



## Article

# Temporal Vegetation Indices and Plant Height from Remotely Sensed Imagery Can Predict Grain Yield and Flowering Time Breeding Value in Maize via Machine Learning Regression

Alper Adak <sup>1</sup>, Seth C Murray <sup>1,\*</sup>, Sofija Božinović <sup>2</sup>, Regan Lindsey <sup>1</sup>, Shakirah Nakasagga <sup>1</sup>, Sumantra Chatterjee <sup>1</sup>, Steven L. Anderson II <sup>3</sup> and Scott Wilde <sup>1</sup>

<sup>1</sup> Department of Soil and Crop Sciences, Texas A&M University, College Station, TX 77843, USA; alperadak@tamu.edu (A.A.); reganlindsey@tamu.edu (R.L.); sn1@tamu.edu (S.N.); sumantra.chatterjee@tamu.edu (S.C.); scwilde@tamu.edu (S.W.)

<sup>2</sup> Maize Research Institute, Zemun Polje, 11185 Belgrade, Serbia; sbozinovic@mrizp.rs

<sup>3</sup> Department of Environmental Horticulture, Institute of Food and Agriculture Science, Mid-Florida Research and Education Center, University of Florida, Apopka, FL 32703, USA; s.anderson1@ufl.edu

\* Correspondence: sethmurray@tamu.edu



**Citation:** Adak, A.; Murray, S.C.; Božinović, S.; Lindsey, R.; Nakasagga, S.; Chatterjee, S.; Anderson, S.L., II; Wilde, S. Temporal Vegetation Indices and Plant Height from Remotely Sensed Imagery Can Predict Grain Yield and Flowering Time Breeding Value in Maize via Machine Learning Regression. *Remote Sens.* **2021**, *13*, 2141. <https://doi.org/10.3390/rs13112141>

Academic Editor: David M Johnson

Received: 14 April 2021

Accepted: 25 May 2021

Published: 29 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Unoccupied aerial system (UAS; i.e., drone equipped with sensors) field-based high-throughput phenotyping (HTP) platforms are used to collect high quality images of plant nurseries to screen genetic materials (e.g., hybrids and inbreds) throughout plant growth at relatively low cost. In this study, a set of 100 advanced breeding maize (*Zea mays* L.) hybrids were planted at optimal (OHOT trial) and delayed planting dates (DHOT trial). Twelve UAS surveys were conducted over the trials throughout the growing season. Fifteen vegetative indices (VIs) and the 99<sup>th</sup> percentile canopy height measurement (CHMs) were extracted from processed UAS imagery (orthomosaics and point clouds) which were used to predict plot-level grain yield, days to anthesis (DTA), and silking (DTS). A novel statistical approach utilizing a nested design was fit to predict temporal best linear unbiased predictors (TBLUP) for the combined temporal UAS data. Our results demonstrated machine learning-based regressions (ridge, lasso, and elastic net) had from 4- to 9-fold increases in the prediction accuracies and from 13- to 73-fold reductions in root mean squared error (RMSE) compared to classical linear regression in prediction of grain yield or flowering time. Ridge regression performed best in predicting grain yield (prediction accuracy = ~0.6), while lasso and elastic net regressions performed best in predicting DTA and DTS (prediction accuracy = ~0.8) consistently in both trials. We demonstrated that predictor variable importance descended towards the terminal stages of growth, signifying the importance of phenotype collection beyond classical terminal growth stages. This study is among the first to demonstrate an ability to predict yield in elite hybrid maize breeding trials using temporal UAS image-based phenotypes and supports the potential benefit of phenomic selection approaches in estimating breeding values before harvest.

**Keywords:** high throughput phenotyping; unoccupied aerial system; temporal vegetation indices; nested design; machine learning regression; phenomic prediction and selection

## 1. Introduction

Estimating breeding values for genotypes (pedigrees) before harvest would be useful for speeding the breeding cycle while reducing combine measurement resources. Early season breeding values can be enabled by implementing temporal field-based high throughput phenotyping (HTP) [1]. Unoccupied aerial systems (UAS; i.e., drones with sensor payloads) can objectively capture the temporal variation of complex traits in crops with limited resources; providing a new tool to dissect complex traits in a time-series manner. UAS provides higher quality raw images with better resolutions than traditional remote sensing platforms such as Landsat-derived images [2]. UAS raw images are processed (e.g., ortho-mosaic and point cloud densification) to generate geographically corrected mosaics over

each time point throughout the growing period [2]. To increase precision of data extracted from the processed images, novel plot-based data extraction pipelines have been developed for breeding nurseries [3,4]. When successfully performed, multiple plot-based temporal phenotypes can be generated that include various vegetation indices (VIs) [5–14] as well as canopy height measurements (CHMs) [15–19]. Temporal VIs and CHMs have been used for downstream analysis in prediction models [15–17,20–26] demonstrating improvement in prediction modeling when utilizing UAS image-based phenotypes over conventional manual measurements. Temporal HTP platforms are enabling researchers to improve the understanding of environmental and growth stage specific interactions.

Genotype-by-environment (GxE) interaction is a special case in plant breeding used to dissect the phenotypic plasticity and reaction norm of complex traits occurring among discriminate environments [27–29]. GxE describes how differences in measured appearances between different genotypes are a result of interactions of each with their environments creating plasticity, which is nearly impossible to predict. However, the examination of temporal phenotypic plasticity (TPP) for complex traits occurring beyond a single growth stage is typically unfeasible due to labor, time, and resource demands. Dissecting the TPP of complex traits at differential growth stages has lagged behind commonly collected terminal phenotypes (e.g., grain yield) resulting in understanding GxE interactions based on the accumulation of the entire growth season. As a result, it is difficult to determine if specific time points caused the significant GxE variation at the end of the season. HTP platforms, such as UAS, enable repeated phenotypic data collection, allowing evaluation of phenotypes at various time points throughout plant lifecycles with large numbers of genotypes at low cost [30]. Recently, UAS have demonstrated that quantitative trait loci (QTL) have varying effect sizes for the same trait (e.g., CHM) at different time points throughout growth periods of maize [16,17,31]. Similarly, both unique and common loci belonging to different time points have been discovered to have associations with vegetative indices and canopy height measurements in maize throughout growth [17,32]. The uniqueness of temporal QTL has also been shown in cotton [33], wheat [34], and sorghum [35]. These findings underline new opportunities in dissecting complex traits using remotely sensed phenotypes belonging to multiple time points instead of one or a few terminal time points, as has been done traditionally.

#### Multicollinearity Challenges in Temporal Predictions

Multiple time point-derived temporal phenotypes have previously improved yield prediction accuracies beyond single time point-derived phenotypes, but come with new challenges [15,20,31]. For instance, when using phenotypic data from many VIs derived across time points as predictors, multicollinearity problems are often revealed in multiple linear regression analysis. Multicollinearity leads to the variance inflation factors (VIFs) among predictors that causes inflation of the variation of the estimated regression coefficients in the prediction model [36]. When multiple time point derived temporal phenotypes are used as predictors, the coefficients of the predictors need to be penalized to increase the prediction accuracy. The reasoning being: (i) the predictors included in a model may have relatively similar and large effects on the predicted variable, (ii) a small number of predictors may have significant and large effects on the predicted variables while others have smaller effects, or (iii) any conditions involving these two possibilities. The three conditions above are underlying phenomena that need to be controlled to accomplish better prediction accuracy in the prediction of complex quantitative terminal traits such as yield and flowering times.

This study aimed to (i) predict the temporal breeding values of maize hybrids using fifteen VIs and CHMs derived from twelve UAS flights over the two trials; (ii) compare performances of linear regression and machine learning based regressions (ridge, lasso, and elastic net) in predicting grain yield and flowering times (days to anthesis; DTA and days to silking; DTS) from temporal breeding values of fifteen VIs and CHMs; and (iii) calculate the temporal variable importance scores of the predictors in the best performing

(i.e., least error with highest prediction accuracy) machine learning regression for yield and flowering times. Understanding these aims will help to better understand how to maximize and improve the results of phenomics-assisted breeding through remote sensing and identify new fundamentally important biological processes for further investigation in plant breeding and quantitative genetic researches.

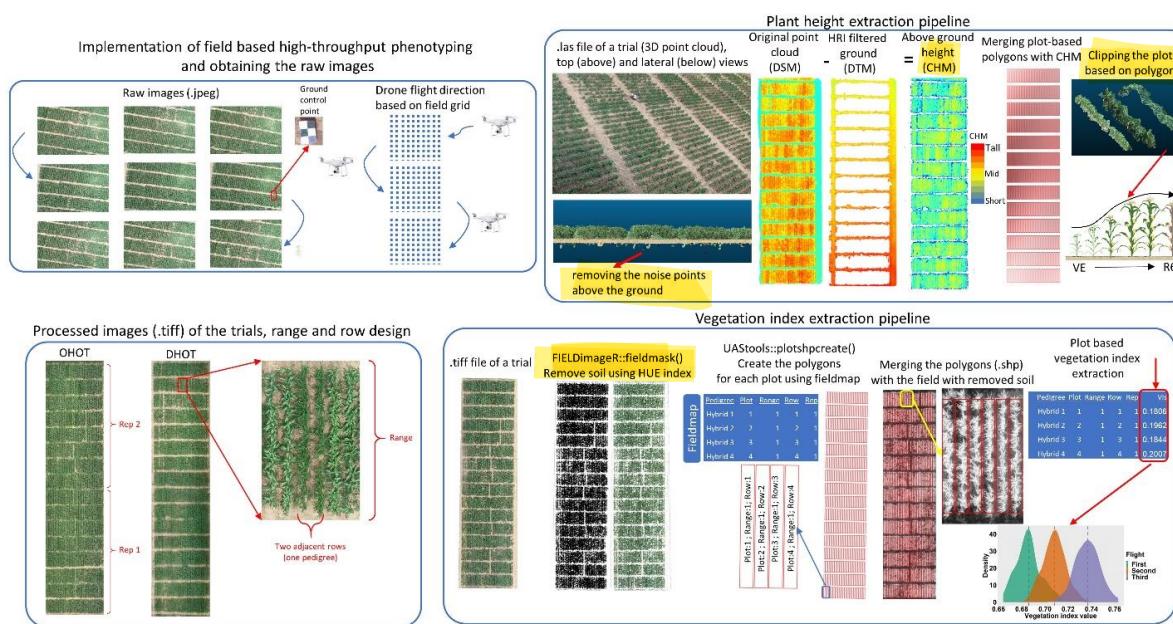
## 2. Materials and Method

### 2.1. Experimental Conditions

The two trials, containing same 100 advanced maize breeding hybrids, were planted in College Station, Texas ( $30^{\circ}32'46.3''$  N  $96^{\circ}26'00.2''$  W) on 21 March 2019 (OHOT trial; irrigated) and 12 April 2019 (DHOT trial; delayed planting) respectively. Each population was planted in a separate randomized completed block design with two replications. Experimental plots consisted of two-row adjacent plantings of the same genotype (0.76 m row spacing, 7 m plot length). Each randomized complete block trial consisted of 13 ranges (from front to back of field, perpendicular to the tractor rows), where each range contained 16 hybrids across tractor rows represented by two-row plots, 32 single rows in total. The last range in each trial had eight two-row plots of commercial check fill to form a rectangle, and commercial checks were planted around the trial as a border.

