Computers and Electronics in Agriculture 162 (2019) 749–758

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Analysis of hyperspectral images for detection of drought stress and recovery in maize plants in a high-throughput phenotyping platform

Mohd Shahrimie Mohd Asaari[a,d,*], Stien Mertens[b,c], Stijn Dhondt[b,c], Dirk Inzé[b,c], Nathalie Wuyts[b,c], Paul Scheunders[a]

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

ARTICLE INFO

Keywords:
Close-range hyperspectral imaging
High-throughput plant phenotyping
Clustering
Spectral similarity measure
Drought stress

ABSTRACT

The study of physiological processes resulting from water-limited conditions in crops is essential for the selection of drought-tolerant genotypes and the functional analysis of related genes. A promising, non-invasive technique for plant trait analysis is close-range hyperspectral imaging (HSI), which has great potential for the early detection of plant responses to water deficit stress. In this work, a data analysis method is described that, unlike vegetation indices, the present method applies spectral similarity on selected bands with high discriminative information, while requiring a careful treatment of uninformative illumination effects. The latter issue is solved by a standard normal variate (SNV) normalization that removes linear effects and a supervised clustering approach to remove pixels that exhibit nonlinear multiple scattering effects. On the remaining pixels, the stress-related dynamics is quantified by a spectral analysis procedure that involves a supervised band selection procedure and a spectral similarity measure against well-watered control plants. The proposed method was validated by a large-scale study of water-stress and recovery of maize plants in a high-throughput plant phenotyping platform. The results showed that the analysis method allows for an early detection of drought stress responses and of recovery effects shortly after re-watering.

1. Introduction

Imaging techniques have improved the precision and throughput of plant phenotyping, and now become a new frontier in phenotypic trait measurement. Current phenotyping platforms include a variety of imaging modalities to obtain high-throughput, non-destructive phenotype data for quantitative assessment of structural and functional plant traits. Plant trait assessment in high-throughput plant phenotyping platforms (HTPP) has recently been studied using close-range hyperspectral imaging (HSI) as a promising non-invasive tool (Ge et al., 2016; Mishra et al., 2017). In particular, HSI has been applied for the assessment of plant responses to biotic and abiotic stress conditions, such as fungal infection, water and nutrient deficits. During the stress development, a number of physiological and biochemical responses happen in plants, including modifications in the functioning of the photosynthetic apparatus, plant organ, water content, leaf surface and internal structure. These modifications alter the leaf optical properties (Sun et al., 2018) that can be measured by HSI. Recent advances in this field encourage studies on plant responses to drought stress, and on the plant's capability to adapt and recover from this stress. Such studies are crucial for the further improvement of crop drought-tolerance in breeding programs.

A common approach for plant trait estimation based on HSI is to utilize vegetation indices (VIs), defined as ratios or linear combinations of reflectances at a few single wavelengths. One advantage of VIs is that they minimize the possible influence of scale factors, including slope effects and variations in illumination conditions (Jay et al., 2017). VIs usually focus on very specific biological traits and processes in plants (Heiskanen et al., 2013; Katsoulas et al., 2016; Ihuoma and Madramootoo, 2017), whereas the complex physiological effects of drought stress alter the reflectance in many different wavelength regions. Thus, VIs may discard significant information leading to a decrease in the discrimination accuracy (Römer et al., 2012).

Another widely used method for retrieving vegetation characteristics from reflectance data is the inversion of radiative transfer models (RTM). In RTM inversion, model parameters such as chlorophyll

* Corresponding author.
E-mail address: MohdShahrimie.MohAsaari@uantwerpen.be (M.S.M. Asaari).

https://doi.org/10.1016/j.compag.2019.05.018
Received 8 November 2018; Received in revised form 21 March 2019; Accepted 7 May 2019
Available online 15 May 2019

concentration, water content, dry matter, and canopy structures are retrieved using look-up-tables and optimization techniques (Sun et al., 2018). A common challenge of these methods is their ill-posedness (Jacquemoud et al., 2009), as various combinations of vegetation parameters may correspond to almost similar spectra. Moreover, this method does not apply well to close-range settings because the physically-based leaf or canopy RTMs are difficult to adapt to the specific close-range illumination problems (Jay et al., 2016).

Data-driven machine learning regression algorithms provide a third way to retrieve plant biophysical variables from the reflectance spectrum (Verrelst et al., 2015; Rapaport et al., 2015). Regression analysis reveals statistical correlations between the spectral variables and biological information. Typically, a flexible learning model is inferred from a training dataset by optimizing the estimation error of the extracted variables. As they implicitly derive the underlying model distribution from a given dataset, these methods are very flexibel. However, they cannot be applied if the required output variables for training the model are not available.

In this work, an alternative data-driven method is proposed. To eliminate scaling effects from leaf orientations and specific allignment of the imaging system in close-range settings, a standard normal variate (SNV) normalization is applied first. To filter out noninformative nonlinear variability induced by multiple scattering and shading in more complex canopy structures, a supervised clustering procedure is proposed and clusters of spectra associated to shadowed and partially occluded areas were discarded. To quantify the dynamics of the water-deficit stress response of a plant, it was characterized by the average SNV spectrum from the retained clusters. An Euclidean distance function was then applied to discriminate stressed from well-watered plants. To optimize the discrimination, a supervised band selection procedure was applied to extract a small subset of top-scoring variables with high class separability. The proposed methodology was validated by a large scale experiment in a HTPP that monitored maize plants during their entire vegetative development period. Six different groups of test plants were monitored: well-watered control plants, and five groups of plants undergoing different water-deficit stress conditions, for which we analyzed their response to the drought stress and their recovery after re-watering.

2. Materials and methods

2.1. Data acquisition

A batch of maize plants was grown in PHENOVISION, the HTPP infrastructure located at VIB, Ghent, Belgium. The plants were divided into six groups udergoing different water irrigation strategy (Fig. 1). All treatments started at the seedling level.

- Group WW (Fig. 1 (a)): the well-watered treatment. Seven plants were irrigated with sufficient water to keep the soil water content at the optimal level of 2.4 g H_2O/g dry soil throughout the entire plant developmental period.
- Group PD-RW1 (Fig. 1 (b)): the progressive drought with re-watering 7 days after the V5-stage 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 (at that time they reach V6 or V7), after which they were re-watered at V6-stage (six leaves developed) with a low amount of water to maintain the soil water content at a deficit level of 1.4 g H_2O/g dry soil until the end of the developmental period.
- Group PD-RW2 (Fig. 1 (c)): the progressive drought with re-watering 7 days after V5-stage and at V12-stage treatment. Seven plants received the PD-RW1 treatment up to V12 vegetation stage (twelve leaves developed). From V12-stage onward, the plants were irrigated with the WW treatment until the end of the developmental period.
- Group SD (Fig. 1 (d)): the severe drought treatment. Four plants were irrigated with a deficit soil water content of 1.4 g H_2O/g dry soil throughout the developmental period.
- Group SD-RW1 (Fig. 1 (e)): the severe drought with re-watering at the V7-stage. 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 according to the WW treatment until the end of the developmental period.
- Group SD-RW2 (Fig. 1 (f)): the severe drought with re-watering in the V12-stage treatment. Seven plants received the SD treatment from the beginning until they reached the V12-stage after which they were irrigated according to the WW treatment until the end of the developmental period.

