Remote Sensing of Environment 196 (2017) 279–292

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Estimation of photosynthesis traits from leaf reflectance spectra: Correlation to nitrogen content as the dominant mechanism

Benjamin Dechant [a,b,*], Matthias Cuntz [a,e], Michael Vohland [d], Elke Schulz [c], Daniel Doktor [b]

[a] Department Computational Hydrosystems, UFZ – Helmholtz Centre for Environmental Research, Permoser Str. 15, 04318 Leipzig, Germany
[b] Department Computational Landscape Ecology, UFZ – Helmholtz Centre for Environmental Research, Permoser Str. 15, 04318 Leipzig, Germany
[c] Department of Soil Ecology, UFZ – Helmholtz Centre for Environmental Research, Theodor-Lieser-Str. 4, 06120 Halle, Germany
[d] Geoinformatics and Remote Sensing, Department for Geography, University of Leipzig, Johannisallee 19a, 04103 Leipzig, Germany
[e] INRA, Université de Lorraine, UMR1137 Ecologie et Ecophysiologie Forestières, 54280 Champenoux, France

ARTICLE INFO

Article history:
Received 6 October 2016
Received in revised form 29 April 2017
Accepted 14 May 2017
Available online 20 May 2017

Keywords:
V_{cmax}
J_{max}
Nitrogen
Photosynthesis
Leaf reflectance
PLS
Hyperspectral

ABSTRACT

Numerous studies have investigated reflectance-based estimations of physico-chemical leaf traits such as the contents of light absorbing pigments, leaf mass per area, or carbon and nitrogen contents. Only few studies, however, attempted to estimate leaf traits that are more directly linked to photosynthesis. We tested the feasibility of estimating two important photosynthesis traits, the maximum carboxylation capacity ($V_{cmax,25}$) and the maximum electron transport rate ($J_{max,25}$), from in-situ leaf reflectance spectra using approaches that are applicable also on larger spatial scales. We conducted measurements of reflectance, CO_2 response curves, leaf mass per area (*LMA*), and nitrogen content per area (N_a) for 37 temperate deciduous tree species and a total of 242 leaves from widely differing light environments. Partial least squares (PLS) regression was used to estimate $V_{cmax,25}$, $J_{max,25}$, *LMA*, and N_a from reflectance spectra.

The results showed that both $V_{cmax,25}$ and $J_{max,25}$ can be estimated from leaf reflectance measurements with good accuracy ($R^2 = 0.64$ for $V_{cmax,25}$, $R^2 = 0.70$ for $J_{max,25}$) even for a large number of tree species and varying light environments. Detailed analysis of reflectance-based PLS and linear regression models with regard to prediction performances and regression coefficients led to the conclusion that the correlation to N_a was the dominating mechanism on which the $V_{cmax,25}$ and $J_{max,25}$ PLS models were based. The PLS regression coefficients of N_a, $V_{cmax,25}$ and $J_{max,25}$ showed clear signatures of nitrogen-related absorption features.

The finding that $V_{cmax,25}$ and $J_{max,25}$ estimations from leaf reflectance are predominantly based on their relationships to N_a has important implications for large scale estimations of these photosynthesis parameters. We suggest that future studies should focus more on large scale estimation of N_a from remote sensing and estimate $V_{cmax,25}$ and $J_{max,25}$ in a separate step using their respective relationships to N_a.

1. Introduction

Many researchers tried to improve the estimates of terrestrial gross primary production (GPP) (Beer et al., 2010) in order to minimize uncertainty in global carbon cycle models. A combination of remote sensing data and either statistical upscaling using machine learning methods (Jung et al., 2011; Papale et al., 2015) or process-based modelling (Bonan et al., 2011; Turner et al., 2004) was mostly used as there is currently no measurement technique that is able to provide GPP estimates on areas larger than the typical footprint of an Eddy Covariance tower, which is approximately 1 km^2. In order to improve the performance of process-based models, spatially and temporally resolved information on plant photosynthetic capacity is required. To our knowledge, only a few studies have tried to estimate the maximum carboxylation capacity, V_{cmax}, and the maximum electron transport rate, J_{max}, directly from leaf reflectance measurements (Ainsworth et al., 2014; Doughty et al., 2011; Serbin et al., 2012). Doughty et al. (2011) obtained poor results ($R^2 \leq 0.5$) for 11 tropical tree species, while Serbin and colleagues obtained very good results ($R^2 = 0.9$) for three species (Ainsworth et al., 2014; Serbin et al., 2012). Serbin et al. (2015) also obtained very good results ($R^2 = 0.9$) for nine crop species using airborne canopy measurements. The strong divergence of the results at the leaf scale leads to the question whether the number or type of species included might affect the results in a negative way, which would be disadvantageous for large scale use of the models. Furthermore, shade leaves were not considered in the studies by Doughty et al. (2011), Serbin et al. (2012) and Ainsworth et al. (2014). Plant traits, however, are known to respond to different light environments, which can result in different

* Corresponding author at: Department Computational Hydrosystems, UFZ – Helmholtz Centre for Environmental Research, Permoser Str. 15, 04318 Leipzig, Germany.
E-mail address: benjamin.dechant@ufz.de (B. Dechant).

http://dx.doi.org/10.1016/j.rse.2017.05.019
0034-4257/

relationships between traits (e.g. Keenan and Niinemets, 2016; Seemann et al., 1987). This can also potentially affect the relationship between leaf reflectance and photosynthesis traits. The results of the canopy scale estimation for crop species (Serbin et al., 2015) seem promising but it is not clear if the method would work equally well for more heterogeneous natural vegetation. Serbin et al. (2012, 2015) raised the question whether there might be a direct link between apparent leaf reflectance (Kim et al., 1993) and V_{cmax} as well as J_{max} via Rubisco or chlorophyll fluorescence.

Many leaf physiological studies have investigated relationships between leaf chemical and morphological traits on the one hand and $V_{cmax,25}$ and $J_{max,25}$ on the other hand (the subscript 25 in $V_{cmax,25}$ and $J_{max,25}$ denotes normalisation to a reference temperature of 25 °C). It is well known that the Rubisco content in the leaf is closely related to $V_{cmax,25}$ (Jacob et al., 1995; Makino et al., 1994; Onoda, 2005) and that cytochrome *f* content has a strong linear relationship to $J_{max,25}$ (Onoda, 2005; Sudo et al., 2003; Terashima and Evans, 1988). These links are not only of statistical nature but represent mechanistic relationships. $V_{cmax,25}$ represents the maximum capacity of carboxylation which is mechanistically linked to the Rubisco content of the leaf, as the Rubisco enzyme catalyses the carboxylation reaction (von Caemmerer, 2000). Cytochrome *f* is known to play an important part in electron transport between the two photosystems (Onoda, 2005). It is unclear, however, if Rubisco and cytochrome *f* content in leaves can be retrieved accurately from reflectance measurements in the range of 400–2500 nm. Their signal in the reflectance spectrum is expected to be very weak compared to dominant absorbers, as these proteins are present in the leaves only in small concentrations and, in contrast to chlorophyll, are not built to strongly absorb light in this part of the spectrum. Therefore, it seems likely that the information related to $V_{cmax,25}$ and $J_{max,25}$ in the reflectance spectra, is not directly related to Rubisco and cytochrome *f* contents but to other leaf traits which have both a stronger absorption signal in this part of the spectrum and a strong but less direct relationship to the target parameters. Potential candidates for these leaf traits are leaf mass per area, *LMA* and nitrogen per area, N_a, as well as phosphorus content, which are highly correlated to both $V_{cmax,25}$ and $J_{max,25}$ (Kattge et al., 2009; Walker et al., 2014). The strong link between N_a and $V_{cmax,25}$ was described in the semi-mechanistic model of Farquhar et al. (1980). Both empirical and mechanistic relationships between N_a and $V_{cmax,25}$ have been used in Earth System Models (Rogers, 2014). The candidate trait with the strongest signal in the reflectance spectrum is *LMA*, but the one with the strongest relationship to $V_{cmax,25}$ and $J_{max,25}$, which also has a relatively strong absorption signal, is N_a (Jacquemoud et al., 1996; Kattge et al., 2009). Phosphorus has been shown to increase the sensitivity of $V_{cmax,25}$ to N_a for high nitrogen contents (Walker et al., 2014). This effect, however, is expected to be rather weak in temperate ecosystems due to the dominance of nitrogen limitation and relatively abundant phosphorus (Elser et al., 2007). Furthermore, phosphorus has been estimated from reflectance with consistently lower accuracy and precision than nitrogen content (Asner et al., 2011, 2014). Based on the above reasoning, we hypothesise that the main mechanism behind reflectance-based $V_{cmax,25}$ and $J_{max,25}$ estimation is the relationship to N_a. We also expect the reflectance-based estimation of $V_{cmax,25}$ and $J_{max,25}$ to be feasible across species, as the relationship to N_a was shown to be valid for a large number of temperate broadleaf tree species (Kattge et al., 2009).

Here, we present the results from a study designed to examine if 1) the estimation of $V_{cmax,25}$ and $J_{max,25}$ from leaf reflectance is possible for many different broadleaf tree species and leaves from different light environments and 2) the mechanism behind the estimation is the correlation to N_a. We studied reflectance-based multivariate regression models to estimate $V_{cmax,25}$ and $J_{max,25}$ and inspected the corresponding regression coefficients closely. The prediction performances of these multivariate regression models were compared to linear regression models based on estimated N_a and *LMA*. We also restricted the range of the reflectance spectrum in order to locate the relevant information contributing to the photosynthesis trait estimation. Furthermore, we inverted a leaf radiative transfer model to estimate chlorophyll, carotenoid and water contents in order to examine if these leaf traits also contributed to the multivariate regression models for $V_{cmax,25}$ and $J_{max,25}$.

2. Materials and methods

2.1. Study site and sampling

Measurements were conducted in the Arboretum of the University of Leipzig located in Großpösna, central Germany (51°15′ N, 12°29′ E, 150 m above sea level). The Arboretum has about 100 European deciduous and coniferous tree species in two different plantations. The main plantation has trees widely spaced with a minimum distance of three metres between trees, and a secondary plantation with minimum distances of 1 m. The trees in both plantations were planted as seedlings in 2012. Measurements were mostly conducted in the secondary plantation due to practical reasons. A total of 242 leaves of 103 trees were measured. For each leaf, the photosynthetic CO_2 response curve and leaf reflectance were measured and the nitrogen content and the dry leaf mass per area were determined destructively. Trees from 37 broadleaf species comprising 21 genera were sampled. Two species, *Morus alba* and *Prunus serotina*, were sampled more intensively with 29 and 42 leaves from 10 and 11 trees, respectively. 4.9 ± 2.6 leaves and 2.3 ± 0.9 trees were measured on average for each of the other 35 species. The list of the measured species including the exact number of leaves and trees per species is given in Table S1 of the supplementary material. The measurements were conducted mainly in August and early September of the year 2014. Observations were made throughout the day, normally between 9:00 CEST and 19:00 CEST.

The leaves were chosen to represent the widest possible range of light environments, from almost complete shade to direct sun light. This was done in order to obtain a wide range of values for the studied leaf traits and to test whether the relationships between leaf reflectance and photosynthesis traits would hold in different light environments. We attempted to cover the range of available light environments in cases where shading by nearby trees restricted the light environments of leaves to diffuse light. A classification of some of the sampled leaves into sun and shade leaves was done by a human observer based on the orientation and exposure to light throughout the day. Leaves exposed mostly to direct sunlight and top-of-canopy leaves exposed to diffuse light were classified as sun leaves, while strongly shaded leaves exposed only to diffuse light were classified as shade leaves.

2.2. Photosynthesis measurements and parameter estimation

2.2.1. In situ measurements

CO_2 response curves were measured in situ for each leaf, using portable Licor 6400 photosynthesis systems (Li-Cor Biosciences, Lincoln, NE, USA). Two of the photosynthesis systems were equipped with fluorescence chambers and one system had the standard blue/red LED light source chamber. All measurements were made with internal LED light sources and custom autoprogrammes. Leaf temperature was always actively stabilized. A set point of 25° C was chosen for most response curves, but lower temperature set points between 20° and 25° C were used for some of the response curves for practical reasons. Relative humidity was not actively stabilized for the sake of higher stability of the photosynthesis and transpiration measurements. The systems with fluorescence chambers had a smaller volume of the sample cell and therefore the flow rate was set to a lower value than for the LED light source chamber in order to obtain measurements with a precision comparable to the measurements with the larger chamber. Each leaf was first adapted to the chamber light, humidity and temperature environment for about 6 min before a CO_2 response curve was measured at a constant photosynthetic photon flux density of 1500 μmol m^{-2} s^{-1} with external CO_2 concentrations in the following order in μmol mol^{-1}:

400, 300, 200, 100, 400, 400, 400, 500, 600, 700, 800, 1000, 1200, 1400, 1600, 1900. Each step in the response curve lasted a minimum of 1 min and a maximum of 2.5 min.

