UAV REMOTE SENSING FOR PLANT TRAITS AND STRESS

EDITED BY: Alessandro Matese, Wenting Han, Peng Fu, Huihui Zhang, Jianfeng Zhou and Frédéric Cointault

PUBLISHED IN: Frontiers in Plant Science

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-333-7
DOI 10.3389/978-2-88976-333-7

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

UAV REMOTE SENSING FOR PLANT TRAITS AND STRESS

Topic Editors:
Alessandro Matese, National Research Council (CNR), Italy
Wenting Han, Northwest A&F University, China
Peng Fu, University of Illinois at Urbana-Champaign, United States
Huihui Zhang, United States Department of Agriculture, United States
Jianfeng Zhou, University of Missouri, United States
Frédéric Cointault, Agrosup Dijon, France

Citation: Matese, A., Han, W., Fu, P., Zhang, H., Zhou, J., Cointault, F., eds. (2022). UAV Remote Sensing for Plant Traits and Stress. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-333-7

Table of Contents

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 February 2021
doi: 10.3389/fpls.2021.640259

Crop Performance Evaluation of Chickpea and Dry Pea Breeding Lines Across Seasons and Locations Using Phenomics Data

*Chongyuan Zhang[1], Rebecca J. McGee[2], George J. Vandemark[2] and Sindhuja Sankaran[1]**

[1] Department of Biological System Engineering, Washington State University, Pullman, WA, United States, [2] USDA-ARS, Grain Legume Genetics and Physiology Research, Washington State University, Pullman, WA, United States

OPEN ACCESS

Edited by:
Jianfeng Zhou,
University of Missouri, United States

Reviewed by:
Abbas Atefi,
University of Nebraska-Lincoln, United States
Xu Wang,
Kansas State University, United States

***Correspondence:**
Sindhuja Sankaran
sindhuja.sankaran@wsu.edu

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

Received: 10 December 2020
Accepted: 11 January 2021
Published: 25 February 2021

Citation:
Zhang C, McGee RJ, Vandemark GJ and Sankaran S (2021) Crop Performance Evaluation of Chickpea and Dry Pea Breeding Lines Across Seasons and Locations Using Phenomics Data. Front. Plant Sci. 12:640259. doi: 10.3389/fpls.2021.640259

The Pacific Northwest is an important pulse production region in the United States. Currently, pulse crop (chickpea, lentil, and dry pea) breeders rely on traditional phenotyping approaches to collect performance and agronomic data to support decision making. Traditional phenotyping poses constraints on data availability (e.g., number of locations and frequency of data acquisition) and throughput. In this study, phenomics technologies were applied to evaluate the performance and agronomic traits in two pulse (chickpea and dry pea) breeding programs using data acquired over multiple seasons and locations. An unmanned aerial vehicle-based multispectral imaging system was employed to acquire image data of chickpea and dry pea advanced yield trials from three locations during 2017–2019. The images were analyzed semi-automatically with custom image processing algorithm and features were extracted, such as canopy area and summary statistics associated with vegetation indices. The study demonstrated significant correlations ($P < 0.05$) between image-based features (e.g., canopy area and sum normalized difference vegetation index) with yield (r up to 0.93 and 0.85 for chickpea and dry pea, respectively), days to 50% flowering (r up to 0.76 and 0.85, respectively), and days to physiological maturity (r up to 0.58 and 0.84, respectively). Using image-based features as predictors, seed yield was estimated using least absolute shrinkage and selection operator regression models, during which, coefficients of determination as high as 0.91 and 0.80 during model testing for chickpea and dry pea, respectively, were achieved. The study demonstrated the feasibility to monitor agronomic traits and predict seed yield in chickpea and dry pea breeding trials across multiple locations and seasons using phenomics tools. Phenomics technologies can assist plant breeders to evaluate the performance of breeding materials more efficiently and accelerate breeding programs.

Keywords: image processing, multispectral imagery, unmanned aircraft vehicle, vegetation indices, yield prediction

INTRODUCTION

Crop cultivars are consistently selected based on their productivity (quantity and/or quality), tolerance to biotic and abiotic stressors, and adaptation to local production systems and environments (Acquaah, 2009; Hatfield and Walthall, 2015). Pulse crops, including pea (*Pisum sativum* L.) and chickpea (*Cicer arietinum* L.), have been bred for their adaptation to the Palouse region in the Pacific Northwest, United States, with the overall goal of developing high-yielding and biotic and abiotic stress-resistant cultivars. The Palouse region, which includes parts of eastern Washington, northern Idaho, and northeastern Oregon, is one of the largest producers of pulse crops in the United States (USDA-NASS, 2020) and is home to several pulse breeding programs. Pulse breeders have developed and released multiple pea and chickpea cultivars with better seed yield, quality, and improved disease resistance (McGee and McPhee, 2012; McGee et al., 2012, 2013; Vandemark et al., 2014, 2015; USDA-ARS, 2018). However, plant breeders have primarily relied on traditional methods to collect phenotypic data on breeding lines. Some of the constraints of these traditional phenotyping approaches are that they are labor-intensive, time-consuming, and subjective with limited availability of data. Therefore, sensing technologies, also referred to as phenomics technologies, are needed to overcome these constraints to facilitate progress of plant breeding and provide data for a more accurate and comprehensive evaluation of breeding lines.

Plant traits evaluated by phenomics technologies in field conditions include early vigor (Kipp et al., 2014; Sankaran et al., 2015), canopy area and temperature (Patrignani and Ochsner, 2015; Bai et al., 2016), plant height (Madec et al., 2017; Wang et al., 2018), heading and flower intensity (Sadeghi-Tehran et al., 2017; Zhang et al., 2020), yield (Donohue et al., 2018; Lai et al., 2018), and phenological stages (Yang et al., 2017). Research using phenomics technologies to monitor or predict crop yield has been conducted for a wide range of crops. Different image-based plant traits, such as flowers, vegetation indices (VIs), plant height, and canopy area (Bai et al., 2016; Tattaris et al., 2016; Thorp et al., 2016; Sun et al., 2018), have been used to monitor and predict crop yield. Thorp et al. (2016) used proximal digital imaging to track *Lesquerella* flowering dynamics and reported that there was a strong correlation between flower cover percentage and seed yield (coefficient of determination or $R^2 \leq 0.84$). Sun et al. (2018) developed a terrestrial light detection and ranging (LiDAR)-based high-throughput phenotyping system and applied it to monitor cotton growth. Their results indicated that canopy height, projected canopy area, and plant volume ($R^2 \leq 0.84$, 0.88, and 0.85, respectively) at 67 and 109 days after planting were positively correlated with yield. In addition to correlating plant traits with yield, researchers have tested models to predict seed yield and biomass of wheat, canola, and corn (Fieuzal et al., 2017; Ballesteros et al., 2018; Donohue et al., 2018; Lai et al., 2018; Anderson et al., 2019). Fieuzal et al. (2017) developed two artificial neural network-based methods (a real-time approach and a diagnostic approach) to estimate corn yield using multi-temporal optical and radar satellite data. The diagnostic approach using reflectance from the red spectral region predicted yield with $R^2 = 0.77$, while the real-time approach using reflectance from the red spectral region and one feature from radar satellite data resulted in a prediction accuracy of $R^2 = 0.69$.

Other performance traits have also been evaluated using sensing technologies, including estimation of phenological stages, 50% flowering, senescence, and maturity (Viña et al., 2004; Yu et al., 2016; Zheng et al., 2016; Yang et al., 2017; Quirós Vargas et al., 2019; Lindsey et al., 2020). Zheng et al. (2016) monitored rice phenology in three growing seasons using a time series of spectral indices obtained using portable spectrometers. They reported that the red-edge chlorophyll index can accurately detect the dates of jointing, middle booting, and soft dough, while the normalized difference vegetation index (NDVI) can detect dates of active tillering, middle heading, and maturity. In our previous study (Quirós Vargas et al., 2019), we found that VIs, including NDVI, green red vegetation index (GRVI), and the normalized difference red-edge index (NDRE), were correlated with days to 50% flowering and physiological maturity in two winter pea experiments.

Although phenomics technologies have been tested on many crops, the evaluation of such technologies across field seasons, locations, and different crop types has been limited for pulse crops. Such efforts are essential to assess the stability and applicability of phenomics technologies to assist breeding programs. Therefore, in this study, we applied sensing technologies to evaluate dry pea and chickpea breeding lines for three growing seasons (2017–2019) for phenotyping applications. Specific objectives were to: 1) monitor yield and other agronomic traits using quadcopter unmanned aircraft vehicle (UAV) multispectral imaging data and 2) predict pulse crop yield with a multivariate regression model.

MATERIALS AND METHODS

Experimental Locations and Plant Materials

The pulse crop (chickpea and dry pea) breeding lines in this study (2017–2019) were evaluated in multiple field locations, near Pullman, WA (46°41′39.0″N, 117°08′53.0″W), Fairfield, WA (47°19′08.0″N, 117°10′05.0″W), and Genesee, ID (46°36′40.0″N, 116°57′39.0"W), United States (**Table 1**). The exact locations of the experiment field sites within an area varied between years due to crop rotation protocols. Advanced yield trials of green pea (panel 01), yellow pea (panel 02), and chickpea (panel 81) breeding lines and relevant commercial check cultivars were planted using a randomized complete block design with three replicates. A seed treatment was applied prior to planting that contained the fungicides fludioxonil (0.56 g kg^{-1}; Syngenta, Greensboro, NC, United States), mefenoxam (0.38 g kg^{-1}; Syngenta), and thiabendazole (1.87 g kg^{-1}; Syngenta), thiamethoxam (0.66 ml kg^{-1}; Syngenta) for insect control, and molybdenum (0.35 g kg^{-1}). Approximately 0.5 g *Mesorhizobium ciceri* inoculant (1×10^8 CFU g^{-1}; Novozyme, Cambridge, MA, United States) was applied to each chickpea seed packet 1 day before planting. Chickpea plots were planted at 6.1 m

TABLE 1 | Summary of the pulse crops' breeding trials and data acquisition using sensing.

Year	Location	Sensing altitude[a]	Crop	Panel[b]	Number of lines/cultivars	Sowing date	Data acquisition dates	Growing degree days[c]
2017[d]	Pullman	45 m	Chickpea	1781	24	5/11	6/26, 7/07, 7/21, 7/28	536, 723, 959, 1087
			Dry pea	1701, 1702	40, 21	5/11	6/26, 7/07, 7/21	536, 723, 959
	Fairfield	45 m	Dry pea	1701, 1702	40, 21	5/11	7/24[e]	1077
2018	Pullman	25 m	Chickpea	1881	21	5/05	6/08, 6/22, 7/03, 7/19, 7/27	423, 582, 721, 991, 1124
			Dry pea	1801, 1802	32, 23	5/05	6/08, 6/22, 7/03, 7/19, 7/27	423, 582, 721, 991, 1124
	Genesee	25 m	Chickpea	1881	21	5/08	6/08, 6/26, 7/09, 7/23, 8/06	372, 576, 752, 997, 1274
			Dry pea	1801	32	5/08	6/08, 6/26, 7/09, 7/23, 8/06	372, 576, 752, 997, 1274
	Fairfield	25 m	Chickpea	1881	21	5/21	6/12, 6/29, 7/12, 7/25	282, 515, 716, 966
			Dry pea	1801, 1802	32, 23	5/21	6/12, 6/29, 7/12, 7/25, 8/07	282, 515, 716, 966, 1274
2019	Pullman	30 m	Chickpea	1981	24	5/04	6/05, 6/17, 7/05, 7/16	387, 549, 772, 939
			Dry pea	1901, 1902	29, 23	5/04	6/05, 6/17, 7/05, 7/16	387, 549, 772, 939
	Genesee	25 m	Chickpea	1981	24	5/03	6/05, 6/18, 7/05, 7/16	384, 559, 756, 923
			Dry pea	1901	29	5/03	6/05, 6/18, 7/05, 7/16	384, 559, 756, 923
	Fairfield	30 m	Chickpea	1981	24	5/06	6/10, 6/28, 7/12, 7/23	425, 679, 902, 1080
			Dry pea	1901, 1902	29, 23	5/06	6/10, 6/28, 7/12	425, 679, 902

[a] *Sensing altitudes (25, 30, and 45 m) as above ground level used in this study resulting in 1.7, 2.0, and 3.1 cm per pixel of ground sampling distance.*
[b] *Panels ending in "01" are advanced green pea, "02" are advanced yellow pea, and "81" are advanced chickpea.*
[c] *Growing degree days: accumulated degree days that are used to estimate temperature-based growing season; degree day = mean temperature - base temperature (base temperature = 3°C in this study) (Bourgeois et al., 2000; Miller et al., 2018).*
[d] *No data were available due to failed trials in Genesee, ID.*
[e] *Data from other time points were not useful (e.g., senesced plants).*

long and 1.5 m wide with approximately 75 cm between plots. After emergence, the plots were trimmed to approximately 4.9 m long, thus leaving alleys of approximately 1.2 m between ranges. Chickpeas were planted at a density of 43 seeds m^{-2} in a 1.5-m × 6.1 m block (≈430,000 seeds ha^{-1}). The chickpea entries had either compound or simple leaf types. Typically, there were four to seven entries each year with simple leaves. The dry pea entries had either normal (cv. 'Columbian' only) leaves or were semi-leafless. Data analysis was conducted without separating the leaf types for each crop, and preliminary analysis indicated that the ranges of the vegetation indices were similar.

Data Acquisition

A quadcopter UAV (AgBot, ATI Inc., Oregon City, OR, United States) and a five-band multispectral camera (RedEdge, MicaSense Inc., Seattle, WA, United States) were deployed to acquire image data during the 3-year study (**Figure 1a**). The multispectral camera mounted on the quadcopter acquired images (12-bit image) with a resolution of 1.2 MP in the spectrum of red (R, 668 ± 5 nm, central band and band width), green (G, 560 ± 10 nm), blue (B, 475 ± 10 nm), near-infrared (NIR, 840 ± 20 nm), and red edge (RE, 717 ± 5 nm). Mission Planner, an open-source ground control station software (ArduPilot Development Team and Community), was used to plan and monitor missions of aerial data acquisition (**Figure 1b**). The UAV was programmed with Mission Planner to fly at a speed of 2–3 m/s and at 25, 30, or 45 m above ground level (AGL), resulting in a ground sampling distance (GSD) of 1.7, 2.0, or 3.1 cm per pixel, respectively, and to acquire images with 80% horizontal and 70% vertical overlaps. A reflectance panel, either a MicaSense reflectance panel (RedEdge, MicaSense Inc.) in 2017 or a Spectralon reflectance panel (99% reflectance; Spectralon, SRS-99-120, Labsphere Inc., North Sutton, NH, United States) in 2018 and 2019, was placed in the field during image acquisition and used for radiometric calibration during image processing. Data were acquired between 10:00 a.m. and 3:00 p.m. local time, and three to five time points of data acquisition were achieved for each season (**Table 1**). The time points for data acquisition were selected to acquire data representing key growth stages, such as early growth, flowering, and seed/pod development stages, and based on suitable weather conditions for UAV flights (e.g., clear sky and low wind). Seed yield data from the dry pea and chickpea trials were collected from each location, while other agronomic and phenological traits were collected only at Pullman each year.

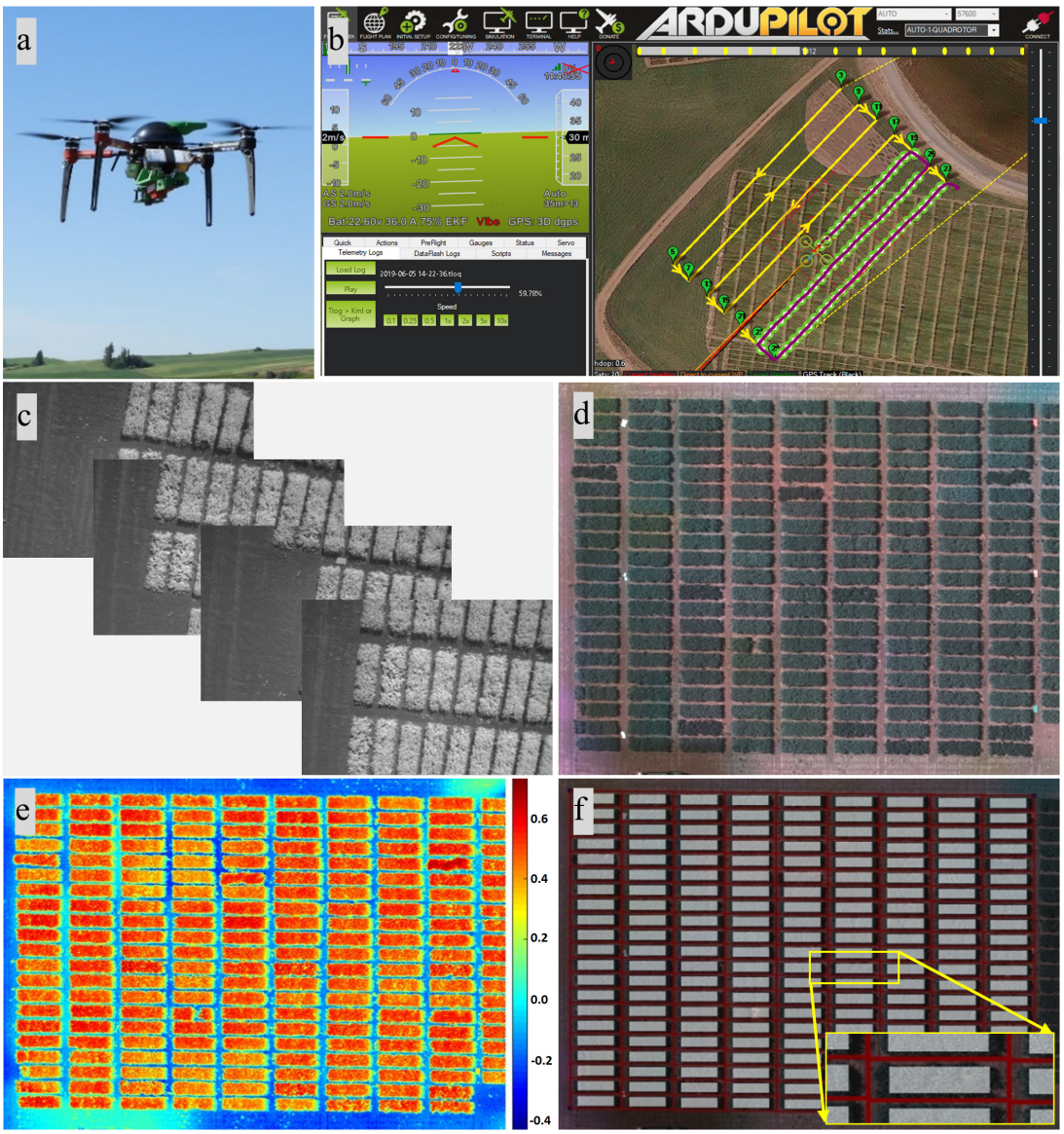

FIGURE 1 | Unmanned aircraft vehicle (UAV)-based data acquisition and image processing. **(a)** UAV and camera. **(b)** Mission Planner showing data acquisition underway. **(c)** Individual images. **(d)** Orthomosaic composite images consisting of red, green, and blue bands. **(e)** Heat map of the soil-adjusted vegetation index. **(f)** Resulting image with the plot separated and the region of interest highlighted.

These traits include days to 50% flowering, days to physiological maturity, pod height, pod height at maturity, overall vine length, canopy height at maturity, node of first flower, and number of reproductive nodes.

Image Processing and Feature Extraction

Images from the multispectral camera (**Figure 1c**) were first processed in Pix4Dmapper (Pix4D Inc., San Francisco, CA, United States) to generate orthomosaic images covering each experimental site. The template used in Pix4Dmapper was based on Ag Multispectral, where the calibration method of "Alternative" was selected in the initial processing. In this type of calibration, the images are optimized for aerial nadir images with accurate geolocation information, low texture content, and relatively flat terrain. Orthomosaic images were imported into custom semi-automated image processing algorithms developed in MATLAB (2018b; MathWorks Inc., Natick, MA, United States) for further processing. The image processing algorithms prompted the user to input a degree to rotate the image, which is prepared for plot segmentation later, and to identify the reflectance panel for radiometric calibration. Following that, composite RGB image and several vegetation index maps were generated using different combinations of orthomosaic images (**Figures 1d,e, 2**). A composite RGB

image was generated for quality inspection. The vegetation indices calculated included normalized difference vegetation index (NDVI), green NDVI (GNDVI), soil-adjusted vegetation index (SAVI), normalized difference red-edge index (NDRE), and triangular vegetation index (TVI) (Harris Geospatial Solutions, 2020). For each dataset (each crop and each data acquisition period), a master mask that separated the canopy from the background, such as soil and flowers (for dry pea only), was generated by setting a threshold on the SAVI index map prior to feature extraction. Threshold data varied between datasets based on canopy vigor and illumination conditions at the time of data acquisition, and the value was selected for each dataset based on visual assessment.

In dry pea, flowers that are white have a different reflectance than stipules and tendrils and were excluded from canopy feature extraction. A similar procedure was not applied to chickpea as chickpea flowers could not be detected in five-band multispectral images at the given resolution due to the small flower size (Zhang et al., 2020). The master mask was overlaid on the composite RGB image for quality inspection and optimization of the threshold for generating a canopy mask. The developed algorithm prompted the user to identify the four corners of the field and automatically separated individual plots with information of the identified corners (**Figure 1f**). Masks for individual plots were then shrunk at the four edges to prevent edge effects, forming regions of interest for each plot that were highlighted with white, as shown in **Figure 1f**. The top and bottom edges of the mask for an individual plot were reduced by 11 (2017) or 20 (2018 and 2019) pixels, while the right and left edges were reduced by 28 (2017) or 50 (2018 and 2019) pixels. More details about the algorithm can be found in Zhang et al. (2019). Image-based features were extracted from regions of interest in each plot, including canopy area (in pixels), and the mean and sum statistic of NDVI, GNDVI, SAVI, NDRE, and TVI plot images. Here, the mean of NDVI, for example, is the average of the NDVI values of the canopy pixels identified by the mask for an individual plot, while the sum of NDVI is the sum of the NDVI values of canopy pixels. At the end of image processing, the algorithms exported the features as Excel files for further analysis. The procedures of image processing for dry pea and chickpea were similar with minor modifications, such as the threshold used to create the master mask.

Data Analysis

Image-based features from the UAV data were analyzed using Pearson's correlation in SAS, University Edition (SAS Institute, Cary, NC, United States). The features were correlated with yield for all locations and with other traits for the Pullman trial only due to availability of data. Plot-by-plot and cultivar-by-cultivar (by averaging the replicates at each field site) correlation analyses were also conducted. Noisy data (e.g., cloud-covered plots) were eliminated prior to analysis.

Yield prediction models were developed using image-based features as predictors to estimate yields in the chickpea (panel 81) and green pea (panel 01) trials. Due to the availability of data across the three locations, only green pea breeding lines were utilized for yield prediction. Yields were predicted using data from each year and each location (2017: Pullman; 2018 and 2019: Pullman, Genesee, and Fairfield) and the combined data for each year (2018 and 2019). The identity of the breeding lines varied from year to year as lines were discarded or added to the panels, and therefore data were only combined within a year. In 2017, no chickpea data were available from

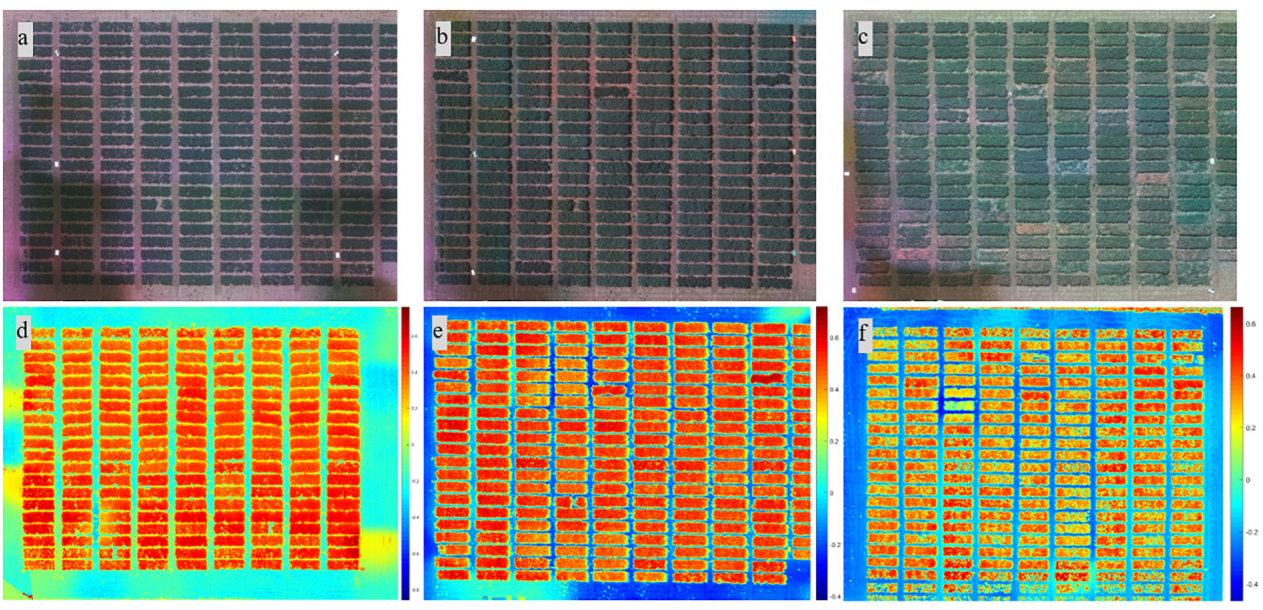

FIGURE 2 | Composite RGB images from different growth stages (2018) and heat map of the soil-adjusted vegetation index from different locations. Images from the **(a)** early, **(b)** flowering, and **(c)** pod/seed development stages. Heat map of the soil-adjusted vegetation index from **(d)** Genesee, **(e)** Pullman, and **(f)** Fairfield.

Genesee and Fairfield, and only one set of pea data was available from Fairfield.

Datasets acquired at flowering, and pod and seed development stages were utilized for yield model development. The data (yield and image features) were normalized using the following formula prior to the model development (**Figure 3**).

$$x_n = \frac{(x_r - m)}{\sigma} \quad (1)$$

where x_n, x_r, m, and σ stand for normalized data, raw data, mean of a feature, and standard deviation of a feature, respectively. Normalization was conducted for each year and each location as well as for the combined data within each year. Yield prediction was conducted using both plot-by-plot data and cultivar-by-cultivar data (average of three replicate data across each field site). Least absolute shrinkage and selection operator (LASSO) regression programmed in MATLAB was used in this study for yield prediction. The parameters used in LASSO included: alpha (weight of lasso *versus* ridge optimization) = 1, MCReps (repetitions of cross-validation) = 3, cross-validation = five-fold, and predictor selection method (for cross-validation) = IndexMinMSE. More details about LASSO in MATLAB can be found at the MathWorks website[1].

For the plot-by-plot analysis, the dataset was divided into the training and the testing datasets with a ratio of 3:1. The training data were further resampled five times (80% of training data to calibrate and 20% of training data to validate) to optimize the models. Finally, the model was evaluated using the test dataset and the process was assessed four times (four iterations) to eliminate effects of randomization. For the cultivar-by-cultivar analysis, due to the limited sample size, the leave-one-out approach for model development and evaluation was utilized (Sammut and Webb, 2010). During model development, a five-fold cross-validation was used, followed by testing the model for as many times as the instances (29–40 lines/cultivars depending on the dataset). The prediction performance was reported in terms of R^2 during the train and test processes and selected image-based features.

[1]https://www.mathworks.com/help/stats/lasso.html

RESULTS

Relationship Between Image Features and Yield

In general, there were significant and positive correlations ($P < 0.05$, r up to 0.74) between image-based features (e.g., canopy area, SAVI, and sum NDVI) and yield with the plot-by-plot chickpea data acquired at the early growth, flowering, and pod/seed development stages across field seasons (2017–2019) and locations (**Figure 4A** and **Supplementary Figure 1**). Chickpea and dry pea flowered between 721 and 772 growing degree days. Plants were considered in the early growth stages before flowering and in the pod/seed development stage between the flowering stage and physiological maturity. Only a few common image features (e.g., NDVI and SAVI) extracted from the data acquired at the early growth stages were significantly correlated with yield across seasons and experimental locations, while more common image features (e.g., canopy area, NDVI, SAVI, sum of NDVI, GNDVI, and SAVI) extracted from the data acquired at the flowering and pod/seed development stages were correlated with yield. A similar trend was found when analyzing the chickpea data using the cultivar-by-cultivar method (r up to 0.93). During the cultivar-by-cultivar data analysis, fewer features were significantly correlated with yield (**Figure 4B** and **Supplementary Figure 2**), which could be due to the smaller dataset compared to the plot-by-plot analysis method. In most cases, the correlations between the image features and yield were greater from data acquired at the flowering and pod/seed development stages in comparison to the early growth stages.

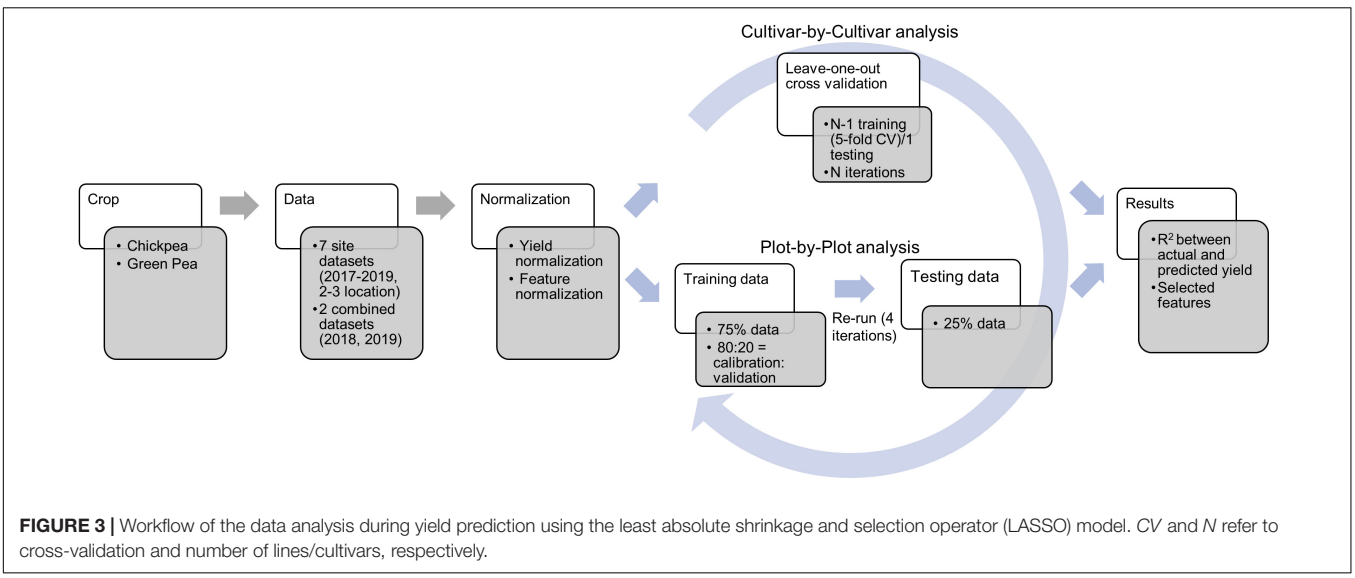

FIGURE 3 | Workflow of the data analysis during yield prediction using the least absolute shrinkage and selection operator (LASSO) model. *CV* and *N* refer to cross-validation and number of lines/cultivars, respectively.

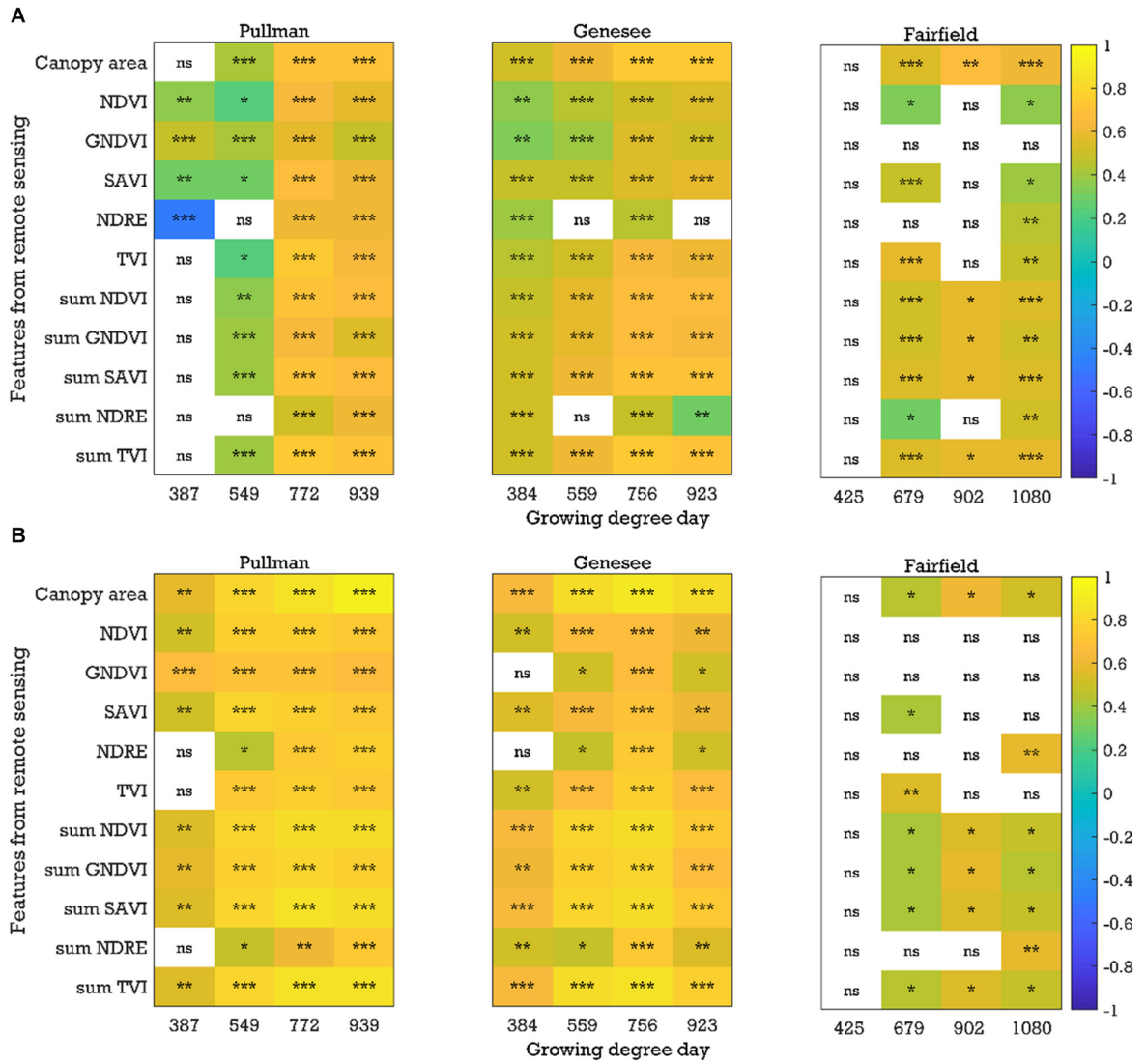

FIGURE 4 | Correlation coefficients between the image-based features and yield for the chickpea yield trial in 2019: **(A)** plot-by-plot analysis and **(B)** cultivar-by-cultivar analysis. *NDVI*, normalized difference vegetation index; *GNDVI*, green *NDVI*; *SAVI*, soil-adjusted vegetation index; *NDRE*, normalized difference red-edge index; *TVI*, triangular vegetation index; *NDVI*, for example, is the average of the NDVI values of canopy pixels, while *sum NDVI* is the sum of the NDVI values of canopy pixels; *ns*, nonsignificant at the 0.05 probability level. Significant probability levels: *0.05, **0.01, and ***0.001.

Similar to chickpea, in the green pea breeding lines, significant positive correlations ($P < 0.05$, r up to 0.83) between the image-based features (e.g., canopy area and sum NDVI) extracted from the plot-by-plot data acquired at the early growth, flowering, and pod/seed development stages and yield were observed across field seasons and experimental locations in most cases (**Figure 5A** and **Supplementary Figure 3**). When analyzing the cultivar-by-cultivar data, typically fewer image features within a time point were significantly correlated with yield, although the r values were up to 0.80 (**Supplementary Figure 4**). It was interesting to note that four of eight trials (field seasons × experimental locations) showed significant negative correlations between yield and image features from the data acquired at the early growth/pre-flowering stages using both analysis methods (plot-by-plot or cultivar-by-cultivar). The potential reason for the negative relationships between the image-based features (e.g., NDVI data) at the early growth stages and yield is unclear and requires further investigation. We have observed that pea cultivars that flower early typically have better early seedling vigor; however, they also have lower seed yields, presumably because the plants do not have an extended vegetative period during which they can produce as much photosynthate (and hence seeds) as later flowering cultivars. In peas, the timing of flowering is dictated by photoperiod response rather than by biomass accumulation. In general, phenotyping the pea trials was more challenging than for chickpea, which may be due to the presence of tendrils in the cultivars. The spectral reflectance of tendrils may be different

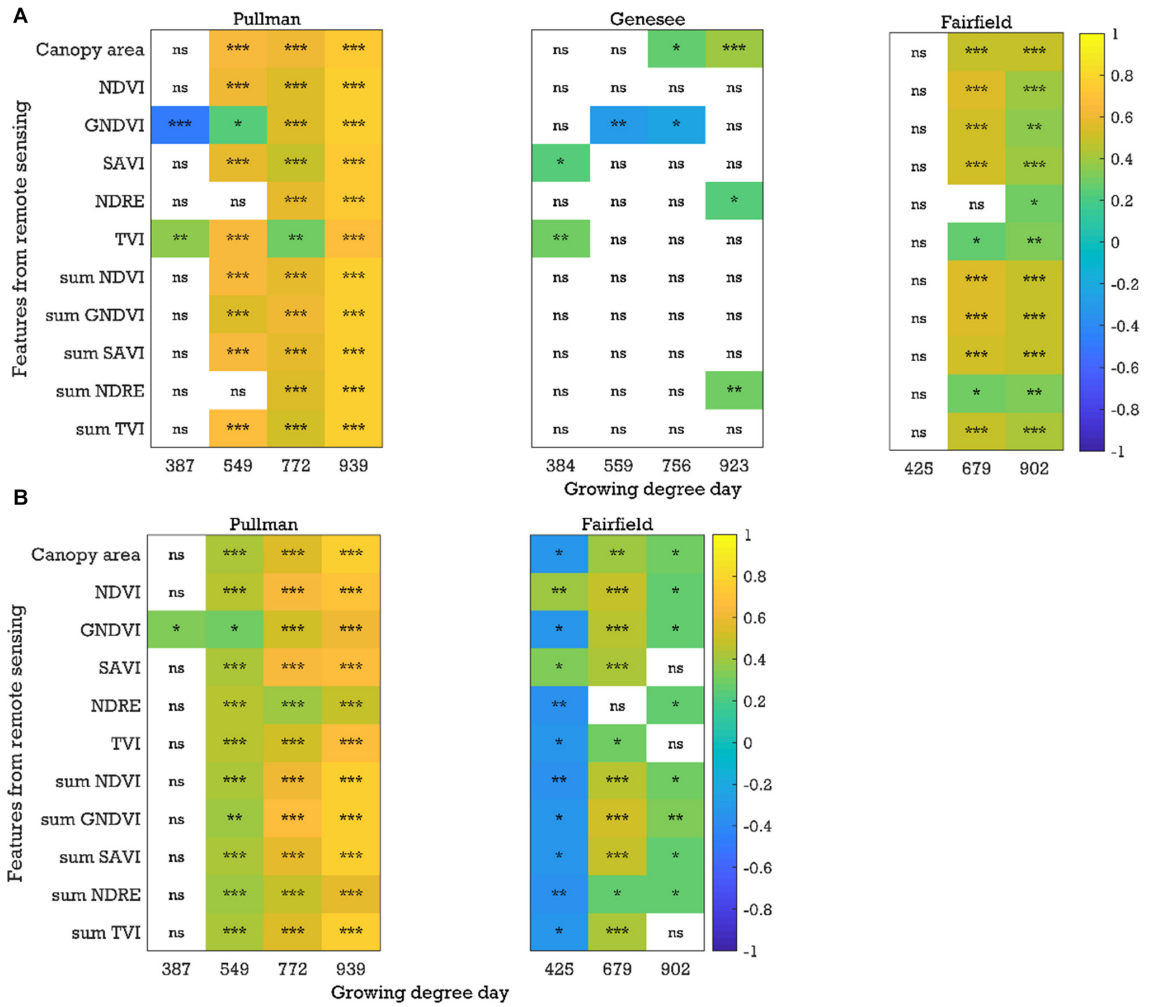

FIGURE 5 | Correlation coefficients between the image-based features and yield for the **(A)** green pea and **(B)** yellow pea trials in 2019 (plot-by-plot analysis). *NDVI*, normalized difference vegetation index; *GNDVI*, green NDVI; *SAVI*, soil-adjusted vegetation index; *NDRE*, normalized difference red-edge index; *TVI*, triangular vegetation index; *NDVI*, for example, is the average of the NDVI values of canopy pixels, while *sum NDVI* is the sum of the NDVI values of canopy pixels; *ns*, nonsignificant at the 0.05 probability level. Significant probability levels: *0.05, **0.01, and ***0.001.

from that of stipules. Similar patterns of significant correlations ($P < 0.05$, r up to 0.85) with data acquired at the early growth, flowering, and pod/seed development stages were found in the yellow pea yield trials (**Figure 5B** and **Supplementary Figures 5,6**). Significant negative correlations at the early stage were rare in the yellow pea yield trials, unlike the green pea yield trials.

Relationship Between Image Features and Other Data Types

Correlations between the image-based features (e.g., NDVI, SAVI, and sum SAVI) and other traits (e.g., days to 50% flowering, days to physiological maturity, plant height, pod length, etc.) acquired from the Pullman trials were analyzed across three field seasons. For the chickpea yield trials, significant ($P < 0.05$) and positive correlations between the image-based features and days to 50% flowering or days to physiological maturity (r up to 0.76 and 0.58, respectively) were found after the flowering stage (**Figure 6** and **Supplementary Figures 7,8**). Most of the negative correlations observed between the image-based features and days to 50% flowering or days to physiological maturity at the early growth stages were not significant. On the other hand, correlations between the image-based features and the remaining traits were not consistent across the three field seasons (data not presented).

With regard to the green pea yield trials, significant and negative correlations ($P < 0.05$, $r > -0.54$) between features (e.g., NDVI, SAVI, and sum SAVI) and days to 50% flowering were observed at the early stages (**Figure 6** and **Supplementary Figure 9**) across 3 years for both analysis methods in most cases. Negative correlations between features and days to physiological maturity were also observed in the early growth stages, although most correlations were not

significant (**Figure 6** and **Supplementary Figure 10**). Early plant vigor (higher vegetation index data) may be associated with early flowering and maturity (early flowering/maturity dates), which would result in negative correlations. As with chickpea, most image-based features acquired after flowering were significantly and positively correlated with days to 50% flowering or days to physiological maturity (r up to 0.75 and 0.72, respectively), especially when the images were acquired close to physiological maturity. No consistent trends in the correlations between the image-based features and the other traits were found. In the yellow pea yield trials, negative correlations between features and days to 50% flowering or days to physiological maturity were also observed in the early growth stages. However, significant positive correlations (r up to 0.85 and 0.84, respectively) between these two traits and most of the image-based features were found in the datasets acquired after flowering (**Supplementary Figures 11,12**). In the yellow pea yield trials, there were some image features (e.g., NDVI, SAVI, and sum SAVI) that were significantly correlated with pod height index (negative) and pod height (positive) across the three field seasons, especially in the pod/seed development stage (data not presented).

Yield Prediction Using LASSO Regression

Chickpea yield can be predicted using multiple image-based features, as summarized in **Table 2**. The prediction accuracy varied across field seasons and locations, regardless of the analysis method, with testing accuracy (for individual locations) of up to 0.84. When the data within a year were combined, the prediction accuracy increased in 2018 and 2019 (testing accuracy of up to 0.91; **Figure 7**). Regardless of whether the data were separated for individual locations or combined within a year, the features selected by LASSO as predictors varied from one to seven. Only features that were selected

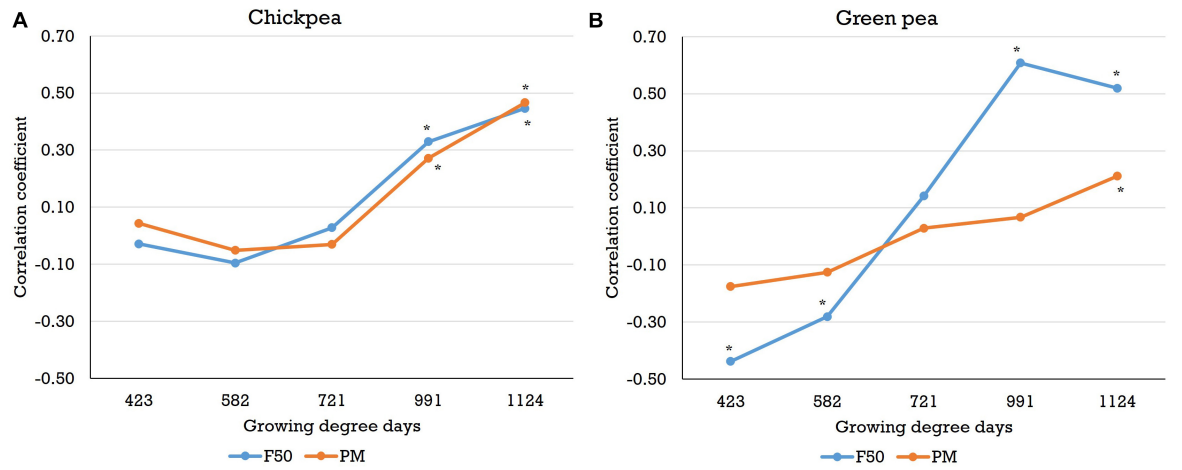

FIGURE 6 | Correlations between the sum normalized difference vegetation index and days to 50% flowering (*F50*) or days to physiological maturity (PM) for **(A)** chickpea and **(B)** green pea in the 2018 field season (plot-by-plot analysis). *Sum NDVI* is the sum of the NDVI values of canopy pixels. Correlations that are significant at the 0.05 probability level are indicated by an *asterisk*.

TABLE 2 | Yield prediction results of the models for chickpea crop.

Year	Location	Plot-by-plot method			Cultivar-by-cultivar method		
		Train R^2	Test R^2	Number of features[a]	Train R^2	Test R^2	Number of features[a]
2017	Pullman	0.55	0.32	4	0.61	0.11	2
2018	Pullman	0.63	0.53	3	0.77	0.42	6
	Genesee	0.45	0.33	3	0.52	0.31	3
	Fairfield	NA	NA	NA	NA	NA	NA
	Combined	0.90	0.89	5	0.93	0.91	4
2019	Pullman	0.74	0.57	3	0.86	0.84	1
	Genesee	0.67	0.51	6	0.83	0.75	2
	Fairfield	0.79	0.70	2	NA	NA	NA
	Combined	0.84	0.82	7	0.86	0.76	5

[a]*Only features with ≥ 75% selection occurrence during multiple iterations/runs during model development were considered.*
NA indicates failure to develop yield prediction model.

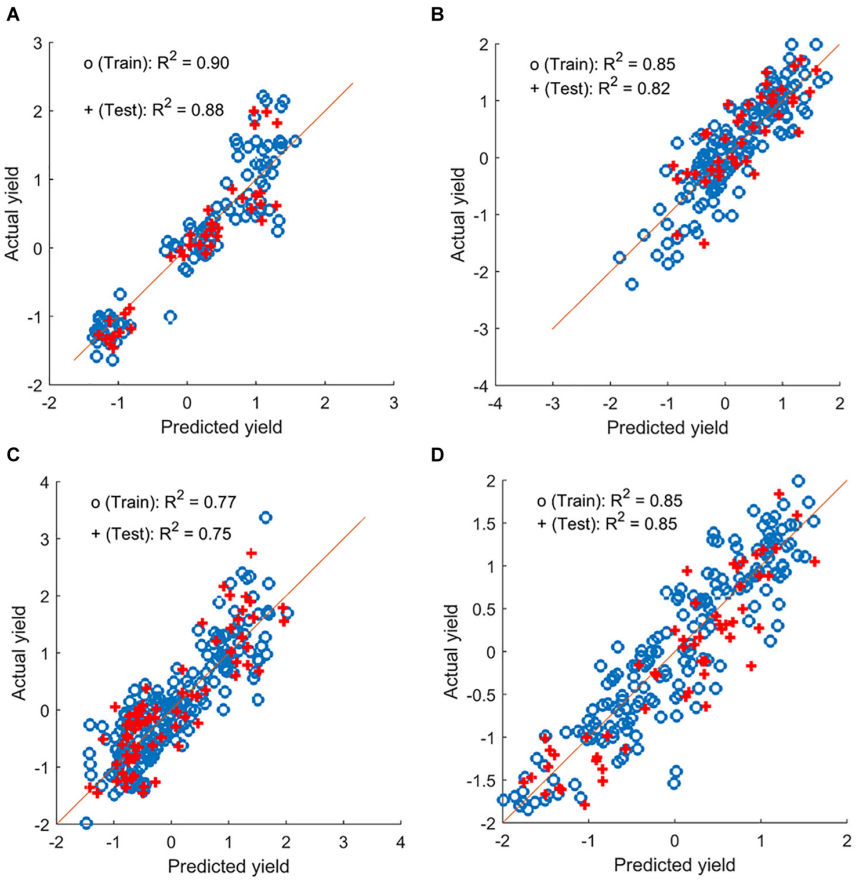

FIGURE 7 | Yield prediction results from a sample iteration using the combined data for chickpea and pea (plot-by-plot analysis). **(A,B)** Chickpea datasets from 2018 and 2019. **(C,D)** Green pea datasets from 2018 and 2019, respectively.

at least 75% of the time during multiple runs of model development were considered. Among these features, canopy area and NDRE or sum of NDRE were usually selected as predictors. Features derived from the data collected at the flowering and pod/seed development stages were both selected during model development, indicating similar importance of these two stages.

Yield in the green pea yield trials can also be estimated by multiple image-based features using data from individual locations and combined within a year (**Supplementary Table 1**). Prediction (testing) accuracy reached up to 0.72 and 0.80 for the data from individual locations and the combined datasets within a year (e.g., **Figure 7**), respectively. Similar to chickpea, the features derived from the images acquired at flowering and at pod/seed development were both selected as predictors. However, more features (3–20) were used in the model development for this panel, and more models performed poorly when analyzing the cultivar-by-cultivar data. These may be related to the complicated pea canopy architecture that comprised stipules and tendrils. Among these features, canopy area, NDRE, and TVI were selected more often as predictors than other features.

DISCUSSION

The study demonstrates that image-based features including canopy area, NDVI, SAVI, and sum NDVI derived from UAV data can be used to monitor performance traits such as yield, days to 50% flowering, or days to physiological maturity across experimental locations and field seasons in two pulse crops, chickpea and pea. Phenomics technologies, especially UAV-based multispectral imaging systems, can be used to acquire data in a standard, rapid, and high-throughput manner, providing plant breeders with information for more informative decision making. Data on other agronomic and phenological traits, such as days to 50% flowering, days to physiological maturity, and plant height, are limited to one location or acquired at a low data acquisition frequency, often due to resource limitations, especially given the number of and the distance to the trial sites. However, using UAV integrated with multispectral camera, image data can be acquired within 30 min per trial (including setting up the UAV system), and multiple locations and crops can be imaged in a single day (depending on the distance to the trial sites). The efficiency of phenomics technologies can improve the availability of such data across multiple locations, which allows plant breeders to study

the interaction between genotypes and the environment on the morphological or phenological traits.

Besides ensuring data availability, phenomics technologies can monitor a wide range of traits in pulse and other crops, including plant height and lodging (Watanabe et al., 2017; Quirós Vargas et al., 2019), disease (Marzougui et al., 2019; Zhang et al., 2019), flower intensity (Yahata et al., 2017; Zhang et al., 2020), and other traits, as discussed in this study. In addition, new traits can be derived from high temporal resolution data, such as crop growth and development curves based on canopy area, vigor, and plant height (Chang et al., 2017; Malambo et al., 2018), allowing plant breeders to assess development of each cultivar quantitatively and intensively. Current and previous studies demonstrated that seed yield or biomass of pulse or other crops can be predicted with image-based features (Fieuzal et al., 2017; Yue et al., 2017; Anderson et al., 2019; Li et al., 2019; Sankaran et al., 2019; Moghimi et al., 2020). Different machine learning models, such as LASSO, SVM, and deep neural networks, have been tested for yield prediction. For example, Moghimi et al. (2020) applied deep neural networks along with aerial hyperspectral images to predict wheat yield, which demonstrated coefficients of determination of 0.79 and 0.41 at the subplot and plot scales, respectively. Yue et al. (2017) selected the 10 most important variables among 172 variables, which were derived from multispectral and RGB images, with random forest and LASSO and used the selected variables to predict wheat yield through support vector machine (SVM) and ridge regression. Their study showed that SVM with random forest-selected variables (r = 0.36–0.77) and ridge regression with LASSO-selected variables (r = 0.40–0.73) slightly outperformed those with all variables (r = 0.25–0.72 and 0.22–0.73, respectively). In this regard, we found similar or better results in the current study, especially with the combined datasets. Such performance monitoring technologies can be applied in agricultural production as well as plant breeding to plan agronomic operations and save labor costs and time.

Although promising results were found in this study, some observations need further investigation. In some high-yielding trials (e.g., 2018 Pullman trial), low correlation coefficients and prediction accuracy with image-based features were observed compared to other seasons and locations. Similar observations were found in dry bean studies (Sankaran et al., 2018, 2019). One possible explanation that Sankaran et al. (2019) proposed may be that low canopy vigor resulted in great differences in the vegetation index values, which led to stronger correlations between ground truth and the vegetation index values.

Further research is also necessary to build more robust yield prediction models and confirm the potential yield predictors. Although it is possible to predict the seed yield of chickpea and dry pea, the image-based features selected in the prediction models varied across locations, years, and analysis methods. Yield prediction should be more consistent across locations and seasons with universal or common features. In general, the performance of machine learning models improves with increased data quantity and quality, which may be exploited in future study. Additional features can also be considered when building such robust prediction models, such as modified chlorophyll absorption ratio index, photochemical reflectance index, normalized difference infrared index (Harris Geospatial Solutions, 2020), plant height (Bendig et al., 2015; Rueda-Ayala et al., 2019), and canopy temperature (Sankaran et al., 2019; Zhang et al., 2019).

One of the challenges of phenotyping dry pea crop is its unique canopy architecture. The canopies of many crops, such as wheat, rice, corn, and soybean, consist of only leaves for a majority of the growing season with flowers among the canopy for a short period of time. In contrast, the pea canopy is made up of stipules and leaflets and/or tendrils for the majority of the growing season, and the tendrils may have different spectral reflectance characteristics from stipules or leaflets, which could have contributed to the lower performance of dry pea than chickpea in this study. Further study is required to identify the spectral reflectance characteristics of tendrils and stipules in pea and its relationship to crop performance.

CONCLUSION

This study was conducted to evaluate phenomics technologies for monitoring performance traits (e.g., seed yield, days to 50% flowering, and days to physiological maturity) and predict the seed yield of chickpea and pea in three growing seasons and three environments (or locations). Significant correlations ($P < 0.05$) between the image features derived from multispectral UAV-based imagery and the yields of chickpea ($r < 0.93$) and pea ($r < 0.85$) were observed at the early growth, flowering, and pod/seed development stages, with a few exceptions. During seed yield prediction with the combined features dataset using LASSO regression, R^2 values up to 0.91 and 0.80 (model testing) were achieved for chickpea and pea, respectively. The image-based features were identified by the LASSO regression models as the yield predictors for chickpea (one to seven features) and pea (3–20 features). The results indicated that phenomics technologies can be employed to collect data and evaluate pulse crop performance in multiple field seasons and environments and save labor and time for plant breeders. With further refinement (e.g., a software platform for data management and image analysis), phenomics technologies can be used to assist plant breeders in evaluating the performance of breeding materials and accelerate the development of new cultivars of pulse and other crops.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

CZ and SS contributed to conceptualization, methodology, formal analysis, software development, visualization, and writing of the original draft. SS, RM, and GV helped with funding acquisition and resources. CZ, SS, RM, and GV did the data

curation and investigation. All authors writing, reviewed, and edited the manuscript.

FUNDING

This activity was funded in part by the United States Department of Agriculture (USDA) – National Institute for Food and Agriculture (NIFA) Agriculture and Food Research Initiative Competitive Projects WNP06825, WNP08532 (accession numbers 1011741, 1008828) and Hatch Project WNP00011 (accession number 1014919).

ACKNOWLEDGMENTS

The authors also appreciate the assistance of Afef Marzougui, Juan Quiros, and Worasit Sangjan during data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.640259/full#supplementary-material

REFERENCES

Acquaah, G. (2009). *Principles of Plant Genetics and Breeding*. Hoboken, NJ: John Wiley & Sons.

Anderson, S. L., Murray, S. C., Malambo, L., Ratcliff, C., Popescu, S., Cope, D., et al. (2019). Prediction of maize grain yield before maturity using improved temporal height estimates of unmanned aerial systems. *Plant Phenome J.* 2, 1–15. doi: 10.2135/tppj2019.02.0004

Bai, G., Ge, Y., Hussain, W., Baenziger, P. S., and Graef, G. (2016). A multi-sensor system for high throughput field phenotyping in soybean and wheat breeding. *Comput. Electron. Agric.* 128, 181–192. doi: 10.1016/j.compag.2016.08.021

Ballesteros, R., Ortega, J. F., Hernandez, D., del Campo, A., and Moreno, M. A. (2018). Combined use of agro-climatic and very high-resolution remote sensing information for crop monitoring. *Int. J. Appl. Earth Obs. Geoinf.* 72, 66–75.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., et al. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Obs. Geoinf.* 39, 79–87. doi: 10.1016/j.jag.2015.02.012

Bourgeois, G., Jenni, S., Laurence, H., and Tremblay, N. (2000). Improving the prediction of processing pea maturity based on the growing-degree day approach. *HortScience* 35, 611–614. doi: 10.21273/HORTSCI.35.4.611

Chang, A., Jung, J., Maeda, M. M., and Landivar, J. (2017). Crop height monitoring with digital imagery from Unmanned Aerial System (UAS). *Comput. Electron. Agric.* 141, 232–237. doi: 10.1016/j.compag.2017.07.008

Donohue, R. J., Lawes, R. A., Mata, G., Gobbett, D., and Ouzman, J. (2018). Towards a national, remote-sensing-based model for predicting field-scale crop yield. *Field Crops Res.* 227, 79–90. doi: 10.1016/j.fcr.2018.08.005

Fieuzal, R., Marais Sicre, C., and Baup, F. (2017). Estimation of corn yield using multi-temporal optical and radar satellite data and artificial neural networks. *Int. J. Appl. Earth Obs. Geoinf.* 57, 14–23. doi: 10.1016/j.jag.2016.12.011

Harris Geospatial Solutions (2020). *Vegetation Indices*. Available online at: http://www.harrisgeospatial.com/docs/VegetationIndices.html (accessed May 20, 2020)

Hatfield, J. L., and Walthall, C. L. (2015). Meeting global food needs: realizing the potential via genetics × environment × management interactions. *Agron. J.* 107, 1215–1226. doi: 10.2134/agronj15.0076

Kipp, S., Mistele, B., Baresel, P., and Schmidhalter, U. (2014). High-throughput phenotyping early plant vigour of winter wheat. *Eur. J. Agron.* 52, 271–278.

Lai, Y. R., Pringle, M. J., Kopittke, P. M., Menzies, N. W., Orton, T. G., and Dang, Y. P. (2018). An empirical model for prediction of wheat yield, using time-integrated Landsat NDVI. *Int. J. Appl. Earth Obs. Geoinf.* 72, 99–108. doi: 10.1016/j.jag.2018.07.013

Li, J., Veeranampalayam-Sivakumar, A.-N., Bhatta, M., Garst, N. D., Stoll, H., Stephen Baenziger, P., et al. (2019). Principal variable selection to explain grain yield variation in winter wheat from features extracted from UAV imagery. *Plant Methods* 15:123. doi: 10.1186/s13007-019-0508-7

Lindsey, A. J., Craft, J. C., and Barker, D. J. (2020). Modeling canopy senescence to calculate soybean maturity date using NDVI. *Crop Sci.* 60, 172–180. doi: 10.1002/csc2.20079

Madec, S., Baret, F., de Solan, B., Thomas, S., Dutartre, D., Jezequel, S., et al. (2017). High-throughput phenotyping of plant height: comparing unmanned aerial vehicles and ground LiDAR estimates. *Front. Plant Sci.* 8:2002. doi: 10.3389/fpls.2017.02002

Malambo, L., Popescu, S. C., Murray, S. C., Putman, E., Pugh, N. A., Horne, D. W., et al. (2018). Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. *Int. J. Appl. Earth Obs. Geoinf.* 64, 31–42. doi: 10.1016/j.jag.2017.08.014

Marzougui, A., Ma, Y., Zhang, C., McGee, R., Coyne, C., Main, D., et al. (2019). Advanced imaging for quantitative evaluation of *Aphanomyces* root rot resistance in lentil. *Front. Plant Sci.* 10:383. doi: 10.3389/fpls.2019.00383

McGee, R. J., Coyne, C. J., Pilet-Nayel, M.-L., Moussart, A., Tivoli, B., Baranger, A., et al. (2012). Registration of pea germplasm lines partially resistant to *Aphanomyces* root rot for breeding fresh or freezer pea and dry pea types. *J. Plant Regist.* 6, 203–207. doi: 10.3198/jpr2011.03.0139crg

McGee, R. J., and McPhee, K. E. (2012). Release of autumn-sown pea germplasm PS03101269 with food-quality seed characteristics. *J. Plant Regist.* 6, 354–357. doi: 10.3198/jpr2011.09.0511crg

McGee, R. J., Nelson, H., and McPhee, K. E. (2013). Registration of 'Lynx' winter pea. *J. Plant Regist.* 7, 261–264. doi: 10.3198/jpr2012.09.0040crc

Miller, P., Lanier, W., and Brandt, S. (2018). *Using Growing Degree Days to Predict Plant Stages*. Bozeman, MT: Montana State University.

Moghimi, A., Yang, C., and Anderson, J. A. (2020). Aerial hyperspectral imagery and deep neural networks for high-throughput yield phenotyping in wheat. *Comput. Electron. Agric.* 172:105299. doi: 10.1016/j.compag.2020.105299

Patrignani, A., and Ochsner, T. E. (2015). Canopeo: a powerful new tool for measuring fractional green canopy cover. *Agron. J.* 107, 2312–2320.

Quirós Vargas, J. J., Zhang, C., Smitchger, J. A., McGee, R. J., and Sankaran, S. (2019). Phenotyping of plant biomass and performance traits using remote sensing techniques in pea (*Pisum sativum*, L.). *Sensors* 19:2031. doi: 10.3390/s19092031

Rueda-Ayala, V., Peña, J., Höglind, M., Bengochea-Guevara, J., and Andújar, D. (2019). Comparing UAV-based technologies and RGB-D reconstruction methods for plant height and biomass monitoring on grass ley. *Sensors* 19:535. doi: 10.3390/s19030535

Sadeghi-Tehran, P., Sabermanesh, K., Virlet, N., and Hawkesford, M. J. (2017). Automated method to determine two critical growth stages of wheat: heading and flowering. *Front. Plant Sci.* 8:252. doi: 10.3389/fpls.2017.00252

Sammut, C., and Webb, G. I. (eds) (2010). "Leave-one-out cross-validation," in *Encyclopedia of Machine Learning*, eds C. Sammut and G. I. Webb (Boston, MA: Springer US), 600–601. doi: 10.1007/978-0-387-30164-8_469

Sankaran, S., Khot, L. R., and Carter, A. H. (2015). Field-based crop phenotyping: multispectral aerial imaging for evaluation of winter wheat emergence and spring stand. *Comput. Electron. Agric.* 118, 372–379. doi: 10.1016/j.compag.2015.09.001

Sankaran, S., Quirós, J. J., and Miklas, P. N. (2019). Unmanned aerial system and satellite-based high resolution imagery for high-throughput phenotyping in dry bean. *Comput. Electron. Agric.* 165:104965. doi: 10.1016/j.compag.2019.104965

Sankaran, S., Zhou, J., Khot, L. R., Trapp, J. J., Mndolwa, E., and Miklas, P. N. (2018). High-throughput field phenotyping in dry bean using small unmanned aerial vehicle based multispectral imagery. *Comput. Electron. Agric.* 151, 84–92. doi: 10.1016/j.compag.2018.05.034

Sun, S., Li, C., Paterson, A. H., Jiang, Y., Xu, R., Robertson, J. S., et al. (2018). In-field high throughput phenotyping and cotton plant growth analysis using LiDAR. *Front. Plant Sci.* 9:16. doi: 10.3389/fpls.2018.00016

Tattaris, M., Reynolds, M. P., and Chapman, S. C. (2016). A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding. *Front. Plant Sci.* 7:1131. doi: 10.3389/fpls.2016.01131

Thorp, K. R., Wang, G., Badaruddin, M., and Bronson, K. F. (2016). *Lesquerella* seed yield estimation using color image segmentation to track flowering dynamics in response to variable water and nitrogen management. *Ind. Crops Prod.* 86, 186–195. doi: 10.1016/j.indcrop.2016.03.035

USDA-ARS (2018). *G. L. G. P. R. U*: *Variety Releases*. Available online at: https://www.ars.usda.gov/pacific-west-area/pullman-wa/grain-legume-genetics-physiology-research/docs/variety-releases/ (accessed April 30, 2020)

USDA-NASS (2020). *Crop Production 2019 Summary 01/10/2020*. Washington, DC: USDA-NASS, 124.

Vandemark, G., Guy, S. O., Chen, W., McPhee, K., Pfaff, J., Lauver, M., et al. (2015). Registration of 'Nash' chickpea. *J. Plant Regist.* 9, 275–278. doi: 10.3198/jpr2014.07.0047crc

Vandemark, G., Muehlbauer, F. J., Mihov, M., Chen, W., McPhee, K., and Chen, C. (2014). Registration of CA0469C025C chickpea germplasm. *J. Plant Regist.* 8, 303–307. doi: 10.3198/jpr2013.09.0057crg

Viña, A., Gitelson, A. A., Rundquist, D. C., Keydan, G., Leavitt, B., and Schepers, J. (2004). Monitoring maize (*Zea mays* L.) phenology with remote sensing. *Agron. J.* 96, 1139–1147. doi: 10.2134/agronj2004.1139

Wang, X., Singh, D., Marla, S., Morris, G., and Poland, J. (2018). Field-based high-throughput phenotyping of plant height in sorghum using different sensing technologies. *Plant Methods* 14:53. doi: 10.1186/s13007-018-0324-5

Watanabe, K., Guo, W., Arai, K., Takanashi, H., Kajiya-Kanegae, H., Kobayashi, M., et al. (2017). High-throughput phenotyping of sorghum plant height using an unmanned aerial vehicle and its application to genomic prediction modeling. *Front. Plant Sci.* 8:421. doi: 10.3389/fpls.2017.00421

Yahata, S., Onishi, T., Yamaguchi, K., Ozawa, S., Kitazono, J., Ohkawa, T., et al. (2017). "A hybrid machine learning approach to automatic plant phenotyping for smart agriculture," in *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN)*, (Anchorage, AK: IEEE), 1787–1793.

Yang, Z., Shao, Y., Li, K., Liu, Q., Liu, L., and Brisco, B. (2017). An improved scheme for rice phenology estimation based on time-series multispectral HJ-1A/B and polarimetric RADARSAT-2 data. *Rem. Sens. Environ.* 195, 184–201. doi: 10.1016/j.rse.2017.04.016

Yu, N., Li, L., Schmitz, N., Tian, L. F., Greenberg, J. A., and Diers, B. W. (2016). Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. *Rem. Sens. Environ.* 187, 91–101. doi: 10.1016/j.rse.2016.10.005

Yue, J., Yang, G., Li, C., Li, Z., Wang, Y., Feng, H., et al. (2017). Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. *Rem. Sens.* 9:708. doi: 10.3390/rs9070708

Zhang, C., Chen, W., and Sankaran, S. (2019). High-throughput field phenotyping of *Ascochyta* blight disease severity in chickpea. *Crop Prot.* 125:104885. doi: 10.1016/j.cropro.2019.104885

Zhang, C., Craine, W. A., McGee, R. J., Vandemark, G. J., Davis, J. B., Brown, J., et al. (2020). Image-based phenotyping of flowering intensity in cool-season crops. *Sensors* 20:1450. doi: 10.3390/s20051450

Zheng, H., Cheng, T., Yao, X., Deng, X., Tian, Y., Cao, W., et al. (2016). Detection of rice phenology through time series analysis of ground-based spectral index data. *Field Crops Res.* 198, 131–139. doi: 10.1016/j.fcr.2016.08.027

Conflict of Interest: 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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 01 April 2021
doi: 10.3389/fpls.2021.628575

UAV-Based Thermal, RGB Imaging and Gene Expression Analysis Allowed Detection of Fusarium Head Blight and Gave New Insights Into the Physiological Responses to the Disease in Durum Wheat

Sara Francesconi[1], Antoine Harfouche[2], Mauro Maesano[2] and Giorgio Mariano Balestra[1]**

[1] Department of Agriculture and Forest Sciences (DAFNE), University of Tuscia, Viterbo, Italy, [2] Department for Innovation in Biological, Agro-Food and Forest Systems (DIBAF), University of Tuscia, Viterbo, Italy

Wheat is one of the world's most economically important cereal crop, grown on 220 million hectares. Fusarium head blight (FHB) disease is considered a major threat to durum (*Triticum turgidum* subsp. *durum* (Desfontaines) Husnache) and bread wheat (*T. aestivum* L.) cultivars and is mainly managed by the application of fungicides at anthesis. However, fungicides are applied when FHB symptoms are clearly visible and the spikes are almost entirely bleached (% of diseased spikelets > 80%), by when it is too late to control FHB disease. For this reason, farmers often react by performing repeated fungicide treatments that, however, due to the advanced state of the infection, cause a waste of money and pose significant risks to the environment and non-target organisms. In the present study, we used unmanned aerial vehicle (UAV)-based thermal infrared (TIR) and red-green-blue (RGB) imaging for FHB detection in *T. turgidum* (cv. Marco Aurelio) under natural field conditions. TIR and RGB data coupled with ground-based measurements such as spike's temperature, photosynthetic efficiency and molecular identification of FHB pathogens, detected FHB at anthesis half-way (Zadoks stage 65, ZS 65), when the percentage (%) of diseased spikelets ranged between 20% and 60%. Moreover, in greenhouse experiments the transcripts of the key genes involved in stomatal closure were mostly up-regulated in *F. graminearum*-inoculated plants, demonstrating that the physiological mechanism behind the spike's temperature increase and photosynthetic efficiency decrease could be attributed to the closure of the guard cells in response to *F. graminearum*. In addition, preliminary analysis revealed that there is differential regulation of genes between drought-stressed and *F. graminearum*-inoculated plants, suggesting that there might be a possibility to discriminate between water stress and FHB infection. This study shows the potential of UAV-based TIR and RGB imaging for field phenotyping of wheat and other cereal crop

OPEN ACCESS

Edited by:
Alessandro Matese, Institute for Bioeconomy, Italian National Research Council, Italy

Reviewed by:
Julia Christine Meitz-Hopkins, Stellenbosch University, South Africa
Marco Camardo Leggieri, Catholic University of the Sacred Heart, Italy
Franz Badeck, Council for Agricultural and Economics Research (CREA), Italy

***Correspondence:**
Sara Francesconi
francesconi.s@unitus.it
Giorgio Mariano Balestra
balestra@unitus.it

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

Received: *12 November 2020*
Accepted: *12 March 2021*
Published: *01 April 2021*

Citation:
Francesconi S, Harfouche A, Maesano M and Balestra GM (2021) UAV-Based Thermal, RGB Imaging and Gene Expression Analysis Allowed Detection of Fusarium Head Blight and Gave New Insights Into the Physiological Responses to the Disease in Durum Wheat. Front. Plant Sci. 12:628575. doi: 10.3389/fpls.2021.628575

species in response to environmental stresses. This is anticipated to have enormous promise for the detection of FHB disease and tremendous implications for optimizing the application of fungicides, since global food crop demand is to be met with minimal environmental impacts.

Keywords: Fusarium head blight, durum wheat, high-throughput plant phenotyping, thermal imaging, RGB imaging, gene expression, early disease detection

INTRODUCTION

Wheat is one of the most cultivated crops worldwide, grown on 220 million hectares and its annual production is estimated to account for more than 700 million tons, providing more than 20% of total human food calories (Khan et al., 2020; Ma et al., 2020). Modern wheat cultivars derive from two species: bread wheat (*Triticum aestivum* L.) and durum-type wheat (*T. turgidum* subsp. *durum* (Desfontaines) Husnache) used for pasta and low-rising bread (Peng et al., 2011; Feldman and Levy, 2012).

The current food demand will double with a projected population of 9 billion in 2050. In response, farmers must grow more on their lands through sustainable intensification of agriculture, an approach to increase yield production without having negative effects on the environment or expanding the agricultural footprint (Giller et al., 2015).

Wheat production is challenged by several abiotic and biotic stresses. Among the plant fungal diseases, Fusarium head blight (FHB) is one of the most destructive diseases leading to 10–70% of yield loss during the epidemic years (McMullen et al., 2012). FHB is caused by the *Fusarium graminearum* species complex (FGSC), which embraces 16 different species. It is considered the most devastating wheat disease due to the lack of resistant cultivars, the significant yield loss and grain quality reduction, and the health risks associated with the contamination of crops with mycotoxin such as deoxynivalenol (DON) and zearalenone (ZEA), produced during the infection progress (Yang et al., 2013; Dweba et al., 2017). The spectrum of *Fusarium* spp. causing FHB on wheat varies at the regional level, depending on weather conditions. Fungal growth is favored by high temperatures and humidity during the growing season. *F. graminearum* Schwabe is the predominant species that causes FHB in many countries, in Asia, North America, South America, and Europe (Dweba et al., 2017; Khan et al., 2020). In field conditions, the inoculum occurs primarily from plant residues and soils while the dissemination of the ascospores and conidia is mainly promoted by water splash and wind. Anthesis is the most susceptible stage to infection. With a warm and humid climate at this stage, airborne spores proliferate and spread intracellularly into the rachial nodes and through the rachis until FHB symptoms are clearly visible. The symptoms include necrosis, premature bleaching of spikes, and shriveled kernels that negatively affect photosynthesis. At favorable conditions, FHB develops rapidly within 3–6 days after the infection. Given the current global warming associated with increased temperatures, major epidemics of FHB are occurring. Thus, proper cultivation schemes and field management are essential to alleviate its threat (Bai and Shaner, 2004; Vaughan et al., 2016; Ma et al., 2020; Rod et al., 2020).

Since 1995, *Fusarium* spp. infect wheat in Italy at various incidence (percentage of spikes showing symptoms) and severity (percentage of areas infected in spikes) depending on the year, the region, and the wheat cultivar involved (Pancaldi et al., 2010). Its incidence and severity are closely related to the amounts of precipitation during wheat anthesis, increasing from the South to the North of Italy, being mostly reported in the Northern-Central regions (Shah et al., 2005; Infantino et al., 2012). Between the two major species, *T. turgidum* is the most widely grown in Italy, but is also more susceptible to FHB (Oliver et al., 2008). Consequently, being mostly used for human consumption, the risk of mycotoxin-contaminated grain entering the food chain is a major concern (Boutigny et al., 2008).

Chemical control of FHB using appropriate effective fungicides and correct application methods and timing can reduce the severity of the disease (Blandino et al., 2012). However, no fully effective FHB fungicide is available (Haidukowski et al., 2012), and the application window is very narrow, spanning just a few days around anthesis (Mesterházy et al., 2003). Recently, an organic strategy to reduce FHB incidence and severity was proposed, but now it needs to be confirmed on wide areas (Francesconi et al., 2020). For these reasons, early detection and control of FHB is a key factor to gain maximum protection of yield (Mahlein et al., 2019) from fungal spread and mycotoxin accumulation; in fact, there is a strong evidence base in the research literature that a prompt and early application of fungicides instead of applying fungicides at late stages of infection drastically reduce the FHB spread and DON accumulation inside the grains, instead of applying fungicides at late stage of infection (Freije and Wise, 2015; Feksa et al., 2019; Bolanos-Carriel et al., 2020).

Recent advances in biological and bioanalytical research enabled genome-scale capturing of biological processes at the molecular level in plants (Weckwerth, 2011). Though transcripts evaluation is fundamental to understand plant responses to pathogens, these techniques are labor-intensive, time-consuming, destructive and slow-down the acquisition of phenotype parameters related to the gene responses, contributing to the phenotyping bottleneck (Furbank and Tester, 2011). Molecular data obtained in greenhouse or field trials combined with phenotypic and environmental data discloses a wealth of information that can be used to improve field management (Alexandersson et al., 2014).

For plant-pathologists, coupling transcriptomics and phenomics to agricultural practices is likely to have a large impact on the understanding of induced plant defenses and pathogen spread. In fact, phenomics and transcripts analysis can reveal important physiological changes in plants in response

to pathogens which can help detecting infections before the appearance of their visible symptoms. Hence, these associated techniques have the potential to be a powerful weapon that optimizes fungicide spraying regimes for plant pathogen management (Alexandersson et al., 2014).

To relieve the phenotyping bottleneck, phenotypic traits should be turned into quantifiable, objective, and consistent measures (Ludovisi et al., 2017). Automated and high-throughput phenotyping (HTP) provides measurements that can track the development of a plant through its life stages and its interaction with the environment, establishing methodologies to detect crucial physiological traits and identify effective genotype-phenotype relationships (Goggin et al., 2015). These methods speed up the phenotyping process and maximize the number of studied plants per experiment (Goggin et al., 2015). Furthermore, they enable automated, non-destructive, and non-invasive screening of large populations and high dimensionality data (Li et al., 2014; Fahlgren et al., 2015).

With regards to FHB, *Fusarium* spp. infects spikelets within wheat spikes, decreasing stomatal conductance, and thus, affecting the chlorophyll and water content (Yang et al., 2016; Francesconi and Balestra, 2020). This allows its early detection by color imaging and by measuring the temperature of spikes and photosynthetic efficiency of plants using thermometers and fluorometers, respectively (Cambaza et al., 2019; Mahlein et al., 2019).

Red-green-blue (RGB) cameras are designed to emulate human vision by sensing light in the visible range of the electromagnetic spectrum (wavelengths from 390 to 700 nm). In this range, the reflectance is predominantly influenced by plant pigments (e.g., chlorophyll) (Mahlein, 2016). This allows the calculation of different vegetation indices (VIs) by computing the reflectance of a certain band of the green and red zone of the electromagnetic spectrum (Barbosa et al., 2019; Brunori et al., 2020). Photosynthetic response of green vegetation to incident light is the basis of VIs where healthy plants exhibit low red reflectance due to absorption of red light by chlorophyll resulting in a high index value, whereas unhealthy, stressed plants with reduced chlorophyll pigment display a low index value (Khan et al., 2018).Therefore, RGB image analysis can serve as an important tool that detects physiological changes in plants caused by pathogen infections (Simko et al., 2016). On the other hand, thermal infrared (TIR) sensors capture images containing information about the energy emitted from body surfaces, such as plant canopies. Plant pathogens affect the loss of water in plants regulated by stomata (Fang and Ramasamy, 2015), altering plant temperature where high values indicate areas with closed stomata (Oerke et al., 2006; Chaerle et al., 2007). Thermography can serve as a tool to measure these changes toward an early detection of plant infections, ideally before symptoms appear (Al Masri et al., 2017).

Unmanned aerial vehicles (UAVs) equipped with cameras and sensors have become advanced field phenomics tools that provide data with high spatial and temporal resolution to bridge the gap between time consuming ground-based measurements and satellite observations (Xie and Yang, 2020; Pineda et al., 2021). UAVs allow rapid and non-destructive measurements and offer much quicker turnaround times than satellites at competitive costs (Ludovisi et al., 2017).

The application of UAV-based imaging techniques has been broadening in several areas of agricultural sciences thanks to their ability to analyze plant temperature and color discrepancies between distinct biological samples (Padmavathi and Thangadurai, 2016; Cambaza et al., 2019). Many recent studies have exploited UAV-based sensors to monitor, detect and phenotype plant stresses in forestry (Sagan et al., 2019), as well as to estimate leaf nitrogen concentration (Lu et al., 2021), water and nitrogen use efficiencies (Yang et al., 2020), and salinity stress (Johansen et al., 2019) in different crops. Plant pathologists are recently also benefiting from the application of UAV-based sensors; in fact, UAV multispectral and hyperspectral imaging have been used to detect *Xylella fastidiosa* (Castrignanò et al., 2021) and FHB (Liu et al., 2020), respectively. Moreover, these techniques are relatively easy to use and are becoming cheaper (Dammer et al., 2011; Cambaza et al., 2019). Their exploitation to monitor FHB can contribute significantly to secure the cereal production systems (Mahlein et al., 2019). In light of these advantages, UAV-mounted cameras and sensors are expected to enable new applications in field-based phenotyping of plant stress traits in large populations rapidly, precisely and accurately (Berni et al., 2009; White et al., 2012; Yang et al., 2017).

In this study, we exploited plant physiological and molecular changes in response to *Fusarium* spp. infection to detect FHB in *T. turgidum* (cv. Marco Aurelio) fields through UAV-based TIR and RGB imaging. Particularly, (i) two VIs as well as spike temperature were calculated using RGB and TIR images, from both uninfected and infected areas and were correlated to photosynthesis perturbation caused by the infection. Additionally, (ii) transcripts of key genes involved in stomatal conductance regulation were investigated in greenhouse experiments, to explore the genotypic changes behind the observed phenotypic perturbations (increase in spike temperature and decrease in photosynthetic efficiency). Furthermore, (iii) transcripts of *F. graminearum* inoculated plants were compared to those obtained from drought-stressed ones in order to investigate a differential response between these two types of stresses.

MATERIALS AND METHODS

Plant Material and Experimental Design of the Field Experiments

The FHB-susceptible *T. turgidum* cv. Marco Aurelio was the cultivar of interest in the present work for both field and greenhouse experiments. This genotype, grown in Central and South Italy, is extensively used for pasta production and it is characterized by excellent protein content and high productivity, thus it is of high economic importance. The experimental fields were located in Amelia (Central Italy, 42°31′22.9″N, 12°25′15.5″E, Umbria Region) and Avigliano Umbro (Central Italy, 42°40′41.1″N, 12°27′44.6″E, Umbria Region). The *T. turgidum* field experiments were

carried out during two consecutive years (2019 and 2020), controlled and drip-irrigated with a nutrient solution containing nitrogen, potassium, phosphorus, and small amounts of other compounds. Therefore, abiotic stresses (nutrient deficiencies and drought stress), diseases, or pests, which cause the same symptoms as FHB, were significantly minimized. On the other hand, fungicides for FHB management (tebuconazole and azoxystrobin) were not applied in order to favor the FHB natural infection. The fields were investigated periodically (one time per day) by experts and farmers to prevent other diseases, pests, or abiotic stress. An experimental plot was allocated within each field, and individuated by a 20 m grid by positioning 16 ground control points (GCPs) to be used for georeferencing. During March, April, and May 2019 and 2020 weather data (minimum, maximum and average temperature, and precipitation) were recorded daily by two meteorological stations installed at 100 m distance from each field to monitor the climate factors influencing the establishment of FHB infection. Historical weather data (from 2010 to 2018) were obtained from the Hydrographic service of Umbria Region[1] to monitor the climatic trend of the last 11 years and to compare minimum, average and maximum temperature values (°C) and average precipitations (mm) recorded during the 9-year seasonal average (2010–2018), 2019, and 2020. The historical weather data were collected from two meteorological stations located in Amelia (42°33′25.0″N, 12°25′01.0″E) and Avigliano Umbro (42°40′39.0″N, 12°26′13.0″E). The experimental design for the field experiments is illustrated in **Figures 1A,B**.

UAV Campaigns

An unmanned DJI Matrice 600 multi-copter (DJI, China) equipped with a Zenmuse X5 RGB camera (DJI, China) and a Zenmuse XT TIR camera (DJI, China) has been used in this study (**Supplementary Figure 1**). DJI Matrice 600 is a hexa-copter (110 cm diagonal size) with a highly resistant carbon fiber frame, offering a 15 kg take-off weight. Its maximum transmission distance is 5 km and its maximum flight time is 40 min (**Supplementary Figure 1**).

The Zenmuse XT was equipped with a 9 mm f1.4 lens. Its thermal sensitivity is less than 50 milliKelvins and the camera enables measurements in the range −25°C to +135°C. The image sensor is a focal plane array (FPA) based on uncooled microbolometers with a spectral band response in the range of 7.5 to 13 μm. The camera field of view is equal to 69° (horizontally) × 56° (vertically), its resolution to 640 × 512 pixels at 30 frames per second (fps), and its spatial resolution to 1.889 milliradians (mrad). The camera captures images at an acquisition frequency of 30 Hz. The camera was radiometrically calibrated to further increase temperature measurement precision by setting external parameters during the flight planning such as air temperature and flat field correction (FFC). This is an offset calibration usually performed at power up, when the camera changes temperature, and periodically during the operation. This calibration compensates for certain errors that build up during the camera operation. During the data acquisition phase, the auto gain mode parameter was applied; the camera automatically selected the optimal gain mode according to the temperature range of the image (Deery et al., 2016; Gómez-Candón et al., 2016; Ludovisi et al., 2017) and pictures were stored as 14-bit digital raw images.

The Zenmuse X5 is a 16-megapixel RGB camera equipped with an M4/3 sensor enabling it to capture detailed images at a resolution of 4608 × 3456 pixels and an ISO range of 100–25600.

The Zenmuse X5 and XT cameras were mounted on a highly reliable gyrostabilized 3-axis gimbal (DJI, China) that automatically stabilizes them in flight. The gimbal constantly communicates with the UAV, and quickly compensates for every minor movement with a precision accuracy of 0.02°.

The 16 GCPs were used to georeference RGB images. For calibration of thermal images, four 60 × 60 cm ground reference panels (GRPs) consisting of two plastic panels covered by black vinyl tape and two white Teflon® were positioned along the borders of each of the two study areas. A real-time kinematic (RTK) global navigation satellite system (GNSS) CS10 model (Leica Geosystems, Switzerland) with an accuracy of 1 mm was used for capturing GCPs and GRPs locations.

UAV campaigns were conducted at the following phenological stages: (i) three-quarters of inflorescence emerged (Zadoks stage 57, ZS 57) (Zadoks et al., 1974) on 6 May 2019 and 7 May 2020; (ii) anthesis half-way (ZS 65) on 17 May 2019 and 18 May 2020; and (iii) kernel watery ripe (ZS 71) on 28 May 2019 and 29 May 2020. Two flights were performed during each campaign. The flight missions have been planned using Ground Station Pro app (DJI GS Pro, China). Each flight lasted 11 min covering an area of 1 ha at a nominal speed of 4 m/s and an altitude of 20 m, during which both cameras acquired nadiral images with 90% frontal and side overlap. To ensure similar solar illumination angles flights were performed between 11:00 and 12:00 local time under stable cloudless and low-wind conditions (**Figures 1D,E**).

Ground Measurements

For each of the 16 GCPs, a circumference of 1 m of diameter was individuated and denominated as sampling area. During the UAV campaigns, we manually measured the spike temperature of sixteen randomly selected plants in each of the 16 sampling areas using a portable infrared thermometer (Fluke 568, Fluke Corporation, United States) with an accuracy of ± 1% or ± 1.0°C (whichever is greater), positioned at 10 cm distance from the spike. At the same time, the FHB severity was calculated by counting the number of diseased spikelets and the total number of spikelets for each of the sixteen plants. Furthermore, the flag leaves of the selected plants were sampled and dark-adapted for 1 h before measuring their photosynthetic efficiency, by quantifying F_v/F_m reflecting the potential quantum efficiency of photosystem II (Maxwell and Johnson, 2000), with a portable fluorometer (V2.00f PAM 2000, Heinz Walz GmbH, Germany). Moreover, the spikes were also sampled to distinguish FHB positive (FHB+) from FHB negative (FHB-) areas by molecular identification of *Fusarium* spp. (**Figure 1C**).

[1] https://annali.regione.umbria.it/

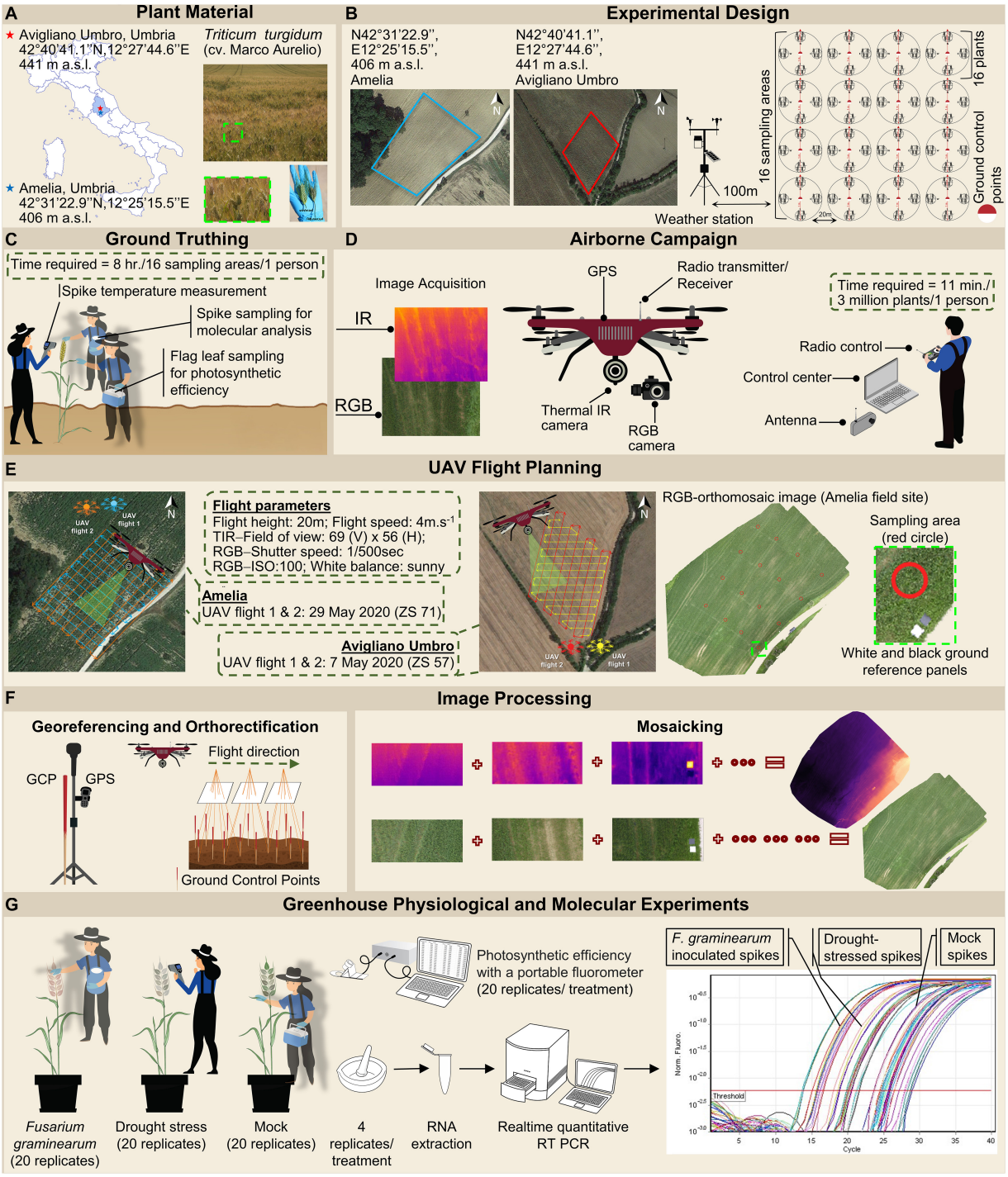

FIGURE 1 | Workflow of the high-throughput field phenotyping methodology and the greenhouse experiment. **(A)** Plant material includes *Triticum turgidum* (cv. Marco Aurelio). Two fields were established in Amelia (Central Italy, 42°31′22.9″N, 12°25′15.5″E, Umbria Region, 406 m above sea level) and Avigliano Umbro (Central Italy, 42°40′41.1″N, 12°27′44.6″E, Umbria Region, 441 m a.s.l) to host the *T. turgidum* plants. Sampled spikes show the symptoms of a Fusarium head blight (FHB)-infected plant, compared to an uninfected plant. **(B)** An experimental plot was allocated within each field, and a 20 m grid was identified by positioning sixteen ground control points (GCPs) to be used for georeferencing. Around each of the GCPs, a circumference of 1 m of diameter was individuated and denominated as a sampling area. Weather data were recorded daily by two meteorological stations installed at 100 m distance from each field. **(C)** Within each sampling area, the spike temperature of sixteen randomly selected spikes was measured, and their flag leaves were sampled to measure their photosynthetic efficiency. Finally, the spikes were also sampled to distinguish FHB positive (FHB+) ones from FHB negative (FHB-) ones by molecular identification of *Fusarium* spp. **(D)** An unmanned DJI Matrice 600 hexacopter was equipped with Zenmuse X5 red-green-blue (RGB) and Zenmuse XT thermal infrared (TIR) sensors.

(Continued)

FIGURE 1 | Continued
The unmanned aerial vehicle (UAV) campaign allowed capturing RGB and TIR images of both experimental plots. **(E)** The flight missions were planned using Ground Station Pro app (DJI GS Pro, China). The UAV was flown in the autonomous mode at a nominal speed of 4 m/s, running parallel to plant rows to ensure complete coverage with good overlaps. **(F)** Image orthorectification, georeferencing, and mosaicking were performed using 16 GCPs captured with a global positioning system (GPS). **(G)** The greenhouse physiological and molecular experiments where plants were subjected to three different treatments: (i) drought stress treatment where plants did not receive water from Zadoks stage (ZS) 51 to 65; (ii) artificial inoculation treatment where spikes were uniformly spray-inoculated at ZS 65 with a suspension of 1×10^5 conidia/mL; and (iii) mock treatment where spikes were uniformly sprayed with a suspension of Tween-20 0.05% resuspended in sterile distilled water at ZS 65. Spikes temperature and photosynthetic efficiency of plants were recorded at 24, 48, and 72 hours post inoculation (hpi) from three independent experiments, each experiment consisting of 20 spikes for each treatment. Four spikes were sampled for each treatment, and were ground with mortar and pestle in liquid nitrogen until a fine powder was obtained, from which, RNA was extracted. Finally, quantitative real-time polymerase chain reaction (RT-qPCR) was performed to investigate transcripts of key genes involved in stomatal conductance regulation; transcripts of *F. graminearum*-inoculated plants were compared to those obtained from drought-stressed ones in order to investigate a differential response between these two types of stresses.

Molecular Diagnostics of *Fusarium* Pathogens

Spikes were stored in a portable fridge immediately after sampling. Afterward, different tissues (palea, lemma, glume, rachis, and kernel when present) from each collected spike were plated on Petri dishes containing potato dextrose agar (PDA) and incubated at 21°C for 72 h. The different morphotypes were subsequentially isolated and cultured on PDA for 1 week. After each morphotype filled the Petri dish, the produced mycelium was gently scraped with a glass rod. 100 mg of mycelium were grinded and DNA was extracted using 80 μL of extraction buffer composed of Tris 100 mM, EDTA 50 mM and NaCl 500 mM; after that, 32 μL of SDS 10% (w/v) were added and the samples were incubated at 65°C for 10 min. 27 μL of potassium acetate 5 M were added and the samples were placed on ice for 20 min, then centrifuged for 20 min at 13.000 rpm. The supernatant was recovered and 80 μL of cold (−20°C) isopropanol were added to each sample; then, the samples were placed on ice for 10 min and centrifuged for 5 min at 13.000 rpm. The supernatant was discarded and 150 μL of cold (−20°C) ethanol (70% v/v) were added to the samples. Samples were centrifuged for 3 min at 13.000 rpm, the supernatant was discarded and the DNA was resuspended in 20 μL of DNase and RNase-free sterile distilled water and stored at −20°C (D'Ovidio and Porceddu, 1996). Total DNA was quantified with Qubit™ fluorometer 1.01 (Invitrogen, United States) using the Qubit™ dsDNA BR Assay Kit (Thermo Fisher Scientific, United States) and diluted to 10 ng/μL. The molecular identification of *Fusarium* spp. was performed by amplifying the translational elongation factor 1-α (*TEF*) sequence using the primer pair EF1_F 5′-ATGGGTAAGGAGGACAAGAC-3′/EF2_R 5′-GGAAGTACCAGTGATCATGTT-3′ designed to identify the FHB complex spp. (Geiser et al., 2004). The polymerase chain reaction (PCR) was performed following the instructions of GoTaq® Green Master Mix (Promega, United States) and prepared in a total volume of 10 μL. The amplification conditions consisted of: (i) an initial denaturation step of 2 min at 95°C; (ii) 35 cycles of 30 s denaturation at 95°C; (iii) 40 s of annealing at 53°C; (iv) 60 s of elongation at 72°C; and (v) a final elongation step of 5 min at 72°C. The amplicon unicity was visualized on 1.5% agarose gel and sequenced by Sanger sequencing (Eurofins Genomics, Germany). The resulted sequences were submitted to BLASTn[2] in order to identify the corresponding *Fusarium* spp.

RGB Image Processing and VIs Calculation

A total of 250 RGB images were acquired using the UAV-mounted Zenmuse X5 camera. The camera exposure mode was set to shutter priority (S) with a shutter speed fixed 1/500 s to ensure minimization of motion blur, ISO of 100 and the white balance to sunny mode.

Geometric camera calibration, orthorectification, and mosaicking of the captured images were performed using Pix4Dmapper (Pix4D, Switzerland) software, specifically designed to process UAV images using techniques rooted in both computer vision and photogrammetry to match conjugate points in overlapped images and to define their relative positions and orientations using bundle block adjustments (Bollard-Breen et al., 2015; Nishar et al., 2016; **Figure 1F**). The output of this step was the RGB orthomosaics of the study areas for each flight mission. No filtering process was applied to the images.

Finally, VIs from the orthomosaics were computed using quantum geographic information system (QGIS) software (version 3.4 Madeira - QGIS Development Team, Open Source Geospatial Foundation). The VIs carried out were: vegetative (VEG) (Hague et al., 2006) and green leaf index (GLI) (Louhaichi et al., 2001). VEG and GLI were calculated using the following formulas:

$$VEG = \frac{G}{R^{a*}B^{(1-a)}} \text{and}$$

$$GLI = \frac{(2^* G - R - B)}{(2^*G + R + B)}$$

where R, G, and B are the reflectance of red, green and blue channels, respectively, and *a* is equal to 0.667.

Inside each sampling area, leaves and spikes were not separated in the images, because the upper part of the fields was uniformly composed by spikes (**Figures 1A,B**), since wheat canopies were very dense.

[2]https://blast.ncbi.nlm.nih.gov/

TIR Image Processing and Temperature Extraction

The UAV-mounted Zenmuse XT TIR camera captured 219 greyscale and georeferenced images. Thermal orthomosaics were generated for each flight, and the radiometric conversion was automatically performed using the Pix4Dmapper software (**Figure 1F**). The removal of bare soil pixels was not necessary, since the wheat planting were extremely dense, reducing the mixed pixel problem. FLIR Tools software (2020© FLIR® Systems, United States) was used to calibrate temperatures in TIR images using the temperature values of GRPs measured immediately after the UAV flights using a portable infrared thermometer (Fluke 568, Fluke Corporation, United States). Calibration checks were performed by comparing the GRPs-measured to the UAV-derived temperatures. No filtering process was applied to the images. The extraction of temperature values was performed for each of the 16 sampling areas. Inside each sampling area, leaves and spikes were not separated in the images, because the upper part of the fields was uniformly composed by spikes (**Figures 1A,B**), since wheat canopies were very dense.

Plant Growth, Inoculation Conditions and Experimental Design of the Greenhouse Experiments

Greenhouse experiments were conducted in a glasshouse located in Viterbo (Central Italy, 42°25′35.8″N, 12°04′49.3″E, Lazio Region). The surface of the *T. turgidum* kernels was sterilized with sodium hypochlorite (0.5% v/v) for 20 min and rinsed twice with sterile distilled water for 5 min. The kernels were germinated in the dark on a paper, soaked in sterile distilled water for 15 days at 4°C to break dormancy, followed by 2 days at room temperature. The seedlings were transferred to 40 × 20 cm pots, filled with TYPical Brill soil (Brill, Germany) and grown at 16–20°C up to the boot stage (ZS 51), 20–24°C during heading and anthesis (ZS 53-69), and 24 –29°C up to maturity (ZS 71-99). The plants were fertilized using ammonium nitrate in the following proportions and at the following stages: 20% at sowing (ZS 00), 40% at tillering (ZS 20), and 40% at heading (ZS 49) (Francesconi et al., 2019). The highly virulent and mycotoxin-producing isolate of *F. graminearum* wild type (WT) 3824 (Tomassini et al., 2009) was cultured at 21°C on synthetic nutrient-poor agar (SNA) to obtain macroconidia (Mandalà et al., 2019). After 10 days on SNA, the conidia were scraped with a glass rod after pipetting 1 mL of sterile distilled water onto the surface of a Petri dish. The conidial suspension was recovered, and the concentration was adjusted to 1×10^5 conidia/mL using a Thoma chamber (0.100 mm depth and 0.0025 mm^2). The inoculum was prepared in sterile distilled water supplemented with 0.05% (v/v) of Tween-20. Plants were subjected to three different treatments: (i) drought stress treatment where plants did not receive water from ZS 51 to 65 (Zadoks et al., 1974; Tambussi et al., 2000); (ii) artificial inoculation treatment where spikes were uniformly spray-inoculated at ZS 65 with a suspension of 1×10^5 conidia/mL; and (iii) mock treatment where spikes were uniformly sprayed with a suspension of Tween-20 0.05% resuspended in sterile distilled water at ZS 65. The spikes were covered with clear plastic bags for 24 h to maintain high humidity levels (>80%). Spikes subjected to drought and mock treatments were sampled after removing the plastic bags, while the inoculated spikes were sampled 24, 48 and 72 hours post inoculation (hpi) to investigate an early response to *F. graminearum*. Collected spikes were immediately stored in liquid nitrogen at −80°C until the extraction of RNA. FHB severity (%) was monitored in the greenhouse by counting the number of bleached spikelets and the total number of spikelets for each spike from 3 to 21 days post inoculation (dpi). In addition, spike temperature and photosynthetic efficiency were recorded at 24, 48, and 72 hpi. The experimental design for the greenhouse trial is illustrated in **Figure 1G**. Data were obtained from three independent experiments, each experiment consisting of 20 spikes for each treatment.

RNA Extraction and cDNA Synthesis

Wheat spikes were ground with mortar and pestle in liquid nitrogen until a fine powder was obtained. The RNA was extracted from 100 mg of powder following the instructions provided by InviTrap® Spin Plant RNA Mini Kit (Stratec Molecular GmbH, Germany), resuspended in RNase-free sterile distilled water, immediately poured onto ice and quantified with Qubit™ fluorometer 1.01 (Invitrogen, United States) using the Qubit™ RNA BR Assay Kit (Thermo Fisher Scientific, United States). To confirm the total quantity and integrity of the RNA, 5 μL of the extracted RNA sample was subjected to a 10-min thermal shock at −80°C, followed by 5 min at 65°C and run on 1.5% denaturing agarose gel. The synthesis of cDNA was performed using 500 ng of RNA following the instructions provided by Xpert cDNA Synthesis Supermix with a gDNA eraser (GriSP Research Solutions, Portugal) in a final volume of 20 μL. To ensure that the synthesis of the cDNA and the elimination of the gDNA had succeeded, a reverse transcription PCR (RT-PCR) of *T. aestivum* Actin (*TaACT*) (containing an intron in the amplified sequence) was performed following the instructions provided by GoTaq® Green Master Mix (Promega, United States) in a total volume of 10 μL. The amplification conditions consisted of: (i) an initial denaturation step of 2 min at 95°C; (ii) 35 cycles of 30 s denaturation at 95°C; (iii) 40 s of annealing at 60°C; (iv) 30 s of elongation at 72°C; and (v) a final elongation step of 5 min at 72°C. The amplification run included a no-template control (NTC) and a genomic DNA (gDNA) control. The amplicons were visualized on 1.5% agarose gel.

Gene Expression by Quantitative Real-Time PCR

Supplementary Table 1 shows the list of target genes, their functions, and the corresponding primer pairs used to perform RT quantitative PCR (RT-qPCR) (Francesconi and Balestra, 2020). Briefly, the primer pair for *T. aestivum* glyceraldeyde-3-phosphate dehydrogenase (*TaGAPDH*) amplification is from Jarošová and Kundu (2010), for *T. aestivum* pathogenesis related protein 1 (*TaPR1*) from Lu et al. (2006), for *TaACT* from Tundo et al. (2016), and *T. aestivum*β-tubulin2 (*TaTUB*) and *T. aestivum* ferredoxin-NADP(H)-oxidoreductase (*TaFNR*) from Tenea et al. (2011). The remaining primers are from Francesconi

and Balestra (2020). The amplification efficiency (E) of RT-qPCR was determined for each primer pair as follows: five 1:10 serial dilutions (1:1-1:10000) were obtained for each cDNA and amplified in four replicates. E and coefficient of determination (R^2) values were calculated by means of the slope of the standard curve obtained by plotting the fluorescence versus the serial dilution concentrations using the equation (Bustin et al., 2009)

$$E \ 10^{\left(-\frac{1}{slope}\right)} - 1$$

Reference genes with closest E values to target genes, highest R^2, and lowest variability were selected for the quantification cycles (Cq). The relative expression levels of target genes were calculated on the basis of the Cq values of four independent biological replicates, each with four technical replicates, for each plant treatment by applying the equation (Bustin et al., 2009)

$$Relative\ expression \ 2^{-\triangle\triangle Cq}$$

using *TaACT*, *TaTUB,* and *TaFNR* as reference genes and the mock treatment to normalize the relative expression levels. Relative expression levels of *TaPR1* and *TaGAPDH* were quantified as internal control of the progression of the infection (Muthukrishnan et al., 2001) and changes in photosynthesis (Zhang et al., 2013). The RT-qPCR was performed following the instructions provided by Rotor-Gene Q (Qiagen, Germany) and Xpert Fast SYBR (uni) MasterMix (GRiSP Research Solutions, Portugal), in a final volume of 10 μL. The amplification conditions consisted of: (i) an initial denaturation step of 3 min at 95°C; (ii) 40 cycles of 5 s denaturation at 95°C; (iii) 30 s of annealing at 60°C; and (iv) 20 s of elongation at 72°C. A final melt cycle (70–99°C) was performed to confirm the unicity of the amplicons. NTC controls were included and the amplification was considered negative when a value of Cq ≥ 38 was detected (Bustin et al., 2009).

