Computers and Electronics in Agriculture 195 (2022) 106806

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Non-destructive analysis of plant physiological traits using hyperspectral imaging: A case study on drought stress

Mohd Shahrimie Mohd Asaari [a,d,*], Stien Mertens [b,c], Lennart Verbraeken [b,c], Stijn Dhondt [b,c], Dirk Inzé [b,c], Koirala Bikram [a], Paul Scheunders [a]

[a] Imec-Vision Lab, University of Antwerp, Belgium
[b] Ghent University, Department of Plant Biotechnology and Bioinformatics, Ghent, Belgium
[c] VIB Center for Plant Systems Biology, Ghent, Belgium
[d] School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Engineering Campus, Nibong Tebal, Penang, Malaysia

ARTICLE INFO

Keywords:
Hyperspectral imaging
Machine learning regression
Plant phenotyping
Physiological traits
Non-destructive measurements
Drought stress

ABSTRACT

Conventional methods to access plant physiological traits are based on destructive measurements by means of biochemical extraction or leaf clipping, thereby limiting the throughput capability. With advances in hyperspectral imaging sensor, fast, non-invasive and non-destructive measurements of a plant's physiological status became feasible. In this work, a non-destructive method for the characterization of a plant's status from hyperspectral images is presented. A supervised data-driven method based on Machine Learning Regression (MLR) algorithms was developed to generate prediction models of four targeted physiological traits: water potential, effective quantum yield of photosystem II, transpiration rate and stomatal conductance. Standard Normal Variate (SNV) transformed reflectance spectra were used as the input variables for building the regression model. Three MLR algorithms: Gaussian Process Regression (GPR), Kernel Ridge Regression (KRR), and Partial Least Squares Regression (PLSR) were explored as candidate methods for building the prediction model of the targeted physiological traits. Validation results show that the non-linear prediction models, developed based on the GPR algorithm produced the best estimation accuracy on all plant traits. The best prediction models were applied to a small-scale phenotyping experiment to study drought stress responses in maize plants. Results show that all estimated traits revealed a significant difference between plants under drought stress and normal growth dynamics as early as after 3 days of drought induction.

1. Introduction

Plants initiate various morphological, biochemical, and complex physiological changes to respond to water deficit stress. In term of physiological response, it usually linked to the changes in hydraulic system, turgor pressure, water potential and stomatal opening, which consequently affect transpiration, photosynthetic activity and other important physiological processes (Farooq et al., 2012). Assessment of psychological traits that limit the plant's productivity due to water deficit and the mechanism of drought tolerance will be crucial to ensure yield stability in a changing climate (Abid et al., 2018). However, in general, this assessment is difficult to realize due to the lack of a simple, rapid and repeatable large-scale screening protocol. For instance, the measurement of water potential by destructive sampling is laborious and time-consuming, which consequently reduces the phenotyping throughput rate.

Hyperspectral sensing is a promising method to realize non-destructive and rapid high-throughput phenotyping. A considerable amount of research has been performed to explore the potential of hyperspectral sensing in plant phenotyping studies such as in physiological and biochemical trait estimation (Silva-Perez et al., 2018; Sun et al., 2019; Feng et al., 2020; Cotrozzi et al., 2020), water stress tolerance (Gerhards et al., 2016; Rehman et al., 2020) plant disease detection (Kuska et al., 2015; Wahabzada et al., 2016) and biomass and yield estimation (Li et al., 2020; Yoosefzadeh-Najafabadi et al., 2021). These works demonstrated that the analysis of the reflectance spectra is promising method for the rapid assessment of multiple plant traits.

The interaction of light with plants differs according to the light frequencies, which are mainly expressed in three regions of the spectrum (Jacquemoud et al., 1996): the visible light (400–700 nm), near-infrared

* Corresponding author at: School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Engineering Campus, Nibong Tebal, Penang, Malaysia.
E-mail address: mohdshahrimie@usm.my (M.S. Mohd Asaari).

https://doi.org/10.1016/j.compag.2022.106806
Received 6 April 2021; Received in revised form 17 February 2022; Accepted 17 February 2022

(NIR) region (700–1100 nm), and the short-wave infrared (SWIR) region (1200–2500 nm). Variations in the leaf optical properties of the reflectance spectrum are linked to specific modification in the structural and biochemical elements of leaves (Jacquemoud and Ustin, 2001). In the visible region, the spectral profile is mainly affected by the leaf pigments that are related to the photosynthetic activity such as chlorophylls, carotenoids, and anthocyanin (Feret et al., 2008). In the NIR part, the reflectance is influenced by the scattering of light within the leaf, which depends on anatomical traits such as the mesophyll thickness and density, and stomata structure (Ustin and Jacquemoud, 2020), while in SWIR region, the reflectance is dominated by water absorption and dry matter (Jacquemoud and Baret, 1990).

Recent studies have demonstrated the potential of hyperspectral sensing for quantifying the effective quantum yield (Jia et al., 2019), water potential (Gonzalez-Dugo et al., 2015; Tosin et al., 2021), stomatal conductance (Jarolmasjed et al., 2018; Vitrack-Tamam et al., 2020), and transpiration rate (Marino et al., 2014; Weksler et al., 2020). Analysis of quantum yield of photosystem II (PSII) has been used in water deficit studies in various plants (Shahenshah, 2010; Ni et al., 2015; Yuan et al., 2016) to understand the photosynthetic behavior of plants. The effect of water deficiency often causes a change in the concentrations of the leave pigments leading to changes in the normal photosynthetic activities of the plants. Additional physiological responses to water deficit conditions are the degradation in water potential and turgor, stomatal closure, which will reduce transpiration and limit gas exchange (Farooq et al., 2012). Reduction in water potential induces stomatal closure when the guard cells surrounding the stomata lose turgor and close the opening, thus reducing stomatal conductance. Reduction in stomatal conductance prevents a further decrease in water potential and subsequently lowers the transpiration rate (Jones, 1998). These physiological traits can be related to the modifications in the leaf internal structure or cell wall composition (Peñuelas et al., 1994), which consequently alter the spectral reflectance (Liu et al., 2004; Tilling et al., 2007).

Retrieval of plant traits from the reflectance spectra has often been achieved through physically-based model inversion (Sun et al., 2018; Shiklomanov et al., 2016; Ali et al., 2016). The inferences made from the retrieval model are generally based on the accepted knowledge of cause-effect relationships embedded in Radiative Transfer Models (RTMs) (Verrelst et al., 2015). In RTM inversion, plant traits such as dry matter, water, and chlorophyll concentration were predicted using optimization techniques and look-up-tables. However, a common problem of this method is the difficulty to handle a situation when similar spectra correspond to multiple combinations of vegetation properties (Combal et al., 2003). Moreover, since all RTMs are defined in terms of reflectance spectra, reflectance data that are transformed by any spectral correction method into a different domain cannot directly be applied for model inversion. This method also lacks flexibility, since the model inversion is limited to only the plant traits that define the RTMs, therefore any vegetation property which was not embedded in the physical model cannot be retrieved.

Alternatively, estimation of plant traits from the reflectance spectra using Machine Learning Regression (MLR) technique is more flexible due to its capability to generate adaptive input–output relationships (Rapaport et al., 2015; Verrelst et al., 2015). Some previous studies have applied Partial Least Squares Regression (PLSR) to develop a relationship between the spectral input and the plant biophysical traits (Ge et al., 2016; Yeh et al., 2016; Weber et al., 2012; Vigneau et al., 2011; Mo et al., 2015; Rapaport et al., 2015; Nguyen and Lee, 2006). This algorithm has gained popularity, as it can deal with the collinearity problem of the high dimensional input variables. Additionally, it performs well, even if the observations are far lower than the predictors. Recently, Kernel Ridge Regression (KRR) and Gaussian Process Regression (GPR), both belonging to the family of non-linear regression methods based on kernels, gained interest in plant traits retrieval (Verrelst et al., 2012; Rivera-Caicedo et al., 2017; Verrelst et al., 2015). Typically, these two methods utilize a kernel-induced feature mapping to cope with non-linear relationships between the biological traits and the observed hyperspectral dataset. These regression techniques have been successfully applied for retrieval of chlorophyll concentration, leaf area index and fractional vegetation cover (Van Wittenberghe et al., 2014).

In this work, the objective was to estimate physiological traits of plants from non-destructive close-range hyperspectral imaging. For this, plant trait estimation models were developed by exploring and validating three recently used MLR algorithms: PLSR, KRR, and GPR. More specific goals were (i) to estimate four targeted physiological traits: water potential, effective quantum yield of photosystem II, stomatal conductance and transpiration rate, and (ii) to apply the developed trait estimation models to a small-scale drought phenotyping study on maize.

2. Materials and Methods

2.1. Targeted Physiological Traits

Physiological responses of plant to water stress are often linked with the modification in water potential, turgor, and stomatal conductance, resulting in reduced transpiration and photosynthetic rates, which will ultimately decrease the plant growth (Farooq et al., 2012). In this work, the following physiological traits were targeted because of their association with drought stress responses: leaf water potential, stomatal conductance, transpiration rate, and effective quantum yield of PSII in the light-adapted state.

Leaf water potential (ψ) is widely accepted as an important measure of a plants water status (Brodribb and Hill, 2000; Boyer, 1968; Elsayed et al., 2011). It is a physical parameter that describes the potential energy, required to move water molecules to a state of free water and is often measured in the unit of pressure. It represents the entire water pressure in a given plant system due to osmotic, matrix, turgor, and gravitational effects (Chavarria and dos Santos, 2012).

Stomatal conductance (g_s) is a physiological trait that relates to the degree of stomatal opening associated with the entering rate of CO_2 or exiting rate of water vapor from the stomata of a leaf. The opening rate of stomatal pores are well-recognized as early responses of a plant to avoid dehydration (Arve et al., 2011). The stomatal action is closely related to the water potential by feedback processes. Reduction in water potential induces turgor loss in the guard cell surrounding the stomata, which reduces the stomata opening and stomatal conductance.

Transpiration rate (T_r) is a physiological trait related to the flow of water liquid from the roots of a plant to the stomata on the leaves, where it is altered into water vapor and is released into the atmosphere (Lee et al., 2018). When the guard cells surrounding the stomatal pores lose turgor, the cells fill the pores and reduce the stomatal opening, and consequently lowering the transpiration rate (T_r) to limit excessive water loss (Bray, 1997; Jones, 1998).

