

DETECTION OF PLANT RESPONSES TO DROUGHT USING CLOSE-RANGE HYPERSPECTRAL IMAGING IN A HIGH-THROUGHPUT PHENOTYPING PLATFORM

Mohd Shahrime Mohd Asaari^{*1,4}, Stien Mertens^{2,3}, Stijn Dhondt^{2,3}, Nathalie Wuyts^{2,3}, Paul Scheunders¹

¹Imec-Vision Lab - University of Antwerp, Belgium

²Department of Plant Systems Biology, Vlaams Instituut voor Biotechnologie (VIB), Gent, Belgium

³Department of Plant Biotechnology and Bioinformatics, Ghent University, Gent, Belgium

⁴School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Penang, Malaysia

ABSTRACT

The detection and characterization of physiological processes in crop plants under water-limited conditions is essential for the selection of drought-tolerant genotypes and the functional analysis of related genes. Close-range hyperspectral imaging (HSI) is a promising, non-invasive technique for sensing of plant traits, and has the potential to detect plant responses to water deficit stress at an early stage. The present study describes a data analysis method to realize this potential. Reflectance spectra of plants in close-range imaging are highly influenced by illumination effects. Standard normal variate (SNV) was applied to reduce linear illumination effects, while non-linear effects were filtered by discarding the affected pixels through a clustering procedure. Once the illumination effects were eliminated, the remaining differences in plant spectra were assumed to be related to changes in plant traits. To quantify stress-related spectral dynamics, a spectral analysis procedure was developed based on a supervised band selection and a direct calculation of a spectral similarity measure against a reference. The proposed method was tested on HSI data of maize plants acquired in a high-throughput plant phenotyping platform for assessment of drought stress responses and recovery after re-watering events. Results show that the spectral analysis method successfully detected the drought stress responses at an early stage and consistently revealed the recovery effects shortly after the re-watering period.

Index Terms— Drought stress, Close-range hyperspectral imaging, Clustering, Spectral distance

1. INTRODUCTION

In recent years, close-range hyperspectral imaging (HSI) has emerged as a promising tool for plant trait assessment in high-throughput plant phenotyping platforms (HTPP) [1]. The goal of HSI within the HTPP is the non-invasive measurement of plant traits based upon reflectance spectra, which permits the

assessment of plant performance over development and in response to stress conditions, such as drought. This advance encourages studies to understand how plants respond to drought stress, and to assess the plant's capability to adapt and recover from the stress. Knowledge gained in this area is crucial for the further improvement of crop drought-tolerance in breeding programs.

Drought stress leads to a wide range of physiological and biochemical responses in plants. This includes the modification in photosynthetic pigment composition, stomatal conductance, surface and internal leaf structure which in turn alter the leaf optical properties. A common approach for plant trait estimation based upon spectral data is to utilize vegetation indices. These are formulated based on the combination of a few single wavelengths, and associated with specific or broad biological traits and processes in plants (reference on review of VIs). However, the complex physiological effects of drought stress cause changes in the reflectance in many different spectral regions. The use of vegetation indices may discard significant information which may lead to a decrease in the discrimination accuracy [2].

Alternatively, data-driven methods have recently been introduced. An important advantage of these methods is their flexibility to easily adapt to any data as they implicitly derive the underlying model distribution from a given dataset. Within this framework, machine learning regression algorithms are usually applied for retrieving plant biophysical variables from the full-spectrum information [3]. Typically, a flexible learning model is established from a training dataset by optimizing the estimation error of the variables extracted. While regression analysis reveals a statistical correlation between the spectral variables and biological information, it cannot be applied if the required output variables for training the model are not available. Due to this constraint, an unsupervised data-driven method is preferred in this work as it requires no prior information for making an explicit inference on a given dataset.

In close-range settings, the reflectance from leaves are highly influenced by the non-solid orientation of the plant

^{*} Corresponding author: MohdShahrime.MohdAsaari@uantwerpen.be

and the specific alignments of the imaging system. To reduce these external effects, the plant spectra were normalized using the standard normal variate (SNV) transformation. However, this transformation does not handle non-linear effects induced by multiple scattering and shading, effects which that are expected to become important in larger plants as the canopy structure becomes more complex. To further filter this non-biological variability, clustering was performed. Spectra from clusters which are affected by this non-linearity (e.g. shadowed and partially occluded areas) were discarded. To quantify the dynamics of plant water-deficit stress, the whole plant was characterized by the average normalized spectrum calculated from all pixels belonging to the retained 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 separation to improve the discrimination between healthy and stressed regimes. The proposed analysis was tested on a series of HSI of maize plants monitored during their whole vegetative period, where the tested plants were divided into four groups: one group was well-watered and the other three groups were treated with different water stress conditions.

