PREDICTION OF YIELD AND LIGNOCELLULOSIC COMPOSITION IN ENERGY CANE USING UNMANNED AERIAL SYSTEMS

A Thesis

by

URIEL CHOLULA RIVERA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Juan Enciso
Co-Chair of Committee,	J. Alex Thomasson
Committee Member,	Jorge A. da Silva
Head of Department,	Stephen W. Searcy

August 2020

Major Subject: Biological and Agricultural Engineering

ABSTRACT

Crop monitoring and appropriate agricultural management practices of elite germplasm will enhance bioenergy's efficiency. Unmanned aerial systems (UAS) may be a useful tool for this purpose. The objective of this study was to assess the use of UAS with true color and multispectral imagery to predict the yield and total cellulosic content (TCC) of newly created energy cane germplasm. A trial was established in the growing season of 2016 at the Texas A&M AgriLife Research and Extension Center in Weslaco, Texas, where 15 energy cane elite lines and three checks were grown on experimental plots, arranged in a randomized complete block design (RCBD) and replicated four times. Four flights were executed at different growth stages in 2018, at the first ratoon crop, using two multi-rotor UAS: the DJI Phantom 4 Pro equipped with RGB camera and the DJI Matrice 100, equipped with multispectral sensor (SlantRange 3p). Canopy cover, canopy height, NDVI (Normalized Difference Vegetation Index), and ExG (Excess Green Index) were extracted from the images and used to perform a stepwise regression to obtain the yield and TCC models. The results showed a good agreement between the predicted and the measured yields (R^2 = 0.88); however, a low coefficient of determination was found between the predicted and the observed TCC (R^2 = 0.30). This study demonstrated the potential application of UAS to estimate energy cane yield with high accuracy, enabling plant breeders to phenotype larger populations and make selections with higher confidence.

DEDICATION

To my parents and siblings

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Juan Enciso, and my committee members, Dr. J. Alex Thomasson, and Dr. Jorge A. da Silva, for their guidance and support throughout the course of this research. Special thanks to Dr. Thiago Marconi for helping me with data collection and providing the equipment for this study. I also thank to Mrs. Gerleene Acuna for her assistance in analyzing energy cane samples for composition analysis.

I am thankful to my parents and siblings for encouraging me to pursue my aspirations and for their unconditional support in my education and endeavors through my career. Thanks also go to my friends, colleagues, and the department faculty and staff for making my time at Texas A&M University a great experience.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professor Juan Enciso and Professor J. Alex Thomasson of the Department of Biological and Agricultural Engineering and Professor Jorge A. da Silva of the Department of Soil and Crop Sciences.

The energy cane germplasm for this study was provided by Professor Jorge A. da Silva, who established the experiment in 2016. Unmanned aerial systems and GPS system were provided by Dr. Thiago Marconi. Mrs. Gerleene Acuna conducted the cell wall composition analyses at the Texas A&M AgriLife Research and Extension Center facilities in Weslaco, Texas.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

Graduate study was supported by a Graduate Research Assistantship from Professor Juan Enciso and a WMHS Academic Scholarship awarded by the Water Management and Hydrological Science program for the academic year 2018-2019.

This work was also made possible in part by USDA-NIFA under Grant Number 554772 and the Hatch project “Development of Engineering Tools for Soil and Water Conservation” under Accession Number 1016863. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the USDA-NIFA.

NOMENCLATURE

AGL	Above Ground Level
ANNs	Artificial Neural Networks
CAI	Cellulose Absorption Index
CC	Canopy Cover
CCS	Commercial Cane Sugar
CH	Canopy Height
CHMs	Canopy Height Models
CIgreen	Green Chlorophyll Index
CIrededge	Red Edge Chlorophyll Index
CMOS	Complementary Metal Oxide Semiconductor
CNN	Convolutional Neural Networks
DSMs	Digital Surface Models
DTM	Digital Terrain Model
ExG	Excess Green Index
GCPs	Ground Control Points
GNDVI	Green Normalized Difference Vegetation Index
GRVI	Green-Red Vegetation Index
LAI	Leaf Area Index
LSTM	Long-Short Term Memory Networks
NDLI	Normalized Difference Lignin Index

NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
PPK	Post Processed Kinematic
RGB	Red Green Blue
RMSE	Root Mean Square Error
RNNs	Recurrent Neural Networks
RRMSE	Relative Root Mean Square Error
RVI	Ratio Vegetation Index
SRPIb	Simple Ratio Pigment Index
TCC	Total Cellulosic Content
UAS	Unmanned Aerial Systems
U.S.	United States

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION AND LITERATURE REVIEW*

When producing biofuels from dedicated bioenergy crops, maintaining high yields in low input conditions is a priority if global environmental change and increase in world population are considered (Takeda & Matsuoka, 2008). Benefits from bioenergy crops can be expanded by deploying varieties adapted for growth on marginal or degraded lands. Promoting high yielding bioenergy crops with positive attributes for water use and soil impact will also expand bioenergy benefits, not to mention the production of bioenergy in land that makes a small contribution to food production.

Perennial C_4 plant species such as energy sorghum (*Sorghum bicolor*), sugarcane (*Saccharum* spp.), and energy cane are promising feedstock species for the South Central and Southern U.S. regions, where favorable weather conditions would allow maximum biomass production rates. The combination of high productivity (~20 dry tons per acre), resulting from the C_4 photosynthesis, with high light, water, and nitrogen use efficiency, drought tolerance, and wide adaptation, make them well suited for marginal lands (Somerville et al., 2010). Because these species are perennial, they can also be ratooned (harvested and allowed to re-sprout from the roots or rhizomes, in the case of *Saccharum* spp.) for several years before replanting is necessary (Ellis & Merry, 2004). These species provide other environmental benefits compared to traditional row crops, including

* Part of this section is reprinted with permission from "Forecasting yield and lignocellulosic composition of energy cane using unmanned aerial systems" by Cholula, U., da Silva, J.A., Marconi, T., Thomasson, J.A., Solorzano J., and Enciso, J. (2020). *Agronomy*, *10*, 718. Copyright 2020 Cholula U., da Silva, J.A., Marconi, T., Thomasson, J.A., Solorzano J., and Enciso, J.

extensive root systems, which can increase nutrient capture, improve soil quality, sequester carbon, reduce erosion, and increase growth rate (Somerville et al., 2010).

Sugarcane is an important crop in the United States, and it is grown commercially in Florida, Hawaii, Louisiana, and Texas. In 2017 the total sugarcane production was 30.16 million tons from 365,844 ha of sugarcane crop grown for sugar and seed (USDA-NASS, 2019). From the total sugarcane production, 28.30 million tons were destined for sugar, and 1.87 million tons were designated for seed. For the same year, the sugarcane production for sugar and seed was valued at $1.03 billion (USDA-NASS, 2019). The value of the production for sugar was $965.76 million, whereas the value of the production for seed was $59.77 million.

Sugarcane is extremely high yielding, producing large quantities of biomass. Nonetheless, minimal efforts have been employed to develop varieties under low-input management specifically. The plant has tremendous yield potential but is restricted to subtropical and tropical regions. However, because modern sugarcane varieties have been derived from a hybridization process involving *S. officinarum* and the wild cane (*S. spontaneum)* (J. A. G. da Silva et al., 1993), which has drought and cold resistance (Park et al., 2015), the creation of energy cane adapted to low input conditions is possible.

Sugarcane genetic breeding is a long process that takes between 10 and 12 years involving hybridization crosses and field selection of new genotypes (Scortecci et al., 2012). Since the crop is highly polyploid and open-pollinated, the hybridization step generates hundreds of thousands of different genotypes to be selected. Because the species is clonally propagated, the biggest challenge for breeders is to identify, among those

genotypes, the very few individuals that will consistently produce higher yields over the years and across different environments. This process that may generate an enormous amount of phenotypic data to be analyzed.

In the past, sugarcane breeding programs were looking for higher sugar productivity, but now some breeding programs are also focused on high yield of fiber (Carvalho-Netto et al., 2014). Different from conventional sugarcane (*Saccharum* spp.), energy cane is selected more for fiber than sucrose composition (Matsuoka et al., 2014). Therefore, as energy cane breeding programs work to develop a high-yield, low-input production system, two important goals for novel germplasm are maximizing productivity and optimizing composition.

Energy cane is a high-fiber crop obtained from the hybridization of sugarcane (*Saccharum* spp.) and wild cane (*Saccharum spontaneum*) (Matsuoka et al., 2014; Salassi et al., 2013). Energy cane fiber is composed of 43% cellulose, 24% hemicellulose and 22% lignin, whereas sugarcane comprises 42% cellulose, 25% hemicellulose and 20% lignin (Kim & Day, 2011). Energy cane has excellent potential as a biofuel feedstock due to its high fiber content. However, one of the current challenges for plant breeders is to expand the production to higher latitudes than sugarcane is presently grown in the U.S. This expansion can potentially be accomplished through conventional breeding techniques to develop varieties more tolerant to stress including cold (Matsuoka et al., 2014; Salassi et al., 2015).

Plants have developed different responses to abiotic stresses, resulting in different traits for avoidance or tolerance (Levitt, 1980). The incorporation of such traits in plant

breeding may be facilitated through phenotyping protocols (Salekdeh et al., 2009). High-throughput phenotyping is particularly important in studies of tolerance to abiotic stresses, such as drought. The highly complex responses of plants to drought require the dissection of such responses into a series of component traits that can be measured most efficiently and accurately with non-destructive image technologies (Berger et al., 2010).

The traditional methods for quantifying crop traits, such as plant height, leaf area index (LAI), biomass and yield, depend on manual sampling, which is time consuming, laborious and inefficient (Berni et al., 2009; Li et al., 2016; Rahaman et al., 2015). Crop yield is commonly estimated with manual surveys or by establishing the relationship between agronomic factors or climatic factors and crop yield based on statistical analysis methods (Swain et al., 2010). However, several observations and sampling of experimental plots are required to determine the parameters of a yield prediction model (Yang et al., 2017).

New approaches have been used for crop yield prediction. You et al. (2017) forecasted soybean yield in the United States at county-level with deep learning models such as convolutional neural networks (CNN) and long-short term memory (LSTM) networks. Their models performed better than traditional remote sensing methods with a 30% reduction of root mean square error (RMSE). Similarly, Khaki et al. (2020) proposed a model to predict corn and soybean yield across the Corn Belt in the United States using convolutional neural networks (CNN) and recurrent neural networks (RNNs) based on environmental data and management practices. The model outperformed other methods tested, achieving a RMSE between 9% and 8% of their corresponding average yields.

