Smart Agricultural Technology 1 (2021) 100006

Contents lists available at ScienceDirect

Smart Agricultural Technology

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

Prediction modeling for yield and water-use efficiency in spinach using remote sensing via an unmanned aerial system

Henry O. Awika[a,1], Jorge Solorzano[a], Uriel Cholula[b,2], Ainong Shi[c], Juan Enciso[a,b], Carlos A. Avila[a,d,*]

[a] Texas A&M AgriLife Research and Extension Center, Weslaco, TX, USA
[b] Department of Agricultural and Biological Engineering, Texas A&M University, College Station, TX, USA
[c] Department of Horticulture, University of Arkansas, Fayetteville, AR, USA
[d] Department of Horticultural Sciences, Texas A&M University, College Station, TX, USA

ARTICLE INFO

Keywords:
Prediction modeling
Remote sensing
Spinach
Unmanned aerial systems
Vegetation indices
Water-use efficiency

ABSTRACT

Predictive modeling (PM) is a useful tool in selecting from a large number of target variables. Given the large quantities of data types generated from unmanned aerial system (UAS) platforms, uncovering the most appropriate candidate UAS data–based phenotype (UASDP) with a strong relationship to target crop traits may be a challenge with the traditional regression models. We hypothesized that by employing machine learning modeling techniques, the contributions of multiple vegetative indices as predictors of biomass and water-use could be defined and ranked. The objective of this study was to apply predictive machine techniques in determining the phase of spinach growth and UASDP variables that best predict yield and water-use efficiency. UASDPs were derived using red-green-blue and multi-spectral sensors mounted on a UAS platform flown weekly on 10 spinach genotypes under well-watered and partial water deficit conditions. Candidate UASDPs including mean and maximum plant height, canopy cover and volume, excess greenness index (ExG), chlorophyll red-edge (ChlRE), normalized difference vegetation index (NDVI) and normalized difference red-edge (NDRE) were generated. UASDPs were used to predict above-ground biomass-based fresh yield, biological dry yield, and field water-use efficiency (WUE_f). This study highlights the use of bootstrap forest partitioning, and partition rank fraction as methods for selecting among many UASDPs as predictors of WUE_f and yield. We also used a weighted geometric mean to integrate various model performance metrics to further refine the UASDPs rankings. UASDPs from mid to late growth stages were better predictors of yield and WUE_f than UASDPs from earlier periods. Canopy volume, NDRE and ExG were the best predictors overall and were useful for distinguishing varieties, particularly in water-stress conditions. We expect that the approaches detailed here will improve reliability of discriminating the importance of UAS-acquired data types thereby improving UAS data-processing efficiency.

1. Introduction

Machine learning based prediction models has greatly improved the handling of multiple factors in defining prediction outcomes [1]. Most traditional regression techniques for crop modeling are sensitive to the influence of outliers and cannot handle data with a large number of predictor combinations [2]. As a result, there is a need for techniques with less strict modeling assumptions and automated methods for identifying informative variables. One area where improved modeling techniques may be useful is the Unmanned Aerial Systems-generated data.

The availability of unmanned aerial (UA) imagery from low-altitude small UA vehicles (UAVs) has facilitated the development of scores of vegetative indices (VI)s and sensor correction methods in recent years, in addition to broadbands refined by hyperspectral filters dedicated to ultra-thin, one- to several-nanometer slices of the electromagnetic spectrum [3–8]. The large variety of indices available poses a methodological challenge for agricultural researchers seeking to select indices useful for predicting traits of interest. Hence there is a need for modeling techniques that can rapidly rank candidate indices according to their predictive performance. We hypothesized that by employing machine

* Correspondence author.
E-mail addresses: henry.awika@uvi.edu (H.O. Awika), Jorge.SolorzanoDiaz@ag.tamu.edu (J. Solorzano), ucholula@nevada.unr.edu (U. Cholula), ashi@uark.edu (A. Shi), Juan.Enciso@ag.tamu.edu (J. Enciso), Carlos.Avila@ag.tamu.edu (C.A. Avila).
[1] Currently: University of the Virgin Islands.
[2] Currently: University of Nevada, Reno, NV.

https://doi.org/10.1016/j.atech.2021.100006
Received 7 June 2021; Received in revised form 25 August 2021; Accepted 29 August 2021

learning modeling techniques, the contributions of multiple VIs as predictors of biomass and water-use could be defined and ranked. An array of more flexible modeling techniques previously used in other disciplines is quickly gaining traction in analyses of agricultural experiments. These techniques use attribute-selection filters for extracting attributes most relevant for vegetation performance [9–11]. Some of the most widely adopted techniques are variations of classification and regression trees (CARTs) [12], which may be generated and combined using a variety of procedures and flavors [13] and often produce excellent results [9,14,15]. CARTs use a sequential partitioning algorithm that recursively selects factors most important for explaining a response variable from a potentially large set of predictors.

The performance of UA system data–based phenotypes (UASDPs) may vary across critical growth phases in a crop. However, it is often difficult to determine this phenological dependency without a temporally continuous set of observations across the entire growth cycle. With small UASs, constraints on equipment, data acquisition, data storage and data processing may make this economically and logistically impossible [16,5,8]. Thus, many researchers have adopted staggered, instantaneous data acquisition flight plans and used multispectral camera types to capture many slices of wave bands at representative interval periods [17,3,18,19]. Such data acquisition plans have improved our understanding of the relationship between crop productivity and crop phenology, biochemical processes, plant–water relations and plant responses to irrigation differences [20,21,5,22].

Vegetation spectra detected by UAS-mounted cameras can be distorted by non-vegetative factors, which in turn can obscure relationships between VIs and plant phenotypes. Most VIs intended for biomass estimation use spectra associated with chlorophyll [23,24,8]. The detection of chlorophyll and most other plant indices is influenced by environmental factors such as atmospheric moisture, illumination and temperature, as well as soil reflectance and soil moisture (S. El-Faki et al., 2000, [25,21,26,7,27]). Such non-vegetative interference can confound the VI selection process and makes automated variable selection approaches all the more important.

In this study we determined the UASDPs based on standing crop phenotypes, including mean height (MeanHt) determined by averaging the heights within all pixels in a quadrant, and maximum plant height (MaxHt) returned within each quadrant. We also determined canopy cover (CCover) and canopy volume (CVol). Vegetation index-based UASDPs were also generated from the visible red-green-blue (RGB) bands, namely excess greenness index (ExG), the near-infrared (NIR) band, namely normalized difference vegetation index (NDVI) and the transition narrow band between the red and NIR spectra and red-edge (RE), namely chlorophyll red-edge (ChlRE) and normalized difference red-edge (NDRE). The UASDPs were used to predict above-ground biomass-based fresh yield (FY), biological yield (BY) as the dry biomass, and field water-use efficiency (WUE_f). These data were collected weekly for 8 weeks from sowing to harvest. The UASDPs and period of growth phases that best predicted FY, BY and WUEf in 10 spinach genotypes grown in well-watered (WW) and partial water deficit (PWD) conditions were evaluated using bootstrap random forest [1] partitioning and rank-fraction [28], as well as weighted geometric means [29]. We expect the approaches detailed here may improve reliability of matching of UAS-acquired signals with green leaf yield traits and WUE_f, thereby improving varietal selection.

2. Materials and methods

2.1. Germplasm, field design and evaluation environment

Four commercial (Bajo, Freja, Sioux and Viroflay) and six Texas A&M AgriLife breeding lines of spinach (Spinacea oleracea) derived from University of Arkansas accessions (08-112, 08-166, 08-290, 08-289, 08-297 and 08-306) were utilized in this study. The 10 varieties were evaluated in sixty 0.743-m^2 plots under two metered irrigation zones of WW and PWD treatments. The plots were arranged in a randomized complete split block design with genotypes grouped in three replicate blocks per water treatment (10 genotypes × 3 replicates × 2 water treatments).

