A machine learning framework for estimating leaf biochemical parameters from its spectral reflectance and transmission measurements.

Bikram Koirala, *Student Member IEEE*, Zohreh Zahiri, *Member IEEE*, Paul Scheunders, *Senior Member IEEE*

***Abstract*—Spectral measurements are commonly applied for the nondestructive estimation of leaf parameters, such as the concentrations of chlorophyll ab, carotenoid, anthocyanin, and brown pigment, the leaf water content, and the leaf mass per area for quantification of vegetation physiology. The most popular way to estimate these parameters is by using spectral vegetation indices. The use of biochemical models allows to employ the full wavelength range (400-2500 nm) and to physically interpret the result. However, their performance is usually lower than that of supervised machine learning regression techniques. Machine learning regression techniques, on the other hand, have the disadvantage that the relation between estimated parameters and the reflectance/transmission spectra is unclear.**

In this paper, a hybrid between a supervised learning method and physical modeling for the estimation of leaf parameters is proposed. In this method, a machine learning regression technique is applied to learn a mapping from the true hyperspectral dataset to a dataset that follows the PROSPECT model. The PROSPECT model then reveals the actual leaf parameters. Two mapping methods, based on gaussian processes (GP) and kernel ridge regression (KRR) are proposed. As an alternative, a mapping onto the leaf absorption spectra is proposed as well. The proposed methodology not only estimates the leaf parameters with a lower error but also solves the interpretation problem of the parameters estimated by the advanced machine learning regression techniques. This method is validated on the ANGERS and LOPEX dataset.

***Index Terms*—Hyperspectral, leaf parameter estimation, machine learning regression**

I. Introduction

Retrieval of leaf parameters, e.g., the concentrations of chlorophyll ab (C_{ab}), carotenoid (C_{xc}), anthocyanin (C_{anth}), the water content (C_w), and the leaf mass per area (C_m) is of great interest due to their direct connection with the vegetation's physiological functions ([1],[2],[3],[4],[5],[6],[7],[8]). The most popular way to relate reflectance/transmission spectra with leaf parameters is by the use of spectral vegetation indices ([9],[10],[11],[12],[13],[14]). As an example, the Normalized Difference Vegetation Index (NDVI) ([12]) uses two bands, one correlated with the chlorophyll (red), and the other uncorrelated (near infrared).

Since techniques based on spectral vegetation indices use a limited number of spectral bands, to extract critical information from a quasi-continuous spectral signal, shape indices were developed. In [13], they were categorized into four classes: a) Red-edge position ([15],[16]), b) Integration-based indices ([17],[18]), c) Derivative-based indices ([19]), and d) Continuum removal ([20],[21],[22],[23]).

Instead of relating a few wavelengths or individual absorption features with the leaf parameters, several deterministic models have been developed to describe the optical properties of plant leaves. These models can be distinguished by the complexity level that is taken into account and the underlying physics. In [24], they are categorized into four classes of models. The simplest class of plate models represents the leaf by absorbing plates with rough surfaces, isotropically diffusing the incident rays of light. N-flux models describe leaves as slabs, diffusing and absorbing the material. Stochastic and radiative transfer models simulate the optical properties of the leaf by using a Markov chain or by directly using a radiative transfer equation. The most complex models are the ray tracing models. They require the optical properties of the leaf material and a detailed description of the internal structure of the leaf.

In the remote sensing community, an improved version of the plate models, i.e., the PROSPECT model ([1],[3],[4]) is widely used. It describes the optical properties of plant leaves in the wavelength range $\lambda \in [400, 2500]$ nm. This model describes the reflectance and transmission spectrum of the leaf as a function of the leaf parameters (C_{ab}, C_{xc}, C_{anth}, C_w, C_m), and their corresponding specific absorption spectra, a wavelength dependent refractive index ($n(\lambda)$) and a parameter characterizing the leaf mesophyll structure (N_{lms}).

In the past two decades, much research has been reported on the estimation of leaf biochemical parameters by inverting the PROSPECT model ([1],[8],[25],[26],[27],[28],[29]). Among the leaf biochemical parameters, C_{ab} and C_w have been studied most extensively because of their strong absorption features in the visible and shortwave infrared ([29]). Quantification of C_{xc}, C_{anth}, and C_m was shown to be much more challenging because their specific absorption spectra overlap with the spectrum of C_{ab} in the visible region and with C_w in the shortwave infrared. To improve the performance of the PROSPECT model for the retrieval of C_m, several strategies have been proposed. In [27], the ill-posedness of the PROSPECT inversion was alleviated by selecting for each leaf biochemical parameter separately the wavelength region to which it is sensitive. In [8], the spectral range from 1700 to 2400 nm was identified as the optimal range for the estimation of C_m. In [30], the model PROSPECT-g was developed that introduced a wavelength-independent factor to represent anisotropic scattering in the elementary layer.

Bikram Koirala, Zohreh Zahiri, and Paul Scheunders are with the Imec-VisionLab, Department of Physics, University of Antwerp, Belgium e-mail: {bikram.koirala, zohreh.zahiri, paul.scheunders}@uantwerpen.be

However, one important drawback of the PROSPECT model is its utilization of fixed specific absorption spectra and refractive index spectrum, hereby assuming that these spectra are the same for the leaves of all plant species.

To account for the spectral variability of the refractive index spectrum of plant leaves and the specific absorption spectra of the leaf biochemical parameters, several advanced machine learning regression algorithms have been used to retrieve biochemical parameters ([2],[8],[29],[31],[32],[33],[34],[35]). The goal of these algorithms is to model the predictive function that best approximates the relationship between spectra and the parameters of interest. These are supervised methods that require a training set of spectra and ground-truth information of leaf parameters. Due to the nonlinear relationship between spectra and the biochemical parameters, kernel methods have been introduced to make the regression algorithms nonlinear. The most popular kernel-based regression methods are Kernel Ridge Regression (KRR) and Gaussian Process Regression (GP).

Opposed to the PROSPECT model that relates the reflectance/transmission spectra to the specific absorption spectra of the leaf parameters, the machine learning regression methods map the reflectance/transmission spectra directly to the leaf parameters. One particular problem with this direct mapping is that the physical relationship between the biochemical parameters and the reflectance/transmission spectra is lost. As a consequence, the estimated values of the leaf parameters do not necessarily fall within their physical range, and even can become negative. A nonnegativity constraint could be enforced on the output variables, but in that case, there is no closed-form solution.

In this paper, an alternative supervised technique for retrieval of leaf parameters is proposed. This method also assumes that a training set of spectra and ground-truth information of leaf parameters is available. The PROSPECT model is used to generate spectra from the ground truth parameters of the training data. Then, a mapping between the actual training spectra and the spectra generated by the PROSPECT model is learned. Two mapping methods are presented, based on KRR ([36],[37]) and GP ([38]). Once the mapping is learned, all test spectra are mapped to the PROSPECT model, and the leaf parameters of the mapped spectra are estimated by inverting the PROSPECT model. As an alternative, the mapping to the leaf absorption spectra which are, according to the PROSPECT model, given by a linear combination of the specific absorption spectra of the leaf parameters, is performed. Inverting the linear model then delivers the parameters.

The proposed methodology combines the physical interpretability of the PROSPECT model with the flexibility and generalizability of the regression methods. The generalization properties of the machine learning regression approaches account for the spectral variability of the refractive index spectrum of plant leaves and the specific absorption spectra of the pigments. The use of the PROSPECT model allows to physically relate the estimated leaf parameters to the reflectance/transmission spectra of the plant leaves.

The remaining of the paper is organized as follows: In section II, the datasets and the different methodologies to estimate leaf parameters from the hyperspectral datasets is described. The PROSPECT model, the different kernel regression methods, and the proposed strategy will be explained. The experimental results are presented in section III and discussed in section IV. Section V concludes this work.

II. Experimental datasets and methods

A. Datasets

1) ANGERS: The ANGERS leaf optical properties database was generated in 2003 at INRA in Angers (France) [1]. This dataset contains transmission and reflectance spectra of 276 leaf samples (43 plant species) and the ground-truth information regarding four parameters (C_{ab}, C_{xc}, C_w, C_m). ASD Field spectroradiometers were used to capture leaf directional-hemispherical reflectance and transmittance spectra (350-2500 nm) with a spectral sampling of 1.4 nm and 2 nm in the VNIR (350-1050 nm) and SWIR (1000-2500) respectively. To extract biochemical information, leaf discs were sampled using a cork borer immediately after the measurement of spectra. The fresh weight of these discs was measured before placing them in a drying oven at 85^0C. After drying them for 48 h, the C_w, and C_m were determined by reweighing. Simultaneously, pigments were extracted using ethanol 95% by grinding fresh leaf discs in a chilled mortar. To prevent acidification, a small amount of $MgCO_3$ and quartz sand was added. The solution of ethanol 95% and pigments were separated from other materials by centrifugation. Further, the absorption spectra of the solution were measured using a dual beam scanning UV-Vis spectrophotometer. C_{ab} and C_{xc} were estimated by using a multi-wavelength analysis ([39]). C_{ab} ranges between 0.78-106.70 μg cm^{-2} and C_{xc} ranges between 0.00-25.28 μg cm^{-2}. The C_w ranges between 0.0044-0.034 cm and C_m ranges between 0.0017-0.0331 g cm^{-2}.

2) LOPEX: This dataset contains transmission and reflection spectra of 330 leaf samples (66 plants) that were captured from 45 plant species and the ground-truth information of four different leaf parameters (C_{ab}, C_{xc}, C_w, and C_m) [1],[40]. The leaf directional-hemispherical reflectance and transmission spectra were captured over the wavelength range 400-2500 nm with 1 nm step size by using Perkin Elmer Lambda 19 spectrophotometers. The spectral resolution of this dataset is 1-2 nm and 4-5 nm in the VNIR (400-1000 nm) and SWIR (1000-2500) respectively [40]. The procedure of estimating leaf parameters from this dataset is similar to the ANGERS dataset, except that acetone 100% was used for extracting leaf pigments. Although this dataset contains four leaf parameters, only C_w and C_m were used for validating the proposed methodology, since the values of C_{ab} and C_{xc} for several leaves of the same plant ([41]) [1] are exactly the same, making these values unreliable. The values of C_w of this dataset range between 0.0021-0.0525 cm while the values for C_m range between 0.0017-0.0157 g cm^{-2}.

B. The PROSPECT model

The PROSPECT model ([1],[3],[4]) is the improved version of a generalized "plate model" ([42],[43]), describing a leaf

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

as a pile of N homogeneous layers separated by $N-1$ air spaces.

