEX, moderate correlation for HYYTIALA and NOURAGUES, and strong correlation for ANGERS. A moderate correlation of 0.44 is measured when pooling all samples together.

Table 1. Summary of the main properties of the experimental datasets. Basic statistics for each dataset (minimum and maximum value, mean and standard deviation) are given for $\boldsymbol{EWT}$ and $\boldsymbol{LMA}$, as well as their correlation *r(*$\boldsymbol{EWT}, \boldsymbol{LMA}$*)*.

	ANGERS	**LOPEX**	**HYYTIALA**	**ITATINGA**	**NOURAGUES**	**PARACOU**
#Samples	308	330	96	415	262	272
#Species/genotypes	43 sp.	46 sp.	10 sp.	4 sp. /16 gt.*	38 sp.	28 sp.
$\boldsymbol{EWT}$ **(mg.cm^{-2})**						
Min – Max	4.40 – 34.00	0.29 –52.48	3.68 – 23.73	2.20 – 20.20	3.20 – 38.10	N/A
Mean ± SD	11.47 ± 4.70	11.13 ± 6.97	9.16 ± 2.98	14.44 ± 2.09	11.73 ± 4.86	N/A
$\boldsymbol{LMA}$ **(mg.cm^{-2})**						
Min – Max	1.66 – 33.10	1.71 – 15.73	2.76 – 15.77	6.90 – 14.70	3.10 – 21.10	5.28 – 25.56
Mean ± SD	5.12 ± 3.53	5.29 ± 2.47	6.27 ± 3.04	10.24 ± 1.62	10.81 ± 3.89	12.32 ± 4.06
r($\boldsymbol{EWT}, \boldsymbol{LMA}$ **)**	0.72	0.28	0.40	0.03	0.51	N/A

* Four species from Eucalyptus genus, corresponding to sixteen genotypes

3. METHODS

a. PROSPECT model: general presentation

PROSPECT is based on the generalized plate model (Allen et al., 1969, 1970) and was initially developed by Jacquemoud and Baret (1990). This model simulates the leaf directional-hemispherical reflectance and transmittance (Schaepman-Strub et al., 2006) with a limited number of input biophysical and biochemical variables, including various absorbing compounds and a unique leaf structure parameter, named N. Many versions have been developed since the first version, in order to include more absorbing compounds (Féret et al., 2017, 2008; Jacquemoud et al., 1996) or to adapt to specific conditions and leaf types, such as needle-shaped leaves (Malenovský et al., 2006). In this study, we used the latest version of PROSPECT, named PROSPECT-D (Féret et al., 2017). As we focused on leaf optical properties in the 900 – 2400 nm range, the capability of PROSPECT in terms of separation of pigments was not critical as no pigment absorbs in this spectral domain, but the refractive index differs from the one used on PROSPECT-5 (Féret et al., 2008). Brown pigments were not retrieved during the inversion, as including them showed no significant difference in the results obtained for any of the strategies tested here.

The N parameter corresponds to the number of uniform compact plates separated by $N-1$ air spaces. The value of N represents the complexity of the leaf internal structure, with low N values corresponding to moderate complexity such as in monocots, and higher N values corresponding to higher complexity, a characteristic of dicots. To date, no protocol exists to experimentally estimate N from leaf samples, other than using leaf optical properties. N influences leaf scattering and shows negligible impact on leaf absorption: increasing N values increase reflectance and decrease transmittance, and N shows particularly strong effects in domains with low absorption, such as the NIR domain. Recently, Qiu et al. (2018) found an extremely strong correlation between N and the ratio between reflectance and transmittance on simulated data.

PROSPECT can be run in forward or inverse mode. The forward mode aims at simulating leaf optical properties based on a full set of biophysical and biochemical properties (leaf chemistry and N). The inverse mode aims at identifying the optimal set of biophysical and biochemical properties that minimize a merit function (or goodness-of-fit criterion) based on a comparison between measured and simulated leaf optics. A common inversion procedure is based on the numerical minimization of the sum of weighted square errors over all spectral bands available. The corresponding merit function M is expressed as follows when using both reflectance and transmittance:

$$M\left(N,\{\ C_i\}_{i=1:p}\right) = \sum_{\lambda=\lambda_1}^{\lambda_n}\left[W_{R,\lambda} \times \left(R_\lambda - \hat{R}_\lambda\right)^2 + W_{T,\lambda} \times \left(T_\lambda - \hat{T}_\lambda\right)^2\right] \quad \text{Eq. 3}$$

with N the leaf structure parameter, p the number of chemical constituents accounted for by PROSPECT and retrieved during the inversion, C_i the biochemical content per leaf surface unit for constituent i, λ_1 and λ_n the first and last wavebands investigated for inversion, R_λ and T_λ the experimental reflectance and transmittance measured at waveband λ, $\hat{R}_\lambda$ and $\hat{T}_\lambda$ the reflectance and transmittance simulated by PROSPECT with $\left\{N,\{\ C_i\}_{i=1:p}\right\}$ as input variables, $W_{R,\lambda}$ the weight applied to the squared difference between experimental and simulated reflectances, and $W_{T,\lambda}$ its equivalent for transmittance. Eq. 3 can be used to estimate the full set of input variables, or a limited subset if prior information or arbitrary value is set for some variables.

b. Estimation of EWT and LMA through iterative optimization

The large majority of the studies focusing on leaf scale model inversions through iterative optimization used Eq. 3 with unweighted merit function over the full spectral domain available ($W_{R,\lambda} = W_{T,\lambda} = 1$).This merit function provides accurate estimates of leaf pigments and EWT (Féret et al., 2017; Jacquemoud et al., 1996; Newnham and Burt, 2001), but several studies reported poor results for LMA estimation (Féret et al., 2008; Riano et al., 2005). Colombo et al. (2008) used an alternative weighting, with $W_{R,\lambda} = (R_\lambda)^{-2}\ and\ W_{T,\lambda} = (T_\lambda)^{-2}$, which is otherwise unused in the literature when inverting leaf models, and not so common when inverting canopy models (Baret and Buis, 2008). In practice, implementing such a merit function requires precaution as high sensor noise (in particular in the SWIR domain) may result in close-to-zero reflectance and transmittance, leading to exaggerated importance of the corresponding spectral bands. This merit function then needs to be adapted to exclude these spectral bands. Colombo et al. (2008) reported fair performances of this merit function for the estimation of EWT, but poor performances for LMA. However, the SWIR domain beyond 1600 nm was not measured for their study, in spite of its importance for the estimation of LMA (Asner et al., 2011, 2009; le Maire et al., 2008). Therefore a fair comparison between this merit function and the unweighted merit function including the full spectral range is required.

As mentioned in the introduction, LMA estimation could also be improved by focusing on optimal spectral ranges (Li and Wang, 2011; Qiu et al., 2018; Wang et al., 2015). This amounts to choosing the weights such that $W_{R,\lambda} = W_{T,\lambda} = 1$ in the considered range, and $W_{R,\lambda} = W_{T,\lambda} = 0$ elsewhere. Note that such a procedure is relatively straightforward and could potentially be applied to the canopy scale in a similar way.

In this study, three inversion procedures were applied to the six independent experimental datasets, and their relative performances were compared. These inversion procedures correspond to "one-step" procedures, aiming at estimating EWT, LMA and N simultaneously from both reflectance and transmittance:

- *Iterative optimization 1 (IO1)* uses an unweighted merit function ($W_{R,\lambda} = W_{T,\lambda} = 1$) with reflectance and transmittance defined from 900 nm to 2400 nm.
- *Iterative optimization 2 (IO2)* uses a weighted merit function as defined by Colombo et al. (2008) ($W_{R,\lambda} = (R_\lambda)^{-2}$ *and* $W_{T,\lambda} = (T_\lambda)^{-2}$) with reflectance and transmittance defined from 900 nm to 2400 nm.
- *Iterative optimization 3 (IO3)* uses a weighted merit function defined by $W_{R,\lambda} = W_{T,\lambda} = 1$ over an optimal contiguous spectral domain $[\lambda_1, \lambda_n]$ defined between 900 and 2400 nm, and

$W_{R,\lambda} = W_{T,\lambda} = 0$ elsewhere. This optimal spectral domain is adjusted in the present study and is the same for both reflectance and transmittance, and for all experimental datasets.

In the case of *IO3*, the exhaustive comparison of all combinations of spectral domains or spectral bands is computationally too demanding and extremely inefficient given the strong correlations between neighboring spectral domains. In order to reduce the computational cost, we focused on contiguous spectral domains defined by partitioning the initial spectral domain into 15 evenly-sized segments of 100 nm from 900 to 2399 nm. The choice of 100 nm segments is driven by constraints in terms of computation and by the ability to identify the main absorption features of EWT and LMA individually. The performances of PROSPECT inversion for the estimation of LMA and EWT were tested with all continuous spectral domains that can be generated from these 15 spectral segments, leading to 120 continuous segments. Finally, the spectral domain leading to the minimum RMSE averaged for all experimental datasets and for the estimation of both LMA and EWT from PROSPECT inversion was selected and defined as the optimal spectral range used in *IO3.*

For *IO1*, *IO2* and *IO3*, N, EWT and LMA were simultaneously estimated using a constrained nonlinear optimization algorithm, i.e., the Sequential Quadratic Programming algorithm implemented within the Matlab function *fmincon*. The lower bounds selected for the three parameters to be optimized were defined to respect the condition of strict positivity and include minimum values observed for experimental data, whereas the upper bounds were set in order to include the maximum values observed for experimental data, with significant margins: EWT values were investigated between 0.01 and 80 mg.cm^{-2}; LMA values were investigated between 0.01 and 40 mg.cm^{-2}; N values were investigated between 0.5 and 4. No correlation constraints between EWT and LMA were included in the inversion procedure, since such correlation was not systematic between datasets.

c. Data-driven estimation of EWT and LMA

The performances of data-driven methods inherently depend on the training data. In most cases, these performances are reported after splitting an experimental dataset into training and validation subsets, and the resulting regression models are not validated on fully independent datasets. In the perspective of operational applications, this raises the question of the possibility to share regression models adjusted with ML algorithms on public experimental datasets, and to use leaf spectroscopy operationally with no destructive measurements required to adjust dataset-specific regression models. With increasing use of machine learning, software packages including already trained regression models may be shared the same way statistical models derived from spectral indices have been proposed in the scientific literature (Féret et al., 2011). We want to answer the following questions related to data-driven methods: do regression models trained with one or several experimental datasets perform well when applied on independent datasets, or should training data systematically include samples from the validation dataset? To answer these questions, three strategies for the composition of a training dataset were tested, and the performances of data-driven methods were compared with PROSPECT inversions:

- *Training sampling 1 (TS1)*: A single dataset was used as training data and the regression model was then applied on each of the remaining datasets.
- *Training sampling 2 (TS2):* All but one experimental datasets were used as training data, and the regression model was then applied on the remaining dataset.
- *Training sampling 3 (TS3)*: All experimental datasets were pooled into a single one, and 300 samples (comparable in size to individual datasets) were randomly selected for training. Validation was then performed on the remaining samples (1668 samples for LMA, and 1396 samples for EWT), and performances (in terms of RMSE) were evaluated per individual dataset and globally. In each case, to account for possible sampling bias, random sampling of training

dataset was repeated 20 times and the distribution of RMSE values across all samplings was calculated.