2.2.2. Parameter estimation

$V_{cmax,25}$ and $J_{max,25}$ were estimated from the response curve measurements described above by fitting the Farquhar-von Caemmerer-Berry photosynthesis model (Farquhar et al., 1980) to the photosynthesis-intercellular CO_2 concentration (A-C_i) curves. The model was programmed according to Knorr (2000) with constants such as activation energies and Michaelis-Menten constants from Knorr (2000) and von Caemmerer (2000). The temperature relationship from Knorr (2000) was used for the CO_2 compensation point and the dark respiration at 25 °C was modelled as a fraction of V_{cmax}. Observed stomatal CO_2 concentrations, C_i, rather than chloroplastic concentrations, C_c, were used as the introduction of mesophyll conductance formulations into the model did not improve the fit results. The actual leaf temperatures for each point of the response curves were used for parameter estimation. The photosynthesis model inversion was conducted using the genetic optimization algorithm Differential Evolution (Price et al., 2005) as implemented in the R package DEoptim (Ardia et al., 2015; Mullen et al., 2011).

2.3. Reflectance measurements and post-processing

Leaf reflectance was measured in situ using an ASD FieldSpec 4 spectrometer with a 2 m long fibre optic and equipped with the ASD plant probe (Analytical Spectral Devices Inc., Boulder, Colorado, USA). Three different parts of the leaf adaxial surface were measured if the leaf size was large enough. A minimum of three scans were used for each measurement and white reference measurements were performed before and after measuring the leaves of each individual tree using the spectral on surface of the leaf clip. The leaf reflectance measurements were normally conducted on the same day as the photosynthesis measurements and in exceptional cases due to bad weather with a delay of 1 day. Delays of the reflectance measurements of the order of a day are not a problem for our study, as $V_{cmax,25}$ and $J_{max,25}$ are slow-changing properties of the leaf (e.g. Xu and Baldocchi, 2003), which are not affected by diurnal changes in stomatal conductance, leaf temperature or light conditions (von Caemmerer, 2000). Leaf reflectance was measured after the photosynthesis measurements in order to avoid potential physiological changes induced by the high light intensity of the ASD plant probe. Mean reflectance spectra were calculated for each leaf by averaging spectra measured on different parts of the leaf. The steps in the spectra located at the transitions (splice points) between different detectors of the spectrometer were removed with a multiplicative correction based on the VNIR detector (350–1000 nm) as reference. This correction method was chosen based on observations of the white reference after measuring the leaves.

2.4. Leaf sample collection and chemical analysis

Measured leaves were typically collected on the same day or 1 day later. They were sealed in plastic bags and transported back to the research institute in a cooling box filled with cooling gel packs and transferred to a refrigerator where they were stored at −18 °C. Leaf area was determined by scanning the leaves with a commercial flat-bed scanner and analysing the images using a custom program written for ImageJ (Schneider et al., 2012), which was based on a simple thresholding approach. The leaves were dried at 80–90 °C for about 72 h for determining the dry weight. The leaves were ground and homogenised using a mixer mill (MM 200, Retsch, Haan, Germany) and analysed by dry combustion using a C/H/N analyser (Vario El III, Elementar, Hanau, Germany) to determine total nitrogen contents. This analysis was conducted for two samples per leaf and the results of the two samples were later averaged. Nitrogen content per area, N_a, was calculated from the chemically determined content per mass by multiplying with leaf mass per area, *LMA*.

2.5. PROSPECT inversion

We did not determine chlorophyll, carotenoid and water content of the leaves chemically as they can be estimated from leaf reflectance with high precision using well established methods (Datt, 1999; Féret et al., 2008; Gitelson et al., 2003; Gitelson et al., 2006). We inverted the leaf radiative transfer model PROSPECT 5B (Féret et al., 2008; Jacquemoud and Baret, 1990) to estimate these leaf chemical contents. The R package hsdar (Lehnert et al., 2015) was used for the PROSPECT model and the R package DEoptim (Ardia et al., 2015; Mullen et al., 2011) was used for the optimisation. We estimated the mesophyll structure parameter, S^{PRO}, the chlorophyll a + b content, C_{ab}^{PRO}, the carotenoid content, C_{car}^{PRO}, the water content, C_w^{PRO}, and the dry matter content, C_m^{PRO}.

2.6. Statistical analysis

2.6.1. Overall framework

The analysis of the measured data was conducted in a way that 1) allowed investigating if the partial least square (PLS) models (described in more detail in section 2.6.2) for the two photosynthesis traits were based on the correlation to N_a and that 2) reflected realistic estimation approaches from a remote sensing perspective. An overview of our approaches is presented in the form of a flowchart in Fig. 1. First, we estimated $V_{cmax,25}$ and $J_{max,25}$ directly from reflectance using PLS regression (method 1 in Fig. 1, abbreviated $\mathscr{R} \rightarrow Ph$). Then, we estimated the photosynthesis traits by taking the reflectance-based PLS estimate N_a^{PLS} as input to the corresponding linear models that were calibrated on the measured N_a values (method 2 in Fig. 1, abbreviated as $\mathscr{R} \rightarrow N_a^{PLS} \rightarrow Ph$, where '$\mathscr{R}$' stands for reflectance and the superscript 'PLS' indicates the method of estimation). The results of these models represent the prediction performance assuming that the PLS models are based on N_a taking into account the imperfect estimation of N_a from reflectance using the PLS method. LMA^{PLS} was added as a second predictor to the models, in order to test its effect on prediction performance. The third method used estimated leaf traits obtained by inverting the leaf radiative transfer model PROSPECT (Jacquemoud and Baret, 1990) and the PLS method to predict the photosynthesis traits (method 3 in Fig. 1, abbreviated $\mathscr{R} \rightarrow T^{PRO} \rightarrow Ph$, where T^{PRO} indicates the leaf traits estimated via PROSPECT inversion). This method was used to determine if 1) it is feasible to estimate photosynthesis traits based on estimated PROSPECT leaf traits and 2) leaf traits such as chlorophyll and water contents, which were not directly measured but potentially contribute to the estimation, could explain a superior performance of method 1 compared to method 2. We used repeated cross-validation to evaluate prediction performance for all methods. The number of repetitions was 200 and in each case a ten-fold cross-validation was performed. The same cross-validation segments were used for all methods. We used bootstrapping to estimate the standard deviation of the median values of prediction performance statistics and used these estimates to calculate the standard error of the median. Further details of the different methods for photosynthesis trait estimation are given in the sections below. All statistical and chemometric analyses were conducted using the programming language R (R Core Team, 2012) with the exception of Gaussian process regression for which Matlab (The MathWorks, Inc., Natick, Massachusetts, United States) was used.

2.6.2. Chemometric analysis with PLS

We used partial least square (PLS) regression (Wold et al., 1984) for the estimation of leaf traits from reflectance. Only the wavelengths from 500 nm to 2350 nm were chosen in order to avoid the parts of the

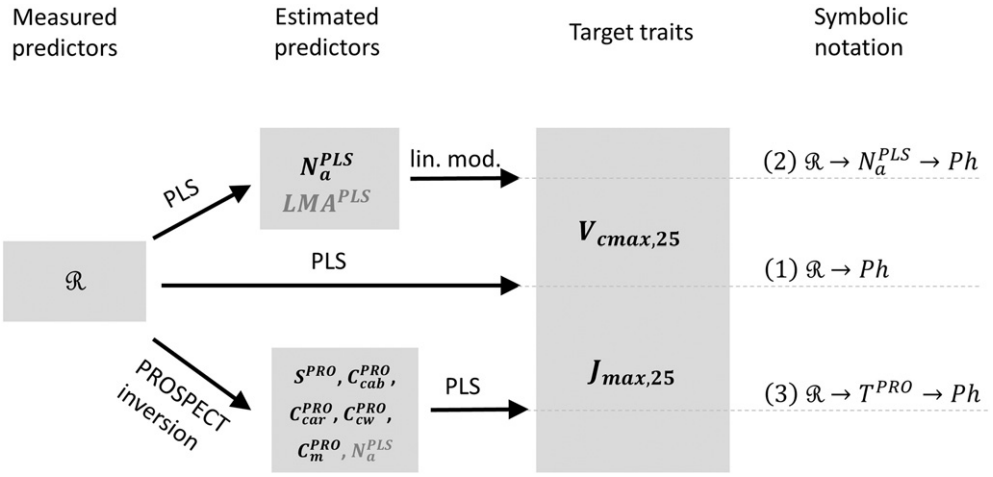

Fig. 1. Flowchart of different models for photosynthesis trait estimation. Measured predictors (first column), predictors estimated from reflectance, $\mathcal{R}$, (second column), target traits which are predicted (third column), and the symbolic notation for each method (fourth column) are shown. (1) are the direct reflectance-based PLS models for the photosynthesis parameters, *Ph*, (2) are the linear models based on PLS-estimated nitrogen content, N_a^{PLS}, and leaf mass per area, LMA^{PLS}, and (3) represent trait-based PLS models with leaf traits, T^{PRO}, estimated by PROSPECT inversion (and N_a^{PLS}). The linear models in method (2) were calibrated on the measured N_a values and then applied to N_a^{PLS}.

spectrum with a low signal to noise ratio. The PLS models for the photosynthesis traits resulted in final models of the form (Eq. 1).

$$T = \sum_{i=1}^{M} a_i \mathcal{R}_i + b \quad (1)$$

where T stands for one of the four measured leaf traits ($V_{cmax,25}$, $J_{max,25}$, LMA, N_a), and $\mathcal{R}_i$ indicates the measured leaf reflectance at channel i. M is the number of wavelength channels, a_i are the regression coefficients at channel i and b is the offset.

PLS regression was run for each leaf trait separately using repeated double cross-validation (rdCV) (Filzmoser et al., 2009). rdCV can be used to both find the optimal number of PLS components in an automated way as well as validate the models by estimating the internal prediction performance. This is done by performing repeated splits into calibration and test data sets using a cross-validation procedure (outer CV loop). Each calibration set is again split into training and validation sets using cross-validation (inner CV loop) to estimate the optimal number of components. Prediction performance is evaluated on the test set (outer loop). rdCV avoids the risk of being misled by splits with particularly good or bad results purely by chance by splitting into calibration and test sets repeatedly, which results in an evaluation of the average prediction performance over many possible splits. The repeated splitting procedure also results in an estimate of the uncertainty for each predicted value and therefore of the prediction performance. The R package chemometrics (Filzmoser and Varmuza, 2012) which uses the PLS algorithm implemented in the package pls (Mevik et al., 2013) was used for this analysis.

We used 200 repetitions for the rdCV and ten-fold cross-validation for both the inner and the outer CV loop rather than the recommended four-fold CV for the outer loop and seven-fold CV for the inner loop (Filzmoser et al., 2009), as we found that for some of the leaf traits, the internal prediction performance was still increasing considerably for a number of calibration data n < 200. In such a case it is necessary to maximise the number of training data in order not to underestimate the prediction performance (Friedman et al., 2001), which can be achieved with a higher number of folds in the cross-validation procedure. Instead of 64% of the dataset used for the training set (3/4 for calibration, 6/7 of the calibration set for training) with the settings recommended by Filzmoser et al. (2009), 81% of the data was used for training (9/10 for calibration, 9/10 of the calibration set for training) which corresponds to almost 200 data points and is therefore close to the training saturation for most of the traits. The optimal number of PLS components was essentially determined by searching for the minimum RMSE of the inner CV loop, but if there was a smaller number of components with a RMSE only higher than the minimum RMSE plus 0.1 times the standard error, this smaller number was selected (Fig. S1). The regression coefficients were calculated for the optimum number of PLS components of each fold and averaged to obtain a mean coefficient spectrum for each repetition.

2.6.3. Outlier removal

An outlier removal was performed as strong residual outliers were detected in the *LMA* model, which had a direct impact on the N_a models. A four standard deviation residual outlier threshold was used, which led to the removal of eight data points clearly separated from the rest of the data (Fig. S2 in the supplementary material). We verified that the outliers were not specific to the PLS method as the same points also appeared as outliers in the PROSPECT inversion results for dry matter content. All analyses in this study were conducted on the dataset without the outliers (n = 242). We verified, however, that including the outliers in the analysis did not change the main conclusions.

2.6.4. Trait-based PLS analysis

Since PLS models are maximizing the covariance between target trait and predictors in each step (Maitra and Yan, 2008), potentially all correlations between target traits and features in the reflectance spectrum are used. The main features in the reflectance spectrum are represented by the parameters of the PROSPECT model. Therefore, we studied models combining the estimated PROSPECT parameters and PLS-estimated N_a. Multiple linear regression models including all PROSPECT traits were studied in a first approach. However, considerable collinearity was detected both when using variable inflation factors and when using condition index analysis for which the R packages car (Fox and Weisberg, 2011) and perturb (Hendrickx, 2012) were used, respectively. Multicollinearity affects the regression in a negative way as it inflates the estimation errors of the coefficients (Farrar and Glauber, 1967; Gunst, 1983), which can lead to erroneous results. In PLS, however, multicollinearity is removed by constructing orthogonal latent variables which are then used for a multiple regression (Dormann et al., 2013; Wold et al., 1984). PROSPECT trait-based PLS models were studied using repeated double cross-validation, rdCV, with the same settings as for the reflectance-based models described above. The estimated PROSPECT traits (and N_a^{PLS}) were mean-centred and scaled to unit variance as the different traits had units differing by up to four orders of magnitude.