### 2.2. Field-Based High Throughput Phenotyping

A DJI Phantom 4 Pro V2.0 (DJI, Shenzhen, China) was flown 25 m above the ground to capturing RGB images at 72 DPI resolution from the standard integrated camera. The raw images were taken using  $5472 \times 3648$  pixels with 90 percent forward and sideward overlap. These images were processed using Agisoft Metashape V15.2 software (Agisoft LLC, Russia) to generate an orthomosaic and point cloud for each flight date (Figure 1). Ground control points (GCPs) were used to aid in data processing during the flights and mitigate aberrations in the resulting georeferenced mosaics. In total, 45 GCPs were used during the flights. A total of 25 UAS surveys were conducted throughout the growing season. Following qualitative QC/QA (Quality Control/Quality Assurance) of the orthomosaics and point clouds, twelve flights were identified to be free of artifacts and blunders (Table 1).



**Figure 1.** Shows the steps of high-throughput phenotyping pipeline including data collection, processing, and extraction from the RGB images.

**Table 1.** Summary table of UAS flights dates and corresponding days after planting (DAP) times for each trial. The optimal planting date trial (OHOT) and delayed planting date trial (DHOT) were sown on 21 March 2019 and 12 April 2019, respectively.

Flight Time (Month)	April				May				June			July	
Flight Time (Day)	20th	22nd	29th	6th	10th	23rd	4th	17th	25th	11th	16th	26th	
Days after planting (DHOT trial)	8	10	17	24	28	41	53	66	74	90	95	105	
Days after planting (OHOT trial)	30	32	39	46	50	63	75	88	96	112	117	127	

The UAS was flown in moderate weather at midday during periods of clear skies and no wind, whenever possible. Consistent time periods were attempted between flights but to obtain more data not all flights could be made in ideal conditions. As the season progressed, issues due to windy conditions became more prevalent in the processing stream with plants swaying further when affected by the wind. Cloud coverage was addressed by a brightness correction measurement from an onboard solar radiance detector or by using images of calibrated reflectance targets in situ prior to and after flights. Large overlaps of images reduced water reflections. Importantly, because varieties were primarily being compared within flights, errors from many of these potential sources of variation are nested within each flight and rendered irrelevant compared to if comparisons are made between flights.

### 2.3. Extracting Temporal Traits from RGB Images and 3D Point Clouds

Fifteen VIs and temporal plant heights were extracted from the orthomosaic images (.tif file extensions) and three-dimensional (3D) point clouds (.las or .laz file extensions), respectively. First, ESRI (Environmental Systems Research Institute, Inc.) shape files (.shp file extensions) were created using *R/UAStools::plotshpcreate* function [3] in R to create the polygons for each consecutive row plot (*nrowplot* = 2, *multirowwind* = T) for each population based on their respective field maps/experimental layouts. The buffer arguments were set to “*rowbuf* = 0.61” and “*rangebuf* = 3.1” to reduce adjacent plot overlap and uninformative data within the walking alleys data extractions. Two fit ESRI shape files were generated separately to extract plot information from each one of the consecutive two row plots in the images and point clouds. These shape files were then used in extracting VIs and plant heights for each time point (Figure 1).

To extract the plot-based VIs the *FIELDImageR* package [4] was used. Before extraction of VIs, the *R/FIELDImageR::fieldMask* function was implemented to remove the soil color from the RGB (Red-Green-Blue, also known as true color)images using HUE index corresponding to soil color [37]. The fifteen VIs were extracted from RGB images using the *R/FIELDImageR::fieldIndex* function for each flight date. The VIs included: the blue green pigment index (BGI) [38], brightness index (BI) [5], blue index, excessive green (EXG) [6], excess green minus excess red index (EXGR) [7], green leaf index (GLI) [8], green index, modified green red VI (MGVRI) [9], normalized difference index (NDI) [10], normalized green-blue difference index (NGBDI) [11], normalized green red difference index (NGRDI) [12], red index, red green blue VI (RGBVI) [9], visible atmospherically resistant index (VARI) [13] and vegetativen (VEG) [14] (Table 2). BI, GLI, NGRDI, VARI and BGI are prebuilt VIs that were used in the *FIELDImageR* package, but remaining VIs were added externally as an additional function into *FIELDImageR* using ‘*R/FIELDImageR::myIndex*’ function.

**Table 2.** Ratios of vegetation indices used in this study.

Vegetation Index	Formula	References
Blue green pigment index (BGI)	$\frac{B}{G}$	[38]
Brightness index (BI)	$\sqrt{\frac{R^2+G^2+B^2}{3}}$	[5]
Excessive green (EXG)	$(2 * G) - R - B$	[6]
Excess green minus excess red index (EXGR)	$(3 * G) - (2.4 * R) - B$	[7]
Green leaf index (GLI)	$\frac{2*G-R-B}{2*G+R+B}$	[8]
Modified green red index (MGVRI)	$\frac{G^2-R^2}{G^2+R^2}$	[9]
Normalized difference index (NDI)	$128 * \left( \left( \frac{G-R}{G+R} \right) + 1 \right)$	[10]
Normalized green-blue difference index (NGBDI)	$\frac{G-B}{G+B}$	[11]
Normalized green red difference index (NGRDI)	$\frac{G-R}{G+R}$	[12]
Red green blue index (RGBVI)	$\frac{G^2-R*B}{G^2+R*B}$	[9]
Visible atmospherically resistant index (VARI)	$\frac{G-R}{G+R-B}$	[13]
Vegetativen (VEG)	$\frac{G}{R^{0.667} * B^{0.334}}$	[14]

R, G and B represent the red, green, and blue bands respectively in the above formulas.

3D point clouds were used to extract the plot-based plant heights of each flight based on the pipeline described in Anderson et al. [15]. First, areas of the two trials (OHOT and DHOT) were clipped, and extreme points above or below the bare ground were removed from the point clouds in CloudCompare (version 2.12 alpha) [39]. After initial anomaly removal, executable functions of batch scripts of FUSION/LDV [40] and LAStools [41,42] software were transported into R to construct the plant height extraction pipeline for each flight. A brief outline of the pipeline contains (i) the sorting the points of clipped point cloud data (LAStools\lasssort.exe), (ii) removing the erratic points around the canopy structure of plants in row plots (LAStools\lasnoise.exe) to construct the digital surface model (DSM), (iii) identifying the ground points using a hierarchical robust interpolation (HRI) [43] ground filtering algorithm (HRI) [43] (FUSION\GroundFilter.exe), (iv) identifying the maximum points from the ground filter to outline digital terrain model (DTM) (LAStools\lasthin.exe), and (v) constructing the DTM model using the maximum points from the previous step (FUSION\GridSurfaceCreate.exe). Next, the DTM model was extracted from the previously constructed noise-filtered point cloud data (DSM model as an output of step ii) to canopy surface models (CSM). Each row of plots in CSM then were clipped (FUSION\PolyClipData.exe), and plot-based plant heights were estimated using the 99th percentile metrics (FUSION\CloudMetrics.exe) (Figure 1).

Plot-based yield was collected using a plot combine. Flowering time was recorded as the point in which 50% of the individuals within a plot were shedding pollen (DTA) or silks emerging (DTS). Days to anthesis/silk were calculated by subtracting the planting date from the recorded flowering date of the respective trait. All temporal phenotypic data as well as R codes are publicly available (see the Data Availability Statement and Supplementary Materials sections).

#### 2.4. Statistical Analysis for High Throughput Phenotyping Data

Statistical analysis of extracted temporal data (VIs and CHM) and visualizations were processed in R studio (version 1.3.959). In the analysis of high-throughput phenotyping data, a nested statistical design was fit. To predict the breeding values for each pedigree in both trials the row, range, (block) and replicates model terms were used to control for spatial variation. This tends to perform as well as other methods of spatial correction for furrow irrigated trials in our environment due to the patterns created by irrigation and tractor-applied fertilizer and cultivation in otherwise fairly uniform soils.

A nested design was used to estimate the temporal breeding values in a time-series manner for fifteen VIs and CHMs. To do this, the statistical Equation (1) below was developed to implement the two trials separately using a mixed linear model in *lme4* package in R [44] to estimate temporal best linear unbiased predictors (TBLUP) of the respective response variable ( $Y$ ) for each pedigree (maize hybrids).

$$Y_{ijklm} = \mu + flight_i + [pedigree(flight)]_{ij} + [range(flight)]_{ik} + [row(flight)]_{il} + [replication(flight)]_{im} + error_{ijklm} \quad (1)$$

where,  $Y$  represents each individual observation of pedigree at each  $i$ th time point (flight dates as DAP unit);  $\mu$  represents the grand mean;  $flight$  represents the effect of  $i$ th flight date,  $i \in (8, 10 \dots 105; DHOT)$  and  $i \in (30, 32 \dots 127; OHOT)$ ;  $[pedigree(flight)]_{ij}$  represents the effect of  $j$ th pedigree within  $i$ th flight date;  $[range(flight)]_{ik} \sim N(0, \sigma_{[range(flight)]_{ik}}^2)$  represents the effect of  $k$ th range within  $i$ th flight date;  $[row(flight)]_{il} \sim N(0, \sigma_{[row(flight)]_{il}}^2)$  represents the effect of  $l$ th row within  $i$ th flight date;  $[replication(flight)]_{im} \sim N(0, \sigma_{[replication(flight)]_{im}}^2)$  represents the effect of  $m$ th replication within  $i$ th flight date, and  $error_{ijklm} \sim N(0, \sigma_{error}^2)$  represents the pooled error comprising residuals of all experimental factors above. Temporal repeatability was calculated for each temporal trait based on the results of explained percent variation of  $[pedigree(flight)]_{ij}$  ( $\sigma_{pedigree(flight)}^2$ ) and  $error_{ijklm}$  ( $\sigma_{error}^2$ ) with numbers of replication (b) (Equation (2)):

$$\text{Temporal repeatability} = \frac{\sigma_{pedigree(flight)}^2}{\sigma_{pedigree(flight)}^2 + \frac{\sigma_{error}^2}{b}} \quad (2)$$

The previous model was adjusted to run without the  $flight$  component to estimate the BLUP for pedigree (maize hybrids) for single time-measured traits that are yield, DTA and DTS (Equation (3)):

$$Y_{ijkl} = \mu + pedigree_i + range_j + row_k + replication_l + error_{ijkl} \quad (3)$$

where  $Y$  represents the each individual observation of pedigree;  $\mu$  represents the grand mean;  $pedigree_i$  represents the effect of  $i$ th pedigree;  $range_j \sim N(0, \sigma_{range_j}^2)$  represents the effect of  $j$ th range;  $row_k \sim N(0, \sigma_{row_k}^2)$  represents the effect of  $k$ th row;  $replication_l \sim N(0, \sigma_{replication_l}^2)$  represents the effect of  $l$ th replication, and  $error_{ijkl} \sim N(0, \sigma_{error_{ijkl}}^2)$  represents the pooled error comprising residuals of all experimental factors above. Repeatability was calculated for Yield, DTA and DTS from Equation (4):

$$\text{Repeatability} = \frac{\sigma_{pedigree}^2}{\sigma_{pedigree}^2 + \frac{\sigma_{error}^2}{b}} \quad (4)$$

Statistical Equation (5) was developed by adding the *management* and  $[pedigree * management]$  components into Equation (1) as nested within  $flight$ . In Equation (5), the management effect contained the OHOT and DHOT trials, indicating optimal and delayed planting, respectively. Equation (5) aimed to contrast each trait value of flight date for each trial. To do that,  $[management * flight]_{ik} \sim N(0, \sigma_{[management * flight]_{ik}}^2)$  effect was included. Here, *management* represents the effect of  $k$ th management,  $k \in (DHOT \text{ and } OHOT)$ ;  $[management * pedigree(flight)]_{ijk} \sim N(0, \sigma_{[management * pedigree(flight)]_{ijk}}^2)$  represents the in-

teraction of  $j$ th pedigree with  $k$ th management within  $i$ th flight date. Spatial variation (*range*, *row* and *rep*) were treated nested within *management* and *flight* in Equation (5).