Some disadvantages of these methods are the lack of specific regression relationships, and the labor-intensive calculation process, which substantially restrict their efficiency and implementation.

Other studies explored remote sensing approaches for predicting crop yield. Rahman and Robson (2016) generated a model based on time series Landsat data to predict sugarcane yield in the Bundaberg region, Queensland, Australia. A significant correlation (R^2 = 0.69 and RMSE = 4.2 t ha^{-1}) was found between the maximum GNDVI (Green Normalized Difference Vegetation Index) provided by the model and the annual harvested yield (t ha^{-1}). Rahman and Robson (2020) improved the accuracies of the previously developed time series model at the individual block level, integrating sugarcane planting or previous harvest dates. High accuracies (R^2 = 0.87 and RMSE = 11.33 (t ha^{-1})) were achieved at the block level when compared to actual harvested yield. Fernandes et al. (2017) predicted sugarcane yield in São Paulo State, Brazil at municipal and regional scales using metrics from normalized difference vegetation index (NDVI) time series from the Moderate Resolution Image Spectroradiometer (MODIS) and ensemble model of artificial neural networks (ANNs). The relative root mean square error (RRMSE) and the coefficient of determination (R^2) for the predicted yield were 6.8% and 0.61, respectively.

Unmanned aerial systems (UAS) equipped with various sensors have been used for rapid and non-destructive high-throughput phenotyping. UAS have the advantage of adaptable and convenient operation, fast access to the data, and high spatial resolution (Yang et al., 2017). Digital and multispectral cameras are frequently used sensors. Some of the applications of these sensors for field-based phenotyping include biomass

estimation, canopy surface modeling, and crop height estimation (Diaz-Varela et al., 2014; Mathews & Jensen, 2013; Zarco-Tejada et al., 2014). Therefore, UAS have been used to predict yield with plant physiological parameters and vegetation indices (Yang et al., 2017). Sanches et al. (2018) predicted sugarcane yield in Brazil using the LAI (Leaf Area Index) and GRVI (Green-Red Vegetation Index). Their results showed that GRVI estimated yield (R^2 = 0.69) with higher accuracy than LAI (R^2 = 0.34), but when both indices were combined, the yield was estimated with greater precision (R^2 = 0.79). However, the authors suggested for future studies the incorporation of plant height and additional indices to improve the results. Chea et al. (2020) developed prediction models for Brix, Pol, fiber, and CCS (Commercial Cane Sugar) value using six vegetation indices (GNDVI, NDVI, RVI, CIgreen, CIrededge, and SRPIb). Their findings indicate that CIrededge is correlated with Pol (R^2 = 0.77) and CCS (R^2 =0.68), independent of variety, whereas Brix models depend on the variety and need different vegetation indices. A weak correlation (R^2 = 0.35 – 0.50) was found between fiber content with the six vegetation indices.

Even though the use of UAS to predict yield and composition of sugarcane has been evaluated, information is missing about the utilization of this technology for energy cane. Additionally, most studies have been focused on the commercial or industrial benefits that UAS can provide. However, this research proposes to use UAS as a tool in plant breeding that facilitates the acquisition of phenotypic information. For this reason, the objective of this study is to assess the use of UAS with true color and multispectral

imagery to predict the yield and total cellulosic content (TCC) of newly created energy cane germplasm.

2. MATERIALS AND METHODS*

2.1. Study Site

The study was conducted on an experimental farm at the Texas A&M AgriLife Research and Extension Center in Weslaco, Texas (26°9'41.96"N, 97°56'30.72"W, 21 m AMSL). This region has a humid subtropical climate (*Cfa*) with an average annual rainfall of 632 mm. A trial was established in the growing season of 2016 in an area of 0.93 ha, where 15 energy cane elite lines and three checks were planted on 9.1 × 3.0 m plots with rows oriented in a north-south direction on a Raymondville clay loam soil. The experimental plots were arranged in a randomized complete block design, replicated four times (Figure 1).

* Part of this section is reprinted with permission from "Forecasting yield and lignocellulosic composition of energy cane using unmanned aerial systems" by Cholula, U., da Silva, J.A., Marconi, T., Thomasson, J.A., Solorzano J., and Enciso, J. (2020). *Agronomy, 10,* 718. Copyright 2020 Cholula U., da Silva, J.A., Marconi, T., Thomasson, J.A., Solorzano J., and Enciso, J.

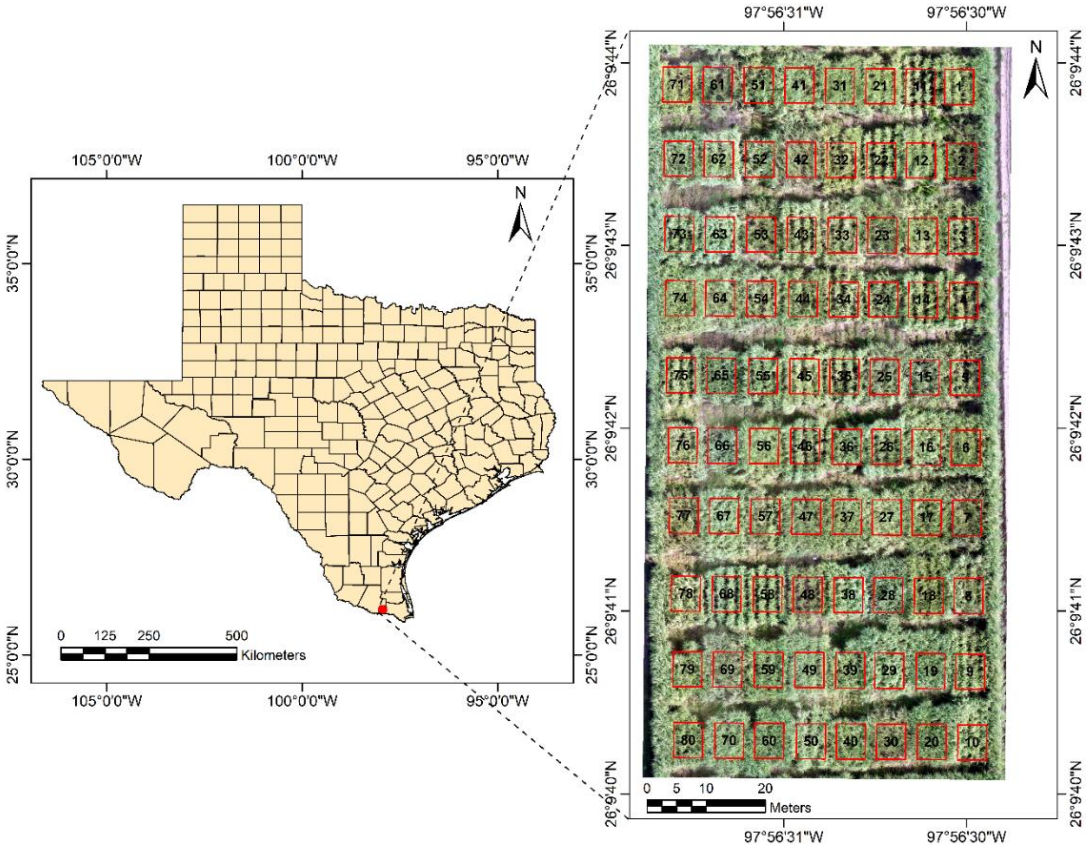

Figure 1. Study area location with the experimental plots (reprinted from Cholula et al., 2020; permissions for reproduction have been obtained from the copyright holders).

2.2. Imagery Acquisition and Processing

Two multi-rotor UAS were used to acquire the data, a DJI Phantom 4 Pro and a DJI Matrice 100 (SZ DJI Technology Co., Ltd., Shenzen, China) (Table 1). The Phantom was equipped with an RGB sensor with a resolution of 20 megapixels (spatial resolution of 0.55 cm/pixel at 20 m), and 1" CMOS (Complementary Metal Oxide Semiconductor) detector. The Matrice was equipped with a SlantRange 3p multispectral sensor (SlantRange Inc., San Diego, CA, USA) and an ambient illumination sensor that can be

used to help calibrate images when sunlight conditions are changing at the time of data collection. The SlantRange 3p sensor has a spatial resolution of 4.8 cm/pixel at 120 m above ground level (AGL).

Table 1. Specifications for the UAS and the sensors used to collect the data (reprinted from Cholula et al., 2020; permissions for reproduction have been obtained from the copyright holders).

	DJI Phantom 4 Pro with RGB sensor	Matrice 100 with SlantRange 3p sensor
Sensor resolution (pixels)	5472 × 3648	1280 × 1024
Spectral resolution	R, G, B	NIR, red edge, R, G
Weight (g)	1388	2781

Images were acquired on July 17, September 18, November 14, and December 19 of 2018, corresponding to 273, 210, 153, and 118 days before harvesting, respectively (Figure 2). The RGB images were collected at 20 m AGL and 80% overlap and sidelap, whereas the multispectral images were collected at 30 m AGL and 70% overlap and sidelap. Flights were conducted between 10:00 AM and 12:00 PM with wind speed less than 8 km/h to avoid image distortion. For georeferencing purposes, eight ground control points (GCPs) were placed uniformly in the study area. The GCPs were surveyed twice with a differential dual frequency, post-processed kinematic (PPK) GPS system, collecting data at 20 Hz (V-map Air model, Micro Aerial Projects L.L.C., Gainesville, FL, USA) (Figure 3).

SlantView (SlantRange Inc., San Diego, CA, USA) software was used to export radiometrically calibrated multispectral images for further analysis. Both RGB and multispectral images were processed in Agisoft Metashape Professional software (Agisoft

LLC, St. Petersburg, Russia) to generate the orthomosaics and digital surface models (DSMs).

(a) (b)

Figure 2. RGB and multispectral sensors used for data collection. The DJI Phantom 4 Pro (a) for RGB and the DJI Matrice 100 platform (b) with the SlantRange 3p sensor for multispectral data collection.