The trial was established in the winter of 2020 in well-cultivated soil at the Texas A&M AgriLife Research and Extension Center at Weslaco, Texas, USA. The experiment was drip-irrigated with the dripline buried at 10 cm. Soil moisture was monitored in the two irrigation zones using Watermark Soil Moisture Sensors, Model 200SS (Irrometer, Riverside CA, USA) installed at a depth of 15.24 cm (6 in). The WW zone was irrigated when the soil moisture was at approximately 20 centibars and the PWD was irrigated when the soil moisture was at 50 centibars. Total water supplied through irrigation per plot for the duration of the experiments was ~31.2 L (WW) and ~18.2 L (PWD). The different irrigation regimens were meant to provide two evaluation 'microenvironments' in close enough proximity to each other to be captured in the same camera overlap from a very-low-flying (20 m) UAV. This eliminated the errors that might have been associated with time lapse when comparing between the two treatments. Hourly rainfall from a weather station located ~200 m NE of the experimental field was tabulated for the period starting 4 d before sowing to 70 days after sowing (DAS) when spinach plots were harvested (December 20, 2019 to March 3, 2020). A total of ~0.023 L of rainfall accumulated per plot during this period. Further details on orientation and location of the experimental field are presented in Fig. 1. Soil type, weather summaries and crop information are presented in Supplementary Materials and Methods.

2.2. UAS, flight plan and high-throughput field data collection

UAS data were acquired on six different days of flight (DOFs) spread across 8 weeks from when seedlings were ~2 cm in height to when plants approached maximum mean height.

UAS data were acquired using a DJI P4 Multispectral platform equipped with a RGB (red, green, and blue) sensor for visible light imaging and five monochrome sensors for multispectral imaging (blue, green, red, red edge, and near-infrared); each sensor has a resolution of 2.08-megapixel. The DJI P4 Multispectral platform also has a spectral sunlight sensor to capture solar irradiance. Both platforms were flown on January 21 and 22, February 11, 19 and 24, and Mach 3, 2020. The DJI Phantom 4 Pro collected 48 images per flight and the DJI multispectral 1219 images per flight.

UAS flight missions were designed using a Pix4D Capture application with a 20-m flying altitude and an 80% overlap for the DJI Phantom 4 Pro and 75% overlap for the DJI multispectral. The timing for the flights were staggered between 11:00 AM to 01:00 PM to target clear cloudless and relatively calm (low wind breeze) intervals to reduce interference form transient clouds and dust. In addition to UAV flights, six permanent and two portable ground control points (GCPs) were surveyed for accurate geo-referencing of images. The coordinates of all GCPs were surveyed using a differential dual-frequency global positioning system manufactured by V-Map (http://v-map.net/). The average accuracy obtained with GCP's was 5.5 mm.

2.3. Image-processing pipeline

RGB images were preprocessed to generate orthomosaic images and digital surface model (DSM) using Agisoft Photoscan Pro software (Agisoft LLC, St. Petersburg, RU). The Multispectral imagery was processed in DJI Terra software to correct for atmospheric effects and geometric variations. The DJI P4 multispectral UAS includes an integrated spectral sunlight sensor that captures solar irradiance on top of the drone. When combined with DJI Terra Software (https://www.dji.com/dji-terra) to generate the multispectral orthomosaic radiometric calibration is performed automatically. The spatial resolutions of the orthomosaic image and DSM on each date were 0.55 cm for the RGB images from the Phantom 4 Pro and 1.05 cm for the multispectral. Spatially georeferenced

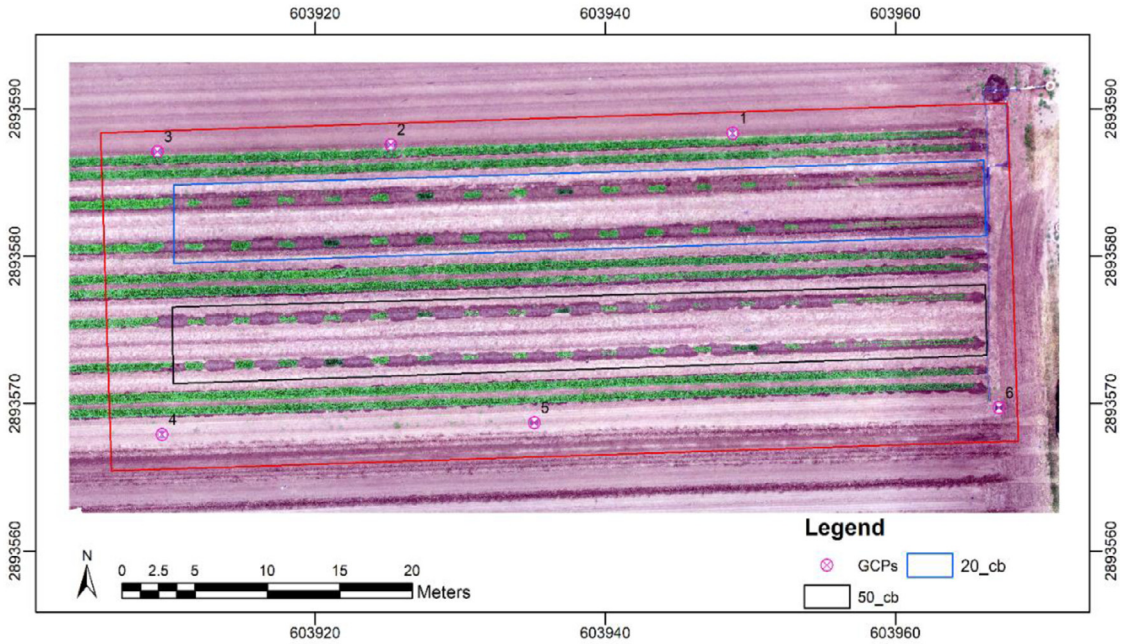

Fig. 1. Experimental spinach field location. Field coordinates are in World Geodetic System 1984 (WGS84) UTM Zone 14N. Blue rectangle, well-watered; grey rectangle, partial water deficit irrigation.

orthomosaics and high-density point clouds were generated using the Structure from Motion (SfM) algorithm [30]. Radiometrically calibrated multispectral orthomosaics were then used to develop VI maps and estimate canopy height, canopy area, canopy volume and spectral VIs. A shapefile of polygons that delimited the plots were created in ArcGIS 10.6.1 (ESRI, Redlands, CA, USA) according to the experimental layout dimensions shown in Fig. 1. Then the boundaries of the polygons were moved inward by 5 cm with the ArcGIS buffer tool to exclude edge effects.

The Canopeo algorithm [31] was used to differentiate canopy pixels from background pixels. Canopy height was calculated by subtracting the initial DSM before planting from the DSM obtained on a flight date.

2.4. Plant growth metrics and indices from RGB and spectral bands

From the DSM height model, plot-wise plant size metrics included minimum canopy height, maximum canopy height (MaxHt), mean canopy height (MeanHt), canopy cover (CCover) and canopy volume (CVol). The MeanHt value was calculated as the mean of all canopy pixel heights per plot. CCover was calculated as the ratio between the number of canopy pixels and the total number of pixels within the plot. CVol was calculated as the accumulated pixel-wise volume (area × DSM height) for each plot.

We generated four VIs from orthomosaics, including NDVI, normalized difference RE (NDRE), chlorophyll RE (ChlRE) and excess greenness index (ExG). These indices, wave bands, their specific derivation and formulae, sensitivities and target utilities are thoroughly reviewed elsewhere [8,17,22,30,32–34]. Average ExG was calculated as described by Woebbecke et al. [34].

2.5. Field-based manual data collection

Three traits were derived from the terminal water use and the terminal yield. The yield was determined per square meter for both fresh and dry above-ground biomass. The fresh (green) biomass component was taken as the material cut at ground level from each plot and weighed immediately in the field. A sample from each plot was also weighed fresh and then reweighed after drying at 60°C until no further change in dry weight was recorded. Each sample dry weight was then extrapolated to represent the whole plot. For each plot, whole-plot above-ground fresh weight and whole-plot above-ground dry weight were denoted as fresh biomass yield (FY) and dry biomass yield or the biological yield (BY), respectively.

Determining FY and BY was important, because spinach is mainly marketed as a leafy green vegetable (hence FY). Yet nutritionally, the tissue matrix containing non-volatile components (dry matter) is more important (hence BY).

Since evapotranspiration was not estimated, we calculated the field water-use efficiency based on above-ground biomass (WUE_f) to account for both irrigation water and rainfall. WUE_f is calculated as the ratio of crop yield to the total amount of water used on a per-plot basis. This value was averaged between replicates for every variety. In this paper, FY, BY and WUE_f are henceforth collectively referred to as terminal traits (TTs).

All phenotypic data used in these analyses are available in Supplementary Table S1.