$$\begin{aligned} Y_{ijklmn} = \mu + & \text{flight}_i + [\text{pedigree}(\text{flight})]_{ij} + \text{management}_k \\ & + [\text{pedigree} * \text{management}(\text{flight})]_{ijk} + [\text{flight} * \text{management}]_{ik} \\ & + [\text{range}(\text{management}, \text{flight})]_{ikl} + [\text{row}(\text{management}, \text{flight})]_{ikm} \\ & + [\text{rep}(\text{management}, \text{flight})]_{ikn} + \text{error}_{ijklmn} \end{aligned} \quad (5)$$

Multi-management based repeatability was calculated by expanding Equation (2) to include the interaction term between *management* and *pedigree* nested within *flight* [*pedigree* \* *management*(*flight*)] with the number of the management (a) and replication (b) (Equation (6)):

$$\text{Repeatability} = \frac{\sigma_{\text{pedigree}(\text{flight})}^2}{\sigma_{\text{pedigree}(\text{flight})}^2 + \frac{\sigma_{\text{pedigree} * \text{management}(\text{flight})}^2}{a} + \frac{\sigma_{\text{error}}^2}{a*b}} \quad (6)$$

## 2.5. Machine Learning Regression

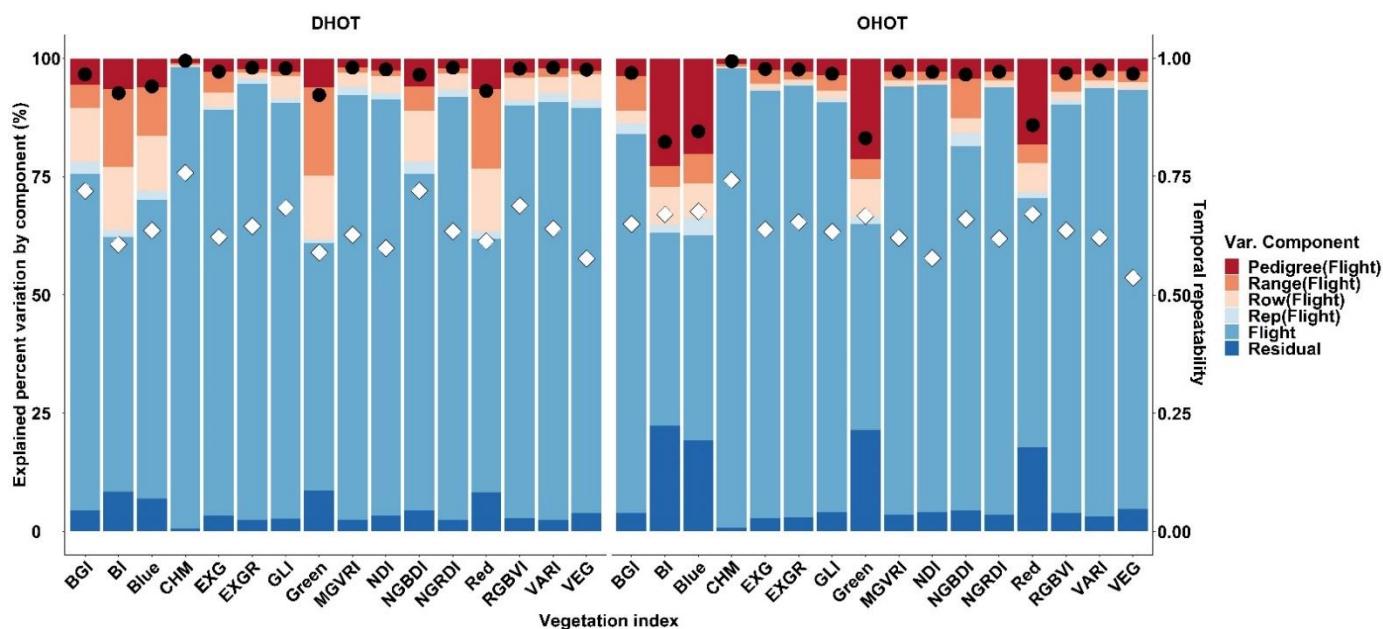
All temporal breeding values (TBLUPs) of VIs and CHMs of pedigree were used to predict three predicted variables (Yield, DTA and DTS) for both populations. Linear, lasso, ridge, elastic net and partial least square regression (PLSR) were run to predict the predicted variables using the *caret* package in R. The data set was split into 60% training and 40% validation data. Cross validation was adjusted using the *R/caret::trainControl()* function, with ten resampling iterations (i.e., folds; numbers = 10) and ten repeated k-fold cross validations (repeats = 10). The adjusted cross validation was then used in each regression model. To run all regression models, *R/caret::train()* function was implemented using *method* = *lm* for linear regression and *method* = *glmnet* for ridge, lasso, and elastic net regressions and “*method* = *pls*” for PLSR. The *tuneGrid* argument was used to tune the “*lambda*” (for elastic net, lasso and ridge regressions) and “*alpha*” (for elastic net regression) using the *expand.grid()* function. Value of “*alpha*” was set as 0 for ridge regression and 1 for lasso regression, while sequential numbers between 0 and 1 by ten equal increment sequences were performed to find the best alpha for elastic net regression. Sequential “*lambda*” values between 0 and 1 by five equal increment sequences were empirically tested to find the best lambda values for lasso, ridge, and elastic net regressions. In PLSR, the *tuneLength* function was used to find the best number of principal components with the lowest cross-validation error. In the prediction of yield in the DHOT trial, *alpha* and *lambda* values used were 1 and 0.25, 0 and 1, 0.1 and 0.50 for lasso, ridge and elastic net regressions, respectively, and numbers of principal component (*ncomp*) used was 8 for PLSR. In the prediction of DTA in the DHOT trial, *alpha* and *lambda* values used were 1 and 0.0001, 0 and 1, 0.1 and 0.0001 for lasso, ridge and elastic net regressions, respectively, and *ncomp* used was 10 for PLSR. In the prediction of DTS in the DHOT trial, *alpha* and *lambda* values used were 1 and 0.0001, 0 and 1, 0.1 and 0.0001 for lasso, ridge and elastic net regressions, respectively, and *ncomp* used was 7 for PLSR. In the prediction of yield in the OHOT trial, *alpha* and *lambda* values used were 1 and 0.25, 0 and 1, 0.1 and 1 for lasso, ridge and elastic net regressions, respectively, and *ncomp* used was 4 for PLSR. In the prediction of DTA in the OHOT trial, *alpha* and *lambda* values used were 1 and 0.25, 0 and 1, 0.1 and 0.50 for lasso, ridge and elastic net regressions, respectively, and *ncomp* used was 10 for PLSR. In the prediction of DTS in the OHOT trial, *alpha* and *lambda* values used were 1 and 0.25, 0 and 1, 0.1 and 0.25 for lasso, ridge and elastic net regressions, respectively, and *ncomp* used was 9 for PLSR. To compare the regression models, coefficient of determination ( $R^2$ ), root mean square errors (RMSE), and mean absolute errors (MAE) were evaluated for each model. To evaluate the prediction accuracy of the linear, ridge, elastic net, lasso and PLSR models, *R/caret::predict()* was used to predict DTA, DTS, and yield of the test data set in each trial using 500 bootstraps iterations. The predicted results of the test data were correlated

(Pearson pairwise correlation) with the actual values of test data set in each bootstrap. Finally, the Tukey HSD (Honest Significant Difference) test was applied to compare the means of correlation results belonging to each model for DTA, DTS, and yield in both trials. Letters were assigned to each model indicating the comparison results of Tukey HSD test to indicate statistically significant differences. The *R/caret::varImp()* function was used to rank the predictors based on their importance (0–100 scale) for lasso, ridge, elastic net and PLSR regression. All R codes are available in Github repository (<https://github.com/alperadak/Supplementary-File-1/blob/main/Supplementary%20file%201.txt>, accessed on 28 May 2021).

### 3. Results

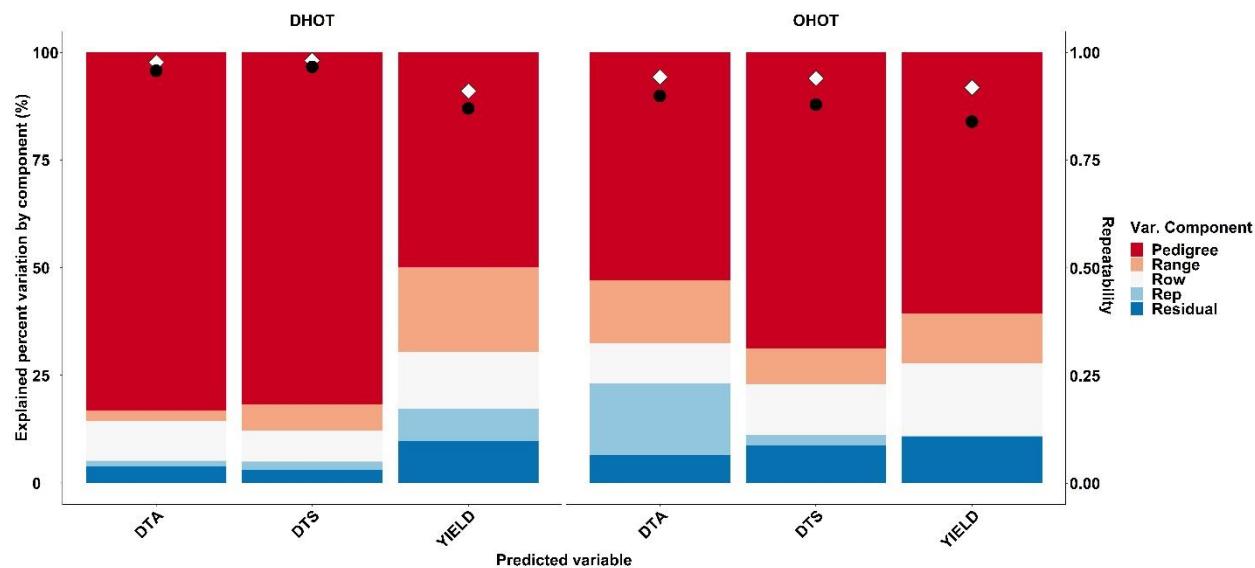
#### 3.1. Explained Percent Variation of Flight and Repeatability

The variation explained by the *flight* component in Equation (1) was the greatest of the variance components and was statistically significant ( $>0.0001$ ) for each UAS trait (fifteen different VIs and CHMs) in each trial. The *flight* component explained up to 97–98% of variation in the CHM for DHOT and OHOT, respectively, while explaining the lowest for true bands (Blue, Red and Green) and BI (41–51%; Figure 2). Similarly, the highest temporal repeatability values were also estimated for CHMs (0.76 in DHOT and 0.74 in OHOT; Equation (2)) and lower repeatability estimates were found for the VI's ( $<0.6$ ; Figure 2).

