Hindawi
BioMed Research International
Volume 2018, Article ID 5713158, 9 pages
https://doi.org/10.1155/2018/5713158

Research Article

Affordable Imaging Lab for Noninvasive Analysis of Biomass and Early Vigour in Cereal Crops

Rita Armonienė,[1] **Firuz Odilbekov**,[1] **Vivekanand Vivekanand**,[2] **and Aakash Chawade**[1]

[1]*Department of Plant Breeding, Swedish University of Agricultural Sciences, 23053 Alnarp, Sweden*
[2]*Centre for Energy and Environment, Malaviya National Institute of Technology, Jaipur, Rajasthan 302017, India*

Correspondence should be addressed to Aakash Chawade; aakash.chawade@slu.se

Received 21 October 2017; Revised 18 February 2018; Accepted 7 March 2018; Published 19 April 2018

Academic Editor: Marco Bazzicalupo

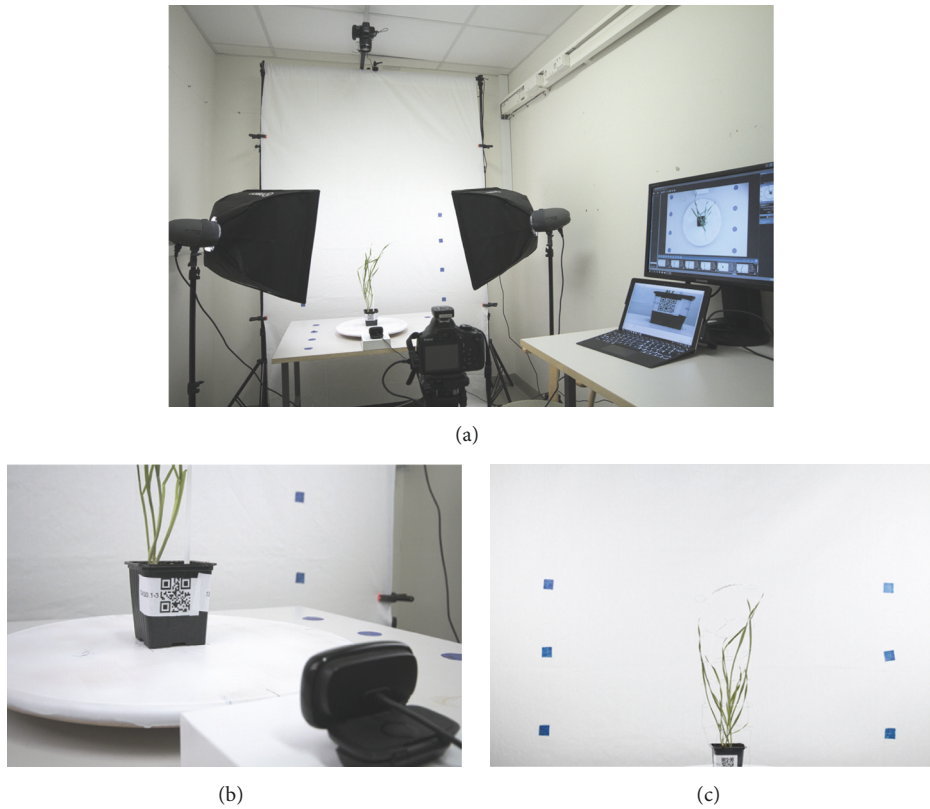
Plant phenotyping by imaging allows automated analysis of plants for various morphological and physiological traits. In this work, we developed a low-cost RGB imaging phenotyping lab (LCP lab) for low-throughput imaging and analysis using affordable imaging equipment and freely available software. LCP lab comprising RGB imaging and analysis pipeline is set up and demonstrated with early vigour analysis in wheat. Using this lab, a few hundred pots can be photographed in a day and the pots are tracked with QR codes. The software pipeline for both imaging and analysis is built from freely available software. The LCP lab was evaluated for early vigour analysis of five wheat cultivars. A high coefficient of determination (R^2 0.94) was obtained between the dry weight and the projected leaf area of 20-day-old wheat plants and R^2 of 0.9 for the relative growth rate between 10 and 20 days of plant growth. Detailed description for setting up such a lab is provided together with custom scripts built for imaging and analysis. The LCP lab is an affordable alternative for analysis of cereal crops when access to a high-throughput phenotyping facility is unavailable or when the experiments require growing plants in highly controlled climate chambers. The protocols described in this work are useful for building affordable imaging system for small-scale research projects and for education.

1. Introduction

Phenotyping morphological and physiological traits of plants is one of the most laborious tasks in plant breeding and thus automated high-throughput plant phenotyping (HTP) facilitates measurement of such traits. Visible light (RGB) imaging facilitates measurement of plant's morphological traits such as biomass, height, width, color, number of leaves, and roots to estimate plant growth rate, health, nutrition status, drought stress, water-use efficiency, nutrient-use efficiency, and early vigour [1–4]. 3D imaging can additionally measure traits such as leaf angle and leaf area which affect photosynthesis efficiency of the plants [5–7]. Hyperspectral, thermal, near-infrared (NIR), and fluorescent imaging are useful for detecting abiotic and biotic stresses [8–12]. While RGB imaging is a common feature in most facilities, some also offer fluorescence, thermal, NIR, or UV imaging. 3D imaging is popular and is available at bigger facilities such as Agrobios Plant Scanalyzer (APS) facility in Italy and The Plant Accelerator-Australian Plant Phenomics Facility. Several freely available software programs are available for image analysis [13] such as HTPheno [14], PlantCV [3], Easy Leaf Area [15], Integrated Analysis Platform [16], ImageHarvest [17], and Canopeo [18].

Whole plant biomass and growth rate at the seedling stage are traits that correlate well with early vigour and can be estimated by HTP in cereal crops. Higher early vigour is associated with higher water-use efficiency [19], nitrogen and phosphate uptake [20, 21], and weed competition [22]. 3D imaging with a NIR camera was used to measure early vigour traits such as leaf length and width and tillering in wheat [5]. Thus, HTP can aid in the evaluation of plants for early vigour based on plant biomass and growth rate both in the controlled conditions [17, 20, 23–26] and in the field [4, 26–28].

Major limiting factors for HTP in the controlled conditions are access to an imaging facility or the costs for establishing one. Thus, such facilities are established with the aim of providing phenotyping as a service and when



(a)

(b)

(c)

Figure 1: Photography setup for low-cost imaging lab. (a) Sample is placed on a rotating disc and is evenly light by two studio strobes. Images are taken with a side-view and a top-view digital camera. (b) QR code on the pot is read by a webcam. All three cameras are tethered to a computer. (c) An exemplified image.

a high-throughput and continuous use of the facility is anticipated. A low-cost phenotyping facility could be a viable alternative when (i) access to a high-throughput facility is unavailable, (ii) phenotyping occasionally for small-scale projects, and (iii) phenotyping is done in climate chambers with constrained spaces. Several custom made affordable imaging systems have been developed for studying drought stress in wheat [29] biotic stress and magnesium deficiency in common beans [30], cold tolerance in pea [31], and biomass in sorghum [32]. A review of various imaging systems and the studied traits was recently published [1]. To be cost-effective, sustainable, and time efficient, a low-cost phenotyping system must produce reproducible results with a reasonable amount of manual labor for setting up, running, and management of the system.

In this work, we have set up a low-cost RGB imaging phenotyping lab that includes automated plant tracking using QR code, imaging, and image analysis. The instructions for setting up a low-cost imaging lab and the data analysis pipeline are described. This system is demonstrated with early vigour analysis in wheat.

2. Materials and Methods

2.1. Setting Up an Imaging Studio. The low-cost phenotyping lab (LCP lab) consists of two studio strobes (Visico ELFIN VL-200 Plus) with an effect of 200 W each and color temperature of 5600 K. The two strobes are fitted with softboxes (50 × 70 cm) on light stands and placed one on each side of the plant at an angle of 45° illuminating both the plant and the background. The strobes contain integrated wireless radio receivers and the images are triggered with a wireless radio transmitter (V801TX, Visico, China) connected to a camera. The white background (1.5 × 2.5 m; FotoBestway Co., Ltd.) is hung on a telescopic boom placed on two light stands one on each side. Blue markers are pasted on the background to aid in framing of the images. Imaging is performed with two entry-level digital single lens reflex (DSLR) cameras (Canon 1300D, Canon, USA) and the 18–55 mm kit lens. The side-view camera is mounted on a tripod, while the top-view camera is mounted on a SpaceArm (Tristar). The distance between the side and the top-view camera and the pot is 1.5 m. A pot is placed on a rotating disc of diameter 38 cm (Snudda, IKEA) which is spray painted to white. Optionally, the pot can also be placed on a rotating turntable of similar diameter. For reading the QR code from the pot, a webcam (Logitech International S.A., USA) is placed 20 cm away from the pot. All three cameras are connected to a computer and tethered (Figure 1).

Pots can be tagged with QR code containing desired metadata such as the cultivar name, replicate number, and treatment. QR code is generated with Bytescout Barcode

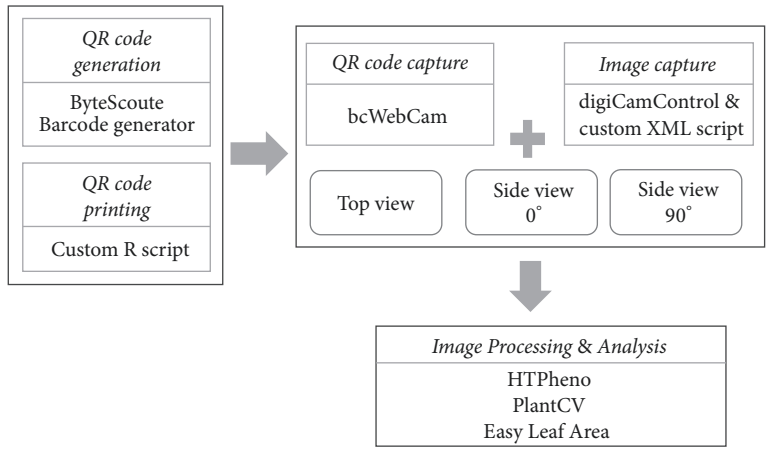


Figure 2: Pipeline for imaging and data processing. QR codes are generated with ByteScoute and printed with a custom R script. Tethered imaging is done with digiCamControl with custom XML script for time-lapse imaging and QR code capture with bcWebCam. Image processing and analysis are done with HTPheno, PlantCV, and Easy Leaf Area software.

generator (https://www.bytescout.com) and printed on self-adhesive labels using a custom R script (Supplementary File 1). Once a pot is placed on a rotating disc, the QR code is read by the webcam placed in front of the pot and operated by the software bcWebCam (http://www.bcwebcam.de). The information read from the QR code is automatically transferred from bcWebCam to the tethering software digiCamControl (http://www.digicamcontrol.com) (Figure 2). The two DSLR cameras are tethered to digiCamControl which takes a series of three images using a custom XML script (Supplementary file 2) such that the first image is taken by the side-view camera followed by a one-second wait and a second image by the top-view camera followed by a six-second wait to manually rotate the disc with pot by 90° and a third image is taken again by the side-view camera. This imaging series is set in a loop in the XML script which can trigger the cameras for a designated number of times. Side-view images are taken with the camera with optimum settings (focal length of 43 mm, ISO 400, F-Stop f/10, and exposure time of 1/100 seconds), while the top-view images are taken with slightly different settings (focal length of 43 mm, ISO 400, F-Stop f/11, and exposure time of 1/60 seconds).

It is important to evenly light the plants and the background to avoid shadows caused by uneven lighting. Strong shadows were avoided in the images and were controlled by adjusting camera settings and exposure. Oversaturation of the images was avoided by referring to the histograms. Upon optimizing lighting, camera settings, and camera distance, custom white balance was obtained by photographing just the background. All images were thereafter captured with the custom white balance to avoid variation in the white balance and light intensity in the images. The images were stored directly to the computer and are of resolution 72 dpi, size 1920 × 1280 pixels in JPEG format, and the files are named with the data read from the QR code. The pots were black in color and did not interfere with image processing.

2.2. Image Processing. Images are first manually inspected to remove those that are improperly lit, plants being out of frame or blurred. Finally, only plants with all three images are retained for further analysis. Thereafter, the images are processed with the software HTPheno [14], PlantCV [3], or Easy Leaf Area [15] on a laptop with a dual-core i7 Intel processor and 16 GB RAM. Optimization of image analysis parameters was carried out for each software program. For HTPheno, the object classes for the side and top-view were background, pot, stickers, and the plant. Thereafter, for each object, the corresponding image coordinates were assigned separately for the top and the side-view. The color ranges for the four objects were assigned with HTPcalib. The PlantCV analysis pipeline was built as described previously [3]. Briefly, (i) conversion of RGB images to HSV color space, (ii) isolation and thresholding the saturation channel, (iii) conversion of RGB to LAB color space and isolation and thresholding of the blue-yellow channel, (iii) joining the saturation and the blue-yellow channels to mask the RGB image followed by extraction and thresholding of green-magenta and blue-yellow channels, and (iv) joining saturation and blue-yellow channels, masking the previously masked image and object identification. For Easy Leaf Area, parameters were leaf minimum green RGB value: 23; G/R: 1.0; G/B: 1.01; scale minimum red RGB value: 33; scale red ratio: 1.0; processing speed: 2.0; minimum leaf pixel: 300.

2.3. Plant Material for the Case Study. Five winter wheat cultivars Stigg, Kranich, Nelson, Nimbus, and Target were chosen for early vigour evaluation as these are popular varieties for commercial cultivation. The seeds were germinated for two days on a moist filter paper in Petri dishes. Germinated seeds were sown in plastic pots (0.4 l) filled with peat substrate Blomjord Exclusive (Emmaljunga Torvmull AB, Sweden). One seed per pot of each genotype was sown in ten replications for each time point. Plants were grown at 20°C in a greenhouse with 16 h photoperiod and light intensity of 250 μmol m^2 s^{-1}. All pots were soaked in equal quantity of water every three days. Seedling images were taken at 10 and 20 days upon sowing. The plants were

Table 1: Summary of results from the three software programs. Coefficient of determination (R^2) obtained from simple linear regression from projected leaf area (PA) from each software programs and the dry weight of the plants. R^2 of relative growth rate is obtained by regression of dry weight RGR (RGR_{DW}) and the projected area RGR (RGR_{PA}).

	10 days	20 days	R^2 of relative growth rate
HTPheno	0.90	0.94	0.90
PlantCV	0.88	0.91	0.89
EasyLeafArea	0.81	0.91	0.86

photographed from two side views and one top view. For each timepoint, the first images were taken by including a ruler to adjust for any changes in the camera distance. The conversion from pixels to centimeter was performed separately for 10 and 20 days timepoint and for two centimeter distance, it was 51 pixels for the side view and 53 pixels for the top view for the two timepoints, respectively. For dry weight analysis, after imaging, the shoots were cut and wrapped in aluminum foil then dried at 100°C for 48 h and weighed. Images were manually filtered out to remove those with plants out of frame. From the 10-day timepoint, no images were removed but at the 20-day timepoint, one plant sample each from Stigg, Target, and Nimbus and two plant samples from Nelson had to be removed due to plants being out of frame in the top view. Thus, for the 10-day timepoint, there were 10 replicates each, while there were 8 replicates for the 20-day timepoint.

2.4. Statistical Analysis. Relative growth rate for each plant based on the dry weight (RGR_{DW}) [23, 33] was estimated with (1), where W_1 and W_2 are dry weights of each plant at timepoints t_1 (10 days) and t_2 (20 days), respectively.

$$RGR_{DW} = \frac{\ln(W_2/W_1)}{t_2 - t_1}. \quad (1)$$

Projected leaf area (PA) is defined by (2), where PA is the projected leaf area obtained from each image and n is the number of angles photographed for a given plant

$$PA = \sum_{1}^{n=3} pa. \quad (2)$$

Relative growth rate for each plant based on the projected leaf area (PA) was estimated with (3) where PA_1 and PA_2 are the projected areas of plants at timepoints t_1 (10 days) and t_2 (20 days), respectively.

$$RGR_{PA} = \frac{\ln(PA_2/PA_1)}{t_2 - t_1}. \quad (3)$$

3. Results

The aim of this work was to build a RGB imaging based phenotyping system that is affordable and portable for low-throughput imaging. LCP lab integrates three cameras, two for imaging and one for reading the QR code (Figure 1). The measurements obtained from the three software programs vary and thus the selection of the software depends on the overall goal of the experiment. HTPheno generates plant height, width, and projected shoot area from the side-view images and x-extent, y-extent, diameter, and projected shoot area from the top-view images. PlantCV generates over 30 different measurements. Easy Leaf Area estimates the projected leaf area in both side- and top-view images. Measurements from the side and top views can be integrated prior to further analysis.

3.1. Case Study: Early Vigour Analysis in Wheat. To estimate the accuracy of imaging with the LCP lab, early vigour analysis was evaluated from 10–20-day-old wheat plants from five cultivars. Images were analyzed with three different software programs HTPheno, plantCV, and Easy Leaf Area. The parameters were adjusted for each software program to maximize the leaf area detection and minimize the detection of other non-plant objects in the images. The three software programs evaluated here have led to an output consisting of a text file with the measurements and images marked with the identified plant regions (Figure 3).

Analysis of variance (ANOVA) from the dry weight data showed that the effect of the genotype (cultivars) on the early vigour was significant ($p < 0.001$) and that there was a significant difference in the mean weight between Stigg and Kranich and Stigg and Nimbus at both timepoints (Tukey's HSD adjusted $p < 0.05$) (Figures 4(a) and 4(c)). ANOVA from HTPheno projected leaf area (PA) showed that the effect of the genotype on the early vigour was significant ($p < 0.001$) and based on the projected leaf area (PA) from HTPheno results, Stigg had significantly different early vigour compared to Kranich and Nimbus at all timepoints (Tukey's HSD adjusted $p < 0.05$) (Figures 4(b) and 4(d)).

A simple linear regression was calculated to estimate the coefficient of determination between the measured dry weights and the projected leaf area (PA) for each plant (Table 1). Results from HTPheno had the highest R^2 at 20-day timepoint. Overall, all three software programs had higher R^2 at 20 days compared to the 10-day timepoint. To evaluate if imaging of all three angles is required, R^2 was obtained for dry weight and images obtained from each of the angles separately or combinations of any two angles (Table 2). The results show that across all timepoints, higher R^2 is obtained from images from all three angles.