The prospect model describes the total reflectance $\mathbf{R}(\lambda)$ and transmission $\mathbf{T}(\lambda)$ of the N layers as a function of the leaf absorption spectrum $\mathbf{k}(\lambda)$ and the leaf refractive index spectrum $\mathbf{n}(\lambda)$ in the wavelength region 400-2500 nm. In its turn, the leaf absorption spectrum is assumed to be a linear combination of the plant biochemical parameters and their corresponding specific absorption spectra:

$$\mathbf{k} = \frac{\sum_{j=1}^{p} \mathbf{k}_{spe,j} c_j}{N_{lms}} \tag{1}$$

where $\mathbf{k}_{spe,j}$ is the specific absorption spectrum of leaf parameter c_j and p is the number of leaf parameters. N_{lms} is the leaf mesophyll structure.

Based on a large number of spectral measurements and ground truth information on the leaf parameters (C_{ab}, C_{xc}, C_w, and C_m), part of which come from the ANGERS and LOPEX dataset, the PROSPECT model has been inverted to obtain an average refractive index spectrum $\mathbf{n}(\lambda)$ and average specific absorption spectra of C_{ab}, C_{xc}, C_w, and C_m [1]. Remark that these are averages over a large number of plant species, and they are assumed to be fixed. The latest version of the PROSPECT model, PROSPECT-D [4] includes two extra leaf parameters: C_{anth}, and the concentration of brown pigment (C_{br}).

With the assumption of fixed $\mathbf{n}(\lambda)$ and $\mathbf{k}_{spe,j}(\lambda)$, the PROSPECT model can now be inverted to estimate the leaf parameters from measured reflectance ($\mathbf{R}_{meas}(\lambda)$) and transmission ($\mathbf{T}_{meas}(\lambda)$) spectra from individual leaves:

$$\begin{aligned} \mathbf{\Theta} &= \arg\min_{\mathbf{\Theta}} \sum_{\lambda} \big[\big(\mathbf{R}_{meas}(\lambda) - \mathbf{R}(\lambda, \mathbf{\Theta})\big)^2 \\ &+ \big(\mathbf{T}_{meas}(\lambda) - \mathbf{T}(\lambda, \mathbf{\Theta})\big)^2 \big] \end{aligned} \tag{2}$$

where $\mathbf{\Theta} = \{N_{lms}, \{c_j\}_{j=1}^{p}\}$, and $\mathbf{R}(\lambda, \mathbf{\Theta})$ and $\mathbf{T}(\lambda, \mathbf{\Theta})$ are the modeled reflectance and transmission spectra by the PROSPECT model.

Although physically sound for the retrieval of leaf parameters, the PROSPECT model has some problems. As mentioned before, the refractive index spectrum is assumed to be constant, while it actually can vary a lot between different leaf samples. Another problem is that the chlorophyll a:b ratio is assumed to be constant, and therefore, the specific absorption spectra of chlorophyll a and b are estimated simultaneously ([4]). Moreover, other pigments are present, the carotenoid group contains xanthophylls and the anthocyanin group contains several different anthocyanins and it is assumed that these don't influence the determination of the specific absorption spectra from these groups. These assumptions lead to the lower performance of the PROSPECT model compared to the advanced machine learning regression algorithms.

C. Machine learning regression algorithms

Machine learning regression algorithms learn the relationship between the high dimensional input (reflectance/transmission spectra) and low dimensional output (leaf parameters) based on a training dataset.

Let us consider a set of N samples $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^{N}$, where $\mathbf{x}_i = \begin{bmatrix} \mathbf{R}_{meas,i} \\ \mathbf{T}_{meas,i} \end{bmatrix}$, $\mathbf{R}_{meas,i}$ and $\mathbf{T}_{meas,i} \in \mathbf{R}_{+}^{d}$ represent the measured reflectance and transmission spectra and $\mathbf{y}_i = \mathbf{C}_i \in \mathbf{R}_{+}^{p}$ is the in-situ measurement of p leaf parameters. The goal of machine learning regression is to learn a mapping function:

$$\mathbf{y} = f(\mathbf{x}) + \boldsymbol{\epsilon} \tag{3}$$

where $\boldsymbol{\epsilon}$ is additive noise. To learn this mapping function, among N samples, n training samples $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^{n}$ are used. After learning this mapping, the performance of the model is tested on the remaining $(N-n)$ samples $\mathbf{X}_* = \{\mathbf{x}_i\}_{i=n+1}^{N}$. Two state-of-the-art machine learning regression algorithms, KRR and GP are presented.

1) Kernel ridge regression: Ridge regression finds a linear relationship between the input $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^{n}$ and output $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^{n}$:

$$\mathbf{y}_i = \mathbf{w}^T \mathbf{x}_i \tag{4}$$

Generally, to tackle the problem of overfitting the training samples, the quadratic cost function J is regularized by the norm of the model weights $\mathbf{w}$:

$$J = 1/2 \left(\left\| \mathbf{Y} - \mathbf{w}^T \mathbf{X} \right\|^2 + \lambda \left\| \mathbf{w} \right\|^2 \right) \tag{5}$$

where λ is the regularization parameter. Minimizing 5 leads to:

$$\mathbf{w} = \left(\mathbf{X}\mathbf{X}^T + \lambda \mathbf{I} \right)^{-1} \left(\mathbf{X}\mathbf{Y}^T \right) \tag{6}$$

where $\mathbf{I}$ is the identity matrix. In the above equation, a matrix with size $(2d \times 2d)$ needs to be inverted. To do the inversion after kernelization, equation 6 has to be re-arranged to contain a matrix of size $(n \times n)$:

$$\begin{aligned} \mathbf{w} &= \left(\mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda \mathbf{I})\mathbf{X}^{-1} \right)^{-1} \left(\mathbf{X}\mathbf{Y}^T \right) \\ &= \mathbf{X} \left(\mathbf{X}^T\mathbf{X} + \lambda \mathbf{I} \right)^{-1} \mathbf{Y}^T \end{aligned} \tag{7}$$

Once the mapping is found, the prediction of the leaf parameters from the test data $\mathbf{Y}_* = \{\mathbf{y}_i\}_{i=n+1}^{N}$ is obtained by:

$$\mathbf{Y}_* = \mathbf{Y} \left(\mathbf{X}^T\mathbf{X} + \lambda \mathbf{I} \right)^{-1} \mathbf{X}^T\mathbf{X}_* \tag{8}$$

To allow nonlinear relationships between input and output, ridge regression needs to be kernelized. A kernelized extension of ridge regression is presented in [36][37]. The original dataset is projected onto an infinite dimensional feature space ($\mathbf{x}_i \rightarrow \boldsymbol{\phi}(\mathbf{x}_i)$). Using the kernel trick, i.e., $k(\mathbf{x}_i, \mathbf{x}_j) = \boldsymbol{\phi}(\mathbf{x}_i)^T \boldsymbol{\phi}(\mathbf{x}_j)$, the mapping of a nonlinear spectra $\mathbf{X}_*$ to the leaf parameters $\mathbf{Y}_*$ is obtained by:

$$\mathbf{Y}_* = f(\mathbf{X}_*) = \mathbf{Y} \left(K(\mathbf{X}, \mathbf{X}) + \lambda \mathbf{I} \right)^{-1} K(\mathbf{X}, \mathbf{X}_*) \tag{9}$$

where $K(\mathbf{X}, \mathbf{X})$ is the matrix of kernel functions between the n training samples (with a dimension of $(n \times n)$) and $K(\mathbf{X}, \mathbf{X}_*)$ is the matrix of kernel functions between the n training samples and the $(N-n)$ test samples (with a dimension of $(n \times (N-n))$).

In this work, a radial basis function (RBF) kernel is applied as the kernel function:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \tag{10}$$

In equation 9, the $(n \times n)$ kernel matrix K that is regularized by λ needs to be inverted. For each test sample, the only computation involved is to determine the kernel function between the n training samples and the test samples. The regularization parameter λ and the parameter of the kernel (σ) were tuned by 10-fold cross-validation of the training samples [44]. To determine the optimal pair ($\hat{\sigma}$, $\hat{\lambda}$), all possible combinations of $\sigma \in \left\{2^{-15}, \cdots, 2^3\right\}$ and $\lambda \in \left\{2^{-15}, \cdots, 2^5\right\}$ were applied and the average mapping error was calculated.

2) Gaussian processes: An alternative strategy to learn the nonlinear relationship between the input $\mathbf{X}$ and the output $\mathbf{Y}$ is given by gaussian process regression (GP). GP is a bayesian approach that estimates the distribution of mapping functions that are consistent with the training set $\{(\mathbf{x}_i,\mathbf{y}_i)|i = 1, 2, \cdots, n\}$.

It is assumed that the observed leaf parameters $(\mathbf{y}_i)$ are related to the input spectra $(\mathbf{x}_i)$ as follows:

$$\mathbf{y}_i = f(\mathbf{x}_i) = \boldsymbol{\phi}(\mathbf{x}_i)^T \mathbf{w} \tag{11}$$

with prior distribution for $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \Sigma_{2d})$. The function $\boldsymbol{\phi}(\cdot)$ maps the input spectrum to an infinite dimensional feature space. The mean and covariance of the outputs can then be computed as follows:

$$\begin{aligned} \mathbb{E}[f(\mathbf{x}_i)] =& \boldsymbol{\phi}(\mathbf{x}_i)^T \mathbb{E}[\mathbf{w}] = \mathbf{0} \\ \mathbb{E}[f(\mathbf{x}_i) f(\mathbf{x}_j)] =& \boldsymbol{\phi}(\mathbf{x}_i)^T \mathbb{E}[\mathbf{w}\mathbf{w}^T] \boldsymbol{\phi}(\mathbf{x}_j) = \boldsymbol{\phi}(\mathbf{x}_i)^T \Sigma_{2d} \boldsymbol{\phi}(\mathbf{x}_j). \end{aligned} \tag{12}$$

GP assumes that the covariance of the outputs is modeled by a squared exponential kernel function:

$$\boldsymbol{\phi}(\mathbf{x}_i)^T \Sigma_{2d} \boldsymbol{\phi}(\mathbf{x}_j)^T = k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\left(-\sum_{b=1}^{2d} \frac{\left(x_i^b - x_j^b\right)^2}{2 l_b^2}\right) \tag{13}$$

where l_b is a characteristic length-scale for each spectral band and σ_f^2 is the variance of the input spectra.