Here, these three strategies used to define the training dataset were used with support vector machine (SVM) regression algorithm corresponding to the *Matlab* implementation of the *LibSVM* library (Chang and Lin, 2011). Reflectance and transmittance measurements from 900 to 2400 nm were stacked in a unique vector, resulting in $n_\lambda = 3002$ predictor spectral variables for each sample. Reflectance and transmittance were scaled between 0 and 1 for each spectral band, as well as leaf chemical constituent of interest (LMA and EWT). The radial basis function (RBF) kernel was selected, which implies optimizing two free parameters, C and γ. C is a cost parameter used to trade error penalty for stability and common to any SVM model. γ is specific to RBF kernels and it corresponds to the inverse of the radius of influence of samples selected by the model as support vectors. The C and γ parameters were optimized using an exhaustive grid search ($C \in [10^{-2}; 10^{-1}; \dots; 10^{+2}], \gamma \in [10^{-5}; 10^{-4}; \dots; 10^{+1}]$ in order to include the default values recommended by Chang and Lin (2011) and a five-fold cross validation over the training data for each combination of C and γ. The optimal C and γ values were then used with the full training data to adjust a regression model.

4. RESULTS

This section is divided into three subsections. The first subsection aims at identifying the optimal spectral domain to be used with *IO3*. This first section is a prerequisite to the second section, which then focuses on the comparison between the three types of iterative optimization, and the two types of training samplings based on the integrality of experimental datasets, *TS1* and *TS2*. Finally, the third section compares the performances of *TS3*, which is based on a random sampling among all experimental datasets, with the performances of *IO3* and *TS2*, when the validation samples are identical to those used in *TS3*.

a. Influence of spectral domain used for the estimation of EWT and LMA with PROSPECT inversion (optimization of *IO3* method)

Figure 1 and Figure 2 show the results obtained for the estimation of EWT and LMA, respectively, when inverting PROSPECT over each dataset and each of the 120 spectral domains defined in Section 3.b with the *IO3* method. For the sake of comparison, for each dataset, the RMSE was normalized by the RMSE obtained when using the spectral information from 900 to 2400 nm, and this normalized RMSE (NRMSE) was expressed as a percentage. In the case of EWT, the optimal spectral domain excluded the NIR domain under 1300 nm for all datasets, but no unique optimal spectral domain common to each dataset could be identified. The relative improvement induced by the reduction of the spectral domain was also strongly dataset-dependent: NRMSE was reduced by 23 % (LOPEX) to 56 % (NOURAGUES).

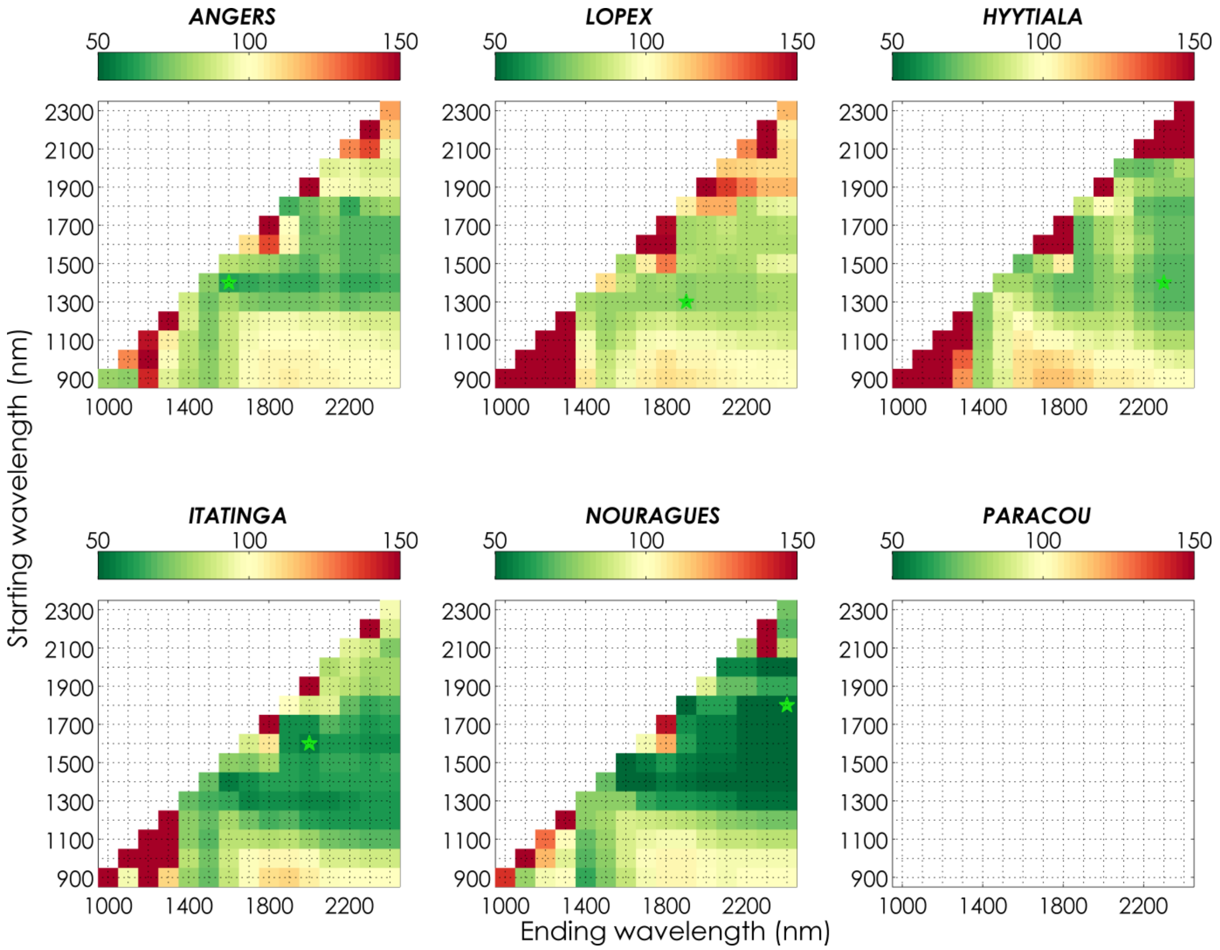

Figure 1. Normalized $RMSE$ ($NRMSE$, in %) obtained for EWT with PROSPECT inversion method *IO3* over each dataset and each reduced spectral domains bounded by a starting wavelength λ_1 (y-axis) and an ending wavelength λ_2 (x-axis). The normalization is specific to each dataset based on the performances of *IO1* (NRMSE=100%, lower right corner). The green star indicates the spectral segment producing the best results.

In the case of LMA, both optimal spectral domain and relative improvement or degradation showed stronger consistency among datasets than for EWT (Figure 2). For all datasets, excluding information from 1500 nm and beyond led to strong degradations of the performances. In the case of LOPEX and HYYTIALA, estimation of LMA could be improved only when using spectral domains with ending wavelength between 2100 and 2400 nm, except when using a narrow spectral domain from 1600 to 1800 nm. For the four other datasets, extended spectral combinations led to improved LMA, as most of the combinations excluding the domain from 900 to 1200 nm led to improved estimation of LMA, except when using a reduced spectral domain ranging from 1800 to 2100 nm only, which corresponds to one of the main absorption features of water. Overall, the optimal spectral range excluded the NIR domain and included spectral information until 2400 nm for all datasets. The relative improvement induced by the selection of an optimal specific for each dataset ranged from 60 (ITATINGA) to 67 % (NOURAGUES).