Relative growth rate estimated from dry weight (RGR_{DW}) and HTPheno projected area (RGR_{PA}) suggests that the cultivar Target has significantly different growth rate (Tukey's HSD $p < 0.05$) from cultivar Nimbus (Figures 5(a) and 5(b)). A significant regression equation was obtained ($p < 0.001$) between the dry weight and HTPheno projected leaf area.

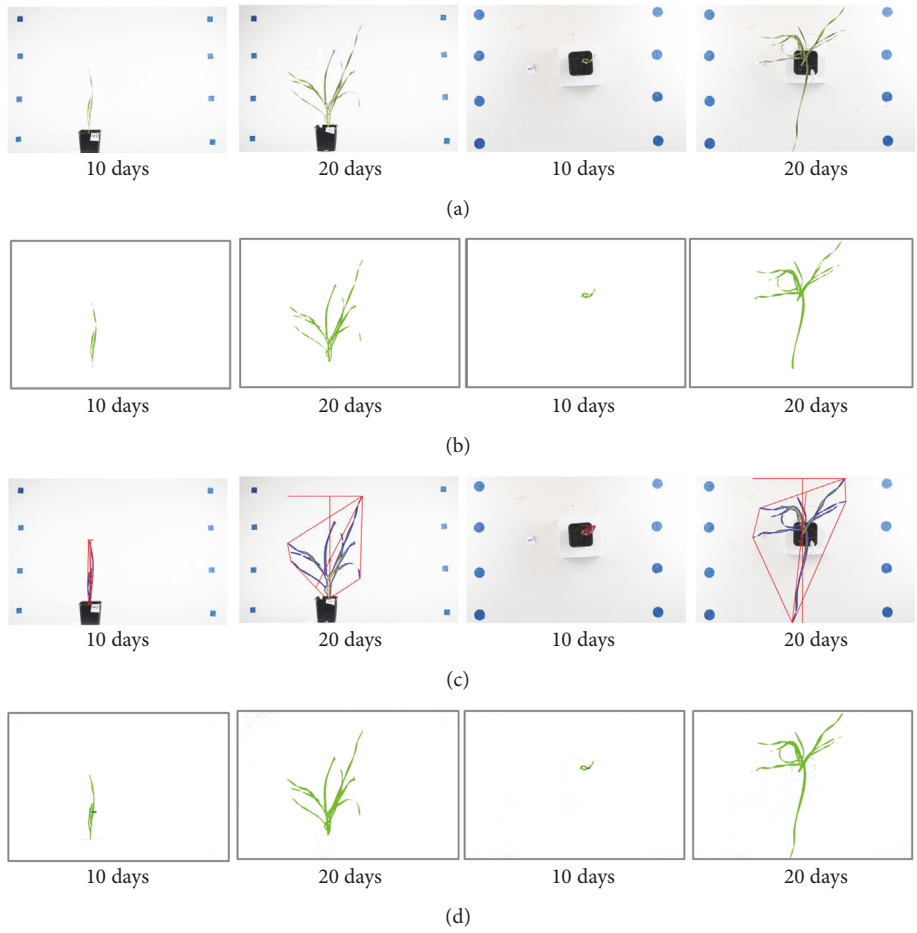


Figure 3: Side- and top-view images of cultivar Nelson at three timepoints after sowing. Images are (a) unprocessed or processed by software (b) HTPheno, (c) plantCV, or (d) Easy Leaf Area.

Table 2: Results from HTPheno software for images from different combinations of angles and coefficient of determination (R^2) obtained from simple linear regression between the projected leaf area (PA) and the dry weight of the plants.

	10 days	20 days
Side 1	0.76	0.85
Side 2	0.80	0.83
Top	0.42	0.72
Side 1 + side 2	0.88	0.89
Side 1 + top	0.77	0.91
Side 1 + side 2 + top	0.90	0.94

Based on overall growth at 20 days and the relative growth rate analysis, it can be suggested that Stigg has the most growth and higher relative growth rate compared to Nimbus.

4. Discussion

Measurement accuracy and reproducibility are the key factors for the evaluation of a phenotyping pipeline. The highest coefficient of determination ($R^2 = 0.94$) obtained in this work was with images from 20-day-old plants analyzed with the HTPheno pipeline (Table 1) and for the relative growth rate, R^2 of 0.9 was obtained for dry weight and HTPheno projected leaf area. However, results from plantCV were only slightly lower. PlantCV offers several customizations, takes relatively less time for analysis, and can be a suitable alternative to HTPheno for larger data sets. In a previous study, phenotyping pipeline consisting of a commercial imaging system Scanalyzer 3D (LemnaTec GmbH, Aachen, Germany) and an open source analysis pipeline IAP was used to photograph maize plants and obtained R^2 of 0.84 and 0.94 with the dry and fresh weight, respectively, with RGB imaging [16]. In another study, 373 rice genotypes were photographed with

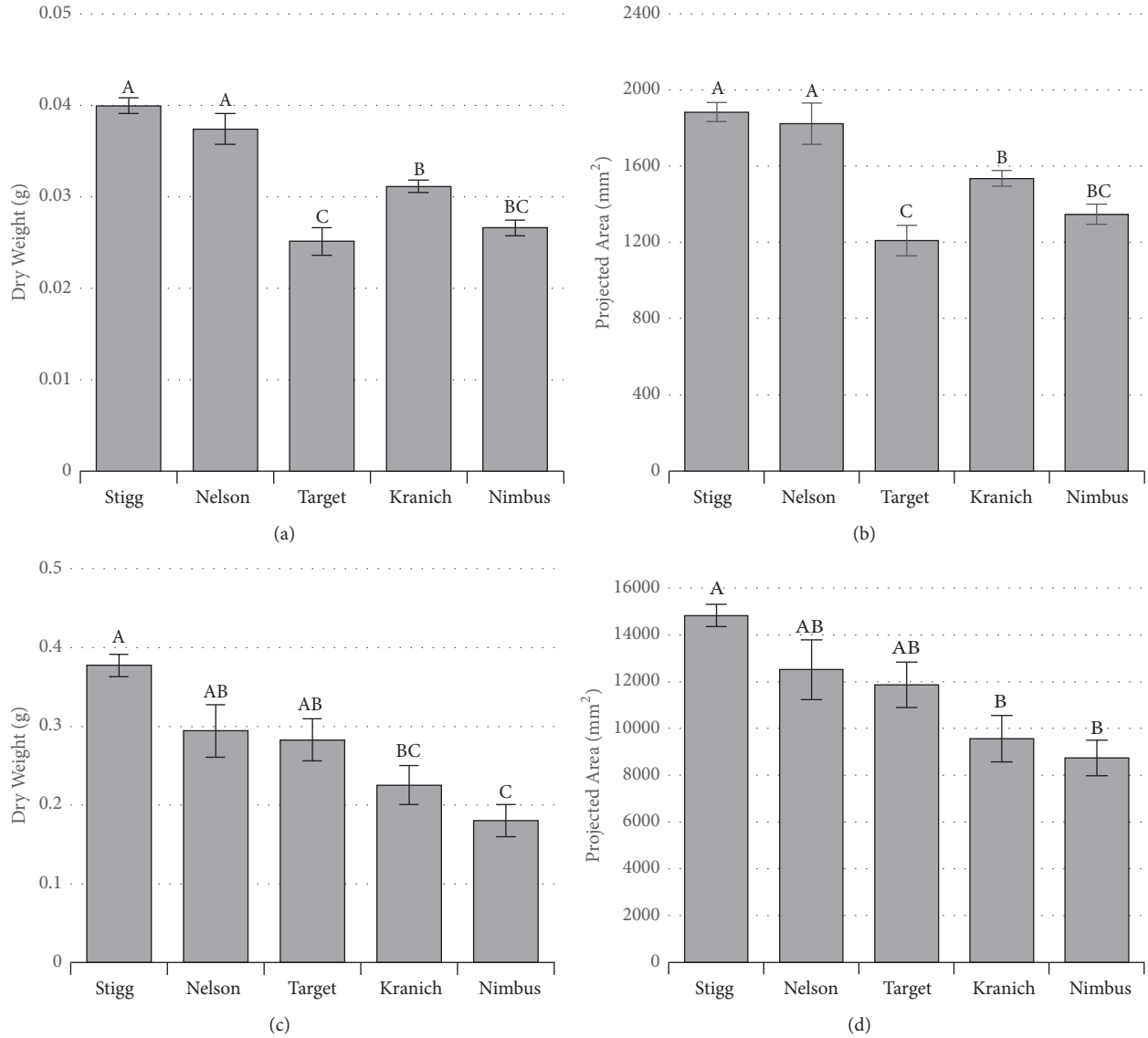


Figure 4: Mean dry weight and mean area of the plants from five cultivars. (a, c) Mean dry weights plants from five cultivars. (b, d) Mean projected leaf area of plants from five cultivars. (a-b) 10 days ($n = 10$); and (c-d) 20 days ($n = 8$) after sowing. Cultivars are sorted based on the mean growth at 20 days. Statistically significant differences (Tukey's HSD adjusted $p < 0.05$) in the means are denoted by different letters above the bars. Error bars are standard error.

Scanalyzer 3D system and analyzed with the open source ImageHarvest analysis pipeline and obtained R^2 of 0.93 with the shoot dry weight, while with the commercial data analysis pipeline LemnaGrid, R^2 of 0.94 was obtained [17]. Imaging of 320 wheat plants with a Scanalyzer 3D system and analysis with LemnaTec 3D Image Analyzer was done and obtained R^2 of 0.96 with a linear model [34]. 3D imaging of rapeseed with PlantEye F300 developed by Phenospex (Heerlen, the Netherlands) resulted in R^2 of 0.97 with shoot dry weight and 3D leaf area [35]. In the current study, lower R^2 at the earlier timepoint can be attributed to technical errors or lower plant to background ratio due to the smaller size of the plants. Focal distance can be adjusted to increase the plant to background ratio but a correction for the field of view needs to be done prior to comparing data from different timepoints [3].

Out of the three software programs tested here, EasyLeafArea is the easiest to set up and has a graphical user interface and a few parameters to be adjusted. It also detects a red object with known dimensions in an image to estimate absolute measurements of the leaf area automatically. PlantCV requires additional dependencies but the installation is well documented. It does not have a graphical user interface and some knowledge of programming is essential to optimize the program for images taken with a new setup. PlantCV also allows detection of objects with known dimensions and a set of computer code can be written to estimate the absolute measurements by including objects with known dimensions. HTPheno is a plugin for the software ImageJ [36]. Optimization of HTPheno for a new set of images is slightly tedious as the color profiles for the plants need to

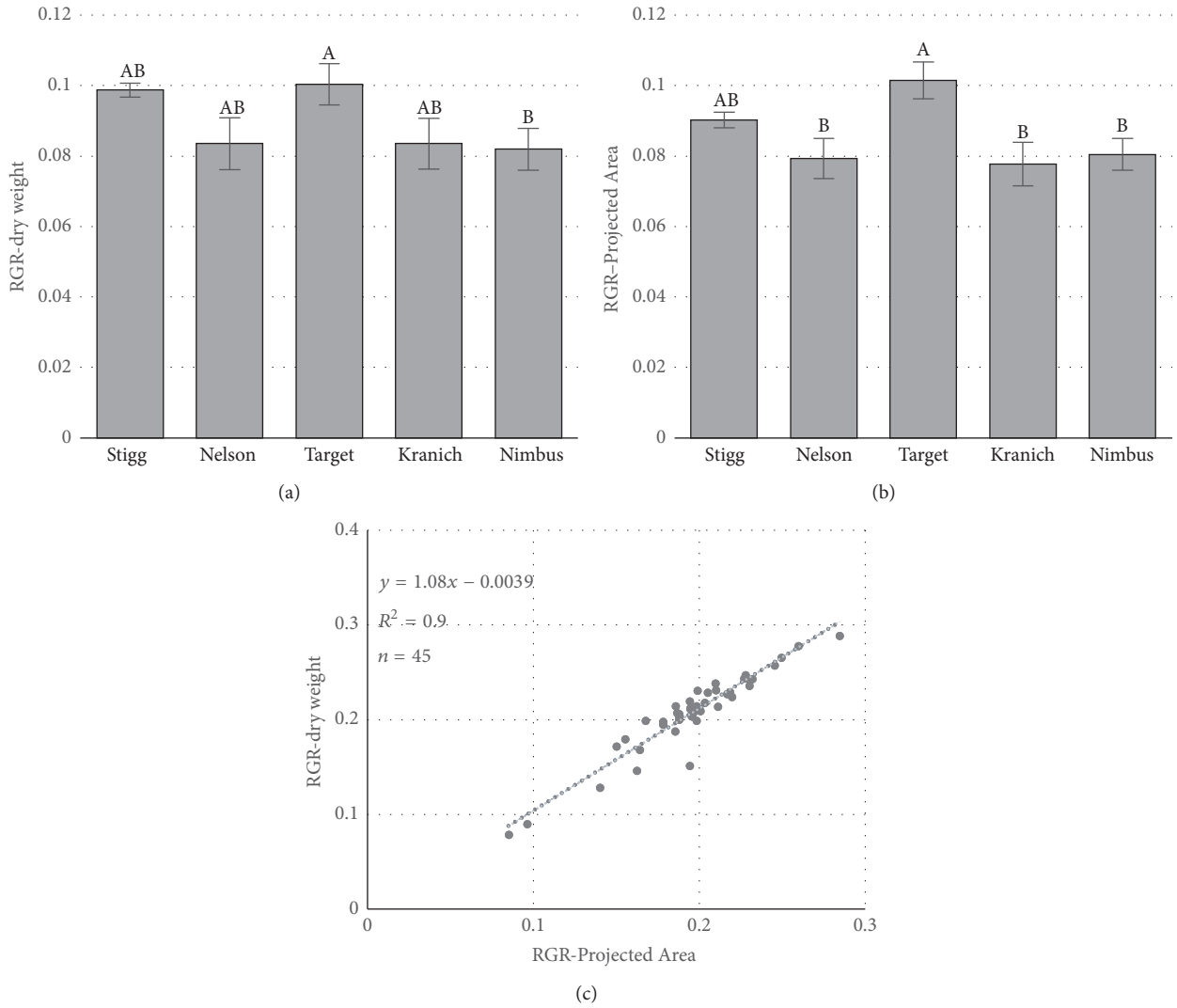


Figure 5: Relative growth rate of the plants from five cultivars based on (a) dry weight and (b) projected leaf area (PA) from HTPheno. Statistically significant differences (Tukey's HSD adjusted $p < 0.05$) in the means are denoted by different letters above the bars. Error bars are standard error. (c) Simple linear regression between RGR_{DW} from dry weight and RGR_{PA} from the projected leaf area from HTPheno.

be manually identified. However, based on this work, the results produced from HTPheno were the best among the three (Table 1). HTPheno also allows estimation of absolute measurements; however, unlike the previous two software programs, in HTPheno, the pixel to centimeter conversion needs to be manually entered. All three software programs are very well documented. Although, in this work, HTPheno performed the best, further optimization of all three software parameters is possible which may improve the results in future studies.

Major advantage of the LCP lab is that it requires just around 10–12 m^2 of working space. The software digiCamControl supports a range of cameras and although we have used an entry-level DSLR camera Canon 1300D (~350 USD), the total cost can be further reduced with an in-expensive consumer camera model such as Nikon Coolpix S5300 (~ 200 USD) supported by the digiCamControl. The studio equipment used in this work is easily available worldwide or can be replaced with the equivalent equipment from other brands. More cameras can be added such as those modified for taking Normalized Difference Vegetative Index (NDVI) pictures. In this work, we used a manual rotating disc, but a semi-automated or a fully automated and programmable rotating base can also be installed for further automation of imaging. Pots were tracked with QR codes and read with a webcam and the metadata is stored in the image filename. This enables automated file naming and classification that simplifies the whole image processing. For the presented case study, QR codes were not used but are described here for ease of implementation in building new imaging systems of this kind.

There are several commercial or custom made HTP facilities available featuring RGB, thermal, infrared, or fluorescence imaging as reviewed earlier [1, 11]. These state-of-the-art HTP facilities are capable of imaging hundreds of plants day and night. In some facilities, plant watering

and imaging are fully automated through pot weighing and conveyor belts, thus requiring minimal manual labour. These facilities are however expensive to install and maintain requiring considerable investment and resources. Also, these HTP facilities, although they can be modular, still require much bigger working space area mainly for the conveyor belt and are not portable. The proposed LCP lab here although lacks many of the key features available in the large-scale systems, the results obtained in this work suggest that a low-cost system can be a viable option in cases where large-scale facilities are not accessible. The smaller size and portability allow the LCP lab to be installed in smaller walk-in growth chambers which enables imaging plants grown in highly controlled growth conditions. It could also be useful for low-throughput phenotyping projects and/or education.

5. Conclusions

We have developed a low-cost RGB imaging phenotyping lab which integrates both imaging and analysis and the detailed description for setting up such a system is provided. LCP lab is a reliable and sustainable option for performing imaging based analysis of morphological traits in cereal crops. LCP lab offers flexibility with the choice of the imaging equipment and the analysis pipeline. It could be a suitable alternative for performing small-scale phenotyping projects or for smaller laboratories or academic institutions with limited resources. Having access to a high-end high-throughput phenotyping facility is still a bottleneck, and thus, a low-cost portable system can help circumvent such limitations.

Disclosure

The funding agencies were not involved in any capacity in the experiments conducted in this work.

Conflicts of Interest

The authors have no conflicts of interest.

Acknowledgments

This work was supported by funding from Lantmännen Research Foundation and Partnerskap Alnarp to Aakash Chawade and a Postdoctoral Fellowship from the Swedish Institute Visby Program to Rita Armonienė.

Supplementary Materials

Description of the two files: Supplementary File 1: R script for preparing barcodes. Supplementary File 2: script for the software digiCamControl for imaging. *(Supplementary Materials)*

References

[1] J. F. Humplík, D. Lazár, A. Husičková, and L. Spíchal, "Automated phenotyping of plant shoots using imaging methods for analysis of plant stress responses - A review," *Plant Methods*, vol. 11, no. 1, article no. 29, 2015.

[2] D. Chen, K. Neumann, S. Friedel et al., "Dissecting the phenotypic components of crop plant growthand drought responses based on high-throughput image analysis w open," *The Plant Cell*, vol. 26, no. 12, pp. 4636–4655, 2014.