Photosynthetic-Related Parameters Measurements

Spike temperature and photosynthetic efficiency were measured for each plant treatment as performed during the field trial, described in section "Ground Measurements." Photosynthetic-related parameters were measured for three independent replicates, each consisting of 20 individual spikes, for every treatment.

Statistical Analyses

One-way analysis of variance (ANOVA) was performed to analyze FHB severity, ground-measurements (temperature and photosynthetic efficiency) of FHB+ and FHB- spikes during the UAV-campaigns, UAV-based VEG, GLI and temperature of FHB+ and FHB- sampling areas during the UAV-campaigns, gene expression values, and temperature and photosynthetic efficiency of drought stressed and inoculated plants during the greenhouse experiments. One level of significance ($p < 0.01$) was computed to assess the significance of the F values. A pairwise analysis was carried out using Tukey's honest significant difference (HSD) test at 0.99 confidence level. Statistical analyses were performed using XLSTAT 2020.4 software (Addinsoft, France). Principal component analysis (PCA) was carried out to classify spike temperature, photosynthetic efficiency and VEG or GLI or UAV-based temperature values coming from FHB+ or FHB- areas and gene expression values, spike temperature and photosynthetic efficiency coming from drought stressed or *F. graminearum* inoculated plants. Heatmap was carried out by computing the *z*-score of the relative gene expression values. PCA and heatmap were computed by using ClustVis software (Tartu, Estonia)[3] (Metsalu and Vilo, 2015).

RESULTS

Weather Conditions Influencing the FHB Severity

Registered data were compared to historical data (2010–2018) (**Figure 2A**), showing that May 2020 was particularly hotter than 2010–2019 (the recorded average temperatures were 20°C, 14°C, and 18°C in May 2020, May 2019 and May 2010–2018, respectively), while May 2019 was characterized by high daily average rainfall (7 mm) compared to May 2020 (2 mm) and the 9-year seasonal average (5 mm). These conditions favored the naturally occurring FHB in the fields of interest. In fact, in 2019, FHB severity (**Figure 2B**) reached 59% and 92% at ZS 65 and 71, respectively, indicating that the wet season (mean relative humidity was > 70%) was particularly favorable for the FHB to spread. Although April and May 2020 were not characterized by frequent rains (average precipitation of 2 mm), low precipitations favored a moderate FHB infection with a severity of 27% and 68% at ZS 65 and 71, respectively.

Ground-Based Measurements During the UAV Campaigns

Figure 3 illustrates the results obtained by ground-based measurements during the UAV campaigns. The molecular identification of FHB was performed by amplifying the *TEF* sequence from a bulk sample obtained from the samples collected in each sampling area, producing an amplicon of 700 bp. At ZS 57, all the sampling areas resulted FHB- in both 2019 and 2020; at ZS 65, 10 of the 16 sampling areas were FHB+ in 2019, while 6 sampling areas were FHB+ in 2020; at ZS 71, all the sampling areas were FHB+ in both 2019 and 2020 (**Figure 3A**). In 2019, thirty *Fusarium* morphotypes were isolated: ten morphotypes were *F. graminearum*, eleven were *F. poae*, eight were *F. avenaceum*, and one was *F. proliferatum*. In 2020, twenty-four morphotypes were identified: ten were *F. graminearum*, ten were *F. poae* and four were *F. avenaceum*. The isolated morphotypes and the data resulted from the BLASTn analyses are listed in **Supplementary Table 2**. Ground-based spike temperature values recorded in 2019 and 2020 revealed that FHB+ spikes had a higher temperature than the FHB- ones at ZS 65 (**Figures 3B,D**). On the other hand, photosynthetic efficiency had an inverse relationship with the FHB severity: variable

[3]https://biit.cs.ut.ee/clustvis/

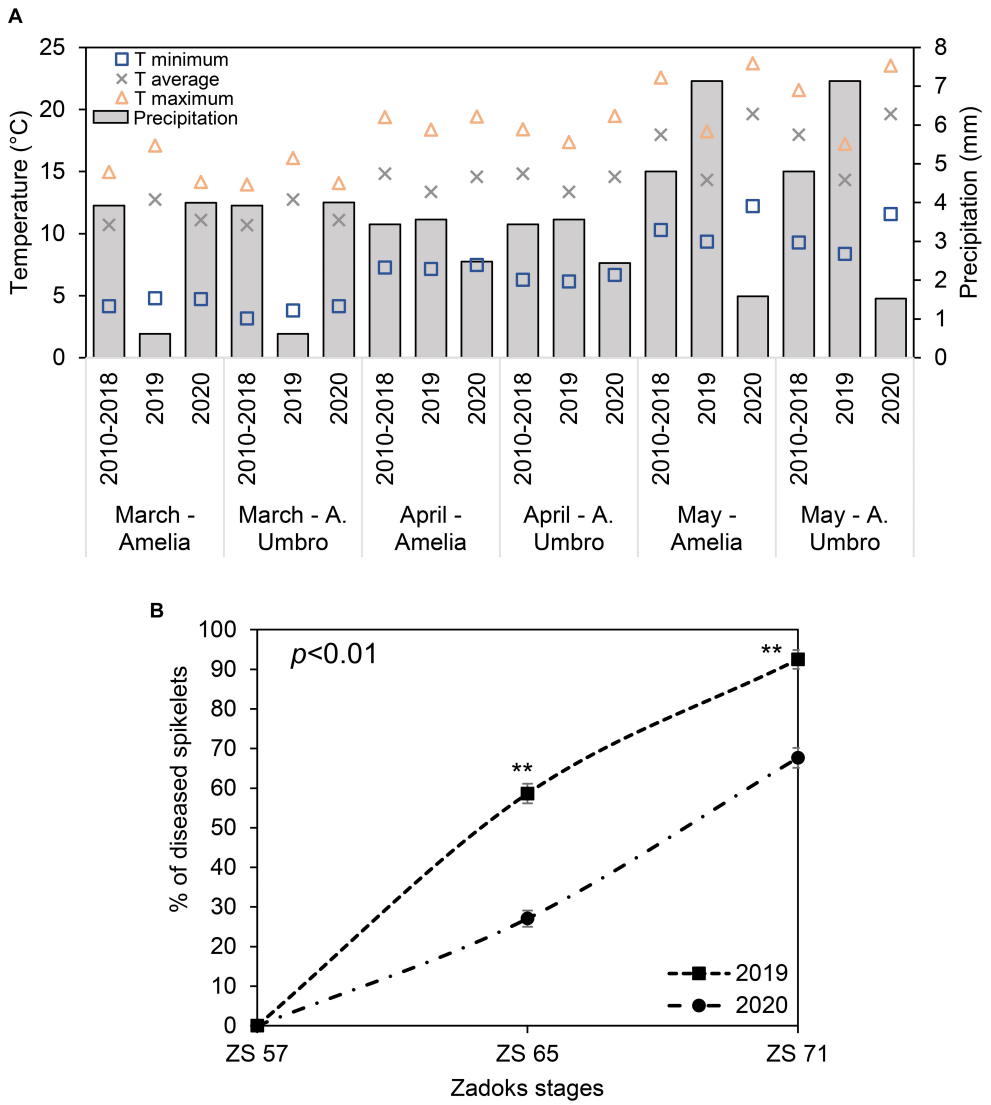

FIGURE 2 | (A) Changes in maximum (yellow triangles), average (grey crosses) and minimum (blue squares) temperature (T), and precipitation (simple bar) of weather data recorded in 2019 in comparison with weather data recorded in 2020 and historical weather data from 2010 to 2018 for March, April, and May. The plotted values were obtained by averaging the mean daily values for each considered period. Weather data for 2019 and 2020 were recorded daily by two meteorological stations installed at 100 m distance from each of the experimental fields located in Amelia (Central Italy, 42°31′22.9″N, 12°25′15.5″E, Umbria Region) and Avigliano Umbro (Central Italy, 42°40′41.1″N, 12°27′44.6″E, Umbria Region). Historical weather data (2010–2018) were collected from the Hydrographic service of Umbria Region (https://annali.regione.umbria.it/). The historical weather data were collected from two meteorological stations located in Amelia (42°33′25.0″N, 12°25′01.0″E) and Avigliano Umbro (42°40′39.0″N, 12°26′13.0″E). **(B)** Severity percentage of Fusarium head blight (FHB) in *Triticum turgidum* (cv. Marco Aurelio) at Zadoks stage (ZS) 57, 65, and 71 in 2019 and 2020. Data represent averages and standard errors of 256 spikes (16 spikes for each of 16 visible targets). Asterisks (**) refer to the statistical analyses performed using one-way analysis of variance (ANOVA) with Tukey's honest significant difference (HSD) post hoc test at 0.99 confidence level and $p < 0.01$.

fluorescence/maximum fluorescence (F_v/F_m) demonstrated to be lower in FHB+ than in FHB- at ZS 65 (**Figures 3C,E**).

UAV-Based TIR and RGB Imaging for FHB Detection

Recorded weather data indicated that during the two UAV-campaigns conducted at ZS 65, the average daily air temperature and humidity values were 15°C (2019), 21°C (2020), and 67% (2019) and 72% (2020), respectively. **Supplementary Figure 2** shows the 16 sampling areas resulting from RGB (A and C) and thermal images (B and D) in Amelia (**Supplementary Figures 2A,B**) and Avigliano Umbro (**Supplementary Figures 2C,D**). PCA (**Figure 4**) demonstrated that VEG (A), GLI (B), and UAV-based temperatures (C) distinguished between FHB+ and FHB- plants at ZS 65 for both the 2019 and 2020 campaigns. Moreover, box-plots indicate that VEG (D), GLI (E), and UAV-based temperature values significantly differed

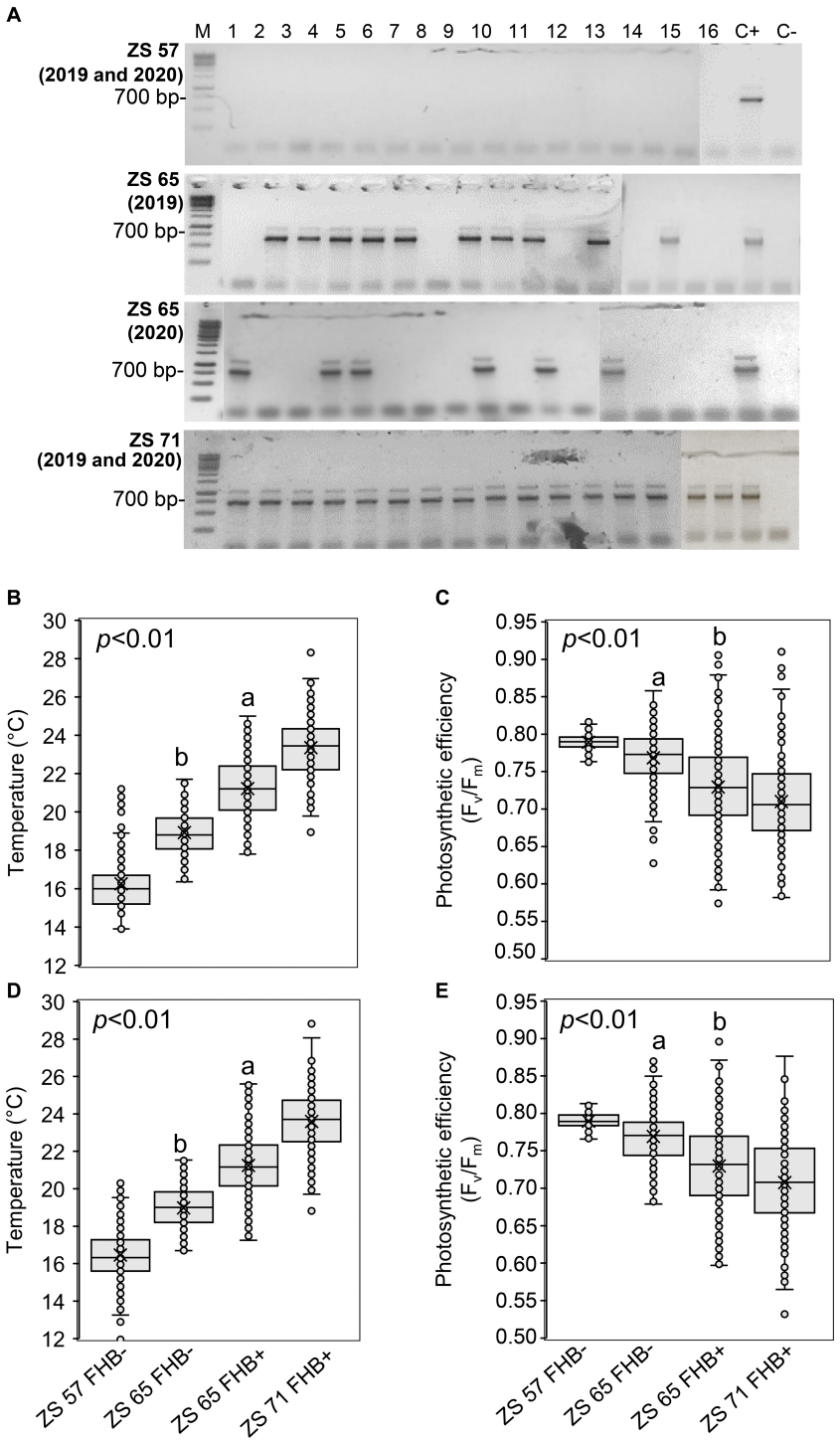

FIGURE 3 | Molecular identification of Fusarium head blight (FHB) and ground-based temperature and photosynthetic efficiency measurements recorded in concurrence with unmanned aerial vehicle (UAV) campaigns conducted at Zadoks stage (ZS) 57, 65, and 71 of *Triticum turgidum* (cv. Marco Aurelio) in 2019 and 2020. **(A)** Molecular identification of *Fusarium* spp. using 1.5% agarose gel of the translational elongation factor 1-α (*TEF*) sequence (700 bp) in spikes of isolates obtained from the 16 sampling areas, sampled in 2019 and 2020 at ZS 57, 65, and 71. At ZS 57, none of the 16 sampling areas resulted FHB positive (FHB+) during both 2019 and 2020 UAV campaigns; at ZS 65, 10 and 6 sampling areas were FHB+ in 2019 and 2020, respectively; at ZS 71, all sampling areas were FHB+ during both 2019 and 2020 UAV campaigns. M represents a 100 bp DNA Ladder (Jena Bioscience); C- represents the negative control and C+ represents the presence of *F. graminearum* wild type (WT) 3824. The figure is obtained from four gels and the original pictures of the gels are available upon request. **(B,D)** Box-plot of the spikes' temperature (°C) and **(C,E)** photosynthetic efficiency (F_v/F_m). The data represent averages and standard errors for 256 spikes (16 spikes for each sampling area) from 2019 **(B,C)** and 2020 **(D,E)**. Different letters refer to the statistical analysis performed using one-way analysis of variance (ANOVA) with the Tukey's honest significant difference (HSD) post hoc test at 0.95 or 0.99 confidence level and $p < 0.05$ or $p < 0.01$.

between FHB+ and FHB− areas during the two UAV-campaigns (2019 and 2020). The average VEG values (D) recorded in FHB− areas were 1.46 and 1.42, while in FHB+ areas were 1.00 and 1.02 in 2019 and 2020, respectively. The average GLI values (E) recorded in FHB− areas were 0.21 and 0.13, while in FHB+ were 0.09 and 0.10 in 2019 and 2020, respectively. The average UAV-based temperature values (F) recorded in FHB− areas were 18.32°C and 18.94°C, while in FHB+ were 21.06°C and 20.94°C in 2019 and 2020, respectively.

Monitoring of FHB in the Greenhouse

The progress of FHB severity was monitored in *F. graminearum*-inoculated plants in the greenhouse from 3 to 21 dpi (**Figure 5A**). The severity gradually increased reaching values close to 100% at 17 dpi, confirming the susceptibility of *T. turgidum* cv. Marco Aurelio. **Figure 5B** shows the FHB symptoms at 21 dpi. While no symptoms were observed on the mock treatment, few non-necrotic bleached spikelets appeared on drought-stressed plants.

Monitoring Spike Temperature and Photosynthetic Efficiency Between Mock, Drought, and Inoculated Treatments

Figure 6 shows the temperature of spikes (**Figure 6A**) and the photosynthetic efficiency (**Figure 6B**) after the three treatments. Compared to the mock treatment, the temperature increased and the photosynthetic efficiency decreased in drought-stressed and inoculated plants, confirming a perturbation of the photosynthetic activity. In fact, the more the infection progressed, the higher the differences in spike temperature and photosynthetic efficiency between the inoculated and mock plants: inoculated spikes reached a temperature of 17.89, 18.48, and 18.67°C while mock spikes measured 17.12, 16.44, and 16.19°C, at 24, 48 and 72 hpi, respectively; the photosynthetic efficiency measured 0.719, 0.695 and 0.618 F_v/F_m for the inoculated spikes and 0.794, 0.793 and 0.801 F_v/F_m for the mock at 24, 48 and 72 hpi, respectively. Notably, *F. graminearum* infection perturbated the photosynthetic parameters more than drought stress, highlighting that, by using these measures, it is possible to distinguish between drought-stressed and FHB−infected plants.

Expression Pattern of the Genes Regulating Stomatal Conductance by RT-qPCR

E, R^2, and the stability of the reference genes were calculated to validate the RT-qPCR results. E ranged from 0.9652 to 1.2741 and R^2 from 0.9651 to 0.9954. Standard errors (SE) among the Cq values of the three reference genes ranged from 0.198 to 0.369 indicating their stable expression under the three different treatments.

Since metrics derived from TIR and RGB images allowed the detection of infected spikes at ZS 65, the greenhouse experiments were designed to investigate the differential stomatal regulation response in proximity of the same phenological stage.

Moreover, drought-stressed plants were studied to observe their gene expression differences with plants under *F. graminearum* inoculation. **Figure 7A** shows a heatmap of the relative expression values of plant genes under drought stress and *F. graminearum* inoculation at 24, 48, and 72 hpi. **Supplementary Table 3** provides the relative expression values, SE, and the HSD test computed at 0.99 confidence level. Under terminal drought stress, *T. aestivum* allene oxide synthase (*TaAOS*), *T. aestivum* terpene synthase (*TaKSL*), *T. aestivum* mitogen-activated protein kinases (*TaMAPK*), *T. aestivum* calcium dependent protein kinase (*TaCDPK*), *T. aestivum* phosphatase (*TaABI*), *T. aestivum* MYB domain transcription factor (*TaPIMP*), *T. aestivum* NADPH oxidase (*TaRBOH*), and *T. aestivum* zeaxanthin epoxidase (*TaZEP*) were slightly up-regulated showing expression values ranging from 1.254-fold to 1.892-fold. Among the different time-points of *F. graminearum* inoculation (24, 48, and 72 hpi), the relative expression values of *TaAOS, T. aestivum* abscisic acid (ABA) aldehyde oxidase (*TaAAO*), *T. aestivum* ABA receptor (*TaREC*), *T. aestivum*β-1,3-glucanase (*TaBG*), *TaMAPK, TaCDPK, T. aestivum* epoxycarotenoid dioxygenase (*TaNCED*), *TaRBOH,* and *TaZEP* gradually increased from 24 to 72 hpi, while *T. aestivum* hydroperoxide lyase (*TaHPL*) and *T. aestivum* cytochrome P450 (*TaCYP450*) were down-regulated. Notably, the expression patterns of *TaKSL, TaAAO, TaREC, TaBG, TaCYP450, TaNCED,* and *TaZEP* were different between drought-stressed and *F. graminearum*-inoculated plants. In particular, in the inoculation treatment, at 72 hpi *TaREC, TaBG, TaNCED,* and *TaZEP* were strongly up-regulated (5.729, 5.143, 4.988, and 4.256-fold change, respectively) while *TaPR1* and *TaGAPDH* were gradually up-regulated from 24 to 72 hpi, indicating that the *F. graminearum* infection perturbated the innate immunity and physiological photosynthesis of the plants. **Figure 7B** represents the PCA obtained by computing relative gene expression values, spike temperature and photosynthetic efficiency values from drought-stressed and *F. graminearum* inoculated plants. PCA demonstrated that these data distinguished between hydric stress and *F. graminearum* inoculation.

DISCUSSION

Changes in temperature and color of spikes are a result of the physiological defensive response of *T. turgidum* to FHB. Indeed, a thickening of the vascular bundles occurs when the infection moves from the floret to the rachilla which causes an increase in temperature, a decrease in photosynthesis efficiency, and a reduced transpiration due to limited water supply and stomatal closure (Kang and Buchenauer, 2000; Kheiri et al., 2019). These physiological changes allow remote sensing techniques to detect and quantify FHB in *T. turgidum* non-destructively.

Previous studies have revealed the potential of remote sensing methods in detecting and assessing plant diseases. For example, Oerke and Steiner (2010) detected FHB using thermal imaging only at a late infection stage, when it was too late for the disease to be controlled. Al Masri et al. (2017) studied the effect of the primary infection site by *F. graminearum* and *F. culmorum* using

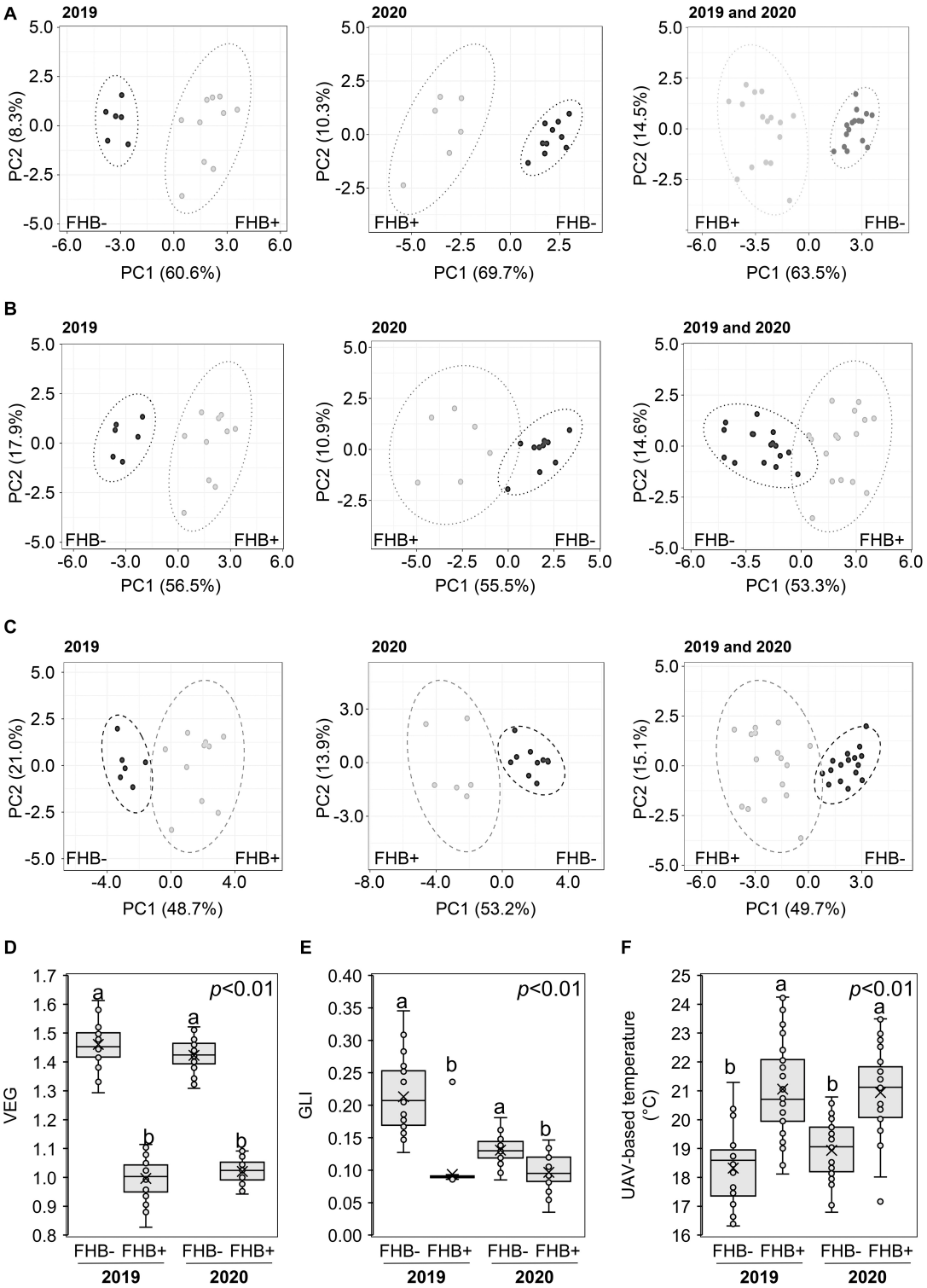

FIGURE 4 | Principal component analysis (PCA) of unmanned aerial vehicle (UAV)-based **(A)** vegetative index (VEG), **(B)** green leaf index (GLI), **(C)** temperature at Zadoks stage (ZS) 65. PCA was performed by using ClustVis Software for $p < 0.05$ to distinguish between FHB+ and FHB– areas. Box-plot of UAV-based **(D)** VEG, **(E)** GLI, and **(F)** temperature from FHB– and FHB+ areas. Data were recorded during 2019 and 2020. The data represent averages and standard errors for four measurements for each sampling area from 2019 **(B,C)** and 2020 **(D,E)**. Different letters refer to the statistical analysis performed using one-way analysis of variance (ANOVA) with the Tukey's honest significant difference (HSD) post hoc test at 0.95 or 0.99 confidence level and $p < 0.05$ or $p < 0.01$. Only the data deriving from 1 year were compared.

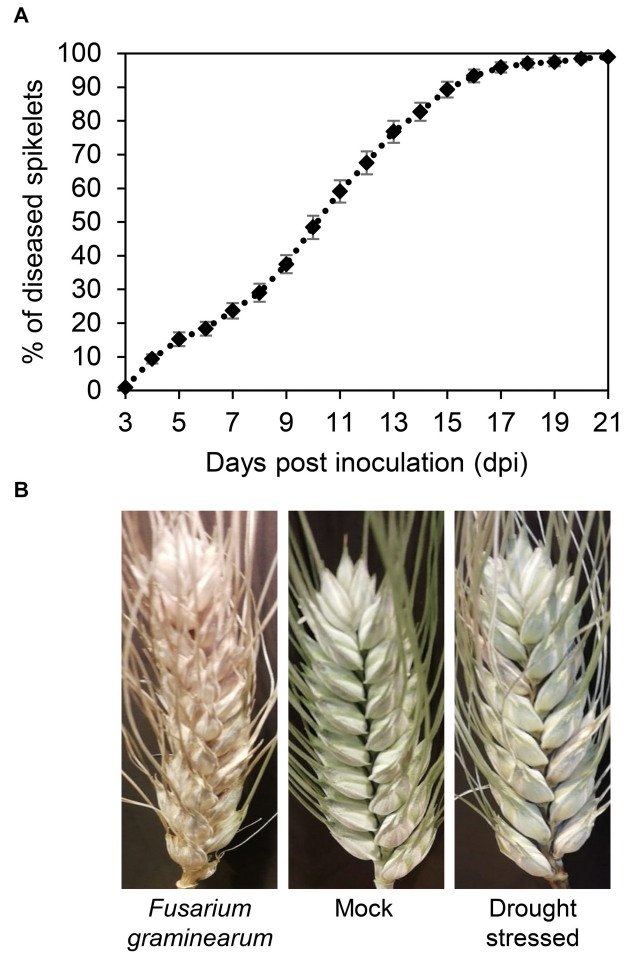

FIGURE 5 | (A) Fusarium head blight (FHB) severity (%) in *Triticum turgidum* (cv. Marco Aurelio) from 3 to 21 days post inoculation (dpi) during the greenhouse experiments. **(B)** FHB symptoms in *T. turgidum* at 21 dpi. The data was derived from the averages and standard errors of three treatments (*F. graminearum*-inoculated, drought-stressed, and mock) with at least 20 plants for each, and three independent experiments for each treatment.

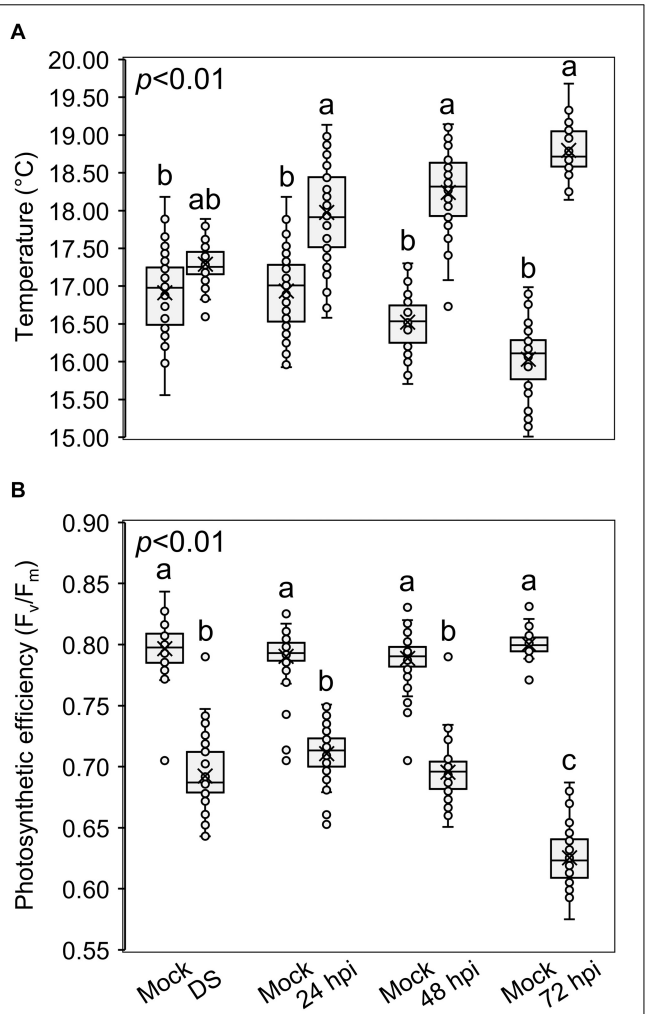

FIGURE 6 | Box-plots of **(A)** spike temperatures (°C) and **(B)** photosynthetic efficiencies (F_v/F_m) of *Triticum turgidum* (cv. Marco Aurelio) plants subjected to mock (M), drought stress (DS) and *Fusarium graminearum* artificial inoculation (24, 48, and 72 hours post inoculation (hpi)) treatments during the greenhouse experiments. The data was derived from the averages and standard errors of three treatments (*F. graminearum*-inoculated, drought-stressed, and mock) with at least 20 plants for each, and three independent experiments for each treatment. Different letters refer to the statistical analysis performed using one-way analysis of variance (ANOVA) with the Tukey's honest significant difference (HSD) post hoc test at 0.99 confidence level and $p < 0.01$.

thermography under controlled conditions and they observed that FHB infection significantly increased the temperature of spikes as it progressed from 6 to 29 dpi. Mahlein et al. (2019) demonstrated that *Fusarium*-infected spikelets showed higher temperatures and lower F_v/F_m values compared to mock control using a digital thermo-camera and a chlorophyll fluorometer. The authors recorded temperature and F_v/F_m values in entire spikes to detect FHB infection at 5 and 7 dpi.

Red-green-blue imaging was widely employed to detect FHB-infected and FHB-damaged kernels (Jaillais et al., 2015; Cambaza et al., 2019; Abbaspour-Gilandeh et al., 2020), but few studies explored this technique to detect FHB on spikes. Huang et al. (2020) proposed an FHB diagnostic model of disease severity based on the fusion of RGB and spectral imaging. The results showed that the model was able to identify FHB severity in plants with an accuracy of 92%, thereby providing a technical basis for timely and effective control of FHB. Dammer et al. (2011) made use of RGB imaging to detect FHB in the field. Experimental plants were artificially infected with a spore suspension and RGB images were captured and analyzed to detect the disease symptoms. The authors found a linear correlation between RGB-derived and visually observed disease levels in plants. Qiu et al. (2019) accurately detected FHB in the field using RGB imaging. RGB-derived data correlated with the number of diseased spikes tallied by manual count. These results are in agreement with those obtained in our study since we observed an increase in temperature and a decrease in photosynthetic efficiency in FHB-infected spikes. Moreover, UAV-based measurements distinguished FHB+ and FHB− areas, confirming that TIR and

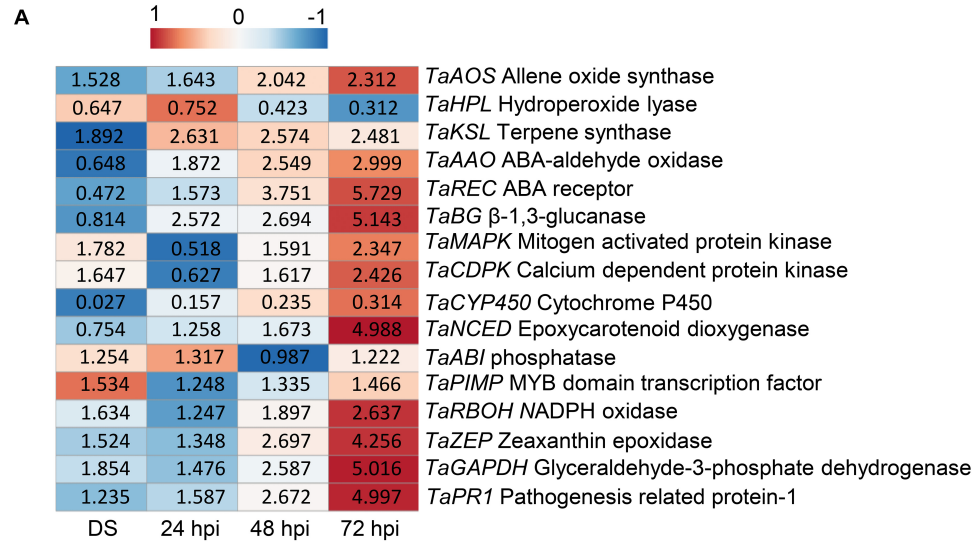

FIGURE 7 | (A) Heatmap of relative expression level of the genes selected in *Triticum turgidum* (cv. Marco Aurelio) exposed to drought stress (DS) and *Fusarium graminearum*-inoculated at 24, 48, and 72 hours post inoculation (hpi)) treatments. The expression values were normalized to the mock treatment and to *T. aestivum* actin (*TaACT*), *T. aestivum*β-tubulin2 (*TaTUB*), and *T. aestivum* ferredoxin-NADP(H)-oxidoreductase (*TaFNR*) as reference genes. The heatmap was constructed by plotting the z-score of the relative gene expression values and it was generated by analyzing data with ClustVis Software. The red color represents the up-regulated genes, while the blue color the down-regulated genes. **(B)** Principial component analysis (PCA) of relative gene expression values, spike temperature and photosynthetic efficiency to distinguish between drought stress and *F. graminearum* infection during the greenhouse experiments. PCA was performed by using ClustVis Software for $p < 0.05$.

RGB imaging are powerful tools for FHB detection. To the best of our knowledge, for FHB detection, TIR and RGB cameras have only been employed on ground-based phenotyping platforms, minimizing their portability and limiting the scale at which they can be used. Thus, this is the first study demonstrating that UAV-mounted TIR and RGB cameras enable rapid characterization of *T. turgidum* and detection of FHB in the field, overcoming the limitations associated with ground-based phenotyping.

In the present study, the relative expression level of genes involved in stomatal regulation was evaluated at ZS 65 to elucidate the genetic mechanism responsible for the phenotypic response to perturbation of photosynthesis, and to establish whether or not a differential gene response exists between drought-stressed and FHB-infected *T. turgidum*. Stomatal closure is the primary response of plants to water deficit, controlled by abscisic acid (ABA), a key hormone involved in

controlling many aspects of plant growth, development, and responses to a variety of biotic and abiotic stresses (Daszkowska-Golec and Szarejko, 2013; Duarte et al., 2019). Our results are in agreement with the literature since the majority of stomatal closure positive regulating genes involved in ABA biosynthesis (*TaKSL, TaZEP, TaCDPK, TaMAPK, TaRBOH,* and *TaPIMP*) were induced, while the negative regulators (*TaCYP450, TaBG,* and *TaREC*) were down-regulated after being exposed to drought stress. In contrast, some positive regulators of stomatal closure (*TaNCED, TaAAO,* and *TaHPL*) were down-regulated while a negative regulator (*TaABI*) was induced. Our findings support the hypothesis stating that hydric stress conditions do not completely induce stomatal closure in drought-tolerant wheat varieties, which correlate with lower level of closure-inducing genes and higher expressions of genes negatively regulators of stomatal closure (Xue et al., 2006; Ji et al., 2011; Rampino et al., 2012; Gallé et al., 2013). Moreover, recorded temperature values of spikes in the greenhouse revealed that Marco Aurelio is moderately tolerant to drought stress since temperature measures of plants exposed to drought stress did not significantly differ from the mock. However, amongst the distinctive responses between drought-stressed and *F. graminearum*-inoculated plants, *TaAAO, TaREC, TaBG,* and *TaNCED* were down-regulated in the former and up-regulated in the latter. On the other hand, *TaMAPK* and *TaCDPK* were up-regulated in drought-stressed but not in *F. graminearum*-inoculated plants at 24 hpi. Spikes temperature and photosynthetic efficiency values of *F. graminearum*-inoculated plants differed significantly from the mock and notably, at 72 hpi, the photosynthetic efficiency allowed the distinction between *F. graminearum*-inoculated and drought-stressed plants. These observations can be extremely helpful to develop further methodologies aimed at distinguishing between drought-stressed and FHB-infected plants in the field. In this regard, biotic and abiotic stresses need to be distinguished in order to optimize practical field management. Shaik and Ramakrishna (2014) segregated biotic and abiotic stresses in rice by applying machine learning approaches to the expression levels of a set of stress-responsive genes. Focusing on the use of imaging, hyperspectral sensors are the most suitable to distinguish biotic and drought stresses in many crops (Jones, 2011; Susič et al., 2018). To date, most of the research studies distinguish between drought and disease infection applied separately, while Ramegowda and Senthil-Kumar (2015) amply reviewed experimental evidence suggesting that, under combined drought and biotic stress, plants exhibit tailored physiological and molecular responses. Such tailored responses occur only in plants exposed to simultaneous stresses and such information cannot be inferred from individual stress studies.

Additionally, in *F. graminearum*-inoculated plants, most of the positive regulators of stomatal closure (*TaAOS, TaKSL, TaAAO, TaNCED, TaPIMP, TaRBOH,* and *TaZEP*) were induced from 24 to 72 hpi while *TaMAPK* and *TaCDPK* were up-regulated at 48 and 72 hpi, confirming that early stomatal closure is the physiological mechanism behind the increasing temperature and decreasing photosynthetic efficiency in spikes. The negative stomatal closure regulators *TaBG* and *TaREC* were also remarkably up-regulated, while *TaCYP450* was down-regulated. Our results are in agreement with the literature data, reporting the induction of *TaBG* and *TaREC* and the down-regulation of *TaCYP450* in FHB-susceptible wheat cultivars. *TaBG* belongs to the pathogenesis-related proteins family (PR2) in wheat, which is known to be induced as a defense mechanism in response to biotic and abiotic stresses (Muthukrishnan et al., 2001). Particularly, De Zutter et al. (2017) investigated the *T. aestivum* response to a combined attack of *F. graminearum* and *Sitobion avenae* aphids, and observed the consistent up-regulation of *PR1* and *PR2*. Another study (Francesconi and Balestra, 2020) demonstrated that *TaPR1* and *TaPR2* were induced in an *F. graminearum*-susceptible *T. aestivum* (cv. Rebelde) but not as much as in the FHB-resistant *T. aestivum* (cv. Sumai3). The up-regulation of *TaREC* could be explained by evidence supporting that it may be involved in FHB susceptibility, since Gordon et al. (2016) found that *REC* silencing in *T. aestivum* (cv. Chinese Spring) resulted in slower progression of FHB symptoms and decreased DON content in wheat heads. On the other hand, *TaCYP450* was down-regulated in *F. graminearum*-infected *T. turgidum*. In fact, much evidence indicated that *CYP450* plays an active role in wheat resistance against FHB and DON accumulation. Strong *CYP450* accumulations were found in *F. graminearum*- and DON-resistant but not in susceptible wheat cultivars (Li et al., 2010; Gunupuru et al., 2018; Francesconi and Balestra, 2020). Several studies also demonstrated that *CYP450* was able to detoxify DON *in vitro* (Ito et al., 2013) and *in vivo* (Gunupuru et al., 2018).

The present study proved that UAV-based TIR and RGB image analysis can detect FHB infections at ZS 65. This can improve different aspects of FHB management and plant breeding. For example, our methodology allows timely detection of FHB and mapping affected locations in the field, thus optimizing the application timing and amount of fungicides needed to control the disease (Oerke and Steiner, 2010). It can also provide valuable information about the severity of FHB and help meet future food traceability requirements. Indeed, the ability to monitor FHB severity before further processing of harvested kernels can help determining whether the grains fit for human or animal consumption, with special regard to mycotoxin content (Dammer et al., 2011). For such purpose, the image-assisted analysis coupled with prediction modeling could be a valuable method to predict and detect the accumulated mycotoxin in the grains (Battilani, 2016; Leplat et al., 2018; Fernando et al., 2021). Several research studies reported also the accumulation of mycotoxin in absence of macroscopic symptoms, while microscopic analysis revealed that the host cells drastically changed after the infection (Brown et al., 2010; Peiris et al., 2011; Alisaac et al., 2021). For such reasons, the detection of mycotoxin in asymptomatic spikes could be successfully achieved by using multispectral imaging (Bauriegel et al., 2011; Dammer et al., 2011; Leplat et al., 2018) to support the mycotoxin traceability performed by the costly techniques based on chromatography (Tittlemier et al., 2021). Furthermore, the presented methodology can help quantifying host resistance to FHB in pre-breeding and commercial breeding trials (Yang et al., 2017), speeding-up breeding programs.

CONCLUSION

The rapid detection of FHB is a key factor to gain maximum, environmentally sustainable protection of yield. To maximize FHB disease control efficiency, we explored the use of UAV-based RGB and TIR imaging supported by ground-truthing to detect the presence of FHB in *T. turgidum* (cv. Marco Aurelio). The present study revealed that: (i) stomatal closure is the physiological mechanism responsible for temperature increase and photosynthetic efficiency decrease in *T. turgidum* spikes during FHB infection; (ii) VIs and temperatures extracted from RGB and TIR imaging data can detect these physiological changes; and (iii) different transcriptional regulations exist between drought-stressed and *F. graminearum*-inoculated plants. These findings provide mechanisms for the detection of FHB in *T. turgidum* and shed light into new valuable genomic information to further develop a phenotyping method able to distinguish between drought-stressed and FHB-infected plants in the field.

Research in plant stress physiology is benefiting from new types of precision disease management technologies based on phenomics, genomics, and transcriptomics data. In the last decade, plant genomics and phenomics have matured to the point where, applied together, they can drastically reduce bottlenecks in phenotypic and genotypic evaluation of plant traits (Flood et al., 2011; Murchie et al., 2018; van Bezouw et al., 2019) and when coupled with artificial intelligence and exascale computing, they can accelerate the development of new crop varieties with improved yield potential and enhanced tolerance to biotic and abiotic environmental stresses (Harfouche et al., 2019; Streich et al., 2020). Their implementation and application will elucidate the architecture of plant physiological mechanisms to develop innovative tools to be applied in a new green revolution (Ray et al., 2013). To date, no studies have been carried out attempting to use UAV-based TIR and RGB imaging data for the detection of FHB in *T. turgidum*. Developing trait measurement methodologies that combine phenomics and genomics to detect plant diseases can provide a timely warning of their imminent threat, allowing decisions to be made in time for fungicides to be effective, reducing the costs and negative environmental impacts of their unnecessary applications. Further research is needed to test the reproducibility of UAV-based phenomics in different environments and to explore their potentiality to distinguish between biotic and abiotic stresses.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

SF, AH, MM, and GB conceived and designed the experiments. SF, AH, and GB led the writing of the manuscript and figure preparation, and contributed to the majority of the critical revision of the manuscript text and figures. SF led the ground-based experiments, data collection, and analysis. MM led the UAV-based experiment, data collection, and analysis. All authors have revised the work for intellectual content and have contributed to and approved the final content.

FUNDING

This research was supported by the Rural Development Program (PSR) of the Regione Umbria (Umbria Region) – 16.2 "SmartAgri Platform" Project (GB), the Italian Ministry for Education, University and Research (MIUR) (Law 232/216, Department of Excellence), the EU 7th Framework Program – WATBIO, Grant No. 311929 (AH), and the Italian Ministry of Education, University and Research Brain Gain Professorship Program to AH.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Federico V. Moresi for technical assistance during the UAV campaigns. The authors would like to thank the reviewers and the editor for their valuable comments and suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.628575/full#supplementary-material

REFERENCES

Abbaspour-Gilandeh, Y., Ghadakchi-Bazaz, H., and Davari, M. (2020). Discriminating healthy wheat grains from grains infected with Fusarium graminearum using texture characteristics of image-processing technique, discriminant analysis, and support vector machine methods. *J. Intell. Syst.* 29, 1576–1586. doi: 10.1515/jisys-2018-0430

Al Masri, A., Hau, B., Dehne, H. W., Mahlein, A. K., and Oerke, E. C. (2017). Impact of primary infection site of Fusarium species on head blight development in wheat ears evaluated by IR-thermography. *Eur. J. Plant Pathol.* 147, 855–868. doi: 10.1007/s10658-016-1051-2

Alexandersson, E., Jacobson, D., Vivier, M. A., Weckwerth, W., and Andreasson, E. (2014). Field-omics-understanding large-scale molecular data from field crops. *Front. Plant Sci.* 5:286. doi: 10.3389/fpls.2014.00286

Alisaac, E., Rathgeb, A., Karlovsky, P., and Mahlein, A. K. (2021). Fusarium head blight: effect of infection timing on spread of *Fusarium graminearum* and spatial distribution of deoxynivalenol within wheat spikes. *Microorganisms* 9:79. doi: 10.3390/microorganisms9010079

Bai, G., and Shaner, G. (2004). Management and resistance in wheat and barley to *Fusarium* head blight. *Annu. Rev. Phytopathol.* 42, 135–161. doi: 10.1146/annurev.phyto.42.040803.140340

Barbosa, B. D. S., Ferraz, G. A. S., Gonçalves, L. M., Marin, D. B., Maciel, D. T., Ferraz, P. F. P., et al. (2019). RGB vegetation indices applied to grass

monitoring: a qualitative analysis. *Agron. Res.* 17, 349–357. doi: 10.15159/AR.19.119

Battilani, P. (2016). Recent advances in modeling the risk of mycotoxin contamination in crops. *Curr. Opin. Food Sci.* 11, 10–15. doi: 10.1016/j.cofs.2016.08.009

Bauriegel, E., Giebel, A., Geyer, M., Schmidt, U., and Herppich, W. B. (2011). Early detection of *Fusarium* infection in wheat using hyper-spectral imaging. *Comput. Electron. Agric.* 75, 304–312. doi: 10.1016/j.compag.2010.12.006

Berni, J. A. J., Zarco-Tejada, P. J., Suárez, L., and Fereres, E. (2009). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Trans. Geosci. Remote Sens.* 47, 722–738. doi: 10.1109/TGRS.2008.2010457

Blandino, M., Haidukowski, M., Pascale, M., Plizzari, L., Scudellari, D., and Reyneri, A. (2012). Integrated strategies for the control of *Fusarium* head blight and deoxynivalenol contamination in winter wheat. *Field Crop. Res.* 133, 139–149. doi: 10.1016/j.fcr.2012.04.004

Bolanos-Carriel, C., Wegulo, S. N., Baenziger, P. S., Funnell-Harris, D., Hallen-Adams, H. E., and Eskridge, K. M. (2020). Effects of fungicide chemical class, fungicide application timing, and environment on *Fusarium* head blight in winter wheat. *Eur. J. Plant Pathol.* 158, 667–679. doi: 10.1007/s10658-020-02109-3

Bollard-Breen, B., Brooks, J. D., Jones, M. R. L., Robertson, J., Betschart, S., Kung, O., et al. (2015). Application of an unmanned aerial vehicle in spatial mapping of terrestrial biology and human disturbance in the McMurdo Dry Valleys, East Antarctica. *Polar Biol.* 38, 573–578. doi: 10.1007/s00300-014-1586-7

Boutigny, A. L., Richard-Forget, F., and Barreau, C. (2008). Natural mechanisms for cereal resistance to the accumulation of Fusarium trichothecenes. *Eur. J. Plant Pathol.* 121, 411–423. doi: 10.1007/s10658-007-9266-x

Brown, N. A., Urban, M., van de Meene, A. M. L., and Hammond-Kosack, K. E. (2010). The infection biology of Fusarium graminearum: defining the pathways of spikelet to spikelet colonisation in wheat ears. *Fungal Biol.* 114, 555–571. doi: 10.1016/j.funbio.2010.04.006

Brunori, E., Maesano, M., Moresi, F. V., Antolini, A., Bellincontro, A., Forniti, R., et al. (2020). Using UAV —based remote sensing to assess grapevine canopy damages due to fire smoke. *J. Sci. Food Agric.* 100, 4531–4539. doi: 10.1002/jsfa.10494

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55, 611–622. doi: 10.1373/clinchem.2008.112797

Cambaza, E., Koseki, S., and Kawamura, S. (2019). Why RGB imaging should be used to analyze Fusarium graminearum growth and estimate deoxynivalenol contamination. *Methods Protoc.* 2:25. doi: 10.3390/mps2010025

Castrignanò, A., Belmonte, A., Antelmi, I., Quarto, R., Quarto, F., Shaddad, S., et al. (2021). Semi-automatic method for early detection of *Xylella fastidiosa* in olive trees using UAV multispectral imagery and geostatistical-discriminant analysis. *Rem. Sens.* 13:14. doi: 10.3390/rs13010014

Chaerle, L., Leinonen, I., Jones, H. G., and Van Der Straeten, D. (2007). Monitoring and screening plant populations with combined thermal and chlorophyll fluorescence imaging. *J. Exp. Bot.* 58, 773–784. doi: 10.1093/jxb/erl257

Dammer, K. H., Möller, B., Rodemann, B., and Heppner, D. (2011). Detection of head blight (*Fusarium* ssp.) in winter wheat by color and multispectral image analyses. *Crop Prot.* 30, 420–428. doi: 10.1016/j.cropro.2010.12.015

Daszkowska-Golec, A., and Szarejko, I. (2013). Open or close the gate – stomata action under the control of phytohormones in drought stress conditions. *Front. Plant Sci.* 4:138. doi: 10.3389/fpls.2013.00138

De Zutter, N., Audenaert, K., Ameye, M., De Boevre, M., De Saeger, S., Haesaert, G., et al. (2017). The plant response induced in wheat ears by a combined attack of Sitobion avenae aphids and *Fusarium graminearum* boosts fungal infection and deoxynivalenol production. *Mol. Plant Pathol.* 18, 98–109. doi: 10.1111/mpp.12386

Deery, D. M., Rebetzke, G. J., Jimenez-Berni, J. A., James, R. A., Condon, A. G., Bovill, W. D., et al. (2016). Methodology for high-throughput field phenotyping of canopy temperature using airborne thermography. *Front. Plant Sci.* 7:1808. doi: 10.3389/fpls.2016.01808

D'Ovidio, R., and Porceddu, E. (1996). PCR-based assay for detecting 1B-genes for low molecular weight glutenin subunits related to gluten quality properties in durum wheat. *Plant Breed.* 115, 413–415. doi: 10.1111/j.1439-0523.1996.tb00944.x

Duarte, K. E., de Souza, W. R., Santiago, T. R., Sampaio, B. L., Ribeiro, A. P., Cotta, M. G., et al. (2019). Identification and characterization of core abscisic acid (ABA) signaling components and their gene expression profile in response to abiotic stresses in Setaria viridis. *Sci. Rep.* 9:4028. doi: 10.1038/s41598-019-40623-5

Dweba, C. C., Figlan, S., Shimelis, H. A., Motaung, T. E., Sydenham, S., Mwadzingeni, L., et al. (2017). Fusarium head blight of wheat: pathogenesis and control strategies. *Crop Prot.* 91, 114–122. doi: 10.1016/j.cropro.2016.10.002

Fahlgren, N., Gehan, M. A., and Baxter, I. (2015). Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Curr. Opin. Plant Biol.* 24, 93–99. doi: 10.1016/j.pbi.2015.02.006

Fang, Y., and Ramasamy, R. P. (2015). Current and prospective methods for plant disease detection. *Biosensors* 5, 537–561. doi: 10.3390/bios5030537

Feksa, H. R., Do Couto, H. T. Z., Garozi, R., De Almeida, J. L., Gardiano, C. G., and Tessmann, D. J. (2019). Pre- and postinfection application of strobilurin-triazole premixes and single fungicides for control of *Fusarium* head blight and deoxynivalenol mycotoxin in wheat. *Crop Prot.* 117, 128–134. doi: 10.1016/j.cropro.2018.12.003

Feldman, M., and Levy, A. A. (2012). Genome evolution due to allopolyploidization in wheat. *Genetics* 192, 763–774. doi: 10.1534/genetics.112.146316

Fernando, W. G. D., Oghenekaro, A. O., Tucker, J. R., and Badea, A. (2021). Building on a foundation: advances in epidemiology, resistance breeding, and forecasting research for reducing the impact of Fusarium head blight in wheat and barley. *Can. J. Plant Pathol.* 4, 1–32. doi: 10.1080/07060661.2020.1861102

Flood, P. J., Harbinson, J., and Aarts, M. G. M. (2011). Natural genetic variation in plant photosynthesis. *Trends Plant Sci.* 16, 327–335. doi: 10.1016/j.tplants.2011.02.005

Francesconi, S., and Balestra, G. M. (2020). The modulation of stomatal conductance and photosynthetic parameters is involved in Fusarium head blight resistance in wheat. *PLoS One* 15:e0235482. doi: 10.1371/journal.pone.0235482

Francesconi, S., Mazzaglia, A., and Balestra, G. M. (2019). Different inoculation methods affect components of Fusarium head blight resistance in wheat. *Phytopathol. Mediterr.* 58, 679–691. doi: 10.13128/Phyto-10942

Francesconi, S., Steiner, B., Buerstmayr, H., Lemmens, M., Sulyok, M., and Balestra, G. M. (2020). Chitosan hydrochloride decreases *Fusarium graminearum* growth and virulence and boosts growth, development and systemic acquired resistance in two durum wheat genotypes. *Molecules* 25, 4752. doi: 10.3390/molecules25204752

Freije, A. N., and Wise, K. A. (2015). Impact of Fusarium graminearum inoculum availability and fungicide application timing on Fusarium head blight in wheat. *Crop Prot.* 77, 139–147. doi: 10.1016/j.cropro.2015.07.016

Furbank, R. T., and Tester, M. (2011). Phenomics – technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Gallé, Á, Csiszár, J., Benyó, D., Laskay, G., Leviczky, T., Erdei, L., et al. (2013). Isohydric and anisohydric strategies of wheat genotypes under osmotic stress: biosynthesis and function of ABA in stress responses. *J. Plant Physiol.* 170, 1389–1399. doi: 10.1016/j.jplph.2013.04.010

Geiser, D. M., Jiménez-Gasco, M. D. M., Kang, S., Makalowska, I., Veeraraghavan, N., Ward, T. J., et al. (2004). FUSARIUM-ID v. 1.0: a DNA sequence database for identifying *Fusarium*. *Eur. J. Plant Pathol.* 110, 473–479. doi: 10.1023/B:EJPP.0000032386.75915.a0

Giller, K. E., Andersson, J. A., Corbeels, M., Kirkegaard, J., Mortensen, D., Erenstein, O., et al. (2015). Beyond conservation agriculture. *Front. Plant Sci.* 6:870. doi: 10.3389/fpls.2015.00870

Goggin, F. L., Lorence, A., and Topp, C. N. (2015). Applying high-throughput phenotyping to plant-insect interactions: picturing more resistant crops. *Curr. Opin. Insect Sci.* 9, 69–76. doi: 10.1016/j.cois.2015.03.002

Gómez-Candón, D., Virlet, N., Labbé, S., Jolivot, A., and Regnard, J. L. (2016). Field phenotyping of water stress at tree scale by UAV-sensed imagery: new insights for thermal acquisition and calibration. *Precis. Agric.* 17, 786–800. doi: 10.1007/s11119-016-9449-6

Gordon, C. S., Rajagopalan, N., Risseeuw, E. P., Surpin, M., Ball, F. J., Barber, C. J., et al. (2016). Characterization of Triticum aestivum abscisic acid receptors and

a possible role for these in mediating Fusarium head blight susceptibility in wheat. *PLoS One* 11:e0164996. doi: 10.1371/journal.pone.0164996

Gunupuru, L. R., Arunachalam, C., Malla, K. B., Kahla, A., Perochon, A., Jia, J., et al. (2018). A wheat cytochrome P450 enhances both resistance to deoxynivalenol and grain yield. *PLoS One* 13:e0204992. doi: 10.1371/journal.pone.0204992

Hague, T., Tillett, N. D., and Wheeler, H. (2006). Automated crop and weed monitoring in widely spaced cereals. *Precis. Agric.* 7, 21–32. doi: 10.1007/s11119-005-6787-1

Haidukowski, M., Visconti, A., Perrone, G., Vanadia, S., Pancaldi, D., Covarelli, L., et al. (2012). Effect of prothioconazole-based fungicides on *Fusarium* head blight, grain yield and deoxynivalenol accumulation in wheat under field conditions. *Phytopathol. Mediterr.* 51, 236–246. doi: 10.14601/PHYTOPATHOL_MEDITERR-9401

Harfouche, A. L., Jacobson, D. A., Kainer, D., Romero, J. C., Harfouche, A. H., Scarascia Mugnozza, G., et al. (2019). Accelerating climate resilient plant breeding by applying next-generation artificial intelligence. *Trends Biotechnol.* 37, 1217–1235. doi: 10.1016/j.tibtech.2019.05.007

Huang, L., Li, T., Ding, C., Zhao, J., Zhang, D., and Yang, G. (2020). Diagnosis of the severity of Fusarium head blight of wheat ears on the basis of image and spectral feature fusion. *Sensors* 20:2887. doi: 10.3390/s20102887

Infantino, A., Santori, A., and Shah, D. A. (2012). Community structure of the *Fusarium* complex on wheat seed in Italy. *Eur. J. Plant Pathol.* 132, 499–510. doi: 10.1007/s10658-011-9892-1

Ito, M., Sato, I., Ishizaka, M., Yoshida, S. I., Koitabashi, M., Yoshida, S., et al. (2013). Bacterial cytochrome P450 system catabolizing the *Fusarium* toxin deoxynivalenol. *Appl. Environ. Microbiol.* 79, 1619–1628. doi: 10.1128/AEM.03227-12

Jaillais, B., Roumet, P., Pinson-Gadais, L., and Bertrand, D. (2015). Detection of *Fusarium* head blight contamination in wheat kernels by multivariate imaging. *Food Control* 54, 250–258. doi: 10.1016/j.foodcont.2015.01.048

Jarošová, J., and Kundu, J. K. (2010). Validation of reference genes as internal control for studying viral infections in cereals by quantitative real-time RT-PCR. *BMC Plant Biol.* 10:146. doi: 10.1186/1471-2229-10-146

Ji, X., Dong, B., Shiran, B., Talbot, M. J., Edlington, J. E., Hughes, T., et al. (2011). Control of abscisic acid catabolism and abscisic acid homeostasis is important for reproductive stage stress tolerance in cereals. *Plant Physiol.* 156, 647–662. doi: 10.1104/pp.111.176164

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. *Front. Plant Sci.* 10:370. doi: 10.3389/fpls.2019.00370

Jones, H. G. (2011). Remote detection of crop water "stress" and distinguishing it from other stresses. *Acta Hortic.* 922, 23–34. doi: 10.17660/ActaHortic.2011.922.2

Kang, Z., and Buchenauer, H. (2000). Ultrastructural and immunocytochemical investigation of pathogen development and host responses in resistant and susceptible wheat spikes infected by *Fusarium culmorum*. *Physiol. Mol. Plant Pathol.* 57, 255–268. doi: 10.1006/pmpp.2000.0305

Khan, M. K., Pandey, A., Athar, T., Choudhary, S., Deval, R., Gezgin, S., et al. (2020). Fusarium head blight in wheat: contemporary status and molecular approaches. *3 Biotech* 10:172. doi: 10.1007/s13205-020-2158-x

Khan, Z., Rahimi-Eichi, V., Haefele, S., Garnett, T., and Miklavcic, S. J. (2018). Estimation of vegetation indices for high-throughput phenotyping of wheat using aerial imaging. *Plant Methods* 14, 1–11. doi: 10.1186/s13007-018-0287-6

Kheiri, A., Moosawi Jorf, S. A., and Malihipour, A. (2019). Infection process and wheat response to *Fusarium* head blight caused by *Fusarium graminearum*. *Eur. J. Plant Pathol.* 153, 489–502. doi: 10.1007/s10658-018-1576-7

Leplat, J., Mangin, P., Falchetto, L., Heraud, C., Gautheron, E., and Steinberg, C. (2018). Visual assessment and computer–assisted image analysis of *Fusarium* head blight in the field to predict mycotoxin accumulation in wheat grains. *Eur. J. Plant Pathol.* 150, 1065–1081. doi: 10.1007/s10658-017-1345-z

Li, L., Zhang, Q., and Huang, D. (2014). A review of imaging techniques for plant phenotyping. *Sensors (Switzerland)* 14, 20078–20111. doi: 10.3390/s141120078

Li, X., Zhang, J. B., Song, B., Li, H. P., Xu, H. Q., Qu, B., et al. (2010). Resistance to *Fusarium* head blight and seedling blight in wheat is associated with activation of a cytochrome P450 gene. *Phytopathology* 100, 183–191. doi: 10.1094/PHYTO-100-2-0183

Liu, L., Dong, Y., Huang, W., Du, X., and Ma, H. (2020). Monitoring wheat *Fusarium* head blight using unmanned aerial vehicle hyperspectral imagery. *Remote Sens.* 12, 3811. doi: 10.3390/rs12223811

Louhaichi, M., Borman, M. M., and Johnson, D. E. (2001). Spatially located platform and aerial photography for documentation of grazing impacts on wheat. *Geocarto Int.* 16, 65–70. doi: 10.1080/10106040108542184

Lu, J., Cheng, D., Geng, C., Zhang, Z., Xiang, Y., and Hu, T. (2021). Combining plant height, canopy coverage and vegetation index from UAV-based RGB images to estimate leaf nitrogen concentration of summer maize. *Biosyst. Eng.* 202, 42–54. doi: 10.1016/j.biosystemseng.2020.11.010

Lu, Z. X., Gaudet, D., Puchalski, B., Despins, T., Frick, M., and Laroche, A. (2006). Inducers of resistance reduce common bunt infection in wheat seedlings while differentially regulating defence-gene expression. *Physiol. Mol. Plant Pathol.* 67, 138–148. doi: 10.1016/j.pmpp.2005.12.001

Ludovisi, R., Tauro, F., Salvati, R., Khoury, S., Mugnozza, G. S., and Harfouche, A. (2017). Uav-based thermal imaging for high-throughput field phenotyping of black poplar response to drought. *Front. Plant Sci.* 8:1681. doi: 10.3389/fpls.2017.01681

Ma, Z., Xie, Q., Li, G., Jia, H., Zhou, J., Kong, Z., et al. (2020). Germplasms, genetics and genomics for better control of disastrous wheat *Fusarium* head blight. *Theor. Appl. Genet.* 133, 1541–1568. doi: 10.1007/s00122-019-03525-8

Mahlein, A. K. (2016). Plant disease detection by imaging sensors – parallels and specific demands for precision agriculture and plant phenotyping. *Plant Dis.* 100, 241–254. doi: 10.1094/PDIS-03-15-0340-FE

Mahlein, A.-K., Alisaac, E., Al Masri, A., Behmann, J., Dehne, H.-W., and Oerke, E.-C. (2019). Comparison and combination of thermal, fluorescence, and hyperspectral imaging for monitoring *Fusarium* head blight of wheat on spikelet scale. *Sensors* 19:2281. doi: 10.3390/s19102281

Mandalà, G., Tundo, S., Francesconi, S., Gevi, F., Zolla, L., Ceoloni, C., et al. (2019). Deoxynivalenol detoxification in transgenic wheat confers resistance to *Fusarium* head blight and crown rot diseases. *Mol. Plant Microbe Interact.* 32, 583–592. doi: 10.1094/MPMI-06-18-0155-R

Maxwell, K., and Johnson, G. N. (2000). Chlorophyll fluorescence – a practical guide. *J. Exp. Bot.* 51, 659–668. doi: 10.1093/jxb/51.345.659

McMullen, M., Bergstrom, G., De Wolf, E., Dill-Macky, R., Hershman, D., Shaner, G., et al. (2012). A unified effort to fight an enemy of wheat and barley: *Fusarium* head blight. *Plant Dis.* 96, 1712–1728. doi: 10.1094/PDIS-03-12-0291-FE

Mesterházy, Á, Bartók, T., and Lamper, C. (2003). Influence of wheat cultivar, species of *Fusarium*, and isolate aggressiveness on the efficacy of fungicides for control of Fusarium head blight. *Plant Dis.* 87, 1107–1115. doi: 10.1094/PDIS.2003.87.9.1107

Metsalu, T., and Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using Principal component analysis and heatmap. *Nucleic Acids Res.* 43, 566–570. doi: 10.1093/nar/gkv468

Murchie, E. H., Kefauver, S., Araus, J. L., Muller, O., Rascher, U., Flood, P. J., et al. (2018). Measuring the dynamic photosynthome. *Ann. Bot.* 122, 207–220. doi: 10.1093/aob/mcy087

Muthukrishnan, S., Liang, G. H., Trick, H. N., and Gill, B. S. (2001). Pathogenesis-related proteins and their genes in cereals. *Plant Cell. Tissue Organ. Cult.* 64, 93–114. doi: 10.1023/A:1010763506802

Nishar, A., Richards, S., Breen, D., Robertson, J., and Breen, B. (2016). Thermal infrared imaging of geothermal environments and by an unmanned aerial vehicle (UAV): a case study of the Wairakei – Tauhara geothermal field, Taupo, New Zealand. *Renew. Energy* 86, 1256–1264. doi: 10.1016/j.renene.2015.09.042

Oerke, E.-C., and Steiner, U. (2010). "Potential of digital thermography for disease control," in *Precision Crop Protection – The Challenge and Use of Heterogeneity*, eds E. C. Oerke, R. Gerhards, G. Menz, and R. Sikora (Dordrecht: Springer), 167–182. doi: 10.1007/978-90-481-9277-9

Oerke, E. C., Steiner, U., Dehne, H. W., and Lindenthal, M. (2006). Thermal imaging of cucumber leaves affected by downy mildew and environmental conditions. *J. Exp. Bot.* 57, 2121–2132. doi: 10.1093/jxb/erj170

Oliver, R. E., Cai, X., Friesen, T. L., Halley, S., Stack, R. W., and Xu, S. S. (2008). Evaluation of Fusarium head blight resistance in tetraploid wheat (*Triticum turgidum* L.). *Crop Sci.* 48, 213–222. doi: 10.2135/cropsci2007.03.0129

Padmavathi, K., and Thangadurai, K. (2016). Implementation of RGB and grayscale images in plant leaves disease detection – comparative study. *Indian J. Sci. Technol.* 9, 1–6. doi: 10.17485/ijst/2016/v9i6/77739

Pancaldi, D., Tonti, S., Prodi, A., Salomoni, D., Dal Prà, M., Nipoti, P., et al. (2010). Survey of the main causal agents of *Fusarium* head blight of durum wheat around Bologna, northern Italy. *Phytopathol. Mediterr.* 49, 258–266.

Peiris, K. H. S., Pumphrey, M. O., Dong, Y., and Dowell, F. E. (2011). Fusarium head blight symptoms and mycotoxin levels in single kernels of infected wheat spikes. *Cereal Chem.* 88, 291–295. doi: 10.1094/CCHEM-08-10-0112

Peng, J. H., Sun, D., and Nevo, E. (2011). Domestication evolution, genetics and genomics in wheat. *Mol. Breed.* 28, 281–301. doi: 10.1007/s11032-011-9608-4

Pineda, M., Barón, M., and Pérez-Bueno, M. L. (2021). Thermal imaging for plant stress detection and phenotyping. *Rem. Sens.* 13:68. doi: 10.3390/rs13010068

Qiu, R., Yang, C., Moghimi, A., Zhang, M., Steffenson, B. J., and Hirsch, C. D. (2019). Detection of *Fusarium* head blight in wheat using a deep neural network and color imaging. *Rem. Sens.* 11:2658. doi: 10.3390/rs11222658

Ramegowda, V., and Senthil-Kumar, M. (2015). The interactive effects of simultaneous biotic and abiotic stresses on plants: mechanistic understanding from drought and pathogen combination. *J. Plant Physiol.* 176, 47–54. doi: 10.1016/j.jplph.2014.11.008

Rampino, P., Mita, G., Fasano, P., Borrelli, G. M., Aprile, A., Dalessandro, G., et al. (2012). Novel durum wheat genes up-regulated in response to a combination of heat and drought stress. *Plant Physiol. Biochem.* 56, 72–78. doi: 10.1016/j.plaphy.2012.04.006

Ray, D. K., Mueller, N. D., West, P. C., and Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS One* 8:e66428. doi: 10.1371/journal.pone.0066428

Rod, K. S., Bradley, C. A., Sanford, D. A., Van, and Knott, C. A. (2020). Integrating management practices to decrease deoxynivalenol contamination in soft red winter wheat. *Front. Plant Sci.* 11:1158. doi: 10.3389/fpls.2020.01158

Sagan, V., Maimaitijiang, M., Sidike, P., Eblimit, K., Peterson, K. T., Hartling, S., et al. (2019). UAV-based high resolution thermal imaging for vegetation monitoring, and plant phenotyping using ICI 8640 P, FLIR Vue Pro R 640, and thermomap cameras. *Rem. Sens.* 11:330. doi: 10.3390/rs11030330

Shah, D. A., Pucci, N., and Infantino, A. (2005). Regional and varietal differences in the risk of wheat seed infection by fungal species associated with *Fusarium* head blight in Italy. *Eur. J. Plant Pathol.* 112, 13–21. doi: 10.1007/s10658-004-6891-5

Shaik, R., and Ramakrishna, W. (2014). Machine learning approaches distinguish multiple stress conditions using stress-responsive genes and identify candidate genes for broad resistance in rice. *Plant Physiol.* 164, 481–495. doi: 10.1104/pp.113.225862

Simko, I., Jimenez-Berni, J. A., and Sirault, X. R. R. (2016). Phenomic approaches and tools for phytopathologists. *Phytopathology* 107, 6–17. doi: 10.1094/PHYTO-02-16-0082-RVW

Streich, J., Romero, J., Gazolla, J. G. F. M., Kainer, D., Cliff, A., Prates, E. T., et al. (2020). Can exascale computing and explainable artificial intelligence applied to plant biology deliver on the United Nations sustainable development goals? *Curr. Opin. Biotechnol.* 61, 217–225. doi: 10.1016/j.copbio.2020.01.010

Susič, N., Žibrat, U., Širca, S., Strajnar, P., Razinger, J., Knapič, M., et al. (2018). Discrimination between abiotic and biotic drought stress in tomatoes using hyperspectral imaging. *Sensors Actuat. B Chem.* 273, 842–852. doi: 10.1016/j.snb.2018.06.121

Tambussi, E. A., Casadesus, J., Munné-Bosch, S., and Araus, J. L. (2000). Photoprotection in water-stressed plants of durum wheat (*Triticum turgidum* var. durum): changes in chlorophyll fluorescence, spectral signature and photosynthetic pigments. *Funct. Plant Biol.* 29, 35–44. doi: 10.1071/PP01104

Tenea, G. N., Peres Bota, A., Cordeiro Raposo, F., and Maquet, A. (2011). Reference genes for gene expression studies in wheat flag leaves grown under different farming conditions. *BMC Res. Notes* 4:373. doi: 10.1186/1756-0500-4-373

Tittlemier, S. A., Brunkhorst, J., Cramer, B., DeRosa, M. C., Lattanzio, V. M. T., Malone, R., et al. (2021). Developments in mycotoxin analysis: an update for 2019-2020. *World Mycotoxin J.* 14, 3–26. doi: 10.3920/wmj2020.2664

Tomassini, A., Sella, L., Raiola, A., D'Ovidio, R., and Favaron, F. (2009). Characterization and expression of Fusarium graminearum endo-polygalacturonases in vitro and during wheat infection. *Plant Pathol.* 58, 556–564. doi: 10.1111/j.1365-3059.2008.02019.x

Tundo, S., Janni, M., Moscetti, I., Mandalá, G., Savatin, D., Blechl, A., et al. (2016). PvPGIP2 accumulation in specific floral tissues but not in the endosperm limits *Fusarium graminearum* infection in wheat. *Mol. Plant Microbe Interact.* 29, 815–821. doi: 10.1094/MPMI-07-16-0148-R

van Bezouw, R. F. H. M., Keurentjes, J. J. B., Harbinson, J., and Aarts, M. G. M. (2019). Converging phenomics and genomics to study natural variation in plant photosynthetic efficiency. *Plant J.* 97, 112–133. doi: 10.1111/tpj.14190

Vaughan, M., Backhouse, D., and Del Ponte, E. M. (2016). Climate change impacts on the ecology of *Fusarium graminearum* species complex and susceptibility of wheat to *Fusarium* head blight: a review. *World Mycotoxin J.* 9, 685–700. doi: 10.3920/WMJ2016.2053

Weckwerth, W. (2011). Green systems biology – from single genomes, proteomes and metabolomes to ecosystems research and biotechnology. *J. Proteomics* 75, 284–305. doi: 10.1016/j.jprot.2011.07.010

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crop. Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003

Xie, C., and Yang, C. (2020). A review on plant high-throughput phenotyping traits using UAV-based sensors. *Comput. Electron. Agric.* 178:105731. doi: 10.1016/j.compag.2020.105731

Xue, G. P., McIntyre, C. L., Chapman, S., Bower, N. I., Way, H., Reverter, A., et al. (2006). Differential gene expression of wheat progeny with contrasting levels of transpiration efficiency. *Plant Mol. Biol.* 61, 863–881. doi: 10.1007/s11103-006-0055-2

Yang, F., Jacobsen, S., Jørgensen, H. J. L., Collinge, D. B., Svensson, B., and Finnie, C. (2013). *Fusarium graminearum* and its interactions with cereal heads: studies in the proteomics era. *Front. Plant Sci.* 4:37. doi: 10.3389/fpls.2013.00037

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. *Front. Plant Sci.* 8:1111. doi: 10.3389/fpls.2017.01111

Yang, M., Hassan, M. A., Xu, K., Zheng, C., Rasheed, A., Zhang, Y., et al. (2020). Assessment of water and nitrogen use efficiencies through UAV-based multispectral phenotyping in winter wheat. *Front. Plant Sci.* 11:927. doi: 10.3389/fpls.2020.00927

Yang, S., Li, X., Chen, W., Liu, T., Zhong, S., Ma, L., et al. (2016). Wheat resistance to *Fusarium* head blight is associated with changes in photosynthetic parameters. *Plant Dis.* 100, 847–852. doi: 10.1094/PDIS-04-14-0398-RE

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zhang, X., Fu, J., Hiromasa, Y., Pan, H., and Bai, G. (2013). Differentially expressed proteins associated with *Fusarium* head blight resistance in wheat. *PLoS One* 8:e82079. doi: 10.1371/journal.pone.0082079

Conflict of Interest: 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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 16 April 2021
doi: 10.3389/fpls.2021.658357

Performance of the Two-Source Energy Balance (TSEB) Model as a Tool for Monitoring the Response of Durum Wheat to Drought by High-Throughput Field Phenotyping

David Gómez-Candón[1]*, *Joaquim Bellvert*[1] and *Conxita Royo*[2]

[1] Efficient Use of Water in Agriculture Program, Institute of Agrifood Research and Technology (IRTA), Fruitcentre, PCiTAL, Parc Científic i Tecnològic Agroalimentari de Gardeny, Lleida, Spain, [2] Sustainable Field Crops Program, Institute of Agrifood Research and Technology (IRTA), Lleida, Spain

OPEN ACCESS

Edited by:
Jianfeng Zhou,
University of Missouri, United States

Reviewed by:
Michael Gomez Selvaraj,
Consultative Group on International Agricultural Research (CGIAR), United States
Ana María Mendez-Espinoza,
Institute of Agricultural Research, Chile

***Correspondence:**
David Gómez-Candón
david.gomez@irta.cat

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

Received: 25 January 2021
Accepted: 22 March 2021
Published: 16 April 2021

Citation:
Gómez-Candón D, Bellvert J and Royo C (2021) Performance of the Two-Source Energy Balance (TSEB) Model as a Tool for Monitoring the Response of Durum Wheat to Drought by High-Throughput Field Phenotyping.
Front. Plant Sci. 12:658357.
doi: 10.3389/fpls.2021.658357

The current lack of efficient methods for high throughput field phenotyping is a constraint on the goal of increasing durum wheat yields. This study illustrates a comprehensive methodology for phenotyping this crop's water use through the use of the two-source energy balance (TSEB) model employing very high resolution imagery. An unmanned aerial vehicle (UAV) equipped with multispectral and thermal cameras was used to phenotype 19 durum wheat cultivars grown under three contrasting irrigation treatments matching crop evapotranspiration levels (ETc): 100%ETc treatment meeting all crop water requirements (450 mm), 50%ETc treatment meeting half of them (285 mm), and a rainfed treatment (122 mm). Yield reductions of 18.3 and 48.0% were recorded in the 50%ETc and rainfed treatments, respectively, in comparison with the 100%ETc treatment. UAV flights were carried out during jointing (April 4th), anthesis (April 30th), and grain-filling (May 22nd). Remotely-sensed data were used to estimate: (1) plant height from a digital surface model (H, R^2 = 0.95, RMSE = 0.18m), (2) leaf area index from multispectral vegetation indices (LAI, R^2 = 0.78, RMSE = 0.63), and (3) actual evapotranspiration (ETa) and transpiration (T) through the TSEB model (R^2 = 0.50, RMSE = 0.24 mm/h). Compared with ground measurements, the four traits estimated at grain-filling provided a good prediction of days from sowing to heading (DH, r = 0.58–0.86), to anthesis (DA, r = 0.59–0.85) and to maturity (r = 0.67–0.95), grain-filling duration (GFD, r = 0.54–0.74), plant height (r = 0.62–0.69), number of grains per spike (NGS, r = 0.41–0.64), and thousand kernel weight (TKW, r = 0.37–0.42). The best trait to estimate yield, DH, DA, and GFD was ETa at anthesis or during grain filling. Better forecasts for yield-related traits were recorded in the irrigated treatments than in the rainfed one. These results show a promising perspective in the use of energy balance models for the phenotyping of large numbers of durum wheat genotypes under Mediterranean conditions.

Keywords: transpiration, remote sensing, plant height, yield, grain number, grain weight, LAI, UAV

INTRODUCTION

Wheat is a staple food for humans, providing 18% of the daily human intake of calories and 20% of protein (http://www.fao.org/faostat/). Durum wheat (*Triticum turgidum* L. subsp. *durum* [Desf.] Husn) represents about 6% of a global wheat production of about 740 million tons per year (FAO, 2017). Wheat production per unit area needs to double by 2050 to meet the projected food demand of a global population forecast to be 9.22 billion. Achieving this objective is a significant challenge that will require increasing the current global yield increase rate of 1.3–2.4% y^{-1} (Ray et al., 2013), whilst at the same time minimizing the use of resources and the environmental impact (Tilman et al., 2011; Lal, 2016). Besides, in the current scenario of global climate change, the success of sustainable agriculture in many regions of the world is totally reliant on water availability. The Mediterranean region –the largest durum wheat producing area worldwide, the largest consumer of durum wheat products and the most important import market–, is one of the most sensitive to the effects of climate change, with projections forecasting a precipitation decrease of 4–27% during the cropping season (Flato et al., 2013). The development of high-yielding cultivars adapted to water-limited conditions is therefore critical to guarantee food security.

There is a general agreement that yield increases can only be achieved by improving the efficiency of large-scale breeding programs, particularly for suboptimal environments (Moshelion and Altman, 2015). One of the major challenges facing breeding programs centered on drought-prone areas is to develop tools capable of quantifying the actual water use of plants under different water regimes. The development of wheat varieties with improved water use efficiency (WUE, yield as a function of water used in transpiration) is seen as a way to increase yield in rainfed environments (Condon, 2004; Condon and Maxwell, 2014). The major challenge for fast genetic progress is to connect genetic variants (genotype) to their expression in observable traits (phenotype), and to predict plant phenotypes from genetic information (Cobb et al., 2013). The enormous advances in the genome sequencing of plants are providing massive genomic datasets, but the lack of efficient methods to rapidly collect large volumes of high quality phenotypic data has become a bottleneck in genomics-assisted breeding (White et al., 2012). Until now, given the complexity of measuring actual transpiration or water status in a large number of plots under field conditions, the difficulty of measuring the phenotypic response of plants to water use constraints has limited the goal of higher yields in breeding programs. Given this difficulty, evaluations of plant transpiration have relied mostly on surrogate traits, although this has most likely resulted in over-dependence on the surrogates (Vadez et al., 2014). Moreover, traditional phenotyping in germplasm evaluation activities under field conditions requires substantial investments in time, labor, and cost.

There is growing scientific interest in the application of remote sensing for high throughput phenotyping (HTP), particularly in breeding and germplasm evaluation activities (Furbank and Tester, 2011, Fiorani and Schurr, 2013; Walter et al., 2015). HTP through remote sensing allows the assessment of plant phenotypes on a scale and with a level of precision and speed that are unattainable with traditional methods (Dhondt et al., 2013). Numerous studies have used either RGB, fluorescent, thermal, hyperspectral, or 3D imaging to estimate morphological traits, biomass, plant growth, yield, water status, canopy temperature, or disease symptoms in many breeding programs and crops (Deery et al., 2014; Haghighattalab et al., 2016; Watanabe et al., 2017; Yang et al., 2017, 2020; Sagan et al., 2019). In addition, crop growth rates and spatial mapping of crop height variations have been obtained in wheat at field scale, as well as in individual plots, from images obtained with an RGB camera mounted on an unmanned aerial vehicle (UAV) (Holman et al., 2016). Madec et al. (2017) obtained a reliable assessment of the height of wheat plants with a digital camera with a 6,000–4,000 pixel sensor mounted on a hexacopter. Shi et al. (2016) developed empirical models to estimate the leaf area index (LAI) and percent canopy cover of winter wheat. Bendig et al. (2014) estimated fresh and dry above-ground biomass in barley from RGB images captured from a small UAV. Chapman et al. (2014) estimated crop lodging in wheat plots of a breeding program from images taken by cameras mounted on a customized robotic helicopter. Detailed reviews on remote sensing tools and platforms available for HTP in a plant breeding context can be found in Araus and Cairns (2014) and Araus et al. (2018).

Water status has been assessed in different crops by HTP thermography (Costa et al., 2013; Leroux et al., 2016; Perich et al., 2019). In many studies, different approaches have been used to calculate the so-called crop water stress index (CWSI) (Jackson et al., 1981; Jones, 1999; Gonzalez-Dugo et al., 2015). However, when the CWSI is calculated either empirically through non-water stress baselines or with reference panels, comparison between cultivars can only be achieved in a relative way. This is because the CWSI depends, among other factors, on the stomatal response to the vapor pressure deficit (VPD), which varies between cultivars and crop developmental stages. Therefore, it is too complex to determine this response for large collections of cultivars. Surface energy balance (SEB) models have also been widely used for assessing the actual evapotranspiration and water status of many crops at different scales, mostly using satellite imagery (Bastiaanssen et al., 1998; Allen et al., 2007; McShane et al., 2017). Among the different SEB models, the two-source energy balance (TSEB) modeling scheme allows the estimation of transpiration and evaporation separately (Norman et al., 1995). However, if very high resolution thermal imagery is available, in which case it is possible to directly retrieve soil (T_{soil}) and canopy (T_c) surface temperatures, the model can also be used, obtaining in some cases higher accuracies (Nieto et al., 2018; Bellvert et al., 2020). As far as we are aware, only Bellvert et al. (2021) have used this model to date for field-based phenotyping (of a collection of almond rootstocks in their work), but this present work is

Abbreviations: ET, evapotranspiration; T, transpiration; H, predicted plant height; PH, observed plant height; DH, days from sowing to heading; DA, days from sowing to anthesis; DM, days from sowing to maturity; DAS, days after sowing; GFD, grain filling duration; NSm^2, number of spikes/m^2; NGS, number of grains/spike; TKW, thousand kernel weight; GFR, grain filling rate; GFD, grain filling duration.

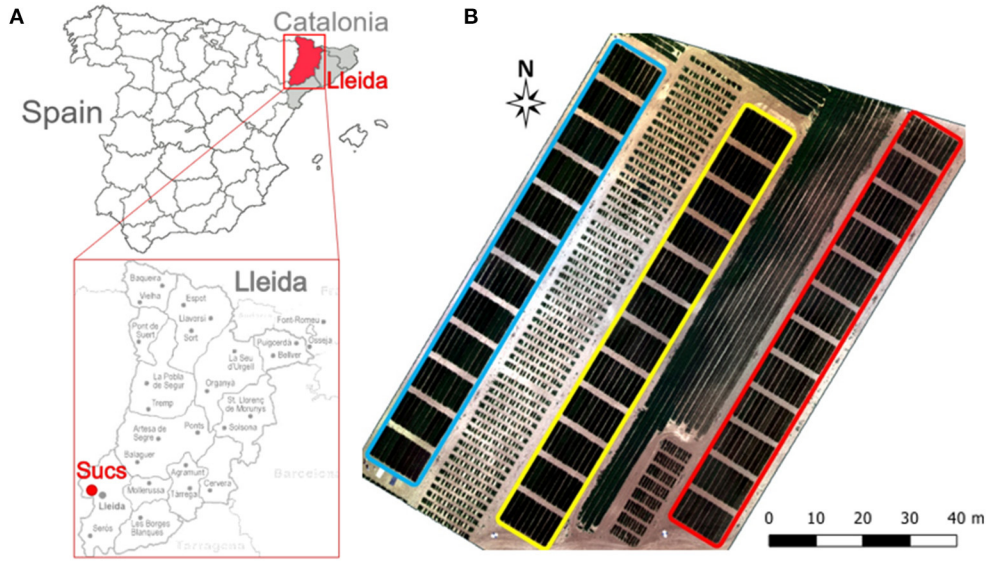

FIGURE 1 | (A) Study site location; **(B)** Layout of the field experiment showing the three irrigation treatments: 100%ETc (blue), 50%ETc (yellow), and rainfed (red).

the first to evaluate the feasibility of TSEB in a set of durum wheat cultivars.

The primary objective of this research was to determine the suitability of using the TSEB model for the assessment of actual evapotranspiration and its components in a collection of spring durum wheat cultivars grown under contrasting water regimes in a Mediterranean environment. The specific objectives were: (1) to quantify the yield penalty caused by a reduction of water availability, (2) to determine the crop growing stage most suitable for assessing important agronomic traits through remote sensing images, and (3) to identify the agronomic traits that can be reliably assessed by remote sensing images and the best performing indicators for them.

MATERIALS AND METHODS

Experimental Setup

The field experiment was conducted at Sucs, Spain (41°41′49″N, 0°25′46″E, 285 m elevation) during the 2018–2019 growing season. The site has a typical Mediterranean climate, with a rainfall and reference evapotranspiration (ET_0) of 177 and 603 mm, respectively, during the growing season. Soil has a fine-loamy texture with a field capacity of 27% and wilting point of 13% as calculated from the Saxton's soil hydraulic calculator (Saxton et al., 1986). Fourteen durum wheat (*Triticum turgidum* ssp. *durum*) commercial varieties (Anvergur, Athoris, Burgos, Calero, Carpio, Claudio, Don Ricardo, Don Sebastián, Eunoble, Euroduro, Grador, Iberus, Sculptur, and Tussur) and five inbred lines from the IRTA durum wheat breeding program (05D278, 07D057, 08D010, 09D066, 09D069) were evaluated under three contrasting irrigation treatments. Irrigation treatments were as follows: (i) 100%ETc, irrigated 100% of the seasonal crop evapotranspiration (ETc), (ii) 50%ETc, irrigated 50% of seasonal ETc, and (iii) Rainfed, which was not irrigated (**Figure 1**). In each irrigation treatment, genotypes were planted following an incomplete block design with four replications and plots of 9.6 m^2 (eight rows 8 m long and 0.15 m apart). Sowing was carried out on December 4th 2018 at a density of 450 seeds/m^2. Due to the low precipitation received from December to February (29 mm) all plots were evenly irrigated on March 1st with 20 mm to guarantee the plants' survival. Irrigation was scheduled on a weekly basis and water was applied during 2–3 days of the week. Sprinklers were installed in a grid of 18 × 18 m and water flow discharge was 7.8 l/h/m^2 for the 100%ETc and 3.9 l/h/m^2 for the 50%ETc treatments. Weekly irrigation was scheduled following a water balance model (Allen et al., 1998). ETc was calculated as a product of the Penman-Monteith ET_0 (Allen et al., 1998) and crop coefficients (Kc). The used crop coefficients were derived from FAO-56 (Allen et al., 1998), and started from 0.7 at the vegetative growth stage to 1.07 at the beginning of the mid-season stage. During the late season (from June 11th), the Kc decreased and reached a value of 0.6. In addition, 0.8 was used as a coefficient of efficiency of the sprinkler irrigation system (Savva et al., 2001). Meteorological data was gathered from an automated weather station belonging to Catalonia's official network of meteorological stations (SMC, www.ruralcat.net/web/guest/agrometeo), which is located around 3 km from the study site. The amount of water applied through irrigation in each treatment during the entire growing season was also measured with digital water meters (CZ2000-3M, Contazara, Zaragoza, Spain). Before sowing, the field was fertilized with 162 u P_2O_5 and 360 u K_2O ha^{-1} and top dressed twice with ammonium nitro-sulfate at rates of 118 kg N/ha at the end of tillering and 50 kg N/ha at mid-jointing. The field was maintained free of weeds, diseases and pests by chemical treatments.

FIGURE 2 | Unmanned aerial vehicle (UAV) and cameras used in the study. **(A)** UAV, **(B)** Multispectral Micasense Rededge, and **(C)** Thermal Flir Vue Pro cameras.

Image Acquisition Campaign

Images were acquired with the Cóndor UAV hexacopter (Dronetools, https://www.dronetools.es/) (**Figure 2**) on April 4th (121DAS, days after sowing), April 30th (147DAS), and May 22nd (169DAS) in 2019, coinciding with the crop developmental stages of mid-jointing, around anthesis and grain filling, respectively (**Figure 3**). The vapor pressure deficit (VPD) and air temperature (T_a) at the moment of image acquisition were respectively, 5.5 KPa and 9.9°C for 121DAS, 12.0 KPa and 18.8°C for 147DAS, and 11.0 KPa and 19.7°C for 169DAS. Flights were always conducted in sunny conditions and with a wind speed below 12 m/s. The UAV was equipped with a multispectral and a thermal camera. The former was a Micasense RedEdge-M (Micasense, 1300 N Northlake Way, Seattle, USA), which has five spectral bands located at the wavelengths 475 ± 20 nm (blue), 560 ± 20 nm (green), 668 ± 10 nm (red), 717 ± 10 nm (red edge), and 840 ± 40 nm (near infrared), and a field of view (FOV) of 47.2°. The thermal camera was a FLIR Vue PRO (FLIR Systems, Wilsonville, OR, USA) with a resolution of 336 × 256 pixels and a 6.8 mm focal length, with a FOV of 45 × 35°. The spectral response was in the range of 7.5–13.5 μm. All flights were conducted at ∼12:00 h solar time. The UAV flew over at a height of 50 m agl (above ground level), capturing images with a resolution of 0.02 and 0.10 m per pixel for the multispectral and thermal cameras, respectively. Flight planning had 80/60 frontal and side overlap, respectively. During image acquisition, *in situ* measurements were conducted for different targets in order to correct the atmospheric contribution to the signal. Temperature measurements were continuously recorded for hot and cold targets (black and white panels, bare soil, and vegetation) with a fixed IR-temperature sensor (Calex PC151LT-O, Pyrocouple series, Calex Electronics Limited, Bedfordshire, UK). The radiometric calibration of the multispectral sensor was conducted through an external incident light sensor that measured the irradiance levels of light at the same bands as the camera. In addition, *in situ* spectral measurements for ground calibration targets were performed using a Jaz spectrometer (Ocean Optics, Inc., Dunedin, FL, USA). The Jaz has a wavelength response from 200 to 1,100 nm and an optical resolution of 0.3 to 10.0 nm. During spectral collection, spectrometer calibration measurements were taken with a reference panel (white color Spectralon^{TM}) and dark current before and after taking readings from radiometric calibration targets. Geometrical correction was conducted using five ground control points (GCP), and measuring the position in each with a handheld global positioning system (GPS) (Geo7x, Trimble GeoExplorer series, Sunnyvale, CA, USA). All images were mosaicked using the Agisoft Photoscan Professional version 1.6.2 (Agisoft LLC., St. Petersburg, Russia) software and geometrically and radiometrically terrain corrected with QGIS 3.4 (QGIS 3.4.15).

Measurements of Agronomic Traits

Crop development was monitored on three replications per treatment on a twice-weekly basis from booting to record the following growth stages (Zadoks et al., 1974): GS55 (heading), GS65 (anthesis), and GS87 (physiological maturity). A plot was considered to have reached a given developmental stage when ∼50% of the plants exhibited the stage-specific phenotypic characteristics. For each UAV flight date, on-ground key crop biophysical parameters were measured as follows. Plant height (PH, cm) was measured in three plants per plot of one replication for each irrigation treatment. PH was also measured at GS87 in three main stems per plot in three replications from the tillering node to the top of the spike, excluding the awns. The LAI was obtained on the same days and in the same plots using a portable linear ceptometer (AccuPAR model LP-80, Decagon Devices Inc., Pullman, WA, USA). Measurements were conducted from 12:00 to 15:00 h (local time) in one replicate of each irrigation treatment. In total, 63 plots were measured for each flight event. Photosynthetically active radiation (PAR) below the wheat was measured placing the ceptometer in a horizontal position at ground level and recording five PAR readings in each plot. A fixed tripod connected to the sensor allowed collection of the incident radiation above the plants. Then, the LAI calculator provided by AccuPAR-L80 (LAI-calculator, METER Group) was used to estimate LAI. Concomitant to image acquisition, in three leaves per plot of one replication in each irrigation treatment, leaf transpiration was also measured with an infrared gas analyzer (IRGA) (LI-7500, LI-COR Inc., Lincoln, NE). Plots were harvested mechanically at ripening and yield (kg/ha) expressed as dry weight. From a random sample of the plants contained in a 0.5-m-long stretch from a central row of each plot of three replications at ripening, the number of spikes/m^2 (NSm^2) and the number of grains/spike (NGS) were assessed. Thousand kernel weight (TKW) was estimated as the mean weight of three sets of 100 grains per plot. Grain filling rate (GFR, mg/day) was obtained as the quotient between grain dry weight and grain-filling duration (GFD) considered to be the number of days between anthesis and physiological maturity.

Remotely-Sensed Estimates of Biophysical Traits

The three-dimensional plant height (H) was estimated from the photogrammetric point cloud of multispectral images. The digital surface model (DSM) and the digital terrain model

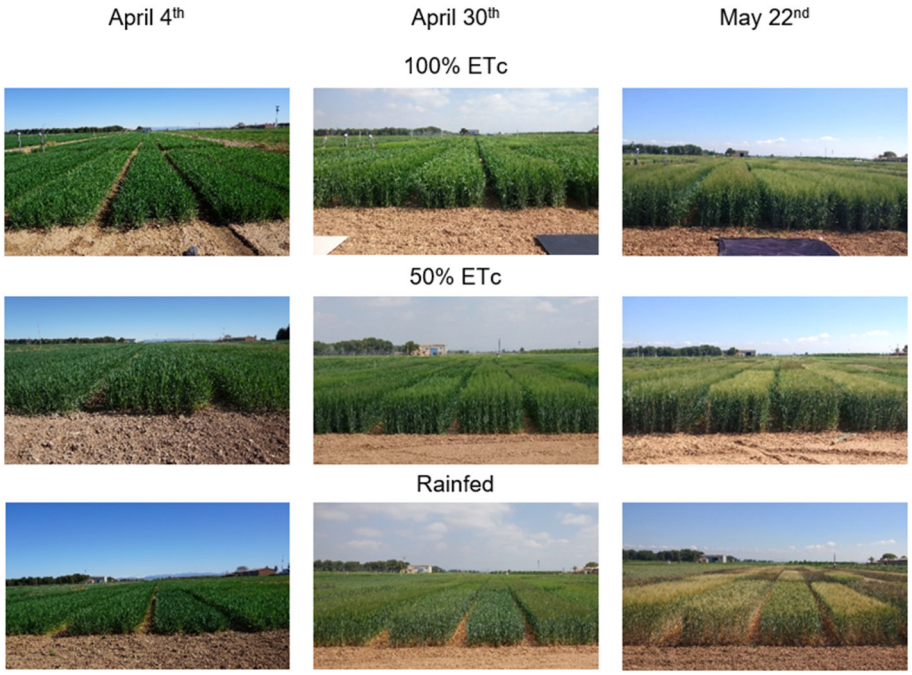

FIGURE 3 | Pictures of each irrigation treatment for the different dates of image acquisition (April 4th, April 30th, and May 22nd).

(DTM, bare soil surface devoid of plants) were both obtained through automatic aerial triangulation, bundle block adjustment, and camera calibration methods using the Agisoft PhotoScan version 1.6.2 (Agisoft, 2020; St. Petersburg, Russia) software. A classification of bare ground pixels was used to obtain the DTM of the field. Then, a raster corresponding to heights was obtained by subtracting the DTM from the DSM using the band math tool of the QGIS software (**Figure 4**). LAI was estimated from spectral vegetation indices. In particular, this study used the improved modified triangular vegetation index ($MTVI_2$) (Yao et al., 2017), which was calculated as:

$$MTVI_2 = \frac{1.5\left[1.2\left(R_{840} - R_{560}\right) - 2.5\left(R_{717} - R_{560}\right)\right]}{\sqrt{\left(2R_{840} + 1\right)^2 - \left(6R_{840} - 5\sqrt{R_{717}}\right) - 0.5}} \quad (1)$$

The fractional vegetation cover (fc) of each plot was also calculated by adapting the equation proposed by Gutman and Ignatov (1998). Instead of the normalized difference vegetation index (NDVI), we used the $MTVI_2$ due to its low saturation at high LAI values:

$$f_c = \frac{MTVI2_i - MTVI2_{soil}}{MTVI2_{veg} - MTVI2_{soil}} \quad (2)$$

where

$MTVI_{2i}$ corresponds to the value on the target plot;
$MTVI_{2soil}$ corresponds to the value of bare soil; and
$MTVI_{2veg}$ corresponds to the value of pure vegetation.

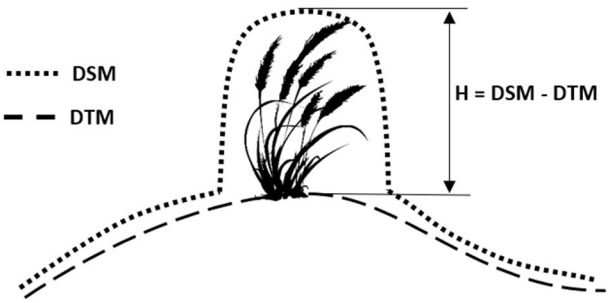

FIGURE 4 | Schematic workflow used to estimate crop height. Digital terrain model (DTM) and digital surface model (DSM) obtained through automatic aerial triangulation. Plant height (H) estimated above ground surface.

Remotely-Sensed Estimates of Evapotranspiration

Actual crop evapotranspiration (ETa) and its partition components were retrieved from the two-source energy balance (TSEB) model (Norman et al., 1995; Kustas and Anderson, 2009). Two-source models partition the surface energy fluxes and the radiometric temperature (T_{rad}) between nominal soil and canopy sources. The approach is therefore able to estimate canopy transpiration (T) and soil evaporation (E) separately. However, because direct measurements of canopy (T_c) and soil (T_{soil}) temperatures are rarely available with satellite imagery, in most applications these component temperatures are estimated in an iterative process in which it is first assumed that green

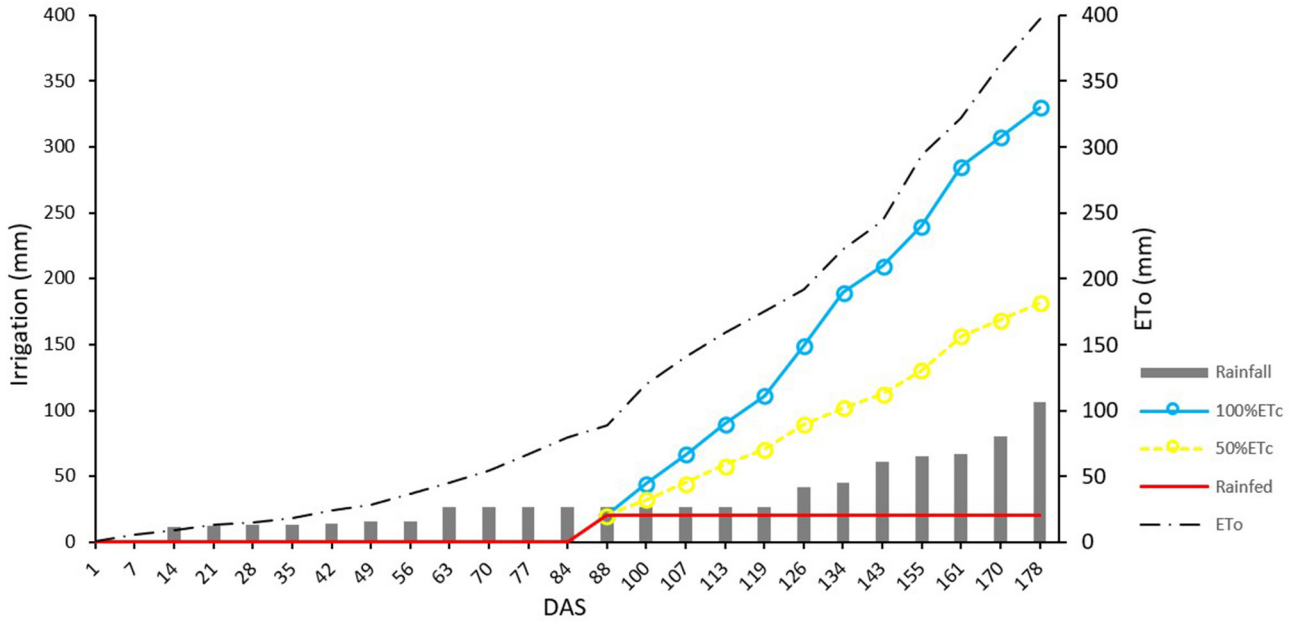

FIGURE 5 | Amount of water applied, rainfall, and reference evapotranspiration (ET_0) corresponding to the different irrigation treatments (100%ETc, 50%ETc, and rainfed) from March 2019 (87 DAS, days after sowing) to May 2020 (177 DAS).

canopy (expressed as the function of LAI that is green) transpires at a potential rate based on the Priestley-Taylor formulation (Priestley and Taylor, 1972). On the other hand, if very high resolution thermal imagery is available, it is possible to obtain T_s and T_c directly, without the need to compute an initial canopy transpiration (Nieto et al., 2018; Bellvert et al., 2021). In this study, T_{soil} and T_c were individually obtained for each plot from the thermal imagery. The model also requires other inputs such as plant height, LAI and f_c, the retrieval of which has been described above. Meteorological inputs were obtained from Catalonia's official network of meteorological stations. For more information, the full python script is available online (https://github.com/hectornieto/pyTSEB, last accessed 20.08.2020) and additional details of the TSEB model are provided by Norman et al. (1995), Kustas and Norman (1999), and Nieto et al. (2018).

Statistical Analyses

Analyses of variance (ANOVAs) were conducted following a split-plot design. Means were compared with a Tukey test at $P < 0.05$. Linear regression equations and Pearson correlation coefficients were used to analyze the relationship between variables.

RESULTS

Effect of Irrigation Treatments on the Agronomic Performance of Durum Wheat

The amount of irrigation water applied throughout the growing season in the 100%ETc and 50%ETc treatments was 340 and 180 mm, respectively (**Figure 5**). When also considering the rainfall from sowing to physiological maturity, the total amount of water received was 450, 285, and 122 mm for the 100%ETc, 50%ETc, and rainfed treatments, respectively. The ANOVAs showed statistically significant differences among irrigation treatments for all of the evaluated agronomic traits (**Table 1**). Yield ranged between 7,274 and 10,446 kg/ha in the 100%ETc, between 5,910 and 8,469 kg/ha in the 50%ETc and between 3,905 and 5,972 kg/ha in the rainfed treatments (**Table 1** and **Supplementary Table 1**). The effect of the total amount of water applied on yield was huge, as the 50%ETc and rainfed treatments reduced yield on average by 18.3 and 48.0%, respectively, in comparison to the treatment meeting all crop water requirements. While the 50%ETc treatment did not diminish the NSm^2, it did decrease the NGS and TKW. The absence of irrigation resulted in larger reductions in NGS than in NSm^2 and TKW. The grain filling rate increased steadily as consequence of water shortage. Plant height was reduced 6.5 and 11.9% in the 50%ETc and rainfed treatments in comparison with 100%ETc.

Reductions in the amount of water applied also significantly shortened the crop cycle (**Table 1**). In comparison with 100%ETc, the 50%ETc treatment resulted in decreases of 3 (2.2%), 4 (2.8%), and 6 (3.3%) days in the length of the periods from sowing to heading, anthesis, and maturity, respectively. The rainfed treatment additionally shortened between 4 and 7 days the periods needed to reach each of these growth stages. In consequence, a significant drop was observed in the duration of the grain-filling period.

TABLE 1 | Mean values ± SE and ranges (between brackets) for yield and yield-related traits of the 19 durum wheat genotypes included in the study.

Trait	Irrigation treatment					
	100% ETc		50% ETc		Rainfed	
Grain yield (kg/ha)	9219 ± 184[a]	(7,274–10,446)	7534 ± 184[b]	(5,910–8,469)	4793 ± 120[c]	(3,905–5,972)
Number of spikes/m^2 (NSm^2)	501 ± 26.1[a]	(289–742)	523 ± 16.0[a]	(387–627)	452 ± 13.6[b]	(356–591)
Number of grains/spike (NGS)	42.1 ± 1.53[a]	(34.3–60.3)	35.8 ± 1.56[b]	(23.6–53.9)	27.9 ± 1.32[c]	(15.3–38.8)
Thousand kernel weight (TKW, g)	55.8 ± 0.69[a]	(45.5–64.8)	53.7 ± 1.01[b]	(44.9–61.3)	51.5 ± 0.78[c]	(45.1–58.4)
Grain filling rate (GFR, mg/day)	1.39 ± 0.04[c]	(1.08–1.64)	1.42 ± 0.03[b]	(1.11–1.67)	1.47 ± 0.03[a]	(1.24–1.68)
Plant height (PH, cm)	92 ± 1.44[a]	(80–102)	86 ± 1.29[b]	(73–95)	81 ± 1.01[c]	(72–92)
Days to heading (DH)	135 ± 0.55[a]	(130–138)	132 ± 0.52[b]	(127–136)	128 ± 0.43[c]	(124–130)
Days to anthesis (DA)	143 ± 0.66[a]	(138–147)	139 ± 0.54[b]	(135–143)	135 ± 0.28[c]	(133–137)
Days to maturity (DM)	183 ± 0.39[a]	(180–186)	177 ± 0.35[b]	(174–143)	170 ± 0.45[c]	(167–174)
Grain filling duration (GFD, days)	40 ± 0.41[a]	(37–43)	38 ± 0.41[b]	(35–40)	35 ± 0.38[c]	(32–38)

Means within rows with different letters are significantly different for a Tukey test at P < 0.05.