The joint distribution of the estimated leaf parameters from the test data $(f(\mathbf{X}_*))$ and the training leaf parameters $(\mathbf{Y})$ is then given by:

$$\begin{aligned} p(f(\mathbf{X}_*^T), \mathbf{Y}^T) &\sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K(\mathbf{X}_*, \mathbf{X}_*) & K(\mathbf{X}_*, \mathbf{X}) \\ K(\mathbf{X}, \mathbf{X}_*) & K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I} \end{bmatrix}\right) \\ &= \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right) \end{aligned} \tag{14}$$

where σ_n^2 is the noise variance of the training spectra, $K(\mathbf{X}_*, \mathbf{X})$ is the matrix of kernel functions between the test samples and the n training samples, and $K(\mathbf{X}_*, \mathbf{X}_*)$ is the matrix of kernel functions between the test samples.

When inverting the partitioned matrix:

$$\begin{aligned} &\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1} \\ =& \begin{pmatrix} \Sigma^{-1} & -\Sigma^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1}\Sigma_{21}\Sigma^{-1} & \Sigma_{22}^{-1} + \Sigma_{22}^{-1}\Sigma_{21}\Sigma^{-1}\Sigma_{12}\Sigma_{22}^{-1} \end{pmatrix} \end{aligned} \tag{15}$$

with $\Sigma = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, (14) can be factorized into the predictive distribution $p(f(\mathbf{X}_*^T)|\mathbf{Y}^T)$ and the marginal $p(\mathbf{Y}^T)$:

$$\begin{aligned} p(f(\mathbf{X}_*^T), \mathbf{Y}^T) =& p(f(\mathbf{X}_*^T)|\mathbf{Y}^T) p(\mathbf{Y}^T) \\ =& \mathcal{N}(\Sigma_{12}\Sigma_{22}^{-1}\mathbf{Y}^T, \Sigma) \mathcal{N}(\mathbf{0}, \Sigma_{22}) \end{aligned} \tag{16}$$

The estimated mapping of the nonlinear spectra $\mathbf{X}_*$ to the leaf parameters $\mathbf{Y}_*$ is then given by:

$$\begin{aligned} \mathbf{Y}_* &= f(\mathbf{X}_*) = \mathbf{Y}\Sigma_{22}^{-1}\Sigma_{12}^T \\ &= \mathbf{Y}(K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I})^{-1} K(\mathbf{X}_*, \mathbf{X})^T \end{aligned} \tag{17}$$

The hyperparameters involved in (13) are automatically optimized by minimizing the log marginal likelihood of the training set: $\log(p(\mathbf{Y}^T|\mathbf{X}^T))$.

D. The proposed method: mapping to the PROSPECT model

The main disadvantage of applying GP and KRR to map the spectra directly to the parameters is that the estimated leaf parameters are not physically related to the spectra. When leaf parameters are estimated by applying equation 9 or 17, there is no guarantee that the estimated leaf parameters are positive. To solve these problems, the PROSPECT model and the machine learning regression techniques are combined in such a way that the estimated parameters are physically interpretable while the accuracy of the parameter estimation is comparable to the unconstrained direct mapping to the leaf parameters. The main idea is to learn a mapping from the actual spectra to spectra that follow the PROSPECT model, after which the leaf parameters can be estimated by inverting the model.

This method consist of the following steps:

1) In the first step, target spectra $\left(\mathbf{x}_i^{\text{target}} = \begin{bmatrix} \mathbf{R}_i \\ \mathbf{T}_i \end{bmatrix}\right)$ are generated by using the ground truth information $(\mathbf{C}_i)$ and the PROSPECT model.

2) In the second step, a mapping between the true spectra $(\mathbf{X})$ and the target spectra $(\mathbf{X}^{\text{target}} = \{\mathbf{x}_i^{\text{target}}\}_{i=1}^n)$ is learned:

$$\mathbf{X}^{\text{target}} = f(\mathbf{X}) + \epsilon \tag{18}$$

The learning of this mapping can be performed using any machine learning regression algorithm. In this work, GP and KRR were used. When the mapping was learned using GP, the mapping between the true test spectra $(\mathbf{X}_* = \{\mathbf{x}_i\}_{i=n+1}^N)$ and the target spectra $(\mathbf{X}_*^{\text{target}} = \{\mathbf{x}_i^{\text{target}}\}_{i=n+1}^N)$ is given by:

$$\begin{aligned} \mathbf{X}_*^{\text{target}} =& f(\mathbf{X}_*) \\ =& \mathbf{X}^{\text{target}}(K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I})^{-1} K(\mathbf{X}_*, \mathbf{X})^T \end{aligned} \tag{19}$$

using equation 13 for computing the kernel functions. We will refer to this method as GP_PROSPECT. The prediction using KRR is given by:

$$\begin{aligned} \mathbf{X}_*^{\text{target}} =& f(\mathbf{X}_*) \\ =& \mathbf{X}^{\text{target}}(K(\mathbf{X}, \mathbf{X}) + \lambda \mathbf{I})^{-1} K(\mathbf{X}_*, \mathbf{X})^T \end{aligned} \tag{20}$$

using equation 10 for computing the kernel functions. We will refer to this method as KRR_PROSPECT.

3) The PROSPECT-D model contains seven leaf parameters. In case the number of ground truth leaf parameters is smaller than those that are involved in the PROSPECT model, parameters that are not a part of the ground truth information are estimated from the true spectra $\left(\{\mathbf{R}_{meas,i}, \mathbf{T}_{meas,i}\}_{i=1}^{N}\right)$ by inverting the PROSPECT model (see 2). Now, the ground truth information becomes:

$$\mathbf{C}_i^{\text{true/est.}} = [\mathbf{C}_i(1:l), \mathbf{C}_i(l+1:p)]^T \tag{21}$$

with l is the number of parameters that are part of the ground truth information and p the total number of parameters that are involved in the PROSPECT model. In step 1, the target spectra are then generated by using $\mathbf{C}_i^{\text{true/est.}}$.

4) In the final step, leaf parameters from the mapped spectra $\left(\hat{\mathbf{x}}_i^{\text{target}} = \begin{bmatrix}\hat{\mathbf{R}}_i \\ \hat{\mathbf{T}}_i\end{bmatrix}\right)$ are estimated by inverting the PROSPECT model.

E. Alternative method: mapping to the leaf absorption spectra

Instead of mapping the hyperspectral dataset onto the PROSPECT spectra, an alternative approach is to estimate the leaf parameters by mapping the hyperspectral dataset $(\mathbf{X})$ onto the absorption spectra $\mathbf{k}_i$. So, this time, in step 1, the target spectra $\mathbf{x}_i^{\text{target}}$ that are generated are absorption spectra by using the ground truth information $(\mathbf{C}_i)$ and equation 1. To maintain consistent notations, we will define $\mathbf{x}_i^{\text{target}} = \begin{bmatrix}\mathbf{k}_i \\ \mathbf{k}_i\end{bmatrix}$.

The mapping between a hyperspectral training set and the ground truth absorption spectra is learned by GP or KRR (step 2). Then, the test hyperspectral data are mapped onto the absorption spectra. The parameters $\mathbf{C}_i^*$ from the mapped spectra $\left(\hat{\mathbf{x}}_i^{\text{target}} = \begin{bmatrix}\mathbf{k}_i^{\mathbf{R}*} \\ \mathbf{k}_i^{\mathbf{T}*}\end{bmatrix}\right)$ are estimated by (step 4):

$$\begin{aligned}\mathbf{C}_i^* &= \arg\min_{\mathbf{C}_i^*} \sum_{\lambda} \left[\left(\mathbf{k}_i^{\mathbf{R}*}(\lambda) - \mathbf{k}(\lambda, \mathbf{C}_i^*)\right)^2\right. \\ &+ \left.\left(\mathbf{k}_i^{\mathbf{T}*}(\lambda) - \mathbf{k}(\lambda, \mathbf{C}_i^*)\right)^2\right] \end{aligned} \tag{22}$$

where $\mathbf{k}_i^{\mathbf{R}*}$ and $\mathbf{k}_i^{\mathbf{T}*}$ are the mapped absorption spectra from the reflectance and the transmission spectrum respectively, and $\mathbf{C}_i^*$ is the estimated leaf parameter of the test spectrum. This estimation can be performed by including physical constraints of the leaf parameters, i.e., lower and upper bounds.

When the mapping between the hyperspectral training set and the ground truth absorption spectra is learned by GP, we will refer to this method as GP_LINEAR. When KRR is used, we refer to the method as KRR_LINEAR.

F. Experimental set-up and Evaluation statistics

To reduce the computational complexity and the dimensionality, the hyperspectral datasets with a 1 nm step-size were resampled to 10 nm. For estimating the performance of the described methods, the ground truth data set was divided into a randomly selected training and a test set. For the ANGERS dataset, five different experiments were performed, by selecting 15, 45, 75, 105 and 135 training samples randomly respectively. For the LOPEX dataset, six different experiments were performed, by selecting 15, 45, 75, 105, 135 and 165 training samples randomly respectively. Each experiment was repeated 100 times.

The following methods were compared:

- The PROSPECT model
- Methods that map the spectra directly to the leaf parameters: KRR and GP
- The proposed methods that map the spectra onto spectra that follow the PROSPECT model: KRR_PROSPECT and GP_PROSPECT
- The proposed methods that map the spectra onto absorption spectra that follow a linear model: KRR_LINEAR and GP_LINEAR

The performance of each regression model for each leaf parameter was evaluated based on the normalized root mean squared error (NRMSE) between the estimated and ground truth leaf parameter to measure the accuracy and the average Pearson's determination coefficient (R^2) to measure the goodness-of-fit:

$$\text{NRMSE } (\%) = \frac{\sqrt{\frac{1}{N-n}\sum_{i=n+1}^{N}(y_{ji} - \hat{y}_{ji})^2} \times 100}{\max(y_{j(n+1)} : y_{jN}) - \min(y_{j(n+1)} : y_{jN})} \tag{23}$$

$$\text{R}^2 = 1 - \frac{\sum_{i=n+1}^{N}(y_{ji} - \hat{y}_{ji})^2}{\sum_{i=n+1}^{N}(y_{ji} - \bar{y}_j)^2} \tag{24}$$

where y_{ji} is the true leaf parameter j and $\hat{y}_{ji}$ the estimated leaf parameter j for test sample i, and $\bar{y}_j$ is the mean of the true leaf parameter j over all test samples.

III. Results

A. ANGERS dataset

Fig. 1 and Fig. 2 show the mean and standard deviation of the NRMSE and R^2 respectively for 100 runs as a function of the applied number of training samples that were selected randomly. From top left to bottom right, results are shown for water content (C_w), the concentration of chlorophyll ab (C_{ab}), leaf mass per area (C_m) and the concentration of carotenoid (C_{xc}).