From all plants involved, hyperspectral images were acquired daily during 50 days from growth stage V2 (two leaves developed), about 2 weeks after the start of the water treatments. A line scan push-broom VNIR-HS camera (ImSpector V10E, Spectral Imaging, Oulu, Finland) was used to capture the hyperspectral images. The complete acquisition process produced 350 hyperspectral images for each WW, PD-RW1, PD-RW2 and SD-RW2 treatment, 300 images for the SD-RW1 and 200 images for the SD treatment, resulting in a total of 1900 images. The acquired images had 510×328 pixels and an average spectral sampling of 3.1 nm which corresponds to 194 bands ranging between 400 and 1000 nm.

All images were radiometrically calibrated by subtracting a dark frame and reflectance was calculated relative to a white reference. Fig. 2 shows a collection of reflectance spectra from plant pixels. The levels of Gaussian noise, present in the spectrum were quantified using the Generalized Cross Validation (GCV) score (Garcia, 2010) (see Table 1). Because of high noise levels below 500 nm and above 850 nm, the images were limited to 111 spectral bands in the range 500–850 nm for further data processing. The plant pixels were then segmented from the background using the normalized difference vegetation index (NDVI). Fig. 3 shows a segmented plant for several threshold values of the NDVI. A threshold of 0.3 was chosen.

All plants were imaged in an indoor environment, inside a closed cabin. The imaging cabin is illuminated with halogen lamps, homogeneously distributed in a 2-dimensional plane of the field of view of the HS camera. Although the illumination is homogeneous, spectral variability occurs due to physical effects of light reflection. In particular, the high spatial resolution of HSI in this close-range setting makes the recorded signal very sensitive to the specific alignment 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. Assuming that the leaf surface is Lambertian, the fraction of the leaf reflectance received by the sensor is largely affected by the inclination of the leaf towards the light source and the distance towards the sensor. These physical effects can be explained by the lambert's cosine law and the inverse square law, which cause multiplicative and additive effects on the reflectance spectra. This induces high uninformative variability in the recorded signals which overlay the subtle effects of the biological traits. Since these effects are linear, a linear pre-treatment technique, the Standard Normal Variate (SNV) was applied (Asaari et al., 2018) to reduce these nuisance variabilities.

2.2. Clustering

The SNV normalization method only accounts for linear scaling effects. In larger plants with a more complex canopy structure, partially occluded leaves, shadowing and multiple reflections at the leaf edges cause unwanted nonlinear variability. To remove this variability, a clustering procedure to discard these regions is proposed.

Typically, unsupervised clustering such as the k-means clustering

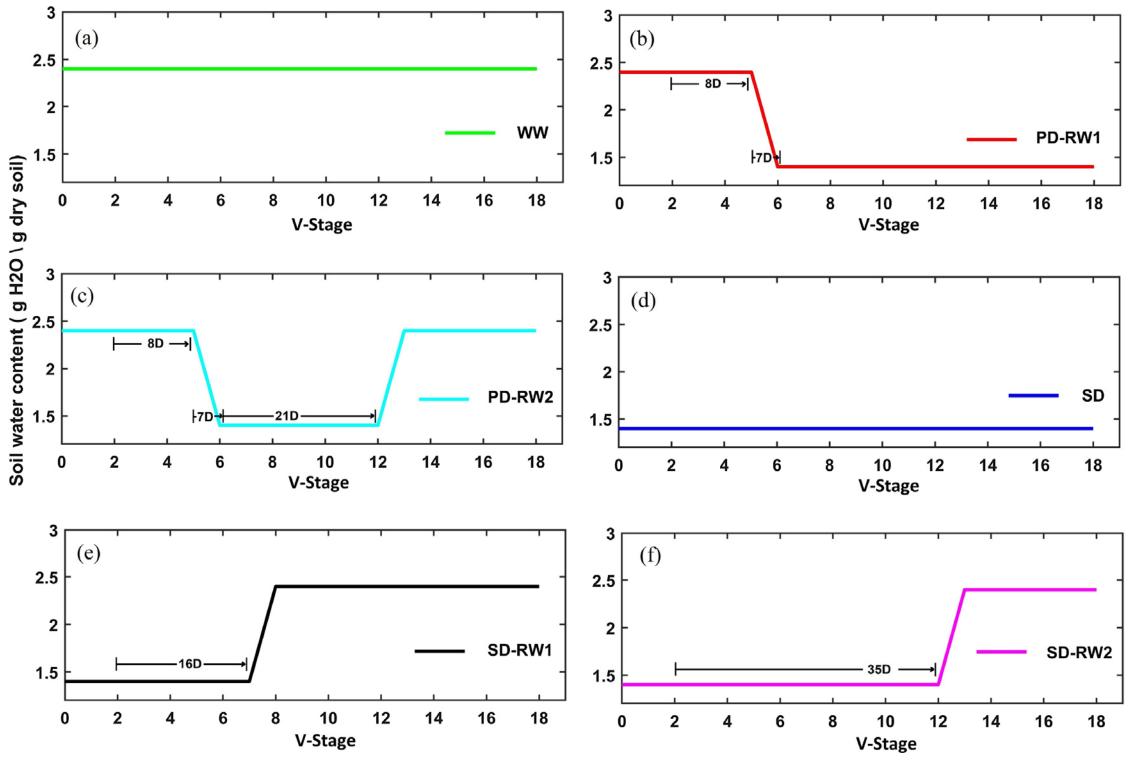

Fig. 1. Six different irrigation strategies applied to maize plants, showing the level of soil water content over the entire vegetative developmental period at different V-stages, indicating the number of developed plant leaves and the number of days to reach a particular V-stage. (a) well-watered treatment (WW), (b) progressive drought with re-watering 7 days after V5-stage treatment (PD-RW1), (c) progressive drought with re-watering 7 days at a water deficit levels after V5-stage and in the V12-stage treatment at a WW level (PD-RW2), (d) severe drought treatment (SD), (e) severe drought with re-watering in the V7-stage at a WW level (SD-RW1) and (f) severe drought with re-watering in the V12-stage at a WW level (SD-RW2).

algorithm can be applied (Asaari et al., 2018; Behmann et al., 2014). In the proposed experiments, tens of millions of spectra are involved. The large-scale data streams in HTPP systems pose computational challenges as the system memory may become saturated. Therefore, in this work, a different clustering strategy is proposed: a supervised method, which combines the Support Vector Machine (SVM) classifier with the k-means clustering algorithm (Li et al., 2004). Since it is a supervised algorithm, it requires labeled instances for training the classifier. To avoid time-consuming manual labeling, unsupervised labeling is performed to create representative spectra for different groups of pixels.