TABLE 2 | *P*-values of the ANOVAs for the traits estimated through remote sensing.

Source of variation	D.F.	ETa	T	H	LAI
Flight date	2	<0.0001	<0.0001	<0.0001	<0.0001
Irrigation treatment	2	<0.0001	<0.0001	<0.0001	<0.0001
Error a	6				
Flight date* Irrigation treatment	4	<0.0001	<0.0001	<0.0001	<0.0001
Error b	12				
Genotype	18	<0.0001	<0.0001	<0.0001	<0.0001
Flight date × Genotype	36	ns	0.0127	ns	0.0453
Irrigation treatment × Genotype	36	0.0005	0.0479	0.0124	0.0018
Residual	396				
Total	512				

ETa, actual evapotranspiration; T, actual transpiration; H, estimated plant height.

Remotely-Sensed Estimates of the Biophysical Parameters and Evapotranspiration Components

The ANOVA showed statistically significant differences between flights, irrigation treatments and their interaction (**Table 2**). This interaction was of a cross-over nature due to the opposite trend observed in the first flight compared with the second and third ones, as shown in **Table 3**. Genotypes also differed for all remotely-sensed traits. The interactions of genotype with flight date and irrigation treatment were significant with the exception of the flight date × genotype interaction for H and ETa (**Table 2**).

The $MTVI_2$ vegetation index (VI) was linearly related with LAI when aggregating data from the three flight dates (R^2 = 0.78, $P < 0.001$, **Figure 6A**). Also, this regression was significant for each specific date, with R^2 values of 0.20, 0.77, and 0.87 for April 4th, April 30th, and May 22nd, respectively. The one-to-one relationship between observed and estimated LAI showed an RMSE of 0.63 (**Figure 6B**). Average remotely-sensed estimated LAI values significantly increased from April 4th to April 30th, but slightly decreased at the third acquisition date (May 22nd) (**Table 3**). Differences in LAI between irrigation treatments were also significant for all image acquisition dates ($P < 0.001$). In contrast with the values observed for flights conducted at anthesis (April 30th) and grain-filling (May 22nd), the LAI values of the rainfed treatment at the jointing stage (April 4th) were the highest. Estimates of plant height through remote sensing were significant, with R^2 of 0.95 and RMSE of 0.18 m when aggregating data from the three dates (**Figure 6C**). Averaged observed values of PH ranged from 0.40 to 0.96 m, respectively, for the first (April 4th) and last (May 22nd) flights. In all dates, the remotely-sensed assessments underestimated the actual PH. Estimates of canopy transpiration (T) obtained through the TSEB model were compared against those measured at leaf level. The regression obtained aggregating data from the three dates was significant ($R^2 = 0.50$, $P < 0.001$) with an RMSE of 0.24 mm/h (**Figure 6D**). Differences in ETa and T between irrigation treatments were also significant for all dates (**Table 3**). Similarly to LAI, the highest and lowest values of ETa and T were respectively identified in the 100%ETc and rainfed treatments, with the partial exception of the flight conducted on April 4th (jointing stage), when the values were inverted.

Relationships Between Agronomic and Remotely-Sensed Traits

Regression analyses were carried out using the aggregated yield of the three irrigation treatments as dependent variable and each of the four traits assessed by remote sensing in each flight event as explanatory ones (**Figure 7**). The results show that LAI, estimated at jointing, could not predict yield. However, the relationships between yield and H, ETa, and T were negative and statistically significant in this first flight (**Figure 7A**). Significant and positive relationships were obtained between yield and remotely-sensed estimated traits on the other two image acquisition dates. With the exception of H, R^2 tended to be slightly higher on the last date (May 22nd), accounting for between 82 and 90% of yield variability (**Figure 7C**). ETa was the parameter which showed the

TABLE 3 | Mean values ± SE for LAI, daily evapotranspiration (ETa), and daily transpiration (T) assessed by remote sensing imagery for each flight date and for each water input treatment.

Irrigation treatment	April 4th		April 30th		May 22nd	
	Mean	Range	Mean	Range	Mean	Range
LAI						
100%ETc	1.66 ± 0.03^{c}	1.24–2.18	3.87 ± 0.05^{a}	2.46–5.11	3.80 ± 0.07^{a}	2.10–5.29
50%ETc	1.84 ± 0.02^{b}	1.47–2.37	3.23 ± 0.05^{b}	2.48–4.21	2.66 ± 0.06^{b}	1.72–3.73
Rainfed	2.03 ± 0.04^{a}	1.11–2.61	2.05 ± 0.04^{c}	1.33–2.98	1.01 ± 0.03^{b}	0.51–1.72
ET_a (mm/day)						
100%ETc	4.88 ± 0.04^{c}	4.11–5.77	6.80 ± 0.03^{a}	6.00–7.22	7.15 ± 0.07^{a}	4.79–7.66
50%ETc	5.41 ± 0.03^{b}	4.72–6.04	6.41 ± 0.04^{a}	5.43–7.03	5.53 ± 0.09^{b}	3.49–7.15
Rainfed	5.66 ± 0.06^{a}	4.63–6.95	4.74 ± 0.08^{b}	3.05–5.85	3.78 ± 0.03^{c}	3.15–4.69
T (mm/day)						
100%ETc	2.45 ± 0.02^{c}	2.06–2.89	4.45 ± 0.03^{a}	3.98–5.01	5.47 ± 0.06^{a}	3.90–6.68
50%ETc	2.56 ± 0.02^{b}	2.23–2.92	4.49 ± 0.03^{a}	3.93–5.03	4.55 ± 0.08^{b}	3.38–6.30
Rainfed	2.67 ± 0.03^{a}	1.86–3.03	3.73 ± 0.06^{b}	2.68–4.89	2.06 ± 0.05^{c}	1.01–3.30

Means within columns and trait with different letters are significantly different for a Tukey test at $P < 0.05$.

highest R^2 with yield for the last two image acquisition dates. Although the R^2 of the yield vs. T regressions were also high, the values were slightly lower in comparison with those obtained between yield and ETa.

For a deeper analysis, the same relationships were examined for each irrigation treatment separately. Significant associations ($P < 0.05$) between remotely-sensed traits and yield were only found for the 100%ETc treatment (**Figure 8**). The non-significant relationships of other irrigation treatments were probably due to the lower range of yield values obtained in them, as shown in **Table 1**. In addition, the relationships between H and yield in the 100%ETc treatment were also not significant (data not shown). The accuracy of fitting yield to LAI, ETa, and T varied between dates (**Figure 8**). The only trait significantly related with yield in the three dates was ETa, which at anthesis and grain filling accounted for 68% of yield variations (**Figures 8B,C**). On the other hand, the relationship between T and yield was also slightly lower in comparison with ETa.

Most of the relationships between traits estimated through remote sensing (ETa, T, H, and LAI) and the agronomic traits other than yield were statistically significant when the data of the three irrigation treatments were aggregated for the analyses (**Table 4**). The second and third flights led to the largest number of significant and positive correlation coefficients, in contrast with the negative associations obtained in the first flight, as observed previously for the relationships with yield. The largest Pearson correlation coefficients ($r > 0.80$, $P < 0.001$) corresponded to the relationships between ETa, T, and LAI with DH, DA, and DM, particularly during the third image acquisition data (grain filling) (**Table 4**). Correlation coefficients between the four remotely-sensed traits and PH were also positive and significant in the two later flights. For the yield components, the largest correlation coefficients appeared for the relationships between ETa, T, and LAI with NGS, and between H and both PH and TKW. Traits assessed from remote sensing could not properly estimate GFR when data of the three irrigation treatments were analyzed at once. Strong relationships were observed between remotely-sensed and phenological traits (**Table 4**).

On the other hand, the correlation coefficients calculated for each irrigation treatment separately showed a completely different picture. The number of statistically significant associations between remotely-sensed traits and yield-related traits was much larger for the 50 and 100%ETc treatments than for the rainfed treatment (**Table 5**). The largest r value obtained for the non-irrigated treatment corresponded to the relationship between LAI and PH during the flight carried out on April 4th ($r = 0.65$, $P < 0.01$), but this relationship was not confirmed in the subsequent image acquisition dates. ETa was negatively and significantly associated with NGS on the first and second image acquisition dates, but this relationship was not significant on the third date (**Table 5**). Moreover, positive and significant correlation coefficients appeared between ETa and PH on the first and third dates, but not the second. For the 50%ETc treatment, significant correlation coefficients were found for all yield-related traits on at least one image acquisition event, with the exception of the NSm^2 and GFD which were not associated with any remotely-sensed estimated trait on any date (**Table 5**). Predicted plant height (H) was significantly and negatively correlated with NGS and positively with GFR on the three image acquisition dates and T was significantly correlated with DH and DA on the second and third dates. No significant relationships were found between remotely-sensed traits assessed at jointing and yield-related traits in the 100%ETc treatment, but H estimated on the second and third dates was significantly related with NGS, TKW, GFR, and PH (**Table 5**). ETa was positively associated with PH, DH, and DA but negatively with GFD on the second and third dates. GFD was significantly and negatively correlated with the four remote sensing traits on the third date, and with ETa and LAI on the second.

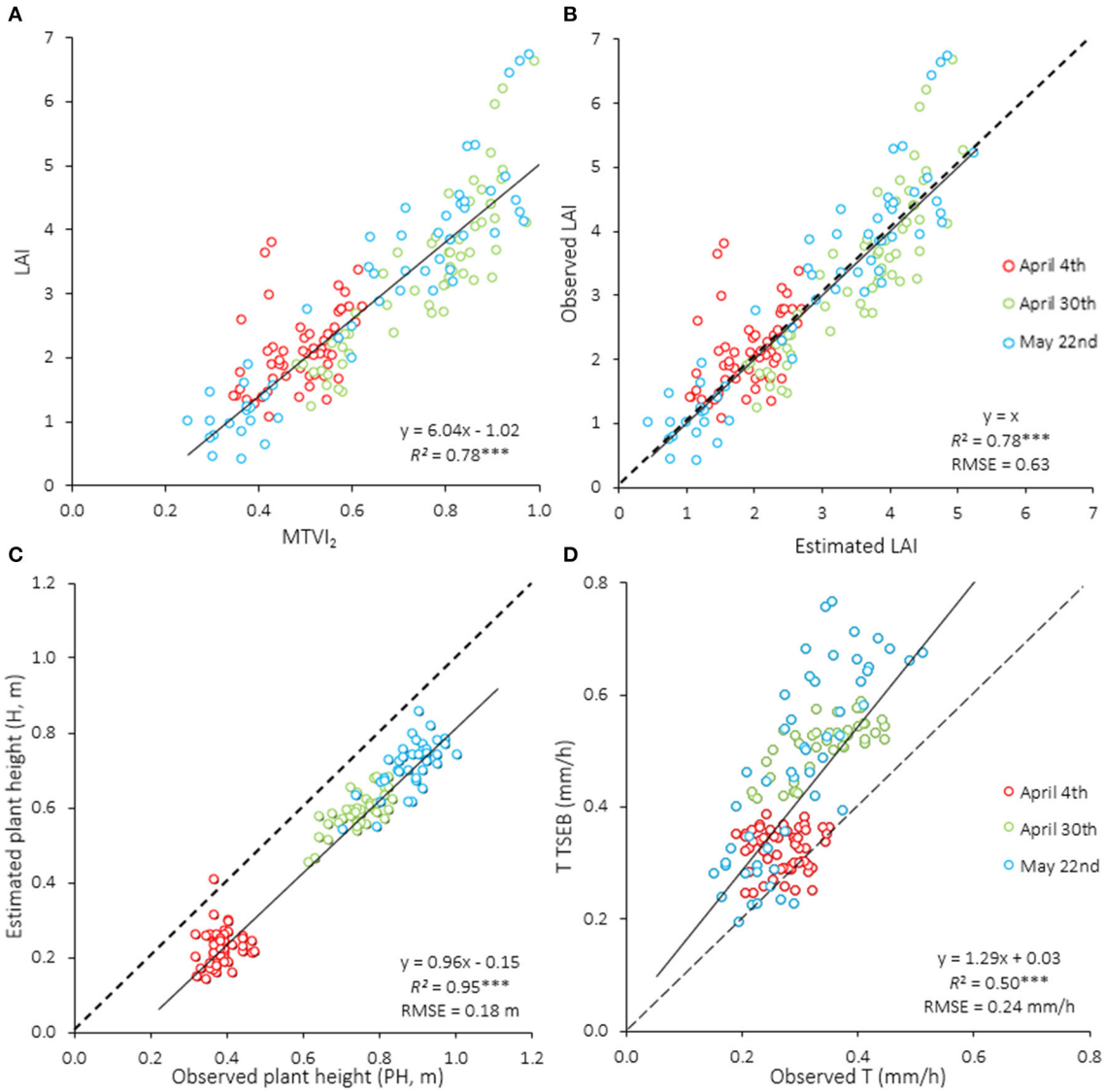

FIGURE 6 | Relationships between **(A)** leaf area index (LAI) and the $MTVI_2$ spectral vegetation index, **(B)** observed and estimated LAI by the $MTVI_2$, **(C)** observed and estimated plant height by the photogrammetric 3D point cloud, and **(D)** observed leaf transpiration with the IRGA device and estimated plant transpiration (T) through the TSEB model. Different colors indicate different dates of image acquisition ($n = 171$). RMSE, root mean square error. ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.

Assessment of Genotypic Differences

Figure 7 shows that the largest difference between genotypes for the remotely-sensed traits was recorded in the third flight. The comparison of genotypic values for each trait and irrigation treatment for that flight event showed that, although genotypes differed in their yield at each irrigation treatment, the discrimination power of the remotely-sensed traits varied depending on the water available for the crop (**Table 6**). Genotypic differences were not statistically significant for H in the rainfed treatment in which significant differences were obtained for LAI, Eta, and T. Genotypes did not differ in ETa and T in either of the irrigated treatments and in LAI in the 50%ETc treatment. In agreement with the positive relationships between yield and either LAI, Eta, and T (**Figure 8**), the genotypes that reached the highest yields in the 100%ETc treatment (Euroduro, Anvergur, Grador, and others shown in **Table 6**) tended to have superior values for these three traits, while the lowest yielding ones showed low values for them. However, this trend was not observed for the 50%ETc treatment, where the significant differences detected for yield between the Sculptur and Claudio cultivars, which obtained the highest yields, and the Don Sebastian cultivar, which gave the least yield, were not associated with specific values of the remotely-sensed traits. Even though the rainfed treatment had the highest discrimination power between cultivars, genotypic differences in remotely-sensed traits were also independent of yield variations.

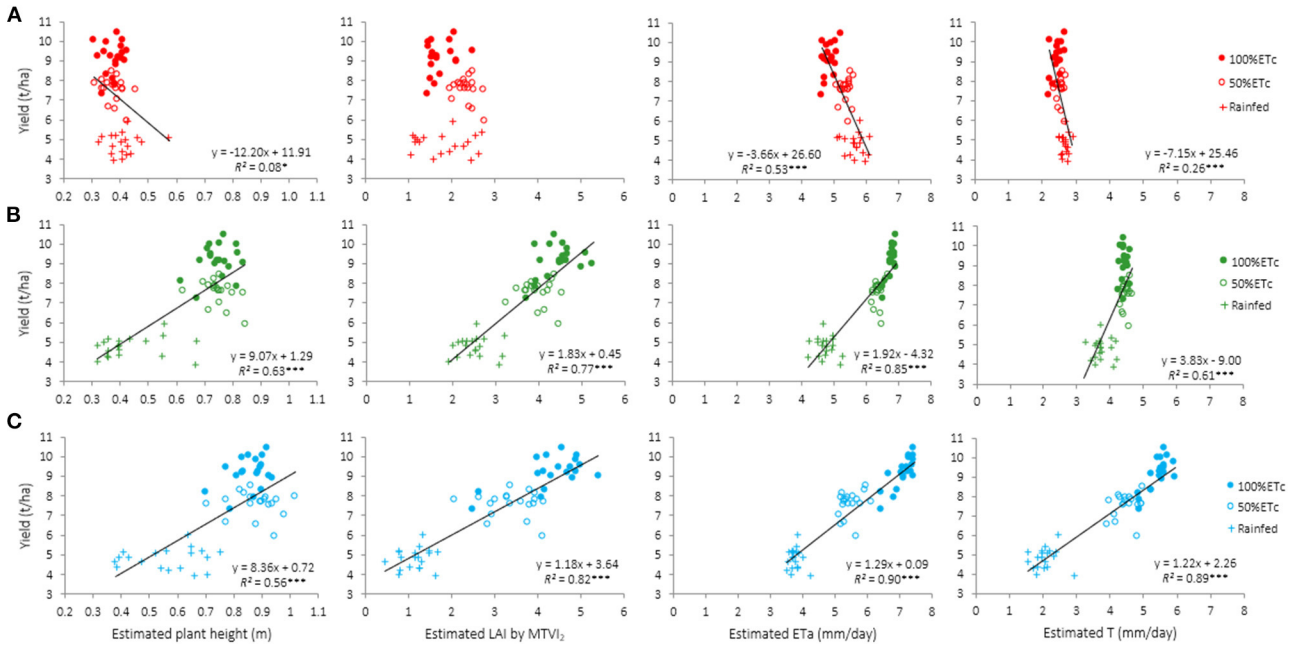

FIGURE 7 | Relationships between estimated plant height, LAI, actual evapotranspiration (ETa), and transpiration (T) assessed from remote sensing imagery and yield for the three irrigation treatments (full circles, 100%ETc; empty circles 50%ETc; +, Rainfed) for each image acquisition event: **(A)** April 4th (jointing), **(B)** April 30th (anthesis), and **(C)** May 22nd (grain filling). Each point represents the mean value of a genotype across replications ($n = 57$). ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.

DISCUSSION

The present study provided a quantifiable assessment of UAV imagery for the purpose of obtaining an accurate estimation of the agronomic performance of durum wheat from the field phenotyping of 19 durum wheat genotypes grown under three contrasting water regimes. The proposed method employs the TSEB model to estimate differences between irrigation treatments and genotypes in actual crop evapotranspiration (ETa) and transpiration (T). The biophysical parameters of the vegetation, such as LAI and canopy height (H) were respectively estimated through spectral vegetation indices and photogrammetry. The feasibility of using this methodology for high-throughput field phenotyping has been demonstrated, since it is robust, repeatable and time, and cost efficient compared with measurements made at ground level.

Effect of Water Availability on Durum Wheat Field Performance

The experimental site is representative of the Mediterranean climate, with a long-term mean temperature of 10.4°C and average rainfall of 248 mm from November to June. This mean temperature was recorded for the 2018–2019 growing season, but rainfall was slightly lower than average. This water scarcity allowed the testing of three contrasting irrigation treatments. Results indicate that a water input of 450 mm (rainfall + irrigation), most of which was supplied during the spring, was enough to cover all the evapotranspiration needs of the durum wheat crop (**Figure 5**). A reduction of 36.7% in the water supplied (285 mm) covered half of these needs (50%ETc), while the non-irrigated treatment (122 mm of rainfall) represented 27.1% of the water needed to meet evapotranspiration needs. The analyses of the effects of water constraints on grain yield revealed that supplying 63.3% (50%ETc) or 27.1% (rainfed treatment) of the water needed to cover the whole crop evapotranspiration needs, resulted in yields that corresponded to 81.7 and 52.0%, respectively, of the yield obtained in the full irrigation treatment (100%ET). Karam et al. (2009) obtained yield decreases between 25 and 28% in rainfed and half-irrigated durum wheat compared with a full-irrigated treatment. These results suggest that durum wheat could be an alternative for irrigated areas with low seasonal water availability, since a reduction of 36.7% in water input decreased yield by only 18.3%, and a water reduction of 72.9% diminished yield by 48.0%. In terms of water productivity (WP), the values were 2.05, 2.64, and 3.93 kg of DM grain/m^3 of water applied in the 100%ETc, 50%ETc, and rainfed treatments, respectively. These results reflect the efforts made by breeders to improve drought tolerance of modern durum wheat cultivars adapted to drought-prone environments (Araus et al., 2003), where yield differences between drought-tolerant and drought-sensitive ideotypes are evident (Senapati et al., 2019).

The 18.3% yield reduction observed in the 50%ETc treatment when compared with the 100%ETc was a consequence of decreases of 15.0% in NGS and 3.8% in TKW, as the NSm^2 was not affected. In addition, the 48.0% yield decrease of the rainfed treatment in comparison with the fully-irrigated one was due to a reduction of 33.7% in NGS, 9.8% in NSm^2, and 7.7% in TKW. In this study, the larger LAI values estimated for the

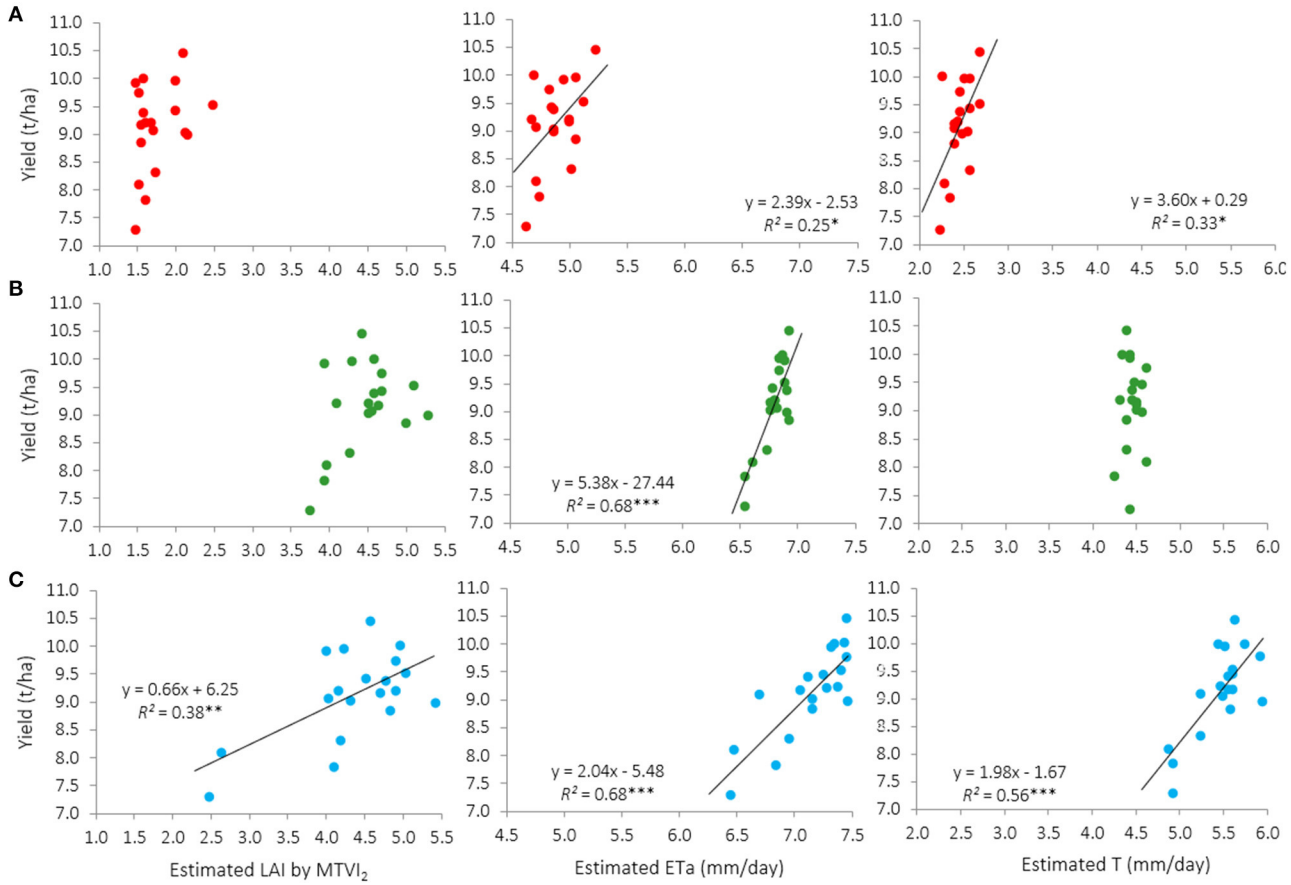

FIGURE 8 | Relationships between estimated LAI, actual evapotranspiration (ETa), and transpiration (T) assessed from remote sensing images and yield for the 100%ETc treatment for each image acquisition event: **(A)** April 4th (jointing), **(B)** April 30th (anthesis), and **(C)** May 22nd (grain filling). Each point represents the mean value of a genotype across replications ($n = 19$). Relationships involving predicted plant height are not shown because $P > 0.05$ in the three image acquisition dates. ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.

TABLE 4 | Pearson correlation coefficients (r) for the relationships between traits assessed through remote sensing and agronomic traits other than yield for each image acquisition event across irrigation treatments ($n = 57$).

Trait	April 4th				April 30th				May 22nd			
	ETa	T	H	LAI	ETa	T	H	LAI	ETa	T	H	LAI
Number of spikes/m^2 (NSm^2)	−0.01	0.08	−0.02	0.25	0.36**	0.38**	0.30*	0.37**	0.28*	0.35**	0.30*	0.35**
Number of grains/spike (NGS)	−0.74***	−0.68***	−0.44***	−0.20	0.59***	0.49***	0.41**	0.54***	0.64***	0.63***	0.41**	0.55***
Thousand kernel weight (TKW)	−0.23	−0.14	0.13	0.11	0.38**	0.27*	0.50***	0.37**	0.39**	0.37**	0.42**	0.38**
Grain filling rate (GFR)	0.28*	0.28**	0.39**	0.07	−0.19	−0.20	−0.03	−0.16	−0.19	−0.18	−0.04	−0.13
Plant height (PH)	−0.40**	−0.20	−0.01	0.11	0.66***	0.55***	0.64***	0.64***	0.69***	0.67***	0.62***	0.65***
Days to heading (DH)	0.65***	−0.48***	−0.24	0.03	0.80***	0.71***	0.66***	0.82***	0.86***	0.85***	0.58***	0.82***
Days to anthesis (DA)	−0.67***	−0.49***	−0.22	−0.02	0.78***	0.67***	0.62***	0.81***	0.85***	0.83***	0.59***	0.81***
Days to maturity (DM)	−0.80***	−0.63***	−0.34	−0.01	0.90***	0.76***	0.74***	0.89***	0.95***	0.92***	0.67***	0.87***
Grain filling duration (GFD)	−0.70***	−0.59***	−0.40*	0.01	0.72***	0.62***	0.63***	0.67***	0.74***	0.71***	0.54***	0.64***

ETa, actual evapotranspiration; T, actual transpiration; H, estimated plant height.
****P < 0.001; **P < 0.01; *P < 0.05.*

TABLE 5 | Significant ($P < 0.05$) Pearson correlation coefficients (r) for the relationships between traits assessed through remote sensing and yield, yield-related traits and crop phenology for each irrigation treatment and each image acquisition event ($n = 19$).

Trait	April 4th				April 30th				May 22nd			
	ET_a	T	H	LAI	ET_a	T	H	LAI	ET_a	T	H	LAI
Irrigation treatment: 100%ETc												
Number of spikes/m^2 (NSm^2)					0.50*					0.48*		
Number of grains per spike (NGS)							−0.53*				−0.69**	
Thousand kernel weight (TKW)							0.63**				0.65**	
Grain filling rate (GFR)							0.59**				0.71***	
Plant height (PH)					0.51*		0.72***		0.59**		0.79***	0.57**
Days to heading (DH)					0.57*			0.47*	0.45*	0.56*		
Days to anthesis (DA)					0.65**			0.49*	0.48*	0.58**		
Grain filling duration (GFD)					−0.71***			−0.49*	−0.59**	−0.64**	−0.59**	−0.57*
Irrigation treatment: 50%ETc												
Number of grains per spike (NGS)	−0.50*	−0.48*	−0.57*	−0.67**			−0.73***				−0.62**	
Thousand kernel weight (TKW)			0.52*	0.48*			0.63**					
Grain filling rate (GFR)			0.49*	0.51*			0.59**				0.53*	
Plant height (PH)											0.52*	
Days to heading (DH)						0.63**		0.48*	0.52*	0.58**		
Days to anthesis (DA)						0.58**		0.47*	0.48*	0.58**		
Days to maturity (DM)					0.52*				0.70***	0.55*		
Rainfed												
Number of spikes/m^2 (NSm^2)									0.59**			
Number of grains per spike (NGS)	−0.57*				−0.56*	−0.46*						
Thousand kernel weight (TKW)							0.46*					
Plant height (PH)	0.50*			0.65**		0.54*			0.54*			
Days to heading (DH)									0.59**			
Days to maturity (DM)								0.63**				
Grain filling duration (GFD)								0.55*				

Traits with nonsignificant r values for any flight have been omitted. ET_a, actual evapotranspiration; T, actual transpiration; H, estimated plant height.
****$P < 0.001$; **$P < 0.01$ *$P < 0.05$.*

rainfed treatment at the jointing stage suggest that the tiller number was probably not strongly affected by drought, which is in agreement with the low reduction of NSm^2 observed in the rainfed treatment. Our results agree with the assumption that NGS is typically the yield component that is most sensitive to drought stress due to severe competition for nutrients during stem elongation (Richards et al., 2001; Kilic and Yagbasanlar, 2010; Liu et al., 2015). Decreases in the NGS from 12.4 to 58.7% have been found in durum wheat under drought stress compared to well-irrigated conditions (Vahamidis et al., 2019).

When compared with the fully-irrigated treatment, the cycle shortening observed in this study ranged between 3 days (2.2%) for DH in the 50%ETc treatment to 13 days (7.1%) for DM in the rainfed treatment. Similar reductions have been reported in the literature (Liu et al., 2015; Varga et al., 2015). The reductions observed in DA and DM in the 50%ETc and rainfed treatments suggest that water stress likely accelerated leaf senescence, which is a common response to water shortage (Ihsan et al., 2016; Pour-Aboughadareh et al., 2020). In relative terms, the greatest shortening was observed in GFD (up to 12.5%), which could not be compensated by the increase of 5.7% in GFR. Decreases of 14% in the duration of grain filling have been reported previously in durum wheat subjected to pre-anthesis drought (Liu et al., 2015). It is well-known that the reduced grain-filling period directly affects grain number and grain size, which largely accounts for the decrease in wheat yields (Dolferus et al., 2011). Plant height decreased 6.5 and 12.0% in the 50%ETc and rainfed treatments, respectively, which is in agreement with the biomass reduction caused by drought shown by previous studies (Pour-Aboughadareh et al., 2020 and references herein).

Predicted vs. Observed Traits

Spectral vegetation indices (VI) assessed from ground level though field spectrometry have been widely used to estimate several wheat traits such as growth status, biomass, yield, or photosynthesis (Aparicio et al., 2000, 2002; Magney et al., 2016). Moreover, UAV-derived VI [e.g., NDVI, soil adjusted VI (SAVI) and optimized soil adjusted VI (OSAVI)] have also been used to estimate the same traits (Yue et al., 2019; Marino and Alvino, 2020), but with the advantages over field spectrometry of generating surface maps in real time, higher flexibility and more convenient operation for estimating plant traits from large

TABLE 6 | Comparison of estimated plant height (H), leaf area index (LAI), actual evapotranspiration (ETa), actual transpiration (T), and yield in each genotype on May 22nd (grain filling).

100%ETc						50%ETc						Rainfed					
Genotype	**H (m)**	**LAI**	**ETa (mm/day)**	**T (mm/day)**	**Yield (t/ha)**	**Genotype**	**H (m)**	**LAI**	**ETa (mm/day)**	**T (mm/day)**	**Yield (t/ha)**	**Genotype**	**H (m)**	**LAI**	**ETa (mm/day)**	**T (mm/day)**	**Yield (t/ha)**
EURODURO	0.94^{a}	4.08^{abc}	7.45	5.63	10.45^{a}	CLAUDIO	0.86^{abc}	2.88	6.00	5.01	8.47^{a}	SCULPTUR	0.61	1.19^{ab}	3.85^{ab}	2.42^{ab}	5.97^{a}
ANVERGUR	0.83^{bcde}	4.18^{abc}	7.45	5.74	10.01^{ab}	SCULPTUR	$0.80a^{bc}$	3.00	5.97	5.25	8.27^{a}	CLAUDIO	0.65	1.02^{ab}	3.83^{ab}	2.08^{ab}	5.40^{ab}
GRADOR	0.92^{ab}	3.74^{abc}	7.35	5.45	9.98^{ab}	ATHORIS	0.78^{bc}	2.50	5.23	4.26	8.05^{ab}	IBERUS	0.59	0.98^{ab}	3.75^{ab}	1.99^{ab}	5.21^{abc}
07D057D4fba	0.84^{abcde}	3.78^{abc}	7.32	5.52	9.95^{abc}	09D066D8cab	0.93^{a}	2.81	5.60	4.63	7.90^{ab}	ANVERGUR	0.59	1.06^{ab}	3.68^{ab}	2.19^{ab}	5.19^{abc}
BURGOS	0.87^{abcd}	4.43^{a}	7.46	5.92	9.75^{abc}	ANVERGUR	0.81^{abc}	2.77	5.20	4.49	7.87^{ab}	CALERO	0.49	0.78^{b}	3.62^{ab}	1.58^{b}	5.13^{abc}
CLAUDIO	0.86^{abcde}	4.14^{abc}	7.41	5.60	9.52^{abcd}	EURODURO	0.90^{ab}	2.35	5.43	4.30	7.85^{ab}	DON RICARDO	0.70	1.16^{ab}	3.82^{ab}	2.38^{ab}	5.11^{abc}
ATHORIS	0.78^{def}	3.76^{abc}	7.26	5.61	9.45^{abcd}	07D057D4fba	0.86^{abc}	2.64	5.75	4.65	7.79^{ab}	BURGOS	0.54	1.07^{ab}	3.91^{ab}	2.16^{ab}	5.08^{abc}
08D010D10cab	0.88^{abcd}	4.07^{abc}	7.13	5.55	9.39^{abcd}	GRADOR	0.86^{abc}	2.18	5.40	4.00	7.75^{ab}	ATHORIS	0.64	0.96^{ab}	3.72^{ab}	1.98^{ab}	5.03^{abcd}
09D066D8cab	0.91^{abc}	3.91^{abc}	7.38	5.47	9.21^{abcd}	CARPIO	0.87^{abc}	2.60	5.33	4.51	7.75^{ab}	08D010D10cab	0.64	1.16^{ab}	4.03^{ab}	2.34^{ab}	4.88^{bcde}
CARPIO	0.86^{abcde}	3.78^{abc}	7.28	5.60	9.20^{abcd}	IBERUS	0.81^{abc}	2.87	6.18	4.95	7.62^{ab}	EURODURO	0.62	0.97^{ab}	3.72^{ab}	1.99^{ab}	4.87^{bcde}
SCULPTUR	0.81^{cdef}	3.99^{abc}	7.06	5.55	9.16^{abcd}	08D010D10cab	0.85^{abc}	2.69	5.33	4.59	7.59^{ab}	07D057D4fba	0.50	0.91^{b}	3.70^{ab}	1.85^{ab}	4.85^{bcde}
09D069D1dcf	0.88^{abcd}	3.73^{abc}	6.70	5.25	9.09^{abcde}	09D069D1dcf	0.86^{abc}	2.79	5.45	4.95	7.59^{ab}	GRADOR	0.52	0.77^{b}	3.56^{ab}	1.57^{b}	4.65^{bcde}
DON RICARDO	0.92^{ab}	3.53^{abc}	7.16	5.48	9.02^{abcde}	CALERO	0.73^{c}	2.52	5.82	4.47	7.58^{ab}	09D069D1dcf	0.53	1.02^{ab}	3.78^{ab}	2.06^{ab}	4.62^{bcde}
IBERUS	0.79^{def}	4.47^{a}	7.46	5.95	8.97^{abcde}	BURGOS	0.84^{abc}	3.12	5.60	4.87	7.54^{ab}	09D066D8cab	0.55	0.92^{b}	3.87^{ab}	1.87^{ab}	4.36^{bcde}
DON SEBASTIAN	0.92^{ab}	4.26^{ab}	7.16	5.58	8.83^{abcde}	DON RICARDO	0.89^{ab}	2.57	5.29	4.40	7.51^{ab}	EUNOBLE	0.60	1.06^{ab}	3.85^{ab}	2.17^{ab}	4.30^{cde}
EUNOBLE	0.89^{abcd}	3.31^{abc}	6.96	5.24	8.31^{bcde}	05D278D1be	0.88^{abc}	2.28	5.18	4.14	7.01^{ab}	CARPIO	0.57	0.98^{ab}	3.70^{ab}	2.01^{ab}	4.29^{cde}
CALERO	0.70^{f}	2.98^{c}	6.48	4.87	8.10^{cde}	TUSSUR	0.73^{c}	2.47	5.34	4.18	6.64^{ab}	TUSSUR	0.48	0.83^{b}	3.51^{b}	1.71^{b}	4.23^{cde}
05D278D1be	0.83^{abcde}	3.11^{bc}	6.85	4.93	7.82^{de}	EUNOBLE	0.86^{abc}	2.24	5.21	3.93	6.48^{ab}	05D278D1be	0.74	0.87^{b}	3.65^{ab}	1.81^{b}	4.00^{de}
TUSSUR	0.76^{ef}	3.01^{c}	6.46	4.91	7.28^{e}	DON SEBASTIAN	0.93^{a}	3.21	5.69	4.83	5.91^{b}	DON SEBASTIAN	0.61	1.48^{a}	4.24^{a}	2.92^{a}	3.91^{e}

Genotypes ordered by yield. Different letters in the same column mean significant differences between genotypes at $p \leq 0.05$ using Tukey's honest significant difference test.

numbers of plots at a time (Lelong et al., 2008; Maimaitijiang et al., 2017). UAV high-resolution VI may detect changes of plant status, thus helping to improve crop monitoring, nitrogen management, and crop yield estimation (Cabrera-Bosquet et al., 2011). However, the relationships between existing VI and biophysical parameters of the vegetation, such as LAI, usually generate an error, in part because some of them saturate at medium-to-high canopy covers, are sensitive to the chlorophyll content or to internal factors such as canopy geometry, leaf and soil optical properties (Baret and Guyot, 1990; Zhou et al., 2017). In this study, the $MTVI_2$ showed a linear relationship with LAI, with an R^2 of 0.78 (**Figure 6A**). This positive linear regression could be attributable to the $MTVI_2$ having a center wavelength located at the red-edge region (717 nm), which is mainly influenced by the plant's structural characteristics and chlorophyll content (Guyot et al., 1992; Yao et al., 2017; He et al., 2020). This suggests that the $MTVI_2$ is not as sensitive to canopy structure as other indices which only use bands at the red and near-infrared regions.

Crop height (H) estimates through photogrammetry were regularly underestimated by 0.18 ± 0.05 m. According to Lechner et al. (2009) and Hengl (2006), the image spatial resolution has to be at least half of the size of the target object to be accurately discriminated though photogrammetric analysis. Therefore, it is possible that the low leaf width of durum wheat, similar to the pixel size (0.02 m), provoked this systematic underestimation (**Figure 6C**). Probably, increasing the number of images acquired from different viewing angles, with a higher overlap, could help to improve H estimates. However, these results are in agreement with those obtained in previous studies at the same spatial resolution in wheat (Holman et al., 2016; Demir et al., 2018), and olive trees (Caruso et al., 2019). Since plant height is one of the necessary ancillary data of the TSEB model, a precise estimation of H is essential to assess plant evapotranspiration.

Estimates of canopy transpiration were validated against leaf transpiration measurements (**Figure 6D**). Although remotely-sensed estimates of T were higher than the measured ones, the relationship had an R^2 of 0.50. The higher T rates assessed through remote sensing were because they were calculated at plot level, whereas the others were calculated only at leaf level. Differences could also be attributable to the likelihood that the partitioning of ET into T and E contains a substantial bias error.

Relationships Between Traits Assessed From Remote Imagery and Agronomic Traits

The sign of the correlation coefficients between remotely-sensed traits and most of the agronomic characteristics were negative in the first flight and positive in the subsequent ones. This was due to the high initial vegetative growth of the plots subjected to the rainfed treatment, as revealed by the LAI, ETa, and T values shown in **Table 3**, which likely reflects the effect of soil variations on the growth of seedlings. It is probable that soil water holding capacity was higher in the area were the rainfed treatment was located. As the season evolved, this trend was reversed and the fully-irrigated treatment showed the highest evapotranspiration rates which previous studies have associated with higher stomatal conductance and photosynthetic rates (Fischer et al., 1998). The highest ETa and T values observed in the 100%ET treatment are in agreement with high yielding wheat cultivars showing higher rates of transpiration (Shimshi and Ephrat, 1975; Reynolds et al., 1994) and with the strong association existing between T and LAI (Blum, 2011).

When the analyses of the relationships between grain yield and the four traits assessed from remote sensing images (H, ETa, T, and LAI) were conducted using the aggregated data of the three irrigation treatments for each flight event, the results clearly show that forecasts were much more accurate at anthesis and grain filling than at jointing (**Figure 7**). Except for H, the correlation coefficients were in general slightly higher in the third flight (May 22nd) than in the second (April 30th), thus suggesting that yield predictions were more accurate when images were captured during grain filling than around anthesis. A lower correlation in LAI was observed at flowering (2nd flight date, 147DAS) in comparison to grain filling (3rd flight date, 169DAS). This can probably be explained by an early senescence reached in some of the genotypes (**Table 1**). Greater variability may explain an increase in the correlation with respect to the second flight date. The analyses conducted for the yield-related traits confirmed that measurements at advanced crop stages were better, as demonstrated in previous studies (Hassan et al., 2018). This was an expected result, as only the potential number of spikes and spikelets per spike are defined at jointing (Simane et al., 1993), while grain setting, grain weight and final yield are determined in subsequent developmental stages (Giunta et al., 1993). NGS, PH, DH, DA, and DM and GFD could be properly assessed through remotely acquired estimates of ETa, T, H, and LAI during grain filling (**Table 4**). The highest R^2 to estimate yield components was observed in ETa rather than with T. It is crucial for ET partitioning to retrieve reliable estimates of canopy and soil temperatures, net radiation, and aerodynamic roughness, with the latter usually obtained from vegetation structural parameters. Therefore, any bias in those estimates could be a source of error when attempting to obtain accurate estimates of T. In addition, the higher range of variability of ETa values in comparison to T contributed to obtaining the highest R^2 when it was regressed with yield. This is because ETa also uses the soil temperature (T_{soil}) of each individual plot and irrigation treatment, with important differences in T_{soil} between irrigated and rainfed plots. On the other hand, predicted plant height (H) was also a good estimator of DH, DA, DM, GFD, PH, and TKW at anthesis, although the values were slightly lower in comparison to the evapotranspiration components. While in this study H was estimated from photogrammetry using multispectral imagery, the advantages of using H instead of ET estimates include the need for fewer inputs, and the lower cost and amount of time needed. Plant height is an essential trait in wheat as it determines the architecture of the plant canopy and has a strong effect on grain number, harvest index and final grain yield (Maccaferri et al., 2008; Liu et al., 2015). The relationships between plant height and yield are environmentally dependent as positive associations have been reported under optimal water conditions and negative associations in water stress environments (Royo

et al., 2008; Dogan, 2009; Talebi et al., 2010). Plant height has been proposed as a potential indicator of tolerance to drought stress under Mediterranean conditions (Liu et al., 2015).

Although the R^2 of the relationships between the assessed parameters and NSm^2 were significant, the weaker relationships obtained for this trait suggest that it cannot be considered a yield component that it is possible to properly estimate through remote imagery. GFR could not be assessed through any remotely-sensed trait, even when considering the aggregated data of the three irrigation treatments (**Table 4**). However, when the relationships between remotely-sensed estimated traits and yield-related traits were analyzed individually for each irrigation treatment, results showed higher accuracy in the irrigated treatments than in the rainfed one (**Table 5**). Although some *r* values obtained from the regressions between remotely-sensed estimated traits and the agronomic ones were significant under rainfed conditions, they did not show consistency among related traits nor across image acquisition dates. This suggests that they could be more casual than causal, and therefore do not demonstrate enough reliability to be recommended for accurate field assessments. This was probably related, as discussed previously, with the wider range of values observed for most traits in the irrigated treatments when compared with the rainfed treatment, as shown in **Table 1**, which increased the predictability of remote sensing imagery. The comparison of the number of significant correlation coefficients obtained in each irrigated treatment in the second and third flights and their values revealed that assessments made in the 100%ETc treatment showed more significant correlation coefficients and with higher values than the ones made in the 50%ETc treatment (**Table 5**). As regards the growth stage most appropriate for predicting yield-related traits, the number of significant correlation coefficients indicated that, as in the case of yield, the second (at anthesis) and third (during grain filling) image acquisition dates were the most suitable, but a higher number of positive associations were found for the third flight.

In relation to the agronomic traits that can be properly assessed by remote sensing imagery, the negative and significant correlation coefficients between H and NGS in the two irrigated treatments and the two later flights (*r* values from −0.53 to −0.73) suggest a causal and consistent association. Similarly, H showed a positive and significant association with GFR in the same four cases, with r values ranging from 0.59 to 0.71, thus indicating a good predictive value. The analysis of the relationships between predicted plant height and TKW showed less consistent results given that the correlation coefficient obtained for the third flight in the 50%ETc treatment was not statistically significant. DH and DA were among the phenological characteristics that were most consistently related with remotely-sensed traits, with ETa being the best predictor for them, mostly during the third flight. GFD was negatively and consistently related with the four remotely-sensed traits estimated from images acquired during the grain filling stage in the fully-irrigated treatment.

Remote sensing imagery has been widely used to assess yield-related traits under a wide range of phenotypical variations (Aparicio et al., 2002; Haghighattalab et al., 2016; Caruso et al., 2019). In the current study, when data of the three irrigation treatments were analyzed together, the yield ranged between 3,905 and 10,446 kg/ha. Previous studies showed that this very wide range of variability is exceptional for durum wheat genotypes grown in the same site where this study was carried out when subjected to a common agronomic management, under both irrigated and rainfed conditions (Aparicio et al., 2000). For this reason, we also decided to assess the suitability of remotely-sensed estimated traits within each irrigation treatment, given that such homogeneous environmental conditions are more representative of real-world cropping systems. In this case, the results showed that yield could only be properly forecasted in the 100%ETc treatment (**Figure 8**). The lack of water restrictions probably allowed the genotypes to express their potentialities, thus maximizing phenotypic differences as shown by the wider range of yields observed in the 100%ETc treatment (3,172 kg/ha) compared with the 50%ETc (2,559 kg/ha) and the rainfed treatment (2,067 kg/ha). Previous studies have also demonstrated that the capacity of spectral reflectance indices to forecast durum wheat grain yield was higher in locations where genotypes reached potential yields (Royo et al., 2003). Under full irrigation conditions (100%ETc), the results of this study also indicate that ETa was the best predictor of yield, particularly when image acquisition was performed around anthesis or during grain filling. On both dates, it accounted for about 68% of yield variations (**Figure 8**).

Capacity of Remotely-Sensed Traits to Discriminate Among Genotypes

The analysis of the data for each genotype provided by the remotely-sensed traits assessed during grain filling gave a wide range of values for all of them. However, in some cases the differences were not wide enough to be statistically significant (**Table 6**). For LAI, ETa, and T, the highest statistical significance was obtained in the rainfed treatment. Although the pattern behind these results was not totally clear, the relatively wider range of values recorded in the rainfed treatment when compared with the irrigated ones could partially explain these differences. Though the absolute values of LAI, ETa, and T were greater in the irrigated treatments than in the rainfed one, in relative terms the differences between the values of the genotypes showing the highest and the lowest value for each trait were largest in the latter. For instance, in the rainfed treatment, the *T* value of cv. Don Sebastian (2.92 mm/day) was 86% superior to that of cv. Grador, which showed the lowest estimate (1.57 mm/day). This relative difference, which was superior to that obtained in the 100%ETc (17.9%) and the 50%ETc (33.5%) treatments was large enough to prove statistically that these two genotypes differed for this trait. Similarly, the relative wider variations among the extreme values for ETa and LAI obtained in the rainfed treatment than in the irrigated ones support the differences obtained in statistical significances. In the case of H, the lack of differences between genotypes in the rainfed treatment could not be attributed to the same reason, as the relative difference in H values was 54%, larger than that observed in the irrigated treatments (23.7 and 27.4% in 100 and 50%ETc, respectively) where statistically significant differences were detected. In this

case, the reason could likely be the low H values in the rainfed treatment resulting from the short plants, associated to drought environments (Madec et al., 2017), and the underestimation of actual plant height occasioned by the methodology employed which was in accordance with previous studies (Holman et al., 2016).

According to other studies that related durum wheat transpiration and yield (Medina et al., 2019), the genotypes with the highest yields in the fully-irrigated treatment showed superior LAI, ETa, and T values. A high LAI in durum wheat genotypes at the milk-grain growth stage denotes a delay of leaf senescence after anthesis, a characteristic that has been positively related with grain yield (Borojevic et al., 1980), thus underlining its importance as a grain yield determining feature. The high values for ETa and T in high-yielding genotypes are in agreement with the positive associations found between T and both leaf area and biomass in wheat grown in well-watered environments (Blum, 2011).

CONCLUSIONS

This study shows the feasibility of using the two-source energy balance (TSEB) with very high resolution imagery to assess differences in the evapotranspiration components of a durum wheat panel. For this purpose, biophysical parameters of the vegetation were successfully estimated from multispectral imagery. Plant height and LAI estimates gave RMSE values of 0.18 m and 0.63, respectively. Significant differences in durum wheat yield and yield components were observed between irrigation treatments. The 50%ETc and rainfed treatments accounted for respective yield reductions of 18.3 and 48.0% in comparison with the treatment that met all crop water requirements (100%ETc). ETa was the remotely-sensed parameter that, when estimated either at anthesis or during grain filling, showed a positive relationship and the highest R^2 with yield, DH, DA, and GFD. When data were analyzed individually for each irrigation treatment, consistent and positive associations were found between ETa and yield, DH and DA and negative associations with GFD in the 100%ETc treatment, but not in the other treatments. The remotely-sensed traits that were assessed were able to discriminate among genotypes, but the significance of the differences depended on the irrigation treatment. As a conclusion, this study demonstrates that remotely-sensed estimates of ETa through the TSEB model are the best predictor of yield components. R^2 values at the grain filling stage were higher in comparison with other remotely-sensed trait estimates such as height, LAI or spectral vegetation indices.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

CR led the project. DG-C, CR, and JB conceived the manuscript, analyzed the data, and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the Operative Group DURCAT (56.21.006.2017 3A) supported by action 16.01.01 (Cooperation for Innovation) of the Catalan Program of Rural Development 2014–2020 and the CERCA programme/Generalitat de Catalunya (http://cerca.cat/). In addition, the manuscript was written in the context of the Marie 550 Sklodowska-Curie Research and Innovation Staff Exchange (RISE) action through the ACCWA project: 551 grant agreement No.: 823965.

ACKNOWLEDGMENTS

The authors would like to thank the skilled technicians of the Sustainable Field Crops and Efficient Use of Water in Agriculture programs of the IRTA. The support of Dr. Fanny Álvaro and Dr. Dolors Villegas in the project is also acknowledged. The authors also acknowledge the invaluable help of Christian Jofre-Cekalović and Ana Pelechá for image processing and of Dr. Hector Nieto for providing access to TSEB model scripts.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.658357/full#supplementary-material

REFERENCES

Agisoft (2020). *Agisoft (Agisoft PhotoScan User Manual Professional Edition) (2020). Version 1.6.2*. AgiSoft LLC. Available online at: http://www.agisoft.com/downloads/user-manuals/ (accessed October 10, 2020).

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). *Crop evapotranspiration guidelines for computing crop water requirements*. Irrigation and Drainage Paper 56. FAO, Rome.

Allen, R. G., Tasumi, M., and Trezza, R. (2007). Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)–Model. *J. Irrig. Drain. Eng.* 133, 380–394. doi: 10.1061/(ASCE)0733-9437(2007)133:4(380)

Aparicio, N., Villegas, D., Araus, J. L., Casadesús, J., and Royo, C. (2002). Relationship between growth traits and spectral reflectance indices in durum wheat. *Crop Sci.* 42,1547–1555. doi: 10.2135/cropsci2002.1547

Aparicio, N., Villegas, D., Casadesús, J., Araus, J. L., and Royo, C. (2000). Spectral vegetation indices as non-destructive tools for determining durum wheat yield. *Agron. J.* 92, 83–91. doi: 10.2134/agronj2000.92183x

Araus, J. L., Bort, J., Steduto, P., Villegas, D., and Royo, C. (2003). Breeding cereals for Mediterranean conditions: ecophysiological clues for biotechnology application. *Ann. Appl. Biol.* 142, 129–141. doi: 10.1111/j.1744-7348.2003.tb00238.x

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Araus, J. L., Kefauver, S. C., Zaman-Allah, M., Olsen, M. S., and Cairns, J. E. (2018). Translating high-throughput phenotyping into genetic gain. *Trends Plant Sci.* 23, 451–466. doi: 10.1016/j.tplants.2018.02.001

Baret, F., and Guyot, G. (1990). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* 35, 161–173. doi: 10.1016/0034-4257(91)90009-U

Bastiaanssen, W. G. M., Pelgrum, H., Wang, J., Ma, Y., Moreno, J. F., Roerink, G. J., et al. (1998). Remote sensing surface energy balance algorithm for land (SEBAL): 2. Validation. *J. Hydrol.* 212, 213–229. doi: 10.1016/S0022-1694(98)00254-6

Bellvert, J., Jofre-Cekalović, C., Pelechá, A., Mata, M., and Nieto, H. (2020). Feasibility of using the two-source energy balance model (TSEB) with Sentinel-2 and Sentinel-3 images to analyze the spatio-temporal variability of vine water status in a vineyard. *Remote Sens.* 12:2299. doi: 10.3390/rs12142299

Bellvert, J., Nieto, H., Pelechá, A., Jofre-Cekalović, C., Zazurca, L., and Miarnau, X. (2021). Remote sensing energy balance model for the assessment of crop evapotranspiration and water status in an almond rootstock collection. *Front. Plant Sci.* 12:608967. doi: 10.3389/fpls.2021.608967

Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., and Bareth, G. (2014). Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. *Remote Sens.* 6, 10395–10412. doi: 10.3390/rs61110395

Blum, A. (2011). *Plant Breeding for Water-Limited Environments*. New York, NY: Springer, 255. doi: 10.1007/978-1-4419-7491-4

Borojevic, S., Cupina, T., and Krsmanovic, M. (1980). Green area parameters in relation to grain-yield of different wheat genotypes. *Zeitschrift fur Pflanzenzuchtung* 84, 265–283.

Cabrera-Bosquet, L., Molero, G., Stellacci, A., Bort, J., Nogués, S., and Araus, J. (2011). NDVI as a potential tool for predicting biomass, plant nitrogen content and growth in wheat genotypes subjected to different water and nitrogen conditions. *Cereal Res. Commun.* 39, 147–159. doi: 10.1556/CRC.39.2011.1.15

Caruso, G., Zarco-Tejada, P. J., González-Dugo, V., Moriondo, M., Tozzini, L., Palai, G., et al. (2019). High-resolution imagery acquired from an unmanned platform to estimate biophysical and geometrical parameters of olive trees under different irrigation regimes. *PLoS ONE* 14: e0210804. doi: 10.1371/journal.pone.0210804

Chapman, S., Merz, T., Chan, A., Jackway, P., Hrabar, S., Dreccer, M. F., et al. (2014). Pheno-copter: a low-altitude, autonomous remote-sensing robotic helicopter for high-throughput field-based phenotyping. *Agronomy* 4, 279–301. doi: 10.3390/agronomy4020279

Cobb, J. N., DeClerck, G., Greenberg, A., Clark, R., and McCouch, S. (2013). Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. *Theor. Appl. Genet.* 126, 867–887. doi: 10.1007/s00122-013-2066-0

Condon, A. G. (2004). Breeding for high water-use efficiency. *J. Exp. Bot.* 55, 2447–2460. doi: 10.1093/jxb/erh277

Condon, L. E., and Maxwell, R. M. (2014). Groundwater-fed irrigation impacts spatially distributed temporal scaling behavior of the natural system: a spatio-temporal framework for understanding water management impacts. *Environ. Res. Lett.* 9:034009. doi: 10.1088/1748-9326/9/3/034009

Costa, J. M., Grant, O. M., and Chaves, M. M. (2013). Thermography to explore plant–environment interactions. *J. Exp. Bot.* 64, 3937–3949. doi: 10.1093/jxb/ert029

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

Demir, N., Sönmez, N. K., Akar, T., and Ünal, S. (2018). Automated measurement of plant height of wheat genotypes using a DSM derived from UAV Imagery. *Proceedings* 2, 1–5. doi: 10.3390/ecrs-2-05163

Dhondt, S., Wuyts, N., and Inzé, D. (2013). Cell to whole-plant phenotyping: the best is yet to come. *Trends Plant Sci.* 18, 428–439. doi: 10.1016/j.tplants.2013.04.008

Dogan, R. (2009). The correlation and path coefficient analysis for yield and some yield components of durum wheat (Triricum turgidum spp. durum L.) in west Anatolia conditions. *Pak. J. Bot* Pakistan Botanical Society. 41, 1081–1089.

Dolferus, R., Ji, X., and Richards, R. A. (2011). Abiotic stress and control of grain number in cereals. *Plant Sci.* 181, 331–341. doi: 10.1016/j.plantsci.2011.05.015

FAO (2017). *Food Outlook: Biannual Report on Global Food Markets*. Available online at: http://www.fao.org/3/a-i7343e.pdf (accessed March 24, 2020).

Fiorani, F., and Schurr, U. (2013). Future scenarios for plant phenotyping. *Annu. Rev. Plant Biol.* 64, 267–291. doi: 10.1146/annurev-arplant-050312-120137

Fischer, R. A., Rees, D., Sayre, K. D., Lu, Z. M., Condon, A. G., and Saavedra, A. L. (1998). Wheat yield progress associated with higher stomatal conductance and photosynthetic rate, and cooler canopies. *Crop Sci.* 38, 1467–1475. doi: 10.2135/cropsci1998.0011183X003800060011x

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W. J., et al. (2013). Evaluation of climate models. *Clim. Change* 5, 741–866. doi: 10.1017/CBO9781107415324.020

Furbank, R., and Tester, M. (2011). Phenomics–technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Giunta, F., Motzo, R., and Deidda, M. (1993). Effect of drought on yield and yield components of durum wheat and triticale in a Mediterranean environment. *Field Crops Res.* 33, 399–409. doi: 10.1016/0378-4290(93)90161-F

Gonzalez-Dugo, V., Hernandez, P., Solis, I., and Zarco-Tejada, P. J. (2015). Using high-resolution hyperspectral and thermal airborne imagery to assess physiological condition in the context of wheat phenotyping. *Remote Sens.* 7, 13586–13605. doi: 10.3390/rs71013586

Gutman, G., and Ignatov, A. (1998). The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *Int. J. Remote Sens.* 19, 1533–1543. doi: 10.1080/014311698215333

Guyot, G., Baret, F., and Jacquemoud, S. (1992). "Imaging spectroscopy for vegetation studies," in *Imaging Spectroscopy: Fundamentals and Prospective Applications*, eds F. Toselli and J. Bodechtel (New York, NY: Kluwer Academic), 145–165.

Haghighattalab, A., Perez, L. G., Mondal, S., Singh, D., Schinstock, D., Rutkoski, J., et al. (2016). Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. *Plant Methods* 12:15. doi: 10.1186/s13007-016-0134-6

Hassan, M., Yang, M., Rasheed, A., Jin, X., Xia, X., Xiao, Y., et al. (2018). Time-series multispectral indices from unmanned aerial vehicle imagery reveal senescence rate in bread wheat. *Remote Sens.* 10:809. doi: 10.3390/rs10060809

He, L., Ren, X., Wang, Y., Liu, B., Zhang, H., Liu, W., et al. (2020). Comparing methods for estimating leaf area index by multi-angular remote sensing in winter wheat. *Sci. Rep.* 10:13943. doi: 10.1038/s41598-020-70951-w

Hengl, T. (2006). Finding the right pixel size. *Comput. Geosci.* 32, 1283–1298. doi: 10.1016/j.cageo.2005.11.008

Holman, F. H., Riche, A. B., Michalski, A., Castle, M., Wooster, M. J., and Hawkesford, M. J. (2016). High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. *Remote Sens.* 8:1031. doi: 10.3390/rs8121031

Ihsan, M., El-Nakhlawy, F. S., Ismail, S. M., Fahad, S., and Daur, I. (2016). Wheat phenological development and growth studies as affected by drought and late season high temperature stress under arid environment. *Front. Plant Sci.* 7:795. doi: 10.3389/fpls.2016.00795

Jackson, R. D., Idso, S. B., Reginato, R. J., and Pinter, P. J. (1981). Canopy temperature as a crop water stress indicator. *Water Resour. Res.* 17, 1133–1138. doi: 10.1029/WR017i004p01133

Jones, H. G. (1999). Use of thermography for quantitative studies of spatial and temporal variation of stomatal conductance over leaf surfaces. *Plant Cell Environ.* 22, 1043–1055. doi: 10.1046/j.1365-3040.1999.00468.x

Karam, F., Kabalan, R., Breidi, J., Rouphael, Y., and Oweis, T. (2009). Yield and water-production functions of two durum wheat cultivars grown under different irrigation and nitrogen regimes. *Agric. Water Manag.* 96, 603–615. doi: 10.1016/j.agwat.2008.09.018

Kilic, H., and Yagbasanlar, T. (2010). The effect of drought stress on grain yield, yield components and some quality traits of durum wheat (Triticum turgidum ssp. durum) cultivars. *Not. Bot. Horti.* Agrobo. 38, 164–170. doi: 10.15835/nbha3814274

Kustas, W., and Anderson, M. (2009). Advances in thermal infrared remote sensing for land surface modeling. *Agric. Forest Meteorol.* 149, 2071–2081. doi: 10.1016/j.agrformet.2009.05.016

Kustas, W. P., and Norman, J. M. (1999). Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover. *Agric. Forest Meteorol.* 94, 13–29. doi: 10.1016/S0168-1923(99)00005-2

Lal, R. (2016). Beyond COP 21: potential and challenges of the "4 per Thousand" initiative. *J. Soil Water Conserv*. 71:20A. doi: 10.2489/jswc.71.1.20A

Lechner, A. M., Stein, A., Jones, S. D., and Ferwerdac, J. G. (2009). Remote sensing of small and linear features: quantifying the effects of patch size and length, grid position and detectability on land cover mapping. *Remote Sens. Environ.* 113, 2194–2204. doi: 10.1016/j.rse.2009.06.002

Lelong, C., Burger, P., Jubelin, G., Roux, B., Labbé, S., and Baret, F. (2008). Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. *Sensors* 8, 3557–3585. doi: 10.3390/s8053557

Leroux, L., Baron, C., Zoungrana, B., Traore, S. B., Lo Seen, D., and Begue, A. (2016). Crop monitoring using vegetation and thermal indices for yield estimates: case study of a rainfed cereal in semi-arid West Africa. *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.* 9, 347–362. doi: 10.1109/JSTARS.2015.2501343

Liu, H., Searle, L. R., Mather, D. E., Able, A. J., and Able, J. A. (2015). Morphological, physiological and yield responses of durum wheat to pre-anthesis water-deficit stress are genotype-dependent. *Crop Pasture Sci*. 66, 1024–1038. doi: 10.1071/CP15013

Maccaferri, M., Sanguineti, M. C., Corneti, S., Ortega, J. L. A., Salem, M. B., and Bort, J. (2008). Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics* 178, 489–511. doi: 10.1534/genetics.107.077297

Madec, S., Baret, F., de Solan, B., Thomas, S., Dutartre, D., Jezequel, S., et al. (2017). High-throughput phenotyping of plant height: comparing unmanned aerial vehicles and ground LiDAR estimates. *Front. Plant Sci*. 8:2002. doi: 10.3389/fpls.2017.02002

Magney, T. S., Eitel, J. U. H., Huggins, D. R., and Vierling, L. A. (2016). Proximal NDVI derived phenology improves inseason predictions of wheat quantity and quality. *Agric. For. Meteorol*. 217, 46–60. doi: 10.1016/j.agrformet.2015.11.009

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 J. Photogramm*. 134, 43–58. doi: 10.1016/j.isprsjprs.2017.10.011

Marino, S., and Alvino, A. (2020). Agronomic traits analysis of ten winter wheat cultivars clustered by UAV-derived vegetation indices. *Remote Sens*. 12:249. doi: 10.3390/rs12020249

McShane, R. R., Driscoll, K. P., and Sando, R. A. (2017). *Review of Surface Energy Balance Models for Estimating Actual Evapotranspiration with Remote Sensing at High Spatiotemporal Resolution over Large Extents*. Scientific Investigations Report. Department of the Interior and U.S. Geological Survey, Reston, VA, United States. doi: 10.3133/sir20175087

Medina, S., Vicente, R., Nieto-Taladriz, M. T., Aparicio, N., Chairi, F., Vergara-Diaz, O., et al. (2019). The plant-transpiration response to vapor pressure deficit (VPD) in durum wheat is associated with differential yield performance and specific expression of genes involved in primary metabolism and water transport. *Front. Plant Sci*. 9:1994. doi: 10.3389/fpls.2018.01994

Moshelion, M., and Altman, A. (2015). Current challenges and future perspectives of plant and agricultural biotechnology. *Trends Biotechnol*. 33, 337–342. doi: 10.1016/j.tibtech.2015.03.001

Nieto, H., Kustas, W. P., Torres-Rúa, A., Alfieri, J. G., Gao, F., Anderson, M. C., et al. (2018). Evaluation of TSEB turbulent fluxes using different methods for the retrieval of soil and canopy component temperatures from UAV thermal and multispectral imagery. *Irrigation Sci*. 37, 389–406. doi: 10.1007/s00271-018-0585-9

Norman, J. M., Kustas, W. P., and Humes, K. S. (1995). Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. *Agric. Forest Meteorol.* 77, 263–293. doi: 10.1016/0168-1923(95)02265-Y

Perich, G., Hund, A., Anderegg, J., Roth, L., Boer, M. P., Walter, A., et al. (2019). Assessment of multi-image UAV based high-throughput field phenotyping of canopy temperature. *Front. Plant Sci*. 11:150. doi: 10.3389/fpls.2020.00150

Pour-Aboughadareh, A., Mohammadi, R., Etminan, A., Shooshtari, L., Maleki-Tabrizi, N., and Poczai, P. (2020). Effects of drought stress on some agronomic and morpho-physiological traits in durum wheat genotypes. *Sustainability* 12:5610. doi: 10.3390/su12145610

Priestley, C. H. B., and Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Rev*. 100, 81–92. doi: 10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2

Ray, D. K., Mueller, N. D., West, P. C., and Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* 8:e66428. doi: 10.1371/journal.pone.0066428

Reynolds, M. P., Balota, M., Delgado, M. I. B., Amani, I., and Fischer, R. A. (1994). Physiological and morphological traits associated with spring wheat yield under hot, irrigated conditions. *Aust. J. of Plant Physiol*. 21, 717–730. doi: 10.1071/PP9940717

Richards, R. A., Condon, A. G., and Rebetzke, G. J. (2001). "Traits to improve yield in dry environments," in *Application of Physiology in Wheat Breeding*, eds M. P. Reynolds, J. I. Ortiz-Monasterio, and A. McNab (Mexico: CIMMYT), 88–100.

Royo, C., Aparicio, N., Villegas, D., Casadesus, J., Monneveux, P., and Araus, J. L. (2003). Usefulness of spectral reflectance indices as durum wheat yield predictors under contrasting Mediterranean environments. *Int. J. Remote Sens*. 24, 4403–4419. doi: 10.1080/0143116031000150059

Royo, C., Martos, V., Ramdani, A., Villegas, D., Rharrabti, Y., and García del Moral, L. F. (2008). Changes in yield and carbon isotope discrimination of Italian and Spanish durum wheat during the 20th century. *Agron. J*. 100, 352–360. doi: 10.2134/agronj2007.0060

Sagan, V., Maimaitijiang, M., Sidike, P., Eblimit, K., Peterson, K. T., Hartling, S., et al. (2019). UAV-based high resolution thermal imaging for vegetation monitoring, and plant phenotyping using ICI 8640 P, FLIR Vue Pro R 640, and thermomap cameras. *Remote Sens*. 11:330. doi: 10.3390/rs11030330

Savva, A. P., Stoutjesdijk, J., Regnier, P. M. A., and Hindkjaer, S. V. (2001). *Irrigation Manual Planning, Development Monitoring and Evaluation of Irrigated Agriculture with Farmer Participation Volume III Module 8*. SAFR/AGLW/DOC/003 Food and Agriculture Organization of the United Nations (FAO) Sub-Regional, Sub-Regional Office for East and Southern Africa (SAFR), Harare, Zimbabwe.

Saxton, K. E., Rawls, W. J., Romberger, J. S., and Papendick, R. I. (1986). Estimating generalized soil-water characteristics from texture. *Soil Sci. Soc. Am. J*. 50:1031. doi: 10.2136/sssaj1986.03615995005000040039x

Senapati, N., Stratonovitch, P., Paul, M. J., and Semenov, M. A. (2019). Drought tolerance during reproductive development is important for increasing wheat yield potential under climate change in Europe. *J. Exp. Bot*. 70, 2549–2560. doi: 10.1093/jxb/ery226

Shi, Y., Thomasson, J. A., Murray, S. C., Pugh, N. A., Rooney, W. L., Shafian, S., et al. (2016). Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLoS ONE* 11:e0159781. doi: 10.1371/journal.pone.0159781

Shimshi, D., and Ephrat, J. (1975). Stomatal behaviour of wheat cultivars in relation to their transpiration, photosynthesis and yield. *Agron. J*. 67, 326–331. doi: 10.2134/agronj1975.00021962006700030011x

Simane, B., Struik, P. C., Nachit, M. M., and Peacock, J. M. (1993). Ontogenic analysis of field components and yield stability of durum wheat in water-limited environments. *Euphytica* 71, 211–219. doi: 10.1007/BF00040410

Talebi, R., Fayyaz, F., Naji, A. M. (2010). Genetic variation and interrelationships of agronomic characteristics in durum wheat under two contrasting water regimes. *Braz. Arch. Biol. Technol.*? 53, 785–791. doi: 10.1590/S1516-89132010000400006

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. U. S. A*. 108, 20260–20264. doi: 10.1073/pnas.1116437108

Vadez, V., Kholova, J., Medina, S., Kakkera, A., and Anderberg, H. (2014). Transpiration efficiency: new insights into an old story. *J. Exp. Bot*. 65, 6141–6153. doi: 10.1093/jxb/eru040

Vahamidis, P., Karamanos, A. J., and Economou, G. (2019). Grain number determination in durum wheat as affected by drought stress: an analysis at spike and spikelet level. *Ann. Appl. Biol*. 174, 190–208. doi: 10.1111/aab.12487

Varga, B., Vida, G., Varga-László, E., Bencze, S., and Veisz, O. (2015). Effect of simulating drought in various phenophases on the water use efficiency of winter wheat. *J. Agro Crop Sci*. 201, 1–9. doi: 10.1111/jac.12087

Walter, A., Liebisch, F., and Hund, A. (2015). Plant phenotyping: from bean weighing to image analysis. *Plant Methods* 11:14. doi: 10.1186/s13007-015-0056-8

Watanabe, K., Guo, W., Arai, K., Takanashi, H., Kajiya-Kanegae, H., Kobayashi, M., et al. (2017). High-throughput phenotyping of sorghum plant height

using an unmanned aerial vehicle and its application to genomic prediction modelling. *Front. Plant Sci.* 8:421. doi: 10.3389/fpls.2017.00421

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crops Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003

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. *Front. Plant Sci.* 8:1111. doi: 10.3389/fpls.2017.01111

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., et al. (2020). Crop phenomics and high-throughput phenotyping: past decades, current challenges and future perspectives. *Mol. Plant* 13, 187–214. doi: 10.1016/j.molp.2020.01.008

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 Sens.* 9:1304. doi: 10.3390/rs9121304

Yue, J., Yang, G., Tian, Q., Feng, H., Xu, K., and Zhou, C. (2019). Estimate of winter-wheat above-ground biomass based on UAV ultrahigh-ground-resolution image textures and vegetation indices. *ISPRSJ. Photogrammetry Remote Sens.* 150, 226–244. doi: 10.1016/j.isprsjprs.2019.02.022

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zhou, H., Wang, J., Liang, S., and Xiao, Z. (2017). Extended data-based mechanistic method for improving leaf area index time series estimation with satellite data. *Remote Sens.* 9:533. doi: 10.3390/rs9060533

Conflict of Interest: 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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 20 May 2021
doi: 10.3389/fpls.2021.609876

Inversion of Winter Wheat Growth Parameters and Yield Under Different Water Treatments Based on UAV Multispectral Remote Sensing

Xin Han[1,2,3], Zheng Wei[1,3], He Chen[1,3], Baozhong Zhang[1,3], Yinong Li[1,3] and Taisheng Du[2]*

[1] State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing, China, [2] College of Water Conservancy and Civil Engineering, China Agricultural University, Beijing, China, [3] National Center of Efficient Irrigation Engineering and Technology Research, Beijing, China

In recent years, the unmanned aerial vehicle (UAV) remote sensing system has been rapidly developed and applied in accurate estimation of crop parameters and yield at farm scale. To develop the major contribution of UAV multispectral images in predicting winter wheat leaf area index (LAI), chlorophyll content (called soil and plant analyzer development [SPAD]), and yield under different water treatments (low water level, medium water level, and high water level), vegetation indices (VIs) originating from UAV multispectral images were used during key winter wheat growth stages. The estimation performances of the models (linear regression, quadratic polynomial regression, and exponential and multiple linear regression models) on the basis of VIs were compared to get the optimal prediction method of crop parameters and yield. Results showed that LAI and SPAD derived from VIs both had high correlations compared with measured data, with determination coefficients of 0.911 and 0.812 (multivariable regression [MLR] model, normalized difference VI [NDVI], soil adjusted VI [SAVI], enhanced VI [EVI], and difference VI [DVI]), 0.899 and 0.87 (quadratic polynomial regression, NDVI), and 0.749 and 0.829 (quadratic polynomial regression, NDVI) under low, medium, and high water levels, respectively. The LAI and SPAD derived from VIs had better potential in estimating winter wheat yield by using multivariable linear regressions, compared to the estimation yield based on VIs directly derived from UAV multispectral images alone by using linear regression, quadratic polynomial regression, and exponential models. When crop parameters (LAI and SPAD) in the flowering period were adopted to estimate yield by using multiple linear regressions, a high correlation of 0.807 was found, while the accuracy was over 87%. Importing LAI and SPAD obtained from UAV multispectral imagery based on VIs into the yield estimation model could significantly enhance the estimation performance. This study indicates that the multivariable linear regression could accurately estimate winter wheat LAI, SPAD, and yield under different water treatments, which has a certain reference value for the popularization and application of UAV remote sensing in precision agriculture.

Keywords: vegetation index, growth parameters, yield, optimal estimation method, different water treatments

OPEN ACCESS

Edited by:
Wenting Han,
Northwest A and F University, China

Reviewed by:
Patompong Johns Saengwilai,
Mahidol University, Thailand
Chunmei Wang,
Aerospace Information Research Institute, Chinese Academy of Sciences, China

***Correspondence:**
Zheng Wei
weiz1983@126.com

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

Received: *24 September 2020*
Accepted: *22 March 2021*
Published: *20 May 2021*

Citation:
Han X, Wei Z, Chen H, Zhang B, Li Y and Du T (2021) Inversion of Winter Wheat Growth Parameters and Yield Under Different Water Treatments Based on UAV Multispectral Remote Sensing. Front. Plant Sci. 12:609876. doi: 10.3389/fpls.2021.609876

INTRODUCTION

The estimation of crop parameters (leaf area index [LAI] and chlorophyll content) is helpful in improving the level of crop monitoring, which is key to accurate monitoring and estimation of crop growth in agricultural management (Huang et al., 2016; Li et al., 2016; Liu et al., 2017; Yebra et al., 2017; Sun et al., 2021). LAI and chlorophyll are often used to describe canopy structure and to predict grain yield (Guo et al., 2020), which requires efficient and rapid measurement of crop LAI and soil and plant analyzer development (SPAD, which is used instead of chlorophyll content). Traditional methods to estimate crop parameter are based on destructive measurement, which not only consume time and manpower but also are difficult to be applied in a large area. In recent decades, remote sensing technology has been successfully applied to crop growth monitoring through satellite platforms, manned airborne platforms, and ground spectral equipment (Michele et al., 2015; Maimaitijiang et al., 2017; Ansar and Muhammad, 2020; Dehkordi et al., 2020). There are two kinds of satellite remote sensing data for crop parameters, namely, optical image and synthetic aperture radar data (Cougo et al., 2015; Castillo et al., 2017; Du et al., 2017; Pham and Yoshino, 2017; Pandit et al., 2018; Li et al., 2019), providing different spatial resolutions, such as SPOT (20 m), MODIS (250 m), Sentinel 1A (10 m), and ALOS-2 PLASAR2 (6 m) (Niu et al., 2019). However, several limitations such as deficient spatiotemporal resolution and cloud cover contamination restrain the application of satellite-based platforms. Relatively, the operation cost of manned airborne platforms is relatively high, and ground-based spectral devices are laborious and suffer from inefficient operations (Zhang and Kovacs, 2012; Yang et al., 2017; Yao et al., 2017; Katja et al., 2018).

In contrast, the rapid development of unmanned aerial vehicle (UAV) platforms provides an economical and efficient method to meet the increasing requirements of spatial, temporal, and spectral resolutions (Yue et al., 2017; Zheng et al., 2018; Heinemann et al., 2020; Qiao et al., 2020). UAV-based multispectral images were adopted to predict crop growth status and to predict grain yield in recent years. For example, Yao et al. (2017) obtained narrowband multispectral images based on UAV and used MTVI2 to estimate wheat LAI effectively, with an accuracy of 0.79 and a relative root mean square error (RMSE) of 24%. Guo et al. (2018) obtained remote sensing images based on UAV and established an inversion model of mangrove LAI by using the vegetation-level interruption index (VLOI), with an inversion accuracy of 0.72 and an RMSE of 0.137. Gao et al. (2016) used a multirotor UAV synchronously carrying a Canon Power Shot G16 digital camera and ADC Lite multispectral sensor to obtain the crown (Tian et al., 2016). Fu et al. (2020) examined the ability of multiple image features derived from UAV RGB images for winter wheat N status estimation across multiple critical growth stages. Another difference of the study on UAV-based prediction of plant LAI was that researchers usually aimed at a few growth periods (Hua et al., 2012). There is still little information on using UAVs to predict plant LAI during the whole important growth stages on a large scale of LAI, and fewer had studied the accuracy comparison of UAV inversion of LAI under different water treatments. The chlorophyll concentration, measured in mass per unit leaf area, is an important biophysical parameter retrievable from reflectance data (Zhu et al., 2020). Tian et al. (2016) used the spectral index of UAV imaging spectrometry to retrieve the chlorophyll concentration of cotton using multiple stepwise regression and partial least squares regression and achieved high accuracy. The Hekou District was selected as the core test area, and 140 ground sampling points were selected. Based on the measured SPAD values and UAV multispectral images, UAV-based SPAD inversion models were constructed, and the most accurate model was selected (Zhang et al., 2019). Cao J. et al. (2020) developed an inversion model that can predict japonica rice chlorophyll content by using the hyperspectral image of the rice canopy collected with a UAV. The inversion model was developed by using an extreme learning machine (ELM), the parameters of which are optimized by using particle swarm optimization (PSO). The PSO-ELM algorithm could accurately model the nonlinear relationship between hyperspectral data and chlorophyll content. The model achieved a coefficient of determination (R^2) of 0.791 and an RMSE of 8.215 mg/L. Furthermore, UAVs are promising remote sensing platforms that is gaining more and more attention for crop studies. For example, Jin et al. (2017) estimated wheat plant density from UAV RGB images. Zhou et al. (2017) estimated grain yield in rice using multitemporal vegetation indices (VIs) from UAV-based multispectral and digital imagery. This parameter has been extensively studied in the field of remote sensing, while research for multispectral data based on UAV is relatively few under different water treatments. The techniques used are mainly based on portable spectrometers, airborne multispectral imagers, and remote sensing satellites (Skudra and Ruza, 2017). Portable spectrometers have difficulty differentiating "point" from "surface," while satellite images have coarser spatial resolution and poor timeliness and are thus prone to "isospectral foreign bodies," resulting in low prediction accuracy. Therefore, the prediction of chlorophyll concentrations and LAI using multispectral sensors on low-altitude UAV remote sensing platforms has gradually become a trend (Cao Y. et al., 2020; Guo et al., 2020; Wan et al., 2020). The existing researchers used the multispectral remote sensing images of medium and low spatial resolutions (such as Landsat 8 and TM) to carry out remote sensing inversion research on winter wheat LAI, SPAD, yield estimation, and other indicators. However, due to the limitations of spatial resolution, revisit period, weather conditions, and other factors, there are still some limitations in the precise monitoring of winter wheat growth.

Therefore, this study is aimed at estimating LAI, SPAD, and yield of winter wheat based on VIs (normalized difference VI [NDVI], soil adjusted VI [SAVI], enhanced VI [EVI], and difference VI [DVI]) derived from UAV multispectral images. The estimation performances of models based on VI alone and VI combinations were also analyzed. According to the obtained optimal estimation of LAI and SPAD values, the crop yield was estimated. More specifically, our study paid attention to the following:

(1) Establishment of winter wheat LAI and SPAD estimation models under different water treatments based on VI alone by using a linear regression model, quadratic polynomial

regression model, and exponential regression analysis and based on VI combinations by multivariable regression (MLR) analysis;

(2) A comparison of the performances of winter wheat LAI and SPAD estimating models and selection of the optimal estimation models of LAI and SPAD;

(3) Estimation of the winter wheat yield by using multivariate regression model based on the optimal LAI and SPAD values obtained in (2) in the flowering stage.

MATERIALS AND METHODS

Study Area

The experiment was carried out on a field in the Daxing District located in the south of Beijing, China (39°37.25′N, 116°25.51′E). The research field with an area of approximately 1.68 km^2 was planted with winter wheat. Thirty 7.8 × 7.5-m^2 fields within each region were chosen as samples for data collection (**Figure 1**). According to the amount of irrigation, the 30 fields were divided into three irrigation levels (low water [0–60 mm], medium water [120–180 mm], and high water [240–300 mm]). The cumulated precipitation rates of DAS210 (days after sowing [DAS], jointing stage), DAS229 (flowering stage), and DAS240 (filling stage) were 134.87, 138.43, and 138.43 mm, respectively (**Figure 2**). The soil type was loam sandy (85% sand, 11.5% power, and 3.5% clay), according to the United States Department of Agriculture taxonomy. Winter wheat was planted on October 5, 2017, and harvested on June 30 with a 265-day life span (**Table 1**).

TABLE 1 | Crop management 2017–2018.

Item	Winter wheat
Sowing date	October 5, 2017
Variety	Zhong Mai 175
Average seeding rate	289 grains/m^2
Harvest date	June 30, 2018
Soil type	Loam sandy

Research Measurements

At the winter wheat key stages (DAS210–DAS240) (Du et al., 2017; Yue et al., 2017; Ji et al., 2020; Jiang et al., 2020; Tao et al., 2020a), crop parameters (LAI and SPAD) were measured on DAS210, DAS229, and DAS240. Sixty sets of samples (LAI [30 sets] and SPAD [30 sets]) were obtained every 7 days.

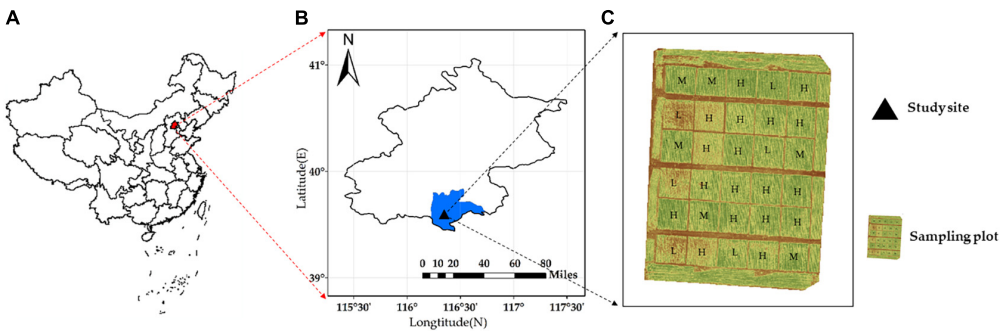

FIGURE 1 | Research area: **(A)** location of the research field in China; **(B)** area view of the research field indicating region division; and **(C)** the location of sampling plots and ground control plots; L, M, and H represent low water, medium water, and high water treatments, respectively.

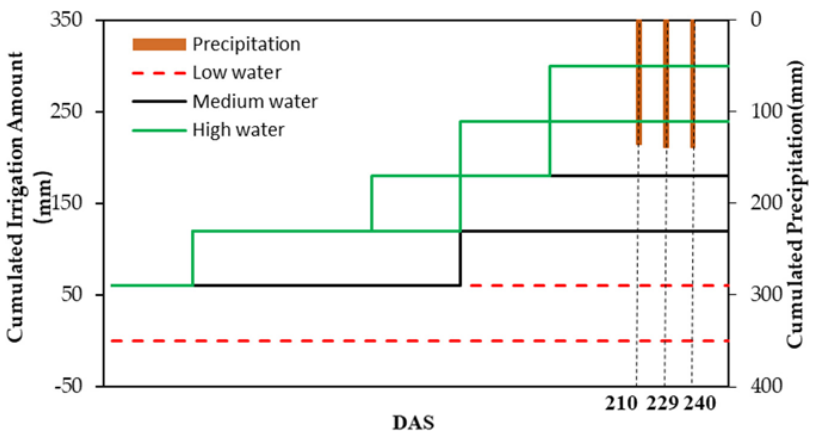

FIGURE 2 | Irrigation and precipitation accumulation under different irrigation treatments (low water, medium water, and high water).

LAI from LAI-2200C: To be synchronous with the imagery, we took three wheat stems from each plot as one sample, separated the green leaves, and used an LI-3000 leaf area meter to scan the green leaf area. We had a total of 18 plots (low water: six plots; medium water: six plots; and high water: six plots), whose length and width are 7.8 and 7.5 m, respectively. The number of stems 1 m in length in each plot was counted manually. Based on formula (1), the LAI of the population was calculated.

$$LAI = \frac{1}{D} \times B \times \frac{A}{C} \times 10^{-4} \quad (1)$$

where D is the distance between two rows of wheat; B is the number of stems in 1 m in length; A is the leaf area of the sample; and C is the number of stems of the sample. A SPAD-502 chlorophyll meter (Minolta Corporation, NJ, United States) was used for *in vivo* measurement of the ratio of light transmittance through the leaf. Instrument readings have been shown to correlate well with laboratory measurements of chlorophyll concentrations in several species (Haie and Keller, 2008). On each sampling campaign, 30 SPAD measurements were collected on average. The chlorophyll meter readings were taken midway on fully expanded top-of-canopy leaves. Each measurement was characterized by the mean of three replicate measurements. The chlorophyll meter measured an area of 2 × 3 mm with an accuracy of ± 1.0 SPAD unit (at room temperature).

Acquisition and Pretreatment of UAV Multispectral Image

In this study, a small six-spin UAV (Nanjing Hepu Aero Science and Technology Co., Ltd.)[1] was used. Multispectral cameras were mounted synchronously on a UAV remote sensing platform (before the camera was used, noise removal and lens distortion correction were carried out). **Table 2** illustrates the UAV specification in more detail.

When the weather was clear and cloudless, three flights (DAS 210, DAS229, and DAS240) were carried out from 12:00 to 13:00, when the solar zenith angle was minimal. Continuous flight monitoring was carried out in 30 plots of the study area. The flight altitude was 60 m, and the spatial resolution was 0.0409 cm.

The image mosaic processing was performed by using the Pix4Dmapper software (Pix4D Inc., Switzerland)[2] (Turner et al., 2012). The preprocessing of mosaic multispectral images included geometric correction and radiation correction, and the geometric correction mainly used the ENVI software. With the Orthophoto image as a reference image, 20 reference points were selected uniformly in different positions of the image to correct the geometric accuracy of the multispectral image. The error of the geometric correction of the image was less than 0.5 pixels after verification. For radiation correction, due to the difference between the time and weather conditions of the multispectral data obtained from different sites, the pseudo-standard ground object radiation correction method was used to convert the multispectral image value into the image reflectance value through the reflectance measured by the ground target (Wang and Liu, 2014).

The five multispectral bands were blue (central wavelength 475 nm, bandwidth 40 nm), green (central wavelength 560 nm, bandwidth 40 nm), red (central wavelength 668 nm, bandwidth 40 nm), red edge (central wavelength 717 nm, bandwidth 10 nm), and near infrared (central wavelength 840 nm, bandwidth 40 nm) (**Figure 3**).

[1]http://www.agrouas.com/

[2]http://www.pix4d.com

TABLE 2 | UAV and multispectral camera specifications.

UAV		Multispectral camera	
Type	**Pixhawk (M600)**	**Type**	**Red Edge-M**
Maximum payload	5 kg	Band range	475–840 nm
Maximum duration	15 min	Terrestrial resolution	0.0409 m

UAV Multispectral VI

Many previous studies have used different VIs in multispectral imagery to estimate the crop parameters (LAI and SPAD). In this study, VIs and one VI combination were calculated by using visible bands, including the NDVI (Rouse et al., 1974), SAVI (Huete, 1988), EVI (Huete et al., 2002), and DVI (Jordan, 1969; **Figure 4**). Their calculation formulas are as follows:

$$NDVI = \frac{R_{nir} - R_{red}}{R_{nir} + R_{red}} \quad (2)$$

$$SAVI = 1.5\frac{R_{nir} - R_{red}}{R_{nir} + R_{red} + 0.5} \quad (3)$$

$$EVI = 2.5\frac{R_{nir} - R_{red}}{R_{nir} + 6R_{red} - 7.5R_{blue} + 1} \quad (4)$$

$$DVI = R_{nir} - R_{red} \quad (5)$$

Note that R_{nir} is the near-infrared reflectance, R_{red} is the red reflectance, and R_{blue} is the blue reflectance.

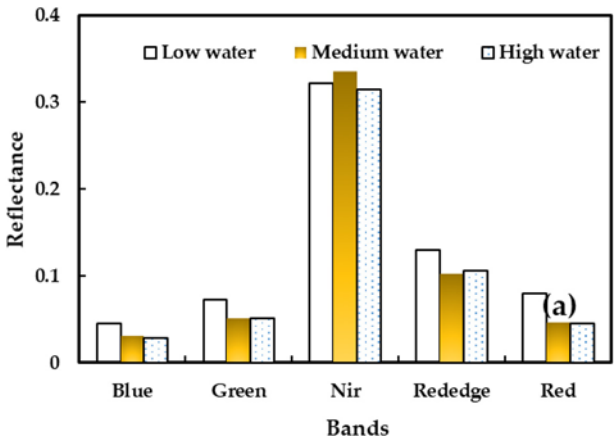

FIGURE 3 | Reflectance of five bands under different water treatments.

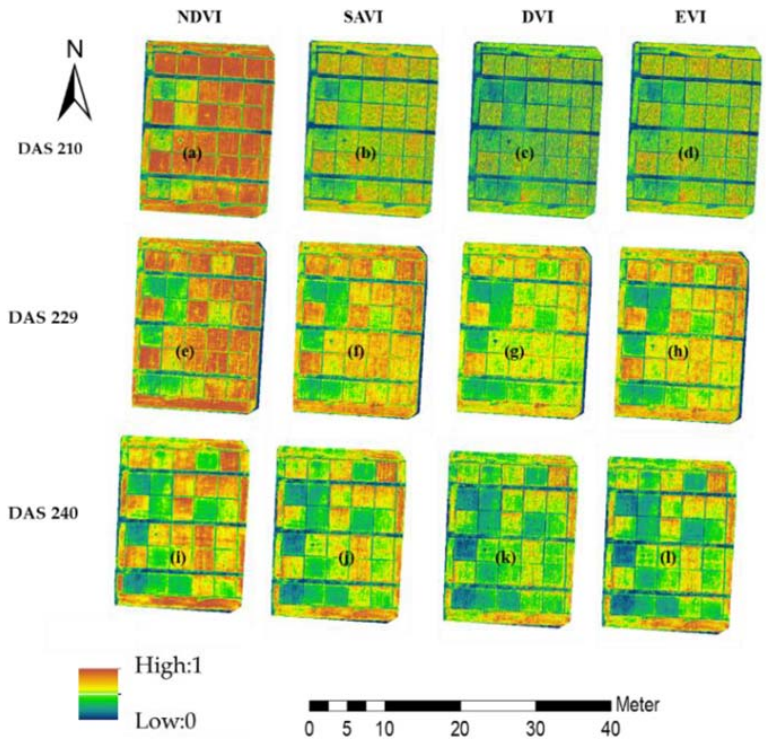

FIGURE 4 | VI of the research area. **(a–d)** were the NDVI, SAVI, SVI and EVI in DAS210 of research area, **(e–h)** were the NDVI, SAVI, SVI and EVI in DAS229 of research area, **(i–l)** were the NDVI,SAVI,SVI and EVI in DAS240 of research area.

Estimation Models of Crop Parameters (LAI and SPAD) and Yield

Figure 5 showed the main procedures of obtaining the optimal estimation model of crop parameters and yield based on winter wheat features derived from UAV multispectral imagery. Four estimation models of winter wheat crop parameters were used in this study, i.e., prediction models: (1) linear regression model, (2) quadratic polynomial regression, (3) exponential model, and (4) multiple linear regression based on VIs. In the establishment of the yield prediction model, multivariable linear regression analysis was adopted.

Statistical Analysis

For statistical analysis, SPSS 23 was adopted. For the spectral reflectance information of winter wheat observed in different plots, the linear regression model, quadratic polynomial regression, exponential model, and multiple linear regression model (Quan et al., 2017) of winter wheat VIs, LAI, and SPAD were established. The validation set was used to fit the predicted and measured values of the model (Marenco et al., 2009; Drusch et al., 2012), and then the multiple linear regression model was used to predict the output.

$$y = ax + b \tag{6}$$

$$y = ax^2 + bx + c \tag{7}$$

$$y = ae^x \tag{8}$$

$$y = y_0 + y_1x_1 + y_2x_2 + \cdots + y_nx_n + m \tag{9}$$

In the formula, $y_0, y_1, y_2, \ldots, y_n$ is the regression coefficient, and m is the model error.

The coefficient of determination (R^2) and RMSE were used to evaluate the performance of each model. Mathematically, a higher R^2 corresponds to a smaller RMSE and thus represents better model accuracy. The following equations were used to calculate R^2 and RMSE (Pandit et al., 2018; Li et al., 2019), respectively:

$$R^2 = \frac{\sum_{i=1}^{n}(y_i - x_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \tag{11}$$

where x_i and y_i are the estimated and measured values, respectively; $\bar{x}$ and $\bar{y}$ are the average estimated and measured values, respectively; and n is the sample number.

RESULTS

Table 3 shows the statistics of LAI, SPAD, and yield measurements for different water treatments. In this study, four VIs (NDVI, SAVI, EVI, and DVI) derived from UAV multispectral imagery were used for the linear regression model, quadratic polynomial regression, exponential model, and multiple linear regression for winter wheat LAI and SPAD under low water, medium water, and high water. In Figures A1–A3 and **Tables 4**, **5**, the four VIs all had significant positive

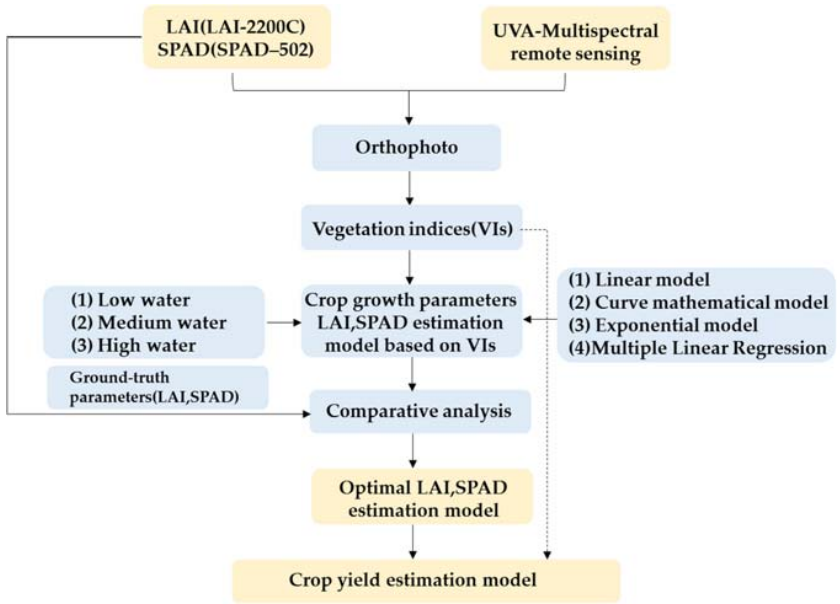

FIGURE 5 | Schematic indicating the main procedures to obtain the optimal estimation model of winter wheat crop parameters and yield.

correlations ($p < 0.01$) with low water, medium water, and high water on LAI and SPAD.

Estimation Models of Winter Wheat LAI on the Basis of UAV Multispectral VIs

As shown in **Table 4**, the NDVI had the highest R^2 values of 0.868 (quadratic polynomial), 0.897 (exponential), and 0.749 (quadratic polynomial), followed closely by the EVI with R^2 of 0.747 (quadratic polynomial), 0.8746 (quadratic polynomial), and 0.741 (quadratic polynomial) and by the and SAVI with R^2 of 0.698 (exponential), 0.852 (quadratic polynomial), and 0.740 (linear). The lowest correlations were observed by using DVI to estimate winter wheat, with R^2 values of 0.620 (exponential), 0.718 (exponential), and 0.607 (exponential). When it comes to the RMSE, a similar observation was obtained. And for low-water LAI, the NDVI also had a lower RMSE of 0.591, followed closely by the EVI and SAVI with RMSE values of 0.607 and 0.615, respectively. The DVI had a larger RMSE of 0.814. For medium and high water treatments, similar observations were found.

TABLE 3 | Descriptive statistics of LAI, SPAD, and yield from the study area.

Treatment	Parameter	Samples	Min	Mean	Max	SD	CV (%)
Low water	LAI	18	2.00	3.68	5.04	0.96	25.98
	SPAD	18	42.08	51.78	63.97	6.37	12.30
	Yield (kg/ha)	6	2,397.0	4,945.4	3,731.2	944.91	25.12
Medium water	LAI	18	2.53	4.22	5.67	0.97	22.93
	SPAD	18	42.36	53.44	63.08	6.97	13.03
	Yield (kg/ha)	6	2,935.3	5,588.6	4,526.4	1,093.4	24.23
High water	LAI	18	2.97	3.97	5.99	0.65	11.35
	SPAD	18	50.66	60.85	72.67	6.91	16.39
	Yield (kg/ha)	6	3,116.3	4,848.6	5,810.7	926.1	19.02

SD, standard deviation; CV, coefficient of variation. Six points were randomly selected from 18 points of high water treatment.

After estimating winter wheat LAI with different water treatments by using a single VI, for LAI with low water treatment (LAI_l), the four VI combinations that had the higher correlations were chosen to estimate winter wheat LAI by adopting MLR analysis. When MLR was used, the estimation performance of winter wheat LAI was improved (**Figure 6A**), with an R^2 value of 0.911, with an increase of 0.0434 compared to the highest R^2 value of 0.8676 (NDVI, quadratic polynomial) for the single VI. The RMSE of LAI decreased to 0.3663, compared to the lowest RMSE value of 0.5914 for single VI, quadratic polynomial. The winter wheat LAI could be calculated based on NDVI, SAVI, EVI, and DVI by using Equation (12). However, for LAI with medium and high water treatments (LAI_m and LAI_h), when MLR was adopted, the estimation performance of winter wheat (LAI_m and LAI_h) was not improved, and the R^2 values were 0.73 and 0.744, with a decrease of 0.167 and 0.005, compared to the highest R^2 values of 0.897 (NDVI, exponential) and 0.748 (NDVI, quadratic polynomial) for the single VI. The RMSE of LAI increased by 0.046 and 0.058, compared to the lowest RMSE values of 0.609 and 0.631 for single VI with single linear regression, quadratic polynomial, and exponential models, respectively. The winter wheat (LAI_m and LAI_h) could be calculated based on NDVI by using Equations (13) and (14) (**Figures 6B,C**).

$$LAI_l = 14.643 \times NDVI - 15.293 \times SAVI + 33.510 \times EVI - 16.431 \times DVI - 0.698 \quad (12)$$

$$LAI_m = 1.3626 \times e^{1.384NDVI} \quad (13)$$

$$LAI_h = 13.567 \times NDVI^2 - 14.567 \times \mathrm{NDVI} + 6.932 \quad (14)$$

TABLE 4 | Correlations between VIs derived from UAV multispectral imagery and LAI with different water treatments.

Model	VIs	Low water		Medium water		High water	
		R^2	RMSE	R^2	RMSE	R^2	RMSE
Linear	NDVI	0.8504**	0.606	0.8943**	0.6222	0.6673**	0.5992
	SAVI	0.6972**	0.5979	0.818**	0.6095	0.7395**	0.6317
	EVI	0.7455**	0.624	0.7643**	0.5924	0.7467**	0.6345
	DVI	0.6609**	0.5915	0.7586**	0.5995	0.7162**	0.6338
Quadratic polynomial	NDVI	0.8676**	0.5914	0.8951**	0.6223	0.7493**	0.6312
	SAVI	0.6973**	0.5979	0.8523**	0.6189	0.7409**	0.6317
	EVI	0.7466**	0.6065	0.8746**	0.6152	0.7261**	0.6298
	DVI	0.6637**	0.6272	0.8208**	0.6116	0.7461**	0.6235
Exponential	NDVI	0.7961**	0.6583	0.8966**	0.6094	0.6504**	0.6081
	SAVI	0.6983**	0.6299	0.8223**	0.6357	0.7249**	0.6551
	EVI	0.6467**	0.6178	0.7241**	0.6112	0.73**	0.6592
	DVI	0.6197**	0.8136	0.7175**	0.6258	0.7038**	0.6610
Multiple	Four VIs	0.911**	0.2251	0.73**	0.6558	0.744**	0.637

*Four VIs, NDVI, SAVI, EVI, and DVI, were used to establish regression estimation models for LAI with different water treatments. **Significant at the level of 0.01.*

Prediction Models of Winter Wheat SPAD Based on UAV Multispectral VIs

In **Table 5**, the NDVI had the highest correlations with R^2 values of 0.808 (quadratic polynomial), 0.874 (exponential), and 0.829 (quadratic polynomial) followed by the SAVI with R^2 values of 0.769 (quadratic polynomial), 0.831 (quadratic polynomial), and 0.725 (quadratic polynomial). For both low and high water treatments, the lowest correlations were observed by using DVI to predict SPAD, with R^2 values of 0.663 (exponential) and 0.623 (exponential), respectively. However, with the medium water treatment, the lowest correlations were observed by using EVI to estimate SPAD, with R^2 values of 0.701 (linear). For low water level, the NDVI also had the lowest RMSE with a value of 7.00, followed by the SAVI with an RMSE value of 7.17; the maximum RMSE of DVI was 7.99. For the medium and high water treatments, similar observations were found. The R^2 was negatively correlated with RMSE.

After estimating winter wheat SPAD with different water treatments by using single VI, for SPAD with low water treatment ($SPAD_l$), the four VI combinations, which had the highest correlations, were chosen to estimate winter wheat SPAD by using MLR analysis. When MLR was used, the estimation performance of winter wheat $SPAD_l$ was improved (**Figure 7A**); the R^2 value was 0.812, with an increase of 0.0038, compared to the highest R^2 value of 0.808 (NDVI, quadratic polynomial) for the single VI. The RMSE of SPAD decreased with a value of 0.004, compared to the lowest RMSE value of 7.022 for single VI, quadratic polynomial. The winter wheat $SPAD_l$ could be calculated based on NDVI, SAVI, EVI, and DVI by using Equation (15). However, with medium and high water treatments ($SPAD_m$ and $SPAD_h$), when MLR was used, the estimation of winter wheat ($SPAD_m$ and $SPAD_h$) was not improved, and the R^2 values were 0.87 and 0.822, with a decrease of 0.0042 and 0.008, respectively, compared to the highest R^2 values of 0.8742 (NDVI, exponential) and 0.830 (NDVI, quadratic polynomial) for the single VI. The RMSE of SPAD increased with values of 0.021 and 0.286, compared to the RMSE of 7.038 and 6.126 for single VI with the single linear regression, quadratic polynomial, exponential models. The winter wheat ($SPAD_m$ and $SPAD_h$) could be calculated based on NDVI by using Equations (16) and (17) (**Figures 7B,C**).

$$SPAD_l = 249.19 \times NDVI - 686.477 \times SAVI + 24.896 \times EVI + 596.61 \times DVI + 24.647 \quad (15)$$

$$SPAD_m = 26.499 \times e^{1.0796 \times NDVI} \quad (16)$$

$$SPAD_h = 108.21 \times NDVI^2 - 103.02 \times NDVI + 74.93 \quad (17)$$

Prediction Models of Winter Wheat Yield on the Basis of Both LAI and SPAD

Remote sensing estimation of winter wheat yield is based on VIs which can reflect crop yield. It is necessary to verify whether the relationship between wheat LAI, SPAD, and yield (measured) is significant. **Figure 8** shows the relationship between LAI, SPAD, and winter wheat yield under different stages. It can be seen that the relationships between LAI, SPAD, and winter wheat yield were the most significant in different models. LAI and SPAD could estimate winter wheat yield well under different stages. For LAI (**Figure 8A**) and SPAD (**Figure 8B**), R^2 values are 0.68 (DAS210), 0.885 (DAS229), and 0.612 (DAS240) and 0.534 (DAS210), 0.949 (DAS229),and 0.566 (DAS240), respectively. This provided a basis for the prediction of winter wheat yield using LAI and SPAD which are estimated by VIs.

Figures 8A,B show that the predicted values of crop growth parameters at three stages (DAS210, DAS229, and DAS240) were consistent with the measured values and that the R^2 values (0.885 for LAI and 0.949 for SPAD) were the highest in DAS229, which was the best estimation period of yield; this result was consistent with that of previous studies (Zhang et al., 2019; Tao et al., 2020b). Under the support of the above, the optimal estimates of winter wheat LAI and SPAD values in DAS229 based on VIs were adopted to estimate the yield of winter wheat by using the quadratic polynomial model (Equations (18) and

TABLE 5 | Correlations between VIs derived from UAV multispectral images and SPAD with different water treatments.