2.6. Data mining and statistical analyses

We first determined the effects of the spinach entry replicates in a combined regression of the three TTs (WUE_f, BY and FY) under the two water treatments. Genotype, treatment and replicates were treated as random effects in a full factorial analysis using a restricted maximum likelihood (REML) model.

We used the predictor screening platform in JMP backend (JMP, Version 14, SAS Institute Inc., Cary, NC, 1989-2019) to discriminate the relative strengths of UASDPs at each DOF as predictors of FY, BY and WUE_f. We applied the principle of bootstrap forest partitioning to model and evaluate the proportional contribution of each of the eight UASDPs in predicting one of three TTs. Bootstrap forests use decision trees to identify potential predictors by comparing and ranking the contribution of each predictor to reducing residual variance. Since the partition models use multiple predictors simultaneously, screening using this model can identify variables that might be weak alone but strong when used in combination with other predictors [1,13,15]. We identified significant UASDPs by DOF from a composite 48 combinations (8

Table 1
Effects of variety, replicates and treatment and their interactions on WUE and yield: variance components and test for model

	Fresh yield				Dry (Biological) yield				Field WUE			
Random Effect	Var Comp	Std Error	Wald p-Value	Pct of Total	Var Comp	Std Error	Wald p-Value	Pct of Total	Var Comp	Std Error	Wald p-Value	Pct of Total
Variety (G)	0.281	0.411	0.4956	12.0	0.0028	0.005	0.5973	7.61	2.71	5.19	0.6011	6.4
Treatment (T)	0.195	0.333	0.5581	8.32	0.0045	0.007	0.5144	12.2	12.0	17.7	0.4999	28.1
Replication (R)	0.183	0.245	0.4556	7.78	0.0036	0.005	0.4231	10.0	3.56	4.66	0.4441	8.4
G x T	0.369	0.305	0.2265	15.7	0.0035	0.004	0.3937	9.55	4.16	4.07	0.307	9.7
G x R	0.559	0.338	0.0977	23.8	0.0085	0.006	0.1298	23.2	8.26	5.16	0.1091	19.4
T x R	-0.068	0.026	0.0098	0	-0.0013	0.001	0.0042	0	-0.73	0.615	0.2335	0
G x T x R	0.760	0.253	<.0001	32.37	0.0138	0.005	<.0001	37.5	12.0	4.00	<.0001	28.1
Residual	Variance absorbed in G x T x R											
Total	2.357	0.600		100	0.0367	0.010		100	42.7	19.2		100
Summary fit												
Rsquare	0.688				0.3815				0.89			
RSquare Adj	0.687				0.3815				0.65			
Root Mean Square Error	0.872				0.1173				3.46			
Mean of Response	3.958				0.4847				15.2			
Number of plots	60				60				60			
AICc	218.9				-27.53				569.6			
BIC	234.24				-12.29				423.2			

Var Comp, variance component; P of T, percent (proportion) of total variance; Rsquare Adj, adjusted R-square; AIC, Akaike's Information Criteria; BIC, Bayesian Information Criteria

UASDPs × 6 DOFs). The contribution of each UASDP × DOF in the bootstrap forest was ranked according to the proportion variance that each UASDP × DOF combination explained out of a possible 100%. The significance of a predictor contribution in the bootstrap was declared at the 20th rank out of the 48 candidate combinations, which accounted for ~42 percentile strength under each treatment (PWD and WW). See Supplementary Table S2.

To find the UASDPs that best predicted each TT, we compared each of the three TTs against each of the eight UASDPs using the robust rank fraction method. This method fits separate regression models for each predictor and then compares each model against a set of P-value test thresholds. The advantage of the rank fraction method is that it allows for tests of a large set of factors across a large number of responses while also effectively controlling for inflated error variance due to outliers and missing values ([28], Inc, 1989-2019, [13]). For each model, we tested whether UASDP coefficients were significantly different from zero using Student's t-test values. Models were fit with REML.

Due to the large number of simultaneous tests, and to support the multiple inferences that are required, we used the Benjamini–Hochberg false discovery rate (FDR) correction [35], which controls for the overall rate of declaring tests significant. All FDR tests were conducted at a 0.05 threshold. Due to large variances between replicates (Table 1), each replicate was treated as a separate (independent) observation in multiple regression. The FDR LogWorth P-values (–Log_{10}Pvalue) were ranked as a fraction of the number of tests, with smaller P-values ranking higher than larger P-values.

Significant TTs by UASDP regressions were re-ranked using a weighted geometric mean (WGM) [29] approach. The WGM is an analytic hierarchy process (APH) [36], which is useful for combining multiple performance metrics. The WGM was calculated as: $WGM = \Pi(Xi \wedge Wi)^{1/n}$, where Π denotes the product is being computed for each of X(performance metric) for each model i raised to the power wi (the weight of each coefficient represented by corresponding DOF); n is sum of the weights w1, w2…..wn. The larger the WGM corresponding to a UASDP, the higher it ranks in predicting yield and WUE_f.

Performance metrics included model effects, model R squares, trend-slopes, and frequencies of significant models (across DOFs). Model effect (Table 2, and detailed factors in Supplementary Table S3) quantified the variation in each TT explained by a UASDP. The R square, adjusted for multiple tests, represents variation in a TT explained by each model (in the form of y by x correlation). Trend-slope is the vector TT:UASDP that relates the rate of temporal change in UASDP across DOFs to a TT.

We used paired t-tests to compare performance metrics used the WGM ranking between WW and PWD treatments.

3. Results

3.1. Genotype by treatment effect for WUE_f and yield

In the REML models of TTs as a function of treatment, genotype and replicate, main effects were not significant (Table 1). This was likely due to the large standard error of variances related to environmental noise, insufficient water deficit and small sample size. However, interactions between genotype, treatment and replicates (G × T × R) were significant and accounted for the largest proportions of variance of ~0.76 (32% of the total by the random factors), 0.01 (37%) and 12.0 (28%) of FY, BY and WUE_f, respectively. This was because the residuals were absorbed in the G × T × R interaction, lowering the actual standard error. The replicates were therefore treated as 'independent' individuals in determining the relationship between TTs and UASDPs across DOFs.

3.2. Predictors combined for WUE_f and yield show variation due to day of flight and treatment

The bootstrap forest models using all the UASDPs identified the DOFs associated with the most accurate predictions of TTs. For both water treatments, predictions tended to be better using data from late DOFs (days 49, 57, 62 and 70) rather than observations up to day 35 (Fig. 1).

In the PWD treatment, the sum proportions of variance from early DOFs were 0.19 (±0.05), 0.14 (±0.03) and 0.23 (±0.05) for FY, BY and WUE_f, respectively (Supplementary Table S3). By contrast, in the later DOFs, variance proportions were 0.81 (±0.29), 0.86 (±0.31) and 0.77 (±0.25) for FY, BY and WUE_f, respectively. In WW conditions, proportions were 0.20 (±0.05), 0.43 (±0.02) and 0.32 (±0.01) in the earlier DOFs, but 0.80 (±0.29), 0.57 (±0.18) and 0.68 (±0.22) in the later DOFs for FY, BY and WUE_f, respectively.

For a given UASDP, the last two to three DOFs tended to have significantly higher proportions of predicted variance compared to earlier DOFs under PWD. However, under the WW treatment, this pattern was

Table 2
Re-ranking of the significant predictors by weighted geometric mean assessment