2.6.5. Analysis of model performance for different parts of the spectrum

Chlorophyll absorption and chlorophyll fluorescence emission are known to be restricted to the visible-near-infrared (VNIR) part of the spectrum (500-900 nm in this study) (Buschmann, 2007), while nitrogen/protein most strongly absorbs in the short-wave-infrared (SWIR) part of the spectrum (900–2350 nm in this study) (Jacquemoud et al., 1996). These spectral regions were therefore studied separately in order to examine their respective contributions to the prediction performance of the full range spectral models. The same methods were used for the VNIR and SWIR as for the full range.

3. Results

3.1. Reflectance spectra

The measured reflectance spectra, presented in Fig. 2, had a wide range of variation in the regions of the spectrum related to chlorophyll content (450–750 nm), leaf mesophyll structure (800–1250 nm), water (1300–2500 nm), and dry matter content (1600–2500 nm). The range of leaf reflectance in the spectral region 450–750 nm was comparable to the ranges measured for a very large dataset including several thousand tropical tree species (Asner et al., 2011). The range in the rest of the spectrum, however, was notably smaller in our dataset. This was due to overall lower values in the 800–1250 nm region, where our measurements did not exceed a value of 0.6 and generally higher values in the region 1300–2500 nm. Nevertheless, the range of our reflectance spectra was considerably larger than the ranges observed in datasets used for other studies investigating the estimation of the maximum carboxylation capacity ($V_{cmax,25}$) and the maximum electron transport rate ($J_{max,25}$), from leaf reflectance (Ainsworth et al., 2014; Serbin et al., 2012).

3.2. Leaf traits

3.2.1. Measured leaf traits

Measured leaf mass per area (*LMA*), and nitrogen content per area (N_a),had considerable ranges of values (Table 1), comparable to or larger than the corresponding ranges in the LOPEX dataset, which is based on >300 leaves of 50 tree, herbaceous and crop species (Hosgood et al., 1995; Jacquemoud et al., 1996). This was also true for the leaf traits estimated via PROSPECT inversion (Table S2), with the exception of leaf water content, which reached higher values in the LOPEX dataset. Our dataset had a much larger range of *LMA* values compared to a previous study by Serbin et al. (2012), but the ranges of N_a were comparable. The range of $V_{cmax,25}$ was comparable to the study of Serbin et al. (2012) and the range of $J_{max,25}$ values was considerably larger due to the presence

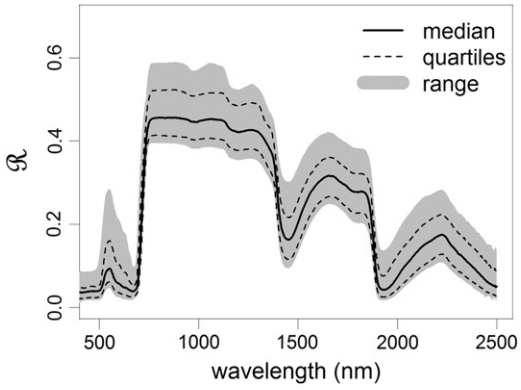

Fig. 2. Overview of leaf reflectance ($\mathfrak{R}$) spectra measured on 242 individual leaves from 37 deciduous tree species. The median reflectance spectrum (continuous black line), the first and third quartiles (dashed black lines) and the range of all measured reflectance spectra (grey shaded area) are shown.

Table 1
Summary statistics of the leaf traits measured in the field. Mean, standard deviation, minimum, and maximum values are given for maximum carboxylation capacity, $V_{cmax,25}$, and maximum electron transport rate, $J_{max,25}$, from A-C_i curves as well as leaf mass per area, *LMA*, and nitrogen content per area, N_a.

Trait	Unit	Mean	Stand. dev.	Min.	Max.
$V_{cmax,25}$	μmol $m^{-2}s^{-1}$	77.5	28.9	24.5	181.1
$J_{max,25}$	μmol $m^{-2}s^{-1}$	140.6	50.9	30.1	295.8
N_a	g m^{-2}	1.53	0.51	0.52	3.16
LMA	g m^{-2}	76.7	21.7	32.5	126.6

of both higher and lower values. Compared to the study of Doughty et al. (2011), the range of both photosynthesis parameters was much larger and the observed values of both photosynthesis traits were consistent with the literature (von Caemmerer, 2000). The ratio $J_{max,25}/V_{cmax,25}$ was 1.85 on average, with a standard deviation of 0.33. These values are in accordance with other studies on the subject (e.g. von Caemmerer, 2000; Onoda, 2005; Wullschleger, 1993). Considerable differences among the species were observed for the photosynthesis traits, N_a and *LMA* (Table S1), but they were probably partly caused by different growing conditions. Clear differences were also observed in the trait values of sun and shade leaves, with a clear tendency towards higher values for the traits for sun leaves (results not shown).

Correlations among the measured leaf traits were all high and significant, as shown in Fig. 3. The strongest correlation was observed between the two photosynthesis traits $V_{cmax,25}$ and $J_{max,25}$ ($\rho = 0.86$) and confirms the relationship known from the literature (Leuning, 1997; Wullschleger, 1993). Both $V_{cmax,25}$ and $J_{max,25}$ were highly correlated to N_a ($\rho = 0.82$, $\rho = 0.83$ respectively), whereas the correlations to *LMA* were considerably lower ($\rho = 0.47$ and $\rho = 0.59$, respectively). The correlations between photosynthesis traits and nitrogen content per mass were also much lower than the corresponding values for N_a ($\rho = 0.48$ and $\rho = 0.38$, respectively). Therefore, nitrogen content per mass was not considered for the rest of the analysis. We observed that the correlation of nitrogen per mass to photosynthesis traits was higher for sun leaves and very low (≤0.2) for shade leaves (results not shown).

3.2.2. Leaf traits estimated by PROSPECT inversion

The inversion of the leaf radiative transfer model PROSPECT worked very well, as measured by the quality of the fit when comparing the simulated and measured reflectance spectra in the range of 400–2500 nm: the median RMSE and its standard error was 0.0090 ± 0.0001 and the median R^2 was 0.9967 ± 0.0001. The good performance of the inversion procedure was also reflected in the results of the comparison of estimated vs. measured *LMA* (after bias correction: $R^2 = 0.71$, RMSE = 11.59 g m^{-2}; Fig. S3). Although there was a global bias of 20.23 g m^{-2}, the correlation was very high ($\rho = 0.92$). Very high values of chlorophyll and carotenoid contents were observed for five leaves (above 100 μg cm^{-2} and 17 μg cm^{-2} for chlorophyll and carotenoids, respectively) but the corresponding RMSEs did not indicate outliers in terms of quality of spectral fit. For the other leaves, the values of the estimated traits were in accordance with the findings of other studies (e.g. Hosgood et al., 1995; Jacquemoud et al., 1996).

None of the leaf traits estimated by PROSPECT inversion had correlations to $V_{cmax,25}$ and $J_{max,25}$ as high as N_a (Fig. S2), but the estimated water content showed a rather strong relationship ($\rho = 0.54$, $\rho = 0.65$ respectively). Chlorophyll content was only moderately correlated to the photosynthesis traits ($\rho = 0.4$).

3.3. Results of regression models

3.3.1. Direct PLS models ($\mathfrak{R} \rightarrow Ph$)

We pooled the data of all species to estimate the measured leaf traits from leaf reflectance using partial least squares (PLS) regression (Eq. 1). The full range PLS models had reasonable predictive performance for

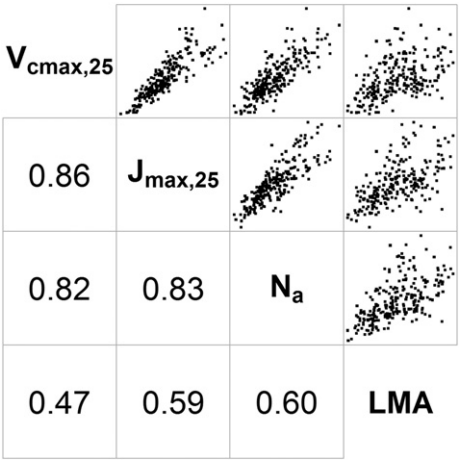

Fig. 3. Correlations between measured maximum carboxylation capacity, $V_{cmax,25}$, the maximum electron transport rate, $J_{max,25}$ nitrogen content per area, N_a, and leaf mass per area, *LMA*. The matrix of Pearson correlation coefficients between measured leaf traits is shown below the diagonal. The corresponding scatterplots are shown above the diagonal in relative units. All correlations were highly significant.

the photosynthesis parameters $V_{cmax,25}$ ($R^2 = 0.64$) and $J_{max,25}$ ($R^2 = 0.70$), as shown in Fig. 4. We observed a tendency for underestimation of the high values, in particular there were three strongly underestimated points with measured values around and above 150 μmol $m^{-2}s^{-1}$. However, these points did not have a particularly large prediction uncertainty over the 200 cross-validation repetitions, indicating a stable estimation. The estimation for $J_{max,25}$ also showed a tendency towards underestimation of large values. Three points with relatively high values were clearly separated from the rest of the data in terms of residuals, but did not show large prediction uncertainties. The PLS models for N_a and *LMA* had very good predictive performances across the full range of measured values ($R^2 = 0.88$ and $R^2 = 0.95$, respectively). The few points with higher residuals did not have unusually large prediction uncertainties. The notable differences in prediction performance of the models for different traits were also clearly reflected in the RMSE normalised to the mean value (NRMSE): NRMSE was 6% for *LMA*, 12% for N_a, while it was 22% for $V_{cmax,25}$ and 20% for $J_{max,25}$.

We also investigated model performances for different spectral regions and present the results in Table 2 in terms of R^2. The patterns for RMSE were analogous and are presented in Table S3. The performance of photosynthesis trait estimation based on the SWIR part of the spectrum was slightly worse than for those based on the full range but still relatively close ($R^2 = 0.61$ for $V_{cmax,25}$ and $R^2 = 0.67$ for $J_{max,25}$). The VNIR models had the worst performances ($R^2 = 0.60$ for $V_{cmax,25}$ and $R^2 = 0.66$ for $J_{max,25}$), but differences to the SWIR results were very small. The N_a and *LMA* models had slightly higher prediction performances in the SWIR than for the full range ($R^2 = 0.889$ vs. $R^2 = 0.881$ for N_a and $R^2 = 0.951$ vs. $R^2 = 0.950$ for *LMA*; p − value $< 2.2 \times 10^{-16}$ for both RMSE and R^2 and both leaf traits). The decrease in performance

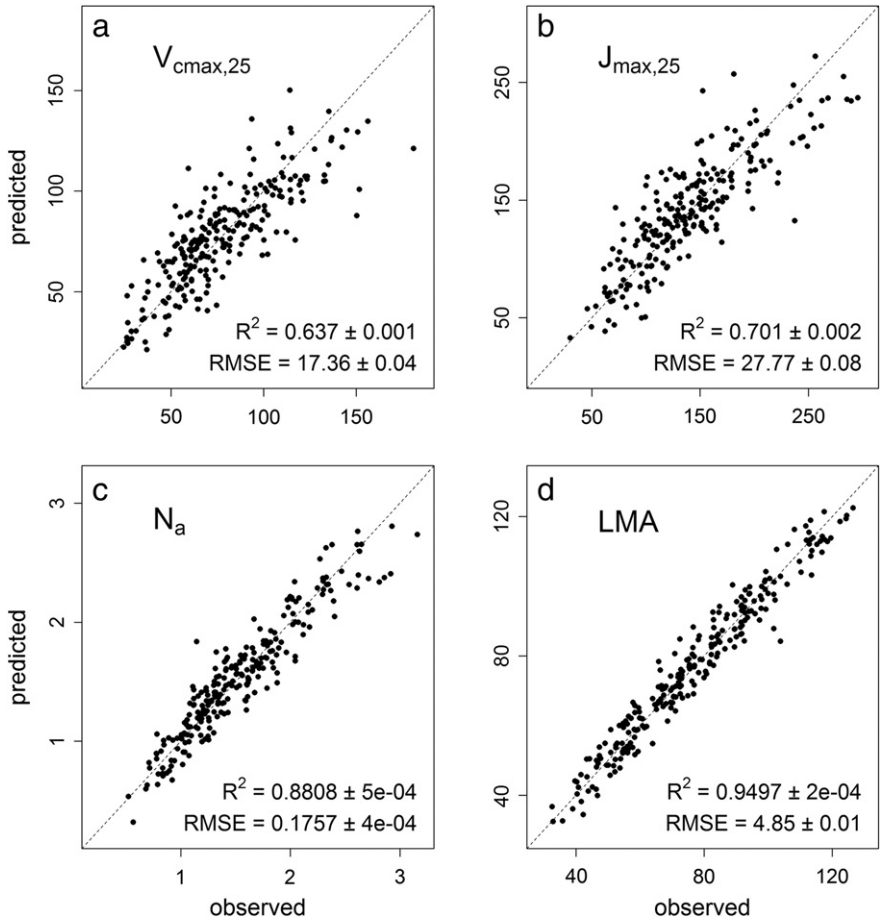

Fig. 4. Results of the PLS models based on the full range reflectance spectra for the four measured leaf traits. The comparisons of predicted vs. observed values of (a) the maximum carboxylation capacity, $V_{cmax,25}$, and (b) the maximum electron transport rate, $J_{max,25}$, in μmol $m^{-2}s^{-1}$, (c) the nitrogen content per area, N_a, and (d) the leaf mass per area, *LMA*, in g m^{-2} are shown. Dashed lines are the one-to-one lines. The coefficient of determination, R^2, and the root mean square error, RMSE, indicate the median values over the 200 repetitions and are given with their corresponding standard errors. The points in the scatterplots are the predictions of the PLS models averaged over the 200 repetitions. The standard errors of the predicted values had medians of 0.30, 0.48, 0.0024, and 0.066 for $V_{cmax,25}$, $J_{max,25}$, N_a, and *LMA*, respectively.