Model	VIs	Low water		Medium water		High water	
		R^2	RMSE	R^2	RMSE	R^2	RMSE
Linear	NDVI	0.7994**	7.205	0.8693**	7.632	0.748**	6.3045
	SAVI	0.7423**	7.089	0.8249**	7.541	0.7061**	6.5237
	EVI	0.7222**	7.048	0.7011**	7.281	0.7112**	6.5348
	DVI	0.6831**	7.968	0.7842**	7.456	0.6549**	6.4787
Quadratic polynomial	NDVI	0.8082**	7.022	0.8697**	7.633	0.8296**	6.1259
	SAVI	0.7689**	7.143	0.8306**	7.553	0.7247**	6.6320
	EVI	0.7485**	7.102	0.803**	7.514	0.7217**	6.6398
	DVI	0.7375**	7.079	0.7491**	7.591	0.6637**	6.5622
Exponential	NDVI	0.7788**	7.341	0.8742**	7.038	0.7369**	6.4004
	SAVI	0.7212**	7.172	0.8243**	8.951	0.6789**	6.6773
	EVI	0.6989**	7.121	0.7034**	7.4651	0.6813**	6.6899
	DVI	0.6627**	7.989	0.7812**	7.599	0.6225**	6.6248
Multiple	Four VIs	0.812**	7.018	0.87**	7.059	0.8221**	6.412

*Four VIs and one VI combination, NDVI, SAVI, EVI, and DVI, were used to establish regression estimation models for SPAD with different water treatments. **Significant at the level of 0.01.*

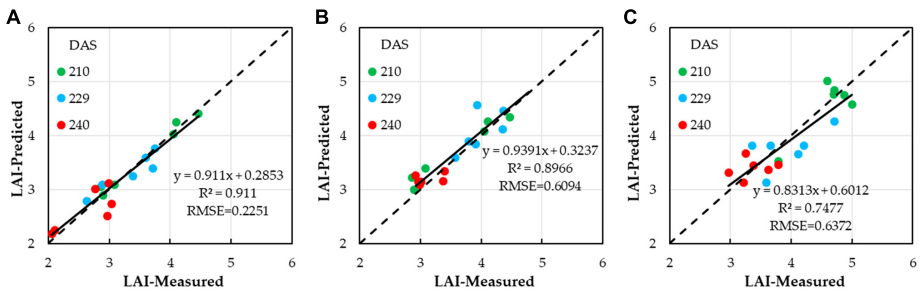

FIGURE 6 | Correlations between winter wheat LAI estimated based on NDVI, SAVI, EVI, and DVI by using optimal regression analysis and ground-truth LAI; **(A)** lower water treatment, **(B)** medium water treatment; **(C)** and higher water treatment.

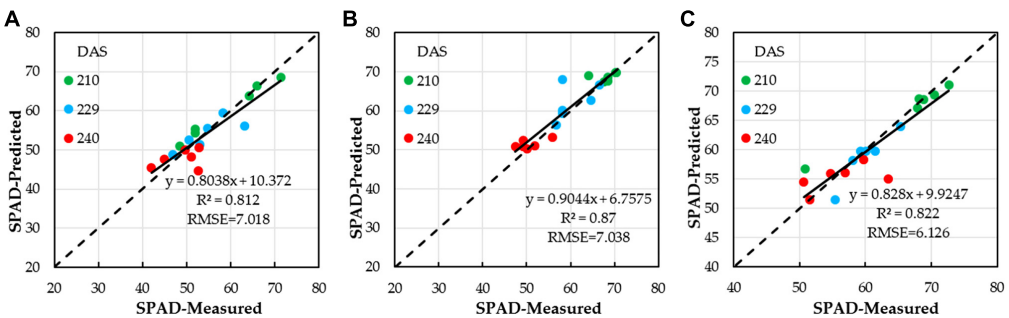

FIGURE 7 | Correlations between winter wheat SPAD estimated based on NDVI, SAVI, EVI, and DVI by using optimal regression analysis and ground-truth SPAD; **(A)** low water treatment, **(B)** medium water treatment; and **(C)** high water treatment.

(19) and **Figure 10**). The accuracy of LAI and SPAD based on UAV multispectral imagery to estimate winter wheat yield was over 87%.

$$\text{Yield} = 1095 \times LAI^2 - 7666.6LAI + 16981 \quad (18)$$

$$\text{Yield} = -2.214 \times SPAD^2 - 366.9SPAD - 9689.2 \quad (19)$$

$$\text{Yield} = 327.44 \times \text{LAI} + 72.15SPAD - 1414.8 \quad (20)$$

After estimating winter wheat yield based on a single parameter (LAI or SPAD) by using quadratic polynomial, when MLR was used in Equation (20), the estimation result of the winter wheat yield was improved (**Figure 10C**), and the R^2 value was 0.807, with the increase of 0.099 and 0.137, respectively, compared to the R^2 values of 0.708 (LAI) and 0.670 (SPAD) for the single parameter. The RMSE of yield decreased by 7.08 and 16.13 kg/ha, compared to the RMSE values of 788.67 kg/ha (LAI, quadratic

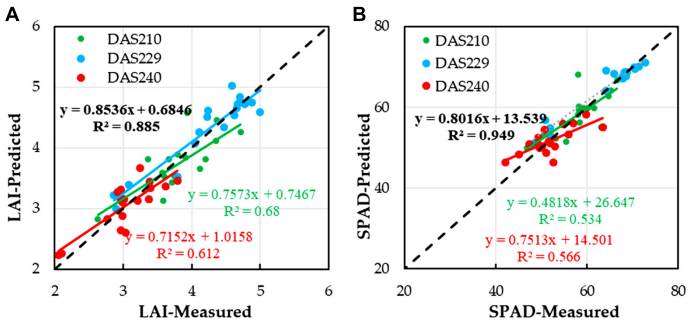

FIGURE 8 | Relationships between LAI, SPAD, and yield of winter wheat under different stages; **(A)** LAI and **(B)** SPAD.

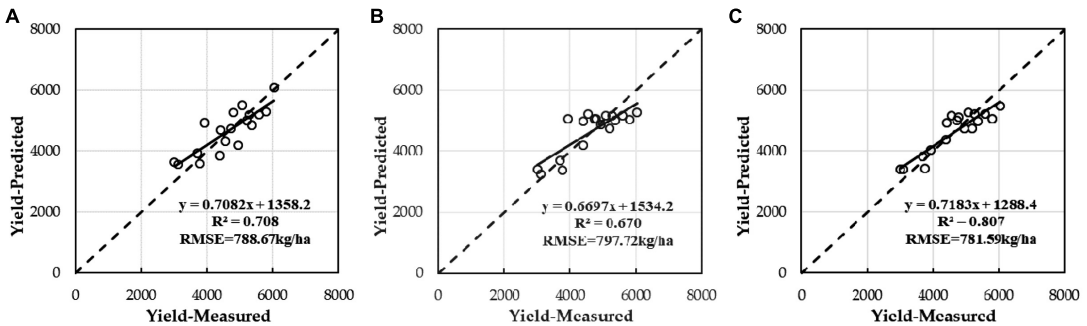

FIGURE 9 | Correlations between yield of winter wheat estimated based on LAI and/or SPAD alone by using optimal regression analysis and multivariable liner regression with ground-truth yield; **(A)** LAI, **(B)** SPAD, and **(C)** LAI and SPAD.

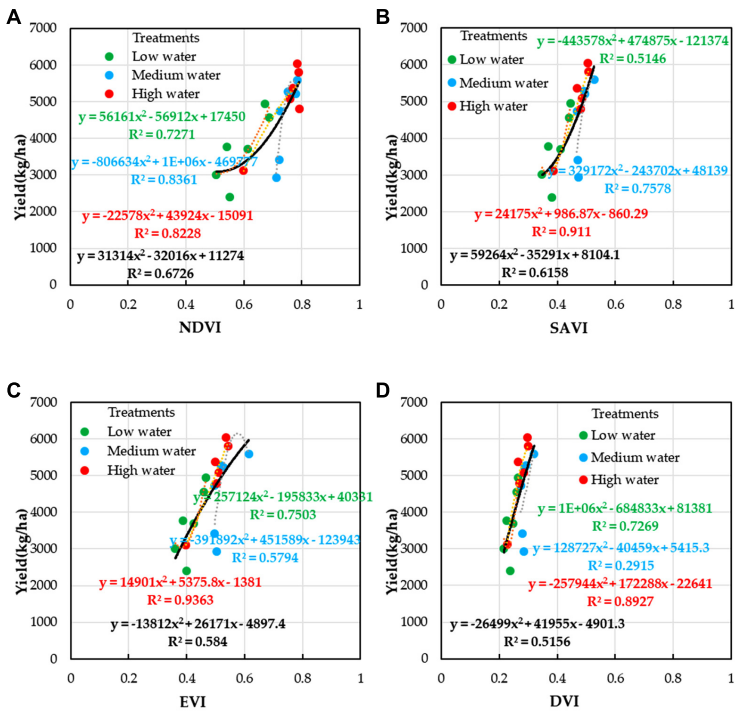

FIGURE 10 | Relationships between VIs and yield of winter wheat under different water treatments; **(A)** NDVI, **(B)** SAVI, **(C)** EVI, and **(D)** DVI.

polynomial) and 797.72 kg/ha (SPAD, quadratic polynomial) for the single parameter, respectively.

DISCUSSION

Estimation of Winter Wheat Parameters at Different Water Treatments

The estimation of winter wheat LAI and SPAD with different water treatments (low water, medium water, and high water) using linear, quadratic polynomial, exponential models in different days is shown in **Appendix Figures A1–A3**. It is shown that the best agreement of predicted winter wheat LAI and SPAD values was for the medium water level (120–180 mm), followed by the low water level (0–60 mm) (LAI); the worst was the high water level (210–240 mm). In **Table 4**, the R^2 values between the LAI obtained from eight estimation models and the measured LAI were more than 0.7 in the medium water level (0.7175–0.8966, at low water level: 0.6197–0.8504 and at high water level: 0.650–0.749); there were few estimation models with R^2 greater than 0.8 between the estimated LAI and the measured LAI. This indicates that different water treatments had an effect on the inversion of LAI-based VIs alone under the same models. However, when multivariable liner regression was adopted (**Table 4** and **Figure 6**), the trend of LAI retrieval from VI-based UAV multispectral imagery had changed (Quan et al., 2017; Berger et al., 2018; Guo et al., 2020). The best inversion result was for the low water level, followed by the medium water level and high water level.

Again, it can be seen from **Table 5** that all the R^2 values between the SPAD obtained estimation models and the measured SPAD were more than 0.7 in medium water level (0.7011–0.8742); in both low water and high water levels, 75% of them have an R^2 greater than 0.7, but in general, the estimation in the low water level was better. However, when the multivariable linear regression was adopted (**Table 5** and **Figure 7**), the results here were not similar to the LAI; the best inversion result was for the medium water level, followed by the high water level and low water level. It is shown that it was necessary to consider multiple VIs to retrieve LAI and SPAD using multivariable linear regression compared to the VI alone (Sun et al., 2012; Boegh et al., 2013; Elarab et al., 2015; Li et al., 2015).

Prediction of Winter Wheat Yield on the Basis of UAV

The crop parameters of winter wheat under different stages could reflect the change of yield (Lin et al., 2009; Sid"ko et al., 2017) (**Figures 8A,B**); the nonlinear values were both very significant (R^2 = 0.885 for LAI and R^2 = 0.949 for SPAD), indicating that the yield could be estimated by measuring LAI and SPAD of winter wheat in the flowering stage (DAS229). The LAI and SPAD of winter wheat were estimated by VI constructed by UAV multispectral imagery (NDVI, EVI, SAVI, and DVI) (Figures A1–A3). **Figure 10** shows the estimation of winter wheat yield with NDVI, SAVI, EVI, and DVI constructed by UAV multispectral imagery under different water treatments, and NDVI had the highest correlation with an R^2 value of 0.673 (medium water > high water > low water), followed by SAVI with an R^2 value of 0.616 (high water > medium water > low water).

According to the principle that LAI and SPAD were closely related to the yield (Simonetta et al., 2009; Fang et al., 2010; Bendig et al., 2013), it was feasible to carry out large-area remote sensing yield estimation based on various VIs constructed by UAV multispectral imagery (**Figures 8–10**). If drought caused wilting of winter wheat, withering of lower leaves, or excessive irrigation, crop development was abnormal, which could be reflected by the dynamic change of crop parameters retrieved by UAV multispectral remote sensing.

FUTURE WORK

There is an increasing need for further raising awareness on the issue of improving estimation performance by using drone multispectral images and constructing crop vision systems to estimate crop parameters and yields. In order to reduce the error in the process of image acquisition and processing, there are still some problems to be solved. First of all, since the pixel value of the image is the reflectivity of the incoming sunlight, the variation of light conditions may lead to the variation of the image-derived features. To further explore the relationship between light changes and camera response (reflection) during UAV flight missions, better integration of light dynamics (UAV attitude, sun position, light scattering, clouds, etc.) is needed to describe light changes. The number and date of data points selected are important issues. In this study, different water treatments (lower water, medium water, and higher water) were selected for the dates from jointing stage to filling stage (DAS210, DAS229, and DAS240). In addition, further research is needed to validate these results for different crops and different sites.

CONCLUSION

The UAV multispectral remote sensing system, as an important farmland-scale data acquisition tool, has great application potential in rapidly, accurately, and economically estimating farmland crop parameters and yields. The results confirmed that the visible light directly derived from UAV multispectral imagery had a high correlation with the measured LAI and SPAD. Compared to the linear regression model, the quadratic polynomial model, and the exponential model based on VIs alone, the MLR based on NDVI, SAVI, DVI, and EVI had higher correlations for both LAI and SPAD under low water treatment with R^2 values of 0.911 and 0.812, respectively. The quadratic polynomial model based on NDVI alone had higher correlations for both LAI and SPAD under medium water treatment, with R^2 values of 0.8996 and 0.87, respectively. However, under high water treatment, the exponential model performance was better than that of the linear model and quadratic polynomial model, with R^2 values of 0.829 and 0.749, respectively. Under different

water treatments, the optimal regression model was different, and with medium water treatment, the estimation was better for both LAI and SPAD. The relationships between the measured crop parameters and the measured yield were verified, and good results were obtained (R^2 = 0.689 for LAI and R^2 = 0.717 for SPAD).

The LAI and SPAD derived from VIs had better potential to estimate winter wheat yield; in the flowering stage, the R^2 values of winter wheat yield estimation based on LAI (quadratic polynomial) and SPAD (quadratic polynomial) were 0.708 and 0.670, respectively. When MLR was used to estimate the yield based on LAI and SPAD, the result of winter wheat yield estimation was improved (R^2 = 0.807, RMSE = 781.59 kg/ha). The ability of VI to identify different aspects of plants is different, which results in improving the prediction performance. Adding LAI and SPAD of UAV multispectral images into the production prediction model based on VIs can significantly improve the performance of production estimation.

In conclusion, this study shows the potential of the UAV multispectral imagery and regression model to estimate the growth parameters and yield of winter wheat. The results provide reference and technical support for the popularization and application of UAV remote sensing in large-scale precision agriculture.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

XH, ZW, and BZ: conceptualization. XH: data curation. YL and ZW: funding acquisition. XH, HC, and ZW: methodology. YL: project administration. TD: supervision. XH: writing—original draft. HC and BZ: writing—review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Key R&D Program of China (nos. 2019YFC0409203 and 2017YFC0403202), the Chinese National Natural Science Fund (no. 51822907), the Special Fund of State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research (no. SKL2018CG03), and the Fund of China Institute of Water Resources and Hydropower Research (nos. 01881910, 0145B742017, 0145B492017, and 0145B102019).

ACKNOWLEDGMENTS

We thank Liwen Bianji (www.liwenbianji.cn/ac) for editing the English text of a draft of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.609876/full#supplementary-material

REFERENCES

Ansar, A., and Muhammad, I. (2020). Evaluating the potential of red edge position (REP) of hyperspectral remote sensing data for real time estimation of LAI & chlorophyll content of kinnow mandarin (Citrus reticulata) fruit orchards. *Sci. Horticu.* 267:109326. doi: 10.1016/j.scienta.2020.109326

Bendig, J., Bolten, A., and Bareth, G. (2013). UAV-based Imaging for Multi-Temporal, very high Resolution Crop Surface Models to monitor Crop Growth Variability Monitoring des Pflanzenwachstums mit Hilfe multitemporaler und hoch auflsender Oberflchenmodelle von Getreidebestnden auf Basis von Bildern aus. *Photogrammetrie-Fernerkundung-Geoinformation* 6, 551–562. doi: 10.1127/1432-8364/2013/0200

Berger, K., Atzberger, C., Danner, M., Urso, G., Mauser, W., Vuolo, F., et al. (2018). Evaluation of the PROSAIL model capabilities for future hyperspectral model environments: a review study. *Remote Sens.* 10:85. doi: 10.3390/rs10010085

Boegh, E., Houborg, R., Bienkowsk, J., Braban, C. F., Dalgaard, T., Van, D. N., et al. (2013). Remote sensing of LAI, chlorophyll and leaf nitrogen pools of crop and grasslands in five European landscapes. *Biogeosciences* 10, 6279–6307. doi: 10.5194/bg-10-6279-2013

Cao, J., Zhang, Z., Tao, F., Zhang, L., and Li, Z. (2020). Identifying the Contributions of Multi-Source Data for Winter Wheat Yield Prediction in China. *Remote Sens.* 12:750. doi: 10.3390/rs12050750

Cao, Y., Jiang, K., Wu, J., and Xu, T. (2020). Inversion modeling of japonica rice canopy chlorophyll content with UAV hyperspectral remote sensing. *PLos One* 15:e0238530. doi: 10.1371/journal.pone.0238530

Castillo, J., Apan, A. A., Maraseni, T. N., and Salmo, S. G. I. (2017). Estimation and mapping of above-ground biomass of mangrove forests and their replacement land uses in the Philippines using Sentinel imagery. *ISPRS J. Photogramm. Remote Sens.* 134, 70–85. doi: 10.1016/j.isprsjprs.2017.10.016

Cougo, M. F., Souza-Filho, P. W., Silva, A. Q., Fernandes M., Santos, J. R., Abreu, M. R. S., et al. (2015). Radarsat-2 Backscattering for the Modeling of Biophysical Parameters of Regenerating Mangrove Forests. *Remote Sens.* 7, 17097–17112. doi: 10.3390/rs71215873

Dehkordi, H., Denis, A., Fouche, J., Burgeon, V., Cornelis, J. T., Tychon, B., et al. (2020). Remotely-sensed assessment of the impact of century-old biochar on chicory crop growth using high-resolution UAV-based imagery. *Int. J. Appl. Earth Obs. Geoinf.* 91:102147. doi: 10.1016/j.jag.2020.102147

Drusch, M., Bello, U. D., Carlier, S., Colin, O., Fernandez, V., Gascon, F., et al. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* 120, 25–36. doi: 10.1016/j.rse.2011.11.026

Du, M., Noboru, N., Atsushi, I., and Yukinori, S. (2017). Multi-temporal monitoring of wheat growth by using images from satellite and unmanned aerial vehicle. *Int. J. Agric. Biol. Eng.* 10, 1–13. doi: 10.25165/j.ijabe.20171005.3180

Elarab, M., Ticlavilca, A. M., Torres-Rua, A. F., Maslova, I., and Mckee, M. (2015). Estimating chlorophyll with thermal and broadband multispectral high-resolution imagery from an unmanned aerial system using relevance vector machines for precision agriculture. *Int. J. Appl. Earth Obs. Geoinf.* 43, 32–42. doi: 10.1016/j.jag.2015.03.017

Fang, Y., Xu, B., Turner, N., and Li, F. (2010). Grain yield, dry matter accumulation and remobilization, and root respiration in winter wheat as affected by seeding rate and root pruning. *Eur. J. Agron.* 33, 257–266. doi: 10.1016/j.eja.2010.07.001

Fu, Y., Yang, G., Li, Z., Song, X., and Zhao, C. (2020). Winter Wheat Nitrogen Status Estimation Using UAV-Based RGB Imagery and Gaussian Processes Regression. *Remote Sens.* 12:3778. doi: 10.3390/rs12223778

Gao, L., Yang, G., Li, H., Li, Z., and He, P. (2016). Winter wheat LAI estimation using unmanned aerial vehicle RGB-imaging. *Chin. J. Ecol. Agric.* 24, 1254–1264.

Guo, X., Wang, L., and Tian, J. (2018). Vegetation Horizontal Occlusion Index (VHOI) from TLS and UAV Image to Better Measure Mangrove LAI. *Remote Sens.* 10:1739. doi: 10.3390/rs10111739

Guo, Y., Wang, H., Wu, Z., Wang, S., and Fu, Y. (2020). Modified Red Blue Vegetation Index for Chlorophyll Estimation and Yield Prediction of Maize from Visible Images Captured by UAV. *Sensors* 20:5055. doi: 10.3390/s20185055

Haie, N., and Keller, A. (2008). Effective Efficiency as a Tool for Sustainable Water Resources Management. *JAWRA J. Am. Water Resour. Assoc.* 44, 961–968. doi: 10.1111/j.1752-1688.2008.00194.x

Heinemann, S., Siegmann, B., Thonfeld, F., Martín, J. M., and Rascher, U. (2020). Land Surface Temperature Retrieval for Agricultural Areas Using a Novel UAV Platform Equipped with a Thermal Infrared and Multispectral Sensor. *Remote Sens.* 12:1075. doi: 10.3390/rs12071075

Hua, S., Ju, H., Zhang, H., Hui, L., and Ling, C. (2012). Partial Least Squares Regression Application in LAI Inversion Using Hyperion Data. *Chin. Agric. Sci. Bull.* 28, 44–52.

Huang, J., Sedano, F., Huang, Y., Ma, H., Li, X., and Liang, S. (2016). Assimilating a synthetic Kalman filter leaf area index series into the WOFOST model to improve regional winter wheat yield estimation. *Agric. For. Meteorol.* 216, 188–202. doi: 10.1016/j.agrformet.2015.10.013

Huete, A. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25, 295–309. doi: 10.1016/0034-4257(88)90106-x

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213. doi: 10.1016/s0034-4257(02)00096-2

Ji, J., Liu, J., Niu, Y., Xuan, K., Jiang, Y., Deng, H., et al. (2020). Comparison of estimation methods for growth parameters of winter wheat based on full-band hyperspectral data. *Crops* 6, 180–188. doi: 10.16035/j.issn.1001-7283.2020.06.027

Jiang, J., Zhang, Z., Cao, Q., Liang, Y., Krienke, B., Tian, Y., et al. (2020). Use of an Active Canopy Sensor Mounted on an Unmanned Aerial Vehicle to Monitor the Growth and Nitrogen Status of Winter Wheat. *Remote Sens.* 12:3684. doi: 10.3390/rs12223684

Jin, X., Liu, S., Baret, F., Hemerlé, M., and Comar, A. (2017). Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. *Remote Sens. Environ.* 198, 105–114. doi: 10.1016/j.rse.2017.06.007

Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology* 50, 663–666. doi: 10.2307/1936256

Katja, B., Clement, A., Martin, D., Guido, D., Wolfram, M., Francesco, V., et al. (2018). Evaluation of the PROSAIL model capabilities for future hyperspectral model environments: a review study. *Remote Sens.* 10:85. doi: 10.3390/rs10010085

Li, M., Wu, J., Song, C., He, Y., Niu, B., Fu, G., et al. (2019). Temporal Variability of Precipitation and Biomass of Alpine Grasslands on the Northern Tibetan Plateau. *Remote Sens.* 11:360. doi: 10.3390/rs11030360

Li, W., Marie, W., Francois, W., Pierre, D., Valerie, D., David, M., et al. (2015). A generic algorithm to estimate LAI, FAPAR and FCOVER variables from SPOT4_HRVIR and landsat sensors: evaluation of the consistency and comparison with ground measurements. *Remote Sens.* 7, 15494–15516. doi: 10.3390/rs71115494

Li, X., Zhang, Y., Luo, J., and Yang, W. (2016). Quanti?cation winter wheat LAI with HJ-1CCD image features over multiple growing seasons. *J. Appl. Earth Obs. Geoinf.* 44, 104–112. doi: 10.1016/j.jag.2015.08.004

Lin, H., Chen, J., Pei, Z., Zhang, S., and Hu, X. (2009). Monitoring sugarcane growth using ENVISAT ASAR data. *IEEE Trans. Geosci. Remote Sens.* 47, 2572–2580. doi: 10.1109/tgrs.2009.2015769

Liu, B., Asseng, S., Wang, A., Wang, S., Tang, L., Cao, W., et al. (2017). Modelling the effects of post-heading heat stress on biomass growth of winter wheat. *Agric. For. Meteorol.* 247, 476–490. doi: 10.1016/j.agrformet.2017.08.018

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. *ISPRSJ Photogramm. Remote Sens.* 134, 43–58. doi: 10.1016/j.isprsjprs.2017.10.011

Marenco, R., Antezanavera, S., and Nascimento, H. (2009). Relationship between specific leaf area, leaf thickness, leaf water content and SPAD-502 readings in six Amazonian tree species. *Photosynthetica* 47, 184–190. doi: 10.1007/s11099-009-0031-6

Michele, C., Pedro, S. F., Arnaldo, S., Marcus, F., João, R. S., Maria, A., et al. (2015). Radarsat-2 Backscattering for the Modeling of Biophysical Parameters of Regenerating Mangrove Forests. *Remote Sens.* 7, 17097–17112. doi: 10.3390/rs71215873

Niu, Y., Zhang, L., Zhang, H., Han, W., and Peng, X. (2019). Estimating Above-Ground Biomass of Maize Using Features Derived from UAV-Based RGB Imagery. *Remote Sens.* 11:1261. doi: 10.3390/rs11111261

Pandit, S., Tsuyuki, S., and Dube, T. (2018). Landscape-Scale Aboveground Biomass Estimation in Bu?er Zone Community Forests of Central Nepal: coupling In Situ Measurements with Landsat 8 Satellite Data. *Remote Sens.* 10, 1–18.

Pham, T., and Yoshino, K. (2017). Aboveground biomass estimation of mangrove species using ALOS-2 PALSAR imagery in Hai Phong City, Vietnam. *J. Appl. Remote Sens.* 11:026010. doi: 10.1117/1.jrs.11.026010

Qiao, L., Gao, D., Zhang, J., Li, M., and Ma, J. (2020). Dynamic Influence Elimination and Chlorophyll Content Diagnosis of Maize Using UAV Spectral Imagery. *Remote Sens.* 12:2650. doi: 10.3390/rs12162650

Quan, X., He, B., Yebra, M., Yin, C., Liao, Z., and Li, X. (2017). Retrieval of forest fuel moisture content using a coupled radiative transfer model. *Environ. Model. Softw.* 95, 290–302. doi: 10.1016/j.envsoft.2017.06.006

Rouse, J., Hass, R. H., schell, J. A., Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ.* 351:309.

Sid"ko, A., Botvich, I., Pisman, T., and Shevyrnogov, A. (2017). Estimation of chlorophyll content and yield of wheat crops from reflectance spectra obtained by ground-based remote measurements. *Field Crops Res.* 207, 24–29. doi: 10.1016/j.fcr.2016.10.023

Simonetta, F., Rosella, M., and Francesco, G. (2009). The effect of nitrogenous fertiliser application on leaf traits in durum wheat in relation to grain yield and development. *Field Crops Res.* 110, 69–75. doi: 10.1016/j.fcr.2008.07.004

Skudra, I., and Ruza, A. (2017). "Effect of nitrogen and sulphur fertilization on chlorophyll content in winter wheat," in *Scientific Journal of Latvia University of Agriculture*, (Germany: Walter de Gruyter).

Sun, H., Ju, H., Zhang, H., Hui, L., and Ling, C. (2012). Partial least squares regression application in LAI inversion using hyperion data. *Chin. Agric. Sci. Bull.* 28, 44–52. doi: 10.1007/s11783-011-0280-z

Sun, Y., Lu, L., and Liu, Y. (2021). Inversion of the leaf area index of rice fields using vegetation isoline patterns considering the fraction of vegetation cover. *Int. J. Remote Sens.* 42, 1688–1712. doi: 10.1080/01431161.2020.1841323

Tao, H., Feng, H., Xu, L., Miao, M., and Fan, L. (2020a). Estimation of Crop Growth Parameters Using UAV-Based Hyperspectral Remote Sensing Data. *Sensors* 20:1296. doi: 10.3390/s20051296

Tao, H., Feng, H., Yang, G., Yang, X., Miao, M., Wu, Z., et al. (2020b). Comparison of winter wheat yields estimated with UAV digital image and hyperspectral data. *Trans. Chin. Soc. Agric. Machinery* 51, 146–155.

Tian, M., Ban, S., Chang, Q., You, M., and Wang, S. (2016). Use of hyperspectral images from UAV-based imaging spectroradiometer to estimate cotton leaf area index. *Agric. Eng. Sci.* 32, 102–108.

Turner, D., Lucieer, A., and Watson, C. (2012). An Automated Technique for Generating Georecti?ed Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (SfM) Point Clouds. *Remote Sens.* 4, 1392–1410. doi: 10.3390/rs4051392

Wan, L., Cen, H., Zhu, J., Zhang, J., Zhu, Y., Sun, D., et al. (2020). Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer-a case study of small farmlands in the South of China. *Agric. For. Meteorol.* 291:108096. doi: 10.1016/j.agrformet.2020.108096

Wang, Z., and Liu, L. (2014). Assessment of Coarse-Resolution Land Cover Products Using CASI Hyperspectral Data in an Arid Zone in Northwestern China. *Remote Sens.* 6, 2864–2883. doi: 10.3390/rs6042864

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. *Front. Plant Sci.* 8:1111. doi: 10.3389/fpls.2017.01111

Yao, X., Wang, N., Liu, Y., Tao, C., Tian, Y., Chen, Q., et al. (2017). Estimation of Wheat LAI at Middle to High Levels Using Unmanned Aerial Vehicle Narrowband Multispectral Imagery. *Remote Sens.* 9:1304. doi: 10.3390/rs9121304

Yebra, M., Liao, Z., Yin, C., Xing, B., and Zhang, X. (2017). A radiative transfer model-based method for the estimation of grassland aboveground biomass. *Int. J. Appl. Earth Obs. Geoinf.* 54, 159–168. doi: 10.1016/j.jag.2016.10.002

Yue, J., Yang, G., Li, C., Li, Z., Wang, Y., Feng, H., et al. (2017). Estimation of Winter Wheat Above-Ground Biomass Using Unmanned Aerial Vehicle-Based Snapshot Hyperspectral Sensor and Crop Height Improved Models. *Remote Sens.* 9:708. doi: 10.3390/rs9070708

Zhang, C., and Kovacs, J. (2012). The application of small-unmanned aerial systems for precision agriculture: a review. *Precis. Agric.* 13, 693–712. doi: 10.1007/s11119-012-9274-5

Zhang, S., Zhao, G., Lang, K., Su, B., Chen, X., Xi, X., et al. (2019). Integrated Satellite, Unmanned Aerial Vehicle (UAV) and Ground Inversion of the SPAD of Winter Wheat in the Reviving Stage. *Sensors* 19:1485. doi: 10.3390/s19071485

Zheng, H., Cheng, T., Li, D., Xia, Y., Tian, Y., Cao, W., et al. (2018). Combining Unmanned Aerial Vehicle (UAV)-Based Multispectral Imagery and Ground-Based Hyperspectral Data for Plant Nitrogen Concentration Estimation in Rice. *Front. Plant Sci.* 9:936. doi: 10.3389/fpls.2018.00936

Zhou, X., Zheng, H., Xu, X., He, J., Ge, X., Yao, X., et al. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS J. Photogramm. Remote Sens.* 130, 246–255. doi: 10.1016/j.isprsjprs.2017.05.003

Zhu, W., Sun, Z., Gong, H., and Zhu, K. (2020). Estimating leaf chlorophyll content of crops via optimal unmanned aerial vehicle hyperspectral data at multi-scales. *Comput. Electron. Agric.* 178:105786. doi: 10.1016/j.compag.2020.105786

Conflict of Interest: 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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 31 May 2021
doi: 10.3389/fpls.2021.637694

GIS-Based Analysis for UAV-Supported Field Experiments Reveals Soybean Traits Associated With Rotational Benefit

Yuya Fukano[1†], Wei Guo[1†], Naohiro Aoki[2], Shinjiro Ootsuka[2], Koji Noshita[3,4,5], Kei Uchida[1], Yoichiro Kato[2], Kazuhiro Sasaki[1‡], Shotaka Kamikawa[1] and Hirofumi Kubota[1]*

[1] Graduate School of Agricultural and Life Sciences, Institute for Sustainable Agro-Ecosystem Services, The University of Tokyo, Tokyo, Japan, [2] Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan, [3] Department of Biology, Kyushu University, Fukuoka, Japan, [4] Plant Frontier Research Center, Kyushu University, Fukuoka, Japan, [5] Japan Science and Technology Agency, PRESTO, Kawaguchi, Japan

OPEN ACCESS

Edited by:
Peng Fu,
University of Illinois at Urbana-Champaign, United States

Reviewed by:
Ana María Mendez-Espinoza, Institute of Agricultural Research (Chile), Chile
Michael Gomez Selvaraj, Consultative Group on International Agricultural Research (CGIAR), United States

***Correspondence:**
Yuya Fukano
yuya.fukano@gmail.com

[†] These authors have contributed equally to this work

[‡] **Present address:**
Kazuhiro Sasaki,
Japan International Research Center for Agricultural Sciences, Tsukuba, Japan

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

Received: 04 December 2020
Accepted: 04 May 2021
Published: 31 May 2021

Citation:
Fukano Y, Guo W, Aoki N, Ootsuka S, Noshita K, Uchida K, Kato Y, Sasaki K, Kamikawa S and Kubota H (2021) GIS-Based Analysis for UAV-Supported Field Experiments Reveals Soybean Traits Associated With Rotational Benefit. Front. Plant Sci. 12:637694. doi: 10.3389/fpls.2021.637694

Recent advances in unmanned aerial vehicle (UAV) remote sensing and image analysis provide large amounts of plant canopy data, but there is no method to integrate the large imagery datasets with the much smaller manually collected datasets. A simple geographic information system (GIS)-based analysis for a UAV-supported field study (GAUSS) analytical framework was developed to integrate these datasets. It has three steps: developing a model for predicting sample values from UAV imagery, field gridding and trait value prediction, and statistical testing of predicted values. A field cultivation experiment was conducted to examine the effectiveness of the GAUSS framework, using a soybean–wheat crop rotation as the model system Fourteen soybean cultivars and subsequently a single wheat cultivar were grown in the same field. The crop rotation benefits of the soybeans for wheat yield were examined using GAUSS. Combining manually sampled data (n = 143) and pixel-based UAV imagery indices produced a large amount of high-spatial-resolution predicted wheat yields (n = 8,756). Significant differences were detected among soybean cultivars in their effects on wheat yield, and soybean plant traits were associated with the increases. This is the first reported study that links traits of legume plants with rotational benefits to the subsequent crop. Although some limitations and challenges remain, the GAUSS approach can be applied to many types of field-based plant experimentation, and has potential for extensive use in future studies.

Keywords: crop rotation, drone, experimental design, legume, wheat, yield

INTRODUCTION

Since R. A. Fisher's initial work on statistical principles of experimental design (Fisher, 1926), field experimentation has played a pivotal role in variety of plant sciences, including ecology, evolutionary biology, forestry, and crop science (Box, 1980; Edmondson, 2005). In field experiments, measured values may vary among the experimental plots owing to the treatments,

but there is always some degree of additional variation caused by both systematic errors, e.g., spatial variation in topography and soil fertility, and random errors, e.g., variations from sampling procedure (Sokal and Rohlf, 1995). Although many studies have addressed methods to minimize these errors, the results of field experiments are still subject to large unwanted and uncontrolled variability (Hurlbert, 1984; Legendre et al., 2004; Payne, 2006; Yang, 2010). In addition, a key factor in field experimentation is the time-consuming nature of sampling coupled with limited availability of time, resulting in small numbers of samples, i.e., small sample size (Hurlbert, 1984). Small sample size can cause significant analysis problems by reducing the statistical power and inferential confidence, especially for data with large systematic and random errors (Nakagawa and Cuthill, 2007). Therefore, the development of cost-effective, high-throughput, and general-purpose measurements and their analytical framework is needed to extend the effectiveness of field experimentation studies (Edmondson, 2005).

It is well known that legume plants such as soybeans can acquire nitrogen from the atmosphere through a mutualistic symbiosis with rhizobia, thereby providing a crucial service to the ecosystem. The use of legumes in crop rotations for their nitrogen-fixing ability has a long history (Chorley, 1981; Stinner et al., 1992); it is widely practiced in both industrialized and developing countries (Giller and Cadisch, 1995; Becker and Johnson, 1998; Biederbeck et al., 2005). Many research projects have quantified this nitrogen contribution and its net effect on subsequent crop yields (Herridge and Rose, 2000; van Kessel and Hartley, 2000; Walley et al., 2007; Anglade et al., 2015; Duc et al., 2015). Interestingly, those studies have found the effects of legume cultivation to be quite variable. Although many studies report positive effects of legume cultivation on subsequent crop yield, others have found neutral or even negative effects (reviewed in Walley et al., 2007; Anglade et al., 2015). Such variations can be partially explained by the extreme variability in nitrogen fixation among different legume crops and among the cultivars used for experimentation. Importantly, however, the large systematic and random errors associated with crop rotation experiments can increase the variability within results, and thereby reduce the statistical power of the data analyses. In addition, the requirement for two different crops to be grown sequentially in the same field in crop rotation trials can magnify the potential for errors during experiments.

To predict the effects of legume cultivation on subsequent crops, to maximize their benefit, and to develop innovative genotypes that can enhance rotational benefits, it is necessary to identify which legume traits are associated with the rotational benefits (Herridge and Rose, 2000). Comparison of rotational effects among cultivars of the same legume species may be the best way to achieve this goal, but no studies have verified differences in rotational benefits among cultivars (Duc et al., 2015). One reason for this lack of research may be the difficulty in detecting cultivar differences using conventional field experiments and statistical methods (e.g., ANOVA), because differences among cultivars are relatively small, whereas the variations in rotational experiments are large. To overcome this challenge, a method is needed to examine the differences in rotational benefits among legume cultivars, and to determine which legume traits are associated with rotational benefits.

Recent advances in technical devices and analytical methods have made it possible to do cost-effective remote sensing of field-cultivated plants (Houle et al., 2010; Furbank and Tester, 2011; Tardieu et al., 2017; Tripodi et al., 2018). Proximal sensing through the use of unmanned aerial vehicles (UAVs) is among the most promising and popular techniques, because it is rapid, non-destructive, cost-effective, and information dense (Sankaran et al., 2015; Yang et al., 2017; Guo et al., 2018; Maes and Steppe, 2019). UAV remote sensing enables the acquisition of a large amount of image data accompanied by location information. Recent studies have shown that UAV sensing and image analysis can be used to estimate several traits of field-cultivated plants, e.g., cover area, volume, height, and normalized difference vegetation index (NDVI) (Guo et al., 2017, 2020; Watanabe et al., 2017; Zhou et al., 2017; Hassan et al., 2019). These techniques allow researchers to cost-effectively and non-destructively obtain large amounts of pixel-level plant canopy data with location information. However, despite these significant benefits, there is no general methodology to integrate the large quantities of image data from UAV remote sensing with the manually collected data from conventional field experimentation.

To investigate these issues, we proposed a simple analytical framework, i.e., geographic information system (GIS)-based analysis for UAV supported field study (GAUSS), to integrate remote sensing data into a conventional field experiment (**Figure 1**). A field cultivation experiment was conducted to examine the effectiveness of the GAUSS framework, using a soybean–wheat [*Glycine max* (L.) Merr.; *Triticum aestivum* L.] crop rotation as the model system. The differences among soybean cultivars in their effect on wheat yield were examined, and the following questions were addressed:

(1) Which type or combination of indices from the UAV imagery best predicted wheat yields?
(2) How did the whole distribution of estimated yields in a plot differ from the actual yield data from manual sample collection at selected locations?
(3) Were there differences in wheat yields associated with different soybean cultivars?
(4) What traits of the soybean cultivars were associated with the wheat yields?

MATERIALS AND METHODS

Field Cultivation of Soybean and Wheat

The study was conducted from June 2018 to June 2019 at the Institute of Sustainable Agro-ecosystem Services (ISAS), the University of Tokyo, Japan (35°43′N, 139°32′E). Soybeans were grown during the summer (June to October), followed by winter wheat in the winter (November to June). The soil was derived from a volcanic ash, classified as a Typic Melanudand (USDA Soil Taxonomy). Climatic data during the soybean and wheat growing seasons are summarized in **Supplementary Table 1**.

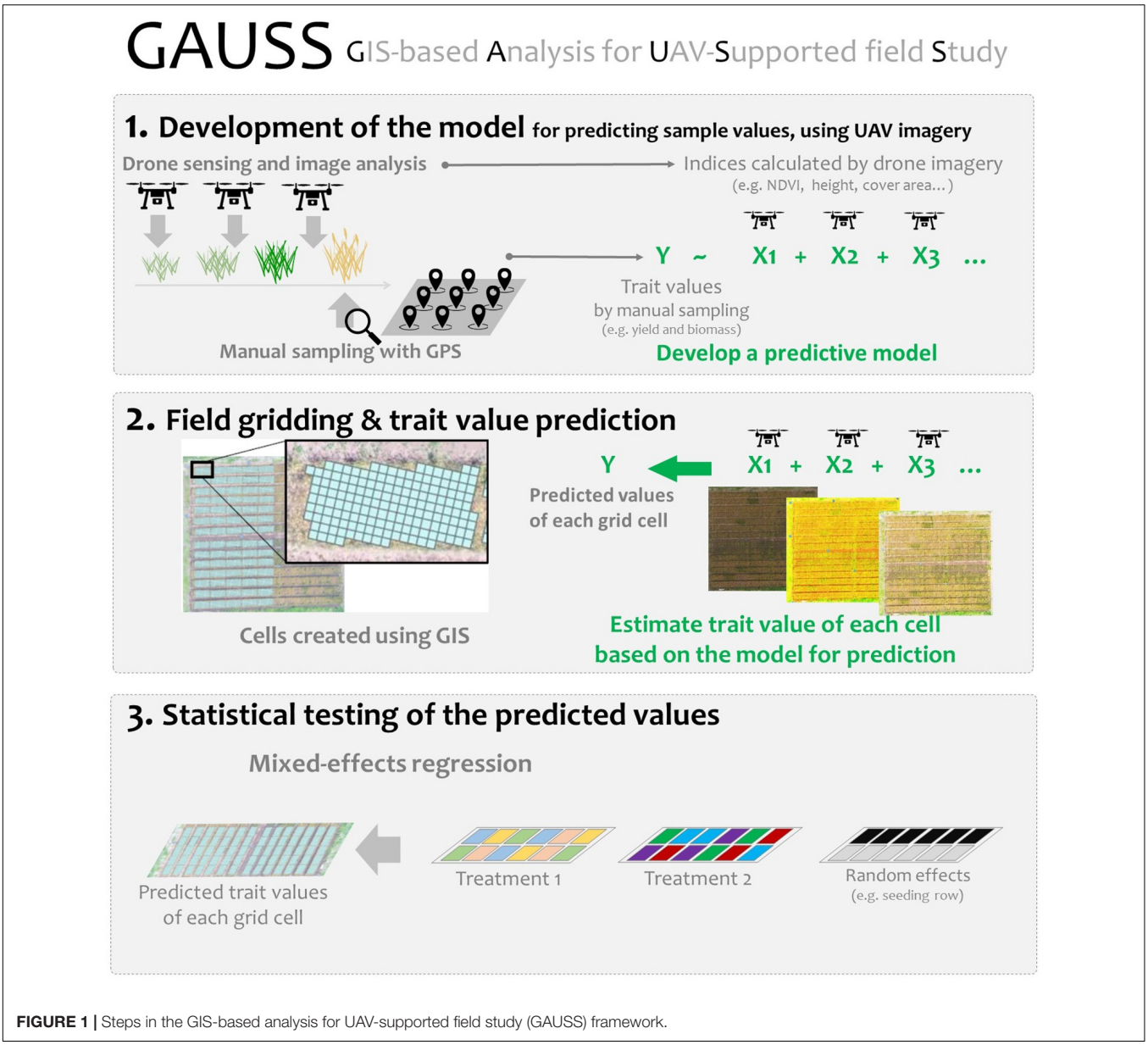

FIGURE 1 | Steps in the GIS-based analysis for UAV-supported field study (GAUSS) framework.

Fourteen soybean cultivars (**Supplementary Table 2**) with different plant types and yield potentials were used (Kaga et al., 2011). Seeds of the GmWMC (Glycine max world mini core-collection) line were obtained from the current Genetic Resource Center, National Agriculture and Food Research Organization, Tsukuba, Japan. There were 70 experimental plots: 56 plots for soybean cultivar testing, 4 plots where the natural weed community was allowed to develop (weedy), 4 plots where the soil was covered with an anti-weed covering (sheet), and 6 plots for destructive sampling of soybean plants (**Figure 2A** and **Supplementary Table 3**). Each plot was 10.08 m^2 (2.4 m × 4.2 m). The cultivars were assigned to the plots in a randomized design, with 2–6 replicates (**Supplementary Table 3**). Three soybean seeds were sown per hill, on 20 June or 20 July 2018, with a row spacing of 60 cm and a hill spacing of 30 cm. Hills were thinned to one seedling at 3 weeks after sowing. A basal fertilizer (N:P:K, 3:10:10) was applied at a rate of 1,000 kg ha^{-1} for soybean cultivation.

The above-ground parts of the soybean plants were manually harvested at physiological maturity during October 2018 and dried completely at 80°C. The whole-plant above-ground dry weight, stem dry weight, seed dry weight, and 100-seed weight were measured. The below-ground parts were left in the soil. The field was tilled twice to a depth of 15 cm using a rotovator at 2 weeks after the soybean harvest. In November 2018, wheat ("*Satonosora*") was uniformly sown over the entire area (40 m × 50 m), including both the location of the 70 soybean plots and the adjacent field area (**Figure 2B**), at 80 kg ha^{-1}. No fertilizer was applied for wheat cultivation. Standard crop protection practices for

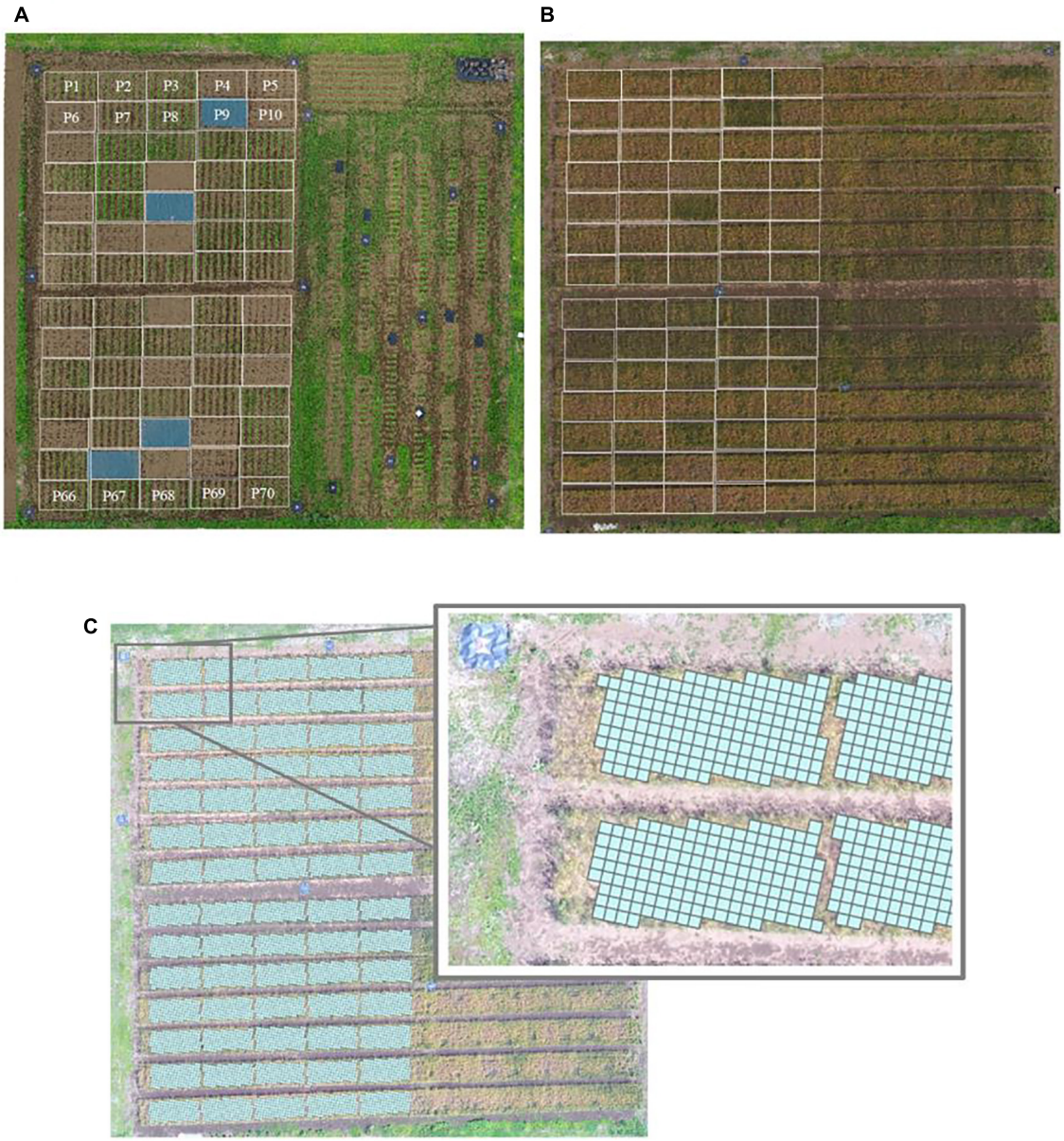

FIGURE 2 | Experimental field P1–70 used in the work reported here. **(A)** Soybean (*Glycine max*) cultivation plots; **(B)** subsequent wheat (*Triticum aestivum*) cultivation, with locations of preceding soybean plots superimposed on the image; **(C)** the field divided into 25-cm × 25-cm cells, using the UAV aerial surveillance data.

soybean (manual weeding, pesticide application, intertillage, and molding) and wheat (herbicide application, manual weeding, and fungicide application) were followed.

During 5 to 10 June 2019, the above-ground parts of the wheat plants were manually harvested from 154 sampling points (1 m × 1 m; **Supplementary Figure 1**) by cutting at the soil surface and placed in a mesh bag. The 154 sampling points comprised one each in the 70 soybean plots and 84 in the adjacent area of the field. A local area RTK-GPS (Real-Time Kinematic Global Positioning System) that was conducted with Hemisphere GNSS devices (Hemisphere GNSS, Scottsdale, AZ, United States) was used to determine the locations of the sampling points. After drying completely at 80°C, the harvested wheat samples were weighed and sorted into immature ears, mature ears, and straw. The dry weights of mature ears and straw were measured separately. The number of mature ears was counted.

GAUSS: An Analytical Framework to Estimate Data Values From UAV Imagery

The GAUSS general framework was used to integrate remote sensing data into a conventional field experiment that

investigated plant growth benefits in a soy–wheat crop rotation. The three main steps of this framework (represented in **Figure 1**) were:

1 Development of the model for predicting sample values, using UAV imagery
2 Field gridding and trait value prediction
3 Statistical testing of the predicted values

Detailed GAUSS Methodologies

Development of the Model for Predicting Sample Values, Using UAV Imagery

Acquisition of Image Data

UAV remote sensing was carried out on 15 February, 14 March, and 12 April 2019, when the target wheat plants were in various stages of growth (tillering, early growth, and lager growth stage). A commercial-grade UAV (DJI Inspire 1, Shenzhen, China) equipped with a multispectral camera was used. The UAV was flown automatically over the field (**Figure 3A**) at an altitude of 30 m under the control of a commercially available flight application (Litchi, VC Technology Ltd., London, England). Two cameras—a Zenmuse X5 (DJI, Shenzhen, China) and a Micasense RedEdge (Micasense, Seattle, WA, United States)—were mounted on the UAV to ensure that RGB and multispectral images were captured during the same flight (**Figure 3B**). Two sets of images of a calibrated reflectance panel placed at about 1 m height were also captured immediately before and after each flight to improve the accuracy of the reflectance data for multi-spectral images. Also, acrylic plates were placed at the four corners of the field and three locations within the field as ground control points (GCPs), and were measured using the Hemisphere RTK differential GNSS device to improve the geolocation accuracy.

The captured images were processed using commercial photogrammetry software (Pix4Dmapper Pro, Pix4D, Lausanne, Switzerland). The pixel-by-pixel values over the entire wheat field were determined by using orthomosaic and digital surface modeling (DSM) that are generated from RGB images to calculate the vegetation cover area and plant height, respectively, and reflectance maps were generated from the multi-spectral images to calculate the NDVI (**Figure 3C**).

Developing the Model for Predicting Wheat Yield

To determine which of the image indices best predicted actual wheat yield, the model selection based on the Akaike's information criteria (AIC) (Akaike, 1974) for results of generalized linear models (GLMs) was used. Accurate geolocation information for both the UAV imagery and the manual sampling points made it possible to map the manually sampled data (actual ground data) into the UAV imagery indices. In the GLM analysis, the dry weight of the harvested wheat ears from each manual sampling location was the response variable. The means of vegetation cover area, plant height, and NDVI at each of the sampling locations, estimated from the UAV images recorded in February, March, and April, were the explanatory variables. Plant height from February was excluded from the analysis owing to low plant height (<10 cm) and consequent low estimation accuracy. The error distribution was Gaussian with an identity link function. The statistical model with the lowest AIC score was selected as the best model (f) for estimating the values of the manually sampled wheat yield data ($y_{\mathbf{p}}$) at a sampling point $\mathbf{p}$ from the UAV images, where:

$$\hat{y}_{\mathbf{p}} = f\left(x_{1,\mathbf{p}}, x_{2,\mathbf{p}}, \cdots, x_{m,\mathbf{p}}\right) + \varepsilon \quad (1)$$

and $\hat{y}_{\mathbf{p}}$ is the estimate of $y_{\mathbf{p}}$, $x_{i,\mathbf{p}}$ represents i-th index (vegetation cover area, plant height, or NDVI in this study) at a sampling point $\mathbf{p}$ derived from the UAV images, m is the total number of indices, and ϵ is measurement error, respectively. The best model was then used in all subsequent steps.

Field Gridding and Trait Value Prediction

With the best model, the wheat yield over the entire field was calculated in ArcGIS Spatial Analyst v. 10.6 software (Esri, Redlands, California, United States) to map the aerial photographs. The field was divided into 25 × 25 cm cells ($c_1, c_2, \cdots, c_{n_c}$, where n_c is the total number of cells) using GIS. With the model (f), the pixel-by-pixel predicted values of the target traits, including wheat yield, were calculated for the entire

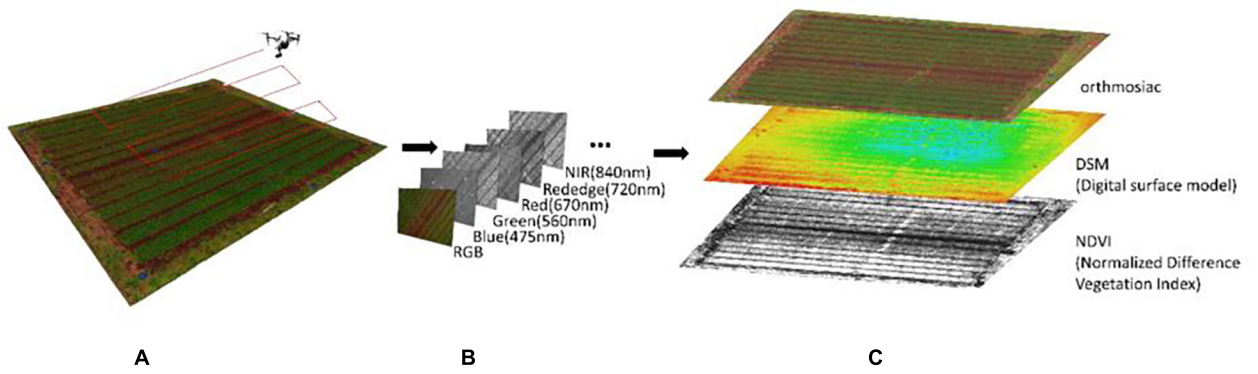

FIGURE 3 | Representation of the image capture and data processing in the GAUSS system. **(A)** UAV overflight to capture image data; **(B)** types of image data collected; **(C)** results of image processing.

field ($\hat{y}_{p_1}$, $\hat{y}_{p_2}$, $\cdots$, $\hat{y}_{p_{n_p}}$, where p_i is the i-th pixel in the UAV image of the field, and n_p is the total number of pixels in the image). Then the average yield of wheat in each cell (c_i) was calculated from the pixel-by-pixel values:

$$\hat{y}_{c_i} = \sum_{p \text{ in } c_i} \hat{y}_p / n_{c_i} \quad (2)$$

where n_{c_i} is the total number of pixel in each cell.

Cells that contained two plots during soybean cultivation and cells containing both corridor and experimental plot areas during wheat cultivation were eliminated (**Figure 2C**). This resulted in 8756 cells of predicted wheat yield values (mean of 125 cells per soybean plot). To assess the relationship between predicted yield and manually sampled yield, the overall distributions of these values for each plot were compared.

Statistical Testing of the Predicted Values

The experimental factors affecting the predicted wheat yields of each cell were analyzed in two ways. The first set of analyses examined whether the spatial variation of wheat yield was affected by the previously grown soybean cultivar, using a generalized linear mixed model (GLMM) with Gaussian distribution and identity link. In this model, predicted wheat yields in each cell were treated as the response variable, and sowing date and cultivar were treated as the explanatory variables. To account for systematic error due to spatial variations in the field and to avoid pseudo-replication caused by repeated observations from the same plots, the sowing row and plot identity (nested within sowing row) were treated as random effects. A significant difference among cultivars was found, so pairwise comparisons between the cultivars that produced the lowest wheat yield and other cultivars were carried out.

The second set of analyses investigated which soybean traits affected the yield of the subsequently grown wheat, again using the GLMM with Gaussian distribution and identity link methodology. Predicted wheat yields of each cell were treated as the response variable, and soybean stem dry weight, seed dry weight, above-ground dry weight, 100-seed weight, and their first-order interactions were treated as explanatory variables. These are typical soybean traits that are measured (Cui et al., 2001; Kaga et al., 2011; Qiu et al., 2013). The sowing row and plot identity (nested within sowing row) were treated as random effects.

The "lme4" package and the "lmer" function in the R software environment for GLMM analyses (R Development Core Team, 2010; Bates et al., 2014) were used. The likelihood ratio test was used to test the significance of the GLMM results. Graphs of GLMM predictions were drawn in the sjPLot package for R (Lüdecke, 2018).

RESULTS

Model Selection for Wheat Yield

AIC showed that the best model included four UAV-based indices (i.e., vegetation cover area on 14 March and 12 April, height on 14 March, and NDVI on 12 April; **Table 1**). There was a high correlation (R^2 = 0.8061) between the values predicted by this model and the observed manually measured yield (**Figure 4A**). Among the models that were explored, all of the top 20 included both cover area on 14 March and NDVI on 12 April (e.g., **Table 1**).

Comparison Between Manually Sampled and Predicted Values

The set of predicted yields in each plot (excluding the 6 plots used for destructive sampling of soybean plants) was determined by using the GAUSS framework. The manually collected wheat yield of a plot often differed from the set of predicted values for that plot (**Figure 4B**). The manually collected values fell outside the quartiles of the predicted values in 54.7% of the plots (35 of 64 plots) and most of these (25) were less than the 25% quartile of the overall distribution of predicted values for the plot.

Differences Among Soybean Cultivars

The measured traits of the soybeans showed considerable variability among cultivars (**Supplementary Table 2**). Interestingly, conventional statistical analysis of the actual wheat yield data did not identify significant variations among the soybean cultivars in their effect on subsequent wheat yield

TABLE 1 | The results of model selection, ranked by Akaike information criteria (AIC), in the search to identify the best model for predicting ear dry weight of wheat (*Triticum aestivum*) from UAV imagery data.

Model No.	Explanatory variables included in the models							R^2	DF	AIC	ΔAIC	Weight
1		C_Mar.14	H_Mar.14		C_Apr.12		N_Apr.12	0.8061	6	2375.4	0	0.082
2		C_Mar.14			C_Apr.12		N_Apr.12	0.8035	5	2375.5	0.03	0.081
3		C_Mar.14	H_Mar.14		C_Apr.12	H_Apr.12	N_Apr.12	0.8083	7	2375.6	0.21	0.074
4		C_Mar.14			C_Apr.12	H_Apr.12	N_Apr.12	0.8058	6	2375.7	0.23	0.073
5	C_Feb.15	C_Mar.14	H_Mar.14		C_Apr.12		N_Apr.12	0.808	7	2375.9	0.47	0.065
16		C_Mar.14	H_Mar.14	N_Mar.14	C_Apr.12	H_Apr.12	N_Apr.12	0.809	8	2377.1	1.66	0.036
17	C_Feb.15	C_Mar.14	H_Mar.14			H_Apr.12	N_Apr.12	0.8045	7	2378.6	3.21	0.016
18		C_Mar.14	H_Mar.14			H_Apr.12	N_Apr.12	0.8016	6	2378.9	3.5	0.014
19		C_Mar.14				H_Apr.12	N_Apr.12	0.7983	5	2379.4	4.01	0.011
20	C_Feb.15	C_Mar.14				H_Apr.12	N_Apr.12	0.8	6	2380.2	4.74	0.008

C, vegetation cover area; H, averaged height; N, averaged normalized difference vegetation index (NDVI) of the sampling points.

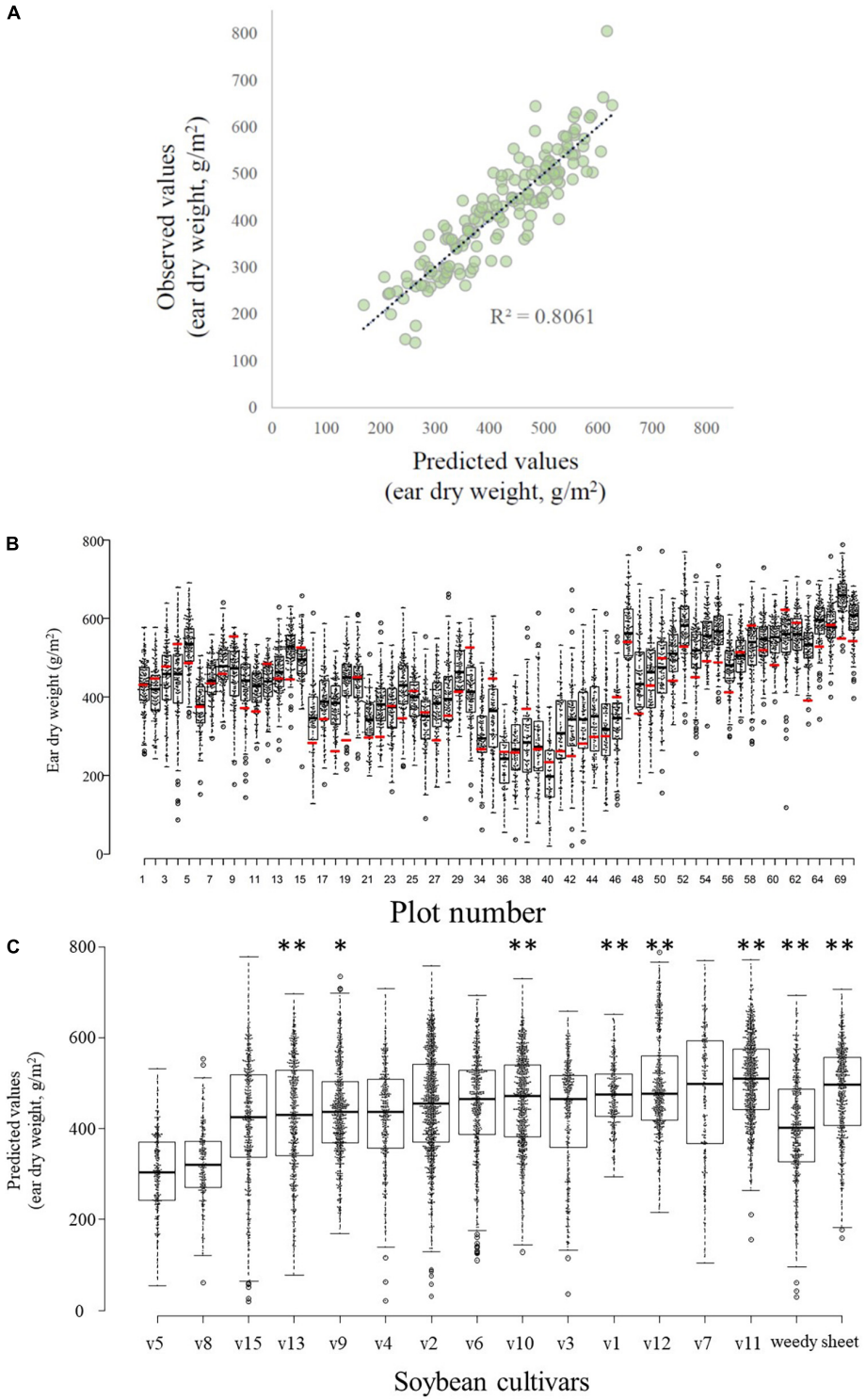

FIGURE 4 | Relationships between predicted and observed values. **(A)** Yields (wheat ear dry weight) predicted by the best model vs. observed values. **(B)** Boxplots and scatter plots of the distribution of predicted wheat yields in each experimental plot (black bar within a box indicates median predicted yield; box bottom and top, 25 and 75% quartiles, respectively, whiskers, 1.5× the interquartile range; open circles outside the box, outliers) and manually sampled yield values (red bar). Note that the 6 field plots (P31–33, P66–68) used for destructive sampling of soybean (*Glycine max*) plants are not included here. A total of 64 box plots are shown, therefore the numbering of the *x*-axis is discontinuous. **(C)** Effects of the different soybean cultivars grown before the wheat crop on the predicted wheat ear dry weights. Boxplot features are as described in **(B)**. Asterisks indicate significant differences (*$P < 0.05$; **$P < 0.01$) between the predicted values for cultivar v5 (which produced the lowest wheat yield) and each other soybean cultivar or weed management method.

(ANOVA, $F = 0.928$, $p = 0.536$, **Supplementary Figure 2**). However, the GAUSS analysis identified significant differences among cultivars in the predicted wheat yields (**Figure 4C** and **Supplementary Table 4**), after the effects of spatial variations were removed statistically. Cultivar v5 was associated with the lowest yield in the subsequent wheat crop, and v11 was associated with the highest yield (**Figure 4C**). There was no significant difference in wheat yields between soybean sowing dates (**Supplementary Table 4**).

Soybean Traits Associated With Wheat Yield

The interaction of soybean stem dry weight × above-ground dry weight had a significant effect on wheat yield (ear dry weight; $P < 0.05$; **Supplementary Table 5**). Increased above-ground dry weight of whole soybean plants apparently reduced subsequent wheat yield, but as the stem weight of soybean increased, the wheat yields also increased (**Figure 5A**). Interestingly, where the above-ground weight of soybean plants was low, the wheat yield was high, regardless of the stem weight of soybean (**Figure 5A**). In fact, the soybean cultivars associated with low yields in wheat were usually those that had relatively large above-ground weight and relatively small stem weight, and vice versa (**Figures 4C**, **5B**). The 100-seed weight of soybean seeds also significantly affected the wheat yield ($P < 0.05$, **Supplementary Table 5** and **Supplementary Figure 3**). The interactions of aboveground dry weight × seed dry weight and of stem dry weight × seed dry weight did not significantly affect wheat yield (**Supplementary Table 5**).

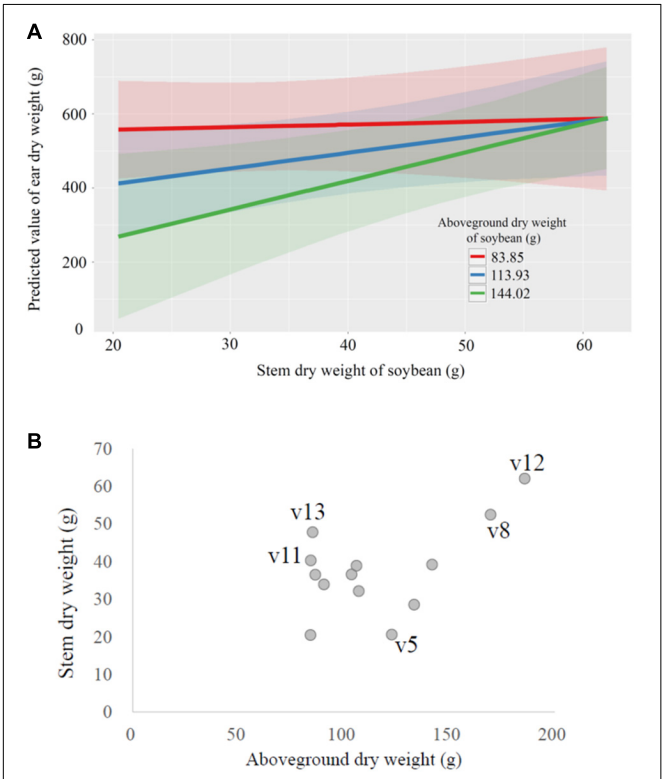

FIGURE 5 | Relationships between selected crop characteristics. **(A)** The marginal effects of the above-ground dry weight and the stem dry weight of soybeans (*Glycine max*) on the predicted ear dry weight of the subsequently grown wheat (*Triticum aestivum*) crop. **(B)** The relationship between above-ground dry weight and stem dry weight of the 14 soybean cultivars.

DISCUSSION

This study used a simple analytical framework, identified as GAUSS, to analyze UAV-supported data from field experiments. The performance of this framework was assessed by analyzing data from a crop rotation trial of soybean and wheat. This framework acquired a large amount of high-spatial-resolution data for predicted wheat yield. Analysis of this data showed significant differences among the soybean cultivars in the yield of the wheat grown after them, and identified the soybean traits associated with increased yield. Despite the long history of legume plants in crop rotations (Chorley, 1981; Stinner et al., 1992), to the best of our knowledge this is the first field study that has identified which traits were associated with the benefits of rotation. Although the conventional analysis of manual sampling data did not identify significant differences among soybean cultivars in their effect on wheat yield (**Supplementary Figure 2**), the GAUSS approach did detect such differences. The large quantity of predicted values with location information generated by this methodology enabled the statistical analysis to include the intra- and inter-plot variations. This suggests that GAUSS has the potential to considerably enhance field experimentation, thereby improving its usefulness.

Soybean cultivars with relatively small above-ground weight, large stem weight, and low 100-seed weight were associated with increased yield in the subsequently grown wheat crop. This suggests several implications for studies of legume-based crop rotations. First, the negative effect of increased above-ground weight of soybean on wheat yield may be attributable to the removal of the above-ground parts of soybean from the field at harvest. Soybean cultivars with large above-ground weight likely absorbed more soil nutrients from the soil than the cultivars that produced small above-ground parts. This removal would have decreased the available nutrient pool for subsequent wheat growth. Second, a study comparing 383 soybean cultivars shows a high correlation between stem and root weight ($r^2 = 0.81$–0.90; Nakamura and Sawahata, 1988). Therefore the soybean cultivars with large stem weights likely produced large amounts of roots, which may have affected the soil physical and chemical conditions; for example, aggregate structure, release of nitrogen compounds, and biological activities (decomposition) may have been enhanced, increasing the yield of the subsequent wheat crop. However, it should be noted that this experiment was conducted in a single location and a single growing season. Multi-year trials at various locations would be necessary for a more convincing conclusion and a cost-effective GAUSS approach would be useful.

This study did not measure any below-ground soybean traits, such as root biomass or number of nodules. Variations among soybean cultivars in the symbiotic performance of rhizobia are known to occur (Appunu et al., 2008). The variations in the effect

of soybean cultivars on wheat yield observed in this experiment may reflect these differences in rhizobial activity. However, preliminary assessments found that the number of rhizobium nodules on the roots of soybeans grown in this field was low. There were no clear differences among the cultivars (data not shown), likely owing to the relatively high soil nutrient content in this field. Future studies are needed to clarify how soybean root residues change soil biophysical properties and increase the yield of subsequent crops. It remains unclear why the 100-seed weight reduced the wheat yield. Chromosome segment substitution lines for 100-seed weight (Liu et al., 2018) may be useful for exploring the causal relationship, but that is beyond the scope of the work reported here. The differences in total seed weight among the soybean cultivars were not associated with differences in wheat yield. This suggests that it may be possible to select soybean cultivars that provide both sufficient soybean yields and crop rotation benefits for the subsequent crop.

In this study, we did not measure the soil nitrate concentration for each rotational treatment. In crop rotations with legumes, the effect of soil nitrogen accumulation on subsequent crops has been reported to be highly variable (Walley et al., 2007; Anglade et al., 2015). One reason for the variation is that non-nitrogen (non-N) factors (such as the bio-physical change in soil properties due to legume residues, other plant nutrients, disease suppression, and weed control) contribute significantly to crop rotational benefits (Stevenson and Van Kessel, 1996; Arcand et al., 2014; Uzoh et al., 2017). In order to verify how the soybean varieties affect the yield of subsequent crops, it will be necessary not only to quantify the change in the soil nitrogen accumulation, but to also examine other bio-physical factors.

The present study applied the GAUSS approach to a crop rotation experiment, but the approach is applicable to a wide range of field experiments. For example, it could be used to study the effects of environmental conditions (water, fertilizer, pesticide, etc.) on yield, with high spatial resolution. GAUSS may also be useful for field experiments in ecology and environmental science. It could be used to measure any plant traits that can be estimated from UAV (or potentially satellite) imagery. For example, UAV imaging and image analysis may enable estimation of important functional traits of complex plant communities, such as biomass, volume, plant height, and photosynthetic activity. The relationship between these functional traits and plant diversity has been examined in grassland field experiments all over the world (Tilman et al., 2006; Cardinale et al., 2007; Zavaleta et al., 2010; Sasaki et al., 2017), but the GAUSS approach has potential to greatly enhance these experiments.

GAUSS also has potential to markedly reduce the time needed for yield surveys, which could facilitate greater numbers of experimental treatments and replicates. For example, in the study reported here, manual collection and measurement of 157 one-square-meter wheat samples required more than 450 person-hours, whereas the UAV drone surveillance, data processing, and GIS analysis took approximately 20 person-hours for GAUSS to estimate the yield of 8,756 cells (total 547.25 m^2) in the same experiment.

The GAUSS approach worked well here, but there are many limitations and challenges that remain to be addressed to facilitate its extensive adoption for field experimentation. First, although the predictive model based on the UAV imagery was relatively accurate (r^2 = 0.8061), this model was based on a relatively small number of UAV imagery datasets on 3 days (a total of seven variables: two height values, two NDVI values, and three cover area values). Increasing the frequency of UAV sensing and adding more explanatory variables could produce an even better predictive model. The method used here to identify the best model was limited by the number of UAV surveys and the number of explanatory variables. If the frequency of UAV sensing is increased and a large number of explanatory variables is included, more flexible analytical methods, such as machine learning, may be useful for estimating models.

Second, the size of grid cells needs to be optimized. In this study, the GIS grid size (25 cm × 25 cm) was based on the size of an individual wheat plant. However, the grid can be any size, depending on the size and scale of the target species.

Third, the appropriate size for experimental plots needs to be determined. Here, the plot shape and size (2.4 m × 4.2 m) was similar to those in typical field experiments. However, GAUSS can detect differences using smaller plot sizes, which could improve the efficiency of field experiments. Future studies will be needed to identify and validate appropriate plot sizes for UAV-supported field experiments.

Fourth, although it is more than 100 times more efficient per unit area than manual measurement, the GAUSS method always needs manually sampled data from which to develop its predictive models.

Fifth, the GAUSS approach is somewhat expensive because it requires a UAV (drone) with a multispectral camera and RTK-GPS. However, these costs are likely to decrease substantially as the technology develops further and becomes more widely used.

Sixth, although development of relatively good predictive models may help overcome the inherent large variations among manually collected samples, examinations of diverse plant species under various field conditions (e.g., rice, potato, and maize, in uniform vs. non-uniform fields) are needed to investigate the variability of GAUSS data within field experiments.

Seventh, in step 1 of GAUSS, we used commercial photogrammetric software (Pix4Dmapper Pro, Pix4D, Lausanne, Switzerland) to run the 3D reconstruction of the field. In step 2 and 3, we also run the separated scripts to sample the field and calculate phenotypic traits, those require several manual operations, which are time-consuming. In the future, building an automated pipeline (e.g., CIAT Pheno-i, Selvaraj et al., 2020) will allow us to build predictive models more cost-effectively, easily-to-sue, and quickly.

In conclusion, a new analytical framework for UAV-supported field experimentation was proposed. This framework may be applicable to a wide range of field experimentation in crops and wild plants. It could improve the way field experiments are conducted, which has not changed much since Fisher's era. Wider usage and resolution of the limitations and challenges will likely enable the proposed GAUSS framework to be used consistently in future field experiments.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

YF, WG, SO, NA, SK, and HK conducted field experiment. WG conducted drone sensing and image analysis. All authors discussed, wrote the manuscript, and designed the experiment.

FUNDING

This work was partially supported by JSPS KAKENHI Grant No. 18KT0087, the Takano Life Science Research Foundation, the JST CREST Program JPMJCR1512, SICORP Program JPMJSC16H2, aXis program JPMJAS2018, and Japan Science and Technology Agency PRESTO Grants JPMJPR17Q8.

ACKNOWLEDGMENTS

We thank K. Ichikawa, D. Ishizuka, and K. Yatsuda for assistance with GPS measurement and crop production management, M. Nakajima and C. Tanaka for weed management, and members of the Institute for Sustainable Agro-ecosystem Services for helpful discussions and support of the research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.637694/full#supplementary-material

Supplementary Figure 1 | The locations of the manual sampling points in the experimental field.

Supplementary Figure 2 | Comparison of the effects of different soybean (Glycine max) cultivars grown before the wheat (Triticum aestivum) crop on ear dry weight of wheat.

Supplementary Figure 3 | The marginal effect of 100-seed dry weight of soybean (Glycine max) on the predicted values of ear dry weight of subsequently grown wheat (Triticum aestivum).

Supplementary Table 1 | Monthly means of mean air temperature, daily solar radiation, and total precipitation.

Supplementary Table 2 | Cultivar identities and values of measured traits of the soybeans (Glycine max) grown in this study.

Supplementary Table 3 | List of experimental plots.

Supplementary Table 4 | Results of generalized linear mixed model (GLMM) analysis of wheat (Triticum aestivum) ear weight in response to previously grown soybean (Glycine max) cultivars.

Supplementary Table 5 | Results of generalized linear mixed model (GLMM) analysis of wheat (Triticum aestivum) ear weight in response to specific traits of previously grown soybeans (Glycine max).

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* 19, 716–723. doi: 10.1109/TAC.1974.1100705

Anglade, J., Billen, G., and Garnier, J. (2015). Relationships for estimating N2 fixation in legumes: Incidence for N balance of legume-based cropping systems in europe. *Ecosphere* 6, 1–24. doi: 10.1890/ES14-00353.1

Appunu, C. C., Sen, D., Singh, M. K., and Dh, B. (2008). Variation in symbiotic performance of *Bradyrhizobium japonicum* strains and soybean cultivars under field conditions. *J. Cent. Eur. Agric.* 9, 185–189. doi: 10.5513/jcea.v9i1.509

Arcand, M. M., Knight, J. D., and Farrell, R. E. (2014). Differentiating between the supply of N to wheat from above and belowground residues of preceding crops of pea and canola. *Biol. Fertil. Soils* 50, 563–570. doi: 10.1007/s00374-013-0877-4

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv*. arXiv:1406.5823

Becker, M., and Johnson, D. E. (1998). Legumes as dry season fallow in upland rice-based systems of West Africa. *Biol. Fertil. Soils* 27, 358–367. doi: 10.1007/s003740050444

Biederbeck, V. O., Zentner, R. P., and Campbell, C. A. (2005). Soil microbial populations and activities as influenced by legume green fallow in a semiarid climate. *Soil Biol. Biochem.* 37, 1775–1784. doi: 10.1016/j.soilbio.2005.02.011

Box, J. F. (1980). R.A. Ffisher and the design of experiments, 1922-1926. *Am. Stat.* 34, 1–7. doi: 10.1080/00031305.1980.10482701

Cardinale, B. J., Wright, J. P., Cadotte, M. W., Carroll, I. T., Hector, A., Srivastava, D. S., et al. (2007). Impacts of plant diversity on biomass production increase through time because of species complementarity. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18123–18128. doi: 10.1073/pnas.0709069104

Chorley, G. P. H. (1981). The agricultural revolution in northern europe, 1750-1880: nitrogen, legumes, and crop productivity. *Econ. Hist. Rev.* 34, 71–93. doi: 10.1111/j.1468-0289.1981.tb02007.x

Cui, Z., Carter, T. E., Burton, J. W., and Wells, R. (2001). Phenotypic diversity of modern Chinese and North American soybean cultivars. *Crop Sci.* 41, 1954–1967. doi: 10.2135/cropsci2001.1954

Duc, G., Agrama, H., Bao, S., Berger, J., Bourion, V., De Ron, A. M., et al. (2015). Breeding a annual grain legumes for sustainable agriculture: new methods to approach complex traits and target new ccultivar ideotypes. *CRC. Crit. Rev. Plant Sci.* 34, 381–411. doi: 10.1080/07352689.2014.898469

Edmondson, R. N. (2005). Past developments and future opportunities in the design and analysis of crop experiments. *J. Agric. Sci.* 143, 27–33. doi: 10.1017/S0021859604004472

Fisher, R. A. (1926). The arrangement of field experiments. *J. Minist. Agric. Gt. Britain* 33, 503–513.

Furbank, R. T., and Tester, M. (2011). Phenomics - technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Giller, K. E., and Cadisch, G. (1995). Future benefits from biological nitrogen fixation: an ecological approach to agriculture. *Plant Soil* 174, 255–277. doi: 10.1007/BF00032251

Guo, W., Fukano, Y., Noshita, K., and Ninomiya, S. (2020). Field-based individual plant phenotyping of herbaceous species by unmanned aerial vehicle. *Ecol. Evol.* 10, 12318–12326. doi: 10.1002/ece3.6861

Guo, W., Zheng, B., Duan, T., Fukatsu, T., Chapman, S., and Ninomiya, S. (2017). EasyPCC: benchmark datasets and tools for high-throughput measurement of the plant canopy coverage ratio under field conditions. *Sensors (Switzerland)* 17, 1–13. doi: 10.3390/s17040798

Guo, W., Zheng, B., Potgieter, A. B., Diot, J., Watanabe, K., Noshita, K., et al. (2018). Aerial imagery analysis – quantifying appearance and number of sorghum

heads for applications in breeding and agronomy. *Front. Plant Sci.* 9:1544. doi: 10.3389/fpls.2018.01544

Hassan, M. A., Yang, M., Rasheed, A., Yang, G., Reynolds, M., Xia, X., et al. (2019). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Sci.* 282, 95–103. doi: 10.1016/j.plantsci.2018.10.022

Herridge, D., and Rose, I. (2000). Breeding for enhanced nitrogen fixation in crop legumes. *Field Crop. Res.* 65, 229–248. doi: 10.1016/S0378-4290(99)00089-1

Houle, D., Govindaraju, D. R., and Omholt, S. (2010). Phenomics: the next challenge. *Nat. Rev. Genet.* 11, 855–866. doi: 10.1038/nrg2897

Hurlbert, S. (1984). Pseudoreplication and the design of ecological field experiments. *Eciol. Monogr.* 54, 187–211.

Kaga, A., Shimizu, T., Watanabe, S., Tsubokura, Y., Katayose, Y., Harada, K., et al. (2011). Evaluation of soybean germplasm conserved in NIAS genebank and development of mini core collections. *Breed. Sci.* 61, 566–592. doi: 10.1270/jsbbs.61.566

Legendre, P., Dale, M. R. T., Fortin, M. J., Casgrain, P., and Gurevitch, J. (2004). Effects of spatial structures on the results of field experiments. *Ecology* 85, 3202–3214. doi: 10.1890/03-0677

Liu, D., Yan, Y., Fujita, Y., and Xu, D. (2018). Identification and validation of QTLs for 100-seed weight using chromosome segment substitution lines in soybean. *Breed. Sci.* 68, 442–448. doi: 10.1270/jsbbs.17127

Lüdecke, D. (2018). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.1.

Maes, W. H., and Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci.* 24, 152–164. doi: 10.1016/j.tplants.2018.11.007

Nakagawa, S., and Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol. Rev.* 82, 591–605. doi: 10.1111/j.1469-185X.2007.00027.x

Nakamura, S., and Sawahata, H. (1988). Ratio of stem weight to root weight of soybean cultivars. *Jpn. J. Crop Sci.* 57, 621–626.

Payne, R. W. (2006). New and traditional methods for the analysis of unreplicated experiments. *Crop Sci.* 46, 2476–2481. doi: 10.2135/cropsci2006.04.0273

Qiu, L. J., Xing, L. L., Guo, Y., Wang, J., Jackson, S. A., and Chang, R. Z. (2013). A platform for soybean molecular breeding: the utilization of core collections for food security. *Plant Mol. Biol.* 83, 41–50. doi: 10.1007/s11103-013-0076-6

R Development Core Team, (2010). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation statistics Computing.

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., et al. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. *Eur. J. Agron.* 70, 112–123. doi: 10.1016/j.eja.2015.07.004

Sasaki, T., Yoshihara, Y., Takahashi, M., Byambatsetseg, L., Futahashi, R., Nyambayar, D., et al. (2017). Differential responses and mechanisms of productivity following experimental species loss scenarios. *Oecologia* 183, 785–795. doi: 10.1007/s00442-016-3806-z

Selvaraj, M. G., Valderrama, M., Guzman, D., Valencia, M., Ruiz, H., Acharjee, A., et al. (2020). Machine learning for high-throughput field phenotyping and image processing provides insight into the association of above and below-ground traits in cassava (*Manihot esculenta* Crantz). *Plant Methods* 16, 1–19. doi: 10.1186/s13007-020-00625-1

Sokal, R. R., and Rohlf, F. J. (1995). *Biometry: the Principles and Practice of Statistics in Biological Research*, 3rd Edn. New York, NY: Freeman.

Stevenson, F. C., and Van Kessel, C. (1996). The nitrogen and non-nitrogen rotation benefits of pea to succeeding crops. *Can. J. Plant Sci.* 76, 735–745. doi: 10.4141/cjps96-126

Stinner, D. H., Glick, I., and Stinner, B. R. (1992). Forage legumes and cultural sustainability: lessons from history. *Agric. Ecosyst. Environ.* 40, 233–248. doi: 10.1016/0167-8809(92)90095-S

Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., and Bennett, M. (2017). Plant phenomics, from sensors to knowledge. *Curr. Biol.* 27, R770–R783. doi: 10.1016/j.cub.2017.05.055

Tilman, D., Reich, P. B., and Knops, J. M. H. (2006). Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* 441, 629–632. doi: 10.1038/nature04742

Tripodi, P., Massa, D., Venezia, A., and Cardi, T. (2018). Sensing technologies for precision phenotyping in vegetable crops: current status and future challenges. *Agronomy* 8:57. doi: 10.3390/agronomy8040057

Uzoh, I. M., Obalum, S. E., Igwe, C. A., and Abaidoo, R. C. (2017). Quantitative separation of nitrogen and non-nnitrogen rotation benefits for maize following velvet bean under selected soil mmanagement practices. *Agric. Res.* 6, 378–388. doi: 10.1007/s40003-017-0272-8

van Kessel, C., and Hartley, C. (2000). Agricultural management of grain legumes: has it led to an increase in nitrogen fixation? *Field Crop. Res.* 65, 165–181. doi: 10.1016/S0378-4290(99)00085-4

Walley, F. L., Clayton, G. W., Miller, P. R., Carr, P. M., and Lafond, G. P. (2007). Nitrogen economy of pulse crop production in the Northern Great Plains. *Agron. J.* 99, 1710–1718. doi: 10.2134/agronj2006.0314s

Watanabe, K., Guo, W., Arai, K., Takanashi, H., Kajiya-Kanegae, H., Kobayashi, M., et al. (2017). High-throughput phenotyping of sorghum plant height using an unmanned aerial vehicle and its application to genomic prediction modeling. *Front. Plant Sci.* 8:421. doi: 10.3389/fpls.2017.00421

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. *Front. Plant Sci* 8:1111. doi: 10.3389/fpls.2017.01111

Yang, R. C. (2010). Towards understanding and use of mixed-model analysis of agricultural experiments. *Can. J. Plant Sci.* 90, 605–627. doi: 10.4141/CJPS10049

Zavaleta, E. S., Pasari, J. R., Hulvey, K. B., and Tilman, G. D. (2010). Sustaining multiple ecosystem functions in grassland communities requires higher biodiversity. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1443–1446. doi: 10.1073/pnas.0906829107

Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., Yao, X., et al. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS J. Photogramm. Remote Sens.* 130, 246–255. doi: 10.1016/j.isprsjprs.2017.05.003

Conflict of Interest: 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.

frontiers in Plant Science

OPINION
published: 22 October 2021
doi: 10.3389/fpls.2021.749374

Bridging the Gap Between Remote Sensing and Plant Phenotyping—Challenges and Opportunities for the Next Generation of Sustainable Agriculture

Miriam Machwitz[1], Roland Pieruschka[2], Katja Berger[3], Martin Schlerf[1], Helge Aasen[4], Sven Fahrner[2], Jose Jiménez-Berni[5], Frédéric Baret[6] and Uwe Rascher[7]*

[1] Department of Environmental Research and Innovation, Luxembourg Institute of Science and Technology, Belval, Luxembourg, [2] Institute of Bio and Geosciences, Plant Sciences, Forschungszentrum Jülich, Helmholtz-Verband Deutscher Forschungszentren, Jülich, Germany, [3] Department of Geography, Ludwig-Maximilians-Universität München, Munich, Germany, [4] Department of Environmental Systems Science, Crop Science, Eidgenössische Technische Hochschule (ETH) Zurich, Zurich, Switzerland, [5] Instituto de Agricultura Sostenible, Consejo Superior de Investigaciones Científicas, Cordoba, Spain, [6] INRAE-EMMAH-CAPTE, Avignon, France, [7] Forschungszentrum Jülich, Institute of Bio- and Geosciences Plant Sciences (IBG-2), Jülich, Germany

Keywords: remote sensing, high-throughput field phenotyping, unmanned aerial vehicles (UAVs), multi-sensor synergies, open-data standards, vegetation traits, radiative transfer models (RTM), smart farming

OPEN ACCESS

Edited by:
Alessandro Matese, National Research Council (CNR), Italy

Reviewed by:
Piero Toscano, National Research Council (CNR), Italy

****Correspondence:***
Miriam Machwitz
miriam.machwitz@list.lu

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

Received: *29 July 2021*
Accepted: *27 September 2021*
Published: *22 October 2021*

Citation:
Machwitz M, Pieruschka R, Berger K, Schlerf M, Aasen H, Fahrner S, Jiménez-Berni J, Baret F and Rascher U (2021) Bridging the Gap Between Remote Sensing and Plant Phenotyping—Challenges and Opportunities for the Next Generation of Sustainable Agriculture. Front. Plant Sci. 12:749374. doi: 10.3389/fpls.2021.749374

INTRODUCTION

Sustainable and resilient agriculture with a low impact on the environment is pivotal to ensure food security for a growing global population. This is of particular importance faced with the unprecedented challenge of climate change (FAO., 2017) for crop production. Sustainable intensification or currently rather the conservation of yield (Rosenqvist et al., 2019) requires the consideration of the entire crop production pipeline, ranging from breeding and identifying varieties adapted to specific environmental conditions, to improving agricultural land management (agriculture 5.0, Saiz-Rubio and Rovira-Más, 2020). An essential aspect of these efforts is the quantitative assessment of the plant traits contributing to increased, reliable production and the efficient use of resources, such as nutrients or water. Faced climate change and the appearance of more frequent and intense stress events, there is a need for resilient breeding lines, as summarized in the review of Razzaq et al. (2021). Besides drought stress, heat stress is expected to have a major negative impact on yield in Europe (Semenov and Shewry, 2011).

In this context, two areas of research, plant phenotyping and remote sensing, are becoming increasingly important. Field phenotyping refers to a quantitative description of a plant's phenotype—i.e., its anatomical, ontogenetical, physiological, and biochemical properties—in its natural environment (Walter et al., 2017). Remote sensing in the agricultural context is the observation of vegetation by a remote device and the retrieval of its qualitative or quantitative properties. While remote sensing and plant phenotyping researchers are both interested in the interaction of plant growth with the environment (including management practices), the two fields have a different focus. Traditionally, remote sensing is used to estimate spatial trends across the landscape, while plant phenotyping aims to remove spatial effects in their data in order to investigate the genetic effects of different plant varieties in response to the prevailing environmental conditions. Nevertheless, both disciplines are united in their efforts to estimate plant traits and explain apparent differences in the phenotype precisely (Aasen and Herrera, (under review)).

Driven by the need for new concepts in sustainable agriculture, an increased use of remote sensing approaches in field phenotyping and vice versa has been observed over the last decade. On one hand, field phenotyping has increasingly deployed imaging instruments traditionally used in remote sensing (Johansen et al., 2019) to meet the need for increased throughput in field phenotyping (Araus and Cairns, 2014). The analysis of remote sensing data by non-experts without full knowledge of the sensing principles hampers the exploitation of the full potential of the methods at hand. On the other hand, remote sensing scientists have started to estimate plant traits and analyze data from breeding experiments (Yang et al., 2017). However, their findings are often not interpreted in light of the physiological processes that shape the relation of a crop within the environment. Additionally, there are differences in input data, acquisition protocols, plant trait definitions, and retrieval models that hinder close cooperation between the two disciplines (**Figure 1**).

Facilitating exchange between the two disciplines offers possibilities to trigger cross-fertilization: An improved understanding of the target traits will allow the remote sensing community to develop more precise and ultimately more useful tools. Likewise, establishing state-of-the-art remote sensing methods as plant phenotyping tools will allow an improved understanding and modeling of crops in dynamic environments. Ultimately, this exchange has the potential to stimulate growth in both communities and their interconnection may lead to new developments toward more sustainable agriculture. There is a need for multiple stress-resilient breeding lines combined with a need for multi-site and multi-regional testing (Rosenqvist et al., 2019). Breeders and plant phenotyping need to provide breeding lines that are able to cope with unprecedented stress conditions. To target sustainable and resilient agriculture, we propose that remote sensing should develop toward near-real time monitoring of certain traits on large scales under several environmental conditions (climate, soil etc.) as a global multi-site experiment. The current work on real-time observations (for example, special issue of MPDI remote sensing in 2021: "Near Real-Time (NRT) Agriculture Monitoring" https://www.mdpi.com/journal/remotesensing/special_issues/NRT_agriculture_monitoring) offers, on the one hand, the possibility to give timely management advice, which again could be optimized through joint research between the two communities. On the other hand, the remotely sensed information on plant traits and their actual condition could be directed back to field phenotyping experts to optimize the breeding lines with respect to certain environmental constraints. A similar concept for forests was proposed by Dungey et al. (2018).

We initiated a discussion between more than 130 experts from the remote sensing and plant phenotyping community in the context of a joint workshop (https://www.senseco.eu/working-groups/wg3-sensor-synergies/) of the COST action SENSECO "Optical synergies for spatiotemporal SENsing of Scalable ECOphysiological traits" (https://www.senseco.eu/) and the ESFRI plant phenotyping infrastructure EMPHASIS "European Infrastructure for Plant Phenotyping" (https://emphasis.plant-phenotyping.eu/). During the discussion, we identified the following key areas for future collaboration:

(i) transferring and harmonizing knowledge on protocols, methods, and data between the two communities;
(ii) optimizing quantitative trait estimation by using new sensors, and integrating data from different spectral domains and spatial resolutions, preferably in real-time;
(iii) linking existing and new modeling approaches and recent developments in artificial intelligence to bridge different observation scales through space and time.

DATA EXCHANGE AND PROTOCOL STANDARDIZATION

Plant phenotyping and remote sensing collect a large amount of data, including spectral observations and biochemical/biophysical plant traits. However, the exchange of these data requires the standardization of measurement protocols and harmonization of measurement procedures. Thus, a broad exchange of measuring concepts for plant trait assessments should be initiated, complemented by an open data policy allowing for the broader use and re-use of data (Fiorani and Schurr, 2013; Reynolds et al., 2019). In particular, plant phenotyping scientists are developing a large number of solutions to address a diversity of crops, traits, and treatments. Since there is no one-size-fits-all solution, existing hardware and software solutions often need to be adapted, even for the same traits of interests in different crops. This has led to both a "Phenotyping Dilemma" as stated by Rosenqvist et al. (2019) and the need for harmonization. This was addressed within COST action FA1306, "The quest for tolerant varieties – phenotyping at plant and cellular level" (Phenomen-All) and the EU-funded projects EPPN and EPPN2020 leading to the ESFRI research infrastructure entitled EMPHASIS. Plant phenotyping will never be able to measure all genotypes under all relevant conditions, thus, further integration of the community and the use and development of existing synergies, such as those between remote sensing and phenotyping, are key to achieving the required impact of improved plant production in times of climate change. It is therefore vital to develop FAIR data approaches (Wilkinson et al., 2016) that link the communities, and sharing phenotyping data (Danilevicz et al., 2021) will benefit plant and crop sciences at large.

In remote sensing, standards for metadata collection by scientists often mainly regard information on sensor performance or geolocation, while auxiliary data about vegetation is often limited to the traits in focus. But, plant status is only a small function of just one individual co-variable and requires additional information on the biotic and abiotic environment and genetic makeup of the plant, thus, such an approach may result in an oversimplified interpretation of the remote sensing signal (Galieni et al., 2021). In contrast, explicit geolocation is essential to link the signal to the field observation and needs to be considered in field phenotyping data collection.

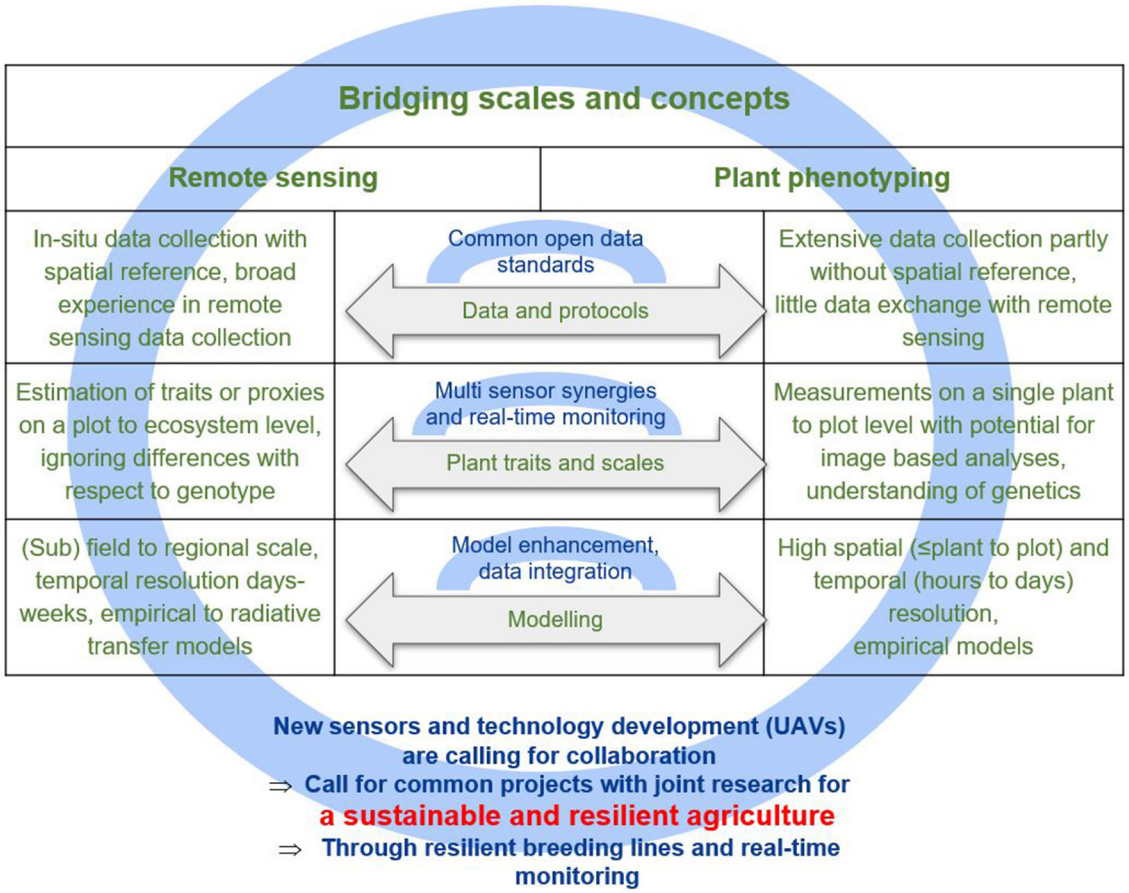

FIGURE 1 | A closer collaboration of remote sensing and plant phenotyping has the potential to foster research for a sustainable and resilient agriculture. Common topics between the two communities (gray arrows) have been identified for collaboration and cutting-edge research. In the left and right columns, differences between remote sensing and plant phenotyping are listed regarding the three topics. In blue, the potential for future collaboration is indicated.

The phenotyping community has developed the Minimum Information About a Plant Phenotyping Experiment (MIAPPE) as a bottom-up standard of metadata required to adequately describe a plant phenotyping experiment with well-defined data models and make the data reusable (Papoutsoglou et al., 2020). Such standards or developed recommendations and guidelines (Manfreda et al., 2018; Tmušić et al., 2020), like in the COST action CA16219 Harmonious (Harmonization of UAS techniques for agricultural and natural ecosystems monitoring, https://www.cost.eu/actions/CA16219/ and https://www.costharmonious.eu/), essentially represent a checklist of how to describe an experiment and could be adapted and extended by including the considerations of both communities to improve the interpretability, reusability, and transferability of data. This would also allow the exchange of data between the two communities and extrapolate results from one experiment to another. In particular, the phenotyping community could benefit from traits estimated within the landscape by remote sensing, while the remote sensing community could use data from the field phenotyping community to improve model development (c.f. section New Sensors for Quantitative Trait Estimation From the Plot and the Ecosystem Scale).

NEW SENSORS FOR QUANTITATIVE TRAIT ESTIMATION FROM THE PLOT AND THE ECOSYSTEM SCALE

The availability of remote sensing data has increased significantly over the last decade. Satellites with a high temporal, spectral, and spatial resolution like the Sentinel-2 sensors or Hyperspectral Precursor and Application Mission (PRISMA) allow for new or improved agricultural applications. Moreover, the advent of nano-satellites further improves revisit times and spatial resolution of satellite systems. Developments in sensor technology, measurement procedures, and data correction workflows have matured UAVs to reliable quantitative remote sensing systems (Aasen et al., 2018) and initiated a new era in

the remote sensing of crops (Zarco-Tejada, 2008; Herrmann and Berger, 2021). Today, a variety of (hyper) spectral, thermal, sun-induced fluorescence, and 3D/LiDAR instruments are available for UAVs. Consequently, they have also become a common tool for (high-throughput) field phenotyping (Yang et al., 2017) and have been proposed as a "game-changer in precision agriculture" (Maes and Steppe, 2019). Traits such as plant height/growth, pigments, canopy cover, and temperature, which are highly relevant for the vitality and performance of crops (Tattaris et al., 2016), can now be derived from UAV remote sensing data (Zarco-Tejada et al., 2012; Aasen and Bareth, 2018; Roth et al., 2018; Perich et al., 2020). Moreover, imaging spectroscopy from UAVs is able to capture data with a viewing geometry closer to the hemispherical-directional reflectance factors of satellite products (Aasen and Bolten, 2018) and thus may bridge the gap between field phenotyping experiments and landscape monitoring (Aasen and Herrera, (under review)). However, UAV flights need careful planning, consideration of regulations, and realistic estimation of manpower (Reynolds et al., 2019).

Another example where we expect that the increasing availability of UAV-based sensors will bridge the spatial gap for a better understanding of plant mechanisms is the assessment of photosynthesis (Quirós-Vargas et al., 2020). The analysis can be performed based on solar-induced fluorescence (SIF) in combination with established hyperspectral indices like the photochemical reflectance index (PRI) from remote sensing platforms (Rascher et al., 2015). SIF imaging provides information that could be used to identify genotypes that maintain a high level of photosynthetic activity under stress conditions. With the start of the FLEX mission (https://earth.esa.int/eogateway/missions/flex), global SIF data will be available to test stress resilience under varying environmental conditions, as stated in the introduction.

Furthermore, a combination of large-scale information with plant phenotyping, using high spatial and temporal trait measurement may help to identify different stress factors, in order to assess the stress stage and underlying mechanisms. For recent examples related to the identification of genotypes capable of tolerating biotic and abiotic stress, see recent reviews (Araus et al., 2018; Watt et al., 2020).

Thanks to the flexibility of remote sensing systems, the dynamic developments of plant traits can be assessed by standardized measurements with little effort. However, for more complex traits, such as the identification of biotic and abiotic crop stress, the selection of the most suitable sensor combination is challenging (Galieni et al., 2021; Berger et al., (in prep)). Nevertheless, remote sensing platforms such as UAVs and micro-satellites are overcoming traditional trade-offs between spatial, temporal, and spectral resolutions. Moreover, applying low-altitude and close-range remote sensing methods in combination with radiative transfer models (RTMs, c.f. section Bridging Observation Scales With Physically-Based Radiative Transfer Models and Machine Learning for Improved Trait Estimation) for field phenotyping allows several of the insights gained to be scaled to the ecosystem where they can be used for more precise field management (Velumani et al., 2021). In conclusion, the combined usage of different sensors may lead to an improved understanding of the actual carbon and water fluxes and to finding cultivars with higher resilience.

BRIDGING OBSERVATION SCALES WITH PHYSICALLY-BASED RADIATIVE TRANSFER MODELS AND MACHINE LEARNING FOR IMPROVED TRAIT ESTIMATION

In field phenotyping studies, parametric regression approaches are typically applied to link vegetation indices derived from multispectral data with plant traits. These models are easy to implement and require little expert knowledge. However, large datasets are needed for calibration and validation and still face the limited transferability of the established models to other crops and different environmental conditions. Moreover, especially when hyperspectral data are used, parametric regressions tend to under-exploit the comprehensive information content hidden in the contiguous spectral data (Verrelst et al., 2019). Therefore, remote sensing scientists have developed radiative transfer models (RTMs) simulating the interactions of the full optical wavelength range with leaves and canopies based on physical laws. Beyond the widely used one-dimensional (1-D) RTMs, which are suitable for homogenous canopy architectures, three-dimensional RTMs open up opportunities to analyze data generated by high-throughput field phenotyping experiments over row crops (Weiss et al., 2020). Thereby, remote sensing could provide spatio-temporal information on specific functional traits of interest. In combination with process modeling and data assimilation strategies, remote sensing could help to understand the processes in plants. One example in the context of sustainable agriculture is the estimation of nitrogen (N) use efficiency. Usually, N content is quantified indirectly from remote sensing data *via* the chlorophyll content (Chlingaryan et al., 2018) and very often still by a parametric regression based on vegetation indices. However, the quantification of the N content is challenging due to the unstable relation of N and chlorophyll, the very subtle spectral signals of proteins, and the dilution phenomenon, which often seems to be neglected in N concentration studies (Bossung et al. (under review)). Novel RTMs developed within the remote sensing community now allow the use of more flexible non-parametric models in the estimation N, which also take into account proteins and provide uncertainty estimations (Berger et al., 2020). Better estimating the plant N by combining information on plant physiology from plant phenotyping with new hyperspectral sensors giving a near-real-time estimation can provide input for optimized management strategies to reduce N applications and thus protect the water resources.