	Partial water deficit treatment (PWD)							Well-watered treatment (WW)							Combined	
	[a]Freq	[b]Mean effect	[c]Mean RSq	[d]Mean slope	[e]Mean total RSq	[f]WGM	[g]Rank	[a]Freq	[b]Mean effect	[c]Mean RSq	[d]Mean slope	[e]Mean total RSq	[f]WGM	[g]Rank	[h]Sum WGM	[i]RAT
Fresh above-ground biomass yield																
MaxHt	4	0.609 (0.204 - 0.643)	0.482 (0.054 - 0.535)	28.8 (12.9 - 54.8)	0.318	2.04bc	5	0	NS (0.054 - 0.380)	NS (0.004 - 0.171)	NS (11.5 - 33.3)	0.105	NA	NA	2.04	8
MeanHt	4	0.563 (0.089 - 0.678)	0.422 (0.010 - 0.595)	62.1 (21.6 - 71.0)	0.296	2.45de	3	0	NS (0.140 - 0.462)	NS (0.023 - 0.253)	NS (34.4 - 55.5)	0.105	NA	NA	2.45	7
Cover	6	0.613 (0.506 - 0.698)	0.480 (0.332 - 0.631)	6.43 (4.92 - 13.2)	0.485	1.24a	8	1	0.698 (0.294 - 0.698)	0.576 (0.102 - 0.576)	19.0 (3.55 - 19.0)	0.223	1.97bc	5	3.21	6
CVol	5	0.600 (0.116 - 0.669)	0.478 (0.018 - 0.593)	59.4 (70.8 – 105)	0.358	2.57e	1	1	0.718 (0.017 - 0.718)	0.610 (0.001 - 0.610)	19.9 (11.1 - 58.2)	0.211	2.06bc	4	4.63	2
ChlRE	4	0.593 (0.187 - 0.666)	0.461 (0.045 - 0.576)	22.9 (18.1 - 48.9)	0.347	1.84b	7	4	0.505 (0.007 - 0.627)	0.313 (0.001 - 0.465)	22.0 (0.91 - 24.7)	0.216	1.51a	6	3.36	5
ExG	6	0.602 (0.543 - 0.652)	0.472 (0.382 - 0.550)	36.2 (22.4 - 72.0)	0.432	2.17cd	4	6	0.652 (0.544 - 0.703)	0.507 (0.350 - 0.585)	44.4 (31.4 - 84.7)	0.453	2.45d	1	4.62	3
NDVI	6	0.606 (0.468 - 0.692)	0.486 (0.284 - 0.620)	22.9 (11.9 - 52.3)	0.486	1.89b	6	5	0.628 (0.015 - 0.646)	0.475 (0.001 - 0.494)	29.6 (4.49 - 66.7)	0.446	2.07bc	3	3.96	4
NDRE	4	0.590 (0.157 - 0.675)	0.459 (0.032 - 0.590)	58.4 (45.7 - 88.1)	0.327	2.51de	2	4	0.514 (0.481 - 0.703)	0.326 (0.273 - 0.584)	58.3 (13.3 - 42.9)	0.225	2.14c	2	4.65	1
Dry above-ground biomass yield																
MaxHt	4	0.678 (0.161 - 0.716)	0.512 (0.029 - 0.569)	3.43 (2.06 - 6.24)	0.336	1.05b	6	0	NS (0.028 - 0.362)	NS (0.001 - 0.172)	NS (0.766 - 3.48)	0.094	na	NA	1.06	8
MeanHt	3	0.685 (0.035 - 0.727)	0.523 (0.001 - 0.587)	6.11 (0.680 - 9.28)	0.317	1.30bc	2	0	NS (0.141 - 0.413)	NS (0.026 - 0.223)	NS (0.392 - 6.68)	0.090	na	NA	1.30	7
Cover	6	0.688 (0.590 - 0.781)	0.532 (0.386 - 0.678)	0.906 (0.740 - 2.07)	0.524	0.692a	8	1	0.528 (0.258 - 0.528)	0.365 (0.087 - 0.365)	1.88 (0.370 - 1.88)	0.142	0.713a	4	1.40	6
CVol	5	0.680 (0.161 - 0.778)	0.519 (0.029 - 0.673)	7.78 (1.48 - 14.9)	0.392	1.40c	1	1	0.559 (0.042 - 0.559)	0.410 (0.002 - 0.410)	2.03 (2.03 - 6.22)	0.161	0.775ab	5	2.17	2
ChlRE	4	0.630 (0.299 - 0.714)	0.444 (0.095 - 0.566)	7.01 (2.01 - 8.50)	0.333	1.25bc	3	2	0.468 (0.012 - 0.509)	0.289 (0.001 - 0.340)	2.37 (0.213 - 2.54)	0.153	0.684a	6	1.94	4
ExG	6	0.625 (0.548 - 0.659)	0.435 (0.333 - 0.482)	4.25 (2.36 - 9.23)	0.408	1.06b	5	5	0.545 (0.436 - 0.547)	0.391 (0.249 - 0.432)	5.08 (3.05 - 9.73)	0.367	1.03c	1	2.08	3
NDVI	5	0.674 (0.505 - 0.777)	0.509 (0.283 - 0.670)	3.00 (1.38 - 6.40)	0.471	1.01b	7	3	0.550 (0.381 - 0.562)	0.397 (0.190 - 0.414)	2.31 (1.32 - 4.85)	0.323	0.797ab	3	1.81	5
NDRE	4	0.630 (0.265 - 0.722)	0.444 (0.078 - 0.579)	7.01 (4.84 - 16.2)	0.322	1.25bc	3	2	0.485 (0.004 - 0.527)	0.311 (0.026 - 0.364)	6.51 (0.14 - 6.85)	0.160	0.994c	2	2.24	1
Field water-use efficiency based on above-ground biomass																
MaxHt	4	0.678 (0.161 - 0.716)	0.512 (0.029 - 0.569	133 (84.0 – 254)	0.336	3.59b	5	0	NS (0.028 - 0.362)	NS (0.001 - 0.172)	NS (18.3 - 83.0)	0.094	na	NA	3.59	8
MeanHt	4	0.685 (0.035 - 0.727)	0.523 (0.001 -0.587)	248 (27.7 – 378)	0.317	4.46cd	3	0	NS (0.059 - 0.413)	NS (0.005 - 0.223)	NS (9.34 – 159)	0.090	na	NA	4.46	6
Cover	6	0.688 (0.590 - 0.781)	0.532 (0.386 - 0.678)	36.9 (27.4 - 56.2)	0.524	2.38a	8	1	0.528 (0.271 - 0.370)	0.365 (0.087 - 0.365)	44.8 (8.83 - 44.8)	0.174	2.05a	5	4.43	7
CVol	5	0.680 (0.161 - 0.778)	0.519 (0.029 - 0.673)	361 (60.1 – 586)	0.392	5.03d	1	1	0.559 (0.042 - 0.559)	0.410 (0.002 - 0.410)	48.4 (48.4 – 148)	0.161	2.23ab	4	7.26	2
ChlRE	4	0.635 (0.299 - 0.714)	0.450 (0.099 - 0.566)	149 (81.7 – 346)	0.333	3.50b	7	2	0.468 (0.012 - 0.509)	0.289 (0.001 - 0.340)	56.5 (24.3 - 60.6)	0.153	1.97a	6	5.46	5
ExG	6	0.625 (0.548 - 0.659)	0.435 (0.333 - 0.482)	173 (91.3 – 376)	0.408	3.61b	4	5	0.545 (0.376 - 0.574)	0.391 (0.185 - 0.432)	121 (78.4 - 485)	0.341	2.96d	1	6.57	3
NDVI	3	0.674 (0.265 - 0.722)	0.509 (0.078 - 0.579)	111 (197 – 662)	0.471	3.37b	6	3	0.550 (0.381 - 0.562)	0.397 (0.190 - 0.414)	55.2 (31.6 – 116)	0.323	2.29ab	3	5.66	4
NDRE	2	0.630 (0.505 - 0.777)	0.444 (0.283 - 0.670)	333 (56.2 – 261)	0.322	4.53cd	2	2	0.485 (0.004 - 0.527)	0.311 (0.001 - 0.364)	155 (3.26 – 163)	0.160	2.86cd	2	7.39	1

[a] Number (Frequency) of DOFs when LogWorth was significant DOFs
[b] Mean (and range) effect assigned to a UASDP as a predictor of TT on significant LogWorth DOFs shown under 'Freq'; range based on all six DOFs
[c] Mean (and range) of total (RSq)Variation explained by LogWorth-significant DOFs
[d] Mean slope (and range) of regression fits for significant DOFs
[e] Mean total RSq, variation explained
[f] WGM, weighted geometric mean based on effect, RSq and slope, with the significant DOFs (Frequency) as the weights
[g] Rank based on the weighted geometric means (Rank 1 for largest mean and so on) for a TT, under a treatment
[h] Sum of weighted geometric means across treatments
[i] RAT, rank across treatmentss based on sum of weighted geometric mean (Rank 1 for largest mean)

Table 3
Summary paired t-tests[ψ] comparing water treatments by model performance metrics

	Frequency			Effect			AdjRsq			Slope			WGM		
Terminal trait	PWD	WW	LWth[ϕ]	PWD	WW	LWth	PWD	WW	LWth	PWD	WW	LWth	PWD	WW	LWth
FY	4.88	3.5	1.19	0.60	0.62	0.51	0.47	0.47	0.33	37.1	32.2	0.40	2.09	2.03	0.31
BY	4.63	2.33	2.43**	0.66	0.52	4.14**	0.49	0.36	3.26**	4.92	3.36	0.95	1.13	0.83	1.57*
WUE_f	4.25	2.33	1.48*	0.66	0.52	4.09**	0.49	0.36	3.23**	193.3	80.2	1.53	3.81	2.39	2.18**

[ψ] Based on mean performance metrics of the eight UASDPs in this study
[ϕ] Conversion to $-Log_{10}$(Pvalue). AdjRsq, adjusted R square; BY, dry above-ground biomass yield; FY, fresh above-ground biomass yield; LWth, LogWorth ($-Log_{10}$Pvalue); PWD, partial water deficit; WGM, weighted geometric mean; WUE_f, field water-use efficiency based on above-ground biomass; WW, well-watered. Significant at * P = 0.05; ** P < 0.01

only observed for ChlRE and NDRE in predicting BY and WUE_f but not FY. This was also observed for NDVI in predicting FY (Supplementary Fig. S1).

