This item is the archived peer-reviewed author-version of:

Close-range hyperspectral image analysis for the early detection of stress responses in individual plants in a high-throughput phenotyping platform

Reference:

Moh Asaari Mohd Shahrime, Mishra Puneet, Mertens Stien, Dhondt Stijn, Inzé Dirk, Wuyts Nathalie, Scheunders Paul, Inze Dirk.- Close-range hyperspectral image analysis for the early detection of stress responses in individual plants in a high-throughput phenotyping platform
ISPRS journal of photogrammetry and remote sensing / ISPRS; International Society for Photogrammetry and Remote Sensing - ISSN 0924-2716 - 138(2018), p. 121-138

Full text (Publisher's DOI): <https://doi.org/10.1016/J.ISPRSJPRS.2018.02.003>

To cite this reference: <https://hdl.handle.net/10067/1504840151162165141>

Close-range hyperspectral image analysis for the early detection of stress responses in individual plants in a high-throughput phenotyping platform

Mohd Shahrimie M. A.^{1,4*}, Puneet Mishra¹, Stien Mertens^{2,3}, Stijn Dhondt^{2,3}, Dirk Inzé^{2,3}, Nathalie Wuyts^{2,3}, Paul Scheunders¹

¹*Imec-Vision Lab, University of Antwerp, Belgium*

²*Ghent University, Department of Plant Biotechnology and Bioinformatics, Ghent, Belgium*

³*VIB Center for Plant Systems Biology, Ghent, Belgium*

⁴*School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Engineering Campus, Nibong Tebal, Penang, Malaysia*

Abstract

The potential of close-range hyperspectral imaging (HSI) as a tool for detecting early drought stress responses in plants grown in a high-throughput plant phenotyping platform (HTPPP) was explored. Reflectance spectra from leaves in close-range imaging are highly influenced by plant geometry and its specific alignment towards the imaging system. This induces high uninformative variability in the recorded signals, whereas the spectral signature informing on plant biological traits remains undisclosed. A linear reflectance model that describes the effect of the distance and orientation of each pixel of a plant with respect to the imaging system was applied. By solving this model for the linear coefficients, the spectra were corrected for the uninformative illumination effects. This approach, however, was constrained by the requirement of a reference spectrum, which was difficult to obtain. As an alternative, the standard normal variate (SNV) normalisation method was

applied to reduce this uninformative variability.

Once the envisioned illumination effects were eliminated, the remaining differences in plant spectra were assumed to be related to changes in plant traits. To distinguish the stress-related phenomena from regular growth dynamics, a spectral analysis procedure was developed based on clustering, a supervised band selection, and a direct calculation of a spectral similarity measure against a reference. To test the significance of the discrimination between healthy and stressed plants, a statistical test was conducted using a one-way analysis of variance (ANOVA) technique.

The proposed analysis techniques was validated with HSI data of maize plants (*Zea mays* L.) acquired in a HTPPP for early detection of drought stress in maize plant . Results showed that the pre-processing of reflectance spectra with the SNV effectively reduces the variability due to the expected illumination effects. The proposed spectral analysis method on the normalized spectra successfully detected drought stress from the third day of drought induction, confirming the potential of HSI for drought stress detection studies and further supporting its adoption in HTPPP.

Keywords: Close-range hyperspectral imaging, linear reflectance model, standard normal variate, spectral similarity measure, plant stress

1. Introduction

In recent years, image sensor technology has led to effective tools for assessing and quantifying plant traits in response to a variety of environmental factors. As imaging is non-invasive, it enables the measurement of traits over time, which is important for monitoring the progression of growth and stress

in individual plants (Fahlgren et al., 2015). Imaging techniques currently used include visible light (RGB), thermal, fluorescence, near-infrared, thermographic, 3D and hyperspectral imaging (Li et al., 2014). Among these modalities, hyperspectral imaging (HSI) is a relatively new technology that combines spatially-resolved data and spectral information (Mishra et al., 2017). HSI is considered as one of the most promising technologies for measuring physiological status and stress-related responses of plants (Matsuda et al., 2012; Mahlein et al., 2012a; Mutka and Bart, 2015). HSI deployed for plant physiological trait assessment is available in PHENOVISION, a newly developed automated and image-based high-throughput plant phenotyping platform (HTPPP) installed in the greenhouse infrastructure of the VIB-UGent Center for Plant Systems Biology (Ghent, Belgium). In this platform, maize plants are grown as a crop model for which a variety of plant traits is studied. For more details regarding PHENOVISION and its technical specifications, readers are advised to visit the web page at <http://www.psb.ugent.be/phenotyping/phenovision>.

In HSI, plant trait information is inferred from the unique profile of reflectance spectra formed by multiple interactions between the plant leaves and electromagnetic radiation from a light source. The reflectance values in the visible spectrum (400-700 nm) are mainly influenced by photosynthetic pigments. In the near infrared region (700-1200 nm), the radiation is primarily reflected due to light scattering with leaf structures, while in the short-wave infrared region (1200-2500 nm), the reflectance is dominated by water absorption and dry matter (Jacquemoud and Baret, 1990). Under stress conditions, plant physiology and structural properties undergo com-

plex changes, which in turn alter the reflectance spectra of leaves (Wahabzada et al., 2016). By applying appropriate spectral analysis, the differences in the reflectance signal can be used to characterise the plant’s physiological state and assess plant genotype-specific responses to biotic and abiotic stresses (Mahlein et al., 2012b; Wahabzada et al., 2015).

In most HSI studies, vegetation indices (VIs) that relate to specific plant parameters such as chlorophyll and water content are commonly derived for quantifying plant traits (Fiorani et al., 2012). VIs are formulated based on the combination of a few single wavelengths, associated with specific biological parameters of the plants. Several close-range studies have successfully applied VIs to characterize plant responses to abiotic and biotic stress factors (Rumpf et al., 2010; Kim et al., 2011; Behmann et al., 2014). Kim et al. (2011) studied drought stress in apple trees by calculating 13 different VIs from the average spectrum of a plant. Their study showed that the normalized difference vegetation index (NDVI) and red-edge NDVI (Sims and Gamon, 2002) were highly correlated with water stress. Rumpf et al. (2010) used the combination of nine VIs, related to different physiological parameters as features in a support vector machine (SVM) model for early detection and classification of plant diseases. Behmann et al. (2014) improved this approach to study drought stress in barley plants. The authors selected an optimized combination of VIs with a high discriminative power and used these as features into their SVM model to distinguish between healthy and stressed plants.

In other studies, the complete reflectance spectrum was used to characterise plant health status (Mahlein et al., 2012b; Römer et al., 2012; Vigneau

et al., 2011; Ge et al., 2016). Mahlein et al. (2012b) performed a pixel-wise mapping of spectral reflectance using a spectral similarity measure called spectral angle mapper (SAM) to detect and quantify disease-specific symptoms at different development stages. Römer et al. (2012) described the full spectrum as a linear combination of archetypes spectra using simplex volume maximisation (SiVM) and applied the extracted coefficients to classify barley plants into healthy and stressed plants. Vigneau et al. (2011) and Ge et al. (2016) built a chemometric model using partial least squares regression (PLSR) to predict nitrogen and leaf water content in cereal plants.

Although different studies show that HSI is a promising technique for quantifying plant responses to stress in an early stage, its potential is not fully realised due to strong influence from uninformative variability due to illumination effects (Behmann et al., 2015a). In particular, before any visible symptoms appear, the progression of stress development is reflected in subtle spectral changes, which interfere with nuisance variability introduced by variations in illumination received and reflected by the plant. The extremely high spatial resolution of HSI in close-range sensing makes the recorded signal very sensitive to specific alignments of the imaging system and the non-solid architecture of the plant. This sensitivity increases further in whole-plant screening scenarios, where the crops are susceptible to complex plant geometry. This complexity induces high unwanted variability, which in turns overlays the signature correlated with plant traits. Therefore, it is important to properly address these variabilities to allow for optimal analysis of the spectral data.

In a recent study, Vigneau et al. (2011) explored the effects of leaf ori-

entation in outdoor HSI settings to improve a prediction model of nitrogen content in wheat plants. They modelled the effects of specular reflection and leaf inclination using linear models, and overcame these effects using the standard normal variate normalisation technique (SNV). Motivated by this work, a model for close-range applications of HSI in indoor settings for a HTPPP is presented here. In this study, a reflectance model is formulated that describes two important sources of uninformative variability: leaf inclinations (i.e. curved and wavy leaf patterns) and differences in leaf distance towards the imaging sensor. The obtained model describes these variabilities as a multiplicative factor that scales the reflectance. Furthermore, differences in leaf inclination cause differences in specular reflection, which is revealed as an offset artefact. Both of these effects can be compensated by utilising a least-squares regression technique. However, one major problem with this strategy is that a prior reference spectrum is required, which is difficult to obtain in practice, especially in the case of whole-plant imaging. For this reason, a common scatter correction method, the SNV normalisation technique, is applied to correct for these effects (Rinnan et al., 2009). To demonstrate the relevance of the developed model, specific illumination experiments are carried out using HSI data of maize plants inside a closed cabin of a HTPPP.