The integration of information from different spectral domains is complex and challenging. Models and tools have been developed to observe photochemistry and energy fluxes of the canopy. These include for example the SCOPE model where VIS/NIR data and fluorescence data are integrated (Van der Tol et al., 2009) but are still not perfect and are not widely used. Models and toolboxes are a big asset for the understanding

of the interaction between vegetation and the environment. Further developments and interdisciplinary work are desirable to optimally combine the information from different sensors to fully describe the different traits, their interactions and the linked environmental triggers.

To obtain functional traits or to indirectly assess crop stress from the diversity of spectral data, a modeling framework should be defined. We propose the use of (shallow) machine learning (ML) regression algorithms combined with RTMs, such as SCOPE coupled with leaf optical properties models (Féret et al., 2021). Within these hybrid methods, training data sets are generated by the RTM and are then learned by the ML algorithm to build the specific retrieval model. Also, deep learning algorithms could be employed, in particular when a large number of different data sets are available and to better describe the highly non-linear relationship between remotely sensed signals and traits of interest. As an additional feature, the quality of training data can be enhanced by implementing active learning heuristics, which recently achieved outstanding results in the estimation of specific traits (Berger et al., 2021; Verrelst et al., 2021). All in all, these hybrid workflows may become a cornerstone for precision agriculture and an essential element for the development of new breeding strategies (Lammerts van Bueren and Struik, 2017).

CONCLUSIONS AND FUTURE PERSPECTIVES

Plant phenotyping and remote sensing work with complementary measurements and concepts, but address the same challenge – namely, the quest for a more sustainable agriculture. Facilitating exchange between the two disciplines offers possibilities to trigger cross-fertilization: An improved understanding of target traits will allow the remote sensing community to develop more focused and precise tools. Likewise, establishing state-of-the-art remote sensing methods as plant phenotyping tools will allow improved understanding and modeling of crops in dynamic environments.

Working on harmonization and implementing open data standards allow the use and re-use of the data for a broader community. Further, bridging scales and concepts offers unique and promising approaches to address major long-term challenges identified both on national and large-scale levels: food security in changing climate conditions, resilient agriculture countering land degradation and erosion, sustaining biodiversity and ecosystem functions, and agro-ecological transition. UAVs are one important common tool that is bridging the technical gap between the two research domains. A huge challenge hereby is the careful (not only short-term price-driven) selection of sensors and appropriate spectral domains to obtain a maximum of information. Along with extensive trait measurements, remote sensing and crop growth models can be advanced, increasing our understanding of plant performance in a dynamic environment. This would result in remote sensing techniques becoming more reliable, increasing their usefulness for practical applications in precision farming.

We anticipate that there is a need to further stimulate cooperation and we advocate initiating projects and network activities between the remote sensing and plant phenotyping communities. Ultimately, this exchange has the potential to stimulate growth in both communities and their interconnection may lead to new developments toward more sustainable agriculture. These interactions may substantially contribute to the European strategic research agenda and the relevant topics are contributing to prominent parts of the EU Green Deal ("From farm to fork" and "EU Biodiversity Strategy").

AUTHOR CONTRIBUTIONS

UR came up with the initial idea of bringing plant phenotyping and remote sensing scientists together. MM initiated the idea of a joint workshop and led the paper writing. MM, RP, KB, MS, SF, and UR jointly organized this workshop in February 2021, which marked the start of this article. HA completely revised the manuscript. All authors contributed to the writing and critical review of the article.

FUNDING

This article is based upon work from COST Action CA17134 Optical synergies for spatiotemporal SENsing of Scalable ECOphysiological traits (SENSECO), supported by COST. It was supported by European Plant Phenotyping Network (EPPN2020: Grant Agreement 731013), from EMPHASIS-PREP (Grant Agreement: 739514), from EOSC-Life (Grant Agreement: 824087). KB was funded within the EnMAP scientific preparation program (DLR Space and BMWi, Grant Number 50EE1923).

ACKNOWLEDGMENTS

We thank Lindsey Auguin for proofreading.

REFERENCES

Aasen, H., and Bareth, G. (2018). "Ground and UAV sensing approaches for spectral and 3D crop trait estimation," in: *Hyperspectral Remote Sensing of Vegetation*, *Vol. II,* eds P. Thenkabail, J. G. Lyon, A. Huete (Milton, ON: Chapman and Hall/CRC).

Aasen, H., and Bolten, A. (2018). Multi-temporal high-resolution imaging spectroscopy with hyperspectral 2D imagers - From theory to application. *Remote Sens. Environ*. 205, 374–389. doi: 10.1016/j.rse.2017.10.043

Aasen, H., Honkavaara, E., Lucieer, A., and Zarco-Tejada, P. (2018). Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows. *Remote Sens*. 10:1091. doi: 10.3390/rs10071091

Araus, J., and Cairns, J. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci*. 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Araus, J., Kefauver, S., Zaman-Allah, M., Olsen, M., and Cairns, J. (2018). Translating high-throughput phenotyping into genetic

gain. *Trends Plant Sci*. 23, 451–466. doi: 10.1016/j.tplants.2018.02.001

Berger, K., Rivera Caicedo, J., Martino, L., Wocher, M., Hank, T., and Verrelst, J. (2021). A survey of active learning for quantifying vegetation traits from terrestrial earth observation data. *Remote Sens*., 13:287. doi: 10.3390/rs13020287

Berger, K., Verrelst, J., Féret, J.-B., Wang, Z., Wocher, M., Strathmann, M., et al. (2020). Crop nitrogen monitoring: recent progress and principal developments in the context of imaging spectroscopy missions. *Remote Sens. Environ*. 242:111758. doi: 10.1016/j.rse.2020.111758

Chlingaryan, A., Sukkarieh, S., and Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. *Comput. Electron. Agric*. 151, 61–69. doi: 10.1016/j.compag.2018.05.012

Danilevicz, M., Bayer, P., Nestor, B., Bennamoun, M., and Edwards, D. (2021). Resources for image-based high-throughput phenotyping in crops and data sharing challenges. *Plant Physiol*. 163, 0–17. doi: 10.1093/plphys/kiab301

Dungey, H. S., Dash, J., Pont, D., Clinton, P., Watt, M., and Telfer, E. (2018). Phenotyping whole forests will help to track genetic performance. *Trends Plant Sci*., 23, 854–864, doi: 10.1016/j.tplants.2018.08.005

FAO. (2017). *The Future of Food and Agriculture Trends and Challenges*. FAO of the UN Rome. Availble online at: http://www.fao.org/3/i6583e/i6583e.pdf (accessed July 19, 2021).

Féret, J.-B., Berger, K., de Boissieu, F., and Malenovský, Z. (2021). PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents. *Remote Sens. Environ*. 252:112173. doi: 10.1016/j.rse.2020.112173

Fiorani, F., and Schurr, U. (2013). Future scenarios for plant phenotyping. *Ann. Rev. Plant Biol*. 64, 267–291. doi: 10.1146/annurev-arplant-050312-120137

Galieni, A., D'Ascenzo, N., Stagnari, F., Pagnani, G., Xie, Q., and Pisante, M. (2021). Past and future of plant stress detection: an overview from remote sensing to positron emission tomography. *Front. Plant Sci*. 11, 1–22. doi: 10.3389/fpls.2020.609155

Herrmann, I., and Berger, K. (2021). Remote and proximal assessment of plant traits. *Remote Sens*. 13:1893. doi: 10.3390/rs13101893

Johansen, K., Morton, M., McCabe, M., Malbeteau, Y. M., Aragon, B., Al-Mashharawi, S. K., et al. (2019). Unmanned aerial vehicle-based phenotyping using morphometric and spectral analysis can quantify responses of wild tomato plants to salinity stress. *Front. Plant Sci*. 10:370. doi: 10.3389/fpls.2019.00370

Lammerts van Bueren, E., and Struik, P. C. (2017). Diverse concepts of breeding for nitrogen use efficiency. A review. *Agron. Sustain. Dev*. 37:50. doi: 10.1007/s13593-017-0457-3

Maes, W., and Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci*. 24, 152–164. doi: 10.1016/j.tplants.2018.11.007

Manfreda, S., McCabe, M., Miller, P. E., Lucas, R., Madrigal, V. P., Mallinis, G., et al. (2018). On the use of unmanned aerial systems for environmental monitoring. *Remote Sens*. 10:641. doi: 10.3390/rs10040641

Papoutsoglou, E., Faria, D., Arend, D., Arnaud, E., Athanasiadis, I. N., Chaves, I., et al. (2020). Enabling reusability of plant phenomic datasets with MIAPPE 1.1. *New Phytol*. 227, 260–273. doi: 10.1111/nph.16544

Perich, G., Hund, A., Anderegg, J., Roth, L., Boer, M., Walter, A., et al. (2020). Assessment of multi-image unmanned aerial vehicle based high-throughput field phenotyping of canopy temperature. *Front. Plant Sci*. 11:150. doi: 10.3389/fpls.2020.00150

Quirós-Vargas, J., Bendig, J., Arthur, A. M., Burkart, A., Julitta, T., Maseyk, K., et al. (2020). Unmanned aerial systems (UAS)-based methods for solar induced chlorophyll fluorescence (SIF) retrieval with non-imaging spectrometers: state of the art. *Remote Sens*. 12:1624. doi: 10.3390/rs12101624

Rascher, U., Alonso, L., Burkart, A., Cilia, C., Cogliati, S., Colombo, R., et al. (2015). Sun-induced fluorescence - a new probe of photosynthesis: First maps from the imaging spectrometer HyPlant. *Global Change Biol*. 21, 4673–4684. doi: 10.1111/gcb.13017

Razzaq, A., Kaur, P., Akhter, N., Wani, S., and Saleem, F. (2021). Next-generation breeding strategies for climate-ready crops. *Front. Plant Sci*. 12:1374. doi: 10.3389/fpls.2021.620420

Reynolds, D., Baret, F., Welcker, C., Bostrom, A., Ball, J., Cellini, F., et al. (2019). What is cost-efficient phenotyping? Optimizing costs for different scenarios. *Plant Sci*. 282, 14–22. doi: 10.1016/j.plantsci.2018.06.015

Rosenqvist, E., Großkinsky, D., Ottosen, C.-O., and van de Zedde, R. (2019). The phenotyping dilemma-the challenges of a diversified phenotyping community. *Front. Plant Sci*. 10:163. doi: 10.3389/fpls.2019.00163

Roth, L., Aasen, H., Walter, A., and Liebisch, F. (2018). Extracting leaf area index using viewing geometry effects-A new perspective on high-resolution unmanned aerial system photography. *ISPRS J. Photogram. Remote Sens*. 141, 161–175. doi: 10.1016/j.isprsjprs.2018.04.012

Saiz-Rubio, V., and Rovira-Más, F. (2020). From smart farming towards agriculture 5.0: a review on crop data management. *Agronomy*. 10:207. doi: 10.3390/agronomy10020207

Semenov, M. A., and Shewry, P. R. (2011). Modelling predicts that heat stress, not drought, win increase vulnerability of wheat in Europe. *Sci. Rep*. 1:66. doi: 10.1038/srep00066

Tattaris, M., Reynolds, M., and Chapman, S. (2016). A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding. *Front. Plant Sci*. 7:1131. doi: 10.3389/fpls.2016.01131

Tmušić, G., Manfreda, S., Aasen, H., James, M. R., Gonçalves, G., Ben-Dor, E., et al. (2020). Current practices in UAS-based environmental monitoring. *Remote Sens*. 12:1001. doi: 10.3390/rs12061001

Van der Tol, C., Verhoef, W., Timmermans, J., Verhoef, A., and Su, Z. (2009). An integrated model of soil-canopy spectral radiances, photosynthesis, fluorescence, temperature and energy balance. *Biogeosciences*. 6, 3109–3129. doi: 10.5194/bg-6-3109-2009

Velumani, K., Lopez-Lozano, R., Madec, S., Guo, W., Gillet, J., Comar, A., et al. (2021). Estimates of maize plant density from UAV RGB images using faster-RCNN detection model: impact of the spatial resolution. *Plant Phenomics*. 2021:984843. doi: 10.34133/2021/9824843

Verrelst, J., Malenovský, Z., Van der Tol, C., Camps-Valls, G., Gastellu-Etchegorry, J.-P., Lewis, P., et al. (2019). Quantifying vegetation biophysical variables from imaging spectroscopy data: a review on retrieval methods. *Surv. Geophys*. 40, 589–629. doi: 10.1007/s10712-018-9478-y

Verrelst, J., Rivera-Caicedo, J. P., Reyes-Muñoz, P., Morata, M., Amin, E., Tagliabue, G., et al. (2021). Mapping landscape canopy nitrogen content from space using PRISMA data. *ISPRS J. Photogram. Remote Sens*. 178, 382–395. doi: 10.1016/j.isprsjprs.2021.06.017

Walter, A., Finger, R., Huber, R., and Buchmann, N. (2017). Opinion: smart farming is key to developing sustainable agriculture. *Proc. Natl. Acad. Sci*. 114, 6148–6150. doi: 10.1073/pnas.1707462114

Watt, M., Fiorani, F., Usadel, B., Rascher, U., Muller, O., and Schurr, U. (2020). Phenotyping: new windows into the plant for breeders. *Ann. Rev. Plant Biol*. 71, 689–712. doi: 10.1146/annurev-arplant-042916-041124

Weiss, M., Jacob,F., and Duveiller, G. (2020). Remote sensing for agricultural applications: a meta-review. *Remote Sens. Environ*. 236:11402. doi: 10.1016/j.rse.2019.111402

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Nat. Sci. Data*. 3:160018. doi: 10.1038/sdata.2016.18

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. *Front. Plant Sci*. 8:1111. doi: 10.3389/fpls.2017.01111

Zarco-Tejada, P. (2008). A new era in remote sensing of crops with unmanned robots. *SPIE Newsroom*. doi: 10.1117/2.1200812.1438

Zarco-Tejada, P., González-Dugo, V., and Berni, J. (2012). Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a

thermal camera. *Remote Sens. Environ.* 117, 322–337. doi: 10.1016/j.rse.2011.10.007

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 29 November 2021
doi: 10.3389/fpls.2021.740322

Estimating Leaf Area Index in Row Crops Using Wheel-Based and Airborne Discrete Return Light Detection and Ranging Data

Behrokh Nazeri[1], Melba M. Crawford[1,2] and Mitchell R. Tuinstra[2]*

[1] Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, United States, [2] Department of Agronomy, Purdue University, West Lafayette, IN, United States

OPEN ACCESS

Edited by:
Alessandro Matese, Institute for Bioeconomy, National Research Council (CNR), Italy

Reviewed by:
Ali Parsaeimehr, Delaware State University, United States
Ajoy Kumar Roy, Indian Council of Agricultural Research (ICAR), India

***Correspondence:**
Behrokh Nazeri
bnazeri@purdue.edu

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

Received: *02 August 2021*
Accepted: *02 November 2021*
Published: *29 November 2021*

Citation:
Nazeri B, Crawford MM and Tuinstra MR (2021) Estimating Leaf Area Index in Row Crops Using Wheel-Based and Airborne Discrete Return Light Detection and Ranging Data. Front. Plant Sci. 12:740322. doi: 10.3389/fpls.2021.740322

Leaf area index (LAI) is an important variable for characterizing plant canopy in crop models. It is traditionally defined as the total one-sided leaf area per unit ground area and is estimated by both direct and indirect methods. This paper explores the effectiveness of using light detection and ranging (LiDAR) data to estimate LAI for sorghum and maize with different treatments at multiple times during the growing season from both a wheeled vehicle and Unmanned Aerial Vehicles. Linear and nonlinear regression models are investigated for prediction utilizing statistical and plant structure-based features extracted from the LiDAR point cloud data with ground reference obtained from an in-field plant canopy analyzer (indirect method). Results based on the value of the coefficient of determination (R^2) and root mean squared error for predictive models ranged from ∼0.4 in the early season to ∼0.6 for sorghum and ∼0.5 to 0.80 for maize from 40 Days after Sowing to harvest.

Keywords: high-throughput phenotyping, remote sensing, LiDAR, leaf area index, machine learning, row crops

INTRODUCTION

Determination of Leaf Area Index (LAI) is essential for modeling the interaction between the atmosphere and the biosphere (Zhu et al., 2020). It is an important biophysical parameter that acts as a primary control for energy, water, and gas exchange within a vegetated ecosystem (Jensen et al., 2008; Zheng and Moskal, 2009). Estimation of LAI is also important for crop modeling (Lobell et al., 2015; Akinseye et al., 2017) and plant breeding (Blancon et al., 2019). Both direct and indirect approaches have been investigated to estimate LAI. Direct methods, which are based on measuring the area of the leaves directly, are accurate but costly, labor-intensive, and time-consuming. In destructive sampling, plants are defoliated within a specific area, and the one-sided leaf surface area is measured from imagery or with an electronic area meter (White et al., 2019) such as an LI-3100C. The average leaf biomass fraction and specific leaf weight, which is defined as leaf dry weight (the oven-dry mass), divided by the one-sided area of the fresh leaves are used to compute LAI, for each plot and sampling date (Yang et al., 2021).

Indirect optical methods estimate LAI from the canopy gap fraction that is defined as the effective LAI (LAI_{eff}). The relationship between LAI_{eff} and true LAI derived from a direct method, which assumes that the leaves are randomly distributed within the canopy, is shown in Eq. 1 (Chen et al., 2005; Ryu et al., 2010).

$$LAI_{eff}(\theta) = \Omega(\theta) \times LAI \quad (1)$$

where $\Omega(\theta)$ is the canopy clumping index that describes the non-randomness of the leaf foliage distribution; it can be estimated through the nonrandom distribution of gap fractions using the logarithmic gap fraction averaging method, and θ is the solar zenith angle (Fang et al., 2019).

Digital cover photography, digital hemispherical photography, and the LAI-2200C plant canopy analyzer are all used to obtain indirect optically-based estimates of LAI (Fournier and Hall, 2017; Fang et al., 2019). Direct measurement methods and some optical methods are also used as references for indirect measurement techniques (Richardson et al., 2009). Indirect methods have been developed for determining LAI over large areas using both active and passive remote sensing. Within the last decade, light detection and ranging (LiDAR) has been used for mapping, modeling, and spatial analysis in many applications, including estimation of LAI. The advantage of LiDAR compared to other remote sensing technologies is that it directly provides three-dimensional coordinates. Promising results have been obtained from LiDAR (Jimenez-Berni et al., 2018) and in combination with hyperspectral imagery (Masjedi et al., 2018, 2019) in modeling biophysical characteristics, including vegetation height and above-ground biomass for agriculture applications (Nie et al., 2016; ten Harkel et al., 2020). LiDAR has also been used to model forest canopy structure (Lefsky et al., 2002) and to estimate LAI in forests (Zhao and Popescu, 2009; Korhonen et al., 2011; Jung and Crawford, 2012; Alonzo et al., 2015).

To estimate LAI from LiDAR, empirical models are developed to represent the relationship between the ground reference LAI and LiDAR-derived metrics. Two types of LiDAR metrics are commonly used in LAI prediction, the Beer-Lambert law based on the laser penetration index (LPI; Richardson et al., 2009) and allometric measurements that are statistically-based features (Pope and Treitz, 2013). Allometric-related features include the mean height and standard deviation, maximum height of all returns, and the coefficient of variation of height. Features based on the Beer-Lambert law include gap fraction and LPI (Nie et al., 2016). Pope and Treitz (2013) demonstrated the combined use of airborne discrete return LiDAR data and WorldView-2 high-resolution imagery to predict LAI in a boreal mixed wood forest. Digital hemispherical photos were used as a ground reference, and statistically significant LiDAR-based inputs for a stepwise linear regression model included the ratio of the first return and total return, the vertical distribution ratio, crown closure, and a vertical complexity index (VCI) that represents structural homogeneity with height (Ludwig et al., 1988; van Ewijk et al., 2011; Pope and Treitz, 2013).

Few studies have focused on estimating LAI for row crops, such as maize, e.g., Nie et al. (2016) and sorghum, e.g., Lang (1986). In addition, in most remote sensing focused studies, discrete return LiDAR data are acquired by manned aircraft and Unmanned Aerial Vehicles (UAVs), which have lower point density and laser penetration than ground-based platforms. Ground-based LiDAR data can acquire data at a very high spatial resolution over shorter crops compared to airborne platforms, and depending on the plant structure, can potentially penetrate deeper into the canopy. Further, these platforms are not subject to localized changes in position, elevation, and look angle that are common with airborne platforms, but are restricted to operation in field conditions during which they can drive and collect data.

Nazeri (2021) investigated the destructive sampling method as a ground reference in estimation of LAI from LiDAR acquired by a UAV over a sorghum field experiment. Three sets of ground reference data collected by the Purdue team in 2019 to parameterize a crop growth model were provided as ground reference data. The relationship between the LiDAR data and LAI computed using destructively sampled ground reference data was weak. The results were not unexpected, as the LiDAR data are physically more closely related to the gap fraction than the assumptions for LAI calculations based on destructive sampling (Hammer et al., 2010; Fang et al., 2019; Yang et al., 2021). The low R^2 of models obtained using the destructive sampling ground reference, coupled with the practical limitations for performing extensive destructive sampling through the season motivated this study of an indirect ground reference method coupled with extensive data acquisitions during the 2020 growing season.

This paper is an exploratory study of LAI prediction using LiDAR point cloud data acquired by a converted high-clearance tractor/sprayer with a custom sensor boom and by low altitude UAVs over sorghum and maize plant breeding experiments. LiDAR platforms and systems with different laser units were evaluated at multiple altitudes for obtaining LAI. Remote sensing acquisitions were matched to the field-based LAI measurements using near-coincident data acquisitions. Multiple strategies for feature extraction were investigated for developing regression-based predictive models, including stepwise multiple linear regression (SMLR), partial least squares regression (PLSR), and support vector regression (SVR). The predictive models were developed based on the indirect ground reference method and evaluated based on the resulting R^2 values and the root mean squared error of the residuals. Contributions of the study include investigation of multiple LiDAR-based features for multitemporal prediction of LAI via regression models and evaluation of the capability of LiDAR sensors and platforms for acquiring data to predict sorghum and maize LAI at multiple times during the growing season.

MATERIALS AND METHODS

Study Area and Experiment Setting

The experiments for this study were conducted at the Agronomy Center for Research and Education at Purdue University, West Lafayette, IN, United States, to evaluate the potential of sorghum varieties for biomass production. Both ground reference and LiDAR data were acquired during the 2020 growing season. In

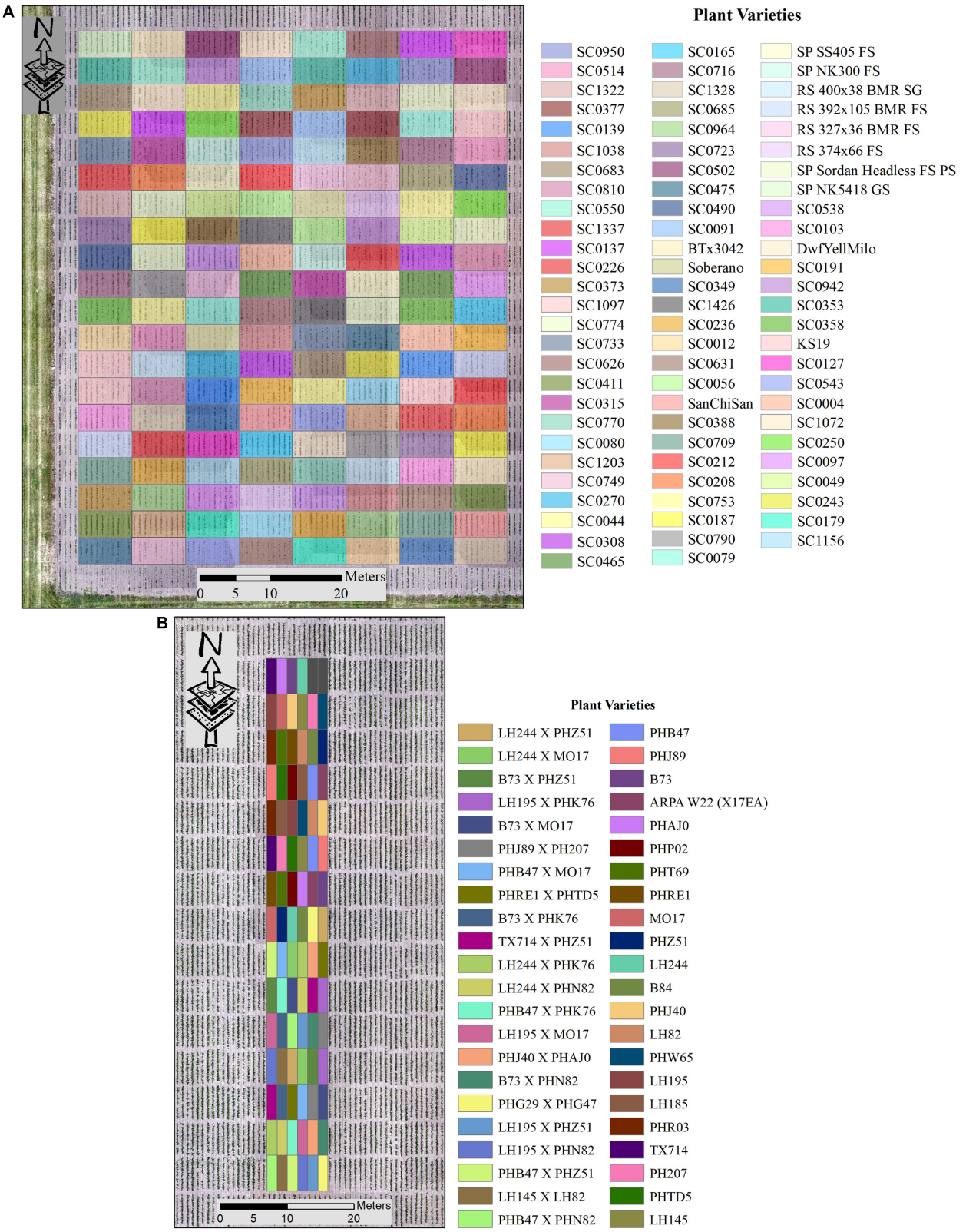

FIGURE 1 | Plot variety layout **(A)** SbDivTc_Cal and **(B)** HIPS.

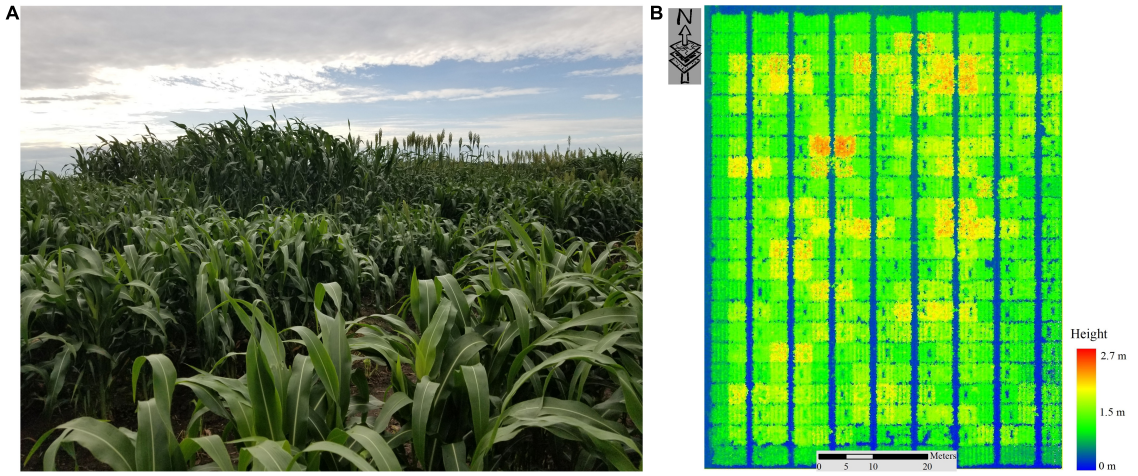

FIGURE 2 | (A) Photograph of the SbDivTc_Cal panel (7/20/2020), **(B)** LiDAR-Based Height Map of SbDivTc_Cal Sorghum Panel (7/20/2020).

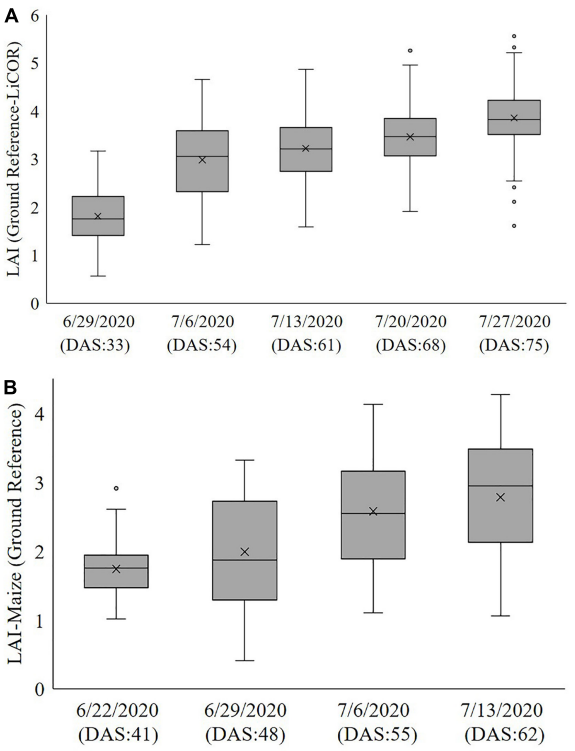

FIGURE 3 | Measured ground reference LAI of **(A)** SbDivTc_Cal and **(B)** HIPS using LAI–2200C (2020).

this study, near concurrent ground-based and UAV LiDAR data were analyzed. The LiDAR data were collected from the Sorghum Biodiversity Test Cross Calibration Panel (SbDivTc_Cal) and the maize High-Intensity Phenotyping Sites (HIPS) Panel. The SbDivTc_Cal experimental design included 80 varieties two replicates in a randomized block design planted in 160 plots (plot size: 7.6 m × 3.8 m), ten rows per plot (row number is counted from the west to east). All 160 plots were included in the analysis for the SbDivTc_Cal data, as LAI ground reference data were acquired for all the plots in the experiment. The HIPS maize experiment contained 44 varieties of maize with two replicas, including hybrids and inbreds. This experiment had 88 plots (plot size: 1.5 m × 5.3 m), two rows per plot. In the early stages, sorghum and maize have very similar plant structures, although sorghum is planted at a higher density (∼200,000 plants/hectare) compared to maize (∼75,000 plants/hectare). During the growing season, the geometric structure of sorghum becomes more complex as tillers develop, decreasing canopy penetration. **Figure 1** shows the layout of the SbDivTc_Cal and HIPS plots based on the respective genotypes.

Differences between varieties can be seen clearly in terms of physical characteristics shown in a photo (**Figure 2A**), and in the LiDAR-based height map acquired by a UAV on 7/20/2020, 68 days after sowing (DAS; **Figure 2B**).

Field Ground Reference Data

In 2020, reference data were collected weekly from June 29 to July 27 for sorghum and from June 22 to July 13 using a handheld plant canopy analyzer (LAI-2200C). The LAI-2200C is a portable instrument for acquiring an indirect measurement of LAI_{eff} based on canopy gap fraction analysis (Welles and Cohen, 1996; Sonnentag et al., 2007; Černý et al., 2019). In sorghum, to avoid the impact of adjacent plots and destructive sampling, LiDAR data from Rows 2 and 3 of each plot were associated with each reference value for developing the predictive models. Two sets of five measurements one measurement above the canopy and four measurements below the canopy near the ground between rows 2 and 3 in the direction of the rows (north-south) were made according to the recommended protocol, then a representative value per plot was calculated using the Field Viewer 2200 (FV2200) software. These values were used as the primary reference data for developing predictive models of LAI based on the LiDAR remote sensing data. The ground reference values ranged from 0.5 to 6 for sorghum and 0.5 to 5 for maize,

increasing during the period of the growing season until sampling was stopped after flowering. The box plots in **Figure 3** show the range of values of ground reference data for both crops within ±1.96 standard deviations for the LAI–2200C based on the date of data collection and corresponding DAS. The values of LAI exceeding 95% were from photoperiod sensitive varieties, whose characteristics increasingly differ from the rest of the experiment as the season progresses. The sequence of 2020 plant canopy analyzer data was used as a ground reference for evaluating the LiDAR-based metrics. Remotely sensed LiDAR data and ground reference acquisitions were separated by no more than 3 days. **Table 1** summarizes the experiment over the SbDivTc_Cal and HIPS 2020 experiments.

Light Detection and Ranging Point Cloud Data Acquisitions

Platforms and Sensors

Remote sensing data were collected by the UAV weekly, first prior to planting to develop the baseline terrain model and at intervals of 1–2 weeks thereafter, depending on the weather, throughout the growing season. Two M600P UAVs were flown over the study area at altitudes of 20 and 40 m and speeds of 3–5 m/s. The UAVs were equipped with a Velodyne VLP-Puck LITE and a Velodyne VLP-32C, respectively. The Velodyne VLP-Puck LITE has 16 channels that are aligned vertically from −15° to +15°, resulting in a total vertical field of view (FOV) of 30°. The point capture rate in single return mode is ∼300,000 points per second. The range accuracy is typically ±3 cm, with a maximum measurement range of 100 m (Velodyne VLP-Puck Lite, 2020). The Velodyne VLP-32C has 32 channels that are aligned vertically from −15° to +25°, in a total vertical FOV of 40°. The point capture rate in a single return mode is ∼600,000 points per second. The range accuracy is typically ±3 cm, with a maximum measurement range of 200 m (Velodyne VLP-32C, 2020). The UAVs were equipped with an integrated global navigation satellite system/inertial navigation system (GNSS/INS) Trimble APX-15v3 for direct georeferencing (Hasheminasab et al., 2020). LiDAR data were acquired by a wheel-based system, a LeeAgra Avenger agricultural high-clearance tractor/sprayer with a custom boom and mounted sensors, referred to in this study as the PhenoRover, on an experimental basis. The boom is constructed from 2.75 m wide T-slot structural aluminum, and the top of the boom can be raised to a maximum of 5.5 m height from the ground. Sensors mounted on the boom include a Headwall hyperspectral VNIR machine vision camera, two FLIR RGB cameras, and a Velodyne VLP-Puck Hi-Res LiDAR, as well as the GNSS/INS navigation system. The VLP-Puck Hi-Res has similar sensor specifications to the VLP-Puck LITE. Its FOV is −10° to +10° (Velodyne VLP-Puck Hi-Res, 2020). The platform speed in the field was 1.5 miles per hour. **Figure 4** shows the PhenoRover and UAV platforms for the 2020 data collection. PhenoRover data were acquired limited times in 2020, subject to field conditions. **Table 2** details the platforms and their mounted sensor specifications for the 2020 data collection.

Table 3 summarizes the LiDAR data collection and the corresponding ground reference measurements in terms of DAS relative to the data collection dates and ground reference measurements.

PhenoRover and Unmanned Aerial Vehicle Light Detection and Ranging Data

The average point densities of the LiDAR data acquired by the sensors on the UAVs depend on the type of sensor, the platform flying height, FOV, and mission characteristics such as the sidelap

TABLE 1 | Experimental design for the 2020 growing season.

Experiment	Genotype	# of plots	# of varieties	Sowing date	Harvest date
HIPS	Hybrid/inbred	88	44	May 12	October 1
SbDivTc_Cal	Hybrid	160	80	May13	August 15

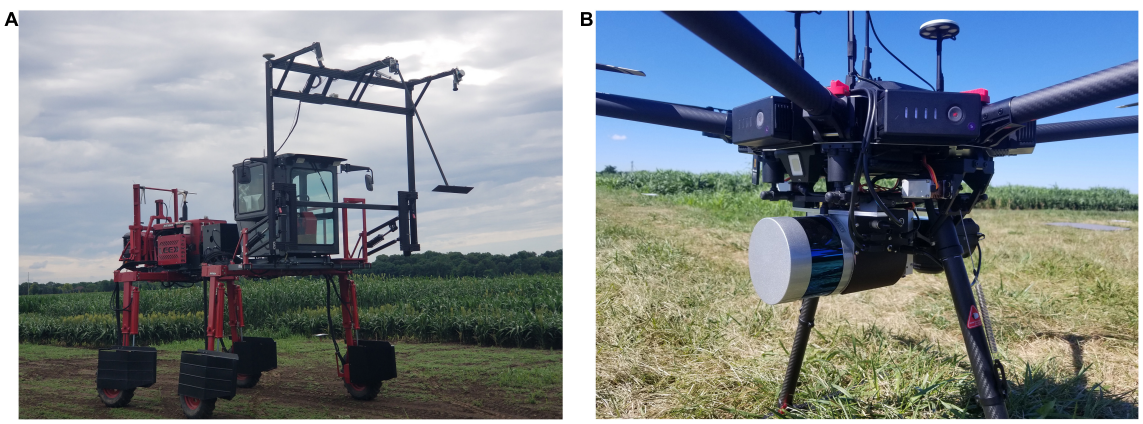

FIGURE 4 | (A) PhenoRover platform with RGB/LiDAR/Hyperspectral/GNSS/INS sensors, **(B)** UAV-2 with RGB/LiDAR/GNSS/INS sensors in 2020.

TABLE 2 | Platforms and mounted sensors specification in 2020.

Platform	Sensor	Unit	Description
UAV-1			
	RGB camera	1	36.4 MP Sony Alpha 7R (ILCE-7R)
	LiDAR sensor	1	Velodyne VLP 16-Puck LITE-range accuracy of ±3 cm
	GNSS/INS	1	Trimble APX-15 v2
	Hyperspectral Camera	1	Nano Hyperspectral (VINIR)
UAV-2			
	RGB camera	1	36.4 MP Sony Alpha 7R (ILCE-7R)
	LiDAR sensor	1	Velodyne VLP 32-range accuracy of ±3 cm
	GNSS/INS	1	Trimble APX-15 v2
PhenoRover			
	RGB camera	2	9.1 MP FLIR Grasshopper3 GigE
	Hyperspectral camera	1	Headwall Machine Vision 270 band line-scanning with 4.8 mm lens
	LiDAR sensors	1	Velodyne VLP-Puck Hi-Res
	GNSS/INS	1	Applanix POS-LV 125

TABLE 3 | Days after sowing (DAS) relative to the available ground reference and LiDAR data in two experiments over SbDivTc_Cal and HIPS.

Experiment	Platform	Flying height	Sowing date	LiDAR data collection date	DAS[1]	Ground reference date	DAS[2]
HIPS	UAV-1	N/A	05/12	06/25	44	06/22	41
	PhenoRover	N/A		06/26	45	06/29	48
	UAV-2	20 m		07/07	56	07/06	55
	UAV-1	20 m		07/11	60	07/13	62
	UAV-2	20 m		07/11	60	07/13	62
	UAV-2	20 m		07/13	62	07/13	62
	PhenoRover	N/A		07/13	62	07/13	62
SbDivTc_Cal	PhenoRover	N/A	05/13	06/26	44	06/29	47
	UAV-1	40 m		07/02	50	06/29	47
	UAV-2	20 m		07/07	55	07/06	54
	UAV-2	20 m		07/13	61	07/13	61
	UAV-1	40 m		07/17	65	07/20	68
	PhenoRover	N/A		07/20	68	07/20	68
	UAV-1	40 m		07/20	68	07/20	68
	UAV-2	20 m		07/20	68	07/20	68
	UAV-1	40 m		07/28	76	07/27	75
	UAV-2	20 m		07/28	76	07/27	75

DAS[1]: DAS with respect to data collection data; DAS[2]: DAS with respect to ground reference data.

TABLE 4 | Point density of sample data on 7/20/2020.

Platform	Flying height	DAS	Point density (Points/ m^2)
UAV-1	40 m	68	70
UAV-2	20 m	68	500
PhenoRover	N/A	68	1,400

of the flightlines. In this study, point density is investigated based on flying height and sensor type, and it is presumed that the rest of the characteristics affecting point density are consistent across the data acquisitions; these values are significantly lower than the LiDAR point density from the PhenoRover because the sensor on the PhenoRover operates at a much lower height (approximately 5 m from the ground). **Table 4** shows the point density of the sensors based on flying height. **Figure 5** illustrates the resulting 3D point cloud from the UAV platforms and PhenoRover over a sorghum sample row. As expected, the canopy penetration achieved by the UAV sensors was lower than the PhenoRover due to the higher platform altitude. UAV-2 with a Velodyne VLP-32C had a higher point density, resulting in greater canopy penetration compared to UAV-1 with a Velodyne VLP-Puck LITE, due to the combined impact of being flown at 20 m and the higher pulse rate of the sensor with more laser beams.

METHODOLOGY

Feature Extraction From Light Detection and Ranging Data

In the HIPS experiment, LiDAR features were extracted at plot level as there were two rows in a plot (**Figure 6A**), while in the SbDivTc_Cal experiment, LiDAR features were extracted at

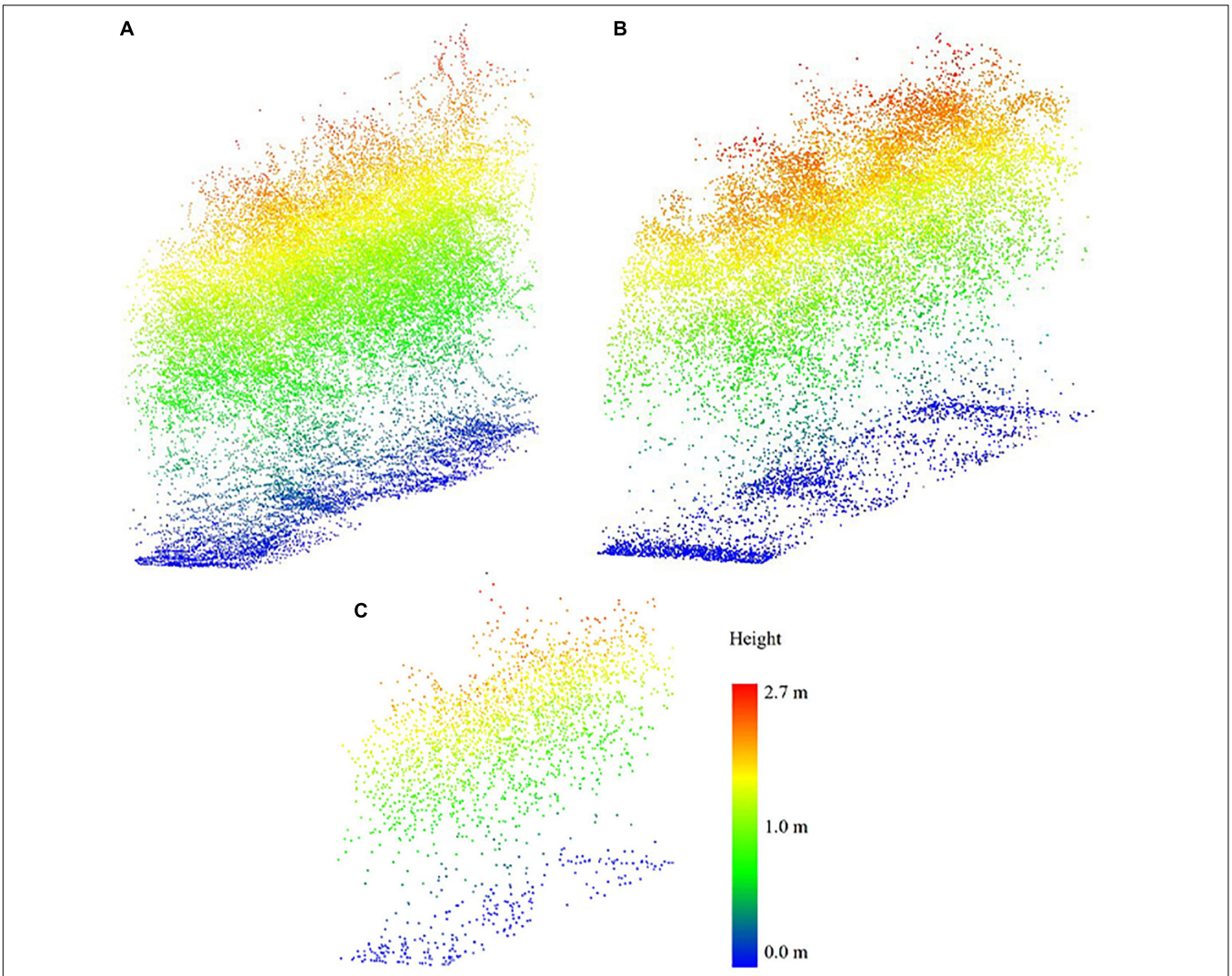

FIGURE 5 | Example sensor point cloud sample data from (7/20/2020) from **(A)** PhenoRover, **(B)** UAV-2, and **(C)** UAV-1.

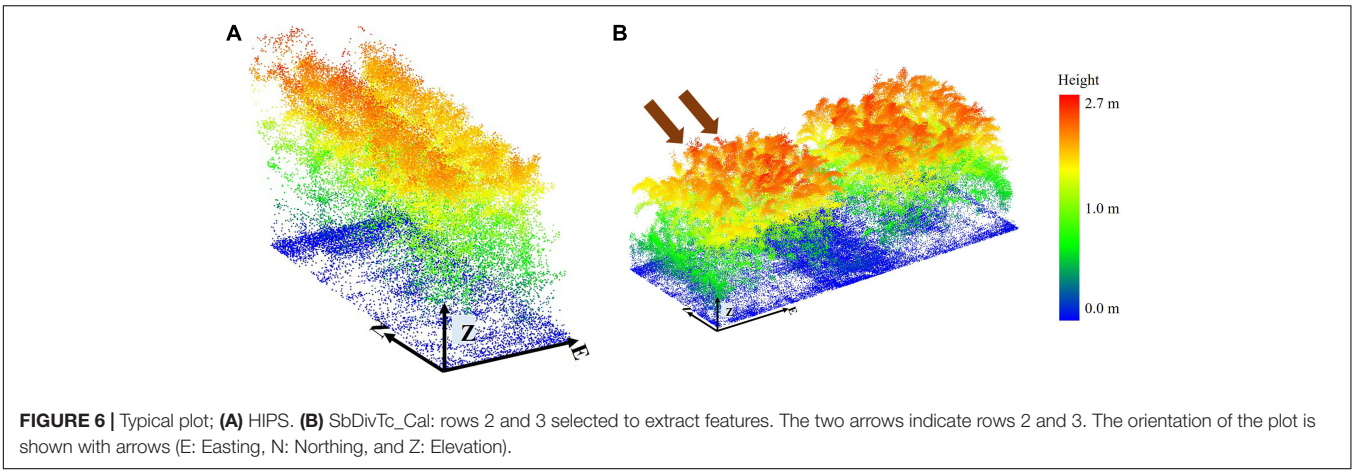

FIGURE 6 | Typical plot; **(A)** HIPS. **(B)** SbDivTc_Cal: rows 2 and 3 selected to extract features. The two arrows indicate rows 2 and 3. The orientation of the plot is shown with arrows (E: Easting, N: Northing, and Z: Elevation).

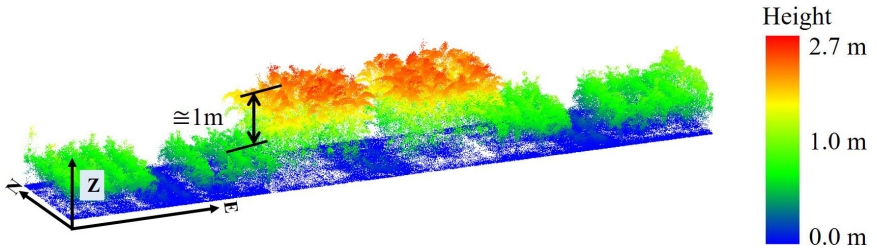

FIGURE 7 | Height of photoperiod sensitive variety SP SS405 FS relative to the surrounding plots 7/28/2020.

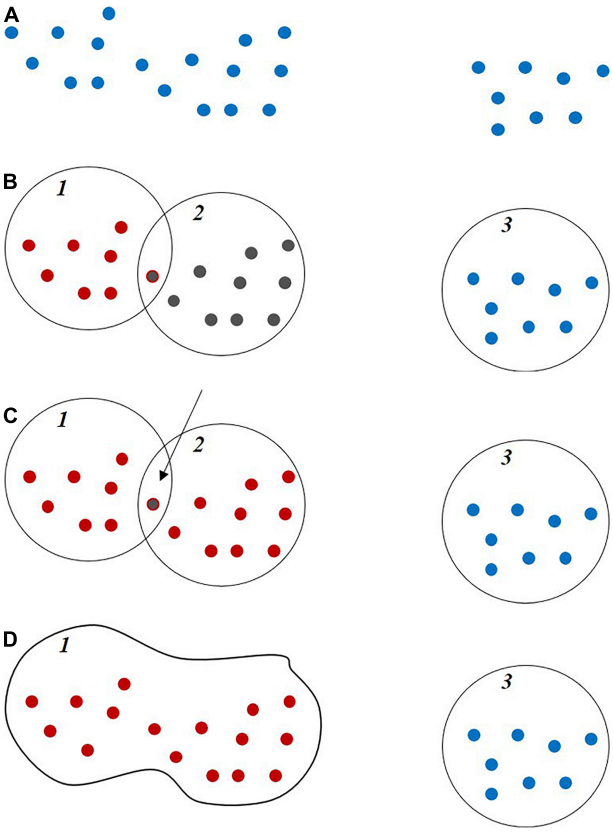

FIGURE 8 | Thematic region growing clustering steps: **(A)** Sample points. **(B)** Initial clustering. **(C)** Finding common points in two close clusters. **(D)** Connecting and joining two clusters.

the row-level within ten-row plots. Rows four, seven, and eight were adjacent to rows that were destructively sampled. Rows one and ten were "border" rows, so they were not necessarily representative of conditions within the plot, particularly for light accessibility when plots with tall varieties were adjacent to plots with short varieties. Rows 2 and 3 were extracted from the remotely sensed data and analyzed for this study. Features were extracted from rows 2 and 3 as a spatially contiguous two-row block (essentially equivalent to a two-row plot) where the ground reference was collected. **Figure 6B** shows a typical plot of the dataset, where rows 5 and 6 were destructively sampled via machine harvesting, and manual destructive harvesting was performed in row 9.

Three varieties of sorgum experiment (ATx623xDwfYellMilo, ATx623xSC0044, and SP SS405 FS) are photoperiod sensitive, as noted previously, and have a different plant structure than the rest of the varieties, especially later in the growing season. For example, "SP SS405 FS" was taller than the surrounding plots by approximately 1.3 m on 7/28/2020 (**Figure 7**). The impact of these varieties on the predictive models was investigated.

As noted in the Introduction, most LiDAR-based features proposed in the literature are based on the height or moments of the histograms of point cloud values in a 3D volume classified as vegetation. The Digital Terrain Model (DTM) required to determine plant heights was derived from a bare earth field using UAV-based LiDAR point cloud data before planting and assumed to be constant throughout the growing season. The height of points was estimated by subtracting the DTM from the "z" coordinate of each point in the dataset. Points with a height of less than 10 cm were considered as ground points and not included in the statistical analysis of the vegetation. The following physically-based features were explored for this study.

Laser penetration index is defined as the fraction of laser points that penetrate the canopy. The index can be calculated in many ways. In this study, it is computed as the ratio between the number of ground points (N_{Ground}) and the total number of points in a given area ($N_{\text{Ground}} + N_{\text{vegetation}}$), which is assumed here to be a row of a plot (Eq. 2). The number of non-ground points is assumed to be equal to the number of points identified as vegetation ($N_{\text{vegetation}}$):

$$\text{LPI}: \frac{N_{\text{Ground}}}{N_{\text{Ground}} + N_{\text{vegetation}}} \tag{2}$$

Features commonly used for allometric relationships include various statistically-based height features extracted from the non-ground point cloud, including plant height at various percent quantiles, mean height, standard deviation of the point cloud height, coefficient of variation of height, skewness of height, and Vegetation Complexity Index (VCI) described in Eq. 3 (van Ewijk et al., 2011).

$$\text{VCI} = \frac{(-\sum_{i=1}^{\text{HB}} [p_i \times \ln(p_i)])}{\ln(\text{HB})} \tag{3}$$

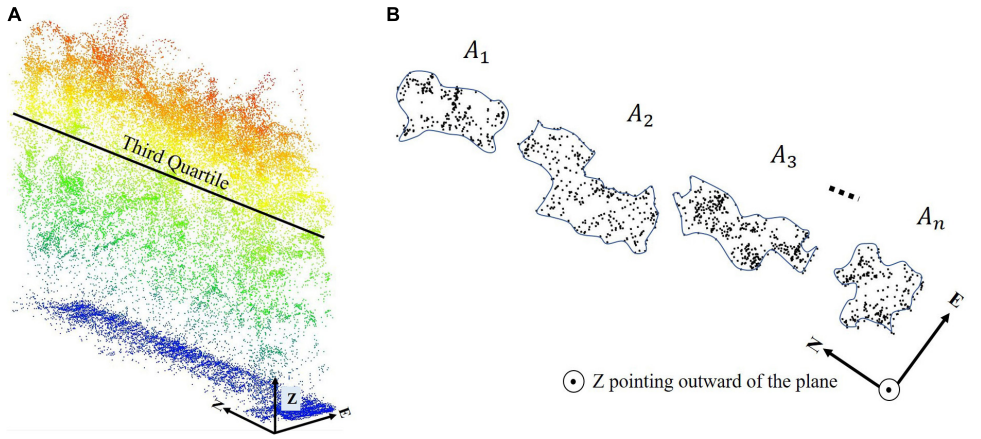

FIGURE 9 | Example of Clusters Area Plane (CAP) feature; **(A)** A typical third quartile of a row and **(B)** cross section at the third quartile.

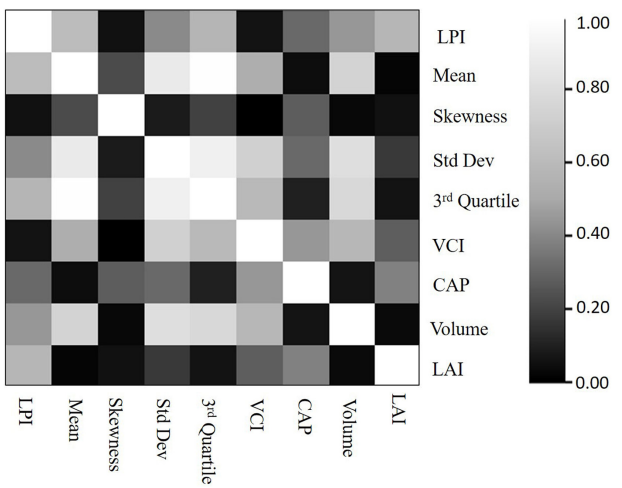

FIGURE 10 | Sensitivity analysis: Feature correlation matrix (zero is the lowest correlation and 1 is the highest correlation).

where HB = total number of height bins, p_i = Proportional abundance ($\frac{\#\text{ of returns}}{\text{Total }\#\text{ of returns}}$) in a height bin ($i$).

A new feature, referred to as the Clusters' Area Plane (CAP), which is based on horizontal characteristics of the point cloud at a given height in a row, was proposed and evaluated in the study. To obtain the CAP feature, a plane is intersected with the point cloud within a row at a given height quartile with ±4 cm thickness of this plate, and the associated points are extracted. The points are clustered using a region-growing approach based on the distance between points and the k-nearest neighbors as follows: the points are represented using a KD tree data structure, and the k-nearest neighbors to each point are determined within a defined radius and assigned to the respective clusters. Then, the clusters with common points are joined, and the cluster number is updated iteratively until no further changes occur in the clusters (**Figure 8**).

Finally, the area of clusters that is larger than a user-defined threshold is calculated, and the total area is defined as the CAP feature (Eq. 4).

$$\text{CAP} = \sum_{i=1}^{n} A_i \tag{4}$$

While the feature does not have a direct physical interpretation, it contains information for predicting LAI based on the horizontal distribution of the plants within the canopy at a given quartile (75% with ±4 cm thickness in this study). The CAP feature was also calculated in other quartiles, e.g., 50% and 25%, but only the 75% quartile provided statistically significant results for the data in these experiments. The 50 and 25% quartiles did not have an adequate number of samples to evaluate the index, both due to penetration of the canopy and its geometric structure. **Figures 9A,B** show a typical example of the CAP feature.

Correlation between features and LAI indicated that LPI has the highest correlation with LAI, and the CAP feature has the second-highest correlation with LAI. The correlation matrix in **Figure 10** also indicates that there is significant correlation between many of the candidate features. For example, the value of the correlation between the standard deviation of height and the mean and third quartile height is greater than 0.9.

Regression-Based Predictive Models

Predictive models were developed using SMLR (Johnsson, 1992), PLSR (Rosipal and Krämer, 2005), and SVR (Feng and Li, 2014). SVR models were investigated with four kernels (linear, polynomial, RBF, and sigmoid), and their hyperparameters were obtained via grid search. Eight features were considered as input variables, including LPI, Height_mean, standard deviation, and skewness, height (3rd Quartile), VCI, Volume of the vegetation in a row based on the convex hull of the points, and CAP. In this study, the training and test data were selected randomly by 75% training and 25% test. Both replicates of each genotype variety were randomly assigned to either training or test. Ten-fold cross-validation was performed on the training set. The

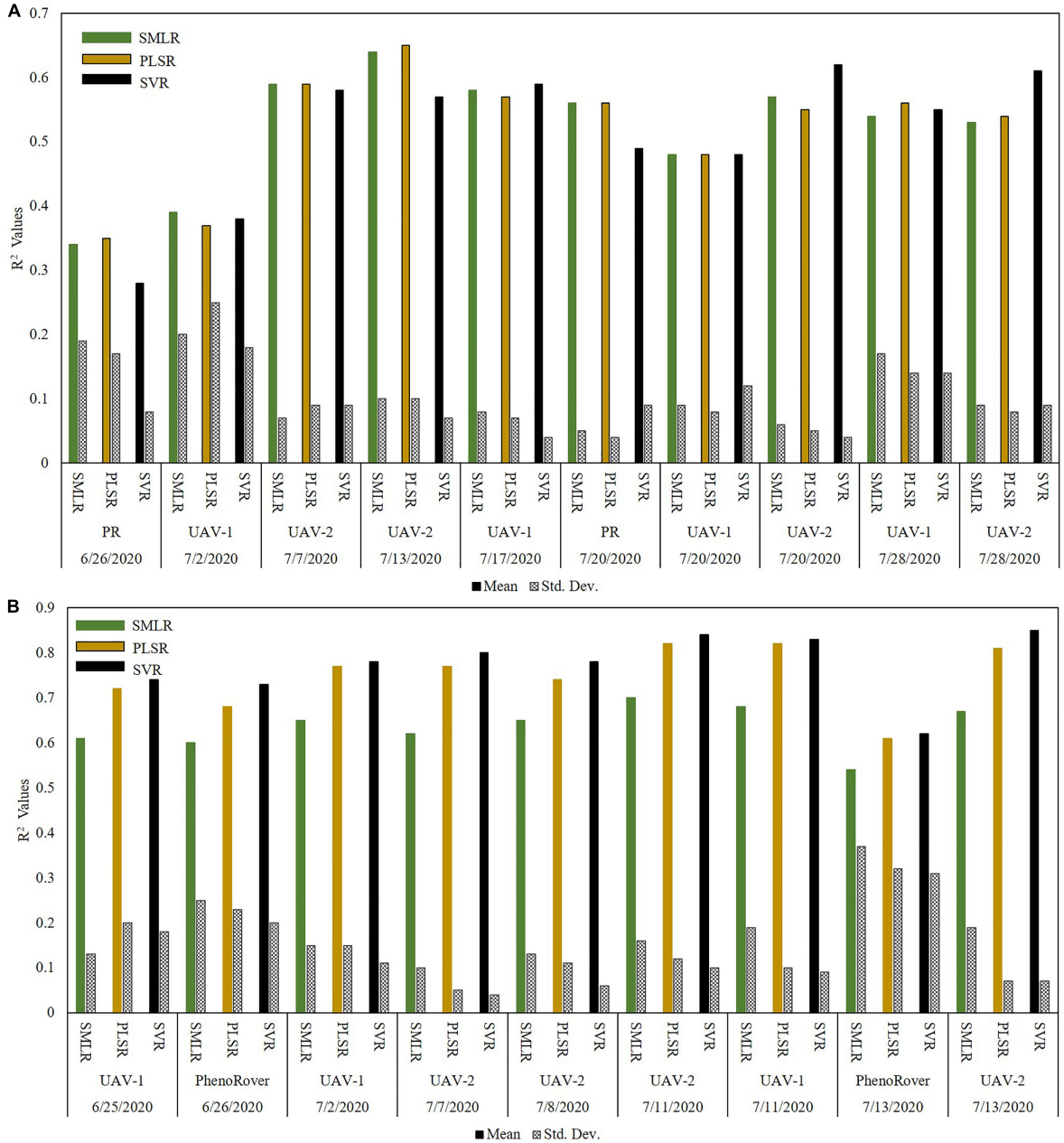

FIGURE 11 | R^2 values for 2020 regression models for LAI estimation **(A)** SbDivTc_Cal and **(B)** HIPS; (PR: PhenoRover).

values of R^2 for the respective models are reported in the results section.

LEAF AREA INDEX PREDICTIVE MODEL RESULTS

The results of the LAI predictive models are included based on the date and the platform. SMLR, PLSR, and SVR with RBF kernel models developed for the 2020 sorghum and maize data are illustrated via bar charts. **Figure 11A** shows the results for sorghum datasets. The models had low R^2 statistics for the first two dates acquired by PhenoRover and UAV-1 (0.28 and 0.38 for the SVR model). The primary reason was the small size of the plants (∼35 and ∼50 cm) for 6/26/2020 and 7/02/2020, respectively. The measurements from the LAI–2200C acquired between the rows were also not representative of the true canopy gap fraction at this height. The values of R^2 for the rest of the dates were consistent throughout the season, even as the plant heights increased rapidly until flowering. **Figure 11B** shows the results for maize datasets. The values of R^2 for all dates were consistent throughout the season and varied from 0.5 to 0.8. The

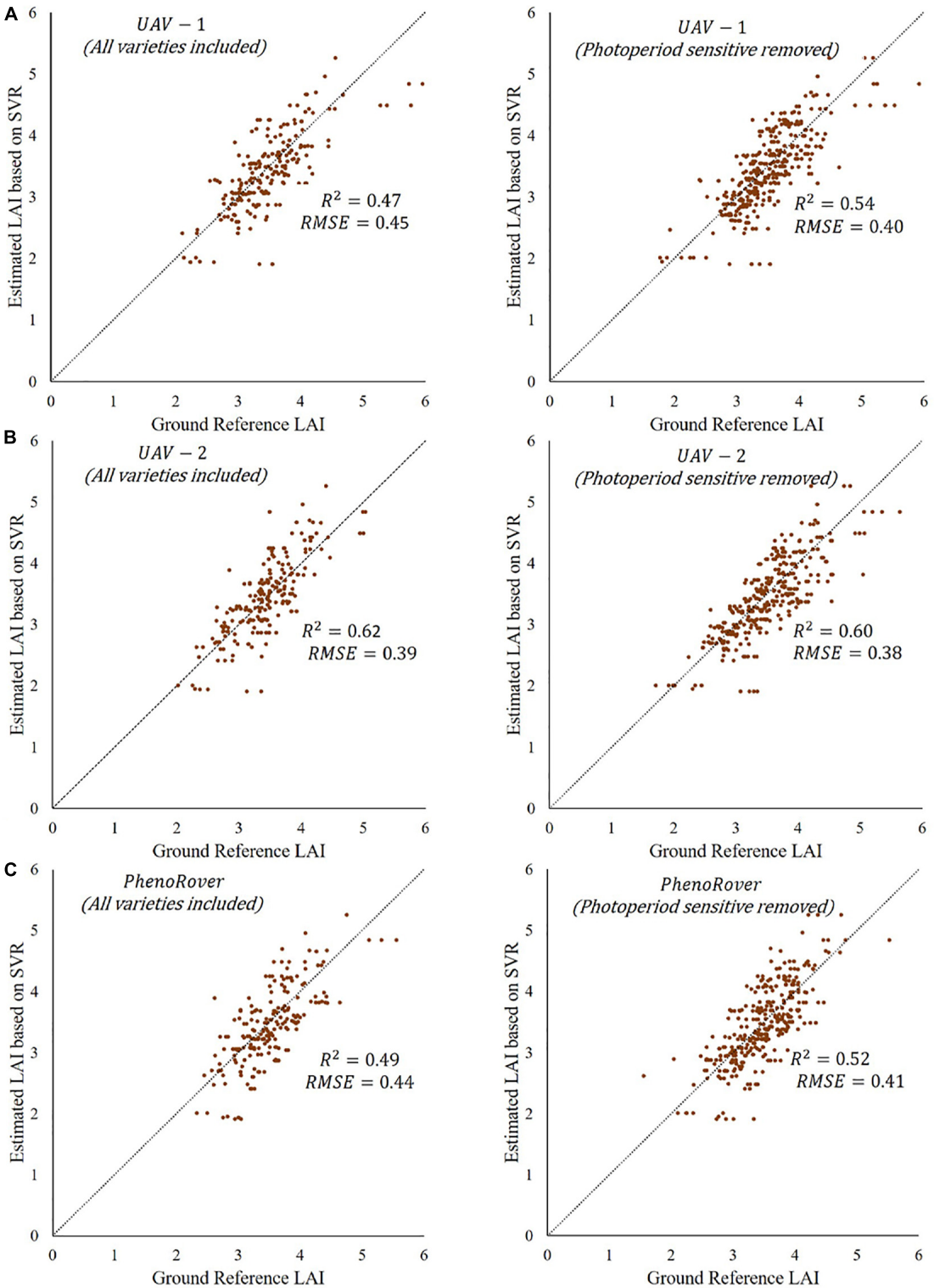

FIGURE 12 | Predictions based on SVR RBF models showing R^2 values and RMSE at midseason (7/20/2020) before and after removing photosensitive varieties for three platforms: **(A)** UAV-1, **(B)** UAV-2, and **(C)** PhenoRover.

results of maize show that the range of R^2 in maize is consistant with sorghum, but generally higher. This is attributed to the maize experiment being planted less dense than sorghum (maize: ∼75,000 plants/hectare vs. sorghum: ∼200,000 plants/hectare) and the lower complexity of the plant structure resulting in greater laser penetration into the canopy later in the season. The p-value from t-test statistics (0.94) showed that differences in the mean of R^2 values from pairwise comparisons of the three regression models were not statistically significant at an α of 0.05. The results also did not indicate significant differences between the mean of R^2 values from pairwise comparisons of combinations of UAV-1 (VLP 16, flown at 40 m), UAV-2 (VLP 32 flown at 20 m), and the PhenoRover.

The three photoperiod sensitive varieties were removed from the sorghum dataset, and R^2 values of all models were calculated. The p-value from t-test statistics (0.57) indicated no significant difference between the mean of R^2 obtained using data prior to and after removing photoperiod sensitive varieties. For example, the plots of one-to-one comparisons of reference vs. the predicted values of SVR model from the UAVs and PhenoRover on 7/20/2020 before and after removing the photoperiod sensitive varieties from the datasets are provided in **Figure 12**. The plots show the model of UAV-1 (**Figure 12A**) and PhenoRover (**Figure 12C**) slightly improved in terms of R^2, but UAV-2 results (**Figure 12B**) were essentially unchanged.

To evaluate the importance of the features, a leave-one-out procedure was used with the SVR-RBF model, which had the highest R^2 value, and the resulting R^2 (R^2_{new}) was calculated (Eq. 5),

$$\text{Weight of feature} = 1 - \frac{R^2_{new}}{R^2_{original}} \quad (5)$$

where R^2_{new} is an R^2 of the model fit without the feature, and $R^2_{original}$ is the R^2 of the model with all features.

Figure 13 shows the feature importance in the models developed for the three platforms on July 20, 2020.

Laser penetration index is the most highly ranked feature, based on the correlation with the plant canopy analyzer data, and the 2nd and 3rd ranked features are CAP and VCI, both of which are also indicative of penetration of the canopy. Additionally, the CAP feature is related to the horizontal distribution of the canopy, as noted previously. The height-related features are correlated and individually have a lower impact on the model, while LPI and CAP represent physically different characteristics. In complex vegetation such as sorghum, which is planted at high density and has tillers, many laser points are concentrated in the upper canopy, and few laser points penetrate deeper in the canopy.

Although the sensor on the PhenoRover was much closer to the canopy, typically between 2 and 5 m depending on the date, and the speed of the PhenoRover was much slower, resulting in increased point density and penetration of the canopy, R^2 values of the models (**Figure 11**) based on data from PhenoRover, UAV-1 (flying height 40 m) and UAV-2 (flying height 20 m) were similar for comparable dates. In most cases, multiple stepwise linear regression models had the lowest R^2 value, and only LPI and VCI features were significant at α = 0.05. As mentioned earlier, the R^2 value for the SMLR, PLSR, and SVR (with an RBF kernel) models are generally similar, and the sample mean of the R^2 values over the season are not statistically different by pairwise comparison in both sorghum and maize.

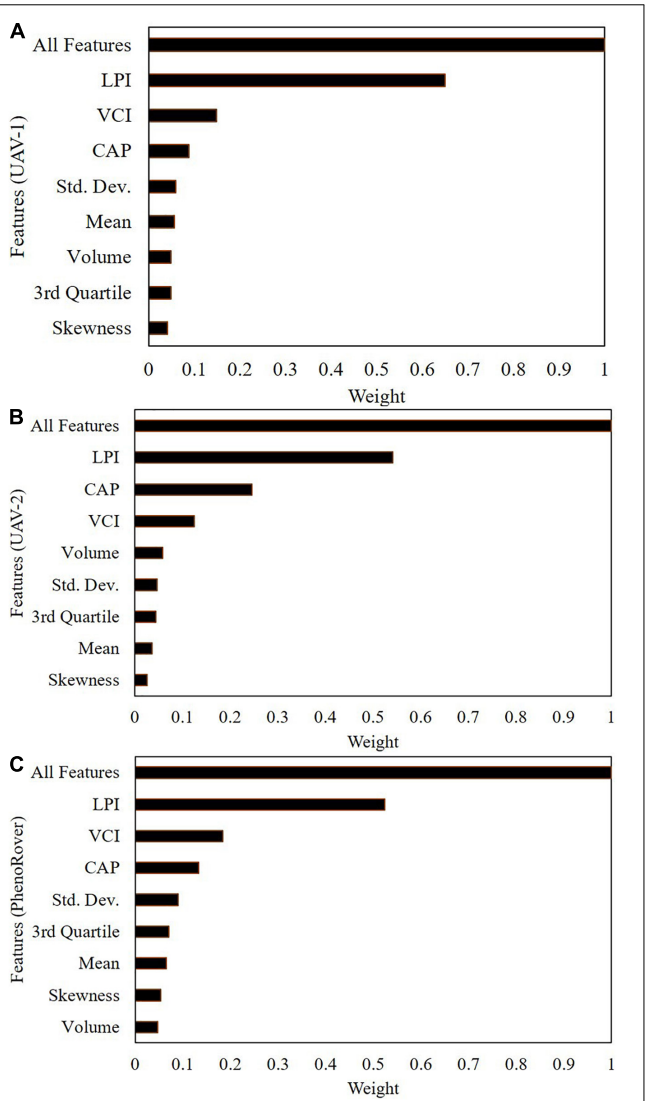

FIGURE 13 | Feature weight evaluation using SVR (RBF) on 7/20/2020: **(A)** UAV-1, **(B)** UAV-2, **(C)** PhenoRover.

SUMMARY AND CONCLUSION

In this exploratory study, the capability of discrete return LiDAR data was investigated for predicting LAI_{eff}. The primary contribution was to develop statistically significant predictive models of LAI over two row crops based on physical features from LiDAR data acquired by multiple platforms during the growing season. In 2020, UAVs and a wheel-based LiDAR dataset were collected and analyzed over two different experiments using

a LAI-2200C plant canopy analyzer. The results based on R^2 values indicate that the LiDAR data are capable of estimating LAI after ∼60 DAS. The R^2 results from maize were compatible with the results from sorghum, and somewhat higher due to less dense planting and complexity in canopy geometry. LiDAR data acquired from the UAV-2 with a Velodyne VLP-32C were higher density, and there was greater penetration of the canopy compared to UAV-1 with a Velodyne VLP-Puck LITE. This was due both to the sensor and the lower flight altitude. However, the R^2 values of the resulting models for LAI were not significantly different. This implies either that the relationship to LAI was dominated by the upper canopy structure or that the penetration associated with more beams and lower flying height was not enough greater to impact the models. Additionally, while the lower height of the boom on the PhenoRover platform was expected to provide improved models due to increased density and penetration, the within-canopy scattering and movement of plants by the platform, especially later in the season, were offsetting problems. As the *t*-test showed, differences in the R^2 values of the models obtained for the different platforms and sensors were not statistically significant. In most datasets, the UAV–based models had higher R^2 values than wheel-based data in 2020, especially later in the growing season when the complex scattering between the near range LiDAR and the canopy appeared to impact the models in both sorghum and maize experiments. The inclusion of data from sorghum photoperiod sensitive varieties did not have a significant impact on the results.

The study encountered multiple challenges, including the limitation of acquiring more wheel-based data subject to weather and field conditions throughout the season. The more frequent remote sensing data acquisition and investigation of the plant canopy analyzer data in 2020 were motivated by the need for more frequent data acquisitions during the vegetative stages of the growth cycle when the plants were growing rapidly and during flowering. The LiDAR data were also impacted by multi-path effects because of the complexity of plants associated with plant density and geometry of sorghum. This motivates further research on denoising approaches. In addition, data encoding approaches may prove useful as an alternative to traditional physical structure-based approaches. The study was conducted in a local environmental condition, and the data were acquired under consistent weather conditions. However, the impact of multiple locations, years, different environmental conditions, soil types, and edaphic factors need to be investigated for the robustness of the models in the application of transfer learning. Finally, further studies are also required, including investigation of other sensor modalities and the sensitivity of the various methods in providing ground reference data and their impact on prediction models.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

BN and MC: conceptualization, formal analysis, and methodology. MC: supervision, writing—review and editing. BN: writing—original draft. MT: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was funded by the Research Projects Agency-Energy (ARPA-E), United States Department of Energy, under Award Number DE-AR0000593.

ACKNOWLEDGMENTS

The authors thank the Purdue TERRA, LARS, and DPRG teams for their work on sensor integration and data collection; and Ayman Habib for his valuable insight throughout this work.

REFERENCES

Akinseye, F. M., Adam, M., Agele, S. O., Hoffmann, M. P., Traore, P. C. S., and Whitbread, A. M. (2017). Assessing crop model improvements through comparison of sorghum (*Sorghum bicolor* L. moench) simulation models: a case study of West African varieties. *Field Crops Res.* 201, 19–31. doi: 10.1016/j.fcr.2016.10.015

Alonzo, M., Bookhagen, B., McFadden, J. P., Sun, A., and Roberts, D. A. (2015). Mapping urban forest leaf area index with airborne lidar usi ng penetration metrics and allometry. *Remote Sens. Environ.* 162, 141–153.

Blancon, J., Dutartre, D., Tixier, M.-H., Weiss, M., Comar, A., Praud, S., et al. (2019). A high-throughput model-assisted method for phenotyping maize green leaf area index dynamics using unmanned aerial vehicle imagery. *Front. Plant Sci.* 10:685. doi: 10.3389/fpls.2019.00685

Černý, J., Pokorný, R., Haninec, P., and Bednář, P. (2019). Leaf area index estimation using three distinct methods in pure deciduous stands. *J. Vis. Exp.* doi: 10.3791/59757

Chen, J. M., Menges, C. H., and Leblanc, S. G. (2005). Global mapping of foliage clumping index using multi-angular satellite data. *Remote Sens. Environ.* 97, 447–457.

Fang, H., Baret, F., Plummer, S., and Schaepman-Strub, G. (2019). An overview of global leaf area index (lai): methods, products, validation, and applications. *Rev. Geophys.* 57, 739–799. doi: 10.1029/2018RG000608

Feng, K., and Li, Q. (2014). "Using stepwise regression and support vector regression to comprise REITs' portfolio," in *Proceedings of the 2014 IEEE 7th Joint International Information Technology and Artificial Intelligence Conference*, (Piscataway, NJ: IEEE), 158–162.

Fournier, R. A., and Hall, R. J. (2017). *Hemispherical Photography In Forest Science: Theory, Methods, Applications*. Dordrecht: Springer.

Hammer, G. L., van Oosterom, E., McLean, G., Chapman, S. C., Broad, I., Harland, P., et al. (2010). Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops. *J. Exp. Bot.* 61, 2185–2202. doi: 10.1093/jxb/erq095

Hasheminasab, S. M., Zhou, T., and Habib, A. (2020). GNSS/INS-assisted structure from motion strategies for UAV-based imagery over mechanized agricultural fields. *Remote Sens.* 12:351. doi: 10.3390/rs12030351

Jensen, J. L., Humes, K. S., Vierling, L. A., and Hudak, A. T. (2008). Discrete return lidar-based prediction of leaf area index in two conifer forests. *Remote Sens. Environ.* 112, 3947–3957. doi: 10.1016/j.rse.2008.07.001

Jimenez-Berni, J. A., Deery, D. M., Rozas-Larraondo, P., Condon, A. T. G., Rebetzke, G. J., James, R. A., et al. (2018). High throughput determination of plant height, ground cover, and above-ground biomass in wheat with LiDAR. *Front. Plant Sci.* 9:237. doi: 10.3389/fpls.2018.00237

Johnsson, T. (1992). A procedure for stepwise regression analysis. *Stat. Pap.* 33, 21–29. doi: 10.1007/BF02925308

Jung, J., and Crawford, M. M. (2012). Extraction of features from LIDAR waveform data for characterizing forest structure. *IEEE Geosci. Remote Sens. Lett.* 9, 492–496. doi: 10.1109/LGRS.2011.2172769

Korhonen, L., Korpela, I., Heiskanen, J., and Maltamo, M. (2011). Airborne discrete-return LIDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index. *Remote Sens. Environ.* 115, 1065–1080. doi: 10.1016/j.rse.2010.12.011

Lang, A. R. G. (1986). Leaf-area and average leaf angle from transmission of direct sunlight. *Aust. J. Bot.* 34, 349–355. doi: 10.1071/bt9860349

Lefsky, M. A., Cohen, W. B., Parker, G. G., and Harding, D. J. (2002). Lidar remote sensing for ecosystem studies: Lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. *BioScience* 52, 19–30.

Lobell, D. B., Thau, D., Seifert, C., Engle, E., and Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sens. Environ.* 164, 324–333. doi: 10.1016/j.rse.2015.04.021

Ludwig, J. A., Quartet, L., and Reynolds, J. F. (1988). *Statistical Ecology: A Primer In Methods And Computing*. Hoboken, NJ: John Wiley & Sons.

Masjedi, A., Carpenter, N. R., Crawford, M. M., and Tuinstra, M. R. (2019). "Prediction of sorghum biomass using uav time series data and recurrent neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Long Beach, CA.

Masjedi, A., Zhao, J., Thompson, A. M., Yang, K.-W., Flatt, J. E., Crawford, M. M., et al. (2018). "Sorghum biomass prediction using uav-based remote sensing data and crop model simulation," in *Proceedings of the IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 7719–7722. doi: 10.1109/IGARSS.2018.8519034

Nazeri, B. (2021). Evaluation of Multi-Platform LiDAR-Based Leaf Area Index Estimates Over Row Crops. Ph.D. Thesis, Purdue University Graduate School, West Lafayette, Indiana, United States.

Nie, S., Wang, C., Dong, P., Xi, X., Luo, S., and Zhou, H. (2016). Estimating leaf area index of maize using airborne discrete-return LiDAR data. *IEEE J. Sel. Top Appl. Earth Obs. Remote Sens.* 9, 3259–3266. doi: 10.1080/2150704X.2015.1111536

Pope, G., and Treitz, P. (2013). Leaf area index (LAI) estimation in boreal mixedwood forest of Ontario, Canada using light detection and ranging (LiDAR) and WorldView-2 imagery. *Remote Sens.* 5, 5040–5063. doi: 10.3390/rs5105040

Richardson, J. J., Moskal, L. M., and Kim, S.-H. (2009). Modeling approaches to estimate effective leaf area index from aerial discrete-return LIDAR. *Agric. For. Meteorol.* 149, 1152–1160.

Rosipal, R., and Krämer, N. (2005). "Overview and recent advances in partial least squares," in *Proceedings of the International Statistical and Optimization Perspectives Workshop" Subspace, Latent Structure and Feature Selection*, (Berlin: Springer), 34–51. doi: 10.1016/j.csbj.2020.07.009

Ryu, Y., Nilson, T., Kobayashi, H., Sonnentag, O., Law, B. E., and Baldocchi, D. D. (2010). On the correct estimation of effective leaf area index: does it reveal information on clumping effects? *Agric. For. Meteorol.* 150, 463–472. doi: 10.1016/j.agrformet.2010.01.009

Sonnentag, O., Talbot, J., Chen, J. M., and Roulet, N. T. (2007). Using direct and indirect measurements of leaf area index to characterize the shrub canopy in an ombrotrophic peatland. *Agric. For. Meteorol.* 144, 200–212.

ten Harkel, J., Bartholomeus, H., and Kooistra, L. (2020). Biomass and crop height estimation of different crops using UAV-based lidar. *Remote Sens.* 12:17. doi: 10.3390/rs12010017

van Ewijk, K. Y., Treitz, P. M., and Scott, N. A. (2011). Characterizing forest succession in central ontario using LiDAR-derived indices. *Photogramm. Eng.Remote Sens.* 77, 261–269. doi: 10.14358/pers.77.3.261

Velodyne VLP-32C (2020). Available online at: http://www.mapix.com/wp-content/uploads/2018/07/63-9378_Rev-D_ULTRA-Puck_VLP-32C_Datasheet_Web.pdf (accessed September 19, 2020).

Velodyne VLP-Puck Hi-Res (2020). Available online at: http://www.mapix.com/wp-content/uploads/2018/07/63-9318_Rev-E_Puck-Hi-Res_Datasheet_Web.pdf (accessed September 19, 2020).

Velodyne VLP-Puck Lite (2020). Available online at: http://www.mapix.com/wp-content/uploads/2018/07/63-9286_Rev-H_Puck-LITE_Datasheet_Web.pdf (accessed September 19, 2020).

Welles, J. M., and Cohen, S. (1996). Canopy structure measurement by gap fraction analysis using commercial instrumentation. *J. Exp. Bot.* 47, 1335–1342. doi: 10.1093/jxb/47.9.1335

White, W. A., Alsina, M. M., Nieto, H., McKee, L. G., Gao, F., and Kustas, W. P. (2019). Determining a robust indirect measurement of leaf area index in California vineyards for validating remote sensing-based retrievals. *Irrig Sci.* 37, 269–280. doi: 10.1007/s00271-018-0614-8

Yang, K.-W., Chapman, S., Carpenter, N., Hammer, G., McLean, G., Zheng, B., et al. (2021). Integrating crop growth models with remote sensing for predicting biomass yield of sorghum. *Silico Plants* 3:diab001.

Zhao, K., and Popescu, S. (2009). Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA. *Remote Sens. Environ.* 113, 1628–1645. doi: 10.1016/j.rse.2009.03.006

Zheng, G., and Moskal, L. M. (2009). Retrieving leaf area index (LAI) using remote sensing: theories, methods and sensors. *Sensors* 9, 2719–2745. doi: 10.3390/s90402719

Zhu, X., Liu, J., Skidmore, A. K., Premier, J., and Heurich, M. (2020). A voxel matching method for effective leaf area index estimation in temperate deciduous forests from leaf-on and leaf-off airborne LiDAR data. *Remote Sens. Environ.* 240:111696. doi: 10.1016/j.rse.2020.111696

Conflict of Interest: 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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 07 December 2021
doi: 10.3389/fpls.2021.756768

Phenomics-Assisted Selection for Herbage Accumulation in Alfalfa (*Medicago sativa* L.)

*Anju Biswas[1], Mario Henrique Murad Leite Andrade[1], Janam P. Acharya[1], Cleber Lopes de Souza[1], Yolanda Lopez[1], Giselle de Assis[2], Shubham Shirbhate[3], Aditya Singh[3], Patricio Munoz[4] and Esteban F. Rios[1]**

[1] Department of Agronomy, University of Florida, Gainesville, FL, United States, [2] EMBRAPA-ACRE, Rio Branco, Brazil, [3] Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL, United States, [4] Department of Horticultural Sciences, University of Florida, Gainesville, FL, United States

OPEN ACCESS

Edited by:
Alessandro Matese, Institute for Bioeconomy, National Research Council, Consiglio Nazionale delle Ricerche (CNR), Italy

Reviewed by:
Roxana Vidican, University of Agricultural Sciences and Veterinary Medicine of Cluj-Napoca, Romania
Raju Bheemanahalli Rangappa, Mississippi State University, United States

***Correspondence:**
Esteban F. Rios
estebanrios@ufl.edu

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

Received: *11 August 2021*
Accepted: *08 November 2021*
Published: *07 December 2021*

Citation:
Biswas A, Andrade MHML, Acharya JP, de Souza CL, Lopez Y, de Assis G, Shirbhate S, Singh A, Munoz P and Rios EF (2021) Phenomics-Assisted Selection for Herbage Accumulation in Alfalfa (Medicago sativa L.). Front. Plant Sci. 12:756768. doi: 10.3389/fpls.2021.756768

The application of remote sensing in plant breeding is becoming a routine method for fast and non-destructive high-throughput phenotyping (HTP) using unmanned aerial vehicles (UAVs) equipped with sensors. Alfalfa (*Medicago sativa* L.) is a perennial forage legume grown in more than 30 million hectares worldwide. Breeding alfalfa for herbage accumulation (HA) requires frequent and multiple phenotyping efforts, which is laborious and costly. The objective of this study was to assess the efficiency of UAV-based imagery and spatial analysis in the selection of alfalfa for HA. The alfalfa breeding population was composed of 145 full-sib and 34 half-sib families, and the experimental design was a row-column with augmented representation of controls. The experiment was established in November 2017, and HA was harvested four times between August 2018 and January 2019. A UAV equipped with a multispectral camera was used for HTP before each harvest. Four vegetation indices (VIs) were calculated from the UAV-based images: NDVI, NDRE, GNDVI, and GRVI. All VIs showed a high correlation with HA, and VIs predicted HA with moderate accuracy. HA and NDVI were used for further analyses to calculate the genetic parameters using linear mixed models. The spatial analysis had a significant effect in both dimensions (rows and columns) for HA and NDVI, resulting in improvements in the estimation of genetic parameters. Univariate models for NDVI and HA, and bivariate models, were fit to predict family performance for scenarios with various levels of HA data (simulated *in silico* by assigning missing values to full dataset). The bivariate models provided higher correlation among predicted values, higher coincidence for selection, and higher genetic gain even for scenarios with only 30% of HA data. Hence, HTP is a reliable and efficient method to aid alfalfa phenotyping to improve HA. Additionally, the use of spatial analysis can also improve the accuracy of selection in breeding trials.

Keywords: high-throughput phenotyping (HTP), normalized difference vegetation index (NDVI), remote sensing (RS), spatial variation, genetic gain, forage, plant breeding

INTRODUCTION

Alfalfa (*Medicago sativa* L.) is the most important perennial forage legume globally because of its relatively high yield and nutritional value (Annicchiarico, 2015). In the United States, alfalfa is the fourth most valued crop behind corn, soybeans, and wheat, with an estimated value of $8.4 billion (USDA-NASS, 2020), playing a critical role in the food supply chain (Feng et al., 2020). In 2018, nearly 53 million tons of alfalfa and alfalfa-grass mixtures were harvested from almost seven million hectares in the United States. Most of the production is concentrated in the mid-east and west coast (USDA-NASS, 2020). Despite its lower presence as a forage crop in the lower southeastern United States and other subtropical regions in the world, breeding efforts are underway to develop non-dormant alfalfa cultivars adapted to these environments (De Assis et al., 2010; Vivela et al., 2018; Adhikari et al., 2019; Acharya et al., 2020).

Alfalfa breeding is typically conducted as phenotypic recurrent selection using among and within half-sib family selection (Casler and Brummer, 2008), although various breeding schemes have been proposed to improve herbage accumulation (HA) in alfalfa (Annicchiarico and Pecetti, 2021). The improvement of HA in alfalfa is challenging due to long selection cycles, tetrasomic inheritance, high inbreeding depression, and significant genotype and environment interaction for this complex trait (Bingham et al., 1994; Brummer, 1999; Annicchiarico, 2015). Additionally, phenotyping for HA requires investment of significant resources (Annicchiarico et al., 2016). In recent years, most alfalfa breeding programs have focused on improving disease/pest resistance, long-term persistence, and other specific traits targeting transgenes for glyphosate tolerance or decreased lignin. The lack of efforts to improve HA can explain the low genetic gain in alfalfa yield observed in the last decades (Brummer and Casler, 2015).

Nevertheless, HA has become a target breeding trait among alfalfa breeders more recently (Sakiroglu and Brummer, 2017; Dos Santos et al., 2018; Adhikari et al., 2019; Acharya et al., 2020; Benabderrahim et al., 2020; He et al., 2020; Ren et al., 2021; Tang et al., 2021). However, traditional field phenotyping for HA is based on the destructive sampling of experimental units at the ground level, weighing fresh samples, drying, and weighing dried samples to estimate dry matter content. The manual phenotyping process for HA is labor-intensive, time-consuming, and costly.

Plant phenotyping plays a central role in plant breeding, and the accurate and rapid acquisition of phenotypic data is valuable for exploring the association between genotypes and phenotypes. In the last few decades, remote sensing has been widely used in agriculture (Maes and Steppe, 2019; Galli et al., 2020), particularly for high-throughput phenotyping (HTP) in breeding applications (Furbank and Tester, 2011; White et al., 2012; Araus and Cairns, 2014; Tattaris et al., 2016; Li et al., 2017; Zhao et al., 2019). Remote sensing offers unprecedented spectral, spatial, and temporal resolution, providing detailed vegetation data (Maes and Steppe, 2019). Several vegetation indices (VIs) such as normalized difference vegetation index (NDVI), green NDVI (GNDVI), normalized difference red edge (NDRE), or Green and Red ratio Vegetation Index (GRVI) have been employed to assess vegetation vigor and canopy cover over multiple crops (Lima-Cueto et al., 2019; Quirós Vargas et al., 2019; Ranjan et al., 2019; Zhang et al., 2019).

Remote sensing techniques have shown to enable efficient and non-destructive estimation of HA in alfalfa (Feng et al., 2020), such as screening large breeding populations (Cazenave et al., 2019). According to Cazenave et al. (2019), HTP can detect small differences in alfalfa yield when screening diverse germplasm. More recently, HTP improved the efficiency of the selection process for biomass in small plots (1.52 m × 0.30 m) in alfalfa breeding populations (Tang et al., 2021), and provided a good prediction of HA in larger plots (6 m × 4 m) (Feng et al., 2020). Remote sensing can mitigate the challenge of measuring HA in large populations for breeding programs focusing on improving alfalfa HA. Therefore, the implementation of HTP can streamline the phenotyping process for HA in alfalfa.

Residual maximum likelihood (REML) is commonly implemented in breeding programs to estimate variance components and calculate genetic parameters using linear mixed models. The use of best linear unbiased prediction (BLUP) is an established technique to predict breeding values, which are then used to guide breeding decisions. BLUP can generate accurate predictions of breeding values even for unbalanced experimental designs (Piepho et al., 2008). Genetic parameters for alfalfa yield are essential to define optimal selection schemes (Casler and Brummer, 2008). Heritability estimates for HA in alfalfa ranged from 0.15 to 0.30 (Bowley and Christie, 1981; Riday and Brummer, 2002; Annicchiarico, 2015; Acharya et al., 2020). These low to moderate estimates are expected in alfalfa and other perennial forage crops with long selection cycles, cross-pollinated breeding schemes, and traits with significant genotype × environment interaction (Annicchiarico, 2015). Most studies in alfalfa have focused on yield in short-term experiments with few harvests (3–4 harvests per year) (Annicchiarico, 2015), except for more recent studies (Acharya et al., 2020; Annicchiarico and Pecetti, 2021). Besides the challenges mentioned above, field trials are associated with intrinsic and extrinsic variations, which can cause some form of spatial variation between experimental units (Sripathi et al., 2017). Local control, such as blocking and randomization, cannot effectively account for all the spatial trends in large experiments (Gilmour et al., 1997). Spatial variation is expected even using complex experimental designs, such as those commonly used in most plant breeding programs. A better way to control the spatial variation is to implement spatial analysis to detect and correct the variation patterns in multiple dimensions. Experimental units close to each other are expected to be higher correlated than those far apart, and improvements in model fitness and higher selection accuracy have been reported in plant breeding programs (Andrade et al., 2020).

The overall objective of our research was to implement HTP, spatial analysis, and linear mixed models to improve the accuracy of the selection process in alfalfa for HA. The specific objectives were: (i) phenotype of an alfalfa breeding population for HA using ground-based manual sampling and utilize a unmanned aerial vehicle (UAV) for HTP; (ii) assess the efficiency of controlling field variation using spatial models for HA and NDVI

in alfalfa, (iii) calculate the genetic parameters based on HA and NDVI using univariate models; (iv) fit bivariate models for HA and NDVI using all data, and for scenarios with different levels (30–90%) of HA data, and (v) quantify the correlation between breeding values, the coincidence of selection of the best families, and genetic gain across the different scenarios for HA data.

MATERIALS AND METHODS

Germplasm Screening and Development of the Reference Breeding Population

Initially, 121 alfalfa populations with different fall dormancy groups were screened for HA in Citra, FL, United States (Acharya et al., 2020). A total of 33 populations were selected based on high HA and persistence across all harvests. Following the screening, controlled crosses were done in the greenhouse to create the alfalfa reference breeding population. A single plant per population was selected based on vigor, and cuttings were made in the summer 2016. A factorial mating design was used to create all possible full-sib combinations; however, some crosses did not produce enough viable seed, and were not included in field trials. Half-sib seed were also harvested from each parental line. All crosses were conducted in controlled conditions in the Forage Breeding and Genetics Lab greenhouse, at University of Florida (Gainesville, FL, United States). Seeds from each full-sib and half-sib families were harvested, threshed individually, planted in 72-cell Styrofoam trays in August 2017, and maintained in the greenhouse until transplanting in November 2017. In total, 145 full-sib and 33 half-sib families were established in the field and this population represents the reference set for the HTP study. Acharya et al. (2020) provide more details and results for the initial screening and crosses.

Experimental Design and Field Management

The breeding trial with the reference population was conducted at Citra, Florida (29°40′ N, 82°167′ W, 48 m above the sea level) following a row and column design with augmented representation of controls. Each experimental unit (1.82 m × 1.82 m) consisted of eight rows spaced at 22.8 cm. The three border rows on each side were seeded with the Bulldog 805 to serve as borders, and twenty alfalfa seedlings were transplanted in the middle two rows. Three rows were seeded with Bulldog 805 on each side of the transplants to serve as borders. The breeding population was composed of 145 full-sib and 34 half-sib families. Three controls were used: the cultivars Bulldog 805, Florida 99, and an advanced breeding line named UF_AP_2015. Eighty-one families were replicated three times, 61 families were replicated two times, and 40 families were used one time due to limited seed availability. The experiment was established in November 2017, and data collection occurred in August 2018, October 2018, December 2018, and January 2019. The field was fertilized with 67.25 kg·K_2O·ha^{-1}, using Muriate of Potash, and with Boron at the rate of 1.12 kg ha^{-1} and herbicide Clethodim (Select, 70.76 g AI/L-1; Valent United States Corporation, Walnut Creek, CA, United States) was applied to control grasses at the rate of 1.05 kg ha^{-1} after each harvest. Manual weeding was done as needed to control broad-leaf weeds after each harvest.

Ground-Based Data Collection

The experimental units were manually harvested to determine HA (kg ha^{-1}) when the control UF_AP_2015 reached 10% blooming (**Figure 1A**). The harvest was performed by mowing the six outer rows (three rows on each side) with a flail mower at 10 cm stubble height, and then the two central rows were cut and weighted to determine the fresh weight (g) by the plot. Approximately 500 g of fresh shoots were collected from each plot and placed in a dryer at 55°C for 7 days to determine dry matter content, and HA per plot was estimated on a dry matter basis (kg ha^{-1}).

Remote Sensing Data Collection

A UAV (DJI Matrice 100) equipped with a multispectral camera (RedEdge, MicaSense, Seattle, WA, United States) was used to obtain imagery over the entire field after the border rows were mowed (**Figure 1B**). AtlasFlight app (MicaSense Inc., Seattle, WA, United States) was used to automatically sample fields at an altitude of 30 m, a flight speed of 6 m/s speed, and enforcing a 75% overlap in collected imagery. A calibration panel (MicaSense Inc., Seattle, WA, United States) was set before starting each flight to allow post-collection calibration of imagery.

Image Processing and Data Acquisition

All images were stitched into orthomosaics using AgiSoft Photoscan (Agisoft LLC, St. Petersburg, Russia; **Figure 2A**). The orthomosaic corresponded to the entire field for a single harvest event and comprised five bands: blue (475 nm), green (560 nm), red (668 nm), near-infrared (NIR, 840 nm), and red edge (717 nm). The orthomosaics were further processed in QGIS 3.14 software (QGIS.org 2020) to refine geolocation using field-collected ground control points. We identified experimental units from imagery and created a spatial reference frame (ESRI Shapefile) for further analysis (**Figure 2B**). The shapefile was edited to include all relevant information on the individual field plot. Subsequently, we masked the canopy from the soil using custom python codes (**Figure 2C**) and used the masked image to calculate the total pixel count in each band (**Figures 2D–F**) and to generate different Vis (NDVI, GNDVI, NDRE, and GRVI) from zonal statistics function in QGIS. We also estimated the sum of all VIs to allow for handling plants that died during the experiment.

Data Analysis

Vegetation Index Selection

We utilized boxplots to assess all VIs and HA distributions across each harvest (**Supplementary Figure 1**). We estimated Pearson correlations between HA and VIs-sum (an integrative indicator of VIs) in R (R Core Team, 2020) to assess prospective best-fitting relationships between HA and VIs. Finally, we used ordinary linear mixed models to model HA using VIs-sum for each harvest (R Core Team, 2020).

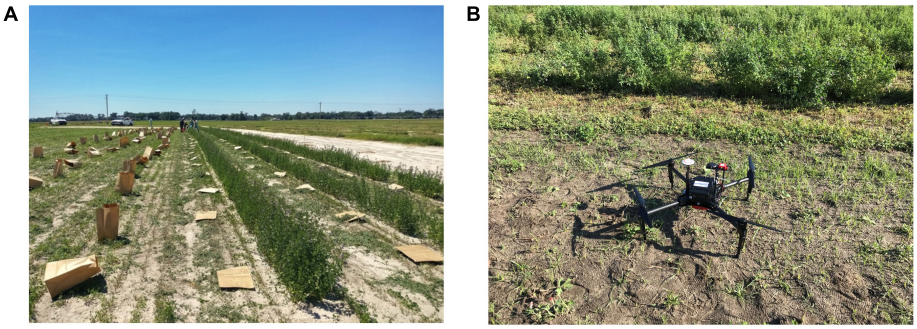

FIGURE 1 | Phenotyping alfalfa for: **(A)** ground-based herbage accumulation (HA); and **(B)** High-throughput phenotyping (HTP) using an unmanned aerial vehicle (Matrice 100) with multispectral camera (MicaSense RedEdge, Seattle, WA, United States) taking off in the experimental area.

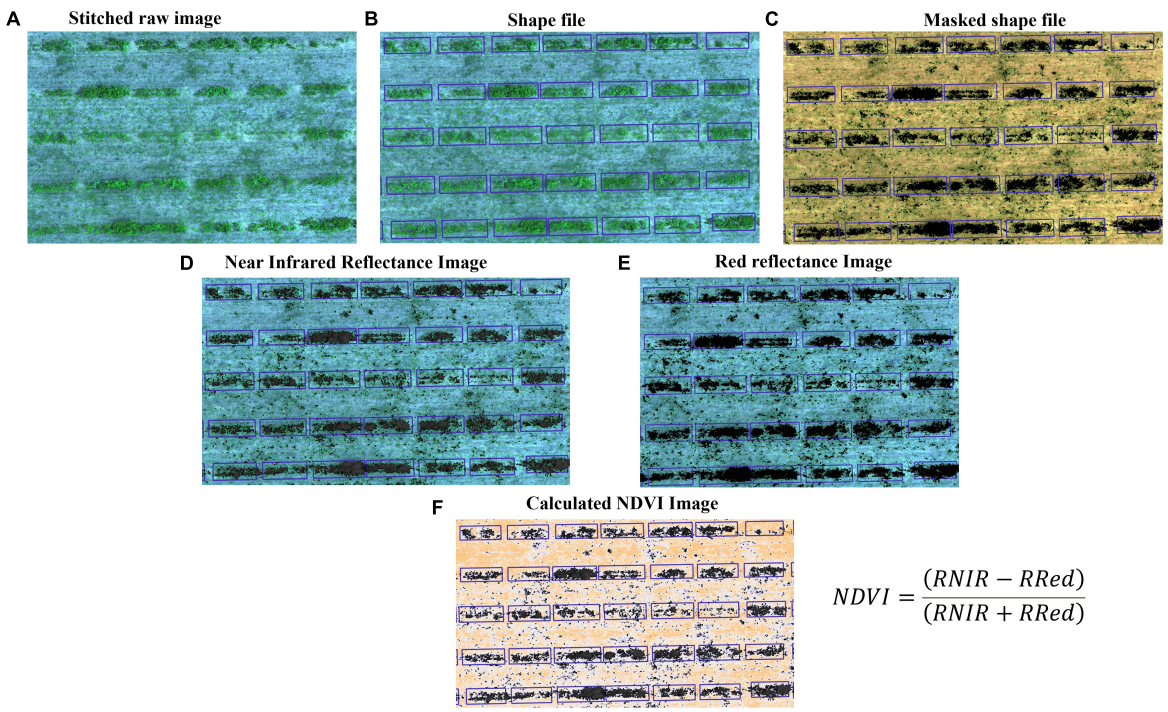

FIGURE 2 | Workflow of image processing and data acquisition for the HTP. **(A)** Stitched raw images (each row is visible) from multispectral camera, **(B)** Shape file: each plot is separated by grid line, **(C)** Masked shapefile: canopy and bare ground places are noticeable, and these images are used to generate vegetation indices, **(D)** Reflectance image of infrared band, **(E)** Reflectance image of red band, and **(F)** Calculated NDVI image using panels **(D,E)**.

Variance Component Estimation: Base Model

Linear mixed models were fit using the package ASReml-R (Butler et al., 2009) in the software R (R Core Team, 2020). The significance of random effects was determined by the likelihood ratio test. Univariate models were fit for NDVI and HA by harvest, as follows:

$$\boldsymbol{y} = \mu + \boldsymbol{Xt} + \boldsymbol{Z_r u_r} + \boldsymbol{Z_c u_c} + \boldsymbol{Z_f u_f} + \boldsymbol{e} \quad (1)$$

where $\boldsymbol{y}$ is the vector of the response variable, μ is the overall mean; t is the fixed effect vector of the check varieties; $\boldsymbol{u_r}$ is the random effect vector of the row, $\boldsymbol{u_r} \sim N(0, \mathbf{I}\,\sigma_r^2)$; $\boldsymbol{u_c}$ is the random effect vector of the column, $\boldsymbol{u_c} \sim N(0, \mathbf{I}\,\sigma_c^2)$; $\boldsymbol{u_f}$ is the random effect vector of the family, assuming that families are independent $\boldsymbol{u_f} \sim N(0, \mathbf{I}\,\sigma_f^2)$, and $\mathbf{e}$ is the independent error random vector of residual, $e \sim N(0, \mathbf{I}\,\sigma_e^2)$. I is the identity matrix associated with the vector, while X, $\boldsymbol{Z_r}$, $\boldsymbol{Z_c}$, and $\boldsymbol{Z_f}$ represent the incidence matrices associated with the vectors t, $\boldsymbol{u_r}$, $\boldsymbol{u_r}$, and $\boldsymbol{u_f}$. The variance components of the effects $\boldsymbol{r}$, $\boldsymbol{c}$, $\boldsymbol{f}$, and $\boldsymbol{e}$ are represented by the σ_r^2, σ_c^2, σ_f^2, and σ_e^2, respectively.

Variance Component Estimation: Spatial Model

Due to the intrinsic variation in the field, we explored spatial models to account for spatial autocorrelation among experimental units. In this model, we assumed that the error term

was auto correlated along the rows and columns, and we used a first-order autoregressive process to fit the error:

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{Xt} + \boldsymbol{Z_r r} + \boldsymbol{Z_c c} + \boldsymbol{Z_f f} + \boldsymbol{\xi} \quad (2)$$

where all terms are the same as the model (1) other than the term $\boldsymbol{\xi}$, which is the independent error random vector of residual, $\boldsymbol{\xi} \sim N(0, R_e\sigma_e^2)$, R_e is the covariance matrix of $\boldsymbol{\xi}$, and it is defined as: $R_e \ \sigma\boldsymbol{\xi}^2\boldsymbol{\Sigma_c}(\boldsymbol{\rho_c}) \bigotimes \boldsymbol{\Sigma_r}(\boldsymbol{\rho_r})$. Where $\boldsymbol{\rho_c}$ and $\boldsymbol{\rho_r}$ are the autocorrelation parameters for the spatial coordinates of row and column; $\boldsymbol{\Sigma_c}(\boldsymbol{\rho_c})$ and $\boldsymbol{\Sigma_r}(\boldsymbol{\rho_r})$ represent the autoregressive correlation matrices; and $\bigotimes$ represents the Kronecker product (Andrade et al., 2020).

From both base and spatial models, we estimated the following genetic and non-genetic parameters: broad-sense heritability (H^2), predicted error variance (PEV), and relative efficiency (RE) between the spatial model and base model. The RE was calculated for the spatial model in relation to the base model based on PEV and values greater than 100 indicate higher efficiency for the spatial model. The RE was measured as follows:

$$RE = 100 \times \left(\frac{PEV_{Base}}{PEV_{Spatial}}\right) \quad (3)$$

The Akaike information criteria (AIC) for each model were used to choose the best model. Additionally, the families were ranked based on their predicted values from each model for the traits HA and NDVI.

