Phenotyping a diversity panel of quinoa using UAV-retrieved leaf area index, SPAD-based chlorophyll and a random forest approach

Jiale Jiang¹ · Kasper Johansen¹ · Clara S. Stanschewski² · Gordon Wellman² · Magdi A. A. Mousa^{3,4} · Gabriele M. Fiene² · Khalid A. Asiry³ · Mark Tester² · Matthew F. McCabe¹

Accepted: 6 December 2021 / Published online: 23 January 2022
© The Author(s) 2021

Abstract

Given its high nutritional value and capacity to grow in harsh environments, quinoa has significant potential to address a range of food security concerns. Monitoring the development of phenotypic traits during field trials can provide insights into the varieties best suited to specific environmental conditions and management strategies. Unmanned aerial vehicles (UAVs) provide a promising means for phenotyping and offer the potential for new insights into relative plant performance. During a field trial exploring 141 quinoa accessions, a UAV-based multispectral camera was deployed to retrieve leaf area index (LAI) and SPAD-based chlorophyll across 378 control and 378 saline-irrigated plots using a random forest regression approach based on both individual spectral bands and 25 different vegetation indices (VIs) derived from the multispectral imagery. Results show that most VIs had stronger correlation with the LAI and SPAD-based chlorophyll measurements than individual bands. VIs including the red-edge band had high importance in SPAD-based chlorophyll predictions, while VIs including the near infrared band (but not the red-edge band) improved LAI prediction models. When applied to individual treatments (i.e. control or saline), the models trained using all data (i.e. both control and saline data) achieved high mapping accuracies for LAI ($R^2=0.977\text{--}0.980$, $\text{RMSE}=0.119\text{--}0.167$) and SPAD-based chlorophyll ($R^2=0.983\text{--}0.986$, $\text{RMSE}=2.535\text{--}2.861$). Overall, the study demonstrated that UAV-based remote sensing is not only useful for retrieving important phenotypic traits of quinoa, but that machine learning models trained on all available measurements can provide robust predictions for abiotic stress experiments.

Keywords Unmanned aerial vehicle (UAV) · Quinoa · Phenotyping · Leaf area index (LAI) · Chlorophyll · Random forest

✉ Jiale Jiang
jjiale.jiang@kaust.edu.sa

Extended author information available on the last page of the article

Introduction

Chenopodium quinoa Willd (quinoa) is a highly nutritious crop that can provide an excellent balance of fiber, lipids, carbohydrates, vitamins, and minerals (Vega-Gálvez et al., 2010). Moreover, quinoa is naturally salt tolerant and has the ability to grow within a range of harsh environments (Adolf et al., 2012; Hariadi et al., 2011). As such, quinoa has the potential to provide a highly nutritious food source that can be grown on marginal lands that are not suitable for other major crops: either now or into the future. Despite its potential, quinoa is still an underutilized crop (Massawe et al., 2016), with active breeding programs (Zurita-Silva et al., 2014) and field management strategies (Alvar-Beltrán et al., 2020) having only recently emerged. To improve agronomic traits and increase the productivity of quinoa, ongoing monitoring of phenotypic traits of different varieties and under a range of field conditions are required to help identify the suitable accessions for cultivation.

In recent years, unmanned aerial vehicles (UAVs) have become an increasingly useful tool for the collection of ultra-high spatial resolution, near real-time sensor-agnostic information (Manfreda et al., 2018; Xiang et al., 2019). While they have been increasingly employed for a variety of plant-phenotyping and crop breeding type applications, to date, relatively few studies have assessed their potential for mapping quinoa phenotypic traits. In a recent example, Sankaran et al. (2019) evaluated correlations between features extracted from multispectral and thermal UAV imagery and quinoa stomatal conductance, plant height and seed yield under non-irrigated and irrigated treatments, and found that the UAV-based green normalized difference vegetation index (GNDVI) and canopy temperature data were able to discriminate between non-irrigated and irrigated quinoa plants. In a study to identify the prediction of quinoa stomatal conductance, Ivushkin et al. (2019) utilized three UAV-mounted sensors, including a light detection and ranging (LiDAR) scanner, thermal camera and hyperspectral camera, to show that a combination of the datasets increased prediction when using a multiple linear regression model. However, apart from the two above-mentioned studies, no other research that uses UAV-based data for predicting quinoa phenotypic traits has been identified.

Among a range of crop phenotypic traits, UAV-derived LAI as a morphological trait and chlorophyll content as a physiological trait provide fundamental information with considerable application potential in crop breeding programs (Jin et al., 2020). While UAV imagery has not previously been used for predicting LAI and chlorophyll content of quinoa plants, it has been used extensively to map LAI and chlorophyll content of other crops. For example, multispectral UAV imagery has been used to predict winter wheat LAI at critical growth stages (Jiang, et al., 2019b), seasonal leaf area dynamics of sorghum breeding lines (Potgieter et al., 2017), soil–plant analysis development (SPAD) measured chlorophyll content values of maize (Deng et al., 2018) and the leaf chlorophyll content of a potato crop (Roosjen et al., 2018). Vegetation indices (VIs) extracted from UAV-based imagery have been the most commonly-used information to predict LAI and chlorophyll content of crops (Jin et al., 2020). VIs represent a combination of spectral bands designed to maximize sensitivity to vegetation characteristics while minimizing confounding factors such as soil background reflectance, directional viewing angles and atmospheric effects (Huete et al., 2002). In addition to VIs, the UAV-based 3D point cloud (including information of canopy thickness, height and leaf density distribution) and canopy textural information (e.g., mean, variance, homogeneity, entropy etc.), representing the structural characteristics of the crop canopy, can also be used for phenotyping, especially the morphological trait of crops such as LAI (Comba et al., 2020; Guillen-Climent et al., 2012; Zheng et al., 2019). However, the

ability of UAV-based imagery for mapping LAI and chlorophyll of quinoa plants remains unexplored.

Three broad types of approaches are commonly employed to assist with mapping plant phenotypic traits using remote sensing or UAV-based data. These comprise parametric, physically-based and non-parametric methods. Parametric methods are the simplest approach for UAV-based crop growth monitoring, allowing the relationship between crop traits and variables to be easily interpreted. However, they are generally not particularly robust (Zheng, et al., 2018b). Physical-based methods (e.g., radiative transfer models) are based on physical laws that can enhance the reliability of predicted results, but the requirement of site-specific information (e.g., biophysical and biochemical variables, sensor viewing and solar angles, and crop structure parameters) for model characterization tends to reduce their implementation and efficiency (Fenghua et al., 2017; Jacquemoud et al., 2009). In more recent times, machine learning methods (which fall under non-parametric methods), have been employed in crop monitoring due to their ability to discover hidden information and their lower sensitivity to data skewness (Maimaitijiang et al., 2017; Yang et al., 2017). Zheng et al. (2018b) compared a simple linear parametric model, a physical-based model and 13 non-parametric models (i.e., three non-parametric linear regressions and 10 machine learning methods) for estimating winter wheat leaf nitrogen content and found that random forest regression achieved the highest accuracy. Many other studies have also applied machine learning methods to improve the estimation of UAV-based phenotypic traits of various crops, such as corn (Lee et al., 2020), maize (Singhal et al., 2019), tomato (Johansen et al., 2020) and rice (Zha et al., 2020). However, the performance of machine learning techniques for predicting phenotypic traits of quinoa using UAV-based imagery is yet to be explored.

Overall, the research aimed to evaluate the utility of UAV-based imagery to estimate two key phenotypic traits: LAI and chlorophyll content. To do this, a random forest machine learning approach was employed to predict LAI and SPAD-based chlorophyll for both control and saline-irrigated quinoa plots based on collected UAV-based multispectral imagery. The specific objectives of this study were: (1) to retrieve the LAI and SPAD-based chlorophyll of quinoa at the plot level using multispectral UAV data and random forest models; and (2) to evaluate the effects of control and saline treatments on prediction accuracies through using training data based on control plots only, saline-irrigated plots only and all available information.