In first instance, k-means clustering was performed on a small subset of all the acquired images from the well-watered control plants and the different stressed groups over the entire development period. The number of clusters k was estimated by analyzing the dispersion of

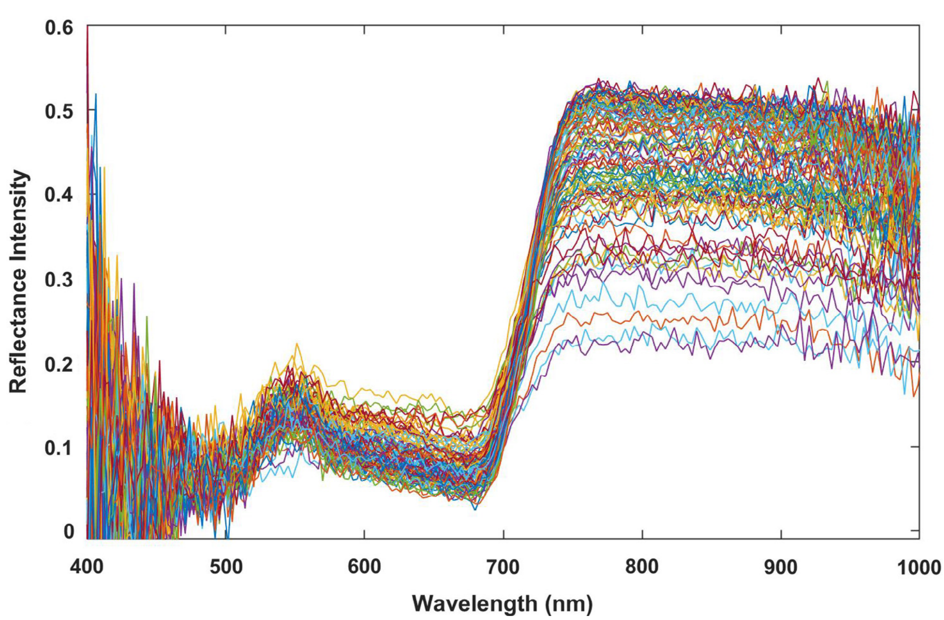

Fig. 2. Reflectance spectra of a collection of plant pixels, covering the spectrum region between 400 nm to 1000 nm. The spectra show high noise levels at wavelength regions below 500 nm and above 850 nm. The noise levels were quantified using the GCV score (see Table 1).

Table 1
Gaussian noise estimation on different wavelength intervals of the spectra from Fig. 2. The estimated variance of this noise is estimated based on the GCV score (Garcia, 2010). Lower GCV scores correspond to lower noise levels.

Wavelength Region	Estimated Noise Variance (GCV score)
400–1000 nm	4.560×10^{-2}
400–500 nm	5.677×10^{-1}
400–850 nm	7.320×10^{-2}
500–850 nm	3.416×10^{-5}
500–980 nm	1.434×10^{-4}
500–1000 nm	3.890×10^{-4}
850–1000 nm	1.600×10^{-3}

the within-groups sum of squares for different values of k (Sarstedt and Mooi, 2014) and was set to 12 (Fig. 4). Then, the resulting cluster centroids were arranged in ascending order, based on the Euclidean norm. In the next step, the training sample size was limited to 100 spectra for each cluster, chosen relatively close to the cluster centroids. This data reduction strategy was aimed at improving the computational efficiency of SVM in both training and prediction phases (Tang et al., 2018). Then, SVM with a radial basis function kernel (Chang and Lin, 2011) was used to train the classifier and all the unlabeled spectra from the entire image collection were classified as belonging to one of the k clusters.

Fig. 5 shows an example of an obtained cluster map, in which the pixels are mapped using a false color representation in accordance with their cluster number. Based on these cluster maps, less-informative clusters were annotated and pixels from these clusters were discarded. Finally, each plant was characterized by one SNV spectrum, obtained by averaging the normalized spectra of all pixels belonging to the retained clusters. The entire development period of each plant is then represented as one spectral time-series.

2.3. Spectral similarity measure

To distinguish stress-related behaviour from control plant growth dynamics, a spectral similarity measure (SSM) was applied between stressed and well-waterd plants. The Euclidean distance measure was applied to calculate the spectral distance between any two spectra $q(\lambda)$ and $r(\lambda)$:

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

where B is the number of bands.

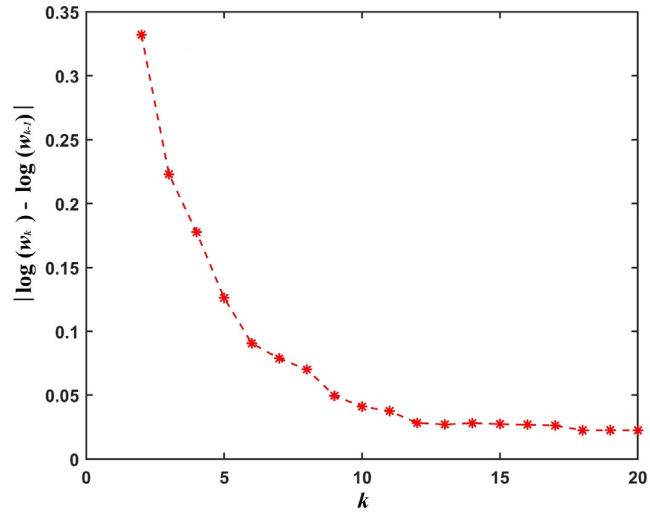

Fig. 4. Choosing the number of clusters by analyzing the dispersion in the within-group sums of squares (w_k). A break point in the curve occurs at $k = 12$.

The similiarity measure allows to compare the dynamics of a plant against a reference. In this work, the reference spectrum at each day was defined as the average spectrum of all plants in the WW group of that particular day. The obtained spectral time-series represents control plant growth and functioning dynamics. The dynamics of a control plant will be very similar to the reference time series (slightly positive since a distance is always positive), while behaviour other than the regular dynamics of the control plants will result in a significant difference with the reference time series.

To increase the disciminative power between stressed and control plants, a supervised band selection procedure was applied. In this work, Fisher's statistics criterion (Grünauer and Vincze, 2015) was applied. It selects a subset of top-scoring bands with high discriminative power that optimise the class separability between two predefined classes (in our case well-watered versus the five groups of stressed plants). The band selection criterion was defined as:

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

where $\tilde{\rho}(\lambda)$ is the selected spectral band, T is a threshold value and $F(\lambda)$ is the ratio of the between-class and the within-class variance. The spectral similarity measure was then applied by only using the selected bands.

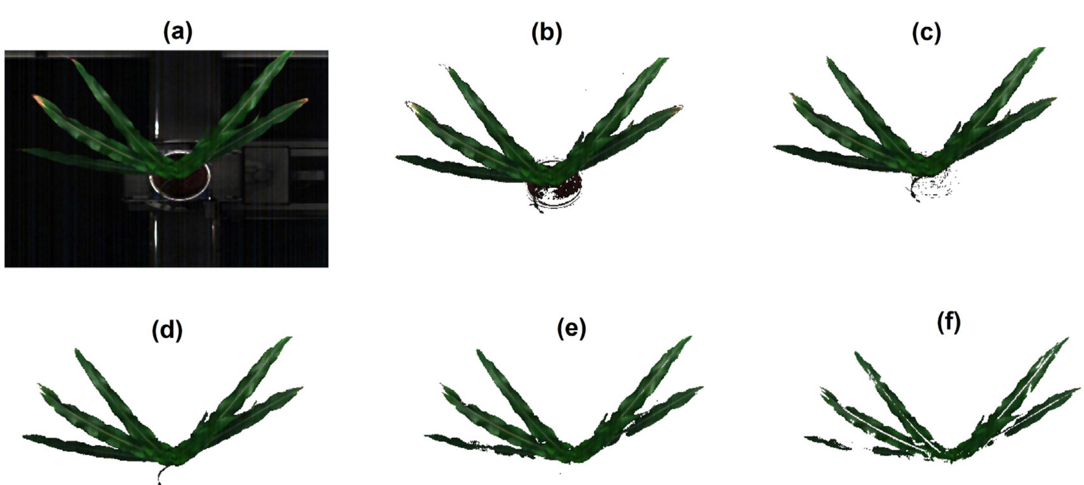

Fig. 3. Segmentation of plant pixels based on NDVI threshold. Full hyperspectral image (a), segmented hyperspectral images based on NDVI threshold of 0.1 (b), 0.2 (c), 0.3 (d), 0.4 (e) and 0.5 (f).