Variance Component Estimation: Bivariate Model and Scenarios for Herbage Accumulation Data

Data for HA and NDVI were combined into a single model to leverage information at both levels (ground-based and HTP). As manual phenotyping for HA is time-consuming and costly, scenarios were simulated *in silico* to quantify how genetic parameter estimates would change when not all experimental units are manually harvested. The simulation was performed by randomly assigning missing values to the full HA dataset to represent hypothetical scenarios when 30, 40, 50, 60, 70, 80, and 90% of the plots would be harvested. The process was repeated 30 times for each scenario. The base model contained 100% of the HA data, and it was used as a baseline to compare with other scenarios. The bivariate model was fitted as follows:

$$\boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{Xt} + \boldsymbol{Z_r u_r} + \boldsymbol{Z_c u_c} + \boldsymbol{Z_f u_f} + \mathbf{e} \quad (4)$$

where $\mathbf{y}$ is a stacked vector of the phenotypic data for traits HA (t1) and NDVI (t2), $\boldsymbol{\mu}$ is the stacked vector of the overall mean for each trait; t is the fixed effect stacked vector of the check varieties for each trait; $\boldsymbol{u_r}$ is the random effect stacked vector of the row for each trait, $\boldsymbol{u_r} \sim N(0, \sigma_r^2)$; $\boldsymbol{u_c}$ is the random effect stacked vector of the column for each trait, $\boldsymbol{u_c} \sim N(0, \sigma_c^2)$; $\boldsymbol{u_f}$ is the random effect stacked vector of the family for each trait, assuming that families are independent $\boldsymbol{u_f} \sim N(0, I_f \bigotimes \mathbf{G})$, $G = \begin{bmatrix} \sigma^2_{ft1} & \sigma_{ft1ft2} \\ \sigma_{ft1ft2} & \sigma^2_{ft2} \end{bmatrix}$; and e is the independent error random stacked vector of residual for each trait, $\mathbf{e} \sim N(0, I_e \bigotimes R)$, where $R = \begin{bmatrix} \sigma^2_{et1} & \sigma^2_{et1et2} \\ \sigma^2_{et1et2} & \sigma^2_{et2} \end{bmatrix}$; $\mathbf{X}$, Z_r, Z_c, and Z_f represent the incidence matrices associated with the vectors t, $\boldsymbol{u_r}$, $\boldsymbol{u_c}$, and $\boldsymbol{u_f}$. The components σ_r^2, σ_c^2, σ^2_{ft1}, σ^2_{ft2}, σ^2_{et1}, and σ^2_{et2} are the variance components for row, columns, family for trait 1, family for trait 2, error for trait 1, and error for trait 2, respectively. The component $\sigma_{ft1,t2}$ is the covariance between trait 1 and 2.

We compared the full bivariate and univariate models utilizing HA, NDVI with the different scenarios using the following calculations: (i) coincidence of selection (%) after applying a 10% selection intensity and (ii) the correlation among predicted values across all families.

Genetic gain (%) was estimated from BLUPs for each family in each harvest (**Supplementary Table 3**) for the bivariate model, and univariate models for HA and NDVI in all scenarios for HA missing data. We calculated genetic gain using the following equation:

$$\textit{Genetic Gain} = \frac{\overline{BLUP_t}}{\overline{Y}} \ \boldsymbol{x}\ \mathbf{100} \quad (5)$$

$\overline{BLUP_t}$ is the mean of the BLUPs of the t selected family for HA, $\overline{Y}$ is the overall mean of all families for HA.

RESULTS

Pearson Correlation and Regression Analysis Between Herbage Accumulation and Vegetation Indices

Herbage accumulation showed variation across harvests (**Supplementary Figure 1**). Harvest one and four showed higher mean HA and variation, while harvest two had the lowest mean HA and variation. All VIs responded similarly to the variation in HA across harvests, as all VIs had a higher mean and variation for harvest one, and the lowest mean and variation in harvest two (**Supplementary Figures 3A–E**). The Pearson correlation between HA and all VIs (NDVI, GNDVI, NDRE, and GRVI) were higher than 0.71 across all four harvests (**Figure 3**). All VIs were able to model HA with moderate and similar accuracy across harvests (**Figure 4** and **Supplementary Figures 2–4**). Harvest one, three, and four showed better prediction ($R^2 > 0.60$) for HA than harvest two ($R^2 \sim 0.51$) (**Figure 4** and **Supplementary Figures 2–4**). Due to the similar results observed among all VIs, NDVI was selected for further analyses.

Spatial Analysis to Control Field Variation

Modeling the spatial variation was necessary for both traits (HA and NDVI) since the autocorrelation in both dimensions was significant in all harvests (**Supplementary Table 1**). Variograms for the base model showed the presence of patterns in the field that may increase error variance (peaks in the variograms indicate trends in the field) (**Figures 5A,C**). The variograms revealed that the spatial model efficiently controlled these patterns (**Figures 5B,D**). The spatial models for HA and NDVI provided better model fitness across all harvests (lower AIC and BIC; **Supplementary Table 2**).

The genotypic variance was significant ($P < 0.001$) for HA and NDVI across harvests, and heritability (H^2) estimates ranged from low (0.12) to moderate (0.31) (**Table 1**). For HA, the spatial

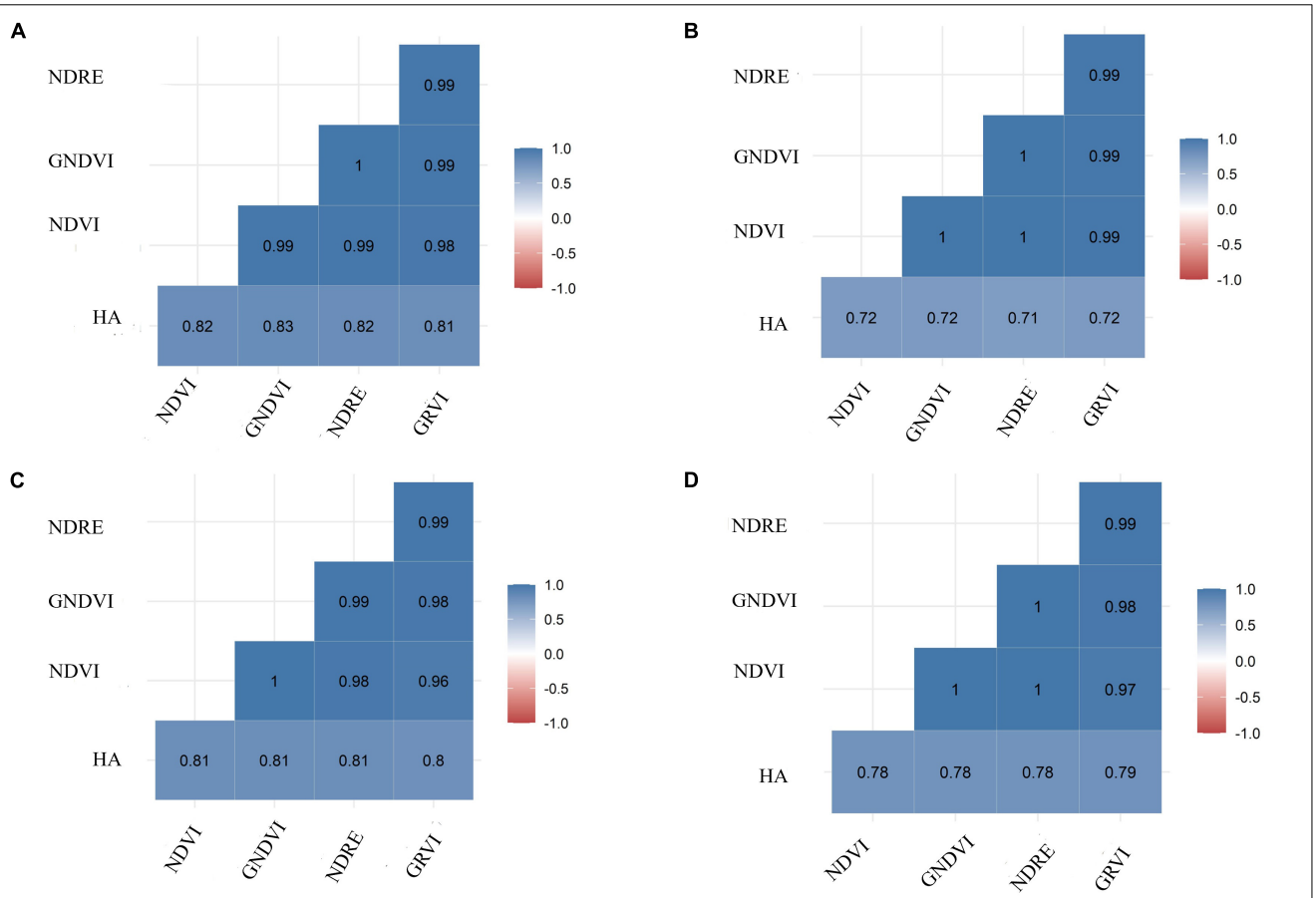

FIGURE 3 | Pearson correlation coefficients between ground-based phenotyping for herbage accumulation (HA—kg ha^{-1}) and four vegetation indices (VIs): NDVI, normalized difference vegetation index; GNDVI, green normalized difference vegetation index; NDRE, normalized difference red edge; GRVI, green and red ratio vegetation index. Harvests: **(A)** harvest one, **(B)** harvest two, **(C)** harvest three, and **(D)** harvest four.

model provided slightly higher H^2 estimates than the base model across the harvests (**Table 1**). Similarly, NDVI models accounting for spatial variation resulted in higher H^2 estimates, except for harvest four (**Table 1**). The H^2 estimates for HA (base and spatial models) were higher than H^2 estimates for NDVI (base and spatial models) across all harvests. As model fitness was greater for spatial models, PEVs were also smaller for spatial models for HA and NDVI across harvests, except for NDVI in harvest four (**Table 1**). The lower PEV in the spatial models yielded higher RE than the base model, for all traits and harvests, except for HA in harvest four. The best model for each trait and harvest was used for further analyses.

Selection of Best Alfalfa Families for Herbage Accumulation and Normalized Difference Vegetation Index

The 179 alfalfa families were ranked based on their predicted values estimated on each harvest for the base and spatial models. Then, a 10% selection intensity was imposed to select the best 17 families (highest HA and NDVI values in each harvest). The coincidence of selection was greater than 75% (13 families out of 17) for the base and spatial models for HA and NDVI in all harvests, except for NDVI in harvest three (**Figure 6**). The coincidence of selection between HA and NDVI, based on the base model, ranged from 36% in harvest 2–65% in harvest one (**Figure 6**). The coincidence of selection between spatial models for HA and NDVI ranged from 41% in harvest 2–71% in harvest one (**Figure 6**).

Harvest one showed the highest coincidence of selection for all model comparisons (**Figure 6**). Base and spatial models for HA and NDVI resulted in 89% coincidence (15 families out of 17) (**Figure 6**). Considering the base model in harvest one, selecting families using HA and NDVI showed 65% coincidence (11 families). After modeling spatial variation, there was a 71% coincidence (12 families) when the best families were selected based on HA and NDVI. Harvest two showed the lowest coincidence when comparing the selected families for HA and NDVI using base and spatial models (**Figure 6**). Considering the base model in harvest two, selecting families using HA and NDVI showed 35% coincidence (five families). There was a 43% coincidence (seven families) for the spatial models to select the best 10% families based on HA and NDVI data. Harvests three

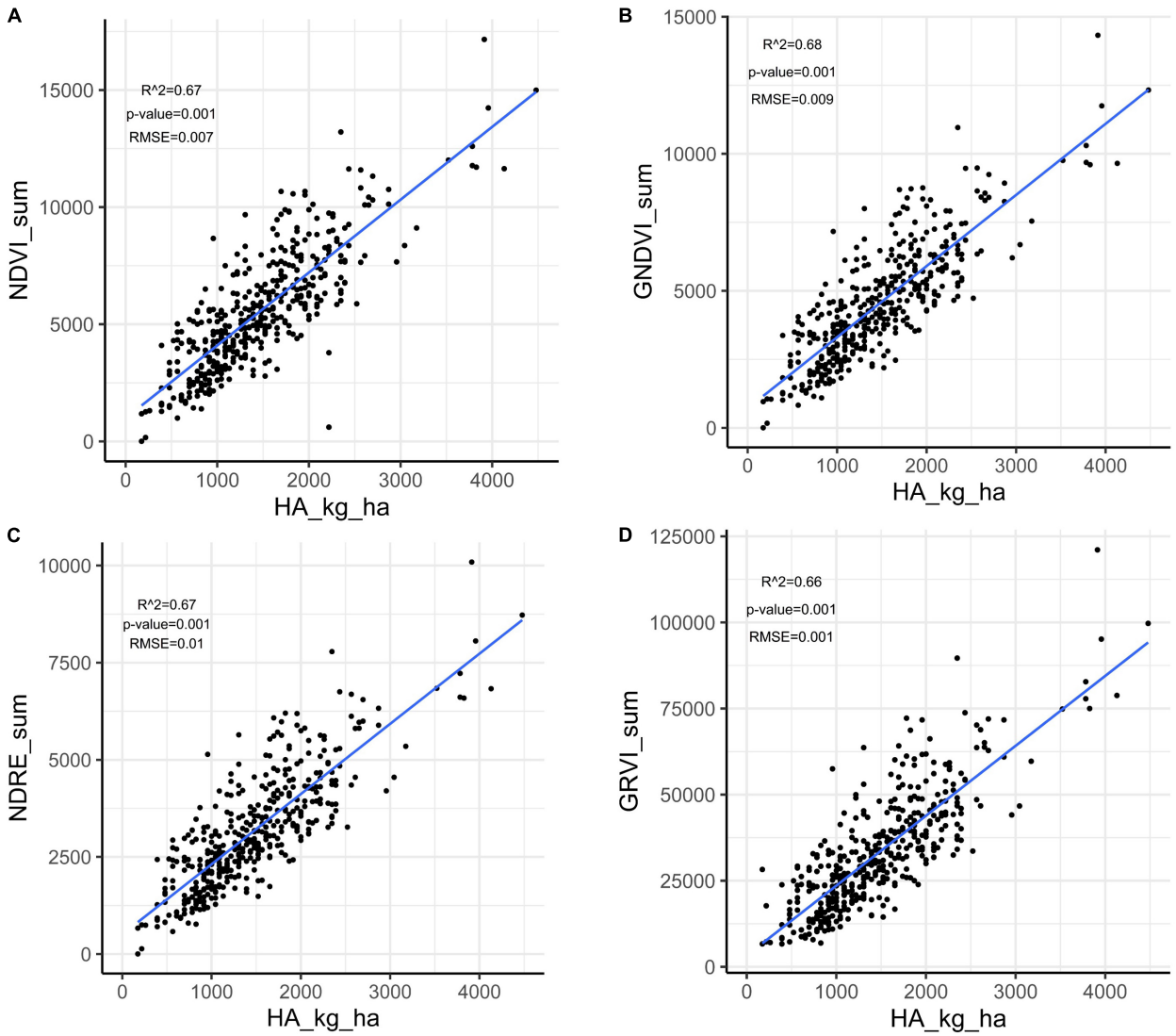

FIGURE 4 | Linear regression between herbage accumulation (HA_kg_ha) and unmanned aerial vehicles-based VIs collected in harvest one in alfalfa families evaluated in Citra, FL. VIs: **(A)** NDVI, normalized difference vegetation index; **(B)** GNDVI, green normalized difference vegetation index; **(C)** NDRE, normalized difference red edge; GRVI, **(D)** green and red ratio VI.

and four showed similar results to harvest one, but slightly lower coincidence when comparing HA and NDVI for base and spatial models.

Univariate and Bivariate Models for Scenarios With Different Levels of Herbage Accumulation Data

The combination of HA and NDVI data into a bivariate analysis was compared to the univariate models for each trait, considering scenarios with various levels of HA data for three parameters: correlation among predicted values across all families, the coincidence of selection for the 10% best families, and genetic gain. In general, all parameters increased as the level of HA data increased in all harvests (**Figures 7**, **8**).

For the bivariate model, the correlation between the predicted values for all families using the complete HA dataset and each scenario with various levels of missing HA data (30–90%) varied between 0.78 (harvest four at 30% HA data collection) and 1 (harvest one at 90% HA data collection) (**Figure 7**). The correlation was consistently higher than 0.90 for three harvests (one, two, and three), even for the scenario when only 30% of HA data were used in the model. The coincidence of selection for the best 10% families varied between 0.64 (harvest four at 30% HA data collection) and 0.85 (harvest one at 90% HA data collection) (**Figure 7**). The correlation and coincidence of selection were consistently higher for the bivariate model than any univariate model across all scenarios (**Figure 7**). The bivariate and univariate models for HA were similar only in harvests three and four, for scenarios when 80 and 90% of the HA data were used

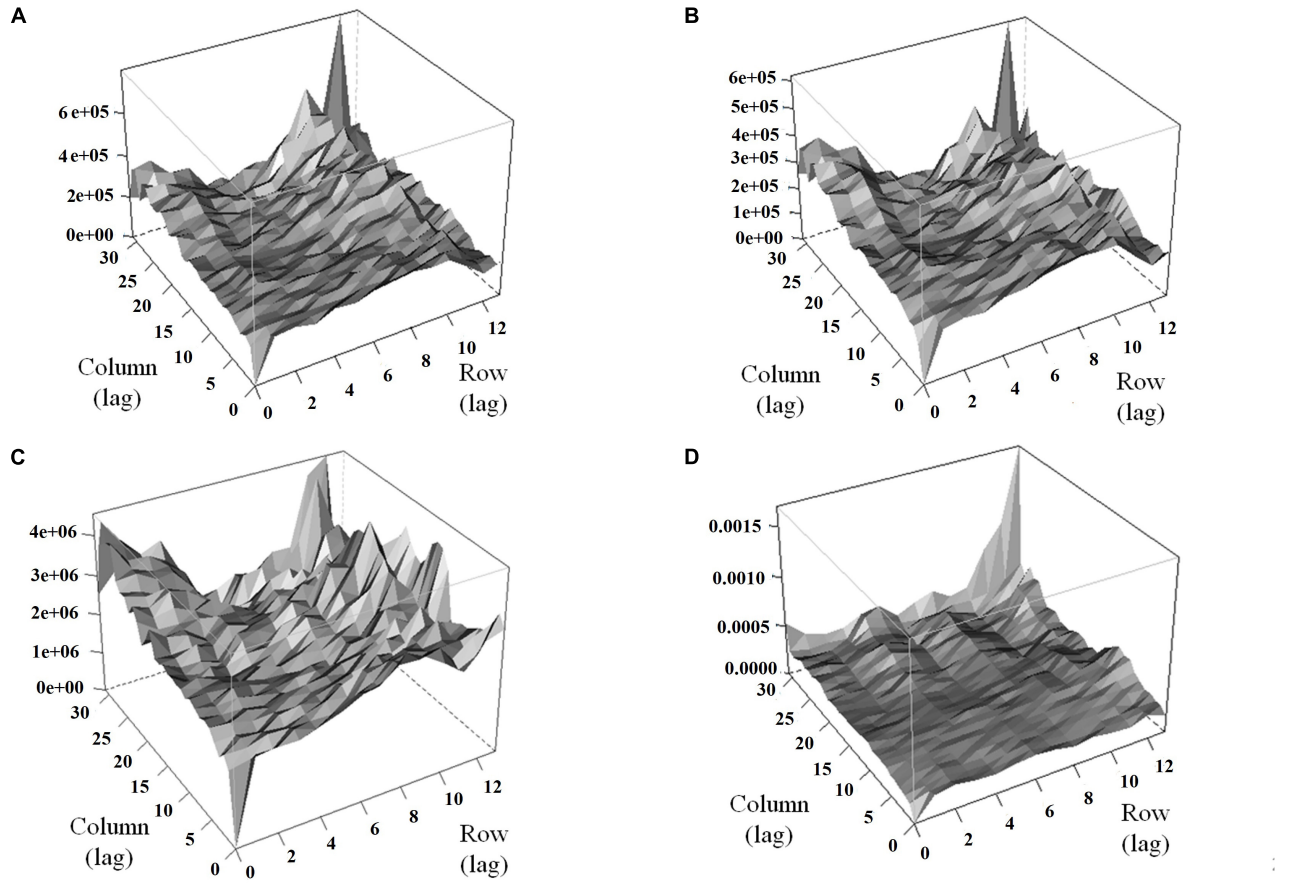

FIGURE 5 | Variograms for HA and normalized difference vegetation index (NDVI) before **(A,C)** and after **(B,D)** the inclusion of terms to control local and global trends for alfalfa yield in harvest one in alfalfa families evaluated in Citra, FL. **(A)** HA-based model, **(B)** HA-spatial model, **(C)** NDVI-based model, **(D)** NDVI-spatial model. Row and column are coordinates for the rows and columns in the experimental area, respectively.

in the model (**Figure 7**). The genetic gain for the bivariate model was higher than the univariate models for HA and NDVI in all harvests (except in harvest three for HA in the scenarios when 80 and 90% of the HA data were used in the model), and it remained stable even for scenarios with low levels of HA data (**Figure 8**).

For the univariate model for HA, the correlation between the genotypic values among all families varied between 0.51 (harvest one at 30% HA data collection) and 0.91 (harvest four for 90% HA data collection) (**Figure 7**). The coincidence of selection varied between 0.37 (harvest three at 30% HA data collection) and 0.84 (harvest two at 90% HA data collection) (**Figure 7**). The genetic gain for the univariate model for HA increased as more HA data was used in the models across all harvests, and higher gains were obtained for HA compared to NDVI for almost all scenarios (**Figure 8**). For the univariate model for NDVI, the correlation between genotypic values among all families varied between 0.45 (harvests two and four at 30% HA data collection) and 0.88 (harvest three for 90% HA data collection) (**Figure 7**). The coincidence for selection varied between 0.33 (harvest four at 30% HA data collection) and 0.63 (harvest four at 90% HA data collection) (**Figure 7**). In general, the univariate model for HA provided higher correlations and % coincidence than univariate models for NDVI (**Figure 7**), and lower genetic gain for HA was obtained when the selection was performed using only NDVI data (**Figure 8**).

DISCUSSION

The ultimate goal in plant breeding is to select superior breeding units (individuals, clones, families, etc.) with the highest accuracy level in a high throughput manner by investing the least possible resources. Alfalfa breeders aim to develop superior cultivars with high yield and quality, exhibiting broad adaptation to various biotic and abiotic stresses. Breeding programs are focusing on the improving HA invest significant resources in collecting and quantifying HA from field trials and drying samples to determine their dry matter content (Annicchiarico, 2015). This process is time-consuming and expensive for large breeding populations. A key component for increasing the efficiency in improving HA yield is the use of fast and precise phenotypic assessment of large breeding populations (Fu, 2015). In this study, 179 alfalfa families were phenotyped for HA across four harvests, totaling 1,792 data points for HA. At the same time, HTP was implemented to assess

TABLE 1 | Estimates of broad-sense heritability (H^2), predicted error variance (PEV) and relative efficiency (RE) for alfalfa families harvested four times in Citra, FL, United States.

Harvest	Parameter	HA		NDVI	
		Base	Spatial	Base	Spatial
1	H^2	0.28***	0.31***	0.21***	0.29***
	PEV	31,145	29,976	413,474	368,663
	RE	–	103.9	–	112.2
2	H^2	0.18***	0.20***	0.14***	0.18***
	PEV	15,503	15,048	2,140	2,099
	RE	–	103.1	–	101.9
3	H^2	0.24***	0.27***	0.13***	0.19***
	PEV	19,785	19,442	54,468	51,576
	RE	–	101.8	–	105.6
4	H^2	0.19***	0.23***	0.12***	0.11***
	PEV	38,223	39,548	27,300	25,200
	RE	–	97.4	–	108.2

Linear mixed models were fitted for herbage accumulation (HA) and normalized difference vegetation index to estimate variance components in a model without accounting for spatial variation (base) and by modeling the spatial variation (spatial) in each harvest.

****denotes significance at $p < 0.001$ for the genetic variance using a Likelihood Ratio Tests (LRT).*

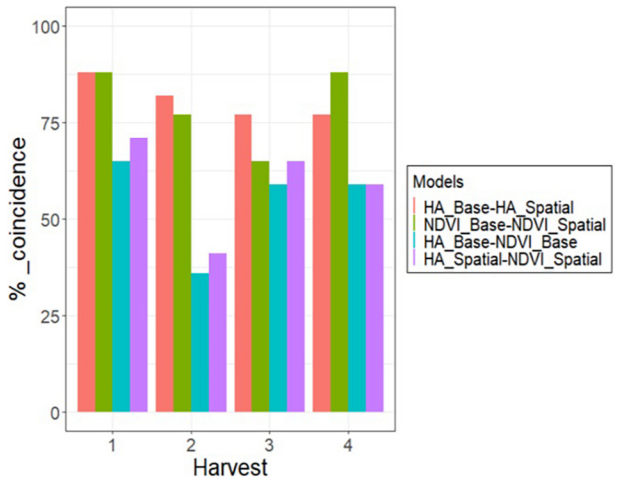

FIGURE 6 | Percent of coincidence in selection for the 10% best alfalfa families between HA and NDVI, using the base and spatial model, across four harvests in alfalfa families evaluated in Citra, FL, United States.

the efficiency of HTP to predict alfalfa HA. All VIs provided a high correlation with HA, and HA in alfalfa was modeled with moderate accuracy ($R^2 > 0.66$ in four harvests). These results follow similar trends from the previous studies evaluating the efficiency of HTP in predicting HA in small plots from alfalfa germplasm and breeding lines (Cazenave et al., 2019; Tang et al., 2021), as well as larger alfalfa plots (Feng et al., 2020).

The progress in plant breeding is measured based on genetic gain, which refers to the amount of increase in performance achieved through cycles of artificial selection (Xu et al., 2017). Several factors affect genetic gain, such as the genetic variation available in breeding populations, trait heritability, selection intensity, and the time required to complete a breeding cycle (Xu et al., 2017). Estimation of heritability can be improved by refining field experiments and statistical approaches, particularly for understanding and controlling spatial variation. One of our goals was to evaluate the effect of spatial models to control field variation and improve the estimation of genetic parameters and family selection. The autocorrelation had a significant impact across rows and columns. The spatial models improved the estimation of genetic and non-genetic parameters for HA in all harvests and NDVI in harvests one, two, and three. After applying spatial analysis, the heritability increased for both HA and NDVI. Similarly, Sripathi et al. (2017) and Andrade et al. (2020) reported high efficiency of spatial analysis in the estimation of genetic parameters in potato and forage breeding populations. These authors reported improvements in model fitness from the base and to spatial, which supports our results. The results presented in our study showed the importance of the spatial model to reduce the PEV and improve selection accuracy. These results reflected higher precision in the selection of the best families. The spatial models for HA and NDVI showed high levels of coincidence of selection in all harvests (>75%, except for NDVI in harvest three), compared to the base model.

Plant breeders can increase the selection intensity through improvements in the scale and precision of genotyping and phenotyping, which will result in higher genetic gain (Xu et al., 2017). One of the strategies to improve selection intensity is by increasing the breeding populations' size, but this comes at the expense of more efforts and resources dedicated to phenotyping. HTP can lead to higher genetic gain by increasing the size of breeding populations and making selections more accurately (Houle et al., 2010; Tang et al., 2021). In our study, H^2 estimates were slightly lower for NDVI than HA, but both traits showed significant genetic variation and moderate to low H^2. Considering only H^2, NDVI was able to detect the genetic variation present in this breeding population and can be used to select breeding lines exhibiting higher NDVI values, which would translate to breeding lines with higher HA ($R^2 > 0.66$). However, the coincidence of selection for the best families with HA and NDVI for both models was low to moderate (0.35 – 0.72), which shows that different families were selected by using NDVI and HA data in univariate models. Moreover, the genetic gain for HA was lower when the selection of the best 10% of the families was performed using only NDVI data. Our results indicated the NDVI data would complement ground-based HA measurements to improve genetic gain for HA in alfalfa.

Costs of field experiments are the limiting factor in alfalfa breeding programs focusing on quantifying HA across multiple harvests in a year and across multiple years and locations. The results presented in this study reported moderate to low H^2 for HA and high correlation coefficients between HA and NDVI across harvests. Multi-trait selection can be applied to take advantage of the correlation between traits and increase selection accuracy for the target trait (Mrode, 2014). HA and NDVI data combined into a bivariate model for each harvest showed a higher correlation among predicted values, a higher coincidence

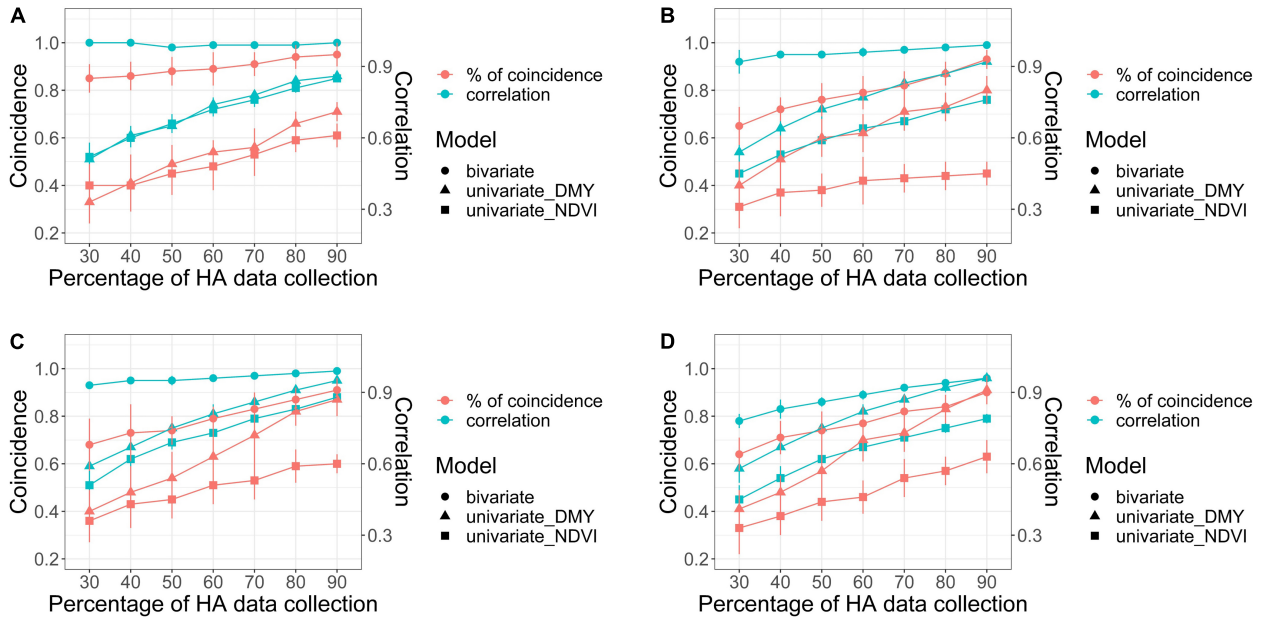

FIGURE 7 | Comparison of bivariate and univariate models (shape) for HA and NDVI for coincidence of selection after applying a 10% selection intensity (red), and correlation among breeding values for all families (green), for scenarios comparing models with increasing levels of HA data collection (30–90%) against a model with 100% HA data, across four harvests in alfalfa families evaluated in Citra, FL, United States. Harvests one **(A)**, two **(B)**, three **(C)**, and four **(D)**.

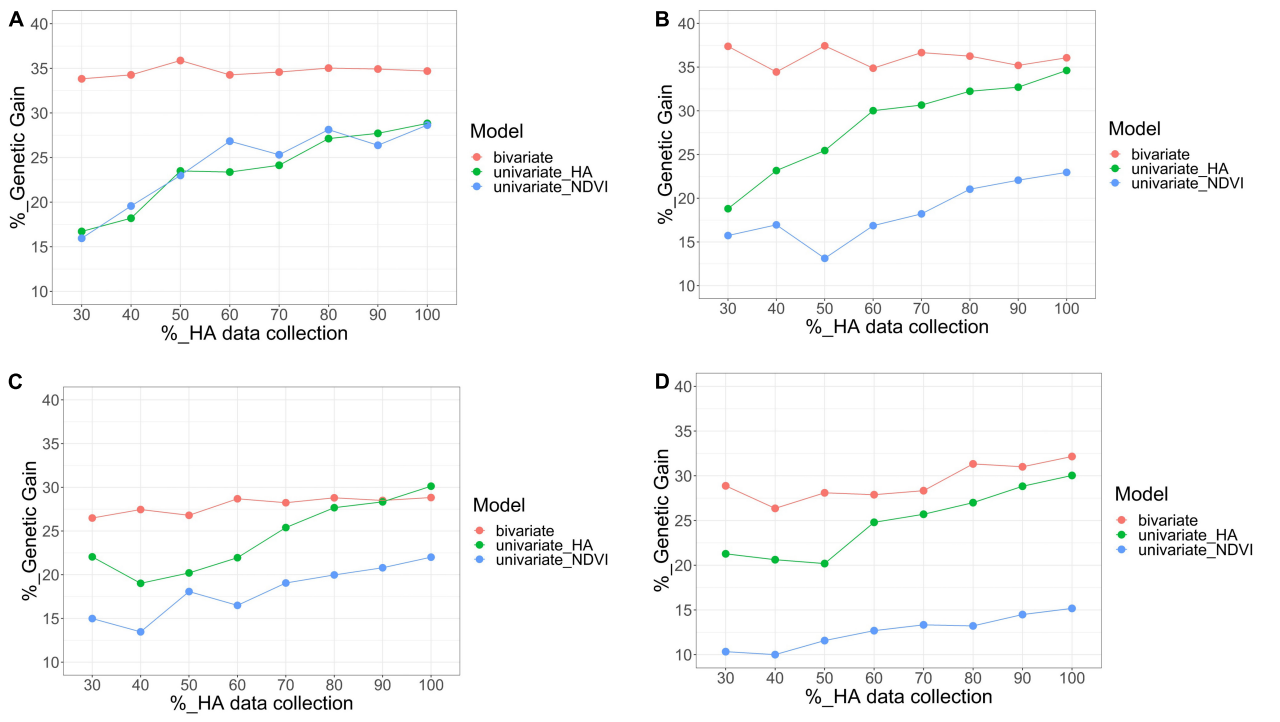

FIGURE 8 | Comparison of bivariate model (HA and NDVI, red), and univariate (HA, green; NDVI, blue) models for genetic gain for HA, after applying a 10% selection intensity for scenarios with increasing levels of HA data collection (30–100%) across four harvests in alfalfa families evaluated in Citra, FL, United States. Harvests one **(A)**, two **(B)**, three **(C)**, and four **(D)**.

of selection, and greater genetic gain than univariate models for HA and NDVI. As the level of HA data used in the models increased (from 30 to 90% of the total data), the correlation, coincidence of selection, and genetic gain increased. These results highlight the importance of collecting HTP data at the harvest time, particularly if breeders are not harvesting all experimental units in large breeding populations. To increase genetic gain for HA, alfalfa breeders could screen more breeding lines by combining HA and HTP phenotyping in their pipeline since the number of plots that need to be harvested will be smaller. In this study, reducing phenotyping efforts by 50% (using only 50% of the available HA data) showed a range between 0.83 and 1 for correlation among all families, 0.75 to 0.89 for the coincidence of selection, and 26.8 to 37.44% genetic gain in bivariate models. Despite the lower correlation and coincidence of selection, the genetic gain remained stable across all scenarios for bivariate models in all harvests. Including NDVI in the phenotyping pipeline for HA in alfalfa could result in greater genetic gains by increasing the size of breeding populations (Xu et al., 2017), while maintaining the resources for HA phenotyping constant.

High-throughput phenotyping is a promising method to develop improved cultivars and achieve high genetic gain. In this study, all VIs showed a high correlation with HA, and the inclusion of NDVI improved the selection accuracy for HA when bivariate models were fitted, even for scenarios with limited HA data. These results suggest that breeders could increase population size while maintaining the same ground-level measurement efforts, and expect increases in genetic gain due to a higher number of breeding candidates. Similar to the previous studies in alfalfa, HTP predicted HA with high accuracy (Feng et al., 2020), and HTP was able to detect differences in biomass production in large breeding populations (Cazenave et al., 2019). The results presented in this study coincide with the report from Tang et al. (2021), where HTP improved the efficiency of the selection process for alfalfa biomass in small plots (1.52 m × 0.30 m). Besides, it was also shown that spatial models controlled field variation and improved the estimation of genetic parameters and the accuracy of family selection.

Despite the improvements in the selection, HTP brings new challenges into the breeding pipeline. HTP data collection, storage, and processing require investments in computer power and storage and programming knowledge for data analysis and interpretation. In conclusion, the investment in time and resources to collect, process, and analyze HTP resulted in a more accurate selection of alfalfa families for HA. The RS data complemented ground-based HA measurements, and the combination of both datasets should result in improvements in alfalfa HA.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon request.

AUTHOR CONTRIBUTIONS

ER and PM designed the experiments. SS and AS collected the UAV images. JA, CS, YL, PM, and ER took the ground measurements. AB processed the images. AB and MA analyzed the data. GA and CS contributed to the data analysis on the ground measurements. AS, PM, and ER provided critical reviews and supported the costs of the study. AB, MA, and ER wrote the article. All authors provided critical reviews.

FUNDING

This research was partially funded by the USDA National Institute of Food and Agriculture, Hatch project 1018058, and by the Agriculture and Food Research Initiative grant number 2014-67013-22418 to PM, J. W. Olmstead and J. B. Endelman from the USDA National Institute of Food and Agriculture. This research was supported by the Plant Molecular Breeding Initiative grant provided through the University of Florida, Institute of Food and Agricultural Science, Plant Breeding Working Group.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.756768/full#supplementary-material

REFERENCES

Acharya, J. P., Lopez, Y., Gouveia, B. T., de Bem Oliveira, I., Resende, M. F. R., Muñoz, P. R., et al. (2020). Breeding Alfalfa (*Medicago sativa* L.) adapted to subtropical agroecosystems. *Agronomy* 10:742. doi: 10.3390/agronomy10050742

Adhikari, L., Makaju, S. O., and Missaoui, A. M. (2019). QTL mapping of flowering time and biomass yield in tetraploid alfalfa (*Medicago sativa* L.). *BMC Plant Biol.* 19:359. doi: 10.1186/s12870-019-1946-0

Andrade, M. H. M. L., Fernandes Filho, C. C., Fernandes, M. O., Bastos, A. J. R., Guedes, M. L., Marçal, T., et al. (2020). Accounting for spatial trends to increase the selection efficiency in potato breeding. *Crop Sci.* 60, 2354–2372. doi: 10.1002/csc2.20226

Annicchiarico, P. (2015). Alfalfa forage yield and leaf/stem ratio: narrow-sense heritability, genetic correlation, and parent selection procedures. *Euphytica* 205, 409–420. doi: 10.1007/s10681-015-1399-y

Annicchiarico, P., and Pecetti, L. (2021). Comparison among nine alfalfa breeding schemes based on actual biomass yield gains. *Crop Sci.* 61, 2355–2371. doi: 10.1002/csc2.20464

Annicchiarico, P., Nazzicari, N., and Brummer, E. C. (2016). *Alfalfa Genomic Selection: Challenges, Strategies, Transnational Cooperation. Breeding in a World of Scarcity*. Berlin: Springer International Publishing. doi: 10.1007/978-3-319-28932-8_22

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61.

Benabderrahim, M. A., Guiza, M., and Haddad, M. (2020). Genetic diversity of salt tolerance in tetraploid alfalfa (*Medicago sativa* L.). *Acta Physiol. Plant.* 42, 1–11. doi: 10.1007/s11738-019-2993-8

Bingham, E. T., Groose, R. W., Woodfield, D. R., and Kidwell, K. K. (1994). Complementary Gene Interactions in Alfalfa are Greater in Autotetraploids than Diploids. *Crop Sci.* 34, 823–829. doi: 10.2135/cropsci1994.0011183X003400040001x

Bowley, S. R., and Christie, B. R. (1981). Inheritance of dry matter yield in a heterozygous population of Alfalfa. *Can. J. Plant Sci.* 61, 313–318. doi: 10.4141/cjps81-044

Brummer, E. C. (1999). Capturing heterosis in forage crop cultivar development. Capturing heterosis in forage crop cultivar development. *Crop Sci.* 39, 943–954. doi: 10.2135/cropsci1999.0011183X003900040001x

Brummer, E. C., and Casler, M. D. (2015). "Cool-Season Forages," in *Yield Gains in Major U.S. Field Crops*, eds S. Smith, B. Diers, J. Specht, and B. Carver (Hoboken, NY: John Wiley & Sons), doi: 10.2135/cssaspecpub33.c3

Butler, D. G., Cullis, B. R., Gilmour, A. R., and Gogel, B. J. (2009). *ASReml-R reference manual*. Brisbane, Qld: Department of Primary Industries and Fisheries.

Casler, M. D., and Brummer, E. C. (2008). Theoretical expected genetic gains for among-and-within-family selection methods in perennial forage crops. *Crop Sci.* 48, 890–902. doi: 10.2135/cropsci2007.09.0499

Cazenave, A., Shah, K., Trammell, T., Komp, M., Hoffman, J., Motes, C. M., et al. (2019). High-throughput approaches for phenotyping alfalfa germplasm under abiotic stress in the field. *Plant Phenome J.* 2, 1–13. doi: 10.2135/tppj2019.03.0005

De Assis, G. M. L., Ruggieri, A. C., Mercadante, M. E. Z., De Camargo, G. M. F., and Carneiro Júnior, J. M. (2010). Selection of alfalfa cultivars adapted for tropical environments with repeated measures using PROC MIXED of SAS® System. *Plant Genet. Resour.* 8, 55–62. doi: 10.1017/S1479262109990153

Dos Santos, I. G., Cruz, C. D., Nascimento, M., Rosado, R. D. S., and De Paula Ferreira, R. (2018). Direct, indirect and simultaneous selection as strategies for alfalfa breeding on forage yield and nutritive value. *Pesqui. Agropecu. Trop.* 48, 178–189. doi: 10.1590/1983-40632018v4851950

Feng, L., Zhang, Z., Ma, Y., Du, Q., Williams, P., Drewry, J., et al. (2020). Alfalfa yield prediction using UAV-based hyperspectral imagery and ensemble learning. *Remote Sens.* 12:2028. doi: 10.3390/rs12122028

Fu, Y. B. (2015). Understanding crop genetic diversity under modern plant breeding. *Theor. Appl. Genet.* 128, 2131–2142. doi: 10.1007/s00122-015-2585-y

Furbank, R. T., and Tester, M. (2011). Phenomics-technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Galli, G., Horne, D. W., Collins, S. D., Jung, J., Chang, A., Fritsche-Neto, R., et al. (2020). Optimization of UAS-based high-throughput phenotyping to estimate plant health and grain yield in sorghum. *Plant Phenome J.* 3:e20010. doi: 10.1002/ppj2.20010

Gilmour, A. R., Cullis, B. R., and Verbyla, A. P. (1997). Accounting for natural and extraneous variation in the analysis of field experiments. *J. Agric. Biol. Environ. Stat.* 2, 269–293. doi: 10.2307/1400446

He, F., Long, R., Zhang, T., Zhang, F., Wang, Z., Yang, X., et al. (2020). Quantitative trait locus mapping of yield and plant height in autotetraploid alfalfa (*Medicago sativa* L.). *Crop J.* 8, 812–818. doi: 10.1016/j.cj.2020.05.003

Houle, D., Govindaraju, D. R., and Omholt, S. (2010). Phenomics: the next challenge. *Nat. Rev. Genet.* 11, 855–866. doi: 10.1038/nrg2897

Li, C., Alseekh, S., Lobos, G. A., Camargo, A. V., Del Pozo, A., Araus, J. L., et al. (2017). Plant phenotyping and phenomics for plant breeding. *Front. Plant Sci.* 8:2181. doi: 10.3389/fpls.2017.02181

Lima-Cueto, F. J., Blanco-Sepúlveda, R., Gómez-Moreno, M. L., and Galacho-Jiménez, F. B. (2019). Using vegetation indices and a uav imaging platform to quantify the density of vegetation ground cover in olive groves (*Olea Europaea L.*) in Southern Spain. *Remote Sens.* 11:2564. doi: 10.3390/RS11212564

Maes, W. H., and Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci.* 24, 152–164. doi: 10.1016/j.tplants.2018.11.007

Mrode, R. A. (2014). *Linear Models For The Prediction Of Animal Breeding Values*. Wallingford: Cabi.

Piepho, H. P., Möhring, J., Melchinger, A. E., and Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161, 209–228. doi: 10.1007/s10681-007-9449-8

Quirós Vargas, J. J., Zhang, C., Smitchger, J. A., McGee, R. J., and Sankaran, S. (2019). Phenotyping of plant biomass and performance traits using remote sensing techniques in pea (*Pisum sativum. L.*). *Sensors* 19:2031. doi: 10.3390/S19092031

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org/

Ranjan, R., Chandel, A. K., Khot, L. R., Bahlol, H. Y., Zhou, J., Boydston, R. A., et al. (2019). Irrigated pinto bean crop stress and yield assessment using ground based low altitude remote sensing technology. *Inform. Process. Agric.* 6, 502–514. doi: 10.1016/J.INPA.2019.01.005

Ren, L., Bennett, J. A., Coulman, B., Liu, J., and Biligetu, B. (2021). Forage yield trend of alfalfa cultivars in the Canadian prairies and its relation to environmental factors and harvest management. *Grass Forage Sci.* 76, 390–399. doi: 10.1111/gfs.12513

Riday, H., and Brummer, E. C. (2002). Forage yield heterosis in alfalfa. *Crop Sci.* 42, 716–723.1.

Sakiroglu, M., and Brummer, E. C. (2017). Identification of loci controlling forage yield and nutritive value in diploid alfalfa using GBS-GWAS. *Theor. Appl. Genet.* 130, 261–268. doi: 10.1007/s00122-016-2782-3

Sripathi, R., Conaghan, P., Grogan, D., and Casler, M. D. (2017). Spatial variability effects on precision and power of forage yield estimation. *Crop Sci.* 57, 1383–1393. doi: 10.2135/cropsci2016.08.0645

Tang, Z., Parajuli, A., Chen, C. J., Hu, Y., Revolinski, S., Medina, C. A., et al. (2021). Validation of UAV-based alfalfa biomass predictability using photogrammetry with fully automatic plot segmentation. *Sci. Rep.* 11:3336. doi: 10.1038/s41598-021-82797-x

Tattaris, M., Reynolds, M. P., and Chapman, S. C. (2016). A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding. *Front. Plant Sci.* 7:1131. doi: 10.3389/fpls.2016.01131

USDA-NASS (2020). *USDA-National Agricultural Statistics Service*. Available online at: www.nass.usda.gov/ (accessed June 21, 2021).

Vivela, D., Basigalup, D. H., Juntolli, F. V., and Ferreira de Paula, R. (2018). "Research priorities and future of alfalfa in Latin America," in *Proceedings of the Second World Alfalfa Congress Global Interaction for Alfalfa Innovation. 11-14 November*, *Cordoba, Argentina* Vol. 2018, 140–143.

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crops Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003

Xu, Y., Li, P., Zou, C., Lu, Y., Xie, C., Zhang, X., et al. (2017). Enhancing genetic gain in the era of molecular breeding. *J. Exp. Bot.* 68, 2641–2666. doi: 10.1093/jxb/erx135

Zhang, X., Zhang, F., Qi, Y., Deng, L., Wang, X., and Yang, S. (2019). New research methods for vegetation information extraction based on visible light remote sensing images from an unmanned aerial vehicle (UAV). *Int. J. Appl. Earth Observ. Geoinform.* 78, 215–226. doi: 10.1016/J.JAG.2019.01.001

Zhao, C., Zhang, Y., Du, J., Guo, X., Wen, W., Gu, S., et al. (2019). Crop phenomics: current status and perspectives. *Front. Plant Sci.* 10:714. doi: 10.3389/fpls.2019.00714

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 20 December 2021
doi: 10.3389/fpls.2021.730181

Entropy Weight Ensemble Framework for Yield Prediction of Winter Wheat Under Different Water Stress Treatments Using Unmanned Aerial Vehicle-Based Multispectral and Thermal Data

Shuaipeng Fei[1,2,3], Muhammad Adeel Hassan[2,4], Yuntao Ma[3], Meiyan Shu[3], Qian Cheng[1], Zongpeng Li[1], Zhen Chen[1] and Yonggui Xiao[2]**

[1] Institute of Farmland Irrigation, Chinese Academy of Agricultural Sciences, Xinxiang, China, [2] National Wheat Improvement Center, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China, [3] College of Land Science and Technology, China Agricultural University, Beijing, China, [4] Dezhou Academy of Agricultural Sciences, Dezhou, China

OPEN ACCESS

Edited by:
Wenting Han,
Northwest A&F University, China

Reviewed by:
Jibo Yue,
Henan Agricultural University, China
Shouyang Liu,
Nanjing Agricultural University, China
Awais Rasheed,
Quaid-i-Azam University, Pakistan

***Correspondence:**
Zhen Chen
chenzhen@caas.cn
Yonggui Xiao
xiaoyonggui@caas.cn

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

Received: *24 June 2021*
Accepted: *08 November 2021*
Published: *20 December 2021*

Citation:
Fei S, Hassan MA, Ma Y, Shu M, Cheng Q, Li Z, Chen Z and Xiao Y (2021) Entropy Weight Ensemble Framework for Yield Prediction of Winter Wheat Under Different Water Stress Treatments Using Unmanned Aerial Vehicle-Based Multispectral and Thermal Data. Front. Plant Sci. 12:730181. doi: 10.3389/fpls.2021.730181

Crop breeding programs generally perform early field assessments of candidate selection based on primary traits such as grain yield (GY). The traditional methods of yield assessment are costly, inefficient, and considered a bottleneck in modern precision agriculture. Recent advances in an unmanned aerial vehicle (UAV) and development of sensors have opened a new avenue for data acquisition cost-effectively and rapidly. We evaluated UAV-based multispectral and thermal images for in-season GY prediction using 30 winter wheat genotypes under 3 water treatments. For this, multispectral vegetation indices (VIs) and normalized relative canopy temperature (NRCT) were calculated and selected by the gray relational analysis (GRA) at each growth stage, i.e., jointing, booting, heading, flowering, grain filling, and maturity to reduce the data dimension. The elastic net regression (ENR) was developed by using selected features as input variables for yield prediction, whereas the entropy weight fusion (EWF) method was used to combine the predicted GY values from multiple growth stages. In our results, the fusion of dual-sensor data showed high yield prediction accuracy [coefficient of determination (R^2) = 0.527–0.667] compared to using a single multispectral sensor (R^2 = 0.130–0.461). Results showed that the grain filling stage was the optimal stage to predict GY with R^2 = 0.667, root mean square error (RMSE) = 0.881 t ha^{-1}, relative root-mean-square error (RRMSE) = 15.2%, and mean absolute error (MAE) = 0.721 t ha^{-1}. The EWF model outperformed at all the individual growth stages with R^2 varying from 0.677 to 0.729. The best prediction result (R^2 = 0.729, RMSE = 0.831 t ha^{-1}, RRMSE = 14.3%, and MAE = 0.684 t ha^{-1}) was achieved through combining the predicted values of all growth stages. This study suggests that the fusion of UAV-based multispectral and thermal IR data within an ENR-EWF framework can provide a precise and robust prediction of wheat yield.

Keywords: UAV, multispectral indices, machine learning, remote sensing, thermal infrared, wheat yield

INTRODUCTION

Bread wheat is one of the most important food crops that feed 40% of the world population (Liu et al., 2020). The timely and accurate evaluation of the grain yield (GY) before harvest can aid the selection of elite genotypes in large breeding programs (Mcbratney et al., 2005; Panda et al., 2010). Yield advocating traits, such as green biomass, leaf area index (LAI), and chlorophyll contents, have been used for within-season yield prediction (Hassan et al., 2018, 2019a). The canopy temperature is another important indicator of transpiration and leaf water potential under drought and heat stress and can help facilitate the selection of resilient genotypes (Zubler and Yoon, 2020). However, phenotyping most of these traits is destructive, time-consuming, and is associated with a high error probability. Therefore, the nondestructive measurements of the above proxy traits of the GY have been employed to increase the prediction accuracy of crop yield cost-effectively (Yu et al., 2016; Elsayed et al., 2017; Hassan et al., 2019a).

In the past few years, low-altitude remote sensing has attracted interest for its application in high-throughput crop phenotyping (Hassan et al., 2019b; Maimaitijiang et al., 2020). The advances in sensor technology have significantly accelerated the use of unmanned aerial vehicles (UAVs) for data collection with high spectral resolution as compared to satellite platforms (Colomina and Molina, 2014; Sidike et al., 2018). Various types of sensors mounted on UAV platforms, such as multispectral, hyperspectral, RGB, and thermal, are being widely used in the phenotypic evaluation of crops, with satisfactory data accuracy. The UAV-based nondestructive multispectral assessments of the LAI (Comba et al., 2020), biomass (Yue et al., 2019), chlorophyll content (Qiao et al., 2020), nitrogen use efficiency (Yang et al., 2020), senescence (Hassan et al., 2021), and GY (Hassan et al., 2019a) have been reported for several crops. These assessments are based on the spectral reflectance from the canopy of plants in the form of light bands with different wavelengths (Li et al., 2014). Thermal remote sensing is also being applied in precision agriculture to detect water stress (Suyoung et al., 2017) and plant resistance (Ludovisi et al., 2017). Recently, the focus has been increased on combining the data from multiple sources, where a group of datasets from multiple sensors is utilized obtained for plant trait estimation. Multi-source data models have the capability to improve crop trait estimations (Maimaitijiang et al., 2020). The use of canopy temperature and spectral information have been demonstrated to improve the model performance in estimating important plant traits for assessment of biotic/abiotic stress (Appeltans et al., 2020; Zubler and Yoon, 2020) and predicting the yield of soybean (Elmetwalli et al., 2020), barely (Rischbeck et al., 2014), and maize (Zhang et al., 2020). For crop yield prediction, flowering to grain filling stages are highly reliable, with good accuracy and repeatability (Hassan et al., 2019a; Hernandez et al., 2015). The predictions made in most studies have been based on the spectral information of an individual growth stage. The accumulation of VIs from jointing to the grain filling stage using a multiple linear regression algorithm has shown good prediction results in rice (Zhou et al., 2017). Since UAV-based temporal information of multispectral vegetation indices (VIs) and temperature can be obtained cost-effectively from multiple growth stages, combining data across the growth stages could help to achieve higher yield prediction accuracy. Machine learning algorithms have been employed with the canopy spectral features as input to construct models for crop trait evaluation, showing high prediction accuracy and adaptability (Wang et al., 2016; Wang J. et al., 2018). The commonly used machine learning algorithms are the random forest (RF) (Breiman, 2001), support vector machine (SVM) (Sain, 1997), and artificial neural network (ANN) (Bradley, 1995), and these have been successfully used for estimating biomass (Wang et al., 2016), LAI (Wang L. et al., 2018), chlorophyll content (Shah et al., 2019), and water content (Tavakoli and Gebbers, 2019). Among the machine learning algorithms, the emerging elastic net regression (ENR) algorithm has been considered one of the most precise prediction method for regression problems (Hui and Hastie, 2005). The ENR algorithm combines the advantages of ridge regression and least absolute shrinkage and selection operator (LASSO) regression to obtain better prediction results (Ogutu et al., 2012). At present, relatively few studies have been conducted on utilizing information obtained by UAV-based sensors as input to the ENR algorithm for the yield prediction of winter wheat.

The entropy weight algorithm is an emerging method in agricultural studies. It works by allocating the weight-based information entropy of the trait in the model (Li et al., 2011). It has been typically used for feature selection and model combination for combining datasets to assess ecosystem health (Cheng et al., 2020), monitor land-use change (Lu et al., 2014), and evaluate the coverage effectiveness of remote sensing satellites (Li et al., 2018). To the best of our knowledge, the entropy weight method has not been used to predict the yield values from multiple growth stages using UAV datasets. The objectives of this study were (1) to evaluate the potential of UAV-based multispectral and thermal sensors for the yield prediction of wheat using the ENR algorithm, (2) to identify the appropriate wheat growth stage for data collection to maximize the yield prediction accuracy, and (3) to investigate the potential of the entropy weight method in combining the predicted GY values from multiple growth stages.

MATERIALS AND METHODS

Germplasm and Experimental Design

Field trials were conducted at the experimental station of the Institute of Farmland Irrigation of Chinese Academy of Agricultural Sciences in Xinxiang (113.8°E, 35.2°N) during the 2019–2020 cropping season (**Figure 1**). In total, 30 winter wheat varieties widely cultivated in the Yellow and Huai Valleys Winter Wheat Zone of China were used in this experiment. Germplasm was planted under three water stress treatments, namely, mild irrigation, moderate irrigation, and high irrigation, to obtain the UAV-based multispectral, thermal, and ground-truth GY data. Irrigation for each treatment was performed in the tillering, wintering, reviving, jointing, heading, and grain filling stages using a laterally moving sprinkler irrigation machine. The

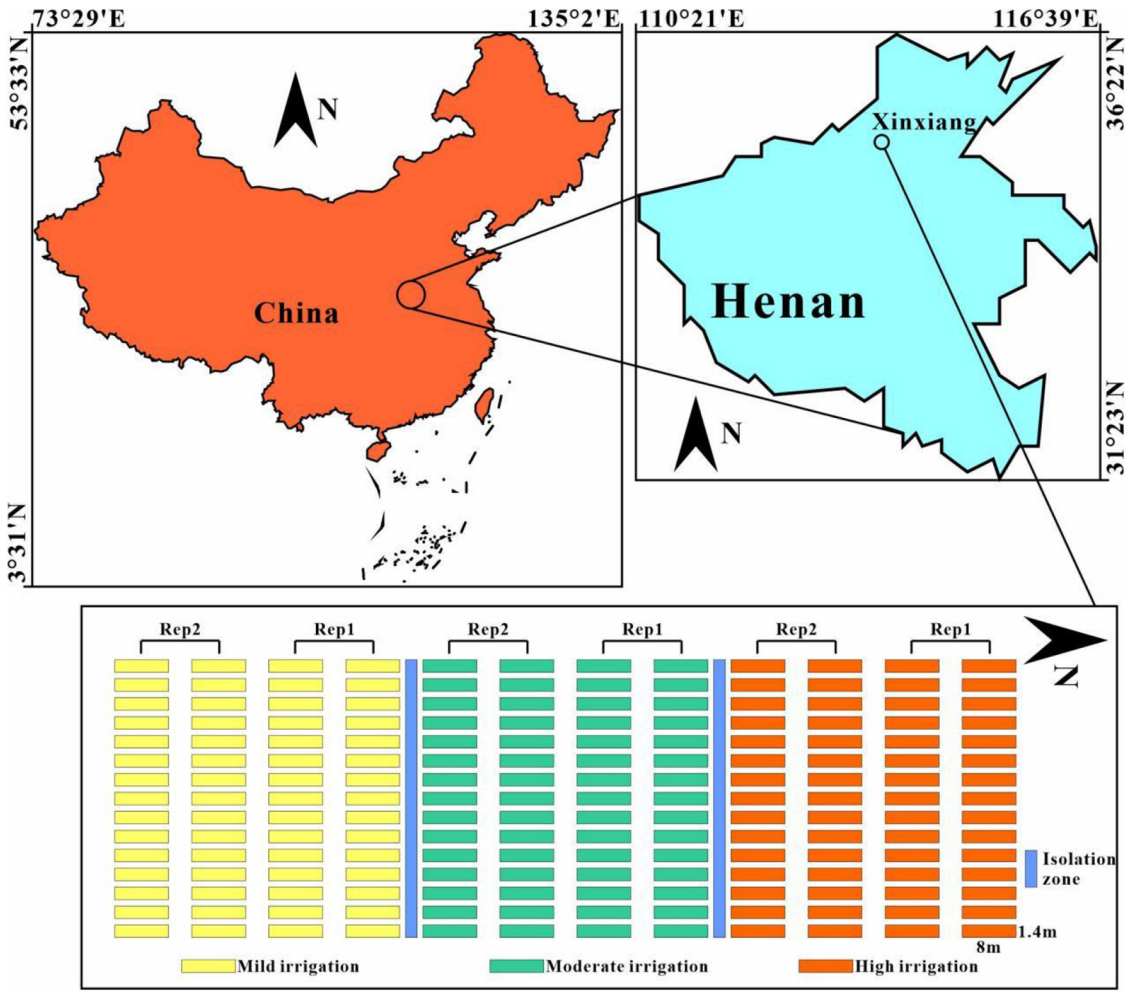

FIGURE 1 | Experimental location, design, and management.

irrigation volume was calculated by the flow rate of the sprinkler nozzle and the duration of irrigation. The total irrigation volume for the mild, moderate, and high irrigation treatments were 145, 190, and 240 mm, respectively (**Table 1**). A completely randomized block design with two replications was adopted for the experiment. The size of each plot was maintained at 11.2 m^2 with the dimensions of 8 m × 1.4 m, representing one cultivar with six rows at a spacing of 0.20 m. Field management (e.g., disease and pest control, fertilizer) was maintained at optimal levels depending on the local conditions. In the 2019–2020 growing season, the total precipitation was 115 mm, and the monthly average temperature was highest (23°C) in July and lowest (−6°C) in January. Wheat was harvested using a plot combine harvester in June 2020. The GY of each plot was weighed at a moisture content of approximately 12.5%.

Data Acquisition and Processing

Figure 2 shows the workflow for the data acquisition. A DJI M210 (DJI Technology Co., Shenzhen, China) carrying a RedEdge MX (MicaSense Parrot, France) multispectral camera and a Zenmuse XT2 (DJI Technology Co., Shenzhen, China) thermal sensor was used to collect high-resolution multispectral and thermal images simultaneously. The RedEdge MX featured five spectral sensors, namely, blue (475 nm), green (560 nm), red (668 nm), red-edge (717 nm), and near-IR (842 nm). The RedEdge MX camera automatically adjusts the ambient light effects through the sunshine sensor, thereby minimizing the

TABLE 1 | Irrigation strategy for each treatment.

Growth stage	Mild irrigation (mm)	Moderate irrigation (mm)	High irrigation (mm)
Tillering	35	35	35
Wintering	35	35	35
Turning green	20	25	35
Jointing	20	35	50
Heading	20	35	50
Grain filling	15	25	35
Total	145	190	240

error in the multispectral images. Zenmuse XT2 contains an 8-mm lens with a 57.12° × 42.44° field of view to record temperature measurements in the 7.5–13.5-μm spectral range with a measurement accuracy of ± 5°C. The DJI ground station was used as an automated flight control system, allowing the user to define the air route and customize the mission plan. Flight mission was executed for all the six growth stages from 11 a.m. to 1 p.m. under a cloudless sky. To avoid the effect of the phenological differences between treatments, the flight missions for each treatment were collected according to the growth stages. To obtain high-resolution images, each flight was set at an altitude of 30 m with 85% front and 80% side image overlapping. Before and after each flight, the calibration board was photographed to convert the digital number (DN) value of the multispectral image into reflectance during subsequent data processing. During the flights, the surface temperature of 12 boards was measured using a handheld thermometer for the radiometric calibration of the thermal images. To obtain the geographic reference of the multisensor UAV image, 18 ground control points (GCPs) were evenly arranged in the field, and their coordinates were measured with a millimeter-level accuracy using a differential global positioning system.

The Pix4Dmapper software (Pix4D SA, Lausanne, Switzerland) was employed for the orthomosaic generation using the UAV-based multispectral and thermal IR images. The geographic coordinates (World Geodetic System, 1984) of the GCPs were used in the photogrammetric workflow of Pix4Dmapper to improve the accuracy of the composite orthomosaics. Dense point clouds were generated using the structure-from-motion (SfM) method in Pix4Dmapper along with the photogrammetric workflow. After radiometric correction, the DN values of the multispectral and thermal IR images were converted to reflectance and temperature (°C). To extract the reflectance and temperature information for each plot, the orthomosaic images were segmented into 180 polygon shapes with assigned IDs defining the cultivars under different irrigation treatments. Polygon shape generation and information extraction are completed in QGIS 3.1.0.[1] A total of 22 indices were used in this study, of which 21 VIs were estimated from multispectral reflectance, and 1 index was calculated from the canopy temperature across the irrigation treatments (**Table 2**).

Gray Relational Analysis

In a gray relational analysis (GRA), a system with incomplete information is called a gray system, meaning that the relationship between the factors is uncertain (Aslan et al., 2012). When the experiment is unclear or when the experimental method cannot be implemented accurately, a gray analysis can help overcome the drawbacks in statistical regression (Jin et al., 2013). For example, there is a close relationship between VIs

[1]https://www.qgis.org/

TABLE 2 | Formulae of multispectral vegetation indices and normalized relative canopy temperature.

Acronym	Index	Formulae	Developer(s)
CIRE	Chlorophyll index RedEdge	$(R_{NIR}/R_{RE})-1$	Gitelson et al., 2003
DVI	Difference vegetation index	$R_{NIR}-R_R$	Tucker et al., 1979
EVI	Enhanced vegetation index	$2.5 \times (R_{NIR}-R_R)/(1R_{NIR}+6 \times R_R-7.5 \times R_B)$	Huete et al., 2002
GNDVI	Green normalized difference vegetation index	$(R_{NIR}-R_G)/(R_{NIR}+R_G)$	Gitelson et al., 1996
MCARI1	Modified chlorophyll absorption in reflectance index 1	$(R_{REG}-R_R-0.2 \times ((R_{REG}-R_G)) \times (R_{REG}/R_R)$	Daughtry et al., 2000
PSRI	Plant senescence reflectance index	$(R_R-R_B)/R_{NIR}$	Merzlyak et al., 1999
MSR	Modified simple ratio index	$((R_{NIR}/R_R)-1)/\sqrt{R_{NIR}/R_R 1}$	Chen, 1996
MTCI	MERIS terrestrial chlorophyll index	$(R_{NIR}-R_{REG})/(R_{RE}-R_R)$	Dash and Curran, 2004
MTVI2	Modified triangular vegetation index 2	$1.5 \times [1.2 \times (R_{NIR}-R_G)-2.5 \times (R_R-R_G)]/\left[\sqrt{2 \times (R_{NIR}+1)^2-6 \times R_{NIR}+5 \times \sqrt{R_R}}-0.5\right]$	Haboudane et al., 2004
NDVI	Normalized difference vegetation index	$(R_{NIR}-R_R)/(R_{NIR}+R_R)$	Rouse, 1972
NDVIRE	Normalized difference vegetation index RedEdge	$(R_{NIR}-R_{REG})/(R_{NIR}+R_{REG})$	Elsayed et al., 2015
NLI	Nonlinear vegetation index	$(R_{NIR} \times R_{NIR}-R_R)/(R_{NIR} \times R_{NIR}+R_R)$	Goel and Qin, 1994
OSAVI	Optimized soil-adjusted vegetation index	$(R_{NIR}-R_R)/(R_{NIR}+R_R+1.6) \times 1.16$	Rondeaux et al., 1996
PPR	Plant pigment ratio	$(R_G-R_B)/(R_G+R_B)$	Metternicht, 2003
RDVI	Re-normalized difference vegetation index	$(R_{NIR}-R_R)/\sqrt{(R_{NIR}+R_R)}$	Wang et al., 1998
RVI	Ratio vegetation index	(R_{NIR}/R_R)	Ba Ret and Guyot, 1991
NRI	Nitrogen reflectance index	$(R_G-R_R)/(R_G+R_R)$	Diker and Bausch, 2003
SAVI	Soil-adjusted vegetation index	$(R_{NIR}-R_R)/(R_{NIR}+R_R+0.5) \times 1.5$	Huete, 1988
SIPI	Structure insensitive pigment index	$(R_{NIR}-R_B)/(R_{NIR}+R_B)$	Penuelas et al., 1995
TCARI	Transformed chlorophyll absorption ratio index	$3[(R_{REG}-R_R)-0.2 \times (R_{REG}-R_G) \times R_{REG}/R_R]$	Eitel et al., 2007
TVI	Triangular vegetation index	$0.5 \times [120 \times (R_{NIR}-R_G)-200 \times (R_R-R_G)]$	Broge and Leblanc, 2001
NRCT	Normalized relative canopy temperature	$\frac{T_i-T_{min}}{T_{max}-T_{min}}$	Elsayed et al., 2015

R_B, R_G, R_R, R_{REG}, and R_{NIR} represent the reflectance of the blue, green, red, red-edge, and near-IR bands of RedEdge MX, respectively. T represents the canopy temperature obtained from Zenmuse XT2.

and yield; however, their detailed relationships remain unclear. Therefore, the main purpose of the GRA is to measure the degree of relationship within this system by analyzing the gray relationships between VIs and GY. The GRA procedure includes the following steps:

1. The reference series reflect the characteristics of the system behavior, and the comparison series influences the system behavior. In this study, the GY was considered the reference series, and each index was considered a comparison series. The reference sequence is represented by the following formula:

$$X_0 = X_0(k)|k = 1, 2...n \tag{1}$$

where n represents the number of samples, and n is 180 in this study. Comparison data series can be expressed as follows:

$$X_i = X_i(k)|k = 1, 2...n,\ i = 1, 2...m \tag{2}$$

There are m comparison data series, each containing n-values.

2. Data in each factor column in the system may have different dimensions, making it difficult to compare or obtain a correct conclusion when comparing. Therefore, to ensure the reliability of the results, the following non-dimensional processing of the data is generally required when performing the GRA:

$$x_i(k) = \frac{X_i(k)}{X_i(l)} \tag{3}$$

$$X_i(l) = \frac{1}{n}\sum_{k=1}^{n} X_i(k) \tag{4}$$

3. The calculation of the difference data series $\triangle_i$ is as follows:

$$\triangle_i = \left(\left|x_{01,} - x_{i1}\right|, \left|x_{02,} - x_{i2}\right|, \ldots \left|x_{0n,} - x_{in}\right|\right) \tag{5}$$

4. The gray relational coefficient $\xi_i(k)$ for the kth data point in the ith difference data series can be expressed as follows:

$$\xi_i(k) = \frac{\triangle_{\min} + \zeta\triangle_{\max}}{\triangle_{i(k)} + \zeta\triangle_{\max}} \tag{6}$$

where $\triangle_{\min}$ and $\triangle_{\max}$ are the global maximum and minimum values in the difference data series, respectively. $\triangle_{i(k)}$ is the kth value in the $\triangle_i$ difference data series, and ζ is the distinguishing coefficient: $\zeta \in [0, 1]$. In this study, the distinguishing coefficient is set to 0.5.

5. Generally, the average value of the gray relational coefficient is taken as the gray relational degree (GRD), which is expressed as follows:

$$\gamma_i = \frac{1}{n}\sum_{k=1}^{n} \xi_i(K). \tag{7}$$

Elastic Net Regression

To avoid the instability of the LASSO solutions when the input features are highly correlated (e.g., a large number of VIs constructed from limited bands), the ENR has been proposed to analyze the high-dimensional data. The ENR is an extension of the LASSO, which is robust to severe multicollinearity among the input features (Ogutu et al., 2012). The ENR combines the penalties of the ridge regression (ℓ_1) and LASSO (ℓ_2) and can be expressed as follows:

$$\widehat{\beta}(\text{enet}) = \left(1 + \frac{\lambda_2}{n}\right)\left\{\underset{\beta}{\arg\min}||y - X\beta||_2^2 + \lambda_2||\beta||_2^2 + \lambda_1||\beta||_1\right\} \tag{8}$$

On setting $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, the ENR is equivalent to the minimizer of the following:

$$\widehat{\beta}(\text{enet2}) = \underset{\beta}{\arg\min}||y - X\beta||_2^2 \tag{9}$$

subject to $P_\alpha(\beta) = (1-\alpha)||\beta||_1 + \alpha||\beta||_2^2 \leq s$ for some s, where $P_\alpha(\beta)$ represents the penalty of the ENR. The ENR can be considered LASSO and ridge regression when $a = 0$ and 1, respectively. The $\ell 1$ part of the ENR is used for automatic variable selection, while the $\ell 2$ part encourages grouped selection and stabilizes the solution paths with respect to the random sampling, thereby improving the prediction results. By introducing a grouping effect when selecting the variable, a group of highly correlated input features tends to have similar coefficients. The ENR can choose the groups of correlated features when these groups are unknown in advance. Notably, the ENR selects more than n variables when $p >> n$, which is different from the LASSO. In this study, there is inevitably a high correlation between the various VIs. Therefore, the ENR will be an ideal choice when using VIs as the input features for yield prediction.

Modeling Framework

In this study, a 10-fold outer cross-validation method was adopted to train and test the model. To avoid contingency, we conducted 50 iterations for the outer cross-validation, resulting in a total of 500 models. The average of the accuracy evaluation index generated from the 500 models was used to evaluate the model performance. In the process of outer cross-validation, the inner cross-validation and the grid search were conducted for parameter tuning of the ENR models (**Figure 2**). In the outer cross-validation, the VIs and GY data were randomly divided into 10 equal subsets. One of them was used for testing each time, and the remaining nine subsets were used for training. Each training set of the outer cross-validation was evenly divided into 10 sets, similar to the outer cross-validation. One of them was used for testing, and the nine subsets were used for training. During the inner cross-validation process, multiple combinations of the candidate parameters were set in the inner training set for model construction and then tested on the inner test set. Each parameter combination was tested 10 times, and the hyperparameter combination with the lowest average test error was set for the outer cross-validation for model training. This study uses the R package "caret"[2] to construct the ENR model

[2]https://CRAN.R-project.org/package=caret

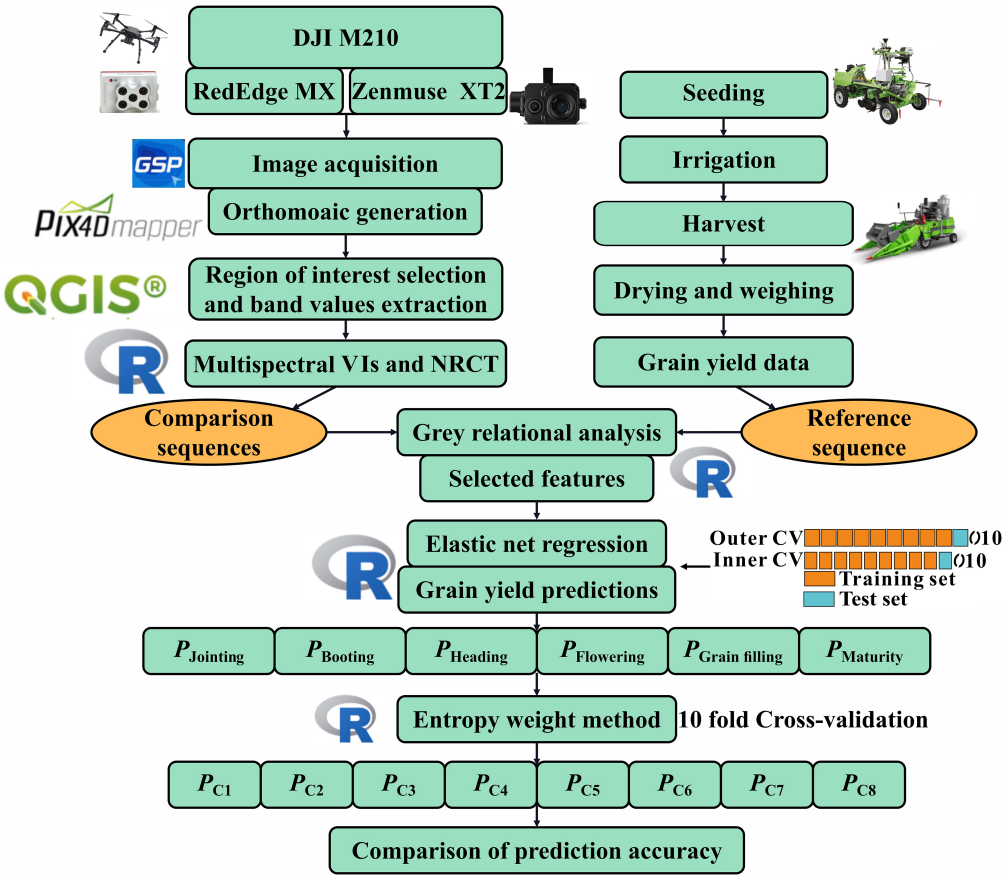

FIGURE 2 | Schematic workflow of the methodology used in this study. *P* denotes the predicted grain yield (GY) value, and C1–C8 indicate the combinations of the values predicted from multiple growth stages. CV, cross-validation; VIs, vegetation indices; NRCT, normalized relative canopy temperature.

for yield prediction. In the "caret" package, the parameters to be adjusted are the *fraction* and quadratic penalty parameter *lambda*. **Table 3** represents the candidate ranges of these two parameters.

TABLE 3 | Candidate hyperparameters for elastic net regression.

Number	*Lambda*	*Fraction*	Number	*Lambda*	*Fraction*
1	0.050	0.000E+00	16	0.541	3.162E-03
2	0.083	1.000E-04	17	0.574	4.047E-03
3	0.116	1.280E-04	18	0.607	5.179E-03
4	0.148	1.638E-04	19	0.640	6.629E-03
5	0.181	2.096E-04	20	0.672	8.483E-03
6	0.214	2.683E-04	21	0.705	1.086E-02
7	0.247	3.433E-04	22	0.738	1.389E-02
8	0.279	4.394E-04	23	0.771	1.778E-02
9	0.312	5.623E-04	24	0.803	2.276E-02
10	0.345	7.197E-04	25	0.836	2.913E-02
11	0.378	9.211E-04	26	0.869	3.728E-02
12	0.410	1.179E-03	27	0.902	4.771E-02
13	0.443	1.509E-03	28	0.934	6.105E-02
14	0.476	1.931E-03	29	0.967	7.814E-02
15	0.509	2.471E-03	30	1.000	1.000E-01

Moreover, we tested the model performance on the test samples in the cross-validation procedure to test the adaptability of the model. The coefficient of determination (R^2), root mean square error (RMSE), relative root-mean-square error (RRMSE), and mean absolute error (MAE) were adopted to evaluate the model performance. The calculation formulae of the parameters are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{10}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{11}$$

$$\mathrm{RRMSE} = \frac{\mathrm{RMSE}}{\bar{y}} * 100\% \tag{12}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{13}$$

where n represents the number of samples, y_i and $\hat{y}_i$ are the measured and predicted GY of sample i, and $\bar{y}$ is the average value of the measured GY. The higher the R^2-value, the lower the

RMSE, RRMSE, and MAE values and the better the performance of the model for GY prediction.

Entropy Weight Method

The ENR algorithm was independently implemented at each growth stage. Instead of using these results to predict the GY individually, we proposed an entropy weight fusion (EWF) model that combines the predicted results from the different growth stages *via* weights obtained during the model training stage. The basic mechanism of the entropy weight method is to use the entropy to characterize the degree of disorder in the system (Farhadinia, 2017). In this method, the relative error between the predicted and measured values of the GY obtained in an individual growth stage by the selected *i*th prediction model can be expressed as follows:

$$E_{ij}\begin{cases} 1, \text{when } \left|\frac{(y_j - y_{ij})}{y_j}\right| \geq 1; \\ \left|\frac{(y_j - y_{ij})}{y_j}\right|, \text{when } 0 \leq \left|\frac{(y_j - y_{ij})}{y_j}\right| \leq 1 \end{cases} \quad (14)$$

where $i = (1, 2, 3...m)$, $j = (1, 2, 3...n)$, and y_{ij} represents the predicted value of the yield forecast model for the *i*th individual growth stage on the *j*th plot. The process for calculating the weights is as follows:

The relative error ratio was calculated between the predicted value of the *i*th individual growth stage and the measured value at plot *j*:

$$P_{ij} = \frac{E_{ij}}{\sum_{i=1}^{n} E_{ij}} \quad (15)$$

where $\sum_{i=1}^{n} P_{ij} = 1$. The entropy value h_i was calculated for the relative error in the *i*th individual growth stage prediction:

$$h_i = -k \sum_{j=1}^{n} \left[P_{ij} \ln(P_{ij})\right] \; k = 1/\ln(n) \quad (16)$$

The relative error variation coefficient was determined based on the principle of the opposite of the entropy value and its degree of variation:

$$d_i = 1 - h_i \quad (17)$$

The weight was then obtained for the predicted output value from a single growth stage:

$$w_i = \frac{1}{n}\left(1 - d_i \Big/ \sum_{i=1}^{n} d_i\right) \quad (18)$$

The weights were obtained by combining the output forecast values from the multiple growth stages. The final output forecast value can be expressed as follows:

$$\hat{y} = \sum_{i=1}^{n} w_i y_{ij} \quad (19)$$

The higher the entropy of the prediction error sequence of a single growth stage, the lower the degree of variation and the greater the weight. The entropy weight method fully considers the relative error in the output prediction value from the different growth stages. Therefore, the predicted results from the multiple growth stages complement each other to improve the accuracy of yield prediction. In this study, eight combinations were created to evaluate the accuracy of the entropy weight method for GY prediction. **Table 4** represents a detailed description of each combination.

RESULTS

Descriptive Statistics of Grain Yield

The distribution of yield from wheat plots is shown in **Figure 3**. The GY was normally distributed under all the irrigation treatments as well as across the treatments. The GY was found to be higher under high irrigation treatments than under the moderate and mild irrigation treatments. The mean GY values for the high, moderate, and mild irrigation treatments were 7.09, 5.99, and 4.40 t ha^{-1}, respectively. The highest coefficient of variation (19.51%) was observed in the mild irrigation treatment and the lowest (12.70%) in the high irrigation treatment. The overall range of the GY data across the irrigation treatment was 2.79–8.64 t ha^{-1}, with a data variation of up to 24.35%. Across treatment data with this type of variation can help evaluate the prediction accuracy of the model.

TABLE 4 | Combination of different growth stages used in the model for grain yield prediction.

Combination	Combination of growth stages
C1	Jointing, booting, heading
C2	Jointing, booting, heading, flowering
C3	Jointing, booting, heading, flowering, grain filling
C4	Jointing, booting, heading, flowering, grain filling, maturity
C5	Booting, heading, flowering
C6	Heading, flowering, grain filling
C7	Flowering, grain filling, maturity
C8	Heading, flowering, maturity

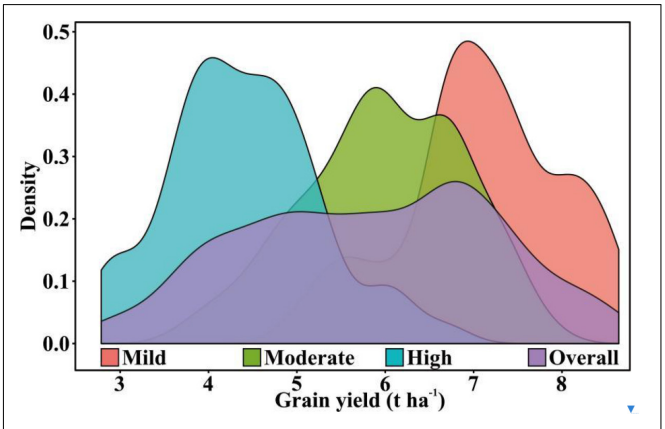

FIGURE 3 | Grain yield distribution curves under various irrigation treatments.

Results of Gray Relational Analysis and Feature Selection

A total of 22 indices were ranked using the GRA method. **Table 5** lists the results for all the growth stages. The GRD of the normalized relative canopy temperature (NRCT) ranked first for the jointing, booting, and flowering stages and relatively high for the heading (rank = 10) and grain filling (rank = 9) stages. However, the NRCT ranked last at maturity. The rankings for most VIs were unstable across all the growth stages. For example, plant pigment ratio (PPR) and difference vegetation index (DVI) had a high ranking at both jointing and booting but ranked low at flowering and grain filling. In accordance with the GRA mechanism, the higher the GRD between the main and the reference sequence, the more closely the sequences are related, which indicates a close relationship between the NRCT and the yield during the multiple growth stages.

To further explore the features with better performance and to reduce the dimensionality of the data, the top feature was iteratively added into the ENR. The performance of the model (i.e., MAE) in the training process was updated until all the features were inputted into the ENR (**Figure 4**). Among the six developmental stages, the grain filling stage yielded the lowest error, and it tended to be stable when the number of features was 19. The highest error was observed in jointing, and the model showed a stable tendency after inputting 16 features. The appropriate numbers of input features for the booting, heading, flowering, and maturity stages were found to be 18, 18, 22, and 22, respectively.

TABLE 5 | Ranking of indices using the gray relational analysis (GRA) for six growth stages.

Rank	Jointing	Booting	Heading	Flowering	Grain filling	Maturity
1	NRCT	NRCT	NLI	NRCT	MSR	TVI
2	PPR	PPR	NDVI	RVI	RVI	NLI
3	OSAVI	SAVI	TCARI	MSR	NDVI	SAVI
4	MTVI2	OSAVI	MSR	NDVI	NLI	RDVI
5	SAVI	RDVI	RVI	NLI	NRI	EVI
6	RDVI	MTVI2	MTVI2	OSAVI	OSAVI	DVI
7	MCARI	PSRI	OSAVI	MTVI2	MTVI2	OSAVI
8	PSRI	TCARI	PPR	TCARI	CIRE	MTVI2
9	TCARI	DVI	SAVI	RDVI	NRCT	PSRI
10	DVI	TVI	NRCT	NRI	NDVIRE	MCARI
11	TVI	MCARI	SIPI	SAVI	GNDVI	MTCI
12	EVI	EVI	RDVI	GNDVI	RDVI	NDVI
13	SIPI	SIPI	MCARI	SIPI	SIPI	RVI
14	GNDVI	NLI	PSRI	PSRI	PSRI	MSR
15	NRI	NDVI	EVI	EVI	SAVI	CIRE
16	NDVIRE	MSR	DVI	CIRE	EVI	NDVIRE
17	NLI	RVI	TVI	NDVIRE	TCARI	SIPI
18	CIRE	MTCI	GNDVI	DVI	TVI	NRI
19	MTCI	CIRE	NRI	TVI	DVI	GNDVI
20	NDVI	NRI	NDVIRE	MTCI	MTCI	TCARI
21	MSR	GNDVI	CIRE	PPR	MCARI	PPR
22	RVI	NDVIRE	MTCI	MCARI	PPR	NRCT

Performance of Elastic Net Regression Model for Individual Growth Stage

To analyze the improvement of the thermal data for yield prediction accuracy, the model was first built using the features extracted from the multispectral images (**Figure 5**). The mean prediction values for the grain filling stage was R^2 = 0.461, followed by flowering (R^2 = 0.432), heading (R^2 = 0.422), maturity (R^2 = 0.417), booting (R^2 = 0.290), and jointing (R^2 = 0.130). **Figure 6** represents the accuracy assessment results of the ENR model for GY predictions by using both thermal and multispectral features. The results show that the dual-sensor data fusion method achieves higher prediction accuracy at all measured stages compared to using single multispectral sensor-based features. As with using only multispectral features, the ENR showed the highest prediction results with a low error at the grain filling (R^2 = 0.667) stage. The mean prediction results at jointing, booting, heading, flowering, and maturity were R^2 = 0.544, R^2 = 0.571, R^2 = 0.602, R^2 = 0.640, and R^2 = 0.527, respectively.

After obtaining the predicted GY using thermal IR and multispectral features, the regression between the predicted GY from the various stages was conducted (**Figure 7**). High correlations ranging from R^2 = 0.59 to R^2 = 0.89 between adjacent growth stages were observed across the growth stages. Moreover, the greater the interval between the growth stages, the lower the R^2-value. For example, the R^2 between the jointing stage and the booting, heading, flowering, grain filling, and maturity stages were 0.78, 0.66, 0.64, 0.52, and 0.41, respectively. In comparison, the correlations between the predicted yield in the maturity and other growth stages were lower, with quite weak regression values ranging from R^2 = 0.41 to R^2 = 0.59. There were differences in the distribution curves of predicted GY values, which provides complementary information.

Performance of Entropy Weight Fusion Method

For comparison with the EWF method, multispectral and thermal features from multiple stages were used as the inputs of ENR to the training model. The results indicated that combining the features of multiple stages increases the accuracy of yield prediction than individual stages (**Figure 8**). The C4 yielded the highest R^2-value of 0.725, followed by C3 (R^2 = 0.717) and C2 (R^2 = 0.691). The remaining combinations achieved similar prediction accuracy (R^2 = 0.669–0.681). However, the obvious fluctuations of accuracy parameters (R^2, RMSE, RRMSE, and MAE) with wide ranges were observed.

Figure 9 represents the performance of the EWF model in predicting the GY using the combined predicted values from the multiple growth stages. Comparing with the individual growth stages, the EWF model also provides a more accurate result regardless of the number of stages adopted. Among the eight combinations, the optimal test results of the EWF model were observed in C4, with a mean R^2 of 0.729. An increase of 0.062 compared with the highest mean R^2-value was observed in the grain filling stage (R^2 = 0.667). Moreover, the RMSE, RRMSE, and MAE values were reduced to 0.831 t ha^{-1}, 14.3%,

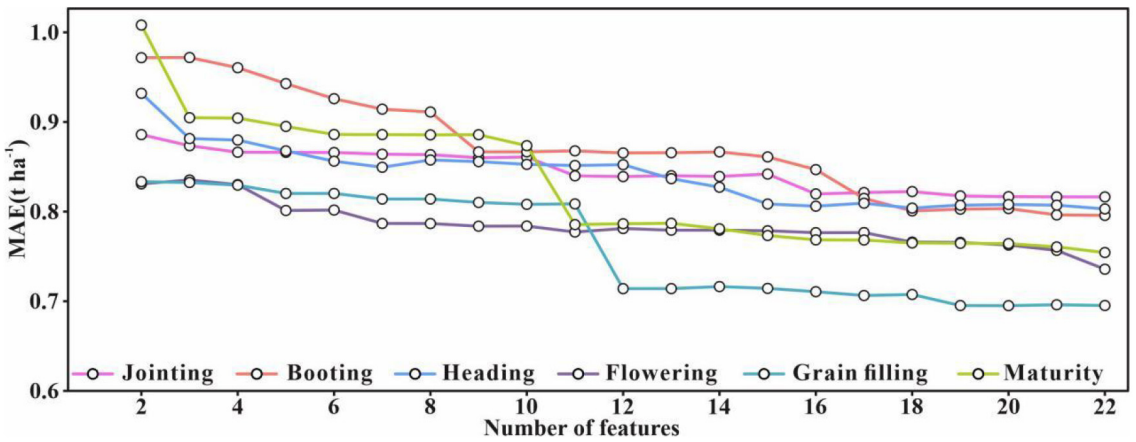

FIGURE 4 | Model training error as a function of the number of features. The order of input of features depends on the gray relational degree (GRD). MAE, mean absolute error.

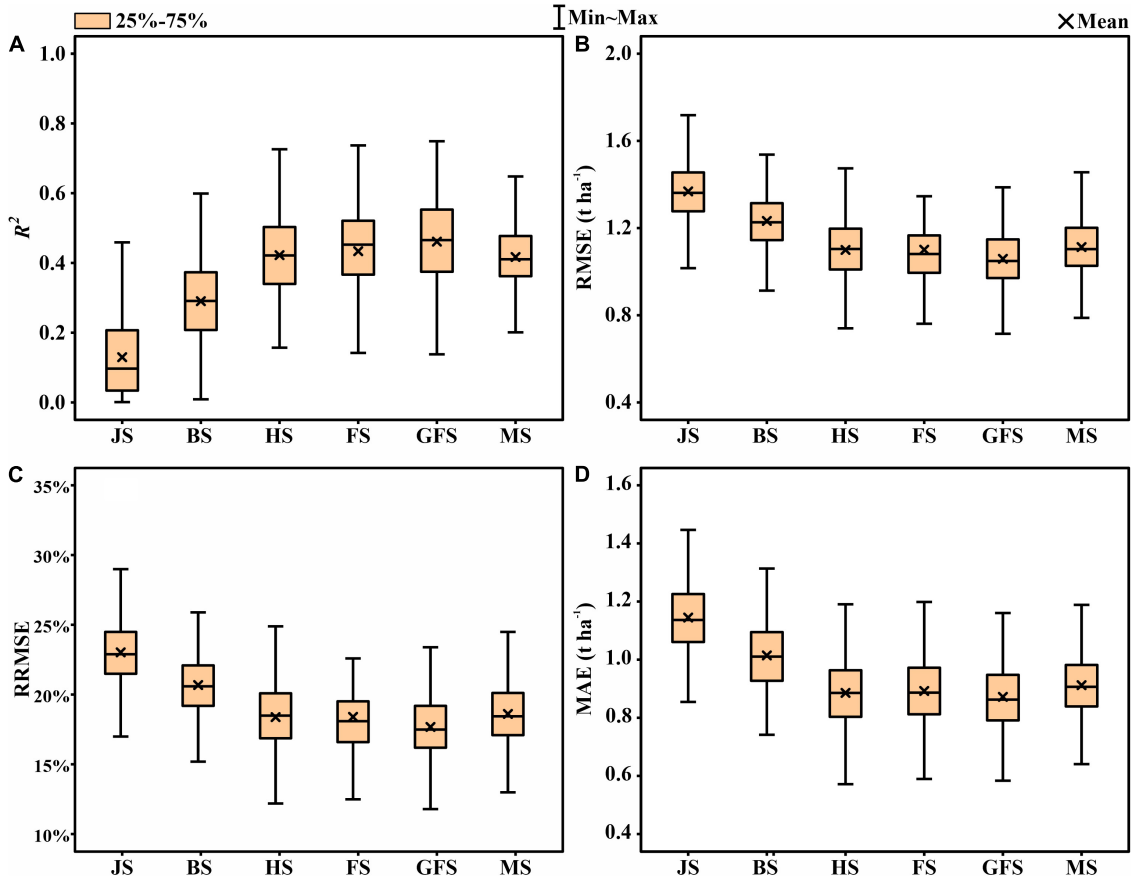

FIGURE 5 | Statistical distributions of **(A)** coefficient of determination (R^2), **(B)** root mean square error (RMSE), **(C)** relative root-mean-square error (RRMSE), and **(D)** mean absolute error (MAE) of the elastic net regression (ENR) algorithm for GY prediction using multispectral features in test phases. JS, jointing stage; BS, booting stage; HS, heading stage; FS, flowering stage; GFS, grain filling stage; MS, maturity stage.

and 0.684 t ha^{-1}, respectively. A low prediction was observed in C1 (R^2 = 0.681). Compared to C5 (R^2 = 0.692), C6 (R^2 = 0.678), C7 (R^2 = 0.677), and C8 (R^2 = 0.688), C2 (R^2 = 0.721) and C3 (R^2 = 0.719) had a more accurate predictions. The fluctuations in the accuracy parameters (R^2, RMSE, RRMSE, and MAE) of the EWF model were more moderate compared to the model

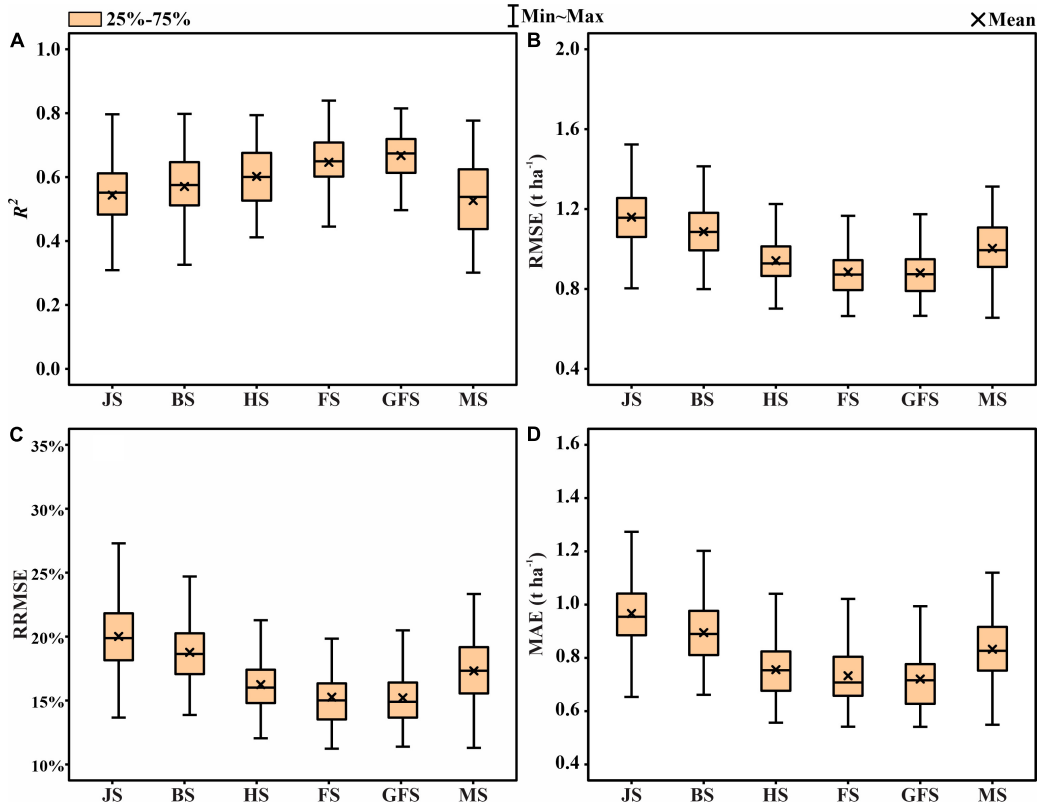

FIGURE 6 | Statistical distributions of **(A)** R^2, **(B)** RMSE, **(C)** RRMSE, and **(D)** MAE of the ENR for GY prediction using both multispectral and thermal features in test phases. JS, jointing stage; BS, booting stage; HS, heading stage; FS, flowering stage; GFS, grain filling stage; MS, maturity stage.

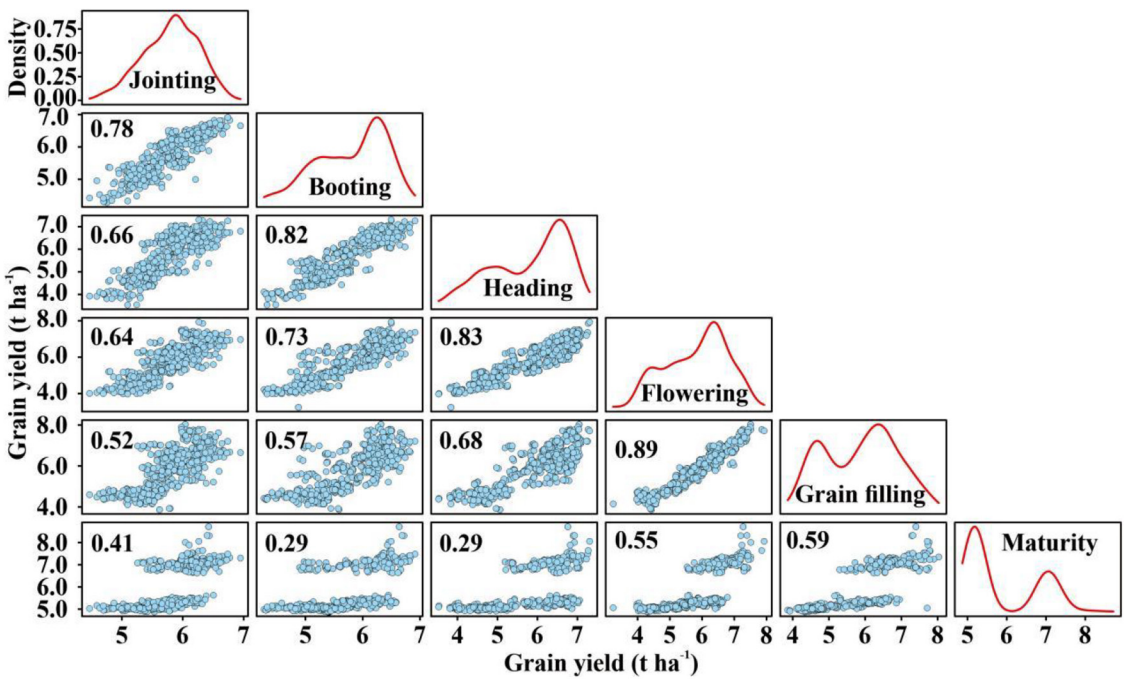

FIGURE 7 | Regression plots, density curve, and R^2-values between predicted GY in six developmental stages.

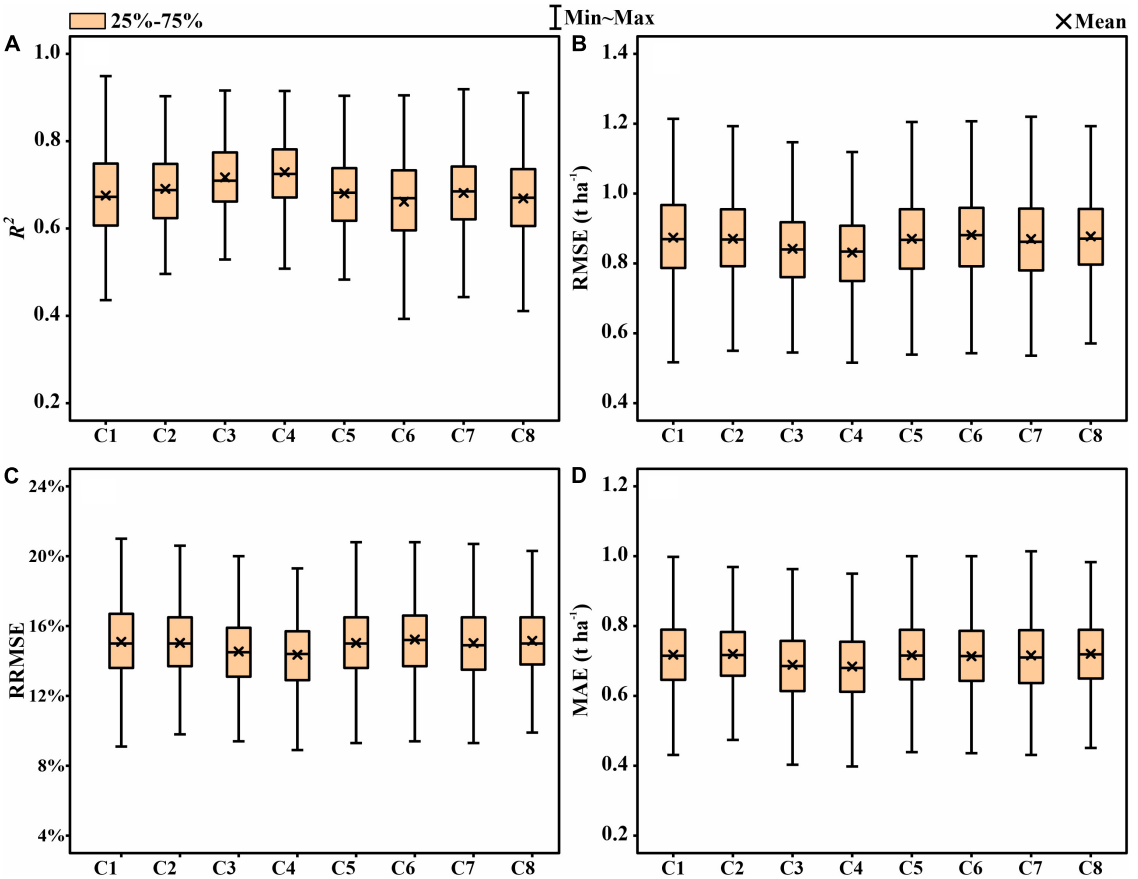

FIGURE 8 | Statistical distributions of **(A)** R^2, **(B)** RMSE, **(C)** RRMSE, and **(D)** MAE of the ENR model that uses both multispectral and thermal features from different stages as inputs.

that directly used multistage features as inputs (**Figures 8**, **9**), which again demonstrates the stability of the EWF method. A paired *t*-test was utilized to assess whether the EWF models performed statistically high in terms of the R^2-values compared with the other models (**Figure 10**). The results showed significantly high R^2-values for the EWF model in all the growth stage combinations.

DISCUSSION

The UAV-based phenotyping is an emerging technique in practical crop breeding. Previous studies have shown that the UAV-based features and the machine learning model can be used together to predict crop yields in breeding work with a large number of crop genotypes (Osval et al., 2017; Fei et al., 2021). In this study, ENR is a relatively new machine learning algorithm being used for yield prediction. ENR combines the properties of ridge regression and LASSO (Ogutu et al., 2012), both of which have been successfully applied to crop yield prediction (Kang et al., 2021; Shafiee et al., 2021). The incorporation of multiple VIs adds collinearity to the models, and the ENR is robust to severe multicollinearity among the input features (Ogutu et al., 2012). Another reason for using ENR was the simplicity of the linear model compared to other machine learning algorithms such as RF or ANNs, which makes the model run less time-consuming and more efficient to train.

Several VIs such as normalized difference vegetation index (NDVI) and green normalized difference vegetation index (GNDVI) have been evaluated to monitor crop health under stress and predict the GY. Most of the multispectral VIs that have been reported were species-specific and easily saturated (Hatfield and Prueger, 2013). Therefore, it is challenging to predict important crop traits using a single VI (Wang et al., 2016). The successful use of multiple VIs to improve the prediction accuracy of important traits in crops has been reported in many studies (Wang et al., 2016; Jin et al., 2020). In this study, 21 multispectral VIs and 1 temperature index (i.e., NRCT) were measured in multiple growth stages to validate the UAV data and ENR and check their accuracy for GY prediction. For accurate yield predictions and to avoid model overfitting, the machine learning algorithms may benefit from using a feature selection algorithm to reduce the dimensionality of the data to an appropriate level (Yoosefzadeh-Najafabadi et al., 2021). The GRA is a widely accepted approach in feature selection (Deris et al., 2013; Lu et al., 2019; Yao et al., 2019; Miswan et al., 2021). The

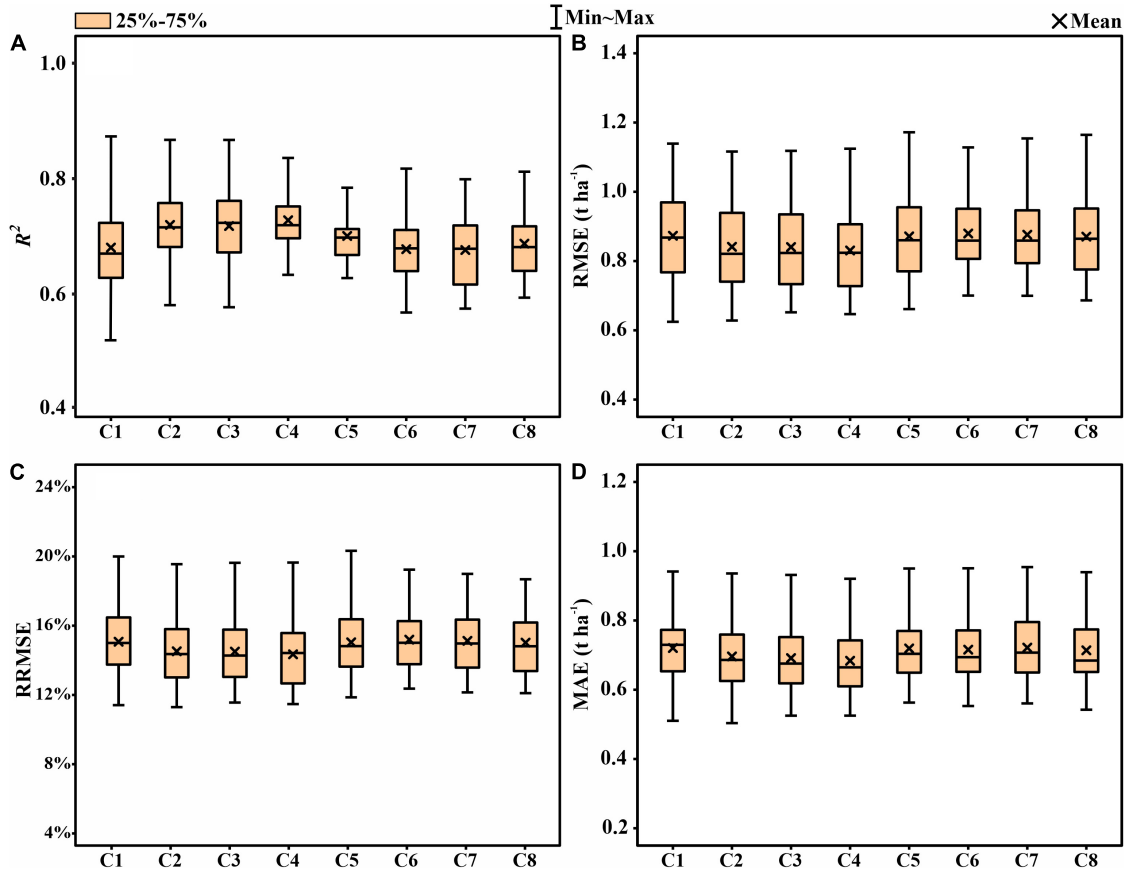

FIGURE 9 | Statistical distributions of **(A)** R^2, **(B)** RMSE, **(C)** RRMSE, and **(D)** MAE of the entropy weight fusion (EWF) method for GY prediction in the test phases.