Finally, effective quantum yield (ϕ_{PSII}) is one of the commonly used chlorophyll fluorescence traits in water deficit stress studies to quantify photosystem II (PSII) activity in order to understand the photosynthetic behavior of plants (Shahenshah, 2010; Ni et al., 2015; Yuan et al., 2016). Due to water shortage, plants initiate metabolic modification along with structural rearrangements and functional alteration of the photosynthesis, associated with the stomatal opening, leaf water potential, and relative water content (Lisar et al., 2012).

2.2. Machine Learning Regression

Consider two blocks of observed variables. The first block represents the vector of input variables (i.e., the hyperspectral profile) $\mathbf{X} \in \Re^{n \times N}$, with n the number of hyperspectral signals and N the number of spectral bands of the hyperspectral signals. The second block represents the output or response variables (i.e. the plant parameters) $\mathbf{Y} \in \Re^{n \times M}$, with

M the number of parameters. The ultimate goal of this work was to retrieve the aforementioned plant traits from the hyperspectral reflectance data. To this end, MLR algorithms were used to model the relationship between the reflectance variables and the targeted plant traits. In this work, both linear and non-linear regression methods were explored. The linear regression model was developed using PLSR. This regression method is the most commonly applied algorithm in hyperspectral data analysis for mapping vegetation properties. For developing non-linear regression models, two recent algorithms: KRR and GPR were selected. The following subsections briefly describe the principles of these three algorithms.

2.2.1. Partial Least Squares Regression (PLSR)

PLSR has been developed as a generalization of multiple linear regression and has been used as an alternative solution to ordinary least squares (OLS) regression in ill-conditioned linear regression problems (Rosipal and Krämer, 2005). PLSR is of particular interest because it can cope with data that suffers from strong collinearity (strongly correlated variables), noise and situations where the number of predictor variables is higher than the number of observations (Wold et al., 1984). In PLSR, the input variables $\mathbf{X}$ and the output variables $\mathbf{Y}$ are decomposed as:

$$\begin{aligned} \mathbf{X} &= \mathbf{T}\mathbf{P}^T + \mathbf{E} \\ \mathbf{Y} &= \mathbf{U}\mathbf{Q}^T + \mathbf{F} \end{aligned} \tag{1}$$

where the matrix $\mathbf{Q}$ and matrix $\mathbf{P}$ represent loading matrices, $\mathbf{U}$ and $\mathbf{T}$ are matrices of the latent vectors, $\mathbf{F}$ and $\mathbf{E}$ are the residual matrices. The latent vectors $\mathbf{t}$ and $\mathbf{u}$ are given by:

$$\begin{aligned} \mathbf{t} &= \mathbf{X}\mathbf{W} \\ \mathbf{u} &= \mathbf{Y}\mathbf{C} \end{aligned} \tag{2}$$

where the weight vectors $\mathbf{W}$ and $\mathbf{C}$ are optimized by using the non-linear iterative partial least squares (NIPALS) algorithm (Wold, 1975). The Decomposition of $\mathbf{X}$ and $\mathbf{Y}$ is done in order to maximize the covariance between $\mathbf{T}$ and $\mathbf{U}$ (Manne, 1987), which then gives

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{F} \tag{3}$$

where $\mathbf{B}$ represents the matrix of regression coefficients. Using the relationship between $\mathbf{T}, \mathbf{U}, \mathbf{W}$ and $\mathbf{P}$ (Manne, 1987; Rännar et al., 1994), the regression coefficients $\mathbf{B}$ are given by:

$$\mathbf{B} = \mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}\mathbf{T} = \mathbf{X}^T\mathbf{U}(\mathbf{T}^T\mathbf{X}\mathbf{X}^T\mathbf{U})^{-1}\mathbf{T}^T\mathbf{Y} \tag{4}$$

2.2.2. Kernel Ridge Regression (KRR)

KRR, also known as least squares support vector machine (Suykens and Vandewalle, 1999), is considered to be the most elementary kernel algorithm which formulation is based on OLS regression and ridge regression (Saunders et al., 1998). In ridge regression, the task is to minimize the following squared loss function:

$$J(\beta) = \min_{\beta} \|\mathbf{Y} - \mathbf{X}\beta\|^2 \tag{5}$$

Due to a limitation in the number of training examples, the variance of the estimated β may be large, causing the predicted response to be unreliable. A simple and yet effective way to prevent this, is to introduce a small value λ as a regularizer into Eq. (5) to bargain the bias and variance of the estimate. The above expression becomes:

$$J(\beta) = \min_{\beta} \|\mathbf{Y} - \mathbf{X}\beta\|^2 + \lambda\|\beta\|^2 \tag{6}$$

The closed form solution is then given by (Rakesh and Suganthan, 2017):

$$\boldsymbol{\beta} = \mathbf{X}^T(\mathbf{X}\mathbf{X}^T + \lambda\mathbf{I})^{-1}\mathbf{Y} \tag{7}$$

where I is identity matrix. The prediction for the new data point $\mathbf{x}_*$ is given by:

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

The KRR algorithm extends the above linear algorithm into a non-linear one using a kernel function. The motivation for introducing a kernel is to discover non-linear relationships between variables in the input space by mapping them onto a high-dimensional feature space, thus allowing to learn a linear function in that new space. This is done by mapping the data $\mathbf{x}_i$ in $\mathbf{X}$ onto the feature vector $\mathbf{x}_i \rightarrow \phi(\mathbf{x}_i)$ using a kernel function, $\kappa(\mathbf{x}, \mathbf{x}')$, which simply returns the dot product of the two vectors. The forecast for the new data points is given by:

$$\mathbf{y}_* = \kappa(\mathbf{K} + \lambda\mathbf{I})^{-1}\mathbf{Y} \tag{9}$$

where κ is the row vector of kernel functions between a test sample and the n training points, and $\mathbf{K}$ is a kernel matrix of inner products between the training vectors. In this work, a Gaussian function was applied as the kernel:

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

2.2.3. Gaussian Process Regression (GPR)

Another algorithm that can be applied to model the non-linear input–output relationship is the GPR (Rasmussen, 2004). GPR is a probabilistic method, based on a non-parametric Bayesian approach to solve regression and classification problems with kernels (Verrelst et al., 2012). For regression problems, the output variables can be described as:

$$y = f(\mathbf{X}) + \mathcal{N}(\mu, \sigma_\epsilon^2) \tag{11}$$

where the function $f(\mathbf{X})$ is assumed to be a random variable which follows a particular distribution, reflecting the uncertainty regarding the function, and is assumed to be distributed as a Gaussian Process (GP). It assumes that the training output $\mathbf{Y}$ and the test output $(f(\mathbf{X}_*))$ are jointly Gaussian:

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

where σ_n^2 is the noise variance of the training data, $K(\mathbf{X}, \mathbf{X}_*)$ is the kernel functions between the n training and test samples, and $K(\mathbf{X}_*, \mathbf{X}_*)$ is the kernel functions between the test samples (Koirala et al., 2019; Koirala et al., 2020). The kernel function used in GP provides the covariance of the output:

$$cov(\mathbf{y}_i, \mathbf{y}_j) = k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\left(- \sum_{b=1}^{N} \frac{\left(\mathbf{x}_i^b - \mathbf{x}_j^b\right)^2}{2l_b^2} \right) \tag{13}$$

where σ_f^2 is the variance of the input spectra, and l_b is a characteristic length-scale for each band. The mean prediction of a new input point $\mathbf{X}_*$ can be derived by:

$$\mathbf{Y}_* = K(\mathbf{X}_*, \mathbf{X})(K(\mathbf{X}, \mathbf{X}) + \sigma_n^2\mathbf{I})^{-1}\mathbf{Y} \tag{14}$$

where the hyperparameters of the kernel function in Eq. (13) are optimized by minimizing the log marginal likelihood of the training dataset (Rasmussen, 2004).

2.3. Plant Cultivation

Greenhouse experiments were conducted on maize plants under drought and well-watered conditions to assess the plant physiological status based on its spectral characteristics. Maize plants were grown in

PHENOVISION platform, the indoor plant phenotyping infrastructure located in VIB-UGent Center for Plant Systems Biology (Ghent, Belgium). The maize seeds from the B104 inbred were sown in 7-liter pots filled with 850 g of peat-based soil with osmocote fertilizer (N.V. Van Israel, Belgium). Once the seedling reached V5 vegetative stage (five leaves developed), the plants were fertilized weekly with 40 ml of 200 ppm N Peters Excel CalMag Grower (Everris, Netherlands) solution. The greenhouse environment was initially set to 22–23 °C and vapor pressure deficit to 1–1.2 kPa until the plants reach V5 stage. After V5 stage, the greenhouse environment was adjusted to provide a diurnal gradient, with temperature ranging from 22 °C at night to 28 °C in the afternoon. Artificial lighting was provided in the greenhouse using high-pressure sodium vapor lamps with a 16/8-h day/night light cycle to achieve an average light intensity of 280 μmol m^{-2} s^{-1}.

In the early stage of cultivation, all maize plants received a well-watered treatment (WW). From growth stage V5 onward, the plants were divided into two groups. In the first group, plants were kept at an optimal soil water content of 2.4 g H_2O g^{-1} dry soil (soil water potential of −10 kPa) throughout the course of the experiment. In the second group, plants underwent a progressive drought treatment (PD). In this treatment, watering was stopped for seven days until a soil water content of 1.4 g H_2O g^{-1} dry soil (soil water potential of −100 kPa) was attained, after which the plants received small amounts of water in order to keep the soil water content constant at the lower level. All available plants were further split into two sets, where each set contained both WW and PD plants. The first set consisted of 108 plants (54 WW and 54 PD plants) were used as ground truth for training the regression models. The second set consisted of 20 plants (10 WW and 10 PD plants) were used to test the obtained prediction models in the experiments to detect drought responses in maize.

2.4. Data Recording

A push-broom line-scan visible near-infrared (VNIR) HSI camera (ImSpector V10E, Spectral Imaging, Oulu, Finland) was used for capturing the hyperspectral images of all plants for 10 days (from 3rd −12th February 2017). The VNIR HSI camera was installed in a dedicated enclosed cabin that is equipped with a lift mechanism, a rotating platform, and a horizontal white reference (the spectralon panel with 99% light reflection). The lift and rotating platform were used to position the sample plant at the optimal distance from the camera and at the level of the white reference surface. The exposure time and the moving speed of the camera were set to 20 ms and 112 mm/s, respectively. The illumination in the cabin was provided by lighting frames that move alongside the camera. The lighting frames consist of 3 × 3 equally spaced 35-Watt halogen lamps placed at either side of the camera.