Table 2
Overview of different trait-based and reflectance-based regression models for the maximum carboxylation capacity, $V_{cmax,25}$, and the maximum electron transport rate, $J_{max,25}$, nitrogen content per area, N_a, and leaf mass per area, *LMA*. Model predictors are either reflectance, $\mathcal{R}$, estimated N_a and *LMA* using a reflectance-based PLS regression model or traits estimate via PROSPECT inversion, indicated by the superscript 'PLS' and 'PRO'. The latter traits are the leaf mesophyll structure parameter, S^{PRO}, the chlorophyll a + b content, C_{ab}^{PRO}, the carotenoid content, C_{car}^{PRO}, the water content, C_w^{PRO}, and the dry matter content, C_m^{PRO}. The median coefficient of determination, R^2, is shown for all models. The standard error of the median was $\leq 1 \times 10^{-3}$ in all cases except the full range reflectance-based model for $J_{max,25}$ where it was 2×10^{-3}. For $V_{cmax,25}$ and $J_{max,25}$, the highest R^2 values in each row are printed in bold, the highest values in each column are underlined and the overall highest values for both traits are highlighted with an asterisk.

Modelled trait	Predictor(s)	Method	R^2		
			Full	VNIR	SWIR
$V_{cmax,25}$	$\mathcal{R}$	PLS	**0.637**	0.596	0.610
	N_a^{PLS}	linear	0.641	0.552	**0.645**
	N_a^{PLS}, LMA^{PLS}	linear	0.639	0.549	**0.643**
	S^{PRO}, C_{ab}^{PRO}, C_{car}^{PRO}, C_w^{PRO}, C_m^{PRO}	PLS	0.395	–	–
	S^{PRO}, C_{ab}^{PRO}, C_{car}^{PRO}, C_w^{PRO}, C_m^{PRO}, N_a^{PLS}	PLS	0.646*	–	–
$J_{max,25}$	$\mathcal{R}$	PLS	**0.701***	0.656	0.667
	N_a^{PLS}	linear	0.639	0.582	**0.647**
	N_a^{PLS}, LMA^{PLS}	linear	0.653	0.610	**0.659**
	S^{PRO}, C_{ab}^{PRO}, C_{car}^{PRO}, C_w^{PRO}, C_m^{PRO}	PLS	0.466	–	–
	S^{PRO}, C_{ab}^{PRO}, C_{car}^{PRO}, C_w^{PRO}, C_m^{PRO}, N_a^{PLS}	PLS	0.644	–	–
N_a	$\mathcal{R}$	PLS	0.881	0.756	**0.889**
LMA	$\mathcal{R}$	PLS	0.950	0.679	**0.951**

of the VNIR models compared to the full range and SWIR models was very clear ($R^2 = 0.756$ for *LMA* and $R^2 = 0.679$ for N_a).

3.3.2. Nitrogen-based ($\mathcal{R} \rightarrow N_a^{PLS} \rightarrow Ph$) and PROSPECT inversion-based models ($\mathcal{R} \rightarrow T^{PRO} \rightarrow Ph$)

We compared different methods of estimating the photosynthesis traits $V_{cmax,25}$ and $J_{max,25}$ in order to gain insights into the mechanism behind the reflectance-based estimation (Fig. 1). The results of this analysis are presented in Table 2 and show that the linear model based on PLS-estimated N_a ($\mathcal{R} \rightarrow N_a^{PLS} \rightarrow Ph$) performed slightly better than the reflectance-based PLS model ($\mathcal{R} \rightarrow Ph$) for $V_{cmax,25}$ ($R^2 = 0.641$ and $R^2 = 0.637$, respectively). The differences were small but significant (p-value $< 3 \times 10^{-6}$). Adding PLS-estimated *LMA* (Table 2) or the leaf traits estimated by PROSPECT inversion did not clearly improve the N_a^{PLS}-based models (Table 2).

The prediction performance for $J_{max,25}$ was best for the reflectance-based PLS model ($\mathcal{R} \rightarrow Ph$; $R^2 = 0.70$). However, N_a explained a large part of the model performance ($R^2 = 0.639$ for $\mathcal{R} \rightarrow N_a^{PLS} \rightarrow J_{max,25}$). Adding *LMA* as predictor slightly improved the results ($R^2 = 0.653$, RMSE 29.9). The performance of the $\mathcal{R} \rightarrow T^{PRO} \rightarrow Ph$ models were clearly worse than the $\mathcal{R} \rightarrow Ph$ and the $\mathcal{R} \rightarrow N_a^{PLS} \rightarrow Ph$ models ($R^2 = 0.40$ for $V_{cmax,25}$ and $R^2 = 0.47$ for $J_{max,25}$; Table 2). Only a small improvement over the $\mathcal{R} \rightarrow N_a^{PLS} \rightarrow J_{max,25}$ was observed, when the PROSPECT traits were added ($R^2 = 0.643$ for $\mathcal{R} \rightarrow (T^{PRO}, N_a^{PLS}) \rightarrow J_{max,25}$, Table 2).

We also analysed the performance of the $\mathcal{R} \rightarrow N_a^{PLS} \rightarrow Ph$ models using either the VNIR or SWIR parts of the spectrum (Table 2). The $\mathcal{R} \rightarrow N_a^{PLS} \rightarrow V_{cmax,25}$ models outperformed the $\mathcal{R} \rightarrow V_{cmax,25}$ models with a clear difference in the SWIR ($R^2 = 0.64$ vs. $R^2 = 0.61$), but had lower performance in the VNIR ($R^2 = 0.55$ vs. $R^2 = 0.60$). In contrast, all $\mathcal{R} \rightarrow N_a^{PLS} \rightarrow J_{max,25}$ models had lower prediction performance than the corresponding $\mathcal{R} \rightarrow J_{max,25}$ models. The results of all $\mathcal{R} \rightarrow N_a^{PLS} \rightarrow J_{max,25}$ improved slightly when LMA^{PLS} was added as predictor. The $\mathcal{R} \rightarrow N_a^{PLS}, LMA^{PLS} \rightarrow J_{max,25}$ model showed similar performance as the $\mathcal{R} \rightarrow J_{max,25}$ model in the SWIR ($R^2 = 0.659$ vs. $R^2 = 0.667$). The full range models using PLS-estimated traits had similar prediction performances as the corresponding SWIR models, reflecting the results of the reflectance-based models for N_a and *LMA*.

3.4. PLS regression coefficients

3.4.1. Correlation analysis

We inspected the regression coefficients of the PLS models, as they contain information on the spectral features used for the estimation. The regression coefficients of $V_{cmax,25}$, $J_{max,25}$ and N_a showed strong similarities in the SWIR (Fig. 5), but the situation was not as clear for the full range and VNIR models (Figs. 6 and S4). The Pearson correlation between regression coefficients was used as similarity measure in order to make a quantitative comparison (Table 3). We found the highest correlation between the PLS model coefficients of $V_{cmax,25}$ and N_a in the SWIR ($\rho = 0.80$), slightly lower values in the VNIR ($\rho = 0.78$) and the lowest correlation for the full range coefficients ($\rho = 0.66$). The situation was similar for $J_{max,25}$ but with generally lower values and the lowest correlation in the VNIR ($\rho = 0.57$, $\rho = 0.44$ and $\rho = 0.76$ for full range, VNIR and SWIR, respectively). The correlation between the coefficients of the $V_{cmax,25}$ and $J_{max,25}$ models was also highest in the SWIR ($\rho = 0.79$), lowest in the VNIR ($\rho = 0.48$), and at an intermediate value for the full range ($\rho = 0.71$). The results of the VNIR and SWIR parts of the full range coefficient spectra were similar to the results of the full spectra (not shown). The only case where the correlation between a photosynthesis trait's coefficients was higher to *LMA* than to N_a coefficients was the VNIR model for $J_{max,25}$ ($\rho = 0.57$ to *LMA* vs. $\rho = 0.44$ to N_a). The $V_{cmax,25}$ model coefficients did not show high correlations to *LMA* coefficients ($\rho < 0.37$ in all cases).

3.4.2. SWIR regression coefficients

As far as the shape of the SWIR regression coefficients (Fig. 5) is concerned, there were distinct features in the $V_{cmax,25}$, $J_{max,25}$ and N_a coefficient spectra at 1690 nm, 1734 nm, 2060 nm and 2180 nm which have been reported to be related to nitrogen and proteins (Curran, 1989) except for the 1734 nm feature. The latter feature was reported to be related to protein absorption without giving the details of the mechanism (Williams, 2006). The $V_{cmax,25}$, $J_{max,25}$ and N_a coefficient spectra also showed a strong and very narrow, negative feature at 2312 nm which was close to the 2300 nm N—H feature reported by Curran (1989). There was a broader negative feature in the $V_{cmax,25}$, $J_{max,25}$ and N_a coefficient spectra at the 1510 nm N—H absorption (Curran, 1989).

The *LMA* regression coefficients did not show clear features at the locations of the nitrogen-related absorptions. The absorption features reported by Curran (1989) and Williams (2006) coincide with local maxima of the specific absorption coefficients for protein in dry leaves estimated with the PROSPECT model (Jacquemoud et al., 1996) and the pseudo-absorption spectrum [$\log(\mathcal{R}^{-1})$] of Rubisco (Elvidge, 1990), shown in Fig. 5.

3.4.3. VNIR regression coefficients

In contrast to the SWIR regression coefficients, the VNIR coefficients had very different shapes for the VNIR model and the VNIR part of the full range model, as shown in Fig. 6. The coefficients of the VNIR models did not show strong features in the red-edge region (650–800 nm) except for $V_{cmax,25}$. All these regression coefficients showed a strong decrease towards 900 nm. Furthermore, all VNIR regression coefficients except the one for *LMA* had very narrow negative features between 500 and 600 nm close to known absorption features of cytochrome b_6f reported by Nelson and Neumann (1972). The VNIR part of the full range model coefficients had the strongest features around the red-edge part of the spectrum, some of which coincided with the peaks of chlorophyll fluorescence emission. The position and strength of the negative features varied slightly for the different leaf traits.

4. Discussion

4.1. Reflectance-based PLS models

Our results showed that the estimation of $V_{cmax,25}$ and $J_{max,25}$ for leaves from a large number of tree species and light environments is

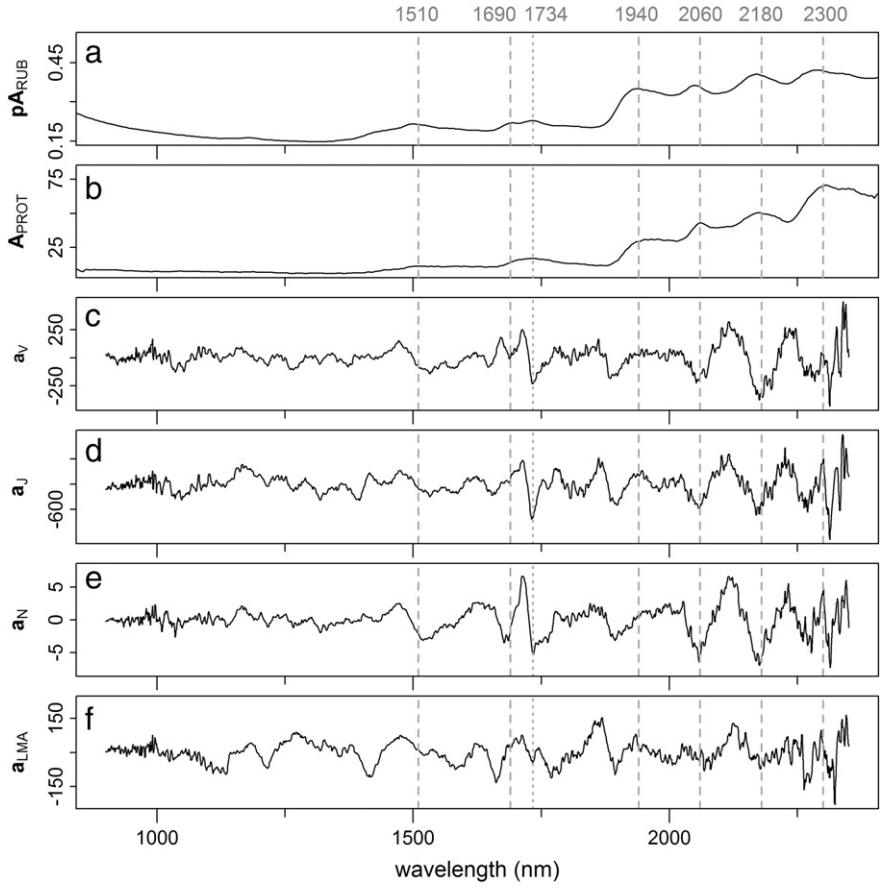

Fig. 5. PLS coefficients of the SWIR models for the four measured leaf traits (c–f) and protein absorption spectra for reference (a, b): a) Rubisco pseudo-absorption, pA_{RUB}, based on measured reflectance (Elvidge, 1990) and b) estimated specific absorption coefficient for proteins in dry leaves, A_{PROT}, using PROSPECT (Jacquemoud et al., 1996), c) PLS coefficients for maximum carboxylation capacity, $V_{cmax,25}$, d) the maximum electron transport rate, $J_{max,25}$, e) nitrogen content per area, N_a, and f) leaf mass per area, *LMA*. The coefficients represent the average values of the 200 repetitions. The standard errors of the mean of the regression coefficients were two orders of magnitude smaller than the maximum coefficient for each trait and are therefore not shown. The dashed grey vertical lines indicate protein/nitrogen absorption features reported in the literature (Curran, 1989). The dotted grey vertical line indicates an additional wavelength reported to be related to protein (Williams, 2006). The exact values of the wavelengths in nanometres are given at the top of the figure in grey.