2. MATERIAL AND METHODS

2.1. Data acquisition

A batch of maize plants was grown in PHENOVISION, the high-throughput plant phenotyping (HTPP) infrastructure located at VIB, Ghent, Belgium. The plants were divided into four groups with different watering treatments (Figure 1). All treatments started at the seedling level. In the well-watered (WW) treatment, seven plants were irrigated with sufficient water to keep the soil water content at the optimal level of 2.4 (g H₂O/g dry soil) throughout the entire developmental period. In the progressive drought (PD) treatment, seven plants received a WW treatment from the beginning (seedling) until they reached the V5-stage (five leaves developed). At the V5-stage, the plants were not irrigated for seven days after which they were re-watered with a low amount of water to maintain the soil water content at a deficit level of 1.4 (g H₂O/g dry soil) until the end of the developmental period. In the severe drought (SD) treatment, four plants were irrigated with a deficit soil water content throughout the developmental period. In the severe drought with early vegetation stage re-watering (SD-RW) treatment, six plants received the SD treatment from the beginning until they reached the V7-stage (seven leaves developed). From this stage onward, the plants were irrigated with the WW treatment until the end of the developmental period.

For all plants involved, hyperspectral images were acquired daily (a 50-day time series) from growth stage V2

(two leaves developed), about 2 weeks after the start of the water treatments using a line scan push-broom VNIR-HS camera (ImSpector V10E, Spectral Imaging, Oulu, Finland), which amounts to a total of 1200 hyperspectral images. The acquired images have 510 × 328 pixels and a spectral sampling of 3.1 nm which corresponds to 194 bands ranging between 400-1000 nm. All images were radiometrically calibrated by subtracting a dark frame and calculating reflectance relative to a white reference. 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. The plant pixels were segmented from the background using the normalized difference index (NDVI) with a threshold of 0.3. Due to the high sensitivity of the reflectance spectra to illumination variability in close range setting, Standard Normal Variate (SNV) normalization was applied to reduce for external effects [4].

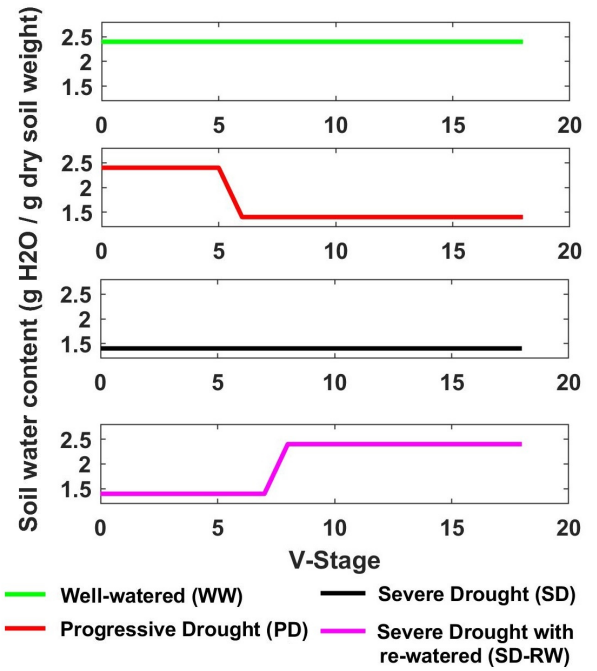


Fig. 1: Four different water treatments applied to maize plants, showing the level of soil water content over the entire developmental period indicated by the V-stage, which represents here the number of developed plant leaves.

2.2. Clustering

The SNV normalization method applied here only accounts for linear effects. In larger plants in particular, non-linear effects may occur which can seriously affect the analysis. These effects are strongly related to the plant canopy structure. To reduce this unwanted variability, less relevant regions such as (partially) occluded leaves and the areas at the leaf edges

which are prone to multiple scattering were discarded. For this, clustering of all SNV normalized spectra from all plants over all time instances was performed.

In previous work [4, 5], clustering of all the input spectra was performed using the k -means clustering algorithm. However, when dealing with a large number of HSI, more than hundreds of millions spectra are involved, which makes the computation extremely heavy and may easily saturate system memory.