Materials and methods

Experimental site and design

The field-based quinoa phenotyping experiment was conducted during the 2019–2020 growing season at the King Abdulaziz University Agricultural Research Station in Hada Al-Sham (21.7961° N, 39.7256° E), Saudi Arabia. The site is located in a tropical arid climate with annual rainfall of < 100 mm and has a predominantly sandy loam soil type. The field trial covered an area of 90 m × 50 m, comprising 756 plots of 1 m × 1 m and distributed within four control blocks and four saline-irrigated blocks (see Fig. 1). Forty seeds were planted in each plot on November 17 and 18, 2019. Across the 756 plots, there were 141 quinoa accessions, of which 2 accessions were replicated four times, 47

Fig. 1 Quinoa phenotyping experiment at the King Abdulaziz University Agricultural Research Station in Hada Al-Sham showing the eight blocks and associated water tanks for control and saline-irrigated treatments. UAV-based MicaSense RedEdge imagery collected on February 23, 2020 shows the 40 selected plots (yellow squares) for focused field analysis and nine ground control points (blue circles) for geometric correction of the unmanned aerial vehicle (UAV) imagery. The photos of plot samples show examples of a plot under the saline and control treatments at three dates during the growing season

accessions were replicated twice, and 92 accessions were replicated three times, in each of the control and saline-irrigated treatments. Using a drip-irrigation delivery system, freshwater irrigation was applied daily for 10 min until December 18, 2019, when the saline treatments were started. Here, the water salinity was 7.5 dS/m with irrigation occurring for 20 min every second day. Each of the 756 plots contained four tubes with 12 drippers located at each 0.3 m. The dripper discharge was 4 l/h. Based on previous experiments of irrigation optimization at the same location, the irrigation times were set to 10 min per day until flowering and subsequently 20 min every second day until harvest, producing a daily average water volume for irrigation of 8 l for each plot. The salt concentration was gradually increased to approximately 17.19 dS/m on January 23, 2020 until harvest, which started on February 17, 2020. Supplementary NaCl was added because it is the major salt component in sea water, as described in Negrão et al. (2017). The electrical conductivity (EC) of the groundwater used for irrigation was measured as a base line. Then, the amount of NaCl was calculated to reach the required salt concentration level. Finally, NaCl was externally dissolved in 25 l of the groundwater, while calibrating the water EC before adding the solution to the irrigation tank.

Field measurements at the plot level

A total of 40 plots, representing five plots from each of the eight blocks, were selected for sampling to ensure that replicated accessions were assessed across different plots (Fig. 1). LAI and chlorophyll measurements were collected from the 40 selected plots on January 23, February 10 and February 23, 2020. The LAI was measured at twilight using an LAI-2200C Plant Canopy Analyzer (LI-COR Biosciences, Lincoln, NE, USA). For each plot, samples were collected next to quinoa plants at the center and from each of the four corners, with an average of the five measurements per plot subsequently calculated (LAI_{avg}). It should be noted that LAI measurements were collected next to quinoa plants, even for plots (especially those under saline-irrigation) with a small number of unevenly distributed plants. To preclude overestimation of LAI for plots with reduced plant density and to allow comparison of LAI between plots with different plant numbers, LAI_{avg} was normalized (LAI_{nor}) based on the mapped vegetation fraction (VF) within each plot:

$$LAI_{nor} = LAI_{avg} \times \frac{VF_p}{VF_{max}}, \quad VF_p = \frac{VF_{veg}}{VF_{all}} \quad (1)$$

where VF_p and VF_{max} represent the VF of each individual plot and the maximum VF of all 40 selected plots, respectively. VF_{veg} and VF_{all} are the number of mapped vegetation pixels and all pixels in each plot, respectively. Further information regarding the extraction of vegetation pixels is presented in the section of data analysis. The plant chlorophyll, hereon referred to as SPAD-based chlorophyll, from each plot was determined using a SPAD-502 Plus Chlorophyll Meter (Konica Minolta Optics Inc., Osaka, Japan). Five different upper canopy leaves of selected plants within each plot were measured three times per leaf. Upper canopy leaves in each plot were selected to provide a representation of the UAV-imaged leaves. The 15 SPAD readings per plot were averaged to derive an overall SPAD-based chlorophyll estimate for each plot. It should be noted that SPAD readings from 13 out of the 40 sampled plots were not available on February 23 as the leaves had become withered and senescent. As such, a total of 107 SPAD-based chlorophyll measurements were used for subsequent analysis.

UAV campaigns and image pre-processing

Data collection

A DJI Matrice 100 quadcopter (DJI Innovations, Shenzhen, China) provided the UAV platform for collection of MicaSense RedEdge-MX (MicaSense, Seattle, WA, USA) multispectral data. With a sensor resolution of 1280×960 pixels, the multispectral data consisted of blue (459–491 nm), green (546.5–573.5 nm), red (660–676 nm), red edge (711–723 nm) and near infrared (NIR; 813.5–870.5 nm) bands. Three UAV campaigns were performed in concert with the in-situ collection of quinoa LAI and SPAD-based chlorophyll data. The MicaSense image data were collected within one hour of solar noon under cloud-free conditions with 80% side-lap and 89% forward overlap (1 photo per second). The Universal Ground Control Station (UgCS) Client application (SPH Engineering, SIA, Riga, Latvia) was used for flight planning to collect image data, with the UAV flying at a speed of 2 m/s with a flight line distance of 4.8 m and at a height of 27 m, which produced a ground

sample distance of 0.018 m. Each flight took approximately 16 min to complete. Nine ground control points (GCPs) were placed within the study area for geometric correction (Fig. 1). The co-ordinates of GCPs were measured using a Leica GS10 base station with an AS10 antenna and a Leica GD15 smart antenna as a rover (Leica Geosystems, St Gallen, Switzerland). All collected global navigation satellite system (GNSS) data were processed using the Leica Geo Office software (Leica Geosystems, St Gallen, Switzerland) to achieve sub-centimeter positional accuracy. Before each flight, eight near-Lambertian reflectance panels were placed within the UAV flight overpass for radiometric calibration. The eight panels were produced from Masonite and had reflectance intensities of approximately 5%, 15%, 25%, 40%, 50%, 55%, 70% and 85%, with measurements taken before each flight using an ASD FieldSpec HandHeld2 spectroradiometer (Malvern Panalytical, Malvern, UK). A spectralon panel with nominal reflectance of 100% was used to optimize the ASD before measurements. For each panel, five ASD reference spectra were recorded and averaged into a single spectral sample. By averaging the ASD reflectance values within the full width at half maximum (FWHM) for each band, the ASD measurements were resampled to match the corresponding spectral bands of the MicaSense camera.

Image Pre-processing

All MicaSense imagery were processed in Agisoft Metashape 1.7 (Agisoft LLC, St. Petersburg, Russia) to produce a georeferenced ortho-mosaic for each data capture. An initial reflectance calibration step was applied to the MicaSense photos in Agisoft Metashape based on the calibration panel with a surface reflectance of 25% (Cao et al., 2019). A region of interest of one of the panel centers within a selected photo was created and the corresponding ASD measurements were imported for that panel for each band. Based on the relationship between image digital numbers and ASD measurements, all photos collected within a single flight were calibrated to ensure consistent brightness between photos before further processing. The camera model was determined automatically when the software loaded the metadata of the imported images. The information of band alignment was included in the metadata as well. After the calibration process, the photos were aligned using a key point limit of 40,000 and a tie point limit of 10,000. Tie points across bands were also generated during this process to strengthen the geometric accuracy between bands. Then, the imagery was georeferenced based on GCPs, achieving root mean square errors ranging from 0.006 to 0.014 m. The dense point cloud was produced at high quality using moderate depth filtering for all the image datasets. Since the focus was on quinoa plants at the plot level rather than the individual plant level, a point density of high quality combined with moderate depth filtering was found to provide both an accurate surface representation and computational efficiency. Subsequently, a digital elevation model (DEM) was generated using the dense point cloud as source data and while enabling hole filling. Finally, an orthomosaic was produced using the DEM surface and the default Mosaic blending mode, producing a pixel size of 0.018 m. The output bands of the MicaSense ortho-mosaics were 16-bit integer values. The average digital number (DN) of the pixels that covered each panel were calculated from the ortho-mosaic. The ASD reflectance of each panel was calculated by resampling the ASD measurements within the full width at half maximum (FWHM) for each band to match the corresponding spectral bands of the MicaSense camera. Finally, a linear empirical function was developed between the average DN of the calibration panels in the orthomosaics and the resampled ASD reflectance

to convert the ortho-mosaics to at-surface reflectance (Barreto et al., 2019; Johansen et al., 2019).

