Published July 11, 2019

The Plant Phenome Journal

Original Research

Prediction of Maize Grain Yield before Maturity Using Improved Temporal Height Estimates of Unmanned Aerial Systems

Steven L. Anderson II, Seth C. Murray,* Lonesome Malambo, Colby Ratcliff, Sorin Popescu, Dale Cope, Anjin Chang, Jinha Jung, and J. Alex Thomasson

Core Ideas

- UAS captured increased genetic variation compared with manual terminal height.
- There were small significant differences in ground filtering methods to extract plant structure.
- Higher resolution did not improve imagery informativeness with regard to plant height.
- Logistic function provides informative phenotypes for temporal maize growth.
- Correlation and prediction accuracy of grain yield increased by ~20% with UAS heights.

Weekly unmanned aerial system (UAS) imagery was collected over the College Station, TX, 2017 Genomes to Fields (G2F) hybrid trial, across three environmental stress treatments, using two UAS platforms. The high-altitude (120-m) fixed-wing platform increased the fraction of variation attributed to genetics and had highly repeatable (R > 60%) height estimates, increasing the genetic variance explained (10–40%) over traditional terminal plant height measurement (PHT_{TRML} ~30%), as well as over the low-altitude rotary-wing UAS platform (10–20%). A logistic function reduced the dimensionality (>20 flights) of each UAS dataset to three parameters (inflection point, growth rate, and asymptote) and produced a more robust predictive model than independent flight dates, effectively summarizing (R^2 > 0.98) the UAS flight dates. The logistic model overcame the need to use specific flight dates when comparing different environments. The UAS height estimates (r = 0.36–0.48) doubled the correlations to grain yield in this G2F experiment compared with PHT_{TRML} (r = 0.23–0.28). Parameters of the logistical function achieved equivalent correlations (r = 0.30–0.46) to individual flight dates (r = 0.36–0.48), improving grain yield prediction by ~400% (R^2 = 0.25–0.34) over PHT_{TRML} (R^2 = 0.06–0.08). Incorporating other UAS-derived parameters beyond plant height may allow yield to be accurately predicted before maturity, speeding breeding programs. A new public R function to generate ESRI shapefiles for plot research is also described.

S.L. Anderson, S.C. Murray, and C. Ratcliff, Dep. of Soil and Crop Sciences, Texas A&M Univ., College Station, TX 77843-2474; L. Malambo and S. Popescu, Dep. of Ecosystem Science and Management, Texas A&M Univ., College Station, TX 77843-2120; D. Cope, Dep. of Mechanical Engineering, Texas A&M Univ., College Station, TX 77843; A. Chang and J. Jung, School of Engineering and Computer Sciences, Texas A&M Univ., Corpus Christi, TX 78412; J.A. Thomasson, Dep. of Biological and Agricultural Engineering, Texas A&M Univ., College Station, TX 77843.

Plant Phenome J. 2:190004 (2019)
doi:10.2135/tppj2019.02.0004

Received 15 Feb. 2019.
Accepted 19 Apr. 2019.
Supplemental material online.
*Corresponding author (sethmurray@tamu.edu).

Genetic variation of terminal plant height (PHT_{TRML}) in maize (*Zea mays* L.) is a highly heritable trait (Anderson et al., 2018; Li et al., 2016b; Mahan et al., 2018; Peiffer et al., 2014; Wallace et al., 2016) and is relatively easy to phenotype, for instance measuring from the ground to the tip of a tassel on a representative plant. However, the labor and time required to collect data is still resource constrained, and height measurements collected in maize research programs are generally taken only once, when the plants have reached maximum growth after the completion of flowering.

Plant height is valuable not only as a phenotype in and of itself, but it has also been shown to be predictive of maize grain yield in some regions and environments (Katsvairo et al., 2003; Machado et al., 2002; Mallarino et al., 1999; Yin et al., 2011). Barrero Farfan et al. (2013) observed positive correlations (r = 0.46) between PHT_{TRML} and grain yield within the semiarid stressed environment of Texas, less correlation (r = 0.19) in the irrigated High Plains, and the highest correlations by combining all Texas environments (r = 0.61). Yin et al. (2011) demonstrated that V10/V12 plant height was highly predictive (R^2 = 0.26–0.87) of grain yield using an exponential regression model. Previous work has shown that early season plant height can be decoupled from PHT_{TRML} and has

Abbreviations: 3D, three-dimensional; ATIN, adaptive triangulated irregular network; BLUP, best linear unbiased predictor; CSF, cloth simulation filter; DAS, days after sowing; DBM, difference-based method; DG2F, unirrigated, optimal planting Genomes to Field trial; DTM, digital terrain model; DSM, digital surface model; G2F, Genomes to Fields; G2FE, irrigated, optimal planting Genomes to Field trial; G2LA, irrigated, delayed planting Genomes to Field trial; HCDH, high canopy density hybrid study site; HRI, hierarchical robust interpolation; LCDH, low canopy density hybrid study site; MAE, mean absolute error; MCDI, medium canopy density inbred study site; PHT_{TRML}, manually measured terminal plant height; RMSE, root mean squared error; SfM, structure from motion; UAS, unmanned aerial system; UAV, unmanned aerial vehicle.

been hypothesized to offer additional insight into yield (Mallarino et al., 1999; Pugh et al., 2018). The relative ease of plant height measurements via remote sensing in the field (Chang et al., 2017; Chu et al., 2018; Han et al., 2018; Malambo et al., 2018) and the potential to predict yield at earlier time points (i.e., before harvest) makes plant height an excellent case study for phenotypic data collection via unmanned aerial systems.

Unmanned aerial systems (UAS) include unmanned aerial vehicles (UAV) that have been equipped with lidar sensors to generate dense, three-dimensional (3D) point clouds (Wallace et al., 2012) or, more commonly, digital RGB or multispectral cameras (Araus and Kefauver, 2018; Hunt and Daughtry, 2017; Sankaran et al., 2015) to collect high-resolution images and 3D point clouds through post-processing of image sets. Specifically, point clouds have been used to estimate aboveground heights of objects and vegetation. Aerial laser scanning technology (i.e., lidar) has been a major source of 3D data sets via manned aerial vehicles but is very expensive. New innovations including low-cost UAS (Reynolds et al., 2018; Sankaran et al., 2015; Shi et al., 2016), optimized image matching software, and graphical processing units have reduced the inefficiency of image-based photogrammetry methods (3D vision) that existed in previous decades (Baltsavias, 1999). The cost and difficulty of lidar–UAS system integration (Wallace et al., 2012) has led to broad adoption of multispectral- and RGB-UAS systems (3D vision UAS) to easily and rapidly produce temporal 3D datasets in agriculture using structure-from-motion (SfM) photogrammetry (Burkart et al., 2018; Holman et al., 2016; Malambo et al., 2018; Pugh et al., 2018; Xavier et al., 2017).

Ground Filtering and Separation Approaches

A critical step in estimating aboveground heights from UAS is the identification of ground points and accurate reconstruction of the digital terrain models (DTMs) to produce digital surface models (DSMs) from the digital elevation model. Ground filtering algorithms have been developed to delineate between points belonging to ground and non-ground classes and have been extensively reviewed, but the field has been dominated by lidar efforts in regard to urban and forested terrains (Chen et al., 2017; Korzeniowska et al., 2014; Meng et al., 2010; Pfeifer and Mandlburger, 2009; Polat and Uysal, 2015; Serifoglu Yilmaz and Gungor, 2016; Sithole and Vosselman, 2004; Weed et al., 2002) with little focus on agricultural cropland. Among comparative ground filter studies, Montealegre et al. (2015) specifically discussed areas covered with cereal crops, concluding that an adaptive triangulated irregular network (ATIN) (Axelsson, 2000) resulted in the most accurate modeling of the terrain within crop- and grassland-dominated study areas. Crop heights are commonly estimated using the difference-based method (DBM) in which DTMs are modeled by pre- or post-season bare earth images (Bareth et al., 2016; Bendig et al., 2013; Chu et al., 2018; Watanabe et al., 2017). Alternatively, the point cloud method identifies ground points within each DSM and creates an independent DTM for each dataset (Malambo et al., 2018; Pugh et al., 2018). Holman et al. (2016) demonstrated that the point cloud method produces reduced root mean square error (RMSE) compared with the DBM due to biased ground representation of the pre- or post-flight ground model. Correct ground modeling is essential to improving estimation accuracy, so further studies are necessary to evaluate the most effective technique to model the terrain specific to using SfM photogrammetry from high-resolution UAS images of a breeding or genetic field trial.

Estimating Maize Height via UAS

Common trends have been demonstrated in past UAS field studies using 3D vision SfM height. Statistical metrics of UAS point clouds have been shown to be significantly correlated to manual phenotyping and lidar datasets in maize (Chu et al., 2018; Hu et al., 2018; Li et al., 2016a; Malambo et al., 2018; Niu et al., 2018; Pugh et al., 2018; Shi et al., 2016; Varela et al., 2017). In many of these studies, plot-level point clouds were divided into quartiles, with the 99th (P99) percentile including the top of the plant and the bottom 1% (P01) representing the soil and aboveground roots. Niu et al. (2018) demonstrated that the use of a higher quantile percentage reduced bias and RMSE in reference to lidar data. Similarly, UAS-derived heights at the higher percentiles commonly found at P95 and P99 in maize, but excluding the P100 and maximum, have shown the greatest correlation to manual plant height measurements and least RMSE (Chu et al., 2018; Malambo et al., 2018; Pugh et al., 2018; van der Voort, 2016). The UAS-derived height estimates are highly repeatable ($R = 0.91–0.99$ for P95) and capable of capturing genotypic variation equivalent to manual height measurements, especially at later dates (>50 d after sowing [DAS]) in the growing season when greater variability is expressed across genotypes (Pugh et al., 2018).

Using high-throughput technologies such as UAS and ground vehicles is rapidly becoming commonplace in agriculture and breeding programs. The majority of the published research has been focused on validation of UAS measurements to manual phenotyping, and it is evident in the literature that UAS-derived phenotypes provide highly accurate measurements, highly correlated to manual phenotyping data. The focus of this study was to expand beyond validation of UAS-estimated heights toward using the data as it is presented (i.e., without validation). The objectives of this study were to: (i) compare multiple methods of ground point filtering for DSM accuracy, (ii) identify sources of variation across UAS platforms and environmental treatments throughout the growing season, (iii) apply nonlinear modeling approaches to identify critical flight dates and capture new growth parameters, and (iv) evaluate UAS height estimates and nonlinear modeling parameters for their ability predict grain yield in maize. To conduct this work efficiently, an improved method was needed to increase the speed of extracting plot information of large studies in the UAS-to-knowledge pipeline; a novel plot boundary delineation function to generate plot boundary ESRI shapefiles automatically given two boundary coordinates, the experimental design, and plot dimensions of the breeding field are also described here.