A common approach in the remote sensing domain to extract biological information from spectral signatures is utilising the inversion of radiative transfer models (RTM) (Ali et al., 2016; Shiklomanov et al., 2016; Sun et al., 2018). In RTM inversion, input parameters such as chlorophyll concentration, water content, dry matter and canopy structures are retrieved using look-up-tables and optimisation techniques. This method however is not

suitable in close-range imaging spectroscopy because the physically-based leaf or canopy RTMs are not adapted to the specific illumination problems (Jay et al., 2016). On the other hand, data-driven methods such as machine-learning regression (Rapaport et al., 2015; Rivera-Caicedo et al., 2017; Houborg and McCabe, 2018) and cluster analysis (Behmann et al., 2014; Schmitter et al., 2017) are more flexible because they can easily adapt to any data as they implicitly derive the underlying model distribution from a given dataset. While regression analysis reveals a statistical correlation between the spectral variables and biological parameters, it cannot be applied if the required output variables for training the model are not available. Due to this constraint, in this work, the unsupervised clustering method is preferred as it requires no prior information for making an explicit inference on a given dataset.

In this work, a direct calculation of a spectral similarity measure of the normalised spectra was used to characterise the plant under drought stress. In the first step, a cluster analysis using the k -means algorithm was performed to obtain a pixel-wise mapping for each HSI. Spectra from clusters that correspond to major leaf parts, excluding edge and vein pixels, were considered for further analysis. To quantify the effect of stress on a particular plant, the whole plant was characterised by the average SNV spectrum calculated from all pixels belonging to the selected clusters. A spectral distance function was defined using Euclidean distance (ED) to allow for a relative comparison between stressed and healthy plants. Further, a supervised band selection procedure was applied to extract a small subset of top-scoring variables with highest class separability to improve the discrimination between

the two regimes. The proposed data analysis method was tested on HSI data of ten maize plants, divided equally into two categories: five well-watered control plants and five plants for which watering was stopped in an early vegetative stage.

2. MATERIALS AND METHODS

This section details the methodology for the early detection of water stress responses in maize plants. Figure 1 summarises the steps involved in the implementation of this methodology. The method contains two parts. The first part involves the data pre-processing. The second part performs spectral analysis on the pre-processed data to characterise the temporal dynamics of the plant health status.

2.1. Image acquisition

All hyperspectral images were recorded by the PHENOVISION HTPPP located in the greenhouse of the VIB-UGent Center for Plant Systems Biology (Ghent, Belgium). The platform has a push-broom visible and near-infrared (VNIR) HSI camera (ImSpector V10E, Spectral Imaging, Oulu, Finland) installed in a dedicated enclosed cabin with lift and rotating platform for the automated presentation of the sample plants. The illumination was provided by two lighting racks with nine (3×3) equally spaced 35 W halogen lamps placed at either side and at the same height of the camera. While recording an image, the lighting racks move along with the camera in the push-broom sense to maintain a homogeneous light distribution in the Field of View (FOV) of the camera. The acquired images have 510×328 pixels and a spectral

Figure 1: Flow diagram of the close-range HSI analysis method for early detection of (drought) stress in (maize) plants.

sampling interval of 3.1 nm leading to 194 spectral bands ranging from 400-1000 nm. The spatial resolution of the image depends on the distance to the HS camera. Since we use a binning of 4 pixels and the location of the white reference is 1.2 m to the camera, the spatial resolution of the HS image is approximately 2.35 mm.

Two sets of HS images were prepared for this study. In the first image set, a maize plant was imaged at five different lift heights (-100, 0, 250, 350 and 450 mm). The images from this data set were used to study the lighting effects on the leaf reflectance spectra. In the second set, ten maize plants were followed during an early drought stress response test. The considered

maize plants are in their early growing stage (V5 to V7 stage) with an average number of 9 leaves per plant, an average leaf area of 450 cm² and an average leaf angle of about 60°. The leaf angle and surface area were determined on the basis of RGB image analysis and manual measurements. In the early stage of cultivation, all maize plants were irrigated according to optimal soil humidity conditions. At a particular stage, the plants were divided into two groups. Five plants were not irrigated for six days to reduce soil water content to a lower level. The other batch of five plants was kept under optimal conditions. HS images were captured daily during nine days for both the control and the stressed plants.

2.2. Image pre-processing

To reduce effects of illumination and detector sensitivity, the camera was calibrated with white and dark references. The radiometric calibration corrects flat Lambertian surfaces for variations due to inhomogeneous illumination conditions. Both reference images were acquired during each plant image acquisition. Therefore, the integration time is almost the same. A dark reference image was acquired by the sensor with closed shutter. The white reference was acquired with a horizontal white reference rectangular board (99 % light reflection) positioned flat at a fixed distance from the sensor. Each acquired spectrum was calibrated as:

$$Reflectance = \frac{Image - Dark}{White - Dark} \quad (1)$$

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. Spectra were smoothed using a rational transformation

function (Oppenheim et al., 1999) with a moving window of 12 elements to remove undesirable variations such as spikes during signal recording at the pixel level.

Since only pixels belonging to the plant contain useful information, plants were segmented from the background by eliminating all non-vegetation spectra. Segmentation was performed by calculating the NDVI for each spectrum. The obtained NDVI images were threshold with a NDVI value of 0.3, and the corresponding mask images were used to segment each HSI individually. The pixels extracted from the segmented HSI were used for further data analysis.

2.3. Linking illumination effects with light reflection phenomena

The radiometric calibration by the white reference only compensates the effects of horizontal in-plane inhomogeneous illumination. The reference was always imaged at a fixed height and inclination (flat). However, the imaging sensor does not receive the same amount of light from all regions in the plant. Figure 2 shows an illustration of light interaction with a maize plant in our indoor HSI system. It can be seen that the reflected radiation captured by the imaging sensor is largely affected by the specific alignment of the imaging system and the geometric structure of the plant. Note that the scope of this work is within the HTPPP environment, where the considered plants are in their early growing stage and the imaging is limited to a single maize plant for every image recording session. Therefore, the influence of canopy structure is presumed to be limited by the effects of leaf inclination towards the light source, leaf distance from the sensor as well as the specular reflection effects.

The aforementioned effects are fairly simple to explain with several assumptions regarding the illumination source and the leaf surface. First, we

Figure 2: Illustration of the interaction between radiation and plant in a close-range indoor hyperspectral imaging setup. N is the leaf surface normal, θ is the angle between the incident light direction and the surface normal and d is the leaf distance towards the sensor.

assume that the light from the nine equally-spaced (3×3) halogen lamps placed next to the camera is homogeneously distributed in a 2-dimensional plane of the FOV of the camera, so that a single incident light direction can be assumed, perpendicular to the horizontal plane. Second, by approximating the leaf as a Lambertian surface (i.e. reflected rays scatter more or less equally in all directions), the fraction of leaf reflectance received by the detector is largely affected by the inclination of the leaf towards the light source and the distance towards the sensor. According to Lambert's cosine law, the incoming energy on the diffuse surface area is proportional to the cosine of the angle between the incident light direction and the surface nor-

mal. This relationship can be combined with the inverse square law, stating that the amount of radiant intensity received by the detector from an object is inversely proportional to the squared distance between the object and the sensor. Both effects are independent of the wavelength, and therefore, scale with the reflectance spectrum as follows:

$$\rho_{meas}(\lambda) = \beta \cdot \rho_{ref}(\lambda) \quad (2)$$

where, $\rho_{meas}(\lambda)$ is the measured leaf reflectance, $\rho_{ref}(\lambda)$ is the leaf reflectance at a reference and $\beta = \frac{1}{d^2} \cdot \cos(\theta)$ is a multiplicative term that models the illumination effects due to leaf distance from the sensor and inclination towards the incident light.

Deviations from the Lambertian law also causes specular reflection. This reflection is a surface phenomenon which is almost wavelength independent and can be higher than the diffusive reflection for wavelengths of strong absorption (Bousquet et al., 2005). It is dependent on leaf surface conditions and varies in magnitude for different species dependent on leaf roughness, undulated shape and the characteristics of upper (*adaxial*) and lower (*abaxial*) surface that differ in colour, hairiness, stomata structure and Epicuticular wax amount. According to Vigneau et al. (2011), the contribution of specular reflections can be modelled by an additive term in the measured reflectance, and Equation (2) can be expanded into:

$$\rho_{meas}(\lambda) = \beta \cdot \rho_{ref}(\lambda) + \alpha \quad (3)$$

where α is a scalar constant. Additionally, a wavelength-dependent additive noise term $n(\lambda)$ describing sensor sensitivities, assumed to be random with

white Gaussian noise can be included in Equation (3) leading to:

$$\rho_{meas}(\lambda) = \beta \cdot \rho_{ref}(\lambda) + \alpha + n(\lambda) \quad (4)$$

Equation (4) describes the effects of plant geometry and specific alignment of the imaging system on the obtained leaf reflectance. These effects cause unwanted variability and need to be corrected for. This can be done by inverting Equation (4), leading to:

$$\rho_{corr}(\lambda) = \frac{\rho_{meas}(\lambda) - a}{b} \quad (5)$$

where $a = \alpha$ and $b = \beta$. This inversion compensates each spectrum for specular reflectance by subtracting a term, and for height and inclination variations by a scaling. The values of α and β can be estimated by performing linear regression analysis. However, the problem with this strategy is that a reference spectrum is required, which is hard to obtain in practical situations. Another way to estimate α and β is from the actual position and orientation of each plant location, which requires a detailed 3D shape model of the plant.