[3] N. Fahlgren, M. Feldman, M. A. Gehan et al., "A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in Setaria," *Molecular Plant*, vol. 8, no. 10, pp. 1520–1535, 2015.

[4] S. Kipp, B. Mistele, P. Baresel, and U. Schmidhalter, "High-throughput phenotyping early plant vigour of winter wheat," *European Journal of Agronomy*, vol. 52, pp. 271–278, 2014.

[5] T. Duan, S. C. Chapman, E. Holland, G. J. Rebetzke, Y. Guo, and B. Zheng, "Dynamic quantification of canopy structure to characterize early plant vigour in wheat genotypes," *Journal of Experimental Botany*, vol. 67, no. 15, pp. 4523–4534, 2016.

[6] S. Paulus, J. Behmann, A.-K. Mahlein, L. Plümer, and H. Kuhlmann, "Low-cost 3D systems: Suitable tools for plant phenotyping," *Sensors*, vol. 14, no. 2, pp. 3001–3018, 2014.

[7] J. Li and L. Tang, "Developing a low-cost 3D plant morphological traits characterization system," *Computers and Electronics in Agriculture*, vol. 143, pp. 1–13, 2017.

[8] D. K. Großkinsky, J. Svensgaard, S. Christensen, and T. Roitsch, "Plant phenomics and the need for physiological phenotyping across scales to narrow the genotype-to-phenotype knowledge gap," *Journal of Experimental Botany*, vol. 66, no. 18, pp. 5429–5440, 2015.

[9] L. Li, Q. Zhang, and D. Huang, "A review of imaging techniques for plant phenotyping," *Sensors*, vol. 14, no. 11, pp. 20078–20111, 2014.

[10] N. Fahlgren, M. A. Gehan, and I. Baxter, "Lights, camera, action: high-throughput plant phenotyping is ready for a close-up," *Current Opinion in Plant Biology*, vol. 24, pp. 93–99, 2015.

[11] M. M. Rahaman, D. Chen, Z. Gillani, C. Klukas, and M. Chen, "Advanced phenotyping and phenotype data analysis for the study of plant growth and development," *Frontiers in Plant Science*, vol. 6, article no. 619, 2015.

[12] F. Fiorani, U. Rascher, S. Jahnke, and U. Schurr, "Imaging plants dynamics in heterogenic environments," *Current Opinion in Biotechnology*, vol. 23, no. 2, pp. 227–235, 2012.

[13] G. Lobet, X. Draye, and C. Périlleux, "An online database for plant image analysis software tools," *Plant Methods*, vol. 9, no. 1, article no. 38, 2013.

[14] A. Hartmann, T. Czauderna, R. Hoffmann, N. Stein, and F. Schreiber, "HTPheno: An image analysis pipeline for high-throughput plant phenotyping," *BMC Bioinformatics*, vol. 12, article no. 148, 2011.

[15] H. M. Easlon and A. J. Bloom, "Easy leaf area: Automated digitalimage analysis for rapid and accurate measurement of leaf area," *Applications in Plant Sciences*, vol. 2, no. 7, Article ID 1400033, 2014.

[16] C. Klukas, D. Chen, and J.-M. Pape, "Integrated analysis platform: An open-source information system for high-throughput plant phenotyping," *Plant Physiology*, vol. 165, no. 2, pp. 506–518, 2014.

[17] A. C. Knecht, M. T. Campbell, A. Caprez, D. R. Swanson, and H. Walia, "Image Harvest: An open-source platform for high-throughput plant image processing and analysis," *Journal of Experimental Botany*, vol. 67, no. 11, pp. 3587–3599, 2016.

[18] A. Patrignani and T. E. Ochsner, "Canopeo: A powerful new tool for measuring fractional green canopy cover," *Agronomy Journal*, vol. 107, no. 6, pp. 2312–2320, 2015.

[19] R. A. Richards, G. J. Rebetzke, A. G. Condon, and A. F. Van Herwaarden, "Breeding opportunities for increasing the efficiency of water use and crop yield in temperate cereals," *Crop Science*, vol. 42, no. 1, pp. 111–121, 2002.

[20] P. R. Ryan, M. Liao, E. Delhaize et al., "Early vigour improves phosphate uptake in wheat," *Journal of Experimental Botany*, vol. 66, no. 22, pp. 7089–7100, 2015.

[21] J. Pang, J. A. Palta, G. J. Rebetzke, and S. P. Milroy, "Wheat genotypes with high early vigour accumulate more nitrogen and have higher photosynthetic nitrogen use efficiency during early growth," *Functional Plant Biology*, vol. 41, no. 2, pp. 215–222, 2014.

[22] N.-O. Bertholdsson, "Early vigour and allelopathy - Two useful traits for enhanced barley and wheat competitiveness against weeds," *Weed Research*, vol. 45, no. 2, pp. 94–102, 2005.

[23] E. H. Neilson, A. M. Edwards, C. K. Blomstedt, B. Berger, B. L. Møller, and R. M. Gleadow, "Utilization of a high-throughput shoot imaging system to examine the dynamic phenotypic responses of a C4 cereal crop plant to nitrogen and water deficiency over time," *Journal of Experimental Botany*, vol. 66, no. 7, pp. 1817–1832, 2015.

[24] M. W. Ter Steege, F. M. Den Ouden, H. Lambers, P. Stam, and A. J. M. Peeters, "Genetic and physiological architecture of early vigor in Aegilops tauschii, the D-genome donor of hexaploid wheat. A quantitative trait loci analysis," *Plant Physiology*, vol. 139, no. 2, pp. 1078–1094, 2005.

[25] M. C. Rebolledo, M. Dingkuhn, B. Courtois et al., "Phenotypic and genetic dissection of component traits for early vigour in rice using plant growth modelling, sugar content analyses and association mapping," *Journal of Experimental Botany*, vol. 66, no. 18, pp. 5555–5566, 2015.

[26] M. L. Maydup, C. Graciano, J. J. Guiamet, and E. A. Tambussi, "Analysis of early vigour in twenty modern cultivars of bread wheat (Triticum aestivum L.)," *Crop & Pasture Science*, vol. 63, no. 10, pp. 987–996, 2012.

[27] J. Svensgaard, T. Roitsch, and S. Christensen, "Development of a mobile multispectral imaging platform for precise field phenotyping," *Agronomy*, vol. 4, no. 3, pp. 322–336, 2014.

[28] F. Liebisch, N. Kirchgessner, D. Schneider, A. Walter, and A. Hund, "Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach," *Plant Methods*, vol. 11, no. 1, article no. 9, 2015.

[29] E. Fehér-Juhász, P. Majer, L. Sass et al., "Phenotyping shows improved physiological traits and seed yield of transgenic wheat plants expressing the alfalfa aldose reductase under permanent drought stress," *Acta Physiologiae Plantarum*, vol. 36, no. 3, pp. 663–673, 2014.

[30] L. Chaerle, D. Hagenbeek, X. Vanrobaeys, and D. Van Der Straeten, "Early detection of nutrient and biotic stress in Phaseolus vulgaris," *International Journal of Remote Sensing*, vol. 28, no. 16, pp. 3479–3492, 2007.

[31] J. F. Humplík, D. Lazár, T. Fürst, A. Husicková, M. Hýbl, and L. Spíchal, "Automated integrative high-throughput phenotyping of plant shoots: A case study of the cold-tolerance of pea (Pisum sativum L.)," *Plant Methods*, vol. 11, no. 1, article no. 20, 2015.

[32] Y. S. Chung, S. C. Choi, R. R. Silva, J. W. Kang, J. H. Eom, and C. Kim, "Case study: Estimation of sorghum biomass using digital image analysis with Canopeo," *Biomass & Bioenergy*, vol. 105, pp. 207–210, 2017.

[33] R. Hunt, *Plant growth curves: the functional approach to plant growth analysis*, Arnold, London, UK, 1982.

[34] M. R. Golzarian, R. A. Frick, K. Rajendran et al., "Accurate inference of shoot biomass from high-throughput images of cereal plants," *Plant Methods*, vol. 7, no. 1, article no. 2, 2011.

[35] K. H. Kjaer and C.-O. Ottosen, "3D laser triangulation for plant phenotyping in challenging environments," *Sensors*, vol. 15, no. 6, pp. 13533–13547, 2015.

[36] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nature Methods*, vol. 9, no. 7, pp. 671–675, 2012.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 23 May 2018
doi: 10.3389/fpls.2018.00685

Proximal Phenotyping and Machine Learning Methods to Identify Septoria Tritici Blotch Disease Symptoms in Wheat

*Firuz Odilbekov[1], Rita Armonienė[1], Tina Henriksson[2] and Aakash Chawade[1]**

[1] Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden, [2] Lantmännen Lantbruk, Svalöv, Sweden

Phenotyping with proximal sensors allow high-precision measurements of plant traits both in the controlled conditions and in the field. In this work, using machine learning, an integrated analysis was done from the data obtained from spectroradiometer, infrared thermometer, and chlorophyll fluorescence measurements to identify most predictive proxy measurements for studying Septoria tritici blotch (STB) disease of wheat. The random forest (RF) models for chlorosis and necrosis identified photosystem II quantum yield (QY) and vegetative indices (VIs) associated with the biochemical composition of leaves as the top predictive variables for identifying disease symptoms. The RF model for chlorosis was validated with a validation set (R^2: 0.80) and in an independent test set (R^2: 0.55). Based on the results, it can be concluded that the proxy measurements for photosystem II, chlorophyll content, carotenoid, and anthocyanin levels and leaf surface temperature can be successfully used to detect STB. Further validation of these results in the field will enable application of these predictive variables for detection of STB in the field.

Keywords: Septoria tritici blotch, wheat, proximal phenotyping, disease detection, machine learning, random forest, machine learning

OPEN ACCESS

Edited by:
Sebastien Christian Carpentier, Bioversity International, Belgium

Reviewed by:
Jin Chen, University of Kentucky, United States
Sébastien Guyader, Institut National de la Recherche Agronomique (INRA), France

****Correspondence:***
Aakash Chawade
aakash.chawade@slu.se

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *09 December 2017*
Accepted: *04 May 2018*
Published: *23 May 2018*

Citation:
Odilbekov F, Armonienė R, Henriksson T and Chawade A (2018) Proximal Phenotyping and Machine Learning Methods to Identify Septoria Tritici Blotch Disease Symptoms in Wheat. Front. Plant Sci. 9:685. doi: 10.3389/fpls.2018.00685

INTRODUCTION

Septoria tritici blotch (STB) caused by *Zymoseptoria tritici* is currently one of the most devastating foliar diseases of wheat in Northwestern Europe causing yield losses every year (Fones and Gurr, 2015; Chawade et al., 2018). It is a hemibiotrophic fungus which penetrates host leaves through stomata and grows very slowly in the intercellular spaces of the mesophyll cells. The latent phase varies between 14–28 days under field conditions and 9–14 days under laboratory conditions (Kema et al., 1996; Shetty et al., 2003; Keon et al., 2007). This symptomless period has been referred to as 'biotrophic' (Kema et al., 2000), however, after more detailed transcriptomic and metabolic analysis, this term has become debatable (Rudd et al., 2015; Sánchez-Vallet et al., 2015). After a latent period, the fungus switches to necrotrophic phase and the infected leaves become chlorotic and develop into necrotic irregularly-shaped blotches (lesions) in which fungal asexual fruiting sporulation structures called pycnidia develop (Steinberg, 2015; Kettles and Kanyuka, 2016).

A cultivar with a high level of resistance can provide an effective mode to control the disease severity, but so far, cultivars with complete resistance are not developed (Chartrain et al., 2004).

STB is typically controlled by fungicides and due to the intensive chemical control, *Z. tritici* populations can rapidly evolve resistance to fungicides (Goodwin et al., 2011). For STB, best practice requires that the fungicides should be sprayed early on in the latent period, as the fungicide application has limited effectiveness in the necrotrophic phase (Fones and Gurr, 2015). Detection of STB in the latent stage can provide more efficient disease control by fungicides, thus minimizing directional selection that favors mutations encoding high level of resistance to fungicides in the *Z. tritici* populations.

Optical imaging techniques such as RGB, thermal, fluorescence, multi- and hyper-spectral imaging were applied to detect various plant diseases (Mahlein, 2016). Colonization of fungal pathogens cause multiple biochemical, physiological and morphological alterations in leaf tissue and it can be inferred from the reflectance of light at visible (VIS, 400–700 nm) and near-infrared (NIR, 700–2000 nm) regions of the electromagnetic spectrum. Hyperspectral imaging can be used to detect foliar diseases in early pathogenesis stage before visible symptoms appear (Kuska et al., 2015; Xie et al., 2016; Thomas et al., 2017). Hyperspectral imaging in VIS/NIR ranges was demonstrated as a powerful tool for detection and/or differentiation of foliar fungal diseases in barley (Kuska et al., 2015; Thomas et al., 2017), cucumber (Berdugo et al., 2014), sugar beet (Mahlein et al., 2012b), wheat (Ashourloo et al., 2014a,b; Cao et al., 2015; Iori et al., 2015), tomato (Xie et al., 2016), oilseed rape (Baranowski et al., 2015), and strawberries (Yeh et al., 2016).

Visible range is mainly influenced by leaf pigments like chlorophyll and carotenoid content and NIR is influenced by leaf structure, internal scattering processes and by leaf water content (Mahlein, 2016). Spectral vegetation indices (VIs) are mathematical equations and transformations derived from two or more wavelengths in the electromagnetic spectrum (Araus et al., 2001). Application of VIs is a common approach to investigate or identify changes in plant physiology and morphology. VIs were developed as proxies to evaluate various plant properties such as leaf area (Rouse et al., 1974), water content (Penuelas et al., 1995), and leaf pigment content (Gitelson et al., 2002; Sims and Gamon, 2002). Among the indices developed, NDVI (Normalized Difference Vegetation Index) is most commonly used as it can estimate nutrient requirements of plants and is thus used for optimizing fertilizer input in the fields (Raun et al., 2002). More than 100 VIs have been developed so far and summarized earlier (Devadas et al., 2008; Agapiou et al., 2012; Pietragalla et al., 2012; Lehnert et al., 2017).

Advances in sensor phenotyping technologies will generate big data. Therefore, extracting patterns and features from this big data requires machine learning (ML) tools (Singh et al., 2016). Application of the ML methods for prediction of various diseases from spectral reflectance data was reviewed recently (Lowe et al., 2017). Using spectral reflectance data, yellow rust in wheat was detected with quadratic discriminant analysis (Bravo et al., 2003; Lowe et al., 2017), multilayer perceptron (Moshou et al., 2004), and regression (Huang et al., 2007). While leaf rust was detected with maximum likelihood classification (Ashourloo et al., 2014b) and powdery mildew with Fishers linear discriminant analysis (Zhang et al., 2012). Thus, spectral reflectance phenotyping and ML methods hold promise for disease detection at early stages of infection.

Early detection of disease symptoms allows taking control measures to avoid further spread of pathogen and consequent yield losses. Disease monitoring methods are time-consuming, can be affected by subjective bias and expensive (Bock et al., 2010). Therefore, there is an increasing demand for innovative and reliable disease monitoring method (Bravo et al., 2003; Mahlein et al., 2012a). The aim of the present study was to evaluate the possibility of identifying disease progression stages of STB on wheat with proximal phenotyping and machine-learning.

MATERIALS AND METHODS

Fungal Inoculation

The *Z. tritici* isolate was isolated from typical STB lesions on leaves of winter wheat collected in 2015 in a field in Lomma, Sweden. The inoculum was obtained from stock conidial suspensions of the isolate stored at −80°C in a sterile 1:1 glycerol–water solution. The fungal isolate was retrieved by adding 10 μl of the spore suspension to Petri dishes containing fresh 4-4-4 agar-malt-yeast medium (YMSA) with antibiotic Kanamycin (50 μg/ml) (Saidi et al., 2012). The isolate was spread on the medium by adding 1 ml of sterile water after 1 day of growth. Petri dishes with the isolate were incubated at 20°C with 12 h photoperiod. Conidial suspensions were prepared by first flooding the surface of the 10-day-old cultures with sterile distilled water and then by scraping the agar surface with a sterilized paint brush to release conidia. The spore concentration was measured using a Neubauer counting chamber. Thereafter, 0.1% TWEEN20 (Sigma) was added to the spore suspension and the final spore concentration was adjusted to 10^7 spores ml^{-1}.

Plant Material

Two independent experiments were conducted under greenhouse conditions. In the first experiment (training and validation set), 10 winter wheat cultivars/breeding lines (Stigg, Oakley, Nelson, Mariboss, Kovas DS, Julius, Hereford, SW05317, SW75638, and Target) were evaluated for resistance to STB. Whereas, in the second experiment (test set) two winter wheat cultivars (Kranich and Nimbus) were evaluated for STB resistance. For both experiments, seeds were germinated for 2 days on a moist filter paper in Petri dishes. Germinated seeds were sown in plastic pots (8 cm × 8 cm × 8 cm) filled with peat substrate Blomjord Exclusive (Emmaljunga Torvmull AB, Sweden). For each genotype, two seeds were sown per pot in three replications in a randomized block design. Plants were grown in a greenhouse at 22°C (day) and 18°C (night) with a 16 h photoperiod.

Inoculation Procedure

Seedlings were inoculated following the full emergence of the third leaf and about 21 days after planting. The conidial suspension was applied to both sides of marked second and third leaf using a flat paintbrush (bristle length 15 mm). The control plants were inoculated with water. Following inoculation, plant

leaves were allowed to dry for 1 h before transferring to the humidity chamber. Plants were kept under the plastic tent at close to 100% humidity for 48 h before being returned to the greenhouse conditions.

Disease Assessment

Disease severity was visually assessed at time-points 15, 17, 18, 20, 21, and 23 for the training and validation set and at 6, 8, 10, 13, 14, 15, and 16 days post-inoculation (dpi) for the test set. Percentage of the inoculated leaf surface (from 0 to 100%) presented the following symptoms: chlorosis [the percentage of chlorotic area (CHL)] and necrosis [the percentage of necrotic area (NEC)]. The symptoms and lesion development over the assessment period were summarized by area under the disease progress curve (AUDPC). Minitab software (Version 17.1.0) was used for statistical calculations. Differences in AUDPC were investigated with ANOVA (PROC GLM) and comparisons of means with Tukey's test.