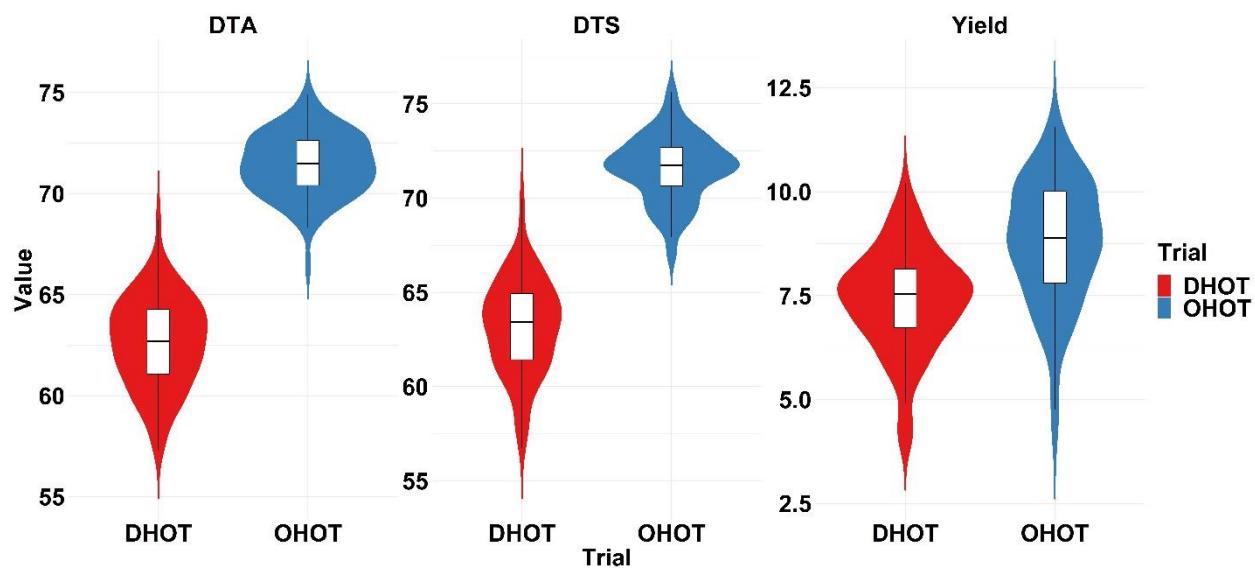


**Figure 2.** Variance component decomposition (Equation (1)) for each predictor trait (fifteen vegetative indices, VIs and canopy height measurement; CHM) for each trial (DHOT on the left and OHOT on the right). X axes, left Y axis, and right Y axes represent the traits, explained percent variation by each variance component, and temporal repeatability, respectively. Black circles and white diamonds represent the  $R^2$  and repeatability values, respectively, for each trait according to the right Y axis scale. Temporal repeatability was calculated based on the Equation (2).

For predicted variables (yield, DTA and DTS) used in the regression models, the explained percent variations of the *pedigree* component in Equation (3) explained the greatest variation (53–83%). Repeatability values for predicted variables were calculated between 0.91 and 0.98 (Figure 3). Predicted values (Equation (3)) of maize hybrids for DTA, DTS, and yield in both trials were given in Figure 4. DHOT trial had shorter days after planting values in DTA and DTS than those of OHOT while grain yield (t/ha) was higher in OHOT than those of DHOT (Figure 4).



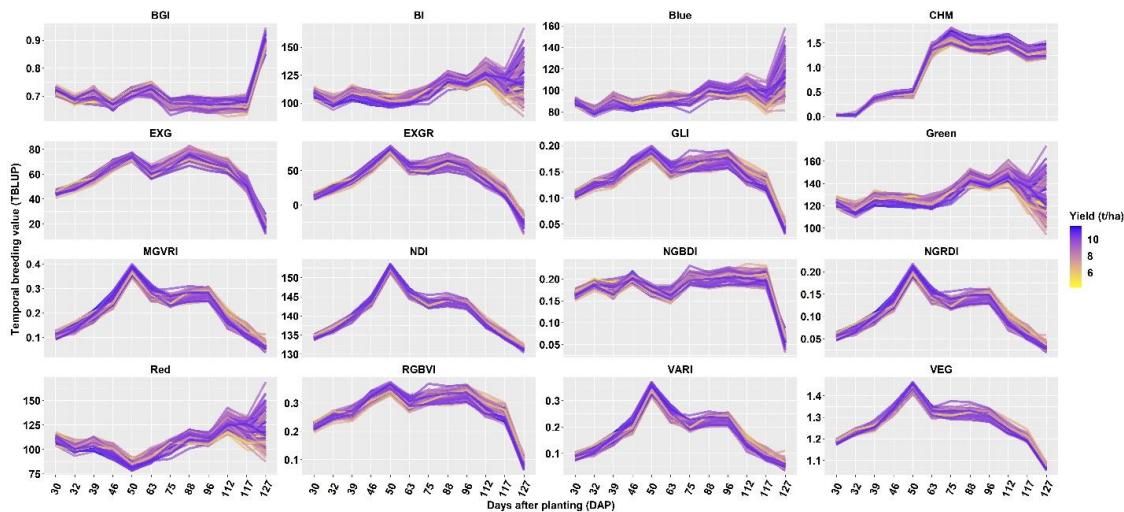
**Figure 3.** Explained percent variation of each component in Equation (3) for each predicted variable (days to anthesis (DTA), days to silking (DTS), and yield ( $t/ha$ )) for each trial (DHOT on the left and OHOT on the right). X axes, left Y, and right Y axes represent the traits, explained percent variation by each variance component, and repeatability, respectively. Black circles and white diamonds represent the  $R^2$  and repeatability values, respectively.



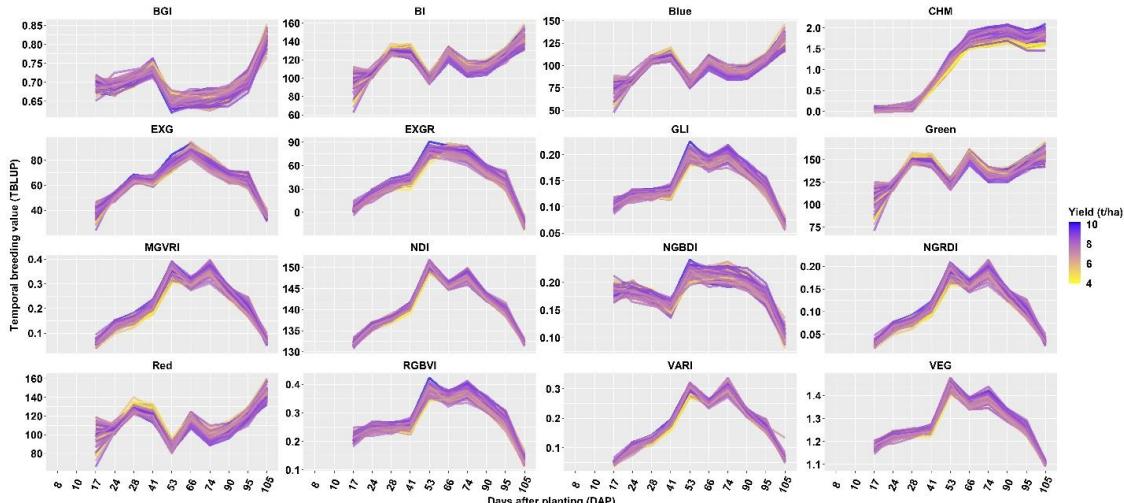
**Figure 4.** The best linear unbiased predictors (BLUPs) values of predicted variables [days to anthesis (DTA), days to silking (DTS), and yield ( $t/ha$ )]. Y axes shows to best linear unbiased predictors (BLUP) of hybrids for DTA, DTS and yield unique to each trait while X axes shows the trials.

### 3.2. Temporal Breeding Values

Effects of the *pedigree(flight)* (Equation (1)) were illustrated to display temporal breeding values (TBLUPs) of predictor traits (fifteen VIs and CHMs) of pedigrees for optimal planting (OHOT) and delayed planting (DHOT) trials separately (Figures 5 and 6). In all, TBLUPs of 16 traits, which were derived from 12 flight time points, were displayed, corresponding to 192 predictors evaluated in regression models for each trial.



**Figure 5.** Temporal best linear unbiased predictions (TBLUPs) of the traits (fifteen vegetation indices, VIs and canopy height measurement; CHM) of the pedigrees in the OHOT (optimal planting trial) population estimated by Equation (1). Each Y axis shows the range of TBLUPs unique to each trait while each X axis shows the flight dates as days after planting (DAP) same to each trait. The heatmap scale was generated from the range of yield (t/ha) values in the OHOT trial, and then applied to each pedigree to show the TBLUPs of each trait through the flight dates along with yield values of pedigree. Blue, white, and red colors in the heatmap scale were used to indicate low, medium, and high yield values, respectively, specific to the OHOT trial.

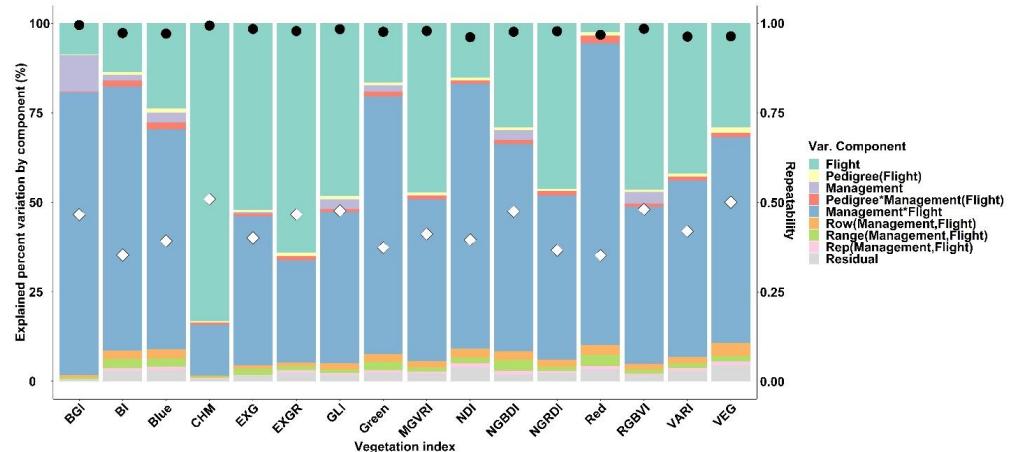


**Figure 6.** Temporal best linear unbiased predictions (TBLUPs) of the traits (fifteen vegetation indices, VIs and canopy height measurement; CHM) of the pedigrees in the DHOT (delayed planting trial) population estimated by Equation (1). Each Y axis shows the range of TBLUPs unique to each trait while each X axis shows the flight dates as days after planting (DAP) same to each trait. The heatmap scale was generated from the range of yield (t/ha) values in the DHOT trial, and then applied to each pedigree to show the TBLUPs of each trait through the flight dates along with yield values of pedigree. Blue, white, and red colors in the heatmap scale were used to indicate low, medium and high yield values, respectively, specific to the DHOT trial.