3.3. Strength of individual UASDPs in predicting WUE_f and yield

Results of the robust rank fraction analysis are shown in Fig. 3.

A total of 144 (6 TTs × 8 UASDPs per DOF × 6 DOFs) combinations were tested for significance in each treatment. In the PWD treatment, models were significant for least one DOF in each TT x UASDP combination (for a total of 111 out of 144 TT × UASDP models) (Table 2). Under the WW treatment there were fewer significant TT × UASDP models (49 out of 144). MeanHt and MaxHt were not significant for any of the DOFs (Supplementary Table S3).

Results of the WGM re-ranking analysis are presented in Table 2. CVol ranked as the strongest predictor for all three TTs under the PWD treatment (Table 2). The second, third and fourth ranked predictors were NDRE, MeanHt and ExG for FY; MeanHt, NDRE and ChlRE (tied), and ExG for BY; and NDRE, MeanHt and ExG for WUE_f. Under the WW treatment, the top three predictors were ExG, NDRE and NDVI for all the traits. CVol was the fourth ranked predictor for FY and WUE_f, while CCover was fourth for BY. MaxHt and MeanHt were not significant in models in the WW treatment.

3.4. UASDPs are more predictive in PWD than in WW (Table 3)

Comparing model performance metrics among water treatments, mean effects and mean adjusted R squares were all significantly larger ($P \leq 0.04$) in PWD than WW for both BY and WUE_f but not FY (Table 3). Mean slope was significantly greater under PWD compared to WW conditions (P = 0.03) only for WUE_f, suggesting that that the aggregate rate of change in factors influencing WUE_f was better represented by the UASDPs. The number (frequency) of DOFs in which the prediction of TTs by UASDPs was significant was also greater under PWD than under WW conditions. This suggests that the UASDPs provided stronger signals under PWD, perhaps due to the likely greater variability between plants or plots. This observation is consistent with the results presented in Fig. 1 and Supplementary Fig. S1.

3.5. Relationship between the effects of genotypes on TTs and UASDPs in later DOFs

The mean spectral values of the most predictive UASDPs at each DOF are plotted against three spinach varieties in Fig. 4. The plots show differential trends between varieties only in the later phases of plant growth. The varieties for which there were statistical differences in UASDPs in the last three DOFs also exhibited corresponding differences for terminal yield and WUE_f (Table 4).

For instance, the top three UASDPs (ExG, CVol and NDRE) exhibited a similar pattern of mean separation between varieties 08-306 and 08-290 and between 08-306 and Viroflay under the PWD treatment. A similar pattern of mean separation is also observed between Viroflay and 08-306 and between Viroflay and 290 under the WW treatment (Fig. 4). The mean UASDP values observed for 08-306 tended to be larger than those of either 08-290 or Viroflay, although this distinction was more subdued for CVol. All other comparisons (e.g., 08-290 vs Viroflay under PWD) were not statistically different. A similar pattern was observed in the mean terminal yield and WUE_f associated with these three varieties (Table 4). These observations suggest that UASDPs in the latter stages of spinach growth may be optimal for predicting yield and WUE_f.

4. Discussion

The goal of this study was to determine the potential of VIs acquired by sensors mounted on a low-flying UAV to predict terminal WUE_f and yield in varieties of spinach in the field.

4.1. WUE_f and yield predictability improves as spinach canopy increases

We evaluated the proportional variance explained by each UASDP at each DOF through predictor screening with a bootstrap forest partitioning model [1,12]. We determined that the composite UASDPs generally became better predictors of yield and WUE_f in the latter three DOFs (Fig. 2), which correspond to the mid to late stages of spinach development [37,38]. Most spectral VIs in the visible broadband and NIR bands are variously affected by multiple non-vegetative environmental factors such as soil [7,21,26], resulting in varying sensitivities. This condition affects their ability to predict primary productivity (reviews [8,39]) on which terminal yield and WUE_f depend. In this study, we did not directly determine such extraneous factors, but they are likely to be less pervasive in late growth stages when spinach canopies were more closed (Supplementary Fig. S2). This suggests that the later flight dates may be better for predicting TTs in fresh leaf spinach (before bolting) using UAS-derived data. Individual differences exist across the DOFs but the general trend across time evident from our analysis supports this interpretation. This result may be important for reducing flight costs by collecting imagery in more informative periods of spinach growth.

4.2. Ranking UASDPs individually may be useful in variable selection

As is evident in Supplementary Fig. S1, greater UASDP predictive importance was associated with later DOFs after decomposing the composite 'model' (described above) into its contributing 'bivariates' (TT × UASDP per DOF). However, the predictive importance of some factors may be masked while others may only show in the presence of another [13,15]. To fully exploit the individual predictive strengths, we implemented a robust rank fraction method ([13,28], Inc, 1989-2019) that automates the modeling of each bivariate individually, producing both the regular linear regression coefficients and a $-Log_{10}$(Pvalue) rank of each regression fit in relation to all others in the batch. Out of a possible 144 (3 TTs × 8 UADP regressions × 6 DOFs) combinations, the robust ranking procedure enabled us to eliminate ~66% and 23% of non-significant combinations in PWD and WW, respectively (Fig. 2 and

Table 4
Means and least square means separation[a] by variety and treatment

	Fresh yield (FY)							Biological yield (BY)					Field water-use efficiency (WUE_f)							
	Least squares mean SE = 0.564; SE Dif = 0.707							Least squares mean SE = 0.067; SE Dif = 0.080					Least squares mean SE = 2.164; SE Dif = 2.590							
08-112,PWD	4.176	A	B	C	D	E	F	0.482	A	B	C	D	19.797	A	B					
08-112,WW	4.713	A	B					0.546	A	B	C		13.029			C	D	E	F	G
08-166,PWD	3.939	A	B	C	D	E	F	0.481	A	B	C	D	19.681	A	B					
08-166,WW	4.234	A	B	C	D	E		0.569	A	B	C		13.647			C	D	E	F	G
08-289,PWD	3.704	A	B	C	D	E	F	0.433	A	B	C	D	17.594	A	B	C				
08-289,WW	4.364	A	B	C	D			0.528	A	B	C	D	12.623				D	E	F	G
08-290,PWD	2.779						F	0.365				D	14.447		B	C	D	E	F	G
08-290,WW	4.849	A	B					0.557	A	B	C		13.478			C	D	E	F	G
08-297,PWD	3.644		B	C	D	E	F	0.427		B	C	D	17.162	A	B	C	D	E		
08-297,WW	4.671	A	B					0.597	A				14.470		B	C	D	E	F	G
08-306,PWD	4.591	A	B	C				0.511	A	B	C	D	20.987	A						
08-306,WW	5.145	A						0.601	A	B			14.455		B	C	D	E	F	G
Banjo,PWD	3.224			C	D	E	F	0.411			C	D	16.527	A	B	C	D	E		
Banjo,WW	4.450	A	B					0.553	A	B	C		13.321			C	D	E	F	G
Freja,PWD	3.741	A	B	C	D	E	F	0.452	A	B	C	D	18.476	A	B	C				
Freja,WW	4.329	A	B	C	D	E		0.519	A	B	C	D	12.362				D	E	F	G
Sioux,PWD	3.491		B	C	D	E	F	0.430	A	B	C	D	17.645	A	B	C	D	E		
Sioux,WW	3.073				D	E	F	0.441	A	B	C	D	10.287						F	G
Viroflay,PWD	2.901					E	F	0.366				D	14.765		B	C	D	E	F	
Viroflay,WW	3.140				D	E	F	0.424	A	B	C	D	9.941							G
Treatment																				
	SE = 0.379; SE Dif = 0.256							SE = 0.046; SE Dif = 0.026					SE = 1.515; SE Dif = 1.044							
PWD	3.673	A						0.439	A				17.626	A						
WW	4.243	A						0.530	A				12.844		B					