Data analysis

Extraction of variables at the plot level

Since spectral data are the most commonly-used and valuable information from UAV-based imagery to predict both LAI and chlorophyll of crops, the five spectral bands and 25 commonly used VIs (Table 1) were extracted from the MicaSense ortho-mosaics for each quinoa plot and used to build prediction models of LAI and SPAD-based chlorophyll at the plot level. To extract variables at the plot level, a shapefile was produced and manually placed for each plot using a square polygon with dimensions of $1.2\text{ m} \times 1.2\text{ m}$ (1.44 m^2). As some plants appeared outside the boundary of the 1 m^2 plots due to plant growth, especially during the last two campaigns, the larger polygon area ensured most plants per plot were included. Calculation of LAI was based on all pixels within this 1.44 m^2 area. Since SPAD-based chlorophyll is specific to the plant leaves, the variables of all pixels representing quinoa plants were extracted by excluding canopy gaps within the 1.44 m^2 area in the comparisons between pixel-derived and field-measured SPAD-based chlorophyll.

Quinoa plant pixels were identified and extracted using a support vector machine (SVM) classification performed in the ENVI software (EXELIS, Boulder, CO, USA). An SVM is a supervised machine learning model that has been widely used for the classification of remotely sensed imagery (Maulik & Chakraborty, 2017; Roli & Fumera, 2001). Regions of interest were selected for two classes i.e. quinoa and non-quinoa features, with all spectral bands used as training data. The SVM settings were set to the default values of the radial basis function (i.e., gamma in kernel function=0.25, penalty parameter=100, pyramid levels=0, and classification probability threshold=0). Based on ground truth data, the Kappa coefficients of the SVM classifications for all UAV imagery were $>97\%$. Consequently, the variables of the pixels representing quinoa plants were extracted for SPAD-based chlorophyll estimation. Additionally, the extracted plant pixels were used to calculate the vegetation fraction values for each plot (Eq. 1).

Estimation and evaluation of LAI and SPAD-based chlorophyll predictions

Here, random forest regression models were used to estimate LAI and SPAD-based chlorophyll. The random forest machine learning approach has been shown to provide a fast and accurate method for estimating crop traits (Houborg & McCabe, 2018; Shah et al., 2019). Three random forest models were developed for each of the LAI and SPAD-based chlorophyll UAV datasets, using a training data set based on: (1) all observations, (2) control observations only and, (3) saline observations only. To develop the models, the optimal number of variables available for splitting at each tree node (mtry) was first identified. The mtry values were tested across a range from 1 to 30 (the number of input variables) via an iterative trial-and-error process to determine the optimal mtry value for each model. The optimal mtry value for each of the three models for both LAI and SPAD-based chlorophyll was identified by the value achieving the smallest mean squared error. The optimal mtry values were 15 and 8 for the LAI and SPAD-based chlorophyll models, respectively. A total of 1000 decision trees (ntree) were used for all random forest models to ensure stable predictions and computational efficiency. The variables used in each random forest

Table 1 Vegetation indices used in this study

Variable	Abbreviation	Full name	Formulation ^a	References
CIgreen		Green chlorophyll index	NIR/G-1	Gitelson et al. (2003)
CIre		Chlorophyll index red-edge	NIR/RE-1	Gitelson et al. (2006)
DVI		Difference vegetation index	NIR-R	Jordan (1969)
EVI		Enhanced vegetation index	$2.5 \times (\text{NIR-R})/(\text{NIR} + 6\text{R} - 7.5\text{G} + 1)$	Huete et al. (1994)
ExG		Excess green vegetation index	$(2\text{G-R-B})/(\text{G} + \text{R} + \text{B})$	Woebbecke et al. (1995)
GBNDVI		Green-Blue normalized difference vegetation index	$(\text{NIR} - (\text{G} + \text{B})) / (\text{NIR} + (\text{G} + \text{B}))$	Wang et al. (2007)
GNDVI		Green normalized difference vegetation index	$(\text{NIR-G})/(\text{NIR} + \text{G})$	Gitelson and Merzlyak (1998)
MEVI		Modified enhanced vegetation index	$2.5 \times (\text{NIR-RE})/(\text{NIR} + 6\text{RE} - 7.5\text{G} + 1)$	Justice et al. (1998)
MSR		Modified simple ratio	$(\text{NIR/RE}-1)/(\text{NIR/RE} + 1)^{0.5}$	Chen (1996)
MTCI		MERIS terrestrial chlorophyll index	$(\text{NIR-RE})/(\text{RE-R})$	Dash and Curran (2004)
MTVI ₂		Modified triangular vegetation index	$1.5 \times (1.2 \times (\text{NIR-G}) - 2.5 \times (\text{RE-G})) / (2 \times (\text{NIR} + 1)^2 - (\text{NIR-5RE})^{0.5} - 0.5)^{0.5}$	Haboudane et al. (2004)
NDRE		Normalized difference red-edge index	$(\text{NIR-RE})/(\text{NIR} + \text{RE})$	Sims and Gamon (2002)
NDVI		Normalized difference vegetation index	$(\text{NIR-R})/(\text{NIR} + \text{R})$	Schell and Deering (1973)
NGRDI		Normalized green red difference index	$(\text{G-R})/(\text{G} + \text{R})$	Gitelson et al. (2002)
NIRRENDVI		NIR-RE normalized difference vegetation index	$((\text{NIR} + \text{RE})/2 - \text{R})/((\text{NIR} + \text{RE})/2 + \text{R})$	Xie et al. (2018)
NPCI		Normalized pigment chlorophyll ratio index	$(\text{R-B})/(\text{R} + \text{B})$	Peñuelas et al. (1994)
PSRI		Plant senescence reflectance index	$(\text{R-G})/\text{NIR}$	Merzlyak et al. (1999)
REDVI		Red-edge difference vegetation index	NIR-RE	Tucker (1979)
RENDVI		Red-edge normalized difference vegetation index	$(\text{RE-R})/(\text{RE} + \text{R})$	Sims and Gamon (2002)
RERVI		Red-edge ratio vegetation index	NIR/RE	Jasper et al. (2009)
RGRI		Red green ratio index	R/G	Gamon and Surfus (1999)
RVI		Ratio vegetation index	NIR/R	Birth and McVey (1968)
SARE		Soil-adjusted red-edge index	$(\text{NIR-RE})/(\text{NIR} + \text{RE} + 0.25) + 0.25$	Ali et al. (2015)
SAVI		Soil-adjusted vegetation index	$(\text{NIR-R})/(\text{NIR} + \text{R} + 0.25) + 0.25$	Huete (1988)

Table 1 (continued)

Variable		Formulation ^a	References
Abbreviation	Full name		
TCVI	True color vegetation index	$1.4 \times (2R - 2B) / 2R - G - 2B + 0.4$	Jiang et al. (2019a)

^aB, G, R, RE, and NIR represent the blue, green, red, red-edge and near infrared bands, respectively

regression model included the five spectral bands and 25 commonly used VIs identified in Table 1. Similar to previous research (e.g. Johansen et al., 2020), training sets were compiled by randomly sampling approximately 2/3 of the entire dataset with replacement, with the remaining data used as the testing set.