feasible with a single model and that this can be explained by the underlying relationships between photosynthetic parameters and N_a. Several other studies investigated the estimation of these leaf traits from reflectance measurements. None of them, however, specifically included shaded leaves despite their importance at large scales and the fact that the relationships between the traits may differ from sun leaves (e.g. Keenan and Niinemets, 2016; Seemann et al., 1987). Doughty et al. (2011) obtained models with rather low prediction performance ($R^2 = 0.4$ and $R^2 = 0.5$ for $V_{cmax,25}$ and $J_{max,25}$, resp.) for 11 different species, while the results of Serbin et al. (2012) for three species grown in a greenhouse were very good ($R^2 = 0.89$ and $R^2 = 0.93$ based on leave-one-out cross-validation for V_{cmax} and J_{max}, resp.; $R^2 = 0.77$ and $R^2 = 0.78$ based on repeated 70%/30% cross-validation for V_{cmax} and J_{max}, resp.). Ainsworth et al. (2014) obtained similarly good results for V_{cmax} estimation of two soybean genotypes ($R^2 = 0.88$ based on leave-one-out cross-validation). Our results were intermediate between those of Doughty et al. (2011) and Serbin et al. (2012) in terms of R^2, but we used a much more diverse dataset both in terms of species and light environments. When comparing prediction performance in terms of RMSE rather than R^2, our results were notably closer to those of Serbin et al. (2012) and Ainsworth et al. (2014) for V_{cmax} than for J_{max}. It should be noted that Serbin et al. (2012) and Ainsworth et al. (2014) estimated the photosynthesis traits at the actual leaf temperature at the time of measurement and did not normalise them to a common reference temperature as done in this study. However, Serbin et al. (2015) obtained very good results for the estimation of temperature-normalised canopy-level $V_{cmax,30}$ for nine crop species ($R^2 = 0.94$ based on repeated 75%/25% cross-validation). It seems, therefore, that the good performance of the leaf-level estimation by Serbin et al. (2012) and Ainsworth et al. (2014) was not due to the estimation of V_{cmax} and J_{max} at the actual temperature instead of temperature-normalised values. This is to be expected also from theoretical considerations, as the temperature dependence of V_{cmax} is not due to a change in Rubisco content but related to the temperature dependence of chemical reactions. The slightly superior predictive performance of $J_{max,25}$ compared to $V_{cmax,25}$ models was consistently observed in all studies that estimated both traits (Doughty et al., 2011; Serbin et al., 2012), including our study.

The prediction performances for *LMA* and nitrogen content per mass in the studies of Doughty et al. (2011) and Serbin et al. (2012) were similar to our results for *LMA* and nitrogen content per area, N_a. Our results showing clearly superior predictive performance of the SWIR models compared to the VNIR models (Table 2) are consistent with the findings of Asner et al. (2011) and further confirm that the strongest absorption signals related to these traits are located in the SWIR part of the spectrum. This is also evident in the shape of specific absorption coefficients of total dry matter content and protein content of the PROSPECT model (Jacquemoud et al., 1996).

4.2. PROSPECT inversion to estimate chlorophyll, carotenoid, water and dry matter contents

The PROSPECT model was validated in inverse mode on independent datasets with very good results for chlorophyll and water contents and promising results for carotenoid and dry matter contents (Féret

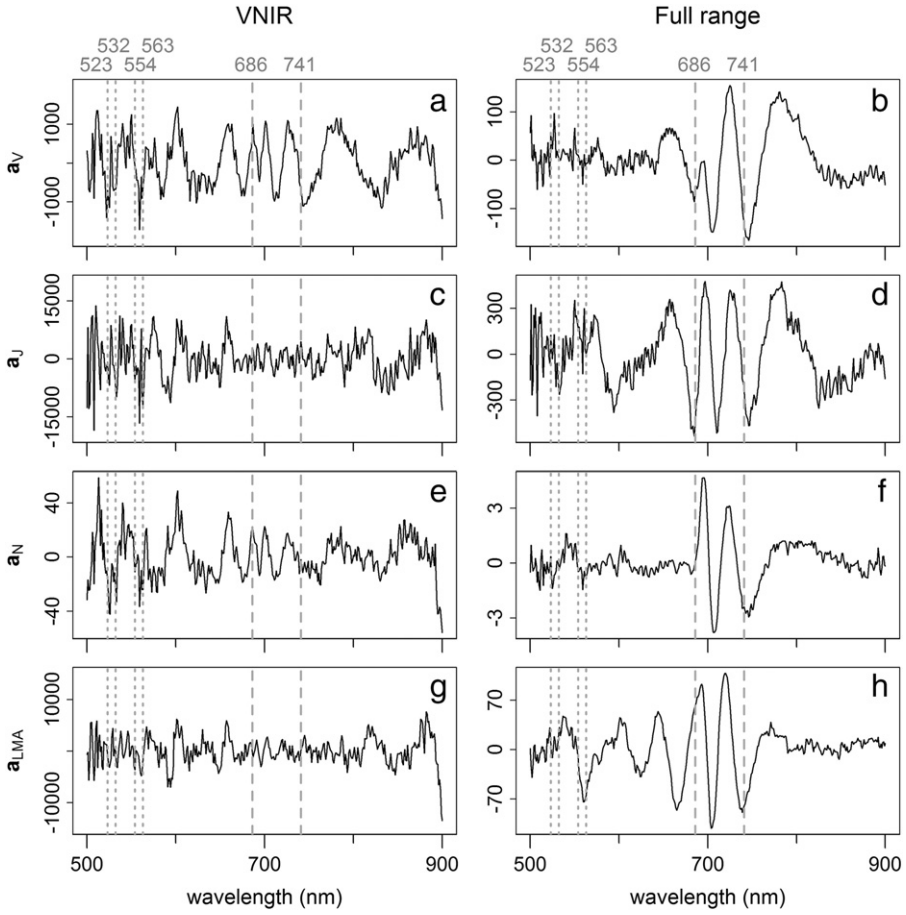

Fig. 6. PLS coefficients spectra for the VNIR model (left column) and the VNIR part of the full range model (right column). The coefficients for (a, b) maximum carboxylation capacity, $V_{cmax,25}$, (c, d) the maximum electron transport rate, $J_{max,25}$, (e, f) nitrogen content per area, N_a, and (g, h) leaf mass per area, *LMA*. The standard errors of the mean of the regression coefficients were two orders of magnitude smaller than the maximum coefficient for each trait. The dashed grey vertical lines indicate the peak positions of chlorophyll fluorescence emission (unpublished results) and the dotted grey vertical lines indicate cytochrome b_6f absorption features (Nelson and Neumann, 1972). The exact values of the wavelengths in nanometres are given at the top of the figure in grey.

et al., 2008, 2011; Li and Wang, 2011; Shiklomanov et al., 2016). Given that our results showed that PROSPECT could not only fit the measured spectra very well, but also could estimate the dry matter content with high precision, we have high confidence in the inversion results for the other parameters, in particular chlorophyll and water contents. Chlorophyll and water contents can be estimated precisely even with simple reflectance ratios or normalised difference indices due to their strong absorption features (e.g. Ceccato et al., 2001; Datt, 1999; Gitelson et al., 2003; Kira et al., 2015). We found high correlations (0.9) between PROSPECT inversion-based and vegetation index-based estimates of chlorophyll, carotenoid and water contents. Although we cannot exclude a global bias of our chlorophyll, carotenoid and water content estimates, as was indeed observed for *LMA*, this does not matter for our analysis as it is invariant to offsets and global scaling.

Table 3
Pearson correlation (ρ) between PLS coefficient spectra. For each pair of leaf traits, the Pearson correlation between their PLS coefficient spectra is shown for the full range model (500–2350 nm), the VNIR (500–900 nm) and SWIR (900–2350 nm) models. The values shown are the median values of the correlation between the coefficients of the 200 repetitions. The standard errors of the median were smaller or equal to 0.01 in all cases. The highest correlations for each trait combination are printed in bold and the highest correlation for each spectral range is underlined.

Traits	ρ		
	Full	VNIR	SWIR
$V_{cmax,25}$, N_a	0.66	0.78	**0.80**
$J_{max,25}$, N_a	0.57	0.44	**0.76**
$V_{cmax,25}$, *LMA*	0.23	0.24	**0.37**
$J_{max,25}$, *LMA*	0.27	**0.57**	0.41
$V_{cmax,25}$, $J_{max,25}$	0.71	0.48	**0.79**
LMA, N_a	0.35	0.36	**0.41**

4.3. Mechanisms of reflectance-based photosynthesis trait estimation

4.3.1. $V_{cmax,25}$

The results of our analysis indicate that the prediction capacity of reflectance-based $V_{cmax,25}$ PLS models was entirely based on the relationship to N_a. The best reflectance-based PLS model for $V_{cmax,25}$ ($\mathcal{R} \rightarrow V_{cmax,25}$) was narrowly outperformed (Table 2) by a model that estimated N_a from reflectance and used the linear relationship between nitrogen and $V_{cmax,25}$ ($\mathcal{R} \rightarrow N_a^{PLS} \rightarrow V_{cmax,25}$). In addition, the analysis of relationships between regression coefficient spectra of reflectance-based $V_{cmax,25}$ and N_a PLS models showed consistently high correlation across both the VNIR and SWIR part of the spectrum (Table 3). The good estimation performance of a single PLS model based on N_a for 37 different temperate deciduous tree species in our study is in accordance with the results of Kattge et al. (2009), who also found a very robust N_a-$V_{cmax,25}$ relationship for a large number of species of temperate broadleaf trees.

4.3.2. $J_{max,25}$

The correlation to N_a explained over 90% of the prediction performance for the full spectral range $J_{max,25}$ models and therefore was the dominating mechanism. The dominance of the N_a contribution to the

estimation could be seen most clearly for the SWIR models, where the small difference in prediction performance between the $\mathcal{R} \rightarrow J_{max,25}$ and the $\mathcal{R} \rightarrow N_a^{PLS} \rightarrow J_{max,25}$ models could be explained by adding LMA^{PLS} as predictor (Table 2). The correlations between PLS regression coefficients (Table 3) also showed strong similarities between N_a and $J_{max,25}$ coefficients for the SWIR but not for the VNIR model. The above results lead to the conclusion that although the relationship to N_a is the dominating mechanism, there is additional information in the VNIR part of the spectrum, contributing to the PLS models. This contribution, however, is very small given that the prediction performances of the SWIR and full range $\mathcal{R} \rightarrow J_{max,25}$ models were very similar ($R^2 = 0.67$ vs. $R^2 = 0.70$, Table 2). This difference could not be explained by adding the chlorophyll, carotenoid and water contents to the reflectance-based nitrogen model (Table 2). Phosphorous content as a candidate to account for the slightly superior performance of the full-range model is highly unlikely, as its absorption signal is mainly located in the SWIR region (Asner et al., 2011; Workman and Weyer, 2007). Two main potential sources of information located in the VNIR region have been mentioned in the literature: chlorophyll fluorescence emission (Serbin et al., 2015; Verrelst et al., 2015), which is part of the apparent reflectance measured in this study (Campbell et al., 2008; Kim et al., 1993; Zarco-Tejada et al., 2000), and cytochrome *f* absorption (Nelson and Neumann, 1972; Onoda, 2005). The regression coefficients of the VNIR PLS model for $J_{max,25}$ (Fig. 6) showed strong features at or close to the cytochrome b_6f absorptions, which were also present in the $V_{cmax,25}$ and N_a PLS regression coefficients. The VNIR part of the corresponding full-range PLS models, however, did not clearly show these absorption features. We could exclude contributions of cytochrome *f* absorption by comparing the VNIR PLS models with models in the restricted range of 600–900 nm. However, we could not investigate a potential contribution of the strongest signal of cytochrome *f* at 421 nm (Nelson and Neumann, 1972), because of the low signal to noise ratio in this part of the reflectance spectra. Although the regression coefficients of the VNIR PLS model for $J_{max,25}$ do not show strong features at the wavelengths most influenced by the chlorophyll fluorescence signal, the VNIR part of the corresponding full-range PLS models appears to indicate a more important role of the red peak of chlorophyll fluorescence at about 690 nm (Fig. 6). This is in agreement with model-based findings of Verrelst et al. (2015) that the red fluorescence peak is more directly related to $V_{cmax,25}$ than the far-red peak (at about 740 nm). In order to further confirm the interpretation of a potential chlorophyll fluorescence contribution, direct measurements of the latter would be required.