An alternative method was proposed to correct for the expected illumination effects using specific spectral pre-treatment techniques. There were two common pre-treatment methods that are suitable to compensate for multiplicative and additive noise terms; the multiplicative scatter correction (MSC) (Geladi et al., 1985) and the SNV (Barnes et al., 1989; Vigneau et al., 2011). Although the mathematical concepts behind MSC and SNV were similar, the MSC method requires a predefined reference spectrum to solve the linear coefficients. Therefore, the SNV method was preferred as it corrects all the spectra independently without any modelling, by the following equation:

$$\mathbf{Z} = \frac{\mathbf{X} - mean(\mathbf{X})}{std(\mathbf{X})} \quad (6)$$

where $\mathbf{X}$ is a row vector containing the original spectrum, $mean(\mathbf{X})$ is the mean of $\mathbf{X}$, $std(\mathbf{X})$ is the standard deviation of $\mathbf{X}$ which is computed over all wavebands and $\mathbf{Z}$ denotes the SNV-transformed spectrum.

It can be easily verified that any two spectra that only differ from each other by a scaling and an offset become equal after SNV transformation. By comparing the SNV transformation to Equation (5), it can be seen that $mean(\mathbf{X})$ takes the role of the specular reflectance, while the effects of inclination and distance towards the detector are eliminated by normalising the spectrum with $std(\mathbf{X})$. One more thing to be noted is that neither the correction method present in Equation (5) or (6) can handle the wavelength-dependent noise term $n(\lambda)$. To reduce the effect of this noise, the measured reflectance was first smoothed before forwarding it to the spectral correction procedure.

Actually, when assuming that the reflectance values are random variables drawn from a certain distribution function, the SNV transformation performs a linear transformation by rescaling these variables into a standard range so that they will have the properties of a standard normal distribution with zero mean and unity standard deviation. While retaining the shape properties of the spectrum, this transformation allows comparing reflectance values from different distributions. The applicability of SNV to handle the expected illumination effects was tested in the experimental part.

2.4. Spectral analysis

Plant responses to stress factors are complex and involve a variety of physiological and biochemical adaptations at cellular and organism levels (Rejeb et al., 2014; Farooq et al., 2012). During the early stage of stress development,

these effects are gradually established within a leaf by very subtle changes, which do not manifest themselves in visible symptoms such as changes in colour or leaf senescence, and therefore are difficult to distinguish from normal variations caused by natural plant growth. The challenge of resolving these effects from normal variation requires an appropriate approach to exploiting the information extracted from the high resolution spatio-temporal spectral data of the plant.

In this section, an analysis method is presented that realises this. It was assumed that two sets of time-series of HSI of plants are available, one set of control plants which follow normal plant growth and another set of plants undergoing some stress. All analyses were performed on the SNV spectra. Some plant parts were assumed to be less relevant for the analysis, such as pixels at the edges or veins of a leaf. Therefore, plant pixels were divided into some distinct groups of pixels, each group containing pixels of similar spectra. For this, clustering of all available SNV spectra, i.e. from all plants over all time instances, was performed. In this work, the k -means clustering algorithm was used, where the appropriate number of clusters k was estimated by means of the Elbow method (Sarstedt and Mooi, 2014; Hackeling, 2014). The Elbow method focuses on reducing the within-cluster error (the cost function) as a result of clustering the data into k groups. As k increases, the average dispersion will decrease monotonically. However, from some value of k on, this decrease starts to flatten markedly, and this point was referred to as the ‘elbow’. In this work, the cost function was defined as:

$$J = |\log(w_k) - \log(w_{k-1})| \quad (7)$$

Where, w_k is the squared ED of a data point to its corresponding cluster centroid, averaged over all data points, given k clusters. The Elbow method then plots the value of the cost function produced by different values of k . Figure 3 shows the implementation of the Elbow method on all maize plants, where it was visually clear that the location of the ‘elbow’ was at $k = 6$, thus suggesting that 6 was the appropriate number of clusters to be used.

Figure 3: Choosing the number of correct cluster based on Elbow method. The point of such an ‘elbow’ occurs at $k = 6$.

After clustering, non-relevant clusters, e.g. corresponding to the edge and vein pixels were manually discarded. However, the clustering procedure in itself was a useful analysis tool. Each plant can be represented by the histogram that contains the proportions of the pixels of the plant belonging to the six different clusters. Any spectral variation over time for a plant

results in a shift in these proportions. As such, the proportion histograms allow for quantitative analysis of the time-series evolution of a plant and can be used to characterise plant growth dynamics or reveal differences between healthy and stressed plants during the concerned period.

In the next step, the difference between normal growth and stress-response was quantified. In this step, each plant was characterised by one SNV spectrum. This spectrum was obtained by averaging the SNV spectra of all pixels belonging to the retained clusters.

In literature, several methods are proposed to quantify biological changes in a plant from its spectral characteristics. These include techniques that directly apply reflectance intensities from the whole spectrum (Mahlein et al., 2012b; Römer et al., 2012), a reconstructed sparse vector from a spectral dictionary (Roscher et al., 2016), and a combination of multiple VIs (Rumpf et al., 2010; Behmann et al., 2014). Since the use of SNV pre-treatment tends to shift the original information elsewhere across the whole spectrum (Fearn, 2009), the calculation of VIs in a similar way will limit the possibility to use VIs for this purpose. Furthermore, the VIs are normally calculated using two or three single bands to relate spectral properties with specific biophysical traits. However, a plant responds to a stress factor in a complex way, which causes changes in different spectral regions spread over the VNIR wavelength interval. VIs may discard significant information which may lead to a decrease in the classification accuracy. Therefore, a technique that uses direct representations from the spectrum was preferred.

In this work, a statistical comparison is proposed between the spectrum of a plant and a reference spectrum by a spectral similarity measure using

distance-metrics (Van Der Meer, 2006). To express the spectral distance between any two spectra $q(\lambda)$ and $r(\lambda)$, the ED, $ED(q, r)$ is used:

$$ED(q, r) = \sqrt{\sum_{\lambda=1}^B (q(\lambda) - r(\lambda))^2} \quad (8)$$

where B the number of bands used.

Measuring spectral distances in HSI data can benefit from higher discrimination accuracy between spectral signatures than in low spectral resolution multispectral data. However, the huge amount of data introduces a challenging problem in the information processing, associated to the high dimensionality of the feature space known as the curse of dimensionality (Hughes, 1968). Thus, a reduction of the dimensionality of hyperspectral data is required so that one can concentrate on the most significant wavelengths in the spectral data to achieve higher discriminant accuracy. For this purpose, a supervised band selection procedure was performed based on a Fisher's statistics (F -value) criterion (Grünauer and Vincze, 2015), in such a way that the selected wavelength bands will optimise the class separability between pre-defined classes (in our case healthy versus stressed). The advantage of this method was that the top-scoring bands could be selected directly from the SNV spectrum without requiring further signal transformation. This may be more appropriate to preserve the biophysical meaning of the variables to be interpreted by a domain expert (Saeys et al., 2007).

Assume that the two classes are stressed and healthy plants, and for each class $i = 1, 2$, a number n_i of plants were available, represented by their SNV spectrum $\rho_{ij}(\lambda), j = 1, \dots, n_i$. For each band, we calculate the F -value:

$$F(\lambda) = \frac{MS_B(\lambda)}{MS_W(\lambda)} \quad (9)$$

where, MS_B reflects the between-group variability, expressed as:

$$MS_B(\lambda) = \sum_{i=1}^2 n_i (\bar{\rho}_i(\lambda) - \bar{\bar{\rho}}(\lambda))^2 \quad (10)$$

where $\bar{\rho}_i(\lambda)$ denotes the sample mean in the corresponding i^{th} class and $\bar{\bar{\rho}}(\lambda)$ denotes the total mean of the data. MS_W refers to the within-group variability, defined as:

$$MS_w(\lambda) = \sum_{ij} (\rho_{ij}(\lambda) - \bar{\rho}_i(\lambda))^2 \quad (11)$$

Each band of the SNV spectrum is then transformed as:

$$\tilde{\rho}(\lambda) = \begin{cases} \rho(\lambda), & \text{if } F(\lambda) \geq T \\ 0, & \text{else} \end{cases} \quad (12)$$

where T is a threshold value.

The transformed SNV spectra can be forwarded to any distance-metric function. In this work, the ED according to Equation (8) was applied to quantify any deviation from regular plant growth dynamics. For this, the dynamics of a plants spectrum was compared against a reference; the average spectrum of all healthy plants at each stage in the evaluation period. In this way, the similarity measure of a healthy plant will only be slightly larger than zero at each time instance. Any dynamics other than the regular plant dynamics of a healthy plant will result in a significant positive value of the similarity measure. The significance of the discrimination result was then tested using a statistical test based on an ANOVA technique. For an evaluation of our proposed method, we performed a comparison with other existing techniques such as SAM, Sequential Forward Selection (SFS) and Sequential Floating Forward Selection (SFFF).