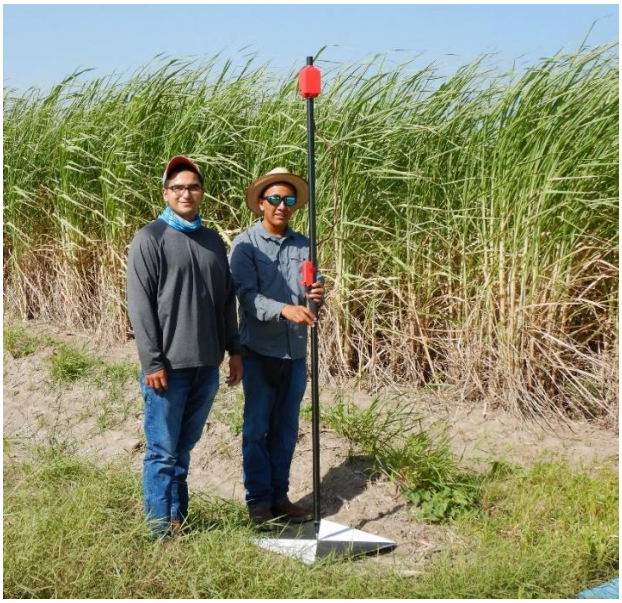

Figure 3. Ground control points (GCPs) survey with post-processed kinematic (PPK) V-map Air Model GPS system.

2.3. Feature Extraction

2.3.1. Canopy Height

Point cloud datasets were imported into Quick Terrain Modeler (Applied Imagery, Chevy Chase, MD, USA) software to generate the canopy height models (CHMs). The ground surface was estimated with the AGL Analyst tool, and a 10 m grid sampling distance was selected. The digital terrain model (DTM) created was subtracted from the digital surface model (DSM) to obtain the CHMs (Figure 4). Additionally, the generated CHMs were processed in ENVI (Harris Geospatial Solutions Inc., Broomfield, CO, USA) to set the negative values to zero. CHMs were imported into ArcGIS 10.6.1 (ESRI, Redlands, CA, USA) in which average canopy height (CH) per plot was extracted.

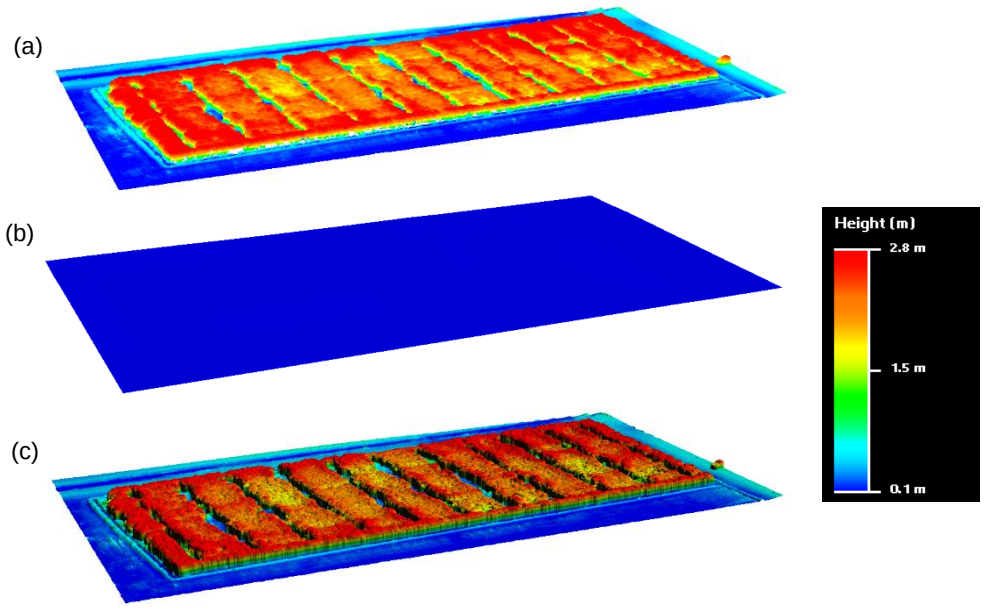

Figure 4. Canopy height model (CHM) generation in Quick Terrain Modeler software. The DTM (b) was subtracted from the DSM (a) to create the CHM (c).

2.3.2. Canopy Cover

RGB orthomosaic images were converted into binary images with the canopeo algorithm (Patrignani & Ochsner, 2015) in QGIS (Open Source Geospatial Foundation Project), where zero designates non-canopy pixels, and one denotes canopy pixels (Figure 5). The zonal statistics tool was used to compute the total number of pixels and the sum of canopy pixels within each plot. Then percentage canopy cover (CC) was calculated as the ratio of canopy pixels to the total number of pixels using Equation (1).

$$CC = \left(\frac{\sum canopy\ pixels}{\sum pixels}\right) \times 100 \quad (1)$$

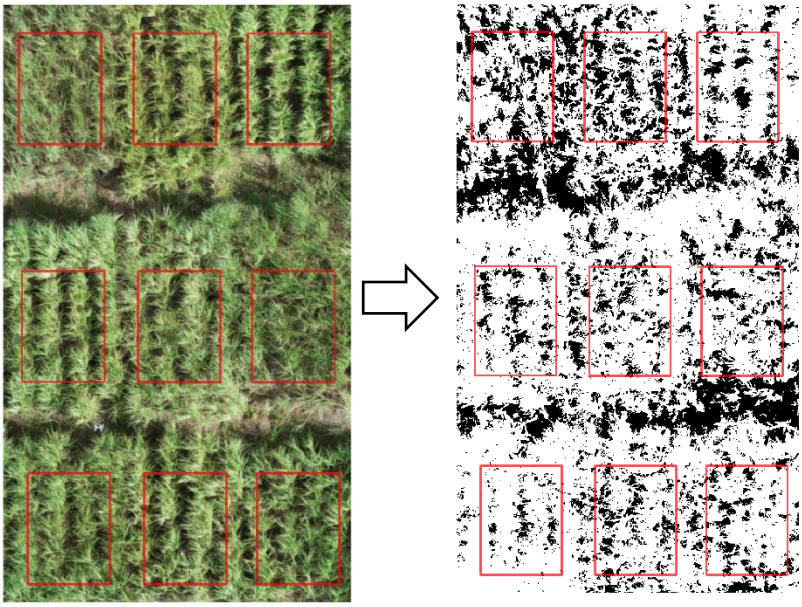

Figure 5. Canopy cover estimation from the orthomosaic images. An RGB orthomosaic image is classified into a binary image with the canopeo algorithm.

2.3.3. Normalized Difference Vegetation Index (NDVI)

The Normalized Difference Vegetation Index (NDVI) was calculated in ArcGIS 10.6.1 based on the multispectral orthomosaic images (Figure 6). In the raster calculator, the near infrared (NIR) and red bands were selected to generate the index (Equation (2) (Rouse et al., 1974)). Then, average NDVI values per plot were extracted with the zonal statistics as a table tool. This index is reported to be well correlated with biomass, and it has been used to describe crop phenology in tomatoes (Thenkabail et al., 2000; Enciso et al., 2019).

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \quad (2)$$

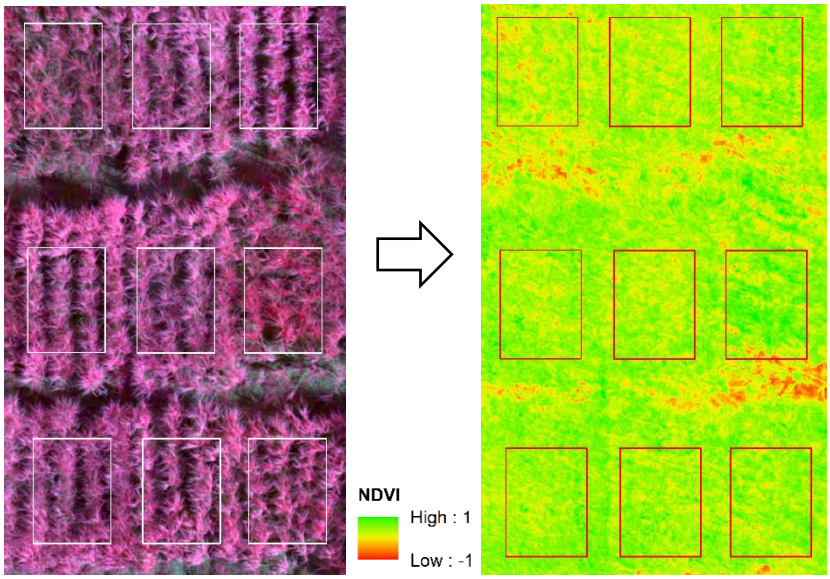

Figure 6. False color composite of a multispectral orthomosaic (left) and NDVI generated (right).

2.3.4. Excess Green Index (ExG)

For the Excess Green Index (ExG) calculation, the RGB orthomosaic images were imported into ArcGIS 10.6.1. The raster calculator tool was used to create ExG by selecting the green, red, and blue bands (Equation (3) (Woebbecke et al., 1995)) (Figure 7). Average ExG values per plot were extracted with the zonal statistics as a table tool.

$$ExG = 2g - r - b \tag{3}$$

where g, r, and b are the normalized spectral components, according to:

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B} \tag{4}$$

R, G, and B denote the values of the red, green, and blue bands, respectively.

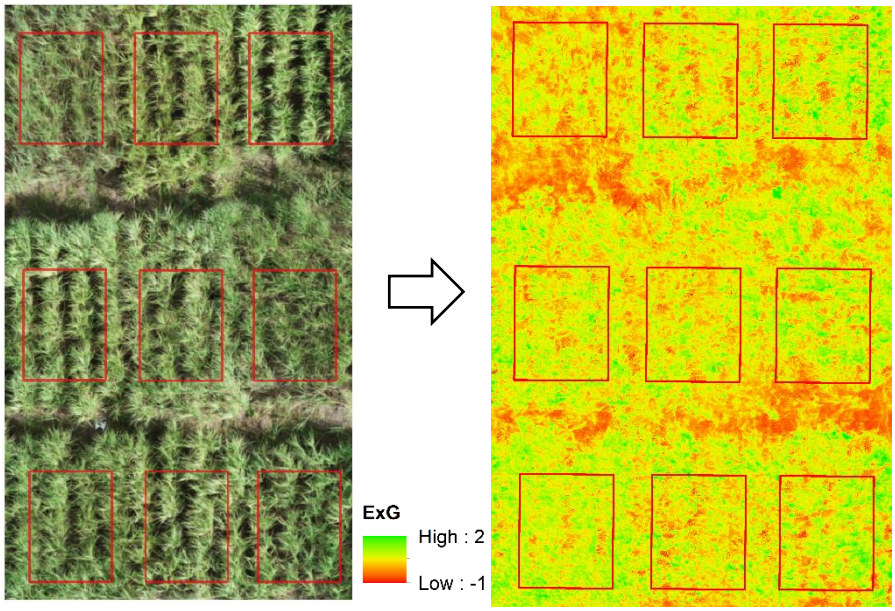

Figure 7. RGB orthomosaic image (left) and ExG generated (right).