One reason for the apparently different behaviour in these two cases (VNIR models and VNIR part of full range models) is that only information complementary to the N_a-based relationship will be used in the full-range model, while the VNIR model does not obtain as much N_a information and might therefore have a different balance of contributions of *LMA* and N_a signals (Table 3). Furthermore, a correction for interfering signals and other interaction effects between SWIR and VNIR could also explain the apparently very different behaviour of the regression coefficients for the full range and the VNIR models.

4.3.3. Potential contributions of PROSPECT estimated leaf traits and influence of different light environments

We did not find evidence for contributions of chlorophyll content or the other PROSPECT estimated leaf traits to the estimation of $V_{cmax,25}$ and $J_{max,25}$ that improved upon the best N_a-based model. The differences in the performances of the VNIR $\mathcal{R} \rightarrow Ph$ and $\mathcal{R} \rightarrow N_a^{PLS} \rightarrow Ph$ models could be explained by small contributions of *LMA* together with chlorophyll and carotenoid content as well as the PROSPECT leaf structure parameter (results not shown). This, however, only reflected the decrease in performance of the N_a estimation in the VNIR (Table 2). For the SWIR models and the full range models, only very small improvements were observed when adding *LMA* and only a marginal improvement was observed when the PROSPECT parameters were added to the reflectance-based N_a estimates for $V_{cmax,25}$ (Table 2). The latter results also exclude a significant contribution of leaf water content to the prediction despite the relatively strong correlation between nitrogen content and leaf water content (Fig. S2; Wang et al., 2015). Furthermore, a contribution from chlorophyll and carotenoids leading to a better model can be strictly excluded for $V_{cmax,25}$ as the best model was based on the estimated nitrogen content in the SWIR (900–2350 nm) where chlorophyll does not absorb at all (e.g. Brown, 1983; Féret et al., 2008; Heldt and Piechulla, 2011; Jacquemoud and Baret, 1990; Taiz and Zeiger, 2010). The best model for $J_{max,25}$ used the full range reflectance, but the improvement over the SWIR and nitrogen based models was marginal.

High correlations between leaf nitrogen and chlorophyll contents have been reported in the literature for certain species and datasets (e.g. Homolová et al., 2013; Houborg et al., 2015; Schlemmer et al., 2013). However, no consistent relationship across many species and large datasets equivalent to that observed in Kattge et al. (2009) for nitrogen content was reported so far. This can be understood by considering the results of several studies (Evans and Poorter, 2001; Niinemets et al., 1998) that found that the chlorophyll content and the fraction of nitrogen invested in the pigment-protein pool strongly react to changes in light environment, while the fraction of nitrogen invested in Rubisco and electron transport is relatively insensitive to these changes. This implies that leaves growing in different light environments can have very different chlorophyll-$V_{cmax,25}$ relationships, which we also observed in our data (Fig. 7). While the relationships between N_a and photosynthesis parameters were quite stable for different light environments, the relationships between chlorophyll content and photosynthesis parameters differed strongly for sun and shade leaves. Even for sun and shade leaves separately, the relationships to photosynthesis parameters were stronger for nitrogen than for chlorophyll. *LMA* showed more consistent relationships to photosynthesis parameters for different light environments than chlorophyll (Fig. S5). Our results were based on estimated rather than measured chlorophyll content.

However, our estimation procedure showed reliable results (sections 3.2.2 and 4.2) and the effect shown in Fig. 7 was also observed in studies that determined chlorophyll content chemically (e.g. Seemann et al., 1987). The corresponding large differences in the chlorophyll-nitrogen content ratios between sun and shade leaves were observed in numerous studies (e.g. Gratani et al., 2006; Poorter et al., 2000; Rozendaal et al., 2006).

For the interpretation of Fig. 7, it is important to note that changes in the chemical composition can be seen in the traits per mass, while differences in the traits per area can be caused purely by changed leaf morphology (*LMA*). We observed that the chlorophyll content per mass was significantly higher for shade leaves than for sun leaves (results not shown), while nitrogen content per mass was similar. This corresponds to the expected higher resource investment in light capture in shade leaves and is in agreement with other studies on the plasticity of plant traits in response to different light environments (e.g. Chazdon et al., 1996; Niinemets, 2010; Poorter et al., 2000).

Croft et al. (2016) found a better correlation of chlorophyll content to V_{cmax} and J_{max} than between N_a and photosynthesis traits over the season for four deciduous tree species. However, we think that it is important to distinguish between correlation of traits over the season (temporal correlation) and across species correlation. Temporal correlation was also observed by Xu and Baldocchi (2003) to be lower between nitrogen per area and V_{cmax} than for nitrogen per mass but the relationship for N_a could be considerably improved by including *LMA* as predictor. Across species correlation between N_a and V_{cmax} was observed to be strong for very large numbers of species (Kattge et al., 2009; Walker et al., 2014). We think it is this property that is the most relevant for large scale application of remote sensing based estimation of photosynthesis traits as it translates into spatial correlation.

4.4. Regression coefficients and signal of Rubisco in leaf reflectance spectra

Although water and other components of the dry matter content such as lignin are interfering with the nitrogen signal in plant leaf

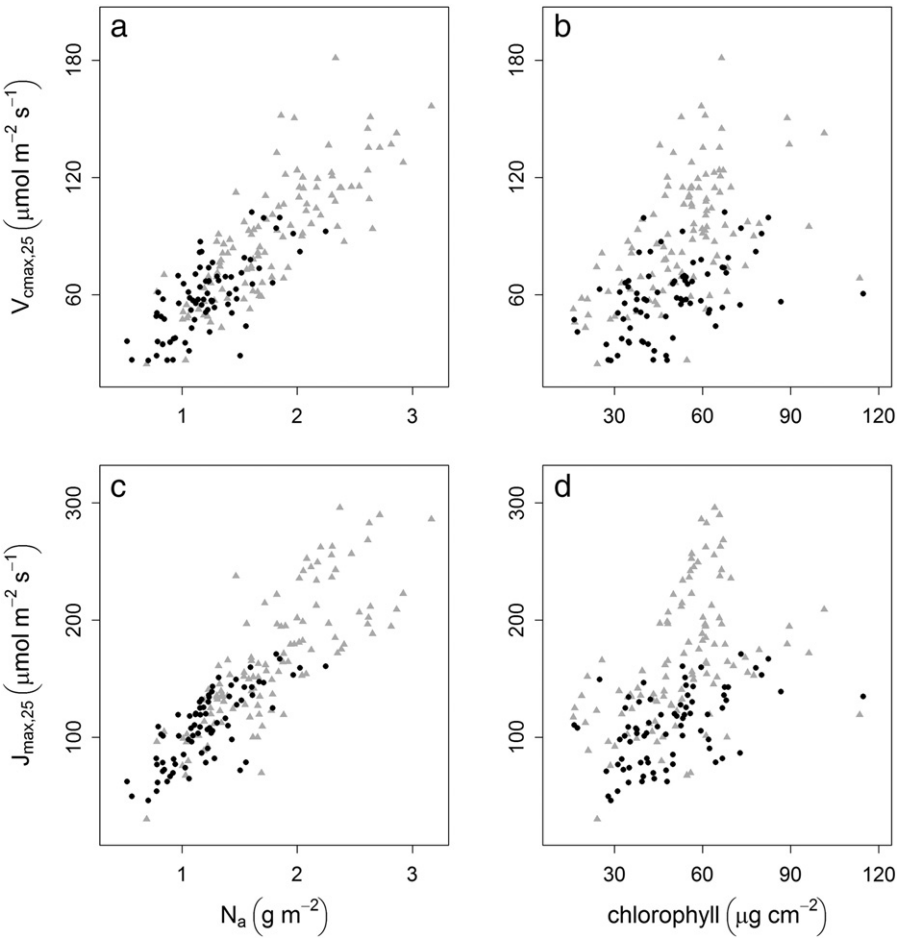

Fig. 7. Relationships between observed nitrogen content, N_a, and photosynthesis parameters $V_{cmax,25}$ and $J_{max,25}$, (a, c) as well as chlorophyll content estimated via PROSPECT inversion and photosynthesis parameters (b, d) for sun (grey triangles) and shade leaves (black circles). The classification in sun and shade leaves was only available for 185 out of the total 242 data points. As leaf light environment was recorded by a human observer, misclassification of some of the points cannot be excluded.

reflectance, some of the wavelength positions reported in the literature as related to nitrogen and/or protein absorption appear as distinct features in the regression coefficients, shown in Fig. 5. No clear feature appeared in the coefficient spectra at the 1940 nm nitrogen-related absorption (Curran, 1989), which is probably linked to the strong water absorption in this region. The overall shape of the specific absorption coefficient for protein in dry leaves estimated with the PROSPECT model (Jacquemoud et al., 1996) and the pseudo-absorption spectrum [log(R^{-1})] of Rubisco (Elvidge, 1990) is very similar (Fig. 5a, b). Small differences in the positions of the local maxima are probably due to either imperfect estimation with the PROSPECT model or differences between pseudo-absorbance and actual absorbance. The latter difference is responsible for the incorrect increase at smaller wavelengths (<1125 nm). The dry leaf specific absorption coefficients were chosen instead of the corresponding fresh leaf coefficients, as the latter showed strong masking effects of water absorption on the estimation of the protein absorption coefficients (Jacquemoud et al., 1996). The correlation between total protein content and N_a was extremely high in the dataset used for the estimation of the protein specific absorption coefficient (Jacquemoud et al., 1996), which implies that the specific absorption coefficient spectrum for N_a has the same shape as that for total protein content. The strong similarity between the nitrogen/protein absorption and the Rubisco absorption can be attributed to the common origin of the absorption features in the spectral range 500–2500 nm related to nitrogen/protein: mostly overtones and combination bands of N—H bond vibrational excitations (Siesler, 2007; Kokaly, 2001). This explains why the reflectance-based PLS models for $V_{cmax,25}$ use the correlation to N_a, as only the N—H signal can be extracted from the reflectance spectrum and not something more directly related to Rubisco content, such as the nitrogen content in Rubisco.

To our knowledge, there is only one leaf-level study related to V_{cmax} or J_{max} estimation that provided the full regression coefficient spectrum (Ainsworth et al., 2014). They obtained very different spectral patterns compared to our results. This raises the question of very different correlation patterns between the chlorophyll, water and dry matter contents and $V_{cmax,25}$ compared to our study, but part of the differences could potentially be explained by the temperature effects included in V_{cmax} in the study of Ainsworth et al. (2014). The latter could potentially also explain the stronger relationship to leaf water content. Both the absence of shade leaves in the study of Ainsworth et al. (2014) and the focus on a crop species, on the other hand, could be an explanation for a higher correlation between chlorophyll and nitrogen contents than we observed in our data (Homolová et al., 2013; Houborg et al., 2015; Schlemmer et al., 2013). Most wavelengths selected for the V_{cmax}, J_{max} and N_{mass} models in the study of Serbin et al. (2012) were located in the SWIR with wavelengths larger than 1500 nm and some in the VNIR with wavelengths smaller than 800 nm. However, it is not clear whether this is consistent with our findings, as the shape of the regression coefficient spectra cannot be reconstructed using this information. The results from the canopy scale study of Serbin et al. (2015) are difficult to compare with our results because of the influence of leaf area index (LAI), soil background and the absence of the water absorption bands. They do not show clear features at similar wavelength positions as observed in this study.

The location of the two strong and broad negative features in the N_a, $V_{cmax,25}$ and $J_{max,25}$ PLS regression coefficients (Fig. 5) at around

2060 and 2180 nm coincide with the 2054 and 2172 nm nitrogen-related features found by Kokaly (2001) for dry leaves.

4.5. Influence of chemometric method and validation procedure

In order to verify the reliability of the repeated-double cross-validation (rdCV) procedure, we compared the results with leave-one-out cross-validation and obtained consistent results. Ideally, our models should be tested on independent datasets to determine if the patterns found in our measurements also hold for different climatic conditions, different soils and other species. Pre-processing methods such as pseudo-absorption, the logarithm of the inverse of reflectance, and first and second derivatives did not show clear improvements in the PLS results for our data. We also verified that our main conclusions were not specific to the PLS method by conducting similar analyses with Random Forests (Breiman, 2001) and Gaussian Process regression (Rasmussen and Williams, 2006) with a squared-exponential kernel. We can, therefore, exclude that our analysis was limited by the linearity of the PLS method or other characteristics of PLS.