Materials and Methods

Germplasm Material and Experimental Design

The Genomes to Fields (G2F) initiative (https://www.genomes2fields.org/) is a multidisciplinary umbrella initiative aimed at understanding the genotype × environment interaction of the maize genome (AlKhalifah et al., 2018; Gage et al., 2017). As of 2018 and beginning in 2014, the G2F collaborators have evaluated more than 94,000 field plots involving more than 1700 hybrids across 77 unique environments in 23 states and provinces in the United States and Canada. For this study, the 2017 G2F trials were evaluated and imaged via UAS in College Station, TX. This trial was comprised of 280 unique hybrids, with 230 common hybrids across three different management environments: irrigated, optimal planting (G2FE); unirrigated, optimal planting (DG2F); and irrigated, delayed (∼30 d) planting (G2LA). Each trial was arranged in a randomized complete block design (two replicates per trial) with two-row plots, 0.76-m row spacing, and 7.62-m plot lengths. The three trials were planted adjacent to each other in a single field of approximately ∼1.4 ha.

Ground Truth Measurements

Manual height measurements were collected on several dates (Supplemental Table S1) throughout the growing season to assess the accuracy of the UAS height estimates. Two heights were taken during manual phenotyping: (i) the apex height, which was either the erect emerging leaf (pre-flowering) or the tip of the tassel, and (ii) the flat plane of the plot during vegetative growth or the flag leaf height during reproductive stages (Supplemental Fig. S1A). Furthermore, manual terminal height measurements were taken on all plots at the tip of tassel height (PHT_{TRML}) after flowering was completed. Manual measurements were collected as a visual plot average by measuring a single representative plant.

UAS Image Collection

Two platforms were used, a rotary-wing and a fixed-wing UAV. The rotary-wing model, a DJI Phantom 3 Professional with a 12 megapixel DJI FC300X camera, was flown at an altitude of 25 m above the ground surface with an 80% forward and side image overlap. Fixed-wing images were collected with a Tuffwing UAV Mapper (http://www.tuffwing.com) equipped with a 24 megapixel Sony a6000 RGB camera. Fixed-wing surveys were conducted at a 120-m altitude with 80% image overlap. The goal was to collect weekly UAS imagery throughout the early growing season and transition to biweekly flights at a 3- to 4-d interval during the exponential growth phase based on observations from earlier studies (Malambo et al., 2018; Pugh et al., 2018). Twenty-two and 19 flights were completed throughout the 2017 growing season by the rotary-wing and fixed-wing platforms, respectively (Supplemental Table S1).

Image Processing

All UAS images were processed using SfM photogrammetry algorithms in either the Pix4Dmapper (https://www.pix4d.com) or Agisoft PhotoScan Professional (AgiSoft, 2016) software. In general, these software packages are equivalent and used to identify common features (tie points) across images followed by triangulation and distortion adjustment optimization to generate densified 3D point clouds, DSM, and orthomosaic images. Due to the large collaborative effort of this project, the preference of the software was based on each group's (fixed wing or rotary wing) capability and familiarity. Ground control points were placed throughout the study sites to ensure correct scale, orientation, and geographic location of generated outputs. All of the fixed-wing flights were processed in Agisoft PhotoScan, while the majority of the rotary-wing flights were processed in Pix4Dmapper (excluding flights on 14 and 27 July 2017). Issues with image matching and tie point identification during stages of canopy closure resulted in large "black holes" within the center of some rotary-wing flight image mosaics. In an attempt to resolve the holes of missing data, Agisoft Photoscan was used in those mosaics with holes and resulted in improved data for some dates (14 and 27 July 2017). Where Agisoft Photoscan did not improve the data quality, manual tie point assignment was performed. All raw and processed image output files from this study are publicly available at Cyverse (Murray et al., 2019).

Data Extraction Pipeline

Following the initial processing of raw images into point clouds, a novel processing pipeline was developed to acquire plot-based height estimates from the point clouds (Fig. 1a). R/UAStools::plotshpcreate (https://github.com/andersst91/UAStools) was developed to construct ESRI shapefiles (.shp) of individual research plots for subsequent plot extraction (Fig. 1b). The initial assignment of these plots is based on the GPS coordinates of an AB line representing the bottom left corner of the first plot (A) and the top left corner of the trial within the same row as the A point (B). Using a data frame containing the experimental design, plot dimensions, and unique plot IDs (i.e., a research "field book"), the script produces an ESRI shapefile that contains all of the plot boundaries necessary to extract plot-level measurements. However, we have found that some manual adjustment is needed when the shapefile is visually overlaid on the mosaics due to subtle variances in tractor rows (even when GPS guided) and in the orthomosaics that are exaggerated when overlaying a precise rectangular grid.

The point clouds were first clipped to the trial level, and large blunders (i.e., serendipitous point anomalies above or below the point cloud) were manually removed using the segment tool in CloudCompare v2.10 (Girardeau-Montaut, 2016). Following manual blunder removal, a custom batch script was run including executable functions from LAStools (Isenburg, 2015; rapidlasso, 2017) and FUSION/LDV (McGaughey, 2016) software (https://github.com/andersst91/UAS_Height_Pipeline). In brief, the pipeline (i) sorted data points (LAStools\lasssort.exe) to improve processing efficiency, (ii) removed additional blunders (LAStools\lasnoise.exe) closer to the canopy structure, (iii) executed a ground filtering algorithm (FUSION\GroundFilter.exe)

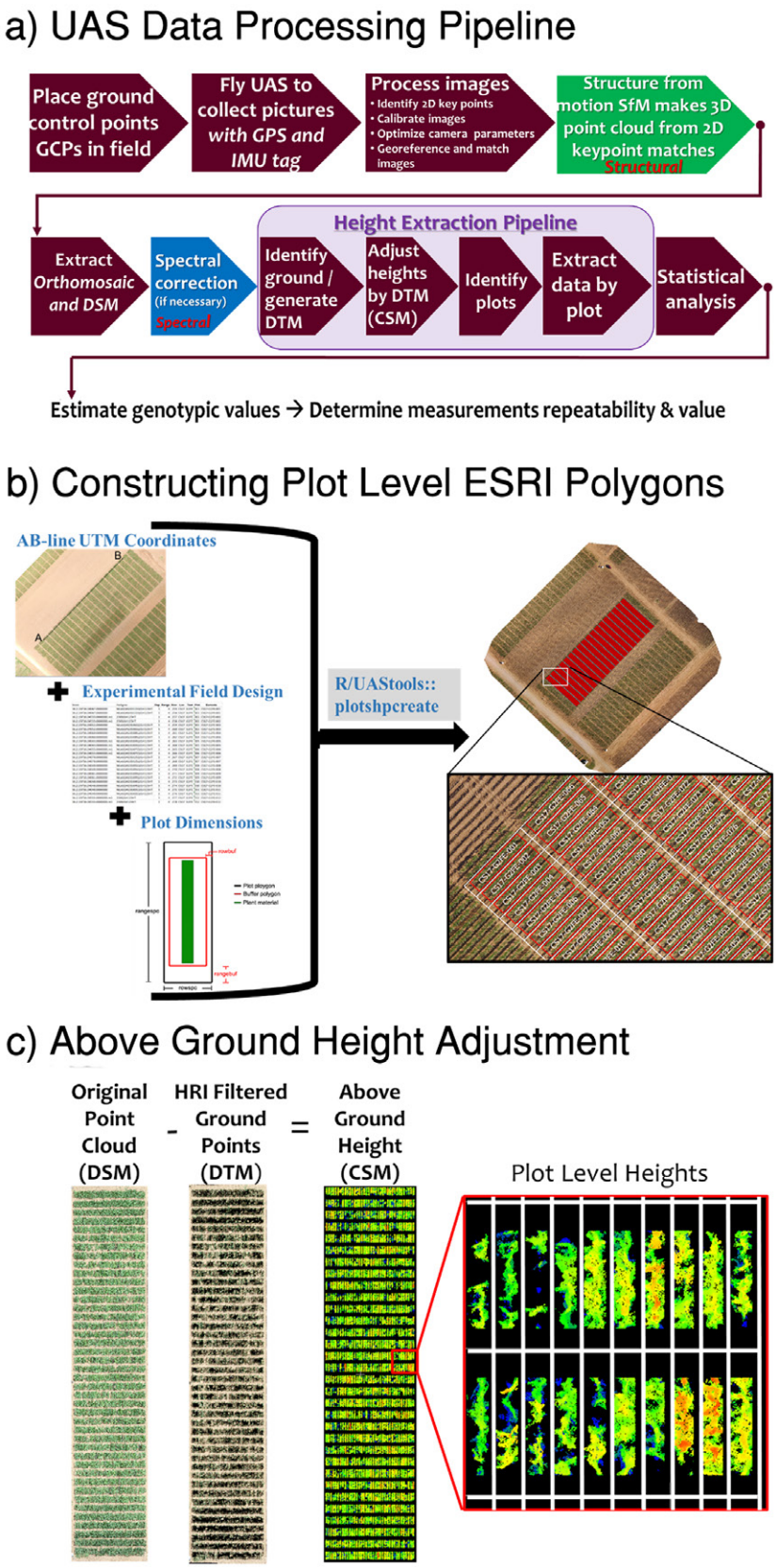

Fig. 1. (a) Flow chart depicting unmanned aerial system (UAS) data curation pipeline from image acquisition to statistical analysis of phenotype estimates; (b) graphical representation of the R/UAStools::plotshpcreate inputs and plot-level polygon ESRI shapefile output; and (c) visual conversion of a digital surface model (DSM) to aboveground canopy surface models (CSM) using digital terrain modeling (DTM) via hierarchical robust interpolation (HRI).

to identify ground points, (iv) identified key points (LAStools\lasthin.exe) on the vertex of the hills from the ground filter for DTM modeling, and (v) interpolated or constructed the DTM from the key points (FUSION\GridSurfaceCreate.exe). Following the DTM construction, the noise-filtered point cloud (step ii output) was adjusted to aboveground height using the DTM (LAStools\lasheight.exe), and points below the digital elevation model (i.e., the ground) were removed in order to not bias the height estimates with negative values (Fig. 1c). Using the adjusted "Z" point cloud, the plot-level ESRI shapefile was used to clip individual plot point clouds (FUSION/PolyClipData.exe) and calculate statistical metrics within each of the plots (FUSION/CloudMetrics.exe), including estimating height from the point clouds. Further data compiling and processing was conducted in R version 3.3.1 (R Core Team, 2016). All manually collected and extracted phenotypes from this study are publicly available (Anderson et al., 2019).

Comparison of Terrain Modeling Methods

Study Areas

Three flight dates were chosen for the purpose of comparing ground filtering algorithms based on available manual height measurements and the collection of images from both UAS platforms (Supplemental Fig. S2). The first site (G2LA 9 May 2017) was characterized with low canopy density, high ground point representation, and young (33 DAS) hybrid maize plants: the low canopy density hybrid (LCDH) site. The second site (G2LA 2 June 2017) was characterized by full canopy closure, minimal ground point representation, and mature (57 DAS) hybrid maize plants: the high canopy density hybrid (HCDH) site. The third site (YYCP 24 May 2017) was characterized by medium canopy density, medium ground point representation, and young vegetative (61 DAS) inbred maize plants: the medium canopy density inbred (MCDI) site. The MCDI study site was separate from G2F, consisting of 533 plots from three biparental recombinant inbred line mapping populations; these plots provide a useful contrast to address other common research needs (e.g., new line development, quantitative trait loci mapping, and trait discovery).