TPP of VIs and CHMs occurring among the flight dates ( $[pedigree(flight)]$ ) component of Equation (1) were found to be statistically significant ( $<0.0001$ ) in both trials. TPP patterns were also found to be unique for many traits across trials. In general, EXG, EXGR, GLI, MGVRI, NDI NGRDI, RGBVI, VARI, and VEG followed concave plasticity patterns while Red followed a slightly convex plasticity pattern (Figures 5 and 6). BI, Blue, Green, and NGBDI slightly increased over time while BGI slightly decreased across the flight dates. CHM followed a sigmoidal plasticity pattern through the flight dates. However,

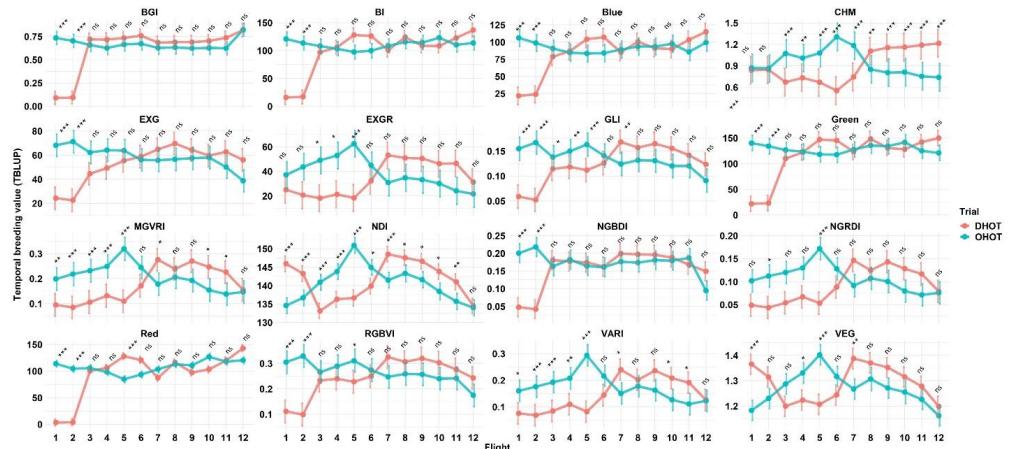
each trial had different slopes, edges, and peaks of the TBLUPs of the traits depending on the flight times (Figures 5 and 6).

Equation (5) was developed to compare the flight means of each trial for each trait by adding the management (DHOT and OHOT trials) effects within flight dates ([management(flight) effects in Equation (5)). The [management(flight)] component explained between 29 and 96% of total depending on VIs and 14 percent for CHM (Figure 7).



**Figure 7.** Explained percent variation of each component in Equation (5) for each predictor trait (fifteen vegetation indices, VIs and canopy height measurement; CHM). X axis, left and right Y axes represent the traits, explained percent variation by each variance component and  $R^2$  respectively. Black circles and white diamonds represent the  $R^2$  and repeatability values, respectively, for each trait according to the right Y axis scale. Repeatability was calculated based on Equation (6).

The majority of the mean comparisons of each time point (flight dates) between the trials were found to be statistically significant (Figure 8). In addition, there were found to be cross over changes in comparisons of flight means varying between one to six times depending on different traits; meaning that the relative performance at one time point and trial could not predict the relative performance in a different time point. However, where there were more cross overs, they were less likely to be significantly different.



**Figure 8.** Temporal comparison results of each time point between each test. The joint analysis of temporal effects of the trials based on means of flight dates was derived from *Management(flight)* component in Equation (5). Each Y axis shows the range of flight means unique to each trait while each X axis shows the flight dates as successive order same to each trait. \*, \*\*, \*\*\* are significance levels at 0.05, 0.01 and 0.001; ns is not significant. Whiskers represent the conditional standard deviations for each time point of each trait.

### 3.3. Temporal Correlations between Predictors and Predicted Variables

Temporal correlation coefficients (Pearson correlation) were calculated between TBLUPs of the traits (predictors) and predicted variables (DTA, DTS, and yield) (Figure 9). To calculate the temporal correlations, BLUPs of pedigrees for each flight date were derived from  $[pedigree(flight)]$  in Equation (1) for each trait and trial. BLUPs of DTA, DTS, and yield were correlated with TBLUPs of predictor traits for each flight date. Temporal correlation results varied from  $-0.6$  to  $0.6$  across the flight dates. The majority of predictor traits fluctuated in correlation temporally (Figure 9). The highest and lowest temporal correlation with DTA, DTS, and yield were found more frequently in DHOT than in OHOT (Figure 9). All traits were found to have both negative and positive correlation coefficients in both trials through the flight dates, except for positive correlation coefficients between CHM and yield in both trials throughout the flight dates (Figure 9).



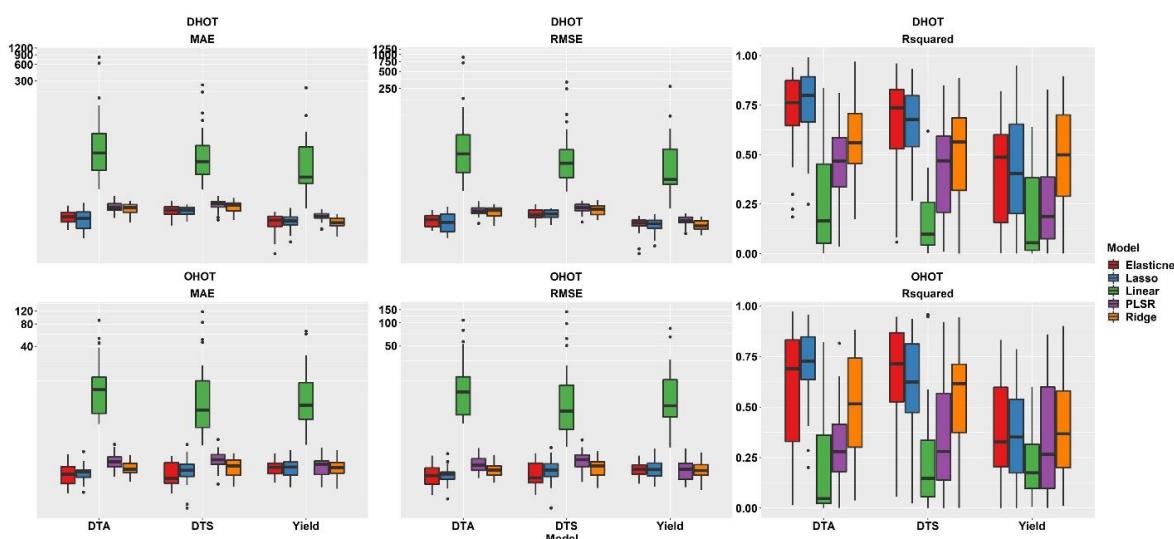
**Figure 9.** Each correlogram chart contains the pairwise correlation coefficients ( $r^2$ ) belonging to time points of each trait with predicted traits [days to anthesis (DTA), days to silking (DTS), and yield (t/ha)]. Correlogram charts above and below the horizontal black dashed line belong to DHOT (delayed planting trial) and OHOT (optimal planting trial) respectively. Time points (days after planting, DAP) were given diagonal and horizontal for each vegetation index and CHM in both trials along with DTA, DTS and yield in each correlogram chart. The correlation coefficient heatmap change from  $-1$  to  $0$  from pink to white and from  $0$  to  $1$  from white to cyan. The cross signs ( $\times$ ) show statistically insignificant pairwise correlations at the  $0.05$  level.

### 3.4. Temporal Correlations among Predictors

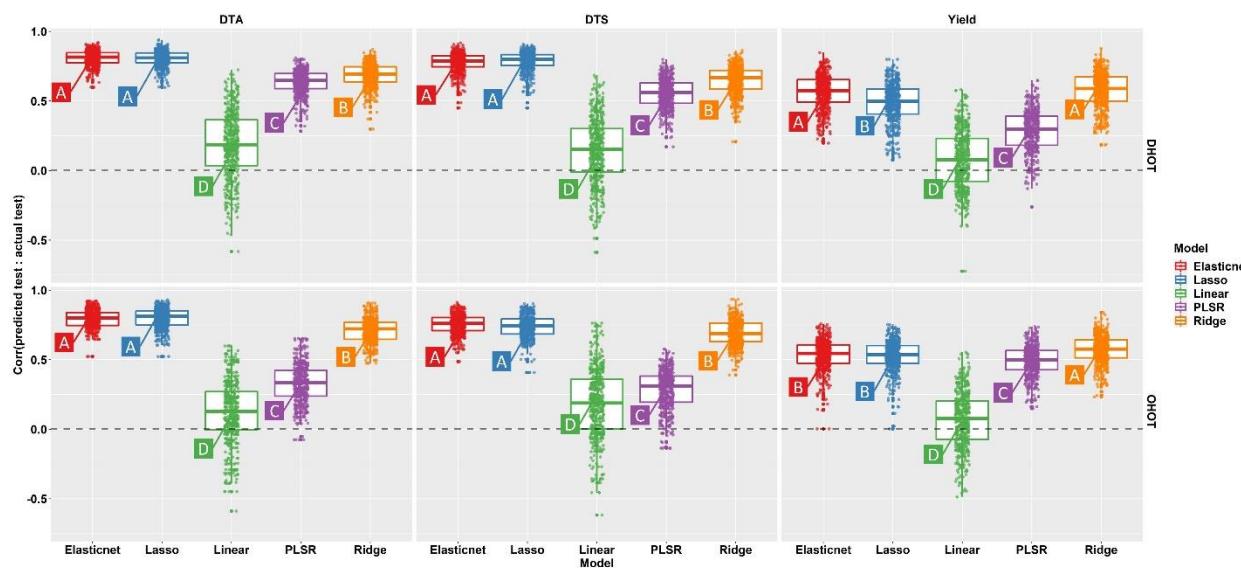
Pairwise correlation coefficients were calculated separately for each time point of each VI and CHM in both trials to show the relationship each image-based phenotype had with itself throughout time (Figure 9). Greater correlation coefficients were found to be more frequent in the OHOT than DHOT trial for majority of VIs as well as CHM, but in both trials, pairwise correlations belonging to certain time points (flight dates) of VIs and CHM were also found to be higher, such as correlations among the later flights in CHM in both trials (Figure 9).

### 3.5. Regression Model Comparisons

Temporal best linear unbiased predictors (TBLUPs) of pedigrees for each predictor trait (fifteen VIs and CHM) were estimated by the component of  $[pedigree(flight)]$  in Equation (1) for each trial separately. BLUPs of predicted variables (DTA, DTS, and yield) were estimated from the *pedigree* component in Equation (3). Predictor traits (fifteen VIs and CHM) were used in four different regressions to predict the predicted variables. TBLUPs, including the individual BLUPs of sixteen predictor traits on twelve flight dates (192 predictors for each trial), were estimated by Equation (1) separately and then used in regression models to predict the predicted variables (DTA, DTS and Yield) for each trial. Elastic net, lasso, ridge, PLSR and linear regressions were used to predict the predicted variables using all 192 predictors and models. Model fit was evaluated based on their root mean square errors (RMSE), mean absolute errors (MAE), and coefficient of determination ( $R^2$ ) as well as prediction accuracies of the models. Linear regression was found to perform the worst among regression models in predicting all three predicted variables in both trials, resulting in the highest RMSE, MAE, and lowest  $R^2$  values and lowest prediction accuracies followed by PLSR (Figures 10 and 11). Slight differences were found among the results of machine learning regressions in terms of predicting predicted variables (Figures 10 and 11). Accordingly, lasso and elastic net regressions were found to be best in predicting flowering times (DTA and DTS) with the prediction accuracy of ~0.75 and ~0.80 in DHOT and OHOT trials (Figure 11). However, ridge regression was found to be best in predicting yield, with prediction accuracy of ~0.60 in DHOT and OHOT trials (Figure 11).



**Figure 10.** Box plots of root mean squares (RMSE), mean absolute errors (MAE), and coefficient of determination ( $R^2$ ) values (from left to right) of linear, elastic net, lasso, and ridge regressions for DHOT (delayed planting trial; above) and OHOT (optimal planting trial; below). Each Y axis has the unique value ranges for RMSE, MAE and  $R^2$  in each trial while each X axis shows the predicted variables used in each regression models [days to anthesis (DTA), days to silking (DTS), and yield (t/ha)] same to RMSE, MAE, and  $R^2$  in each trial. Whiskers represent the standard errors. Y axes of RMSE and MAE were scaled based on the  $\log_2$  to show the outliers belonging to the linear model.