3. RESULTS

3.1. Reducing illumination effects

Two experiments were performed to demonstrate the ability of the SNV to reduce the expected illumination effects. In the first experiment, a single maize plant was imaged at five different lift heights (-100, 0, 250, 350, 450 mm). The reflectance spectra, averaged over the whole plant, are shown in Figure 5(a). The figure demonstrates that the mean reflectance globally increases as the plant gradually moves closer to the sensor. This increase in reflectance was mainly caused by the difference in light intensity received by the detector.

To study the effect of local leaf inclination, the image acquired at 450 mm elevation from the same maize plant was used. A fully grown leaf depicting a local inclination pattern was considered for extracting the spectra. On a manually identified local region, six small patches (6×6 pixels) were defined (see Figure 4), and their corresponding reflectance spectra were extracted and averaged over each patch for further analysis.

Figure 5(b) shows the averaged reflectance spectra extracted from the six patches. The plot shows that the reflectance changes considerably with the local inclinations. At patch R_1 , the leaf is more or less oriented along the horizontal plane. Moving towards patch R_3 , the angle with the horizontal plane grows, resulting in a lower spectral reflectance. Moving onward from patch R_4 to R_6 , the patch surfaces orient progressively back towards a horizontal plane with an increasing reflectance.

In Figure 5(a), it can be observed that when the plant is imaged at different heights, the spectral variability is mainly characterised by scaling.

Figure 4: Six patches (R_1 - R_6) on a selected leaf region with local inclination pattern.

The contribution of specular reflection is angular dependent, varies slowly at a lower incident angle and becomes more prominent at larger angles (Walter-Shea et al., 1989; Bousquet et al., 2005). Since the orientation of the leaf facet remained almost unchanged in the height experiment, the changes in incident angle and the viewing direction were probably minor and therefore the contribution of additive artefacts was very small (see Table 1). On the other hand, in Figure 5(b), in the local leaf inclination experiment, the scaling effect was almost the same as in the height experiment, but the additive artefact was almost 50 times larger (see Table 2). It was also observed that this additive effect was predominant for wavelengths of strong absorption, which is consistent with what has been reported in Bousquet et al. (2005).

To demonstrate the linear model described in Equation (4) and the observation seen in Figure 5, multivariate linear regression was performed independently on the two different data sets. For the regression, a reference spectrum was defined by choosing one of the available mean reflectance spec-

Figure 5: The mean reflectance spectra of a whole plant measured at: (a) five different height and (b) six patches in a local inclination region of Figure 4.

tra from each measurement set. For the elevation experiment, the mean spectrum extracted from the plant at 0 mm was used as the reference, while for the inclination experiment, the mean spectrum extracted from patch R_1 in Figure 4 was used as the reference. The unknown linear coefficients for each spectrum, i.e. α and β , were estimated from ordinary least-squares regression by fitting the rest of the measurements against the reference spectra. The estimated coefficients are summarised in Tables 1 and 2. In the elevation experiment (Table 1), one could observe a gradual increase in β as the leaf moves closer to the camera. In the inclination experiment (Table 2), β decreased with increasing local inclination. With respect to the effect of specular reflection, we refer to the α coefficient in Table 1 and 2. The regression analysis showed that the additive artefacts caused by local inclination effects were far stronger than what was observed for the height effect.

Figure 6(a and c) shows the results of the spectral correction obtained using Equation (5) for the spectra in Figure 5. However, the major problem with this strategy was the need for a reference spectrum. Alternatively, the

Table 1: Regression coefficients for spectral measurements at different plant heights.

Reference	0mm	0mm	0mm	0mm
Measured	-100mm	250mm	350mm	450mm
α	-0.0004	0.0001	0.0002	0.0003
β	0.9356	1.1756	1.2481	1.3111

Table 2: Regression coefficients for spectral measurements at different patches in a local inclination region.

Reference	R ₁	R ₁	R ₁	R ₁	R ₁
Measured	R ₂	R ₃	R ₄	R ₅	R ₆
α	-0.0201	-0.0268	-0.0279	-0.0164	-0.0013
β	0.9315	0.8243	0.8292	0.8860	0.9502

differences in reflectance spectra in both experiments can be eliminated by the SNV normalisation, described in Section 2.3. Figures 6(b) and 6(d) show the SNV spectra from the two experiments. One can see that the spectra almost perfectly overlap each other. To quantify the accuracy of the SNV-transformed spectra, the signal-to-noise ratio (SNR) was used as a measure of signal quality. In this step, the average fitting error for each spectral correction (i.e. the difference between measured spectra against a reference) was calculated and shown in Figure 6. The SNR was then obtained by computing the ratio of sum squared magnitude of the average spectra (excluding the reference) to that of residual. The SNR of the SNV transformed spectra was found to be comparable to the SNR values of the spectral correction in Figure 6 (a and c). Furthermore, the obtained high SNR values for both spectral correction approaches indicate that the unexplained remaining systematics (i.e. the residual due to sensor noise) was relatively low.

Figure 7(a) shows the reflectance spectra from one maize plant (at 0 mm elevation), accumulated over all pixels. The huge variability was mainly

Figure 6: Corrected spectra based on model coefficients from (a) the elevation experiment and (c) the inclination experiment; corrected spectra based on SNV from (b) the elevation experiment and (d) the inclination experiment.

caused by differences in elevation and inclination of the different plant pixels. Figure 7(b) shows the SNV spectra. It can be observed that the SNV normalisation reduced most of the variability observed in the reflectance spectra. Figure 7(c) shows the standard deviation from the spectra in reflectance (Blue) and after the SNV (Red). It can be seen that in reflectance, the standard deviation only captured the global shape of the spectra. This was an indication that the reflectance was hugely affected by the envisioned illumination effects which overlays the information related to particular plant

traits. However, in the standard deviation of the SNV spectra, three clear peaks appeared in the regions between 500-600 nm, 600-700 nm and after 700-750 nm. The spectral profile in these wavelength regions is generally associated with some leaf pigments that control the plant photosynthetic and light use efficiency (Demmig-Adams and Adams, 1996; Sims and Gamon, 2002; Blackburn, 2007).

Figure 7: Reflectance spectra from a single maize plant accumulated over all pixels. Spectra in (a) reflectance, (b) SNV spectra, and (c) standard deviation of reflectance (blue) and SNV (red).

To further demonstrate the effectiveness of the SNV correction at pixel-level, and to demonstrate the usefulness of the clustering procedure, a cluster analysis was performed on the HS images from the elevation experiment. Figure 8 shows the results obtained by globally clustering all pixels of the maize plant imaged at five different heights. The first row represents the RGB images reconstructed from the HS images. As the plant moved towards the camera, the plant appeared bigger. The second row in Figure 8 presents the cluster map and the corresponding proportion histograms obtained from the reflectance clustering based on ED. It can be observed that the majority of the pixels gradually shifted from a lower to a higher cluster label as the

Figure 8: Clustering results of a single plant acquired at four different heights. The first row presents the RGB colour images. The second row presents the reflectance cluster maps and their corresponding histograms using ED. The third row presents the reflectance cluster maps and their corresponding histograms using cosine distance. The fourth row presents the SNV cluster maps and their corresponding histograms using ED.

Figure 9: Results of cluster analysis on local variation from a high contrast region. (a) RGB image, (b) clustering on reflectance using Euclidean distance, (c) clustering on reflectance using cosine distance, and (d) clustering on SNV using Euclidean distance.

elevation of the plant increased. There was no reason why the same plant would have dissimilar cluster distributions at different heights, other than the differences in light intensity received by the plant at different heights. Therefore, it was of utmost importance to homogenise these illumination effects before performing any data processing.

In the remote sensing domain, SAM has been widely used as it is considered to be insensitive to illumination and albedo effects. This algorithm determines the similarity between two non-zero vectors as an inner product that measures the angle between them. The angle calculation uses only the vector direction and not the vector length. In this sense, it is also known as the vector-included cosine distance method which normalises each vector

to a unit length. To compare the accuracy of SNV with this normalisation approach, we performed a k -means clustering based on the cosine distance measure. The third row in Figure 8 presents the cluster map and the corresponding proportion histograms obtained from the clustering of reflectance spectra based on the cosine distance and the fourth row presents the cluster map and the corresponding histograms obtained from the clustering of SNV spectra based on ED. It can be observed that, regardless of the distance variability, the proportion histograms were very similar. Both normalisation approaches seemed to be able to handle global scaling of the signal.