Terrain Modeling Methods

The DBM of terrain modeling was compared with more advanced point cloud methods commonly used with lidar data. The DBM relies on a preseason (i.e., pre-plant) or postseason flight of the

bare ground to model the terrain. This terrain model is then subtracted from the DSMs in season to obtain a crop surface model (Supplemental Fig. S1B). The point cloud based methods are algorithms that work iteratively through point clouds of each flight and identify ground points based on classification tuning parameters that are set by the user. Three point cloud methods were selected for evaluation including hierarchical robust interpolation (HRI) (Kraus and Pfeifer, 1998), cloth simulation filter (CSF) (Zhang et al., 2016), and ATIN (Axelsson, 2000) based on (i) open access software, (ii) computational efficiency, and (iii) accuracy performance as indicated in the literature. Optimized filter parameters were identified through minimization of the RMSE and mean absolute error (MAE) between UAS height estimates and manual ground truth measurements taken the same day as the UAS surveys. Optimized algorithm parameters were then used to compare ground filtering methods across UAS platforms and study sites. Details on the point cloud based methodology and optimized filtering parameters can be found in Supplement 3.

Statistical Inference

Variance Component Estimates

From the extracted point cloud derived canopy height metrics (P90, P95, P99, and Max), we fit mixed linear models using restricted maximum likelihood in JMP version 14.0.0 (SAS Institute, 2018) to define the best linear unbiased predictors (BLUPs) of the hybrids by their pedigree. Models were fit on a per flight date basis by UAS platform. The individual G2F trials were evaluated as a randomized complete block design including spatial regression (range and row [what furrow irrigation runs down], also called row and column, respectively, where furrow irrigation is not used):

$$Y = \mu + \sigma^2_G + \sigma^2_r + \sigma^2_i + \sigma^2_j + \sigma^2_\varepsilon \quad [1]$$

with terms genotype ($\sigma_G{}^2$), replicate ($\sigma_r{}^2$), range ($\sigma_i{}^2$), row ($\sigma_j{}^2$), and residual error ($\sigma_\varepsilon{}^2$). By flying all three trials within the same flight dates, we were able to evaluate the variance components of UAS plant height as a multi-environment randomized complete block design:

$$Y = \mu + \sigma^2_G + \sigma^2_E + \sigma^2_{G\times E} + \sigma^2_{E(r)} + \sigma^2_{E(i)} + \sigma^2_{E(j)} + \sigma^2_\varepsilon \quad [2]$$

with terms genotype ($\sigma_G{}^2$), environment ($\sigma_E{}^2$), genotype × environment interaction ($\sigma_{G\times E}{}^2$), replicate nested within environment ($\sigma_{E(r)}{}^2$), range nested within environment ($\sigma_{E(i)}{}^2$), and row nested within environment ($\sigma_{E(j)}{}^2$).

Repeatability

Repeatability (R) estimates represent the percentage of genetic variation explained by the data compared with the experimental variation explained excluding identifiable environmental effects. Repeatability was calculated on an entry means basis similar to broad-sense heritability (H^2) with the key differentiation of presence (H^2) or absence (R) of familial structure. Within-environment repeatability estimates were calculated on single environments with the number of replicates (r):

$$R = \frac{\sigma^2_G}{\sigma^2_G + \sigma^2_\varepsilon / r} \quad [3]$$

Multi-environment repeatability was calculated by expanding Eq. [3] to include the entry × environment interaction variation component (and the number of environments, E):

$$R = \frac{\sigma^2_G}{\sigma^2_G + \sigma^2_{G\times E} / E + \sigma^2_\varepsilon / rE} \quad [4]$$

Nonlinear Logistic Function

Implementation of nonlinear modeling was assessed to further reduce the dimensionality of the dataset of multiple flights throughout the growing season. Maize being an annual crop, we assumed that plant height should follow an asymptotic model that begins with zero at planting and concludes its lifespan with a terminal growth parameter (Archontoulis and Miguez, 2015). The three-parameter logistic model best followed these assumptions:

$$f(x) = \frac{L}{1 + \exp\left[-k\left(x - x_0\right)\right]} \quad [5]$$

where modeling height is a function of DAS (x) with the asymptote (L, m), inflection point (x_0, DAS), and the growth rate (k, DAS^{-1}) of the fitted curve (Verhulst, 1838). The asymptote is the maximum value of the curve, which represents terminal PHT. The inflection point indicates the DAS when the rate of growth is maximized. The growth rate parameter defines the steepness of the logistic curve. Logistic curves were fit using the Fit Curve tool in JMP 14 (Analyze → Specialized Modeling → Fit Curve) and parameters were estimated on UAS height estimates on a plot basis as well as on a pedigree basis using the BLUPs of the individual environment restricted maximum likelihood models (Eq. [1]). Significance of the logistic parameters was evaluated using the chi squared (χ^2) test ($\alpha = 0.05$, df = 1) to identify logistical curves with poor fits to UAS height estimates. Plots with nonsignificant parameter fits were excluded in further analysis because the logistical function would not accurately represent that plot's or pedigree's growth model.

Stepwise Regression of Predictive Models

Forward and reverse stepwise regression were performed in JMP 14 using the Fit Model function to identify the most predictive UAS height parameters with respect to grain yield (t ha^{-1}). Parameters identified by the stepwise regression procedure were then fit as continuous effects in a linear model to assess their ability to predict yield based on their coefficient of determination (R^2) and RMSE. The parameters tested for each UAS platform included three sets of predictors: (i) the logistic parameters, (ii) pedigree BLUPs by flight date, and (iii) the combination of logistic parameter and pedigree BLUPs by flight date. Predictors were removed if they were not significant in the fit model. Due to the time series nature of our dataset, collinearity between the predictor variables was evaluated using the variance inflation factor (VIF).

The VIF [$1/(1 - R^2)$] cutoff was set to VIF ≤ 4.0, and the variable that caused the least reduction in R^2 of the model was removed.

Results and Discussion

Extraction of informative UAS height data from SfM photogrammetry point clouds required the optimization of terrain modeling and selection of the optimal point cloud metric to be implemented within the data extraction pipeline. We first optimized the terrain modeling procedure through the comparison of four ground modeling methods (HRI, ATIN, CSF, and DBM) across three survey sites varying in canopy structure (LCDH, MCDH, and HCDH) and two UAV platforms (rotary wing and fixed wing). Based on these results, comparison across UAS platforms at each flight date were made using HRI to identify sources of variation throughout the growing season. Following the comparison of UAS platforms by flight date, nonlinear logistic functions were fit to identify critical flight dates and capture new growth parameters. Finally, we evaluated UAS height estimates and nonlinear modeling parameters for their ability predict maize grain yield.

Optimizing Terrain Modeling and Point Cloud Metric

A subset of three flight dates (LCDH: G2LA 9 May 2017; HCDH: G2LA 2 June 2017; and MCDI: YYCP 24 May 2017) were chosen to evaluate terrain modeling methods and point cloud metric comparisons across different maize canopy structures to optimize the data extraction pipeline prior to processing the complete season datasets. Selection of the three flight dates was based on the availability of manual height measurements while maintaining high qualitative appearance (i.e., minimal noise) from both UAS platforms all on the same date (Pugh et al., 2018). Comparisons were made between the DBM and three point cloud methods (HRI, ATIN, and CSF) to identify the optimal terrain modeling method to be implemented within the data extraction pipeline. Further comparisons were made between four point cloud metrics (P90, P95, P99, and Max) to identify the most informative metric based on RMSE, MAE, percentage of genetic variance explained, and repeatability.

Accuracy of Ground Filtering Methods vs. Ground Truth Measurements

Across both UAS platforms and canopy structures, all of the algorithms performed similarly (Fig. 2a) when their parameters were optimized (Sithole and Vosselman, 2004), probably due to the relatively flat plane of the study sight (irrigation furrows notwithstanding) compared with the more varied natural terrain these algorithms were designed around. Across both UAS platforms, halving the resolution (fixed wing $\sim$2 cm pixel^{-1}; rotary wing $\sim$1 cm pixel^{-1}; Supplemental Table S1) via fixed-wing flights had a noticeable impact on the MAE (FW_{P95} 19–40 cm; RW_{P95} 10–21 cm) of the height estimates compared with ground truth (Fig. 2a). The fixed wing achieved its best MAE to ground truth across the canopy structures 16, 20, and 11 cm within LDCH-DBM-Max, MDCI-HRI-Max, and HCDH-HRI-Max, respectively. The rotary wing achieved its best MAE to ground truth across the canopy structures within 6, 8, and 10 cm for LDCH-CSF-P99, MDCI-DBM-Max, and HCDH-HRI-P95, respectively. Within the LCDH and MCDI sites, the MAE of the fixed wing ranged from $\sim$18 to 45 cm, whereas the rotary wing ranged from $\sim$8 to 25 cm, depending on the filter method and metric.

These results demonstrated that sparse canopy structure (e.g., tassels, young plants) were better captured by the low-altitude rotary wing rather than the fixed wing. We hypothesize that the reduced resolution results in triangulating pixels (i.e., smoothing of the canopy structure) at lower elevations in the canopy as well as failure to capture less dense features at the canopy apex (e.g., tassel, erect leaf, flag leaf). In general, plant height was consistently underestimated by UAS estimates from the high-altitude fixed wing, and overall accuracy improved with increased crop maturity. In contrast, the low-altitude rotary wing transitioned from underestimated to overestimated heights as the metric percentile was increased, indicating increased error blunders above the canopy surface within the rotary-wing datasets (Supplemental Fig. S3).

Genetic Variation and Repeatability of Terrain Model Comparisons

Absolute accuracy relative to traditional manual measurements as tested above is important to validate plant height estimates. However, plant breeders (focused on selecting the best cultivar) and geneticists (focused on distributions for mapping) can sufficiently use and are more interested in relative rankings, genetic variation captured, and repeatability across germplasm. Genetic variation (σ^2_G) explained and repeatability (R) are two metrics that have been used to compare the precision of different point cloud percentiles (Pugh et al., 2018) but can also be used to compare the precision of different UAS platforms, different canopy structures, and different ground filtering algorithms. Overall, both genetic variation and repeatability showed similar results among each factor individually (UAS platforms, canopy densities, ground filtering algorithms, and point cloud percentiles) when looking at only a single factor (Fig. 2b); however, specific interactions of these factors are notable and can inform best practices.