The spectral data of the sample plant consists of top-view image of a single plant (Fig. 1(a)), the white reference surface, and a black reference (the camera with closed shutter). The acquired hypespectral images were radiometrically calibrated to compensate for variation in pixel-to-pixel sensor response and spatial non-uniformities in illumination. The acquired images have 510 × 328 pixels with a spectral sampling interval of 3.1 nm leading to 194 spectral bands ranging from 400 to 1000 nm (Fig. 1(c)). The spatial resolution of the image is approximately 2.35 mm when sample distance is 1.2 m away from the camera. Due to high noise levels below 500 nm and above 850 nm, the spectral range was limited to 500–850 nm for further data processing, leading to 111 spectral bands (Asaari et al., 2019).

The plant pixels were segmented from the background image (Fig. 1 (b)) based on the NDVI threshold (Asaari et al., 2018). Standard Normal Variate (SNV) normalization was applied to reduce the illumination effects (Fig. 1(d)). The clustering algorithm was applied on the SNV normalized spectra to further filter out non-linear variability due to multiple scattering effects. Non-relevant clusters such as pixel belonging to vein and shaded regions were discarded. Each plant then was characterized by a single SNV spectrum by averaging all pixels from the retained clusters (Asaari et al., 2019). Matlab software (R2017b) was used to extract and pre-process the spectral data, and also used to construct the PLSR, KRR and GPR prediction models.

The ground truth data, required for training the regression models was generated right after the hyperspectral image acquisition process. Non-destructive measurements were performed to obtain effective

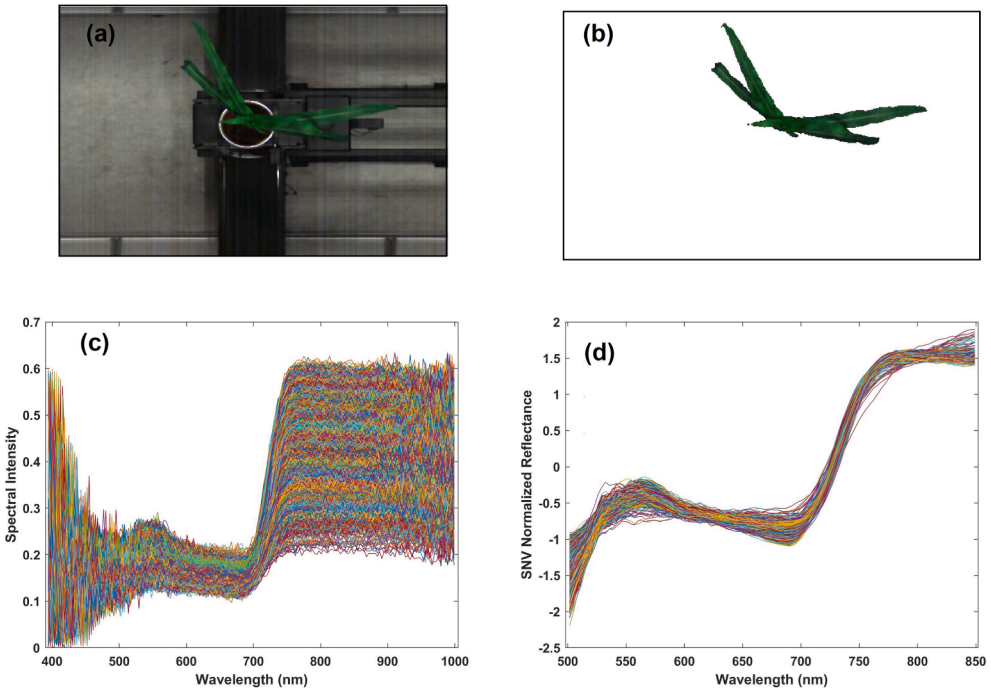

Fig. 1. The Hyperspectral data from a sample plant. (a) The RGB image (for visualization purposes) of the plant sample was reconstructed from the radiometrically calibrated Hyperspectral data. (b) The plant pixels after segmentation. (c) The reflectance spectra of the segmented pixels and (d) The SNV normalized spectra of (c).

quantum yield (ϕ_{PSII}), stomatal conductance of H_2O (g_s) and transpiration rate (T_r). The measurements were performed on Day-0, Day-5, Day-7 and Day-9 of the drought period, four to five times a day on five WW and five PD plants. A portable LI-COR 6400-XT Infrared Gas Analyzer (LI-COR, Lincoln, Nebraska, USA) was used for collecting the gas-exchange and fluorescence measurements. During the measurement, a steady state CO_2 level was set to 400 μmol mol^{-1}, the photosynthetic active radiation flux density was set to 230–360 μmol photons m^{-2} s^{-1} and the greenhouse temperature was maintained between 25–31 °C. The measurement of water potential was performed using a destructive sampling procedure with a pressure chamber (PMS Instrument Company, Albany, Oregon, USA). The destructive measurements were collected three times per day on three to five plants on Day-0, Day-3, Day-5, Day-7 and Day-9. For this destructive sampling, the top leaves (leaf 5–9) were chosen because these leaves were mostly visible in the hyperspectral image. The selected leaves were punched for water potential (ψ) measurements. For more information about these procedures, we refer to Mertens et al. (2021).

3. Results and Discussion

3.1. Model Development and Validation

The model accuracy of the MLR algorithms strongly depends on the selection of the model parameters. For example, the accuracy of the PLSR model is strongly depending on the number of retained latent variables (LV). When this number is insufficient, some useful information will be lost, leading to under-fitting. On the other hand, if too many LVs are retained, the model will risk over-fitting which affects the robustness of the model. For KRR, the estimation model highly depends on the regularization parameter λ that controls the trade-off between bias and variance, and the kernel parameter σ that controls the smoothness of the prediction function. In this work, the optimization of the parameters of PLSR and KRR was done through standard cross-validation. For the cross-validation, each dataset was split into a training set (60%) and an independent validation set (40%). The estimation of the number of LVs (for PLSR), λ and σ (for KRR) was based on the quality of each regression (i.e., lowest normalized root mean squared error (NRMSE)). GPR automatically optimizes the hyperparameters θ, based on the maximization of the marginal likelihood in the training set (Verrelst et al., 2012).

Fig. 2 shows the plots of the average NRMSE as a function of the number of LVs of the PLSR model, for each plant physiological parameter. The average NRMSE and its standard deviation were calculated using 10-fold cross-validation. From these plots, the optimal number of LVs was found to be 5 for water potential and quantum yield efficiency and 7 for stomatal conductance and transpiration rate.

The tuning of the KRR hyperparameters λ and σ was done by cross-validation with a grid search (Exterkate, 2013). The grid for selecting λ was given by the values $\left\{\frac{1}{2^{10}}, \frac{1}{2^9}, \ldots, \frac{1}{2^2}, \frac{1}{2^1}\right\}$ and the grid for σ was $\{1, 2, 3, 4, 5, 6\}$. Fig. 3 shows the 10-fold cross-validation results of the KRR model for the water potential. It can be observed that increasing σ resulted in a global increase of NRMSE for all the λ values. The NRMSE gradually increased for values of λ larger than $\frac{1}{2^5}$. The optimal values were $\sigma = 1$ and $\lambda < \frac{1}{2^5}$. A fine tuning of λ at the selected value of $\sigma = 1$ is presented in Fig. 4, where the optimal value of λ was found to be $\frac{1}{2^7}, \frac{1}{2^8}, \frac{1}{2^9}$ and $\frac{1}{2^{10}}$ for water potential, quantum yield, stomatal conductance and transpiration rate, respectively.

In the next experiment, the aim was to apply the best hyperparameters obtained from the results presented in Fig. 2 - Fig. 4 to search for the best model for each plant trait. In order to reduce skewed results, a 20-fold cross-validation approach with a different random set of training:validation data with a 60:40% ratio was performed during the model training. The regression models were developed independently for each trait, and therefore a different random set was applied in each cross-validation cycle. The random selection of the data instances was performed to ensure that the training and validation sets are

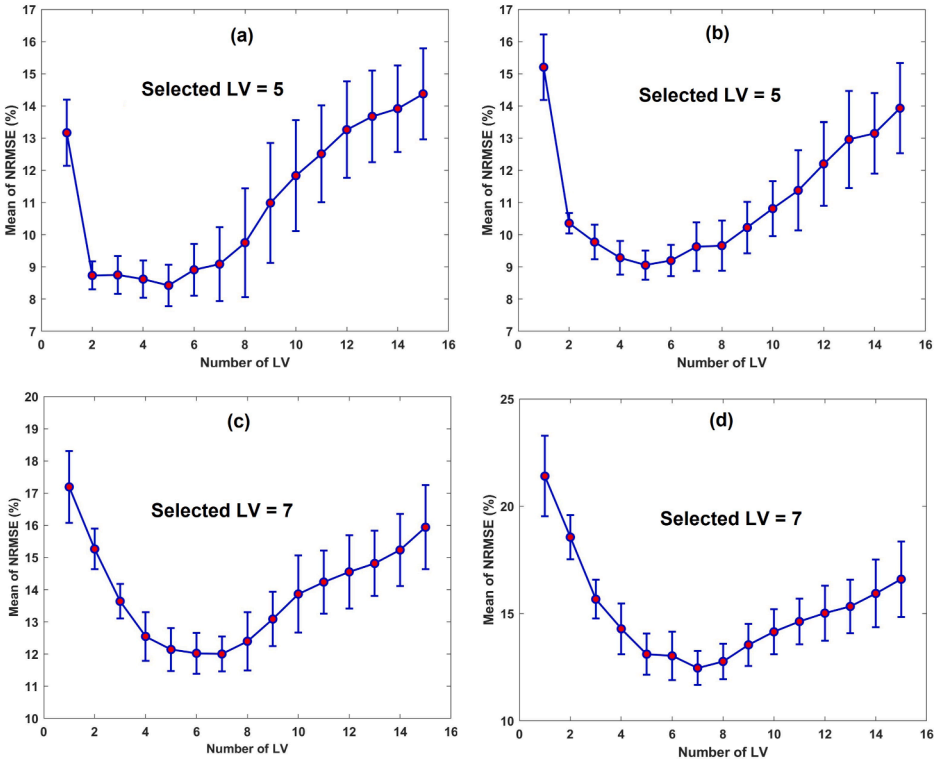

Fig. 2. Estimation of the appropriate number of LVs using PLSR, based on a 10-fold cross-validation for (a) water potential, (b) quantum yield, (c) stomatal conductance, (d) transpiration rate.