In the next experiment, we further justified the effectiveness of SNV to handle illumination variability in a local region. In this procedure, we considered a local region from two pieces of leaves from the recorded image at 450 mm height as shown in Figure 9. In the RGB image, a high contrast illumination region due to leaf orientation was annotated. From the reflectance cluster maps (Figure 9(b)), it is clear that the assignment of cluster labels was mainly related to illumination effects. Clusters 5 and 6 were strongly associated with higher light intensity regions whereas lower clusters corresponded to lower light intensity regions. When the cosine distance was used (Figure 9(c)), the obtained reflectance cluster maps did not show much improvement over the result of Figure 9(b), except that most of the vein part was consistently classified into a specific cluster label. In the high contrast illumination region, the result still showed huge variation in the cluster assignments. As previously explained in section 2.3, the reflectance observed in this local region was not only spoiled by a scaling effect, but it was also affected by offset artefacts due to specular reflection. The spectral normal-

isation attained by the cosine distance only took care of the scaling effect. Therefore, the obtained cluster map (Figure 9(c)) was still largely affected by illumination effects. On the SNV cluster map in Figure 9(d), it can be observed that leaves were consistently associated with cluster 5 and 6 across the selected region regardless of the light intensity level, the veins were described by cluster 1 and 2, and some parts of the edges were mapped by clusters 3 and 4. One more thing to note is that, from the SNV cluster map in Figure 9(d), the distribution of cluster labels allowed a more significant interpretation related to biological properties of the plants. Although the two leaves visually looked quite similar, their biological characteristics were not the same. Actually, these leaves appear from a different node. The smaller one is the leaf that appeared from the upper node while the bigger leaf was the leaf from the lower node of the stalk. In general, for a maize plant, the leaf from the lower node reaches the maturity state much earlier than the upper one. The leaf from the upper node is much more juvenile and still in the growing stage.

3.2. Early drought stress detection

In this section, the proposed spectral analysis methodology was tested in an experiment on early drought stress detection in maize plants. For this, the progressive stress development of the plants was determined from a series of HSI, during nine days of water shortage. After pre-processing, a global clustering of all available images was performed. Figure 10 shows the results of a cluster analysis performed on a single well-watered and drought maize plant, respectively. The average histogram and the standard deviation of the five plants from both groups are depicted in Figure 11. It can be

Figure 10: Evolution of a single plant for the selected duration of the drought response. The top row presents the SNV cluster maps and their corresponding proportion histograms for a well-watered plant. The bottom row presents the SNV cluster maps and their corresponding proportion histograms for a plant under progressive drought.

Figure 11: Evolution of all plant for the selected duration of the drought response. The top row presents the average histograms for a well-watered plant. The bottom row presents the average histograms for a plant under progressive drought.

observed that for the plant under drought stress, there was a very significant evolution during the drought stress, especially in clusters 4, 5 and 6. Cluster 5, which was negligible in the early days of drought became dominant later on. On the other hand, clusters 4 and 6 which were predominant at the beginning, progressively reduced in proportion at the end of the drought period. This rapid evolution pattern was seemingly related to changes in the traits of the plant under drought stress. The well-watered plant also showed relative similar evolution patterns, but these were much less pronounced and probably due to natural plant growth.

Figure 12: RGB image (a) and cluster map (b) from the SNV spectra of a healthy maize plant.

To quantify these drought-related effects, each plant was characterised using its average SNV spectrum. In this step, clusters 1,2 and 3 were discarded. To explain, the obtained cluster map of one healthy plant is shown in Figure 12. The clusters were arranged based on the Euclidean norm of the cluster centroids in ascending order. One can see and visually compare with the RGB image that clusters 1, 2 and some pixels from cluster 3 were strongly associated with spectra from the leaves edges, and may have been

influenced by the background spectrum. A small number of pixels that reside at the very narrow limb area (the region where the reflectance is strongly influenced by non-linear effects) were also mapped to cluster 3. One can also notice that the remaining spectra associated with cluster 3 corresponded to the vein parts of the leaves. To maximally concentrate on the most relevant information related to possible changes in plant traits, spectra belonging to clusters 4, 5 and 6 were considered. It was already clear from the cluster analysis results (see Figure 10) that most of the transitions during the drought period appeared in these clusters.

By taking as a reference the average of the spectra of the five well-watered plants at any particular day, any plant on the same day could be evaluated against this reference by the proposed similarity measure method. Figure 13 plots the spectral distances of plants against this reference during the whole experiment. Each data point represents the average spectral distance obtained from the five plants for each class, i.e., blue = well-watered (WW); red = progressive drought (PD), on one particular day of the experiment. Standard deviations of these average spectral distances over the five plants are given as well. Figure 13(a) and (b) shows results using the entire spectrum without the band selection procedure. Figure 13(a) shows the results using the original reflectance spectra, while in Figure 13(b), the SNV spectra were used. The results on the reflectance spectra showed a high within-class variability, mainly caused by the illumination effects, which made the distinction between WW plants and plants under drought stress difficult. When the illuminations effects are normalised using SNV, a lower within class variability was obtained, leading to a clear distinction between the two regimes.

The separation between the two regimes could be further enlarged by employing the proposed band selection procedure to select only the top-scoring bands with high discriminative power. The F -values were obtained using Equation (9). The threshold was set empirically as the one that finds a significant difference between the well-watered and drought plants at the earliest stage. Figures 13(c) and (d) show the obtained F -values plotted in function of wavelength for reflectance spectra and SNV spectra respectively. Noticeably, the F -value in SNV is larger than the values obtained in the original reflectance domain, illustrating that the plant traits overlaid by the illumination effects now appear more noticeable in the normalised spectra leading to a greater distinction between drought and healthy plants. Moreover, it is also noted that after normalisation, the spectrum followed a systematic form with several variation peaks occurring in the 600-700 nm, 700-780 nm and 780-800 nm regions. This specific pattern could be linked to the biological processes that change the biophysical properties of the plant under drought stress. In the visible and the red-edge regions, the changes in the reflectance may be related to the modification of pigments that control the plant photosynthetic functioning (Sims and Gamon, 2002; Blackburn, 2007). In the NIR, the changes may be due to light scattering of the internal properties of the cell structure related to leaf thickness, brown pigment, protein and water content (Knippling, 1970; Curran, 1989; Carter and Knapp, 2001; Scoffoni et al., 2014).

Figure 13(e-f) show the spectral distances after using the band selection method in reflectance and SNV respectively. The selected threshold of 70% of the max F -score was used, leading to a selection of 8 and 12 bands for

the reflectance and SNV spectra, respectively. The significance of the separation between healthy and drought plants was quantified using the one-way ANOVA test (Table 3). Results using the band selection had improved over the use of all bands for each day of the experiment. This was especially the case for the SNV spectra, where a significant drought detection was obtained from the third day of drought induction. One more thing to be noted is that after day 7, the obtained spectral distances seemed to stabilise and even decrease again. This was probably because from that day on, the plants were re-watered, and the plants started recovering again.

To further evaluate the performance of our methodology, we performed the comparison with several established approaches. First, the aim was to compare the SNV normalisation procedure with SAM, a popular method used for spectral classification. As explained earlier, the SAM method applies the cosine distance, which normalises the data to a unit length, in this way becoming less sensitive to illumination effects. Figure 14 shows the drought stress detection using SAM applied on the full reflectance and the selected bands, respectively. The p -values of the ANOVA test are given in Table 3. The results using SAM show a higher within-class variability compared to our results in Figure 13 (b) and (f). SAM hardly managed to discriminate the drought stress plants from the healthy plants. This was mainly because SAM only accounts for illumination effects that scale the reflectance, whereas the offset was not regarded. However, using SNV normalisation, both of these effects were significantly reduced, and a clear distinction between healthy and stressed plants was realised.

In the next comparison, the aim was to evaluate and compare the perfor-

Figure 13: Evolution of the spectral distance during the drought stress experiment. The Euclidean distance and standard deviation using all bands on (a) reflectance spectra, (b) SNV spectra; the F -value calculated on (c) reflectance spectra, (d) SNV spectra; ED and standard deviation calculated using selected bands with threshold value set at 70% of the maximum F value on (e) reflectance, and (f) SNV spectra.

Figure 14: Drought stress detection using SAM classification method (Reflectance +SAM) based on (a) full reflectance spectra, (b) selected reflectance bands as used in Figure 13(e)

Figure 15: Drought stress detection based on the implementation of sequential subset selection on SNV spectra using (a) sequential forward selection with Euclidean distance (SNV+SFS+ED) and (b) sequential floating forward selection with Euclidean distance (SNV+SFFS+ED).

mance of the Fisher's criterion subset selection with other well-known feature selection methods. For this, SFS and SFFS were considered. In the Fisher's criterion approach, the relevance of the features was assessed based on the intrinsic characteristic properties of the data. In this case, the score (i.e., F -statistic) was calculated, and low-scoring features were removed. This

Figure 16: Drought stress detection by including water absorption region. (a) the F -value, (b) evolution of the spectral distance during the drought stress experiment.

approach, however, has problems with redundancy, since it risks to mainly select groups of neighbouring bands and to ignore feature interactions with other parts of the spectrum. In sequential selection approaches, the evaluation is obtained by finding the “best” combination of features either based on a classifier’s performances or class separability measures. Figure 15 shows the results of the drought stress detection using the SFS and SFFS methods based on the scatter-matrix evaluation criterion. When compared with our result using the Fisher’s criterion approach in Figure 13 (f) these methods show nearly the same performance (see also the p -value in Table 3). Although the Fisher’s criterion included some redundant attributes, the selection of the optimal threshold (Figure 13(d)) allowed for interactions between high-discriminative features from different parts of the spectrum as well, thus reducing the redundancy effect. In this regard, the use of Fisher’s criterion has additional advantages such as its scalability to high-dimensional datasets, computational efficiency, independence of classification models and low risk

of over-fitting (Saeys et al., 2007).