2.4. Cell Wall Composition Analysis

Before harvesting, one sample of 10 stalks per variety was collected per replicate for composition analyses (Figure 8). In the laboratory, the chemical composition of the energy cane bagasse was determined (cellulose, hemicellulose, and lignin) with near infrared spectroscopy (Berding et al., 1991).

Figure 8. Energy cane sampling for composition analyses.

2.5. Harvest Data

At the end of the season the trial was harvested with a Cameco sugarcane harvester on April 16, 2019. The plot weights were measured with a small capacity weigh wagon (3-Ton Weigh Wagon, CAMECO) instrumented with three load cells (Bischoff et al., 2001). Then the yield was calculated in megagrams per hectare (Mg ha^{-1}).

2.6. Data Analysis

The data were analyzed in JMP 14 software (SAS Institute Inc., Cary, NC, USA) to identify outliers with the quantile range outliers method. A tail quantile value of 0.25 was defined, which means that the interquantile range is between 0.25 and 0.75 quantiles of the data. Then a multiplier (Q) of 1.5 was selected to identify values as outliers. An outlier was calculated as any value more than Q times the interquantile range from the lower and upper quantiles.

A multivariate analysis was performed in JMP 14 to find the pairwise and higher relationships among the CC, CH, NDVI, ExG, yield, and total cellulosic content (TCC). This analysis allowed summarization of the linear relationships between flights to check for data consistency. A correlation probability report was generated, which showed the *p*-values that correspond to a test of the null hypothesis that the true correlation between the variables is zero.

2.7. Yield and Total Cellulosic Content (TCC) Models

A stepwise multiple regression approach with k-fold cross-validation was implemented in JMP 14 software to obtain the yield and total cellulosic content (TCC) models with the data from the varieties Ho02-113, TH16-19, and TH16-25. Ho02-113 is a high fiber sugarcane (energy cane) variety released for use as a biofuel feedstock by the USDA-ARS Sugarcane Research Unit in Houma, Louisiana; TH16-19 and TH16-25 are two energy cane varieties from Dr. Jorge da Silva's Sugarcane Breeding Program. Six folds were selected to split the data into groups of an equal number of observations. The stopping rule adopted was Max K-Fold RSquare, which attempted to find a model to maximize the coefficient of determination for the validation set. Stepwise multiple linear regression is a commonly implemented empirical statistical method for high-throughput field phenotyping (Richards, 1990). It is used to improve the prediction performance of the models by eliminating unnecessary factors and selecting significant factors (Yu et al., 2016).

3. RESULTS*

3.1. Statistics

After an initial analysis, no outliers were found in the variables CC, CH, NDVI, and ExG extracted from the UAS images, and in the observed yield and TCC (Table 2). The highest mean CC, CH, and ExG values were observed for the second flight (85.74, 3.73, and 0.23, respectively), while the lowest values of CC and ExG were for the fourth flight; the lowest mean CH was observed during the first flight. Similarly, the highest mean NDVI values were obtained for the third flight (0.61), whereas the lowest NDVI values were for the fourth flight.

* Part of this section is reprinted with permission from "Forecasting yield and lignocellulosic composition of energy cane using unmanned aerial systems" by Cholula, U., da Silva, J.A., Marconi, T., Thomasson, J.A., Solorzano J., and Enciso, J. (2020). *Agronomy, 10*, 718. Copyright 2020 Cholula, U., da Silva, J.A., Marconi, T., Thomasson, J.A., Solorzano J., and Enciso, J.

Table 2. Descriptive statistics of canopy cover (CC), canopy height (CH), NDVI, ExG, yield, and total cellulosic content (TCC) (reprinted from Cholula et al., 2020; permissions for reproduction have been obtained from the copyright holders).

	N	Mean	Median	SD	Min	Max
CC – 1st	12	61.35	61.85	16.03	37.75	84.02
CC – 2nd	12	85.74	86.30	7.80	70.37	97.02
CC – 3rd	12	69.78	71.25	5.96	61.69	78.34
CC – 4th	12	33.18	31.84	16.93	12.59	60.79
CH– 1st	12	2.22	2.27	0.33	1.74	2.86
CH– 2nd	12	3.73	3.73	0.35	3.23	4.39
CH – 3rd	12	3.60	3.50	0.89	2.03	4.96
CH – 4th	12	2.76	2.62	0.34	2.32	3.35
NDVI – 1st	12	0.54	0.54	0.06	0.41	0.62
NDVI – 2nd	12	0.58	0.58	0.03	0.51	0.61
NDVI – 3rd	12	0.61	0.61	0.02	0.58	0.65
NDVI – 4th	12	0.47	0.45	0.05	0.39	0.58
ExG – 1st	12	0.18	0.17	0.04	0.13	0.26
ExG – 2nd	12	0.23	0.23	0.02	0.20	0.27
ExG – 3rd	12	0.18	0.19	0.02	0.14	0.23
ExG – 4th	12	0.12	0.13	0.04	0.05	0.17
Yield (Mg ha^{-1})	12	61.57	67.30	15.74	31.36	80.49
TCC (%)	12	62.18	61.87	3.67	55.99	66.91

1st–first flight (July 17, 2018), 2nd–second flight (September 18, 2018), 3rd–third flight (November 14, 2018), 4th–fourth flight (December 19, 2018), N–sample size, SD–standard deviation, Min–minimum value, Max–maximum value.

The canopy cover variability was greater during the third flight, evidenced by the negatively skewed distribution (Figure 9a). Canopy height variability was higher during the first, third, and fourth flights, which presented a negatively skewed distribution for the first flight and a positively skewed distribution for the third and fourth flights (Figure 9b). NDVI showed higher variation in the fourth flight with a positively skewed distribution (Figure 9c). ExG variability was observed in the first flight with a positively skewed distribution and the third flight with a negatively skewed distribution (Figure 9d).

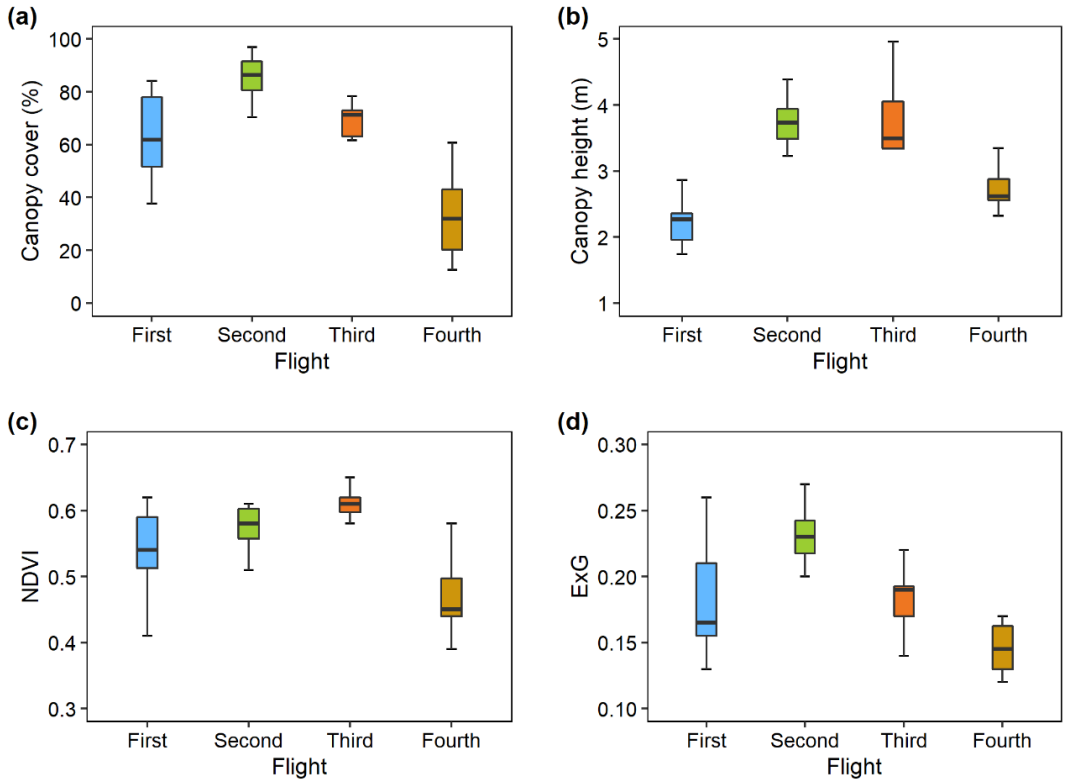

Figure 9. Box-plot of canopy cover (a), canopy height (b), NDVI (c), and ExG (d) (reprinted from Cholula et al., 2020; permissions for reproduction have been obtained from the copyright holders).

3.2. Relationship Analysis

The correlations among the independent variables were analyzed between flights (Table 3). CC showed high temporal consistency between the first and the fourth flight (r = 0.73); however, a weak correlation was found between the other flights. Positive relationships associated with canopy height were consistent throughout the evaluation period, being the lowest correlation between the first and fourth flight (r = 0.73). In contrast, the highest correlation was between the first and second flight (r = 0.93). A strong positive relationship was found between NDVI values of the first and second flight (r =

0.76), and between the first and fourth flights ($r = 0.60$); nevertheless, for the rest of flights, weak positive or negative correlations were observed. The highest ExG correlation was between the first and fourth flight ($r = 0.63$), followed by the correlation between the first and second flight ($r = 0.39$). On the other hand, weak positive relationships were found for the remaining flights.

From each flight were identified the variables with a moderate or strong correlation with yield and TCC. For the first flight, CC, CH, and NDVI were positively related to yield, whereas CC and ExG were associated with TCC. For the second flight, CC, CH, and NDVI were positively correlated to yield; CC, CH, and ExG were related to TCC. For the third flight, the variables associated with yield were CH and NDVI: the first variable was positively correlated, and the second negatively correlated. In contrast, CH, NDVI, and ExG were related to TCC: CH was positively correlated, and NDVI and ExG negatively correlated. For the fourth flight, all the variables were moderately related to yield; CC, CH, and ExG were associated with TCC.