[a] For each terminal trait (FY, BY and WUE_f), variety by treatment connected by the same letter are not statistically different. Means separated by Student's t-tests. SE, standard error of mean; SE Dif, standard error of mean difference PWD, partial water deficit; WW, well-watered

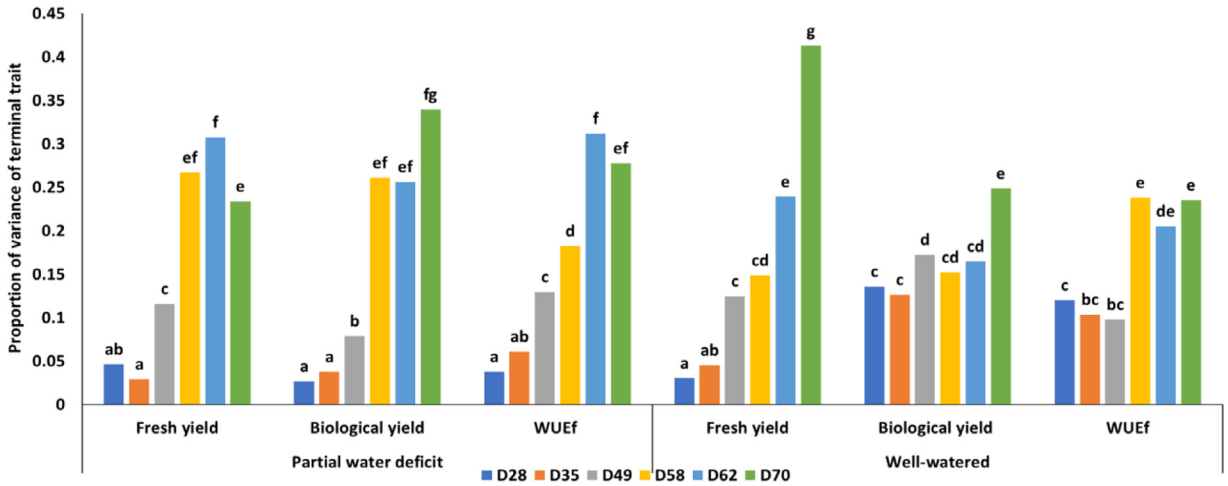

Fig. 2. Predictor proportional ranking by day of flight (D) using all unmanned aerial system data–based phenotypes. Each bar represents the sum proportion of variance predicted. Sum proportions not sharing a letter are statistically different. WUE_f is field water-use efficiency. Fresh yield is above-ground fresh weight determined in the field at harvest. Biological yield is above-ground dry mass.

Table 2 columns 1, 2 & 9; each combination is identified in Supplementary Table S3). We suggest that this technique is useful in situations with a large number of potential indices and/or measurements from multi- and hyperspectral imagery.

4.3. WGM ranking UASDPs with performance metrics improves prediction of terminal yield and water-use efficiency

We have shown that the proportions of variance in terminal traits (FY, BY and WUE_f) explained by the UASDPs were generally higher in the second half of the sampling dates when the canopy was closing in (Fig. 1 and Supplementary Fig. S1; additional information in Supplementary Table S2). Using the principle of WGM, we reranked the UASDPs.

Growth-associated parameters, including lateral growth (CCover) and vertical growth (MaxHt), individually were either the least significant (lowest ranked, CCover) or the most inconsistent predictors of TTs across treatments (Table 3). CCover relies on a two-dimensional perspective [31] that does not account for tiered overlaps of leaves, while MaxHt overlooks undergrowth, which is important in defining harvestable spinach total volume. It is thus not surprising that CVol, which incorporates both lateral and vertical growth, was ranked highest among the UASDPs overall.

As RE light is more permeable through leaves, it has better canopy penetrance than red light [8,40]. RE-based indices like NDRE are better indicators of vegetation health/vigor than NDVI for mid- to late-season crops that have accumulated high levels of chlorophyll in their leaves [40]. NDVI is more sensitive to soil interference and often loses sensitivity after plants accumulate a critical level of leaf cover or chlorophyll. This is because NDVI is a function of the red spectra and chlorophyll has a maximum absorption in the red waveband, which does not penetrate very far past a few leaf layers [8,39,40]. We also speculate that the deeper canopy penetrance of RE and its derived indices (e.g., NDRE) are better at predicting yield and WUE_f. In this study, the WGM re-ranking found NDRE and CVol to be the two best overall UASDPs for predicting TTs. NDRE was superior to the widely used NDVI and to indices relying on the visible spectra encompassing blue light, red light, or a mixture of visible and NIR light. The green chromatic hue-based ExG index, even though it is based on RGB, is well suited for both nonshaded and shaded sunlit conditions [34]. We speculate that the ability to detect signals from the shaded parts of the canopy made ExG a stronger predictor of TTs in both treatments. However, the penetrance of RE seems to be relatively more important under our experimental conditions.

4.4. Consistency between random forest and WGM rank fraction rankings

UASDP × DOF model performance metrics were generally better in PWD than WW conditions. Since performance metrics were calculated for each UASDP separately, we expected that the patterns observed would generally mirror the trends observed in the random forest ranking (Fig. 1), even though these two methods were completely different. We found evidence that rankings from these methods were similar for BY and WUE_f. The differences between WW and PWD were statistically significant for these TTs in the random forest analysis for most DOFs (Fig. 1 and Supplementary Table S3, compared to Table 3). These observations also mirrored the paired t-test for the WGM (insignificant LogWorth 0.31 [$\sim P = 0.49$] for FY, significant LogWorth 1.58 [$\sim P = 0.03$] for BY and significant LogWorth 2.18, [$\sim P < 0.01$ for WUE_f]; Table 3). These results suggest that the two parallel methods used in this study have the potential for enhancing analysis of multiple UAS-acquired plant data. To our knowledge, this may be among the first reports using WGM to improve the selection of prediction factors in plant studies. Further studies in this area are needed.

4.5. Temporal trends in UASDPs parallel separation of TTs associated with spinach varieties

We examined how variation in UASDPs across DOFs was associated with WUE_f and yield. The top three ranked UASDPs (i.e., NDRE, CVol and ExG, were used to detect differences in TTs between treatments and between varieties (Table 4). We evaluated whether varieties with significant mean separation also had detectable differences in UASDP time sequence signals.

Varieties that had differences in TTs also had corresponding differences in UASDP values in latter DOFs. Statistical differences in UASDPs were detectable between varieties only in later growth stages (Fig. 3 and Table 4). UASDP differences corresponded to statistically significant differences in WUE_f and yield between some varieties. This suggests remotely sensed plant growth and health indices can be used to predict yield and water-use efficiency, as well as to discriminate between varieties, particularly in water-limited situations. We suggest that predictive accuracy is dependent on the choice of the correct UASDP, the DOF (plant growth phase), and the reflectance spectral bands applied. We demonstrated that variable selection could be improved by integrating and fine-tuning different modeling techniques. We expect the approaches detailed here may improve reliability of matching of UAS-acquired signals with green leaf yield traits and WUE_f, thereby improving varietal selection.

