SORGHUM BIOMASS PREDICTION USING UAV-BASED REMOTE SENSING DATA AND CROP MODEL SIMULATION

Ali Masjedi[1*], *Jieqiong Zhao*[2], *Addie M. Thompson*[3], *Kai-Wei Yang*[4], *John E. Flatt*[5],
Melba M. Crawford[6], *David S. Ebert*[7], *Mitchell R. Tuinstra*[8], *Graeme Hammer*[9], *and Scott Chapman*[10]

[1,5,6]Dept. of Civil Engineering, [2,7]Dept. of Electrical and Computer Engineering, [3,4,8]Dept. of Agronomy
Purdue University, W. Lafayette, IN, USA
[9,10] The University of Queensland, Gatton 4343, QLD, Australia
E-mail: [1]amasjedi@purdue.edu; [2]zhao413@purdue.edu; [3]thomp464@purdue.edu;
[4]yang1301@purdue.edu; [5]flatt@purdue.edu; [6]mcrawford@purdue.edu; [7]ebert@purdue.edu;
[8]drmitch@purdue.edu; [9]g.hammer@uq.edu.au; [10]scott.chapman@uq.edu.au

ABSTRACT

Accurate phenotyping with unmanned aerial vehicles is a remote sensing application that has received recent attention as plant breeders seek to automate the expensive and time consuming traditional manual acquisition of measurements of plant traits. This paper focuses on the prediction of sorghum biomass utilizing high temporal and spatial resolution remote sensing data. Two methods are investigated for biomass prediction. The first uses nonlinear regression models to predict biomass directly from remote sensing data, based on features from Light Detection And Ranging (LiDAR) point clouds and hyperspectral data. The second strategy focuses on the biophysical sorghum crop model, APSIM, first, using remote sensing data to parametrize the crop model, and then simulating the biomass. Results from both approaches are provided and evaluated for an agricultural test field at the Agronomy Center for Research and Education (ACRE) at Purdue University.

Index Terms— Hyperspectral, LiDAR, phenotyping, APSIM, crop model, biomass prediction, machine learning

1. INTRODUCTION

Obtaining plant biophysical traits through automated phenotyping using remote sensing is advantageous because traditional field measurements of different phenotypes including plant height, leaf counts, and above ground biomass are expensive and time consuming to obtain. Multiple studies have demonstrated that these phenotypes can be estimated using remote sensing data for both large and small agricultural fields [1][2][3]. Biomass is an important phenotype that indicates the crop condition that is appropriate for crop monitoring and yield estimation [4].

This work was funded by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy under Grant DE-AR0000593.

Two primary approaches are used to estimate biomass using remote sensing data. The first method focuses on regression or neural network based models to find an empirical relationship between the biomass and remote sensing features. In [5], a partial least squares regression model was used to estimate above ground biomass of grasslands. Multiple features were extracted from hyperspectral data, including the original reflectance, first order derivative reflectance, and band-depth indices to estimate biomass. A random forest regression algorithm was used in [6] to model and predict wetland biomass in a high density, vegetated wetland. The prediction accuracy using the regression models depends on many factors. The importance of sample size, data type, and prediction method for remote sensing-based estimations of above ground forest biomass were studied in [7]. The authors demonstrated that the accuracy of the biomass estimates was highly dependent on the prediction method and data type, and less dependent on sample size for their data. The two main drawbacks of empirical methods are that they usually require a large number of samples to build an accurate regression model [3], and that they are local models, that is, they are not applicable to other areas and dates.

Alternatively, remote sensing data are incorporated into crop simulation models by calibration, forcing, or updating methods [8]. The **A**gricultural **P**roduction **S**ystems s**IM**ulator (APSIM) dynamic simulation model is capable of predicting the growth and productivity of plant species based on plant genetics, environmental conditions, and management practices [9]. The main drawback of simulation models like APSIM is that they require numerous input parameters. However, if the input parameters are accurate, they predict phenotypes with high accuracy [9].