4.6. Implications for large scale estimation of $V_{cmax,25}$ and $J_{max,25}$

We believe that our findings have important implications for large scale estimation of $V_{cmax,25}$ and $J_{\max,25}$ from canopy reflectance measurements. First, the indirectness of the estimation should translate into a markedly lower prediction performance of canopy scale multivariate regression models for $V_{cmax,25}$ and $J_{\max,25}$ compared to N_a. The results of Serbin et al. (2015) for different crop types do not appear to show this effect, but we expect to see a clear decrease in prediction performance for natural vegetation and a large number of species. Secondly, our results suggest that a two-step estimation process might be advantageous: estimation of nitrogen content from canopy reflectance observations in a first step and estimation of the photosynthesis parameters based on the relationships to nitrogen content in a second step. This approach would be both more transparent in terms of data processing and would help to use different nitrogen-photosynthesis parameter relationships for different ecosystems or soil conditions. The necessity for the latter was shown in a study by Kattge et al. (2009), as the N_a — $V_{cmax,25}$ relationships had clearly different slopes for tropical trees growing on oxisols, C_3 herbaceous species, C_3 crop species and coniferous trees compared to temperate broadleaved trees. Hence, we expect that a direct estimation of $V_{cmax,25}$ and $J_{\max,25}$ from canopy reflectance data using multivariate regression or machine learning methods on large scales would also require different models depending on the ecosystem or soil conditions. The underlying reasons for this, however, would probably remain obscure because of the difficulties of model interpretation and lack of mechanistic understanding. Lastly, we do not recommend the large scale application of the equivalent approach to the $R \rightarrow T^{PRO} \rightarrow Ph$ estimation on the canopy scale, where PROSPECT inversion would be replaced by PROSAIL inversion. We expect worse results on the canopy scale than on the leaf scale as the considerable uncertainties of PROSAIL inversion would add to the relatively weak relationships between PROSPECT traits and photosynthesis traits. Applications on agricultural fields, however, may find the PROSAIL-inversion-based approach useful (Houborg et al., 2013) as the correlation between leaf chlorophyll and nitrogen content was observed to be high in some studies (e.g. Homolová et al., 2013; Schlemmer et al., 2013; Houborg et al., 2015). Obviously, the nitrogen-based approach suggested above requires above all a reliable estimation of canopy nitrogen content on a large scale. Several studies have demonstrated the feasibility of estimating canopy nitrogen content using hyperspectral observations from both airborne (Asner et al., 2011; Singh et al., 2015; Wang et al., 2016) and even spaceborne platforms (Martin et al., 2008). Once upcoming hyperspectral satellite missions such as EnMAP (Stuffler et al., 2007) provide reliable data, canopy nitrogen content could be estimated globally and could be the basis for global estimations of $V_{cmax,25}$, which might then be used to improve the parametrisation of process-based carbon cycle models. Even if this approach did not result in a good seasonal estimation of $V_{cmax,25}$, or otherwise not enough cloud-free satellite observations were available to capture the seasonal dynamics, the estimated nitrogen content at some point(s) in the season could still be used to constrain the modelling of the seasonal development of $V_{cmax,25}$ and $J_{\max,25}$ by using an optimal allocation approach such as the one presented in Quebbeman and Ramirez (2016).

5. Conclusions

We showed that the photosynthesis traits $V_{cmax,25}$ and $J_{\max,25}$ can be estimated reasonably well from leaf reflectance with a single multivariate regression model for a large number of deciduous tree species and leaves in widely varying light environments. We further demonstrated that the reflectance-based PLS models for $V_{cmax,25}$ and $J_{\max,25}$ were clearly based on the correlation to nitrogen content per area, N_a. The separate estimation of N_a from reflectance combined with a known linear $V_{cmax,25}$-N_a relationship yielded slightly better results than direct estimates of $V_{cmax,25}$ from reflectance spectra. Furthermore, we confirmed that the N_a-related information is mainly located in the short-wave-infrared (SWIR) part of the reflectance spectrum and therefore the best PLS models for N_a are those restricted to the SWIR. The $J_{\max,25}$ models were also almost entirely based on the correlation to N_a. We found that *LMA* had a small contribution to the $J_{\max,25}$ models and that additional information is located in the VNIR part of the spectrum, which is most probably related to the contribution of chlorophyll fluorescence. We also found that correlations to other main absorbers such as chlorophyll, carotenoids and water did not contribute useful information complementary to N_a to the estimation of $V_{cmax,25}$ and only marginally to $J_{\max,25}$, if at all. This was also reflected in the low accuracy and precision of the models based on leaf traits estimated via PROSPECT inversion.

In summary, we did not find indications that any information more directly related to $V_{cmax,25}$ and $J_{\max,25}$ than the total nitrogen content per leaf area can be estimated from leaf reflectance spectra. This has important implications for large scale estimation of these photosynthesis parameters. On the one hand, it clearly indicates limitations with regard to achievable accuracy and precision of the estimates (with relative errors on the order of 20%), on the other hand it suggests the need to focus on the improvement of reflectance-based nitrogen content estimation on the canopy scale combined with a better understanding of the relationship between photosynthesis traits and nitrogen content. This approach would not only reflect the actual mechanism of reflectance-based estimation of $V_{cmax,25}$ and $J_{\max,25}$ but could also facilitate ground validation and could help to use different relationships between nitrogen content and photosynthesis parameters for different ecosystems and soil conditions.

Acknowledgements

We thank the following interns and technical assistants for contributing to the sample preparation for chemical analysis: L. Dienstbach, I. Quirós, F. Kirsten, B. Haller, M. Rossi, B. Galinsky, and T. Franz. We thank G. Schuhmann for coordinating the sample processing and preparation for chemical analysis and G. Henning for conducting the chemical analysis. We also thank G. Schulz, S. Lehmann, H. Zöphel, S. Gimper, and J. Welles for technical advice and support. We are grateful to A. Jung from the University of Leipzig, K. Grosser from iDiv as well as H. Borsdorf and C. Rebmann from UFZ for providing measurement equipment. We thank M. Lange for support with IT issues, A. Schmidt for help with Gaussian Process regression and A. Piayda who helped with the analysis of leaf scan images. We also thank S. Jacquemoud for providing the protein specific absorption coefficient data. Finally, we are indebted to C. Wirth, B. Reu, and B. Ohse for making the measurement campaign in the Arboretum of the University of Leipzig possible. We also thank three anonymous reviewers who helped to improve the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2017.05.019.

References

Ainsworth, E.A., Serbin, S.P., Skoneczka, J.A., Townsend, P.A., 2014. Using leaf optical properties to detect ozone effects on foliar biochemistry. Photosynth. Res. 119:65–76. http://dx.doi.org/10.1007/s11120-013-9837-y.

Ardia, D., Mullen, K.M., Peterson, B.G., Ulrich, J., 2015. 'DEoptim': differential evolution in 'R'. R Package Version 2.2-2 http://CRAN.R-project.org/package=DEoptim.

Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C., Carranza, L., ... Weiss, P., 2011. Spectroscopy of canopy chemicals in humid tropical forests. Remote Sens. Environ. 115:3587–3598. http://dx.doi.org/10.1016/j.rse.2011.08.020.

Asner, G.P., Martin, R.E., Carranza-Jiménez, L., Sinca, F., Tupayachi, R., Anderson, C.B., Martinez, P., 2014. Functional and biological diversity of foliar spectra in tree canopies throughout the Andes to Amazon region. New Phytol. 204:127–139. http://dx.doi.org/10.1111/nph.12895.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., ... Papale, D., 2010. Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. Science 329:834–838. http://dx.doi.org/10.1126/science.1184984.

Bonan, G.B., Lawrence, P.J., Oleson, K.W., Levis, S., Jung, M., Reichstein, M., ... Swenson, S.C., 2011. Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. J. Geophys. Res. 116. http://dx.doi.org/10.1029/2010JG001593.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brown, J.S., 1983. A new evaluation of chlorophyll absorption in photosynthetic membranes. Photosynth. Res. 4, 375–383.

Buschmann, C., 2007. Variability and application of the chlorophyll fluorescence emission ratio red/far-red of leaves. Photosynth. Res. 92:261–271. http://dx.doi.org/10.1007/s11120-007-9187-8.

Campbell, P.K.E., Middleton, E.M., Corp, L.A., Kim, M.S., 2008. Contribution of chlorophyll fluorescence to the apparent vegetation reflectance. Sci. Total Environ. 404:433–439. http://dx.doi.org/10.1016/j.scitotenv.2007.11.004.

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.

Chazdon, R.L., Pearcy, R.W., Lee, D.W., Fetcher, N., 1996. Photosynthetic Responses of Tropical Forest Plants to Contrasting Light Environments. Springer US.

Croft, H., Chen, J.M., Luo, X., Bartlett, P., Chen, B., Staebler, R.M., 2016. Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. Glob. Chang. Biol. http://dx.doi.org/10.1111/gcb.13599.

Curran, P., 1989. Remote sensing of foliar chemistry. Remote Sens. Environ. 30, 271–278.

Datt, B., 1999. Remote sensing of water content in Eucalyptus leaves. Australian Journal of Botany]–>Aust. J. Bot. 47, 909–923.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27–46. http://dx.doi.org/10.1111/j.1600-0587.2012.07348.x.

Doughty, C.E., Asner, G.P., Martin, R.E., 2011. Predicting tropical plant physiology from leaf and canopy spectroscopy. Oecologia 165:289–299. http://dx.doi.org/10.1007/s00442-010-1800-4.

Elser, J.J., Bracken, M.E.S., Cleland, E.E., Gruner, D.S., Harpole, W.S., Hillebrand, H., ... Smith, J.E., 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. Ecol. Lett. 10:1135–1142. http://dx.doi.org/10.1111/j.1461-0248.2007.01113.x.

Elvidge, C.D., 1990. Visible and near infrared reflectance characteristics of dry plant materials. Int. J. Remote Sens. 11:1775–1795. http://dx.doi.org/10.1080/01431169008955129.

Evans, J., Poorter, H., 2001. Photosynthetic acclimation of plants to growth irradiance: the relative importance of specific leaf area and nitrogen partitioning in maximizing carbon gain. Plant Cell Environ. 24, 755–767.

Farquhar, G.D., von Caemmerer, S., Berry, J.A., 1980. A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. Planta 149, 78–90.

Farrar, D.E., Glauber, R.R., 1967. Multicollinearity in regression analysis: the problem revisited. Rev. Econ. Stat. 92–107.

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

Féret, J.-B., François, C., Gitelson, A., Asner, G.P., Barry, K.M., Panigada, C., ... Jacquemoud, S., 2011. Optimizing spectral indices and chemometric analysis of leaf chemical properties using radiative transfer modeling. Remote Sens. Environ. 115:2742–2750. http://dx.doi.org/10.1016/j.rse.2011.06.016.

Filzmoser, P., Varmuza, K., 2012. Multivariate statistical analysis in chemometrics. R Package Version 1.3.8 http://CRAN.R-project.org/package=chemometrics.

Filzmoser, P., Liebmann, B., Varmuza, K., 2009. Repeated double cross validation. J. Chemom. 23:160–171. http://dx.doi.org/10.1002/cem.1225.

Fox, J., & Weisberg, S. (2011). An {R} Companion to Applied Regression, Second Edition. Thousand Oaks CA: Sage. Retrieved from https://cran.r-project.org/web/packages/car/car.pdf. (URL: Http://Socserv.socsci.mcmaster.ca/Jfox/Books/Companion)

Friedman, J., Hastie, T., Tibshirani, R., 2001. The Elements of Statistical Learning. Springer series in statistics Vol. 1. Springer, Berlin Retrieved from. http://statweb.stanford.edu/~tibs/book/preface.ps.

Gitelson, A.A., Gritz, Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. J. Plant Physiol. 160:271–282. http://dx.doi.org/10.1078/0176-1617-00887.

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. http://dx.doi.org/10.1029/2006GL026457.

Gratani, L., Covone, F., Larcher, W., 2006. Leaf plasticity in response to light of three evergreen species of the Mediterranean maquis. Trees 20:549–558. http://dx.doi.org/10.1007/s00468-006-0070-6.

Gunst, R.F., 1983. Regression analysis with multicollinear predictor variables: definition, direction, and effects. Communications in Statistics - Theory and Methods 12:2217–2260. http://dx.doi.org/10.1080/03610928308828603.

Heldt, H.-W., Piechulla, B., 2011. Plant Biochemistry. fourth ed. Academic Press.

Hendrickx, J., 2012. perturb: tools for evaluating collinearity. R Package Version 2.05 http://CRAN.R-project.org/package=perturb.

Homolová, L., Malenovský, Z., Clevers, J.G.P.W., García-Santos, G., Schaepman, M.E., 2013. Review of optical-based remote sensing for plant trait mapping. Ecol. Complex. 15:1–16. http://dx.doi.org/10.1016/j.ecocom.2013.06.003.