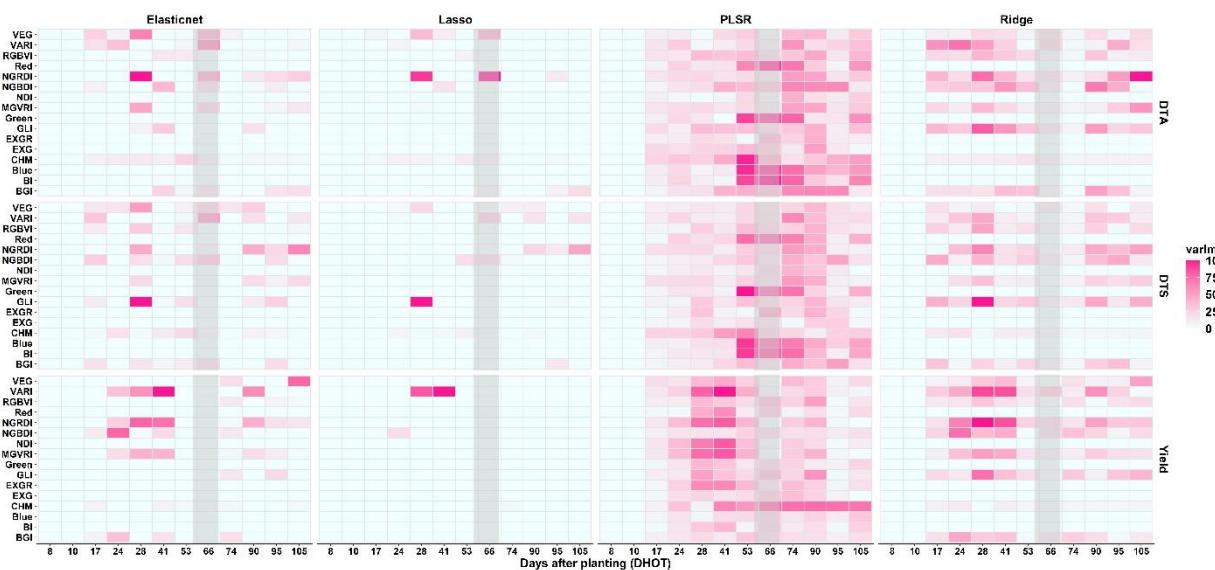
**Figure 11.** Prediction accuracy results of each model for days to anthesis (DTA), days to silking (DTS), and yield ( $t/ha$ ), respectively. Y axes represents the correlation coefficients ( $r^2$ ) between predicted value and actual value of days to anthesis (DTA), days to silking (DTS) and yield ( $t/ha$ ) (from left to right) in DHOT (delayed planting trial) and OHOT (optimal planting trial) (From top to bottom). X axis shows the predictive models used in this study. Each box represents letters of Tukey HSD comparison results for each predicted trait in each trial.

### 3.6. Variable Importance

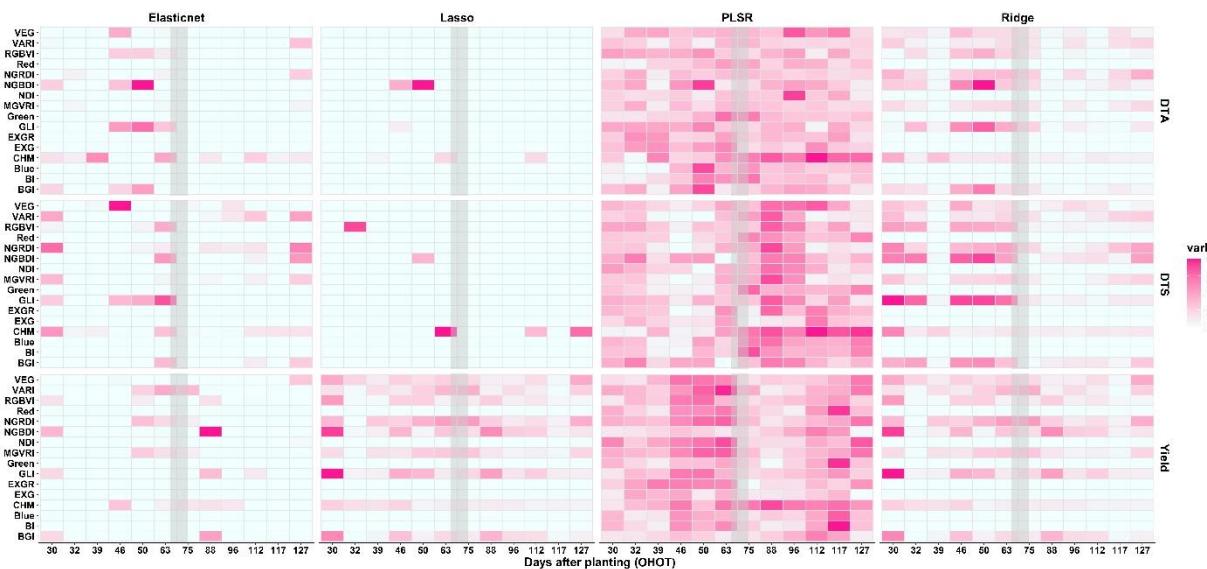
The relative importance of the TBLUPs were calculated to show the temporal importance of predictor traits in the prediction of the predicted variables (DTA, DTS, and yield) used in machine learning regressions. Results showed that some of the VIs were found to contribute predictions of predicted variables throughout the flight dates, while others were found to contribute predictions at specific flight dates, depending on the predicted variables as well as the trials (Figures 12 and 13).

The best-performing predictors of yield nominated in ridge (in both trials) and elastic net (in DHOT only) regressions (Figure 11) were NGRDI, NGBDI VARI and GLI, traits with fluctuating variable importance scores across the flight dates in both trials (Figures 12 and 13). Their variable importance scores were frequently higher before flowering than after flowering time (Figures 12 and 13). The remaining traits were found to have a lower importance score than NGRDI, NGBDI, VARI and GLI in the prediction of yield across both trials (Figures 12 and 13); however, their importance scores were also higher before the plants flowered in both trials.

Lasso and elastic net regressions performed best in predicting the flowering times (DTA and DTS), in the both trials (Figure 11). The methods nominated fluctuating variable importance of the temporal traits across the flight dates (Figures 12 and 13); many of the flight dates belong to the period before flowering. For instance, GLI and NGRDI on 28th DAP were important date/predictor combinations for flowering time in the DHOT trial (Figure 12). NGBDI and GLI on 46th, 50th, 63rd DAP for DTA, (Figure 13), and NGRDI, VARI, CHM, GLI and MGVRI on 30th DAP for DTS (Figure 13) were important date/predictor combinations in the OHOT trial.



**Figure 12.** Heatmap temporal variable importance of predictor traits across the flight dates generated by the ‘varImp’ function of each machine learning regression in the DHOT trial. X axes are identical and represent the flight dates of DHOT as days after planting (DAP). Left Y-axes are identical and represent the predictor traits for each regression model and predicted variables. The heatmap illustrations were divided into nine for the elastic net, lasso and ridge regressions top to bottom and DTA, DTS and Yield predicted variables from left to right. Heatmap scales show the variable importance scores as 0–100 scales calculated from the ‘varImp’ functions of each machine learning regression. Flowering times varied from 57 to 69 for DTA and 57 to 70 for DTS (grey boxes) in DHOT.



**Figure 13.** Heatmap of temporal variable importance of predictor traits across the flight dates generated by the ‘varImp’ function of each machine learning regression in the OHOT trial. X axes are identical and represent the flight dates of OHOT as days after planting (DAP). Left Y axes are identical and represent the predictor traits for each regression model and predicted variables. The heatmap illustrations were divided into nine for the elastic net, lasso and ridge regressions top to bottom and DTA, DTS and Yield predicted variables from left to right. Heatmap scales show the variables importance scores as 0–100 scales calculated from the ‘varImp’ functions of each machine learning regression. Flowering times varied from 66 to 75 for DTA and 67 to 76 for DTS (grey boxes) in OHOT.

## 4. Discussion

Implementing field-based high-throughput phenotyping (HTP) tools such as unoccupied aerial systems (UAS) in agriculture has been recognized to be an efficient way to monitor temporal variation of plant growth [2,15,30,31]. Temporal variations throughout growth are novel sources of variation that have not previously been measured and now can be used for plant breeding and genetics in prediction models. Estimating the degree of phenotypic variation controlled by the genetic variation of tested pedigrees for HTP-derived traits (e.g., measuring repeatability for VIs and CHMs) is a first fundamental step to determine how heritable or consistent trait measurements are before using VIs and CHMs in prediction models. The sixteen traits (fifteen VIs and CHM) used in this study had temporal repeatability values of 0.60 to 0.76 in delayed and optimal planting trials (DHOT and OHOT trials, respectively) (Figure 2), suggesting that estimating the breeding values from HTP-derived traits of tested pedigrees has potential in prediction models as well as in the decision making process for plant breeding. Temporal variation of each HTP-derived trait, estimated by Equation (1) for each trial separately, was the source of the biggest proportion of total experimental variation. This temporal resolution, derived from twelve time points, supported the existence of TPP across pedigrees for fifteen VIs and CHMs in both trials. Joint analysis of both trials, estimated by Equation (5), showed that different reaction norms to compare each time point (flight means) under different environmental conditions characterized the dissimilar TPP patterns of VIs and CHMs. Importantly, NGBDI and BGI were found to have the least TPP under two distinct environmental conditions and are proposed as the most robust VIs from this study. Moreover, NGBDI was found to be an important variable in yield prediction at 30 DAP in the OHOT trial, as well as for predicting flowering times at 28 DAP in both OHOT and DHOT; if this can be further validated it would be valuable and exciting for early-season predictions of grain yield. The above, along with the high repeatability values (~0.7 in Figure 2) nominate NGBDI, NGRDI, and GLI as the most effective and robust predictors of yield and flowering times in this study. However, it remains worthwhile to estimate multiple VI's since the same collected data is needed and many indices had non-redundant information.

### 4.1. Physiological Basis of These Predictions

Vegetation indices (VIs) are biologically based remote sensing measurements that can assess the physiological processes of maize hybrids quantitatively using the reflected wavelengths. More specifically, VIs are ratios that belong to the green apparatus of maize where photosynthetic activity occurs. The visible spectrum of solar radiation (400 nm to 700 nm) is absorbed by plants pigments, of which chlorophyll *a* and *b* absorb the blue-violet and red-blue regions of the visible spectrum to provide the light for photosynthetic activity and reflect far red wavelengths. Photoassimilate is produced by photosynthetic activity containing the light dependent reactions that are likely affected by different environmental conditions and contributed to yield differences. As a result of these processes, temporal correlations, and predictive ability of VIs with yield in maize likely depends on environmental variability [45]; in agreement with our findings (Figure 9). Differences between the wavelengths belonging to different varieties tested in different environmental conditions relate to how the plants handle multiple stresses [46].

Although this study showed the surprising predictive ability of VIs belonging to time points earlier than the reproductive (flowering) stage, for flowering times later time points were also important, likely relating to senescence after grain fill. Grain yield can be adversely affected because of heat stress or early senescence after flowering. VIs belonging to time points after flowering lose their ability to predict yield as senescence accelerates, which was found for both trials. Early leaf senescence, especially, shortens the grain-filling duration in maize, and would be expected to reduce measured grain yield. A leaf senescence index has been proposed to be well correlated with VIs, as well as yield in maize under low nitrogen stress [46]. Slow reduction in photosynthetic pigments throughout plant growth is associated with greater yield, as well as with the stay-green

trait [47]. As such, using temporal breeding values of the VIs near the end of growth may quantitatively select maize hybrids where the higher yield is a result of slow reduction in the VIs across growth.