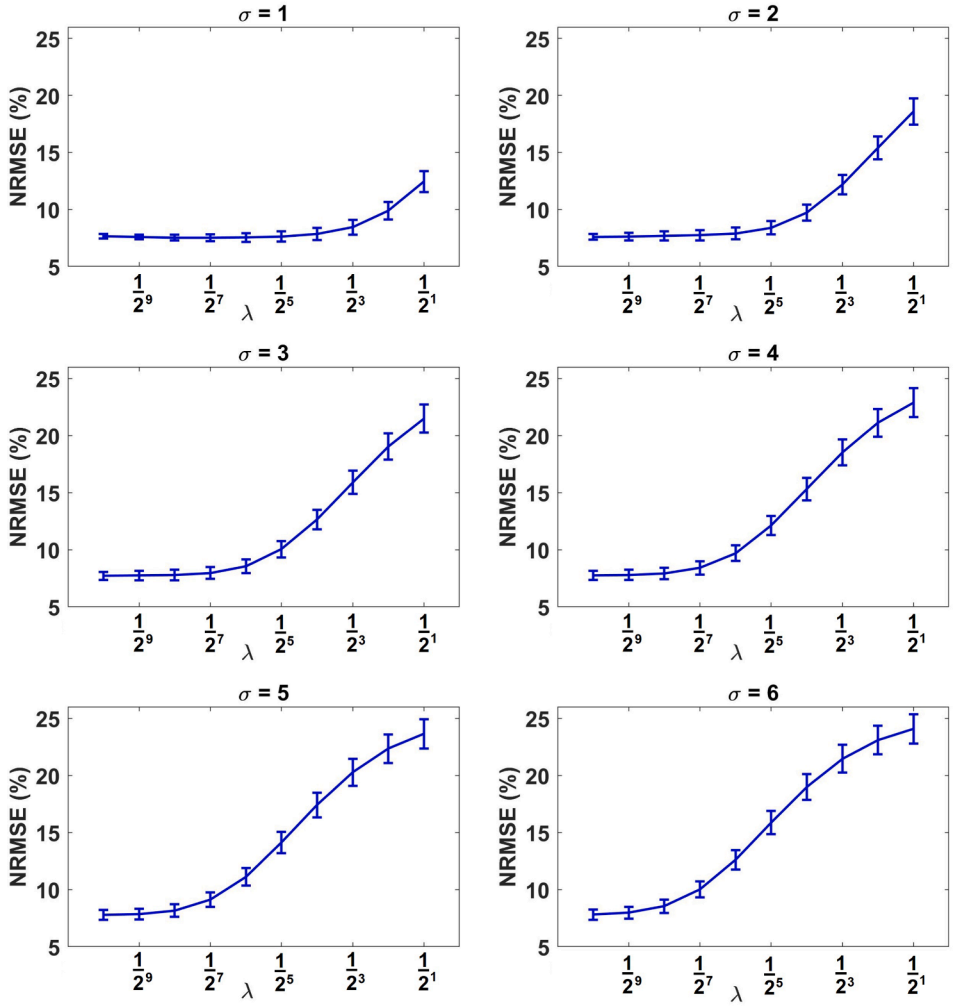

Fig. 3. KRR parameter tuning based on grid search for the water potential case. The plots show NRMSE as a function of λ for different values of σ.

representative of the overall distribution (i.e., from both WW and PD treatments) so that the models remain general and less overfit.

Table 1 summarizes the performance of all regression methods for each plant trait. The average Pearson's coefficient (R^2) and the average NRMSE was calculated to evaluate the performance of each model. For each regression method, the best prediction model was selected based on the lowest NRMSE. Validation of the prediction models for each plant trait, based on the best NRMSE and R^2- values is shown in Fig. 5.

Results in Table 1 demonstrated that all regression methods performed very well for water potential and quantum yield estimation with an average NRMSE, less than 10% and an average R^2 above 0.82. The retrieval of stomatal conductance and transpiration rate showed a slightly lower accuracy, with an average NRMSE between 10–13% and an average R^2 between 0.7–0.79. The non-linear models (i.e., KRR and GPR) outperformed the linear model (i.e., PLSR) for all plant traits. Although the accuracy of PLSR is still considered to be acceptable, its full potential is limited by its restriction to linear attribute relationships, which is probably not always correct in real life scenarios. The good performance of the non-linear models was due to the involvement of the kernel method, which was able to cope with strong non-linear relationships between the plant traits and the input spectrum. Therefore, compared to the linear PLSR algorithm, the KRR and GPR regression methods produced more flexible non-linear input–output relations, thus leading to higher estimation accuracies.

The highest model accuracy (i.e., average NRMSE) was obtained by GPR, while KRR had a slightly lower performance than GPR. This lower performance was probably caused by the suboptimal manual hyperparameter tuning (see Fig. 5). The good performance of GPR due to the following reasons: First, along with predictive mean, the GPR also provide the predictive variance, a measure of uncertainty, which can be important when making predictions far away from the input data. Second, the hyperparameters of the kernel functions in the GPR model can be optimized efficiently by maximizing the marginal likelihood of the training set (Verrelst et al., 2012). Intuitively, this sophisticated framework avoids the time-consuming grid search of the cross-validation procedure and can avoid suboptimal performance in case grids are not set up appropriately. Finally, after optimization, the relevance of each spectral band can be determined from the inverse of the parameter l_b of the kernel function (Eq. (13)) (Verrelst et al., 2012; Verrelst et al., 2013; Van Wittenberghe et al., 2014). Hence, this can gain insight in the most relevant normalized spectral bands.

From the 20 runs of GPR, a histogram plot of the probability occurrence of the top 20 ranked wavelengths with lowest l_b was created, and visualized in Fig. 6. For the estimation of the water potential, important bands appeared in the regions 500–520 nm, 600–680 nm, and 720–800 nm. For quantum yield efficiency, bands in the red (560–720 nm) and the near-infrared (760–820 nm) wavelength regions contributed most to the prediction model. For stomatal conductance, the top performing wavelengths were located in the regions 580–660 nm and 780–820 nm, and for transpiration rate, the predominant bands were in the green (500–560 nm), near the red-edge (720–760 nm) and in the NIR (800–820 nm) wavelength regions.

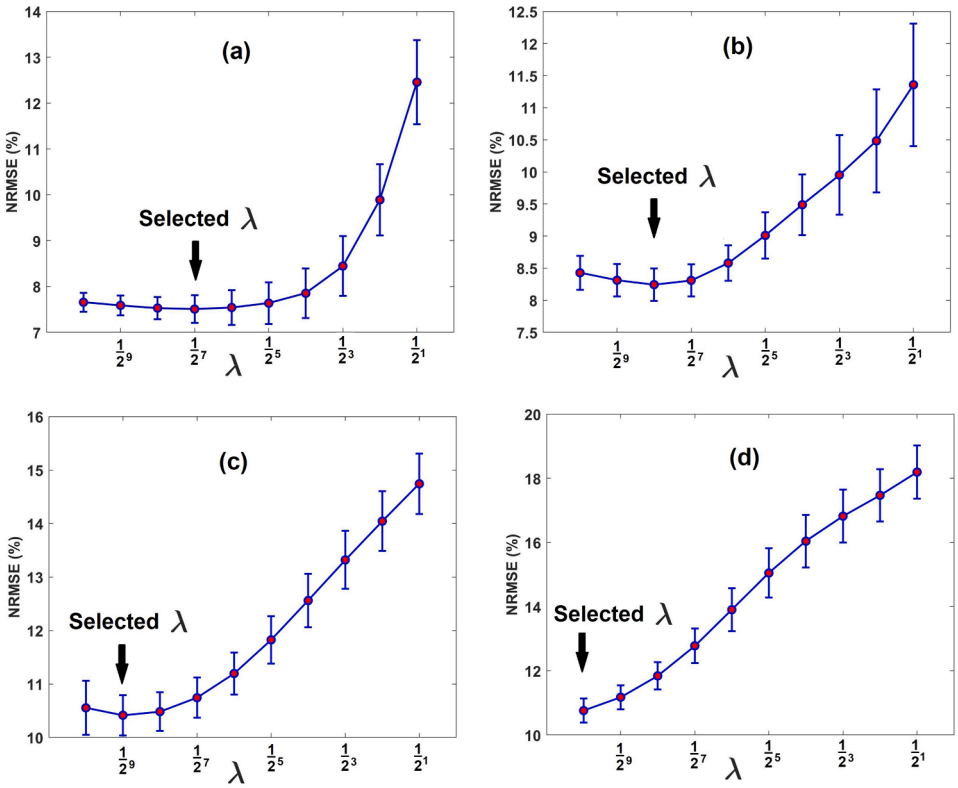

Fig. 4. KRR parameter tuning based on grid search. A fine-tuning for λ at the fixed value of $\sigma = 1$ for (a) water potential, (b) quantum yield, (c) stomatal conductance, (d) transpiration rate.

Table 1
Performance of regression models on plant traits estimation; normalised root mean squared error (NRMSE) and Pearson's determination coefficient (R^2) ± standard deviation for 20-fold cross-validation over the full SNV spectrum.

Plant Parameter	Regression Model	Average NRMSE (%)	Best NRMSE (%)	Average R^2	Best R^2
Water Potential (ψ)	PLSR	8.71 ± 0.681	7.10	0.89 ± 0.020	0.91
	KRR	7.87 ± 0.416	6.65	0.91 ± 0.015	0.92
	GPR	7.84 ± 0.397	6.85	0.91 ± 0.014	0.92
Quantum yield (ϕ_{PSII})	PLSR	9.78 ± 0.45	8.94	0.82 ± 0.025	0.84
	KRR	8.72 ± 0.421	8.10	0.86 ± 0.023	0.87
	GPR	8.62 ± 0.370	7.95	0.86 ± 0.022	0.88
Stomatal conductance (g_s)	PLSR	12.67 ± 0.630	11.30	0.70 ± 0.036	0.78
	KRR	10.82 ± 0.382	9.94	0.78 ± 0.031	0.84
	GPR	10.39 ± 0.531	9.22	0.79 ± 0.030	0.86
Transpiration rate (T_r)	PLSR	13.11 ± 0.689	12.25	0.71 ± 0.029	0.76
	KRR	12.17 ± 0.707	11.26	0.76 ± 0 034	0.82
	GPR	11.69 ± 0.619	10.92	0.77 ± 0.033	0.83

3.2. Drought Stress Detection

In this experiment, the developed regression models were applied in a drought response test. The best non-linear GPR estimation models were considered, because these models had the highest prediction accuracy of all non-linear models. For comparison, the best linear prediction models from PLSR were applied as well (see Fig. 5). Fig. 7 shows the physiological trait predictions of the GPR models during ten days of the drought response test. Each plot visualizes the mean predicted trait values of ten WW and ten PD plants. The standard deviations are given as well to assess the effect of sample size and error variance. For the quantification of the significance difference between the two regimes, one-way ANOVA was performed to reveal the significant differences between each group (Römer et al., 2012). Referring to Table 2, a p-value highlighted in bold indicates the successful separation between WW and PD plants at the corresponding day. Note that the water withholding period of the PD plants started on Day-1, and that from Day-7 onward, the plants were re-watered, albeit to a lower soil water content.