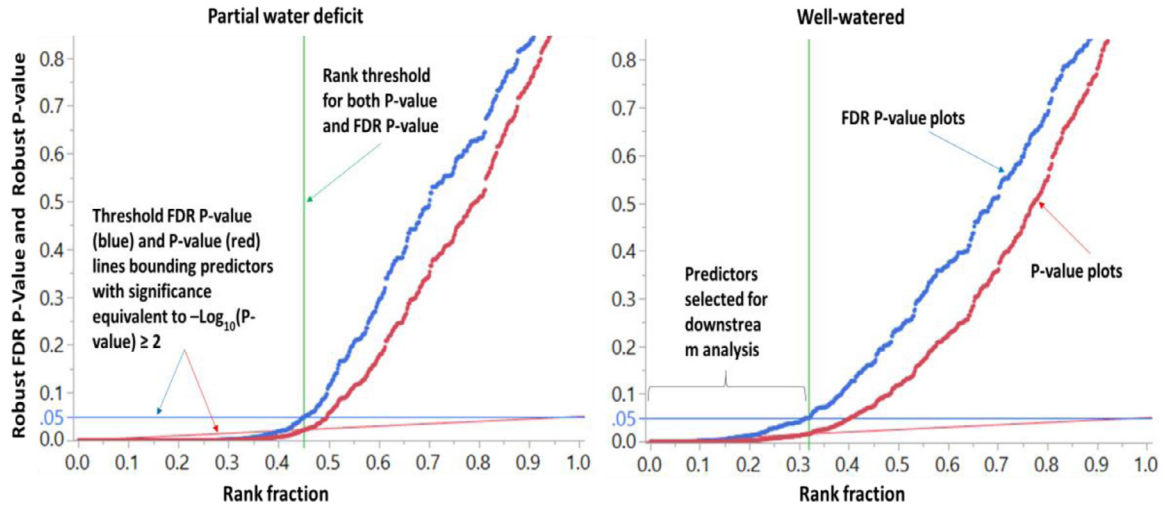

Fig. 3. Discriminating predictor ranks by P-value and false discovery rate (FDR) thresholds. FDR P-values and P-values are plotted against the rank fraction (RF), which ranks the P-values for each unmanned aerial system–derived trait. P-values are from regressions against terminal yield and water-use traits.

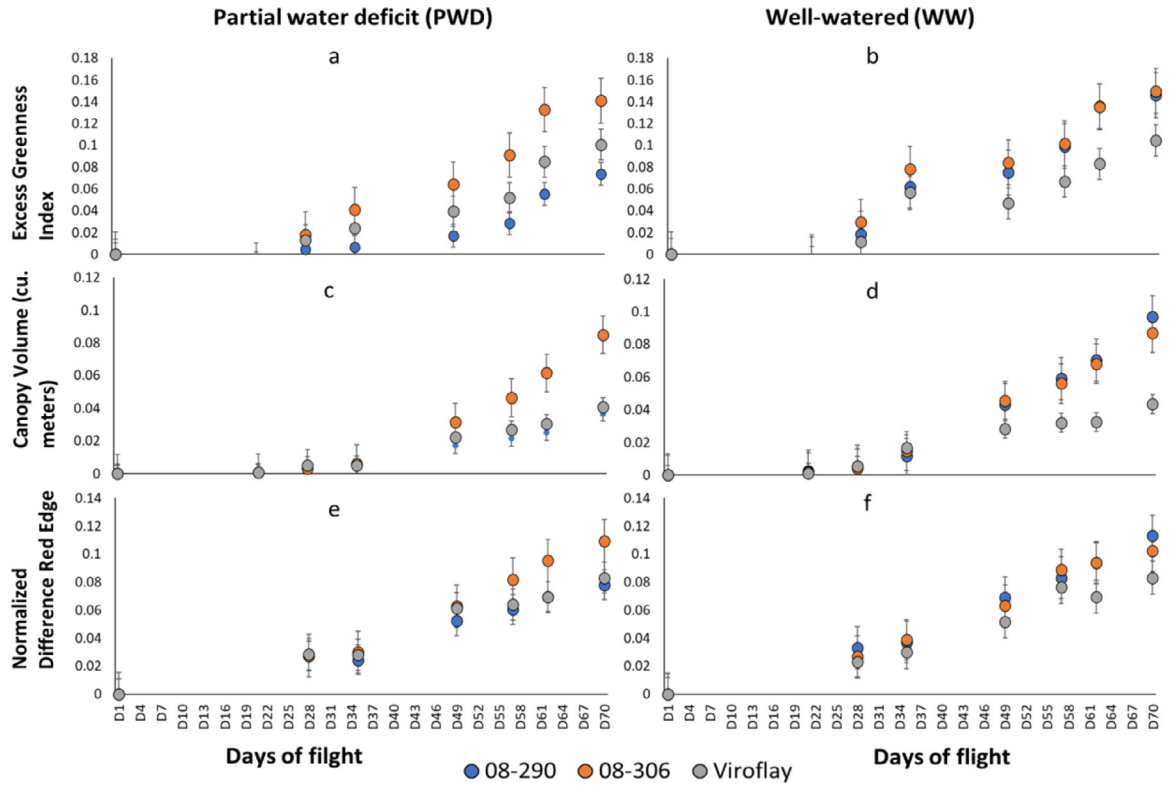

Fig. 4. Trends in top-ranked three unmanned aerial system data–based phenotypes (UASDPs) across time for three spinach varieties. UASDPs are on the vertical axis, and UAS days of flight (D) are on the horizontal axis. Each marker point represents the average UASDP value for three plots (each containing 28 plants) within a water treatment (partial water deficit [PWD] or well-watered [WW]). Varieties having statistically different least square means for WUE_f and yield are shown (PWD, 08-306 vs 08-290 and 08-306 vs Viroflay; WW, Viroflay vs 08-306 and Viroflay vs 290; any other combination [e.g., 08-290 vs Viroflay under PWD] had no statistical difference).

5. Conclusions

We determined the phases of spinach growth and UASDP variables most useful for predicting yield and water-use efficiency. UASDPs were derived from RGB and multispectral sensors mounted on a UAS platform. Composite UASDPs from later growth stages were better predictors of yield and WUE_f. Several UASDP indices were statistically different between varieties and treatments in the late phases of spinach growth. These statistical differences in UASDPs corresponded to the measured differences in terminal yield and water-use efficiencies between the varieties. The top three VIs and plant growth parameters across treatments were NDRE, ExG and CVol.

In this study, prediction models were dependent on the choice of the correct UASDP, the DOF (plant growth phase), and the reflectance spectral bands. This study also highlights the potential for integrating multi-factor modeling and the principle of geometric mean for prioritizing TT–predictor relationships. Applying different modeling techniques can enhance decision-making regarding the choice

of appropriate VI for predicting traits in spinach and other leafy greens.

Supplementary Materials: Fig. S1 Unmanned aerial system data (UASD) by day of flight (DOF) importance scores from the random forest partitioning models., Fig. S2 Time sequence images obtained from UAV-mounted sensors, Supplementary Materials and Methods, Table S1 Phenotypic data, Table S2 Predictor screening of the unmanned aerial system–derived phenotypes, Table S3 Robust rank fraction evaluation of terminal water-use efficiency and yield against unmanned aerial system–derived predictors

Declaration of Competing Interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Henry O. Awika: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Jorge Solorzano:** Data curation, Writing – original draft. **Uriel Cholula:** Data curation, Writing – original draft. **Ainong Shi:** Writing – original draft. **Juan Enciso:** Conceptualization, Writing – original draft. **Carlos A. Avila:** Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing, Funding acquisition.

Acknowledgments

We thank Ayrton Laredo, Alexandra Hernandez, Samantha Serna, and Leticia Rodriguez at the Texas A&M AgriLife Research and Extension Center-Weslaco for assistance with field work, irrigation, fertilization, harvesting, and yield data collection. Also, we thank Dr. Juan Landivar for technical support and data analysis advice.

Funding

This study was supported in part by funds from USDA-National Institute of Food and Agriculture Specialty Crops Research Initiative 2017-51181-26830 assigned to C.A. and A.S., and Texas A&M AgriLife Research Seed Grants to C.A. This research was also partially funded by USDA-NIFA, grant number 2017-68007-26318 “Diversifying the Water Portfolio for Agriculture in the Rio Grande Basin”, the Texas Water Development Grant “Agricultural Water Management Strategies and Education Programs for Water Conservation in South Texas”, and the Hatch projects: “Combining Molecular and High-throughput Phenotyping Breeding Approaches to Improve Yield, Quality and Disease Resistance in Tomato and Spinach”, grant number 1026646 To C.A.A. and “Development of Engineering Tools for Soil and Water Conservation”, grant number 1016863 to J.E.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.atech.2021.100006.