The results indicate that the estimation error is reduced when the number of applied training samples is increased. Also, almost all methods outperform the PROSPECT model from a certain number of training samples on. For each of the two regression methods, the proposed strategy of mapping onto the PROSPECT model or the leaf absorption spectra outperforms the direct mapping onto the leaf parameters. In general, mapping onto the leaf absorption spectra delivers the best results. Except for WC, GP_LINEAR outperforms all other methods. It outperforms the PROSPECT model already when only 15-45 training samples are applied.

Fig. 3 shows the validation of the prediction models for each of the four leaf parameters for the case of 75 training samples. Each time, from the 100 experiments, the result with the best NRMSE and R^2 is depicted. It can be seen that all methods accurately estimated C_{ab}, C_{xc}, C_w and C_m for a large range of values. Both KRR and GP predicted negative values

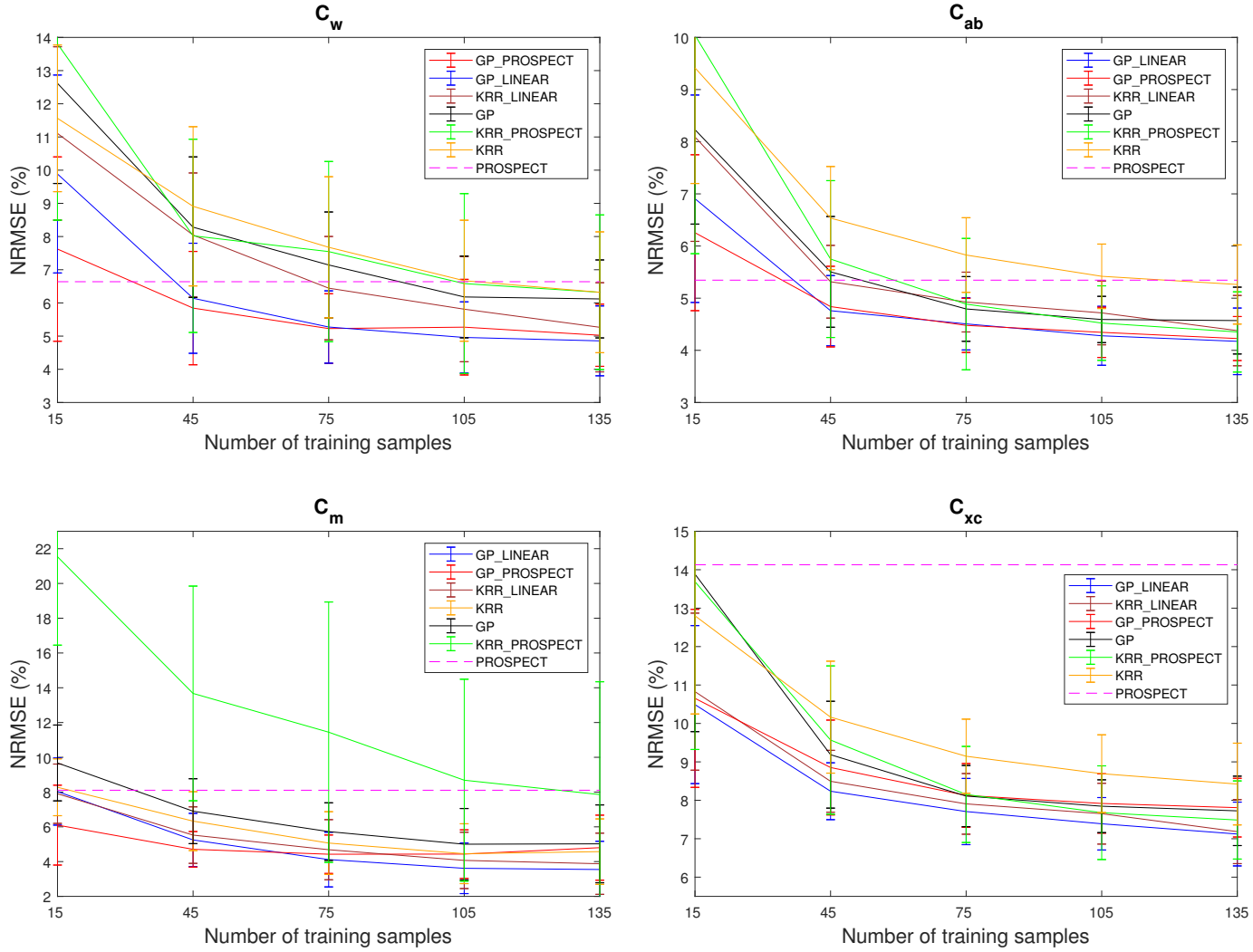


Fig. 1: NRMSE (100 runs) obtained by the PROSPECT model, GP, GP_PROSPECT, GP_LINEAR, KRR, KRR_PROSPECT and KRR_LINEAR, in function of applied number of training samples (the ANGERS dataset). C_w, C_{ab}, C_m, and C_{xc} refer to water content, the concentration of chlorophyll ab, leaf mass per area, and the concentration of carotenoid respectively.

for C_{ab} and C_{xc}. These results demonstrate that mapping to the PROSPECT model or the leaf absorption spectra avoids negative values and obtains a sound physical interpretation of the estimated parameters.

B. LOPEX dataset

Fig. 4 and Fig. 5 show the mean and standard deviation of the NRMSE and R^2 respectively for 100 runs as a function of the applied number of training samples that were selected randomly for the LOPEX dataset. On the left, the results are shown for C_w, and on the right for C_m. Fig. 6 shows the validation of the prediction models for C_w and C_m for the case of 75 training samples. Each time, from the 100 experiments, the result with the best NRMSE and R^2 is shown.

Similar results are obtained as with the ANGERS dataset, although the advantage of the machine learning regression methods over the PROSPECT model are not so clear in case of C_w. In case of C_m, mapping onto the leaf absorption spectra outperforms the other methods.

C. Training on the ANGERS and tested on the LOPEX dataset

To test the generalization capability of the proposed methodology, the models were trained by using the ANGERS dataset and were validated on the LOPEX dataset. Although the ANGERS dataset contains ground truth of four different leaf parameters, only C_w and C_m were used to make it compatible with the LOPEX ground truth leaf parameters. The experiment was limited to the wavelength region 900-2500 nm, because leaf pigments do not have absorption features in that region, and thus will not influence the results. The hyperspectral dataset (reflectance/transmission) with a 1 nm step-size was resampled to 8 nm resulting in 201 wavebands. Fig. 7 shows the validation of the prediction models for C_w and C_m when the models were trained by using 276 training samples from the ANGERS dataset. The proposed methods outperformed

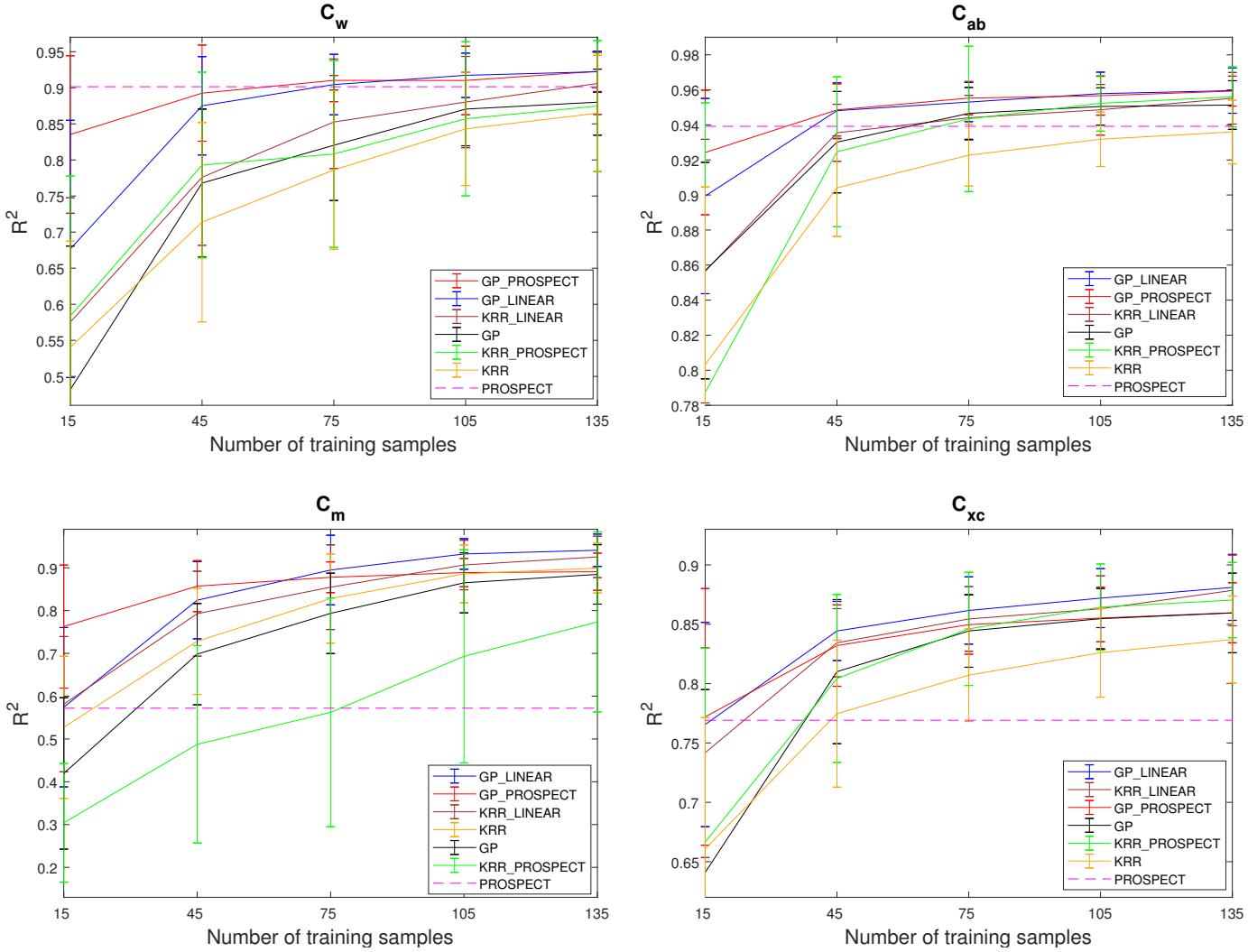


Fig. 2: R^2 (100 runs) obtained by the PROSPECT model, GP, GP_PROSPECT, GP_LINEAR, KRR, KRR_PROSPECT and KRR_LINEAR, in function of applied number of training samples (the ANGERS dataset). C_w, C_{ab}, C_m, and C_{xc} refer to water content, the concentration of chlorophyll ab, leaf mass per area, and the concentration of carotenoid respectively.

the direct mapping onto the leaf parameters. KRR_LINEAR was the best performer for the estimation of both C_w and C_m.