From these results, it can be observed that all four plant traits were able to differentiate between the two regimes at a very early stage of drought induction. For the WW control plants, all of the estimated traits showed an increasing trend during the course of the monitoring period. For the effective quantum yield, the increasing trend was less significant than for the other traits. This can be attributed to the fact that the underlying training data did not show a clear increasing trend (see the plots in the second row of Fig. 5, and most of the values of effective quantum yield of WW plants were concentrated between 0.58 and 0.65). The overall increasing trend of the estimated traits is expected as the plants transitioned from a juvenile to a more mature developmental stage during the time span of the experiment, which is associated with morphological and physiological changes.

For the drought stressed plants, all traits decreased during the water shortage period, as the plants responded to the stress development. Under drought conditions, water movement into the root cells drops.

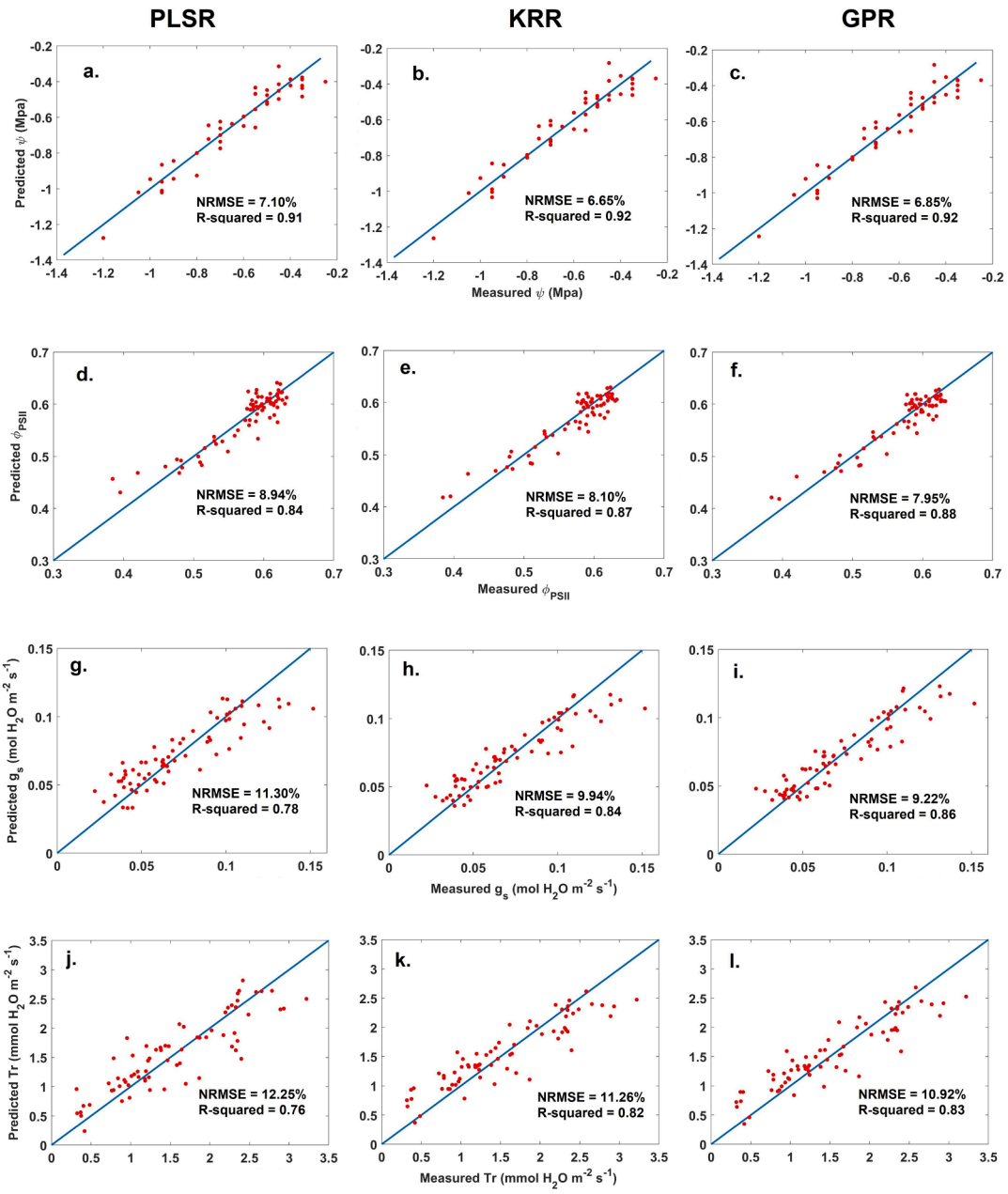

Fig. 5. The measured versus the estimated values of water potential (a-c), effective quantum yield of PSII (d-f), stomatal conductance to H_2O (g-i), and transpiration rate (j-l), based on PLSR (first column), KRR (second column) and GPR (third column). The presented results are the best predictions with lowest NRMSE and highest R^2 from a 20-fold cross-validation (for averages see Table 1).

The absorbed water will be lost due to various metabolic processes, and can not be fully replaced, causing a loss in turgor. As a defensive response, the stomatal aperture closes, thus reducing the transpiration rate. Stomatal closure will also reduce CO_2 uptake, thus decreasing the carbon assimilation rate.

The drought induced decrease in the physiological traits observed in this experiment agrees with results reported in some previous works on maize (Efeoğlu et al., 2009; Sousa et al., 2017; Xu et al., 2008; Grzesiak et al., 2006). Remark that, immediately after re-watering, it can be observed that the predicted traits start to increase again. However, to confirm that the plants adapt to the drought or manage to recover from the stress, more measurements over a longer time span are needed.

To demonstrate the relevance of using non-linear GPR models for plant trait estimation, a comparison with a widely used linear regression technique, the PLSR was performed. Fig. 8 shows the physiological trait predictions of the PLSR models during ten days of the drought response test. For the estimation of water potential and quantum yield, prediction models from both techniques were able to detect drought as early as on the third day of the drought induction. From the *p*-values presented in Table 2, a slight difference was observed on Day-3, where the separation between the two plant regimes was larger for the GPR predictions than for the PLSR predictions. In the case of stomatal conductance and transpiration rate, the discriminating margin between the two groups was lower with PLSR, and a significant discrimination between the 2 regimes was delayed by one day. These results confirm that the use of non-linear GPR models was justified.

The performed experiments were based on the retrieval of vegetation properties from the entire plant. However, it is also interesting to consider the estimation at pixel level. In that case, the trait prediction models need to be applied on each pixel of the high-resolution images.

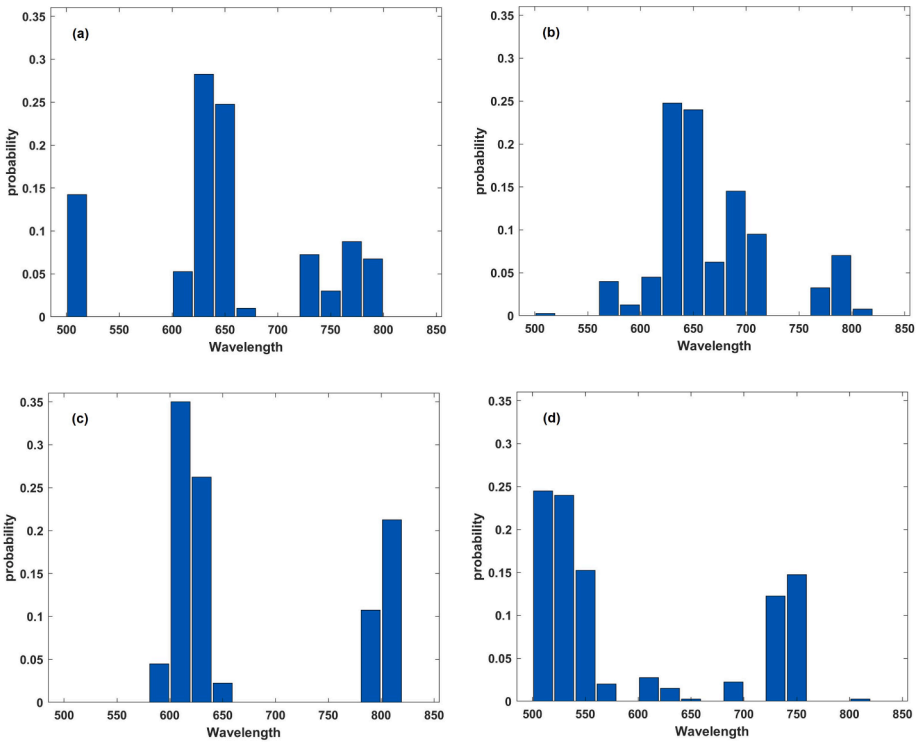

Fig. 6. Probability of 20 top ranked bands with the lowest l_b from 20-fold cross-validation of GPR on (a) water potential, (b) quantum yield, (c) stomatal conductance, (d) transpiration rate.