The P90 and P95 metrics most consistently captured the greatest genetic variation across study sites and ground filtering methods (Fig. 2b), consistent with other findings using a different experiment (Pugh et al., 2018). The P_{max} metric captured the least genetic variation and had increased noise in low-altitude flights despite showing the most consistency with the ground data (Fig. 2b). The P90 and P95 metrics of the HRI and ATIN methods consistently explained greater genetic variation and repeatability than the ground truth measurements (Fig. 2b; red bar) across all sites and platforms, with HRI tending to outperform ATIN. Although the DBM outperformed HRI in genetic variation and repeatability at some sites (FW-LDCH and RW-MDCI), the majority of situations resulted in lower genetic variation and repeatability (less

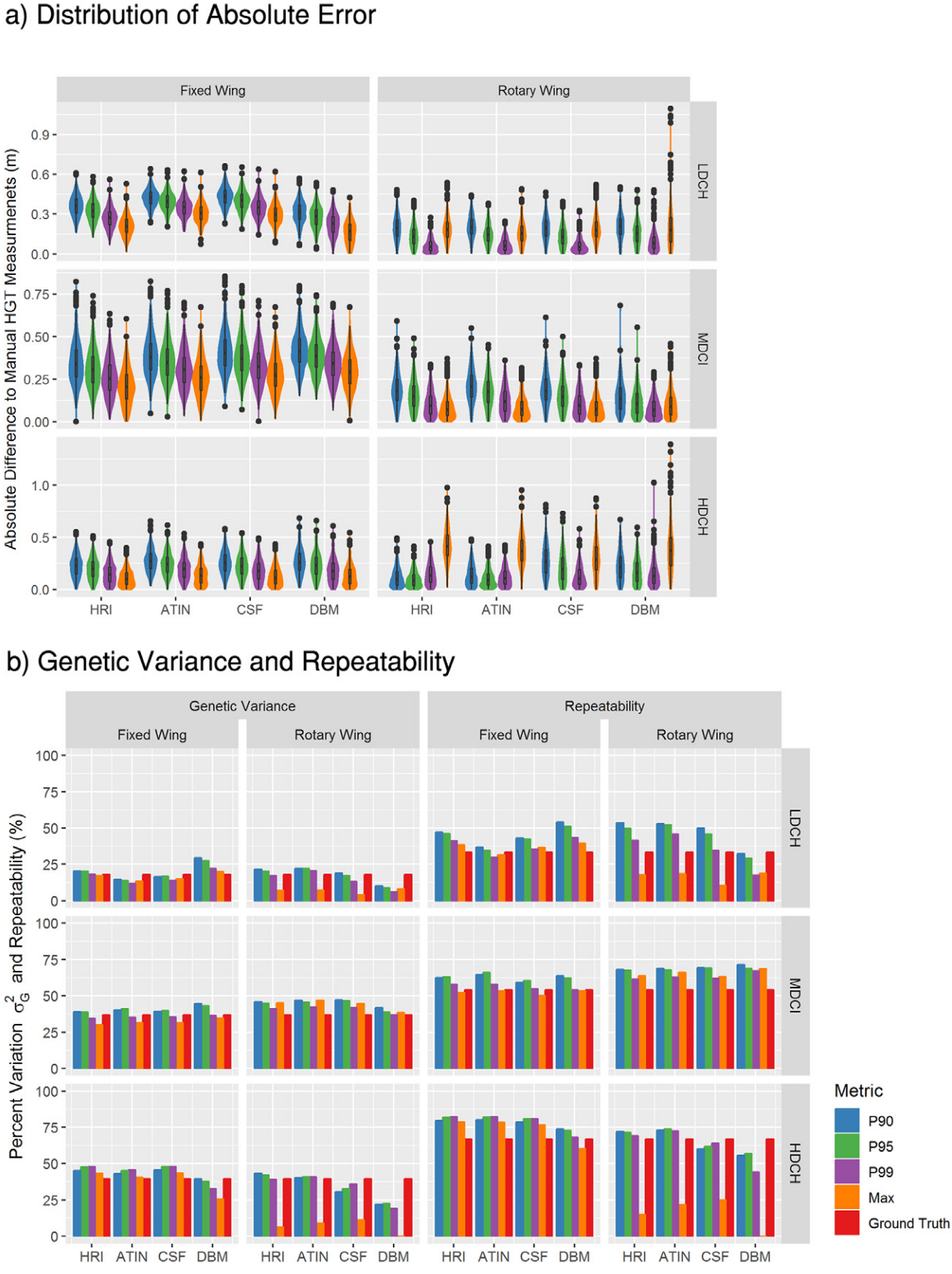

Fig. 2. (a) Violin plots comparing the distribution of absolute difference between unmanned aerial system (UAS) height metrics and ground truth measurements across UAS platforms (fixed wing and rotary wing) and study sites (LCDH, low canopy density hybrids; MCDI, medium canopy density inbred; HCDH, high canopy density hybrids) for each of the ground filtering methods (HRI, hierarchical robust interpolation; ATIN, adaptive triangulated irregular network; CSF, cloth simulation filter; DBM, difference based method); and (b) comparison of the percentage of genetic variation explained (left) and repeatability (right) across UAS platforms and study sites for each of the ground filtering methods.

desirable) compared with the three point cloud based methods and lower variation than ground truth data (FW-LCDH, RW-MDCI, and FW-HDCH) (Fig. 2b).

Several important discoveries were made from this comparative study for implementing UAS SfM height estimates. First, high accuracy compared with manual measurements did not result in genetic variation and repeatability being maximized (e.g., CSF and DBM approaches; Fig. 2; Supplemental Fig. S3) because the ground measurements themselves are probably flawed (biased across data curators' consistency and experience phenotyping a trait of interest). Second, although specific point cloud percentiles had greater accuracy, that did not always correlate to the highest repeatability or genetic variation (e.g., P99 vs. P95 and P90; Fig. 2). Third, one of the greatest benefits of UAS height estimates was the ability to substantially improve repeatability over manual measurements. While genetic variation was improved somewhat across study sites and platforms, repeatability increased by reducing error and better partitioning spatial variance; for example, the HRI method across all canopy densities and UAS platforms (R_{P95}: 50–80%; σ_G^2: 20–50%; Fig. 2b) outperformed ground truth measurements (R: 30–60%; σ_G^2: 18–40%; Fig. 2b) with more useful variance decomposition (Supplemental Fig. S4). Finally, if adequate ground representation is available throughout the study area (alleys between plots for example) in each flight, point cloud filtering (specifically HRI) methods are a more robust alternative to the difference-based method. The HRI method was easy to optimize, robust across study sites and UAS platforms, and improved repeatability over manual measurements. Based on these results, digital terrains were modeled off ground points identifies with the HRI ground filtering method, and the P95 metric was used to estimate plant height from point clouds in our data extraction pipeline for the rest of the study.

Comparison of UAS Platforms across Flight Dates

Statistics of UAS Survey Flight Dates

Throughout the growing season, most UAS surveys had either no difference or 1 d difference between flight dates of the two UAS platforms (Supplemental Table S1). During the beginning of the growing season, minimal plant structure was captured by UAS imagery due to a sparse canopy density and small physical size of the maize seedlings. Plant structure was not represented within the fixed-wing point clouds until 48 DAS for G2FE and DG2F and 35 DAS for G2LA (later plantings have faster germination and growth), while the rotary wing first detected plant structure at 27 DAS for G2FE and DG2F and at 21 DAS for G2LA (Fig. 3). The early plantings (G2FE and DG2F) demonstrate that higher flight altitudes require increased canopy structure before being represented in the SfM point clouds. This 48-d delay was probably due to a 13-d gap in fixed-wing flights during early growth stages (biweekly) in which the date when structure became capturable was missed. Understanding the date at which structure can be captured is important to reduce resources from UAS surveys of non-informative dates but is also critical for nonlinear modeling of growth.

The goal of increasing flights to twice weekly (every 3 to 4 d) rather than once a week was to capture the exponential growth period of maize, when a few days has been shown to make a large difference (Pugh et al., 2018). Unfortunately, the complete exponential stage was missed for the fixed-wing flights in the first plantings (G2FE and DG2E) due to limited knowledge of when this stage would begin. With the delayed planting of G2LA, surveys were collected biweekly and the exponential growth stage was captured effectively by both fixed and rotary wings (Fig. 3). The fixed-wing surveys of G2LA captured the exponential growth stage beginning around 35 DAS. and P95 height effectively increased by ~42 cm wk^{-1} at a rate of ~6 cm d^{-1} (Supplemental Table S2; Fig. 3). Within the rotary-wing surveys, the exponential growth stage began around 35 DAS, and P95 height effectively increased by 37, 37, and 42 cm wk^{-1} at a rate of 5, 5, and 7 cm d^{-1} across the G2FE, DG2F, and G2LA trials, respectively (Supplemental Table S2; Fig. 3). The higher resolution of the rotary wing coupled with weekly flights rather than once every 2 wk resulted in better observations of the exponential growth phase via temporal flight dates in all trials.

Analysis of temporal P95 height data indicated that a combination of survey methods should be used to successfully capture the growth patterns of maize hybrids. We have identified that weekly or fortnightly UAS surveys should begin 3 wk after sowing and continue through the flowering stage to accurately model the exponential growth stage and may require the combination of different flight altitudes based on the maturity of the trial. Early season flights should be flown at lower altitudes (≤25 m) to increase the detection of sparse plant structure by SfM photogrammetry, while later season flights should be flown at higher altitudes (>25 m) to ensure image matching, tie point identification, and point cloud densification. The ability to capture early season plant structure is still limited and will require improved SfM functionality or methods that do not rely on SfM photogrammetry (e.g., lidar or stereo sensors).

Variance Components and Repeatability of UAS Flight Dates

As the crop grew, total variance throughout the growing season increased in a quadratic manner across both platforms and all trials, although the trend was less consistent for the low-altitude rotary wing (black circles; Fig. 4). The repeatability estimates (white triangles; Fig. 4) were moderate (>60%) to very high (>90%), excluding uninformative image sets (e.g., flight dates with noticeably increased total variance like the rotary-wing flight on Day of the Year 128 of the optimal planted trials, the DSM of which was also visibly distorted). We determined that distorted flights were caused by a failure to identify key tie points in the mature canopies of the early plantings (DG2F and G2FE), leading to poor modeling of the canopy structure; this resulted in increased error variance, reduced consistency between replicates, and reduced genetic variance. Fixed-wing surveys captured ~10 to 40% greater genetic variation than PHT_{TRML} (σ_G^2: ~30%). The rotary wing