In this paper, both strategies using remote sensing data for biomass prediction are implemented and evaluated. Figure 1 illustrates the general workflow of this paper for biomass prediction. For empirical models, two nonlinear statistical learn-

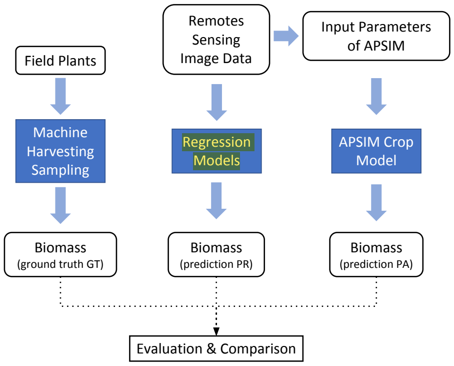
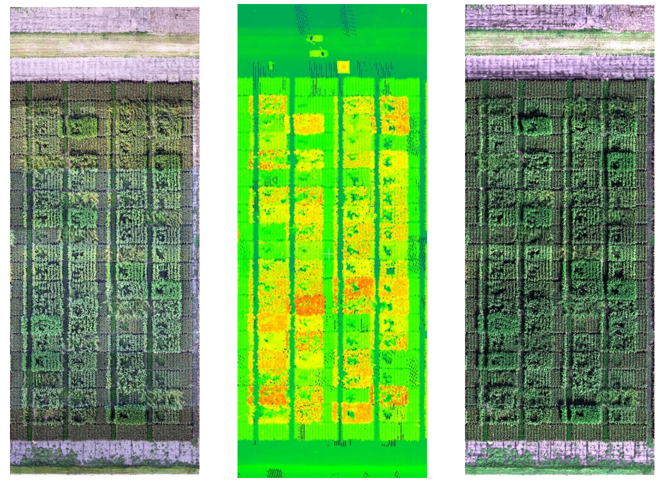


Fig. 1. The biomass prediction analysis flow.



(a) RGB image (b) LiDAR data (c) Hyperspectral data

Fig. 3. Remote sensing data acquired in August 30th.

ing algorithms are used. To build the APSIM crop model, leaf counts per plant, plant stand count, leaf area index (LAI), tiller number per plant, leaf size distribution, extinction coefficient of canopy (K), and radiation use efficiency (RUE) are needed [9]. Some parameters were estimated using field measurements in 2015 for eighteen hybrid sorghum genotypes. Leaf appearance rate, one of the ground based inputs to APSIM, was instead estimated using Unmanned Aerial Vehicle (UAV) RGB images acquired weekly during the 2017 growing season at ACRE. The details of estimating leaf appearance rate are described in Section 3.2.

2. DATA SET AND FEATURE EXTRACTION

2.1. Field Ground Truth Data

Eighteen genotypes of sorghum were planted in twelve row plots in the test field. Figure 2 shows the ground reference map for the study field. Colored boxes indicate 72 field plot boundaries, with eighteen in the N-S direction and four in the E-W direction. Ground truth biomass data were acquired using a biomass harvester for all 72 plots on July 31st and September 27th in 2017.

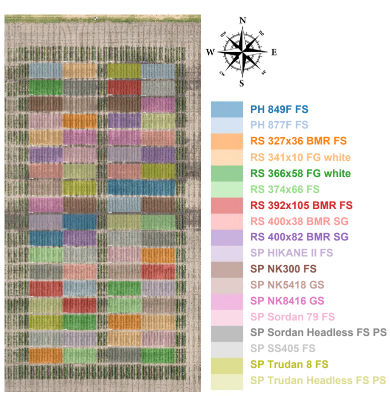


Fig. 2. Eighteen hybrid sorghum variates color coded in ground reference map with 72 plots. Each variety was planted

2.2. Remote Sensing Data

The RGB images were collected on a weekly basis between June 21st and August 30th in 2017 by a Sony Alpha camera on-board a DJI M600 multi-rotor platform at an altitude of 40 meters and velocity of 8m/s. Figure 3 (a) shows an example RGB image acquired on Aug 30th. The ground resolution for RGB images is 1 cm per pixel.

The LiDAR data were collected on the same day as the RGB images at an altitude of 20 meters and UAV velocity of 8m/s. Both the RGB and LiDAR systems were integrated with an Applanix APX-15 GNSS system to provide precise positioning. Figure 3 (b) shows one of the LiDAR data sets acquired on Aug 30th. From LiDAR data, the 50, 75, and 95 percentile heights for each plot were extracted and used in the regression models.