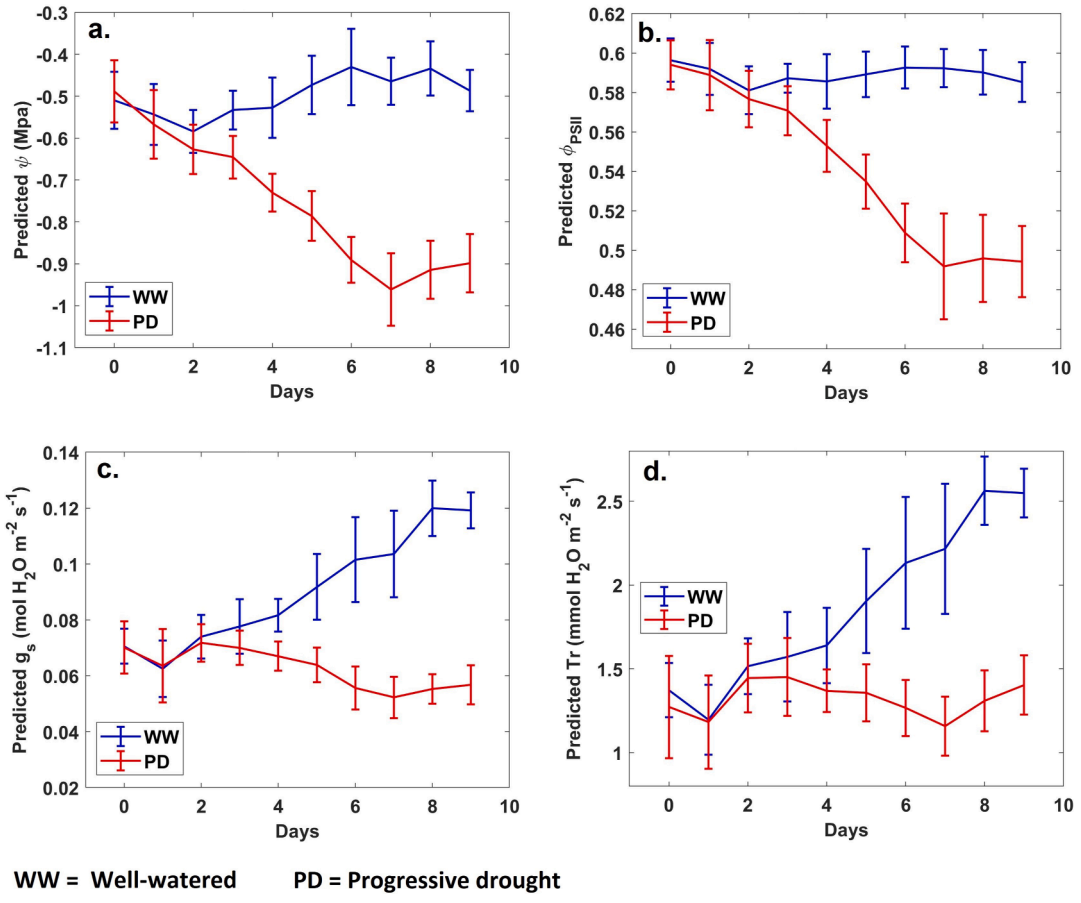

Fig. 7. Drought stress detection results based on the prediction (mean values and standard deviations of ten plants) of (a) water potential, (b) effective quantum yield of PSII, (c) stomatal conductance to H_2O, and (d) transpiration rate using parameter estimation models developed from GPR.

Table 2
The *p*-values of a one-way ANOVA test for the comparison between well-watered and progressive drought plants obtained from GPR and PLSR prediction models on water potential (ψ), effective quantum yield of PSII (ϕ_{PSII}), stomatal conductance (g_s) and transpiration rate (T_r). Bold indicate significance with *p*-value < 0.05 (*), *p*-value < 0.01 (**), *p*-value < 0.001 (***).

Day	GPR				PLSR			
	ψ	ϕ_{PSII}	g_s	T_r	ψ	ϕ_{PSII}	g_s	T_r
0	0.5109	0.6592	0.9038	0.4230	0.8924	0.4790	0.7056	0.4076
1	0.5007	0.6649	0.8340	0.7746	0.4148	0.7027	0.8128	0.8745
2	0.1012	0.4658	0.5007	0.4963	0.1930	0.5972	0.3506	0.4580
3	**0.0001*****	**0.0012****	**0.0491***	0.2884	**0.0003*****	**0.0303***	0.0920	0.3617
4	**0.0000*****	**0.0000*****	**0.0000*****	**0.0183***	**0.0000*****	**0.0017****	**0.0005*****	0.1427
5	**0.0000*****	**0.0000*****	**0.0000*****	**0.0011****	**0.0000*****	**0.0000*****	**0.0001*****	**0.0280***
6	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0001*****
7	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****
8	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****
9	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****	**0.0000*****

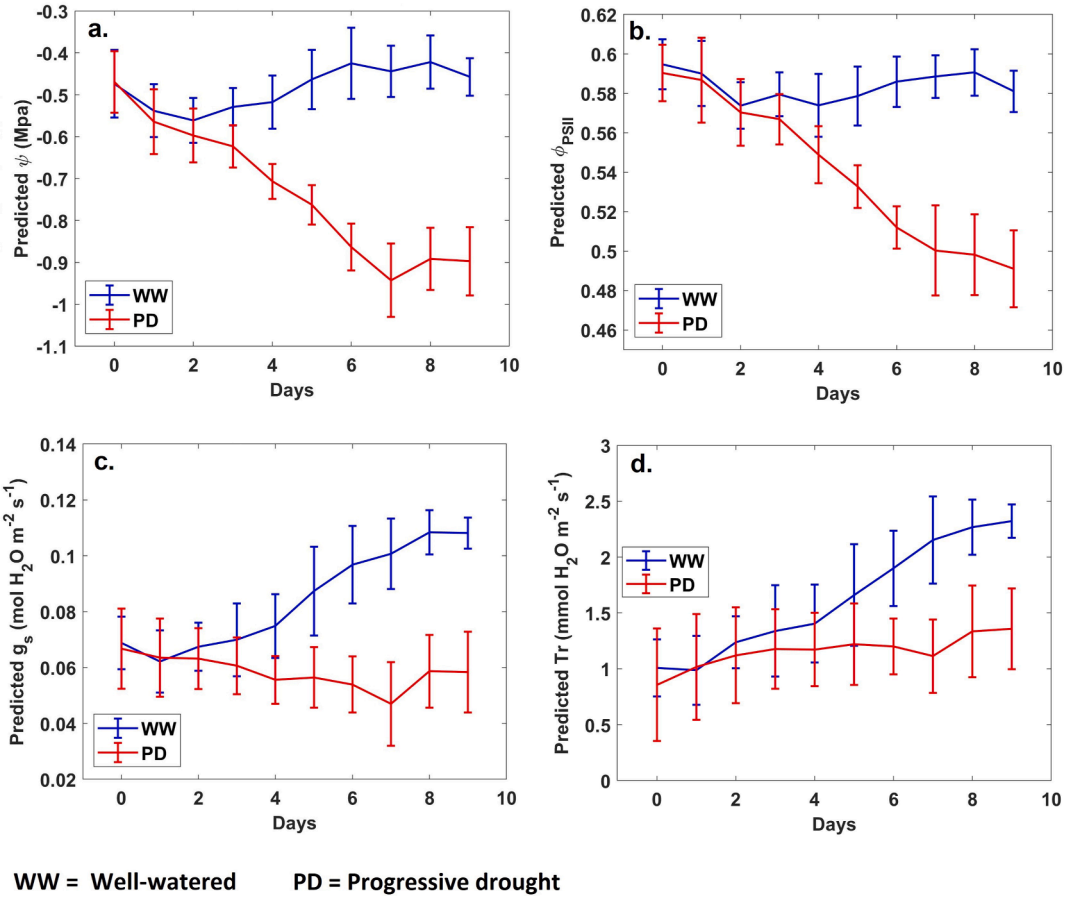

Fig. 8. Drought stress detection results based on the prediction (mean values and standard deviations of ten plants) of (a) water potential, (b) effective quantum yield of PSII, (c) stomatal conductance to H_2O, and (d) transpiration rate using parameter estimation models developed from PLSR.

Fig. 9 shows an example of the pixel-wise estimation of effective quantum yield, obtained from the GPR model. When assigning a false color to the pixels, according to the predicted values of the parameter, the spatial distribution of the estimated effective quantum yield during the drought stress conditions is revealed. Apart from the color mapping, a histogram of the pixel-level predictions for the entire plant can be created to characterize the physiological state of the plant. The histograms in Fig. 8 show that more PD pixels obtained low effective quantum yield values than WW pixels. This observation is in agreement with the results presented in Fig. 7(b). Nevertheless, directly applying the created model, based on the average spectrum of an entire plant may lead to additional errors. We anticipate that the trait values could vary within the leaves area, and therefore perforation size may affect traits measurement. In this case, the development of prediction models should consider the actual pixels that correspond to the location of the trait measurements. This will be an interesting direction to discover in future work, especially if one wants to discover what is the best physiological traits to illustrate within plant variation (i.e., within the leaves area) corresponding to the stress level.

4. Conclusions

For experimental research in the biotechnology domain, rapid and non-destructive measurements of physiological traits is important to

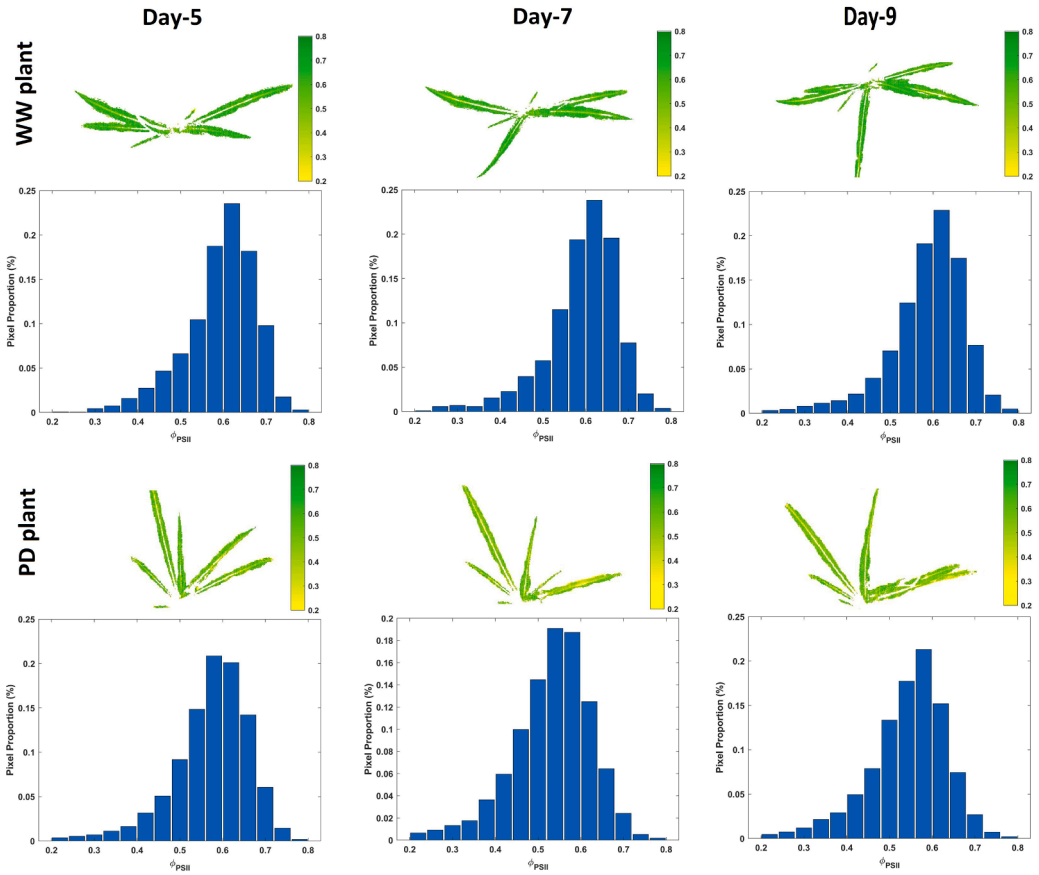

Fig. 9. Visualization of spatial distribution of the predicted effective quantum yield for WW and PD plants, and histograms of the predicted traits. The predicted values were obtained from the GPR model.