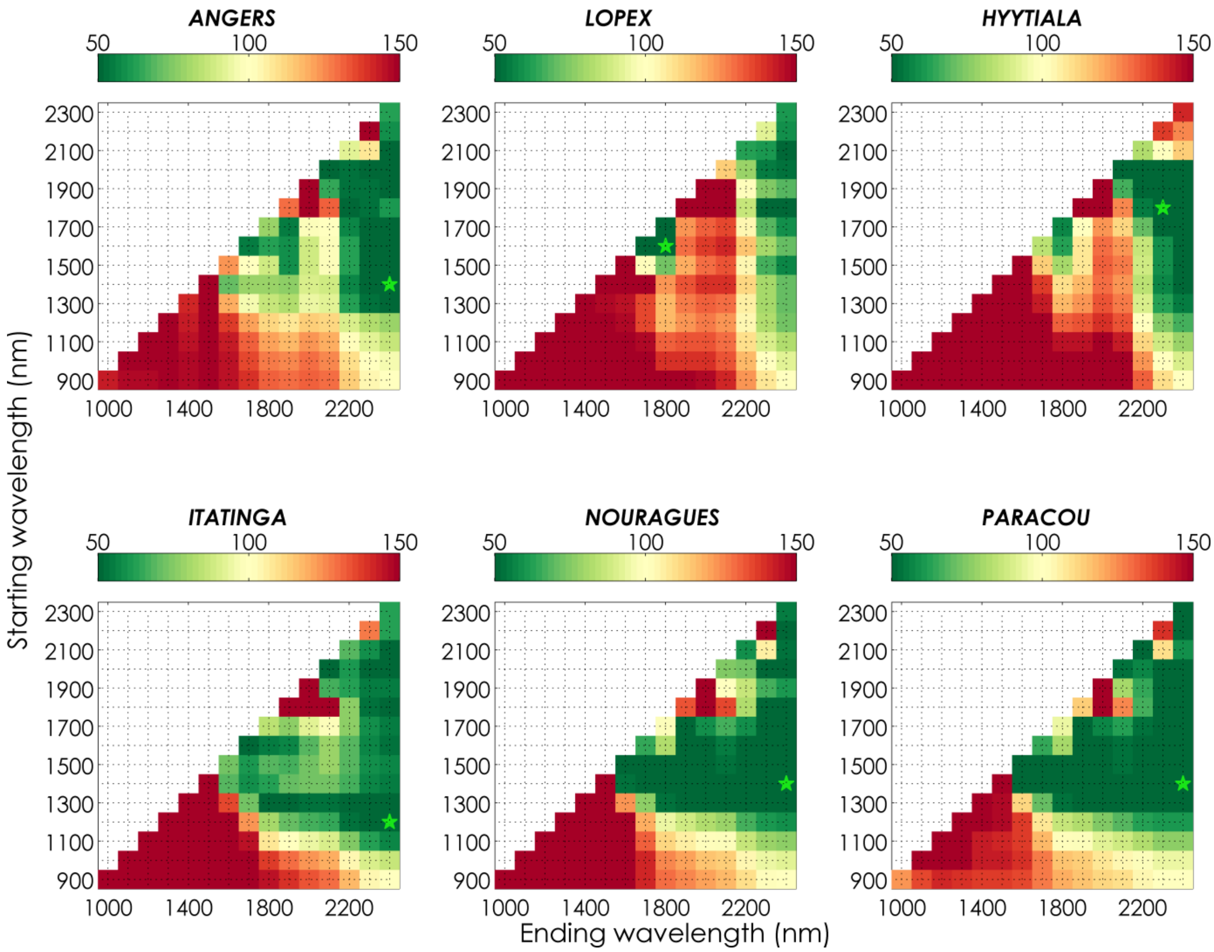

Figure 2. Normalized $RMSE$ (NRMSE, in %) obtained for LMA estimation with PROSPECT inversion method *IO3*, over each dataset and each reduced spectral domains bounded by a starting wavelength λ_1 (y-axis) and an ending wavelength λ_2 (x-axis). The normalization is specific to each dataset based on the performances of *IO1* (NRMSE=100%, lower right corner). The green star indicates the spectral segment producing the best results.

These figures provide a visual representation of the spectral domains leading to improved or decreased performances compared to full spectral information. They confirm that selecting the appropriate spectral information during inversion strongly influences for the estimation of leaf constituents.

Figure 3 provides NRMSE for the estimation of EWT and LMA averaged over all datasets, and confirms suboptimal performances obtained when using NIR information only. Overall, the spectral domain ranging from 1700 to 2400 nm was found to be optimal when estimating EWT and LMA simultaneously (mean NRMSE was reduced by 33% for EWT and by 55 % for LMA), and was used hereafter within the *IO3* method.

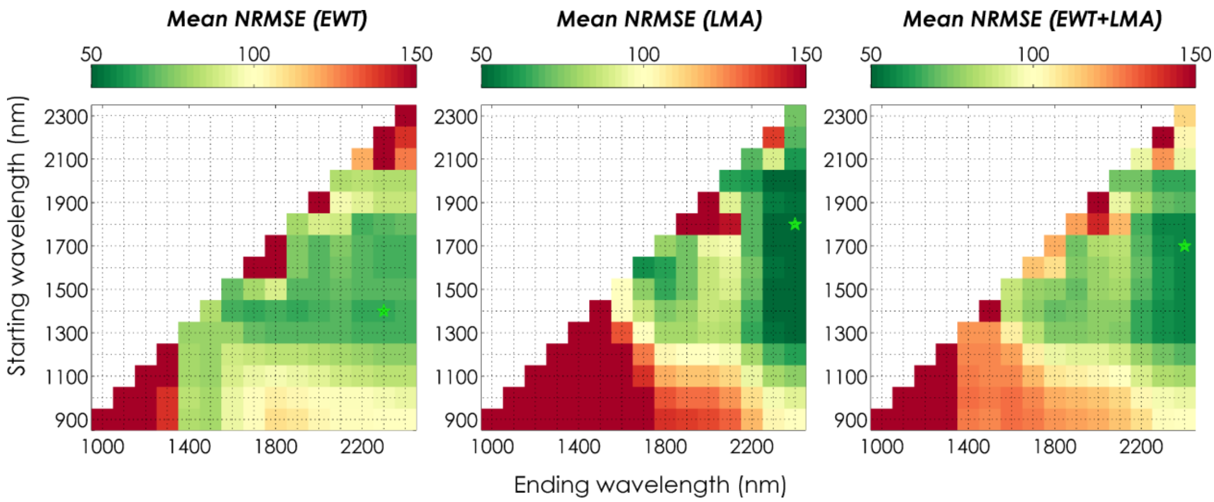

Figure 3. Mean normalized RMSE values (NRMSE, in %) obtained for the estimation of EWT (left), LMA (center), and both constituents (right), after PROSPECT inversion over all experimental datasets pooled and each of the 120 spectral domains defined in Section 3.b. The green star indicates the spectral segment producing the best results.

b. Comparison of PROSPECT inversion methods and ML algorithms for the estimation of LMA and EWT: training ML with independent datasets

The performances obtained for the estimation of EWT when using *TS1* and *TS2* for ML regression, and *IO1*, *IO2* or *IO3* (with the 1700 – 2400 nm spectral range) for PROSPECT inversion are reported in Table 2. Overall, *IO2* and *IO3* produced the most consistent results, systematically outperforming the other methods. ML regressions performed particularly poorly compared to *IO2* and *IO3*, and *TS2* led to the better results than *TS1* (except form HYYTIALA). *TS1* led to very inconsistent results, with 175% increase compared to *IO2* and *IO3* on average, and up to 500% increase in RMSE compared to PROSPECT inversion *IO2* when estimating EWT from ITATINGA after training with LOPEX.

Table 2. RMSE values (in $mg.cm^{-2}$) obtained for the estimation of EWT with SVM and training strategies *TS1* and *TS2*, and with *IO1*, *IO2* and *IO3*. For each column (validation dataset), the minimum RMSE is indicated in bold, and colors correspond to the level of performances, from green color for minimum RMSE to red color for maximum RMSE.

Method	Train \ Valid	ANGERS	LOPEX	HYYTIALA	ITATINGA	NOURAGUES	PARACOU
TS1	ANGERS	-	4.82	1.90	3.31	2.49	-
	LOPEX	3.14	-	3.23	6.73	2.32	-
	HYYTIALA	3.79	5.40	-	2.84	3.92	-
	ITATINGA	3.38	5.82	3.03	-	3.43	-
	NOURAGUES	2.54	5.04	3.15	2.47	-	-
	PARACOU	-	-	-	-	-	-
TS2	All but 1	2.47	4.54	2.68	2.08	2.10	-
IO1	PROSPECT	2.07	2.03	1.72	1.93	3.44	-
IO2	PROSPECT	1.48	**1.68**	1.44	**1.13**	**1.21**	-
IO3	PROSPECT	**1.41**	1.70	**1.21**	1.20	1.66	-

Figure 4 provides scatterplots for the results showed in Table 2 and corresponding to *IO1*, *IO2*, *IO3* and SVM regression with sampling strategy *TS2*. Overall, *IO2* showed the best performances for the estimation of EWT, and SVM regression produced the lowest performances, mainly because of the strong error obtained for extreme values on LOPEX.

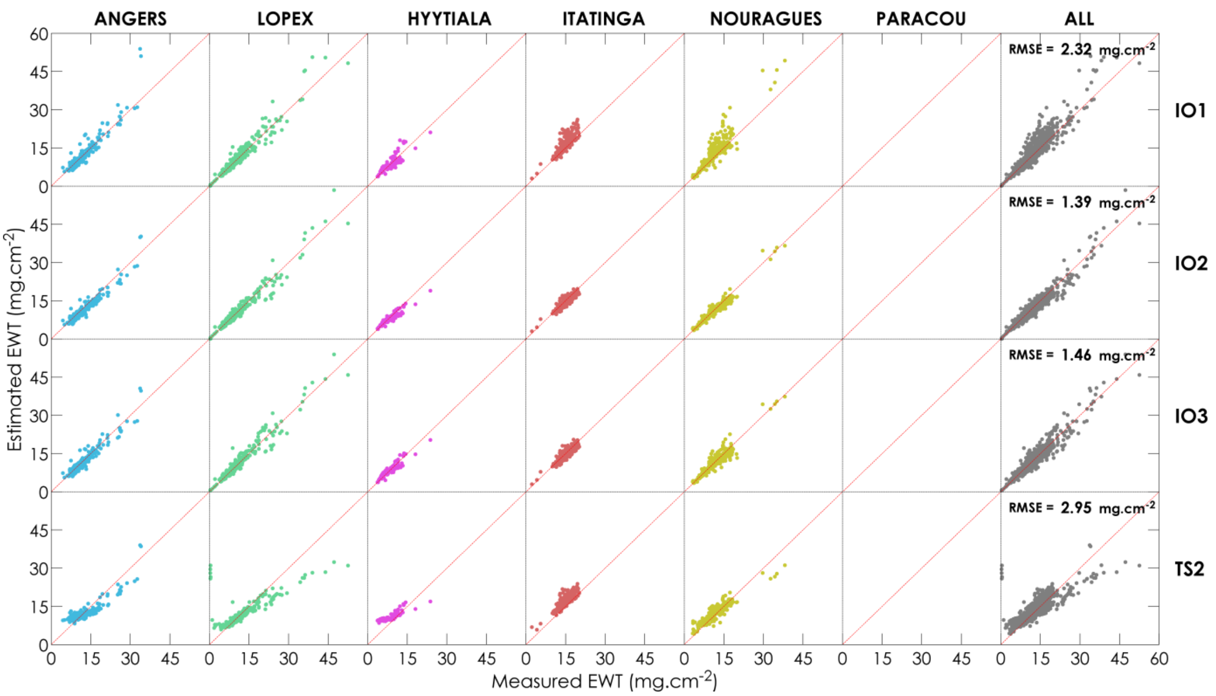

Figure 4. EWT estimation results obtained using PROSPECT inversion (*IO1*, *IO2*, *IO3*) and ML regression (training sampling *TS2*).