The hyperspectral data were acquired by a Headwall Nano VNIR push-broom scanner on-board a multirotor UAV platform approximately weekly between June 21st and September 10th in 2017. Each image has 136 bands, 4 cm spatial resolution, and 5 nm spectral resolution. An Applanix APX-15 GNSS system was fixed to the hyperspectral line camera for georeferencing the data. Figure 3 (c) shows the RGB bands of hyperspectral data acquired on August 30th. In this study,

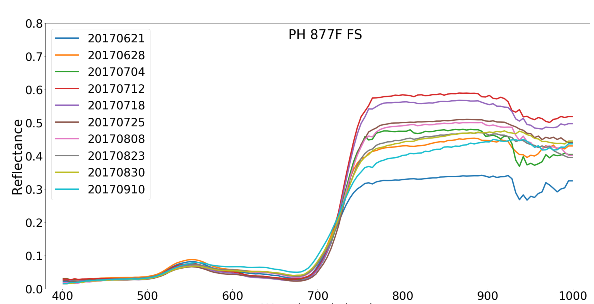


Fig. 4. Reflectance of a sorghum variety at multiple times during the season.

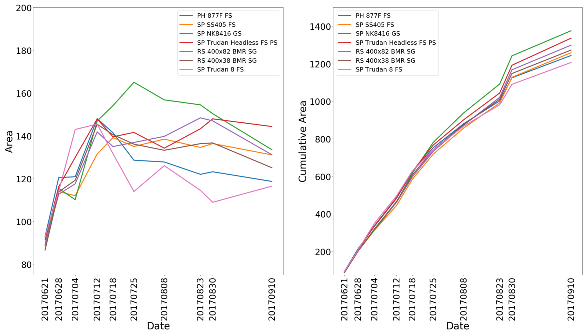


Fig. 5. Area under mean reflectance signature of some varieties (left) and the cumulative area over the season (right).

three types of features were extracted from the hyperspectral data and used in the regression models: original reflectance data, vegetation indices, and the area under the average of the spectral curves for each variety on a given date. Figure 4 shows the average for one variety for ten dates during the summer of 2017. Figure 5 shows the area under mean spectral curve for each date (left) and the cumulative area over the time (right).

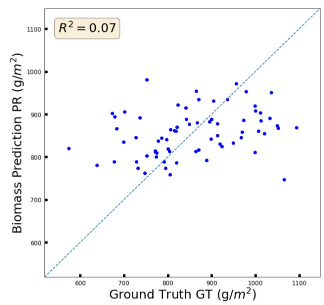
3. METHODOLOGY

3.1. Biomass Prediction using Regression Models

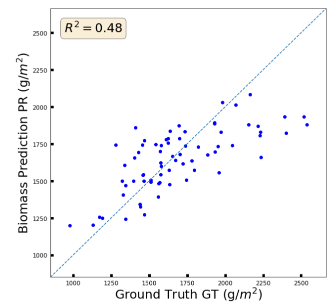
In this work, two nonlinear regression models are developed, Support Vector Regression (SVR) and Multi-Layer Perceptron (MLP). The parameter setting for SVR and MLP is described in our previous work [3]. Three-fold cross validation was used to test all the regression models.

3.2. Biomass Estimation using APSIM

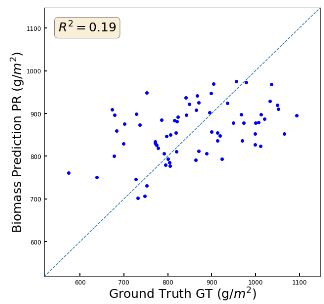
The APSIM crop model simulates the crop growth in a given environmental condition by incorporating weather data (actual or predicted), soil conditions, water information, and plant management strategies. All the parameters required by the APSIM model are available through the standard sorghum model except those which are cultivar specific, including leaf appearance rate, tiller number per plant, leaf size distribution, extinction coefficient of the canopy, and radiation use efficiency. This work focuses on the calculation of leaf appearance rate for 18 hybrids, based on ground reference data and remote sensing data, with other parameters being estimated from manual observations. Leaf appearance rate is the number of days required for appearance of successive leaf tips. The daily temperature has a significant impact on sorghum growth [10]. The accumulated growing degree days (GDDs) are associated with the measurements of the appearance of two consecutive leaves. The leaf counts derived by image processing algorithms described in [11] for two days, June 16th and June 21st, were used to calculate leaf appearance rate, assuming the relationship between accumulated GDD and number of visible leaves is linear.