D. Training on the LOPEX and tested on the ANGERS dataset

Similarly, the models were trained by using the LOPEX dataset and were validated on the ANGERS dataset. Fig. 8 shows the validation of the prediction models for C_w and C_m when the models were trained by using 330 training samples from the LOPEX dataset. From the figure, it can be observed that KRR_PROSPECT was the best performer for estimating C_w with the lowest NRMSE and the highest R^2 while both direct mapping methods (KRR and GP) could not perform better than the PROSPECT model. For the estimation of C_m, the PROSPECT model was the best performer with the lowest NRMSE but R^2 of KRR_LINEAR was the highest. Both KRR and GP estimated negative values for C_m for several spectra.

IV. Discussion

From the experimental results, the following general conclusions can be drawn:

- The supervised methods outperform the use of the PROSPECT model for estimating leaf biochemical parameters from both the LOPEX and the ANGERS dataset. This is partially because these methods make use of a training dataset. However, the generic nature of the regression algorithms allows them to account for the spectral variability of the specific absorption spectra and refractive index spectrum. It also demonstrates that a few training samples (15-45) are enough to outperform the PROSPECT model.
- The strategy of mapping reflectance/transmission spectra onto either the PROSPECT model or to the linear model outperforms methods that directly map to the leaf parameters. The main difference is that the direct mapping techniques lose the physical relation between the

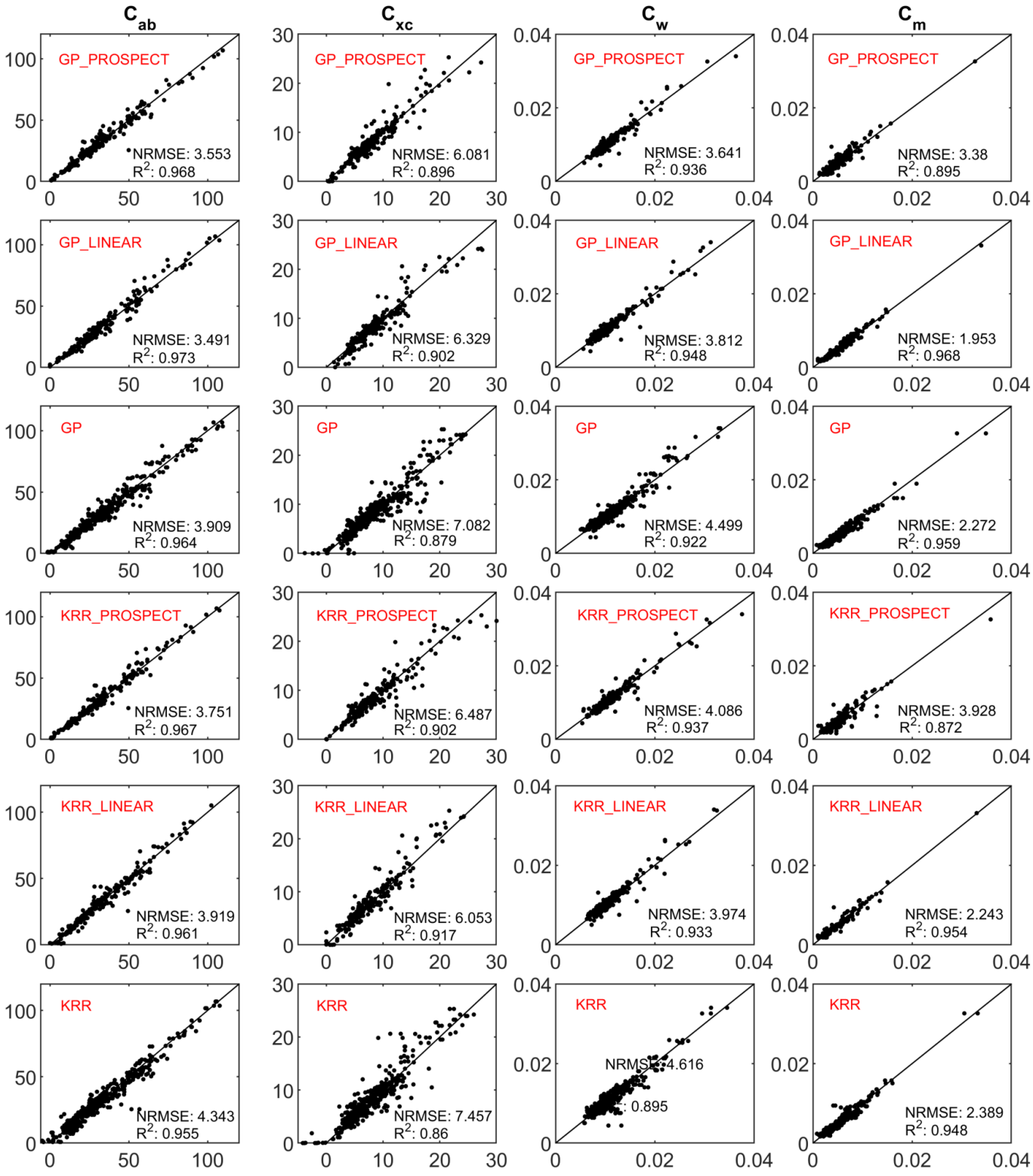


Fig. 3: Validation between the measured (Y-axis) and the estimated (X-axis) values of the concentration of chlorophyll ab (μg cm^{-2}), carotenoid (μg cm^{-2}), water content (cm), and leaf mass per area (g cm^{-2}). The presented results are the best prediction from the 100 runs using 75 training samples (the ANGERS dataset).

reflectance/transmission spectra and the leaf parameters. In Fig. 3 (C_{ab} and C_{xc}) and Fig. 8 (C_m), negative values can be observed for the estimated leaf biochemical parameters by direct mapping (GP and KRR).

- The performance of both the PROSPECT model and the supervised techniques is affected by the uncertainty

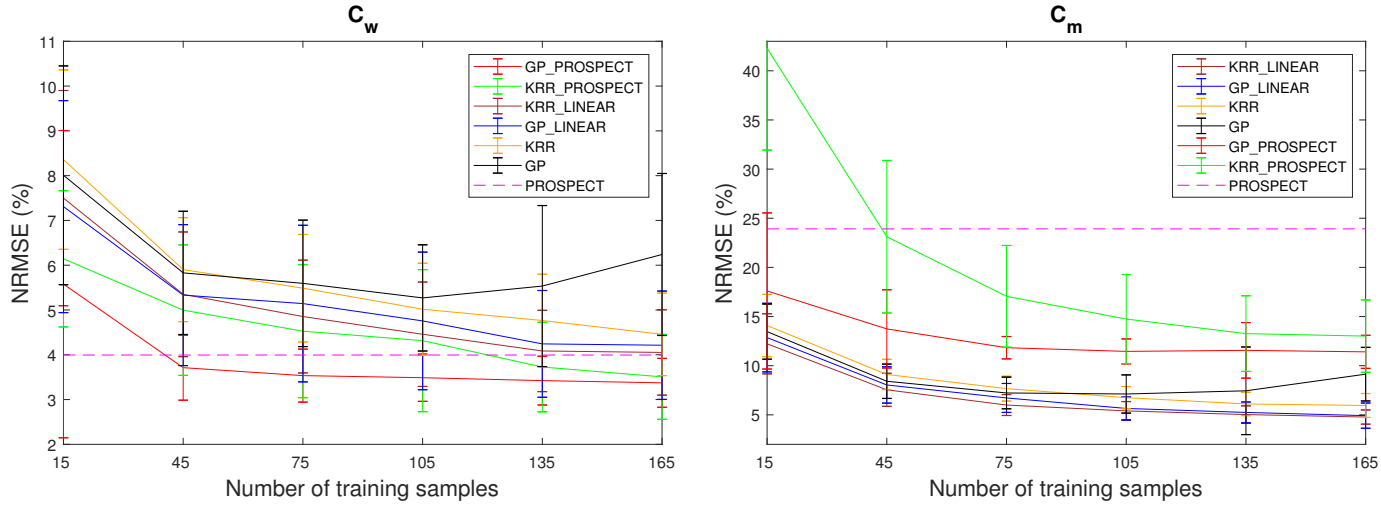


Fig. 4: NRMSE (100 runs) obtained by PROSPECT, GP, GP_PROSPECT, GP_LINEAR, KRR, KRR_PROSPECT and KRR_LINEAR in function of applied number of training samples (the LOPEX dataset). C_w and C_m refer to water content and leaf mass per area respectively.

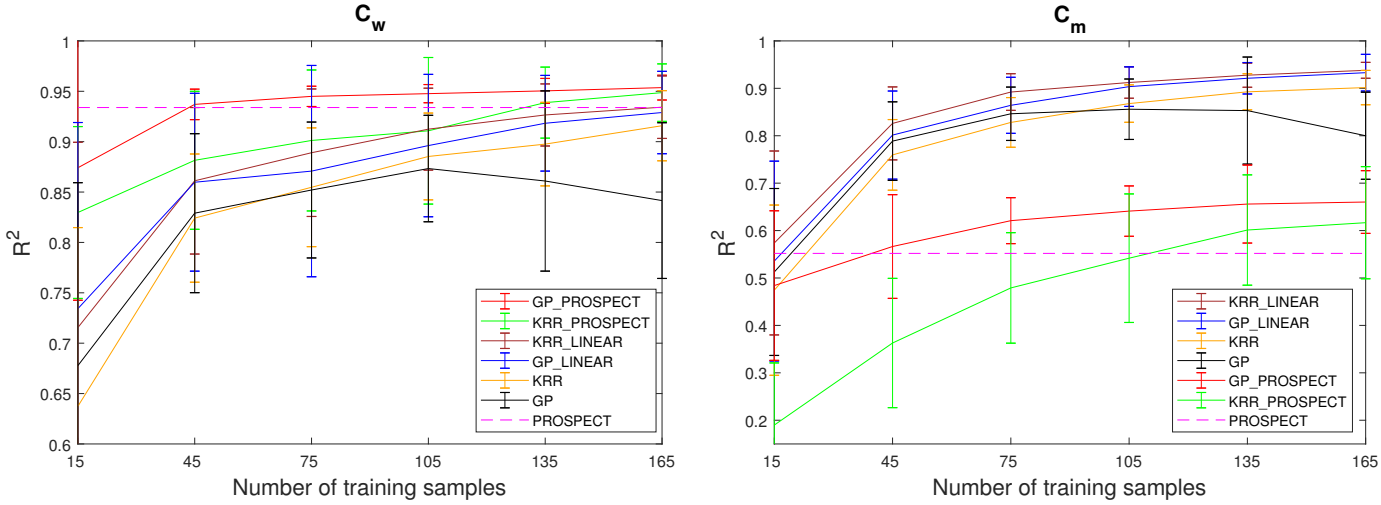


Fig. 5: R^2 (100 runs) obtained by PROSPECT, GP, GP_PROSPECT, GP_LINEAR, KRR, KRR_PROSPECT and KRR_LINEAR in function of applied number of training samples (the LOPEX dataset). C_w and C_m refer to water content and leaf mass per area respectively.

in the ground truth measurements of the biochemical parameters. The uncertainty on C_w and C_m is mainly due to improper measurements of the weight and area of the leaf discs. The uncertainty in the pigment concentrations and in particular in C_{xc} is expected to be high. In the ANGERS dataset, C_{xc} was estimated from the absorption spectra of the solution (ethanol 95% and pigments) by using the equation of Lichtenthaler (1987) ([39]). Due to the complexity of the mixture of chlorophyll and carotenoid pigments, a chromatography technique (high-pressure liquid chromatography) would be required to prepare high-quality ground truth of C_{xc}. Moreover, to predict C_{xc} accurately by using the PROSPECT model, the specific absorption spectrum for each carotenoid pigment from the carotenoid group is required. In the ANGERS dataset, no distinction is made between these different pigments.