The performances obtained for the estimation of LMA when using training samplings *TS1* and *TS2* for ML regression, and *IO1*, *IO2* or *IO3* for PROSPECT inversion are reported in Table 3. *IO3* outperformed the other methods for all datasets except HYYTIALA and ITATINGA: *IO2* slightly outperformed *IO3* for ITATINGA only and *TS2* outperformed *IO3* for HYYTIALA and ITATINGA. However, the difference in RMSE between *IO3* and the optimal method remained less than 20% for these two datasets. The relative performances obtained with *IO1* and *IO2* differed among datasets:

while using *IO2* led to significantly improved estimation of LMA compared to *IO1* for five datasets (from a 26% decrease in RMSE for LOPEX to more than 50% for ITATINGA, NOURAGUES and PARACOU), and slightly degraded estimation compared to *IO3* for four datasets, the performances obtained for HYYTIALA were degraded by more than 75% compared to *IO1*, with systematic strong overestimation (Figure 5). On the other hand, the RMSE corresponding to estimation of LMA using *IO3* decreased by 60% compared to *IO1*. ML regression trained with *TS2* performed better than *IO1* overall but was outperformed by *IO2* and *IO3*. As for EWT, ML trained with *TS1* led to very inconsistent results, and was strongly outperformed by *IO2*, *IO3* and ML regressions trained with strategy *TS2* in most cases.

Table 3. RMSE values (in mg.cm^{-2}) obtained for the estimation of LMA with SVM and training samplings *TS1* and *TS2*, and with *IO1*, *IO2* and *IO3*. For each column (validation dataset), the minimum RMSE is indicated in bold, and colors correspond to the level of performances, from green color for minimum RMSE to red color for maximum RMSE.

Method	Valid / Train	ANGERS	LOPEX	HYYTIALA	ITATINGA	NOURAGUES	PARACOU
TS1	ANGERS	-	4.91	2.49	3.73	2.76	2.70
	LOPEX	2.92	-	2.18	4.33	4.85	5.86
	HYYTIALA	2.51	2.47	-	3.19	3.32	4.00
	ITATINGA	6.27	5.57	5.08	-	3.86	4.50
	NOURAGUES	4.04	4.82	3.74	1.40	-	2.21
	PARACOU	2.96	3.96	2.30	1.25	2.11	-
TS2	All but 1	2.31	4.06	**1.33**	1.23	2.14	2.41

IO1	PROSPECT	2.48	3.36	3.49	2.60	3.95	4.75
IO2	PROSPECT	1.24	2.48	6.12	**1.20**	1.71	2.25
IO3	PROSPECT	**0.93**	**1.99**	1.52	1.44	**1.59**	**1.73**

Figure 5 provides scatterplots for the results showed in Table 3 and corresponding to *IO1*, *IO2*, *IO3* and SVM regression with training sampling *TS2*. Overall, *IO3* produced the most accurate estimation of LMA. *IO2*, *IO3* and *TS2* respectively resulted in 33%, 55% and 27% decreases in RMSE for the estimation of LMA when compared to *IO1*.

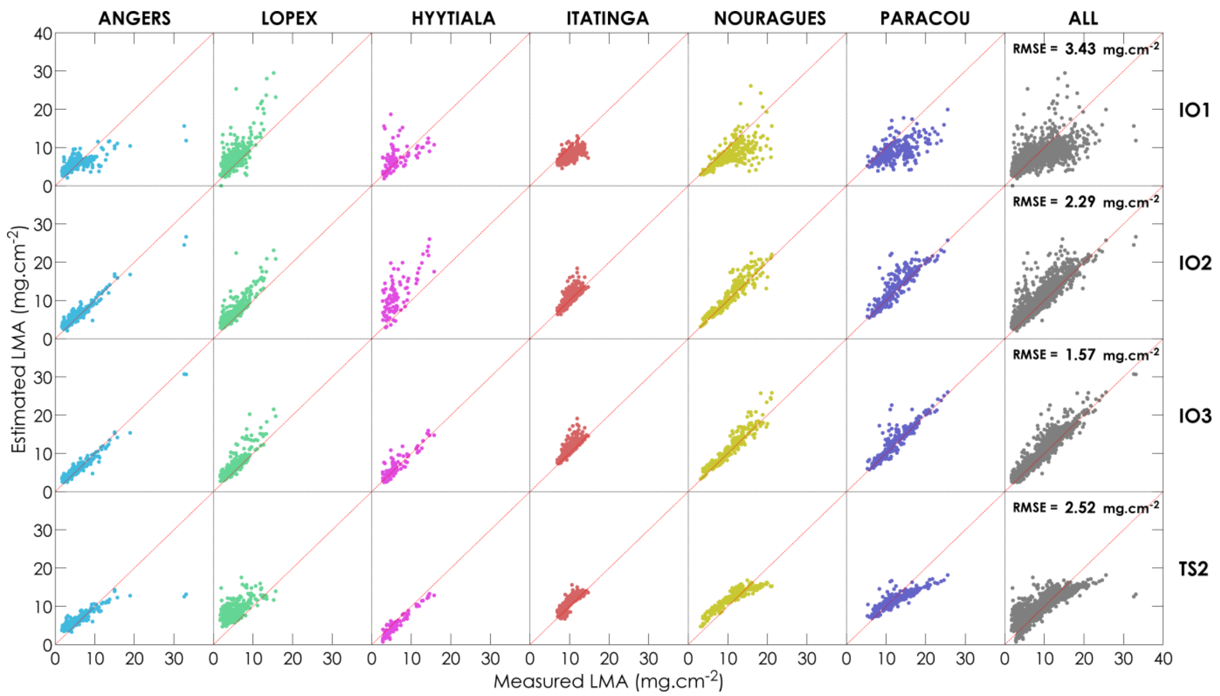

Figure 5. LMA estimation results obtained using PROSPECT inversion (*IO1*, *IO2*, *IO3*) and SVM regression (training sampling *TS2*).

c. Comparison of PROSPECT inversion methods and ML algorithms for the estimation of LMA and EWT: training ML with pooled datasets

Table 4 and Table 5 summarize the performances of SVM regression for the estimation of EWT and LMA when *TS3* is selected as training strategy (i.e. all dataset are pooled together and 300 calibration samples are randomly selected). The performances corresponding to *IO3* and *TS2* were computed for the same validation samples as with *TS3* for each of the 20 repetitions in order to ensure fair comparison.

The mean performances reported in Table 4 and Table 5 were very similar to those reported in Table 2 and Table 3 for both *IO3* and *TS2*, which means that *IO3* systematically outperformed *TS2* on individual datasets, except for the estimation of LMA for HYYTIALA and ITATINGA. *TS3* outperformed *TS2* in most cases for the estimation of both EWT and LMA. Still, *TS3* was outperformed by *IO3* when estimating EWT, the overall RMSE increasing by 44% (and by 99% when using TS2). When estimating LMA, *TS3* and *IO3* showed very similar overall performances, with less than 6% increase of RMSE for *TS3* when compared to *IO3*. *IO3* and *TS3* showed very similar average RMSE for LOPEX, HYTTIALA and NOURAGUES, *TS3* showed higher RMSE for ANGERS and PARACOU, and lower RMSE for ITATINGA. However, the standard deviations associated with these performances highlight the strong effect of training and validation samplings on the performances of the ML algorithm: the standard deviation computed over 20 repetitions was 5 to 20 times higher for *TS3* than *IO3* when estimating EWT, while it was 2.5 to 10 times higher when estimating LMA. The standard deviations related to the performances of *TS2* were generally similar to those obtained for *IO3*, suggesting that the strong differences in performance between regression models were induced by the selection of the training samples.

Table 4. Mean RMSE and standard deviation of RMSE (both in mg.cm^{-2}) of the estimation of EWT using SVM regression (*TS2* and *TS3*) and PROSPECT inversion (*IO3*) on the validation samples used for *TS3*. Best mean performances are indicated in bold.

	ANGERS	LOPEX	HYYTIALA	ITATINGA	NOURAGUES	PARACOU	Total
TS3	1.76±0.24	2.77±0.47	2.08±0.37	1.64±0.33	2.18±0.37	-	2.12±0.26
TS2	2.47±0.04	4.55±0.28	2.66±0.07	2.08±0.04	2.1±0.05	-	2.97±0.10
IO3	**1.43±0.04**	**1.70±0.07**	**1.21±0.04**	**1.21±0.02**	**1.65±0.05**	-	**1.47±0.02**

Table 5. Mean RMSE and standard deviation of RMSE (both in mg.cm^{-2}) of the estimation of $\boldsymbol{LMA}$ using SVM regression (*TS2* and *TS3*) and PROSPECT inversion (*IO3*) on the validation samples used for *TS3*. Best mean performances are indicated in bold (differences in mean RMSE < 1% are considered equivalent).