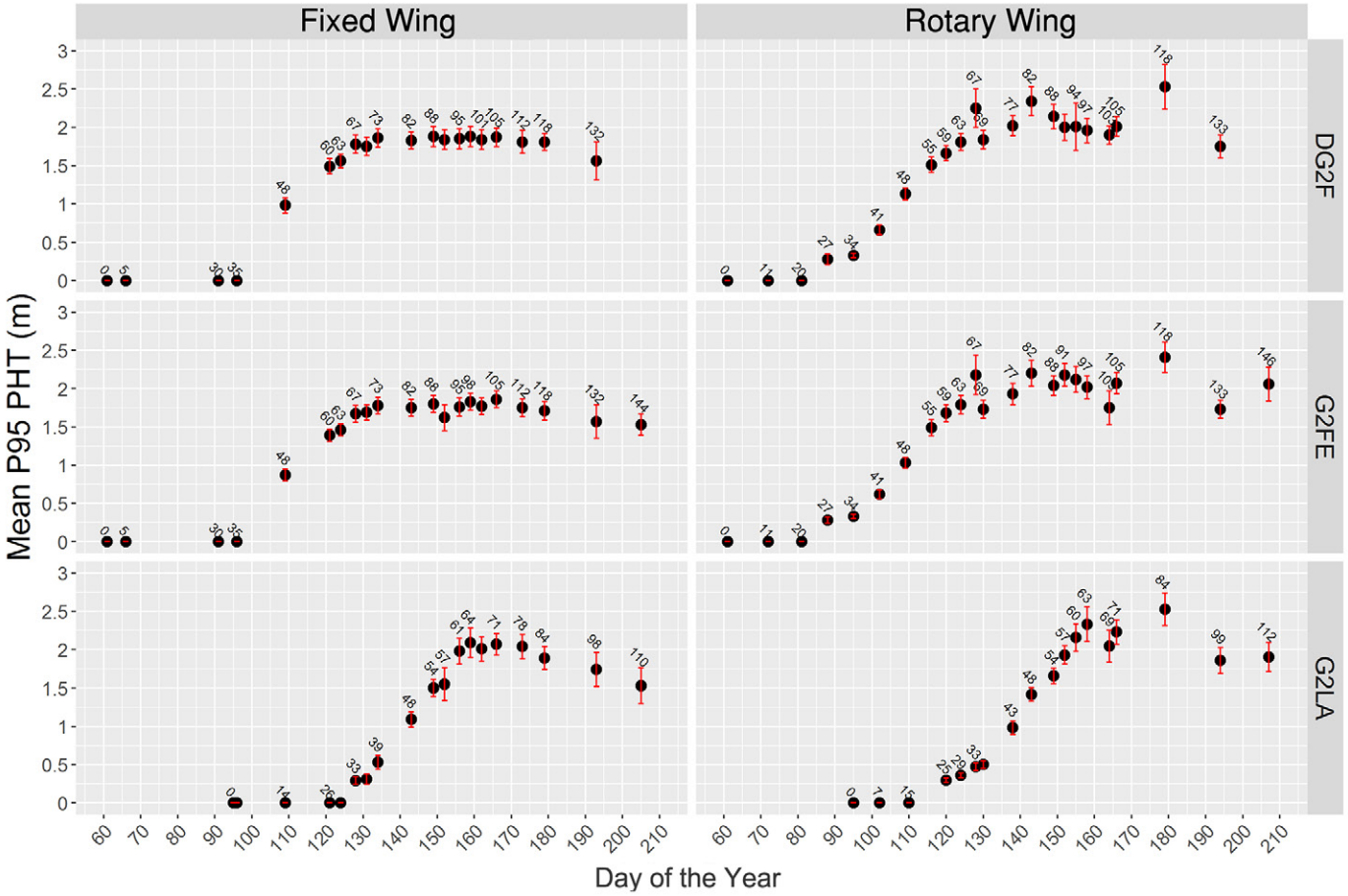

Fig. 3. Mean P95 height estimate (PHT) on a plot basis across the Genomes to Fields trials (DG2F, optimal planted, unirrigated trial; G2FE, optimal planted, irrigated trial; G2LA, delay planted, irrigated trial) and unmanned aerial system (UAS) platforms (fixed wing and rotary wing). Red error bars indicate the 95% confidence intervals scaled by one order of magnitude for visualization purposes. Numbers above the confidence intervals indicate days after sowing.

itself did not improve the explanation of genetic variance within the early planted trials (Fig. 4) and increased genetic variance by ~10% in the late planting (G2LA) compared with PHT_{TRML}. Specifically, estimates from low-altitude images became inconsistent during canopy closure accompanied by serendipitous spikes and dips in P95 height estimates and shrinkage of pedigree BLUP variance (e.g., rotary wing in Supplemental Fig. S5).

Variance component decomposition demonstrated that the majority of UAS surveys were informative. We recommend temporal data collection, which allows identification of flights that deviate from a normal trend (e.g., spikes in total variance, reduced genetic variance, increased residual error). It would be difficult to identify if images collected from a single UAS survey should be used in downstream analysis without temporal comparison. Continued research is required to develop tools and methodologies for classifying an individual UAS flight as informative without a comparison group.

Nonlinear Logistic Growth Curves

These and previous UAS surveys of plant height captured appeared as a sigmoidal growth pattern (Fig. 3), which is commonly applied to plant growth (Archontoulis and Miguez, 2015; Wardhani and Kusumastuti, 2013). While these data are highly informative, completing UAS surveys more than 20 times in a season is resource intensive and impractical, and the data can be redundant (assuming quality data are collected on every flight) for some dates. Furthermore, it is not possible to compare data across environments with different planting dates. A model that can both reduce the number of flights needed and predict the optimal flight dates after sowing would be valuable to maximize flight efficiency. Nonlinear models that capture the sigmoidal growth, specifically the logistical function (Eq. [5]), provide tools to model temporal crop growth and reduced dimensionality. Nonlinear models were fit on a plot-level basis, and BLUPs of the logistical parameters were extracted on a pedigree basis within each trial.

The fit of the logistic function had a RSME of 0.06 to 1.13 m across the trial environments, with the fixed wing (0.06–0.10 m) having a slightly better fit than the rotary wing (0.10–0.13 m). Similarly, the mean R^2 across plots ranged from 0.98 to 0.99, demonstrating that the logistic function accurately explained the variation in P95 height, regardless of environmental conditions or UAS platform (Supplemental Table S3).

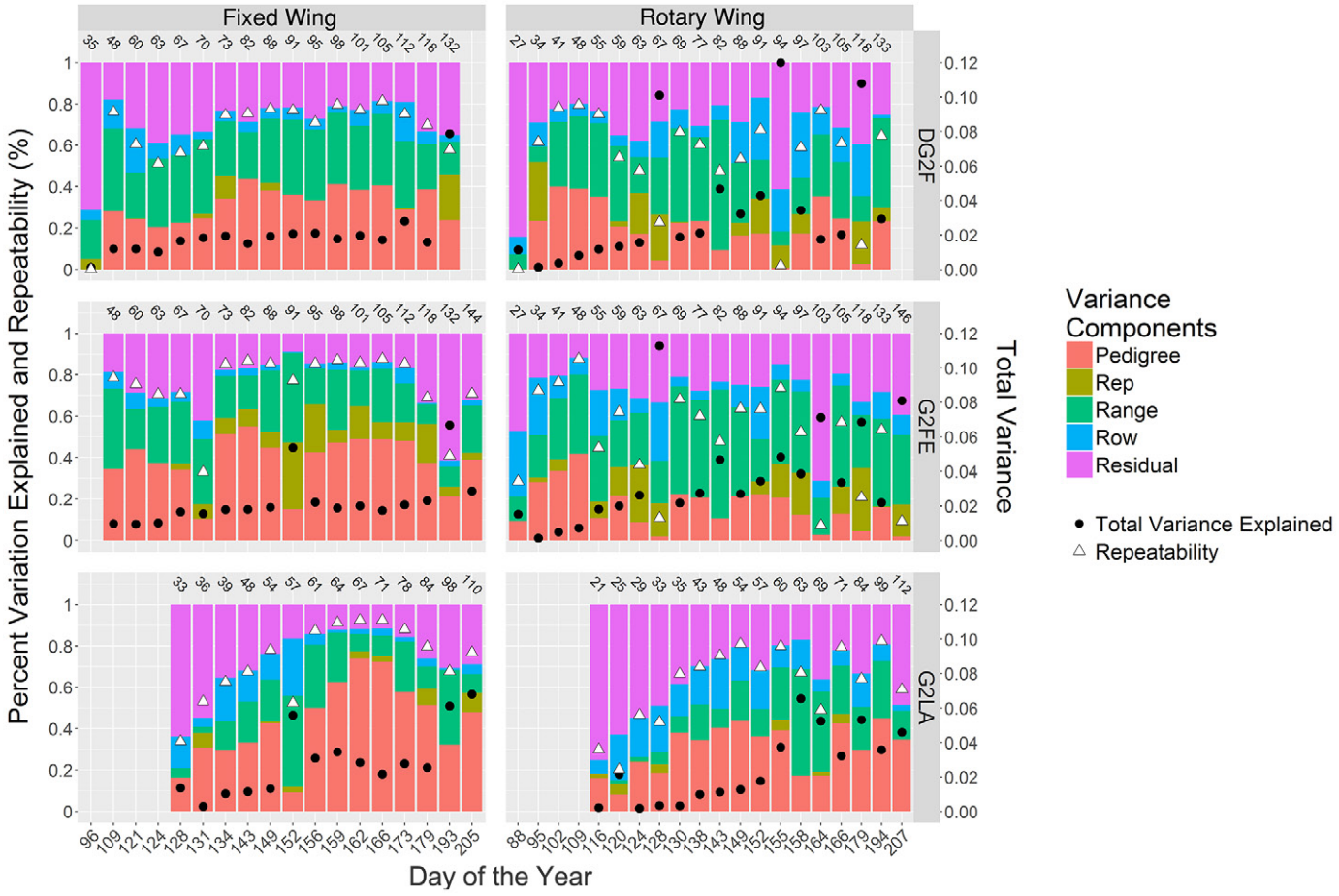

Fig. 4. Stack bar graphs of the percentage of variation explain by the variables in Eq. [3] for individual unmanned aerial system (UAS) surveys of individual UAS platforms (fixed wing and rotary wing) and experimental trials (DG2F, optimal planted, unirrigated trial; G2FE, optimal planted, irrigated trial; G2LA, delay planted, irrigated trial). Days of the year of the UAS image collection are indicated on the *x* axis and days after sowing (DAS) are indicated by the numbering above the bars. Total variance captured (black circle) per image set, defined by the right *y* axis, puts repeatability and genetic variance explained into perspective of other flight dates. Repeatability is indicated by the white triangles.

Variance component and repeatability estimates demonstrated that the three parameters of the plot-level logistic function captured equivalent or greater genetic variation then PHT_{TRML} ($\sigma_G^2 = \sim 30\%$, $R = 60–67\%$ within environments). The inflection point explained the greatest genetic variation in the early planted trials, ranging from 43 to 65% of the total variation, and met or exceeded the genetic variance captured by any single flight date or PHT_{TRML} (Supplemental Fig. S6). High genetic variation of the inflection point demonstrated that there was wide genetic variability in PHT at mid-season growth (halfway between zero and asymptote). The inflection point is a novel predictive phenotype not captured by single height estimates alone. The variance explained by the asymptote exceeded the genetic variation of PHT_{TRML}, excluding the G2FE and DG2F rotary wing, where tie point identification was poor. The asymptote, indicating terminal growth of the logistic curve, should be equivalent to the PHT_{TRML} measurement. Growth rate explained greater genetic variation than the asymptote in three data sets (fixed-wing DG2F, fixed-wing G2FE, and rotary-wing DG2F) and was never greater than the genetic variation explained by the inflection point. The fixed-wing comparisons demonstrated that variation in the growth rate is reduced if planting is delayed (G2LA trial vs. early planting). Reduced variation is explainable by increased growing degree days later in the season leading to more consistent, rapid growth across genetic backgrounds. The rotary wing effectively captured the genetic variation in the inflection point, which occurred during periods of lower canopy density when tie points could be better identified.