Because of the high uncertainty in the ground truth of C_{xc}, both the supervised approaches and the PROSPECT model performed low at the estimation of this parameter. Although the results on the other parameters are better, one cannot expect errors to be lower than the uncertainties in the ground truth.

- The low performance of the PROSPECT model for estimating C_m is reported in many studies [8],[45],[46],[47]. This can also be observed in the results of the LOPEX dataset where the error (NRMSE) on the estimation of C_m was higher than 20%. This is because a single specific absorption spectrum is defined for the leaf mass per area, i.e., an average spectrum for dry matter.

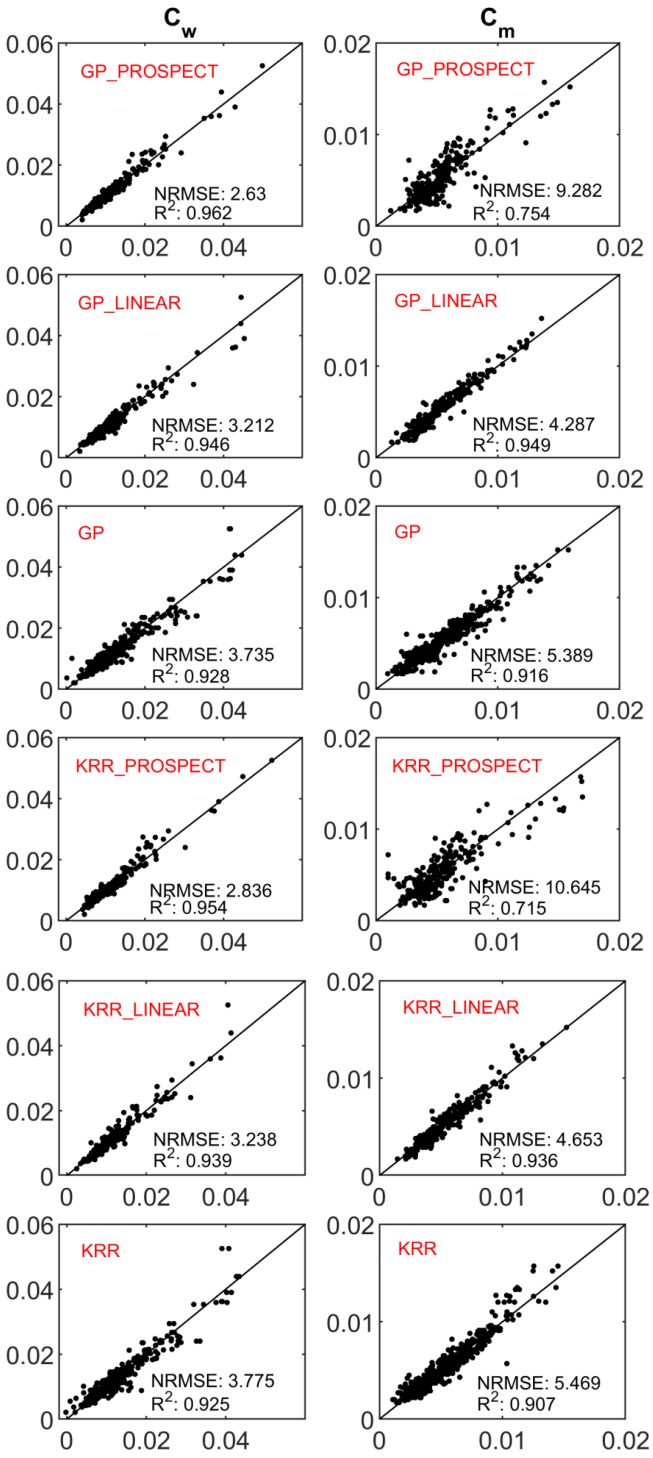


Fig. 6: Validation between the measured (Y-axis) and the estimated (X-axis) values of water content, and leaf mass per area. The presented results are the best prediction with best NRMSE and R^2 from the 100 runs using 75 training samples (the LOPEX dataset).

However, dry matter contains various organic materials (cellulose, hemicellulose, lignin, proteins, starch), each with their absorption spectrum. When using a single specific absorption spectrum, it is implicitly assumed that the relative proportion of each of these single constituents are constant among the leaves. The main reason for the better performance of the PROSPECT model for this parameter on the ANGERS dataset is that the specific absorption spectrum of dry matter was better adapted to the ground truth measurements. All supervised methods perform better on the estimation of C_m, because they adapt to the spectral variability of the dry matter.

- The performance of the proposed methodologies was equivalent to or better than the PROSPECT model when an independent dataset was used for the model validation. This demonstrates their generalization capability.
- An advantage of the proposed methodology is that any nonlinear regression algorithm can be applied for learning the mapping. In this work, two different kernel methods were compared. Gaussian processes generally seem to outperform kernel ridge regression when training and testing samples are from the same dataset. But kernel ridge regression outperformed gaussian processes when training and testing samples were independent from each other. On the other hand, gaussian process was computationally expensive when spectra were either mapped to the PROSPECT model or the linear model.
- Although the leaf mesophyll structure (N_{lms}) impacts simulated spectra, there is no protocol to experimentally estimate N_{lms} from leaf samples. Generally, it is determined by inverting the PROSPECT model. N_{lms} has a maximum effect in the NIR from 800-1000 nm where absorption is at its minimum ([1]). To estimate it, in [1], only three wavelengths were used to invert the PROSPECT model, corresponding to the maximum reflectance, the maximum transmittance, and the minimum absorptance respectively. In this work, the values were obtained by inverting the PROSPECT model using the entire spectrum (400-2500 nm) and were very close to the optimal ones provided by the datasets (ANGERS and LOPEX).
- To investigate the impact of the number of ground truth leaf parameters on the retrieved leaf biochemical parameters, C_{ab} was disregarded from the ANGERS dataset. The estimation error of C_{xc} (highly correlated with C_{ab}) was increased by only 0.1-1% when 75 training samples were used for learning the mapping.
- Although BRF spectra are more practical for leaf-level applications, they contain a significant specular component. The PROSPECT model is calibrated with directional-hemispherical reflectance and transmission factor spectra, captured by spectrometers equipped with an integrating sphere, and cannot simulate BRF spectra. To account for the specular component, in [48], a physically-based method called PROCOSINE was proposed. Similarly, in [49], the PROSPECT model was coupled with a continuous wavelet transform (PROCWT) to suppress the effect of the specular component. The generic nature of the proposed methodology allows replacing the PROSPECT model either with PROCOSINE or PROCWT. The PROCOSINE or PROCWT model can tackle the specular component, while the advanced

machine learning algorithms can account for the spectral variability of the specific absorption spectra and refractive index spectrum.

- The use of the PROSPECT model assumes that both reflectance and transmission spectra are available to estimate leaf biochemical parameters. However, when only reflectance or BRF spectra are available, as is the case at the regional or the global level, the proposed regression methodology can be combined with canopy models (such as PROSAIL) to estimate both leaf area index and chlorophyll content.
  To demonstrate the potential of the proposed methodology for canopy level applications, we performed an experiment, only using the reflectance spectra of both LOPEX and ANGERS data sets. Also, in this case, we observed that the proposed methodology outperformed the PROSPECT model, while the direct mapping methods performed worse or only slightly better than the PROSPECT model in most cases.
- All methods were developed in Matlab and ran on an Intel Core $i7\text{-}8700K$ CPU, 3.20 GHz machine with 6 cores. The runtimes of the proposed methods on the LOPEX dataset (see III-D) and the ANGERS dataset (see III-C) are shown in Table I. As can be seen, the runtime of GP_PROSPECT and GP_LINEAR is relatively high due to the involvement of 2×203 hyperparameters. KRR_PROSPECT, KRR_LINEAR, and KRR have a lower runtime compare to GP_PROSPECT and GP_LINEAR because it involves only two free parameters.

TABLE I: The runtime in seconds.

Method	$\text{time}_{\text{LOPEX}}$ (s)	$\text{time}_{\text{ANGERS}}$ (s)
KRR_PROSPECT	25.34	25.85
GP_PROSPECT	384.02	384.12
KRR_LINEAR	25.17	25.89
GP_LINEAR	384.78	382.29
KRR	20.96	20.96
GP	4.99	5.16

V. Conclusion

In this work, a hybrid between a model-based and supervised data-driven method for leaf parameter estimation from spectral reflectance/transmission measurements was proposed. The proposed method is based on the learning of a mapping between a true hyperspectral dataset and the PROSPECT model. Two kernel-based mapping methods are proposed. As an alternative, mapping to the leaf absorption spectra is proposed as well. The proposed methods are shown to outperform the PROSPECT model and supervised machine learning regression methods that map directly to the leaf parameters.

The procedure to map reflectance/transmission spectra to the PROSPECT model can be extended to any biochemical/physical model. The main limitation of this method is that it cannot be applied to estimate parameters which are not part of existing radiative transfer models.

Acknowledgement

The research presented in this paper is funded by BELSPO (Belgian Science Policy Office) in the frame of the STEREO III programme – project GEOMIX (SR/06/357).