Table 3. Correlation coefficient matrix between CC, CH, NDVI, ExG, yield, and TCC (reprinted from Cholula et al., 2020; permissions for reproduction have been obtained from the copyright holders).

	CC[1]	CC[2]	CC[3]	CC[4]	CH[1]	CH[2]	CH[3]	CH[4]	NDVI[1]	NDVI[2]	NDVI[3]	NDVI[4]	ExG[1]	ExG[2]	ExG[3]	ExG[4]	Yield	TCC
CC[1]	1.00																	
CC[2]	0.46	1.00																
CC[3]	0.01	0.14	1.00															
CC[4]	0.73**	0.47	-0.19	1.00														
CH[1]	0.56	0.54	-0.32	0.79	1.00													
CH[2]	0.56	0.63*	-0.32	0.71	0.93***	1.00												
CH[3]	0.55	0.83***	0.04	0.55	0.78**	0.83***	1.00											
CH[4]	0.56	0.71**	0.00	0.53	0.73**	0.83***	0.76**	1.00										
NDVI[1]	0.29	0.81**	-0.01	0.33	0.72**	0.79**	0.85***	0.85***	1.00									
NDVI[2]	0.08	0.82**	0.14	0.00	0.29	0.38	0.63*	0.48	0.76**	1.00								
NDVI[3]	-0.44	-0.58*	-0.06	-0.16	-0.13	-0.21	-0.47	-0.19	-0.31	-0.56	1.00							
NDVI[4]	0.35	0.48	0.18	0.53	0.71**	0.66*	0.66*	0.67*	0.60*	0.36	0.09	1.00						
ExG[1]	0.85***	0.43	-0.15	0.81**	0.57	0.57	0.38	0.55	0.29	0.00	-0.23	0.20	1.00					
ExG[2]	0.42	0.71**	0.50	0.26	-0.01	0.05	0.37	0.26	0.28	0.53	-0.70*	0.02	0.39	1.00				
ExG[3]	-0.09	-0.08	0.45	0.12	0.11	0.00	-0.04	-0.06	0.08	-0.12	0.25	0.14	0.10	0.05	1.00			
ExG[4]	0.65*	0.49	-0.09	0.89***	0.82**	0.73**	0.64*	0.46	0.43	0.14	-0.32	0.59*	0.63*	0.26	0.31	1.00		
Yield	0.32	0.70*	-0.09	0.44	0.75**	0.79**	0.90***	0.63*	0.84***	0.55	-0.44	0.44	0.26	0.24	0.13	0.60*	1.00	
TCC	0.49	0.30	-0.10	0.53	0.28	0.35	0.39	0.32	0.07	-0.08	-0.36	0.07	0.37	0.31	-0.40	0.33	0.32	1.00

[1] First flight, [2] second flight, [3] third flight, and [4] fourth flight. * $p < 0.05$, ** $p < 0.01$, ***$p < 0.001$

3.3. Energy Cane Yield and TCC Models

The yield models were obtained for the four flights after performing a stepwise regression. The coefficient of determination (R^2), *p*-value, and RMSE were used to evaluate the model's performance (Table 4). The variables that most influenced energy cane yield were NDVI and canopy height.

The models for the first and third flights showed a good coefficient of determination to estimate energy cane yield. The lowest R^2 was for the fourth flight, while the highest R^2 corresponded to the third flight (Table 4).

Table 4. Energy cane yield models with the coefficient of determination, *p*-value, and RMSE for the flight campaigns (reprinted from Cholula et al., 2020; permissions for reproduction have been obtained from the copyright holders).

Flight	Model	R^2	*p*-value	RMSE
07/17/18	yield = 222.08 NDVI – 58.39	0.71	0.0006	8.89
09/18/18	yield = 30.66 CH + 148.40 NDVI – 138.34	0.69	0.0049	9.64
11/14/18	yield = – 0.68 CC + 16.26 CH + 177.04 ExG + 18.25	0.88	0.0004	6.26
12/19/18	yield = – 0.70 CC + 26.14 CH + 381.68 ExG – 32.95	0.62	0.0432	11.39

In the same way, total cellulosic content (TCC) models for the fourth flights presented different coefficients of determination, *p*-values, and RMSE (Table 5). Nonetheless, since the variables were not significantly correlated to TCC, then low R^2 were obtained for all the flights.

Table 5. Energy cane TCC models with the coefficient of determination, *p*-value, and RMSE for the flight campaigns (reprinted from Cholula et al., 2020; permissions for reproduction have been obtained from the copyright holders).

Flight	Model	R^2	*p*-value	RMSE
07/17/18	TCC = 0.11 CC + 55.26	0.24	0.1032	3.35
09/18/18	TCC = 3.58 CH + 47.19 ExG + 38.00	0.21	0.3519	3.61
11/14/18	TCC = 1.54 CH – 57.14 ExG + 67.15	0.30	0.1996	3.39
12/19/18	TCC = 0.11 CC + 58.39	0.28	0.0779	3.27

The relationship between observed and predicted yield is presented in Figure 10. This figure shows that the best relationship between observed and predicted yield is given by the model for the third flight with a coefficient of determination of 0.88 (Figure 10c). In contrast, the lowest relationship of yields is found in the model for the fourth flight (Figure 10d).

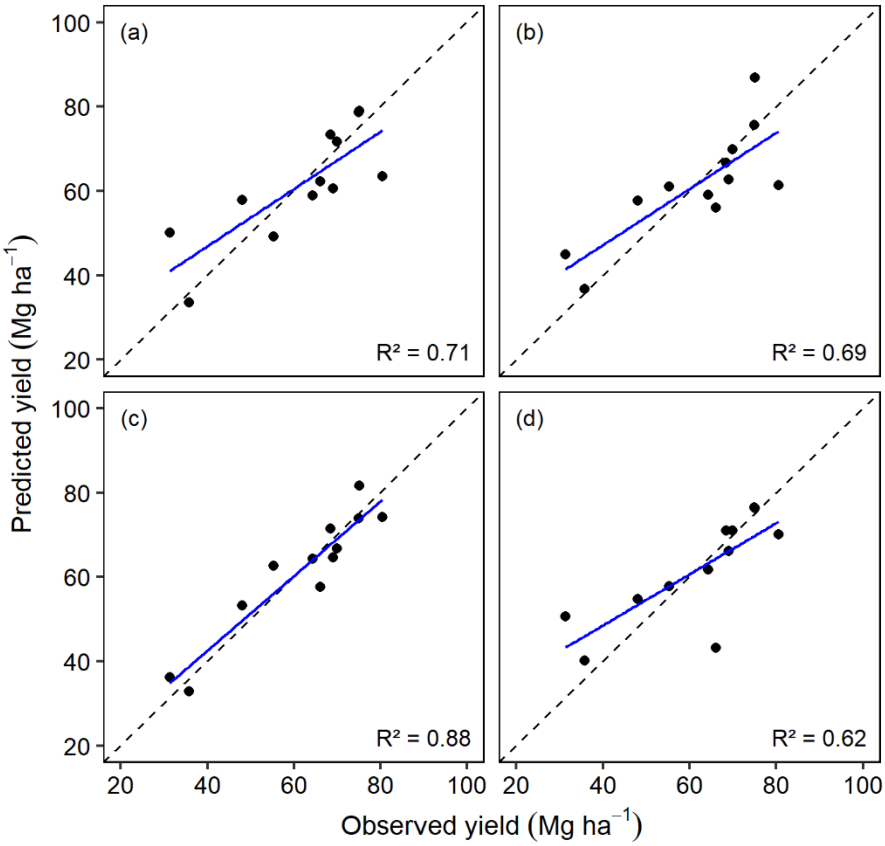

Figure 10. Relationship between observed and predicted yields for the first (a), second (b), third (c), and fourth (d) flights (reprinted from Cholula et al., 2020; permissions for reproduction have been obtained from the copyright holders).

The relationship between observed TCC and predicted TCC is shown in Figure 11. The best agreement between the observed and predicted TCC is provided by the model for the third flight with a coefficient of determination of 0.30 (Figure 11c). On the other hand, the lowest relationship of TCC is observed in the model for the second flight (Figure 7d).

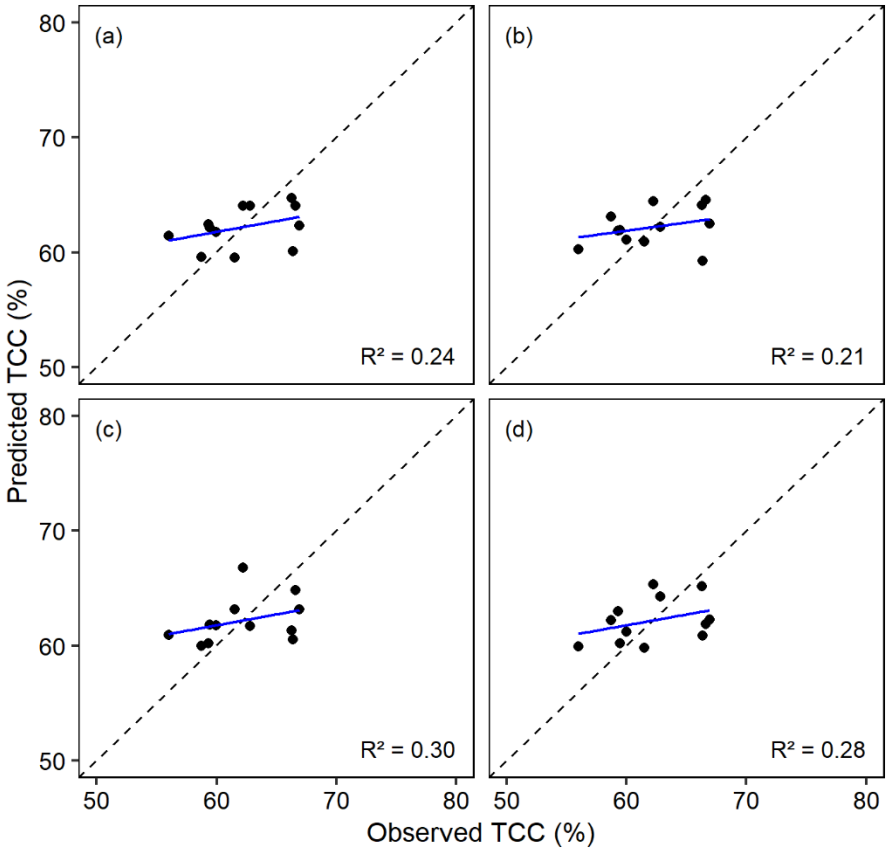

Figure 11. Relationship between the observed and predicted TCC for the first (a), second (b), third (c), and fourth (d) flights (reprinted from Cholula et al., 2020; permissions for reproduction have been obtained from the copyright holders).