	ANGERS	LOPEX	HYYTIALA	ITATINGA	NOURAGUES	PARACOU	Total
TS3	1.70±0.28	**1.98±0.56**	1.56±0.22	**1.12±0.29**	**1.59±0.19**	2.01±0.18	1.64±0.18
TS2	2.24±0.21	4.05±0.06	**1.33±0.04**	1.23±0.02	2.13±0.03	2.43±0.05	2.31±0.05
IO3	**0.92±0.03**	**2.00±0.07**	1.54±0.08	1.45±0.04	**1.58±0.06**	**1.77±0.07**	**1.54±0.03**

5. DISCUSSION

a. Differences in performances among merit functions

Our study shows that *IO1,* the most commonly used merit function, is actually outperformed by a less common merit function (*IO2*) when estimating EWT and LMA from PROSPECT inversion using reflectance and transmittance in the NIR/SWIR domain (900-2400 nm). These results are in agreement with the results obtained when investigating the optimal spectral domain to be used with

IO3: Figure 3 shows that, in most cases, selecting a spectral domain including NIR information leads to suboptimal estimation of both EWT and LMA. Therefore, the application of a weight inversely proportional to the square of the reflectance and transmittance (*IO2*) reduce the importance of spectral domains showing higher reflectance and transmittance values such as the NIR domain. The improvement is particularly strong for the estimation of LMA, as reported in Figure 3. The particularly low performances obtained for the estimation of LMA on HYYTIAA were also investigated. The leaf optical properties measured for this dataset showed low SNR, particularly in the SWIR domain for wavelengths of 2300 nm and beyond. The estimation of LMA with *IO2* was strongly improved on this dataset when applying a Savitzky-Golay smoothing filter and restricting the spectral domain from 1700 to 2300 nm. The exclusion of the spectral domain beyond 2300 nm was responsible for the strongest improvement. Finally, the RMSE obtained for HYYTIAA when using the merit function used in *IO2* and these preprocessing reached 1.97 $mg.cm^{-2}$, which is still 30% higher than the RMSE obtained with *IO3*. Therefore using *IO2* is strongly discouraged when the signal to noise ratio of leaf optical properties is not sufficient, while *IO3* based on the 1700-2400 nm spectral range appears to be reliable even with low SNR.

b. Physical interpretation of the performances obtained with PROSPECT inversion

As highlighted in the previous section, the SNR of leaf optical properties can become a strong limitation when estimating leaf constituents using PROSPECT inversion if the spectral domain and merit functions are not carefully chosen. However, this SNR is not the main limiting factor explaining the poor performances of *IO1* for the estimation of LMA and its suboptimal performances for the estimation of EWT. Indeed, the NIR domain is theoretically characterized by a higher signal to noise ratio for leaf material but still appears to be the main limitation for an accurate estimation of these leaf constituents. Therefore, we attempt here to list possible explanations for such poor performances.

i. Predominant water absorption

The main reason cited to explain the poor retrieval of LMA is the predominant water absorption in the SWIR domain. Indeed, Figure 3 **Erreur ! Source du renvoi introuvable.**shows that LMA is poorly estimated when the spectral domains used for inversion mainly include domains with strong water absorption, such as the domain from 1800 to 2100 nm. However Figure 3 also shows that LMA can still be estimated accurately even if most of the spectral information corresponds to domains with predominant water absorption. Our results show that the main limitation with *IO1* is actually caused by the NIR domain between 900 and 1300 nm: most of the spectral domains excluding such wavebands resulted in improved estimation of LMA. The 900-1300 nm range does not show predominant water or dry matter absorption, so the poor retrieval of LMA cannot be explained by absorption features hidden by water absorption or any other constituent.

ii. Approximations of PROSPECT

As any model, PROSPECT is based on a number of approximations. Although some of these approximations are possible sources of inaccuracy in specific situations, they guarantee good overall performances given a minimum number of descriptors of leaf biophysical properties. Model discrepancies in the simulation of leaf optical properties may be explained by inaccurate physical description at three levels: surface effects, volume scattering and volume absorption.

Surface effects strongly depend on the presence of waxes or trichomes, and Barry and Newnham (2012) reported how epicuticular waxes affect PROSPECT inversion. Surface effects mostly influence leaf reflectance in the domains characterized by strong absorption where the leaf reflectance is minimum (Bousquet et al., 2005; Jay et al., 2016). In the NIR/SWIR spectral range, these domains mainly depend on water absorption. The sensitivity analysis performed by Jay et al. (2016) with similar EWT values showed that surface effects have the largest influence beyond 1800 nm, this domain being close to the one leading to optimal PROSPECT inversion results with *IO3* (1700-2400 nm). Such a result thus tends to indicate that surface effects had a limited detrimental influence on estimation performance.

Volume scattering is modeled by multiple factors in PROSPECT, including leaf structure with the N parameter, and the refractive index. The unique value of the refractive index is a well-identified simplification of PROSPECT, as it does not agree with the Kramers-Kronig relations stating that the real (refractive index) and imaginary (absorption coefficient) parts of the complex refractive index of a medium are physically linked (Lucarini et al., 2005). Qiu et al. (2018) developed PROSPECT-g, a modified version of PROSPECT including an additional wavelength-independent factor specific to each leaf and aiming at representing first-order effects of anisotropic scattering, which are not included through the N structural parameter of the original PROSPECT model. They also proposed a multistage inversion to be used with PROSPECT-g. This inversion procedure may strongly increase computing time, and the applicability of PROSPECT-g inversion at the canopy scale does not seem straightforward as additional parameters may increase the ill-posedness of canopy models such as PROSAIL (Jacquemoud et al., 2009). However, they reported promising results, including improved estimation of LMA and improved simulation of both reflectance and transmittance in the NIR domain when compared to PROSPECT-5.

Volume absorption is defined by the SACs which are adjusted based on experimental data during the calibration of PROSPECT (Féret et al., 2008, 2017). We attempted a recalibration of the SAC for LMA in order to reduce the inaccuracies observed between experimental and simulated data, and improve the estimation of LMA. This did not lead to any improvement when including the NIR domain. Moreover, the incorrect definition of the SAC corresponding to LMA would lead to systematic underestimation or overestimation of absorption when running PROSPECT in direct mode. However, the analysis of the residuals between measured leaf optical properties and their simulated counterparts obtained with PROSPECT in direct mode did not result in systematic errors (results not shown). The SAC corresponding to LMA in PROSPECT integrates the optical influences of various organic constituents, which may also lead to inaccuracies if leaf samples include strong variations in stoichiometry. However, the data required to test this possible source of inaccuracy was not available.

iii. Bias in the leaf optical measurements

As highlighted in the introduction, the uncertainty associated to leaf optical measurements in the NIR domain may be increased because of the incomplete collection of the light leaving the highly scattering tissue (Merzlyak et al., 2002). Merzlyak et al. (2004) proposed a correcting factor for transmittance based on the hypothesis that leaf absorption in the NIR domain from 780 to 900 nm is negligible for healthy leaves. However this correcting factor is not adopted as a standard correction by the community. In order to detect possible uncertainty in the optical measurements in the NIR domain with our data, we tested our ML approach with *TS1* (training with a unique dataset) and spectral information either from 1700 to 2400 nm or from 1400 to 2400 nm (results not showed). For both LMA and EWT**,** the regression models applied on independent datasets performed similarly for the two spectral domains considered, but systematically performed better than the regression models trained with the spectral information from 900 to 2400 nm. However, they were still outperformed by PROSPECT inversion. Such a result thus tends to confirm that leaf optical measurements in the NIR domain might be affected by some experimental uncertainty.

The poor performances reported for the estimation of LMA with PROSPECT inversion using *IO1* are therefore mainly explained by the use of the NIR domain, which is subject to inaccuracies, from a modeling and/or from an experimental point of view. Based on our study, we cannot conclude on the relative importance of one or the other factor. These two possibilities should then be considered and tested using the methods proposed in the literature (Merzlyak et al., 2004; Qiu et al., 2018). Finally, the difference between directional hemispherical measurements and bidirectional measurements should be systematically accounted for and appropriate physical models should be used with the type of data they are expected to simulate.

c. Influence of the sampling of the training dataset on machine learning algorithms

Our results highlight the strong influence of the training dataset on the performances of ML methods, which is not an original result *per se*. However, the different training strategies tested here show that regression models should be used with extreme care when they are applied on data which were not collected in the exact same conditions as training data. Finally, the optimal training strategy in our case, *TS3*, requires that each campaign aiming at collecting leaf optical properties in order to estimate constituent content based on statistical/ML methods should include destructive measurements to be used during the training step. This means that publicly available datasets such as ANGERS and LOPEX should not be used as the only training datasets for the estimation of leaf chemistry based on spectroscopy from independent datasets. The origin of the suboptimal performances obtained in particular with *TS2* and *TS3* should also be investigated. ML algorithms are currently mainly used for their predictive capacity. However, they can also as part of a descriptive framework. Feilhauer et al. (2015) proposed an interesting illustration as they suggested combining multiple methods in order to identify the most relevant spectral bands related to leaf chemistry, based on both experimental and simulated data. Following the same method, the identification of the spectral bands maximizing the generalization ability of ML algorithms by discarding spectral domains prompt to experimental uncertainty or model approximations could be considered. Finally, hybrid methods using simulated data during the training stage of a ML algorithm appear as an interesting alternative to data-driven methods purely based on experimental data, and further investigation is needed in order to define the proper strategy to generate such training dataset and combine the generalization ability of physically-based approaches with the computational efficiency of data-driven approaches.

d. Relevance of these results for leaf trait monitoring

The results obtained in this study contribute to a better understanding of the optimal remotely-sensed monitoring of LMA and EWT, two key vegetation traits that convey multiple information about the spatial and temporal variation in ecological and functional diversity of terrestrial ecosystems. This can possibly contribute to facilitating the study of plant functions and their interactions with and responses to the environment. As an example, Feilhauer et al. (2018) provide a good illustration of the interest of remotely-sensed LMA for ecological analysis of wetland vegetation, in particular for the better understanding of the effect of long-term drought on ecosystem functions. They focused on LMA because of its plasticity in response to variable environmental conditions, and its relationship with potential growth rate.

The estimation of these traits at the leaf scale now needs to be further investigated at the canopy scale. In order to test the applicability of our approach at the canopy scale, the first step will consist in working with a simulated dataset obtained with canopy reflectance models such as SAIL (Jacquemoud et al., 2009; Verhoef, 1984) and DART (Gastellu-Etchegorry et al., 1996, 2015). The direct application of model inversion based on iterative optimization restricts the complexity of the canopy model, hence the type of vegetation to be investigated: the adaptation of our method should be relatively straightforward when using PROSAIL on homogeneous canopy covers, but hybrid methods should be considered when using DART simulations and working on heterogeneous canopy covers.