Sensor Phenotyping

In the training and validation set, sensor phenotyping was done at time points 14, 15, 17, and 18 dpi for the infected plants while for the mock-inoculated plants, sensor phenotyping was done at 14 dpi. In the test set, sensor phenotyping was done for both mock-inoculated and infected plants at 6, 8, 10, 13, 14, 15, and 16 dpi. Earlier time-points were additionally included in the test set to evaluate the possibility to detect disease symptoms earlier in the disease progression with the developed computational models. A handheld active light fluorometer (FluorPen FP 100-MAX, Photon Systems Instruments, Czechia) with detachable leaf-clips was used for measuring QY (Photosystem II quantum yield). For QY measurements, for each plant, two leaf-clips were attached to the control or infected leaves and were dark adapted for 15 min prior to the measurements. Thereafter, the leaves were removed from the plants and spectral reflectivity (350–1150 nm) of the leaves were recorded with a resolution of 1 nm with a handheld spectroradiometer sensor Apogee PS-100 (Apogee Instruments, Inc., United States) using a reflectance probe with an internal light source (AS-003, Apogee Instruments, Inc., United States). The spectroradiometer was calibrated against a white reference standard Apogee AS-004 (Apogee Instruments, Inc., United States) prior to the measurements. The leaf temperature was measured with a infrared thermometer Apogee MI-210 (Apogee Instruments, Inc., United States). Finally, the leaves were scanned with Epson Perfection V200 scanner.

Spectral Data Analysis and Machine Learning

The obtained raw spectral reflectance data files were analyzed further to remove noise, detect outliers and calculate VIs using the open-source software Specalyzer[1] and the hsdar R package (Lehnert et al., 2017). Data quality of the spectral files was inspected manually in Specalyzer. Each replicate consisting of two plants was considered as a sample. For spectral measurements from all samples, areas around the edges of the spectra were trimmed due to low signal-to-noise ratio and the region from 420 to 1000 nm was retained for further analysis. Finally, 119 previously known VIs were estimated from the spectral data in Specalyzer for further analysis (Supplementary File 1). Thus, in total, 121 variables were obtained for each sample consisting of 119 VIs, QY, and leaf surface temperature. PCA (principal component analysis) was performed in the software Simca 14.1 (Umetrics, Sweden) and the data was scaled by unit variance (UV) scaling method for PCA.

Random forest (RF) regression models were built from the training set with 121 samples from eight cultivars/breeding lines and was validated on the validation set of 30 samples from two lines (SW75638 and Target) consisting 6 uninfected and 24 infected samples. Recursive feature elimination algorithm (RFE) from the R package Caret (Kuhn et al., 2016) was used for feature selection with parameters: function "rfFuncs," method "repeatedcv" and repeats 10. Separate prediction models were thereafter built with the features selected with RFE for chlorosis and necrosis using the R package Caret. Common parameters for building the models were the variables selected by RFE, eight cultivars from the training set, ntrees 2000, resampling method "repeatedcv," number 10, repeats 10 and importance "True." Percentage of chlorosis and necrosis were used as scoring parameters for model training and testing. The models were tested on a test set consisting of 94 samples with equal number of infected and uninfected samples from the winter wheat cultivars Kranich and Nimbus.

RESULTS

Genotype Variation for STB Severity

In the training set, 10 genotypes showed good variation in STB severity across all time points upon infection (**Figure 1**). Cultivar Stigg had less chlorotic and necrotic symptoms compared to the cultivar Hereford and the breeding line SW75638. The chlorotic symptoms were visible by 14 dpi but were more pronounced by 16 dpi in most genotypes. Necrotic symptoms appeared by 16 dpi in the susceptible genotype but Stigg did not show many necrotic symptoms even at 18 dpi. AUDPC was calculated based on scoring for chlorosis and necrosis at all timepoints to quantify resistance (**Figure 2**). A significant difference ($p < 0.01$) among the 10 winter wheat genotypes was observed in AUDPC of NEC. Target, Stigg, and Nelson revealed highest level of resistance and the most susceptible cultivars in this experiment were Hereford and SW75638 (**Figure 2**). A significant correlation ($r = 0.86$, $p < 0.001$) was also found between the AUDPC of CHL and NEC. In the test set, clear differences in the chlorotic and necrotic symptoms were observed among the two winter wheat cultivars Nimbus and Kranich (**Figure 3**). Disease symptoms appeared much earlier in Nimbus (at 13 dpi) compared to Kranich. Quantification of CHL and NEC by AUDPC in the test set showed significant differences between the two cultivars ($p < 0.01$, **Figure 4**). Cultivar Kranich exhibited a higher level of resistance compared to Nimbus in both STB disease development stages (CHL and NEC).

[1] www.specalyzer.org

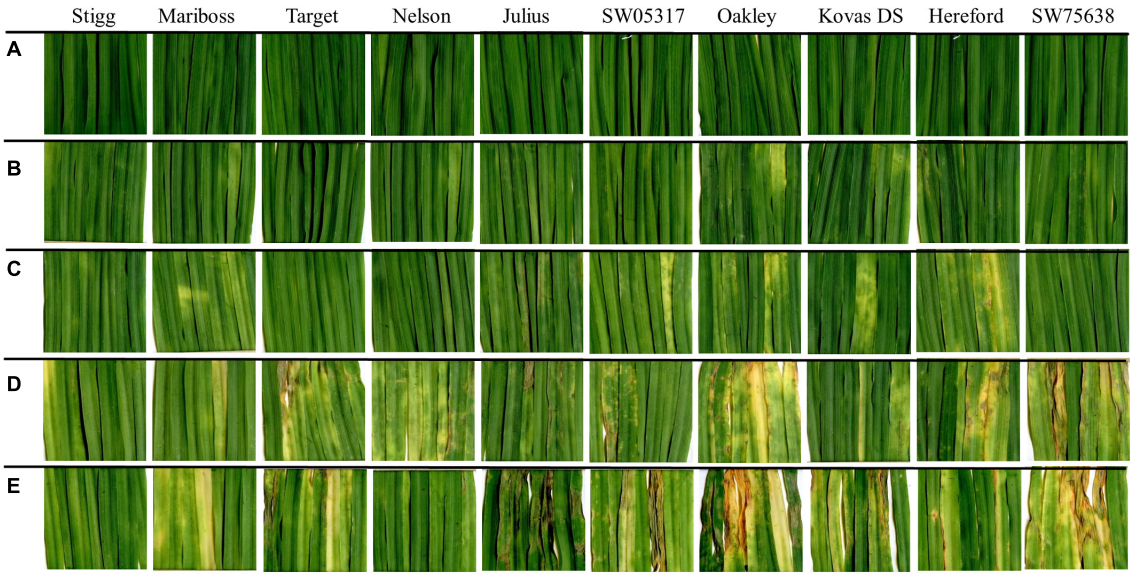


FIGURE 1 | Septoria tritici blotch (STB) symptoms on 10 winter wheat cultivars at **(A)** 14 dpi control; **(B)** 14 dpi infected; **(C)** 15 dpi infected; **(D)** 16 dpi infected; and **(E)** 18 dpi infected. Cultivars are sorted based on the necrotic symptoms.

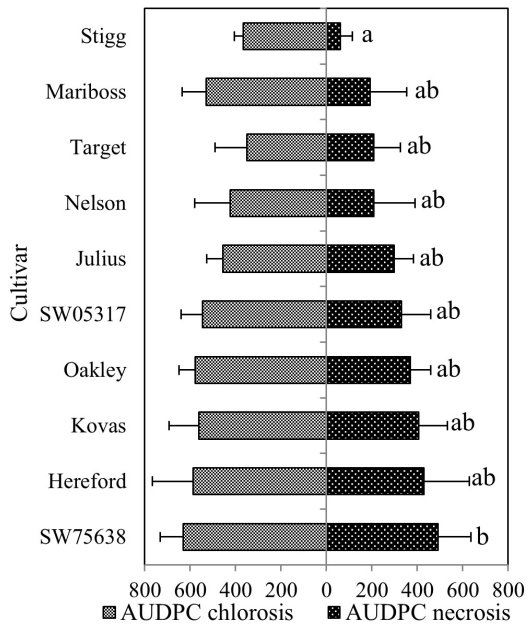


FIGURE 2 | The area under the disease progress curve (AUDPC) for STB among the 10 different cultivars and breeding lines. Means that do not share a letter are significantly different at $p < 0.05$ (Tukey multiple comparison test). Ordered based on the necrotic symptoms.

Multivariate Analysis

Clustering of samples from the training and the test set were studied by PCA of 121 variables consisting of 119 VIs, leaf surface temperature and QY measurements from each sample. In the PCA of the training set, the first and the second components explained 60.2 and 16.3% of the variation respectively (**Figure 5**). Most of the samples from the early time points clustered together (14 and 15 dpi), whereas, samples from the later time points (17 and 18 dpi) were more scattered. The first component explained the variability in the disease progression over time while the second component explained the inter-cultivar variation during disease progression. Thus, disease progression over time was the major variability in the data explained by the PCA. In **Figure 5**, it can be observed that the variability in the data increases with the disease progression. The control samples at 14 dpi have the lowest variability and are thus relatively tightly clustered followed by increasing separation among the infected samples from 14 to 18 dpi. The results from PCA indicates that the 10 cultivars in the training set have physiological differences in their response to the infection which is also corroborated by disease symptoms evaluated with AUDPC analysis (**Figure 2**). In the PCA plot from the test set, the two PCA components explained 56.6 and 18.3% variation respectively (**Figure 6**). Similar to the PCA from the training set, in the test set, the first component explained the variability in the disease progression over time, additionally, some separation was also observed among the control samples as the control samples from the later time-points separated from earlier time-points. This suggests physiological differences in the control plants occurring over a duration of 10 days (6–16 dpi). Among the infected plants, sample separation in Nimbus (susceptible) is detected at 13 dpi whereas in Kranich (resistant) the separation was at 15 dpi. Also, a clear and distinct separation of Nimbus samples at 16 dpi is observed. This indicates distinct differences in the physiological status of Kranich and Nimbus genotypes upon STB infection and these differences become apparent after 13 dpi with the sensor measurements.

A heatmap was prepared for the test set from the relative intensities of the sensor phenotyping data obtained from the

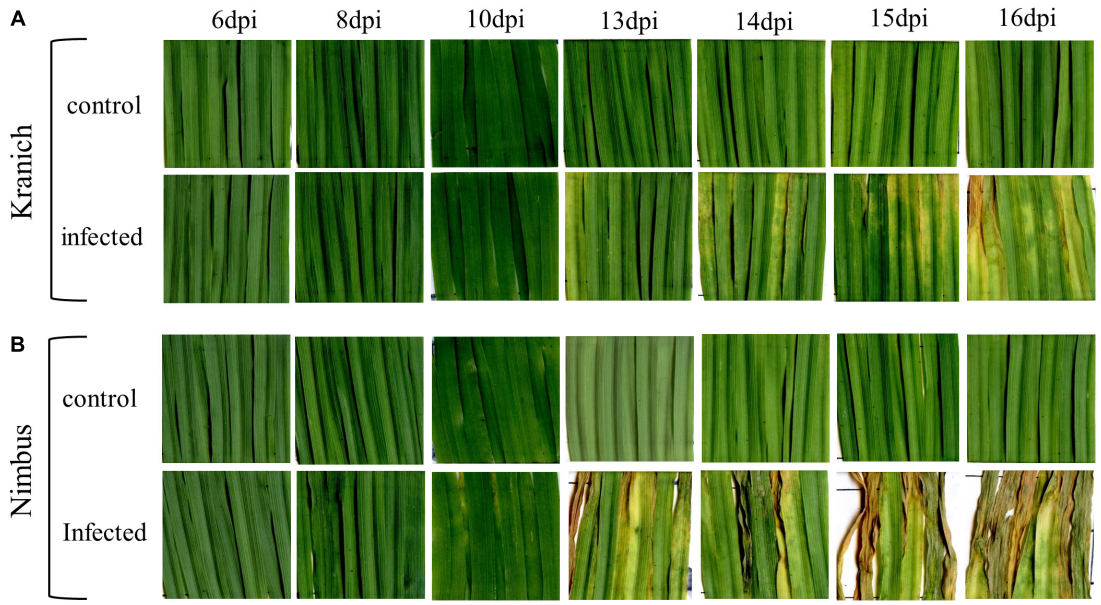


FIGURE 3 | Symptoms of STB on two winter wheat cultivars. **(A)** Kranich, **(B)** Nimbus, with representative pictures of the disease symptoms after 6, 8, 10, 13, 14, 15, and 16 dpi.

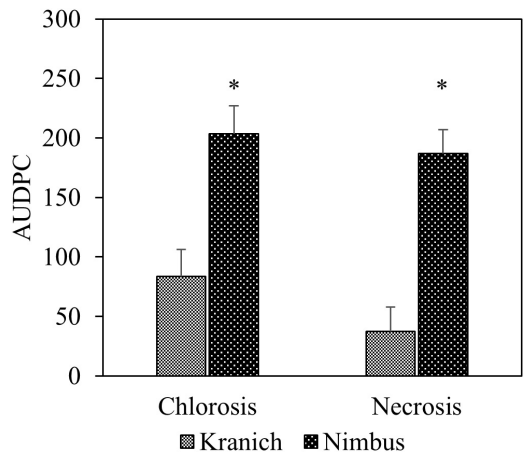


FIGURE 4 | The area under the disease progress curve (AUDPC) for STB for the two cultivars Kranich and Nimbus (two sample *t*-test, $*p < 0.01$).

ratio of the data from the infected plants to that of the mock-inoculated plants (**Figure 7**). Based on the dendrogram, two clusters were obtained at the highest level of tree branching, cluster-I consisted of 30 variables and cluster II 91 variables. As can be seen from the heatmap, lower relative intensities were recorded for several VIs in cluster-II at the later time-points in both cultivars and distinctly lower intensities were observed in the susceptible cultivar Nimbus. Variables in cluster-II negatively correlate with STB symptoms as the intensities of these variables decrease with increase in necrosis. Several variables in cluster-II were affected by necrosis already upon the first visible symptoms of necrosis at 13 dpi in the susceptible cultivar Nimbus. In the resistant cultivar Kranich, relatively less pronounced changes in intensities of variables from cluster-II were seen. Furthermore, unlike in Nimbus, variables in cluster-II were not affected at earlier time-points in Kranich which is in accordance with the delayed necrosis symptoms observed in Kranich.

Correlation analysis of the sensor data was done between the training and the test set to analyze the reproducibility of the measurements over time and genotypes. At first, for each experiment, correlation analysis was done separately for each of the 121 variables and the respective CHL measurement of the sample, thereafter, correlations obtained for each variable from the two experiments were compared. A high coefficient of determination ($R^2 = 0.91$) was obtained indicating good reproducibility of the measurements under similar conditions across time and genotypes (**Figure 8**).

STB Detection With Random Forest

To identify and evaluate key predictive variables for STB infection, automated feature selection was done followed by building RF models with the selected features. For chlorosis, the feature elimination algorithm RFE identified four variables QY, D2 (derivative index), LRDSI1 (leaf rust disease severity index) and LRDSI2 as most predictive. While for necrosis, five variables namely QY, MCARI2/OSAVI2, ARI (anthocyanin reflectance index), SR8 (simple ratio 8) and D1 were identified by the RFE algorithm as important. RF regression models were developed separately for CHL and NEC from the training set and the selected variables. The percentage of variation explained by the models was 45.41% for chlorosis and 21.04% for necrosis with a mean of squared residuals of 581 and 284 respectively. The variables QY was identified as predictive in both chlorosis and necrosis models and

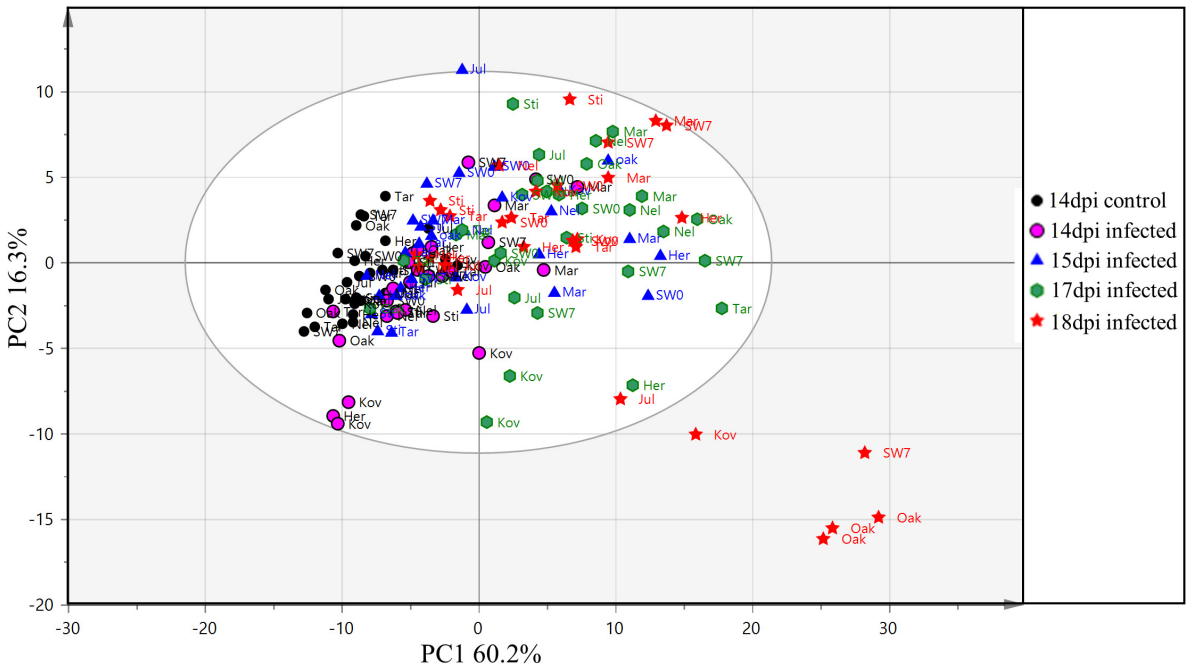


FIGURE 5 | Principal component analysis (PCA) from the vegetation indices, QY and leaf temperature of 10 winter wheat cultivars and breeding lines (training and validation set) after 14, 15, 17, and 18 days post-infection. Elliptical region represents the normal operating area by Hotelling's T2 tolerance with 95% confidence. Sti: Stigg; Oak: Oakley; Nel: Nelson; Mar: Mariboss; Kov: Kovas DS; Jul: Julius; Her: Hereford; SW0: SW05317; SW7: SW75638; and Tar: Target.