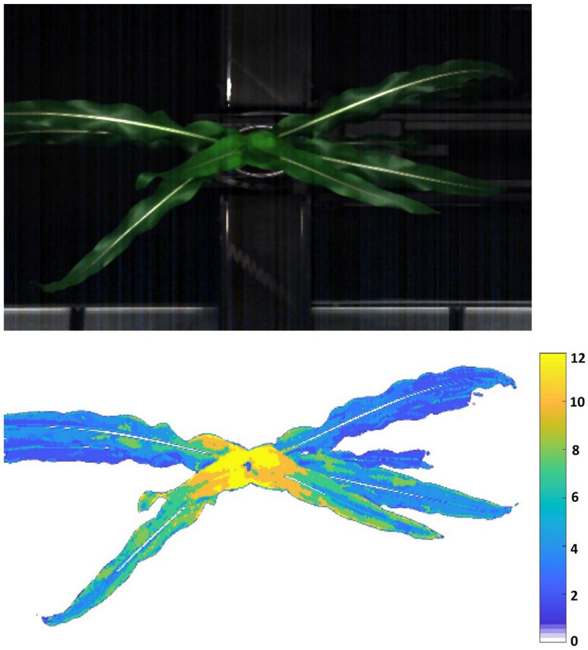

Fig. 5. RGB image and cluster map from a maize plant at the V13 growing stage.

3. Results and discussion

In the first experiment, we validated the clustering strategy of Section 2.2. To do so, we evaluated the performance of the proposed technique against the original k-means clustering algorithm. For this, a fraction (25%) of the spectral data was proportionally distributed to five test data sets, to conduct five independent experiments. The ground truth labels for this test data was obtained using the k-means clustering algorithm. Then, for each of the five experiments, 100 spectra of each cluster ($k = 12$) were randomly chosen to train the SVM. The remaining spectra acted as validation data, for which the obtained label was compared against the ground truth obtained by k-means clustering. Table 2 shows the SVM classification accuracy on this validation dataset. The overall agreement between the proposed and the k-means clustering was above 96%, confirming that the use of the supervised clustering approach was justified.

The proposed clustering algorithm was applied to label every pixel in each individual plant and the resulting cluster map was further analyzed to filter out less-informative spectra. Fig. 5 shows an example of a cluster map of a single maize plant at developmental stage V13 (13 leaves developed). At this stage, the complex 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 that method only reduces the linear effects (i.e. scaling and offset due to leaf 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 regions from which the sensor receives a low level of illumination because they are more distant from the light source or that contain shading and partially occluded leaves. Leaf edges belong to these lower clusters as well. The spectra in these regions are expected to be influenced by multiple scattering and were therefore discarded from further analysis.

Table 2
Classification accuracy of the proposed supervised clustering approach in five independent experiments. Ground truth labeling was obtained from the k-means clustering algorithm. The processing time is based on the experiment running on Matlab R2018a with 4.00 GHz Intel Core i7 CPU and 32.0 GB system memory.

Data set	Number of test spectra	Match cluster between SVM and k-means (%)	Processing time (s)	
			SVM	k-means
1	2.0007×10^6	95.83	101.55	405.41
2	2.0929×10^6	96.23	105.94	439.98
3	2.3254×10^6	96.33	118.98	619.30
4	2.2473×10^6	96.32	113.82	616.38
5	2.1135×10^6	96.26	108.26	483.52
	Overall performance:	96.19	109.71	512.92

The next experiment was an actual experiment with well-watered control and water-deficit stress treatments to monitor the growth dynamics of the plants from the six different watering treatment groups, and to analyse the response to drought and recovery after re-watering. The proposed method from Section 2.3 was applied to obtain the spectral distance of each plant from the reference spectra, during the entire experiment (53 days). The well-watered group acts as a control group. Fig. 6 shows the plots for the five different stressed groups, each time compared to the plot of the WW control group. Each data point is an average over all plants of the group; standard deviations are given as well. Note that there were no measurements available on days 8, 32 and 33.

Fig. 6(a) shows the results of the group PD-RW1 versus the WW control group. The drought stress was detected as early as the third day of the drought induction (at T1, irrigation was completely stopped). The difference with the control group gradually increased as the plants were withheld from water. At T2, 7 days after T1, the plants were watered again albeit to a lower soil water content than the well-watered treatment, after which the difference started to decrease, indicating that the plants were recovering. About 15 days after re-watering, the plants seemed to have completely recovered. However, this situation did not persist until the end of the developmental period, as after day 40, the difference with the control started to grow again. Apparently, the plants initially adapted to the lower soil water content, but at a later development stage, they seemed to re-experience drought stress.

Fig. 6(b) shows the results of the group PD-RW2 versus the control group. The water treatment of this group is identical to the one of PD-RW1 up to day 37 (T3). As expected, the behaviour is very similar to the behaviour of the PD-RW1 group. After that day, the plants were irrigated again with higher water levels equivalent to the WW treatment. From day 40 on (3 days after starting the WW treatment), a significant deviation from PD-RW1 group was observed, as the PD-RW2 group seemed to have fully recovered from the drought stress.

Fig. 6(c) shows the results of the group SD versus the control group. Since the irrigation for SD plants was limited from the start (i.e. two weeks before day 1), the effect of drought stress was visible from the first day of observation. From that day on, the difference with the control group decreases monotonically until day 10, indicating that the drought plants were adapting to the water stress environment. From the literature, it is known that plants can adapt through various biological mechanisms (Xu et al., 2010; Zegada-Lizarazu and Monti, 2013; Sun et al., 2016). After this, the plants seemed to behave as WW control plants until day 35, after which the plants start to re-experience drought stress. This effect seemed to start earlier and to be more severe than for the plants in the progressive drought treatment (PD-RW1), indicating a very serious impairment in the plant development of the SD group.

For the remaining two groups, SD-RW1 and SD-RW2, the goal was to evaluate to what extent plants have the capacity to recover from severe drought stress when re-watering is performed. The SD-RW1 group was fully re-watered after severe drought induction, at an early vegetative state (V7), while SD-RW2 was fully re-watered at a later development stage (V12). Fig. 6(d) and (e) show the results of these groups versus the control group. For the SD-RW1 group, the plant health status stabilizes shortly after re-watering (at point T4) and