An important challenge for the applicability of our results at the canopy scale is the low intensity of the solar radiation in the optimal SWIR domain identified in this study, which usually leads to low signal to noise ratio. Currently, hyperspectral information is mainly available from airborne imaging spectroscopy (Asner et al., 2012; Schaepman et al., 2015). Asner et al. (2015) obtained accurate estimation of LMA based on multivariate statistical methods applied on imaging spectroscopy for heterogeneous canopies in tropical ecosystems, and they also concluded on the importance of the spectral domain from 2000 nm to 2500 nm for a proper calibration of the regression models. Recently, Feilhauer et al. (2018) reported good suitability of airborne imaging spectroscopy analyzed with a hybrid method (Random forest trained with PROSAIL simulations) for LMA mapping in natural ecosystems. Hyperion is the only spaceborne sensor, but the signal to noise ratio is known to be relatively low (le Maire et al., 2008). The contribution of modeling through sensitivity studies performed at canopy scale may therefore provide insightful information for the instrumental specifications of future satellites dedicated to the monitoring of vegetation and environment such as EnMAP, and for the development of algorithms (Jetz et al., 2016; Lee et al., 2015; Leitão et al., 2015).

6. CONCLUSIONS

In this paper, we compared the performances of various methods for the estimation of EWT and LMA based on leaf reflectance and transmittance in the spectral domain ranging from 900 to 2400 nm. These methods included PROSPECT inversion based on iterative optimization with various merit functions and machine learning (ML) algorithms with different training strategies. Six independent datasets acquired from various vegetation types, including temperate, boreal and tropical ecosystems were used in order to validate our results.

Our results showed that the poor performances of PROSPECT inversion reported in many studies for the estimation of LMA could be dramatically improved when excluding spectral information in the

NIR domain from 900 to 1300 nm. We investigated the performances of PROSPECT inversion for the estimation of EWT and LMA using multiple spectral subdomains, and identified an optimal spectral domain ranging from 1700 to 2400 nm. Overall, PROSPECT inversion performed on this spectral domain provided more accurate LMA and EWT estimates than ML algorithms trained on experimental datasets. Unlike ML algorithms, PROSPECT inversion showed strong generalization ability. Despite numerous studies showing the poor performances of PROSPECT for the estimation of LMA, our study shows that model inversion using iterative optimization can outperform other methods with an appropriate merit function, with no need for recalibration or training stage. By this study, we therefore confirm the strong potential and accuracy of PROSPECT on critical spectral domains. We also identified weaknesses which can be attributed either to physical modeling and experimental acquisition of leaf optical properties in the NIR domain.

These results motivate further investigation involving hybrid methods for the estimation of LMA and EWT, in order to take advantage of the computational efficiency of data-driven algorithms and overcome limitations inherent to suboptimal experimental sampling of training data. Implications of these results for the optimal estimation of LMA and EWT at the canopy scale will also be investigated, as LMA and EWT are both key traits when monitoring ecosystem functions.

7. ACKNOWLEDGMENTS

The authors warmly thank Luc Bidel, Christophe François and Gabriel Pavan who collected the ANGERS dataset. This work was funded by the TOSCA program grant of the French Space Agency (CNES) (HyperTropik project); the International Network for Terrestrial Research and Monitoring in the Arctic (INTERACT) (FLUO-SYNTHESIS project); the Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES). ITATINGA genotype test is funded in part by the EUCFLUX project of Forestry Science and Research Institute (IPEF) and in part by SOERE F-ORE-T, which is supported annually by Ecofor, Allenvi, and the ANAEE-F. The authors also thank Valentine Alt, Samuel Counil, and Philippe Gaucher for tree climbing and data collection, as well as Anna Grandchamp for her help during data collection in Nouragues field station.

8. REFERENCES

Ali, A.M., Darvishzadeh, R., Skidmore, A.K., Duren, I. van, Heiden, U., Heurich, M., 2016. Estimating leaf functional traits by inversion of PROSPECT: Assessing leaf dry matter content and specific leaf area in mixed mountainous forest. Int. J. Appl. Earth Obs. Geoinformation 45, 66–76. https://doi.org/10.1016/j.jag.2015.11.004

Allen, W.A., Gausman, H.W., Richardson, A.J., 1970. Mean effective optical constants of cotton leaves. J. Opt. Soc. Am. 60, 542–547. https://doi.org/10.1364/JOSA.60.000542

Allen, W.A., Gausman, H.W., Richardson, A.J., Thomas, J.R., 1969. Interaction of isotropic light with a compact plant leaf. J. Opt. Soc. Am. 59, 1376–1379. https://doi.org/10.1364/JOSA.59.001376

Antúnez, I., Retamosa, E.C., Villar, R., 2001. Relative growth rate in phylogenetically related deciduous and evergreen woody species. Oecologia 128, 172–180. https://doi.org/10.1007/s004420100645

Asner, G.P., Knapp, D.E., Boardman, J., Green, R.O., Kennedy-Bowdoin, T., Eastwood, M., Martin, R.E., Anderson, C., Field, C.B., 2012. Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion. Remote Sens. Environ. 124, 454–465. https://doi.org/10.1016/j.rse.2012.06.012

Asner, G.P., Martin, R.E., 2016. Spectranomics: Emerging science and conservation opportunities at the interface of biodiversity and remote sensing. Glob. Ecol. Conserv. 8, 212–219. https://doi.org/10.1016/j.gecco.2016.09.010

Asner, G.P., Martin, R.E., Anderson, C.B., Knapp, D.E., 2015. Quantifying forest canopy traits: Imaging spectroscopy versus field survey. Remote Sens. Environ. 158, 15–27. https://doi.org/10.1016/j.rse.2014.11.011

Asner, G.P., Martin, R.E., Ford, A.J., Metcalfe, D.J., Liddell, M.J., 2009. Leaf chemical and spectral diversity in Australian tropical forests. Ecol. Appl. 19, 236–253. https://doi.org/10.1890/08-0023.1

Asner, G.P., Martin, R.E., Tupayachi, R., Emerson, R., Martinez, P., Sinca, F., Powell, G.V., Wright, S.J., Lugo, A.E., 2011. Taxonomy and remote sensing of leaf mass per area (LMA) in humid tropical forests. Ecol. Appl. 21, 85–98.

Baret, F., Buis, S., 2008. Estimating Canopy Characteristics from Remote Sensing Observations: Review of Methods and Associated Problems, in: Liang, S. (Ed.), Advances in Land Remote Sensing. Springer Netherlands, Dordrecht, pp. 173–201.

Barry, K., Newnham, G., 2012. Quantification of chlorophyll and carotenoid pigments in eucalyptus foliage with the radiative transfer model PROSPECT 5 is affected by anthocyanin and epicuticular waxes., in: Proc. Geospatial Science Research 2 Symposium, GSR 2012, Melbourne, Australia, December 10-12, 2012.

Bousquet, L., Lachérade, S., Jacquemoud, S., Moya, I., 2005. Leaf BRDF measurements and model for specular and diffuse components differentiation. Remote Sens. Environ. 98, 201–211. https://doi.org/10.1016/j.rse.2005.07.005
Bowyer, P., Danson, F.M., 2004. Sensitivity of spectral reflectance to variation in live fuel moisture content at leaf and canopy level. Remote Sens. Environ. 92, 297–308. https://doi.org/10.1016/j.rse.2004.05.020
Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324
Brown, M., Lewis, H.G., Gunn, S.R., 2000. Linear spectral mixture models and support vector machines for remote sensing. IEEE Trans. Geosci. Remote Sens. 38, 2346–2360. https://doi.org/10.1109/36.868891
Carter, G.A., Knapp, A.K., 2001. Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration. Am. J. Bot. 88, 677–684.
Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., Grégoire, J.-M., 2001. Detecting vegetation leaf water content using reflectance in the optical domain. Remote Sens. Environ. 77, 22–33. https://doi.org/10.1016/S0034-4257(01)00191-2
Chang, C.-C., Lin, C.-J., 2011. LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2, 1–27. https://doi.org/10.1145/1961189.1961199
Chapin, F.S., 2003. Effects of Plant Traits on Ecosystem and Regional Processes: a Conceptual Framework for Predicting the Consequences of Global Change. Ann. Bot. 91, 455–463. https://doi.org/10.1093/aob/mcg041
Chuvieco, E., Riaño, D., Aguado, I., Cocero, D., 2002. Estimation of fuel moisture content from multitemporal analysis of Landsat Thematic Mapper reflectance data: Applications in fire danger assessment. Int. J. Remote Sens. 23, 2145–2162. https://doi.org/10.1080/01431160110069818
Colombo, R., Meroni, M., Marchesi, A., Busetto, L., Rossini, M., Giardino, C., Panigada, C., 2008. Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling. Remote Sens. Environ. 112, 1820–1834. https://doi.org/10.1016/j.rse.2007.09.005
Conejo, E., Frangi, J.-P., de Rosny, G., 2015. Neural network implementation for a reversal procedure for water and dry matter estimation on plant leaves using selected LED wavelengths. Appl. Opt. 54, 5453. https://doi.org/10.1364/AO.54.005453
Cornelissen, J.H.C., Grootemaat, S., Verheijen, L.M., Cornwell, W.K., van Bodegom, P.M., van der Wal, R., Aerts, R., 2017. Are litter decomposition and fire linked through plant species traits? New Phytol. 216, 653–669. https://doi.org/10.1111/nph.14766
Cornelissen, J.H.C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D.E., Reich, P.B., Steege, H. ter, Morgan, H.D., Heijden, M.G.A. van der, Pausas, J.G., Poorter, H., 2003. A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Aust. J. Bot. 51, 335. https://doi.org/10.1071/BT02124
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297. https://doi.org/10.1007/BF00994018
Datt, B., 1999. Remote Sensing of Water Content in Eucalyptus Leaves. Aust. J. Bot. 47, 909–923.
Dawson, T.P., Curran, P.J., Plummer, S.E., 1998. The biochemical decomposition of slash pine needles from reflectance spectra using neural networks. Int. J. Remote Sens. 19, 1433–1438. https://doi.org/10.1080/014311698215540
de la Riva, E.G., Olmo, M., Poorter, H., Ubera, J.L., Villar, R., 2016. Leaf Mass per Area (LMA) and Its Relationship with Leaf Structure and Anatomy in 34 Mediterranean Woody Species along a

Water Availability Gradient. PLOS ONE 11, e0148788. https://doi.org/10.1371/journal.pone.0148788

Diaz, S., Cabido, M., 2001. Vive la diff?rence: plant functional diversity matters to ecosystem processes. Trends Ecol. Evol. 16, 646–655. https://doi.org/10.1016/S0169-5347(01)02283-2

Drucker, H., Burges, C.J.C., Kaufman, L., C, C.J., Kaufman, B.L., Smola, A., Vapnik, V., 1996. Support Vector Regression Machines.