In this work, a different clustering strategy is proposed. It combines the k -means clustering algorithm with a supervised classifier, a Support Vector Machine (SVM) [6]. First, the k -means clustering was performed on a small subset of the time series of the HSI from plants of the healthy as well as the stressed groups. The resulting centroids were arranged in ascending order based on the Euclidean norm. Then, a limited number of spectra, relatively close to the centroids were selected as the labeled instances for the training of the SVM. Once the SVM was trained, all the unlabeled spectra from the whole image collection were classified into k clusters. Based on the obtained cluster map, less-informative clusters were annotated and pixels from these regions were discarded. In the next step, each plant was characterized by one SNV spectrum. This spectrum was obtained by averaging the normalized spectra of all pixels belonging to the retained clusters.

2.3. Spectral similarity measure

To distinguish the stress-related phenomena from regular plant process dynamics, a direct representation of a spectral similarity measure was used for a relative comparison between stressed and healthy plants. In this procedure, the spectral distance between any two spectra $q(\lambda)$ and $r(\lambda)$ was determined using the Euclidean distance (ED):

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

where B is the number of bands. By applying the similarity measure, the dynamics of a plant spectrum can be compared against a reference spectrum (e.g. from well-watered plants).

To increase the discrimination accuracy between the two spectra, a supervised band selection procedure was applied. In this work, Fisher's statistics criterion [7] was applied to select a subset of top-scoring bands with high discriminative power in such a way that the selected bands optimise the class separability between two predefined classes (in our case well-watered versus the three groups of stressed plants). The band selection criterion was defined as:

$$\tilde{\rho}(\lambda) = \begin{cases} \rho(\lambda), & \text{if } F(\lambda) \geq T \\ 0, & \text{else} \end{cases} \quad (2)$$

where $\tilde{\rho}(\lambda)$ is the selected spectral band, T is a threshold value and $F(\lambda)$ is the ratio of the between-class variance over

the within-class variance. The spectral similarity measure was then applied by only using the selected bands.

3. RESULTS AND DISCUSSION

Figure 2 shows an example of a cluster map of a single maize plant. This cluster map was obtained from a large maize plant at developmental stage V13 (13 leaves developed). At this stage, the canopy structure may lead to non-linear illumination effects, particularly due to multiple scattering. These non-linearities cannot be corrected by the applied SNV normalization as the method only reduces the linear effects (i.e., scaling and offset due to inclination and elevation variability). From visual comparison of the cluster map with the RGB image, one can notice that the lower clusters (1-3) are mostly associated with pixels at regions that receive a low level of illumination because of higher distance from the light source and shading by partial-occluded leaves, and also some pixels which are near the leaf edges. The spectra in these regions are expected to be influenced by multiple scattering and were therefore discarded from further analysis.

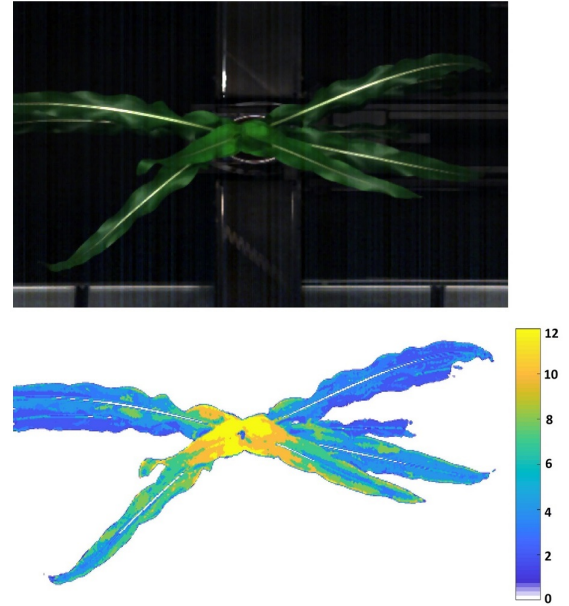


Fig. 2: RGB image and cluster map from a maize plant at the V13 growing stage.

To quantify the drought-related effects, each plant was characterized by a single SNV spectrum, by averaging over the pixels from the retained clusters. By taking as a reference spectrum, the average of the spectra of all the 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 3 plots the spectral distances of stressed plants against this reference during the whole experiment. Each data point represents the average spectral distance