This study used VIs calculated by wavelength of ranges of visible light, not affected by additional factors that affect near-infrared (NIR) reflectance such as soil brightness, moisture and color, canopy and cell structure, or cloud shadows [13]. As such, these VIs belonging to false positive color reflectance should result in greater precision in extraction of VIs, compared to NIR reflectance [13,48]. For example, normalized difference vegetation index (NDVI) which contains the NIR reflectance, was found to be very sensitive towards background brightness since NDVI increases accordingly if the background brightness increases [49]. It might be concluded that the VIs without NIR used here are more sensitive in detecting high amounts of absorption of red referring to the presence of the chlorophyll and green vegetation, disregarding noisy reflectance.

The delayed planting trial (DHOT) had higher growing degree days as well as higher photo-thermal time accumulations than the optimal planting trial (OHOT) across plant growth stages in maize breeding nurseries in Texas. The temporal differences in weather station measurements between the two trials suggest a primary reason for the TPP of the pedigrees of several VIs likely stemmed from the rapid accumulation of growing units in DHOT, a combination of both higher photoperiod and increased heat. Taken together, despite the fact that VIs and CHMs were shown to be controlled by genetic variation to a certain extent, monitoring of TPP belonging to VIs and CHMs can be more appropriate to evaluate performance of maize hybrids under diverse environmental conditions including their different time points. Because correlations between VIs and yield, or CHMs and yield, are dynamic, and different time points had different correlation coefficients (Figure 9), this suggests that correlations in earlier time points could be considered to assess the performance of maize hybrids. In genomic selection, selections are accomplished through estimating the large and small marker effects on phenotype belonging to terminal growth stages. Now, with temporal phenotypes estimated by Equation (1), it is possible to include the genotypic variation occurring at early stages into phenomic selection [50] and phenomic selection can contain temporal marker effects on VIs and CHMs that give rise to better prediction accuracy. Analyzing multiple time points of different management conditions can supply better confidence levels, as well as, more statistical power in predicting temporal breeding values; this will allow temporal breeding values to be later used in downstream analysis, such as, more accurate prediction modeling. In addition, using VIs with high repeatability values as predictive variables will lead to more effective estimates, especially under stress conditions, largely because the signal-to-noise ratio is higher [21,51].

#### 4.2. Model Comparisons

Machine learning-based prediction models performed best in comparison with linear models for predicting flowering times and yield using the temporal data. The predictive ability of ridge regression in genomic prediction has been demonstrated [52]. Similarly, ridge regression excelled in predicting yield in phenomic prediction compared to other models in this study (Figure 11), signifying that prediction accuracy of ridge regression can be increased in selecting the high yielding maize hybrids when a greater number of phenotypic predictors are extracted from UAV images. Ridge regression models have previously been used in yield prediction using reflectance wavelengths (between 350 to 2500 nm) in wheat and performed best in prediction under drought [53]. This result is in agreement with our finding that ridge regression performed best under the stress trial (DHOT) in yield prediction by using VIs. Hernandez et al. [53] in wheat used only reflectance at anthesis and grain filling time periods; this study, however, discovered important effects of VIs at many earlier time points in maize, even before flowering and grain filling such as at 28 and 30 DAP in the DHOT and OHOT trial, respectively. If this can be validated in further work it is an exciting and important finding not only for speeding up breeding decisions earlier in the season, and the associated physiological ramifications,

but also for identifying important growth stages for precision agricultural management by farmers. Our findings also support the value of temporal VIs measurement for predicting yield; not only during specific time periods of growth, but that the entire growth period has value with different effect sizes at each time point. In this way these VI's may be better used like a genomic selection approach of incorporating and weighting all time points, than a genetic mapping approach of identifying and monitoring only the time point with the largest effect size; as supported in phenomic selection approaches, using laboratory near-infrared reflectance spectroscopy data of grain [54,55].

This study also revealed that flowering times (DTA and DTS) can be best predicted before flowering using the temporal VIs, which is critical where earliness traits for maize can be selected based on the values of certain VIs at specific early time points, such as NGRDI and VARI at 28 DAP in DHOT; NGRDI and GLI at 30 DAP in OHOT (Figures 12 and 13).

Single time point-derived CHMs [15] and reflectance bands [20] have previously been shown to perform worse than using fewer parameters estimated from modeling combined data derived across multiple time points. Our findings suggested that temporal variations of VIs and CHMs should be trained together in a predictive model using machine learning based regressions to predict grain yield with more accuracy, when the number of predictors are high, which otherwise may lead to overfitting and multicollinearity problems.

Given the long history of satellite remote sensing to predict yield [56,57], the results of the temporal and visible aspects of this study were surprising. However, it must be emphasized that to date there have been comparatively few studies using remote sensing predictors across segregating genetic populations to predict differences within species. These findings similarly suggest that unquantified genetic variation might exist that would affect satellite yield predictions. This also suggests new temporal measurements that farmers might be able to use to identify plant stress and intervene earlier than previously to predict yield loss. In summary, the strong predictive power of temporal VIs and CHMs for predicting yield was shown in this study which could serve as an alternative to genomic prediction methods in the near future.

## 5. Conclusions

Field-based high-throughput phenotyping identified that the phenotypic variation occurring early in developmental stages of plants can be used as novel predictors to predict yield and flowering times in phenomic selection. This research extracted fifteen different vegetation indices and canopy height measurements from the aerial drone imageries belonging to twelve time points across the growth of tested maize hybrids. Our study demonstrated that (i) predicted temporal vegetation indices and plant heights for each pedigree (maize hybrid) have discrimination power of low-yielding and high-yielding maize hybrids and that these temporal vegetation indices and plant heights are heritable traits in maize; (ii) machine learning-based regressions have better prediction accuracy in prediction yield and flowering times than linear regression when temporal vegetation indices and plant heights are used as predictors in phenomics selection in maize; and (iii) variable importance scores of vegetation indices are higher at earlier time points than later time points, indicating that variation in earlier in growth should be monitored to predict yield and flowering times with higher accuracy in maize.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/rs13112141/s1>, Spreadsheet S1: Excel sheets contain the predictors (temporal vegetation indices and canopy height measurements) and predicted traits (days to anthesis, silking and yield) belonging to OHOT (optimal planting trial) and DHOT (late planting trial) predicted by Equation (1) and used in prediction models.

**Author Contributions:** Conceptualization, A.A. and S.C.M.; methodology, A.A., S.C.M., S.B., S.C. and S.L.A.II; software, A.A., S.C.M. and S.L.A.II; validation, A.A., S.C.M., S.L.A.II and S.C.; formal analysis, A.A. and S.C.M.; investigation, A.A., S.C.M., S.B., R.L., S.N., S.C. and S.W.; resources, A.A. and S.C.M.; data curation, A.A., S.C.M., S.B., S.L.A.II, S.C. and S.W.; writing—original draft preparation, A.A. and S.C.M.; writing—review and editing, A.A., S.C.M., S.B., R.L., S.N., S.C., S.L.A.II and S.W.; visualization, A.A.; supervision, S.C.M.; project administration, S.C.M.; funding acquisition, S.C.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** USDA-NIFA-AFRI with awards numbers of 2017-67013-26185, 2020-68013-32371, and 2021-67013-33915, USDA-NIFA Hatch funds, Texas A&M AgriLife Research, the Texas Corn Producers Board and the Eugene Butler Endowed Chair in Biotechnology financially support this research.

**Data Availability Statement:** Data set containing predictors (temporal vegetation indices and canopy height measurements) and predicted traits (days to anthesis, silking and yield) belonging to OHOT (optimal planting trial) and DHOT (late planting trial) used in prediction models was given as supplementary data 1 set. The R code used in prediction models was given in Github repository (<https://github.com/alperadak/Supplementary-File-1/blob/main/Supplementary%20file%201.txt>, accessed on 28 May 2021).

**Acknowledgments:** We would like to acknowledge Stephen Labar for technical support and graduate and undergraduate students as well as employees belonging to the Texas A&M Quantitative Genetics and Maize Breeding Program for their hard works and efforts maintaining fields and collecting phenotypic data. We also would like to thank to Texas AgriLife community for their technical support and supplying open field areas. Alper Adak was supported by a fellowship from Ministry of National Education of Turkey and Republic of Turkey Ministry of Agriculture and Forestry.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Krause, M.R.; Mondal, S.; Crossa, J.; Singh, R.P.; Pinto, F.; Haghatalab, A.; Shrestha, S.; Rutkoski, J.; Gore, M.A.; Sorrells, M.E.; et al. Aerial high-throughput phenotyping enables indirect selection for grain yield at the early generation, seed-limited stages in breeding programs. *Crop Sci.* **2020**, *60*, 3096–3114. [[CrossRef](#)]
2. Shi, Y.; Thomasson, J.A.; Murray, S.C.; Pugh, N.A.; Rooney, W.L.; Shafian, S.; Rajan, N.; Rouze, G.; Morgan, C.L.S.; Neely, H.L.; et al. Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLoS ONE* **2016**, *11*, e0159781. [[CrossRef](#)]
3. Anderson, S.L.; Murray, S.C. R/UAStools: Plotshpcreate: Create multi-polygon shapefiles for extraction of research plot scale agriculture remote sensing data. *Front. Plant Sci.* **2020**, *11*, 511768. [[CrossRef](#)] [[PubMed](#)]
4. Matias, F.I.; Caraza-Harter, M.V.; Endelman, J.B. FIELDImageR: An r package to analyze orthomosaic images from agricultural field trials. *Plant Phenome J.* **2020**, *3*, e20005. [[CrossRef](#)]
5. Richardson, A.J.; Wiegand, C.L. Distinguishing vegetation from soil background information. *Photogramm. Eng. Remote Sens.* **1977**, *43*, 1541–1552.
6. Woebbecke, D.M.; Meyer, G.E.; Von Bargen, K.; Mortensen, D.A. Color indices for weed identification under various soil, residue, and lighting conditions. *Trans. ASABE* **1995**, *38*, 259–269. [[CrossRef](#)]
7. Meyer, G.E.; Neto, J.C. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* **2008**, *63*, 282–293. [[CrossRef](#)]
8. Louhaichi, M.; Borman, M.; Johnson, D. Spatially located platform and aerial photography for documentation of grazing impacts on wheat. *Geocarto Int.* **2001**, *16*, 65–70. [[CrossRef](#)]
9. Bendig, J.; Yu, K.; Aasen, H.; Bolten, A.; Bennertz, S.; Broscheit, J.; Gnyp, M.L.; Bareth, G. 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.* **2015**, *39*, 79–87. [[CrossRef](#)]
10. Hamuda, E.; Glavin, M.; Jones, E. A survey of image processing techniques for plant extraction and segmentation in the field. *Comput. Electron. Agric.* **2016**, *125*, 184–199. [[CrossRef](#)]
11. Hunt, E.R., Jr.; Cavigelli, M.; Daughtry, C.S.T.; McMurtrey, J.E., III; Walhall, C.L. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Prec. Agric.* **2005**, *6*, 359–378. [[CrossRef](#)]
12. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
13. Gitelson, A.A.; Kaufman, Y.J.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87. [[CrossRef](#)]
14. Hague, T.; Tillett, N.D.; Wheeler, H. Automated crop and weed monitoring in widely spaced cereals. *Precis. Agric.* **2006**, *7*, 21–32. [[CrossRef](#)]