Eviner, V.T., Chapin III, F.S., 2003. Functional Matrix: A Conceptual Framework for Predicting Multiple Plant Effects on Ecosystem Processes. Annu. Rev. Ecol. Evol. Syst. 34, 455–485. https://doi.org/10.1146/annurev.ecolsys.34.011802.132342

Feilhauer, H., Asner, G.P., Martin, R.E., 2015. Multi-method ensemble selection of spectral bands related to leaf biochemistry. Remote Sens. Environ. 164, 57–65. https://doi.org/10.1016/j.rse.2015.03.033

Feilhauer, H., Schmid, T., Faude, U., Sánchez-Carrillo, S., Cirujano, S., 2018. Are remotely sensed traits suitable for ecological analysis? A case study of long-term drought effects on leaf mass per area of wetland vegetation. Ecol. Indic. 88, 232–240. https://doi.org/10.1016/j.ecolind.2018.01.012

Féret, J.-B., François, C., Asner, G.P., Gitelson, A.A., Martin, R.E., Bidel, L.P.R., Ustin, S.L., le Maire, G., Jacquemoud, S., 2008. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. Remote Sens. Environ. 112, 3030–3043. https://doi.org/10.1016/j.rse.2008.02.012

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

Féret, J.-B., Gitelson, A.A., Noble, S.D., Jacquemoud, S., 2017. PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. Remote Sens. Environ. 193, 204–215. https://doi.org/10.1016/j.rse.2017.03.004

Gastellu-Etchegorry, J., Demarez, V., Pinel, V., Zagolski, F., 1996. Modeling radiative transfer in heterogeneous 3-D vegetation canopies. Remote Sens. Environ. 58, 131–156. https://doi.org/10.1016/0034-4257(95)00253-7

Gastellu-Etchegorry, J.-P., Yin, T., Lauret, N., Cajgfinger, T., Gregoire, T., Grau, E., Féret, J.-B., Lopes, M., Guilleux, J., Dedieu, G., Malenovský, Z., Cook, B., Morton, D., Rubio, J., Durrieu, S., Cazanave, G., Martin, E., Ristorcelli, T., 2015. Discrete anisotropic radiative transfer (DART 5) for modeling airborne and satellite spectroradiometer and LIDAR acquisitions of natural and urban landscapes. Remote Sens. 7, 1667–1701. https://doi.org/10.3390/rs70201667

Gitelson, A.A., Keydan, G.P., Merzlyak, M.N., 2006. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. Geophys. Res. Lett. 33, L11402. https://doi.org/10.1029/2006GL026457

Gratani, L., Varone, L., 2006. Long-time variations in leaf mass and area of Mediterranean evergreen broad-leaf and narrow-leaf maquis species. Photosynthetica 44, 161–168. https://doi.org/10.1007/s11099-006-0001-1

Gualtieri, J.A., 2009. The Support Vector Machine (SVM) Algorithm for Supervised Classification of Hyperspectral Remote Sensing Data, in: Camps-Valls, G., Bruzzone, L. (Eds.), Kernel Methods for Remote Sensing Data Analysis. John Wiley & Sons, Ltd, Chichester, UK, pp. 49–83.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Netw. 2, 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Hosgood, B., Jacquemoud, S., Andreoli, G., Verdebout, J., Pedrini, A., Schmuck, G., 1994. Leaf Optical Properties Experiment 93 (LOPEX93) (European Commission No. EUR 16095 EN). Joint Research Centre, Institute for Remote Sensing Applications.

Jacquemoud, S., Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. Remote Sens. Environ. 34, 75–91. https://doi.org/10.1016/0034-4257(90)90100-Z

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 Sens. Environ. 56, 194–202. https://doi.org/10.1016/0034-4257(95)00238-3

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., Ustin, S.L., 2009. PROSPECT+ SAIL models: A review of use for vegetation characterization. Remote Sens. Environ. 113, S56–S66. https://doi.org/doi:10.1016/j.rse.2008.01.026

Jay, S., Bendoula, R., Hadoux, X., Féret, J.-B., Gorretta, N., 2016. A physically-based model for retrieving foliar biochemistry and leaf orientation using close-range imaging spectroscopy. Remote Sens. Environ. 177, 220–236. https://doi.org/10.1016/j.rse.2016.02.029

Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F.W., Asner, G.P., Guralnick, R., Kattge, J., Latimer, A.M., Moorcroft, P., Schaepman, M.E., Schildhauer, M.P., Schneider, F.D., Schrodt, F., Stahl, U., Ustin, S.L., 2016. Monitoring plant functional diversity from space. Nat. Plants 2, 16024. https://doi.org/10.1038/nplants.2016.24

Lardeux, C., Frison, P.-L., Tison, C., Souyris, J.-C., Stoll, B., Fruneau, B., Rudant, J.-P., 2009. Support Vector Machine for Multifrequency SAR Polarimetric Data Classification. IEEE Trans. Geosci. Remote Sens. 47, 4143–4152. https://doi.org/10.1109/TGRS.2009.2023908

le Maire, G., François, C., Dufrêne, E., 2004. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. Remote Sens. Environ. 89, 1–28. https://doi.org/10.1016/j.rse.2003.09.004

le Maire, G., François, C., Soudani, K., Berveiller, D., Pontailler, J.-Y., Bréda, N., Genet, H., Davi, H., Dufrêne, E., 2008. 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. Remote Sens. Environ. 112, 3846–3864.

le Maire, G., Marsden, C., Nouvellon, Y., Grinand, C., Hakamada, R., Stape, J.-L., Laclau, J.-P., 2011. MODIS NDVI time-series allow the monitoring of Eucalyptus plantation biomass. Remote Sens. Environ. 115, 2613–2625. https://doi.org/10.1016/j.rse.2011.05.017

Lee, C.M., Cable, M.L., Hook, S.J., Green, R.O., Ustin, S.L., Mandl, D.J., Middleton, E.M., 2015. An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. Remote Sens. Environ. 167, 6–19. https://doi.org/10.1016/j.rse.2015.06.012

Leitão, P., Schwieder, M., Suess, S., Okujeni, A., Galvão, L., Linden, S., Hostert, P., 2015. Monitoring Natural Ecosystem and Ecological Gradients: Perspectives with EnMAP. Remote Sens. 7, 13098–13119. https://doi.org/10.3390/rs71013098

Li, D., Cheng, T., Jia, M., Zhou, K., Lu, N., Yao, X., Tian, Y., Zhu, Y., Cao, W., 2018. PROCWT: Coupling PROSPECT with continuous wavelet transform to improve the retrieval of foliar chemistry from leaf bidirectional reflectance spectra. Remote Sens. Environ. 206, 1–14. https://doi.org/10.1016/j.rse.2017.12.013

Li, P., Wang, Q., 2011. Retrieval of leaf biochemical parameters using PROSPECT inversion: A new approach for alleviating ill-posed problems. IEEE Trans. Geosci. Remote Sens. 49, 2499–2506. https://doi.org/10.1109/TGRS.2011.2109390

Lucarini, V., Saarinen, J.J., Peiponen, K.-E., Vartiainen, E.M. (Eds.), 2005. Kramers-Kronig relations in optical materials research, Springer series in optical sciences. Springer, Berlin.

Main, R., Cho, M.A., Mathieu, R., O'Kennedy, M.M., Ramoelo, A., Koch, S., 2011. An investigation into robust spectral indices for leaf chlorophyll estimation. ISPRS J. Photogramm. Remote Sens. 66, 751–761. https://doi.org/10.1016/j.isprsjprs.2011.08.001

Malenovský, Z., Albrechtová, J., Lhotáková, Z., Zurita-Milla, R., Clevers, J.G.P.W., Schaepman, M.E., Cudlín, P., 2006. Applicability of the PROSPECT model for Norway spruce needles. Int. J. Remote Sens. 27, 5315–5340. https://doi.org/10.1080/01431160600762990

Merzlyak, M.N., Chivkunova, O.B., Melø, T.B., Naqvi, K.R., 2002. Does a leaf absorb radiation in the near infrared (780–900 nm) region? A new approach to quantifying optical reflection, absorption and transmission of leaves. Photosynth. Res. 72, 263–270. https://doi.org/10.1023/A:1019823303951

Merzlyak, M.N., Melø, T.B., Razi Naqvi, K., 2004. Estimation of leaf transmittance in the near infrared region through reflectance measurements. J. Photochem. Photobiol. B 74, 145–150. https://doi.org/10.1016/j.jphotobiol.2004.03.003

Mobasheri, M.R., Fatemi, S.B., 2013. Leaf Equivalent Water Thickness assessment using reflectance at optimum wavelengths. Theor. Exp. Plant Physiol. 25, 196–202. https://doi.org/10.1590/S2197-00252013005000001

Newnham, G.J., Burt, T., 2001. Validation of a leaf reflectance and transmittance model for three agricultural crop species. IEEE, pp. 2976–2978. https://doi.org/10.1109/IGARSS.2001.978227

Oliveira, J. de C., Feret, J.-B., Ponzoni, F.J., Nouvellon, Y., Gastellu-Etchegorry, J.-P., Campoe, O.C., Stape, J.L., Rodriguez, L.C.E., le Maire, G., 2017. Simulating the Canopy Reflectance of Different Eucalypt Genotypes With the DART 3-D Model. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 1–9. https://doi.org/10.1109/JSTARS.2017.2690000