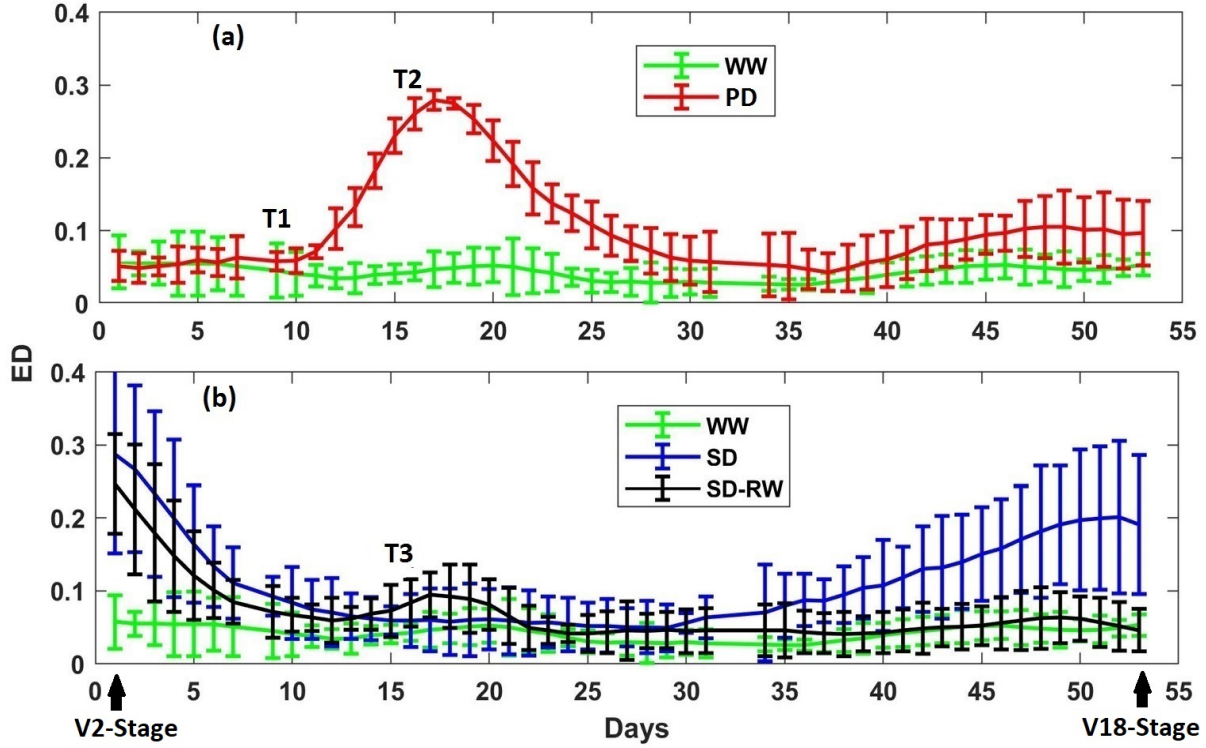


Fig. 3: Evolution of the spectral distance throughout the drought stress experiment. Plants grew from the V2 until the V18-stage.

obtained from all plants of each of the four treatments. Standard deviations of these average spectral distances over the number of plants in each class are given as well. Note that there were no measurements available on days 8, 32 and 33.

Figure 3(a) shows the results of the spectral comparison between plants of classes WW and PD. The drought stress was detected as early as the third day of the drought induction (at T1, irrigation was fully stopped). The differences in spectral distance between these two groups gradually increased as the plants were withheld from water. At T2, the plants were watered again albeit to a lower soil water content than the well-watered treatment, after which the obtained spectral distance start to decrease and even stabilize again, as the plants recovered. However, the stabilization did not persist until the end of the developmental period, as after day 40, the spectral distance start to deviate again. A possible explanation might be that although plants adapted to the lower level of available water, at later developmental stages, they re-experienced drought stress, meaning that adaptation was not complete.

Figure 3(b) presents the comparison between plants of treatment SD and SD-RW, where the WW plants were used as the reference. Since the irrigation was limited for both SD and SD-RW plants from the start (i.e. two weeks earlier than the start of the image acquisition), the effect of drought stress was visible from the first day of observation. From that day on, the spectral distance decreases monotonically, indicating that the drought plants were adapting to the water stress environment

through various biological mechanisms. When the plants progressed towards a later developmental stage, a significant discrimination between the two treatments was detected after day 35, indicating an impairment in the plant development of the SD group. As the SD-RW plants were fully well-watered again at an early vegetation stage (T3), these plants were able to fully recover and regain their optimal growth pattern. This was not achieved by the plants in the SD group.

The discrimination between the healthy and drought-stressed plants was achieved solely by determining differences in plant spectra. In this context, the obtained spectral characterization is referred to as non-targeted determination, with no direct link between the spectra and specific phenotypic traits. In order to obtain a possible biological interpretation, the information from the band selection strategy may provide useful indicators to correlate the spectral variations with specific plant traits. The obtained top-scoring bands with high discriminative power are shown in Figure 4. The spectrum follows a systematic shape with several peaks occurring in the 600-700 nm, 700-780 nm and 800-850 nm spectral regions. This specific pattern may be linked to the biological processes that change the biological properties of the plant during the stress and recovery period, such as the leaf biochemical composition, morphology of leaf surface and the internal cell structure [8]. The changes of reflectance in the visible and the red-edge regions are mainly related to the modification of photosynthetic pigments, while in the

NIR, the changes are due to light scattering of the internal properties of the cell structure related to leaf thickness and plant dry matter [9].