The presented HSI analysis methodology was tested to study the effects of plant water stress. However, previous literature (Peñuelas et al., 1993, 1996, 1997; Serrano et al., 2000) suggested that a direct assessment of water stress symptoms could be achieved by studying the spectral signature around 900-970 nm, one of the water absorption regions. To test whether the information from this spectral region can improve our earlier results, we reapplied our methodology by considering the spectral range up to 1,000 nm. Figure 16 (a) shows the F -value calculated for this wavelength range. Compared to Figure 13(d), the F -value differed, mainly due to the increase in the spectral dimensionality. Nevertheless, the systematic pattern remained similar, indicating that the locations of the important information did not change. It can also be observed that a slight peak appeared around 950 nm, indicating that there was probably a change related to water status of the plant, but this information was not as dominant as the spectral variations at 600-700 nm, 700-780 nm and the 780-800 nm regions. A possible explanation is that the water content per unit leaf area may not change much due to moderate water stress during the early stage because the plant tries to maintain a level compatible with its basic functioning such as by reducing transpiration rates (Farooq et al., 2012). Moreover, a direct biological control such as a reduction in stomatal conductance may protect the plant from the stress by allowing plant water saving and improving water-use efficiency by the plant (Chaves et al., 2009).

Figure 16 (b) shows the result of the drought stress detection after including the information at the water band region. Note that when using the

threshold method as before, no features from the water band region were selected. As an alternative, the SFS method was used from which 15 bands were selected (one band from 550-600 nm region, four bands from 600-700 nm region, three bands from 700-770 nm region, six bands from 770-900 nm region and one band from 900-970 nm region). As expected, there was no further improvement observed when compared to the earlier results in Figure 13 (f) and Figure 15 (a).

Notice that the spectral correction method applied here did not handle non-linear effects. Such effects may appear due to multiple light scattering. Since these effects are related to structure, either plant canopy structure or internal leaf structure, they may have a large impact on the distinction between healthy and stress plants. Particularly, when there are significant differences in term of leaf inclination, the discrimination between healthy and stress plants may not only be related to biological differences but may also be attributed to plant canopy structure effects. To investigate if the canopy structure effects could affect the discrimination between stressed and healthy plants, the leaf angles of all plants of the drought experiment were determined using RGB imaging analysis. For this, all leaves from the sixth developed leaf and younger were considered. As shown in Figure 17, the leaf angles were calculated in two modes, the angle from the base to the stem and the angle from the tip to the stem. The difference between these two angles was defined as the leaf bending angle and it was used to describe changes in the canopy structure of the plants.

Figure 18 plots the evolution of plant architecture based on the leaf bending angles calculated from the sixth leaf and upwards, and averaged over the

five healthy and stressed plants. It can be observed that there were small differences in leaf bending angle in some leaves, however these occurred mostly at the end of the drought period. This was probably due to a delay in leaf development, especially in leaves 10 and 11 that were smaller in the drought plants, and therefore seemed to bend less as compared to the leaves from healthy plants. The leaves that contributed most to the HSI analysis (the largest ones) were leaves 8 and 9, which only showed small differences between angles towards the end stage. Therefore, we can conclude that at the beginning of the drought development, plant architecture were small and may have only had a minor influence on the discrimination between healthy and drought plants.

Figure 17: Illustration of leaf angle measurements. The leaf base angle was defined as the angle between the bottom part of the leaf (black lines) and the stem. The leaf angle was defined as the angle between the tips (red lines) and the stem.

Figure 18: Mean of leaf bending angles of six leaves over the drought experiment period.

Table 3: P -values of one-way ANOVA at 0.05 significance level

Day	Full spectrum			Selected bands (F -Criterion)			Selected band (Sequential Algorithm)	
	Reflectance+SAM	Reflectance+ED	SNV+ED	Reflectance+SAM	Reflectance+ED	SNV+ED	SNV+SFS+ED	SNV+SFFS+ED
1	0.4508	0.0810	0.1875	0.2461	0.1913	0.1852	0.3293	0.2435
2	0.2910	0.8931	0.4228	0.5717	0.8967	0.6128	0.6691	0.5521
3	0.2392	0.0969	0.0644	0.3216	0.1479	0.0129	0.0281	0.0441
4	0.0233	0.3461	0.0059	0.0406	0.7471	0.0123	0.0055	0.0060
5	0.2273	0.1304	0.0020	0.0056	0.0081	0.0001	0.0008	0.0007
6	0.2487	0.3705	0.0013	0.0779	0.0130	0.0001	0.0008	0.0006
7	0.0424	0.0144	0.0003	0.0170	0.0007	0.0000	0.0001	0.0002
8	0.0003	0.0427	0.0001	0.0001	0.0007	0.0000	0.0001	0.0001
9	0.0002	0.1398	0.0001	0.0001	0.0038	0.0000	0.0001	0.0000

4. DISCUSSION

4.1. Illumination effects

The close-range application of HSI is a complex problem and suffers largely from high unwanted variability mainly arising from illumination effects (Section 2.3). This variability masks the actual biological changes arising from any particular plant trait. In remote sensing, some physical models such as PROSPECT (Jacquemoud and Baret, 1990) and PROSAIL (Jacquemoud et al., 2009) have been developed. The applicability of these models to a close-range HSI system is limited since they cannot take into account the local geometric aspects, i.e. the distance and orientation of each local object in the scene with respect to the light source and the observer. Recently, a new physics-based model (PROCOSINE) was proposed for close-range HSI (Jay et al., 2016). The idea behind PROCOSINE is the integration of specular surface reflection and local leaf orientations into the directional hemispherical reflectance model from PROSPECT. However, the distance from the light source and towards the camera was not taken into account. The model was tested on detached leaves, which were all at the same distance from the light source, and for which the incident light angle was known a priori. However,

in the context of plant phenotyping, the integration of HSI in a HTPPP requires a non-destructive imaging procedure.

In another recent work, the effects of local inclination of leaves on reflectance spectra were studied in an application of sugar beet monitoring (Behmann et al., 2015b). A 3D model of the plant was obtained by a point cloud map from a laser triangulation scanner, and later on Roscher et al. (2016) used this for the radiometric calibration of local inclination effects. However, apart from the HS camera, this approach requires specific sensors to generate high-resolution 3D images. Moreover, the spectral information and the 3D image are captured in different coordinate systems, and a projection of 2D spectral information onto a 3D image plane is required. The generation of the 3D HS model is not straightforward; it requires a proper calibration model and suitable computer vision methods. This approach can only be fully realised if all the sensors have successfully been calibrated in a unique reference coordinate system with sufficient accuracy.

In our approach, the spectral correction was performed based on a normalisation technique, which relies only on the HS data, without the need of a 3D model. In this work, the illumination effects were explained by a linear model that takes into account local leaf orientations, the distance to the detector, and specular surface reflections. The illumination effects are corrected with a scatter correction pre-treatment technique, the SNV transformation, which does not require a common reference spectrum to solve for the linear coefficients. This transform allows to compare reflectance values coming from distinct distributions and is very effective in reducing within-class variation (Barnes et al., 1989). Our experiments confirmed the effectiveness of

the SNV correction method. However, this pre-treatment still presents some disadvantages. Since the SNV is very effective in removing scaling and offset effects, the transformation might remove valuable linear effects. Furthermore, it tends to shift or spread the original information elsewhere across the spectrum (Fearn, 2009). This effect is due to the mean-centred procedure of SNV, in which a change of wavelengths in one direction will cause other wavelengths to change in opposite directions.

In this work, the applied reflectance model does not consider other challenging optical phenomena such as multiple scattering from the surrounding elements. For example, referring to the cluster map in Figure 10, the area in the middle of the plant may be strongly influenced by multiple scattering from the stem and other leaves. In our case, multiple scattering from the plant canopy structure is not seriously affecting the discrimination results because the considered plants are still in the early growing stage, where the canopy architecture is rather simple for a single plant. However, for the case of higher, more mature plants or for applications that focus on groups of multiple plants with dense foliage, the influence of this multiple scattering has to be carefully addressed. In such situation, complex illumination factors will require more descriptive models (i.e. polynomial models) to describe the non-linear behaviour of the phenomena. Furthermore, the integration of such improved models with 3D plant models or the implementation of advanced machine-learning method such as neural networks can also be considered to attain more accurate correction result.

4.2. Early drought stress detection

In plant-related studies using HS data, a common approach to understanding plants responses to environmental factors is to calculate VIs to relate the spectral profiles with biological parameters such as pigments and water content. VIs, however, only consist of a very limited amount of spectral bands. To make use of the wealth of information contained in spectral reflectances, more data-driven analysis approaches may be useful.