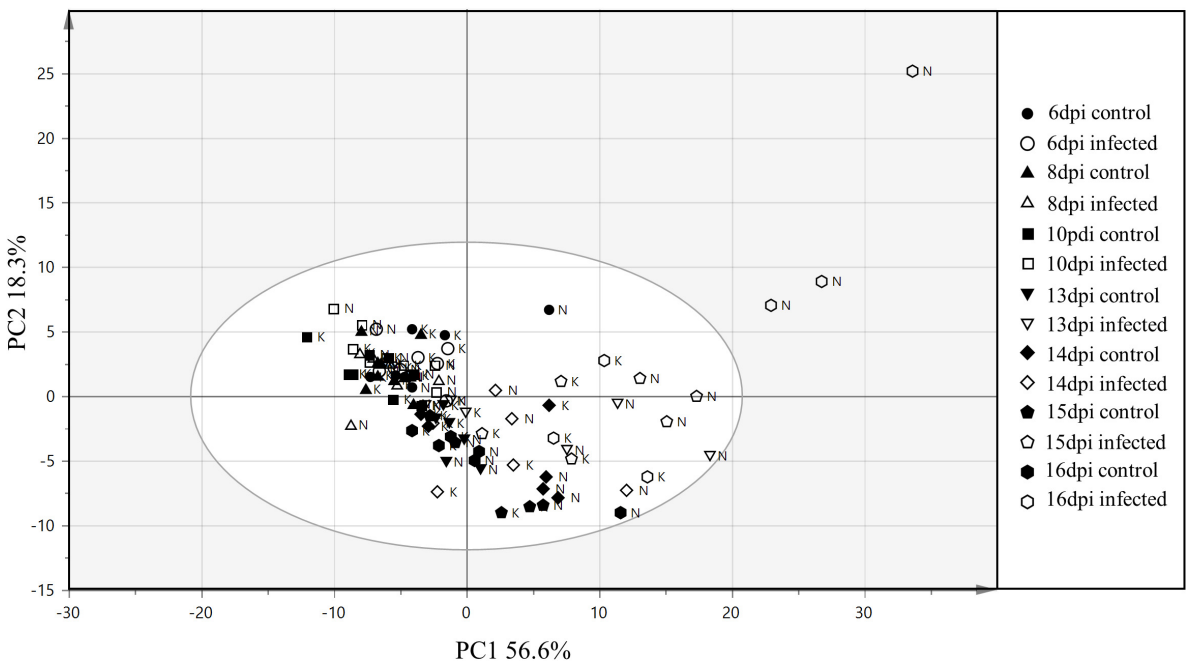


FIGURE 6 | Principal component analysis from the vegetation indices, QY and leaf temperature of two winter wheat cultivars (test set) N: Nimbus and K: Kranich across different time points. Elliptical region represents the normal operating area by Hotelling's T2 tolerance with 95% confidence.

had significantly different ($p < 0.05$) levels in the infected plants compared to the mock-inoculated plants (**Figure 9**). Leaf surface temperature was not selected as a predictive variable although there were significant differences in the surface temperature of the infected and mock-inoculated plants (**Figure 9**).

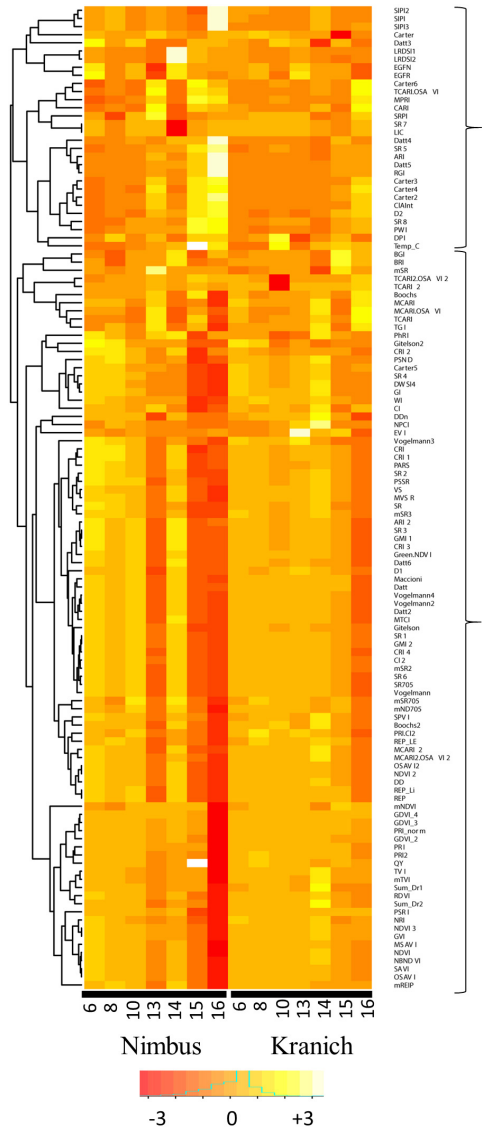


FIGURE 7 | The relative intensities of various vegetation indices of two winter wheat cultivars after infection with STB across different time points.

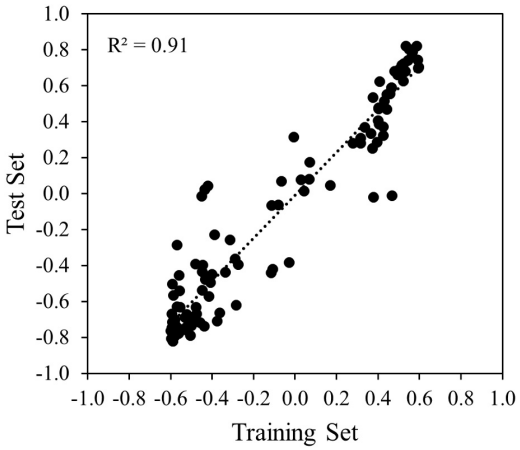


FIGURE 8 | Scatter plot and the simple linear regression line between training and test set.

The two RF models were tested on a validation set consisting of 30 samples from the genotypes SW75638 and Target and an independent test set of 94 samples from two cultivars Kranich and Nimbus. The samples were predicted separately with the RF models built for CHL and NEC. A simple linear regression was calculated to estimate the relationship between the observed and the predicted STB infection. For the validation set, chlorosis was predicted with R^2: 0.80 and necrosis with R^2: 0.92, while for the test set, chlorosis was predicted with R^2: 0.55 (**Figure 10**).

DISCUSSION

In the present work, sensor phenotyping with spectral reflectance, chlorophyll fluorescence, and leaf temperature was performed to evaluate the possibility of detecting different STB developmental stages. In this study, the number of indices affected by the disease increased upon disease progression (**Figure 7**), an observation which was also previously discussed (Ashourloo et al., 2014b). This is due to the magnitude of changes in the leaf morphology and physiology brought upon by disease progression. Sensor phenotyping clearly separated control and infected plants based on the progression of the disease (**Figures 5**, **6**). This separation is influenced by the underlying genetic resistance of the genotypes to STB.

In this work, different indices were identified as top predictive indices for chlorosis and necrosis except for QY which was common in both. This can be due to the distinctly different leaf composition in these two stages. QY (PSII) was affected by STB in the susceptible genotype Nimbus at an early stage of disease progression but was not affected to the same extent in Kranich (**Figure 9**). Previously, chlorophyll fluorescence kinetics were studied for powdery mildew and leaf rust infection in wheat and early detection of infection was possible with chlorophyll fluorescence measurements but not with NDVI (Kuckenberg et al., 2008). In this work, leaf infrared temperature was not selected as a top predictive variable by RF. Leaf temperature was significantly different in the control and infected plants of the two cultivars from the test set at 15 dpi. In the susceptible cultivar Nimbus, statistically significant difference of 1.5°C ($p < 0.05$) was observed between the control and the infected plants at time point 15 dpi with higher temperature recorded from the infected plants (**Figure 9**). These results confirm the results from a previous work where the canopy temperature measured with an infrared thermometer positively correlated (r = 0.48–0.74) with STB coverage in the field (Eyal and Blum, 1989).

The top predictive VIs for chlorosis were D2, LRDSI1, and LRDSI2. The derivative index D1 and D2 correlated well with the natural steady state chlorophyll fluorescence emission by photosystem I and II in the range 639–730 nm (Zarco-Tejada et al., 2003). Both LRDSI1 and LRDSI2 were developed for detecting wheat leaf rust with prediction accuracies of >85%

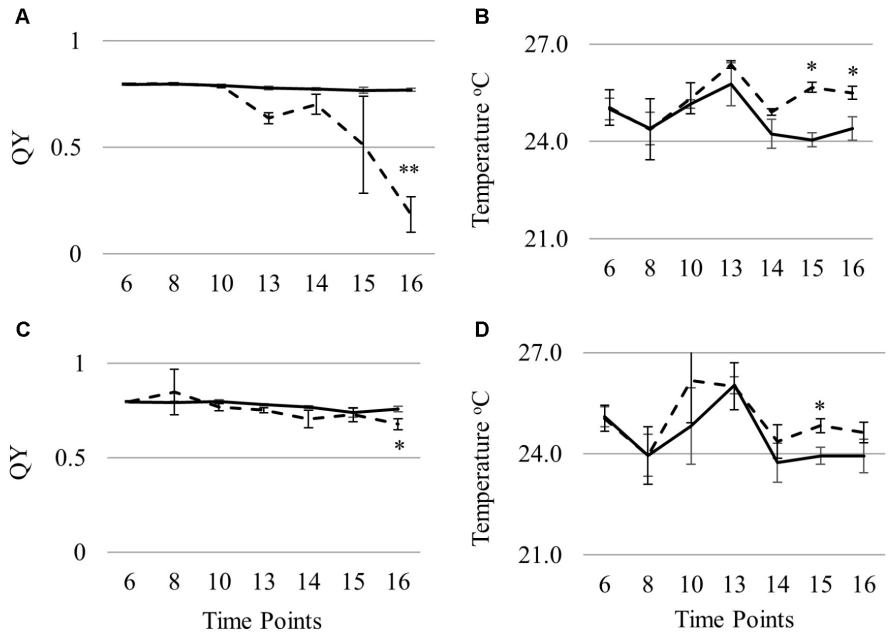


FIGURE 9 | Differences in QY **(A,C)** and leaf surface temperature **(B,D)** in Nimbus **(A,B)** and Kranich **(C,D)** across different time points (two sample *t*-test, $*p < 0.05$).

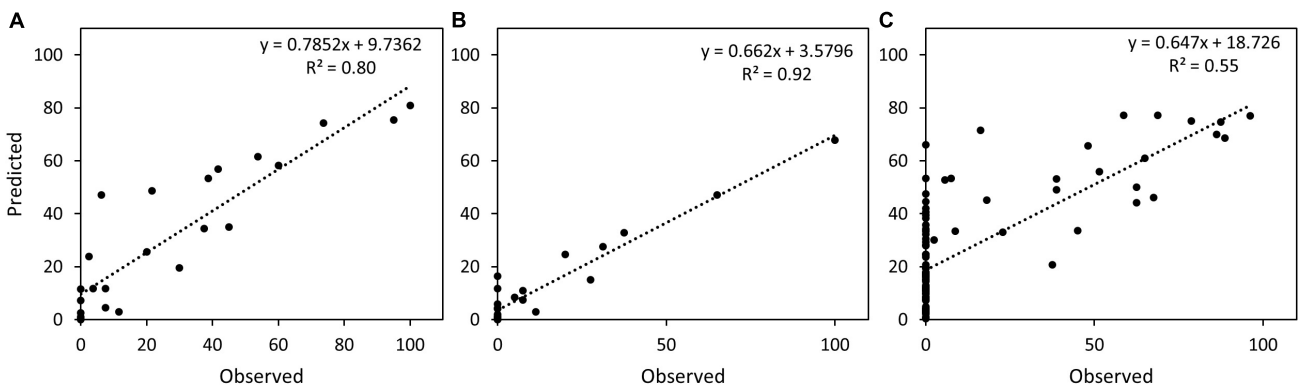


FIGURE 10 | Scatter plot and the simple linear regression line between predicted and observed STB infection scores. **(A)** Chlorosis in the validation set, **(B)** necrosis in the validation set, **(C)** chlorosis in the test set.

(Ashourloo et al., 2014a). Thus, the predictive VIs for chlorosis detect the levels of chlorophyll fluorescence and anthocyanin levels in the leaves.

Top predictive VIs for necrosis were MCARI2.OSAVI2, ARI, SR8, and D1. In the previous work, MCARI2.OSAVI2 was developed for measuring chlorophyll content while tolerating leaf area index (Wu et al., 2008). ARI was developed for estimating anthocyanin reflectance in senescing and stressed leaves (Gitelson et al., 2007). SR8 was developed to estimate carotenoid content in conifer forest (Hernández-Clemente et al., 2012). Thus the top predictive indices for necrosis identified in this work suggests differing levels of chlorophyll, anthocyanin, and carotenoid content in the infected leaves.

Spectral reflectance was previously used to develop spectral indices for detection of different plant diseases. In wheat, leaf rust was detected for plants in a controlled environment with the VIs NBNDVI, NDVI, PRI, GI, and RVSI with an accuracy of over 60% (Ashourloo et al., 2014b). Ashourloo et al. (2014a) developed two new VIs LRDSI1 and LRDSI2 to detect leaf rust in a controlled environment with the R^2: 0.9. By proximal and airborne hyperspectral phenotyping, Huang et al. (2007) identified photochemical reflectance index (PRI) as the most predictive VI (R^2: 0.97) for yellow rust detection in wheat. Cao et al. (2015) studied 17 VIs for prediction of powdery mildew in wheat under field conditions and reported that difference vegetation index (DVI), triangular vegetation index (TVI) and the area of red edge peak significantly correlated with powdery mildew severity. The RF models developed in this work for detection of STB utilize several variables and thus provide a possibility to uniquely detect STB in wheat with various sensors.

STB was detected by NDVI and land surface temperature using satellite imaging with MODIS (moderate-resolution imaging spectroradiometer) and the spatial modeling conducted with linear regression trees and boosted regression trees was suggested as a promising approach for studying STB spread using satellite imaging (Wakie et al., 2016). In this work, NDVI was not detected as a top predictive index, however, the predictive VIs identified in this work can be further validated in the field conditions by proximal and remote sensing.

Further work with parallel measurements of different diseases with various sensors under both greenhouse and field conditions is required to understand the overlap of various predictive VIs.

Septoria tritici blotch resistance traits can be further combined with STB escape and tolerance traits to further reduce the disease progression and maintain yields under disease pressure. Some of the traits identified for STB escape are early vigor, growth rate, plant height, leaf length, leaf spacing, prostrate leaves, leaf insertion angle, flag leaf emergence, and heading time (Arraiano et al., 2009; Brown et al., 2015). Disease tolerance can be explained by the maintenance of yield in spite of the disease pressure. The tolerance trait was studied in a susceptible wheat cultivar 'Miriam' which maintained yield even under disease pressure by increasing photosynthesis in the residual green area of the infected leaves (Kuckenberg et al., 2008). Using affordable high-throughput phenotyping (Armonienė et al., 2018), and genotyping, the escape and tolerance traits can be combined with the resistance traits for developing improved wheat varieties. Furthermore, the sensor data can be integrated with transcriptomics, proteomics, and metabolomics data to improve our understanding of the underlying biological mechanisms. RF was selected in this work as it is equally effective for classification and regression (Díaz-Uriarte and Alvarez de Andrés, 2006) and has been successfully used earlier for disease detection in plants (Chawade et al., 2016). RF is efficient at identifying predictive information from data integrated from various sources and thus it can be explored further for systems-level understanding of responses of plants to various stresses.

CONCLUSION

This work has resulted in identifying top predictive indices for detecting STB. Different predictive variables were selected by the RF model for prediction of chlorosis and necrosis in leaves. From this work, it can be concluded that precision phenotyping with proximal sensors holds the potential for detecting STB and further validation of the identified indices in the field conditions will enable implementation of these precision phenotyping techniques in the field for detection of STB.

AUTHOR CONTRIBUTIONS

AC planned and designed the project. TH recommended the genotypes. FO and RA performed the experiments. AC and FO analyzed the data. All authors interpreted the results and contributed to the writing.

FUNDING

This work was funded through funding from Lantmännen Research Foundation (2016F010) and Partnerskap Alnarp.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.00685/full#supplementary-material

REFERENCES

Agapiou, A., Hadjimitsis, D., and Alexakis, D. (2012). Evaluation of broadband and narrowband vegetation indices for the identification of archaeological crop marks. *Remote Sens.* 4, 3892–3919. doi: 10.3390/rs4123892

Araus, J. L., Casadesus, J., and Bort, J. (2001). "Recent tools for the screening of physiological traits determining yield," in *Application of Physiology in Wheat Breeding*, eds M. P. Reynolds, J. I. Ortiz-Monasterio, and A. Mc Nab (Mexico City: CIMMYT), 55–77.