Hosgood, B., Jacquemoud, S., Andreoli, G., Verdebout, J., Pedrini, A., Schmuck, G., 1995. Leaf Optical Properties Experiment 93 (LOPEX93) (No. EUR-16095-EN). Institute for Remote Sensing Applications, Ispra, Italy: European Commission, Joint Research Centre Retrieved from. http://bookshop.europa.eu/en/leaf-optical-properties-experiment-93-lopex93–pbCLNA16095/downloads/CL-NA-16-095-EN-C/CLNA16095ENC_001.pdf;pgid=y8dIS7GUWMdSR0EAlMEUUsWb000080-ExZe4;sid=MxztItc_PSft94d09OVzMbUakd-RxnX8T0w=?FileName=CLNA16095ENC_001.pdf&SKU=CLNA16095ENC_PDF&CatalogueNumber=CL-NA-16-095-EN-C.

Houborg, R., Cescatti, A., Migliavacca, M., Kustas, W.P., 2013. Satellite retrievals of leaf chlorophyll and photosynthetic capacity for improved modeling of GPP. Agric. For. Meteorol. 177:10–23. http://dx.doi.org/10.1016/j.agrformet.2013.04.006.

Houborg, R., McCabe, M.F., Cescatti, A., Gitelson, A., A., 2015. Leaf chlorophyll constraint on model simulated gross primary productivity in agricultural systems. Int. J. Appl. Earth Obs. Geoinf. 43:160–176. http://dx.doi.org/10.1016/j.jag.2015.03.016.

Jacob, J., Greitner, C., Drake, B.G., 1995. Acclimation of photosynthesis in relation to Rubisco and non-structural carbohydrate contents and in situ carboxylase activity in *Scirpus olneyi* grown at elevated CO2 in the field. Plant Cell Environ. 18, 875–884.

Jacquemoud, S., Baret, F., 1990. PROSPECT: a model of leaf optical properties spectra. Remote Sens. Environ. 34, 75–91.

Jacquemoud, S., Ustin, S.L., Verdebout, J., Schmuck, G., Andreoli, G., Hosgood, B., 1996. Estimating leaf biochemistry using the PROSPECT leaf optical properties model. Remote Sens. Environ. 56, 194–202.

Jung, M., Reichstein, M., Margolis, H.A., Cescatti, A., Richardson, A.D., Arain, M.A., ... Williams, C., 2011. Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. J. Geophys. Res. 116. http://dx.doi.org/10.1029/2010JG001566.

Kattge, J., Knorr, W., Raddatz, T., Wirth, C., 2009. Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. Glob. Chang. Biol. 15:976–991. http://dx.doi.org/10.1111/j.1365-2486.2008.01744.x.

Keenan, T.F., Niinemets, Ü., 2016. Global leaf trait estimates biased due to plasticity in the shade. Nature Plants 3:16201. http://dx.doi.org/10.1038/nplants.2016.201.

Kim, M.S., Chappelle, E.W., Corp, L., McMurtrey III, J.E., 1993. The contribution of chlorophyll fluorescence to the reflectance spectra of green vegetation. Geoscience and Remote Sensing Symposium, 1993. IGARSS'93. Better Understanding of Earth Environment., International. IEEE :pp. 1321–1324 Retrieved from. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=322673.

Kira, O., Linker, R., Gitelson, A., 2015. Non-destructive estimation of foliar chlorophyll and carotenoid contents: focus on informative spectral bands. Int. J. Appl. Earth Obs. Geoinf. 38:251–260. http://dx.doi.org/10.1016/j.jag.2015.01.003.

Knorr, W., 2000. Annual and interannual CO2 exchanges of the terrestrial biosphere: process-based simulations and uncertainties. Glob. Ecol. Biogeogr. 9, 225–252.

Kokaly, R.F., 2001. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. Remote Sens. Environ. 75, 153–161.

Lehnert, L.W., Meyer, H., Bendix, J., 2015. hsdar: manage, analyse and simulate hyperspectral data in R. R package version 0.3.1. http://CRAN.R-project.org/package=hsdar.

Leuning, R., 1997. Scaling to a common temperature improves the correlation between the photosynthesis parameters Jmax and Vcmax. J. Exp. Bot. 48, 345–347.

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. http://dx.doi.org/10.1109/TGRS.2011.2109390.

Maitra, S., Yan, J., 2008. Principle component analysis and partial least squares: two dimension reduction techniques for regression. Applying multivariate statistical models. 79 Retrieved from. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.473.4340&rep=rep1&type=pdf#page=81.

Makino, A., Nakano, H., Mae, T., 1994. Responses of ribulose-1, 5-bisphosphate carboxylase, cytochrome f, and sucrose synthesis enzymes in rice leaves to leaf nitrogen and their relationships to photosynthesis. Plant Physiol. 105, 173–179.

Martin, M.E., Plourde, L.C., Ollinger, S.V., Smith, M.-L., McNeil, B.E., 2008. A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. Remote Sens. Environ. 112:3511–3519. http://dx.doi.org/10.1016/j.rse.2008.04.008.

Mevik, B.-H., Wehrens, R., Liland, K.H., 2013. pls: partial least squares and principal component regression. R Package Version 2.4-3 http://CRAN.R-project.org/package=pls.

Mullen, K.M., Ardia, D., Gil, D.L., Windover, D., Cline, J., et al., 2011. DEoptim: an R package for global optimization by differential evolution. J. Stat. Softw. 40, 1–26.

Nelson, N., Neumann, J., 1972. Isolation of a cytochrome b6-f particle from chloroplasts. J. Biol. Chem. 247, 1817–1824.

Niinemets, Ü., 2010. A review of light interception in plant stands from leaf to canopy in different plant functional types and in species with varying shade tolerance. Ecol. Res. 25:693–714. http://dx.doi.org/10.1007/s11284-010-0712-4.

Niinemets, Ü., Kull, O., Tenhunen, J.D., 1998. An analysis of light effects on foliar morphology, physiology, and light interception in temperate deciduous woody species of contrasting shade tolerance. Tree Physiol. 18, 681–696.

Onoda, Y., 2005. Seasonal change in the balance between capacities of RuBP carboxylation and RuBP regeneration affects CO2 response of photosynthesis in Polygonum cuspidatum. J. Exp. Bot. 56:755–763. http://dx.doi.org/10.1093/jxb/eri052.

Papale, D., Black, T.A., Carvalhais, N., Cescatti, A., Chen, J., Jung, M., ... Ráduly, B., 2015. Effect of spatial sampling from European flux towers for estimating carbon and water fluxes with artificial neural networks: sampling effect on fluxes upscaling. J. Geophys. Res. Biogeosci.:n/a http://dx.doi.org/10.1002/2015JG002997.

Poorter, L., Kwant, R., Hernandez, R., Medina, E., Werger, M.J.A., 2000. Leaf optical properties in Venezuelan cloud forest trees. Tree Physiol. 20, 519–526.

Price, K.V., Storn, R.M., Lampinen, J.A., 2005. Differential Evolution: A Practical Approach to Global Optimization. Springer, Berlin; New York.

Quebbeman, J.A., Ramirez, J.A., 2016. Optimal allocation of leaf-level nitrogen: Implications for covariation of V_{cmax} and J_{max} and photosynthetic downregulation. J. Geophys. Res. Biogeosci. 121:2464–2475. http://dx.doi.org/10.1002/2016JG003473.

R Core Team, 2012. R: a language and environment for statistical computing. R Found. Stat. Comput. Vienna Austria. http://www.R-project.org/ (ISBN 3-900051-07-0).

Rasmussen, C.E., Williams, C.K.I., 2006. Gaussian Processes for Machine Learning. MIT Press, Cambridge, Mass.

Rogers, A., 2014. The use and misuse of V c,max in Earth System Models. Photosynth. Res. 119:15–29. http://dx.doi.org/10.1007/s11120-013-9818-1.

Rozendaal, D.M.A., Hurtado, V.H., Poorter, L., 2006. Plasticity in leaf traits of 38 tropical tree species in response to light; relationships with light demand and adult stature. Funct. Ecol. 20:207–216. http://dx.doi.org/10.1111/j.1365-2435.2006.01105.x.

Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., Rundquist, D., 2013. Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. Int. J. Appl. Earth Obs. Geoinf. 25:47–54. http://dx.doi.org/10.1016/j.jag.2013.04.003.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9:671–675. http://dx.doi.org/10.1038/nmeth.2089.

Seemann, J.R., Sharkey, T.D., Wang, J., Osmond, C.B., 1987. Environmental effects on photosynthesis, nitrogen-use efficiency, and metabolite pools in leaves of sun and shade plants. Plant Physiol. 84, 796–802.

Serbin, S.P., Dillaway, D.N., Kruger, E.L., Townsend, P.A., 2012. Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. J. Exp. Bot. 63:489–502. http://dx.doi.org/10.1093/jxb/err294.

Serbin, S.P., Singh, A., Desai, A.R., Dubois, S.G., Jablonski, A.D., Kingdon, C.C., ... Townsend, P.A., 2015. Remotely estimating photosynthetic capacity, and its response to temperature, in vegetation canopies using imaging spectroscopy. Remote Sens. Environ. 167:78–87. http://dx.doi.org/10.1016/j.rse.2015.05.024.

Shiklomanov, A.N., Dietze, M.C., Viskari, T., Townsend, P.A., Serbin, S.P., 2016. Quantifying the influences of spectral resolution on uncertainty in leaf trait estimates through a Bayesian approach to RTM inversion. Remote Sens. Environ. 183:226–238. http://dx.doi.org/10.1016/j.rse.2016.05.023.

Siesler, H.W., 2007. Basic principles of near-infrared spectroscopy. In: Burns, D.A., Ciurczak, E.W. (Eds.), Handbook of near-infrared analysis, third ed. CRC press, pp. 7–19.

Singh, A., Serbin, S., McNeil, B.E., Kingdon, C.C., Townsend, P.A., 2015. Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. Ecol. Appl. 25, 2180–2197.

Stuffler, T., Kaufmann, C., Hofer, S., Förster, K.P., Schreier, G., Mueller, A., ... Haydn, R., 2007. The EnMAP hyperspectral imager—an advanced optical payload for future applications in Earth observation programmes. Acta Astronautica 61:115–120. http://dx.doi.org/10.1016/j.actaastro.2007.01.033.

Sudo, E., Makino, A., Mae, T., 2003. Differences between rice and wheat in ribulose-1, 5-bisphosphate regeneration capacity per unit of leaf-N content. Plant Cell Environ. 26, 255–263.

Taiz, L., Zeiger, E., 2010. Plant Physiology. fifth ed. Sinauer Assoc, Sunderland, USA.

Terashima, I., Evans, J.R., 1988. Effects of light and nitrogen nutrition on the organization of the photosynthetic apparatus in spinach. Plant Cell Physiol. 29, 143–155.

Turner, D.P., Ollinger, S.V., Kimball, J.S., 2004. Integrating remote sensing and ecosystem process models for landscape-to regional-scale analysis of the carbon cycle. Bioscience 54, 573–584.

Verrelst, J., Rivera, J.P., van der Tol, C., Magnani, F., Mohammed, G., Moreno, J., 2015. Global sensitivity analysis of the SCOPE model: What drives simulated canopy-leaving sun-induced fluorescence? Remote Sens. Environ. 166:8–21. http://dx.doi.org/10.1016/j.rse.2015.06.002.

von Caemmerer, S., 2000. Biochemical Models of Leaf Photosynthesis. CSIRO, Collingwood.

Walker, A.P., Beckerman, A.P., Gu, L., Kattge, J., Cernusak, L.A., Domingues, T.F., ... Woodward, F.I., 2014. The relationship of leaf photosynthetic traits - V_{cmax} and J_{max} - to leaf nitrogen, leaf phosphorus, and specific leaf area: a meta-analysis and modeling study. Ecology and Evolution 4:3218–3235. http://dx.doi.org/10.1002/ece3.1173.

Wang, Z., Skidmore, A.K., Darvishzadeh, R., Heiden, U., Heurich, M., Wang, T., 2015. Leaf nitrogen content indirectly estimated by leaf traits derived from the PROSPECT model. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8:3172–3182. http://dx.doi.org/10.1109/JSTARS.2015.2422734.

Wang, Z., Wang, T., Darvishzadeh, R., Skidmore, A., Jones, S., Suarez, L., ... Hearne, J., 2016. Vegetation indices for mapping canopy foliar nitrogen in a mixed temperate forest. Remote Sens. 8:491. http://dx.doi.org/10.3390/rs8060491.

Williams, P., 2006. Near-infrared spectroscopy of cereals. Handbook of Vibrational Spectroscopy.

Wold, S., Ruhe, H., Wold, H., Dunn, W.J., 1984. The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM J. Sci. Stat. Comput. 5, 735–743.

Workman Jr., J., Weyer, L., 2007. Practical Guide to Interpretive near-Infrared Spectroscopy. CRC press.

Wullschleger, S.D., 1993. Biochemical limitations to carbon assimilation in C3 plants—a retrospective analysis of the A/Ci curves from 109 species. J. Exp. Bot. 44, 907–920.

Xu, L., Baldocchi, D.D., 2003. Seasonal trends in photosynthetic parameters and stomatal conductance of blue oak (*Quercus douglasii*) under prolonged summer drought and high temperature. Tree Physiol. 23, 865–877.

Zarco-Tejada, P.J., Miller, J.R., Mohammed, G.H., Noland, T.L., 2000. Chlorophyll fluorescence effects on vegetation apparent reflectance: I. Leaf-level measurements and model simulation. Remote Sens. Environ. 74, 582–595.