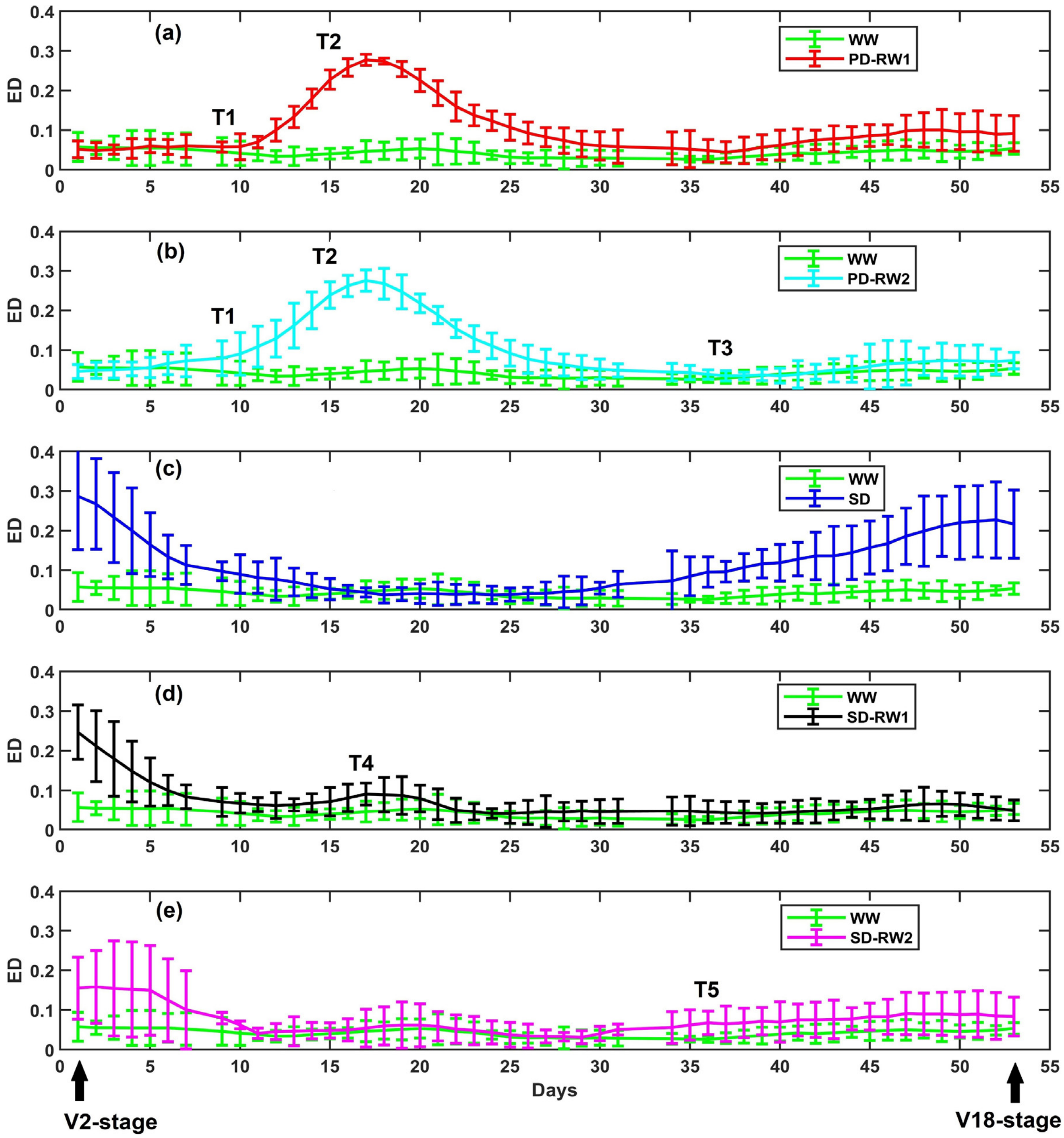

Fig. 6. Evolution of the spectral distance with respect to the control group throughout the drought stress experiment for the WW control group, the PD-RW1 group (acute drought between T1 and T2), the PD-RW2 group (acute drought between T1 and T2 and re-watering to WW level at T3), the SD group and the SD-RW1 (re-watering to WW level at T4), and SD-RW2 groups (re-watering to WW level at T5). Plants grew from the V2 until the V18-stage.

remains undifferent from the control group until the end of the vegetative development stage. This indicates that these plants were able to fully recover and regain their optimal growth and functioning pattern. However, this was not achieved by the SD-RW2 group, that deviates from the control group after the late re-watering period (T5). This indicates that re-watering at a later development stage does not allow plants to entirely recover from severe drought stress.

In the next experiment, the aim was to study the positive effect of the cluster procedure on the results. For this, the same experiment on the WW and the PD-RW1 groups was repeated but then without performing the clustering. As a consequence, all pixels of the plants, including the ones that were influenced by nonlinear effects, were included in the experiment. All other procedures, i.e. SNV normalization and band selection were performed as before.

Fig. 7 plots the evolution of the plants in the PD-RW1 group against the WW control group. From this plot, it can be observed that in general, the standard deviations were larger than in the original experiment. This effect remained rather small at the early vegetation stages, but became larger at the later vegetation stages, where the canopies were larger and more complex, leading to more serious effects of multiple scattering and shading. Because of this, during the early vegetation stage, not performing the clustering had only a minor effect on the discrimination between control and drought plants. The only difference that was observed was that the onset of the water stress was detected only on the fourth day after the drought induction, one day later than the case where clustering was applied. However, at later vegetation stages, the high standard deviations hindered the distinction between healthy and drought plants, such that the re-experience of drought stress after 40 days remained entirely unnoticed.

Many past and recent studies have applied VIs to characterize the biophysical and physiological plant status in response to drought stress (Rumpf et al., 2010; Kim et al., 2011; Amatya et al., 2012; Sun et al.,

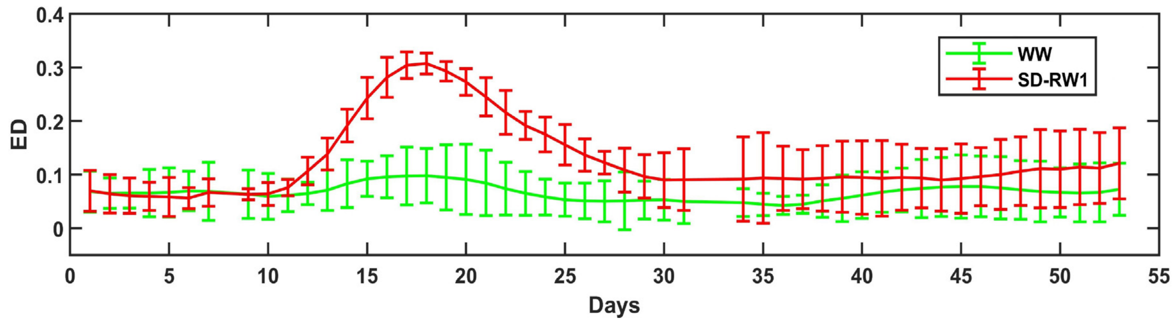

Fig. 7. The obtained spectral distance when no cluster treatment was performed. The evolution plots show the comparison between plants in PD-RW1 group versus control plants throughout the drought stress experiment.