Armonienė, R., Odilbekov, F., Vivekanand, V., and Chawade, A. (2018). Affordable imaging lab for noninvasive analysis of biomass and early vigour in cereal crops. *BioMed Res. Int.* 2018:5713158. doi: 10.1155/2018/5713158

Arraiano, L. S., Balaam, N., Fenwick, P. M., Chapman, C., Feuerhelm, D., Howell, P., et al. (2009). Contributions of disease resistance and escape to the control of *Septoria tritici* blotch of wheat. *Plant Pathol.* 58, 910–922. doi: 10.1111/j.1365-3059.2009.02118.x

Ashourloo, D., Mobasheri, M., and Huete, A. (2014a). Developing two spectral disease indices for detection of wheat leaf rust (*Pucciniatriticina*). *Remote Sens.* 6, 4723–4740. doi: 10.3390/rs6064723

Ashourloo, D., Mobasheri, M., and Huete, A. (2014b). Evaluating the effect of different wheat rust disease symptoms on vegetation indices using hyperspectral measurements. *Remote Sens.* 6, 5107–5123. doi: 10.3390/rs6065107

Baranowski, P., Jedryczka, M., Mazurek, W., Babula-Skowronska, D., Siedliska, A., and Kaczmarek, J. (2015). Hyperspectral and thermal imaging of oilseed rape (*Brassica napus*) response to fungal species of the genus *Alternaria*. *PLoS One* 10:e0122913. doi: 10.1371/journal.pone.0122913

Berdugo, C., Zito, R., Paulus, S., and Mahlein, A. K. (2014). Fusion of sensor data for the detection and differentiation of plant diseases in cucumber. *Plant Pathol.* 63, 1344–1356. doi: 10.1111/ppa.12219

Bock, C., Poole, G., Parker, P., and Gottwald, T. (2010). Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. *Crit. Rev. Plant Sci.* 29, 59–107. doi: 10.1080/07352681003617285

Bravo, C., Moshou, D., West, J., McCartney, A., and Ramon, H. (2003). Early disease detection in wheat fields using spectral reflectance. *Biosyst. Eng.* 84, 137–145. doi: 10.1016/s1537-5110(02)00269-6

Brown, J. K., Chartrain, L., Lasserre-Zuber, P., and Saintenac, C. (2015). Genetics of resistance to *Zymoseptoria tritici* and applications to wheat breeding. *Fungal Genet. Biol.* 79, 33–41. doi: 10.1016/j.fgb.2015.04.017

Cao, X., Luo, Y., Zhou, Y., Fan, J., Xu, X., West, J. S., et al. (2015). Detection of powdery mildew in two winter wheat plant densities and prediction of grain yield using canopy hyperspectral reflectance. *PLoS One* 10:e0121462. doi: 10.1371/journal.pone.0121462

Chartrain, L., Brading, P., Widdowson, J., and Brown, J. (2004). Partial resistance to *Septoria tritici* blotch (*Mycosphaerella graminicola*) in wheat cultivars Arina and Riband. *Phytopathology* 94, 497–504. doi: 10.1094/PHYTO.2004.94.5.497

Chawade, A., Alexandersson, E., Bengtsson, T., Andreasson, E., and Levander, F. (2016). Targeted proteomics approach for precision plant breeding. *J. Proteome Res.* 15, 638–646. doi: 10.1021/acs.jproteome.5b01061

Chawade, A., Armonienė, R., Berg, G., Brazauskas, G., Frostgård, G., Geleta, M., et al. (2018). A transnational and holistic breeding approach is needed for sustainable wheat production in the Baltic Sea region. *Physiol. Plant.* doi: 10.1111/ppl.12726 [Epub ahead of print].

Devadas, R., Lamb, D. W., Simpfendorfer, S., and Backhouse, D. (2008). Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. *Precis. Agric.* 10, 459–470. doi: 10.1007/s11119-008-9100-2

Díaz-Uriarte, R., and Alvarez de Andrés, S. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics* 7:3. doi: 10.1186/1471-2105-7-3

Eyal, Z., and Blum, A. (1989). Canopy temperature as a correlative measure for assessing host response to *Septoria tritici* blotch of wheat. *Plant Dis.* 73, 468–471. doi: 10.1094/pd-73-0468

Fones, H., and Gurr, S. (2015). The impact of *Septoria tritici* blotch disease on wheat: an EU perspective. *Fungal Genet. Biol.* 79, 3–7. doi: 10.1016/j.fgb.2015.04.004

Gitelson, A. A., Kaufman, Y. J., Stark, R., and Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* 80, 76–87. doi: 10.1016/S0034-4257(01)00289-9

Gitelson, A. A., Merzlyak, M. N., and Chivkunova, O. B. (2007). Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochem. Photobiol.* 74, 38–45. doi: 10.1562/0031-865520010740038opaneo2.0.co2

Goodwin, S. B., M'barek, S. B., Dhillon, B., Wittenberg, A. H., Crane, C. F., Hane, J. K., et al. (2011). Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genet.* 7:e1002070. doi: 10.1371/journal.pgen.1002070

Hernández-Clemente, R., Navarro-Cerrillo, R. M., and Zarco-Tejada, P. J. (2012). Carotenoid content estimation in a heterogeneous conifer forest using narrow-band indices and PROSPECT+DART simulations. *Remote Sens. Environ.* 127, 298–315. doi: 10.1016/j.rse.2012.09.014

Huang, W., Lamb, D. W., Niu, Z., Zhang, Y., Liu, L., and Wang, J. (2007). Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. *Precis. Agric.* 8, 187–197. doi: 10.1007/s11119-007-9038-9

Iori, A., Scala, V., Cesare, D., Pinzari, F., D'Egidio, M., Fanelli, C., et al. (2015). Hyperspectral and molecular analysis of *Stagonospora nodorum* blotch disease in durum wheat. *Eur. J. Plant Pathol.* 141, 689–702. doi: 10.1007/s10658-014-0571-x

Kema, G., Annone, J., Sayoud, R., Van Silfhout, C., Van Ginkel, M., and De Bree, J. (1996). Genetic variation for virulence and resistance in the wheat-*Mycosphaerella graminicola* pathosystem. I: interactions between pathogen isolates and host cultivars. *Phytopathology* 86, 200–212. doi: 10.1094/Phyto-86-200

Kema, G. H., Verstappen, E. C., and Waalwijk, C. (2000). Avirulence in the wheat *Septoria tritici* leaf blotch fungus *Mycosphaerella graminicola* is controlled by a single locus. *Mol. Plant Microbe Interact.* 13, 1375–1379. doi: 10.1094/MPMI.2000.13.12.1375

Keon, J., Antoniw, J., Carzaniga, R., Deller, S., Ward, J. L., Baker, J. M., et al. (2007). Transcriptional adaptation of *Mycosphaerella graminicola* to programmed cell death (PCD) of its susceptible wheat host. *Mol. Plant Microbe Interact.* 20, 178–193. doi: 10.1094/MPMI-20-2-0178

Kettles, G. J., and Kanyuka, K. (2016). Dissecting the molecular interactions between wheat and the fungal pathogen *Zymoseptoria tritici*. *Front. Plant Sci.* 7:508. doi: 10.3389/fpls.2016.00508

Kuckenberg, J., Tartachnyk, I., and Noga, G. (2008). Temporal and spatial changes of chlorophyll fluorescence as a basis for early and precise detection of leaf rust and powdery mildew infections in wheat leaves. *Precis. Agric.* 10, 34–44. doi: 10.1007/s11119-008-9082-0

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., et al. (2016). *Classification and Regression Training. R Package Version 6.0–73*. Available at: https://cran.r-project.org/package=caret

Kuska, M., Wahabzada, M., Leucker, M., Dehne, H.-W., Kersting, K., Oerke, E.-C., et al. (2015). Hyperspectral phenotyping on the microscopic scale: towards automated characterization of plant-pathogen interactions. *Plant Methods* 11:28. doi: 10.1186/s13007-015-0073-7

Lehnert, L. W., Meyer, H., and Bendix, J. (2017). *Hsdar: Manage, Analyse and Simulate Hyperspectral Data in R. R Package Version 0.7.0*.

Lowe, A., Harrison, N., and French, A. P. (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods* 13:80. doi: 10.1186/s13007-017-0233-z

Mahlein, A.-K. (2016). Plant disease detection by imaging sensors–parallels and specific demands for precision agriculture and plant phenotyping. *Plant Dis.* 100, 241–251. doi: 10.1094/PDIS-03-15-0340-FE

Mahlein, A.-K., Oerke, E.-C., Steiner, U., and Dehne, H.-W. (2012a). Recent advances in sensing plant diseases for precision crop protection. *Eur. J. Plant Pathol.* 133, 197–209. doi: 10.1007/s10658-011-9878-z

Mahlein, A.-K., Steiner, U., Hillnhütter, C., Dehne, H.-W., and Oerke, E.-C. (2012b). Hyperspectral imaging for small-scale analysis of symptoms caused by different sugar beet diseases. *Plant Methods* 8:3. doi: 10.1186/1746-4811-8-3

Moshou, D., Bravo, C., West, J., Wahlen, S., McCartney, A., and Ramon, H. (2004). Automatic detection of 'yellow rust' in wheat using reflectance measurements and neural networks. *Comput. Electron. Agric.* 44, 173–188. doi: 10.1016/j.compag.2004.04.003

Penuelas, J., Baret, F., and Filella, I. (1995). Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica* 31, 221–230.

Pietragalla, J., Mullan, D. M., and Mendoza, R. S. (2012). "Spectral reflectance," in *Physiological Breeding II: A Field Guide to Wheat Phenotyping*, eds A. J. D. Pask, J. Pietragalla, D. M. Mullan, and M. P. Reynolds (Mexico City: CIMMYT), 32–36.

Raun, W. R., Solie, J. B., Johnson, G. V., Stone, M. L., Mullen, R. W., Freeman, K. W., et al. (2002). Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. *Agron. J.* 94, 815–820. doi: 10.2134/agronj2002.8150

Rouse, J. W. Jr., Haas, R., Schell, J., and Deering, D. (1974). "Monitoring vegetation systems in the Great Plains with ERTS," in *Proceedings of the Third Earth Resources Technology Satellite- 1 Symposium*, Greenbelt, MD.

Rudd, J. J., Kanyuka, K., Hassani-Pak, K., Derbyshire, M., Andongabo, A., Devonshire, J., et al. (2015). Transcriptome and metabolite profiling of the infection cycle of *Zymoseptoria tritici* on wheat reveals a biphasic interaction with plant immunity involving differential pathogen chromosomal contributions and a variation on the hemibiotrophic lifestyle definition. *Plant Physiol.* 167, 1158–1185. doi: 10.1104/pp.114.255927

Saidi, A., Eslahi, M., and Safaie, N. (2012). Efficiency of *Septoria tritici* sporulation on different culture media. *Trakia J. Sci.* 10, 15–18.

Sánchez-Vallet, A., McDonald, M. C., Solomon, P. S., and McDonald, B. A. (2015). Is *Zymoseptoria tritici* a hemibiotroph? *Fungal Genet. Biol.* 79, 29–32. doi: 10.1016/j.fgb.2015.04.001

Shetty, N. P., Kristensen, B., Newman, M.-A., Møller, K., Gregersen, P. L., and Jørgensen, H. L. (2003). Association of hydrogen peroxide with restriction of *Septoria tritici* in resistant wheat. *Physiol. Mol. Plant Pathol.* 62, 333–346. doi: 10.1016/S0885-5765(03)00079-1

Sims, D. A., and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* 81, 337–354. doi: 10.1016/s0034-4257(02)00010-x

Singh, A., Ganapathysubramanian, B., Singh, A. K., and Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. *Trends Plant Sci.* 21, 110–124. doi: 10.1016/j.tplants.2015.10.015

Steinberg, G. (2015). Cell biology of *Zymoseptoria tritici*: pathogen cell organization and wheat infection. *Fungal Genet. Biol.* 79, 17–23. doi: 10.1016/j.fgb.2015.04.002

Thomas, S., Wahabzada, M., Kuska, M. T., Rascher, U., and Mahlein, A.-K. (2017). Observation of plant–pathogen interaction by simultaneous hyperspectral imaging reflection and transmission measurements. *Funct. Plant Biol.* 44, 23–34. doi: 10.1071/FP16127

Wakie, T. T., Kumar, S., Senay, G. B., Takele, A., and Lencho, A. (2016). Spatial prediction of wheat Septoria leaf blotch (*Septoria tritici*) disease severity in Central Ethiopia. *Ecol. Inform.* 36, 15–30. doi: 10.1016/j.ecoinf.2016.09.003

Wu, C., Niu, Z., Tang, Q., and Huang, W. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: modeling and validation. *Agric. For. Meteorol.* 148, 1230–1241. doi: 10.1016/j.agrformet.2008.03.005

Xie, C., Yang, C., and He, Y. (2016). "Detection of grey mold disease on tomato leaves at different infected stages using hyperspectral imaging," in *Proceedings of the 2016 ASABE Annual International Meeting* (St. Joseph, MI: American Society of Agricultural and Biological Engineers), 1.

Yeh, Y.-H., Chung, W.-C., Liao, J.-Y., Chung, C.-L., Kuo, Y.-F., and Lin, T.-T. (2016). Strawberry foliar anthracnose assessment by hyperspectral imaging. *Comput. Electron. Agric.* 122, 1–9. doi: 10.1016/j.compag.2016.01.012

Zarco-Tejada, P. J., Pushnik, J. C., Dobrowski, S., and Ustin, S. L. (2003). Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. *Remote Sens. Environ.* 84, 283–294. doi: 10.1016/s0034-4257(02)00113-x

Zhang, J.-C., Pu, R.-L., Wang, J.-H., Huang, W.-J., Yuan, L., and Luo, J.-H. (2012). Detecting powdery mildew of winter wheat using leaf level hyperspectral measurements. *Comput. Electron. Agric.* 85, 13–23. doi: 10.1016/j.compag.2012.03.006

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Specalyzer—an interactive online tool to analyze spectral reflectance measurements

Alexander Koc[1], Tina Henriksson[2] and Aakash Chawade[1]

[1] Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden
[2] Lantmännen Lantbruk, Svalöv, Sweden

ABSTRACT

Low-cost phenotyping using proximal sensors is increasingly becoming popular in plant breeding. As these techniques generate a large amount of data, analysis pipelines that do not require expertise in computer programming can benefit a broader user base. In this work, a new online tool Specalyzer is presented that allows interactive analysis of the spectral reflectance data generated by proximal spectroradiometers. Specalyzer can be operated from any web browser allowing data uploading, analysis, interactive plots and exporting by point and click using a simple graphical user interface. Specalyzer is evaluated with case study data from a winter wheat fertilizer trial with two fertilizer treatments. Specalyzer can be accessed online at http://www.specalyzer.org.

Subjects Agricultural Science, Bioinformatics, Data Mining and Machine Learning, Data Science
Keywords Specalyzer, Spectroradiometer, Proximal phenotyping, Online tool, Spectral reflectance, Vegetation indices

BACKGROUND

High-throughput plant phenotyping (HTPP) is becoming increasingly popular with the development of new low-cost phenotyping technologies and sensors. HTPP can aid in the detection of plant traits for applications in breeding and farming and for gaining fundamental understanding of molecular mechanisms underlying the trait of interest (*Furbank & Tester, 2011*). HTPP can be performed on individual plants, trial plots or big farms. Sensors are available for estimating spectral reflectance in the leaves of individual plants or small plots to large scale phenotyping of big farms with unmanned aerial vehicles mounted with hyperspectral cameras or with satellite imaging (*Muñoz Huerta et al., 2013*; *Tattaris, Reynolds & Chapman, 2016*).

Phenotyping by proximal spectroradiometers can be performed to estimate various traits in several different crops. Canopy biomass and nitrogen status in wheat was estimated with a proximal spectrometer with a wavelength range of 400–900 nm mounted on a tractor (*Hansen & Schjoerring, 2003*), leaf area index in rice was measured with a handheld spectrometer with a wavelength range of 250–2,500 nm (*Wang et al., 2007*), nitrogen uptake in winter wheat was estimated with a handheld spectrometer with a wavelength range of 350–1,000 nm (*Yao et al., 2013*), grain yield and protein content in winter wheat was also measured with a handheld spectroradiometer with a wavelength range of 447–1,752 nm

Submitted 3 March 2018
Accepted 31 May 2018
Published 25 June 2018

Corresponding author
Aakash Chawade,
aakash.chawade@slu.se

Academic editor
Budiman Minasny

Additional Information and Declarations can be found on page 11

DOI **10.7717/peerj.5031**

(*Xue, Cao & Yang, 2007*) and wheat yield under irrigation was estimated with a portable spectroradiometer with a wavelength range of 350–1,100 nm (*Babar et al., 2006*).

A large number of vegetation indices (VIs) have been developed using the visible and near-infrared spectral wavelengths for estimation of various traits of interest in plants (*Agapiou, Hadjimitsis & Alexakis, 2012*). These VIs can be estimated from the spectral reflectance from the proximal spectroradiometers (*Deery et al., 2014*; *Gizaw, Garland-Campbell & Carter, 2016*; *Hansen & Schjoerring, 2003*), and with the sensors from unmanned aerial vehicles and satellites (*Franke & Menz, 2007*; *Haghighattalab et al., 2016*; *Tattaris, Reynolds & Chapman, 2016*; *Tucker & Sellers, 1986*). As the VIs and their association with a trait of interest is often known, the VIs estimated from the new measurements can aid in detection of the associated traits. VIs have proven to be effective in estimation of leaf area index (*Tucker & Sellers, 1986*; *Wang et al., 2007*), radiation use efficiency (*Peñuelas, Filella & Gamon, 1995*), water status (*Peñuelas et al., 1993*), leaf pigments (*Sims & Gamon, 2002*), grain yield (*Babar et al., 2006*; *Cao et al., 2015*; *Gizaw, Garland-Campbell & Carter, 2016*; *Xue, Cao & Yang, 2007*) and diseases (*Cao et al., 2015*; *Mahlein, 2016*; *Odilbekov et al., 2018*). VIs have also been used in conjunction with multivariate and machine learning techniques to build advanced models for improved detection of complex traits (*Mahlein, 2016*; *Odilbekov et al., 2018*; *Singh et al., 2016*). VIs have been most successfully used for efficient application of fertilizers in the field. Leaching of fertilizer leads to ground and water pollution and wastage of resources, and thus, improving the nitrogen use efficiency of the crops and the efficient application of fertilizers is important for sustainable agriculture (*Chawade et al., 2018*; *Muñoz Huerta et al., 2013*).

Spectral data has to be processed upon acquisition and the steps involve pre-processing to remove outlier samples, normalization, trimming of edges to remove low signal-to-noise ratio wavelengths and estimation of VIs. The R packages hsdar (*Lehnert, Meyer & B, 2017*) and pavo (*Maia et al., 2013*) provide most of the features required for analysis of the collected spectral data in the R command line environment. The package pavo additionally provides features for visualization of the data. While the two packages are efficient in analyzing the data through a command line interface, they lack a graphical user interface and require users to be familiar with the R environment. This can be challenging for users unfamiliar with computer programming. The target users of spectral data analysis are primarily biologists and plant breeders who are interested in studying the trait of interest using spectral reflectance. Thus, a tool with a graphical user interface for spectra data analysis will allow a broader user base to work with such data.