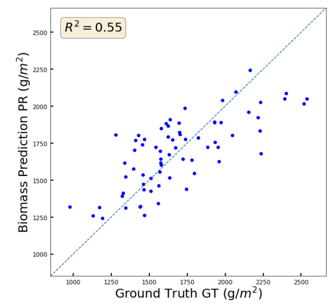
(a) SVR July 31st



(b) SVR Sep. 27th



(c) MLP July 31st



(d) MLP Sep. 27th

Fig. 6. Biomass prediction results using hyperspectral and LiDAR data.

4. RESULTS AND DISCUSSION

Figure 6 shows the results of predicting biomass for 72 plots based on hyperspectral and height features extracted from the remote sensing hyperspectral and LiDAR height features on dates prior to the biomass harvest. R^2 values of 0.07 and 0.19 were obtained for the SVR and MLP for the July 31st biomass data. The R^2 values for biomass prediction at the end of the season, September 27th, are 0.48 and 0.55 by SVR and MLP, respectively. Given that there are eighteen genotypes in the current field and some have very different characteristics, the

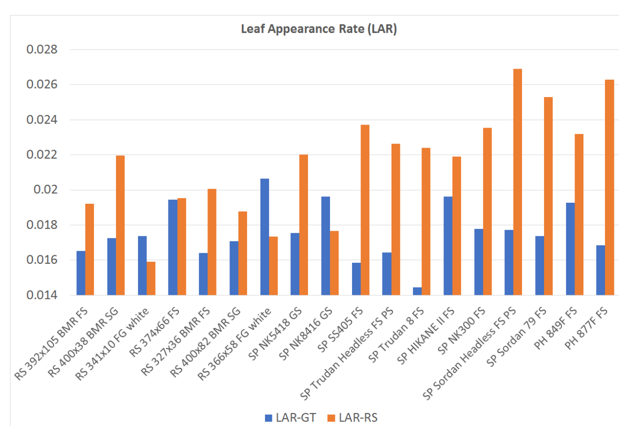


Fig. 7. Leaf appearance rate estimated by field measurements (LAR-GT) vs. RGB images (LAR-RS).

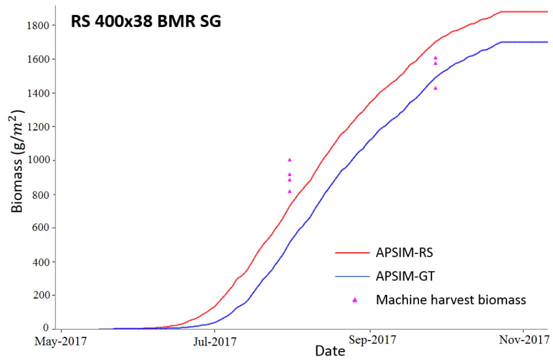


Fig. 8. Biomass simulation of RS 400x38 BMR SG by APSIM crop model.

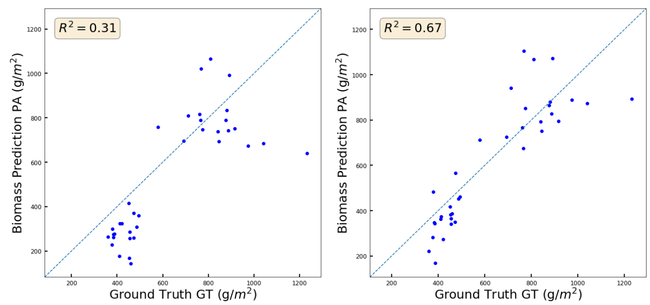


Fig. 9. Biomass simulation results from two models, APSIM-GT (left) and APSIM-RS (right).