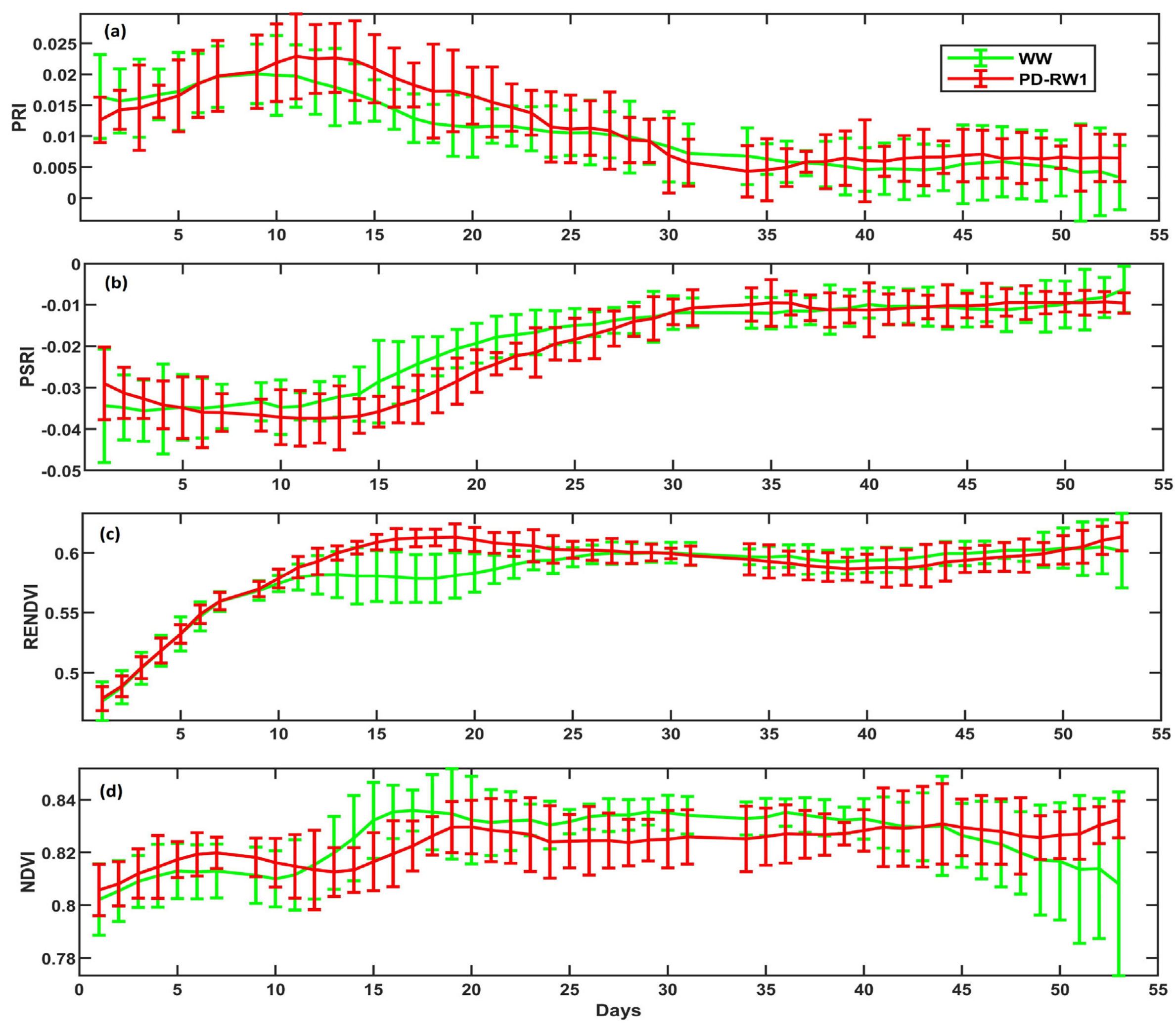

Fig. 8. Evolution of spectra for the plants in PD-RW1 group versus plants in the WW group based on the calculation of vegetation indices (a) PRI (b) PSRI and (c) RENDVI (d) NDVI.

2014; Behmann et al., 2014; Gago et al., 2015). The photochemical reflectance index (PRI) and the normalized difference vegetation index (NDVI) are the most commonly used VIs for crop water stress assessment. Other reflectance indices like the red-edge normalized difference vegetation index (RENDVI) and plant senescence reflectance index (PSRI) have also been used with varying results. In Sun et al. (2014), a significant correlation between PRI and water content was found, while in Kim et al. (2011) it was shown that RENDVI and NDVI are two indices that are highly correlated with plant water stress. In addition to these indices, Behmann et al. (2014) reported PSRI as a relevant indicator for detecting plant stress.

To test the relevance of the proposed spectral analysis method, a comparison with the aforementioned VIs on the drought stress experiments was performed. To calculate the VIs, no SNV normalization was applied, because VI's need to be obtained directly from reflectance spectra, and because VI's take scaling effects automatically into account. However, the same clustering treatment as in the proposed method was applied to account for nonlinear illumination effects.

Fig. 8 shows the plots of PRI, PSRI, RENDVI, and NDVI of the PD-RW1 versus the control group. In general, deviations from the control

Table 3
The p-values of a one-way ANOVA at the 0.05 significance level for the proposed method and the four VIs. The obtained p-values are based on the comparison between plants from the WW group and the SD-RW1 group.

Early vegetative stage						Later vegetative stage					
Day	Proposed method	PRI	PSRI	RENDVI	NDVI	Day	Proposed method	PRI	PSRI	RENDVI	NDVI
1	0.7372	0.2236	0.4022	0.7630	0.3992	27	0.0005	0.3441	0.5796	0.6976	0.0645
2	0.9277	0.5799	0.5979	0.8586	0.3863	28	0.1679	0.8147	0.2396	0.3780	0.0023
3	0.2696	0.8332	0.3635	0.9205	0.9096	29	0.0625	0.0918	0.4345	0.9798	0.0080
4	0.8635	0.4295	0.8261	0.9744	0.5388	30	0.0766	0.2084	0.3618	0.8588	0.0592
5	0.5378	0.5926	0.4745	0.9652	0.3278	31	0.1401	0.3700	0.3038	0.5027	0.1399
6	0.4600	0.5100	0.3235	0.8707	0.3206	34	0.1100	0.0652	0.4076	0.6012	0.2394
7	0.2191	0.7718	0.2154	0.9692	0.0367	35	0.2523	0.9366	0.3583	0.2406	0.1635
9	0.1683	0.2129	0.6387	0.0416	0.0737	36	0.1718	0.4621	0.3819	0.4515	0.1000
10	0.2091	0.9063	0.6815	0.1186	0.2872	37	0.0486	0.5161	0.0197	0.2593	0.0392
11	0.0366	0.8479	0.0611	0.4353	0.8184	38	0.3959	0.3895	0.6715	0.2187	0.0117
12	0.0216	0.0392	0.9530	0.0776	0.8894	39	0.6940	0.7205	0.1211	0.9861	0.5397
13	0.0000	0.0619	0.1098	0.0228	0.4294	40	0.0968	0.6887	0.4539	0.3899	0.6709
14	0.0000	0.0510	0.0314	0.0211	0.0505	41	0.0161	0.0706	0.9326	0.2079	0.8077
15	0.0000	0.0937	0.1318	0.0056	0.0092	42	0.0378	0.8634	0.9499	0.1208	0.4192
16	0.0000	0.3341	0.0569	0.0034	0.0025	43	0.0466	0.7792	0.4542	0.4301	0.7397
17	0.0000	0.0112	0.0127	0.0006	0.0042	44	0.0236	0.0442	0.5086	0.4659	0.7123
18	0.0000	0.0667	0.0073	0.0005	0.3626	45	0.0452	0.4044	0.8618	0.2917	0.9515
19	0.0000	0.0252	0.0013	0.0015	0.7318	46	0.0076	0.4155	0.3545	0.2486	0.8008
20	0.0000	0.1660	0.0891	0.0015	0.9838	47	0.0258	0.4004	0.3391	0.4173	0.5864
21	0.0001	0.1221	0.0490	0.0028	0.7882	48	0.0560	0.8625	0.2016	0.6139	0.3713
22	0.0001	0.0889	0.1008	0.0130	0.5704	49	0.0101	0.8138	0.5188	0.3303	0.3720
23	0.0001	0.9523	0.0315	0.1362	0.1930	50	0.0214	0.0148	0.2856	0.8201	0.2923
24	0.0000	0.1845	0.0538	0.1999	0.2497	51	0.0383	0.5005	0.9895	0.9784	0.3237
25	0.0005	0.5216	0.1688	0.1383	0.1961	52	0.1411	0.7394	0.9408	0.6797	0.2370
26	0.0001	0.0725	0.9700	0.2241	0.2348	53	0.0730	0.2188	0.1773	0.3711	0.0913