References

[1] L. BREIMAN, Random forests, Mach. Learn. 45 (2001) 5–32.
[2] D. MCCAULEY, Predictive Agriculture: Crop Modeling for the Future, CSA News 65 (2020) 3–9.
[3] M. GHAZAL, Y. ALKHALIL, H. HAJJDIAB, UAV-based remote sensing for vegetation cover estimation using NDVI imagery and level sets method, in: 2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2015, pp. 332–337.
[4] A.A. GITELSON, Y.J. KAUFMAN, R. STARK, D. RUNDQUIST, Novel algorithms for remote estimation of vegetation fraction, Remote Sens. Environ. 80 (2002) 76–87.
[5] P.J. HARDIN, R.R JENSEN, Small-Scale Unmanned Aerial Vehicles in Environmental Remote Sensing: Challenges and Opportunities, GIScience Remote Sens. 48 (2011) 99–111.
[6] J. KELCEY, A. LUCIEER, Sensor Correction of a 6-Band Multispectral Imaging Sensor for UAV Remote Sensing, Remote Sens. 4 (2012) 1462–1493.
[7] S.H.A. MOHAMMAD HOSSEIN SHAHROKHNIA, 2 - Remotely Sensed Spatial and Temporal Variations of Vegetation Indices Subjected to Rainfall Amount and Distribution Properties, in: C.G. HAMID REZA POURGHASEMI (Ed.), Spatial Modeling in GIS and R for Earth and Environmental Sciences, Elsevier, 2019.
[8] J. XUE, B. SU, Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications, J. Sens. (2017) 1353691 2017.
[9] W. AERTSEN, V. KINT, J. VAN ORSHOVEN, K. ÖZKAN, B. MUYS, Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests, Ecol. Modell. 221 (2010) 1119–1130.
[10] Marinković, B., Crnobarac, J., Brdar, S., Antić, B., Jacimovic, G., Crnojević, V. 2009. Data Mining Approach for Predictive Modeling of Agricultural Yield Data.
[11] N. TILLY, H. AASEN, G. BARETH, Fusion of Plant Height and Vegetation Indices for the Estimation of Barley Biomass, Remote Sens. 7 (2015) 11449–11480.
[12] L. BREIMAN, J. FRIEDMAN, C.J. STONE, R.A OLSHEN, Classification and regression trees, CRC press, 1984.
[13] C. STROBL, J. MALLEY, G. TUTZ, An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests, Psychol. Methods 14 (2009) 323–348.
[14] J. HOLLOWAY, K. MENGERSEN, Statistical Machine Learning Methods and Remote Sensing for Sustainable Development Goals: A Review, Remote Sens. 10 (2018) 1365.
[15] D.S. SIROKY, Navigating Random Forests and related advances in algorithmic modeling, Stat. Surv. 3 (2009) 147–163.
[16] J. FAN, F. HAN, H. LIU, Challenges of big data analysis, Natl. Sci. Rev. 1 (2014) 293–314.
[17] H. AASEN, E. HONKAVAARA, A. LUCIEER, P. ZARCO-TEJADA, Quantitative Remote Sensing at Ultra-High Resolution with UAV Spectroscopy: A Review of Sensor Technology, Measurement Procedures, and Data Correction Workflows, Remote Sens. (2018) 10.
[18] A.B. POTGIETER, B. GEORGE-JAEGGLI, S.C. CHAPMAN, K. LAWS, L.A. SUÁREZ CADAVID, J. WIXTED, J. WATSON, M. ELDRIDGE, D.R. JORDAN, G.L HAMMER, Multi-Spectral Imaging from an Unmanned Aerial Vehicle Enables the Assessment of Seasonal Leaf Area Dynamics of Sorghum Breeding Lines, Front. Plant Sci. 8 (2017).
[19] K. ZHANG, X. GE, P. SHEN, W. LI, X. LIU, Q. CAO, Y. ZHU, W. CAO, Y. TIAN, Predicting Rice Grain Yield Based on Dynamic Changes in Vegetation Indexes during Early to Mid-Growth Stages, Remote Sens. 11 (2019) 387.
[20] R. DAROYA, M RAMOS, NDVI image extraction of an agricultural land using an autonomous quadcopter with a filter-modified camera. 2017 7th IEEE International Conference on Control System, Comput. Eng. (ICCSCE) (2017) 110–114 24-26 Nov2017.
[21] M.F. GARBULSKY, J. PEÑUELAS, J. GAMON, Y. INOUE, I FILELLA, The photochemical reflectance index (PRI) and the remote sensing of leaf, canopy and ecosystem radiation use efficiencies: A review and meta-analysis, Remote Sens. Environ. 115 (2011) 281–297.
[22] J. JORGE, M. VALLBÉ, J.A SOLER, Detection of irrigation inhomogeneities in an olive grove using the NDRE vegetation index obtained from UAV images, Eur. J. Remote Sens. 52 (2019) 169–177.
[23] G. HAILEMICHAEL, A. CATALINA, M.R. GONZALEZ, P. MARTIN, Relationships between water status, leaf chlorophyll content and photosynthetic performance in Tempranillo vineyards, S. Afr. J. Enol. Vitic. 37 (2016) 149–156.
[24] J.L. HATFIELD, A.A. GITELSON, J.S. SCHEPERS, C.L WALTHALL, Application of Spectral Remote Sensing for Agronomic Decisions, Agron. J. 100 (2008) S-117-S-131.
[25] R. DARVISHZADEH, A. SKIDMORE, C. ATZBERGER, S. VAN WIEREN, Estimation of vegetation LAI from hyperspectral reflectance data: Effects of soil type and plant architecture, Int. J. Appl. Earth Obs. Geoinf. 10 (2008) 358–373.
[26] T.J. MALTHUS, B. ANDRIEU, F.M. DANSON, K.W. JAGGARD, M.D STEVEN, Candidate high spectral resolution infrared indices for crop cover, Remote Sens. Environ. 46 (1993) 204–212.
[27] J. TORRES-SÁNCHEZ, F. LÓPEZ-GRANADOS, I.A. DE CASTRO, J.M PEÑA-BARRAGÁN, Configuration and Specifications of an Unmanned Aerial Vehicle (UAV) for Early Site Specific Weed Management, PLoS One 8 (2013) e58210.
[28] P. ROYSTON, W. SAUERBREI, Multivariable Model - Building: A Pragmatic Approach to Regression Anaylsis based on Fractional Polynomials for Modelling Continuous Variables, Wiley, 2008.
[29] J. KREJČÍ, J STOKLASA, Aggregation in the analytic hierarchy process: Why weighted geometric mean should be used instead of weighted arithmetic mean, Expert Syst. Appl. 114 (2018) 97–106.
[30] M.J. WESTOBY, J. BRASINGTON, N.F. GLASSER, M.J. HAMBREY, J.M REYNOLDS, Structure-from-Motion’ photogrammetry: A low-cost, effective tool for geoscience applications, Geomorphology 179 (2012) 300–314.
[31] A. PATRIGNANI, T.E OCHSNER, Canopeo: A powerful new tool for measuring fractional green canopy cover, Agron. J. 107 (2015) 2312–2320.
[32] J.U.H. EITEL, L.A. VIERLING, M.E. LITVAK, D.S. LONG, U. SCHULTHESS, A.A. AGER, D.J. KROFCHECK, L. STOSCHECK, Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland, Remote Sens. Environ. 115 (2011) 3640–3646.
[33] Y. FURUKAWA, J. PONCE, Accurate, dense, and robust multiview stereopsis, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2009) 1362–1376.
[34] D. WOEBBECKE, G. MEYER, K. VON BARGEN, D MORTENSEN, Color indices for weed identification under various soil, residue, and lighting conditions, Trans. Am. Soc. Agric. Eng. 38 (1995) 259–269.

[35] Y. BENJAMINI, Y. HOCHBERG, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, J. R. Stat. Soc.. Series B (Methodological) 57 (1995) 289–300.
[36] T.L. SAATY, Decision making with the analytic hierarchy process, Int. J. Serv. Sci. 1 (2008) 83–98.
[37] H.O. AWIKA, R. BEDRE, J. YEOM, T.G. MARCONI, J. ENCISO, K.K. MANDADI, J. JUNG, C.A AVILA, Developing Growth-Associated Molecular Markers Via High--Throughput Phenotyping in Spinach, Plant Genome 12 (2019) 190027.
[38] E.B.J. SMOLDERS, R MERCKX, Growth analysis of soil-grown spinach plants at different N-regimes. Optimization of Plant Nutrition, Springer, Dordrecht, 1993.
[39] S. HUANG, L. TANG, J.P. HUPY, Y. WANG, G. SHAO, A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing, J. Forestry Res. (2020).
[40] B.H.H. BOIARSKII, Comparison of NDVI and NDRE Indices to Detect Differences in Vegetation and Chlorophyll Content, J. Mech. Cont. Math. Sci., Special Issue 20 (2019) 20–29.