Although the repeatability of logistic parameters was reduced (did not exceed 60%; Supplemental Fig. S6) compared with the best individual flight date UAS P95 estimates, the logistic parameters provided an opportunity to use multi-environment UAS data sets in a combined analysis. Specifically, logistical parameters do not confine UAS surveys to similar DAS or calendar dates across environments or years (e.g., P95 at 60 DAS). Combined analysis of PHT_{TRML} measurements ($\sigma_G^2 = 48\%$, $R = 70\%$, Table 1) was exceeded only by the fixed-wing P95 inflection point ($\sigma_G^2 = 50\%$, $R = 82\%$, Table 1). Although limited improvement was made in capturing greater genetic variation of the logistic parameter over PHT_{TRML}, a noticeable reduction (23–77%) in the residual variation (excluding rotary-wing growth rate) was observed (Table 1). Specifically, variance was partitioned to a greater extent within

Table 1. Combined analysis (Eq. [2]) across Genomes to Fields trials for manual terminal height logistic curve parameters for each unmanned aerial system (UAS) platform. The UAS estimates used hierarchical robust interpolation ground modeling and P95 height estimates.

		Fixed wing			Rotary wing		
Component	PHTTRML†	Asymptote	Growth rate	Inflection point	Asymptote	Growth rate	Inflection point
Genotype (G)	9.9×10^{-3} (48)‡	1.3×10^{-2} (29)	1.3×10^{-4} (23)	1.7 (50)	7.9×10^{-3} (37)	4.1×10^{-5} (16)	1.3 (16)
Environment (E)	2.1×10^{-3} (10)	2.5×10^{-2} (55)	1.8×10^{-4} (33)	0.5 (16)	3.4×10^{-3} (16)	5.4×10^{-5} (21)	5.5 (70)
G × E	0 (0)	8.0×10^{-4} (2)	5.4×10^{-5} (10)	0.1 (4)	1.8×10^{-4} (1)	1.9×10^{-6} (1)	5.4×10^{-2} (1)
E(replicate)	0 (0)	1.5×10^{-3} (3)	4.5×10^{-5} (8)	5.9×10^{-2} (2)	4.3×10^{-5} (0)	3.2×10^{-5} (12)	0.2 (2)
E(range)	2.1×10^{-3} (10)	1.1×10^{-3} (2)	4.7×10^{-5} (8)	0.2 (7)	3.2×10^{-3} (15)	2.4×10^{-5} (9)	0.1 (1)
E(row)	3.2×10^{-3} (2)	7.9×10^{-4} (2)	1.7×10^{-5} (3)	0.3 (7)	2.1×10^{-3} (10)	2.2×10^{-5} (8)	0.2 (2)
Residual	6.1×10^{-3} (30)	3.2×10^{-3} (7)	8.9×10^{-5} (16)	0.5 (14)	4.4×10^{-3} (21)	8.3×10^{-5} (32)	0.6 (7)
Repeatability	0.70	0.85	0.62	0.82	0.73	0.42	0.77

† Manually measured terminal plant height.
‡ Values are raw variance component estimates, with percentage of genetic variation explained by each model variable and entry means in parentheses.

environment, genotype × environment interaction, and spatial variables, resulting in a 3 to 15% increase in repeatability estimates over PHT_{TRML} (excluding growth rate). The results demonstrated that nonlinear logistic modeling could provide highly repeatable, genetically informative phenotypes, which would alleviate the need for capturing of UAS surveys at equivalent DAS across trials, years, or locations, allowing more efficient targeting of flight dates, as well as providing novel phenotypes beyond simple height measurements. Incorporation of growing degree days, weather patterns, or other time-dependent parameters as the dependent factor (x) of the growth curve could improve comparisons of growth curves across sites and warrants further investigation.

Correlation to Grain Yield

While the plant height trait is of interest in and of itself, it is of greater interest as a phenotype correlated with and an aid in predicting the highest yielding genotypes. Pearson's correlations between PHT_{TRML} and grain yield demonstrated 0.28, 0.25, and 0.23 for DG2F (Fig. 5), G2FE, and G2LA trials, respectively (Table 2; Supplemental Fig. S7, S8, and S9). These correlations are slightly lower than generally seen in the Texas A&M breeding program and substantially lower than that found by Barrero Farfan et al. (2013), probably due to the G2F experiment including hybrids of diverse origins that contained a variety of unadapted factors that affect the yield and plant height relationship in different ways (e.g., photoperiod sensitivity, temperature stress, drought stress, etc.). The UAS P95 height estimates showed higher correlations to grain yield than PHT_{TRML} beginning ~70 DAS for DG2F and G2FE, while flights after ~50 DAS showed higher correlations to yield in the G2LA trial. Furthermore, fixed-wing UAS P95 heights had maximum yield correlations of 45, 42, and 42%, while the rotary wing reached 41, 36, and 46% correlation to yield for DG2F, G2FE, and G2LA, respectively (Table 2). The ~20% increase in correlation to yield from UAS P95 estimates over PHT_{TRML} measurements demonstrate that UAS P95 height estimates can serve as an improved method for collecting phenotypes to improve genetic gain.

Correlation between temporal measures of UAS P95 height and yield increased with time and were least informative for grain yield prior to the reproductive growth and grain fill stages. If using only plant height to predict yield, late season flights are more informative than flights prior to the vegetative-to-growth transition. However, both plant height and grain yield are sculpted by daily interactions between the genetics of the plants and the environment up to that point. This lack of correlation between early season UAS P95 height and yield suggests that the genetic variation in early season height is under independent genetic control. We hypothesize that vigorous early season growth could

Table 2. Pedigree best linear unbiased predictor (BLUP) correlation between grain yield and manual terminal plant height (PHT_{TRML}), the flight date with the highest correlation, and the logistic parameters across trials (DG2F, optimal planted, unirrigated trial; G2FE, optimal planted, irrigated trial; G2LA, delay planted, irrigated trial). Combined columns indicated correlations based on the pedigree BLUPs of a combined trial analysis.

	Fixed wing				Rotary wing			
Parameter	DG2F	G2FE	G2LA	Combined	DG2F	G2FE	G2LA	Combined
PHT_{TRML}	0.28**	0.25***	0.23***	0.27***	0.28**	0.25***	0.23***	0.27***
Best flight date	0.45***	0.42***	0.43***	–	0.41***	0.36***	0.47***	–
Asymptote	0.44***	0.42***	0.42***	0.39***	0.44***	0.38***	0.45***	0.41***
Growth rate	−0.46***	−0.42***	−0.34***	−0.42***	−0.30***	−0.13***	−0.13**	−0.29***
Inflection point	0.46***	0.42***	0.18	0.36***	0.42***	0.36***	0.15*	0.36***

* Significant at $\alpha < 0.05$.
** Significant at $\alpha < 0.01$.
*** Significant at $\alpha < 0.001$.

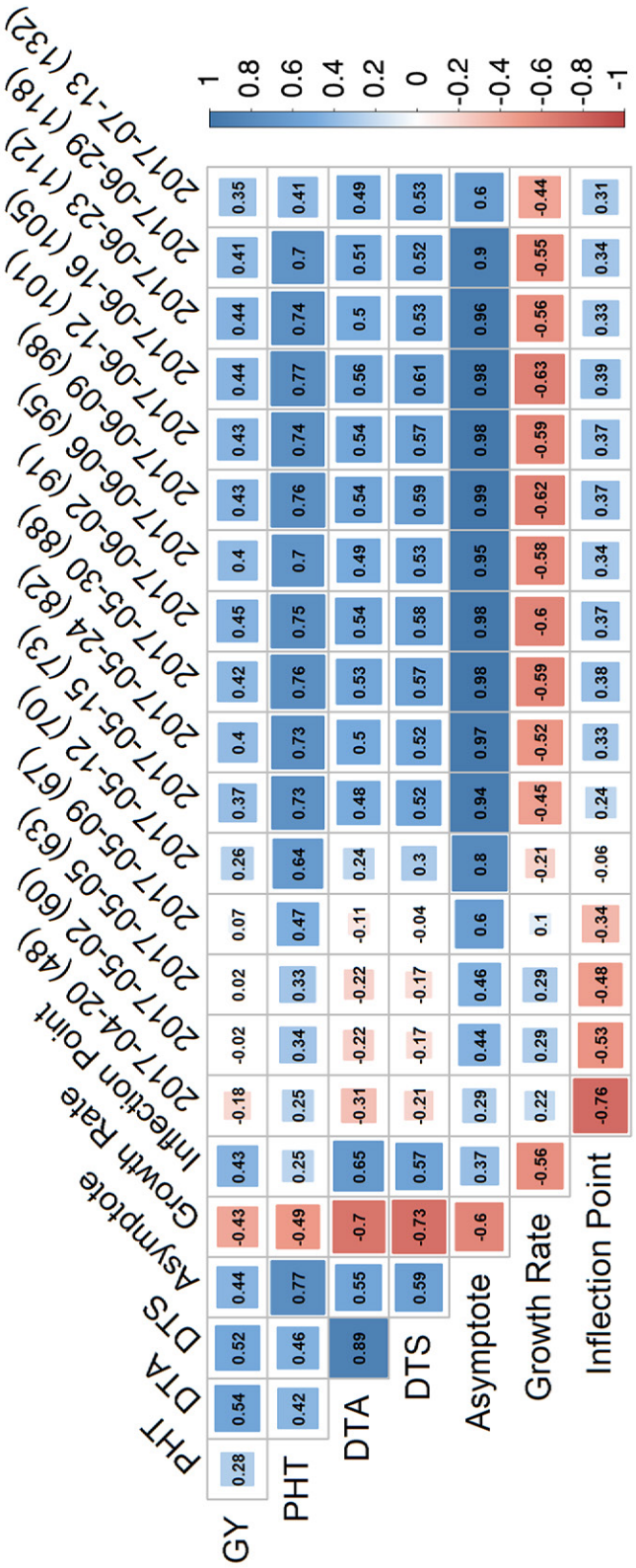

Fig. 5. Heat map comparing correlations between grain yield (GY), manual terminal plant height (PHT), flowering time (days to anthesis [DTA] and days to silking [DTS]), logistic parameters (asymptote, growth rate, inflection point), and unmanned aerial system (UAS) P95 estimates by flight date for the optimal planted, unirrigated trial (DG2F) surveyed via the fixed-wing UAS.

be pyramided into the terminally taller, higher yielding plants to develop cultivars improved across all growth stages; UAS P95 height estimates would be critical for practically testing this hypothesis.

While individual flight dates were highly correlated to grain yield, the correlation of the logistic parameters to grain yield equaled the highest correlated single flight date P95 height measurement. The asymptote parameter was 38 to 45% correlated to grain yield across trials and platforms (Table 2). The asymptote parameter describes the maximum plant height of the logistic curve and was 69 to 76% correlated to PHT_{TRML} (Fig. 5; Supplemental Fig. S7, S8, and S9). The inflection point was also 36 to 46% correlated to grain yield in DG2F and G2FE, while becoming less informative (15–18%) in the later G2LA trial (Table 2). We speculate that the high correlation of inflection point and asymptote to grain yield may be equivalent to the previously reported predictive power of V6 and V10/12 height to grain yield (Machado et al., 2002; Yin et al., 2011), although a leaf counting study would be necessary to validate this hypothesis. Growth rate depicted a negative trend to grain yield while showing a significant reduction in correlation from fixed wing (−30 to −46%) to rotary wing (−13 to −30%). The negative correlations between growth rate and both grain yield and plant height relates to the negative correlation between flowering time (37–73%) and growth rate (Supplemental Fig. S7, S8, and S9). Early maturity results in less vegetative growth and shorter plants. In addition, early hybrids are more likely to originate from the far northern United States and may be maladapted to Texas conditions.

Predicting Grain Yield from UAS Height Phenotypes

While simple correlations provided relationships between grain yield and a single P95 estimate or logistic parameter, it is possible that multiple height factors could be combined to make more robust predictions of yield. Forward and reverse stepwise regression was performed to identify the most predictive UAS height parameters for grain yield (t ha^{-1}); the best of these were then fit as continuous effects in a linear model. Models were developed for each UAS platform separately, with three sets of possible predictors: (i) the logistic parameters, (ii) pedigree P95 height BLUPs by flight date, and (iii) the combination of logistic parameters and pedigree P95 height BLUPs by flight date.

Initially, PHT_{TRML} was used as the predictor of grain yield, which resulted in R^2 values of 0.08, 0.06, and 0.07 with RMSE values of 1.02, 1.05, and 0.68 t ha^{-1} for DG2F, G2FE, and G2LA, respectively, and $R^2 = 0.16$, RMSE = 0.50 t ha^{-1} from a combined trial analysis. Excluding logistic parameters, the fixed-wing flights

identified the two flight dates most informative in yield prediction at ~40 to 50 DAS and ~110 to 120 DAS, significantly increasing R^2 (0.28–0.38) and reducing RMSE by ~0.06 t ha^{-1} (Table 3; Supplemental Table S4). Similar R^2 and RSME values were obtained with rotary-wing flights with the addition of a third predictor around 70 DAS. Using only the logistic parameters, the predictive power was slightly reduced (fixed-wing R^2: 0.26–0.34; rotary-wing R^2: 0.25–0.33) compared with the best sets of individual flight dates (Table 3) but were significantly improved compared with PHT_{TRML}. Combining individual flight data and logistic parameters showed no improvement in predictive power and was inconsistent in the inclusion of only flight dates or a combination of flight dates and logistic parameters in the selected models (Supplemental Table S4). The individual flight dates identified through stepwise regression co-localize with the inflection point period (40–53 DAS) and terminal plant height (i.e., asymptote), which indicates why the logistic parameter model achieved equivalent predictive power to the specific flight date model.

The comparison of relative rankings of hybrids from the models using logistic parameters vs. PHT_{TRML} measurements demonstrated improved selection accuracy of UAS-derived logistic curves over PHT_{TRML} measurements (Supplemental Table S5). Prediction of grain yield using logistic parameters improved ranking error by 7 to 10 ranks over PHT_{TRML} prediction (MAE = 59–78 ranks; Supplemental Table S5). Although improvement in relative ranking is ideal, plant breeders generally select a subset (e.g., the top 10% yielding hybrids) of their evaluated material to advance in evaluation trials. The UAS logistic-based prediction improved the selection accuracy of the top 10% yielding hybrids by ~50 to 150% over PHT_{TRML} predictions (21%), and the combined analysis demonstrated 7 and 12% increases in selection accuracy for the fixed wing and rotary wing, respectively (Supplemental Table S5).

For a plant breeding program, selecting material to advance by UAS a month or more before maturity can speed the breeding cycle, substantially decrease the cost and time compared with combine harvesting, and allow more environments to be screened. However, we acknowledge that the prediction of grain yield solely on height measurements is not an acceptable model of yield prediction in these trials; nevertheless, significant improvements in predictive power were obtained by using UAS technologies temporally. Additional UAS-estimated phenotypes (vegetation indices, canopy cover, plant population, etc.) need to be developed and included with height for better predictions of yield if plant breeders will ever be able to select based on remote sensing data.

For fundamental research into plant physiology, genetics, and development, these UAS findings open up interesting avenues to identify differences in growth trajectories, impractical to measure previously. Most importantly, such studies can be conducted on mature plants, nondestructively, in a field setting, which is important if discoveries are ever to be used in practical crop or agronomic improvement.

Conclusion

This study is one of the first applications of UAS phenotyping of agricultural research at a representative scale (>1500 plots) of a breeding or agricultural research program. The comparisons of different UAS platforms and flight altitudes have provided additional insights toward reliable application of UAS imagery within an agricultural field trial setting, specifically within crops with dense canopy structure yet sparse apex canopy features (e.g., tassels). To our knowledge, this is one of the first empirical studies to move beyond UAS phenotype validation toward phenotypic predictive modeling across a large set of plant material (280 hybrids), while validating a previous finding (Pugh et al., 2018) in a different germplasm pool and environment. Four of the most important findings were: (i) the dense canopy structure at later growth stages of maize restricts execution of SfM photogrammetry, returning inconsistent data quality, specifically at low flight altitudes; (ii) increased genetic variation (10–40%) was captured by UAS P95 compared with conventional manual terminal plant height measurements, accompanied with reduced residual error, resulting in increased measurement repeatability; (iii) logistic functions accurately model UAS maize height estimates, which can be used in place of independent flight dates to develop robust prediction models and allow execution of combined environment analysis with relative ease; and (iv) predictive modeling of grain yield via UAS height estimates or logistic function parameters demonstrated substantial improvements in the proportion of grain yield variation explained and overall selection accuracy compared with traditional PHT_{TRML} in model selection accuracy.

Table 3. Coefficient of determination (R^2) for the best prediction models of yield defined by stepwise regression (Supplemental Table S4) by unmanned aerial system (UAS) platform (fixed wing and rotary wing) and Genomes to Fields trial (DG2F, optimal planted, unirrigated trial; G2FE, optimal planted, irrigated trial; G2LA, delay planted, irrigated trial). Combined columns indicated the combined trial analysis.

	Fixed wing				Rotary wing			
Grain yield predictor†	DG2F	G2FE	G2LA	Combined	DG2F	G2FE	G2LA	Combined
Logistic parameters	0.34 (0.92)‡	0.26 (0.95)	0.27 (0.62)	0.32 (0.49)	0.33 (0.94)	0.32 (0.99)	0.25 (0.62)	0.32 (0.49)
Flight dates	0.38 (0.92)	0.35 (0.97)	0.28 (0.60)	NA§	0.39 (0.90)	0.33 (0.99)	0.26 (0.61)	NA
Logistic parameters and flight dates	0.37 (0.92)	0.35 (0.97)	0.28 (0.60)	NA	0.38 (0.92)	0.33 (0.98)	0.26 (0.61)	NA

† Logistic parameters, prediction model defined using logistic parameters in the stepwise regression; flight dates, prediction model defined using UAS estimates by flight date in the stepwise regression; logistic parameters and flight dates, prediction model defined using logistic parameters and UAS estimates by flight date in the stepwise regression.

‡ Values in parentheses are the root mean squared errors of grain yield (t ha^{-1}).

§ NA, the same flight date was at different growth stages for early and late (G2LA) plantings, so they could not be combined.

Supplemental Material

Supplement 1 contains supplemental tables.

Supplement 2 contains supplemental figures.

Supplement 3 contains a discussion of point-cloud-based ground filtering algorithms.

Funding

This project was made possible by financial support from USDA–NIFA–AFRI Award no. 2017-67013-26185, USDA–NIFA Hatch funds, Texas A&M AgriLife Research, the Texas Corn Producers Board, the Iowa Corn Promotion Board, and the Eugene Butler Endowed Chair in Biotechnology. Steven Anderson was funded for one year by the Texas A&M College of Agriculture and Life Sciences Tom Slick Senior Graduate Fellowship.

Author Contributions

S.L. Anderson: conceptualization, data curation, formal analysis, investigation, methodology, original draft, review and editing (lead), supervision, validation, visualization; S.C. Murray: conceptualization, funding acquisition, methodology, project administration, supervision, resources, review and editing (supporting); L. Malambo: conceptualization, data curation, formal analysis, investigation, methodology, review and editing (supporting), software; A. Chang: data curation, formal analysis, investigation, methodology, review and editing (supporting), software; C. Ratcliff: data curation, investigation, methodology, review and editing (supporting); S. Popescu: conceptualization, funding acquisition, resources, review and editing (supporting), software, supervision; D. Cope: conceptualization, funding acquisition, resources, review and editing (supporting), supervision; J. Jung: data curation, formal analysis, investigation, methodology, resources, review and editing (supporting), software, supervision; A. Thomasson: conceptualization, resources, review and editing (supporting), and supervision.

Acknowledgments

We would like to acknowledge Andrew Vree and Ian Gates for conducting the Tuffwing UAV Mappers surveys in 2017; Cody Bagnall and Dalton Askew for constructing the GCPs used to georeference survey images; Misty Miles for coordinating the flight teams and providing administrative support; Dr. N. Ace Pugh and David Horne for their collaborative efforts in assuring that UAS surveys were completed; Jacob Pekar, David Rooney, and Stephen Labar for their agronomic and technical support; graduate students and undergraduate and high school employees of the Texas A&M Quantitative Genetics and Maize Breeding Program for their hard work and effort maintaining fields and collecting phenotypic data, specifically, Dalton Askew and Raul Ramirez for collection of ground truth measurements throughout the growing season; and all members of the Texas A&M UAS project for their collaboration, without which implementation of UAS at Texas A&M would not have been so successful.

References

AgiSoft. 2016. PhotoScan Professional. Version 1.2.6. AgiSoft, St. Petersburg, Russia. http://www.agisoft.com/downloads/installer/.

AlKhalifah, N., D.A. Campbell, C.M. Falcon, J.M. Gardiner, N.D. Miller, M.C. Romay, et al. 2018. Maize Genomes to Fields: 2014 and 2015 field season genotype, phenotype, environment, and inbred ear image datasets. BMC Res. Notes 11:452. doi:10.1186/s13104-018-3508-1

Anderson, S.L., A.L. Mahan, S.C. Murray, and P.E. Klein. 2018. Four parent maize (FPM) population: Effects of mating designs on linkage disequilibrium and mapping quantitative traits. Plant Genome 11:170102. doi:10.3835/plantgenome2017.11.0102

Anderson, S.L., S.C. Murray, L. Malambo, C. Ratcliff, S. Popescu, D. Cope, et al. 2019. Data from: Prediction of maize grain yield before maturity using improved temporal height estimates of unmanned aerial systems. Dryad Digital Repository. doi:10.5061/dryad.3295k54

Araus, J.L., and S.C. Kefauver. 2018. Breeding to adapt agriculture to climate change: Affordable phenotyping solutions. Curr. Opin. Plant Biol. 45B:237–247. doi:10.1016/j.pbi.2018.05.003

Archontoulis, S.V., and F.E. Miguez. 2015. Nonlinear regression models and applications in agricultural research. Agron. J. 107:786–798. doi:10.2134/agronj2012.0506

Axelsson, P. 2000. DEM generation from laser scanner data using adaptive TIN models. Int. Arch. Photogramm. Remote Sens. 33:110–117.

Baltsavias, E.P. 1999. A comparison between photogrammetry and laser scanning. ISPRS J. Photogramm. Remote Sens. 54:83–94. doi:10.1016/S0924-2716(99)00014-3

Bareth, G., J. Bendig, N. Tilly, D. Hoffmeister, H. Aasen, and A. Bolten. 2016. A comparison of UAV- and TLS-derived plant height for crop monitoring: Using polygon grids for the analysis of crop surface models (CSMs). Photogramm. Fernerkund. Geoinf. 2016:85–94. doi:10.1127/pfg/2016/0289

Barrero Farfan, I.D., S.C. Murray, S. Labar, and D. Pietsch. 2013. A multi-environment trial analysis shows slight grain yield improvement in Texas commercial maize. Field Crops Res. 149:167–176. doi:10.1016/j.fcr.2013.04.017

Bendig, J., A. Bolten, and G. Bareth. 2013. UAV-based imaging for multi-temporal, very high resolution crop surface models to monitor crop growth variability. Photogramm. Fernerkund. Geoinf. 2013:551–562. doi:10.1127/1432-8364/2013/0200

Burkart, A., V. Hecht, T. Kraska, and U. Rascher. 2018. Phenological analysis of unmanned aerial vehicle based time series of barley imagery with high temporal resolution. Precis. Agric. 19:134–146. doi:10.1007/s11119-017-9504-y

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

Chen, Z., B. Gao, and B. Devereux. 2017. State-of-the-art: DTM generation using airborne LIDAR data. Sensors 17:150. doi:10.3390/s17010150

Chu, T., M.J. Starek, M.J. Brewer, S.C. Murray, and L.S. Pruter. 2018. Characterizing canopy height with UAS structure-from-motion photogrammetry: Results analysis of a maize field trial with respect to multiple factors. Remote Sens. Lett. 9:753–762. doi:10.1080/2150704X.2018.1475771

Gage, J.L., D. Jarquin, C. Romay, A. Lorenz, E.S. Buckler, S. Kaeppler, et al. 2017. The effect of artificial selection on phenotypic plasticity in maize. Nature Commun. 8:1348. doi:10.1038/s41467-017-01450-2

Girardeau-Montaut, D. 2016. CloudCompare. Version 2.8. https://www.danielgm.net/cc/.

Han, X., J.A. Thomasson, G.C. Bagnall, N. Pugh, D.W. Horne, W.L. Rooney, et al. 2018. Measurement and calibration of plant-height from fixed-wing UAV images. Sensors 18:4092. doi:10.3390/s18124092

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

Hu, P., S.C. Chapman, X. Wang, A. Potgieter, T. Duan, D. Jordan, et al. 2018. Estimation of plant height using a high throughput phenotyping platform based on unmanned aerial vehicle and self-calibration: Example for sorghum breeding. Eur. J. Agron. 95:24–32. doi:10.1016/j.eja.2018.02.004

Hunt, E.R., Jr., and C.S.T. Daughtry. 2017. What good are unmanned aircraft systems for agricultural remote sensing and precision agriculture? Int. J. Remote Sens. 39:5345–5376. doi:10.1080/01431161.2017.1410300

Isenburg, M. 2015. LAStools: Efficient tools for LiDAR processing. Version 130506. Dep. of Comput. Sci., Univ. of North Carolina, Chapel Hill.

Katsvairo, T.W., W.J. Cox, and H.M. Van Es. 2003. Spatial growth and nitrogen uptake variability of corn at two nitrogen levels. Agron. J. 95:1000–1011. doi:10.2134/agronj2003.1000

Korzeniowska, K., N. Pfeifer, G. Mandlburger, and A. Lugmayr. 2014. Experimental evaluation of ALS point cloud ground extraction tools over different terrain slope and land-cover types. Int. J. Remote Sens. 35:4673–4697. doi:10.1080/01431161.2014.919684

Kraus, K., and N. Pfeifer. 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS J. Photogramm. Remote Sens. 53:193–203. doi:10.1016/S0924-2716(98)00009-4

Li, W., Z. Niu, H. Chen, D. Li, M. Wu, and W. Zhao. 2016a. Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. Ecol. Indic. 67:637–648. doi:10.1016/j.ecolind.2016.03.036

Li, X., Z. Zhou, J. Ding, Y. Wu, B. Zhou, R. Wang, et al. 2016b. Combined linkage and association mapping reveals QTL and candidate genes for plant and ear height in maize. Front. Plant Sci. 7:833. doi:10.3389/fpls.2016.00833

Machado, S., E. Bynum, T. Archer, R. Lascano, L. Wilson, J. Bordovsky, et al. 2002. Spatial and temporal variability of corn growth and grain yield. Crop Sci. 42:1564–1576. doi:10.2135/cropsci2002.1564

Mahan, A.L., S.C. Murray, and P.E. Klein. 2018. Four-parent maize (FPM) population: Development and phenotypic characterization. Crop Sci. 58:1106–1117. doi:10.2135/cropsci2017.07.0450

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

Mallarino, A., E. Oyarzabal, and P. Hinz. 1999. Interpreting within-field relationships between crop yields and soil and plant variables using factor analysis. Precis. Agric. 1:15–25. doi:10.1023/A:1009940700478

McGaughey, R. 2016. FUSION/LDV: Software for LIDAR data analysis and visualization. Version 3.60+. US For. Serv., Pac. Northw. Res. Stn., Corvallis, OR.

Meng, X., N. Currit, and K. Zhao. 2010. Ground filtering algorithms for airborne LiDAR data: A review of critical issues. Remote Sens. 2:833–860. doi:10.3390/rs2030833

Montealegre, A.L., M.T. Lamelas, and J. de la Riva. 2015. A comparison of open-source LiDAR filtering algorithms in a Mediterranean forest environment. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8:4072–4085. doi:10.1109/JSTARS.2015.2436974

Murray, S.C., L. Malambo, S. Popescu, D. Cope, S.L. Anderson, A. Chang, et al. 2019. G2F Maize UAV Data, College Station, Texas 2017. CyVerse Data Commons. doi:10.25739/4ext-5e97

Niu, Q., H. Feng, G. Yang, C. Li, H. Yang, B. Xu, and Y. Zhao. 2018. Monitoring plant height and leaf area index of maize breeding material based on UAV digital images. Trans. Chin. Soc. Agric. Eng. 34:73–82.

Peiffer, J.A., M.C. Romay, M.A. Gore, S.A. Flint-Garcia, Z. Zhang, M.J. Millard, et al. 2014. The genetic architecture of maize height. Genetics 196:1337–1356. doi:10.1534/genetics.113.159152

Pfeifer, N., and G. Mandlburger. 2009. LiDAR data filtering and DTM generation. CRC Press, Boca Raton, FL. p. 307–334.

Polat, N., and M. Uysal. 2015. Investigating performance of airborne LiDAR data filtering algorithms for DTM generation. Measurement 63:61–68. doi:10.1016/j.measurement.2014.12.017

Pugh, N., D.W. Horne, S.C. Murray, G. Carvalho, L. Malambo, J. Jung, et al. 2018. Temporal estimates of crop growth in sorghum and maize breeding enabled by unmanned aerial systems. Plant Phenome J. 1:170006. doi:10.2135/tppj2017.08.0006

R Core Team. 2016. R: A language and environment for statistical computing. Version 3.3.1. R Found. Stat. Comput., Vienna. https://www.R-project.org.

rapidlasso. 2017. LAStools: Efficient LiDAR processing software. Version 170628. rapidlasso GmbH, Gilching, Germany. http://rapidlasso.com/LAStools.

Reynolds, D., F. Baret, C. Welcker, A. Bostrom, J. Ball, F. Cellini, et al. 2018. What is cost-efficient phenotyping? Optimizing costs for different scenarios. Plant Sci. 282:14–22. doi:10.1016/j.plantsci.2018.06.015

Sankaran, S., L.R. Khot, C.Z. Espinoza, S. Jarolmasjed, V.R. Sathuvalli, G.J. Vandemark, et al. 2015. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. Eur. J. Agron. 70:112–123. doi:10.1016/j.eja.2015.07.004

SAS Institute. 2018. JMP. Version 14.0.0. SAS Inst., Cary, NC.

Serifoglu Yilmaz, C., and O. Gungor. 2016. Comparison of the performances of ground filtering algorithms and DTM generation from a UAV-based point cloud. Geocarto Int. 33:522–537. doi:10.1080/10106049.2016.1265599

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

Sithole, G., and G. Vosselman. 2004. Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. ISPRS J. Photogramm. Remote Sens. 59:85–101. doi:10.1016/j.isprsjprs.2004.05.004

van der Voort, D. 2016. Exploring the usability of unmanned aerial vehicles for non-destructive phenotyping of small-scale maize breeding trials. Wageningen Univ. and Res. Ctr., Wageningen, the Netherlands.

Varela, S., Y. Assefa, P.V. Prasad, N.R. Peralta, T.W. Griffin, A. Sharda, et al. 2017. Spatio-temporal evaluation of plant height in corn via unmanned aerial systems. J. Appl. Remote Sens. 11:036013. doi:10.1117/1.JRS.11.036013

Verhulst, P.-F. 1838. Notice sur la loi que la population suit dans son accroissement. Corresp. Math. Phys. 10:113–126.

Wallace, J.G., X. Zhang, Y. Beyene, K. Semagn, M. Olsen, B.M. Prasanna, and E.S. Buckler. 2016. Genome-wide association for plant height and flowering time across 15 tropical maize populations under managed drought stress and well-watered conditions in sub-Saharan Africa. Crop Sci. 56:2365–2378. doi:10.2135/cropsci2015.10.0632

Wallace, L., A. Lucieer, C. Watson, and D. Turner. 2012. Development of a UAV–LiDAR system with application to forest inventory. Remote Sens. 4:1519–1543. doi:10.3390/rs4061519

Wardhani, W.S., and P. Kusumastuti. 2013. Describing the height growth of corn using Logistic and Gompertz model. Agrivita 35:237–241.

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

Weed, C.A., M.M. Crawford, A.L. Neuenschwander, and R. Gutierrez. 2002. Classification of LIDAR data using a lower envelope follower and gradient-based operator. In: Geoscience and Remote Sensing Symposium, Toronto, ON, Canada. 24–28 June 2002. IEEE Publ., Piscataway, NJ. doi:10.1109/IGARSS.2002.1026124

Xavier, A., B. Hall, A.A. Hearst, K.A. Cherkauer and K.M. Rainey. 2017. Genetic architecture of phenomic-enabled canopy coverage in *Glycine max*. Genetics 206:1081–1089. doi:10.1534/genetics.116.198713

Yin, X., M.A. McClure, N. Jaja, D.D. Tyler, and R.M. Hayes. 2011. In-season prediction of corn yield using plant height under major production systems. Agron. J. 103:923–929. doi:10.2134/agronj2010.0450

Zhang, W., J. Qi, P. Wan, H. Wang, D. Xie, X. Wang, and G. Yan. 2016. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. Remote Sens. 8:501. doi:10.3390/rs8060501