4. DISCUSSION*

True color and multispectral imagery can be used to extract plant measurements and vegetation indices to estimate yield. However, it is necessary to know the best time for data collection to make the predictions, so there is a need to assess the accuracy of the models generated for each of the flights. A previous study by Sanches et al. (2018) found that the inflection point of biomass accumulation by the crop is a useful reference in the estimation of sugarcane yield with UAS images.

NDVI has been used in several studies of yield prediction, and it has shown good results (Casadesús et al., 2007; Kyratzis et al., 2017; Pinheiro Lisboa et al., 2018; Zhou et al., 2017). In this study, it was significantly or moderately correlated to yield for the first (r = 0.84) and second (r = 0.55) assessments. Nevertheless, the higher correlation was observed during the first flight, which indicates that this crop stage can be used to collect UAS imagery for energy cane yield prediction.

Sugarcane height is highly correlated with biomass and nitrogen uptake (Portz et al., 2012). Additionally, height can be an indicator of yield and other parameters since it is highly influenced by soil, total sugar content, leaf nitrogen, temperature, and light intensity (De Souza et al., 2017; S. Rahman, 2012). Silva et al. (2008) suggested that stalk height is a useful trait for sugarcane breeding to accelerate and reduce the costs of the process. Nevertheless, when sugarcane plants attain a certain height, they start to lodge,

* Part of this section is reprinted with permission from "Forecasting yield and lignocellulosic composition of energy cane using unmanned aerial systems" by Cholula, U., da Silva, J.A., Marconi, T., Thomasson, J.A., Solorzano J., and Enciso, J. (2020). *Agronomy, 10*, 718. Copyright 2020 Cholula, U., da Silva, J.A., Marconi, T., Thomasson, J.A., Solorzano J., and Enciso, J.

and environmental factors such as wind can cause plant breakage (De Souza et al., 2017). In this experiment, plant height was also affected by wind, and for the last two assessments, lodging was evident. The yield models for the third and fourth flights contain canopy height as one of their predictors; however, since the measurements after lodging are not representative, these models could have some limitations estimating energy cane yield.

Canopy cover is related to crop growth and water use (Trout et al., 2008); hence, it is an essential parameter in crop monitoring. Wiedenfeld and Enciso (2008) found that sugarcane can compensate for differences in water levels producing maximum yields (Wiedenfeld & Enciso, 2008). In this study, canopy cover did not play an essential role in yield prediction; it was moderately correlated with the yield for the first (r = 0.32) and the fourth (r = 0.44) assessments. However, it was strongly correlated with the yield for the second assessment (r = 0.70). These results are due mainly to the thresholds used to separate green vegetation from non-vegetation according to the canopeo algorithm (Patrignani & Ochsner, 2015).

ExG is useful for discriminating between green and non-green vegetation (Woebbecke et al., 1995). This index has been used coupled with crop classified mean heights to predict corn grain yield and showed good results (Geipel et al., 2014). Opposite to these findings, in the present study, the mean ExG values for the first, second, and third assessments were not significantly correlated to energy cane yield; it was only significantly correlated to yield for the fourth flight (r = 0.60).

In this study, a stepwise regression analysis was implemented to obtain the yield and TCC models, and the accuracy was assessed by means R^2, p-value, and RMSE. Energy cane yield was satisfactorily estimated by two models, built with UAS data collected 273 and 153 days before harvest. The predictor for the first model is NDVI, which implies that the acquisition of multispectral imagery is required, whereas, for the second model, the predictors are CC, CH, and ExG, which can be extracted from RGB images. The performance of these models was also satisfactory since RMSE ranged from 6.26 to 8.89 Mg ha^{-1}. These findings are similar to the model performances found by Pinheiro Lisboa et al. (2018), of between 0.24 and 10.34 Mg ha^{-1} and the RMSE reported by Fernandes et al. (2017), ranging from 7.20 to 11.0 Mg ha^{-1}. Nevertheless, the performance of the model for the fourth flight was not satisfactory, since it was 11.39 Mg ha^{-1}.

The previous findings may apply to other locations if the environmental conditions at the time of the data collection are appropriate. It can also be applied to other crops such as sorghum (Shafian et al., 2018), maize (Maresma et al., 2016), wheat (Fu et al., 2020), and rice (Zhou et al., 2017). Furthermore, other variables that may be included in future yield prediction models could be abiotic stresses, such as temperature, drought, and soil salinity, which would increase the importance of this technology as a valuable tool for the breeding of stress tolerant varieties (Ashapure et al., 2019).

TCC was not successfully estimated by the models created, and the main reason for this was the lack of correlation of TCC with the variables used. The coefficient of determination ranged from 0.21 to 0.30. These results agree with those reported by Chea et al. (2020), who found a weak correlation (R^2 = 0.35 – 0.50) between fiber content with

GNDVI, NDVI, RVI (Ratio Vegetation Index), CIgreen (Green Chlorophyll Index), CIrededge (Red Edge Chlorophyll Index), and SRPIb (Simple Ratio Pigment Index). Moreover, Roberts et al. (2011) highlighted that lignocellulosic content indices rely on short-wave-infrared (SWIR) wavelengths, and wavelengths from 1500 to 1800 nm and 2000 to 2350 nm, which compare reflectance at an absorbing wavelength to a non-absorbing wavelength. In this case, hyperspectral imagery was not available to calculate indices such as CAI (Cellulose Absorption Index) (Daughtry, 2001) or NDLI (Normalized Difference Lignin Index) (Serrano et al., 2002), so this could be a limiting factor to predict TCC in energy cane.

5. CONCLUSIONS

This study assessed the utilization of UAS in the prediction of yield and composition of energy cane. Crop parameters and vegetation indices (NDVI and ExG) were extracted from true color and multispectral imagery to generate the models. The yield was satisfactorily estimated by two of the models created, the first model with data collected 273 days before harvest (R^2 = 0.71) and the second with data collected 153 days before harvest (R^2 = 0.88). TCC was not estimated satisfactorily; the highest coefficient of determination was 0.30. This study demonstrated the potential application of UAS to estimate energy cane yield with high accuracy, enabling plant breeders to phenotype larger populations and make selections with higher confidence. Further investigation is required fort TCC prediction considering the use of hyperspectral imagery.

REFERENCES

Ashapure, A., Oh, S., Marconi, T. G., Chang, A., Jung, J., Landivar, J., & Enciso, J. (2019). Unmanned aerial system based tomato yield estimation using machine learning. In J. A. Thomasson, M. McKee, & R. J. Moorhead (Eds.), *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IV* (Vol. 11008, p. 22). SPIE. https://doi.org/10.1117/12.2519129

Berding, N., Brotherton, G. A., le Brocq, D. G., & Skinner, J. C. (1991). Near infrared reflectance spectroscopy for analysis of sugarcane from clonal evaluation trials: I. Fibrated Cane. *Crop Science*, *31*(4), 1017–1023. https://doi.org/10.2135/cropsci1991.0011183x003100040035x

Berger, B., Parent, B., & Tester, M. (2010). High-throughput shoot imaging to study drought responses. *Journal of Experimental Botany*, *61*(13), 3519–3528. https://doi.org/10.1093/jxb/erq201

Berni, J. A. J., Zarco-Tejada, P. J., Suárez, L., & Fereres, E. (2009). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(3), 722–738. https://doi.org/10.1109/TGRS.2008.2010457

Bischoff, K. P., Gravois, K. A., Schexnayder Jr., H. P., & Hawkins, G. L. (2001). The effect of harvest method and plot size on the estimation of sugarcane yield. *Journal American Society of Sugar Cane Technologists*, *21*, 51–60.

Carvalho-Netto, O. V., Bressiani, J. A., Soriano, H. L., Fiori, C. S., Santos, J. M.,

Barbosa, G. V. S., Xavier, M. A., Landell, M. G. A., & Pereira, G. A. G. (2014). The potential of the energy cane as the main biomass crop for the cellulosic industry. *Chemical and Biological Technologies in Agriculture*, *1*(1), 1–8. https://doi.org/10.1186/s40538-014-0020-2

Casadesús, J., Kaya, Y., Bort, J., Nachit, M. M., Araus, J. L., Amor, S., Ferrazzano, G., Maalouf, F., Maccaferri, M., Martos, V., Ouabbou, H., & Villegas, D. (2007). Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. *Annals of Applied Biology*, *150*(2), 227–236. https://doi.org/10.1111/j.1744-7348.2007.00116.x

Chea, C., Saengprachatanarug, K., Posom, J., Wongphati, M., & Taira, E. (2020). Sugar yield parameters and fiber prediction in sugarcane fields using a multispectral camera mounted on a small unmanned aerial system (UAS). *Sugar Tech*, 1–17. https://doi.org/10.1007/s12355-020-00802-5

Cholula, U., da Silva, J. A., Marconi, T., Thomasson, J. A., Solorzano, J., & Enciso, J. (2020). Forecasting yield and lignocellulosic composition of energy cane using unmanned aerial systems. *Agronomy*, *10*(5), 718. https://doi.org/10.3390/agronomy10050718

Daughtry, C. S. T. (2001). Discriminating crop residues from soil by shortwave infrared reflectance. *Agronomy Journal*, *93*(1), 125–131. https://doi.org/10.2134/agronj2001.931125x

De Souza, C. H. W., Lamparelli, R. A. C., Rocha, J. V., & Magalhães, P. S. G. (2017). Height estimation of sugarcane using an unmanned aerial system (UAS) based on

structure from motion (SfM) point clouds. *International Journal of Remote Sensing*, *38*(8–10), 2218–2230. https://doi.org/10.1080/01431161.2017.1285082

Diaz-Varela, R. A., Zarco-Tejada, P. J., Angileri, V., & Loudjani, P. (2014). Automatic identification of agricultural terraces through object-oriented analysis of very high resolution DSMs and multispectral imagery obtained from an unmanned aerial vehicle. *Journal of Environmental Management*, *134*, 117–126. https://doi.org/10.1016/j.jenvman.2014.01.006