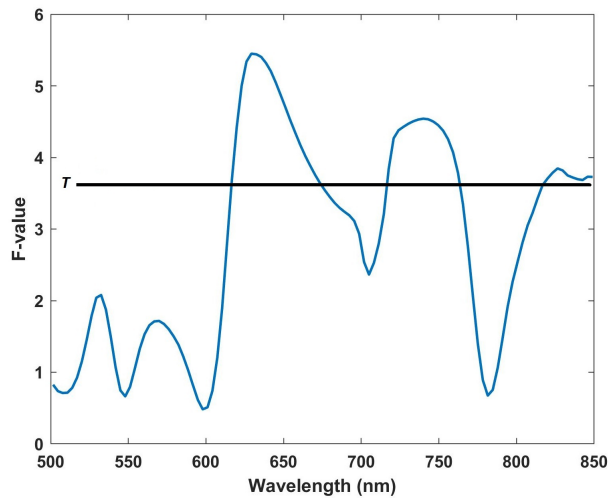


Fig. 4: The F -value obtained from the band selection procedure. The threshold was set at 70% of the maximum F value.

4. CONCLUSION

In this study, it was demonstrated that close-range HSI is a promising rapid and nondestructive technique for detection of drought stress responses of individual plants over time. The proposed method uncovers biological-related traits from spectral reflectance by a data-driven method that combines clustering, band selection, and spectral similarity measures. The analysis method was applied to detect plant responses to drought and recovery after a re-watering period in maize plants during the whole vegetative development. Experimental results show that the method can clearly discriminate plants under water-deficit stress from healthy plants at an early stage of stress development. The method also clearly reveals the recovery of plants after a re-watering period. This demonstrates the usefulness of close-range HSI as a novel technology for high-throughput phenotyping studies that may boost the understanding of the genetics of drought tolerance in breeding research. Further research and practical optimization are however needed to fully realize its potential for the phenotypic exploration of novel traits based upon prevailing spectra in groups of genotypes, or differences in spectra between genotypes.

5. REFERENCES

- [1] Puneet Mishra, Mohd Shahrime Mohd Asaari, Ana Herrero-Langreo, Santosh Lohumi, Belén Diezma, and Paul Scheunders, "Close range hyperspectral imaging of

plants: A review," *Biosystems Engineering*, vol. 164, no. Supplement C, pp. 49 – 67, 2017.

- [2] Christoph Römer, Mirwaes Wahabzada, Agim Ballvora, Francisco Pinto, Micol Rossini, Cinzia Panigada, Jan Behmann, Jens Léon, Christian Thureau, Christian Bauckhage, et al., "Early drought stress detection in cereals: simplex volume maximisation for hyperspectral image analysis," *Functional Plant Biology*, vol. 39, no. 11, pp. 878–890, 2012.
- [3] Jochem Verrelst, Gustau Camps-Valls, Jordi Muñoz-Marí, Juan Pablo Rivera, Frank Veroustraete, Jan GPW Clevers, and José Moreno, "Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties—a review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 108, pp. 273–290, 2015.
- [4] Mohd Shahrime Mohd Asaari, Puneet Mishra, Stien Mertens, Stijn Dhondt, Dirk Inzé, Nathalie Wuyts, and Paul Scheunders, "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*, vol. 138, pp. 121 – 138, 2018.
- [5] Jan Behmann, Jörg Steinrücken, and Lutz Plümer, "Detection of early plant stress responses in hyperspectral images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, pp. 98–111, 2014.
- [6] Maokuan Li, Yusheng Cheng, and Honghai Zhao, "Unlabeled data classification via support vector machines and k-means clustering," in *Proceedings of the International Conference on Computer Graphics, Imaging and Visualization, CGIV*, 2004, pp. 183–186.
- [7] Andreas Grünauer and Markus Vincze, "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*, 2015.
- [8] Rita Linke, Katja Richter, Judith Haumann, Werner Schneider, and Philipp Weihs, "Occurrence of repeated drought events: can repetitive stress situations and recovery from drought be traced with leaf reflectance?," *Periodicum biologorum*, vol. 110, no. 3, pp. 219–229, 2008.
- [9] Josep Peñuelas and Iolanda Filella, "Visible and near-infrared reflectance techniques for diagnosing plant physiological status," *Trends in plant science*, vol. 3, no. 4, pp. 151–156, 1998.