Analysis tools with graphical user interface independent of the operating system enable ease of use and a broader acceptance of spectral reflectance techniques. Visualization tools allow users with little or no skills in computer programming to analyze multidimensional phenotypic data and identify dominant patterns relevant to their research question. Two aspects are of fundamental importance for building a visualization tool, namely, data analysis capabilities and user experience (UX). Data readability and interpretability can be improved with graphical representation with several different types of plots including boxplots, barplots, heatmaps, histograms and scatter plots (*Calatroni & Wildfire, 2017*; *Chawade, Alexandersson & Levander, 2014*). Appropriate selection of colors and sizes of

various data points in the plots connects the graphic to the real world (*Yau, 2013b*). Although automation of data analysis is desired for high-throughput experiments, a balance between automation and visualization based decision making is at times necessary. The field of visual analytics allows the achievement of this goal and is the subject of ongoing research (*O'Donoghue et al., 2010*). UX aspects necessitate developing a user friendly graphical interface (*Pavelin et al., 2012*) which could be simplistic and web-based (*Chawade, Alexandersson & Levander, 2014*) or require a local installation offering advanced customization possibilities (*Kerren et al., 2017*; *Shannon, 2003*).

In this work, a new online tool Specalyzer is proposed which enables analysis and visualization of the collected spectral data in a web browser and thus can be used on any device with a web browser and an internet connection. Various features in Specalyzer are described and evaluated with a case study on fertilizer treatment in winter wheat.

METHODS AND MATERIALS

Implementation

Specalyzer is a web application implemented in the statistical programming language R v3.4.2 (*R Development Core Team, 2016*) and built using the Shiny v1.0.5 web application framework (*Chang et al., 2017*). Specalyzer uses the the hsdar v0.5.1 package (*Lehnert, Meyer & B, 2017*) and the included Speclib function for managing spectral data and calculating VIs. Plotly v4.7.1 (*Sievert et al., 2017*) is used to generate interactive visualizations shown in the web application. Specalyzer also uses dplyr v0.7.4 (*Wickham et al., 2017a*) and reshape2 v1.4.2 (*Wickham, 2007*) for extracting and transforming data, and asdreader v0.1-3 (*Roudier, 2017*) for reading binary ASD FieldSpec® data files. Together, these packages are used for data input, processing, and visualization functionality for spectral data. The package shinyjs v1.0 (*Attali, 2018*) is used for additional user-interface functionality, and readr v1.1.1 (*Wickham, Hester & Francois, 2017b*) for reading data from disk and generating tables for export. Specalyzer code is available at https://github.com/alkc/specalyzer.

Data input

The spectral data formats supported by Specalyzer are (a) ASD FieldSpec® binary spectral data files, (b) SpectraWiz® data files and (c) data merged in a single generic text file. Continuous and/or discrete attribute data in the form of a tab-delimited table can also be uploaded. The attribute data should be organized so that each column is a trait and each of the rows correspond to a spectral sample. The first column should be labelled "filename" in the header row and should include the filenames of the corresponding spectral data files. Additionally, a comma-delimited matrix of filenames can be uploaded that includes the spatial distribution of the samples in the field. The spatial matrix is used for visualization of the VI or attributes of samples in the field.

Estimation of vegetation indices

Specalyzer estimates 140 VIs within the spectral range of 400–1,000 nm and are summarized in File S1. If the reflectance measurements of the wavelengths required for estimating a

VI is unavailable, missing values are reported for the given index. Thus, to estimate all indices provided by Specalyzer, data acquisition with spectroradiometers with a resolution of one nanometer is recommended. Specalyzer does not perform any data interpolation and thus such data transformation can be necessary when the data is collected with spectral instruments with lower resolutions. Such data transformation should be performed prior to using Specalyzer.

Case study data

A winter wheat field trial was conducted in Svalöv, in Southern Sweden in 2015–2016 with two fertilizer treatments (140 and 180 kg ha^{-1}). The spectral reflectance from 10 breeding lines was measured with the handheld Apogee PS-100 spectroradiometer (Apogee Instruments Inc., Logan, UT, USA). The spectroradiometer was calibrated against the white reference at every tenth reading and the measurements were made in the range of 339–1,100 nm. Due to a low signal-to-noise ratio in the areas around the edges of the measured spectral interval, the data in the range of 400–1,000 nm was considered for further analysis. The measurements were made in June 2016 around midday under a clear sky at the post-anthesis growth stage (Zadoks 71–77). The spectroradiometer was held approximately 1 m above the canopies for reflectance measurements.

RESULTS

Features available in Specalyzer

The aim of Specalyzer is to aid in the quality control, pre-processing, estimation of VIs and visualization of the spectral reflectance data (Fig. 1). This is achieved with a web application with an interactive user interface capable of processing and visualizing raw and processed data (Fig. 2). Specalyzer is platform independent and can be used in a web browser on any computer or mobile device. The available pre-processing features are removing outlier samples, trimming the spectral range and calculation of 140 VIs. Information about the replicates can be optionally included in the attribute file and various plots can be created by grouping the samples by the provided attribute(s). VIs relevant to a trait of interest can be identified by performing correlation analysis (quantitative trait) or one-way ANOVA analysis (quantitative trait).

The available plotting features for the spectral data are scatterplots for individual samples, mean and median of all samples, and the variance component. For the VIs, boxplots can be created for ordinal attributes and scatterplots for continuous attributes. Optionally, the data points (samples) can be overlaid over the boxplots. Additionally, a custom VI can also be manually added to the list of VIs. Samples can be grouped by any provided attribute for further analysis and plotting. PCA plots can be created from the spectral data to analyze sample grouping, data structure and outliers. In the PCA plot, automatic sample outlier detection is available and is based on standard deviation of a sample from the mean of the loadings of principal components 1 and 2. Fieldmaps can be created to visualize the spatial distribution of samples in the field and the spatial variation in the intensities of a given VI. All plots in Specalyzer are interactive providing further information on each datapoint in the plot by hovering the mouse pointer over it. Finally, various plots can be saved in

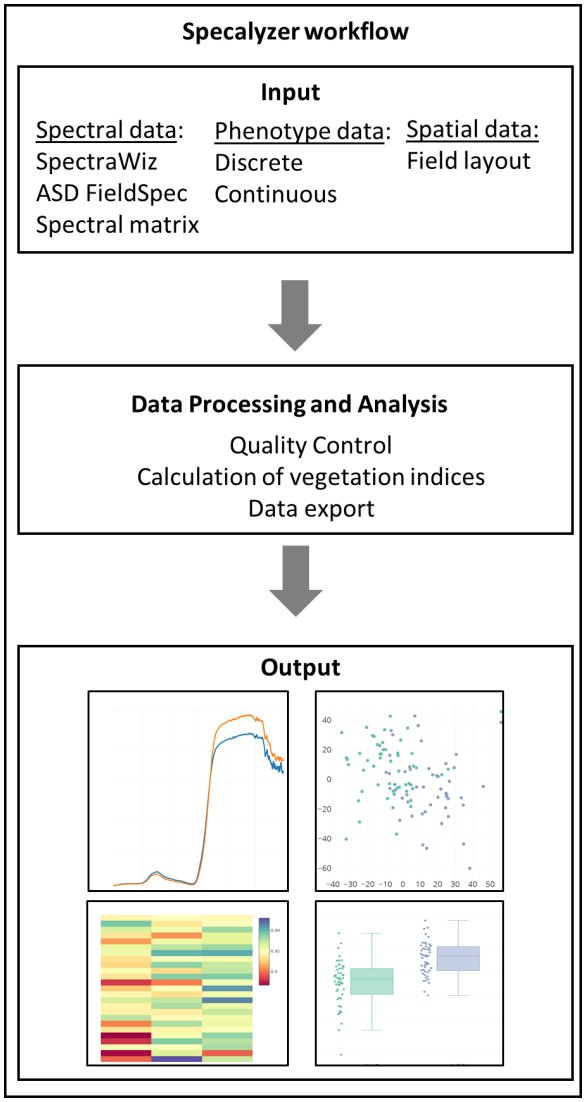


Figure 1 Specalyzer workflow. Specalyzer workflow for data input, different aspects of data processing and analysis, and output visualization.

Full-size DOI: 10.7717/peerj.5031/fig-1

portable network graphics (PNG) format and the spectral and the VI data can be exported for further analysis. While exporting, samples can also be averaged based on any provided attributes.

Evaluation of Specalyzer

The features in Specalyzer were evaluated in a case study from a fertilizer field trial with two treatments of fertilizer levels.

Case study

A field trial was conducted with replications and two fertilizer treatments (140 and 180 kg ha^{-1}). Spectral data was collected as described in the Methods section. Figure 3A illustrates

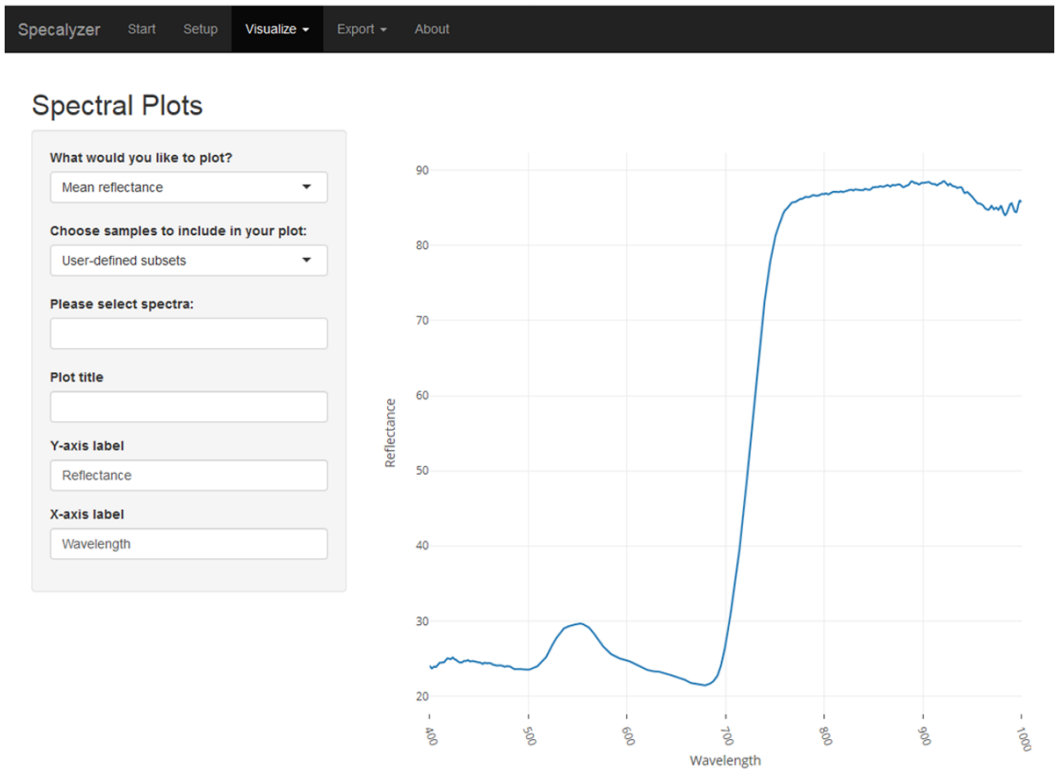


Figure 2 Specalyzer web application screenshot. A screenshot of the Specalyzer web-application displaying the menu for generating spectral plots from spectral reflectance data.

Full-size DOI: 10.7717/peerj.5031/fig-2

spectral plots from Specalyzer where regions between 300–400 nm and 1,000–1,200 nm have low signal-to-noise ratio. These regions can be filtered away by masking unwanted regions and new plots can be generated for further analysis (Fig. 3B). Another important quality control (QC) feature in Specalyzer is outlier detection with PCA plots (Fig. 4). For the case study data, an outlier is detected in the PCA plot (Fig. 4B), and the sample label is identified by hovering the mouse over the outlier sample. A spectral plot with the outlier sample together with another randomly chosen sample shows drastically different spectral reflectance profiles (Fig. 4A). The outlier sample in this case study was a control sample of dry leaves. The new PCA plot after filtering away the outlier sample shows uniform distribution of samples (Fig. 4C).

The plots shown in Figs. 3 and 4 can also be used to compare samples by attributes for investigating the spectral reflectance in response to different attributes. For example, in the case study, the spectral plots in Fig. 3B shows the difference in mean reflectance in samples from two fertilizer treatments. The mean reflectance of samples vary for the two treatments with the samples receiving more fertilizer showing increased reflectance in the near-infrared spectrum. Samples can also be colored by attributes in the PCA plot (Fig. 4C).

Specalyzer also calculates 140 previously known VIs from the spectral data. These indices can be calculated for individual spectral samples, or can be aggregated by attribute for scatterplots and boxplots (Fig. 5). Boxplots for a few selected indices are shown for

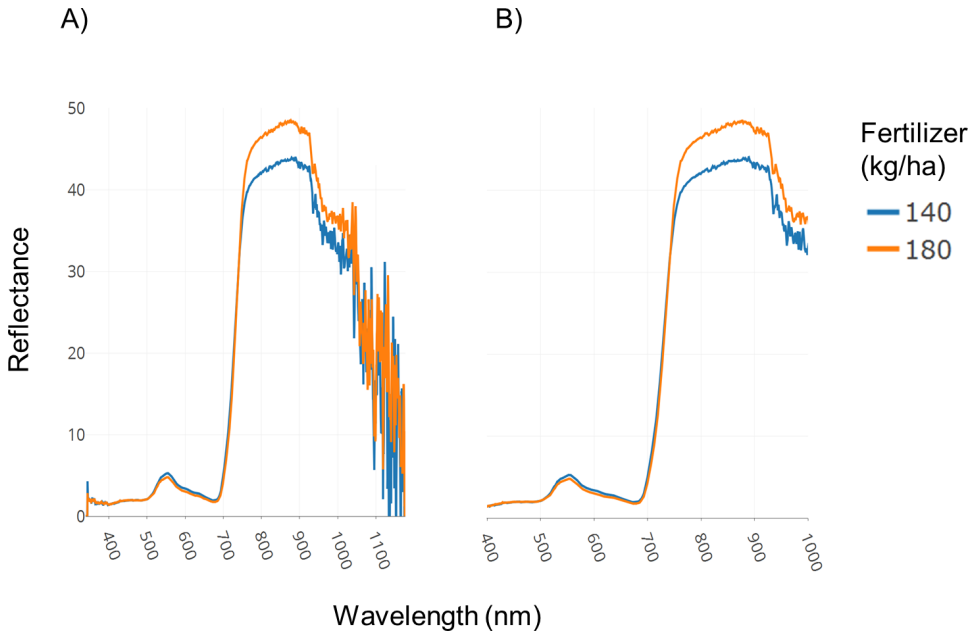


Figure 3 Filtering regions with low signal-to-noise ratio. Demonstration of the effect of removing regions with low signal-to-noise ratio in spectral data within Specalyzer. (A) Aggregated spectral data showing low signal-to-noise ratio in the intervals 300–400 and 1,000–1,200 nm (B) Aggregated spectral data after trimming away noisy regions.

Full-size DOI: 10.7717/peerj.5031/fig-3

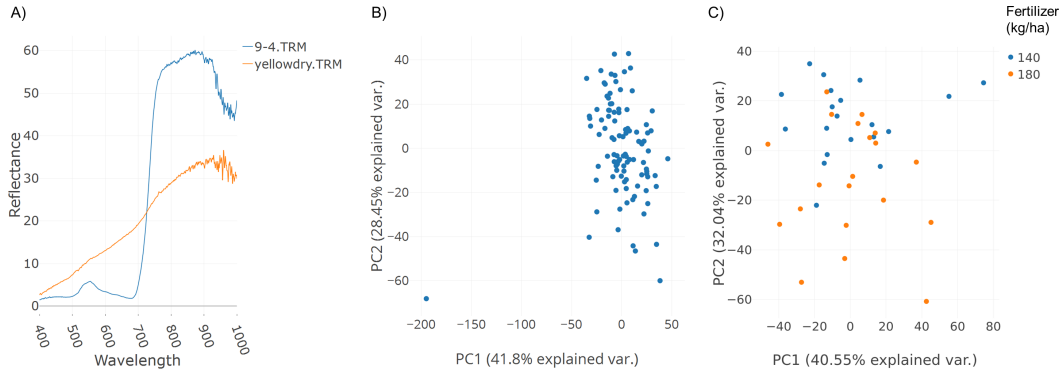


Figure 4 Outlier detection in Specalyzer. Outlier detection and removal in spectral data using Specalyzer. (A) Reflectance data collected from a winter wheat canopy (blue) compared to reflectance data collected from dry leaves (orange) (B) PCA clustering analysis of the same spectral data set with spectral samples from several wheat canopies clustering to the right and the dry leaf sample appearing to the left as an outlier (C) Recalculated PCA plot with outlier dry leaf spectral sample removed, and with the remaining wheat canopy spectral samples colored by fertilizer treatment (140 and 180 kg ha^{-1}).

Full-size DOI: 10.7717/peerj.5031/fig-4

the case study data where treatment differences are observed for the indices such as NDVI and TCARI while no differences in the treatment can be seen for the indices EVI and WI (Fig. 5). Boxplots can also be created for individual samples or for sample groupings. In a boxplot with samples grouped by replication, variation in NDVI can be seen in the

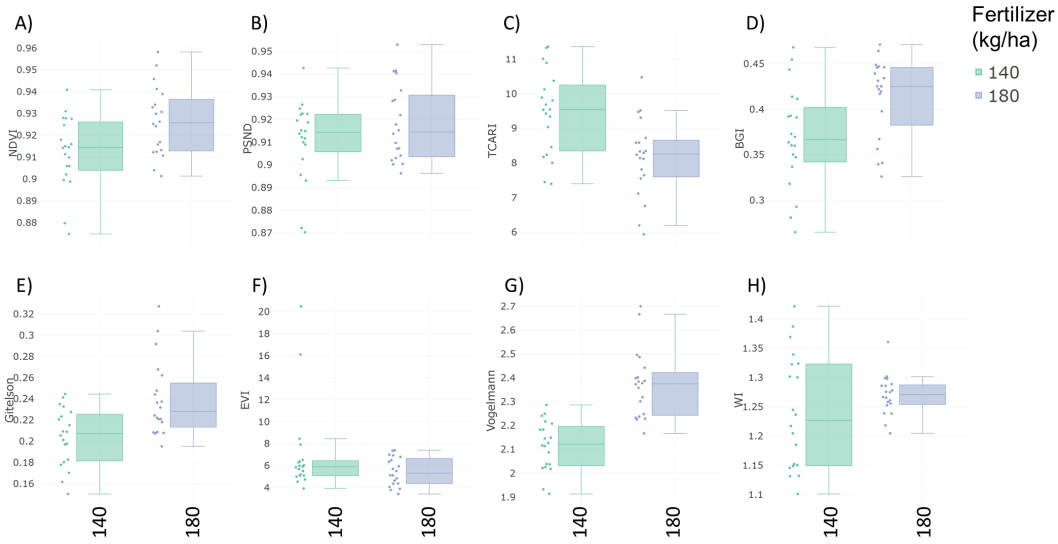


Figure 5 Boxplots of vegetation indices. Vegetation index boxplots from spectral data collected from winter wheat canopies aggregated by two fertilizer treatments (140 and 180 kg ha^{-1}). (A) Normalized Difference Vegetation Index (NDVI) (B) Pigment Specific Normalized Difference (PSND) (C) Transformed Chlorophyll Absorption Reflectance Index (TCARI) (D) BGI (E) Gitelson Index (F) Enhanced Vegetation Index (EVI) (G) Vogelmann index (H) Water Index (WI).