Ellis, R. D., & Merry, R. E. (2004). Sugarcane Agriculture. In *Sugarcane* (pp. 101–142). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470995358.ch5

Enciso, J., Avila, C. A., Jung, J., Elsayed-Farag, S., Chang, A., Yeom, J., Landivar, J., Maeda, M., & Chavez, J. C. (2019). Validation of agronomic UAV and field measurements for tomato varieties. *Computers and Electronics in Agriculture*, *158*, 278–283. https://doi.org/10.1016/j.compag.2019.02.011

Fernandes, J. L., Ebecken, N. F. F., & Esquerdo, J. C. D. M. (2017). Sugarcane yield prediction in Brazil using NDVI time series and neural networks ensemble. *International Journal of Remote Sensing*, *38*(16), 4631–4644. https://doi.org/10.1080/01431161.2017.1325531

Fu, Z., Jiang, J., Gao, Y., Krienke, B., Wang, M., Zhong, K., Cao, Q., Tian, Y., Zhu, Y., Cao, W., & Liu, X. (2020). Wheat growth monitoring and yield estimation based on multi-rotor unmanned aerial vehicle. *Remote Sensing*, *12*(3), 508. https://doi.org/10.3390/rs12030508

Geipel, J., Link, J., & Claupein, W. (2014). Combined spectral and spatial modeling of

corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. *Remote Sensing*, *6*(11), 10335–10355. https://doi.org/10.3390/rs61110335

Khaki, S., Wang, L., & Archontoulis, S. V. (2020). A CNN-RNN framework for crop yield prediction. *Frontiers in Plant Science, 10*, 1750. https://doi.org/10.3389/fpls.2019.01750

Kim, M., & Day, D. F. (2011). Composition of sugar cane, energy cane, and sweet sorghum suitable for ethanol production at Louisiana sugar mills. *Journal of Industrial Microbiology and Biotechnology*, *38*(7), 803–807. https://doi.org/10.1007/s10295-010-0812-8

Kyratzis, A. C., Skarlatos, D. P., Menexes, G. C., Vamvakousis, V. F., & Katsiotis, A. (2017). Assessment of vegetation indices derived by UAV imagery for durum wheat phenotyping under a water limited and heat stressed Mediterranean environment. *Frontiers in Plant Science, 8*, 1114. https://doi.org/10.3389/fpls.2017.01114

Levitt, J. (1980). II Water, radiation, salt and other stresses. In *Responses of plants to environmental stresses* (2nd ed., Issue Ed. 2, p. 607). Academic Press.

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

Maresma, Á., Ariza, M., Martínez, E., Lloveras, J., & Martínez-Casasnovas, J. (2016).

Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (*Zea mays* L.) from a standard UAV Service. *Remote Sensing*, *8*(12), 973. https://doi.org/10.3390/rs8120973

Mathews, A. J., & Jensen, J. L. (2013). Visualizing and quantifying vineyard canopy LAI using an unmanned aerial vehicle (UAV) collected high density structure from motion point cloud. *Remote Sensing*, *5*(5), 2164–2183. https://doi.org/10.3390/rs5052164

Matsuoka, S., Kennedy, A. J., Santos, E. G. D. dos, Tomazela, A. L., & Rubio, L. C. S. (2014). Energy cane: Its Concept, development, characteristics, and prospects. *Advances in Botany*, *2014*, 1–13. https://doi.org/10.1155/2014/597275

Park, J.-W., Benatti, T. R., Marconi, T., Yu, Q., Solis-Gracia, N., Mora, V., & Da Silva, J. A. (2015). Cold responsive gene expression profiling of sugarcane and Saccharum spontaneum with functional analysis of a cold inducible Saccharum homolog of NOD26-like intrinsic protein to salt and water stress. *PLoS ONE*, *10*(5). https://doi.org/10.1371/journal.pone.0125810

Patrignani, A., & Ochsner, T. E. (2015). Canopeo: A powerful new tool for measuring fractional green canopy cover. *Agronomy Journal*, *107*(6), 2312–2320. https://doi.org/10.2134/agronj15.0150

Pinheiro Lisboa, I., Melo Damian, J., Roberto Cherubin, M., Silva Barros, P. P., Ricardo Fiorio, P., Cerri, C. C., & Pellegrino Cerri, C. E. (2018). Prediction of sugarcane yield based on NDVI and concentration of leaf-tissue nutrients in fields managed with straw removal. *Agronomy*, *8*(9), 196.

https://doi.org/10.3390/agronomy8090196

Portz, G., Amaral, L. R., & Molin, J. P. (2012). Measuring sugarcane height in complement to biomass. In *Proceedings of the 10th International Conference on Precision Agriculture (ISPA)*.

Rahaman, M. M., Chen, D., Gillani, Z., Klukas, C., & Chen, M. (2015). Advanced phenotyping and phenotype data analysis for the study of plant growth and development. *Frontiers in Plant Science*, *6*, 619. https://doi.org/10.3389/fpls.2015.00619

Rahman, M. M., & J. Robson, A. (2016). A Novel approach for sugarcane yield prediction using Landsat time series imagery: A case study on Bundaberg region. *Advances in Remote Sensing*, *5*(2), 93–102. https://doi.org/10.4236/ars.2016.52008

Rahman, M. M., & Robson, A. (2020). Integrating Landsat-8 and Sentinel-2 time series data for yield prediction of sugarcane crops at the block level. *Remote Sensing*, *12*(8), 1313. https://doi.org/10.3390/RS12081313

Rahman, S. (2012). *Growth, yield and quality of plant and ratoon crops of sugarcane as affected by plant material and management practices* (Doctoral dissertation). University of Rajshahi, Department of Agronomy and Agricultural Extension.

Richards, J. (1990). Computer processing of remotely-sensed images: An introduction. *Earth-Science Reviews*, *27*(4), 392–394. https://doi.org/10.1016/0012-8252(90)90075-7

Roberts, D. A., Roth, K. L., & Perroy, R. L.(2016). Hyperspectral vegetation indices. In P. S. Thenkabail, A. Huete, & J. G. Lyon (Eds.), *Hyperspectral Remote Sensing of*

Vegetation (pp. 309–328). CRC Press. https://doi.org/10.1201/b11222-20

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Goddard Space Flight Center 3rd ERTS-1 Symposium*, 309–317. https://ntrs.nasa.gov/search.jsp?R=19740022614

Salassi, M. E., Brown, K., Hilbun, B. M., Deliberto, M. A., Gravois, K. A., Mark, T. B., & Falconer, L. L. (2014). Farm-scale cost of producing perennial energy cane as a biofuel feedstock. *Bioenergy Research*, *7*(2), 609–619. https://doi.org/10.1007/s12155-013-9390-8

Salassi, M. E., Falconer, L. L., Mark, T. B., Deliberto, M. A., Hilbun, B. M., & Cooper, T. L. (2015). Economic potential for energy cane production as a cellulosic biofuel feedstock in the southeastern United States. *AIMS Energy*, *3*(1), 25–40. https://doi.org/10.3934/energy.2015.1.25

Salekdeh, G. H., Reynolds, M., Bennett, J., & Boyer, J. (2009). Conceptual framework for drought phenotyping during molecular breeding. *Trends in Plant Science*, *14*(9), 488–496. https://doi.org/10.1016/j.tplants.2009.07.007

Sanches, G. M., Duft, D. G., Kölln, O. T., Luciano, A. C. dos S., De Castro, S. G. Q., Okuno, F. M., & Franco, H. C. J. (2018). The potential for RGB images obtained using unmanned aerial vehicle to assess and predict yield in sugarcane fields. *International Journal of Remote Sensing*, *39*(15–16), 5402–5414. https://doi.org/10.1080/01431161.2018.1448484

Scortecci, K. C., Creste, S., Calsa, T., Xavier, M. A., Landell, M. G. A., Figueira, A., & Benedito, V. A. (2012). Challenges, opportunities and recent advances in sugarcane

breeding. In I. Abdurakhmonov (Ed.), *Plant Breeding* (pp. 267–296). InTech.

Serrano, L., Peñuelas, J., & Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, *81*(2–3), 355–364. https://doi.org/10.1016/s0034-4257(02)00011-1

Shafian, S., Rajan, N., Schnell, R., Bagavathiannan, M., Valasek, J., Shi, Y., & Olsenholler, J. (2018). Unmanned aerial systems-based remote sensing for monitoring sorghum growth and development. *PLoS ONE*, *13*(5). https://doi.org/10.1371/journal.pone.0196605

Silva, J. A. G. da, Sorrells, M. E., Burnquist, W. L., & Tanksley, S. D. (1993). RFLP linkage map and genome analysis of Saccharum spontaneum. *Genome*, *36*(4), 782–791. https://doi.org/10.1139/g93-103

Silva, M. D. A., Da Silva, J. A. G., Enciso, J., Sharma, V., & Jifon, J. (2008). Yield components as indicators of drought tolerance of sugarcane. *Scientia Agricola*, *65*(6), 620–627. https://doi.org/10.1590/S0103-90162008000600008

Somerville, C., Youngs, H., Taylor, C., Davis, S. C., & Long, S. P. (2010). Feedstocks for lignocellulosic biofuels. *Science*, *329*(5993), 790–792.

Swain, K. C., Thomson, S. J., & Jayasuriya, H. P. W. (2010). Adoption of an unmanned helicopter for low-altitude remote sensing to estimate yield and total biomass of a rice crop. *Transactions of the ASABE*, *53*(1), 21–27.

Takeda, S., & Matsuoka, M. (2008). Genetic approaches to crop improvement: Responding to environmental and population changes. *Nature Reviews Genetics*,

9(6), 444-457. https://doi.org/10.1038/nrg2342

Thenkabail, P. S., Smith, R. B., & De Pauw, E. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment, 71*(2), 158–182. https://doi.org/10.1016/S0034-4257(99)00067-X

Trout, T. J., Johnson, L. F., & Gartung, J. (2008). Remote sensing of canopy cover in horticultural crops. *HortScience, 43*(2), 333–337. https://doi.org/10.21273/hortsci.43.2.333

USDA-NASS. (2019). *Agricultural statistics 2019*. Washington, DC: USDA-NASS. Retrieved from https://www.nass.usda.gov/Publications/Ag_Statistics/2019/2019_complete_publication.pdf.