References

[1] J.-B. Feret, C. François, G. P. Asner, A. A. Gitelson, R. E. Martin, L. P. Bidel, S. L. Ustin, G. le Maire, and S. Jacquemoud, "Prospect-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments," *Remote Sensing of Environment*, vol. 112, no. 6, pp. 3030 – 3043, 2008.

[2] S. V. Wittenberghe, J. Verrelst, J. P. Rivera, L. Alonso, J. Moreno, and R. Samson, "Gaussian processes retrieval of leaf parameters from a multi-species reflectance, absorbance and fluorescence dataset," *Journal of Photochemistry and Photobiology B: Biology*, vol. 134, pp. 37 – 48, 2014.

[3] S. Jacquemoud and F. Baret, "Prospect: A model of leaf optical properties spectra," *Remote Sensing of Environment*, vol. 34, no. 2, pp. 75 – 91, 1990.

[4] J.-B. Féret, A. Gitelson, S. Noble, and S. Jacquemoud, "Prospect-d: Towards modeling leaf optical properties through a complete lifecycle," *Remote Sensing of Environment*, vol. 193, pp. 204 – 215, 2017.

[5] Y. Zhang, J. M. Chen, J. R. Miller, and T. L. Noland, "Leaf chlorophyll content retrieval from airborne hyperspectral remote sensing imagery," *Remote Sensing of Environment*, vol. 112, no. 7, pp. 3234 – 3247, 2008.

[6] M. Schlerf, C. Atzberger, J. Hill, H. Buddenbaum, W. Werner, and G. Schüler, "Retrieval of chlorophyll and nitrogen in norway spruce (picea abies l. karst.) using imaging spectroscopy," *International Journal of Applied Earth Observation and Geoinformation*, vol. 12, no. 1, pp. 17 – 26, 2010.

[7] S. L. Ustin, A. Gitelson, S. Jacquemoud, M. Schaepman, G. P. Asner, J. A. Gamon, and P. Zarco-Tejada, "Retrieval of foliar information about plant pigment systems from high resolution spectroscopy," *Remote Sensing of Environment*, vol. 113, pp. S67 – S77, 2009, imaging Spectroscopy Special Issue.

[8] J.-B. Féret, G. le Maire, S. Jay, D. Berveiller, R. Bendoula, G. Hmimina, A. Cheraiet, J. Oliveira, F. Ponzoni, T. Solanki, F. de Boissieu, J. Chave, Y. Nouvellon, A. Porcar-Castell, C. Proisy, K. Soudani, J.-P. Gastellu-Etchegorry, and M.-J. Lefèvre-Fonollosa, "Estimating leaf mass per area and equivalent water thickness based on leaf optical properties: Potential and limitations of physical modeling and machine learning," *Remote Sensing of Environment*, vol. 231, p. 110959, 2019.

[9] J. PEÑUELAS, I. FILELLA, C. BIEL, L. SERRANO, and R. SAVÉ, "The reflectance at the 950–970 nm region as an indicator of plant water status," *International Journal of Remote Sensing*, vol. 14, no. 10, pp. 1887–1905, 1993.

[10] A. A. Gitelson, Y. G. †, and M. N. Merzlyak, "Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves," *Journal of Plant Physiology*, vol. 160, no. 3, pp. 271 – 282, 2003.

[11] Q. Wang and P. Li, "Hyperspectral indices for estimating leaf biochemical properties in temperate deciduous forests: Comparison of simulated and measured reflectance data sets," *Ecological Indicators*, vol. 14, no. 1, pp. 56 – 65, 2012.

[12] J. W. Rouse, R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS." in *Proceedings of the Third ERTS Symposium*, vol. 1. NASA, Dec. 1973, pp. 309–317.

[13] J. Verrelst, G. Camps-Valls, J. Muñoz-Marí, J. P. Rivera, F. Veroustraete, J. G. Clevers, and J. Moreno, "Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – a review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 108, pp. 273 – 290, 2015.

[14] S. Jay, N. Gorretta, J. Morel, F. Maupas, R. Bendoula, G. Rabatel, D. Dutartre, A. Comar, and F. Baret, "Estimating leaf chlorophyll content in sugar beet canopies using millimeter- to centimeter-scale reflectance imagery," *Remote Sensing of Environment*, vol. 198, pp. 173 – 186, 2017.

[15] J. Delegido, J. Verrelst, L. Alonso, and J. Moreno, "Evaluation of sentinel-2 red-edge bands for empirical estimation of green lai and chlorophyll content," *Sensors*, vol. 11, no. 7, pp. 7063–7081, 2011.

[16] W. J. Frampton, J. Dash, G. Watmough, and E. J. Milton, "Evaluating the capabilities of sentinel-2 for quantitative estimation of biophysical variables in vegetation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 82, pp. 83 – 92, 2013.

[17] Z. Malenovský, C. Ufer, Z. Lhotáková, J. Clevers, M. E. Schaepman, J. Albrechtová, and P. Cudlín, "A new hyperspectral index for chlorophyll estimation of a forest canopy: Area under curve normalised to maximal band depth between 650-725 nm," 2006.

[18] J. Delegido, L. Alonso, G. González, and J. Moreno, "Estimating chlorophyll content of crops from hyperspectral data using a normalized area over reflectance curve (naoc)," *International Journal of Applied Earth Observation and Geoinformation*, vol. 12, no. 3, pp. 165 – 174, 2010.

[19] P. J. Zarco-Tejada, J. R. Miller, G. H. Mohammed, T. L. Noland, and P. H. Sampson, "Vegetation stress detection through chlorophyll a + b estimation and fluorescence effects on hyperspectral imagery." *Journal of environmental quality*, vol. 31 5, pp. 1433–41, 2002.

[20] P. J. Curran, J. L. Dungan, and D. L. Peterson, "Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: Testing the kokaly and clark methodologies," *Remote Sensing of Environment*, vol. 76, no. 3, pp. 349 – 359, 2001.

[21] Z. Malenovský, L. Homolová, R. Zurita-Milla, P. Lukeš, V. Kaplan, J. Hanuš, J.-P. Gastellu-Etchegorry, and M. E. Schaepman, "Retrieval of spruce leaf chlorophyll content from airborne image data using continuum removal and radiative transfer," *Remote Sensing of Environment*, vol. 131, pp. 85 – 102, 2013.

[22] J. Wang, R. Xu, and S. Yang, "Estimation of plant water content by spectral absorption features centered at 1,450 nm and 1,940 nm regions," *Environmental Monitoring and Assessment*, vol. 157, pp. 459–469, 2008.

[23] A. B. González-Fernández, J. R. Rodríguez-Pérez, M. Marabel, and F. Álvarez Taboada, "Spectroscopic estimation of leaf water content in commercial vineyards using continuum removal and partial least squares regression," *Scientia Horticulturae*, vol. 188, pp. 15 – 22, 2015.

[24] S. Jacquemoud and S. L. Ustin, "Leaf optical properties: A state of the art," 2000.

[25] J. Sun, S. Shi, J. Yang, B. Chen, W. Gong, L. Du, F. Mao, and S. Song, "Estimating leaf chlorophyll status using hyperspectral lidar measurements by prospect model inversion," *Remote Sensing of Environment*, vol. 212, pp. 1 – 7, 2018.

[26] K. Barry, G. Newnham, and C. Stone, "Estimation of chlorophyll content in eucalyptus globulus foliage with the leaf reflectance model prospect," *Agricultural and Forest Meteorology*, vol. 149, no. 6, pp. 1209 – 1213, 2009.

[27] P. Li and Q. Wang, "Retrieval of leaf biochemical parameters using prospect inversion: A new approach for alleviating ill-posed problems," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 7, pp. 2499–2506, July 2011.

[28] Q. Miao, W. Zhao, X. Guo, K. Liu, J. Han, and Z. Wang, "Inversion of reed leaf chlorophyll content based on prospect model," in *2011 19th International Conference on Geoinformatics*, June 2011, pp. 1–4.

[29] J.-B. Féret, C. François, A. Gitelson, G. P. Asner, K. M. Barry, C. Panigada, A. D. Richardson, and S. Jacquemoud, "Optimizing spectral indices and chemometric analysis of leaf chemical properties using radiative transfer modeling," *Remote Sensing of Environment*, vol. 115, no. 10, pp. 2742 – 2750, 2011.

[30] F. Qiu, J. M. Chen, W. Ju, J. Wang, Q. Zhang, and M. Fang, "Improving the prospect model to consider anisotropic scattering of leaf internal materials and its use for retrieving leaf biomass in fresh leaves," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 6, pp. 3119–3136, June 2018.

[31] L. Li, Y.-B. Cheng, S. Ustin, X.-T. Hu, and D. Riaño, "Retrieval of vegetation equivalent water thickness from reflectance using genetic algorithm (ga)-partial least squares (pls) regression," *Advances in Space Research*, vol. 41, no. 11, pp. 1755 – 1763, 2008.

[32] B. Liu, Y.-M. Yue, R. Li, W.-J. Shen, and K.-L. Wang, "Plant leaf chlorophyll content retrieval based on a field imaging spectroscopy system," *Sensors*, vol. 14, no. 10, pp. 19 910–19 925, 2014.

[33] S. Ullah, A. K. Skidmore, A. Ramoelo, T. A. Groen, M. Naeem, and A. Ali, "Retrieval of leaf water content spanning the visible to thermal infrared spectra," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, pp. 56 – 64, 2014.

[34] S. H. Shah, Y. Angel, R. Houborg, S. Ali, and M. F. McCabe, "A random forest machine learning approach for the retrieval of leaf chlorophyll content in wheat," *Remote Sensing*, vol. 11, no. 8, 2019.

[35] C. R. Yendrek, T. Tomaz, C. M. Montes, Y. Cao, A. M. Morse, P. J. Brown, L. M. McIntyre, A. D. Leakey, and E. A. Ainsworth, "High-throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reflectance," *Plant Physiology*, vol. 173, no. 1, pp. 614–626, 2017.

[36] M. Welling, *Kernel ridge regression*. Toronto: Max Welling's Classnotes in Machine Learning, 2013.

[37] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*. Cambridge: Cambridge University Press, 2004.

[38] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. New York: The MIT Press, 2006.

[39] H. K. Lichtenthaler, "[34] chlorophylls and carotenoids: Pigments of photosynthetic biomembranes," in *Plant Cell Membranes*, ser. Methods in Enzymology. Academic Press, 1987, vol. 148, pp. 350 – 382.

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

[41] S. Jacquemoud, S. Ustin, J. Verdebout, G. Schmuck, G. Andreoli, and B. Hosgood, "Estimating leaf biochemistry using the prospect leaf optical properties model," *Remote Sensing of Environment*, vol. 56, no. 3, pp. 194 – 202, 1996.