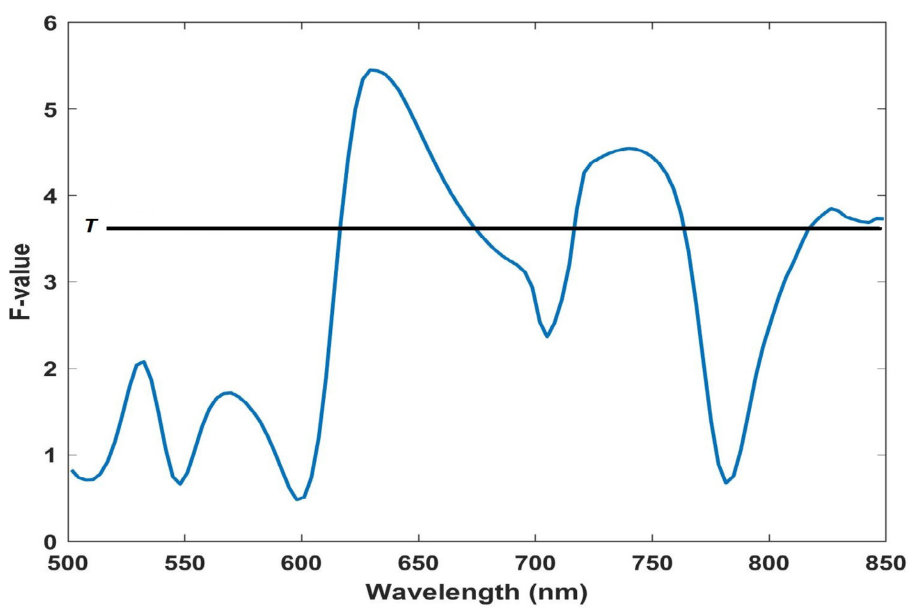

Fig. 9. The F-value obtained from the band selection procedure. The threshold was set to 70% of the maximum F value.

seem to appear at the same time intervals as in the proposed method (between day 10 an day 30 and from day 40 on), but less clear. To quantify this, a statistical significance test was conducted using analysis of variance (ANOVA). Table 3 presents the p-values obtained from the ANOVA test at 0.05 significance level for the proposed method and the four VIs.

Among the four VIs tested, RENVI was the best index for the detection of the water stress. Nevertheless, when compared to the proposed method the result was far less significant. None of the VIs was able to significantly determine the recovery at the later development stage. Clearly, the limited amount of spectral information provided by the VIs was not sufficient for a proper analysis of the drought stress and recovery after re-watering. The proposed method is capable of revealing these subtle differences by making optimal use of the most discriminative spectra from the entire wavelength range.

In the proposed method, the discrimination between control and drought-stressed plants was achieved solely by determining differences in plant spectra. Such spectral characterization is referred to as non-targeted, since it reveals no direct link between the spectral reflectance and specific phenotypic traits. For a possible biological interpretation, the information from the band selection strategy may provide useful indicators to correlate the spectral variations to specific plant traits. In Fig. 9, the F-score from the band selection procedure is shown. The curve follows a systematic shape with several peaks of top-scoring

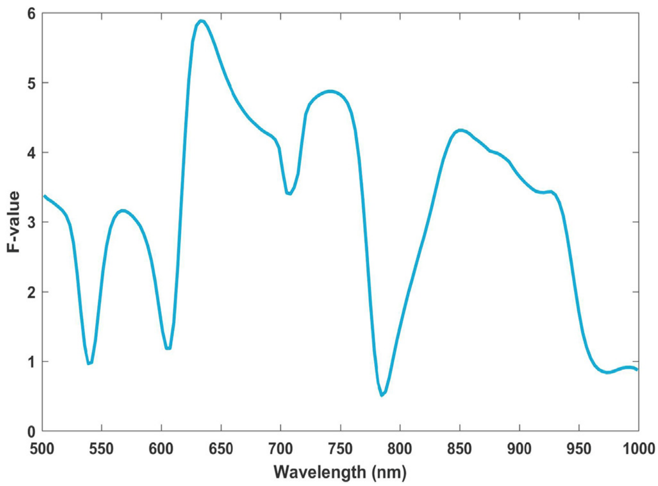

Fig. 10. The F-value obtained from the band selection procedure, for the wavelength range up to 1000 nm. The threshold was set to 70% of the maximum F value.

bands with high discriminative power, occurring in the 600–700 nm, 700–780 nm and 800–850 nm spectral regions. The position of these peaks are quite relevant when compared with the wavelengths used in the calculation RENDVI and NDVI, the best two indices proposed in a study of plant responses to drought by Kim et al. (2011). This specific pattern may be linked to the changes in the biological properties of the plant during the stress and recovery period, such as the leaf biochemical composition, the morphology of the leaf surface and the internal cell structure (Linke et al., 2008). Changes in reflectance in the visible and the red-edge regions are mainly related to the modification of photosynthetic pigments, while in the NIR, the reflectance is influenced by light scattering of the internal properties of the cell structure that is related to leaf thickness and plant dry matter(Peñuelas Filella, 1998).

The proposed method avoids the wavelength regions below 500 nm and above 850 nm because of noise. However, in (Peñuelas et al., 1993; Serrano et al., 2000), it is suggested that spectral reflectance beyond 850 nm is also useful for a direct assessment of plant stress. 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 1000 nm. Fig. 10 shows the F-values calculated for this wavelength range. Compared to Fig. 9, the systematic pattern remained similar, indicating that the locations of the important information did not change. The value of F-score decreases beyond the 850 nm region. It can also be observed that around the water absorption region at 900–950 nm a small peak occurs, smaller than the spectral variations at 600–700 nm, 700–800 nm and 800–850. In order to include information from the bands beyond 850 nm, the threshold for the F-value would need to be reduced, which is expected not to increase the discrimination results. A possible explanation is that the biological changes beyond the 850 nm wavelength region are overlaid by the extreme levels of signal noise in this region.

4. Conclusions and future perceptive

In this study, it was demonstrated that HSI is a promising rapid and nondestructive technique for the detection of drought stress responses of individual plants over time. The proposed method is able to reveal drought stress and recovery from drought stress from spectral reflectance by a data-driven method that combines clustering, band selection, and a spectral similarity measure. In the experiments, the analysis method was validated in a HTPP in a study of maize plants udergoing different types of drought stress during their entire vegetative development. Experimental results showed that the method clearly discriminated plants under water-deficit stress from healthy plants at an early stage of stress development. The method also clearly revealed the recovery of plants after a re-watering period. This demonstrates the usefulness of HSI as a novel technology for high-throughput phenotyping studies that can boost the understanding of the genetics of drought tolerance in breeding research. It is also to be noticed that the presented method is general and not limited to drought stress, and whenever there is an interest for monitoring plant process dynamics at the plant scale, it can be applied to different types of systemic stress.

Further research and practical optimization are however required 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. The compensation of illumination effects can be further improved by adopting more descriptive illumination models, such as dichromatic reflection models (Uto Kosugi, 2013) or digital surface models (Friman et al., 2011). To attain a more accurate estimation of geometry-related parameters, the integration of the 3D scene (Behmann et al., 2016) and the use of machine learning algorithms can be considered. An interesting approach is to render the 3D plant model using multiple viewpoints with a full frame snapshot hyperspectral camera system that captures all bands simultaneously (Aasen et al., 2015). With the release of high resolution snapshot hyperspectral cameras, such as the Specim IQ sensor (Behmann et al., 2018), the generation of highly accurate 3D plant models becomes possible. Another benefit of such 3D plant models is that the physiological traits extracted from the spectral information can be fused with morphological traits extracted from the 3D plant structural information. This is the research direction for our future work.