Oren, R., Schulze, E.-D., Matyssek, R., Zimmermann, R., 1986. Estimating photosynthetic rate and annual carbon gain in conifers from specific leaf weight and leaf biomass. Oecologia 70, 187–193. https://doi.org/10.1007/BF00379238

Osnas, J.L.D., Lichstein, J.W., Reich, P.B., Pacala, S.W., 2013. Global Leaf Trait Relationships: Mass, Area, and the Leaf Economics Spectrum. Science 340, 741–744. https://doi.org/10.1126/science.1231574

Pérez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., Bret-Harte, M.S., Cornwell, W.K., Craine, J.M., Gurvich, D.E., Urcelay, C., Veneklaas, E.J., Reich, P.B., Poorter, L., Wright, I.J., Ray, P., Enrico, L., Pausas, J.G., de Vos, A.C., Buchmann, N., Funes, G., Quétier, F., Hodgson, J.G., Thompson, K., Morgan, H.D., ter Steege, H., Sack, L., Blonder, B., Poschlod, P., Vaieretti, M.V., Conti, G., Staver, A.C., Aquino, S., Cornelissen, J.H.C., 2013. New handbook for standardised measurement of plant functional traits worldwide. Aust. J. Bot. 61, 167. https://doi.org/10.1071/BT12225

Poorter, H., Niinemets, Ü., Poorter, L., Wright, I.J., Villar, R., 2009. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. New Phytol. 182, 565–588. https://doi.org/10.1111/j.1469-8137.2009.02830.x

Puglielli, G., Crescente, M.F., Frattaroli, A.R., Gratani, L., 2015. Leaf Mass Per Area (LMA) as a Possible Predictor of Adaptive Strategies in Two Species of *Sesleria* (Poaceae): Analysis of Morphological, Anatomical and Physiological Leaf Traits. Ann. Bot. Fenn. 52, 135–143. https://doi.org/10.5735/085.052.0201

Qiu, F., Chen, J.M., Ju, W., Wang, J., Zhang, Q., Fang, M., 2018. Improving the PROSPECT Model to Consider Anisotropic Scattering of Leaf Internal Materials and Its Use for Retrieving Leaf Biomass in Fresh Leaves. IEEE Trans. Geosci. Remote Sens. 56, 3119–3136. https://doi.org/10.1109/TGRS.2018.2791930

Rees, M., Osborne, C.P., Woodward, F.I., Hulme, S.P., Turnbull, L.A., Taylor, S.H., 2010. Partitioning the Components of Relative Growth Rate: How Important Is Plant Size Variation? Am. Nat. 176, E152–E161. https://doi.org/10.1086/657037

Reich, P.B., Walters, M.B., Ellsworth, D.S., 1997. From tropics to tundra: global convergence in plant functioning. Proc. Natl. Acad. Sci. U. S. A. 94, 13730–13734.

Reich, P.B., Walters, M.B., Ellsworth, D.S., Vose, J.M., Volin, J.C., Gresham, C., Bowman, W.D., 1998. Relationships of leaf dark respiration to leaf nitrogen, specific leaf area and leaf life-span: a test across biomes and functional groups. Oecologia 114, 471–482. https://doi.org/10.1007/s004420050471

Réjou-Méchain, M., Tymen, B., Blanc, L., Fauset, S., Feldpausch, T.R., Monteagudo, A., Phillips, O.L., Richard, H., Chave, J., 2015. Using repeated small-footprint LiDAR acquisitions to infer spatial and temporal variations of a high-biomass Neotropical forest. Remote Sens. Environ. 169, 93–101. https://doi.org/10.1016/j.rse.2015.08.001

Riano, D., Vaughan, P., Chuvieco, E., Zarco-Tejada, P.J., Ustin, S.L., 2005. Estimation of fuel moisture content by inversion of radiative transfer models to simulate equivalent water thickness and dry matter content: analysis at leaf and canopy level. IEEE Trans. Geosci. Remote Sens. 43, 819–826. https://doi.org/10.1109/TGRS.2005.843316

Romero, A., Aguado, I., Yebra, M., 2012. Estimation of dry matter content in leaves using normalized indexes and PROSPECT model inversion. Int. J. Remote Sens. 33, 396–414. https://doi.org/10.1080/01431161.2010.532819

Schaepman, M.E., Jehle, M., Hueni, A., D'Odorico, P., Damm, A., Weyermann, J., Schneider, F.D., Laurent, V., Popp, C., Seidel, F.C., Lenhard, K., Gege, P., Küchler, C., Brazile, J., Kohler, P., De Vos, L., Meuleman, K., Meynart, R., Schläpfer, D., Kneubühler, M., Itten, K.I., 2015. Advanced radiometry measurements and Earth science applications with the Airborne Prism Experiment (APEX). Remote Sens. Environ. 158, 207–219. https://doi.org/10.1016/j.rse.2014.11.014

Schaepman-Strub, G., Schaepman, M.E., Painter, T.H., Dangel, S., Martonchik, J.V., 2006. Reflectance quantities in optical remote sensing—definitions and case studies. Remote Sens. Environ. 103, 27–42. https://doi.org/10.1016/j.rse.2006.03.002

Schimel, D., Pavlick, R., Fisher, J.B., Asner, G.P., Saatchi, S., Townsend, P., Miller, C., Frankenberg, C., Hibbard, K., Cox, P., 2015. Observing terrestrial ecosystems and the carbon cycle from space. Glob. Change Biol. 21, 1762–1776. https://doi.org/10.1111/gcb.12822

Schmitter, P., Steinrücken, J., Römer, C., Ballvora, A., Léon, J., Rascher, U., Plümer, L., 2017. Unsupervised domain adaptation for early detection of drought stress in hyperspectral images. ISPRS J. Photogramm. Remote Sens. 131, 65–76. https://doi.org/10.1016/j.isprsjprs.2017.07.003

Stumpf, A., Kerle, N., 2011. Object-oriented mapping of landslides using Random Forests. Remote Sens. Environ. 115, 2564–2577. https://doi.org/10.1016/j.rse.2011.05.013

Sun, J., Shi, S., Yang, J., Du, L., Gong, W., Chen, B., Song, S., 2018. Analyzing the performance of PROSPECT model inversion based on different spectral information for leaf biochemical properties retrieval. ISPRS J. Photogramm. Remote Sens. 135, 74–83. https://doi.org/10.1016/j.isprsjprs.2017.11.010

Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. Remote Sens. Environ. 16, 125–141. https://doi.org/10.1016/0034-4257(84)90057-9

Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G.P.W., Moreno, J., 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – A review. ISPRS J. Photogramm. Remote Sens. 108, 273–290. https://doi.org/10.1016/j.isprsjprs.2015.05.005

Verrelst, J., Rivera, J.P., Gitelson, A., Delegido, J., Moreno, J., Camps-Valls, G., 2016. Spectral band selection for vegetation properties retrieval using Gaussian processes regression. Int. J. Appl. Earth Obs. Geoinformation 52, 554–567. https://doi.org/10.1016/j.jag.2016.07.016

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., Garnier, E., 2007. Let the concept of trait be functional! Oikos 116, 882–892. https://doi.org/10.1111/j.2007.0030-1299.15559.x

Wang, L., Qu, J.J., Hao, X., Hunt, E.R., 2011. Estimating dry matter content from spectral reflectance for green leaves of different species. Int. J. Remote Sens. 32, 7097–7109. https://doi.org/10.1080/01431161.2010.494641

Wang, Z., Skidmore, A.K., Wang, T., Darvishzadeh, R., Hearne, J., 2015. Applicability of the PROSPECT model for estimating protein and cellulose+lignin in fresh leaves. Remote Sens. Environ. 168, 205–218. https://doi.org/10.1016/j.rse.2015.07.007

Weng, E., Farrior, C.E., Dybzinski, R., Pacala, S.W., 2017. Predicting vegetation type through physiological and environmental interactions with leaf traits: evergreen and deciduous forests in an earth system modeling framework. Glob. Change Biol. 23, 2482–2498. https://doi.org/10.1111/gcb.13542

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., Villar, R., 2004. The worldwide leaf economics spectrum. Nature 428, 821–827. https://doi.org/10.1038/nature02403

Yebra, M., Dennison, P.E., Chuvieco, E., Riaño, D., Zylstra, P., Hunt, E.R., Danson, F.M., Qi, Y., Jurdao, S., 2013. A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. Remote Sens. Environ. 136, 455–468.

Zhang, X.M., He, G.J., Zhang, Z.M., Peng, Y., Long, T.F., 2017. Spectral-spatial multi-feature classification of remote sensing big data based on a random forest classifier for land cover mapping. Clust. Comput. 20, 2311–2321. https://doi.org/10.1007/s10586-017-0950-0

9. LIST OF FIGURE CAPTIONS

Figure 1. Normalized $RMSE$ ($NRMSE$, in %) obtained for EWT with PROSPECT inversion method *IO3* over each dataset and each reduced spectral domains bounded by a starting wavelength λ_1 (y-axis) and an ending wavelength λ_2 (x-axis). The normalization is specific to each dataset based on the performances of *IO1* (NRMSE=100%, lower right corner). The green star indicates the spectral segment producing the best results.

Figure 2. Normalized $RMSE$ (NRMSE, in %) obtained for LMA estimation with PROSPECT inversion method *IO3*, over each dataset and each reduced spectral domains bounded by a starting wavelength λ_1 (y-axis) and an ending wavelength λ_2 (x-axis). The normalization is specific to each dataset based on the performances of *IO1* (NRMSE=100%, lower right corner). The green star indicates the spectral segment producing the best results.

Figure 3. Mean normalized RMSE values (NRMSE, in %) obtained for the estimation of EWT (left), LMA (center), and both constituents (right), after PROSPECT inversion over all experimental datasets pooled and each of the 120 spectral domains defined in Section 3.b. The green star indicates the spectral segment producing the best results.

Figure 4. EWT estimation results obtained using PROSPECT inversion (*IO1*, *IO2*, *IO3*) and ML regression (training sampling *TS2*).

Figure 5. LMA estimation results obtained using PROSPECT inversion (*IO1*, *IO2*, *IO3*) and SVM regression (training sampling *TS2*).