Full-size DOI: 10.7717/peerj.5031/fig-5

breeding lines with breeding line 5 having the highest NDVI in the treatment group with 180 kg fertilizer ha^{-1} (Fig. 6). Similar plots can be created for over 140 VIs allowing detailed analysis of the VIs and the treatments. Specalyzer also supports visualizing indices against continuous traits in scatterplots. Furthermore, fieldmaps can be created to visualize the spatial distribution of the measurements in the field and the corresponding intensities of the indices. In Fig. 7, the sample number 5 can be identified with higher NDVI levels in the treatment group with 180 kg fertilizer ha^{-1} and from the fieldmap it can be seen that higher NDVI is specific for sample 5 indicating that this sample might have higher nutrient use efficiency. The processed spectral and the VI data can be exported for further analysis in a statistical software.

DISCUSSION

Effectiveness of data visualization can be estimated based on user interactivity, ability to integrate data from different sources and ease of access to the tool (*Yau, 2013a*; *Zhu, Hoon & Teow, 2015*). Visualization allows ease of use and greater insight into the data in a way that is not obvious from descriptive statistics (*Calatroni & Wildfire, 2017*). Interactivity in data visualization facilitates overview of data followed by zooming and filtering the plots when required for greater details (*Shneiderman, 1996*). The graphical user interface of Specalyzer makes it an easy-to-use web application for exploring spectral reflectance, attributes and spatial information of datasets in any web browser, without requiring programming skills or having to install a software. The data visualization functions in Specalyzer are useful for both QC and for analyzing spectral data in relation to the attribute data. The plots

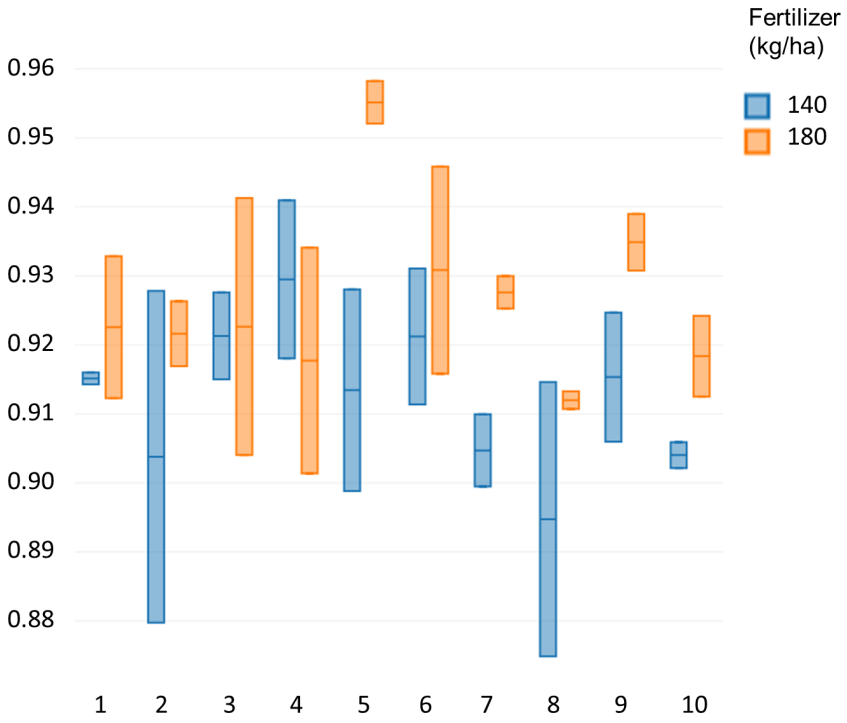


Figure 6 NDVI of the winter wheat cultivars. Comparison of NDVI values between 10 different winter wheat cultivars for two fertilizer treatments (140 and 180 kg ha^{-1}).

Full-size DOI: 10.7717/peerj.5031/fig-6

in Specalyzer are interactive, allowing users to explore the data interactively to identify dominant patterns, new insights and decision making for the underlying traits of interest. Specalyzer currently accepts data from two spectroradiometer vendors, ASD FieldSpec® (Malvern Panalytical, Malvern, Worcestershire, UK) and Apogee SpectraWiz® (Apogee Instruments Inc., Logan, UT, USA) format. Additionally, it also accepts tabulated data in a generic text file format. This allows broader application of Specalyzer for data obtained from various spectroradiometer devices.

Large data sets have hundreds of variables increasing the data complexity and thus making them difficult to visualize. Dimensionality reduction methods such as PCA allow visualizing data in a two-dimensional plane where samples with similar profiles are clustered closer together while dissimilar samples are separated in space which allows visualization of the clustering patterns and the underlying similarity matrices (*Calatroni & Wildfire, 2017*). In Specalyzer, the interactive PCA plot has the features to zoom and modify color and size of the data points based on user provided attributes, enabling analysis of trait of interest and detecting outliers.

Specalyzer is implemented as an online tool and thus can be used with any mobile platform with a web browser and access to internet. It is built with the R programming language with an interactive graphical user interface using the Shiny web application. The zooming, panning and tooltips features in charts are provided by plotly R package which is a high-level interface to the JavaScript plotting library plotly.js. This allows interactive

Figure 7 Fieldmap with NDVI. NDVI intensities visualized as a spatial grid corresponding to plots in a field trial. Each rectangle corresponds to an NDVI intensity from a specific wheat cultivar plot in a field trial where the cultivars were subjected to two different fertilizer treatments (140 and 180 kg fertilizer ha^{-1}).

Full-size DOI: 10.7717/peerj.5031/fig-7

data analysis where the output is continuously updated based on changes to the parameters by the user. There are several advantages to using Shiny and plotly in Specalyzer, (a) User-friendliness for analyzing big datasets; (b) Platform independence allowing flexibility in using devices; (c) Customized charts allowing greater control; (d) Interactivity to easily identify outliers and data points of interest and (e) Publication-quality figures. This enables, for example, analysis of the collected data with a mobile phone while in the field which can facilitate identifying individual plots using Specalyzer for further manual inspection in the field. This can save time for germplasm evaluation in the field thus reducing costs. Phenotyping carts are mounted with proximal sensors such as RGB and hyperspectral cameras, infrared thermometers and spectroradiometers which are being developed and are operated from a computer (*Deery et al., 2014*). Specalyzer can be further modified to be used with spectroradiometers on these carts enabling instantaneous analysis of the acquired data in the field.

Future work on Specalyzer will involve expanding the data visualization toolkit and improving existing data visualization functionality. For example, an important improvement is to enable users to aggregate spectral measurements by more than one attribute. Another important improvement in the data visualization menus would be adding plot layout and output controls for users to get customized publication-ready figures out of the application. Currently, VIs estimated by Specalyzer can be exported for further analysis. In a parallel project on wheat, we estimated VIs in Specalyzer and thereafter using machine learning, identified key VIs to detect the fungal disease Septoria tritici blotch of wheat (*Odilbekov et al., 2018*). Thus, another beneficial feature in Specalyzer would be to include

various machine learning methods to classify samples and identify key VIs underlying a trait of interest. Based on the case study presented here and the previous work (*Odilbekov et al., 2018*) we suggest that Specalyzer can be a useful tool for analyzing spectral reflectance data.

CONCLUSION

Efficient management and analysis of the phenotypic data is crucial and thus there is a clear need for development of new tools that allow users with broader expertise to analyze and interpret the acquired data. As this work demonstrated, Specalyzer provides an interactive graphical user interface for spectral data analysis and for estimation of several previously known VIs. Analyzing big datasets is a challenging task and thus Specalyzer can help facilitate this process. Further work is required to introduce additional features such as machine learning for variable selection and spatial analysis.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The project was supported by funding from Partnerskap Alnarp and NordForsk (No. 84597). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Partnerskap Alnarp and NordForsk: 84597.

Competing Interests

Tina Henriksson is employed by Lantmännen Lantbruk, Svalöv, Sweden.

Author Contributions

- Alexander Koc performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Tina Henriksson performed the experiments, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.
- Aakash Chawade conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The data used in this manuscript is available from the the software webpage http://www.specalyzer.org.

Specalyzer code can be accessed at https://github.com/alkc/specalyzer.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.5031#supplemental-information.

REFERENCES

Agapiou A, Hadjimitsis D, Alexakis D. 2012. Evaluation of broadband and narrowband vegetation indices for the identification of archaeological crop marks. *Remote Sensing* **4**:3892–3919 DOI 10.3390/rs4123892.

Attali D. 2018. shinyjs: easily improve the user experience of your shiny apps in seconds. R package version 1.0 ed. *Available at https://cran.r-project.org/web/packages/shinyjs/* (accessed on 11 June 2018).

Babar MA, Reynolds MP, Van Ginkel M, Klatt AR, Raun WR, Stone ML. 2006. Spectral reflectance indices as a potential indirect selection criteria for wheat yield under irrigation. *Crop Science* **46**:578–588 DOI 10.2135/cropsci2005.0059.

Calatroni A, Wildfire JJ. 2017. Graphic depiction of bioinformatics data. *Journal of Allergy and Clinical Immunology* **140**:1519–1522 DOI 10.1016/j.jaci.2017.05.043.

Cao X, Luo Y, Zhou Y, Fan J, Xu X, West JS, Duan X, Cheng D. 2015. Detection of powdery mildew in two winter wheat plant densities and prediction of grain yield using canopy hyperspectral reflectance. *PLOS ONE* **10(3)**:1–14 DOI 10.1371/journal.pone.0121462.

Chang W, Joe C, Allaire JJ, Xie Y, McPherson J. 2017. Shiny: web application framework for R. *Available at https://cran.r-project.org/web/packages/shiny/* (accessed on 11 June 2018).

Chawade A, Alexandersson E, Levander F. 2014. Normalyzer: a tool for rapid evaluation of normalization methods for omics data sets. *Journal of Proteome Research* **13**:3114–3120 DOI 10.1021/pr401264n.

Chawade A, Armoniené R, Berg G, Brazauskas G, Frostgård G, Geleta M, Gorash A, Henriksson T, Himanen K, Ingver A, Johansson E, Jørgensen LN, Koppel M, Koppel R, Makela P, Ortiz R, Podyma W, Roitsch T, Ronis A, Svensson JT, Vallenback P, Weih M. 2018. A transnational and holistic breeding approach is needed for sustainable wheat production in the Baltic Sea region. *Physiologia Plantarum* DOI 10.1111/ppl.12726.

Deery D, Jimenez-Berni J, Jones H, Sirault X, Furbank R. 2014. Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy* **4**:349–379 DOI 10.3390/agronomy4030349.

Franke J, Menz G. 2007. Multi-temporal wheat disease detection by multi-spectral remote sensing. *Precision Agriculture* **8**:161–172 DOI 10.1007/s11119-007-9036-y.

Furbank RT, Tester M. 2011. Phenomics—technologies to relieve the phenotyping bottleneck. *Trends in Plant Science* **16**:635–644 DOI 10.1016/j.tplants.2011.09.005.

Gizaw SA, Garland-Campbell K, Carter AH. 2016. Use of spectral reflectance for indirect selection of yield potential and stability in Pacific Northwest winter wheat. *Field Crops Research* **196**:199–206 DOI 10.1016/j.fcr.2016.06.022.

Haghighattalab A, González Pérez L, Mondal S, Singh D, Schinstock D, Rutkoski J, Ortiz-Monasterio I, Singh RP, Goodin D, Poland J. 2016. Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. *Plant Methods* **12(35)**:1–15 DOI 10.1186/s13007-016-0134-6.

Hansen PM, Schjoerring JK. 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment* **86**:542–553 DOI 10.1016/s0034-4257(03)00131-7.

Kerren A, Kucher K, Li Y-F, Schreiber F. 2017. BioVis explorer: a visual guide for biological data visualization techniques. *PLOS ONE* **12(11)**:1–14 DOI 10.1371/journal.pone.0187341.

Muñoz Huerta R, Guevara-Gonzalez R, Contreras-Medina L, Torres-Pacheco I, Prado-Olivarez J, Ocampo-Velazquez R. 2013. A review of methods for sensing the nitrogen status in plants: advantages, disadvantages and recent advances. *Sensors* **13**:10823–10843 DOI 10.3390/s130810823.

Lehnert LW, Meyer H, Bendix J. 2017. Hsdar: manage, analyse and simulate hyperspectral data in R. R package version 0.5.1. *Available at https://cran.r-project.org/web/packages/hsdar/* (accessed on 11 June 2018).

Mahlein A-K. 2016. Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease* **100**:241–251 DOI 10.1094/pdis-03-15-0340-fe.

Maia R, Eliason CM, Bitton P-P, Doucet SM, Shawkey MD, Tatem A. 2013. pavo: an R package for the analysis, visualization and organization of spectral data. *Methods in Ecology and Evolution* **4**:906–913 DOI 10.1111/2041-210x.12069.

Odilbekov F, Armoniené R, Henriksson T, Chawade A. 2018. Proximal phenotyping and machine learning methods to identify septoria tritici blotch disease symptoms in wheat. *Frontiers in Plant Science* **9**:1–11 DOI 10.3389/fpls.2018.00685.

O'Donoghue SI, Gavin A-C, Gehlenborg N, Goodsell DS, Hériché J-K, Nielsen CB, North C, Olson AJ, Procter JB, Shattuck DW, Walter T, Wong B. 2010. Visualizing biological data—now and in the future. *Nature Methods* **7**:S2–S4 DOI 10.1038/nmeth.f.301.

Pavelin K, Cham JA, De Matos P, Brooksbank C, Cameron G, Steinbeck C. 2012. Bioinformatics meets user-centred design: a perspective. *PLOS Computational Biology* **8(7)**:e1002554 DOI 10.1371/journal.pcbi.1002554.

Peñuelas J, Filella I, Biel C, Serrano L, SavÉ R. 1993. The reflectance at the 950–970 nm region as an indicator of plant water status. *International Journal of Remote Sensing* **14**:1887–1905 DOI 10.1080/01431169308954010.

Peñuelas J, Filella I, Gamon JA. 1995. Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytologist* **131**:291–296 DOI 10.1111/j.1469-8137.1995.tb03064.x.

R Development Core Team. 2016. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Roudier P. 2017. asdreader: reading ASD binary files in R. R package version 0.1-3 ed. *Available at https://cran.r-project.org/web/packages/asdreader/* (accessed on 11 June 2018).

Shannon P. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**:2498–2504 DOI 10.1101/gr.1239303.

Shneiderman B. 1996. The eyes have it: a task by data type taxonomy for information visualizations. In: *Proceedings 1996 IEEE symposium on visual languages*. 336–343.

Sievert C, Parmer C, Hocking T, Chamberlain S, Ram K, Corvellec M, Despouy P. 2017. plotly: create interactive web graphics via 'plotly.js'. R package version 4.7.1 ed.

Sims DA, Gamon JA. 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**:337–354 DOI 10.1016/s0034-4257(02)00010-x.

Singh A, Ganapathysubramanian B, Singh AK, Sarkar S. 2016. Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science* **21**:110–124 DOI 10.1016/j.tplants.2015.10.015.

Tattaris M, Reynolds MP, Chapman SC. 2016. A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding. *Frontiers in Plant Science* **7**:1–9 DOI 10.3389/fpls.2016.01131.

Tucker CJ, Sellers PJ. 1986. Satellite remote sensing of primary production. *International Journal of Remote Sensing* **7**:1395–1416 DOI 10.1080/01431168608948944.

Wang F-M, Huang J-F, Tang Y-L, Wang X-Z. 2007. New vegetation index and its application in estimating leaf area index of rice. *Rice Science* **14**:195–203 DOI 10.1016/s1672-6308(07)60027-4.

Wickham H. 2007. Reshaping data with the reshape package. *Journal of Statistical Software* **21(12)**:1–20 DOI 10.18637/jss.v021.i12.

Wickham H, Francois R, Henry L, Müller K. 2017a. dplyr: a grammar of data manipulation. R package version 0.7.4 ed. *Available at https://cran.r-project.org/web/packages/dplyr/* (accessed on 11 June 2018).

Wickham H, Hester J, Francois R. 2017b. readr: read rectangular text data. R package version 1.1.1 ed. *Available at https://cran.r-project.org/web/packages/readr/* (accessed on 11 June 2018).

Xue L-H, Cao W-X, Yang L-Z. 2007. Predicting grain yield and protein content in winter wheat at different N supply levels using canopy reflectance spectra. *Pedosphere* **17**:646–653 DOI 10.1016/s1002-0160(07)60077-0.

Yao X-F, Yao X, Tian Y-C, Ni J, Liu X-J, Cao W-X, Zhu Y. 2013. A new method to determine central wavelength and optimal bandwidth for predicting plant nitrogen uptake in winter wheat. *Journal of Integrative Agriculture* **12**:788–802 DOI 10.1016/s2095-3119(13)60300-7.

Yau N. 2013a. *Data points: visualizations that mean something*. Indianapolis: John Wiley & Sons, Inc.

Yau N. 2013b. Visualizing with clarity. In: Long C, ed. *Data points: visualization that means something*. Indianapolis: John Wiley & Sons, Inc., 201–240.

Zhu Z, Hoon HB, Teow K-L. 2015. Interactive data visualization techniques applied to healthcare and decision making. In: *Big data analytics in bioinformatics and healthcare*. Pennsylvania: IGI global, 46–59 DOI 10.4018/978-1-4666-6611-5.ch003.