Individual random forest models were developed based on all quinoa plots, the control plots only, and the saline plots only. In each dataset, there were 30 variables, comprising the five spectral bands and the 25 VIs (see Table 1). To assess the performance of the 25 VIs, they were divided into three categories: (1) VIs with only the visible bands (i.e., the blue, green and red bands); (2) VIs with the NIR band, but excluding the red-edge band; and (3) VIs including the red-edge band. The importance of each variable was evaluated based on how much the accuracy decreased when a variable was excluded (Liaw & Wiener, 2002): determined here using the percentage increase in mean squared error (%IncMSE). The accuracy of the least-squares and random forest models were evaluated using coefficient of determinations (R^2) and the root-mean-squared error (RMSE) between predicted and measured LAI and SPAD-based chlorophyll. The best performing models were subsequently applied to all 756 plots to predict LAI and SPAD-based chlorophyll for each of the three UAV dates and to assess differences between control and saline-irrigated plots. In all cases, the random forest models were implemented with the R package “randomForest” (<https://CRAN.R-project.org/package=randomForest>).

Results

Evaluation of variables importance

The importance of different variables for estimation of LAI and SPAD-based chlorophyll using random forest regression was evaluated, with results presented in Figs. 2 and 3. As can be seen, individual spectral bands were generally less important than VIs. For instance, the percentage increase in mean squared error (%IncMSE) values were close to 0 for the green and red-edge bands of the MicaSense camera for LAI estimation (Fig. 2) and the NIR band for SPAD-based chlorophyll estimation (Fig. 3), indicating little contribution of those bands for the predictions.

The top 20% of important variables were identified for each model, representing 6 out of 30 variables (see boxed outlines in Figs. 2 and 3). For the LAI models, most VIs with the NIR band (but without the red-edge band) were highly ranked in their importance. Five VIs including the NIR band (but excluding the red-edge band) were among the six highest ranked variables when predicting LAI of the control plots only (Fig. 2a), whereas the LAI models based on the saline plots only (Fig. 2b) and all plots (Fig. 2c) comprised three and four NIR VIs, respectively, among their six highest ranked variables. Apart from RENDVI, which was ranked in the top 20% most important variables in the LAI model developed on the control plots only (Fig. 2a), the contributions of VIs including the red-edge band for LAI predictions were limited (Fig. 2). In contrast, VIs with the red-edge band were significantly more important in the SPAD-based chlorophyll models (Fig. 3). This was particularly the case for the SPAD-based chlorophyll model using the control plots only, where the top 20% most important variables consisted entirely of VIs with the red-edge band (Fig. 3a). Overall, the results of variable importance showed that the red-edge band performed well for estimating SPAD-based chlorophyll ($R^2 > 0.7$), while it correlated

Fig. 2 Variables importance, as measured using the percentage increase in mean squared error (%IncMSE), in the random forest regression models constructed on: **a** the control plots only; **b** the saline plots only; and **c** all plots for predicting leaf area index (LAI). Vegetation index abbreviations are expanded on in Table 1

poorly with LAI ($R^2 < 0.2$). On the other hand, the NIR band was correlated well with LAI ($R^2 > 0.45$) but showed almost no correlation with chlorophyll ($R^2 < 0.08$).

To compare variable importance in the random forest models including different datasets (all, control only, saline only) for predicting LAI and SPAD-based chlorophyll, the variable importance was sorted into five ranked categories (Fig. 4). The variables with the highest importance ranking in Fig. 4 corresponded to the top 20% most important variables outlined in Figs. 2 and 3. For the LAI prediction models developed on the control plot only, saline plot only and all plots, the variable importance ranking was the same in 13 out of 30 cases, while the remaining variables (except NIR, Clgreen, and MTVI2) were within a single importance ranking category of each other. Conversely, the ranking of variable importance for predicting chlorophyll showed larger variation in the three models (control only, saline only, all). Only 7 out of 30 variables had the same importance ranking across the three models, while 11 variables were within a single importance ranking of

Fig. 3 Variables importance, as measured using the percentage increase in mean squared error (%IncMSE), in the random forest regression models constructed on: **a** the control plots only; **b** the saline plots only; and **c** all plots for predicting SPAD-based chlorophyll. Vegetation index abbreviations are expanded on in Table 1

each other. The remaining 12 variables showed larger variation in importance between the three models. For instance, NPCI was one of the top 20% most important variables in the model developed on saline plots only, but was ranked 24th out of 30 variables in the model trained on control plots only (Figs. 3 and 4). Based on the observations of importance and the variation of importance variables between different models, it is clearly important not to exclude variables in one model based on findings from another model.

Assessment of model predictions

To explore the generality of the different model constructions, the random forest models developed on all quinoa plots were applied to all plots, the control plots only and the saline plots only, whereas the models developed on the control plots only and the saline plots only

Fig. 4 The rankings of variable importance in the random forest regression models developed on the control plots only (C), the saline plots only (S), and all plots (All) for the estimation of leaf area index (LAI) and SPAD-based chlorophyll. Abbreviations for the vegetation indices are expanded upon in Table 1

Fig. 5 Scatterplots, coefficient of determination (R^2), root mean square error (RMSE) and number of samples (N) of predicted measurements of (a, c) leaf area index (LAI) and (d, f) SPAD-based chlorophyll based on the MicaSense RedEdge unmanned aerial vehicle (UAV) imagery using random forest models developed on (a, d) all quinoa plots or only control plots, applied only to the control plots, b, e all quinoa plots or only saline plots, applied only to the saline plots, and c, f all plots, applied to all plots

were only applied to those specific plots. Compared to the models trained on only control or saline plots, the models trained using all plots slightly improved the LAI and SPAD-based chlorophyll predictions when applied to the control or saline plots (Fig. 5). This may be attributed to the larger range of measurements in the dataset of all plots (including both control and saline plots), as the lowest and highest measures of LAI and SPAD-based chlorophyll belonged to the control and saline datasets, respectively. For predicting LAI, a reduction in RMSE of 9.58% and 36.13% was observed when using the model developed

Fig. 6 Box plots of the variation in the leaf area index (LAI) and SPAD-based chlorophyll estimates from the MicaSense RedEdge camera using the random forest regression models developed on all measured plots for all 756 control and saline plots on January 23, February 11 and February 23, 2020

on all plots and applied to the control plots and saline plots, respectively, while the increase in R^2 was negligible (≤ 0.01) (Fig. 5a, b). For predicting SPAD-based chlorophyll, the models developed on and applied to the control or saline plots had very similar values of R^2 (difference ≤ 0.002) and RMSE (difference < 0.18) to the results produced when using the model developed on all plots (Fig. 5d, e). These results indicate that the models developed on all plots could improve the LAI predictions of the control or saline plots, whereas predicting SPAD-based chlorophyll of control or saline plots were not influenced by the inclusion of additional plots from both treatments within the models.

Mapping of LAI and SPAD-based chlorophyll

Since the models developed on all measured quinoa plots provided the best results, they were applied to predict the variations of LAI and SPAD-based chlorophyll for all 756 quinoa plots across each of the three dates. As shown in Fig. 6, values of LAI and SPAD-based chlorophyll for both control and saline quinoa plots generally decreased from January 23 to February 23, reflecting the observations that plant leaves tended to wither and senesce towards harvest. Compared to SPAD-based chlorophyll, the LAI predictions showed considerably more outliers, with high values located outside the whiskers of the boxplot. These outliers were likely due to lodging of plants, which is the permanent displacement of crop stems from their vertical position as a result of stem buckling and/or root displacement. Lodging was found in 25 control and 17 salt plots and appeared to be wind-related rather than influenced by treatment type. The LAI values (except the minimum value) of saline plots were slightly higher than those of the control plots on January 23 and February 11, while the control plots showed higher LAI values (e.g., the maximum, median and mean values) than the saline plots on February 23. The SPAD-based chlorophyll values (except the maximum value) of the control plots were higher than the saline plots on all three dates. The differences in LAI and SPAD-based chlorophyll values between control and saline plots for the last image campaign on February 23 likely responded to the final field irrigation occurring on February 19 (as quinoa plants need to dry out prior to harvest; Aguilar and Jacobsen (2003) and the accumulating salinity stress in the saline plots. It is noted that the range of SPAD-based chlorophyll values increased through the growing season, likely reflecting the varying salt tolerance of the different quinoa accessions.

Fig. 7 MicaSense RedEdge based maps of leaf area index (LAI) and SPAD-based chlorophyll using the random forest regression model trained on all plots i.e. for all 756 control (C) and salinity-treated (S) plots on February 23, 2020. Frequency plots show the distribution of LAI and SPAD-based chlorophyll estimates

Figure 7 shows the spatial pattern of the control and saline-irrigated quinoa plots in terms of LAI and SPAD-based chlorophyll as derived from the MicaSense imagery collected on February 23. Most of the control plots (59.0%) and saline plots (59.8%) had LAI values between 0.3 and 0.9. LAI values higher than 0.9 were mainly distributed in the control plots (with 4.5% more plots than the saline plots), while relatively lower values ($LAI < 0.3$) mostly appeared in the saline plots (with 3.7% more plots than the control plots). Similarly, the predicted distribution in terms of SPAD-based chlorophyll also showed more control plots in the high value range (> 50) and more saline plots in the low value range (< 10). Although 4.0% more plots under saline treatment than control plots occurred with SPAD-based chlorophyll value between 40 and 50, SPAD-based chlorophyll values higher than 50 mostly appeared in the control plots (37.8%) with 13.8% more occurrences than the saline plots. Conversely, 6.6% fewer plots with a SPAD-based chlorophyll value < 10 were distributed in the control plots than the saline plots. While the low LAI or SPAD-based chlorophyll values in the saline plots were likely due to salinity stress, 51 out of 141 plant accessions in the saline plots had similar or even larger average LAI and SPAD-based chlorophyll values than their corresponding replicates in the control plots, possibly indicating high salt tolerance.

Discussion

To predict LAI and SPAD-based chlorophyll in quinoa at the plot level, random forest machine learning models were used and evaluated how different variables derived from a UAV-based multispectral camera affected the prediction accuracy. The results illustrate

that vegetation indices (VIs) were generally more important than the individual spectral bands for predicting LAI and SPAD-based chlorophyll, which is in line with the findings of previous studies using VIs to mitigate the effects of illumination geometry and enhance the spectral information of crops (Fernandez-Gallego et al., 2019; Lillesand et al., 2015; Xue & Su, 2017). Many studies have indicated that using red-edge reflectance bands can improve the prediction of chlorophyll in crops such as rice (Evri et al., 2008), winter wheat (Kanke et al., 2012), and maize (Deng et al., 2018). Likewise, in this research on quinoa, the red-edge band and VIs with the red-edge band proved useful for predicting SPAD-based chlorophyll. The result is likely due to the sensitivity of the red-edge wavelength region (680–760 nm) to chlorophyll contents. In the red-edge range, the reflectance of green plants increases fast in response to the increased chlorophyll absorption of light in the red portion of the spectrum (Horler et al., 1983; Le Maire et al., 2008; Zarco-Tejada et al., 2019). Thus, the red-edge band and associated VIs are relatively sensitive to the SPAD-chlorophyll value. In addition, Deng et al. (2018) found that the red-edge VIs were less sensitive to areas with low vegetation coverage. Hence, non-vegetation features (e.g., soil and shadows) have limited effects on the prediction of the SPAD-chlorophyll based on the red-edge VIs. Measurements of other biophysical and biochemical parameters of crops have also been found to be highly correlated with the red-edge reflectance, with examples including nitrogen contents in winter wheat (Zheng et al., 2018a), biomass in corn and soybean (Kross et al., 2015), and LAI in wheat and maize (Sun et al., 2019). On the other hand, the results showed that predictions of LAI were not improved by using the red-edge band or from the VIs that included the red-edge band. Instead, the NIR band and the VIs based on the NIR band generally had higher importance in the random forest models. This is consistent with a number of previous studies that have detailed that NIR reflectance (700–1000 nm) can be strongly influenced by LAI, while being less sensitive to variations in chlorophyll content (Delegido et al., 2013; Sun et al., 2019; Yao et al., 2017). Overall, the findings demonstrate the utility of using multispectral cameras that include both a red-edge and NIR band when mapping phenotypic traits such as chlorophyll and LAI. Therefore, the VIs including red-edge and NIR bands are recommended for the predictions of LAI- or chlorophyll-related phenotypic information, such as biomass, nitrogen and yield. While this study focused on spectral information (i.e. spectral bands and VIs) and achieved high accuracies for mapping LAI and SPAD-chlorophyll, the use of 3D point clouds and canopy textural information from UAV-based imagery and their combination with VIs should be explored in future research for quinoa phenotyping, particularly of morphometric characteristics.

Results also identified that separate random forest models specifically developed for LAI estimation of control or saline plots improved predictions of plots under different treatment regimes. Conversely, SPAD-based chlorophyll of control and saline plots could be predicted with similar accuracies using models developed on either plots of the specific treatments or all plots including both treatments, possibly due to the control (0.82–70.22 based on field data) and saline plots (0.80–69.76 based on field data) having a similar range of SPAD-based chlorophyll values. Considering the high accuracies of predicting LAI ($R^2 > 0.97$ and $RMSE < 0.19$) and SPAD-based chlorophyll ($R^2 > 0.98$ and $RMSE < 3.04$), the developed random forest models did not seem to be affected by the large variety of accessions (34 out of 141 accessions included in the training data). Other studies have also achieved good results when using a random forest approach on predicting phenotypic traits in crops of different accessions. For example, Sun et al. (2021) found that random forest regression resulted in better accuracies than linear regressions for predicting canopy chlorophyll content in 24 varieties of winter wheat and one variety of soybean. Yuan et al.

(2017) compared random forest, artificial neural network, support vector machine and a partial least-squares regression models for retrieving LAI in 46 varieties of soybeans and found that random forest models achieved the best performance. In line with the current study, these previous efforts demonstrate that the random forest approach is suitable for predicting phenotypic traits from a large selection of sample plots even when variations in accessions, treatments and progression in crop growth are present. An advantage of random forest is its ability to accommodate large datasets through non-linear regression based on a large number of decision trees and high-dimensional data training (Breiman, 2001). However, it is worth noting that a random forest cannot generally predict beyond the training data range, as it only uses the values of the training dataset to grow regression trees (Iqbal et al., 2018). Therefore, it is important to carefully select training datasets covering a wide and representative variable range to improve model accuracy (Maxwell et al., 2018).

Using the LAI and SPAD-based chlorophyll results from the random forest models, traits likely to contribute to and/or reflect salinity tolerance in quinoa plants could be quantified. Approximately 36% of the 141 quinoa accessions in the saline plots showed LAI and SPAD-based chlorophyll values that were either similar or higher than their corresponding replicates in the control plots. These specific accessions might indicate a genetic tolerance to environmental stress, and provide an avenue for further identification of salinity tolerant accessions for breeding studies. However, more phenotypic information may further facilitate UAV image-derived identification of salt tolerance in quinoa accessions, e.g. plant height, canopy temperature, biomass and yield. Based on the results of quinoa phenotyping in this study, it is likely that multispectral UAV data combined with random forest modeling can be applied for deriving information on additional quinoa traits. In addition, observations from multiple sensors may assist in identifying accessions with high salinity tolerance. For example, Awlia et al., (2016, 2021) used red–green–blue (RGB) and chlorophyll fluorescence (ChlF) imaging to capture the growth, morphology, color and photosynthetic performance for identifying salt tolerance in *Arabidopsis thaliana* plants. Malbeteau et al. (2021) and Stutsel et al. (2021) used UAV-based thermal data to monitor canopy temperature of wild tomato species in response to salt stress. Ivushkin et al. (2019) combined plant height from a LiDAR scanner, the UAV-based hyperspectral physiological reflectance index (PRI) and canopy temperature from UAV-based thermal cameras to increase the discrimination of quinoa plants in the control and salt-treated plots. While all of these studies showcase the potential of UAV data to detect salt stress in crop, the underlying genetic insights that deliver the observable phenotypic features need to be investigated (McCabe & Tester, 2021). Such investigations will be assisted by the collection of additional phenotypic traits and identification of the most suitable image datasets and modeling approaches from which to derive these observable features.