maximize the genetic resources for crop improvement. In this work, we provided evidence that physiological traits that are affected by stress can be accurately estimated with rapid, non-destructive HSI measurements. Four physiological traits: water potential, effective quantum yield, stomatal conductance, and transpiration rate were targeted as the proxies for drought stress responses, as plants have different ways and mechanisms to react to drought. Three machine learning regression algorithms: PLSR, KRR, and GPR were explored to determine the relationships between the targeted physiological traits and the SNV normalized hyperspectral profiles. Model validation revealed that the non-linear parameter estimation models, developed based on GPR delivered the best prediction performance. The GPR prediction models were applied to a drought stress study of maize. Statistical significance tests proved that the estimated plant traits were able to distinguish drought from well-watered plants at a very early stage even before visible symptom appear. The obtained results reveal the promising capability of HSI for water stress detection in plants and confirm the full potential for high-throughput and non-destructive phenotyping studies.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study has been conducted by a research collaboration between Imec-Vision Lab and VIB-UGent Center for Plant Systems Biology, and was funded by BASF, Hercules Foundation (ZW1101), and the "Bijzonder Onderzoeksfonds Methusalem Project" (No. BOFMET201 5000201) of Ghent University. The authors also fully acknowledged Ministry of Higher Education Malaysia (MOHE) for Fundamental Research Grant Scheme with Project Code: FRGS/1/2020/TK0/USM/02/13 and Universiti Sains Malaysia for supporting Mohd Shahrimie Mohd Asaari.

References

Abid, M., Ali, S., Qi, L.K., Zahoor, R., Tian, Z., Jiang, D., Snider, J.L., Dai, T., 2018. Physiological and biochemical changes during drought and recovery periods at tillering and jointing stages in wheat (Triticum aestivum L.). Scient. Rep. 8 (1), 4615.

Ali, A.M., Skidmore, A.K., Darvishzadeh, R., van Duren, I., Holzwarth, S., Mueller, J., 2016. Retrieval of forest leaf functional traits from hyspex imagery using radiative transfer models and continuous wavelet analysis. ISPRS J. Photogram. Remote Sens. 122, 68–80.

Arve, L., Torre, S., Olsen, J., Tanino, K., 2011. Stomatal responses to drought stress and air humidity. In: Shanker, A., Venkateswarlu, B. (Eds.), Abiotic stress in plants-Mechanisms and adaptations. InTech, pp. 268–280.

Asaari, M.S.M., Mertens, S., Dhondt, S., Inzé, D., Wuyts, N., Scheunders, P., 2019. Analysis of hyperspectral images for detection of drought stress and recovery in maize plants in a high-throughput phenotyping platform. Comput. Electron. Agric. 162, 749–758.

Asaari, M.S.M., Mishra, P., Mertens, S., Dhondt, S., Inzé, D., Wuyts, N., Scheunders, P., 2018. Close-range hyperspectral image analysis for the early detection of stress responses in individual plants in a high-throughput phenotyping platform. ISPRS J. Photogram. Remote Sens. 138, 121–138.

Boyer, J.S., 1968. Relationship of water potential to growth of leaves. Plant Physiol. 43 (7), 1056–1062.

Bray, E.A., 1997. Plant responses to water deficit. Trends Plant Sci. 2 (2), 48–54.

Brodribb, T.J., Hill, R.S., 2000. Increases in water potential gradient reduce xylem conductivity in whole plants. Evidence from a low-pressure conductivity method. Plant Physiol. 123 (3), 1021–1028.

Chavarria, G., dos Santos, H.P., 2012. Plant water relations: Absorption, transport and control mechanisms. In: Montanaro, G., Dichio, B. (Eds.), Advances in selected plant physiology aspects. InTech, pp. 105–132.