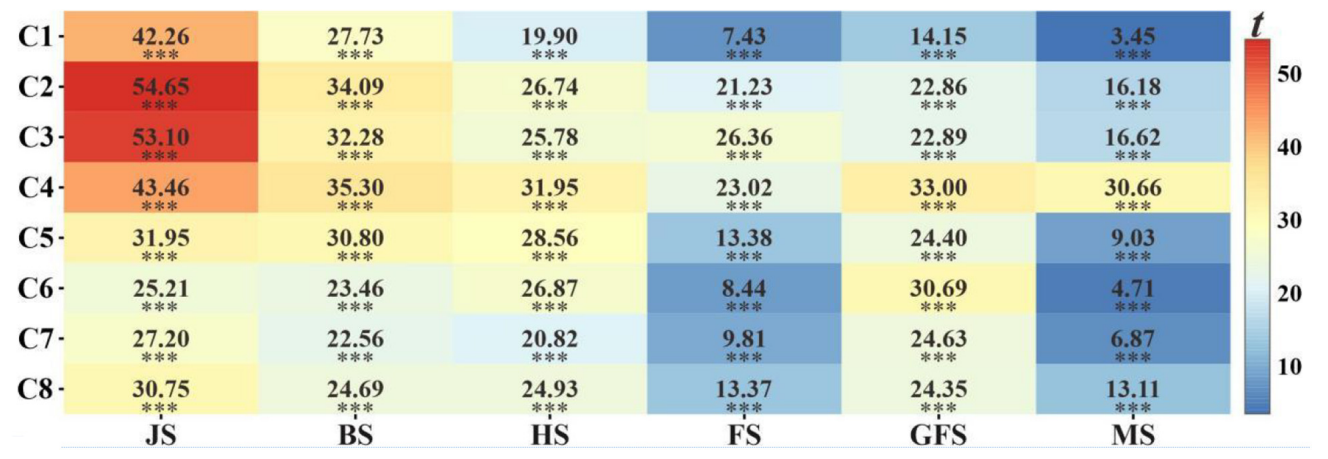

FIGURE 10 | Results from paired *t*-test between model R^2 obtained from the EWF method and the individual stages. *** significant at $P \leq 0.001$; JS, jointing stage; BS, booting stage; HS, heading stage; FS, flowering stage; GFS, grain filling stage; MS, maturity stage.

results in this study show that GRA can reduce the dimensionality of the input features to some extent.

The model performed poorly when using multispectral VIs to predict yield. This could be due to the saturation issue associated with the visible-near-infrared (Vis-NIR) sensor for dense vegetation such as wheat, soybean, and rice (Thenkabail et al., 2000; Tilly et al., 2015). The fusion of multispectral and thermal features includes canopy spectral and temperature information outperformed for yield prediction, which was consistent with previous reports (Maimaitijiang et al., 2017, 2020). Temperature

is closely related to plant physiological processes such as transpiration, leaf water potential, and photosynthesis (Sagan et al., 2019). Generally, high canopy temperature is negatively correlated with crop yield (Tattaris et al., 2016; Sagan et al., 2019). Previously, the UAV-based thermal IR data has been successfully applied to plant trait evaluation (Gonzalez-Dugo et al., 2013; Ludovisi et al., 2017; Liu et al., 2018; Raeva et al., 2019; Crusiol et al., 2020). Previous studies have also shown that combining thermal data with multispectral data outperformed as compared to the fusion of spectral and structural information from RGB and multispectral images for the prediction of LAI, biomass, chlorophyll, and nitrogen in soybean (Maimaitijiang et al., 2017).

The spatial heterogeneity of the ground changes among the developmental stages of the crop could lead to a significant difference in the prediction accuracy across the growth stages (Juliane et al., 2014). The results of yield prediction based on the individual growth stages are similar to previous reports, i.e., wheat yield prediction accuracy was higher at grain filling stages under different growth environments (Hernandez et al., 2015; Hassan et al., 2019a). During the grain filling stage, the starch, protein, and organic matter produced by photosynthesis are transported to the grain (Guan et al., 2017), and this stage is closely linked to the thousand-grain weight. Therefore, the accuracy of yield estimation was highest at the grain filling stage. In addition, the reduction in the greenness and chlorophyll level after the grain filling stages due to the decrease in the degree of dry matter accumulation in the leaves of plants could influence the detection accuracy of VIs based on the red and near-IR light (Yue et al., 2017). This reduces the model performance in the late developmental stage, which causes a decrease in yield prediction accuracy at maturity. In addition, crop canopy information at varying growth stages is associated with different yield elements, and a combination of remote sensing data from multiple growth stages can effectively improve the yield prediction accuracy.

Another main objective of this study was to use an appropriate method to acquire the prediction values from a combination of temporal remote sensing data across the growth cycle. Although previous studies used temporal VIs for yield prediction, most of them used a single VI (Wang et al., 2014; Zhou et al., 2017), which can be influenced by different degrees of saturability or soil background (Wang et al., 2016). We first directly used the multispectral and thermal features from multiple stages as inputs to ENR, and this method was able to obtain higher yield prediction accuracy than individual stages, but the accuracy parameters fluctuated more compared to EMF and were slightly lower than the prediction accuracy of the EMF method for some combinations. This may be due to the redundant information generated by the accumulation of features from multiple growth stages. In addition, the excessive dimensionality of input features also poses the risk of overfitting the machine learning model (Feng et al., 2017; Coolen et al., 2020). Among the combinations of EMF, the prediction accuracy of C2 was comparable to a combination with the highest prediction accuracy of C4. The data required for C2 can be obtained at the flowering stage, which is appropriate for application in practical management. The results of this study suggest that the fusion of multispectral and thermal features within an entropy weight ensemble framework can provide accurate wheat yield predictions. However, more comprehensive studies, such as studies of different crop varieties in different environments, are needed to determine the most accurate and efficient multistage data for combination.

CONCLUSION

A rapid and nondestructive method for an accurate GY prediction of wheat is desired in breeding programs. In this study, an ensemble framework was developed to increase the GY prediction accuracy by integrating the predicted values from multiple stages using the UAV-based multispectral and thermal IR imagery. The test results showed that the prediction accuracy of the grain filling stage was the highest among the six growth stages. The ensemble method outperformed the individual stage-based GY prediction in terms of accuracy. Combining the features of the first four growth stages allows for early and accurate yield prediction to aid in decision-making. This study offers a new method for GY prediction through UAV-based remote sensing, and it can help in large breeding activities.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

SF analyzed the data and wrote the manuscript. ZC and YX directed the trial and provided the main idea. QC and ZL helped to collect data. MH, YM, and MS provided comments and suggestions to improve the manuscript. All authors read and agreed to the published version of the manuscript.

FUNDING

This study was supported by the Technology Innovation Program of the Chinese Academy of Agricultural Sciences (CAAS-ZDXT-2019002) and the Key Grant Technology Project of Xinxiang (ZD2020009).

ACKNOWLEDGMENTS

We would like to acknowledge all the crew at the Institute of Farmland Irrigation of Chinese Academy of Agricultural Sciences.

REFERENCES

Appeltans, S., Guerrero, A., Nawar, S., Pieters, J., and Mouazen, A. M. (2020). Practical recommendations for hyperspectral and thermal proximal disease sensing in potato and leek fields. *Remote Sens.* 12:1939. doi: 10.3390/rs12121939

Aslan, N., Shahrivar, A. A., and Abdollahi, H. (2012). Multi-objective optimization of some process parameters of a lab-scale thickener using gray relational analysis. *Sep. Purif. Technol.* 90, 189–195. doi: 10.1016/j.seppur.2012.02.033

Ba Ret, F., and Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* 35, 161–173. doi: 10.1016/0034-4257(91)90009-U

Bradley, J. B. (1995). Neural networks: A comprehensive foundation. *Inf. Process. Manage.* 31, 786. doi: 10.1016/0306-4573(95)90003-9

Breiman, L. (2001). Random forest. *Mach. Learn.* 45, 5–32. doi: 10.1023/a:101093340432

Broge, N. H., and Leblanc, E. (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* 76, 156–172. doi: 10.1016/S0034-4257(00)00197-8

Chen, J. M. (1996). Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Can. J. Remote. Sens.* 22, 229–242. doi: 10.1080/07038992.1996.10855178

Cheng, W., Xi, H., Sindikubwabo, C., Si, J., and Wu, T. (2020). Ecosystem health assessment of desert nature reserve with entropy weight and fuzzy mathematics methods: A case study of Badain Jaran Desert. *Ecol. Indic.* 119:106843. doi: 10.1016/j.ecolind.2020.106843

Colomina, I., and Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS J. Photogramm. *Remote Sens.* 92, 79–97. doi: 10.1016/j.isprsjprs.2014.02.013

Comba, L., Biglia, A., Aimonino, D. R., Tortia, C., and Gay, P. (2020). Leaf Area Index evaluation in vineyards using 3D point clouds from UAV imagery. *Precis. Agric.* 21, 881–896. doi: 10.1007/s11119-019-09699-x

Coolen, A. C. C., Sheikh, M., Mozeika, A., Aguirre-Lopez, F., and Antenucci, F. (2020). Replica analysis of overfitting in generalized linear regression models. *J. Phys. A Math. Theor.* 53, doi: 10.1088/1751-8121/aba028

Crusiol, L. G. T., Nanni, M. R., Furlanetto, R. H., Sibaldelli, R. N. R., Cezar, E., Mertz-Henning, L. M., et al. (2020). UAV-based thermal imaging in the assessment of water status of soybean plants. *Int. J. Remote Sens.* 41, 3243–3265. doi: 10.1080/01431161.2019.1673914

Dash, J., and Curran, P. J. (2004). "Evaluation of the meris terrestrial chlorophyll index" in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. (United States: IEEE). 1–257.

Daughtry, C., Walthall, C. L., Kim, M. S., Colstoun, E., and McMurtrey, M. M. III. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* 74, 229–239. doi: 10.1016/S0034-4257(00)00113-9

Deris, A. M., Zain, A. M., and Sallehuddin, R. (2013). Hybrid GR-SVM for prediction of surface roughness in abrasive water jet machining. *Meccanica* 48, 1937–1945. doi: 10.1007/s11012-013-9710-2

Diker, K., and Bausch, W. C. (2003). Potential use of nitrogen reflectance index to estimate plant parameters and yield of maize. *Biosyst. Eng.* 85, 437–447. doi: 10.1016/S1537-5110(03)00097-7

Eitel, J. U. H., Long, D. S., Gessler, P. E., and Smith, A. M. S. (2007). Using in-situ measurements to evaluate the new RapidEye^{TM} satellite series for prediction of wheat nitrogen status. *Int. J. Remote Sens.* 28, 4183–4190. doi: 10.1080/01431160701422213

Elmetwalli, A. H., El-Hendawy, S., Al-Suhaibani, N., Alotaibi, M., Tahir, M. U., Mubushar, M., et al. (2020). Potential of hyperspectral and thermal proximal sensing for estimating growth performance and yield of soybean exposed to different drip irrigation regimes under arid conditions. *Sensors.* 20:6569. doi: 10.3390/s20226569

Elsayed, S., Elhoweity, M., Ibrahim, H. H., Dewir, Y. H., Migdadi, H. M., and Schmidhalter, U. (2017). Thermal imaging and passive reflectance sensing to estimate the water status and grain yield of wheat under different irrigation regimes. *Agric. Water Manag.* 189, 98–110. doi: 10.1016/j.agwat.2017.05.001

Elsayed, S., Rischbeck, P., and Schmidhalter, U. (2015). Comparing the performance of active and passive reflectance sensors to assess the normalized relative canopy temperature and grain yield of drought-stressed barley cultivars. *Field Crop. Res.* 177, 148–160. doi: 10.1016/j.fcr.2015.03.010

Farhadinia, B. (2017). A multiple criteria decision making model with entropy weight in an interval-transformed hesitant fuzzy environment. *Cogn Comput.* 9, 513–525. doi: 10.1007/s12559-017-9480-6

Fei, S., Hassan, M. A., He, Z., Chen, Z., Shu, M., Wang, J., et al. (2021). Assessment of ensemble learning to predict wheat grain yield based on UAV-Multispectral reflectance. *Remote Sens.* 13:2338. doi: 10.3390/rs13122338

Feng, X., Liang, Y., Shi, X., Xu, D., Wang, X., and Guan, R. (2017). Overfitting reduction of text classification based on AdaBELM. *Entropy* 19:330. doi: 10.3390/e19070330

Gitelson, A. A., Kaufman, Y. J., and Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* 58, 289–298. doi: 10.1016/S0034-4257(96)00072-7

Gitelson, A. A., Vina, A., Arkebauer, T. J., Rundquist, D. C., Keydan, G., and Leavitt, B. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys. Res. Lett.* 30:1248. doi: 10.1029/2002GL016450

Goel, N. S., and Qin, W. (1994). Influences of canopy architecture on relationships between various vegetation indices and LAI and Fpar: A computer simulation. *Remote Sens. Environ.* 10, 309–347. doi: 10.1080/02757259409532252

Gonzalez-Dugo, V., Zarco-Tejada, P., Nicolás, E., Nortes, P. A., Alarcón, J. J., Intrigliolo, D. S., et al. (2013). Using high resolution UAV thermal imagery to assess the variability in the water status of five fruit tree species within a commercial orchard. *Precis. Agric.* 14, 660–678. doi: 10.1007/s11119-013-9322-9

Guan, K., Wu, J., Kimball, J. S., Anderson, M. C., Frolking, S., Li, B., et al. (2017). The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields. *Remote Sens. Environ.* 199, 333–349. doi: 10.1016/j.rse.2017.06.043

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., and 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 Sens. Environ.* 90, 337–352. doi: 10.1016/j.rse.2003.12.013

Hassan, M. A., Yang, M., Awais, R., Jin, X., Xia, X., Xiao, Y., et al. (2018). Time-series multispectral indices from unmanned aerial vehicle imagery reveal senescence rate in bread wheat. *Remote Sens.* 10:809. doi: 10.3390/rs10060809

Hassan, M. A., Yang, M., Awais, R., Yang, G., Reynolds, M., Xia, X., et al. (2019a). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Sci.* 282, 95–103. doi: 10.1016/j.plantsci.2018.10.022

Hassan, M. A., Yang, M., Fu, L., Rasheed, A., Zheng, B., Xia, X., et al. (2019b). Accuracy assessment of plant height using an unmanned aerial vehicle for quantitative genomic analysis in bread wheat. *Plant Methods* 15:37.

Hassan, M. A., Yang, M., Rasheed, A., Tian, X., Reynolds, M., Xia, X., et al. (2021). Quantifying senescence in bread wheat using multispectral imaging from an unmanned aerial vehicle and QTL mapping. *Plant Phys.* 2021, kiab 431,

Hatfield, J. L., and Prueger, J. H. (2013). Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices. *Remote Sens.* 2, 562–578. doi: 10.3390/rs2020562

Hernandez, J., Lobos, G. A., Matus, I., Del Pozo, A., Silva, P., and Galleguillos, M. (2015). Using ridge regression models to estimate grain yield from field spectral data in bread wheat (Triticum Aestivum L.) grown under three water regimes. *Remote Sens.* 7, 2109–2126. doi: 10.3390/rs70202109

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213. doi: 10.1016/S0034-4257(02)00096-2

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25, 295–309. doi: 10.1016/0034-4257(88)90106-X

Hui, Z., and Hastie, T. (2005). Regularization and variable selection via the elastic net. *J. R. Statist. Soc. B* 67:768. doi: 10.1111/j.1467-9868.2005.00527.x

Jin, X., Li, Z., Feng, H., Ren, Z., and Li, S. (2020). Deep neural network algorithm for estimating maize biomass based on simulated Sentinel 2A vegetation indices and leaf area index. *Crop J.* 8, 87–97. doi: 10.1016/j.cj.2019.06.005

Jin, X., Xu, X., Song, X., Li, Z., Wang, J., and Guo, W. (2013). Estimation of leaf water content in winter wheat using gray relational analysis–partial least squares modeling with hyperspectral data. *Agron. J.* 105, 1385–1392. doi: 10.2134/agronj2013.0088

Juliane, B., Andreas, B., Simon, B., Janis, B., Silas, E., and Georg, B. (2014). Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. *Remote Sens.* 6, 10395–10412. doi: 10.3390/rs61110395

Kang, Y., Nam, J., Kim, Y., Lee, S., Seong, D., Jang, S., et al. (2021). Assessment of regression models for predicting rice yield and protein content using unmanned aerial Vehicle-Based multispectral imagery. *Remote Sens.* 13:1508. doi: 10.3390/rs13081508

Li, H., Li, D., and Li, Y. (2018). A multi-index assessment method for evaluating coverage effectiveness of remote sensing satellite. *Chinese J. Aeronaut.* 151, 98–108. doi: 10.1016/j.cja.2018.05.015

Li, L., Zhang, Q., and Huang, D. (2014). A review of imaging techniques for plant phenotyping. *Sensors*. 14, 20078–20111. doi: 10.3390/s141120078

Li, X., Wang, K., Liu, L., Jing, X., Yang, H., and Gao, C. (2011). Application of the entropy weight and topsis method in safety evaluation of coal mines. *Procedia Eng.* 26, 2085–2091. doi: 10.1016/j.proeng.2011.11.2410

Liu, H., Zhang, X., Xu, Y., Ma, F., Zhang, J., Cao, Y., et al. (2020). Identification and validation of quantitative trait loci for kernel traits in common wheat (Triticum Aestivum L.). *BMC Plant Biol.* 20:529. doi: 10.1186/s12870-020-02661-4

Liu, T., Li, R., Zhong, X., Jiang, M., Jin, X., Zhou, P., et al. (2018). Estimates of rice lodging using indices derived from UAV visible and thermal infrared images. *Agr. Forest Meteorol.* 252, 144–154. doi: 10.1016/j.agrformet.2018.01.021

Lu, D., Yu, H. L., and Yu, G. M. (2014). Assessing the land use change and ecological security based on RS and GIS: A case study of Pingdingshan city. *China. Adv. Mater. Res.* 905, 329–333. doi: 10.4028/www.scientific.net/AMR.905.329

Lu, S. X., Lin, G., Que, H., Li, M. J. J., Wei, C. H., and Wang, J. K. (2019). Grey relational analysis using Gaussian process regression method for dissolved gas concentration prediction. *Int J Mach Learn Cyb.* 10, 1313–1322. doi: 10.1007/s13042-018-0812-y

Ludovisi, R., Tauro, F., Salvati, R., Khoury, S., and Harfouche, A. (2017). UAV-based thermal imaging for high-throughput field phenotyping of black poplar response to drought. *Front. Plant Sci.* 8:1681. doi: 10.3389/fpls.2017.01681

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 J. Photogramm. Remote Sens.* 134, 43–58. doi: 10.1016/j.isprsjprs.2017.10.011

Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., and Fritschi, F. B. (2020). Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens. Environ.* 237:111599. doi: 10.1016/j.rse.2019.111599

Mcbratney, A., Whelan, B., Ancev, T., and Bouma, J. (2005). Future directions of precision agriculture. *Precis. Agric.* 6, 7–23. doi: 10.1007/s11119-005-0681-8

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., and Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiol. Plant.* 106, 135–141. doi: 10.1034/j.1399-3054.1999.106119.x

Metternicht, G. (2003). Vegetation indices derived from high-resolution airborne videography for precision crop management. *Int. J. Remote Sens.* 24, 2855–2877. doi: 10.1080/01431160210163074

Miswan, N. H., Chan, C. S., and Ng, C. G. (2021). *Hospital Readmission Prediction Based on Improved Feature Selection Using Grey Relational Analysis and LASSO*. United Kingdom: Emerald Group Publishing. doi: 10.1108/GS-12-2020-0168

Ogutu, J. O., Schulz-Streeck, T., and Piepho, H. (2012). Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. *BMC Proc.* 6:S10. doi: 10.1186/1753-6561-6-S2-S10

Osval, A., Montesinos López, A. M. L., José, C., Gustavo, D. L. C., Alvarado, G., Mondal, S., et al. (2017). Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. *Plant Methods* 13:4. doi: 10.1186/s13007

Panda, S. S., Ames, D. P., and Panigrahi, S. (2010). Application of vegetation indices for agricultural crop yield prediction using neural network techniques. *Remote Sens.* 2:673. doi: 10.3390/rs2030673

Penuelas, J., Filella, I., and Gamon, J. A. (1995). Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytol.* 131, 291–296. doi: 10.1111/j.1469-8137.1995.tb03064.x

Qiao, L., Gao, D., Zhang, J., Li, M., and Ma, J. (2020). Dynamic influence elimination and chlorophyll content diagnosis of maize using UAV spectral imagery. *Remote Sens.* 12:2650. doi: 10.3390/rs12162650

Raeva, P. L., Šedina, J., and Dlesk, A. (2019). Monitoring of crop fields using multispectral and thermal imagery from UAV. *Eur. J. Remote Sens.* 52, 192–201. doi: 10.1080/22797254.2018.1527661

Rischbeck, P., Baresel, P., Elsayed, S., Mistele, B., and Schmidhalter, U. (2014). Development of a diurnal dehydration index for spring barley phenotyping. *Funct. Plant Biol.* 41, 1249–1260. doi: 10.1071/FP14069

Rondeaux, G., Steven, M., and Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* 55, 95–107. doi: 10.1016/0034-4257(95)00186-7

Rouse, J. Jr. (1972). "Monitoring the vernal advancement and retrogradation (green WaveEffect) of natural vegetation" in *Nasa/Gsfct Type Final Report*. (Maryland: NASA Goddard Space Flight Center).

Sagan, V., Maimaitijiang, M., Sidike, P., Eblimit, K., Peterson, K., Hartling, S., et al. (2019). UAV-based high resolution thermal imaging for vegetation monitoring, and plant phenotyping using ICI 8640 P, FLIR Vue Pro R 640, and thermoMap Cameras. *Remote Sens.* 11:330. doi: 10.3390/rs11030330

Sain, S. R. (1997). The nature of statistical learning theory. *Technometrics* 38:409. doi: 10.1080/00401706.1996.10484565

Shafiee, S., Lied, L. M., Burud, I., Dieseth, J. A., Alsheikh, M., and Lillemo, M. (2021). Sequential forward selection and support vector regression in comparison to LASSO regression for spring wheat yield prediction based on UAV imagery. *Comput. Electron. Agr.* 183:106036. doi: 10.1016/j.compag.2021.106036

Shah, S. H., Angel, Y., Houborg, R., Ali, S., and McCabe, M. F. (2019). A random forest machine learning approach for the retrieval of leaf chlorophyll content in wheat. *Remote Sens.* 11:920. doi: 10.3390/rs11080920

Sidike, P., Sagan, V., Qumsiyeh, M., Maimaitijiang, M., Essa, A., and Asari, V. (2018). *Adaptive Trigonometric Transformation Function with Image Contrast and Color Enhancement: application to Unmanned Aerial System Imagery*. United Kingdom: IEEE. 1–5. doi: 10.1109/LGRS.2018.2790899

Suyoung, P., Dongryeol, R., Sigfredo, F., Hoam, C., Esther, H. M., and Mark, O. (2017). Adaptive estimation of crop water stress in nectarine and peach orchards using high-resolution imagery from an unmanned aerial vehicle (UAV). *Remote Sens.* 9:828. doi: 10.3390/rs9080828

Tattaris, M., Reynolds, M. P., and Chapman, S. C. (2016). A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding. *Front. Plant Sci.* 7:1131. doi: 10.3389/fpls.2016.01131

Tavakoli, H., and Gebbers, R. (2019). Assessing nitrogen and water status of winter wheat using a digital camera. *Comput. Electron. Agr.* 157, 558–567. doi: 10.1016/j.compag.2019.01.030

Thenkabail, P. S., Smith, R. B., and De Pauw, E. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sens. Environ.* 71, 158–182.

Tilly, N., Aasen, H., and Bareth, G. (2015). Fusion of plant height and vegetation indices for the estimation of barley biomass. *Remote Sens.* 7, 17291–17296.

Tucker, C. J., Elgin, J. H., McMurtrey, M. M. III., and Fan, C. J. (1979). Monitoring corn and soybean crop development with hand-held radiometer spectral data. *Remote Sens. Environ.* 8, 237–248. doi: 10.1016/0034-4257(79)90004-X

Wang, J., Chen, Y., Chen, F., Shi, T., and Wu, G. (2018). Wavelet-based coupling of leaf and canopy reflectance spectra to improve the estimation accuracy of foliar nitrogen concentration. *Agr. Forest Meteorol.* 248, 306–315. doi: 10.1016/j.agrformet.2017.10.017

Wang, L., Chang, Q., Yang, J., Zhang, X., Li, F., and Xu, Y. (2018). Estimation of paddy rice leaf area index using machine learning methods based on hyperspectral data from multi-year experiments. *PLoS One* 13:e0207624. doi: 10.1371/journal.pone.0207624

Wang, K., Shen, Z. Q., and Wang, R. C. (1998). Effects of nitrogen nutrition on the spectral reflectance characteristics of rice leaf and canopy. *J. Zhejiang Agric. Univ.* 24, 93–97.

Wang, L., Tian, Y., Yao, X., Zhu, Y., and Cao, W. (2014). Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. *Field Crop. Res.* 164, 178–188. doi: 10.1016/j.fcr.2014.05.001

Wang, L., Zhou, X., Zhu, X., Dong, Z., and Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop J.* 4, 212–219. doi: 10.1016/j.cj.2016.01.008

Yang, M., Hassan, M. A., Xu, K., Zheng, C., Rasheed, A., Zhang, Y., et al. (2020). Assessment of water and nitrogen use efficiencies through UAV-based

multispectral phenotyping in winter wheat. *Front. Plant Sci.* 11:927. doi: 10.3389/fpls.2020.00927

Yao, Z., Zhang, T., Wang, J., and Zhu, L. (2019). A feature selection approach based on grey relational analysis for within-project software defect prediction. *J. Grey Syst.* 31, 105–116.

Yoosefzadeh-Najafabadi, M., Earl, H. J., Tulpan, D., Sulik, J., and Eskandari, M. (2021). Application of machine learning algorithms in plant breeding: Predicting yield from hyperspectral reflectance in soybean. *Front. Plant Sci.* 11:624273. doi: 10.3389/fpls.2020.624273

Yu, N., Li, L., Schmitz, N., Tiaz, L. F., Greenberg, J. A., and Diers, B. W. (2016). Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. *Remote Sens. Environ.* 187, 91–101. doi: 10.1016/j.rse.2016.10.005

Yue, J., Yang, G., Li, C., Li, Z., Wang, Y., Feng, H., et al. (2017). Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. *Remote Sens.* 9:708. doi: 10.3390/rs9070708

Yue, J., Yang, G., Tian, Q., Feng, H., Xu, K., and Zhou, C. (2019). Estimate of winter-wheat above-ground biomass based on UAV ultrahigh-ground-resolution image textures and vegetation indices. *ISPRS J. Photogramm. Remote Sens.* 150, 226–244. doi: 10.1016/j.isprsjprs.2019.02.022

Zhang, L., Zhang, Z., Luo, Y., Cao, J., and Tao, F. (2020). Combining optical, fluorescence, thermal satellite, and environmental data to predict county-level maize yield in china using machine learning approaches. *Remote Sens.* 12:21. doi: 10.3390/rs12010021

Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., Yao, X., et al. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS J. Photogramm. Remote Sens.* 130, 246–255. doi: 10.1016/j.isprsjprs.2017.05.003

Zubler, A. V., and Yoon, J. (2020). Proximal methods for plant stress detection using optical sensors and machine learning. *Biosensors* 10:193. doi: 10.3390/bios10120193

Conflict of Interest: 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.

The reviewer AR declared a past co-authorship with the authors MH, YX to the handling editor.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 11 March 2022
doi: 10.3389/fpls.2022.722442

Machine Learning Strategies for the Retrieval of Leaf-Chlorophyll Dynamics: Model Choice, Sequential Versus Retraining Learning, and Hyperspectral Predictors

Yoseline Angel and Matthew F. McCabe*

Hydrology, Agriculture and Land Observation Group, Water Desalination and Reuse Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

Monitoring leaf Chlorophyll (Chl) *in-situ* is labor-intensive, limiting representative sampling for detailed mapping of Chl variability at field scales across time. Unmanned aeria-l vehicles (UAV) and hyperspectral cameras provide flexible platforms for observing agricultural systems, overcoming this spatio-temporal sampling constraint. Here, we evaluate a customized machine learning (ML) workflow to retrieve multi-temporal leaf-Chl levels, combining sub-centimeter resolution UAV-hyperspectral imagery (400–1,000 nm) with leaf-level reflectance spectra and SPAD measurements, capturing temporal correlations, selecting relevant predictors, and retrieving accurate results under different conditions. The study is performed within a phenotyping experiment to monitor wild tomato plants' development. Several analyses were conducted to evaluate multiple ML strategies, including: (1) exploring sequential versus retraining learning; (2) comparing insights gained from using 272 spectral bands versus 60 pigment-based vegetation indices (VIs); and (3) assessing six regression methods (linear, partial-least-square regression; PLSR, decision trees, support vector, ensemble trees, and Gaussian process; GPR). Goodness-of-fit (R^2) and accuracy metrics (MAE, RMSE) were determined using training/testing and validation data subsets to assess the models' performance. Overall, while equally good performance was obtained using either PLSR, GPR, or random forest, results show: (1) the retraining strategy improved the ability of most of the approaches to model SPAD-based Chl dynamics; (2) comparative analysis between retrievals and validation data distributions informed the models' ability to capture Chl dynamics through SPAD levels; (3) VI predictors slightly improved R^2 (e.g., from 0.59 to 0.74 units for GPR) and accuracy (e.g., MAE and RMSE differences of up to 2 SPAD units) in specific algorithms; (4) feature importance examined through these methods, revealed strong overlaps between relevant bands and VI predictors, highlighting a few decisive spectral ranges and indices useful for retrieving leaf-Chl levels. The proposed ML framework allows the retrieval of high-quality spatially distributed and multi-temporal SPAD-based chlorophyll maps at an ultra-high pixel resolution (e.g., 7 mm).

Keywords: chlorophyll, hyperspectral image, SPAD – leaf greenness, machine learning, UAV, multitemporal analyses, vegetation indices, digital phenotyping

OPEN ACCESS

Edited by:
Frédéric Cointault, Agrosup Dijon, France

Reviewed by:
John Louis Van Hemert, Corteva Agriscience™, United States
Lingling Ma, Chinese Academy of Sciences (CAS), China

****Correspondence:***
Yoseline Angel
yoseline.angellopez@kaust.edu.sa

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

Received: *08 June 2021*
Accepted: *08 February 2022*
Published: *11 March 2022*

Citation:
Angel Y and McCabe MF (2022) Machine Learning Strategies for the Retrieval of Leaf-Chlorophyll Dynamics: Model Choice, Sequential Versus Retraining Learning, and Hyperspectral Predictors. Front. Plant Sci. 13:722442. doi: 10.3389/fpls.2022.722442

INTRODUCTION

Chlorophyll (Chl) is the primary pigment that drives the exchange of energy required for sugar production through photosynthesis, which ultimately sustains life, produces oxygen, and regulates CO_2 for the entire planet. From the interaction of visible solar radiation with leaves (approximately 400–750 nm), around 85% is absorbed by leaf pigments to fuel the photosynthesis processes, 10% is reflected, 2% is emitted as fluorescence, and the rest is transmitted (Lambers and Oliveira, 2019). However, this balance can vary depending on the chlorophyll content and concentration throughout the plant developmental phases, which itself is subject to environmental factors that influence physiological responses like growth, structural changes, and stress. The importance of Chl quantification, beyond its inherent ecosystem value, is widely documented in the agricultural literature, with efforts exploring its role in underpinning gross primary productivity (Houborg et al., 2015a), leaf nitrogen monitoring (Schlemmer et al., 2013), assessing health status (López-López et al., 2016), supporting fertilization management practices (Gabriel et al., 2017), and senescence detection (Noodén et al., 1997). Despite the importance of chlorophyll for phenotyping and agricultural purposes, accurately quantifying its temporal dynamics at different spatial scales (i.e., leaf, canopy, or field) remains a significant challenge, given the laborious and time-consuming sampling procedures required for its accurate characterization. From the diversity of methods available for examining leaf chlorophyll content, two of the most widely used include a destructive laboratory procedure based on *in vitro* spectrophotometric techniques (Wellburn, 1994; Porra, 2002; Netto et al., 2005) and a non-destructive method based on *in-situ* observations collected *via* chlorophyll meters, such as the Soil Plant Analysis Development (SPAD) system (Yuan et al., 2016; Shah et al., 2017; Dong et al., 2019).

However, despite the high accuracy provided by the laboratory method, or the portability offered by handheld sensors, both procedures face limitations when covering large study areas, where numerous samples are required to assess entire plant populations. An alternative and complementary approach to tackle this limitation is through combining field-based sampling (or scouting) with remote sensing based observations. Total chlorophyll can be tracked by its reflectance response using optical sensors (Curran, 1989), which can detect spectral absorption peaks within the visible wavelengths, centered at the 400–450 nm range and around 680 nm for Chl-a, and at the 450–500 nm range and around 650 nm for Chl-b. In recent decades, progress has been made in using multispectral satellite observations in combination with field data to estimate a range of "greenness" indices and gross primary productivity using MODIS (Wang et al., 2020), Landsat (Croft et al., 2015; Houborg et al., 2015b) and Sentinel-2 (Clevers and Gitelson, 2013; Delloye et al., 2018) platforms. Other initiatives have explored space-borne hyperspectral imagery from the EO-1 Hyperion sensor, tracking yield dynamics based on Chl content and leaf area index (Wu et al., 2008; Houborg et al., 2016). However, recent advances in miniaturizing optical sensors and systems, which can capture high spatial, spectral and temporal resolution data, offer new research opportunities to progress open questions in retrieval models and dynamics of Chl at both leaf and canopy scales. For example, studies based on unmanned aerial vehicles (UAVs) coupled with hyperspectral cameras have examined pigments content estimation (Zarco-Tejada et al., 2013a) by replicating modeling approaches already implemented with satellite and airborne-base data. With the enhanced spatial and temporal resolution afforded by such systems (Aasen et al., 2018), these technologies also bring new challenges in terms of the computational efficiency required to process, model, and analyze the large volumes of data collected.

Translating these massive quantities of hyperspectral imagery and *in-situ* data into useable information and knowledge requires improved and targeted modeling strategies. Early studies using UAV-based imaging spectroscopy were often focused on monitoring and characterizing croplands, retrieving Chl content, and other specific physiological properties using a range of methods. Broadly speaking, these approaches can be grouped into parametric, machine learning (ML), radiative transfer models (RTM), or hybrid methods (see Verrelst et al., 2019 for a full review). Parameterized relationships between spectral bands sensitive to physiological traits, more generally referred to as vegetation indices (VIs), are probably the most common approach to map pigments content (Haboudane et al., 2002), with examples including the Photochemical Reflectance Index (PRI; Zarco-Tejada et al., 2013b), the optimized soil-adjusted vegetation index (OSAVI) and the modified chlorophyll absorption in reflectance index (MCARI; Domingues Franceschini et al., 2017), among many others. Statistical regression approaches are routinely employed to capture relationships between spectral features and biophysical traits. For example, one of the more widely used linear methods is partial least squares regression (PLSR), which has been implemented to simultaneously estimate Chl and LAI (Kanning et al., 2018). ML regression algorithms have become increasingly popular due to their diversity of model types and utility for analyzing large datasets, with examples including random forest, and support vector machines. Bayesian algorithms, such as the Gaussian process regression (GPR; Rasmussen and Williams, 2006; Verrelst et al., 2012; Camps-Valls et al., 2016), have gained popularity in remote sensing applications due to their capacity to measure uncertainty and include prior knowledge about the modeled variables by using kernel functions. Together with GPR, an ensemble of multiple algorithms (Feilhauer et al., 2015; Vanbrabant et al., 2019) has been shown to outperform what can be achieved from application of any single method. Finally, more recent developments have sought to combine elements from the approaches mentioned above, resulting in hybrid methods that have the advantage of complementing the biophysical properties of VIs and RTMs with the computational efficiency and flexibility of non-parametric models, especially when dealing with large datasets (Capolupo et al., 2015). Hybrid-combinations remain an open and promising research path for phenotyping at canopy and leaf–level, with

applications including training ML regression approaches with simulated VIs retrieved by RTMs (Liang et al., 2016; Houborg and McCabe, 2018) or producing ensembles of dimensionality reduction (DR) and MLR methods able to filter critical spectral predictors (Rivera-Caicedo et al., 2017; Shah et al., 2019) to boost hyperspectral derived results.

Machine learning has shown considerable potential for delivering novel insights in leaf Chl retrieval, yet there are numerous implementation challenges that can frustrate application, including algorithm choice, training data, learning strategies, and predictors selection (Verrelst et al., 2019). Identifying the right algorithm among the many available depends on evaluating elements such as accuracy, interpretability, complexity, scalability, and computational cost: pre-analysis steps that are not always followed. Some approaches make particular assumptions about the data structure (i.e., distribution), demanding an exhaustive exploratory data analysis prior to modeling. Establishing a learning strategy has important implications for making the most of limited training data and prediction purposes. For instance, a data integration strategy for multi-temporal observations is required to understand how dynamics in Chl content combine to affect spectral responses. Likewise, the idea of strengthening the predictive power by using hundreds of spectral bands as predictors may result in computationally expensive models and multicollinearity issues, thus requiring coupled DR methods or testing transformed variables that bolster the spectral features sensitive to Chl (i.e., VIs). Overall, although there is not a generic recipe that can be applied to most ML problems, a modeling framework that integrates the above-highlighted aspects provides a much-needed road-map for retrieving biophysical variables from hyperspectral data.

The present study aims to assess the robustness of a machine learning framework to map a metric of leaf-Chl through the use of multi-temporal ultra-high spatial resolution (e.g., order of millimeters) UAV-based hyperspectral imagery. This is done using a training dataset composed of multi-temporal *in-situ* SPAD observations, together with VI estimates from field-based spectra measurements (350–2,500 nm) at the leaf level. Coincident and high-spatial resolution UAV-based hyperspectral scans (400–1,000 nm across 272 continuous bands) were also collected to provide a spatially distributed extension to the point scale *in-situ* training collections. A novel aspect of this work is that the modeling framework provides strategies for selecting the best-suited training/retrieval combinations based on accuracy assessment, using multiple learning models, spectral bands, and VIs predictors, and performed under both sequential and retraining learning techniques. Three particular research objectives are explored in this study: (1) examining the ability of different training strategies (retraining versus sequential) to capture and exploit temporal correlations in leaf-Chl; (2) quantify any potential gain (*via* a feature selection method) from using pigment-based VIs as Chl content predictors versus individual spectral bands; and (3) evaluate the performance of different ML regression approaches to accurately model and retrieve a SPAD-based Chl metric under dynamic training and field conditions.

MATERIALS AND METHODS

Study Area and Experimental Design

As part of a phenotyping study of a wild tomato (*Solanum pimpinellifolium*) crop, data were acquired over an experimental farm located in the valley of Hada Al-Sham, approximately 250 m above sea-level and 60 km east of Jeddah, Saudi Arabia (**Figure 1**). The regional climate is tropical and subtropical desert, with annual rainfall averages of around 100 mm. Although there was no rainfall during the study period, several sandstorms occurred through the growing cycle. 1,200 individual tomato plants were planted across the field, spaced equally at 1.5 m intervals and comprising 60 rows aligned along the north-east direction at approximately 2 m separation. The area was divided into four plots, containing a total of 300 plants each. The substrate was predominantly sandy loam soil. Five field campaigns were conducted between the fall and winter seasons (from November 2017 to January 2018), capturing crop growth stages corresponding to establishment, development, flowering, fruiting, and pre-harvest. As the primary purpose of the original phenotyping study was to identify salinity tolerance within the chosen selections, the experiment included duplicate saline and freshwater irrigation across four sub-areas (see Barreto et al., 2019 for further details). Here we focus our analyses on a single quadrant in order to reduce the computational burden involved in multi-quadrant processing (considering the terabytes of imagery involved). Ultimately, results can be expanded to the rest of the field to explore the impacts of salt-stress.

Field Spectra Data Collection

During each campaign, field-based reflectance spectra were collected close to solar noon using an ASD FieldSpec-4 (Analytical Spectral Devices Inc., Boulder, CO, United States) spectroradiometer, which samples data in the visible (VIS) and shortwave infrared (SWIR) spectral range (from 350 nm to 2,500 nm), with a resampled spectral resolution of 1 nm. From the total population of 1,200 plants, 36 individuals were randomly selected, and the reflectance response from three of their top leaflets measured (i.e., 108 samples for each campaign). Eleven sampling plants died before the last campaign (20180114, pre-harvest) due to strong winds, reducing the number of samples from 108 to 75 leaflets. An 8-degree fore optic lens was attached to a pistol grip to limit the field of view (FOV) diameter to 1.5 cm, measuring at a constant 10 cm zenith distance from each leaflet, which was placed on a black background. A white spectralon reference panel was used to calibrate the spectral measurements during the collection process. Five reflectance measurements were recorded for every leaflet, averaged, and spectrally resampled from 400 nm to 1,000 nm to match the spectral resolution of the UAV-based hyperspectral imagery (272 bands; see "Hyperspectral Imagery Collection and Calibration") by using a Gaussian model based on the FWHM spacings and wavelengths information from the hyperspectral camera in the software ENVI.

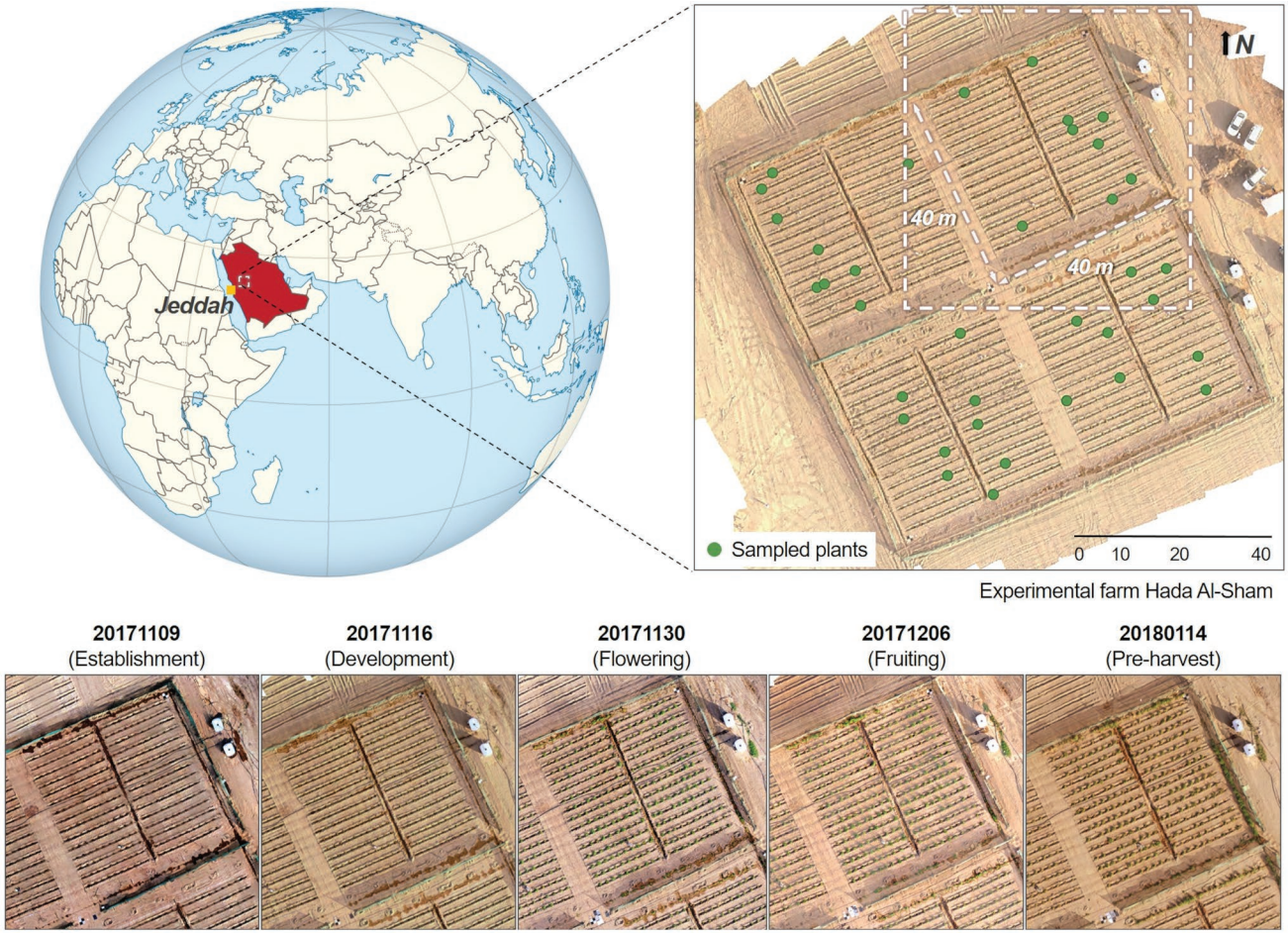

FIGURE 1 | Temporal maps of the studied quadrant at the Hada Al-Sham experimental facility (Lat. = 21.797°, Long. = 39.725°), where a wild tomato species was cultivated. Five different collections were performed between November 2017 and January 2018. The globe map is part of a series of SVG locator maps of countries including elements that have been adapted from the file: Afro-Eurasia on the globe (red).svg, and distributed under CC-BY-SA-3.0 license.

Ground-Truth Data Sampling

Non-destructive measurements of relative chlorophyll content (Chl) per leaf surface area were collected from the same spectrally sampled leaflets each day between 9:00 and 11:00 am local time, using a handheld SPAD-502 optical chlorophyll meter (Konica Minolta, Inc., Osaka, Japan). The operation of the SPAD meter is based on light transmittance at red and near-infrared wavelengths through a plant leaf. The instrument has two LEDs, one of which emits red radiance at 650 nm, with the other emitting near-infrared radiance at 940 nm. Most of the red light is absorbed by plants for photosynthesis, whereas longer near-infrared light passes through the leaf or is reflected. The ratio of transmittance at the near-infrared and red wavelengths is estimated and expressed as a unitless indicator, commonly referred to as SPAD units, which can range between 0 and 50 under standard measurement conditions (relative humidity <85% at <35°C) with a ± 1 unit accuracy, and up to 70 under high humidity/temperature conditions, with a drift of ±0.04 units per °C (Konica Minolta, 2009). Several studies have demonstrated near-linear and mostly exponential correlations between SPAD values and leaf chlorophyll content, although they can vary among species and growth habit groups (Markwell et al., 1995; Uddling et al., 2007; Cerovic et al., 2012; Parry et al., 2014; Shah et al., 2017). Since chlorophyll is not uniformly distributed in leaves and the device covers a small area of 6 mm^2 (2 × 3 mm) per measurement, the SPAD average of five different locations across each leaflet surface was considered as a metric of its chlorophyll content.

Hyperspectral Imagery Collection and Calibration

Spatially dense hyperspectral imagery was collected using a Nano-Hyperspec (Headwall Photonics, 2020a) push-broom camera integrated onboard a DJI Matrice 600 (M600) hexacopter (DJI, 2020). The Nano-Hyperspec was fitted with a 12 mm lens that afforded a horizontal field of view (FOV) of 21.1°, and collected data across the 400–1,000 nm spectral range in 272 continuous bands, with a 6 nm full-width half-maximum (FWHM). Flights were performed close to solar noon under clear sky conditions for all campaigns (see **Figure 1** for specific dates), with a view

zenith angle of zero and at an altitude of 15 m above the ground. Raw data were translated into radiance values using Headwall's SpectralView package (Headwall Photonics, 2020b), including the specific sensor calibration files for each band (Barreto et al., 2019). Automated georectification and mosaicking were performed to obtain geometrically accurate data-cubes with a ground sampling distance (GSD) of 0.007 m. Further details on the geometric calibration process can be found in (Angel et al., 2020). For the spectral calibration, at-sensor radiance data-cubes were converted into surface reflectance by performing an empirical line correction method (Wang and Myint, 2015), which estimates a linear regression for each band, matching ground-truth reflectance with its correspondent radiance spectra in the hyperspectral image. Following the procedure of (Barreto et al., 2019), reflectance data were collected from six near-Lambertian gray-scale panels (60×80 cm) placed in the middle of the field before each hyperspectral scanning. The ASD FieldSpec-4 bare fiber optic (25°) was attached to the pistol grip measuring at a constant 50 cm zenith distance, limiting the FOV diameter to approximately 22 cm. A chess-patterned target and soil reflectance measurements taken from across the field were also used for validation. Finally, reflectance mosaics were spectrally enhanced by applying a pixel-based Savitzky–Golay smoothing filter (Savitzky and Golay, 1964) and running a de-noising process with the minimum noise fraction (MNF) transformation approach (Green et al., 1988), in order to attenuate any artifacts that may lead to distorted spectra shapes affecting the data reliability (Ruffin and King, 1999).

Extraction of Vegetation Indices

Vegetation indices (VIs) are mathematical formulations of spectral bands that are widely used to quantify structural, physiological, and biochemical plant characteristics. These relationships are based on established correlations between reflectance spectra features and specific phenotypic traits. In addition to band specific ratios, narrowband (or hyperspectral) VIs often combine many continuous bands to capture spectral profile features, such as slopes, curvatures, and absorption depths (Thenkabail et al., 2019). Our study investigates 60 significant VIs that have previously been reported in the hyperspectral literature to be correlated with leaf chlorophyll content at various stages. For instance, from the Index DataBase (Henrich et al., 2009; one of the most comprehensive online resources) we selected VIs that were reported in studies conducted with spectrally similar sensors (i.e., covering the 400–1,000 nm range, including CASI550 and PHI, with 288 and 244 bands, respectively). We also include the Chl indices used by (Zarco-Tejada et al., 2019) to generate large-scale chlorophyll content maps, as well as the summary of derivative VIs in (Thenkabail et al., 2019). In addition, the VIs explored by (Shah et al., 2019) to retrieve leaf Chl in wheat, and those studied by (Houborg et al., 2016) to detect leaf Chl dynamics from hyperspectral satellite imagery were added. In total, 145 unique spectral bands were used in the formulations of the 60 indices, which were arranged into nine groups based on their calculation of similar phenotypic properties (see **Supplementary Table 1**, including formulations and key citations, and **Figure 2**). The field spectral profiles and the hyperspectral imagery were used to calculate the VIs at a leaf scale and pixel level, respectively.

MACHINE LEARNING MODELING WORKFLOW

In this study, a variety of regression approaches were evaluated under a proposed machine learning framework for multi-temporal mapping of leaf chlorophyll content (in SPAD units). The retrieval process combines five steps, including feature selection, learning different methods, cross-validating each model, assessing their performance, and mapping the SPAD predictions. SPAD data described in "Ground-Truth Data Sampling" are used as the response variable y, and the predictor variables x are derived from the field spectra samples ("Field Spectra Data Collection"). Two training strategies are tested: one considers sequential learning, and the other a time-series or retraining prediction (Dietterich, 2002). In sequential learning, the entire sequence of ground-truth observations is used to train each model and make all the predictions. In retraining prediction, models are cumulatively trained or retrained, and predictions are retrieved with the sampled data starting from the first stage t_1 up to a time t_1 (i.e., $i=5$ growth stages). Retraining implies a repetition of the workflow that generated the previously fitted model, but based on a new training dataset that reflects the most recent and current status of the plants, which is composed of the previous data and the new data (i.e., t_1, t_1+t_2, t_2+t_3, t_3+t_4, t_4+t_5), thus re-fitting the model while keeping its underlying architectural components (i.e., predictor variables, hyperparameters). This is an important strategy to explore because models can be retrained progressively with newly sampled data.

The learning workflow (see **Figure 3**) starts with the selection of predictor features, where each model is trained by using either all spectral bands (272 bands) or the set of vegetation indices (60 VIs), derived from the field spectra samples, allowing an investigation of the correlation and relevance of these variables as predictors. In addition, the subset of 145 bands that are used in the calculation of the various VIs were considered to examine any gain from transforming the spectral bands into VIs and the capability of the models to capture relevant and unknown relations between these selected bands and SPAD levels. The framework is evaluated with the most common nonparametric ML regression methods reported in retrieving biophysical variables from remote sensing applications (Verrelst et al., 2019). It is worth noting that the word nonparametric does not imply the lack of parameters, but that such parameters are adjustable and can be tuned by minimizing the estimation error while training takes place. In order to identify an optimal model structure, a total of 17 algorithms from three main categories are trained (e.g., linear-based, decision tree-based, and kernel-based). These include multivariate linear regression, partial-least-square regression (PLSR), decision trees, ensemble trees, support vector machines (SVM), and Gaussian processes regression (GPR). Kernel-based methods allow for an exploration of different types of mathematical functions (or kernels) to model the unknown or non-explicit relationships in the input data under a specific kind of function. For instance, those most commonly used in Earth Observation (EO) studies include linear,

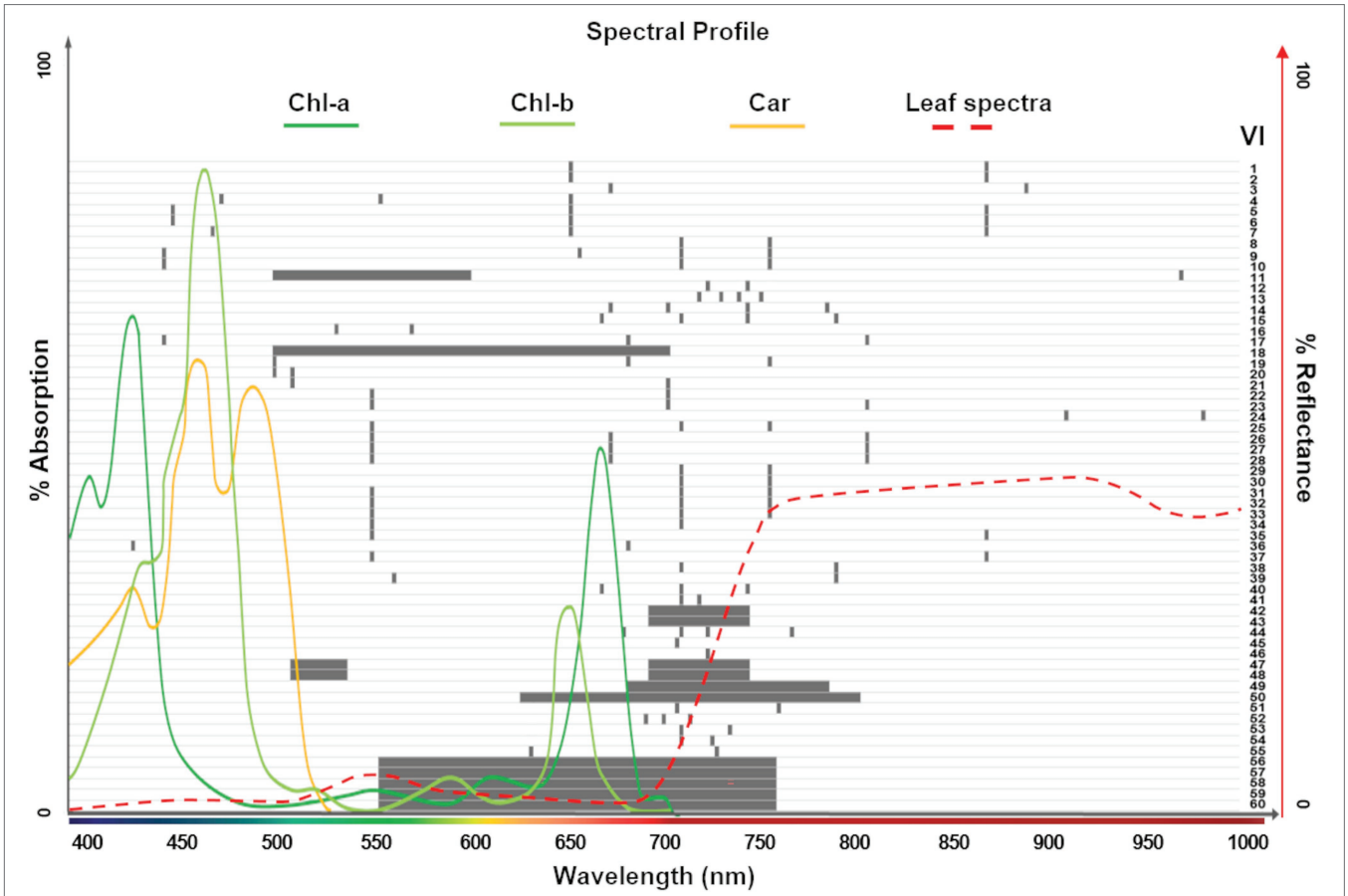

FIGURE 2 | Spectral ranges covered by the 60 vegetation indices (VIs) explored in this study (gray bars; see **Supplementary Table 1** for complete list). The absorption spectrum for photosynthetic pigments (Chl-a, Chl-b, and Car; dark green, light green and yellow lines, respectively; Lambers and Oliveira, 2019) and the typical leaf reflectance spectra (red dashed line) overlapping these ranges show the dominant pigments considered for each index. A total of 145 unique spectral bands were used in the formulations of the 60 VIs.

polynomial, and radial basis function (RBF) for SVM, and covariance-kernels for GPR (e.g., exponential, rational quadratic, RBF, and Matern; Camps-Valls et al., 2016; See more details in **Supplementary Material**). The architecture of some of these nonparametric approaches have embedded automated dimensionality reduction (DR) mechanisms, or band analysis tools (BAT; Rivera-Caicedo et al., 2017), to select relevant predictors, which is critical when dealing with hundreds of variables. For instance, the PLSR reduces the predictors to a smaller set of uncorrelated components, while the decision and ensembles of trees rely on pruning strategies, and GPR implicitly infers the feature's relevance from a length-scale parameter enclosed in the covariance functions. 80% of the input dataset is used to train and test each model under a 5-folds cross-validation routine, with their goodness of fit estimated using the R^2 metric (1) The remaining 20% of observations are used to identify the best performing model by assessing two prediction accuracy metrics: the root mean square error (RMSE), (2) the mean absolute error (MAE), and (3) Using these metrics, the most accurate model per method is employed to retrieve the multi-temporal SPAD predictions for each data-cube at a pixel-level. Since the pixel size is in the order of millimeters (e.g., 7 mm), the spectral profile of a pixel vector is assumed at a leaf scale. Using a plant delineation mask to exclude soil background (see Barreto et al., 2019 for further details), the best overall ML models are feed with the masked datacube to retrieve the SPAD maps and determine the most relevant predictors.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \hat{y}\right)^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \quad (2)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \quad (3)$$

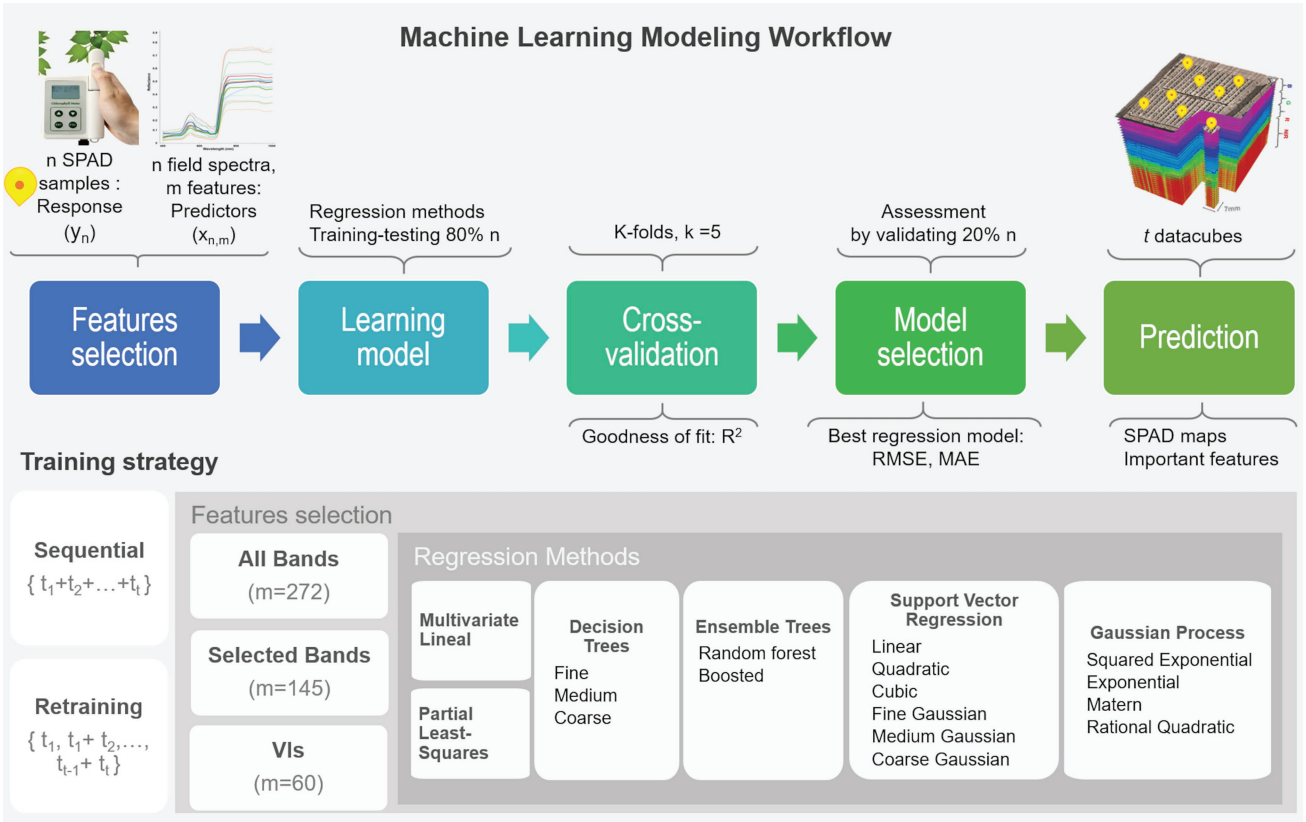

FIGURE 3 | Machine learning workflow for the retrieval of multi-temporal leaf chlorophyll dynamics using ultra-high-resolution UAV-hyperspectral imagery. The retraining loop should operate using the selected algorithms by starting from the first dataset (i.e., $t = 1$), then re-running the fitting process on the next training datasets, but using the previously fitted predictor variables and hyper-parameters (i.e., $t_1 + t_2$, $t_2 + t_3$). In this way, the model is updated as new training data is used in the learning process, and predictions are estimated accordingly.

In addition to retrieving SPAD maps, another useful output to examine from the selected methods is the importance of the features estimated by each approach while fitting the models. Feature selection approaches used for dimensionality reduction (DR), allow for the determination of each predictor's relevance in any particular model by scoring the features with a relevancy metric. They also help to better understand the dynamics between dependent and independent variables and enable a subset of less redundant features that could lead to model improvement (Mladenić, 2006).

RESULTS

Exploratory Input Data Analysis

The various modeling trials performed in this study, as outlined in the workflow described in **Figure 3**, evaluate both sequential and retraining learning strategies. Ground-truth data comprise a total of 108 observations of SPAD samples and leaf spectra per field campaign, randomly split into two subsets: training/testing (80%) and validation (20%), assuring their distributions are as similar as possible. SPAD observations, considered as the prediction response, are variable, symmetrically distributed, and rising across the growing season (**Figure 4A**). For the first date, 50% of the data ranged between 29 and 38 SPAD units, with a median of 33. Half of the samples reached a higher median of 49 for the second date, within a range of 45–54 SPAD units. For the third campaign, the observations have a narrower distribution than the previous collection, with a minimum and maximum value of 38 and 67 units, respectively, although reaching a slightly higher median of 52 units. SPAD data for the fourth campaign were more widely distributed, with half ranging between 48 and 59 units, with a median value of 53 units. For the last date, observations were less variable, spanning between 40 and 66 SPAD units, with a median of 55 units. Although training/testing and validation datasets show a slightly different distribution, their median values follow a similar trend over time (**Figure 4A**).

Each of the 272 bands from the resampled field spectra data are considered as an individual prediction feature. **Figure 4B** shows the multi-temporal spectra mean and their standard deviation, which in general follow a similar pattern in the blue (450–510 nm) and red edge (660–730 nm) regions, although differing along the green (510–660 nm) and near-infrared (740–1,000 nm) wavelengths. For the first date (20171109), the maximum average green and NIR reflectance reached 18 and

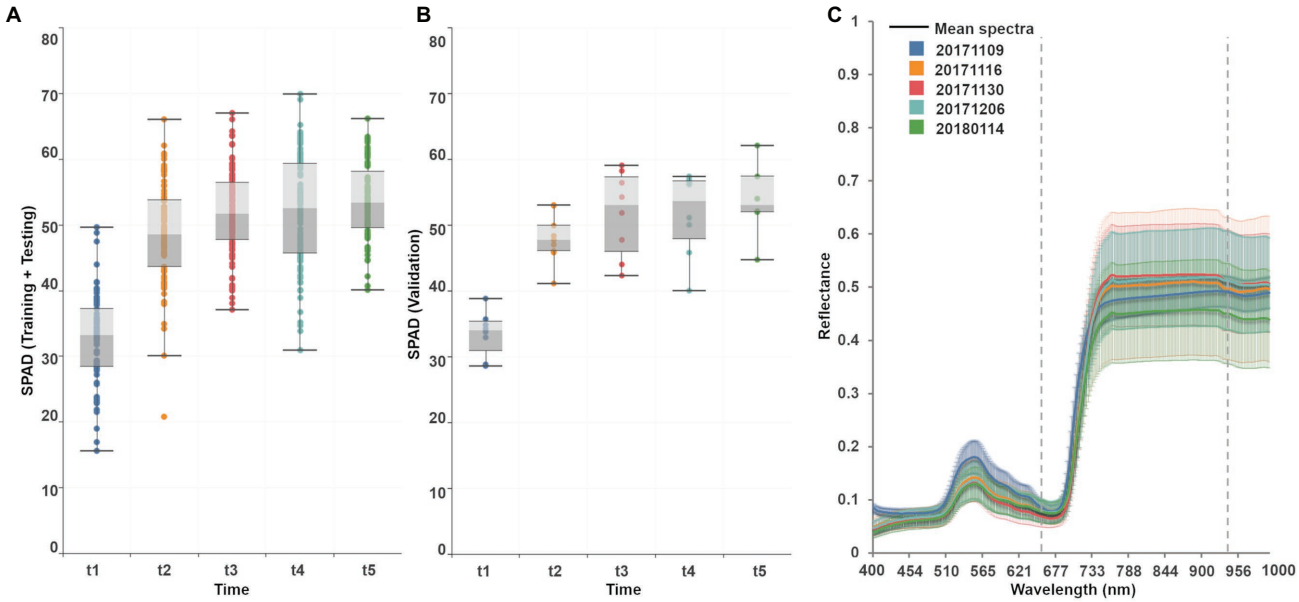

FIGURE 4 | (A) Traditional boxplots show the multi-temporal distribution of SPAD observations divided into training/testing (left) and validation sets (right). Median values are shown as the horizontal lines near the box centers, and the quartiles are delimited by the horizontal lines above and below the median. Whiskers indicate the variability outside upper and lower quartiles, locating the maximum and minimum scores at the top and bottom ends, respectively. An outlier observation (21 SPAD units) is shown during the second campaign (t_2) in the training/testing dataset. **(B)** Reflectance profiles collected per field campaign (format YYYYMMDD). The continuous lines indicate the mean spectra, and the filled areas define the range of reflectance measured each time. Dashed lines indicate the wavelengths where reflectance data reaches the lowest (~650 nm) and the highest (~950 nm) variability.

49%, with a standard deviation of 2 and 3%, respectively. For the second, third, and fourth campaigns, the mean green peak decreased to ~15%, with a standard deviation of ~3%, and the mean NIR response increased to ~50%, with a standard deviation of ~10%: although the NIR response during the second date is up to ~13% higher. For the last date, green reflectances were similar to previous collection times, yet the average NIR responses dropped to 45%, with a standard deviation of ~10%.

Pearson correlation matrices were calculated to visualize any multicollinearity in both predictor sets, bands and VIs (**Figure 5**). In the case of spectral bands, SPAD data has a low negative correlation ($-0.2 < r < -0.3$) with bands in the blue (400–510 nm) and red ranges (650–680 nm), but more highly negatively correlated ($r < -0.6$) with the green spectral window (510–650 nm) and the red-edge bands (680–730 nm). In contrast, near-infrared bands (730–1,000 nm) are weakly correlated ($r \approx 0$) with SPAD, but with high multicollinearity between them ($r > 0.8$): recognized as the Hughes phenomenon, which is thoroughly documented in the literature (Thenkabail and Lyon, 2012). Concerning the VIs (**Supplementary Table 1**), the narrowband greenness indices have a stronger correlation with SPAD than the broadband greenness indices. In contrast, photosynthesis efficiency, senescence, and pigments–based indices are weakly correlated with both SPAD and the other indices. For the set of VIs that evaluate leaf chlorophyll based on reflectance and derivative spectra, most show a high correlation with SPAD, while several (10 out of 31) have low inter-correlation with the rest. Finally, the group of continuum-removed VIs exhibit a low correlation with SPAD, with two negatively correlated with the other indices. Overall, although some of the indices are strongly inter-correlated ($r > 0.8$), this is only a measure of the association between them, not their causation. All the bands and VIs were included as predictor variables to let the MLR methods evaluate their relevance in predicting leaf chlorophyll.

Multiple Model Regression Assessment

Three different metrics, R^2, RMSE, and MAE, were used to undertake a comparative accuracy assessment of the models. R^2 was used to assess the performance of the models from the *k*-fold cross-validation, and RMSE and MAE to evaluate their actual accuracy by using ground-truth validation data. **Figure 6** summarizes these metrics for all the algorithms tested under the two training scenarios and three sets of predictors: all bands, selected bands used for calculating the VIs, and VIs.

For the sequential strategy case, PLSR coupled with band predictors achieved the best fitted models, either using all the bands ($R^2_{\text{AllBands}} = 0.80$) or the selected subset ($R^2_{\text{SelectedBands}} = 0.89$); although retrieving more accurate results when using all of them ($\text{RMSE}_{\text{SelectedBands}} = 5.45$, $\text{MAE}_{\text{SelectedBands}} = 4.27 > \text{RMSE}_{\text{AllBands}} = 4.40$, $\text{MAE}_{\text{AllBands}} = 3.50$). As can be seen in **Supplementary Table 2**, the various GPR models, along with SVRLinear and SVRMediumGaussian, were the best-performing algorithms across both types of predictors, reaching low RMSE and MAE around

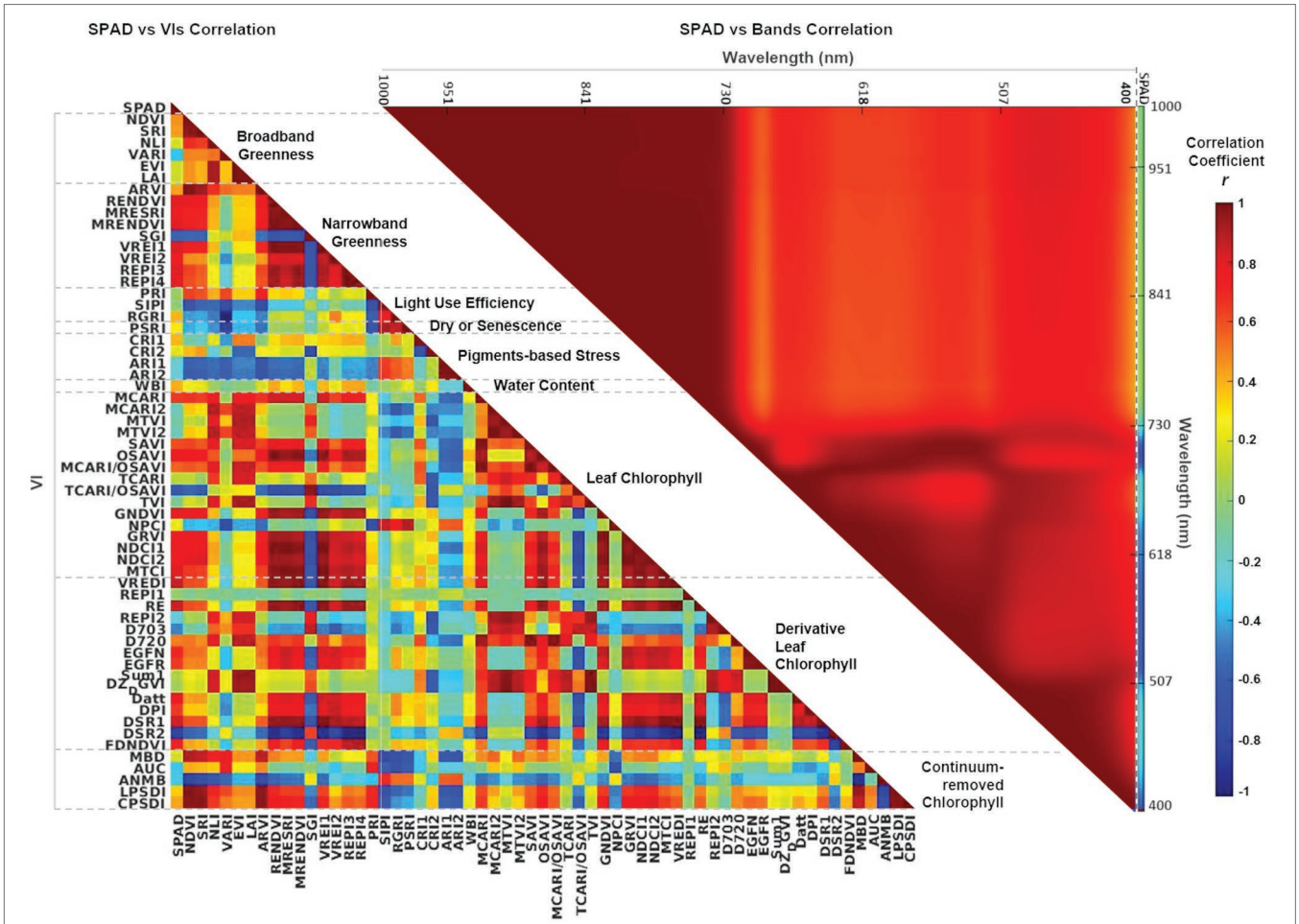

FIGURE 5 | Correlation matrices showing Pearson's *r* coefficient between SPAD and individual VIs (left-hand side), and between SPAD and individual spectral bands (right-hand side). The first column in the SPAD versus VIs matrix (left) and the last column in the SPAD versus Bands matrix (right) show the coefficients for the dependent variable SPAD. The color scale represents correlation coefficients between 1 and – 1. VIs are grouped and organized according to **Supplementary Table 1**.

5 SPAD units when using spectral bands and below 3 SPAD units for the VIs case (and with an R^2 above 0.7). In comparison, multivariate linear regression and SVRCubic models produced the poorest results, with higher errors (RMSE > 5, MAE > 6) and R^2 below 0.3. Overall, the GPRSquaredExponential (i.e., with a squared exponential kernel) achieved the second-best scores when considering both types of prediction features ($RMSE_{AllBands} = 5.17$, $RMSE_{SelectedBands} = 5.09$, $RMSE_{VIs} = 2.28$, $MAE_{AllBands} = 4.06$, $MAE_{SelectedBands} = 3.96$, $MAE_{VIs} = 1.98$, $R^2_{AllBands} = 0.76$, $R^2_{SelectedBands} = 0.76$, $R^2_{VIs} = 0.77$). In general, most of the models coupled with band predictors reached comparable R^2 and accuracy results, except for PLSR and the multivariate linear approaches; whereas modeling with the set of VI predictors was more accurate and better fitted than performing with the spectral bands.

For the retraining strategy, accuracy and goodness-of-fit of the models were assessed date by date. Overall, improved RMSE and MAE metrics were achieved using band predictors when compared to the sequential strategy case, although the R^2 decreased gradually over time (**Supplementary Table 2**). Results using VI predictors provided models that were comparable (i.e., similar R^2) to those produced by band features, but also recorded a similar drop in R^2 through time and producing slightly less accurate predictions. Particularly for the last campaign (t_5), a considerable drop ($R^2 > 0.4$ for all models) can be explained by the reduced number of spectral leaflet samples used to train/test the models (see "Field Spectra Data Collection"). For the first date (t_1), the GPRSquaredExponential was the most accurate ($RMSE_{AllBands} = 2.14$, $MAE_{AllBands} = 1.73$) when paired with all the bands, and the best fitted ($R^2_{AllBands} = 0.86$, $R^2_{SelectedBands} = 0.75$, $R^2_{VIs} = 0.83$) using the three sets of predictors: although beaten by PLSR when coupled with VIs ($RMSE_{VIs} = 1.93$, $MAE_{VIs} = 0.86$). For the second campaign (t2), the GPRSquaredExponential again produced the most accurate results from the all bands-based case ($RMSE_{AllBands} = 2.71$, $MAE_{AllBands} = 2.13$), while the PLSR model was better fitted using the selected bands ($R^2_{SelectedBands} = 0.95$). In the VIs-based case, the PLSR was the most accurate ($RMSE_{VIs} = 2.56$, $MAE_{VIs} = 0.98$),

FIGURE 6 | A comparative assessment was performed between the 17 trained models by comparing three different metrics: R^2, RMSE, and MAE. Model-fit and accuracy were evaluated under both training strategies (sequential and retraining) and considering different sets of prediction features (all spectral bands, selected bands from VIs, and VIs). An average prediction error threshold of 5 SPAD units was established to evaluate individual model accuracy (dashed line). Base-10 log scale is used for the *x*-axis (RMSE).

but GPRSquaredExponential was better fitted ($R^2_{VIs} = 0.74$). For the third date (t3), the GPRSquaredExponential model achieved the highest accuracy ($RMSE_{AllBands} = 3.62$, $MAE_{AllBands} = 3.01$) under the all bands-based setup, and RandomForest the best fitting ($R^2_{AllBands} = 0.67$); whereas in the selected bands-based and the VIs-based case, PLSR was the best fitted ($R^2_{SelectedBands} = 0.75$, $R^2_{VIs} = 0.79$), and the most accurate ($RMSE_{SelectedBands} = 3.77$, $MAE_{SelectedBands} = 2.66$, $RMSE_{VIs} = 1.72$, $MAE_{VIs} = 0.78$). For the

fourth campaign (t4), the PLSR trained with the selected bands was the best-fitted model ($R^2_{SelectedlBands} = 0.97$), although it was poorly accurate ($RMSE_{SelectedBands} = 6.37$, $MAE_{SelectedBands} = 5.30$) to the others performance. In contrast, the GPRSquaredExponential was the second best-fitted model ($R^2_{AllBands} = 0.61$, $R^2_{VIs} = 0.64$), reaching the highest accuracy when trained with all the band features ($RMSE_{AllBands} = 4.49$, $MAE_{AllBands} = 3.99$). However, this model was exceeded by the RandomForest ($RMSE_{VIs} = 3.35$, $MAE_{VIs} = 1.53$) when using VI predictors. For the last date (t5), the RandomForest model shown the highest accuracy ($RMSE_{AllBands} = 5.59$, $MAE_{AllBands} = 4.37$) and best-fitting ($R^2_{AllBands} = 0.41$) when coupled with all the bands; whereas PLSR was the most accurate when trained with VIs ($RMSE_{VIs} = 2.89$, $MAE_{VIs} = 0.96$). Overall, sub-optimal results were retrieved using the selected set of bands, and all the models were poorly fitted with R^2 below 0.1.

Under both training strategies, the three sets of predictors showed comparable performance, yet most of the VIs-based models achieved better scores than the band-based ones. Slightly better results were reached by most of the models when trained with all the bands than with the selected ones (i.e., between ~0.1 and 0.4 units in R^2 scores), except for the PLSR. Also, despite being the weakest approach (with the highest RMSE and MAE, and the lowest R^2) across the different strategies and predictors, the multivariate linear (or linear) model scores were considerably improved when using the subset of selected bands (see **Supplementary Table 2**; **Figure 6**). For instance, the linear model improved relative to the medium tree model performance and was comparable to the random forest model results under the sequential strategy. Considering the slightly better scores retrieved by the full set of spectral bands than the set of selected bands, the models trained with all 272 bands were considered together with the VIs-based ones to perform the predictions and select the best among them to produce the SPAD maps.

Multiple Model SPAD Predictions

Based on the assessment metrics, a criteria model selection was established to map the SPAD predictions. Different factors, such as leaf water content and irradiance changes, can introduce between 2 and 4 unit biases in SPAD readings (Martínez and Guiamet, 2004), in addition to the instrumental accuracy of ±1 units. Accounting for these influences and equating the quality of the SPAD predictions with the SPAD readings, we defined an error threshold of up to 5 units for both metrics (RMSE and MAE) under the assumption that all of the errors would have the same magnitude, which is the only theoretical case when RMSE and MAE would be equal. Under these criteria, models with average prediction errors above 5 SPAD units were excluded from the final selection, and the best-fitting model from each approach was used to retrieve the predictions from the hyperspectral imagery (see **Figure 6**). Accordingly, PLSR, the medium trees, random forest, SVR linear, and GPR squared exponential models were selected together with the linear regression, which was only included for comparison purposes despite it showing inferior performance among all model configurations. These trained models were used to retrieve the SPAD values at a pixel level on the multi-temporal hyperspectral data-cubes, then averaged at a plant level to evaluate the spatial and temporal distribution of the predictions across the study area. **Figure 7** shows an array of the results organized by learning strategy, type of predictors, time, and types of models, where each box comprises a matrix of cells that represent the mean predicted SPAD per plant, following the same sowing arrangement of the field (rows × columns).

In the sequential strategy (**Figure 7A**), most of the models retrieved homogeneous SPAD maps across time, hence not showing significantly different changes during the growing season. Of note, the multivariate linear model produced results at the extremes when using different predictors, underestimating (≤ 20 units) when using bands, and overestimating (≥ 70 units) when using VI predictors. Similarly, PLSR estimates diverge under different predictor scenarios, realizing homogeneous result series (~30–50 units) through the band predictors while overestimating with VIs (≥ 60 units). In contrast, the medium tree model retrieved almost identical results for both cases, with SPAD values ranging around ~40 units: comparable with the performance of random trees when only using VI predictors. However, when using band features, random forest yielded similar results to SVR linear and GPR squared exponential models, with more variable retrievals between ~30 to 50 units (although the GPR-based map for the fourth date shows higher values around ~60 units). Using VI features, SVR-based results were more heterogeneous (around ~45 units) during the first two campaigns than the retrievals (≥ 60 units) during the last three dates. Overall, following a sequential learning strategy, GPR squared exponential predictions were the most dynamic across time, indicating its flexibility to learn and model temporal dynamics. However, this performance was degraded when using a more straightforward set of predictors like vegetation indices.

In general, for the retraining strategy (**Figure 7B**), the performance of the learning algorithms over both types of predictors was similar, with SPAD estimates varying over time and space - except those calculated by the multivariate linear regression and PLSR, which led to poor retrieval performance. Again, medium tree and random forest methods, combined with both predictor sets, reached matching results, increasing across time between ~30 and ~60 SPAD units during the first and last campaigns, respectively (although the VIs-based maps present slightly lower values than the bands-based ones). In the case of the SVR linear method, contrasting results were produced by each class of predictors, showing an ascendant trend across the time when using band features, comparable to the decision trees and GPR results. However, the SVR linear method underestimated results using VIs, even yielding values below 20 SPAD units for some plants during the last three campaigns (t_3, t_4, t_5). In contrast, the GPR squared exponential results are congruent with the decision trees-based predictions under both types of predictors, although different estimates were achieved for the second and last dates. When using band features for the last campaign, the GPR-based map overestimated results, with values above ~60 SPAD units.

Together with the mean SPAD estimates, GPR reports the standard deviation at 95% ($\sigma_{95\%}$) confidence interval for each

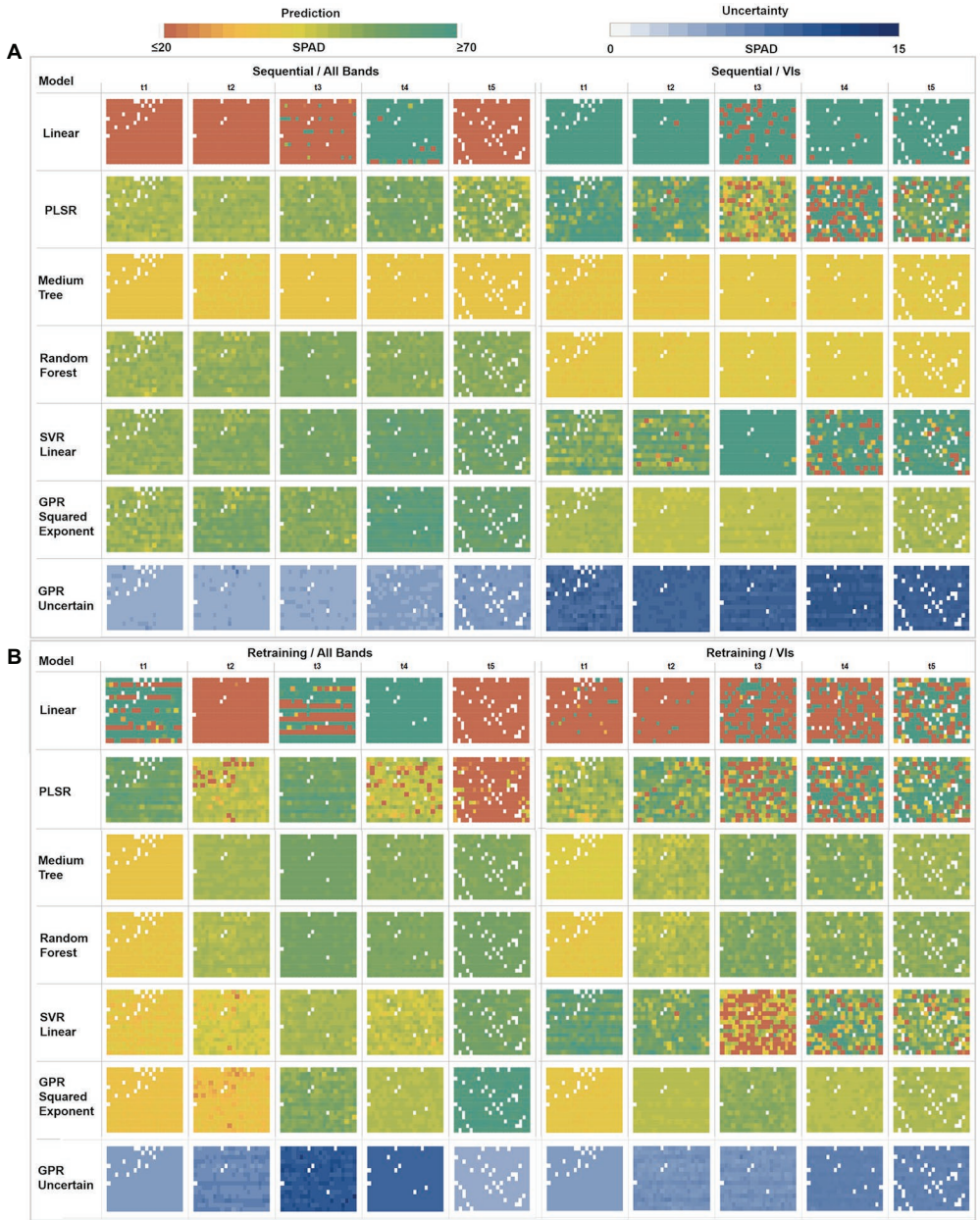

FIGURE 7 | Multi-model comparison of averaged SPAD predictions and GPR uncertainties at a plant level. **(A)** Sequential strategy results using All Bands versus VIs. **(B)** Retraining strategy results using All Bands versus VIs. Each box comprises a matrix of cells representing the mean predicted SPAD per plant, following the field sowing arrangement (rows × columns).

prediction, which is used as an uncertainty metric to assess the variance of the retrievals. For the sequential learning case (**Figure 7A**), homogeneous variances were achieved across the whole series, with low uncertainties between 4 and 6 SPAD units, when training with the band predictors, but high uncertainties ranging between 10 and 15 SPAD units when using VI features. In contrast, under the retraining strategy (**Figure 7B**), heterogeneous variance levels were observed throughout the series. For instance, an increase in the uncertainty was reported when using band features, starting from 5 to 6 units in the first date (20171109), then between 6 and 7 units during the second date (t_2), until reaching standard deviations between 9 and 11 units during the third collection (t_3). After this, the uncertainty levels decreased slightly to around 10 units on the fourth date (t_4), achieving the lowest variance in the last date (t_5) with 4 SPAD units. More stable variances were achieved using VI predictors, starting with low uncertainty levels between 5 and 6 SPAD units in the first three dates, with a minor rise of variance for the last two dates, with values between 6 and 8 units.

Model Selection

The identification of suitable learning algorithms and predictors is a critical step to develop accurate SPAD retrieval maps. Therefore, an additional evaluation of the previously selected methods was performed by comparing the retrieved and original distributions of the validation dataset. **Figure 8** presents the variability of SPAD predictions and *in-situ* collected measurements used for validation, with colored and gray delineated box-plots, respectively. As can be seen, the multivariate linear regression results were widely dispersed, either overestimated or underestimated, and divergent from the measured data distribution. In the case of PLSR, the distribution of the estimates reflected the validation reference only under the sequential training strategy using band predictors. Overall, most of the pre-selected algorithms achieved relatively consistent results, except SVR linear when combined with VI predictors.

For the sequential approach (**Figure 8A**) with band features, predictions from all models, except for the medium trees, followed the ascending trend across time observed in the *in-situ* data, although exceeding them by up to 15 units on the first date, if comparing their medians. The best distribution matches with the observations were obtained by random forest, and GPR squared exponential only on the third date (t_3). In contrast, when using VI predictors, overall median estimates were below the ground-truth data distribution by between 10 and 20 SPAD units for the decision trees models, and between 3 and 10 units for the GPR model.

For the retraining case (**Figure 8B**), the performance was generally higher relative to the sequential strategy and more consistent between the assessed algorithms. In particular, the distribution of the results from the medium trees and random forest were comparable, reaching similar medians, although with different dispersion, over the different campaigns. The GPR squared exponential model achieved better results when combined with VI predictors, following the SPAD observations trend, although with less dispersed distributions and reaching median values between 1 and 5 SPAD units below the *in-situ* data. The best predictions were achieved by combining random forest with VI features, retrieving minor discrepancies between 1 and 3 units in the median values compared to the SPAD measurements, and yielding matching spread distributions, as shown by the minimum and maximum values of each campaign dataset.

Based on the accuracy assessment and the analysis of the prediction distributions achieved by all the evaluated methods, three out of the 17 regression models were selected to retrieve the multi-temporal SPAD maps at a pixel scale. The selected methods included the PLSR using all bands under the sequential strategy, and random forest and GPR squared exponential, using vegetation indices as predictor variables under the retraining strategy.

Multi-Temporal Spatial Predictions of SPAD

The hyperspectral mosaics, and the vegetation index data-cubes derived from them for each campaign, were used as input to feed the three selected regression models and to retrieve the multi-temporal SPAD maps at a pixel scale. **Figure 9** depicts a comparison of the results achieved by each method over some of the sowed furrows, showcasing the differences between PLSR, random forest and GPR squared exponential estimates throughout the study period. For the first three campaigns, the PLSR model reached different results than the other two methods, with a slight increase from an average of 48 units in the first stage (t_1) to 50 units in the third (t_3), whereas random forest and GPR squared exponential reached similar estimates, with an increase in SPAD values from an average of 30 units in the first stage (t_1) to 55 units in the third (t_3). The PLSR and GPR-based estimates were uniformly distributed in leaves, changing over time without marked differences between stages, especially during the last two dates. However, some negative retrievals from the PLSR approach during the third and fourth campaigns can be seen as gap pixels in the showcased plants in **Figure 9**. On the other hand, random forest-based predictions differed from time to time, with a slight decrease in the last stage and showing more clustered estimates toward the center of the plants surface.

The plant growth dynamics can be described from the multi-temporal SPAD retrievals detailed in **Figure 9**. As can be observed, leaf chlorophyll content increases as plants grow and increase their leaf density from the establishment stage ($t = 1$), reaching a maximum value at the start of flowering ($t = 3$). Plants reach a mature state and produce fruits ($t = 4$), where SPAD estimates from PLSR increase slightly, whereas random forest results remain at the same level, and retrievals from GPR decrease slightly. At the pre-harvesting stage ($t = 5$), leaves and stems gradually age, turning yellowish, which is evident from the low SPAD levels predicted by all the methods. These dynamics are consistent with the distribution analysis described previously (**Figure 8B**), and illustrate how SPAD results from the three methods follow a similar temporal trend, albeit with GPR retrievals presenting a more homogeneous and tighter distribution than PLSR and random forest estimates. The uncertainty maps retrieved by the GPR algorithm were plotted to assess the variance of the predictions at the pixel level. As previously noted, the uncertainty of the estimates varies across the map series, starting with low (~4–6) standard deviations during the first three stages and slightly higher values during the fruiting and pre-harvesting phases (~6–8 units). The uncertainty of the predictions could come from either the propagation of uncertainties through time or high variations in the leaf-Chl levels' dynamics. Moreover, when zooming into the maps at a plant level, the spatially distributed uncertainty levels can be observed over the plant projected areas, although with some higher variances associated with either bright or shadowed pixels.

Important Features

The PLSR, random forest, and GPR approaches include feature selection mechanisms in their architecture to score the band and VI predictors base on their relevance toward the SPAD variable (see **Supplementary Material**). Various metrics are reported in each model: for instance, feature importance for the random forest, weakness index for GPR, and weight index

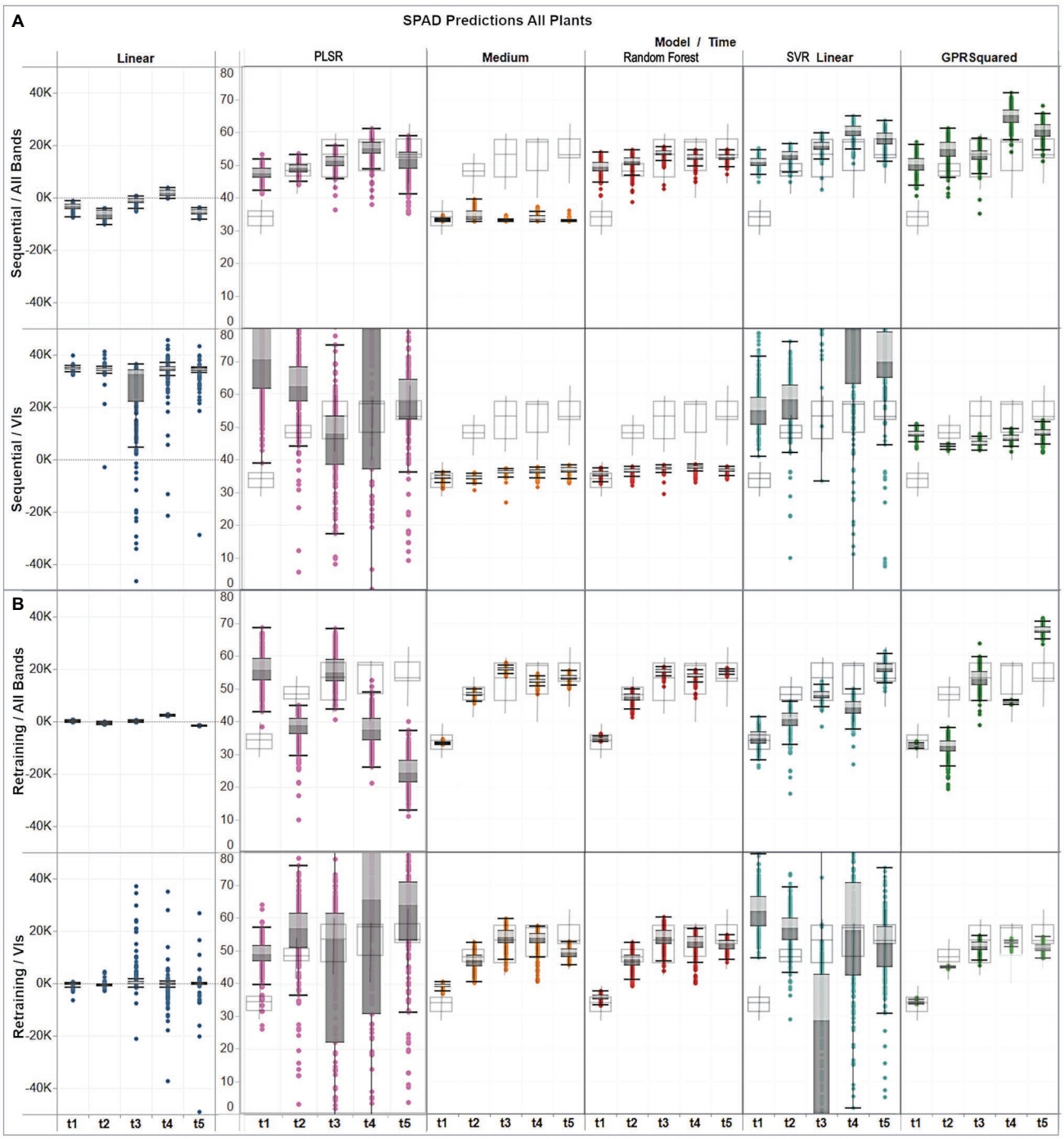

FIGURE 8 | Assessment of multi-temporal predictions distribution (gray colored boxes) against the actual distribution (gray delineated boxes) of the validation data, showing **(A)** distribution per model under the sequential strategy using band and VI predictors; and **(B)** distribution per model under the retraining strategy using band and VI predictors.

for PLSR (**Figure 10**). The PLSR weights are retrieved for each of the components used to fit the model, which describes how strongly each component depends on the original predictors. The number of components (i.e., seven) was tuned by minimizing the error of the predictions through cross-validation during the training/validation stage (**Figure 10D**). The predictor scores were extracted from the three selected models when trained under the strategy they perform the best (**Figure 8**). Thus, the fitted models under the sequential strategy were used for comparing the spectral

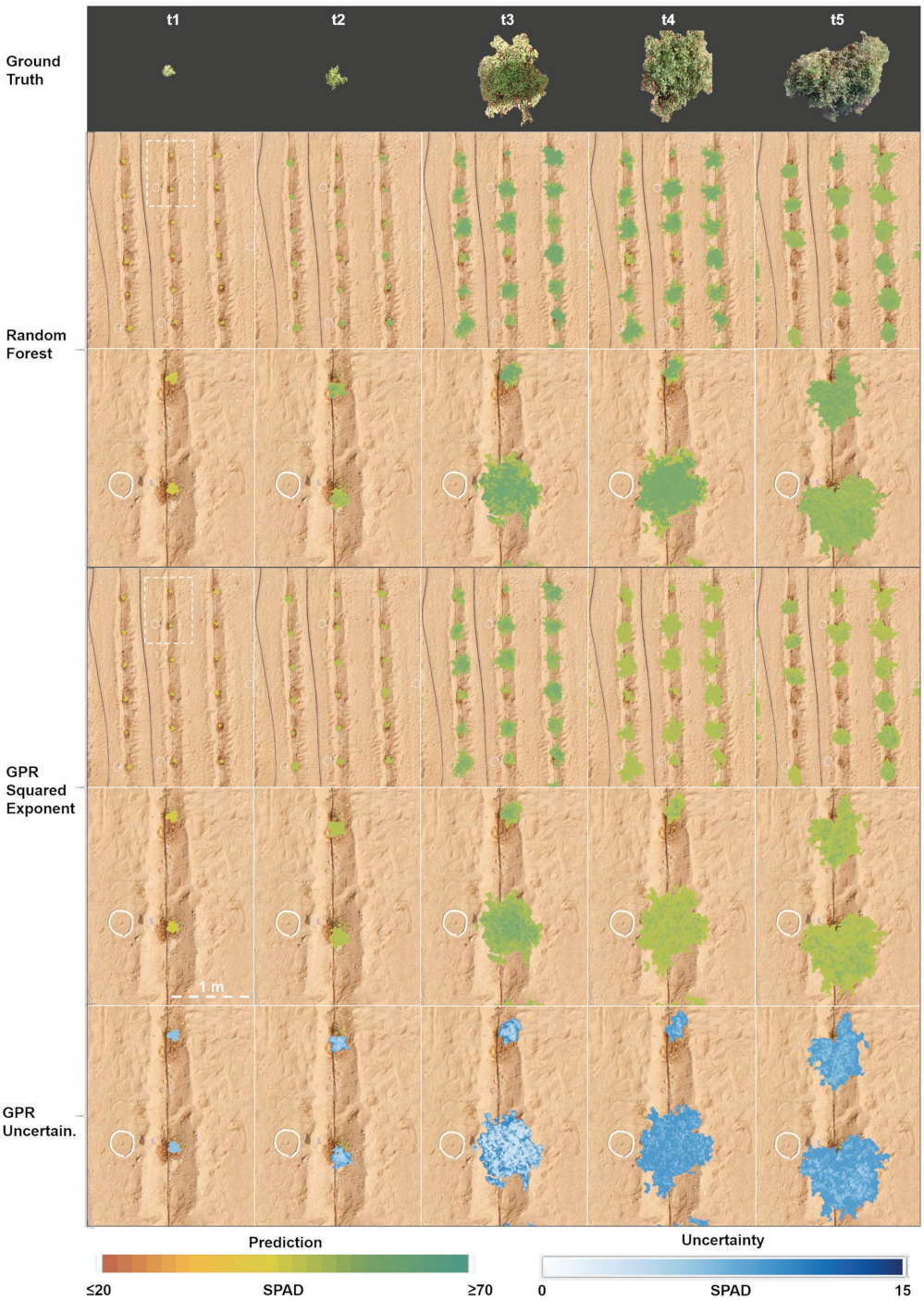

FIGURE 9 | Comparison of multi-temporal SPAD prediction maps generated with PLSR coupled with band predictors, random forest and GPR squared exponential models, using vegetation indices predictors (lower panels). True color-balanced pictures of a showcase plant (upper panel) depict changes throughout the growing cycle. The gradient-colored bars represent the estimated SPAD values in the range between 20 to 70 units, with a bin size of 2. The SPAD uncertainties are shown in the range between 0 to 15 units, with a bin size of 1.

bands' metrics, whereas their results under the retraining strategy were gathered for the VIs case.

Since each model has its own metric, their predictors' scores cannot be directly compared. Thus, we use a quartile classification to rank the top features of each method, where each quartile contains 25% of the total predictors. For PLSR (**Figure 10D**) and random forest (**Figure 10C**), the highest quartile (Q_3) was set as the threshold to denote the most important features,

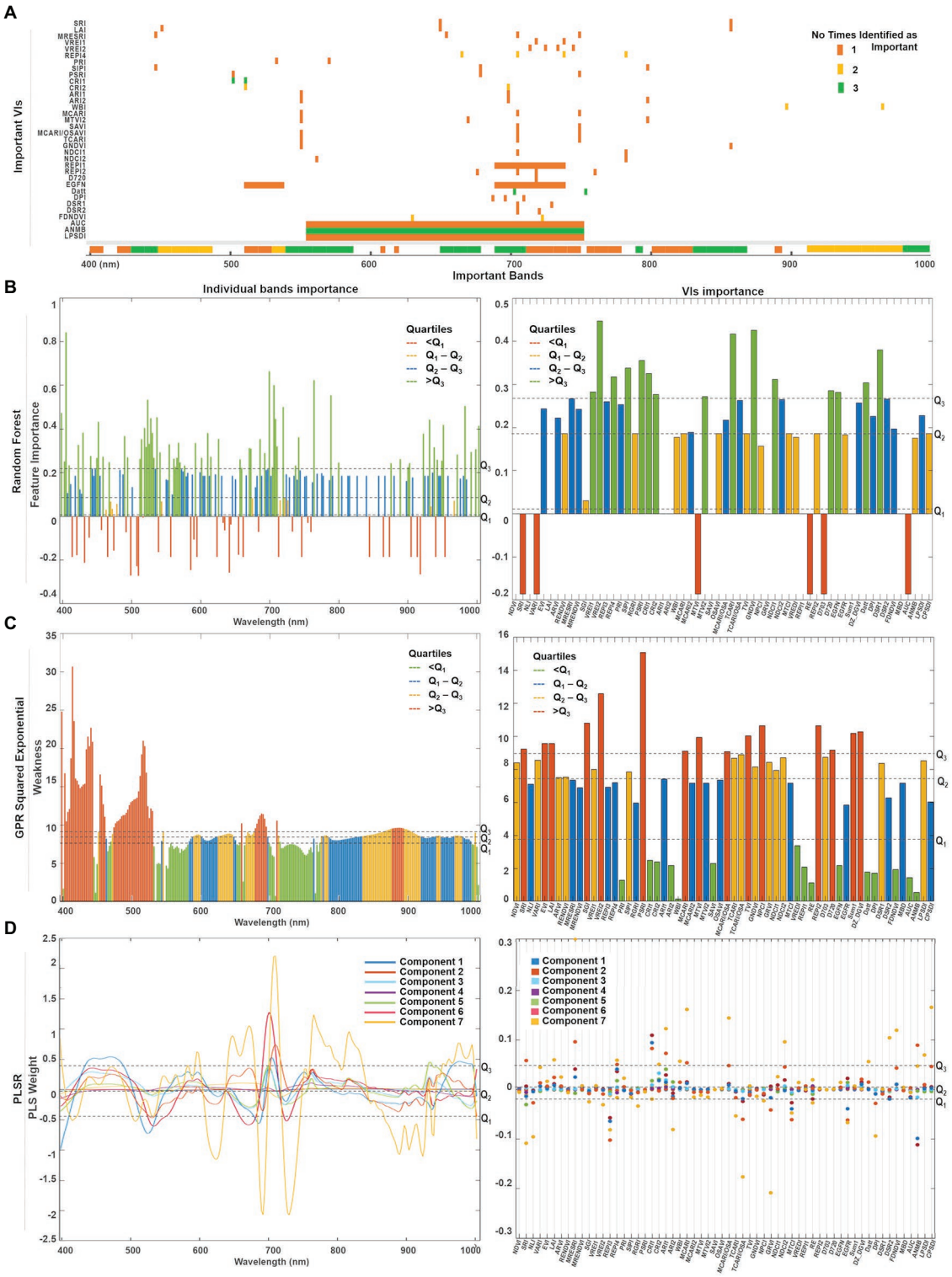

FIGURE 10 **(A)** Summary of the top predictors (bands and VIs) that were identified by either one (orange), two (yellow), or three (green) of the selected methods. A quartile classification was used to rank the top features of each method. **(B)** For random forest, the highest quartile was set as threshold to identify the most important bands ($Q_3 = 0.21$) and VIs ($Q_3 = 0.27$). **(C)** For GPR squared exponential, the lowest quartile was set as threshold to classify the less weak bands ($Q_1 = 8.02$) and VIs ($Q_2 = 7.16$). **(D)** The predictor weights were retrieved for the first seven components used to fit the PLSR model. The highest quartile was set as the threshold to denote the most relevant bands ($Q_3 = 0.48$) and VIs ($Q_3 = 0.04$).

whereas the lowest quartile (Q_1) was set for the GPR squared exponential (**Figure 10B**) to indicate the less weak (or most relevant) variables. A summary of the top predictors retrieved by either one, two, or three of the selected methods were highlighted (see **Figure 10A**), allowing the identification of those variables that might play a physically meaningful role in predicting Chl levels. The spectral information contained by the highlighted VIs can be traced in **Figure 10A**. Only three VIs contain information from the blue spectral region (~450 nm), followed by a small group of VIs that gather information from the green wavelengths (500 nm −550 nm). However, most of the important indices comprise the information from the red edge region (650–750 nm), and few individual VIs collect the information from narrow spectral ranges along the near-infrared (~800, 870, and 970 nm).