Conclusion

The potential of multispectral UAV imagery to map LAI and SPAD-based chlorophyll and then to discriminate plant response to control and salt treatments in a quinoa phenotyping experiment was explored using a random forest machine learning approach. An examination of the relative information content delivered through using individual spectral bands and VIs derived from multispectral UAV imagery was also detailed. Results illustrated that VIs were generally more informative than individual bands for predicting LAI and SPAD-based chlorophyll. Most of the VIs with the red-edge band were found to be important

predictor variables in the SPAD-based chlorophyll models, whereas the VIs including the NIR band (but without the red-edge band) had the highest importance for predicting LAI. In order to capture the largest number of observable phenotypes, it is recommended to use a camera with both NIR and red-edge bands for mapping quinoa traits. The performance of different training datasets for predicting LAI and SPAD-based chlorophyll of control and saline-irrigated plots were also evaluated. Compared to those models trained on control or saline plots only, the models developed using all plots decreased the RMSE by up to 9.58% and 36.13% when predicting LAI of the control and saline plots, respectively. Limited RMSE reduction (<5.8%) occurred in the results for predicting SPAD-based chlorophyll. Overall, the results demonstrate the practical utility of employing multispectral UAV-based sensing for the estimation of key phenotypic traits of quinoa and provide a foundation for further exploration and identification of potential salinity tolerances. In addition to LAI and SPAD-based chlorophyll, other observable traits related to yield and salt tolerance such as plant height, canopy temperature and growth rate, can be integrated into the analysis to identify desirable commercial traits of the quinoa accessions studied.

Acknowledgements Research reported herein was supported through a collaborative project between the King Abdallah University of Science and Technology (KAUST) and King Abdulaziz University (KAU) under grant numbers KAUST-OSR-3936 and KAUST-JP-19-005. The authors, therefore, acknowledge with thanks KAU & KAUST for their technical and financial support. We would also like to thank Omar Lopez Camargo, Samir AlMashharawi, Alejandra Barreto and Chunfeng Ma for their assistance during the data and field collection campaigns.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adolf, V. I., Shabala, S., Andersen, M. N., Razzaghi, F., & Jacobsen, S.-E. (2012). Varietal differences of quinoa's tolerance to saline conditions. *Plant and Soil*, 357(1–2), 117–129. <https://doi.org/10.1111/j.1439-037X.2010.00446.x>
- Aguilar, P. C., & Jacobsen, S.-E. (2003). Cultivation of quinoa on the Peruvian Altiplano. *Food Reviews International*, 19, 31–41. <https://doi.org/10.1081/FRI-120018866>
- Ali, M., Montzka, C., Stadler, A., Menz, G., Thonfeld, F., & Vereecken, H. (2015). Estimation and validation of RapidEye-based time-series of leaf area index for winter wheat in the Ruhr catchment (Germany). *Remote Sensing*, 7(3), 2808–2831. <https://doi.org/10.3390/rs70302808>
- Alvar-Beltrán, J., Fabbri, C., Verdi, L., Truschi, S., Dalla Marta, A., & Orlandini, S. (2020). Testing proximal optical sensors on Quinoa growth and development. *Remote Sensing*. <https://doi.org/10.3390/rs12121958>
- Awlia, M., Alshareef, N., Saber, N., Korte, A., Oakey, H., Panzarová, K., et al. (2021). Genetic mapping of the early responses to salt stress in *Arabidopsis thaliana*. *Plant Journal*, 107(2), 544–563. <https://doi.org/10.1111/tpj.15310>
- Awlia, M., Nigro, A., Fajkus, J., Schmoeckel, S. M., Negrão, S., Santelia, D., et al. (2016). High-throughput non-destructive phenotyping of traits that contribute to salinity tolerance in *Arabidopsis thaliana*. *Frontiers in Plant Science*, 7, 1414. <https://doi.org/10.3389/fpls.2016.01414>
- Barreto, M. A. P., Johansen, K., Angel, Y., & McCabe, M. F. (2019). Radiometric assessment of a UAV-based push-broom hyperspectral camera. *Sensors*, 19(21), 4699. <https://doi.org/10.3390/s19124699>

- Birth, G. S., & McVey, G. R. (1968). Measuring the color of growing turf with a reflectance spectrophotometer. *Agronomy Journal*, 60(6), 640–643. <https://doi.org/10.2134/agronj1968.00021962006000060016x>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Cao, S., Danielson, B., Clare, S., Koenig, S., Campos-Vargas, C., & Sanchez-Azofeifa, A. (2019). Radiometric calibration assessments for UAS-borne multispectral cameras: Laboratory and field protocols. *ISPRS Journal of Photogrammetry and Remote Sensing*, 149, 132–145. <https://doi.org/10.1016/j.isprsjprs.2019.01.016>
- Chen, J. M. (1996). Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Canadian Journal of Remote Sensing*, 22(3), 229–242. <https://doi.org/10.1080/07038992.1996.10855178>
- Comba, L., Biglia, A., Ricauda Aimonino, D., Tortia, C., Mania, E., Guidoni, S., et al. (2020). Leaf Area Index evaluation in vineyards using 3D point clouds from UAV imagery. *Precision Agriculture*, 21, 881–896. <https://doi.org/10.1007/s11119-019-09699-x>
- Dash, J., & Curran, P. J. (2004). The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing*, 25(23), 5403–5413. <https://doi.org/10.1080/0143116042000274015>
- Delegido, J., Verrelst, J., Meza, C., Rivera, J., Alonso, L., & Moreno, J. (2013). A red-edge spectral index for remote sensing estimation of green LAI over agroecosystems. *European Journal of Agronomy*, 46, 42–52. <https://doi.org/10.1016/j.eja.2012.12.001>
- Deng, L., Mao, Z., Li, X., Hu, Z., Duan, F., & Yan, Y. (2018). UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 124–136. <https://doi.org/10.1016/j.isprsjprs.2018.09.008>
- Evri, M., Akiyama, T., & Kawamura, K. (2008). Spectrum analysis of hyperspectral red edge position to predict rice biophysical parameters and grain weight. *Journal of the Japan Society of Photogrammetry and Remote Sensing*, 47(2), 4–15. https://doi.org/10.4287/jsprs.47.2_4
- Fenghua, Y., Tongyu, X., Wen, D., Hang, M., Guosheng, Z., & Chunling, C. (2017). Radiative transfer models (RTMs) for field phenotyping inversion of rice based on UAV hyperspectral remote sensing. *International Journal of Agricultural and Biological Engineering*, 10(4), 150–157. <https://doi.org/10.25165/j.ijabe.20171004.3076>
- Fernandez-Gallego, J. A., Kefauver, S. C., Vatter, T., Gutiérrez, N. A., Nieto-Taladriz, M. T., & Araus, J. L. (2019). Low-cost assessment of grain yield in durum wheat using RGB images. *European Journal of Agronomy*, 105, 146–156. <https://doi.org/10.1016/j.eja.2019.02.007>
- Gamon, J., & Surfus, J. (1999). Assessing leaf pigment content and activity with a reflectometer. *The New Phytologist*, 143(1), 105–117. <https://doi.org/10.1046/j.1469-8137.1999.00424.x>
- Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, 160(3), 271–282. <https://doi.org/10.1078/0176-1617-00887>
- Gitelson, A. A., Kaufman, Y. J., Stark, R., & Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, 80(1), 76–87. [https://doi.org/10.1016/S0034-4257\(01\)00289-9](https://doi.org/10.1016/S0034-4257(01)00289-9)
- Gitelson, A. A., Keydan, G. P., & Merzlyak, M. N. (2006). Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophysical Research Letters*. <https://doi.org/10.1029/2006GL026457>
- Gitelson, A. A., & Merzlyak, M. N. (1998). Remote sensing of chlorophyll concentration in higher plant leaves. *Advances in Space Research*, 22(5), 689–692. [https://doi.org/10.1016/S0273-1177\(97\)01133-2](https://doi.org/10.1016/S0273-1177(97)01133-2)
- Guillen-Climent, M. L., Zarco-Tejada, P. J., Berni, J. A. J., North, P. R. J., & Villalobos, F. J. (2012). Mapping radiation interception in row-structured orchards using 3D simulation and high-resolution airborne imagery acquired from a UAV. *Precision Agriculture*, 13, 473–500. <https://doi.org/10.1007/s11119-012-9263-8>
- Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90(3), 337–352. <https://doi.org/10.1016/j.rse.2003.12.013>
- Hariadi, Y., Marandon, K., Tian, Y., Jacobsen, S.-E., & Shabala, S. (2011). Ionic and osmotic relations in quinoa (*Chenopodium quinoa* Willd.) plants grown at various salinity levels. *Journal of Experimental Botany*, 62(1), 185–193. <https://doi.org/10.1093/jxb/erq257>
- Horler, D., Dockray, M., & Barber, J. (1983). The red edge of plant leaf reflectance. *International Journal of Remote Sensing*, 4(2), 273–288. <https://doi.org/10.1080/01431168308948546>

- Houborg, R., & McCabe, M. F. (2018). A hybrid training approach for leaf area index estimation via Cubist and random forests machine-learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 135, 173–188. <https://doi.org/10.1016/j.isprsjprs.2017.10.004>
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83, 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- Huete, A., Justice, C., & Liu, H. (1994). Development of vegetation and soil indices for MODIS-EOS. *Remote Sensing of Environment*, 49(3), 224–234. [https://doi.org/10.1016/0034-4257\(94\)90018-3](https://doi.org/10.1016/0034-4257(94)90018-3)
- Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
- Iqbal, F., Lucieer, A., & Barry, K. (2018). Poppy crop capsule volume estimation using UAS remote sensing and random forest regression. *International Journal of Applied Earth Observation and Geoinformation*, 73, 362–373. <https://doi.org/10.1016/j.jag.2018.06.008>
- Ivushkin, K., Bartholomeus, H., Bregt, A. K., Pulatov, A., Franceschini, M. H. D., Kramer, H., et al. (2019). UAV based soil salinity assessment of cropland. *Geoderma*, 338, 502–512. <https://doi.org/10.1016/j.geoderma.2018.09.046>
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., et al. (2009). PROSPECT+ SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, 113, S56–S66. <https://doi.org/10.1016/j.rse.2008.01.026>
- Jasper, J., Reusch, S., & Link, A. (2009). Active sensing of the N status of wheat using optimized wavelength combination: impact of seed rate, variety and growth stage. In E. J. van Henten, D. Goense, & C. Lokhorst (Eds.), *Precision agriculture '09, Proceedings of the 7th European Conference on Precision Agriculture* (pp. 23–30). Wageningen, The Netherlands: Wageningen Academic Publishers.
- Jiang, J., Cai, W., Zheng, H., Cheng, T., Tian, Y., Zhu, Y., et al. (2019a). Using digital cameras on an unmanned aerial vehicle to derive optimum color vegetation indices for leaf nitrogen concentration monitoring in winter wheat. *Remote Sensing*, 11(22), 2667. <https://doi.org/10.3390/rs11222667>
- Jiang, J., Zheng, H., Ji, X., Cheng, T., Tian, Y., Zhu, Y., et al. (2019b). Analysis and evaluation of the image preprocessing process of a six-band multispectral camera mounted on an unmanned aerial vehicle for winter wheat monitoring. *Sensors*, 19(3), 747. <https://doi.org/10.3390/s19030747>
- Jin, X., Zarco-Tejada, P., Schmidhalter, U., Reynolds, M. P., Hawkesford, M. J., Varshney, R. K., et al. (2020). High-throughput estimation of crop traits: A review of ground and aerial phenotyping platforms. *IEEE Geoscience and Remote Sensing Magazine*, 9(1), 200–231. <https://doi.org/10.1109/MGRS.2020.2998816>
- Johansen, K., Morton, M. J. L., Malbeteau, Y. M., Aragon, B., Al-Mashharawi, S. K., Ziliani, M. G., et al. (2019). Unmanned aerial vehicle-based phenotyping using morphometric and spectral analysis can quantify responses of wild tomato plants to salinity stress. *Frontiers in Plant Science*, 10, 370. <https://doi.org/10.3389/fpls.2019.00370>
- Johansen, K., Morton, M. J. L., Malbeteau, Y., Aragon, B., Al-Mashharawi, S., Ziliani, M. G., et al. (2020). Predicting biomass and yield in a tomato phenotyping experiment using UAV imagery and random forest. *Frontiers in Artificial Intelligence*. <https://doi.org/10.3389/frai.2020.00028>
- Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, 50(4), 663–666. <https://doi.org/10.2307/1936256>
- Justice, C. O., Vermote, E., Townshend, J. R., Defries, R., Roy, D. P., Hall, D. K., et al. (1998). The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4), 1228–1249. <https://doi.org/10.1109/36.701075>
- Kanke, Y., Raun, W., Solie, J., Stone, M., & Taylor, R. (2012). Red edge as a potential index for detecting differences in plant nitrogen status in winter wheat. *Journal of Plant Nutrition*, 35(10), 1526–1541. <https://doi.org/10.1080/01904167.2012.689912>
- Kross, A., McNairn, H., Lapen, D., Sunohara, M., & Champagne, C. (2015). Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *International Journal of Applied Earth Observation and Geoinformation*, 34, 235–248. <https://doi.org/10.1016/j.jag.2014.08.002>
- Le Maire, G., François, C., Soudani, K., Berveiller, D., Pontailier, J.-Y., Bréda, N., et al. (2008). Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sensing of Environment*, 112(10), 3846–3864. <https://doi.org/10.1016/j.rse.2008.06.005>
- Lee, H., Wang, J., & Leblon, B. (2020). Using linear regression, random forests, and support vector machine with unmanned aerial vehicle multispectral images to predict canopy nitrogen weight in corn. *Remote Sensing*, 12(13), 2071. <https://doi.org/10.3390/rs12132071>

- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18–22.
- Lillesand, T., Kiefer, R. W., & Chipman, J. (2015). *Remote sensing and image interpretation*. New Jersey, USA: Wiley.
- Maimaitijiang, M., Ghulam, A., Sidike, P., Hartling, S., Maimaitiyiming, M., Peterson, K., et al. (2017). Unmanned Aerial System (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134, 43–58. <https://doi.org/10.1016/j.isprsjprs.2017.10.011>
- Malbeteau, Y., Stutsel, B., Morton, M., Aragon, B., Angel, Y., Johansen, K., et al. (2021). Overcoming the challenges of thermal infrared orthomosaics using a swath-based approach to correct for dynamic temperature and wind effects. *Remote Sensing*, 13(16), 3255.
- Manfreda, S., McCabe, M. F., Miller, P. E., Lucas, R., Pajuelo Madrigal, V., Mallinis, G., et al. (2018). On the use of unmanned aerial systems for environmental monitoring. *Remote Sensing*, 10(4), 641. <https://doi.org/10.3390/rs10040641>
- Massawe, F., Mayes, S., & Cheng, A. (2016). Crop diversity: An unexploited treasure trove for food security. *Trends in Plant Science*, 21(5), 365–368. <https://doi.org/10.1016/j.tplants.2016.02.006>
- Maulik, U., & Chakraborty, D. (2017). Remote Sensing Image Classification: A survey of support-vector-machine-based advanced techniques. *IEEE Geoscience and Remote Sensing Magazine*, 5(1), 33–52. <https://doi.org/10.1109/MGRS.2016.2641240>
- Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 39(9), 2784–2817. <https://doi.org/10.1080/01431161.2018.1433343>
- McCabe, M. F., & Tester, M. (2021). Digital insights: Bridging the phenotype-to-genotype divide. *Journal of Experimental Botany*, 72(8), 2807–2810. <https://doi.org/10.1093/jxb/erab108>
- Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., & Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum*, 106(1), 135–141. <https://doi.org/10.1034/j.1399-3054.1999.106119.x>
- Negrão, S., Schmöckel, S., & Tester, M. (2017). Evaluating physiological responses of plants to salinity stress. *Annals of Botany*, 119(1), 1–11. <https://doi.org/10.1093/aob/mcw191>
- Peñuelas, J., Gamon, J. A., Fredeen, A. L., Merino, J., & Field, C. B. (1994). Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sensing of Environment*, 48(2), 135–146. [https://doi.org/10.1016/0034-4257\(94\)90136-8](https://doi.org/10.1016/0034-4257(94)90136-8)
- Potgieter, A. B., George-Jaeggli, B., Chapman, S. C., Laws, K., Suárez Cadavid, L. A., Wixted, J., et al. (2017). Multi-spectral imaging from an unmanned aerial vehicle enables the assessment of seasonal leaf area dynamics of sorghum breeding lines. *Frontiers in Plant Science*, 8, 1532. <https://doi.org/10.3389/fpls.2017.01532>
- Roli, F., & Fumera, G. (2001). Support vector machines for remote sensing image classification. In *Image and signal processing for remote sensing VI*. Barcelona, Spain: International Society for Optics and Photonics, pp 160–166. <https://doi.org/10.1117/12.413892>
- Roosjen, P. P., Brede, B., Suomalainen, J. M., Bartholomeus, H. M., Kooistra, L., & Clevers, J. G. (2018). Improved estimation of leaf area index and leaf chlorophyll content of a potato crop using multi-angle spectral data—potential of unmanned aerial vehicle imagery. *International Journal of Applied Earth Observation and Geoinformation*, 66, 14–26. <https://doi.org/10.1016/j.jag.2017.10.012>
- Sankaran, S., Espinoza, C. Z., Hinojosa, L., Ma, X., & Murphy, K. (2019). High-throughput field phenotyping to assess irrigation treatment effects in Quinoa. *Agrosystems, Geosciences & Environment*, 2(1), 1–7. <https://doi.org/10.2134/age2018.12.0063>
- Schell, J., & Deering, D. (1973). NDVI. *NASA Special Publication*, 351, 309.
- Shah, S. H., Angel, Y., Houborg, R., Ali, S., & McCabe, M. F. (2019). A random forest machine learning approach for the retrieval of leaf chlorophyll content in wheat. *Remote Sensing*, 11(8), 920. <https://doi.org/10.3390/rs11080920>
- Sims, D. A., & Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, 81(2–3), 337–354. [https://doi.org/10.1016/S0034-4257\(02\)00010-X](https://doi.org/10.1016/S0034-4257(02)00010-X)
- Singhal, G., Bansod, B., Mathew, L., Goswami, J., Choudhury, B., & Raju, P. (2019). Chlorophyll estimation using multi-spectral unmanned aerial system based on machine learning techniques. *Remote Sensing Applications: Society and Environment*, 15, 100235. <https://doi.org/10.1016/j.rsase.2019.100235>
- Stutsel, B., Johansen, K., Malbeteau, U. M., & McCabe, M. F. (2021). Detecting plant stress using thermal and optical imagery from an unoccupied aerial vehicle. *Frontiers in Plant Science*. <https://doi.org/10.3389/fpls.2021.734944>

- Sun, Q., Jiao, Q., Qian, X., Liu, L., Liu, X., & Dai, H. (2021). Improving the retrieval of crop canopy chlorophyll content using vegetation index combinations. *Remote Sensing*, 13(3), 470. <https://doi.org/10.3390/rs13030470>
- Sun, Y., Qin, Q., Ren, H., Zhang, T., & Chen, S. (2019). Red-edge band vegetation indices for leaf area index estimation from sentinel-2/msi imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 58(2), 826–840. <https://doi.org/10.1109/TGRS.2019.2940826>
- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2), 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0)
- Vega-Gálvez, A., Miranda, M., Vergara, J., Uribe, E., Puente, L., & Martínez, E. A. (2010). Nutrition facts and functional potential of quinoa (*Chenopodium quinoa* willd.), an ancient Andean grain: a review. *Journal of the Science of Food and Agriculture*, 90(15), 2541–2547. <https://doi.org/10.1002/jsfa.4158>
- 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(3), 195–203. [https://doi.org/10.1016/S1672-6308\(07\)60027-4](https://doi.org/10.1016/S1672-6308(07)60027-4)
- Woebbecke, D. M., Meyer, G. E., Von Bargen, K., & Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE*, 38(1), 259–269. <https://doi.org/10.13031/2013.27838>
- Xiang, T.-Z., Xia, G.-S., & Zhang, L. (2019). Mini-unmanned aerial vehicle-based remote sensing: Techniques, applications, and prospects. *IEEE Geoscience and Remote Sensing Magazine*, 7(3), 29–63. <https://doi.org/10.1109/mgrs.2019.2918840>
- Xie, Q., Dash, J., Huang, W., Peng, D., Qin, Q., Mortimer, H., et al. (2018). Vegetation indices combining the red and red-edge spectral information for leaf area index retrieval. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(5), 1482–1493. <https://doi.org/10.1109/JSTARS.2018.2813281>
- Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*. <https://doi.org/10.1155/2017/1353691>
- Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., et al. (2017). Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives. *Frontiers in Plant Science*, 8, 1111. <https://doi.org/10.3389/fpls.2017.01111>
- Yao, X., Wang, N., Liu, Y., Cheng, T., Tian, Y., Chen, Q., et al. (2017). Estimation of wheat LAI at middle to high levels using unmanned aerial vehicle narrowband multispectral imagery. *Remote Sensing*, 9(12), 1304. <https://doi.org/10.3390/rs9121304>
- Yuan, H., Yang, G., Li, C., Wang, Y., Liu, J., Yu, H., Feng, H., Xu, B., Zhao, X., & Yang, X. (2017). Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: Analysis of RF, ANN, and SVM regression models. *Remote Sensing*, 9, 309. <https://doi.org/10.3390/rs9040309>
- Zarco-Tejada, P. J., Hornero, A., Beck, P., Kattenborn, T., Kempeneers, P., & Hernández-Clemente, R. (2019). Chlorophyll content estimation in an open-canopy conifer forest with Sentinel-2A and hyperspectral imagery in the context of forest decline. *Remote Sensing of Environment*, 223, 320–335. <https://doi.org/10.1016/j.rse.2019.01.031>
- Zha, H., Miao, Y., Wang, T., Li, Y., Zhang, J., Sun, W., et al. (2020). Improving unmanned aerial vehicle remote sensing-based rice nitrogen nutrition index prediction with machine learning. *Remote Sensing*. <https://doi.org/10.3390/rs12020215>
- Zheng, H., Cheng, T., Li, D., Zhou, X., Yao, X., Tian, Y., et al. (2018a). Evaluation of RGB, color-infrared and multispectral images acquired from unmanned aerial systems for the estimation of nitrogen accumulation in rice. *Remote Sensing*. <https://doi.org/10.3390/rs10060824>
- Zheng, H., Li, W., Jiang, J., Liu, Y., Tao, C., Tian, Y., et al. (2018b). A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle. *Remote Sensing*. <https://doi.org/10.3390/rs10122026>
- Zheng, H., Cheng, T., Zhou, M., Li, D., Yao, X., Tian, Y., et al. (2019). Improved estimation of rice above-ground biomass combining textural and spectral analysis of UAV imagery. *Precision Agriculture*, 20, 611–629. <https://doi.org/10.1007/s11119-018-9600-7>
- Zurita-Silva, A., Fuentes, F., Zamora, P., Jacobsen, S.-E., & Schwember, A. R. (2014). Breeding quinoa (*Chenopodium quinoa* Willd.): Potential and perspectives. *Molecular Breeding*, 34(1), 13–30. <https://doi.org/10.1007/s11032-014-0023-5>

Authors and Affiliations

Jiale Jiang¹ · **Kasper Johansen¹** · **Clara S. Stanschewski²** · **Gordon Wellman²** · **Magdi A. A. Mousa^{3,4}** · **Gabriele M. Fiene²** · **Khalid A. Asiry³** · **Mark Tester²** · **Matthew F. McCabe¹**

¹ Hydrology, Agricultural and Land Observation, Water Desalination and Reuse Center, Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

² Center for Desert Agriculture and Division of Biological & Environmental Sciences & Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

³ Department of Arid Land Agriculture, Faculty of Meteorology, Environment & Arid Land Agriculture, King Abdulaziz University, Jeddah 80208, Saudi Arabia

⁴ Department of Vegetable Crops, Faculty of Agriculture, Assiut University, Assiut 71526, Egypt