In recent studies, (Kersting et al., 2012; Römer et al., 2012), unsupervised learning methods were applied to analyse HS data to uncover drought stress responses in barley plants. In these studies, the archetypal analysis was applied to find the extreme data points to represent the level of senescence. The archetypes were obtained using SiVM, a matrix factorisation technique that computes the low-rank approximations of the input data matrix as convex combinations of extreme data points. Once the archetypes were calculated, the resulting coefficient in the reconstruction matrix was used as a measure of similarity between each spectrum and the archetype classes. To ensure that the obtained archetypes are highly correlated to stress, an input from the expert domain is required to exclude non-biotic spectra and to assign archetypes to correct senescence classes. Behmann et al. (2014) analysed the same data set of barley plants in Römer et al. (2012) using a different data-driven approach. In their work, a combination of VIs, unsupervised clustering and a supervised classification method using ordinal SVM was applied. Drought stress was detected on the 8th day after the drought induction. In that study, no illumination corrections were applied.

In this study, the SNV was applied as pre-treatment procedure to handle

the expected linear illumination effects from the recorded reflectance signal. However, the spectral correction method applied here did not handle multiple scattering effects and therefore they still remained in the spectra after SNV normalisation. Since the considered plants had a simple canopy structure, these influences are assumed to have minor effects on the differentiation between control and drought plants unless there is a significant difference in terms of canopy structure between the two regimes. The changes in canopy structure are probably not the main factor that contribute to the differentiation between healthy and stress plants because at very early stages of drought, these changes are not significantly large (as verified by our measurements of the leaf angles). These effects, however are expected to have a vital contribution under more severe cases of drought stress, since at this stage the plant may experience a critical decline in water potential, largely affecting the rate of leaf growth (Boyer, 1968) and subsequently will induce a significant change in the leaf angles.

On the other hand, the influence of multiple scattering is quite important in the NIR. Given the limited changes in canopy architecture (i.e. leaf angles), we believe that the changes in cluster proportions observed in Figure 10 and 11 may be related to changes in internal leaf structure during dehydration. The contribution of spectral information from the NIR region, other than the water band range, could therefore be more interesting. From the result of the band selection procedure (Figure 16 (a)), we found that the F value at the NIR band (800-850 nm) is much higher than the value at the water band region. This indicates that, other than by water concentration, the leaf reflectance in the NIR region was also determined by another factor, which

we believe to be mainly related to multiple reflection and scattering in the internal cellular structure (Knipling, 1970).

When the NIR radiation passes through the leaf, it is diffused and scattered as it undergoes multiple reflections and refractions at several refractive discontinuities among cell membranes, cell walls, intercellular air spaces, protoplasts, nuclei and stomata (Gausman, 1977). During leaf dehydration, changes in the internal structure may occur as a consequence of a reduction of leaf conductance and loss of cell turgor pressure. The physiological impact of leaf dehydration on leaf structure has been observed in previous studies through leaf shrinkage in both area and thickness (Tang and Boyer, 2007; Sancho-Knapik et al., 2011; Scoffoni et al., 2014). Leaf shrinkage may consequently alter the morphological characteristics in the internal structure such as the mesophyll thickness and density, cell wall elasticity, intercellular air-space area and stomata size and density (Franks and Beerling, 2009; Niinemets et al., 2009; Sancho-Knapik et al., 2011; Galmes et al., 2013). Due to these morphological adjustments, many of the reflective interfaces are eliminated as internal air space is reduced and cell walls come together after cell contents shrink away from the interior cell wall. A new arrangement of leaf tissue is formed after reorientation, which modified their radiation-diffusing capacity and thus accounts for the changes in the leaf reflectance profile (Knipling, 1970).

The buildup of biological stress symptoms due to water deficit can also be observed from other physiological traits such as a reduction in photosynthetic and transpiration rate (Miyashita et al., 2005) and declining leaf water potential (Correia et al., 2006), which mainly due to stomatal closure and

other metabolic impairment (Tezara et al., 1999; Farooq et al., 2012). A direct linkage of the spectral information to these physiological traits can be obtained using a statistical correlation, for example using regression methods. However, in this paper, no actual measurements of these traits were performed. Therefore, our stress detection result cannot quantitatively be related to the aforementioned biological factors. This will be an interesting research direction for us in the future.

5. CONCLUSION

Present study demonstrates the potential of non-destructive close-range HSI under controlled conditions in a HTPPP for early drought stress detection in plants. The proposed method accounts for uninformative variability caused by the interaction of light with the non-solid architecture of the plant structure. This variability induces multiplicative and additive effects on the reflectance spectra which can be explained by a linear model by considering local leaf orientations and distance towards the sensor. Experimental results show that these effects can be effectively reduced by using the SNV transformation. We further developed a data analysis method to recover the biophysical information from the SNV spectra. The combination of unsupervised clustering and a supervised band selection procedure with the use of a distance metric has shown to effectively reveal the stress effects from the spectral information at an early stage. This method was applied in a drought stress experiment on maize plants and revealed a clear distinction between the drought plants and the controls as early as the third day of drought induction. This demonstrates that close-range HSI has a high possibility to

become a rapid and non-destructive novel technology for high throughput phenotype studies. Remark that, although the stress detection was achieved by analysing the mean spectrum of plants, the clustering analysis to select relevant plant pixels took spatial information into account. In the future, we will further investigate the potential of this HSI technology to detect plant stress at the pixel-level scale and extend this work to study different biophysical parameters in plants. The presented spectral analysis method is general, and whenever there is an interest for monitoring the plant dynamics at a whole-plant scale, it can be applied to different types of systemic stress.

Acknowledgments

The research presented in this paper is funded by the BAHAMAS project of the Flemish research center IMEC and partly supported by the Academic Staff Training Scheme (ASTS) of the Universiti Sains Malaysia and the Ministry of Higher Education of Malaysia in supporting Mr. Mohd Shahrimie Mohd Asaari for his research residency at Imec-Vision Lab University of Antwerp, Belgium.

References

- Ali, A. M., Skidmore, A. K., Darvishzadeh, R., van Duren, I., Holzwarth, S., Mueller, J., 2016. Retrieval of forest leaf functional traits from hypspx imagery using radiative transfer models and continuous wavelet analysis. *ISPRS Journal of Photogrammetry and Remote Sensing* 122, 68–80.
- Barnes, R. J., Dhanoa, M. S., Lister, S. J., 1989. Standard normal variate

- transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* 43 (5), 772–777.
- Behmann, J., Mahlein, A.-K., Paulus, S., Kuhlmann, H., Oerke, E.-C., Plümer, L., 2015a. Calibration of hyperspectral close-range pushbroom cameras for plant phenotyping. *ISPRS Journal of Photogrammetry and Remote Sensing* 106, 172–182.
- Behmann, J., Mahlein, A.-K., Paulus, S., Kuhlmann, H., Oerke, E.-C., Plümer, L., 2015b. Generation and application of hyperspectral 3d plant models. In: Agapito L., Bronstein M., Rother C. (Ed.), *Computer Vision - ECCV 2014 Workshops. Lecture Notes in Computer Science*, vol 8928. Springer, pp. 117–130.
- Behmann, J., Steinrücken, J., Plümer, L., 2014. Detection of early plant stress responses in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing* 93, 98–111.
- Blackburn, G. A., 2007. Hyperspectral remote sensing of plant pigments. *Journal of experimental botany* 58 (4), 855–867.
- Bousquet, L., Lachérade, S., Jacquemoud, S., Moya, I., 2005. Leaf brdf measurements and model for specular and diffuse components differentiation. *Remote Sensing of Environment* 98 (2), 201–211.
- Boyer, J. S., 1968. Relationship of water potential to growth of leaves. *Plant physiology* 43 (7), 1056–1062.
- Carter, G. A., Knapp, A. K., 2001. Leaf optical properties in higher plants:

- linking spectral characteristics to stress and chlorophyll concentration. *American journal of botany* 88 (4), 677–684.
- Chaves, M. M., Flexas, J., Pinheiro, C., 2009. Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Annals of Botany* 103 (4), 551–560.
- Correia, M. J., Osório, M. L., Osório, J., Barrote, I., Martins, M., David, M. M., 2006. Influence of transient shade periods on the effects of drought on photosynthesis, carbohydrate accumulation and lipid peroxidation in sunflower leaves. *Environmental and experimental botany* 58 (1), 75–84.
- Curran, P. J., 1989. Remote sensing of foliar chemistry. *Remote sensing of Environment* 30 (3), 271–278.
- Demmig-Adams, B., Adams, W. W., 1996. The role of xanthophyll cycle carotenoids in the protection of photosynthesis. *Trends in Plant science* 1 (1), 21–26.
- Fahlgren, N., Gehan, M. A., Baxter, I., 2015. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Current opinion in plant biology* 24, 93–99.
- 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 Press, Berlin-Heidelberg, pp. 1–33.
- Fearn, T., 2009. The effect of spectral pre-treatments on interpretation. *NIR news* 20 (6), 15–16.