15. Anderson, S.L.; Murray, S.C.; Malambo, L.; Ratcliff, C.; Popescu, S.; Cope, D.; Chang, A.; Jung, J.; Thomasson, J.A. Prediction of Maize Grain Yield before Maturity Using Improved Temporal Height Estimates of Unmanned Aerial Systems. *Plant Phenome J.* **2019**, *2*, 1–15. [[CrossRef](#)]
16. Anderson, S.L.; Murray, S.C.; Chen, Y.; Malambo, L.; Chang, A.; Popescu, S.; Popescu, S.; Cope, D.; Jung, J. Unoccupied aerial system enabled functional modeling of maize height reveals dynamic expression of loci. *Plant Direct* **2020**, *4*, e00223. [[CrossRef](#)]
17. Adak, A.; Conrad, C.; Chen, Y.; Wilde, S.C.; Murray, S.C.; Anderson, S.; Subramanian, N.K. Validation of Functional Polymorphisms Affecting Maize Plant Height by Unoccupied Aerial Systems (UAS) Discovers Novel Temporal Phenotypes. *Genes Genomes Genet.* **2021**, jkab075. [[CrossRef](#)]
18. Pugh, N.A.; Horne, D.W.; Murray, S.C.; Carvalho, G.; Malambo, L.; Jung, J.; Chang, A.; Maeda, M.; Popescu, S.; Chu, T.; et al. Temporal estimates of crop growth in sorghum and maize breeding enabled by unmanned aerial systems. *Plant Phenome J.* **2018**, *1*. [[CrossRef](#)]
19. Tirado, S.B.; Hirsch, C.N.; Springer, N.M. UAS-based imaging platform for monitoring maize growth throughout development. *Plant Direct* **2020**, *4*, e00230. [[CrossRef](#)]
20. Aguate, F.M.; Trachsel, S.; Pérez, L.G.; Burgueño, J.; Crossa, J.; Balzarini, M.; Gouache, D.; Bogard, M.; De los Campos, G. Use of hyperspectral image data outperforms vegetation indices in prediction of maize yield. *Crop Sci.* **2017**, *57*, 2517–2524. [[CrossRef](#)]
21. Montesinos-López, O.A.; Montesinos-López, A.; Crossa, J.; de los Campos, G.; Alvarado, G.; Suchismita, M.; Rutkoski, J.; González-Pérez, L.; Burgueño, J. Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. *Plant Methods* **2017**, *13*, 4. [[CrossRef](#)] [[PubMed](#)]
22. Maresma, Á.; Ariza, M.; Martínez, E.; Lloveras, J.; Martínez-Casasnovas, J.A. Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (*Zea mays* L.) from a standard UAS service. *Remote Sens.* **2016**, *8*, 973. [[CrossRef](#)]
23. Shanahan, J.F.; Schepers, J.S.; Francis, D.D.; Varvel, G.E.; Wilhelm, W.W.; Tringe, J.M.; Schlemmer, M.R.; Major, D.J. Use of remote-sensing imagery to estimate corn grain yield. *Agron. J.* **2001**, *93*, 583–589. [[CrossRef](#)]
24. García-Martínez, H.; Flores-Magdaleno, H.; Ascencio-Hernández, R.; Khalil-Gardezi, A.; Tijerina-Chávez, L.; Mancilla-Villa, O.R.; Vázquez-Peña, M.A. Corn Grain Yield Estimation from Vegetation Indices, Canopy Cover, Plant Density, and a Neural Network Using Multispectral and RGB Images Acquired with Unmanned Aerial Vehicles. *Agriculture* **2020**, *10*, 277. [[CrossRef](#)]
25. Peng, Y.; Gitelson, A.A. Application of chlorophyll-related vegetation indices for remote estimation of maize productivity. *Agric. For. Meteorol.* **2011**, *151*, 1267–1276. [[CrossRef](#)]
26. Wu, G.; Miller, N.D.; De Leon, N.; Kaepller, S.M.; Spalding, E.P. Predicting Zea mays flowering time, yield, and kernel dimensions by analyzing aerial images. *Front. Plant Sci.* **2019**, *10*, 1251. [[CrossRef](#)] [[PubMed](#)]
27. Li, X.; Guo, T.; Mu, Q.; Li, X.; Yu, J. Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6679–6684. [[CrossRef](#)]
28. Gage, J.L.; Jarquin, D.; Romay, C.; Lorenz, A.; Buckler, E.S.; Kaepller, S.; Alkhailafah, N.; Bohn, M.; Campbell, D.A.; Edwards, J.; et al. The effect of artificial selection on phenotypic plasticity in maize. *Nat. Commun.* **2017**, *8*, 1–11. [[CrossRef](#)]
29. Rogers, A.R.; Dunne, J.C.; Romay, C.; Bohn, M.; Buckler, E.S.; Ciampitti, I.A.; Edwards, J.; Ertl, D.; Flint-Garcia, S.; Gore, M.C.; et al. The Importance of Dominance and Genotype-by-Environment Interactions on Grain Yield Variation in a Large-Scale Public Cooperative Maize Experiment. *Genes Genomes Genet.* **2021**, *11*, jkaa050.
30. Araus, J.L.; Cairns, J.E. Field high-throughput phenotyping: The new crop breeding frontier. *Trends Plant Sci.* **2014**, *19*, 52–61. [[CrossRef](#)] [[PubMed](#)]
31. Adak, A.; Murray, S.C.; Anderson, S.L.; Popescu, S.C.; Malambo, L.; Romay, M.C.; de Leon, N. Unoccupied aerial system (UAS) discovered overlooked loci capturing the variation of entire growing period in maize. *Plant Genome* **2021**. [[CrossRef](#)] [[PubMed](#)]
32. Adak, A.; Murray, S.C.; Anderson, S.L.; Popescu, S.C.; Lonesome, M.; Dale, C. Discovery of temporal loci controlling segregation of vegetation Indices through maize hybrid growth. In Proceedings of the 63rd Annual Maize Genetics Meeting, Virtual. 8–12 March 2021; p. 89.
33. Pauli, D.; Andrade-Sánchez, P.; Carmo-Silva, A.E.; Gazave, E.; French, A.N.; Heun, J.; Hunsaker, D.J.; Lipka, A.E.; Setter, T.L.; Strand, R.J.; et al. Field-based high-throughput plant phenotyping reveals the temporal patterns of quantitative trait loci associated with stress-responsive traits in cotton. *Genes Genomes Genet.* **2016**, *6*, 865–879. [[CrossRef](#)] [[PubMed](#)]
34. Singh, D.; Wang, X.; Kumar, U.; Gao, L.; Noor, M.; Imtiaz, M.; Singh, R.P.; Poland, J. High-throughput phenotyping enabled genetic dissection of crop lodging in wheat. *Front. Plant Sci.* **2019**, *10*, 394. [[CrossRef](#)]
35. Miao, C.; Xu, Y.; Liu, S.; Schnable, P.S.; Schnable, J.C. Increased power and accuracy of causal locus identification in time series genome-wide association in sorghum. *Plant Physiol.* **2020**, *183*, 1898–1909. [[CrossRef](#)]
36. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: New York, NY, USA, 2013; Volume 112.
37. Escadafal, R. Remote sensing of soil color: Principles and applications. *Remote Sens. Rev.* **1989**, *7*, 261–279. [[CrossRef](#)]
38. Zarco-Tejada, P.; Berjon, A.; Lopezlozano, R.; Miller, J.; Martin, P.; Cachorro, V.; Gonzalez, M.; Defrutos, A. Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sens. Environ.* **2005**, *99*, 271–287. [[CrossRef](#)]
39. Girardeau-Montaut, D. CloudCompare. Version 2.8. 2016. Available online: <https://www.danielgm.net/cc/> (accessed on 21 May 2021).
40. McGaughey, R. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*. Version 3.60+; US Forest Service Pacific Northwest Research Station: Corvallis, OR, USA, 2016.

41. Isenburg, M. *LASTools: Efficient Tools for LiDAR Processing*. Version 130506; Department of Computer Science, University of North Carolina: Chapel Hill, NC, USA, 2015.
42. Rapidlasso. *LASTools: Efficient LiDAR Processing Software*. Version 170628; Rapidlasso GmbH: Gilching, Germany, 2017; Available online: <http://rapidlasso.com/LASTools> (accessed on 21 May 2021).
43. Kraus, K.; Pfeifer, N. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS J. Photogramm. Remote Sens.* **1998**, *53*, 193–203. [[CrossRef](#)]
44. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. *arXiv* **2014**, arXiv:1406.5823.
45. Smith, M.R.; Rao, I.M.; Merchant, A. Source-sink relationships in crop plants and their influence on yield development and nutritional quality. *Front. Plant Sci.* **2018**, *9*, 1889. [[CrossRef](#)]
46. Zaman-Allah, M.; Vergara, O.; Araus, J.L.; Tarekegne, A.; Magorokosho, C.; Zarco-Tejada, P.J.; Hornero, A.; Albà, A.H.; Das, B.; Craufurd, P.; et al. Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* **2015**, *11*, 35. [[CrossRef](#)] [[PubMed](#)]
47. Lee, E.A.; Tollenaar, M. Physiological basis of successful breeding strategies for maize grain yield. *Crop Sci.* **2007**, *47*, 202–215. [[CrossRef](#)]
48. Zhang, X.; Zhang, F.; Qi, Y.; Deng, L.; Wang, X.; Yang, S. New research methods for vegetation information extraction based on visible light remote sensing images from an unmanned aerial vehicle (UAV). *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *78*, 215–226. [[CrossRef](#)]
49. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *J. Sens.* **2017**, *2017*, 1–17. [[CrossRef](#)]
50. Rincent, R.; Charpentier, J.P.; Faivre-Rampant, P.; Paux, E.; Le Gouis, J.; Bastien, C.; Segura, V. Phenomic selection is a low-cost and high-throughput method based on indirect predictions: Proof of concept on wheat and poplar. *Genes Genomes Genet.* **2018**, *8*, 3961–3972. [[CrossRef](#)]
51. Weber, V.S.; Araus, J.L.; Cairns, J.E.; Sanchez, C.; Melchinger, A.E.; Orsini, E. Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes. *Field Crops Res.* **2012**, *128*, 82–90. [[CrossRef](#)]
52. Endelman, J.B. Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Gen.* **2011**, *4*, 250–255. [[CrossRef](#)]
53. Hernandez, J.; Lobos, G.A.; Matus, I.; Del Pozo, A.; Silva, P.; Galleguillos, M. 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.* **2015**, *7*, 2109–2126. [[CrossRef](#)]
54. Lane, H.M.; Murray, S.C.; Montesinos-López, O.A.; Montesinos-López, A.; Crossa, J.; Rooney, D.K.; Barrero-Farfán, I.D.; De La Fuente, G.N.; Morgan, C.L.S. Phenomic selection and prediction of maize grain yield from near-infrared reflectance spectroscopy of kernels. *Plant Phenome J.* **2020**, *3*, e20002. [[CrossRef](#)]
55. Rincent, R.; Laloë, D.; Nicolas, S.; Altmann, T.; Brunel, D.; Revilla, P.; Rodriguez, V.M.; Moreno-Gonzales, J.; Melchinger, A.; Bauer, E.; et al. Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: Comparison of methods in two diverse groups of maize inbreds (*Zea mays* L.). *Genetics* **2012**, *192*, 715–728. [[CrossRef](#)]
56. Peng, B.; Guan, K.Y.; Zhou, W.; Jiang, C.Y.; Frankenberg, C.; Sun, Y.; He, L.Y.; Köhler, P. Assessing the benefit of satellite-based Solar-Induced Chlorophyll Fluorescence in crop yield prediction. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *90*, 102126. [[CrossRef](#)]
57. Peng, B.; Guan, K.; Pan, M.; Li, Y. Benefits of seasonal climate prediction and satellite data for forecasting US maize yield. *Geophys. Res. Lett.* **2018**, *45*, 9662–9671. [[CrossRef](#)]