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.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.compag.2019.05.018.

References

Aasen, H., Burkart, A., Bolten, A., Bareth, G., 2015. Generating 3D hyperspectral information with lightweight uav snapshot cameras for vegetation monitoring: from camera calibration to quality assurance. ISPRS J. Photogramm. Remote Sens. 108, 245–259.

Amatya, S., Karkee, M., Alva, A.K., Larbi, P., Adhikari, B., 2012. Hyperspectral imaging for detecting water stress in potatoes. In: ASABE Annual International Meeting, Paper number: 12-1345197.

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

Behmann, J., Acebron, K., Emin, D., Bennertz, S., Matsubara, S., Thomas, S., Bohnenkamp, D., Kuska, M.T., Jussila, J., Salo, H., et al., 2018. Specim IQ: Evaluation of a new, miniaturized handheld hyperspectral camera and its application for plant phenotyping and disease detection. Sensors 18 (2), 441.

Behmann, J., Mahlein, A.-K., Paulus, S., Dupuis, J., Kuhlmann, H., Oerke, E.-C., Plümer, L., 2016. Generation and application of hyperspectral 3D plant models: methods and challenges. Mach. Vis. Appl. 27 (5), 611–624.

Behmann, J., Steinrücken, J., Plümer, L., 2014. Detection of early plant stress responses in hyperspectral images. ISPRS J. Photogramm. Remote Sens. 93, 98–111.

Chang, C.-C., Lin, C.-J., 2011. LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2, 27:1–27:27. Software available at. <http://www.csie.ntu.edu.tw/cjlin/libsvm>.

Friman, O., Tolt, G., Ahlberg, J., 2011. Illumination and shadow compensation of hyperspectral images using a digital surface model and non-linear least squares estimation. Proceedings of SPIE 8180. In Image and Signal Processing for Remote Sensing XVII, Online: https://doi.org/10.1117/12.898084.

Gago, J., Douthe, C., Coopman, R., Gallego, P., Ribas-Carbo, M., Flexas, J., Escalona, J.,

Medrano, H., 2015. UAVs challenge to assess water stress for sustainable agriculture. Agric. Water Manag. 153, 9–19.
Garcia, D., 2010. Robust smoothing of gridded data in one and higher dimensions with missing values. Comput. Stat. Data Anal. 54 (4), 1167–1178.
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. Comput. Electron. Agric. 127, 625–632.
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.
Heiskanen, J., Rautiainen, M., Stenberg, P., Mõttus, M., Vesanto, V.-H., 2013. Sensitivity of narrowband vegetation indices to boreal forest lai, reflectance seasonality and species composition. ISPRS J. Photogramm. Remote Sens. 78, 1–14.
Ihuoma, S.O., Madramootoo, C.A., 2017. Recent advances in crop water stress detection. Comput. Electron. Agric. 141, 267–275.
Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., Ustin, S.L., 2009. Prospect+ sail models: a review of use for vegetation characterization. Remote Sens. Environ. 113, 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 Sens. Environ. 177, 220–236.
Jay, S., Gorretta, N., Morel, J., Maupas, F., Bendoula, R., Rabatel, G., Dutartre, D., Comar, A., Baret, F., 2017. Estimating leaf chlorophyll content in sugar beet canopies using millimeter-to centimeter-scale reflectance imagery. Remote Sens. Environ. 198, 173–186.
Katsoulas, N., Elvanidi, A., Ferentinos, K.P., Kacira, M., Bartzanas, T., Kittas, C., 2016. Crop reflectance monitoring as a tool for water stress detection in greenhouses: A review. Biosyst. Eng. 151, 374–398.
Kim, Y., Glenn, D.M., Park, J., Ngugi, H.K., Lehman, B.L., 2011. Hyperspectral image analysis for water stress detection of apple trees. Comput. Electron. Agric. 77 (2), 155–160.
Li, M., Cheng, Y., Zhao, H., 2004. Unlabeled data classification via support vector machines and k-means clustering. In: Proceedings of the International Conference on Computer Graphics, Imaging and Visualization, CGIV, pp. 183–186.
Linke, R., Richter, K., Haumann, J., Schneider, W., Weihs, P., 2008. Occurrence of repeated drought events: can repetitive stress situations and recovery from drought be traced with leaf reflectance? Periodicum biologorum 110 (3), 219–229.
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.
Peñuelas, J., Filella, I., 1998. Visible and near-infrared reflectance techniques for diagnosing plant physiological status. Trends Plant Sci. 3 (4), 151–156.
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. Int. J. Remote Sens. 14 (10), 1887–1905.
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 J. Photogramm. Remote Sens. 109, 88–97.
Römer, C., Wahabzada, M., Ballvora, A., Pinto, F., Rossini, M., Panigada, C., Behmann, J., Léon, J., Thurau, C., Bauckhage, C., et al., 2012. Early drought stress detection in cereals: simplex volume maximisation for hyperspectral image analysis. Funct. Plant Biol. 39 (11), 878–890.
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. Comput. Electron. Agric. 74 (1), 91–99.
Sarstedt, M., Mooi, E., 2014. Cluster analysis. In: A Concise Guide to Market Research. Springer, pp. 273–324.
Serrano, L., Ustin, S.L., Roberts, D.A., Gamon, J.A., Penuelas, J., 2000. Deriving water content of chaparral vegetation from aviris data. Remote Sens. Environ. 74 (3), 570–581.
Sun, C., Gao, X., Chen, X., Fu, J., Zhang, Y., 2016. Metabolic and growth responses of maize to successive drought and re-watering cycles. Agric. Water Manag. 172, 62–73.
Sun, C., Li, C., Zhang, C., Hao, L., Song, M., Liu, W., Zhang, Y., 2018. Reflectance and biochemical responses of maize plants to drought and re-watering cycles. Ann. Appl. Biol. 172 (3), 332–345.
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 J. Photogramm. Remote Sens. 135, 74–83.
Sun, P., Wahbi, S., Tsonev, T., Haworth, M., Liu, S., Centritto, M., 2014. On the use of leaf spectral indices to assess water status and photosynthetic limitations in olea europaea l. during water-stress and recovery. PloS One 9 (8), e105165.
Tang, T., Chen, S., Zhao, M., Huang, W., Luo, J., 2018. Very large-scale data classification based on k-means clustering and multi-kernel SVM. Soft. Comput. https://doi.org/10.1007/s00500-018-3041-0.
Uto, K., Kosugi, Y., 2013. Leaf parameter estimation based on leaf scale hyperspectral imagery. IEEE J. Sel.Top. Appl. Earth Observ. Remote Sens. 6 (2), 699–707.
Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G., Moreno, J., 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties–a review. ISPRS J. Photogramm. Remote Sens. 108, 273–290.
Xu, Z., Zhou, G., Shimizu, H., 2010. Plant responses to drought and rewatering. Plant Signal. Behav. 5 (6), 649–654.
Zegada-Lizarazu, W., Monti, A., 2013. Photosynthetic response of sweet sorghum to drought and re-watering at different growth stages. Physiol. Plant. 149 (1), 56–66.