[42] W. A. Allen, H. W. Gausman, A. J. Richardson, and J. R. Thomas, "Interaction of isotropic light with a compact plant leaf∗," *J. Opt. Soc. Am.*, vol. 59, no. 10, pp. 1376–1379, Oct 1969.

[43] W. A. Allen, H. W. Gausman, A. J. Richardson, and C. L. Wiegand, "Mean effective optical constants of thirteen kinds of plant leaves," *Appl. Opt.*, vol. 9, no. 11, pp. 2573–2577, Nov 1970.

[44] P. Exterkate, "Model selection in kernel ridge regression," *Computational Statistics and Data Analysis*, vol. 68, p. 16 pages, 2013.

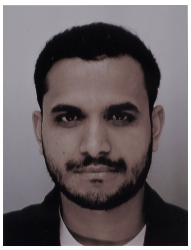
[45] D. Riano, P. Vaughan, E. Chuvieco, P. J. Zarco-Tejada, and S. L. Ustin, "Estimation of fuel moisture content by inversion of radiative transfer models to simulate equivalent water thickness and dry matter content: analysis at leaf and canopy level," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 819–826, April 2005.

[46] R. Colombo, M. Meroni, A. Marchesi, L. Busetto, M. Rossini, C. Giardino, and C. Panigada, "Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling," *Remote Sensing of Environment*, vol. 112, no. 4, pp. 1820 – 1834, 2008, remote Sensing Data Assimilation Special Issue.

[47] G. le Maire, C. François, K. Soudani, D. Berveiller, J.-Y. Pontailler, N. Bréda, H. Genet, H. Davi, and E. Dufrêne, "Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass," *Remote Sensing of Environment*, vol. 112, no. 10, pp. 3846 – 3864, 2008.

[48] S. Jay, R. Bendoula, X. Hadoux, J.-B. Féret, and N. Gorretta, "A physically-based model for retrieving foliar biochemistry and leaf orientation using close-range imaging spectroscopy," *Remote Sensing of Environment*, vol. 177, pp. 220 – 236, 2016.

[49] D. Li, T. Cheng, M. Jia, K. Zhou, N. Lu, X. Yao, Y. Tian, Y. Zhu, and W. Cao, "Procwt: Coupling prospect with continuous wavelet transform to improve the retrieval of foliar chemistry from leaf bidirectional reflectance spectra," *Remote Sensing of Environment*, vol. 206, pp. 1 – 14, 2018.



Bikram Koirala received the B.S. degree in Geomatic Engineering from the Purbanchal University, Nepal, and the M.S. degree in Geomatic Engineering from the University of Stuttgart, Germany in 2011 and in 2016 respectively. In 2017, he joined Vision Lab, Department of Physics, the University of Antwerp as a Ph.D. researcher. He is a student IEEE member. His research interest includes machine learning and hyperspectral image processing.

Zohreh Zahiri received the bachelor degree in Civil Engineering from Shahid Ashrafi Esfahani University, Isfahan, Iran, and the M.S. degree in Architectural Conservation from the Isfahan University of Art in 2009 and in 2013 respectively. She received her Ph.D. degree in "Characterization of facade materials using close-range hyperspectral data" from University College Dublin (UCD), Department of Civil Engineering, Dublin, Ireland. She is currently a postdoctoral researcher at Vision Lab, Department of Physics, University of Antwerp. Her research interests are the application of hyperspectral imaging in civil engineering and related areas as well as data analysis methods.

Paul Scheunders (M'98) received the B.S. degree and the Ph.D. degree in physics, with work in the field of statistical mechanics, from the University of Antwerp, Antwerp, Belgium, in 1983 and 1990, respectively. In 1991, he became a research associate with the Vision Lab, Department of Physics, University of Antwerp, where he is currently a full professor. His current research interest includes remote sensing and hyperspectral image processing. He has published over 200 papers in international journals and proceedings in the field of image processing, pattern recognition, and remote sensing. Paul Scheunders is Associate Editor of the IEEE Transactions on Geoscience and Remote Sensing and has served as a program committee member in numerous international conferences. He is a senior member of the IEEE Geoscience and Remote Sensing Society.

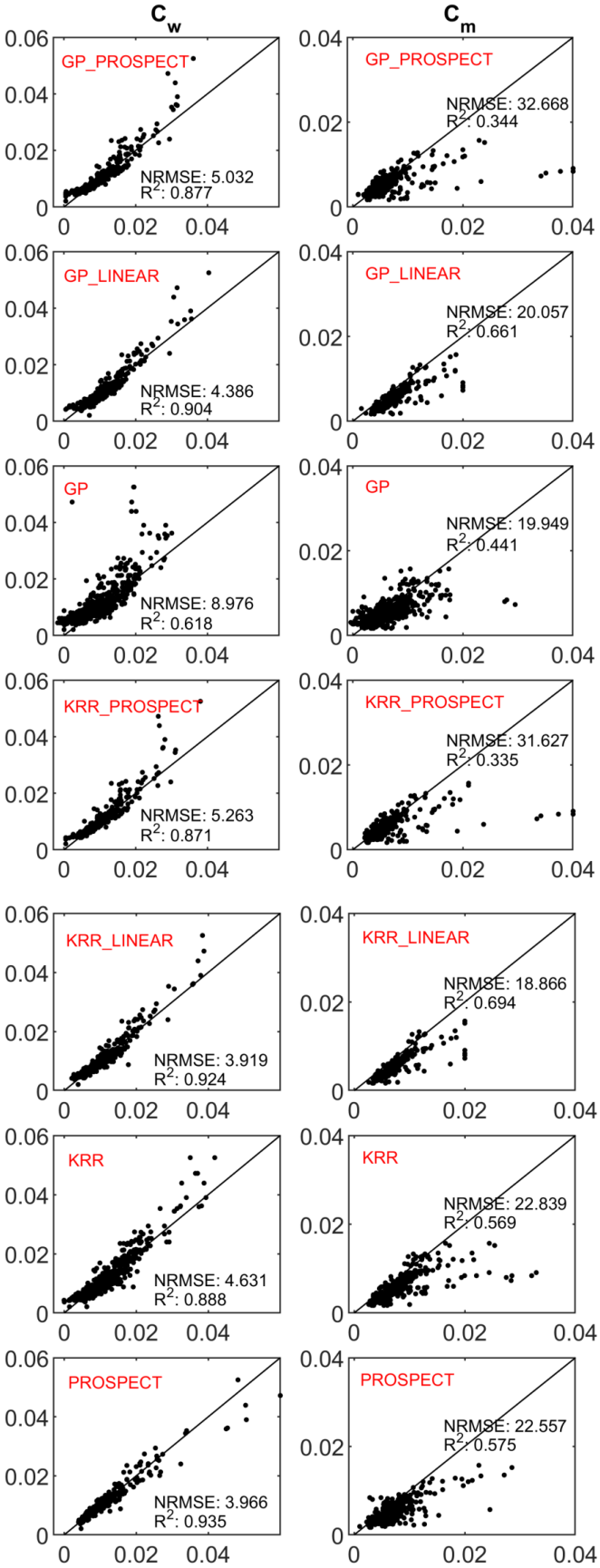


Fig. 7: Validation between the measured (Y-axis) and the estimated (X-axis) values of water content and leaf mass per area (the LOPEX dataset). The training was performed by applying the ANGERS dataset. C_w and C_m refer to water content and leaf mass per area respectively.

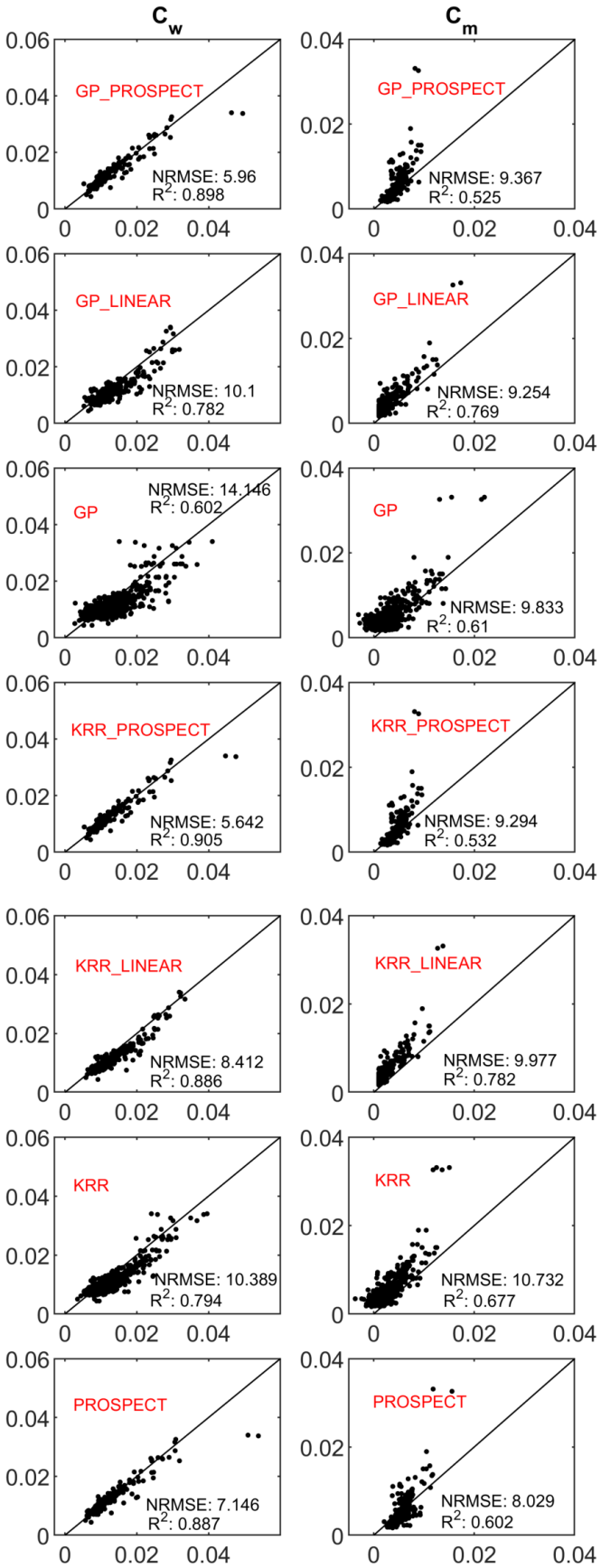


Fig. 8: Validation between the measured (Y-axis) and the estimated (X-axis) values of water content and leaf mass per area (the ANGERS dataset). The training was performed by applying the LOPEX dataset. C_w and C_m refer to water content and leaf mass per area respectively.