R^2 value for the combined data is relatively good.

Figure 7 shows the leaf appearance rate estimated from field measurements (LAR-GT) and RGB images (LAR-RS). The LAR-RS is higher for most of the varieties. Figure 8 shows the biomass simulations for one of the varieties obtained by the APSIM crop model using the leaf appearance rate estimated from weekly field measurements (APSIM-GT), and RGB images (APSIM-RS). The field measurements of biomass for both July 31st and September 27th are also shown. In this case, the APSIM-RS under-estimates the July 31st biomass data, but over-estimates for September 27th, when the APSIM-GT yielded better results. The R^2 values for both simulations are shown in Figure 9. Examining eighteen hybrids for three replicates and two dates, the simulated biomass tends to be under-estimated at the early date, and is poorly estimated at the later date and particularly for highly productive varieties. The leaf appearance rate derived from the RGB images is generally greater than the value from the manual measurements, resulting faster leaf area growth and higher predicted biomass, as shown in Figure 8.

5. CONCLUSION AND FUTURE WORK

In this paper, two methods for biomass prediction were explored. The MLP regression model predicted the end of the season biomass with relatively high accuracy. Further studies will focus on multi-temporal feature extraction/selection. The APSIM crop simulation model was based on parameters derived from ground based and remote sensing data. The leaf appearance rate derived from images would have been more accurate if additional early season data had been acquired. Further, late season values were impacted by the complexity of the canopy, potentially resulting in counting leaf segments, and redundancy on multiple dates. Use of other remote sensing derived inputs to APSIM is also being investigated.

6. REFERENCES

[1] Y. Chen, J. Ribera, C. Boomsma, and E. Delp, Locating Crop Plant Centers from UAV-Based RGB Imagery, in 2017 IEEE International Conference on Computer Vision Workshops (ICCVW). IEEE, 10 2017, pp. 2030 2037.

[2] A. L. Nguy-Robertson, Y. Peng, A. A. Gitelson, T. J. Arkebauer, A. Pimstein, I. Herrmann, A. Karnieli, D. C. Rundquist, and D. J. Bonfil, Estimating green LAI in four crops: Potential of determining optimal spectral bands for a universal algorithm, Agricultural and Forest Meteorology, 2014.

[3] Z. Zhang, A. Masjedi, J. Zhao, and M. M. Crawford, Prediction of Sorghum Biomass Based on Image Based Features Derived From Time Series of Uav Images, pp. 61546157, 2017.

[4] W. Feng, B.-B. Guo, H.-Y. Zhang, L. He, Y.-S. Zhang, Y.-H. Wang, Y.-J. Zhu, and T.-C. Guo, Remote estimation of above ground nitrogen uptake during vegetative growth in winter wheat using hyperspectral red-edge ratio data, Field Crops Research, 2015.

[5] J. Chen, S. Gu, M. Shen, Y. Tang, and B. Matsushita, Estimating aboveground biomass of grassland having a high canopy cover: An exploratory analysis of in situ hyperspectral data, International Journal of Remote Sensing, 2009.

[6] O. Mutanga, E. Adam, and M. A. Cho, High density biomass estimation for wetland vegetation using worldview-2 imagery and random forest regression algorithm, International Journal of Applied Earth Observation and Geoinformation, 2012.

[7] F. E. Fassnacht, F. Hartig, H. Latifi, C. Berger, J. Hernandez, P. Corval an, and B. Koch, Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass, Remote Sensing of Environment, 2014.

[8] X. Jin, L. Kumar, Z. Li, H. Feng, X. Xu, G. Yang, and J. Wang, A review of data assimilation of remote sensing and crop models, 2018.

[9] G. L. Hammer, E. Van Oosterom, G. McLean, S. C. Chapman, I. Broad, P. Harland, and R. C. Muchow, Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops, 2010.

[10] T. Gerik, B. W. Bean, R. Vanderlip et al., Sorghum growth and development, Texas FARMER Collection, 2003.

[11] J. Ribera, F. He, Y. Chen, A. F. Habib, and E. J. Delp, Estimating phenotypic traits from uav based rgb imagery, in Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Workshop on Data Science for Food, Energy, and Water, 2016.