- Fiorani, F., Rascher, U., Jahnke, S., Schurr, U., 2012. Imaging plants dynamics in heterogenic environments. *Current opinion in biotechnology* 23 (2), 227–235.
- Franks, P. J., Beerling, D. J., 2009. Maximum leaf conductance driven by co₂ effects on stomatal size and density over geologic time. *Proceedings of the National Academy of Sciences* 106 (25), 10343–10347.
- Galmes, J., OCHOGAVÍA, J., Gago, J., ROLDÁN, E. J., Cifre, J., CONESA, M. À., 2013. Leaf responses to drought stress in mediterranean accessions of *solanum lycopersicum*: anatomical adaptations in relation to gas exchange parameters. *Plant, cell & environment* 36 (5), 920–935.
- Gausman, H. W., 1977. Reflectance of leaf components. *Remote sensing of Environment* 6 (1), 1–9.
- 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.
- Geladi, P., MacDougall, D., Martens, H., 1985. Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Appl. Spectrosc.* 39 (3), 491–500.
- Grünauer, A., Vincze, M., 2015. Using dimension reduction to improve the classification of high-dimensional data. In: 39th Annual Workshop of the Austrian Association for Pattern Recognition (OAGM 2015), arXiv:1505.01065v1.

- Hackeling, G., 2014. Mastering Machine Learning with scikit-learn. Packt Publishing Ltd, pp. 124–127.
- Houborg, R., McCabe, M. F., 2018. A hybrid training approach for leaf area index estimation via cubist and random forests machine-learning. *ISPRS Journal of Photogrammetry and Remote Sensing* 135, 173–188.
- Hughes, G., 1968. On the mean accuracy of statistical pattern recognizers. *IEEE transactions on information theory* 14 (1), 55–63.
- Jacquemoud, S., Baret, F., 1990. Prospect: A model of leaf optical properties spectra. *Remote sensing of environment* 34 (2), 75–91.
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., Franois, C., Ustin, S. L., 2009. PROSPECT + SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment* 113 (SUPPL. 1), S56–S66.
- Jay, S., Bendoula, R., Hadoux, X., Féret, J.-B., Gorretta, N., 2016. A physically-based model for retrieving foliar biochemistry and leaf orientation using close-range imaging spectroscopy. *Remote Sensing of Environment* 177 (Supplement C), 220–236.
- Kersting, K., Wahabzada, M., Römer, C., Thureau, C., Ballvora, A., Rascher, U., Léon, J., Bauckhage, C., Plümer, L., 2012. Simplex distributions for embedding data matrices over time. In: Joydeep Ghosh, Huan Liu, I. D. C. D., Kamath, C. (Eds.), *Proceedings of the 2012 SIAM International Conference on Data Mining*. SIAM, pp. 295–306.

- Kim, Y., Glenn, D. M., Park, J., Ngugi, H. K., Lehman, B. L., 2011. Hyperspectral image analysis for water stress detection of apple trees. *Computers and Electronics in Agriculture* 77 (2), 155–160.
- Knipling, E. B., 1970. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote sensing of environment* 1 (3), 155–159.
- Li, L., Zhang, Q., Huang, D., 2014. A review of imaging techniques for plant phenotyping. *Sensors* 14 (11), 20078–20111.
- Mahlein, A.-K., Oerke, E.-C., Steiner, U., Dehne, H.-W., 2012a. Recent advances in sensing plant diseases for precision crop protection. *European Journal of Plant Pathology* 133 (1), 197–209.
- Mahlein, A.-K., Steiner, U., Hillnhütter, C., Dehne, H.-W., Oerke, E.-C., 2012b. Hyperspectral imaging for small-scale analysis of symptoms caused by different sugar beet diseases. *Plant methods* 8, 3.
- Matsuda, O., Tanaka, A., Fujita, T., Iba, K., 2012. Hyperspectral imaging techniques for rapid identification of arabidopsis mutants with altered leaf pigment status. *Plant and Cell Physiology* 53 (6), 1154–1170.
- Mishra, P., Asaari, M. S. M., Herrero-Langreo, A., Lohumi, S., Diezma, B., Scheunders, P., 2017. Close range hyperspectral imaging of plants: A review. *Biosystems Engineering* 164 (Supplement C), 49 – 67.
- Miyashita, K., Tanakamaru, S., Maitani, T., Kimura, K., 2005. Recovery responses of photosynthesis, transpiration, and stomatal conductance in kid-

- ney bean following drought stress. *Environmental and experimental botany* 53 (2), 205–214.
- Mutka, A. M., Bart, R. S., 2015. Image-based phenotyping of plant disease symptoms. *Frontiers in plant science* 5, 734.
- Niinemets, Ü., Díaz-Espejo, A., Flexas, J., Galmés, J., Warren, C. R., 2009. Role of mesophyll diffusion conductance in constraining potential photosynthetic productivity in the field. *Journal of Experimental Botany* 60 (8), 2249–2270.
- Oppenheim, A. V., Schafer, R. W., Buck, J. R., 1999. *Discrete-time Signal Processing* (2Nd Ed.). Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
- Peñuelas, J., Filella, I., Biel, C., Serrano, L., Save, R., 1993. The reflectance at the 950–970 nm region as an indicator of plant water status. *International journal of remote sensing* 14 (10), 1887–1905.
- Peñuelas, J., Filella, I., Serrano, L., Save, R., 1996. Cell wall elasticity and water index ($r_{970} \text{ nm}/r_{900} \text{ nm}$) in wheat under different nitrogen availabilities. *International Journal of Remote Sensing* 17 (2), 373–382.
- Peñuelas, J., Pinol, J., Ogaya, R., Filella, I., 1997. Estimation of plant water concentration by the reflectance water index wi (r_{900}/r_{970}). *International Journal of Remote Sensing* 18 (13), 2869–2875.
- 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.

- Rejeb, I. B., Pastor, V., Mauch-Mani, B., 2014. Plant responses to simultaneous biotic and abiotic stress: molecular mechanisms. *Plants* 3 (4), 458–475.
- Rinnan, Å., van den Berg, F., Engelsen, S. B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry* 28 (10), 1201–1222.
- 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., Thureau, C., Bauckhage, C., et al., 2012. Early drought stress detection in cereals: simplex volume maximisation for hyperspectral image analysis. *Functional Plant Biology* 39 (11), 878–890.
- Roscher, R., Behmann, J., Mahlein, A.-K., Dupuis, J., Kuhlmann, H., Plümer, L., 2016. Detection of disease symptoms on hyperspectral 3d plant models. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* III-7, 89–96.
- Rumpf, T., Mahlein, A.-K., Steiner, U., Oerke, E.-C., Dehne, H.-W., Plümer, L., 2010. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture* 74 (1), 91–99.

- Saeys, Y., Inza, I., Larrañaga, P., 2007. A review of feature selection techniques in bioinformatics. *bioinformatics* 23 (19), 2507–2517.
- Sancho-Knapik, D., Álvarez-Arenas, T. G., Peguero-Pina, J. J., Fernández, V., Gil-Pelegrín, E., 2011. Relationship between ultrasonic properties and structural changes in the mesophyll during leaf dehydration. *Journal of Experimental Botany* 62 (10), 3637–3645.
- Sarstedt, M., Mooi, E., 2014. Cluster analysis. In: *A Concise Guide to Market Research*. Springer Berlin Heidelberg, pp. 273–324.
- Schmitter, P., Steinrücke, J., Römer, C., Ballvora, A., Léon, J., Rascher, U., Plüemer, L., 2017. Unsupervised domain adaptation for early detection of drought stress in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing* 131, 65–76.
- Scoffoni, C., Vuong, C., Diep, S., Cochard, H., Sack, L., 2014. Leaf shrinkage with dehydration: Coordination with hydraulic vulnerability and drought tolerance. *Plant Physiology* 164 (4), 1772–1788.
- Serrano, L., Ustin, S. L., Roberts, D. A., Gamon, J. A., Penuelas, J., 2000. Deriving water content of chaparral vegetation from aviris data. *Remote Sensing of Environment* 74 (3), 570–581.
- 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 Sensing of Environment* 183, 226–238.

- Sims, D. A., Gamon, J. A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote sensing of environment* 81 (2), 337–354.
- 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.
- Tang, A.-C., Boyer, J. S., 2007. Leaf shrinkage decreases porosity at low water potentials in sunflower. *Functional plant biology* 34 (1), 24–30.
- Tezara, W., Mitchell, V., Driscoll, S., Lawlor, D., 1999. Water stress inhibits plant photosynthesis by decreasing coupling factor and atp. *Nature* 401, 914–917.
- Van Der Meer, F., 2006. The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. *International journal of applied earth observation and geoinformation* 8 (1), 3–17.
- Vigneau, N., Ecartot, 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.
- Wahabzada, M., Mahlein, A.-K., Bauckhage, C., Steiner, U., Oerke, E.-C., Kersting, K., 2015. Metro maps of plant disease dynamics automated mining of differences using hyperspectral images. *PLoS One* 10 (1), e0116902.
- 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, Article number: 22482.

Walter-Shea, E., Norman, J., Blad, B., 1989. Leaf bidirectional reflectance and transmittance in corn and soybean. Remote Sensing of Environment 29 (2), 161–174.