A combination of both individual bands and band ranges were identified as the most relevant for each model (**Supplementary Table 3**). Groups of continuous bands comprising less than 10 nm spectral range were counted as a single variable around the central band. For instance, the bands around 405 nm were identified as relevant by random forest and GPR, but weighted low by PLSR, although with a broader range for the random forest (i.e., 400–410 nm) where Chl-a is highly absorbed. The spectral variables between 450 and 490 nm were ranked as relevant by all the methods, coinciding with the most substantial Chl-b absorption. Bands from the green spectral range (i.e., 530 and 590 nm), where Chl reflectance peaks, were also highlighted as relevant by random forest and GPR, but not highly weighted by the first and seventh PLSR components. Toward the red wavelengths, random forest identified two narrow ranges (i.e., 610 and 620 nm), that agreed with GPR and PLSR in the Chl-a and Chl-b absorption crest (i.e., 650–670 nm). From 690 nm to 750 nm, the red-edge region was also highlighted as relevant by all models, although GPR identified two specific ranges: one spanning from 695 to 710 nm at the beginning of the red-edge, and another from 720 to 760 nm, where the red-edge inflection point shifts accordingly to Chl content. The three methods identified a critical thin region between 760 and 770 nm at the end of the red-edge. Near-infrared (NIR) bands indicated variable relevancy levels among the methods: PLSR highly ranked the region between 800 and 830 nm, while the flat sill on the plant reflectance spectra (i.e., 910–1,000 nm) was identified by random forest and PLSR, whereas GPR highlighted the very end of the NIR (i.e., 980–1,000 nm).

A total of 15 VIs were scored as relevant by each method (**Supplementary Table 3**; **Figure 10**). From the greenness indices, PLSR identified two of the broadband VIs (i.e., SRI and LAI) as significant and corresponding with random forest in selecting the narrowband red-edge position index REPI4. However, random forest highlighted two other indices of this category as relevant (i.e., VREI1, VRE2). The light use efficiency indicators (PRI and SIPI) were considered by GPR and random forest as critical, while the senescence index (PSRI) was only highlighted by random forest. In contrast, the stress-on-pigments production indices were highly scored by the three methods (i.e., CRI1, CRI2, ARI1, and ARI2). The water content index (WBI) was classified as relevant by the GPR and the PLSR models. From the 15 leaf Chl indices, random forest selected four of them (i.e., MTVI2, TCARI, GNDVI, and NDCI1), GPR selected just one (SAVI), while the PLSR model selected three (i.e., MCARI, MCARI/OSAVI, and NDCI2). From the derivative-based leaf Chl indices, all the methods highlighted the Datt index as critical, although random forest identified three more (i.e., D720, EGFN, and DSR1) and GPR identified another two (i.e., DPI and FDNDVI)—coinciding with PLSR in selecting the FDNDVI index. Finally, GPR and PLSR both identified ANMB from the continuum removed-based indices, whereas only GPR highlighted the area under the curve index (AUC) and the leaf plant stress detection index (LPSDI).

The feature importance analysis was extended to the models trained with the subset of 145 bands (i.e., the source of the VI predictors) to investigate how the feature selection operates on a smaller dataset. The same training strategies and ML approaches were used to retrieve, score, and classify the 145 selected bands, following the rationale presented previously (**Supplementary Figure 1**). **Supplementary Table 4** summarizes the relevant bands and spectral ranges identified from the 145 bands subset.

DISCUSSION

Sequential Versus Retraining Learning

Modeling physiological traits such as leaf Chl content throughout a crop growing season requires treating plant traits as continuous processes across time, which can be accounted for through implementing sequential and retraining learning strategies. Sequential learning is a common practice in remote sensing, wherein the full observed series is used to fit a single model assuming that the relationships between the prediction features and the independent variable remain fixed through time (Dietterich, 2002). In contrast, the retraining strategy uses a loop to learn a model progressively as new data is collected. Deciding whether to follow one or the other relies on the modeling problem and the data itself, since both are data-driven strategies after all. This study followed a simple and useful diagnostic suggested in the machine learning literature by examining the target variable distribution (Sculley et al., 2014).

Leaf chlorophyll and reflectance response change significantly through time, which is evident in this study by analyzing the distributions of sampled SPAD and spectral data (**Figure 4A**). Hence, the correlation between SPAD and predictors, either by bands or VIs, are dynamic as well. The temporal distribution of the SPAD validation dataset can be used as a reference to assess the coherence of the SPAD retrievals (**Figure 8**). In doing this, the accuracy metrics (MAE, RMSE) can be used jointly to determine the best candidate models to retrain (i.e., PLSR, random forest, GPR squared exponential). That is, the retraining loop should operate using the selected algorithms by starting from the first dataset (i.e., $t = 1$), then re-running the fitting process on the next training datasets (i.e., $t = 2$, $t = 3$), but using the previously fitted predictor variables and hyper-parameters. In this way, the model is updated as new

training data is used in the learning process, and predictions are estimated accordingly.

The predicted SPAD maps averaged at a plant level and retrieved under the retraining strategy (**Figure 7B**) are coherent with the SPAD distributions of the *in-situ* validation dataset (**Figure 8B**). Such a result demonstrates the capability of the selected models to learn from *in-situ* data using a retraining routine, and thereby enhances the capacity to turn UAV-based hyperspectral imagery into valid multi-temporal SPAD maps. It also proves the capability and flexibility of the retraining strategy to capture temporal dynamics in chlorophyll levels from series of hyperspectral imagery by fitting multi-temporal regression models and advancing, for instance, uni-temporal approaches that develop individual growth-stage models (Aasen and Bolten, 2018). Such a strategy also offers a solution to open questions raised in some related studies, where sequential learning was implemented to map Chl series from satellite (Houborg et al., 2016) and UAV (Vanbrabant et al., 2019) hyperspectral images, and advising further investigation in learning regression approaches capable of capturing subtle temporal dynamics linked to short-term variations in plant traits.

Most of the selected algorithms performed significantly better under the retraining learning strategy, although the goodness-of-fit (R^2) estimated from the training/testing dataset can be affected as new training data is introduced. Further experiments can be conducted to incorporate a time component (i.e., hyper-parameter, kernel) in the regression algorithms definition, which can be fully dedicated to capturing temporal correlations, and non-stationary behavior associated to Chl content dynamics. One candidate to consider for advancing more specialized modeling structures is the Gaussian process, which can be composed of temporal and spectral covariance kernels, as already demonstrated by other applications in modeling solar irradiation predictions (Camps-Valls et al., 2016).

Leaf Chl Retrieval Using Spectral Bands Versus Vegetation Indices

Feature transformation is a critical task in any machine learning framework, especially when involving datasets comprising of hundreds of variables. This aspect should be carefully reviewed by evaluating the types of variables that are part of the dataset and exploring possible transformations and reductions to optimize the model performance (Guyon and Elisseeff, 2003). In this study, three elements can be highlighted regarding evaluating VIs as transformed variables out of reflectance spectral bands. First, using VIs slightly improves the goodness-of-fit and prediction accuracy of different types of ML models. Second, the VI predictors approach provide an alternative way to use spectral variables without affecting the capturing of temporal dynamics. Third, hyperspectral VI predictors add specific biophysical background to the training knowledge, hence enriching model interpretability.

Few studies in the literature have explored retrieving SPAD-based Chl levels using VI predictors derived from hyperspectral datasets. For example, linear regressions (Qi et al., 2020), random forest (Shah et al., 2019), or Cubist (Houborg and McCabe, 2018) have been combined with different types of VIs, reaching more accurate results than using spectral bands, which has also been achieved in this study. Although higher R^2 scores were reached when using all bands as predictors than the 145 selected bands (i.e., all except for the PLSR model). In terms of accuracy, higher accuracies (i.e., up to 3 SPAD units below) were reported when using the VI features. Indeed, these results follow what has been suggested in other studies regarding the low (or no) impact of the number of variables in the accuracy estimates, but rather the importance of identifying the marginal effect of the explanatory variables in the dependent variable (Alin, 2010).

Dimensional reduction can be made *via* pruning spectral bands or transforming them into new variables related to plant biochemical traits (as is done herein). (Feilhauer et al., 2015) pruned bands based on regression coefficients (R^2) and metrics that measure band importance, managing to reduce predictors to dozens of Chl absorption channels within the range of 500–750 nm. The same spectral region was fully covered in our study by 15 derivative and continuum-removal based VIs that also inform on Chl content. If we apply this pruning approach to our dataset, it would require approximately 110 bands (**Figure 2**) to train the models, which is still a large number compared to the available observations (i.e., three samples per plant and 36 plants, for a total of about 108 samples per campaign), leading to the question: what is the impact of not having much larger samples than predictors? A clear example of the impact is evident in the performance of the multivariate linear regression approach presented in this study ("Multiple Model Regression Assessment" and "Multiple Model SPAD Predictions"). When observations do not sufficiently exceed the number of predictors, the least square cost function may overfit the training set, consequently producing poor retrievals. While the other algorithms can cope with this dimensionality issue, a lower-dimensional dataset is desired to improve the computational efficiency of the workflow, and thus feature transformation is a suitable alternative to follow.

Rivera-Caicedo et al. (2017) have investigated dimensionality reduction approaches such as principal components transformation and partial least squares (PLSR), among others. Although some of their trials led to better-fitted models than using all bands, only slight improvements were achieved in terms of accuracy. Similar results were reached in our study when comparing the different ML methods against PLSR (**Supplementary Table 2**; **Figure 6**). PLSR reached the best fitting and accuracy scores under the sequential strategy by using band predictors; however, it was exceeded by the random forest and GPR models when using the VIs features under the retraining strategy. Multicollinearity causes this performance by increasing the vulnerability of the predictor weights to vary whenever there is a small change in data, resulting in unstable model performances. Based on the ML algorithm designs (i.e., their mathematical formulations), some are inherently able to handle multicollinearity better than others. For instance, random forest deals well with large dimensional problems due to its pruning strategy, which uses bootstrapping and feature sampling to pick different sets of data and features, and estimate relative

importance while training each tree (**Supplementary Material**; **Figure 10B**). Alternatively, GPR kernels are coupled with a length scale parameter that measures how strong is each predictor variable in a model (**Supplementary Material**; **Figure 10C**). It was possible to perform a comparative analysis through such feature engineering methods to account for coincident relevant variables identified under each approach (**Figure 10A**). However, although few VIs and spectral ranges were classified as relevant by all three methods, results could still be affected by multicollinearity, especially when features reach similar scores, leading to difficulties in ranking their importance. Developing reduction and transformation dimensionality approaches for hyperspectral data remains a challenge (Thenkabail and Lyon, 2012). As such, new ways to construct alternative prediction variables require continued investigation.

VIs can be considered as a transformed version of the spectral bands that involve known relationships between spectral response and biophysical traits, and hence, are suited to track temporal dynamics along the phenological growth stages. At first glance, such patterns are traceable in our results. For instance, evident Chl dynamics were retrieved by retraining random forest and GPR squared exponential coupled either with VIs or spectral bands under the same strategy (**Figure 7B**). Such agreement among methods reveals that VIs can be used alternatively as predictor variables. Even more, an exhaustive comparative analysis on the importance of the predictors ("Important Features") allowed us to trace the spectral information in the most relevant bands and indices (**Supplementary Tables 3**, **4**), revealing that although most of the assessed indices contain the same spectral information covered by the relevant bands, VIs transform the spectral data into new explanatory variables. As such, the VIs approach should be considered as a feature transformation strategy more than a dimensionality reduction method, with the advantages of providing interpretable results and being straightforward to implement in production.

Which ML Model to Use for Multi-Temporal Retrieving Leaf Chl?

Considering the wide gamut of non-parametric ML methods, the present study sought to examine the most commonly used types of supervised algorithms (Verrelst et al., 2019). Beyond indicating which, if any, particular method could be identified as being the best in estimating Chl metrics from high spatial, spectral and temporal data, some findings can be highlighted based on this study's data characteristics, the learning strategies, and the subsequent results.

One of the explored methods was support vector regression (SVR) using three different types of kernels and scales (**Supplementary Material**). Our results showed that SVR linear was the best performing kernel when using the 272 reflectance bands as predictors, although retrieving poorer results when the number of predictors was reduced to 60 VIs. Such behavior suggests that SVR algorithms require a preliminary kernel and feature engineering to fit the regression relationship. In a previous study, (Malenovský et al., 2017) found that the Gaussian SVR outperformed random forest in retrieving total chlorophyll (C_{ab}) for Antarctic moss by using uni-temporal UAV-based hyperspectral data at a sub-decimeter resolution, training with continuum-removed bands as predictors, and advising to optimize feature selection if intending to use fewer predictors.

Another approach explored in this study was the ensemble of trees. Two general algorithms, boosted and bagged, were tested by training multiple individual medium trees (i.e., ntrees = 60). Specifically, the random forest was explored from the bagged approach, which in general outperformed the boosted ensemble, reaching higher accuracies and better-fitted regressions (**Figure 6**). However, some of these differences were minor: for instance, when using VI predictors under sequential learning, and for the first and fourth stages under retraining learning, which suggest both methods are suitable for modeling and retrieving multi-temporal Chl content dynamics. The choice of random forest over boosted trees was based on ease of use, since it relies on less tuning parameters than boosted ensembles, and is less prone to overfit when training highly variable or noisy data (Breiman, 2001). It is advised to perform a comparative analysis by training both algorithms, carefully tuning the shrinkage or learning rate parameter in boosting trees, which is decisive in its performance. Similar results have been reported in studies that followed a sequential learning strategy (Shah et al., 2019), finding slightly improved accuracies (i.e., from 5.5 to 3.5 $\mu g/cm^2$ in the RMSE) and better-fitted models (i.e., up to 0.89–0.95 units in the R^2) when training random forest with VIs than when using spectral bands. However, random forest retrievals and validation distributions did best with the retraining learning routine.

Finally, we also examined one of the most promising approaches in hyperspectral remote sensing data analysis: a Bayesian kernel-based method referred to as Gaussian process regression (GPR; Camps-Valls et al., 2016). Four different kernels or covariance functions were compared: exponential, squared exponential (SE), Matern, and rational quadratic, with all of them integrated with a maximum likelihood technique for auto-tuning their parameters. Any of these covariance functions captures the similarity between pairs of observations under the assumption that if the input predictors are close to each other, it is expected that their SPAD values will also be close. Although marginal differences were found among the tested formulations in terms of accuracy with MAE and RMSE (<2 SPAD units), the SE kernel stood out from the rest (**Figure 6**). Moreover, the assessment metrics for the SE kernel outperformed most of the models examined in this study: only barely surpassed by random forest in some trials using VI predictors and retraining learning. However, SE estimates were associated with high uncertainties when using VIs under sequential learning (~10–15 SPAD units), and for the last two stages when using band features under the retraining routine (~10–12 SPAD units). Accordingly, distributions of the same trials showed discrepancies between the estimates and actual SPAD values from the validation samples, indicating that the GPR model should be subject to optimization routines (Verrelst et al., 2016; Rivera-Caicedo et al., 2017), despite its flexibility and robustness to deal with multi-temporal hyperspectral data.

Few studies in the vegetation spectroscopy literature have extensively compared GPR with other non-parametric regressions (Ashourloo et al., 2016). In two notable examples, (Verrelst et al., 2012); (Rivera-Caicedo et al., 2014) found SE kernel performance exceeded decision trees, neural networks, support vector regression, and kernel ridge methods. To date, most of the ML implementations in hyperspectral applications have tended toward random forest implementations to retrieve biophysical variables (Shah et al., 2019; Vanbrabant et al., 2019). Based on this previous research and the analysis above, the GPR squared exponential, PLSR and random forest were moved to production in the last stage of our workflow, in order to plot the multi-temporal SPAD maps at a pixel level, and including the uncertainty maps produced by GPR, which provides an additional level of information regarding the quality of the estimates. Further analysis based on deep learning and neural networks is recommended to compare their performance under the same dataset.

Practical Considerations on the Learning Workflow for Chl Retrieval

In general, any non-linear ML approach can be implemented to model and retrieve multi-temporal physiological traits such as Chl content or SPAD levels. However, different strategies and techniques should be explored in order to ensure the efficiency, accuracy, and transferability of model selections. The first general task to pursue in a ML framework should include an exploratory comparative assessment of different models within the training workflow, using the full training/testing data, but predicting over a validation subset. Numerous learning libraries and toolboxes (Rivera-Caicedo et al., 2014) are available for both open-source and commercial applications, including multiple regression algorithms that can be easily implemented to run a preliminary comparative analysis as soon as the first collection of data is available. Selecting the methods to further explore should rely on the data characteristics, such as spatial, spectral, and temporal resolution, and the number of ground-truth samples available.

For chlorophyll monitoring applications specifically, non-destructive *in-situ* sampling can be conducted using chlorophyll meters that provide a relative indicator of leaf Chl content (i.e., SPAD), which can be considered the dependent variable to estimate. However, if the formulations to translate relative units to physical units (i.e., $\mu mol/m^2$) are available, it is advised to translate the data and use Chl content as the dependent variable to estimate (Parry et al., 2014). It is also preferable to use field spectra data to train and fit the models, assuring its comparability with the UAV hyperspectral imagery, which must be radiometrically calibrated and processed in advance (Angel et al., 2020). For structurally vertical, complex, and mixed-species, e.g., orchards or cereals, it is also essential to include BDRF corrections (Aasen and Bolten, 2018). If field spectra data are not available, it may potentially be replaced by synthetic spectral datasets generated through inverting radiative transfer models, although field data is necessary for validation (Feilhauer et al., 2015).

A final consideration relies on quality assessment tasks. Ensuring sufficient observations to split between training, testing, and validation will improve the learning routines and the assessment and model selection stages. In particular, validation is a decisive phase in the ML workflow, and it can be performed by cross-validation when few data are available. When multi-temporal data is involved, poor results can be overlooked if evaluating the validation dataset across time is skipped. Since phenotypic data is dynamic, modeling should be treated as a continuous process by periodically retraining and validating the models, particularly if the new incoming data distribution varies significantly from the first dataset. Such a phenomenon is known as model drift in machine learning literature (Webb et al., 2016) and has to be continuously assessed through different metrics. Complementary to analyzing quality metrics (i.e., RMSE, MAE, R^2), some algorithms like GPR can provide the uncertainty associated with each prediction, allowing uncertainties to be mapped together with model estimates. Accounting for uncertainties associated with SPAD retrieval is of particular interest, since small variations in SPAD units will lead to exponential variations in the actual Chl content (Parry et al., 2014). Multiple evaluation techniques can be further explored to assess different metrics and make improvements accordingly until accurate and coherent results are obtained.

CONCLUSIONS

An innovative machine learning retrieval framework for mapping leaf chlorophyll content across a crop cycle was developed using ultra-high-resolution UAV-based hyperspectral imagery, *in-situ* SPAD observations, and field-based leaf spectra. The workflow evaluates a range of model scenarios to determine the best-performing methods based on the production of accurate and coherent multi-temporal retrievals. The intercomparison of six different ML approaches, including some variations and kernel formulations, accounted for a total of 17 different models. *In-situ* observations were split into a training/testing subset used to fit the models through cross-validation and estimate R^2, and a validation subset was employed to assess the accuracy of the models *via* RMSE and MAE. Three main aspects can be highlighted as innovations from the proposed framework, which match the particular research objectives explored in this study:

- Strategies for selecting the best-suited training/retrieval combinations based on accuracy assessment.
- Evaluation of sequential versus retraining learning strategies.
- Comparison of VIs and spectral band predictors in explaining the SPAD variable.

It was determined that a retraining learning strategy, whereby a model is updated as new data becomes available, proved superior in capturing the temporal dynamics of SPAD-based Chl. In contrast, if models are trained using the full data series at any instance in time (i.e., a sequential learning strategy), only a few model combinations could yield results close to the validation data, vanishing Chl-level changes over time. It was determined that the best combination of training conditions was achieved by coupling sequential learning with spectral bands, and retraining learning with VI predictors. However, given the link that VIs establish between plant traits and

spectral responses along the phenology stages, VI predictors may be preferred over spectral bands in order to add interpretability to the models without deteriorating their performance or accuracy. In this direction, PLSR, GPR and random forest were selected as the most promising approaches to optimize and estimate the SPAD predictions. Overall, PLSR and GPR squared exponential outperformed the other models in terms of accuracy and goodness-of-fit when operated under the sequential and retraining strategies, respectively. However, random forest estimates were closer than GPR to the actual validation data distribution, which was used as a reference to evaluate the multi-temporal coherence of the results. An additional assessment element is provided by uncertainty metrics that are included as part of the GPR results. Filtering of the most relevant predictors (bands and VIs) resulted from the inherent feature importance mechanisms of the PLSR, random forest, and GPR approaches. By scoring and classifying the predictors, the selected models reached some agreement on strong individual bands and VIs that highlighted a few decisive spectral ranges and indices useful for retrieving Chl levels.

While a comprehensive assessment of factors contributing to model accuracy and performance was evaluated herein, there remain several further opportunities to advance upon the evaluated approaches, considering the wide range of learning strategies, optimization, and assessment techniques available in open source and commercial applications. Of particular note, there is a need to develop approaches capable of capturing non-evident relationships within large, high-spectral, −temporal, and -spatial datasets to cover canopy scales, and solving prediction problems even under limited *in-situ* training data. Hybrid machine learning with radiative transfer models is a further option to explore the particular leaf and canopy optical properties of agricultural species, integrating other information from UAV systems (e.g., lidar-based canopy height), allowing to account for biochemical and structural traits simultaneously. As data-collection technology evolves, producing ever-larger volumes of data, identifying how best to retrieve accurate informatics quickly and efficiently is an area of critical and much needed research interest. If such techniques and approaches are not developed, we risk being overwhelmed by information, thereby losing the capacity for process insight and knowledge advancement.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

YA: conceptualization, methodology, software, formal analysis, investigation, visualization, and writing original draft. MFM: conceptualization, methodology, validation, writing review and editing, and supervision. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the King Abdullah University of Science and Technology (KAUST).

ACKNOWLEDGMENTS

We acknowledge Dr. Mitchell Morton, Dr. Sonia Negrao, Gabriele Fiene, and Dinara Utarbayeva from the Salt Lab led by Prof. Mark Tester, with whom field experiments were undertaken. We thank Dr. Kasper Johansen, Dr. Yoann Malbeteau, Samer Al-Mashharawi, Maria Alejandra Perea, Ting Li, Bruno Aragon, Matteo Ziliani, and Cristhian Andrade from the Hydrology, Agriculture and Land Observation (HALO) group, who were also participants of the field data collection. We thank Prof. Magdi A. Mousa and all the workers at the King Abdulaziz University Agricultural Research Station in Hada Al-Sham for their extensive assistance with field maintenance and harvesting.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.722442/full#supplementary-material

REFERENCES

Aasen, H., and Bolten, A. (2018). Multi-temporal high-resolution imaging spectroscopy with hyperspectral 2D imagers – From theory to application. *Remote Sens. Environ.* 205, 374–389. doi: 10.1016/j.rse.2017.10.043

Aasen, H., Honkavaara, E., Lucieer, A., and Zarco-Tejada, A. P. J. (2018). Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows. *Remote Sens.* 10:1091. doi: 10.3390/rs10071091

Alin, A. (2010). Multicollinearity. WIREs comp. *Stat* 2, 370–374. doi: 10.1002/wics.84

Angel, Y., Turner, D., Parkes, S., Malbeteau, Y., Lucieer, A., and McCabe, M. F. (2020). Automated Georectification and mosaicking of UAV-based Hyperspectral imagery from push-broom sensors. *Remote Sens.* 12:34. doi: 10.3390/rs12010034

Ashourloo, D., Aghighi, H., Matkan, A. A., Mobasheri, M. R., and Rad, A. M. (2016). An investigation into machine learning regression techniques for the leaf rust disease detection using Hyperspectral measurement. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 9, 4344–4351. doi: 10.1109/JSTARS.2016.2575360

Barreto, M. A. P., Johansen, K., Angel, Y., and McCabe, M. F. (2019). Radiometric assessment of a UAV-based push-broom Hyperspectral camera. *Sensors* 19:4699. doi: 10.3390/s19214699

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Camps-Valls, G., Verrelst, J., Munoz-Mari, J., Laparra, V., Mateo-Jimenez, F., and Gomez-Dans, J. (2016). A survey on Gaussian processes for earth-observation data analysis: a comprehensive investigation. *IEEE Geosci.Remote Sens. Mag.* 4, 58–78. doi: 10.1109/MGRS.2015.2510084

Capolupo, A., Kooistra, L., Berendonk, C., Boccia, L., and Suomalainen, J. (2015). Estimating plant traits of grasslands from UAV-acquired hyperspectral images: a comparison of statistical approaches. *ISPRS Int. J. Geo-Inf.* 4, 2792–2820. doi: 10.3390/ijgi4042792

Cerovic, Z. G., Masdoumier, G., Ghozlen, N. B., and Latouche, G. (2012). A new optical leaf-clip meter for simultaneous non-destructive assessment of leaf chlorophyll and epidermal flavonoids. *Physiol. Plant.* 146, 251–260. doi: 10.1111/j.1399-3054.2012.01639.x

Clevers, J. G. P. W., and Gitelson, A. A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. *Int. J. Appl. Earth Obs. Geoinf.* 23, 344–351. doi: 10.1016/j.jag.2012.10.008

Croft, H., Chen, J. M., Zhang, Y., Simic, A., Noland, T. L., Nesbitt, N., et al. (2015). Evaluating leaf chlorophyll content prediction from multispectral remote sensing data within a physically-based modeling framework. *ISPRS J. Photogramm. Remote Sens.* 102, 85–95. doi: 10.1016/j.isprsjprs.2015.01.008

Curran, P. J. (1989). Remote sensing of foliar chemistry. *Remote Sens. Environ.* 30, 271–278. doi: 10.1016/0034-4257(89)90069-2

Delloye, C., Weiss, M., and Defourny, P. (2018). Retrieval of the canopy chlorophyll content from Sentinel-2 spectral bands to estimate nitrogen uptake in intensive winter wheat cropping systems. *Remote Sens. Environ.* 216, 245–261. doi: 10.1016/j.rse.2018.06.037

Dietterich, T. G. (2002). "Machine learning for sequential data: a review," in *Structural, Syntactic, and Statistical Pattern Recognition. SSPR/SPR 2002. Lecture Notes in Computer Science. Vol. 2396*. eds. T. Caelli, A. Amin, R. P. W. Duin, D. de Ridder, M. Kamel (Berlin, Heidelberg: Springer), 15–30.

DJI (2020) Matrice 600, Available at: https://www.dji.com/matrice600 (Accessed June, 2021).

Domingues Franceschini, M. H., Bartholomeus, H., Van Apeldoorn, D., Suomalainen, J., Kooistra, L., Correction, L., et al. (2017). Intercomparison of unmanned aerial vehicle and ground-based narrow band spectrometers applied to crop trait monitoring in organic potato production. *Sensors* 17:2265. doi: 10.3390/s17102265

Dong, T., Shang, J., Chen, J. M., Liu, J., Qian, B., Ma, B., et al. (2019). Assessment of portable chlorophyll meters for measuring crop leaf chlorophyll concentration. *Remote Sens.* 11:2706. doi: 10.3390/rs11222706

Feilhauer, H., Asner, G. P., and Martin, R. E. (2015). Multi-method ensemble selection of spectral bands related to leaf biochemistry. *Remote Sens. Environ.* 164, 57–65. doi: 10.1016/j.rse.2015.03.033

Gabriel, J. L., Zarco-Tejada, P. J., López-Herrera, P. J., Pérez-Martín, E., Alonso-Ayuso, M., and Quemada, M. (2017). Airborne and ground level sensors for monitoring nitrogen status in a maize crop. *Biosyst. Eng.* 160, 124–133. doi: 10.1016/j.biosystemseng.2017.06.003

Green, A. A., Berman, M., Switzer, P., and Craig, M. D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sens.* 26, 65–74. doi: 10.1109/36.3001

Guyon, I., and Elisseeff, A. (2003). An introduction to variable and feature selection. *J. Mach. Learn. Res.* 3, 1157–1182. doi: 10.5555/944919.944968

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., and Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* 81, 416–426. doi: 10.1016/S0034-4257(02)00018-4

Headwall Photonics (2020a). Hyperspectral imaging sensors. Available at: https://www.headwallphotonics.com/hyperspectral-sensors (Accessed June, 2021).

Headwall Photonics (2020b). SpectralView. Available at: http://www.headwallphotonics.com/software (Accessed June, 2021).

Henrich, V., Jung, A., Götze, C., Sandow, C., Thürkow, D., and Gläßer, C. (2009). "Development of an online indices database: motivation, concept and implementation," in *Proceedings of the 6th EARSeL Imaging Spectroscopy SIG Workshop Innovative Tool for Scientific and Commercial Environment Applications*, March 16, 2009; Tel Aviv, Israel

Houborg, R., and McCabe, M. F. (2018). A hybrid training approach for leaf area index estimation via cubist and random forests machine-learning. *ISPRS J. Photogramm. Remote Sens.* 135, 173–188. doi: 10.1016/j.isprsjprs.2017.10.004

Houborg, R., McCabe, M.F., Angel, Y., and Middleton, E.M. (2016). "Detection of chlorophyll and leaf area index dynamics from sub-weekly hyperspectral imagery," in *Proceedings SPIE 9998. Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII 999812*, October 25, 2016.

Houborg, R., McCabe, M.F., and Cescatti, A., A., A.A. Gitelson (2015b). Leaf chlorophyll constraint on model simulated gross primary productivity in agricultural systems. *Int. J. Appl. Earth Obs. Geoinf.*, 43, 160–176, doi:10.1016/j.jag.2015.03.016

Houborg, R., McCabe, M. F., Cescatti, A., Gao, F., Schull, M., and Gitelson, A. (2015a). Joint leaf chlorophyll content and leaf area index retrieval from Landsat data using a regularized model inversion system (REGFLEC). *Remote Sens. Environ.* 159, 203–221. doi: 10.1016/j.rse.2014.12.008

Kanning, M., Kühling, I., Trautz, D., and Jarmer, T. (2018). High-resolution UAV-based Hyperspectral imagery for LAI and chlorophyll estimations from wheat for yield prediction. *Remote Sens.* 10:2000. doi: 10.3390/rs10122000

Konica Minolta (2009). Chlorophyll meter SPAD-502 Plus. Available at: https://www.konicaminolta.com/instruments/download/catalog/color/pdf/spad502plus_catalog_eng.pdf (Accessed June, 2021).

Lambers, H., and Oliveira, R.S. (2019). *Plant Energy Budgets: The Plant's Energy Balance, Plant Physiological Ecology. 2nd Edn*. Switzerland: Springer Nature, 265–269.

Liang, L., Qin, Z., Zhao, S., Di, L., Zhang, C., Deng, M., et al. (2016). Estimating crop chlorophyll content with hyperspectral vegetation indices and the hybrid inversion method. *Int. J. Remote Sens.* 37, 2923–2949. doi: 10.1080/01431161.2016.1186850

López-López, M., Calderón, R., González-Dugo, V., Zarco-Tejada, P. J., and Fereres, E. (2016). Early detection and quantification of almond red leaf blotch using high-resolution Hyperspectral and thermal imagery. *Remote Sens.* 8:276. doi: 10.3390/rs8040276

Malenovský, Z., Lucieer, A., King, D. H., Turnbull, J. D., and Robinson, S. A. (2017). Unmanned aircraft system advances health mapping of fragile polar vegetation. *Methods Ecol. Evol.* 8, 1842–1857. doi: 10.1111/2041-210X.12833

Markwell, J., Osterman, J. C., and Mitchell, J. L. (1995). Calibration of the Minolta SPAD-502 leaf chlorophyll meter. *Photosynth. Res.* 46, 467–472. doi: 10.1007/BF00032301

Martínez, D. E., and Guiamet, J. J. (2004). Distortion of the SPAD 502 chlorophyll meter readings by changes in irradiance and leaf water status. *Agronomie* 24, 41–46. doi: 10.1051/agro:2003060

Mladenić, D. (2006). "Feature selection for dimensionality reduction," in *Subspace, Latent Structure and Feature Selection. SLSFS 2005. Lecture Notes in Computer Science, Vol. 3940*. eds. C. Saunders, M. Grobelnik, S. Gunn and J. Shawe-Taylor (Berlin, Heidelberg: Springer), 84–102.

Netto, A. T., Campostrini, E., Oliveira, J. G. D., and Bressan-Smith, R. E. (2005). Photosynthetic pigments, nitrogen, chlorophyll a fluorescence and SPAD-502 readings in coffee leaves. *Sci. Hortic.* 104, 199–209. doi: 10.1016/j.scienta.2004.08.013

Noodén, L. D., Guiamét, J. J., and John, I. (1997). Senescence mechanisms. *Physiol. Plantarum* 101, 746–753. doi: 10.1111/j.1399-3054.1997.tb01059.x

Parry, C., Blonquist, J. M. Jr., and Bugbee, B. (2014). In situ measurement of leaf chlorophyll concentration: analysis of the optical/absolute relationship. *Plant Cell Environ.* 37, 2508–2520. doi: 10.1111/pce.12324

Porra, R. J. (2002). The chequered history of the development and use of simultaneous equations for the accurate determination of chlorophylls a and b. *Photosynth. Res.* 73, 149–156. doi: 10.1023/A:1020470224740

Qi, H., Zhu, B., Kong, L., Yang, W., Zou, J., Lan, Y., et al. (2020). Hyperspectral inversion model of chlorophyll content in Peanut leaves. *Appl. Sci.* 10:2259. doi: 10.3390/app10072259

Rasmussen, C. E., and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*, The MIT Press, New York.

Rivera-Caicedo, J. P., Verrelst, J., Muñoz-Marí, J., Camps-Valls, G., and Moreno, J. (2017). Hyperspectral dimensionality reduction for biophysical variable statistical retrieval. *ISPRS Int. J. Geo-Inf.* 132, 88–101. doi: 10.1016/j.isprsjprs.2017.08.012

Rivera-Caicedo, J. P., Verrelst, J., Muñoz-Marí, J., Moreno, J., and Camps-Valls, G. (2014). Toward a semiautomatic machine learning retrieval of biophysical parameters. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 7, 1249–1259. doi: 10.1109/JSTARS.2014.2298752

Ruffin, C., and King, R. L. (1999). The analysis of hyperspectral data using Savitzky-Golay filtering-theoretical basis. 1, in *Proceedings 1999 IEEE International Geoscience and Remote Sensing Symposium*, June–July, 1999. IGARSS'99 2, 756–758.

Savitzky, A., and Golay, M. J. E. (1964). Soothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36, 1627–1639. doi: 10.1021/ac60214a047

Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., et al. (2013). Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *Int. J. Appl. Earth Obs. Geoinf.* 25, 47–54. doi: 10.1016/j.jag.2013.04.003

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., et al. (2014). Machine Learning: The High Interest Credit Card of Technical Debt, SE4ML: Software Engineering for Machine Learning (NIPS 2014 Workshop). Available at: https://research.google/pubs/pub43146/

Shah, S. H., Angel, Y., Houborg, R., Ali, S., and McCabe, M. F. (2019). A random Forest machine learning approach for the retrieval of leaf chlorophyll content in wheat. *Remote Sens.* 11:920. doi: 10.3390/rs11080920

Shah, S. H., Houborg, R., and McCabe, M. F. (2017). Response of chlorophyll, carotenoid and SPAD-502 measurement to salinity and nutrient stress in wheat (Triticum aestivum L.). *Agronomy* 7:61. doi: 10.3390/agronomy7030061

Thenkabail, P., and Lyon, J. (eds.) (2012). *Hyperspectral Remote Sensing of Vegetation*. Boca Raton: CRC Press, 83–84.

Thenkabail, P., Lyon, J., and Huete, A. (eds.) (2019). *Hyperspectral Indices and Image Classifications for Agriculture and Vegetation*. Boca Raton:CRC Press, 3–23.

Uddling, J., Gelang-Alfredsson, J., Piikki, K., and Pleijel, H. (2007). Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynth. Res.* 91, 37–46. doi: 10.1007/s11120-006-9077-5

Vanbrabant, Y., Tits, L., Delalieux, S., Pauly, K., Verjans, W., and Somers, B. (2019). Multitemporal chlorophyll mapping in pome fruit orchards from remotely piloted aircraft systems. *Remote Sens.* 11:1468. doi: 10.3390/rs11121468

Verrelst, J., Malenovský, Z., Van der Tol, C., Camps-Valls, G., Gastellu-Etchegorry, J. P., Lewis, P., et al. (2019). Quantifying vegetation biophysical variables from imaging spectroscopy data: a review on retrieval methods. *Surv. Geophys.* 40, 589–629. doi: 10.1007/s10712-018-9478-y

Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera-Caicedo, J. P., Camps-Valls, G., et al. (2012). Machine learning regression algorithms for biophysical parameter retrieval: opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* 118, 127–139. doi: 10.1016/j.rse.2011.11.002

Verrelst, J., Rivera-Caicedo, J. P., Gitelson, A., Delegido, J., Moreno, J., Camps-Valls, G., et al. (2016). Spectral band selection for vegetation properties retrieval using Gaussian processes regression. *Int. J. Appl. Earth Obs. Geoinf.* 52, 554–567. doi: 10.1016/j.jag.2016.07.016

Wang, R., Gamon, J. A., Emmerton, C. A., Springer, K. R., Yu, R., and Hmimina, G. (2020). Detecting intra- and inter-annual variability in gross primary productivity of a North American grassland using MODIS MAIAC data. *Agric. For. Meteorol.* 281:107859. doi: 10.1016/j.agrformet.2019.107859

Wang, C., and Myint, S. W. (2015). A simplified empirical line method of radiometric calibration for small unmanned aircraft systems-based remote sensing. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 8, 1876–1885. doi: 10.1109/JSTARS.2015.2422716

Webb, G. I., Hyde, R., Cao, H., Nguyen, H. L., and Petitjean, F. (2016). Characterizing concept drift. *Data Min. Knowl. Disc.* 30, 964–994. doi: 10.1007/s10618-015-0448-4

Wellburn, A. R. (1994). The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *J. Plant Physiol.* 144, 307–313. doi: 10.1016/S0176-1617(11)81192-2

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

Yuan, Z., Cao, Q., Zhang, K., Ata-Ul-Karim, S. T., Tian, Y., Zhu, Y., et al. (2016). Optimal leaf positions for spad meter measurement in rice. *Front. Plant Sci.* 7:719. doi: 10.3389/fpls.2016.00719

Zarco-Tejada, P. J., González-Dugo, V., Williams, L. E., Suárez, L., Berni, J. A. J., Goldhamer, D., et al. (2013b). A PRI-based water stress index combining structural and chlorophyll effects: assessment using diurnal narrow-band airborne imagery and the CWSI thermal index. *Remote Sens. Environ.* 138, 38–50. doi: 10.1016/j.rse.2013.07.024

Zarco-Tejada, P. J., Guillén-Climent, M. L., Hernández-Clemente, R., Catalina, A., González, M. R., and Martín, P. (2013a). Estimating leaf carotenoid content in vineyards using high resolution hyperspectral imagery acquired from an unmanned aerial vehicle (UAV). *Agric. For. Meteorol.* 171-172, 281–294. doi: 10.1016/j.agrformet.2012.12.013

Zarco-Tejada, P. J., Hornero, A., Beck, P. S. A., Kattenborn, T., Kempeneers, P., and 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 Sens. Environ.* 223, 320–335. doi: 10.1016/j.rse.2019.01.031

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 23 March 2022
doi: 10.3389/fpls.2022.716506

Detecting Intra-Field Variation in Rice Yield With Unmanned Aerial Vehicle Imagery and Deep Learning

Emily S. Bellis[1,2,3†], Ahmed A. Hashem[2,3,4†], Jason L. Causey[1,2], Benjamin R. K. Runkle[5], Beatriz Moreno-García[5], Brayden W. Burns[4], V. Steven Green[3,4], Timothy N. Burcham[3], Michele L. Reba[6] and Xiuzhen Huang[1,2]*

[1] Department of Computer Science, Arkansas State University, Jonesboro, AR, United States, [2] Center for No-Boundary Thinking, Arkansas State University, Jonesboro, AR, United States, [3] University of Arkansas System Division of Agriculture, Little Rock, AR, United States, [4] College of Agriculture, Arkansas State University, Jonesboro, AR, United States, [5] U. S. Department of Agriculture, Department of Biological and Agricultural Engineering, University of Arkansas, Fayetteville, AR, United States, [6] USDA Agricultural Research Service Delta Water Management Research Unit, Jonesboro, AR, United States

OPEN ACCESS

Edited by:
Alessandro Matese, Institute for Bioeconomy (CNR), Italy

Reviewed by:
Tiebiao Zhao, University of California, Merced, United States
Shangpeng Sun, McGill University, Canada

***Correspondence:**
Emily S. Bellis
ebellis@astate.edu

†These authors share first authorship

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

Received: *28 May 2021*
Accepted: *18 February 2022*
Published: *23 March 2022*

Citation:
Bellis ES, Hashem AA, Causey JL, Runkle BRK, Moreno-García B, Burns BW, Green VS, Burcham TN, Reba ML and Huang X (2022) Detecting Intra-Field Variation in Rice Yield With Unmanned Aerial Vehicle Imagery and Deep Learning. Front. Plant Sci. 13:716506. doi: 10.3389/fpls.2022.716506

Unmanned aerial vehicles (UAVs) equipped with multispectral sensors offer high spatial and temporal resolution imagery for monitoring crop stress at early stages of development. Analysis of UAV-derived data with advanced machine learning models could improve real-time management in agricultural systems, but guidance for this integration is currently limited. Here we compare two deep learning-based strategies for early warning detection of crop stress, using multitemporal imagery throughout the growing season to predict field-scale yield in irrigated rice in eastern Arkansas. Both deep learning strategies showed improvements upon traditional statistical learning approaches including linear regression and gradient boosted decision trees. First, we explicitly accounted for variation across developmental stages using a 3D convolutional neural network (CNN) architecture that captures both spatial and temporal dimensions of UAV images from multiple time points throughout one growing season. 3D-CNNs achieved low prediction error on the test set, with a Root Mean Squared Error (RMSE) of 8.8% of the mean yield. For the second strategy, a 2D-CNN, we considered only spatial relationships among pixels for image features acquired during a single flyover. 2D-CNNs trained on images from a single day were most accurate when images were taken during booting stage or later, with RMSE ranging from 7.4 to 8.2% of the mean yield. A primary benefit of convolutional autoencoder-like models (based on analyses of prediction maps and feature importance) is the spatial denoising effect that corrects yield predictions for individual pixels based on the values of vegetation index and thermal features for nearby pixels. Our results highlight the promise of convolutional autoencoders for UAV-based yield prediction in rice.

Keywords: convolutional autoencoder, remote sensing, UAS—unmanned aerial system, grain crop, precision agriculture

INTRODUCTION

Rice (*Oryza sativa*) is one of the most important staple food crops globally (Khush, 2001). However, efficient production remains a major challenge, and there is a growing need to increase yield gains per unit land area while conserving natural resources to meet current and future demands (Grassini et al., 2013). For example, nitrogen fertilization is one of the most challenging aspects of rice production, with recommended rates and timing depending significantly on cultivar, soil type, and other factors (Hardke, 2018). To optimize production while minimizing inputs and environmental impacts, real-time monitoring could enable more efficient identification of crop stress, yield projection, and decision-making throughout the season.

Remotely sensed images acquired by Unmanned Aerial Vehicles (UAVs) provide a flexible means to monitor crop stress and other production factors throughout the growing season. UAVs equipped with thermal sensors are sensitive to longwave infrared radiation (7,000–12,000 nm) and since transpiration rates and evaporative cooling decrease under water-limited conditions, thermal sensors are particularly suitable for early detection of drought stress (Maes and Steppe, 2019; Burns et al., 2022). UAVs can also be equipped with multispectral sensors that capture multiple spectral regions in relatively broad bands. In addition to red, green, and blue (RGB) bands, multispectral sensors capture wavelengths in the near-infrared (NIR) range (730–900 nm). A healthy vegetative canopy typically has very high reflectance in the NIR spectrum. Thus, multispectral imagery is particularly adept at assessing nutrient status for yield prediction (Maes and Steppe, 2019). Recently, state-of-the-art deep learning approaches are proving to be highly useful for yield prediction using analysis of images acquired by UAVs (Nevavuori et al., 2019), outperforming other methods.

An important consideration for the design of deep learning models from UAV-derived data is how to account for temporal variations in the spectral signatures of a developing crop. Rice canopy structure changes rapidly during vegetative growth, with early-season images mostly comprised of bare soil during seed germination, emergence, and seedling development. Approximately thirty days after planting at about the five-leaf stage, the first rice tiller appears (Hardke, 2018). At this time, flooding is initiated in the delayed-flood system used in Arkansas. Increased tillering coincides with an increase in green biomass, when the normalized difference vegetation index (NDVI), calculated based on reflectance in NIR and red bands (**Table 1**), begins to increase rapidly (Wang et al., 2014). Panicle initiation marks the beginning of reproduction. The developing panicle eventually emerges from the stem and is fully visible at heading when flowering begins. After pollination, the panicle develops, and the rice kernels fill, changing in color from light green to yellow and, ultimately tan, as the grains ripen and leaves senesce. Thus, spectral signatures steadily change with the development and maturation of the rice crop.

One strategy to account for variation in spectral and thermal indices across development is to let the model learn important features (such as changes in NDVI associated with developmental stage) during training. For example, Nevavuori et al. (2019) used Convolutional Neural Networks (CNNs) on wheat and malting barley fields to predict crop yield from derived vegetation indices and raw RGB data acquired from UAVs (∼0.3 m resolution). These CNNs were trained on data combined from nine fields, split into "early" and "late" growing season datasets based on the image collection date. Mean absolute percentage error was lower for models trained on early season (8.8%) compared to late season data (11.4%). These results suggest that relatively high performance can be achieved for yield prediction at the intra-field scale, even without more fine-grained consideration of plant developmental stage.

An alternative approach explicitly accounts for temporal aspects of variation in plant development in the model architecture. Recurrent neural networks (RNNs) are well-suited for sequential data due to the use of hidden states to capture relevant information from prior states. RNNs have been particularly successful for classification of land cover data from satellite imagery, due to the ability to leverage temporal patterns across image time series (Minh et al., 2018; Sun et al., 2019). Temporal data structures can also be considered with CNNs, when convolutions occur across images in the temporal dimension as well as in the spatial dimensions, and are called 3D-CNNs or temporal CNNs. Temporal CNNs demonstrated slightly improved performance compared to RNNs for land cover classification when considering spectral and temporal dimensions of the data only (Pelletier et al., 2019) and also when temporal, spectral, and spatial dimensions were considered (Li et al., 2017; Ji et al., 2018). While their utility is well-demonstrated for the task of land cover classification from satellite imagery, it is unknown whether temporal network architectures could also demonstrate improved accuracies for tasks such as intra-field prediction of crop yield based on higher pixel count images (as compared to satellite images) from UAVs.

In this study, we assume that spatial variation in nutrient and water availability drives intra-field variation in spectral indices, and predict this variation will manifest as deviations from average conditions, observable from UAV imagery. We hypothesize that a model architecture that accounts for complex spatio-temporal patterns (e.g., 3D-CNN architecture) will be more informative for predicting intra-field yield variation compared to a spatial-only model (e.g., 2D-CNN architecture, **Figure 1**). We further determine whether deviations from average conditions matter most at certain time points, or if images taken during particular developmental stages are equally predictive of future yield. Finally, we characterize the nature of the benefit of the tested deep learning architectures for our dataset.

MATERIALS AND METHODS

Study Site

Our study focuses on a single study site located in the state of Arkansas, which contributes approximately half of

TABLE 1 | Vegetation indices.

Index	Abbreviation	Equation[a]	References
Normalized difference vegetation index	NDVI	$\frac{(\rho_{NIR}-\rho_r)}{(\rho_{NIR}+\rho_r)}$	Tucker, 1979; Hatfield and Prueger, 2010; Sharma et al., 2015
Chlorophyll index green	CI_{green}	$\frac{(\rho_{NIR})}{\rho_g} - 1$	Gitelson et al., 2005, 2003; Hatfield and Prueger, 2010
Red-edge normalized difference vegetation index	RENDVI	$\frac{(\rho_{NIR}-\rho_{RE})}{(\rho_{NIR}+\rho_{RE})}$	Gitelson and Merzlyak, 1997; Hatfield and Prueger, 2010
Green normalized difference vegetation index	GNDVI	$\frac{(\rho_{NIR}-\rho_g)}{(\rho_{NIR}+\rho_g)}$	Gitelson and Merzlyak, 1997; Hatfield and Prueger, 2010
Normalized area vegetation index	NAVI	$1 - \frac{(\rho_r)}{\rho_{NIR}}$	Carmona et al., 2015
Triangle greenness index	TGI	$-0.5[(670-480)*(\rho_r-\rho_g) - (670-550)(\rho_r-\rho_b)]$	Hunt et al., 2011, 2013

[a] *Reflectance (ρ) is measured at the wavelength denoted by the subscript: red (r), green (g), blue (b), red-edge (RE), and near-infrared (NIR).*

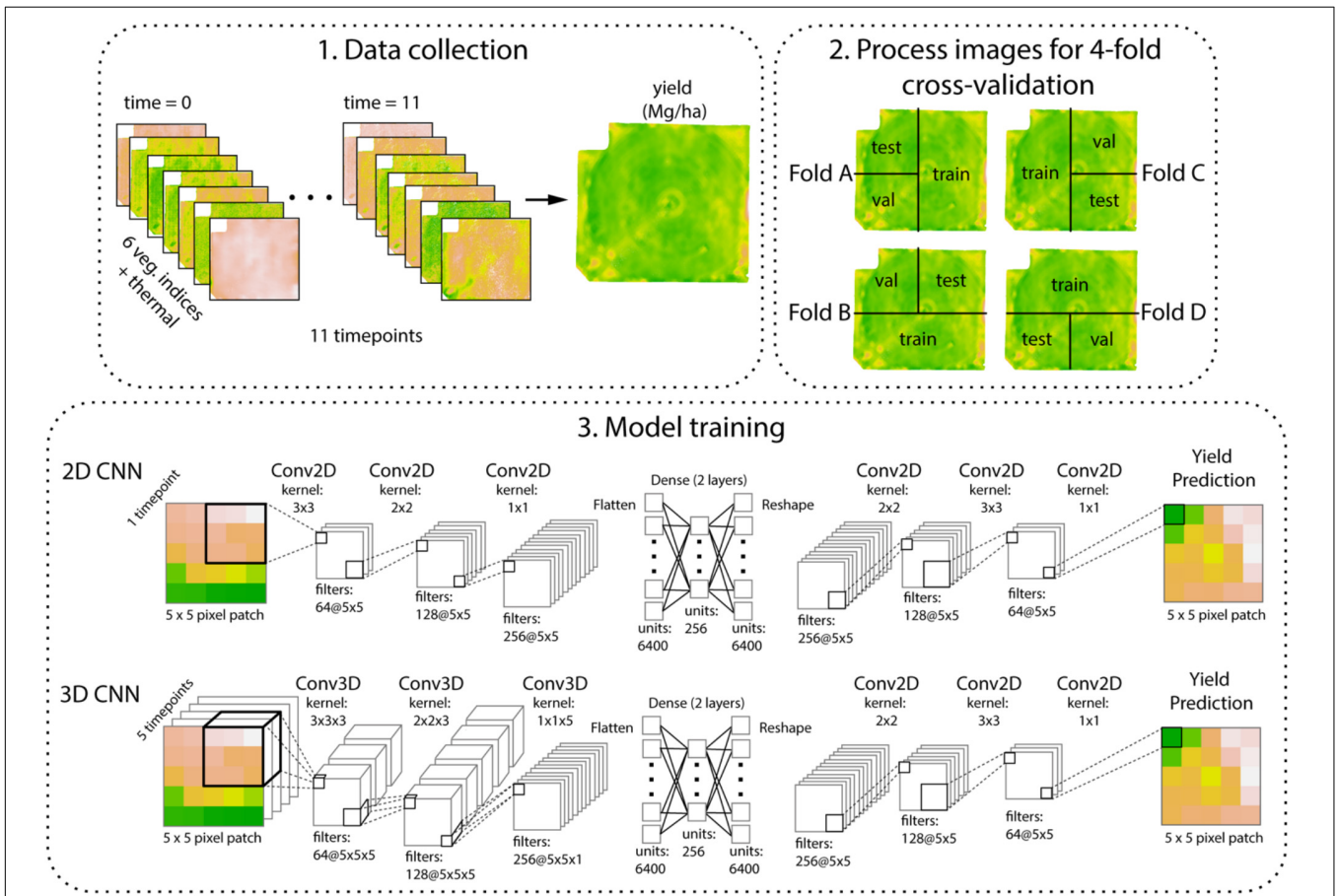

FIGURE 1 | Study design. UAV images were collected at 11 time points during the season. Predictions were based on thermal data and six vegetation indices derived from red, green, blue, red-edge, and near-infrared bands (CIgreen, GNDVI, NAVI, NDVI, RENDVI, and TGI). The field was divided into training, test, and validation images as shown for four-fold cross-validation. These larger regions were divided into 5 × 5 pixel patches (50 cm resolution pixels). Convolutional neural networks (CNNs) used an autoencoder-like structure to predict yield on an output 5 × 5 pixel patch based on 5 × 5 pixel input images, each with seven features. Model training for the 2D-CNN was based on a single time point; 3D-CNN used the five time points centered around the reproductive phase of crop growth.

the agricultural land area harvested for rice grown in the United States (United States Department of Agriculture Economic Research Service [USDA], 2021). The study site is a 16-ha, zero-grade (0% slope) field within a large farm operation in Lonoke County (e.g., Runkle et al., 2019). The farm produces rice using a rice-after-rice (i.e., continuous rice) production system and a drill-seeded, delayed flood program and burns rice straw after harvest. Field soil is classified as silt loam: 33% Calhoun silt loam (Fine-silty, mixed, active, thermic Typic Glossaqualfs) and 66% Calloway silt loam (Fine-silty, mixed, active, thermic Aquic Fraglossudalfs) (United States Department of Agriculture Natural Resources Conservation Service, 2020).

Rice Agronomics

The rice hybrid Gemini 214CL (Rice Tec, Inc., Alvin, TX) was drill seeded on 16 May 2019 using a seeding rate of 25 kg ha^{-1}.

TABLE 2 | Flyover dates for the 2019 season.

Date	Days after planting	Growth stage	Average canopy height (cm) (n = 5)	Leaf area index (LAI) (n = 2)	Floodwater depth (cm) (n = 5)
Apr. 04	NA	NA	NA	NA	NA
May 21	6	Pre-emergence	NA	NA	0
June 13	29	Vegetative —tillering	25	0.26–0.30	0
June 29	45	Vegetative – Tillering	55	2.42–5.43	9–15
July 11	57	Reproductive—R0—panicle initiation	73	6.79–7.17	0–5
Aug. 01	78	Reproductive—R2—booting	82	7.16–7.27	0–3
Aug. 13	90	Reproductive—R4—flowering anthesis	111	6.51–6.57	15–23
Aug. 21	98	Grain filling—maturation—hard stage	104	6.23–6.47	10–14
Aug. 28	105	Grain filling—maturation—hard stage	104	NA	0
Sep. 07	115	Grain filling—maturation—hard stage	104	6.50–6.83	0
Sep. 13	121	One day before harvest	NA	NA	0

Floodwater depth and LAI are given as ranges for the minimum and maximum values. For flyover days that occurred between dates when LAI or floodwater depth were measured, values are from the nearest day (within ± 4 days). Measurements were performed in the north side of the field within a 15-m radius.

Fertilizer was applied on 03 June (20 kg ha^{-1} N and 52 kg ha^{-1} P_2O_5 using diammonium phosphate), 11 June (101 kg ha^{-1} K_2O using potassium chloride), 12 June (101 kg ha^{-1} N using urea), and 25 June (50 kg ha^{-1} N using urea). To conserve water resources, the field was irrigated using alternate wetting and drying flood management (Bouman and Tuong, 2001; Henry et al., 2017).

The field was mechanically harvested on 14 September 2019 using a commercial combine and a circular harvest pattern with an 8.5-m header width. Rough rice yield was measured using a calibrated yield monitor (GPS-enabled John Deere Greenstar 3 2630 harvest monitor). Yield data were excluded from a 10-m buffer surrounding the field perimeter and associated drainage ditch. The data were checked and filtered using Yield Editor software (Sudduth and Drummond, 2007), and the harvest grain moisture content was 14.6%. ArcGIS software was used to develop a raster layer with a spatial resolution of 50 cm, using the spherical model, within the Kriging/CoKriging tool (Burrough, 2001). To further characterize the study site, throughout the growing season, growth, floodwater depth, canopy height and Leaf Area Index (LAI) were measured within 4 days of each flyover date (**Table 2**). Canopy height and floodwater depth were measured in five locations, while LAI was measured in two flagged locations with a total area of 1 m^2 for each location using the LAI-2200C (LI-COR Biosciences). These measurements were performed on the north side of the field within a 15-m radius of each other, and the northern field edge.

For further evaluation of trained models, we also considered a separate 27-ha field within the same farm in the 2020 growing season. This field was water seeded (seeds broadcasted from an airplane over a flooded field) with CL XL745 rice hybrid cultivar (Rice Tec, Inc., Alvin, TX) on 02 April 2020 using a seeding rate of 32.5 kg ha^{-1}. Fertilizer was applied on 01 June (22 kg ha^{-1} N and 57 kg ha^{-1} P_2O_5 using diammonium phosphate), 11 June (52 kg ha^{-1} N using urea), 18 June (52 kg ha^{-1} N using urea) and 25 June (52 kg ha^{-1} N using urea). The field was also irrigated using alternate wetting and drying flood management to conserve water resources, and the rice residue was also burnt. The field was harvested on 17 August 2020 using the same combine previously described, and the harvest moisture was 15.6%. Field soil is classified as silty clay: Perry Silty Clay (Very-fine, smectitic, thermic Chromic Epiaquerts) (United States Department of Agriculture Natural Resources Conservation Service, 2020).

Unmanned Aerial Vehicle Data Collection

A UAV with an Altum sensor (multispectral and thermal) was used for image data collection. Data were collected at approximately 7-day intervals, weather permitting (**Table 2**). The Matrice 210 V-2 quadcopter (DJI, Shenzhen, Nanshan District, China) was used and equipped with an Altum sensor (MicaSense, Seattle, Washington) to collect blue (B, 475 nm), green (G, 560 nm), red (R, 668 nm), red edge (RE, 717 nm), near-infrared (NIR, 840 nm), and thermal (11,000 nm) data. Data collection occurred within 2 h of solar noon local time. Prior to each flight, radiometric calibration images were captured (MicaSense, Seattle, Washington). Flight design parameters were calculated using the MicaSense flight calculator, while the Atlas Flight application was used to deploy flight missions (MicaSense, Seattle, Washington). The flight altitude was 120 m above ground level (AGL), and horizontal velocity was 10 m s^{-1} with 75% front/side overlap. The Pix4D mapper software (Pix4D Inc., Prilly, Switzerland) was used to stitch the raw imagery, producing orthomosaics. The model builder tool within ArcMap 10.7.1 (ESRI, 2011, Redlands, California) was used to calculate six vegetation indices, including CI_{green}, Normalized Area Vegetation Index (NAVI), NDVI, Red-Edge NDVI (RENDVI), Green NDVI (GNDVI), and Triangular Greenness Index (TGI) based on the equations in **Table 1**. The six derived vegetation indices and the thermal layer were used as the input features for model training.

For the 2020 growing season, UAV data were collected on 05 July 2020 for the 27-ha field only, during booting stage.

Image Processing

After producing orthomosaics and generating vegetation indices, images were further processed in R ver. 4.0 (R Core Team, 2020). Images were downsampled from 5 to 50 cm resolution, using the "aggregate" function of the raster package, and then split into

5 × 5 pixel tiles. This enabled faster processing of image data and a match to the spatial resolution of yield data. Images were then stacked across time. After cropping out the edges of the field and removing tiles with missing values, tiles were partitioned for four-fold cross-validation. Tiles were split into training (∼50%), test (∼25%), and validation (∼25%) datasets, in the four-fold shown in **Figure 1**. This strategy was used so that some field regions were never seen during training, rather than randomly assigning images to the test set, which would inflate model performance. For the two deep learning models, the validation set is used during model training, where model weights are updated each epoch if performance on the validation set improves; the test set is held out for the final evaluation after training. For the statistical models (linear, null, and XGBoost models), the training and validation sets can be considered equivalently. Non-overlapping tiles of 5 × 5 pixels were output as .csv files and then converted into .npy arrays for faster reading in Python 3.

Model Training

Evaluation Metrics

To evaluate each model, four statistical parameters were used to assess model performance: Root Mean Squared Error (RMSE), coefficient of determination (R^2), mean absolute error (MAE), and mean bias error (MBE), calculated as follows:

$$\mathrm{RMSE} = \left[\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2\right]^{0.5}$$

$$R^2 = \frac{[\sum_{i=1}^{n}(Y_i - \bar{Y}_{obs})(\hat{Y}_i - \bar{Y}_{exp})]^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_{obs})^2 \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y}_{exp})^2}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{Y}_i - Y_i|$$

$$\mathrm{MBE} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i - Y_i\right)$$

where Y_i is the observed yield for pixel i, $\hat{Y}_i$ is the predicted yield for pixel i, n is the total number of pixels in the dataset, $\bar{Y}_{obs}$ is the mean observed yield for all pixels, and $\bar{Y}_{exp}$ is the mean predicted yield for all pixels. MAE and RMSE measure the average magnitude of difference in the observed and predicted response, with RMSE placing greater penalty on large errors. MBE is also a measurement of the error between the predicted and observed response but takes into account the sign of the errors. However, MBE should be interpreted with caution as large errors may cancel each other out if they are in the opposite direction. R^2 represents the proportion of variance in the dataset that is explained by the model.

Null Models

As a baseline for comparison, we considered the difference between each pixel and a constant layer assigned the value of the mean yield calculated from all pixels assigned to the training set. Evaluation metrics for the null model on the training set vary slightly across time points as a result of differences in the number of missing values on different days.

Linear Models

We fit linear models to predict yield using seven predictors (all six vegetation indices and thermal rasters) using the lm() function from R version 4.0 (R Core Team, 2020).

XGBoost

We trained gradient boosted decision trees using the R implementation of XGBoost (Chen and Guestrin, 2016; Chen et al., 2021). This model was designed to capture complex interactions among predictor values, but did not consider spatial or temporal dimensions of our data structure. We did not perform extensive parameter tuning for every individual model, but chose parameter values that gave similar performance on training sets as observed for 2D-CNN models on a subset of data. Specifically, we used default settings with the exception of a slower learning rate (eta = 0.2), a maximum tree depth of 2 (max_depth = 2) to capture only pairwise interactions among predictors, and 200 rounds of training (nrounds = 200). The same parameter values were used to train all XGBoost models.

2D-Convolutional Neural Networks

We included "spatial models" (2D-CNNs) to determine whether considering information from nearby pixels improved yield prediction. We developed a 2D-CNN with an autoencoder-like structure. An autoencoder is a neural network trained to encode data into a compressed representation and then reconstruct the original data from the encoded representation (**Figure 1**). Here, we take advantage of this type of architecture to predict an output 5 × 5 pixel image of yield, based on an input 5 × 5 image patch for the same location acquired by a UAV. Each image was associated with seven input features, corresponding to values from the thermal sensor or for a different vegetation index (**Figure 1**). We did not train models using data for 13 June 2020 due to a malfunction in the thermal sensor.

Our 2D-CNN was implemented in Python 3.8 using Keras with a TensorFlow v2.2.0 backend (Abadi et al., 2015; Chollet, 2015). The final architecture involved one sub-network of three convolutional layers for encoding, followed by two fully connected layers, and a second sub-network of three convolutional layers for decoding. The parameter specifications for each layer are shown in **Table 3**. A "ReLU" activation function was used for each layer in the network besides the last layer, which used a linear activation function. We used the "adam" optimizer and quantified loss based on the mean squared error. CNNs were trained for 50 epochs, and weights for models with the best performance on the validation set were saved to evaluate performance on the test set. Preliminary models were trained for up to 200 epochs, but only minor improvements in model performance were observed with additional training.

3D-Convolutional Neural Networks

To determine whether considering information from nearby time points improved yield prediction, we developed "spatial-temporal models" (3D-CNNs). For this analysis, we used 5 days beginning just prior to the reproductive phase (flyover dates from 29 June 2019 through 21 August 2019), which ended

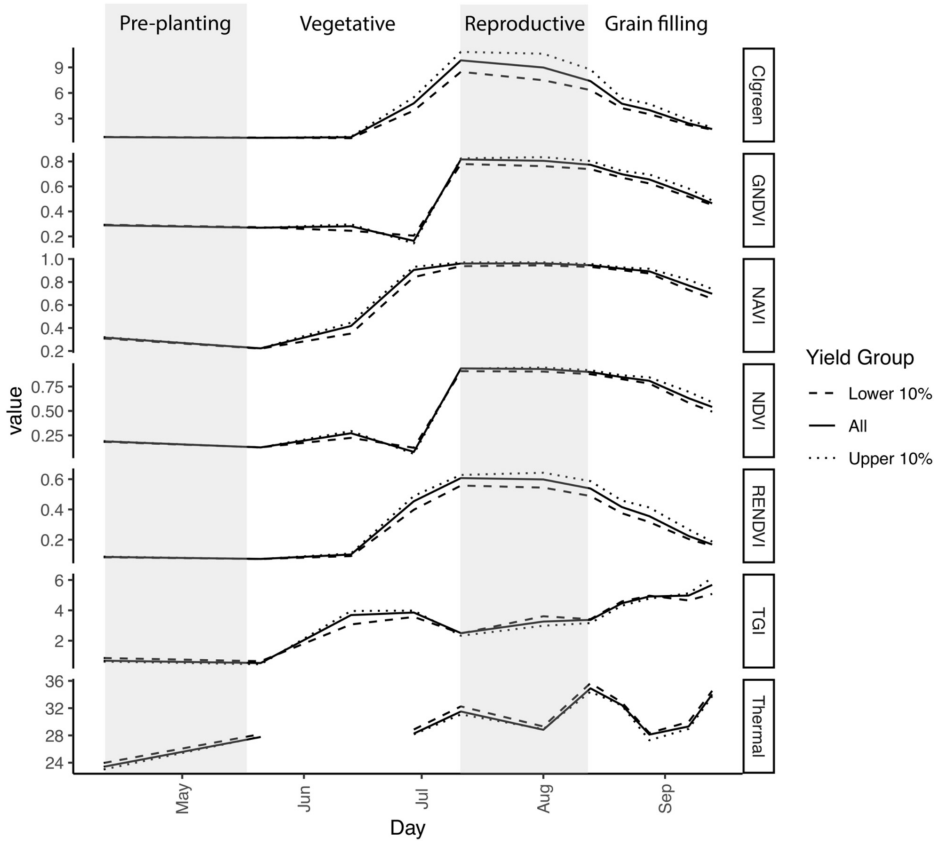

FIGURE 2 | Variation in temperature and six vegetation indices over time in 2019. Gray shaded areas indicate time prior to seeding and the reproductive phase (panicle initiation through flowering anthesis). Additional details regarding growth stages for each flyover date are provided in **Table 2**. First flooding occurred on 29 June, resulting in a dip in NDVI and GNDVI. Plotted values are the mean for 50 cm × 50 cm pixels with yield in the lower 10% (n = 52,649), the upper 10% (n = 52,537), or all pixels (n = 526,735). Thermal was not available for the flyover on 13 June.

approximately 3 weeks prior to harvest and also included the days that we anticipated to be most informative with respect to variation in vegetation indices (**Figure 2**). We also tested 3D-CNNs that included all 11 time points, but found early on during testing that they primarily learned to weight features from the final time point, just prior to harvest. Our 3D-CNNs were designed to have a parallel structure to our 2D-CNNs with the exception that convolutions occurred in three dimensions in the encoding stage of the network (**Figure 1**).

Computational Infrastructure

The XGBoost, linear, and null models were trained in minutes or less on a personal desktop computer (16 Gb RAM; Intel Core i5 3 GHz processor). Each 2D-CNN was trained on a single node of the Trestles cluster at the University of Arkansas High Performance Computing Center (AHPCC). Each of these nodes is equipped with 64 Gb of memory and four AMD 6136 2.4 GHz CPUs for a total of 32 cores; 2D-CNNs required approximately five hours to train (eight CPU hours) and a maximum of 10 Gb of virtual memory. We used the same computing infrastructure for training 3D-CNNs as for 2D-CNNs; each 3D-CNN required approximately 18 hours (24 CPU hours) to train.

Model Comparison

In comparing our models, we sought to answer three questions: (1a) Do model architectures that capture spatial information improve yield prediction over traditional statistical learning approaches? (1b) If so, do models that also include data from multiple time points improve yield prediction over models that only include spatial information? (2) Which day(s) have the strongest signal for deep-learning based yield prediction? and (3) What are the most important spectral features for prediction?

For the first two questions, we compared average test set RMSE across time points for all models. To qualitatively determine the impact of different model architectures on yield predictions, we also projected models to field scale. Input images containing vegetation indices and thermal layers were processed in R as described for model training. CNN models were loaded into R using the "reticulate" package to enable interoperability between R and python codes (Ushey et al., 2020). Predictions for each tile were generated iteratively and tiled together for the prediction map for the field.

For the last question, we determined the relative importance of each feature for the trained 2D-CNNs by removing variation observed for that feature in input images from the test set. To

TABLE 3 | Parameters for CNNs.

Layer	Parameters (2D-CNN)	Parameters (3D-CNN)
Conv2D (for 2D-CNN) or Conv3D (for 3D-CNN)	Filters = 64 Kernel_size = (3,3) Padding = "same" Activation = "relu" Input_shape = (5,5,7)	Filters = 64 Kernel_size = (3,3,3) Padding = "same" Activation = "relu" Input_shape = (5,5,5,7)
Conv2D (for 2D-CNN) or Conv3D (for 3D-CNN)	Filters = 128 Kernel_size = (2,2) Padding = "same" Activation = "relu"	Filters = 128 Kernel_size = (2,2,3) Padding = "same" Activation = "relu"
Conv2D (for 2D-CNN) or Conv3D (for 3D-CNN)	Filters = 256 Kernel_size = (1,1) Padding = "valid" Activation = "relu"	Filters = 256 Kernel_size = (1,1,5) Padding = "valid" Activation = "relu"
Reshape	Target_shape = (5,5,256)	
Flatten	NA	
Dense	Units = 256 Activation = "relu"	
Dense	Units = 5,400	
Reshape	Target_shape = (5,5,256)	
Conv2D	Filters = 128 Kernel_size = (2,2) Padding = "same" Activation = "relu"	
Conv2D	Filters = 64 Kernel_size = (3,3) Padding = "same" Activation = "relu"	
Conv2D	Filters = 1 Kernel_size = (1,1) Padding = "same" Activation = "linear"	

Each Keras layer refers to a building block of the neural network, including convolution layers (Conv2D and Conv3D), reshaping layers (Flatten and Reshape), and fully connected layers (Dense). Besides the last layer, all layers used a rectified linear unit ("relu") activation function that directly outputs the input, if positive, or zero otherwise.

"blank" variation in a feature, all actual values for that feature in each tile were replaced by the mean value observed across all pixels in the test set. Test set RMSE was then determined using the function call to Keras "evaluate" (Chollet, 2015).

Code Availability

Python and R code used to process data, train and evaluate models, and recreate **Figures 2–5**, is available at https://github.com/em-bellis/XASU_rice.

RESULTS

Rice Yield Variation

We first characterized intra-field variation in yield in the 2019 growing season. Rough rice yield was 9.06 ± 0.9 Mg ha^{-1} (mean ± std. dev.) measured across 526,735 grid cells at 50-cm resolution. As expected, vegetation indices varied with rice crop developmental stage and differed between high- and low-yielding areas of the field (**Figure 2**). CI_{green} and RENDVI showed the greatest contrast during reproduction, peaking at booting stage (CI_{green}) or flowering (RENDVI) in the highest-yielding areas of the field. TGI values also differed among high- and low-yielding areas of the field, particularly during vegetative growth and booting stage (**Figure 2**).

Spatial vs. Non-Spatial Models

We next evaluated the ability of deep learning-based, spatially explicit models to predict yield from vegetation index and thermal feature information. Compared to the null model, all models showed improved performance during training, indicating that vegetation indices and thermal features provided useful information for predicting yield (**Figure 3**). Linear models performed worst for the training set data for eight out of 10 days. Non-spatial (XGBoost) models performed best on training set data for six of 10 days, reaching the best performance on images acquired a week prior to harvest (**Figure 3**).

Performance on test sets, however, revealed a clear benefit of our deep learning-based spatial models for predicting grain yield both in terms of higher accuracy and lower variability in predictions across folds (**Figure 3** and **Table 4**). Similar ranking of models was observed for all metrics (**Table 4**). 2D-CNNs trained on images taken during booting stage (01 Aug) or later showed the best performance (RMSE: 7.4–8.2% of mean yield; **Figure 3**). Average test RMSE of XGBoost models during these same developmental stages was higher, ranging from 8.5 to 10.3% of the mean yield. Performance of XGBoost models was also highly variable across folds, with standard deviation up to 7.4% of the mean yield vs. 4.2% in 2D-CNNs (n = 4 folds, based on observations over all time points). The difference in performance of XGBoost models on training and test sets may be indicative of overfitting. However, even for days on which performance of CNNs and XGBoost models on the training set was nearly identical (i.e., 29 June, 21 Aug, and 28 Aug), 2D- and 3D-CNNs showed markedly better performance on the test set (**Figure 3**). 2D-CNNs also outperformed other models with respect to MAE and R^2, though not with respect to MBE, suggesting that although they were more accurate, 2D-CNNs tended to overpredict slightly more than other models (**Table 4**). Models trained on images collected from the booting stage or later performed best, though 2D-CNNs trained on images from earlier time points also performed notably better than other models (RMSE: 8.7–9.3% of mean yield; **Figure 3**).

To further evaluate the benefit of our spatial models, we projected predictions from 2D-CNNs trained during the booting stage to field scale (**Figure 4**). This analysis suggested that a main benefit of the 2D-CNN model, compared to models that do not incorporate information from nearby pixels, may be a spatial denoising effect of the 2D-CNN. Compared to less complex models, CNNs were less likely to underpredict yield, particularly where yields were higher (**Figures 4A,B**).

Spatial vs. Spatial-Temporal Models

We observed comparable performance for the two deep learning models using the tested architectures. Average test RMSE for

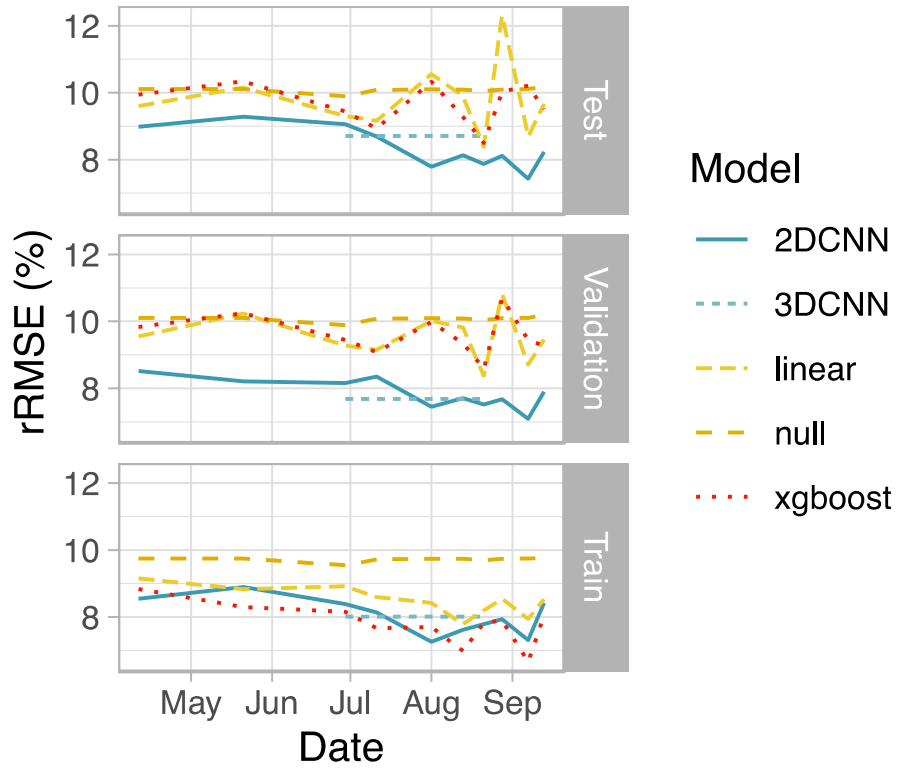

FIGURE 3 | Summary of model performance. Root mean squared error (RMSE) relative to the mean yield for the field is shown as the average across all four data folds. Note that results for the 3DCNN are for a single model based on input from five time points together, with relative RMSE (rRMSE) shown as a constant value across the five included time points.

3D-CNNs only exceeded that of 2D-CNNs on 29 June and 11 July, likely due to the fact that the 3D-CNN model also included data from the more informative, later time points (**Figure 3**). Our results suggest that 2D-CNN models provide a benefit for the task of yield prediction in rice over simpler models and may offer similar performance to some deep learning architectures that incorporate data from multiple timepoints. Future studies may find further benefit of temporal network architectures relative to the 2D-CNNs tested here, for example by altering the intervals of the selected time points.

Spectral Feature Importance

The cost of a UAV increases with the number of sensors it carries and sensor complexity. To assess if it is possible to achieve similar prediction accuracy with fewer sensors or bands, we determined the importance of each index on model performance of booting stage 2D-CNNs. Booting stage is early enough to be useful to the farmer, such as for determining the need for late boot nitrogen fertilization of rice hybrids (Hardke, 2018). It was also found to have one of the lowest RMSE values (**Figure 3**).

At booting stage, CI_{green} was the most important feature (index) for predicting rice yield with 2D-CNNs (**Figure 5**). Depending on the fold, test-set RMSE increased by 0.03–0.17 Mg ha^{-1} when variation among pixels in CI_{green} was removed, consistent with high variation in CI_{green} among yield groups at this time point (**Figure 2**). TGI and thermal information were also important, but only for some data folds (**Figure 5**). Other indices appeared to matter little to overall model performance, with negligible or even positive effects on model performance when observed variation in those features was removed (**Figure 5**). However, since calculation of TGI relies on three bands (red, green, blue), CI_{green} relies on two bands (green, NIR), and thermal information was also useful for some models (**Figure 5**), a UAV equipped with all sensors is recommended to achieve levels of performance reported here on other datasets.

Generalization to New Datasets

To explore the extent to which our findings may generalize to new contexts, we evaluated performance of late booting stage models from 2019 (**Figure 4**) on a separate, nearby field imaged in the 2020 growing season. All 2019 models underpredicted yield in 2020 (**Figure 6**), consistent with substantially higher mean yield for the 27-ha field compared to the training dataset (11.4 vs. 9.1 Mg/ha). Among all single-day UAV-based models, the 2D-CNN model had the highest accuracy, indicating it was also more translatable to a different field and growing season compared to the other models (**Figure 6**).

Further improving performance in new contexts will require a greater diversity of training images for different rice cultivars, growing seasons, soil types, and management conditions. To inform future experimental design, we determined the extent to

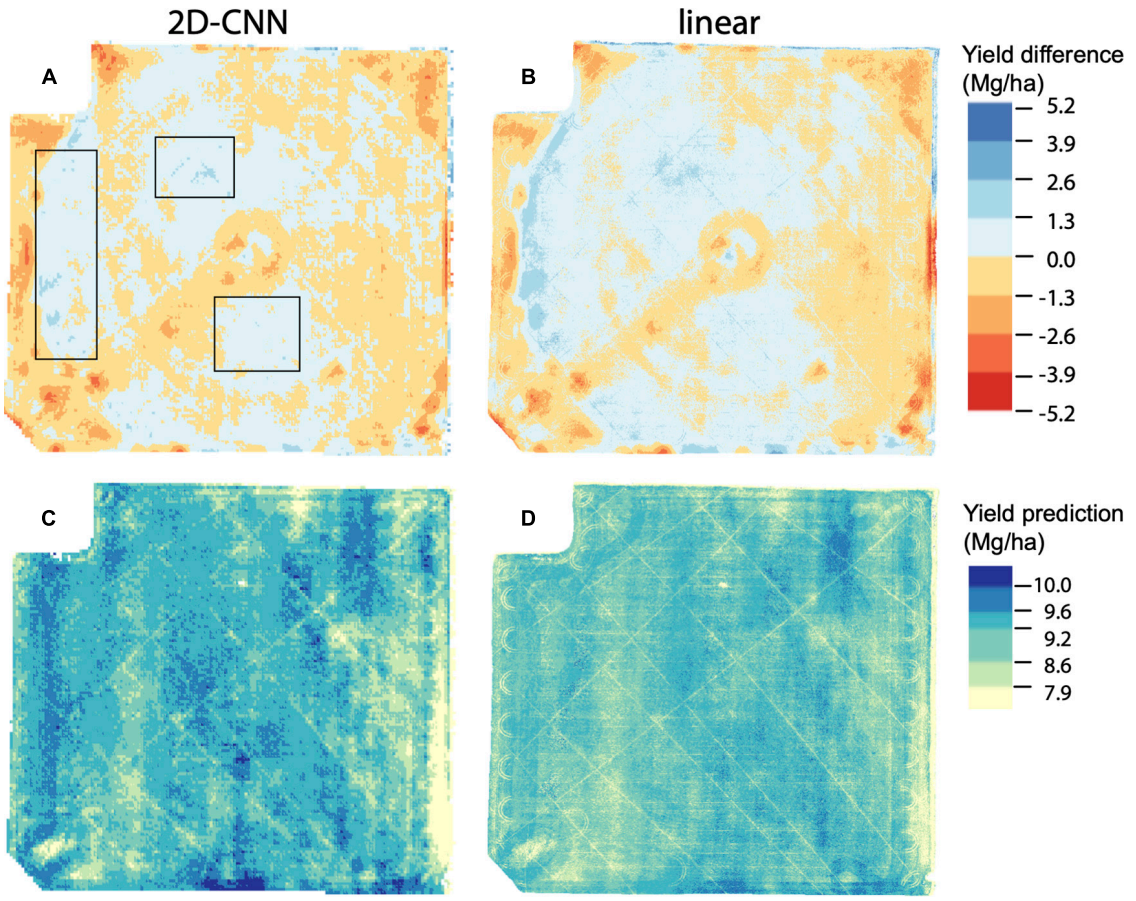

FIGURE 4 | Yield prediction maps based on data for 01 Aug 2019, for models trained on the fold B dataset. Data for 2D-CNN is shown in **(A,C)** whereas data for linear model is shown in **(B,D)**. Prediction error (yield difference) is shown in **(A,B)** and represents observed yield minus predicted yield; **(C,D)** show predicted yield. Black boxes in **(A)** indicate regions described in main text where yield was underpredicted to a considerably lesser degree in the 2D-CNN model compared to the linear model. For **(A,B)**, breakpoints for color scale are evenly spaced. For **(C,D)**, breakpoints for the color scale are chosen based on the 10th, 25th, 50th, 75th, and 90th percentile values of observed yield. Mg/ha: megagrams/hectare.

which similar performance could be expected for models trained on smaller datasets. A subset of 1,000 tiles was randomly selected from the fold B dataset (**Figure 1**; ∼10% of tiles compared to full-scale training). After 50 epochs, RMSE as low as 0.77 Mg/ha was observed for the smaller training set; in contrast, lower RMSE (0.68 Mg/ha) was achieved within 50 epochs for the full training set. Given the modest increase in RMSE with the smaller dataset, it may be prudent to train future deep learning models using at least a similar-sized training set (∼7 ha) as the full-scale training set used here.

DISCUSSION

In this study, we present an autoencoder-like CNN architecture for intra-field prediction of rice yield. The best single-day model showed improved performance compared to simpler models trained on the same data, and comparable or improved performance to similar UAV-based studies in wheat and barley (Nevavuori et al., 2019), rice (Yang et al., 2019; Wan et al., 2020; Duan et al., 2021), and soybean (Maimaitijiang et al., 2020; **Table 5**). With respect to yield prediction in rice, we report slightly better performance of our late booting stage 2D-CNN compared to Yang et al. (2019), after accounting for higher average yield in our study [RMSE of 0.72 (**Table 4**) vs. 0.76 Mg/ha (Yang et al., 2019)]. With respect to RMSE, we report slightly lower performance for rice yield prediction compared to two other studies; however, these studies tested performance using leave-one-out cross-validation (Duan et al., 2021) or random samples distributed throughout the field (Wan et al., 2020), which could inflate performance compared to the spatially explicit strategy for cross-validation used here (**Figure 1**). Compared to other studies, R^2 values for our model were relatively low, likely because of greater amount of overall yield variation in other studies due to experimental nitrogen treatment (Wan et al., 2020) and differing management practices (Yang et al., 2019). Our findings additionally suggest a benefit of autoencoder-like 2D-CNNs for spatial denoising of yield predictions by incorporating information from nearby pixels. With the exponential rise in

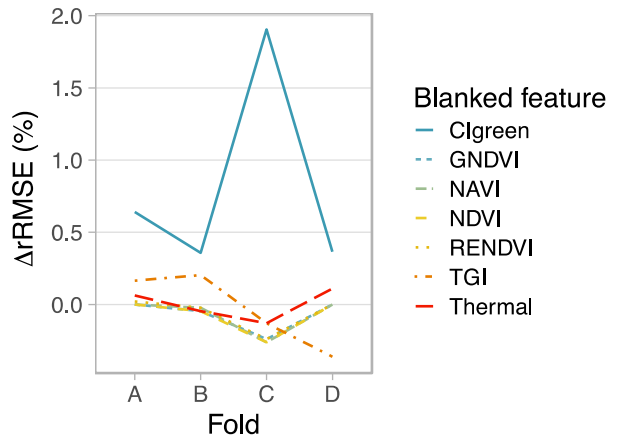

FIGURE 5 | Feature importance for model training on images from 01 Aug 2019. To blank observed variation in a feature, actual values were replaced by the mean value across all pixels observed for that feature. ΔRMSE was calculated by subtracting the test set RMSE of the original model with no feature blanked from the blanked feature model. ΔRMSE is given relative to the mean yield (ΔrRMSE). Larger, more positive ΔrRMSE values indicate higher feature importance for the respective feature.

adoption of UAVs for remote sensing in agriculture (Maes and Steppe, 2019), this study provides timely guidance for future large-scale training data collection efforts and their integration with development of deep-learning models.

Surprisingly, we observed similar performance for yield prediction for 2D-CNNs as for a 3D-CNN architecture using data from multiple time points. However, 3D-CNN architectures may show a greater increase in performance if trained on diverse datasets that include multiple rice cultivars and environments, particularly if there are significant cultivar- or environment-specific differences in the pattern of vegetation index change over time (Duan et al., 2021). Exploring the benefit of 3D-CNNs for better generalization across climate zones and cultivars is a promising area for future work, since a primary benefit of these architectures may be the ability to take into account shifts in phenology across different climates and cultivars. The dataset utilized here, which focuses on fine-scale yield prediction across a large, heterogeneous field for a single year, minimizes variation due to cultivar and environmental differences, and so any temporal variation in vegetation indices associated with yield may not contribute to a strong spectral signature in the dataset. Our pre-processing pipeline also does not include any explicit classification of soil- or weed-derived pixels, or inclusion of canopy structure/texture features (e.g., Maimaitijiang et al., 2020), which could also impact the relationships among timepoints and the relative performance of 3D-CNNs. Use of vegetation index features that are less sensitive to saturation and soil background effects (e.g., Yang et al., 2019) is another strategy which might influence 3D-CNN performance relative to 2D-CNNs.

Our results also highlight the potential for UAVs to support management recommendations even during early growth stages (Nevavuori et al., 2019). Although the best single-day models were obtained during booting stage or later, the 2D-CNNs showed considerably better performance than other models even when trained on data acquired during vegetative growth stages (**Figure 3**). This difference in prediction for 2D-CNNs vs. other models was observable even prior to planting when the performance of single-day models was surprisingly competitive with models based on information from later in the season (**Figure 3**). Other studies demonstrate the success of 2D-CNNs for corn yield prediction based only on pre-season variables, including soil electroconductivity maps and satellite imagery acquired after soil tillage (Barbosa et al., 2020). Bare-soil images taken by UAVs prior to planting may also capture features that correlate with soil properties important to yield (Khanal et al., 2018).

For future large-scale efforts on a greater diversity of rice cultivars from different fields, regions, years, and management conditions, our results suggest it may be worthwhile to focus data collection at time points just prior to common crop management intervention points. In turn, growth-stage specific single-day models can be trained using these data. For

TABLE 4 | Evaluation of model performance on the test set for single-day models (trained on data at late booting stage) and 3D-CNN.

Model	Fold	RMSE (Mg/ha)	MBE (Mg/ha)	MAE (Mg/ha)	R^2
Null	A	0.93	−0.50	0.79	n.d.
	B	0.75	0.08	0.57	n.d.
	C	0.80	0.30	0.60	n.d.
	D	0.83	0.13	0.62	n.d.
	Mean	*0.83*	*0.00*	*0.65*	*n.d.*
Linear	A	1.06	−0.78	0.93	0.17
	B	0.70	0.07	0.51	0.18
	C	0.94	0.59	0.70	0.15
	D	0.83	0.10	0.63	0.04
	Mean	*0.88*	*−0.01*	*0.69*	*0.14*
XGBoost	A	0.88	−0.45	0.72	0.11
	B	0.74	−0.02	0.53	0.19
	C	0.83	0.50	0.61	0.22
	D	0.83	0.11	0.63	0.04
	Mean	*0.82*	*0.04*	*0.63*	*0.14*
2D-CNN	A	0.73	−0.19	0.57	0.18
	B	0.68	0.13	0.47	0.29
	C	0.63	0.02	0.45	0.30
	D	0.84	0.28	0.62	0.10
	Mean	*0.72**	*0.06**	*0.53*	*0.22*
3D-CNN	A	0.80	−0.37	0.67	0.08
	B	0.67	0.17	0.48	0.27
	C	0.77	−0.24	0.61	0.37
	D	0.90	0.45	0.67	0.06
	Mean	*0.79*	*0.00*	*0.61*	*0.20*

**Indicates significant difference in mean value between 2D-CNN and linear model only ($p \leq 0.05$; one-way ANOVA).*
RMSE, root mean squared error; MBE, mean bias error; MAE, mean absolute error;R^2, coefficient of determination. n.d., not defined (observed ~ predicted yield is a vertical line for the null model).

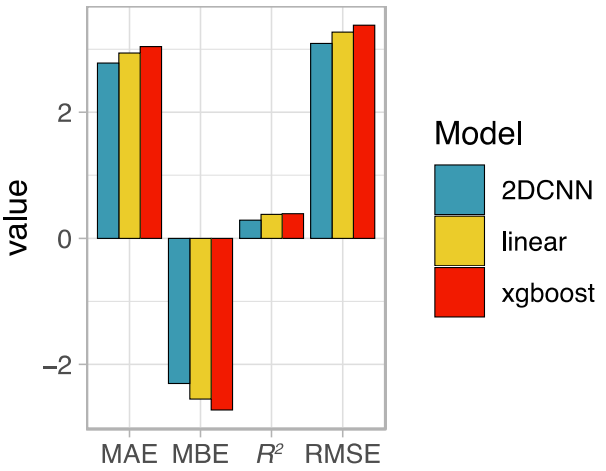

FIGURE 6 | Average performance of late booting stage models from 2019, evaluated on a separate nearby field in the 2020 growing season. MAE, mean absolute error; MBE, mean bian error; R^2, coefficient of determination; RMSE, root mean squared error; 2DCNN, two-dimensional convolutional neural network.

example, Arkansas currently recommends nitrogen fertilization after internode elongation (for some cultivars) or at late booting (for hybrid cultivars) (Hardke, 2018); the most robust models might be explicitly trained for optimal performance at those stages. The use of growth-stage specific models may be particularly valuable because the importance of different vegetation indices for yield prediction varies over time (**Figure 2**; Duan et al., 2019). Compared to 3D-CNNs, 2D-CNNs would also require less computational and environmental resources for training (Strubell et al., 2019; Henderson et al., 2020; Bender et al., 2021) and fewer flyovers to generate yield predictions when models are deployed.

Further contributing to the simplicity of our deep learning models is the lower resolution of images used for the models in our study (50 cm) compared to the resolution of images available from the Altum sensor (5 cm). Using down-sampled images, our 2D-CNNs reached maximum performance relatively early during training. Contributing to this, the true relationship between yield and vegetation indices may be relatively simple; high linear correlations with yield are often reported (Duan et al., 2019). Furthermore, higher resolution of input images would not match the scale of accuracy of yield maps generated using data collected by commercial harvesters (**Figure 1**). Conversely, without a combine yield monitor, it would be very difficult to acquire a sufficient volume of labeled data needed to train deep learning models.

If the relationship between vegetation indices and yield is relatively simple, and the resolution of imagery used here precludes automated detection of individual objects in images, what is the utility of our CNN architectures for yield prediction? One of the primary benefits may be an image denoising effect of the autoencoder-like model architecture. Autoencoders have been widely successful for image denoising for a variety of applications (Xie et al., 2012). Robustness to partial destruction of the input is a characteristic of particular interest for denoising autoencoders (Vincent et al., 2008). Our study suggests that similar architectures are also useful for denoising "outputs." For example, although yield maps used for training included noise (e.g., circular impressions due to the driving pattern of the combine harvester), these patterns are absent in prediction maps (**Figures 4C,D**). Future models trained to predict yield using higher resolution images from UAVs might benefit from a two-stage approach, where yield maps from a combine harvester

TABLE 5 | Model performance for comparable studies using UAV imagery for yield prediction.

References	Crop	Model	Performance	Description
This study	Rice	2D-CNN	7.9% (rRMSE)	Yield predicted from thermal and six VIs using data at late booting stage.
			5.8% (MAPE)	Performance based on 4-fold cross validation from the same field and season.
			0.22 (R^2)	
Duan et al. (2021)	Rice	Neural network	5.3–7.1% (rRMSE)	Yield predicted on two individual VIs from 6 or more imaging days.
			0.48–0.62 (R^2)	Performance based on leave-one-out cross validation from the same field and season.
Wan et al. (2020)	Rice	random forest	2.75% (rRMSE)	Yield predicted from four RGB- and multispectral-derived features. Data set included substantial yield variation due to experimental nitrogen treatment.
			0.83 (R^2)	Performance based on random held-out set from the same field and season.
Yang et al. (2019)	Rice	2D-CNN	26.6% (MAPE)	Yield predicted from raw RGB and multispectral imagery at ripening stage.
			0.49 (R^2)	Performance based on held-out set of independently managed plots from the same season.
Maimaitijiang et al. (2020)	Soybean	2D-CNN	15.9% (rRMSE)	Yield predicted from 72 features derived from multispectral, thermal, and RGB sensors on a single day. Data set included substantial yield variation due to cultivar-specific differences.
			0.72 (R^2)	Performance based on held-out set from the same field and season.
Nevavuori et al. (2019)	Wheat/barley	2D-CNN	8.8–12.6% (MAPE)	Yield predicted from RGB or a single VI measured on a single day. UAV data were combined for two crops, nine fields, and multiple imaging dates. Images for "early" or "late" season models were sub-sampled, shuffled, and split into test and train sets.

rRMSE, relative root mean squared error; MAPE, mean absolute percentage error.

first undergo error correction using the model architecture presented here (**Figure 1**).

Taken together, our study highlights the benefits of relatively simple CNN architectures for yield prediction in rice using remotely sensed images. Incorporating such models into data analysis pipelines could balance the overall costs of data collection and model training and demonstrates the potential benefits of deep learning for sustainable agriculture and precision management.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

AH, EB, JC, BR, TB, MR, and XH designed the study. BR and MR directed the experimental study. BR directed collection of yield data and other study site agronomic information, with help from BM-G. AH directed collection and processing of UAV images, with help from BB. EB developed CNNs with input from XH and JC. The manuscript was drafted by EB with input from AH, BR, MR, VG, JC, and BM-G. All authors contributed to manuscript revisions and read and approved the final version.

FUNDING

This work was supported by an award from the University of Arkansas System Division of Agriculture Research and Extension Service to XH and funding from the Arkansas Biosciences Institute (the major research component of the Arkansas Tobacco Settlement Proceeds Act of 2000) and the Judd Hill Foundation. Field data collection was supported by the USDA-NRCS under Cooperative Agreements 69-7103-17-004 and 68–7103–17–119, and the NSF under CBET Award 1752083. This research was supported by the Arkansas High Performance Computing Center which was funded through multiple National Science Foundation grants and the Arkansas Economic Development Commission.

ACKNOWLEDGMENTS

We thank Juan Arguijo, Anthony Zadoorian, Colby Reavis, and Bennett Barr for their assistance in field data collection and Mark Isbell of Isbell Farms / Zero Grade Farms for hosting and helping manage our experiment. We also thank Joseph Massey for helpful feedback on early versions of the manuscript.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Available online at: http://tensorflow.org/ (accessed May 01, 2020).

Barbosa, A., Trevisan, R., Hovakimyan, N., and Martin, N. F. (2020). Modeling yield response to crop management using convolutional neural networks. *Comput. Electron. Agric.* 170:105197. doi: 10.1016/j.compag.2019.105197

Bender, E. M., Gebru, T., McMillan-Major, A., and Shmitchell, S. (2021). "On the dangers of stochastic parrots: Can language models be too big?," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.FAccT '21*, (New York: Association for Computing Machinery), 610–623. doi: 10.1145/3442188.3445922

Bouman, B. A. M., and Tuong, T. P. (2001). Field water management to save water and increase its productivity in irrigated lowland rice. *Agric. Water Manag.* 49, 11–30. doi: 10.1016/S0378-3774(00)00128-1

Burns, B. W., Green, V. S., Hashem, A. A., Massey, J. H., Shew, A. M., Adviento-Borbe, M. A. A., et al. (2022). Determining nitrogen deficiencies for maize using various remote sensing indices. *Precis. Agric.* doi: 10.1007/s11119-021-09861-4

Burrough, P. A. (2001). GIS and geostatistics: essential partners for spatial analysis. *Environ. Ecol. Stat.* 8, 361–377. doi: 10.1023/A:1012734519752

Carmona, F., Rivas, R., and Fonnegra, D. C. (2015). Vegetation Index to estimate chlorophyll content from multispectral remote sensing data. *Eur. J. Remote Sens.* 48, 319–326. doi: 10.5721/EuJRS20154818

Chen, T., and Guestrin, C. (2016). "XGBoost: A scalable tree boosting system" in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data*. New York: ACM digital library.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., et al. (2021). *xgboost: eXtreme Gradient Boosting*. Available online at: https://CRAN.R-project.org/package=xgboost (accessed May 01, 2020).

Chollet, F. (2015). *Keras*. Available online at: https://github.com/fchollet/keras (accessed May 01, 2020).

Duan, B., Fang, S., Gong, Y., Peng, Y., Wu, X., Zhu, R., et al. (2021). Remote estimation of grain yield based on UAV data in different rice cultivars under contrasting climatic zone. *Field Crops Res.* 267:108148. doi: 10.1016/j.fcr.2021.108148

Duan, B., Fang, S., Zhu, R., Wu, X., Wang, S., Gong, Y., et al. (2019). Remote estimation of rice yield with unmanned aerial vehicle (UAV) data and spectral mixture analysis. *Front. Plant Sci.* 10:204. doi: 10.3389/fpls.2019.00204

ESRI (2011). *ArcMap: Release 10*. Redlands, CA: Environmental Systems Research Institute.

Gitelson, A. A., Gritz, Y., and Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* 160, 271–282. doi: 10.1078/0176-1617-00887

Gitelson, A. A., and Merzlyak, M. N. (1997). Remote estimation of chlorophyll content in higher plant leaves. *Remote Sens.* 18, 2691–2697.

Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., and Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* 32, 1–4. doi: 10.1029/2005GL022688

Grassini, P., Eskridge, K. M., and Cassman, K. G. (2013). Distinguishing between yield advances and yield plateaus in historical crop production trends. *Nat. Commun.* 4:2918. doi: 10.1038/ncomms3918

Hardke, J. T. (2018). *Arkansas Rice Production Handbook*. Little Rock: University of Arkansas Division of Agriculture Cooperative Extension Service.

Hatfield, J. L., and Prueger, J. H. (2010). Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices. *Remote Sens.* 2, 562–578. doi: 10.3390/rs2020562

Henderson, P., Hu, J., Romoff, J., Brunskil, E., Jurafsky, D., and Pineau, J. (2020). Towards the systematic reporting of the energy and carbon footprints of machine learning. *J. Mach. Learn. Res.* 21, 1–43.

Henry, C., Joseph, M., Jarrod, H., Jason, K., Michele, R., and Arlene, A. B. (2017). *Using Alternate Wetting & Drying (AWD) Rice Flooding Management*. Arkansas: Arkansas Experiment Station Research Series, 1–5.

Hunt, E. R., Daughtry, C. S. T., Eitel, J. U. H., and Long, D. S. (2011). Remote sensing leaf chlorophyll content using a visible band index. *Agronom. J.* 103, 1090–1099. doi: 10.2134/agronj2010.0395

Hunt, E. R., Doraiswamy, P. C., McMurtrey, J. E., Daughtry, C. S. T., and Perry, E. M. (2013). A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *Int. J. Appl. Earth Observ. Geoinf.* 21, 103–112. doi: 10.1016/j.jag.2012.07.020

Ji, S., Zhang, C., Xu, A., Shi, Y., and Duan, Y. (2018). 3D Convolutional Neural Networks for crop classification with multi-temporal remote sensing images. *Remote Sens.* 10:75. doi: 10.3390/rs10010075

Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., and Shearer, S. (2018). Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Comput. Electron. Agric.* 153, 213–225. doi: 10.1016/j.compag.2018.07.016

Khush, G. S. (2001). Green revolution: the way forward. *Nat. Rev. Genet.* 2, 815–822. doi: 10.1038/35093585

Li, Y., Zhang, H., and Shen, Q. (2017). Spectral–spatial classification of hyperspectral imagery with 3D Convolutional Neural Network. *Remote Sens.* 9:67. doi: 10.3390/rs9010067

Maes, W., and Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends Plant Sci.* 24, 152–164. doi: 10.1016/j.tplants.2018.11.007

Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., and Fritschi, F. B. (2020). Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens. Environ.* 237:111599. doi: 10.1016/j.rse.2019.111599

Minh, D. H. T., Ienco, D., Gaetano, R., Lalande, R., Ndikumana, E., Osman, F., et al. (2018). Deep Recurrent Neural Networks for winter vegetation quality mapping via multitemporal SAR Sentinel-1. *IEEE Geosci. Remote Sens. Lett.* 15, 464–468. doi: 10.1109/LGRS.2018.2794581

Nevavuori, P., Narra, N., and Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. *Comput. Electron. Agric.* 163:104859. doi: 10.1016/j.compag.2019.104859

Pelletier, C., Webb, G. I., and Petitjean, F. (2019). Temporal Convolutional Neural Network for the classification of satellite image time series. *Remote Sens.* 11:523. doi: 10.3390/rs11050523

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Runkle, B. R. K., Suvočarev, K., Reba, M. L., Reavis, C. W., Smith, S. F., Chiu, Y. L., et al. (2019). Methane emission reductions from the alternate wetting and drying of rice fields detected using the eddy covariance method. *Environ. Sci. Technol.* 53, 671–681. doi: 10.1021/acs.est.8b05535

Sharma, L. K., Bu, H., Denton, A., and Franzen, D. W. (2015). Active-optical sensors using red NDVI compared to red edge NDVI for prediction of corn grain yield in North Dakota. *U.S.A. Sens.* 15, 27832–27853. doi: 10.3390/s151127832

Strubell, E., Ganesh, A., and McCallum, A. (2019). "Energy and policy considerations for deep learning in NLP," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence: Association for Computational Linguistics.

Sudduth, K. A., and Drummond, S. T. (2007). Yield Editor: software for removing errors from crop yield maps. *Agronom. J.* 99, 1471–1482. doi: 10.2134/agronj2006.0326

Sun, Z., Di, L., and Fang, H. (2019). Using Long Short-Term Memory Recurrent Neural Network in land cover classification on Landsat and Cropland data layer time series. *Int. J. Remote Sens.* 40, 593–614. doi: 10.1080/01431161.2018.1516313

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. doi: 10.1016/0034-4257(79)90013-0

United States Department of Agriculture Economic Research Service [USDA] (2021). *Rice Yearbook*. Economic Research Service, Department of Agriculture. Available online at: https://www.ers.usda.gov/data-products/rice-yearbook/rice-yearbook (accessed May 01, 2020).

United States Department of Agriculture Natural Resources Conservation Service (2020). *Web Soil Survey, SSURGO Database*. Available online at: https://websoilsurvey.sc.egov.usda.gov/App/WebSoilSurvey.aspx (accessed May 01, 2020).

Ushey, K., Allaire, J. J., and Tang, Y. (2020). *reticulate: Interface to 'Python'*. Available online at: https://CRAN.R-project.org/package=reticulate (accessed May 01, 2020).

Vincent, P., Larochelle, H., Bengio, Y., and Manzagol, P.-A. (2008). "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th International Conference on Machine Learning.ICML '08*, (New York: Association for Computing Machinery), 1096–1103. doi: 10.1145/1390156.1390294

Wan, L., Cen, H., Zhu, J., Zhang, J., Zhu, Y., Sun, D., et al. (2020). Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer – a case study of small farmlands in the South of China. *Agric. For. Meteorol.* 291:108096. doi: 10.1016/j.agrformet.2020.108096

Wang, L., Zhang, F. C., Jing, Y. S., Jiang, X. D., Yang, S. D., and Han, X. M. (2014). Multi-temporal detection of rice phenological stages using canopy spectrum. *Rice Sci.* 21, 108–115. doi: 10.1016/S1672-6308(13)60170-5

Xie, J., Xu, L., and Chen, E. (2012). "Image denoising and inpainting with deep neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1.NIPS'12*, (Red Hook: Curran Associates Inc.), 341–349. doi: 10.3390/e23111481

Yang, Q., Shi, L., Han, J., Zha, Y., and Zhu, P. (2019). Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. *Field Crops Res.* 235, 142–153. doi: 10.1016/j.fcr.2019.02.022

Conflict of Interest: 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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership