Elucidating Sorghum Biomass, Nitrogen and Chlorophyll Contents With Spectral and Morphological Traits Derived From Unmanned Aircraft System

Jiating Li1, Yeyin Shi1*, Arun-Narendhiran Veeranampalayam-Sivakumar1 and Daniel P. Schachtman2

1 Department of Biological Systems Engineering, University of Nebraska-Lincoln, Lincoln, NE, United States, 2 Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, United States

Unmanned aircraft systems (UAS) provide an efficient way to phenotype crop morphology with spectral traits such as plant height, canopy cover and various vegetation indices (VIs) providing information to elucidate genotypic responses to the environment. In this study, we investigated the potential use of UAS-derived traits to elucidate biomass, nitrogen and chlorophyll content in sorghum under nitrogen stress treatments. A nitrogen stress trial located in Nebraska, USA, contained 24 different sorghum lines, 2 nitrogen treatments and 8 replications, for a total of 384 plots. Morphological and spectral traits including plant height, canopy cover and various VIs were derived from UAS flights with a true-color RGB camera and a 5-band multispectral camera at early, mid and late growth stages across the sorghum growing season in 2017. Simple and multiple regression models were investigated for sorghum biomass, nitrogen and chlorophyll content estimations using the derived morphological and spectral traits along with manual ground truthed measurements. Results showed that, the UAS-derived plant height was strongly correlated with manually measured plant height ($r = 0.85$); and the UAS-derived biomass using plant height, canopy cover and VIs had strong exponential correlations with the sampled biomass of fresh stalks and leaves (maximum $r = 0.85$) and the biomass of dry stalks and leaves (maximum $r = 0.88$). The UAS-derived VIs were moderately correlated with the laboratory measured leaf nitrogen content ($r = 0.52$) and the measured leaf chlorophyll content ($r = 0.69$) in each plot. The methods developed in this study will facilitate genetic improvement and agronomic studies that require assessment of stress responses in large-scale field trials.

Keywords: UAV, phenotyping, multispectral, nitrogen stress, biomass, chlorophyll, plant height, canopy cover

INTRODUCTION

Following rice, wheat, corn, and barley, sorghum is the fifth most important cereal crop worldwide (Ramatoulaye et al., 2016). It is widely used in human consumption, animal feed, and biofuel production (Stanton et al., 2017). As reported, in 2016, the sorghum production in the U.S. was about 12.2 million tons which is approximately 20% of the world sorghum production.
Sorghum (FAOSTAT, 2017). Serving as the biomass crop for biofuel production, sorghum has the advantages of an annual growth cycle, high caloric value, and low management cost (Fernandes et al., 2018). An efficient and timely method for the prediction of sorghum biomass will help to speed the development of higher biomass varieties. The benefits of sorghum as a biomass crop could be further enhanced if genotypes with high tolerance to stresses such as reduced nitrogen or water deficit can be more easily identified, which will be facilitated by integrating sorghum genotyping and phenotyping technologies.

In the past decade, gene sequencing technology has advanced, allowing the crop genomic information to be collected much easier and more cheaply (Furbank and Tester, 2011). However, genomic selection is still hampered by the speed and ease of obtaining large amounts of phenotypic information. Traditionally, in-field phenotyping has been conducted manually, which consumes a great deal of labor and time. High-throughput phenotyping technology developed in recent years opens opportunities to automate and speed up breeding pipelines. Depending on the traits of interest and growth stages, high-throughput phenotyping can be conducted either in the lab or in the field. For the field-based phenotyping, the ground-based systems and the aerial-based systems usually work as complementary platforms to achieve the final goal of rapid and accurate trait collection. Ground-based systems such as the gantry systems (Virlet et al., 2017), cable-suspended systems (Kirchgessner et al., 2017) and mobile cart or robotic systems (Svensgaard et al., 2014) conduct proximal sensing over or under the plant canopy with little limitation on sensor weight or size. The aerial high-throughput phenotyping usually implemented with unmanned aircraft systems (UAS) operated at low altitudes which have limited sensor payloads or weight and can only detect traits remotely over the canopy. However, they are capable of covering a larger area in a shorter period of time which minimizes the measurement error caused by changes in environmental factors, and are independent of the soil condition which may hamper movement of ground based systems. Typical types of UAS are fixed-wing, rotary-wing, and hybrid systems. A rotary-wing platform was selected in this study to conduct slow speed, low altitude and more stable phenotypic data collection for sorghum.

UAS technology has been widely used to study various traits in different crops including sorghum. Morphological traits are often measured from natural color images, i.e., RGB images, or estimated from spectral images. Sorghum and corn plant height is a trait that has been investigated in several studies using the structure from motion technique and RGB images (Shi et al., 2016; Watanabe et al., 2017; Hu et al., 2018; Malambo et al., 2018; Pugh et al., 2018). Sorghum ground cover (Duan et al., 2017; Potgieter et al., 2017; Shafian et al., 2018) and leaf area index (Potgieter et al., 2017; Shafian et al., 2018) were directly calculated from the RGB images or estimated using spectral information from multispectral camera. Visible morphological traits are easier to measure than physiological traits, such as chlorophyll content, nitrogen concentration, and water content. The physiological traits are often hard to be assessed by the human eye but can be detected in the infrared spectra and the variations in those important traits become more obvious if they are depicted using vegetation indices (VIs). For example, normalized difference red edge (NDRE) was used to differentiate stay-green and senescent lines in sorghum breeding (Potgieter et al., 2017). Sorghum grain yield was well correlated with the normalized difference vegetation index (NDVI) derived from a modified three-band camera (green, red, and NIR) (Stanton et al., 2017) and a multispectral camera (Shafian et al., 2018). Sorghum panicle volume was estimated from RGB orthomosaic, DSM and point cloud (Chang et al., 2017a).

Biomass and nitrogen status of sorghum is particularly important for the development of new higher yielding nitrogen use efficiency sorghum varieties for lignocellulosic production. While grain sorghum is easy to harvest, energy sorghum is not because of its very large size. The main interest of growing energy sorghum is in the biomass which may be used to produce cellulosic ethanol. The crop is over 3.5 to 4.5 m high and specialized equipment which is not usually readily available is required for harvest. Therefore, the use of UAS to estimate biomass and nitrogen status of sorghum provides a highly efficiency way for breeders to improve the crop. Most of the UAS related studies on sorghum were focused on plant height, ground cover, leaf area index and grain yield estimation so far. The only study we found for sorghum biomass estimation was using a UAS based hyperspectral and RGB system and machine learning modeling (Zhang et al., 2017). The results were promising which inspired us to move forward to investigate alternative low-cost method based on multispectral and RGB cameras for sorghum biomass, nitrogen, and chlorophyll content estimation. As for sorghum nitrogen and chlorophyll estimation, most of the previous studies focused on qualitative differentiation between treatments such as high and low nitrogen treatment or stay-green and senescent lines; while no study was found that investigated the quantitative relationship between sorghum nitrogen or chlorophyll content and UAS-derived traits.

The objective of this study was to investigate the potential of using UAS-derived multispectral and morphological traits for sorghum biomass, nitrogen and chlorophyll content estimates. Three specific objectives were:

1. Obtain sorghum spectral and morphological traits from UAS based remote sensing, including various vegetation indices, plant height and canopy cover;
2. Establish predictive models for sorghum biomass, nitrogen and chlorophyll contents using the obtained morphological and spectral traits; and
3. Evaluate how predictions compared with ground truth measurements.

MATERIALS AND METHODS

A flowchart has been provided in Figure 1, summarizing the main steps of this study: image data collection, image preprocessing, morphological and spectral trait extraction, and statistical analysis.
**Field Experimental Design**

Field location and plot design—The field experiment was conducted over a 1.38 ha sorghum nitrogen stress trial located in Central City, Nebraska, US (41°12′3.0″ N, 97°56′40.56″ W), in the growing season of 2017. The field was planted on May 26, 2017 with 24 sorghum lines (see Supplementary Table 1) in two nitrogen treatments (for the low nitrogen treatment no nitrogen was applied and for the high treatment 85 pounds of nitrogen per acre were added) and eight replications in a randomized complete block design (Figure 2). Each plot was 3 m by 3 m containing four rows with 0.10 m within-row spacing and 0.76 m row spacing delineated by a rectangular in Figure 2. This field trail was located on a commercial farm with center pivot irrigation. Nine inches of irrigation was added contain 0.9 ppm nitrate.

Collection of hyperspectral radiometer data—The most recently fully expanded leaf was taken from two random plants within a plot. The ASD FieldSpec 4 Standard Res (Analytical Spectral Devices, Colorado) was then used to measure hyperspectral reflectance readings in the 350 to 2,500 nm wavelength range of the leaf. The methods used here for the ASD readings are described in Yendrek et al. (2017). Both
The ends of the leaf were removed leaving roughly six to eight inches of the mid-section of the leaf. The leaf tissue on both sides of the midrib were removed from the midrib. One side was put into a paper envelope and dried at 50°C for nitrogen analysis and the other was placed into an aluminum foil packet which was then put on dry ice and subsequently stored at −80°C.

Laboratory analysis of leaf tissue—Frozen leaf issue was removed from −80°C and placed on dry ice. In a darkened room leaf disks (6 mm diameter) were punched from frozen leaves in the weigh boat, on dry ice until there was approximately 90 mg of leaf tissue which was about 18 to 27 disks. Prior to extraction 2.5 mL of 100% methanol was added to 15 mL tubes. Three replicates containing 30 mg or six to nine disks were taken and submerged in the methanol. The tubes were then placed in a rack in the dark and placed on a rotary shaker at 250 RPM for 24 h. After 24 h 200 mL of each sample was used to fill a 96 well black sided plate (Corning™ Costar) with a clear flat bottom, with one of the wells being filled with 200 µL of 100% methanol to be used as the blank. The top of the plate is secured with sealing film to prevent methanol evaporation. The plate was then read three times at 666 nm for chlorophyll A, 653 nm for chlorophyll B, 470 nm for carotenoids using a BioTek Synergy H1 Hybrid Reader. The chlorophyll/methanol equation (Lichtenthaler and Wellburn, 1983) was then used in to calculate chlorophyll A, chlorophyll B, and carotenoids of the extracts. The average of the three replicates was calculated for each sample. For nitrogen analysis the leaves were roughly chopped with a stainless-steel scissors in the envelopes and then sent to Ward Labs (Kearney, NE) for analysis of total nitrogen.

Field ground truth measurements—Plant height was measured on September 7, 2017 and October 9, 2017 from all eight replicates for each treatment. Plant heights were measured as the average height of plants in one of the center rows of the 4-row plot. It was measured on plants in the middle of the 3-meter row with a telescoping measuring stick which allows you to look up to align the top of the stick with top of the plants then record the height at eye level. Total above ground biomass was harvested on October 9, 2017 at which time fresh and dry above ground biomass (leaf and stem) were sampled in 363 plots. Eight replicates from each treatment were measured. A 0.91 m section of row in the middle of the plots was identified and plants were cut down at the bottom of the stem at soil level. Stalks with leaves and panicles were weighed and recorded separately on scales in the field but only the weight of stalks with leaves were used as the fresh biomass in this study. Subsamples of three stalks with leaves were reweighed to get the fresh weight and then bagged, oven dried and used to calculate the dry to fresh weight ratios which was then used to calculate the dry weights of the plots.

UAS, Sensors and Flights

The system used for image capture was a Matrice 600 Pro multi-rotor UAS platform (DJI, Shenzhen, China), equipped with a Zenmuse X3 RGB camera (DJI, Shenzhen, China) and a multispectral camera RedEdge (MicaSense, Seattle, UAS). The RGB camera has 4000 by 2250 effective pixels. The multispectral camera system has five spectral bands blue, green, red, red edge, and near infrared (Table 1), each with 1280 by 960 effective pixels, and a downwelling light sensor system installed horizontally on top of the UAS used to measure the environmental irradiance and post-calibrate reflectance readings. As another source of radiometric calibration data, a standard calibration panel came with the multispectral camera was imaged on the ground before or after each flight.
TABLE 1 | Center wavelength and full width at half maximum (FWHM) bandwidth of each spectral band of the RedEdge multispectral camera.

| Spectral band | Center wavelength (nm) | Bandwidth FWHM (nm) |
|---------------|------------------------|---------------------|
| Blue          | 475                    | 20                  |
| Green         | 560                    | 20                  |
| Red           | 668                    | 10                  |
| Red edge      | 717                    | 10                  |
| Near infrared | 840                    | 40                  |

Three flights were conducted on July 17, August 19, and September 11 in 2017. Flights were auto-piloted using DJI GO and DJI GS Pro applications with 92% forward and side overlap between images at 30 m above ground level. The resulting ground sampling distance (GSD) was 1.3 cm/pixel for RGB image and 2.0 cm/pixel for the multispectral image. The flight altitude and image acquisition parameters were tested and determined to optimize the flight duration and the quality of mosaicked maps. Eleven ground control points (GCPs) were distributed along the edges and inside the field each time before the flight for geometric calibration in image processing. Their geo-coordinates were accurately measured by a survey-grade RTK-GPS with less than 3 cm level accuracy.

Image Pre-processing
Two main tasks were completed in the pre-processing stage: the orthomosaic map and digital surface model (DSM) generation with proper geometric and radiometric calibrations; and the plot delineation to prepare for later processing.

RGB images were mosaicked by Pix4Dmapper software (Pix4D, Lausanne, Switzerland). Basically, there were three steps in Pix4Dmapper: Initial Processing; Point Cloud and Mesh; DSM, Orthomosaic and Index. In the initial processing, RGB raw images were imported into to extract and match key-points among neighboring images to form a rough mosaic. Geo-coordinates of the centers of GCPs were imported for geometric calibration and improving the initial mosaicking to form 3D point cloud and mesh. The final outputs were the 2D orthomosaic and DSM.

Multispectral images were mosaicked by Atlas Cloud service (MicaSense, Seattle, USA). Radiometric calibration was automatically addressed during this process using the irradiance measured in the field from the standard calibration panel and downwelling light sensor. The five-layer, 16-bit GeoTIFF output from Atlas was converted to five-layer reflectance GeoTIFF following the sensor instruction with a pixel value of 32768 equal to 100% reflectance. The multi-layer reflectance orthomosaic generated from the multispectral images were used later to estimate canopy cover and calculated various vegetation indices.

In order to conduct plot-based analysis, each plot boundary was delineated in the multispectral orthomosaic and DSM maps with unique plot ID as shapefiles in ArcGIS (Figure 2). The shapefiles, DSMs and multispectral orthomosaics were exported to R software for further data analysis.

Shaded area would affect the reflectance recorded by camera, which was more pronounced when the plant was bigger. For the data set collected on the last date when sorghum plants were at their maximum height, more shadows were cast over neighboring shorter vegetation rows and soil. This was noticed in the multispectral images where shadowed vegetation and soil pixels had abnormally higher VI values than sunlit vegetation pixels (Woebbecke et al., 1995). Those pixels with abnormal VI values were filtered out in this study and only sunlit vegetation pixels were used for VI calculation in each plot. To eliminate the shadow interference on plant VI calculation, the ExG index (Table 2) map, which was applied in other studies in distinguishing vegetative areas from soil or residue background (Woebbecke et al., 1995), was calculated. In this study, both the soil and shaded vegetation pixels in the ExG index map had lower values than the sunlit vegetation pixels (Figure 3C) so that they were filtered out and only the sunlit vegetation pixels were retained for VI calculation. A threshold of 0.046 was determined by trial and error and used to segment the vegetation pixels from the soil pixels in the ExG map (Figure 3D). The segmented vegetation pixels formed a mask which was applied in the further processing for ground cover estimate and VI calculations.

Morphological and Spectral Traits Extraction
Morphological and spectral plant traits were extracted from the pre-processed data to estimate sorghum biomass and nitrogen and chlorophyll contents, including plant height, canopy cover, and various VIs at the individual plot level.
Plant height was derived by subtracting the digital terrain model (DTM) from the digital surface model (DSM). The DSM was generated along with the orthomosaic from mosaicking the RGB images in Pix4D, and was geometrically calibrated with the GCP’s coordinates surveyed by the RTK-GPS during field data collections. The DTM was generated by linearly interpolating the soil surface on east and west side of the field, assuming little elevation changes within this 1.38 ha field. The derived plant height map had the same spatial resolution of 1.3 cm as the RGB orthomosaic. The 90th, 93rd, 95th, and 98th percentiles of all pixels falling into a plot boundary were calculated and compared with the manually sampled plant height in the same plot to find the one with highest correlation. After shadow removal, canopy cover was calculated as the ratio of the number of segmented sunlit vegetation pixels to the total number of pixels in a plot (Lee and Lee, 2011).

Calibrated reflectance in each multispectral band was extracted, and various VISs were calculated for each plot by averaging the VI values of all pixels of interest within the plot boundary. These VIs included normalized difference vegetation index (NDVI), renormalized difference vegetation index (RDVI) (Roujean and Breon, 1995), green normalized difference vegetation index (GNDVI) (Gitelson et al., 1996), green chlorophyll index (CI_{Green}) and red edge chlorophyll index (CI_{RedEdge}) (Schlemmera et al., 2013), normalized difference red edge index (NDRE) (Fitzgerald et al., 2006), and RGB vegetation index (RGBVI) (Bendig et al., 2015) (Table 2). NDVI is one of the most commonly used indices for estimating crop physiological traits such as chlorophyll. RDVI uses the same spectral bands as NDVI; however, RDVI is less sensitive to the variation of soil background. Given the saturation problem of NDVI after canopy closure, RDVI may be considered superior (Fu et al., 2013). GNDVI was found to have wider dynamic range than NDVI and is more sensitive to chlorophyll-a concentration (Gitelson et al., 1996). Similar to NDVI, NDRE is a good indicator of chlorophyll or nitrogen status (Fitzgerald et al., 2006); however, the replacement of red band with the red edge band makes NDRE more sensitive to the biomass change than NDVI after canopy closure. Using bands in the visible spectra, RGBVI can be used to estimate biomass (Bendig et al., 2015).

**Statistical Modeling for Biomass, Nitrogen, and Chlorophyll Contents**

Since the biomass was sampled late in the season close to the last flight date, only the UAS data collected on September 11, 2017 was used for the biomass analysis. To estimate the fresh and dry biomass using remotely sensed plant traits, simple exponential regression (SER) models were first built using univariate morphological or spectral trait. Given that the biomass is intuitively related to multiple traits such as plant height, and stalk diameter, it is also worth investigating the integration of more than one trait using multiple exponential regression (MER) models to see if the estimation of biomass can be improved. To select predictors for the MER models, a correlation matrix was first calculated to avoid including predictors that were highly correlated (Table 3). In this study, Pearson’s correlation coefficient r was used and the strength of the correlation was determined as summarized by Asuero et al. (2006): r ranging from 0 to 0.29 was interpreted as little if any correlation, r ranging from 0.30 to 0.49 was regarded as low correlation, and r ranging from 0.50 to 0.69 was moderate correlation, and r greater than 0.69 was high to very high correlation.

**TABLE 2 | Formulas of vegetation indices used in this study.**

| VIs     | Formula                                                                 | Feature/application                                                                 |
|---------|-------------------------------------------------------------------------|------------------------------------------------------------------------------------|
| ExG     | 2 × (Pgreen − PRed − PBlue)                                            | Distinguishes vegetation from soil background                                      |
| NDVI    | \(\frac{(P_{\text{NIR}} − P_{\text{Red}})}{(P_{\text{NIR}} + P_{\text{Red}})}\) | Correlates with green biomass, chlorophyll                                        |
| RDVI    | \(\frac{(P_{\text{NIR}} − P_{\text{Red}})}{\sqrt{(P_{\text{NIR}} + P_{\text{Red}})}}\) | Less sensitive to the interfering effects of soil background                        |
| GNDVI   | \(\frac{(P_{\text{NIR}} − P_{\text{Green}})}{(P_{\text{NIR}} + P_{\text{Green}})}\) | Correlates with Chlorophyll-a                                                    |
| CI_{Green} | \(\frac{(P_{\text{NIR}})}{(P_{\text{Green}})} \times 1\) | Correlates with chlorophyll and nitrogen                                           |
| CI_{RedEdge} | \(\frac{(P_{\text{NIR}})}{(P_{\text{RedEdge}})} \times 1\) | Correlates with chlorophyll and nitrogen                                           |
| NDRE    | \(\frac{(P_{\text{NIR}} − P_{\text{RedEdge}})}{(P_{\text{NIR}} + P_{\text{RedEdge}})}\) | Correlates with chlorophyll or nitrogen                                            |
| RGBVI   | \(\frac{(P_{\text{Green}})}{(P_{\text{Red}})} \times \frac{P_{\text{Red}}}{P_{\text{Blue}}} \times P_{\text{Red}}\) | Estimates biomass                                                                  |

*p* means spectral reflectance.

**TABLE 3 | Correlation matrix of candidate predictors in biomass prediction.**

| Plant height | Canopy cover | NDVI | NDRE | RDVI | RGBVI |
|--------------|--------------|------|------|------|-------|
| Plant height | 1            |      |      |      |       |
| Canopy cover | 0.61         | 1    |      |      |       |
| NDVI         | 0.74         | 0.54 | 1    |      |       |
| NDRE         | 0.68         | 0.25 | 0.90 | 1    |       |
| RDVI         | 0.80         | 0.57 | 0.68 | 0.89 | 1     |
| RGBVI        | 0.54         | 0.75 | 0.64 | 0.26 | 0.59  | 1 |

r ranging from 0 to 0.29 was interpreted as little if any correlation (highlighted in blue), r ranging from 0.30 to 0.49 was regarded as low correlation, and r ranging from 0.50 to 0.69 was moderate correlation (highlighted in yellow), and r greater than 0.69 was high to very high correlation.
from the 10-fold cross validation. The established regression models were further tested using the testing set, and were evaluated using the Pearson correlation coefficients ($r$) (Equation 3) and RMSE.

$$ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - M_i)^2} $$

$$ \text{STD} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} $$

$$ r = \frac{\sum_{i=1}^{n} (M_i - \bar{M})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{n} (M_i - \bar{M})^2 \sum_{i=1}^{n} (P_i - \bar{P})^2}} $$

where $n$ is the number of samples, $P_i$ stands for predicted value, $M_i$ stands for manually measured value, $\bar{P}$ is the mean of predicted values, $\bar{M}$ is the mean of manually measured values. $x_i$ is the observed values, and $\bar{x}$ is the mean value of these observations.

In order to examine the effect of nitrogen treatments, $t$-tests were conducted using three VIs (CI$_{\text{Green}}$, CI$_{\text{RedEdge}}$, and NDRE) calculated from the last flight (September 11, 2017) between high nitrogen (192 plots) and low nitrogen (192 plots) treatments.

Furthermore, to evaluate the relationship between various VIs and sorghum chlorophyll and nitrogen contents, $r$ was calculated between the various VIs and the manually measured leaf chlorophyll and nitrogen contents for sampled sorghum plants from each plot. For chlorophyll, 70 plots in July, 68 plots in August, and 112 plots in September had valid samples (250 samples in total); for nitrogen, 50 plots in July, 50 plots in August, and 69 plots in September had valid samples (169 samples in total).

RESULTS

Plant Height Estimation

Compared to other percentile values, a stronger linear correlation was obtained ($r = 0.85$) between the 90th percentile of estimated plant height from the RGB orthomosaic and the manually sampled plant height in 363 plots in September 2017 (Figure 4). The RMSE was 49.8 cm and the $r$ was 0.85 between UAS derived plant height and manually measured plant height. The coefficient of variation (CV) for aerial data estimated plant height was 32.11% and for manually measured plant height was 27.92%.

Fresh and Dry Biomass Estimation

The correlations between a single UAS-derived morphological or spectral trait and the manually sampled fresh or dry biomass tended to be exponential rather than linear in this study (Figure 5).

For the simple exponential model of fresh biomass (Table 4), plant height gave higher correlation than canopy cover ($r = 0.81$), whereas RDVI and NDVI provided better results than the other VIs ($r = 0.83$ for RDVI, and $r = 0.80$ for NDVI).

Slightly better correlations were obtained when multiple traits were combined into the fresh biomass regression model, with the outcome being that $r$ was greater than 0.80 for all combinations. Interestingly, when used individually in the simple exponential models, either NDRE or RGBVI resulted in lower correlations ($r = 0.66$ for NDRE, $r = 0.57$ for RGBVI); however, the combination of them using the multiple exponential model largely improved the correlation with fresh biomass ($r = 0.82$). Similar results were found in prediction of dry biomass. The morphological trait plant height ($r = 0.87$), spectral traits RDVI ($r = 0.78$) and NDVI ($r = 0.78$) individually exhibited better correlations with the dry biomass using simple exponential models; while the combination of them with other traits did not significantly improve the results in this case (Table 5).

Nitrogen Treatment Effect on UAS-Derived Traits

As shown in Table 6, significant differences were found in the remotely sensed three VIs between high and low nitrogen treatments ($p < 0.0001$). It is shown in Figure 6 that the nitrogen effect can be clearly distinguished with the three selected VIs—CI$_{\text{Green}}$, CI$_{\text{RedEdge}}$, and NDRE—derived from the late season UAS data.

Chlorophyll and Nitrogen Content Estimation

Moderate to strong correlations ($r > 0.5$) were found between the chlorophyll content of leaf samples and the corresponding VIs, except NDVI, of same plots calculated from multispectral aerial data over July, August, and September in 2017 (Table 6). Similar results were found for the nitrogen content (Table 7). These VIs included CI$_{\text{Green}}$, CI$_{\text{RedEdge}}$, GNDVI, and NDRE. VIs were also calculated from the leaf-level hyperspectral measurements sampled in the same plots. Similar correlations ($r > 0.4$) were found when the specific spectral bands were taken from the hyperspectral radiometer data as those calculated from the multispectral aerial data (Table 7). Although correlation using NDVI improved with the hyperspectral radiometer, the
FIGURE 5 | Exponential correlations between some UAS-derived traits and the manually sampled biomass of stalks and leaves (fresh or dry), over 363 plots: (A) RDVI and fresh biomass, (B) plant height and fresh biomass, (C) RDVI and dry biomass, (D) plant height and dry biomass.

NDVI index had lower correlations with chlorophyll and nitrogen contents than other VIs, while NDRE remained the VI with highest correlation with nitrogen and chlorophyll in this case.

**DISCUSSION**

The moderate to strong correlations ($r$ varied from 0.55 to 0.88) found between the UAS-derived plant morphological and spectral traits and the sorghum late-season biomass, nitrogen, and chlorophyll contents in this study indicates that UAS should be useful for phenotyping. Compared with the hyperspectral reflectance that was manually sampled at the leaf level using hyperspectral radiometer with a leaf clip, the UAS-derived VIs using the five-band multispectral camera resulted in similar correlations with nitrogen and chlorophyll contents when the same VIs were calculated from the measured hyperspectral reflectance (Table 7). This not only demonstrates the fidelity of the UAS-based remote spectral sensing, but also indicates the potential for scaling up the high-throughput phenotyping from ground-based leaf level to UAS-based canopy level assessment.

When a single trait was used for prediction with simple exponential regression models, estimated plant height, RDVI, and NDVI indices individually had the strongest correlations with both fresh and dry sorghum biomass among the various remotely sensed traits. The high and robust correlation derived from plant height was also found in a previous study in barley (Bendig et al., 2015). Interestingly, the NDRE index showed a little lower correlation to fresh and dry biomass ($r$ within 0.66 and 0.70) than NDVI and NDRE but significantly outperformed NDVI in chlorophyll and nitrogen content estimations (Table 7) which showed NDRE’s known advantage over NDVI after canopy closure due to the saturation in the red spectral band at the mid to late growth stages (Mutanga and Skidmore, 2004). Late season canopy cover had moderate correlations with biomass ($r$ within 0.59 and 0.62) but was not shown to be superior to vegetation indices in our study. The inferior correlation of RGBVI index with biomass compared with other traits was also reported in barley (Tilly et al., 2015). However, it is noteworthy that, if no near-infrared spectral data and only the RGB information was available, the UAS-derived RGBVI index alone still provided low to moderate correlations with fresh and dry biomass ($r$ within 0.48 and 0.57).
TABLE 4 | Fresh biomass estimation results in 10-fold cross validation as well as in the testing set, based on simple exponential regression (SER) and multiple exponential regression (MER) models.

| Model | Predictors                        | Training set: average of 10-fold cross validation | Formula * | Testing set |
|-------|-----------------------------------|--------------------------------------------------|-----------|-------------|
|       | RMSE (kg) | STD (kg) | Y $= 18.70 \times e^{0.004 \times PH + 0.246 \times CC}$ | RMSE (kg) | r           |
| SER   | Plant height                        | 23.78   | 4.06                              | 26.25     | 0.81        |
|       | Canopy cover                       | 32.30   | 5.16                              | 35.78     | 0.62        |
|       | NDVI                               | 23.67   | 3.11                              | 25.89     | 0.80        |
|       | NDRE                               | 28.19   | 3.40                              | 31.88     | 0.66        |
|       | RDVI                               | 18.10   | 3.60                              | 24.26     | 0.83        |
|       | RGBVI                              | 32.73   | 4.69                              | 36.79     | 0.57        |
| MER   | Plant height and canopy cover      | 23.73   | 4.10                              | 25.98     | 0.81        |
|       | NDVI and RGBVI                     | 23.69   | 3.11                              | 25.62     | 0.80        |
|       | NDRE and RGBVI                     | 23.89   | 3.57                              | 24.92     | 0.82        |
|       | Plant height and NDRE              | 23.32   | 3.81                              | 23.70     | 0.84        |
|       | Plant height and RGBVI             | 23.31   | 3.92                              | 25.54     | 0.81        |
|       | Canopy cover and NDVI              | 23.01   | 3.28                              | 24.18     | 0.82        |
|       | Canopy cover and NDRE              | 24.35   | 3.89                              | 25.76     | 0.80        |
|       | Canopy cover and RDVI              | 23.10   | 3.81                              | 23.81     | 0.83        |
|       | Plant height, canopy cover, and NDRE | 22.82  | 3.89                              | 24.47     | 0.83        |

1STD is standard deviation of 10 RMSE values of the total plot weight from the 10-fold cross validation; Y, predicted dry biomass (kg/plot); PH, plant height; CC, canopy cover; NDVI, normalized difference vegetation index; NDRE, normalized difference red edge index; RDVI, renormalized vegetation index; and RGBVI is the RGB vegetation index.

TABLE 5 | Dry biomass estimation results in 10-fold cross validation as well as in the testing set, based on simple exponential regression (SER) and multiple exponential regression (MER) model.

| Model | Predictor(s)                        | Training set: average of 10-fold cross validation | Formula * | Testing set |
|-------|-----------------------------------|--------------------------------------------------|-----------|-------------|
|       | RMSE (kg) | STD (kg) | Y $= 5.63 \times e^{0.004 \times PH}$ | RMSE (kg) | r           |
| SER   | Plant height                        | 4.83    | 0.70                              | 4.89      | 0.87        |
|       | Canopy cover                       | 7.20    | 0.93                              | 7.89      | 0.59        |
|       | NDVI                               | 5.73    | 0.67                              | 5.92      | 0.78        |
|       | NDRE                               | 6.34    | 0.78                              | 6.70      | 0.70        |
|       | RDVI                               | 5.58    | 0.69                              | 5.83      | 0.78        |
|       | RGBVI                              | 7.51    | 0.74                              | 8.41      | 0.48        |
| MER   | Plant height and canopy cover      | 4.82    | 0.70                              | 4.88      | 0.87        |
|       | NDVI and RGBVI                     | 5.74    | 0.67                              | 5.92      | 0.78        |
|       | NDRE and RGBVI                     | 5.73    | 0.68                              | 5.74      | 0.79        |
|       | RDVI and RGBVI                     | 5.58    | 0.69                              | 5.82      | 0.78        |
|       | Plant height and NDRE              | 4.81    | 0.68                              | 4.78      | 0.87        |
|       | Plant height and RGBVI             | 4.82    | 0.68                              | 4.87      | 0.87        |
|       | Canopy cover and NDVI              | 5.59    | 0.53                              | 5.71      | 0.79        |
|       | Canopy and NDRE                    | 5.52    | 0.56                              | 5.40      | 0.82        |
|       | Canopy cover and RDVI              | 5.49    | 0.69                              | 5.74      | 0.78        |
|       | Plant height, canopy cover, and NDRE | 4.76  | 0.67                              | 4.68      | 0.88        |

1STD is standard deviation of 10 RMSE values of the total plot weight from the 10-fold cross validation; Y, predicted dry biomass (kg/plot); PH, plant height; CC, canopy cover; NDVI, normalized difference vegetation index; NDRE, normalized difference red edge index; RDVI, renormalized vegetation index; and RGBVI is the RGB vegetation index.

When several traits with multiple exponential regression models were used, similar correlations were achieved for fresh and dry biomass prediction as compared to the results derived from single traits in this study. The correlation using several traits was best when using the traits that had the strongest correlations individually, i.e., the plant height and RDVI indices in this case.
Similar data fusion models that were investigated in previous studies varied in the ability to predict biomass over the single metric predictions and this depended on the traits that were added at different growth stages and the correlations between the traits. The integration of RGBVI and plant height resulted in small improvement in the biomass prediction in barley ($r$ from 0.89 to 0.92) but no improvement in the biomass prediction was found with the integration of plant height and other VIs derived from visible and near-infrared spectra (Bendig et al., 2015). This can probably be explained by the moderate to high correlations between the plant height and various VIs found in this study (Table 3).

Further work will be needed to improve biomass prediction through the inclusion of additional morphological traits such as stalk diameter and additional or customized spectral bands. Some other traits that were not included in this study also showed ability to increase the estimation accuracy when combined with some of the UAS-derived traits in this study. When adding manually measured stem diameter on the UAS-derived plant height, the biomass prediction using a volumetric cylinder equation in corn was significantly improved ($r$ from 0.56 to 0.93) (Varela et al., 2017). However, automating stem diameter measurements is challenging and may not be useful in many energy sorghum varieties that do not flower in North America (SD Kresovich, personal communication). Another case would be the combination of hyperspectral canopy reflectance and plant height which improved the accuracy of estimating winter wheat biomass ($r$ from 0.73 to 0.91) (Yue et al., 2017). In addition, customization of spectral bands of the UAS-based multispectral sensor based on the feature spectral bands derived from the leaf-level or ground-based hyperspectral sensing (Yendrek et al., 2017) for specific applications can scale up the throughput of phenotyping capabilities in the field while reducing the sensor instrument cost.

 Improvement can also be achieved by including temporal data during the growing season and using more sophisticated statistical models. The late-season biomass predicted by single or multiple UAS-derived traits had strong exponential correlations with the sampled fresh biomass (maximum $r = 0.84$) and dry biomass (maximum $r = 0.88$); however, no significant improvement was found if multiple traits collected on the same date were used to build models (Tables 4, 5). Similar results were reported in study on sorghum biomass prediction (Zhang et al., 2017) showing that more data on additional traits measured on the same day provided no significant improvement for biomass prediction. However, significant improvements were found when measurements from multiple time points with either a single trait or multiple traits were used (Zhang et al., 2017). Also, the exponential relationships found between UAS-derived traits and biomass in this study were similar to a previous finding in barley (Bendig et al., 2014).

We recommend segmenting vegetation pixels and shaded soil pixels for some VI calculations to avoid the interference of shaded soil pixels in the calculations. Shadows that were cast on the canopy and soil were identified in this study to have much higher NDVI values than the sunlit vegetation pixels (Figure 3). If an averaged VI value was calculated for all pixels encompassed within a plot boundary, plots with more shaded vegetation and soil areas may result with higher VI values than those without much shaded areas even though the NDVI of the actual leaves may be lower. Shading may not be a problem for production...
agricultural applications when the whole field was planted with same variety and population; however, in the application of phenotyping when many small plots that contained different varieties or uneven stands this cause substantial errors in estimating the true canopy VIs. Accurate segmentation of sunlit vegetation and shaded vegetation and soil pixels in the image processing is important to ensure the reliability of VI values. In this study, we used the ExG index map for segmentation which was effective but still resulted with some mis-classification of shaded vegetation and shaded soil pixels. This also resulted in a lower estimation of canopy cover especially for those plots with significant shadows. Future research will be needed to investigate the hyperspectral reflectance patterns of the sunlit and shaded vegetation and soil pixels and corresponding classification algorithms with proper band selection techniques (Sun et al., 2015) to customize multispectral cameras for high-throughput applications.

The results obtained using UAS-derived DSM to estimate plant height are very promising ($r = 0.85$) and are similar to other studies (Geipel et al., 2014; Chang et al., 2017b; Watanabe et al., 2017; Hu et al., 2018; Malambo et al., 2018; Pugh et al., 2018) Considering the strong correlation between the late-season plant height and the sampled fresh and dry biomass ($r = 0.81$), this method can be used to quickly estimate plant biomass. The UAS-derived plant height in this study was only investigated using the data collected on a single date late in the season since the main purpose was end-of-season biomass prediction. The accuracy of the height estimation achieved in this study (RMSE = 49.8 cm) needs to be improved in order to be applied to plant height estimation in earlier growth stages when plants are smaller. Ideally, the structure from motion (SfM) algorithm used behind this technology to generate the point clouds or the structure of a targeted object can achieve reprojection error at only about one pixel (Snavely et al., 2008) which in our case would be about 1.3 cm accuracy. However, in real world agricultural applications, errors are induced due to the movement of the plant canopy by the wind (Chang et al., 2017b). During our data collections, the wind speed was 5 m/s (10 mph) which caused the top canopy to sway at decimeter level. This rendered the SfM algorithm difficult to use because matched keypoints among images taken from different angles were hard to find and therefore errors were generated. Moreover, variation in leaf angle, canopy structure and presence or absence of panicles among genotypes caused the discrepancy between manually sampled plant height and plant height estimated from the UAS-derived point clouds data.

In addition, the accuracy of UAS-derived plant height also depends on the accuracy of the derived digital terrain model (DTM) or the elevation of the field (Malambo et al., 2018). In this study, we assumed a constant elevation change of the field since the field was flat and relatively small (1.38 ha); however, ignoring the within-field unevenness may induce a small amount of error. A pre-planting elevation mapping can largely reduce such errors.

Improving prediction accuracy to develop a more generally applicable model for energy sorghum will be a future goal. The models developed in this study were based on a single season and so they may be further improved through the incorporation of multi-season and multi environment data. Incorporating additional data sets such as growing degree days, precipitation, soil physicochemical properties, planting dates and other agronomic practices may all allow for the further improvement of predictive models using data from UAS.

**AUTHOR CONTRIBUTIONS**

JL, A-NV-S, YS conducted the aerial data collections. DS provided the field, all the ground truth and hyperspectral radiometer data. JL and YS performed the aerial data analysis and wrote the manuscript. DS edited and reviewed the manuscript. All authors reviewed the manuscript and agreed with the submission.

**ACKNOWLEDGMENTS**

We gratefully acknowledge startup funds provided by the Agricultural Research Division and the Office of Research and Economic Development of the University of Nebraska-Lincoln, provided to YS, and a grant from Office of Science (BER), U.S. Department of Energy, Grant no DE-SC0014395 to DS. In addition, the authors wish to thank Stephanie Futrell and Bert Devilbiss who were involved in the field ground truth data collection, lab analysis, and hyperspectral data collection and analysis and the laboratories of Dr. Stephen Kresovich (Clemson) and Dr. Ismail Dweikat (UNL) who provided seed for these experiments.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01406/full#supplementary-material

**REFERENCES**

Asuero, A. G., Sayago, A., and González, A. G. (2006). The correlation coefficient: an overview. Crit. Rev. Anal. Chem. 36, 41–59. doi: 10.1080/10408340500526766

Bendig, J., Andreas, B., Simon, B., Janis, B., Silas, E., and Georg, B. (2014). Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. Remote Sens. 6, 10395–10412. doi: 10.3390/rs61110395

Bendig, J., Kang, Y., Helge, A., Andreas, B., Simon, B., Janis, B., et al. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. Int. J. Appl. Earth Obs. Geoinform. 39, 79–87. doi: 10.1016/j.jag.2015.02.012

Chang, A., Jinha, J., Junho, Y., Murilo, M., and Juan, L. (2017a). “Sorghum panicle extraction from unmanned aerial system data,” in 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (Fort Worth, TX: IEEE), 4350–4353. doi: 10.1109/IGARSS.2017.8127965
Roujean, J.-L., and Breon, F. M. (1995). Estimating PAR absorbed by vegetation.

Pugh, N. A., Horne, D. W., Murray, S. C., Carvalho, G., Malambo, L., Jung, Potgieter, A. B., George-Jaeggli, B., Chapman, S. C., Laws, K., Suárez Cadavid, Malambo, L., Popescu, S. C., Murray, S. C., Putman, E., Pugh, N. A., Horne, D. W., Mutanga, O., and Skidmore, A. K. (2004). Narrow band vegetation indices for Sorghum breeders. J. Appl. Remote Sens. 11:0198035. doi: 10.1117/1.JRS.11.0198035

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

Snively, N., Seitz, M. S., and Szeli, R. (2008). Modeling the world from internet photo collections. Int. J. Comput. Vis. 80, 189–210. doi: 10.1007/s11263-007-0107-3

Stanton, C., Sterck, J. M., Elliott, N., Brewer, M., Maeda, M. M., and Chu, T. (2017). Unmanned aircraft system-derived crop height and normalized difference vegetation index metrics for Sorghum yield and aphid stress assessment. J. Appl. Remote Sens. 11:026035. doi: 10.1117/1.JRS.11.026035

Sun, W., Zhang, L., Du, B., Li, D., and Lai, M. L. (2015). Band selection using improved sparse subspace clustering for hyperspectral imagery classiﬁcation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8, 2784–2797. doi: 10.1109/JSTARS.2015.2417156

Varela, S., Assefa, Y., Vara Prasad, P. V., Peralta, N. R., Griffin, T. W., Sharda, A., 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

Virlet, N., Saberianesh, K., Sadeghi-Tehrani, P., and Hawkesford, M. J. (2017). Field scanner: an automated robotic ﬁeld phenotyping platform for detailed crop monitoring. Funct. Plant Biol. 44, 143–153. doi: 10.1071/FP16163

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

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., and Mortensen, D. A. (1995). Color indices for weed identiﬁcation under various soil, residue, and lighting conditions. Transac. ASAE 38, 259–269. doi: 10.13031/2013.27838

Yendrek, C. R., Tomaz, T., Montes, M. C., Cao, Y., Morse, M. A., Brown, P. J., et al. (2017). High-throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reﬂectance. Plant Physiol. 173, 614–626. doi: 10.1093/plysphz/173.2.614

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

Zhang, Z., Masjedi, A., Zhao, J., and Crawford, M. M. (2017). “Prediction of Sorghum biomass based on image based features derived from time series of UAV images,” in International Geoscience and Remote Sensing Symposium (IGARSS) (Fort Worth, TX), 6154–6157.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or ﬁnancial relationships that could be construed as a potential conﬂict of interest.

Copyright © 2018 Li, Shi, Veeranamalayam-Sivakumar and Schachtman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.