Combal, B., Baret, F., Weiss, M., Trubuil, A., Mace, D., Pragnere, A., Myneni, R., Knyazikhin, Y., Wang, L., 2003. Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem. Remote Sens. Environ. 84 (1), 1–15.
Cotrozzi, L., Peron, R., Tuinstra, M.R., Mickelbart, M.V., Couture, J.J., 2020. Spectral phenotyping of physiological and anatomical leaf traits related with maize water status. Plant physiology 184 (3), 1363–1377.
Efeoğlu, B., Ekmekci, Y., Cicek, N., 2009. Physiological responses of three maize cultivars to drought stress and recovery. South African Journal of Botany 75 (1), 34–42.
Elsayed, S., Mistele, B., Schmidhalter, U., 2011. Can changes in leaf water potential be assessed spectrally? Funct. Plant Biol. 38 (6), 523–533.
Exterkate, P., 2013. Model selection in kernel ridge regression. Computational Statistics and Data Analysis 68, 1–16.
Farooq, M., Hussain, M., Wahid, A., Siddique, K., 2012. Drought stress in plants: An overview. In: Aroca, R. (Ed.), Plant Responses to Drought Stress: From Morphological to Molecular Features. Springer, pp. 1–33.
Feng, X., Zhan, Y., Wang, Q., Yang, X., Yu, C., Wang, H., Tang, Z., Jiang, D., Peng, C., He, Y., 2020. Hyperspectral imaging combined with machine learning as a tool to obtain high-throughput plant salt-stress phenotyping. Plant J. 101 (6), 1448–1461.
Feret, J.-B., François, C., Asner, G.P., Gitelson, A.A., Martin, R.E., Bidel, L.P., Ustin, S.L., le Maire, G., Jacquemoud, S., 2008. Prospect-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. Remote Sens. Environ. 112 (6), 3030–3043.
Ge, Y., Bai, G., Stoerger, V., Schnable, J.C., 2016. Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput rgb and hyperspectral imaging. Computers and Electronics in Agriculture 127, 625–632.
Gerhards, M., Rock, G., Schlerf, M., Udelhoven, T., 2016. Water stress detection in potato plants using leaf temperature, emissivity, and reflectance. Int. J. Appl. Earth Obs. Geoinf. 53, 27–39.
Gonzalez-Dugo, V., Hernandez, P., Solis, I., Zarco-Tejada, P.J., 2015. Using high-resolution hyperspectral and thermal airborne imagery to assess physiological condition in the context of wheat phenotyping. Remote Sensing 7 (10), 13586–13605.
Grzesiak, M., Grzesiak, S., Skoczowski, A., 2006. Changes of leaf water potential and gas exchange during and after drought in triticale and maize genotypes differing in drought tolerance. Photosynthetica 44 (4), 561–568.
Jacquemoud, S., Baret, F., 1990. Prospect: A model of leaf optical properties spectra. Remote Sens. Environ. 34 (2), 75–91.
Jacquemoud, S., Ustin, S., Verdebout, J., Schmuck, G., Andreoli, G., Hosgood, B., 1996. Estimating leaf biochemistry using the prospect leaf optical properties model. Remote sensing of environment 56 (3), 194–202.
Jacquemoud, S., Ustin, S.L., 2001. Leaf optical properties: A state of the art. In: 8th International Symposium of Physical Measurements & Signatures in Remote Sensing, pp. 223–332.
Jarolmasjed, S., Sankaran, S., Kalcsits, L., Khot, L.R., 2018. Proximal hyperspectral sensing of stomatal conductance to monitor the efficacy of exogenous abscisic acid applications in apple trees. Crop Protection 109, 42–50.
Jia, M., Li, D., Colombo, R., Wang, Y., Wang, X., Cheng, T., Zhu, Y., Yao, X., Xu, C., Ouer, G., et al., 2019. Quantifying chlorophyll fluorescence parameters from hyperspectral reflectance at the leaf scale under various nitrogen treatment regimes in winter wheat. Remote Sensing 11 (23), 2838.
Jones, H.G., 1998. Stomatal control of photosynthesis and transpiration. J. Exp. Bot. 49, 387–398.
Koirala, B., Khodadadzadeh, M., Contreras, C., Zahiri, Z., Gloaguen, R., Scheunders, P., 2019. A supervised method for nonlinear hyperspectral unmixing. Remote Sensing 11 (20), 2458.
Koirala, B., Zahiri, Z., Scheunders, P., 2020. A machine learning framework for estimating leaf biochemical parameters from its spectral reflectance and transmission measurements. IEEE Trans. Geosci. Remote Sens. 58 (10), 7393–7405.
Kuska, M., Wahabzada, M., Leucker, M., Dehne, H.-W., Kersting, K., Oerke, E.-C., Steiner, U., Mahlein, A.-K., 2015. Hyperspectral phenotyping on the microscopic scale: Towards automated characterization of plant-pathogen interactions. Plant Methods 11 (1), 1.
Lee, S.J., Kim, J., Kim, H., Ryu, J., 2018. Enhancement of plant leaf transpiration with effective use of surface acoustic waves: Effect of wave frequency. RSC Advances 8 (27), 15141–15148.
Li, B., Xu, X., Zhang, L., Han, J., Bian, C., Li, G., Liu, J., Jin, L., 2020. Above-ground biomass estimation and yield prediction in potato by using uav-based rgb and hyperspectral imaging. ISPRS Journal of Photogrammetry and Remote Sensing 162, 161–172.
Liu, L., Wang, J., Huang, W., Zhao, C., Zhang, B., Tong, Q., 2004. Estimating winter wheat plant water content using red edge parameters. Int. J. Remote Sens. 25 (17), 3331–3342.
Manne, R., 1987. Analysis of two partial-least-squares algorithms for multivariate calibration. Chemometrics and Intelligent Laboratory Systems 2 (1), 187–197.
Marino, G., Pallozzi, E., Cocozza, C., Tognetti, R., Giovannelli, A., Cantini, C., Centritto, M., 2014. Assessing gas exchange, sap flow and water relations using tree canopy spectral reflectance indices in irrigated and rainfed olea europaea l. Environ. Exp. Bot. 99, 43–52.
Mertens, S., Verbraeken, L., Sprenger, H., Demuynck, K., Maleux, K., Cannoot, B., De Block, J., Maere, S., Nelissen, H., Bonaventure, G., Crafts-Brandner, S.J., Vogel, J.T., Bruce, W., Inze, D., Wuyts, N., 2021. Proximal hyperspectral imaging detects diurnal and drought-induced changes in maize physiology. Frontiers in plant science 12, 240.
Mo, C., Kim, M.S., Kim, G., Cheong, E.J., Yang, J., Lim, J., 2015. Detecting drought stress in soybean plants using hyperspectral fluorescence imaging. Journal of Biosystems Engineering 40 (4), 335–344.
Nguyen, H.T., Lee, B.-W., 2006. Assessment of rice leaf growth and nitrogen status by hyperspectral canopy reflectance and partial least square regression. Eur. J. Agron. 24 (4), 349–356.
Ni, Z., Liu, Z., Huo, H., Li, Z.-L., Nerry, F., Wang, Q., Li, X., 2015. Early water stress detection using leaf-level measurements of chlorophyll fluorescence and temperature data. Remote Sensing 7 (3), 3232–3249.
Peñuelas, J., Gamon, J., Fredeen, A., Merino, J., Field, C., 1994. Reflectance indices associated with physiological changes in nitrogen-and water-limited sunflower leaves. Remote sensing of Environment 48 (2), 135–146.
Rakesh, K., Suganthan, P.N., 2017. An ensemble of kernel ridge regression for multi-class classification. Procedia Computer Science 108, 375–383.
Rännar, S., Lindgren, F., Geladi, P., Wold, S., 1994. A PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: Theory and algorithm. J. Chemom. 8 (2), 111–125.
Rapaport, T., Hochberg, U., Shoshany, M., Karnieli, A., Rachmilevitch, S., 2015. Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. ISPRS Journal of Photogrammetry and Remote Sensing 109, 88–97.
Rasmussen, C.E., 2004. Gaussian processes in machine learning. In: Bousquet, O., Von Luxburg, U., Ratsch, G. (Eds.), Advanced Lectures on Machine Learning, Vol. 3176. Springer, pp. 63–71.
Rehman, T.U., Ma, D., Wang, L., Zhang, L., Jin, J., 2020. Predictive spectral analysis using an end-to-end deep model from hyperspectral images for high-throughput plant phenotyping. Computers and Electronics in Agriculture 177, 105713.
Rivera-Caicedo, J.P., Verrelst, J., Muñoz-Marí, J., Camps-Valls, G., Moreno, J., 2017. Hyperspectral dimensionality reduction for biophysical variable statistical retrieval. ISPRS Journal of Photogrammetry and Remote Sensing 132, 88–101.
Römer, C., Wahabzada, M., Ballvora, A., Pinto, F., Rossini, M., Panigada, C., Behmann, J., Léon, J., Thurau, C., Bauckhage, C., et al., 2012. Early drought stress detection in cereals: Simplex volume maximisation for hyperspectral image analysis. Funct. Plant Biol. 39 (11), 878–890.
Rosipal, R., Krämer, N., 2005. Overview and recent advances in partial least squares. In: International Statistical and Optimization Perspectives Workshop Subspace, Latent Structure and Feature Selection. pp. 34–51.
Saunders, C., Gammerman, A., Vovk, V., 1998. Ridge regression learning algorithm in dual variables. In: Proceedings of the 15th International Conference on Machine Learning, pp. 515–521.
Lisar, Seyed Y.S., Rouhollah Motafakkerazad, M.M.H., Rahman, I.M.M., 2012. Water stress in plants: Causes, effects and responses. In: Rahman, I.M.M., Hasegawa, H. (Eds.), Water Stress. InTech.
Shahenshah, Isoda A., 2010. Effects of water stress on leaf temperature and chlorophyll fluorescence parameters in cotton and peanut. Plant Production Science 13 (3), 269–278.
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.
Silva-Perez, V., Molero, G., Serbin, S.P., Condon, A.G., Reynolds, M.P., Furbank, R.T., Evans, J.R., 2018. Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. J. Exp. Bot. 69 (3), 483–496.
Sousa, C.A.F., Paiva, D.S., Casari, R.A. d. C.N., Oliveira, N.G., Molinari, H.B.C., Kobayashi, A.K., Magalhães, P.C., Gomide, R.L., Souza, M.T., 2017. A procedure for maize genotypes discrimination to drought by chlorophyll fluorescence imaging rapid light curves. Plant methods 13 (1), 61.
Sun, D., Cen, H., Weng, H., Wan, L., Abdalla, A., El-Manawy, A.I., Zhu, Y., Zhao, N., Fu, H., Tang, J., et al., 2019. Using hyperspectral analysis as a potential high throughput phenotyping tool in gwas for protein content of rice quality. Plant methods 15 (1), 1–16.
Sun, J., Shi, S., Yang, J., Du, L., Gong, W., Chen, B., Song, S., 2018. Analyzing the performance of prospect model inversion based on different spectral information for leaf biochemical properties retrieval. ISPRS Journal of Photogrammetry and Remote Sensing 135, 74–83.
Suykens, J.A., Vandewalle, J., 1999. Least squares support vector machine classifiers. Neural Process. Lett. 9 (3), 293–300.
Tilling, A.K., O'Leary, G.J., Ferwerda, J.G., Jones, S.D., Fitzgerald, G.J., Rodriguez, D., Belford, R., 2007. Remote sensing of nitrogen and water stress in wheat. Field Crops Research 104 (1–3), 77–85.
Tosin, R., Pôças, I., Novo, H., Teixeira, J., Fontes, N., Graça, A., Cunha, M., 2021. Assessing predawn leaf water potential based on hyperspectral data and pigment's concentration of vitis vinifera l. in the douro wine region. Sci. Hortic. 278, 109860.
Ustin, S.L., Jacquemoud, S., 2020. How the optical properties of leaves modify the absorption and scattering of energy and enhance leaf functionality. In: Cavender-Bares, J., Gamon, J.A., Townsend, P.A. (Eds.), Remote sensing of plant biodiversity. Springer, Cham, pp. 349–384.
Van Wittenberghe, S., Verrelst, J., Rivera, J.P., Alonso, L., Moreno, J., Samson, R., 2014. Gaussian processes retrieval of leaf parameters from a multi-species reflectance, absorbance and fluorescence dataset. J. Photochem. Photobiol., B 134, 37–48.
Verrelst, J., Alonso, L., Camps-Valls, G., Delegido, J., Moreno, J., 2012. Retrieval of vegetation biophysical parameters using gaussian process techniques. IEEE Trans. Geosci. Remote Sens. 50 (5), 1832–1843.
Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G., Moreno, J., 2015. Optical remote sensing and the retrieval of terrestrial vegetation

bio-geophysical properties–A review. ISPRS Journal of Photogrammetry and Remote Sensing 108, 273–290.
Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera, J.P., Camps-Valls, G., Moreno, J., 2012. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for sentinel-2 and-3. Remote Sens. Environ. 118, 127–139.
Verrelst, J., Rivera, J.P., Moreno, J., Camps-Valls, G., 2013. Gaussian processes uncertainty estimates in experimental Sentinel-2 LAI and leaf chlorophyll content retrieval. ISPRS Journal of Photogrammetry and Remote Sensing 86, 157–167.
Verrelst, J., Rivera, J.P., Veroustraete, F., Muñoz-Marí, J., Clevers, J.G., Camps-Valls, G., Moreno, J., 2015. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods–A comparison. ISPRS Journal of Photogrammetry and Remote Sensing 108, 260–272.
Vigneau, N., Ecarnot, M., Rabatel, G., Roumet, P., 2011. Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in wheat. Field Crops Research 122 (1), 25–31.
Vitrack-Tamam, S., Holtzman, L., Dagan, R., Levi, S., Tadmor, Y., Azizi, T., Rabinovitz, O., Naor, A., Liran, O., 2020. Random forest algorithm improves detection of physiological activity embedded within reflectance spectra using stomatal conductance as a test case. Remote Sensing 12 (14), 2213.
Wahabzada, M., Mahlein, A.-K., Bauckhage, C., Steiner, U., Oerke, E.-C., Kersting, K., 2016. Plant phenotyping using probabilistic topic models: uncovering the hyperspectral language of plants. Scientific Reports 6, 22482.
Weber, V.S., Araus, J.L., Cairns, J.E., Sanchez, C., Melchinger, A.E., Orsini, E., 2012. Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes. Field Crops Research 128, 82–90.
Weksler, S., Rozenstein, O., Haish, N., Moshelion, M., Walach, R., Ben-Dor, E., 2020. A hyperspectral-physiological phenomics system: Measuring diurnal transpiration rates and diurnal reflectance. Remote Sensing 12 (9), 1493.
Wold, H., 1975. Path models with latent variables: The NIPALS approach. In: Blalock, H., Aganbegian, A., Borodkin, F., Boudon, R., Capecchi, V. (Eds.), Quantitative Sociology: International Perspectives on Mathematical and Statistical Modeling. pp. 307–357.
Wold, S., Ruhe, A., Wold, H., Dunn III, W., 1984. The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM Journal on Scientific and Statistical Computing 5 (3), 735–743.
Xu, Z., Zhou, G., Wang, Y., Han, G., Li, Y., 2008. Changes in chlorophyll fluorescence in maize plants with imposed rapid dehydration at different leaf ages. J. Plant Growth Regul. 27 (1), 83–92.
Yeh, Y.-H., Chung, W.-C., Liao, J.-Y., Chung, C.-L., Kuo, Y.-F., Lin, T.-T., 2016. Strawberry foliar anthracnose assessment by hyperspectral imaging. Computers and Electronics in Agriculture 122, 1–9.
Yoosefzadeh-Najafabadi, M., Earl, H.J., Tulpan, D., Sulik, J., Eskandari, M., 2021. Application of machine learning algorithms in plant breeding: predicting yield from hyperspectral reflectance in soybean. Frontiers in plant science 11, 2169.
Yuan, X., Yang, Z., Li, Y., Liu, Q., Han, W., 2016. Effects of different levels of water stress on leaf photosynthetic characteristics and antioxidant enzyme activities of greenhouse tomato. Photosynthetica 54 (1), 28–39.