Wiedenfeld, B., & Enciso, J. (2008). Sugarcane responses to irrigation and nitrogen in semiarid South Texas. *Agronomy Journal*, *100*(3), 665-671. https://doi.org/10.2134/agronj2007.0286

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., & Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE, 38*(1), 259–269. https://doi.org/10.13031/2013.27838

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., Xu, B., Yang, X., Zhu, D., Zhang, X., Zhang, R., Feng, H., Zhao, X., Li, Z., Li, H., & Yang, H. (2017). Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives. *Frontiers in Plant Science, 8,* 1111. https://doi.org/10.3389/fpls.2017.01111

You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep gaussian process for crop yield prediction based on remote sensing data. In *Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*.

Yu, X., Liu, Q., Wang, Y., Liu, X., & Liu, X. (2016). Evaluation of MLSR and PLSR for estimating soil element contents using visible/near-infrared spectroscopy in apple orchards on the Jiaodong peninsula. *Catena*, *137*, 340–349. https://doi.org/10.1016/j.catena.2015.09.024

Zarco-Tejada, P. J., Diaz-Varela, R., Angileri, V., & Loudjani, P. (2014). Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods. *European Journal of Agronomy*, *55*, 89–99. https://doi.org/10.1016/j.eja.2014.01.004

Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., Yao, X., Cheng, T., Zhu, Y., Cao, W. X., & Tian, Y. C. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *130*, 246–255. https://doi.org/10.1016/j.isprsjprs.2017.05.003

APPENDIX A

CROP PARAMETERS AND VEGETATION INDICES

Table 6. Data extracted from RGB and multispectral orthomosaic images.

Variety	Rep	Plot ID	CC[1]	CC[2]	CC[3]	CC[4]	CH[1]	CH[2]	CH[3]	CH[4]	NDVI[1]	NDVI[2]	NDVI[3]	NDVI[4]	ExG[1]	ExG[2]	ExG[3]	ExG[4]
Ho02-113	1	6	84.02	92.24	73.24	59.28	2.35	3.67	3.71	2.78	0.53	0.57	0.59	0.45	0.26	0.27	0.20	0.17
Ho02-113	2	36	77.94	94.97	72.91	30.36	2.26	4.06	4.68	3.31	0.62	0.60	0.58	0.45	0.21	0.25	0.17	0.12
Ho02-113	3	44	62.72	89.28	77.88	33.78	2.28	3.69	4.60	2.61	0.55	0.58	0.60	0.52	0.14	0.24	0.19	0.15
Ho02-113	4	72	77.92	97.02	61.69	60.79	2.86	4.39	4.96	3.35	0.62	0.61	0.60	0.58	0.21	0.23	0.14	0.17
TH16-19	1	3	77.92	80.83	62.64	51.67	2.31	3.92	3.34	2.57	0.48	0.53	0.62	0.44	0.24	0.22	0.19	0.16
TH16-19	2	34	57.54	77.88	62.23	24.73	2.39	3.77	3.43	2.53	0.54	0.58	0.59	0.44	0.16	0.20	0.19	0.13
TH16-19	3	55	38.51	91.41	71.39	33.33	2.25	3.85	3.56	2.78	0.59	0.61	0.61	0.49	0.16	0.24	0.22	0.14
TH16-19	4	78	63.51	82.36	72.73	40.29	2.60	4.01	3.87	3.17	0.59	0.55	0.65	0.54	0.20	0.20	0.23	0.13
TH16-25	1	11	42.90	70.37	71.10	21.72	1.74	3.23	2.03	2.32	0.41	0.51	0.63	0.43	0.13	0.21	0.17	0.07
TH16-25	2	22	54.48	79.90	70.01	13.51	1.79	3.24	2.21	2.57	0.49	0.56	0.62	0.44	0.17	0.23	0.17	0.07
TH16-25	3	58	60.99	86.90	78.34	16.05	1.89	3.44	3.42	2.63	0.52	0.61	0.61	0.48	0.16	0.25	0.19	0.08
TH16-25	4	68	37.75	85.69	63.22	12.59	1.98	3.50	3.34	2.50	0.54	0.59	0.62	0.39	0.14	0.22	0.16	0.05

[1] First flight, [2] second flight, [3] third flight, and [4] fourth flight.

Table 7. Observed yield at the end of the season and predicted yield with the aerial data collected at different growth stages.

Variety	Rep	Plot ID	Observed yield (Mg ha^{-1})	Predicted yield (Mg ha^{-1})			
				First flight	Second flight	Third Flight	Fourth flight
Ho02-113	1	6	64.29	58.89	59.12	64.35	61.70
Ho02-113	2	36	80.49	63.43	61.44	74.09	70.06
Ho02-113	3	44	75.11	79.00	86.95	81.54	76.34
Ho02-113	4	72	74.92	78.66	75.63	73.89	76.42
TH16-19	1	3	55.32	49.20	61.00	62.68	57.80
TH16-19	2	34	69.08	60.61	62.74	64.59	66.08
TH16-19	3	55	69.93	71.77	69.85	66.65	70.93
TH16-19	4	78	68.47	73.35	66.75	71.43	70.96
TH16-25	1	11	35.72	33.55	36.70	32.77	40.20
TH16-25	2	22	31.36	50.14	44.90	36.15	50.65
TH16-25	3	58	48.09	57.94	57.78	53.22	54.64
TH16-25	4	68	66.12	62.35	56.04	57.54	43.11

Table 8. Observed TCC before harvest and predicted TCC with the aerial data collected at different growth stages.

Variety	Rep	Plot ID	Observed TCC (%)	Predicted TCC (%)			
				First flight	Second flight	Third Flight	Fourth flight
Ho02-113	1	6	66.29	64.74	64.09	61.32	65.17
Ho02-113	2	36	66.91	62.34	62.46	63.15	62.25
Ho02-113	3	44	62.23	64.05	64.45	66.78	65.34
Ho02-113	4	72	66.61	64.06	64.57	64.86	61.86
TH16-19	1	3	62.82	64.05	62.22	61.70	64.30
TH16-19	2	34	59.98	61.75	61.08	61.74	61.22
TH16-19	3	55	58.73	59.61	63.11	59.98	62.20
TH16-19	4	78	59.32	62.43	61.86	60.22	63.00
TH16-25	1	11	66.36	60.10	59.27	60.53	60.87
TH16-25	2	22	55.99	61.41	60.27	60.94	59.94
TH16-25	3	58	59.45	62.14	61.92	61.81	60.23
TH16-25	4	68	61.51	59.52	60.90	63.18	59.83

APPENDIX B

NDVI AND ExG MAPS

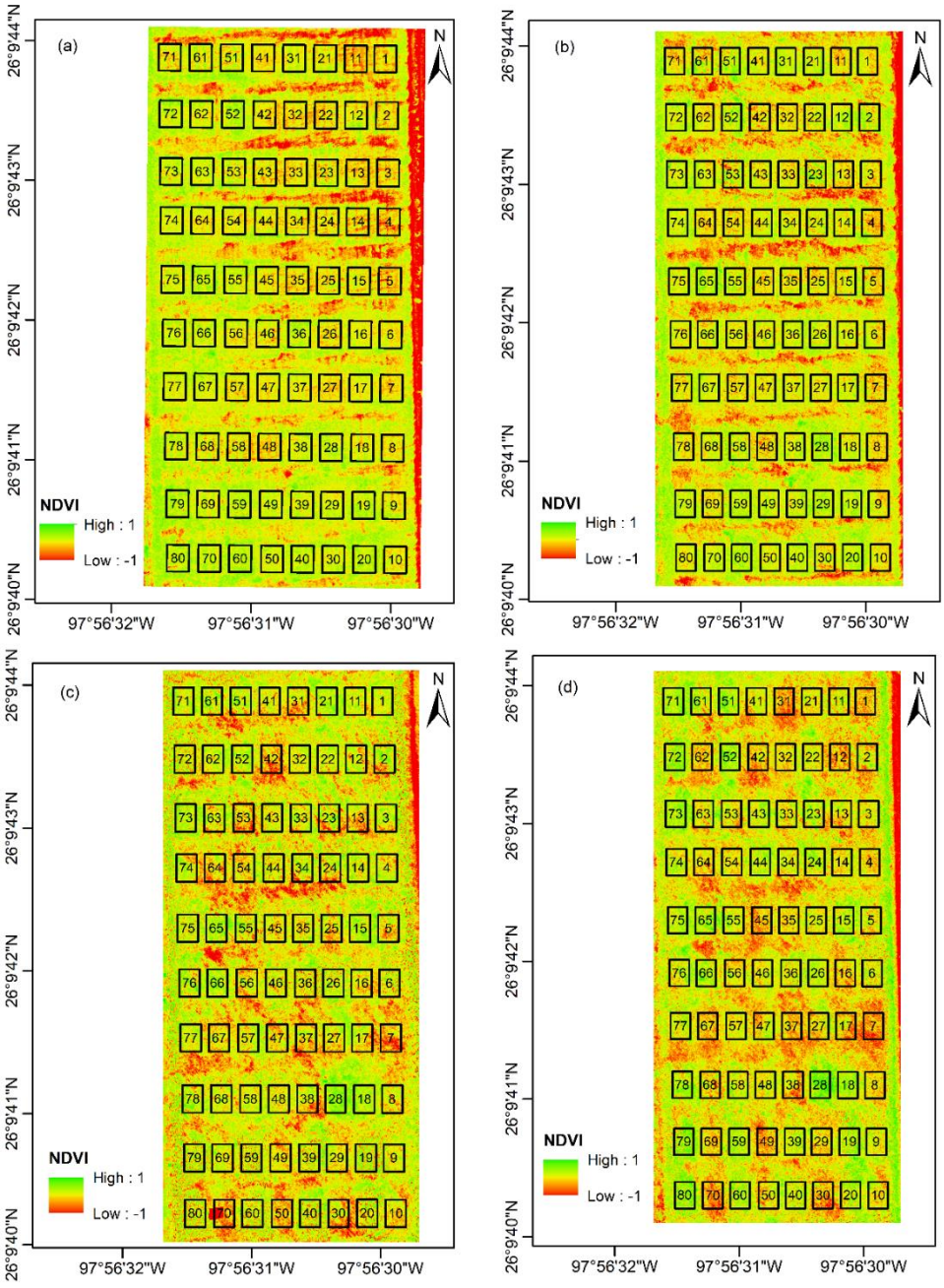

Figure 12. NDVI maps for the first (a), second (b), third (c), and fourth (d) flights.

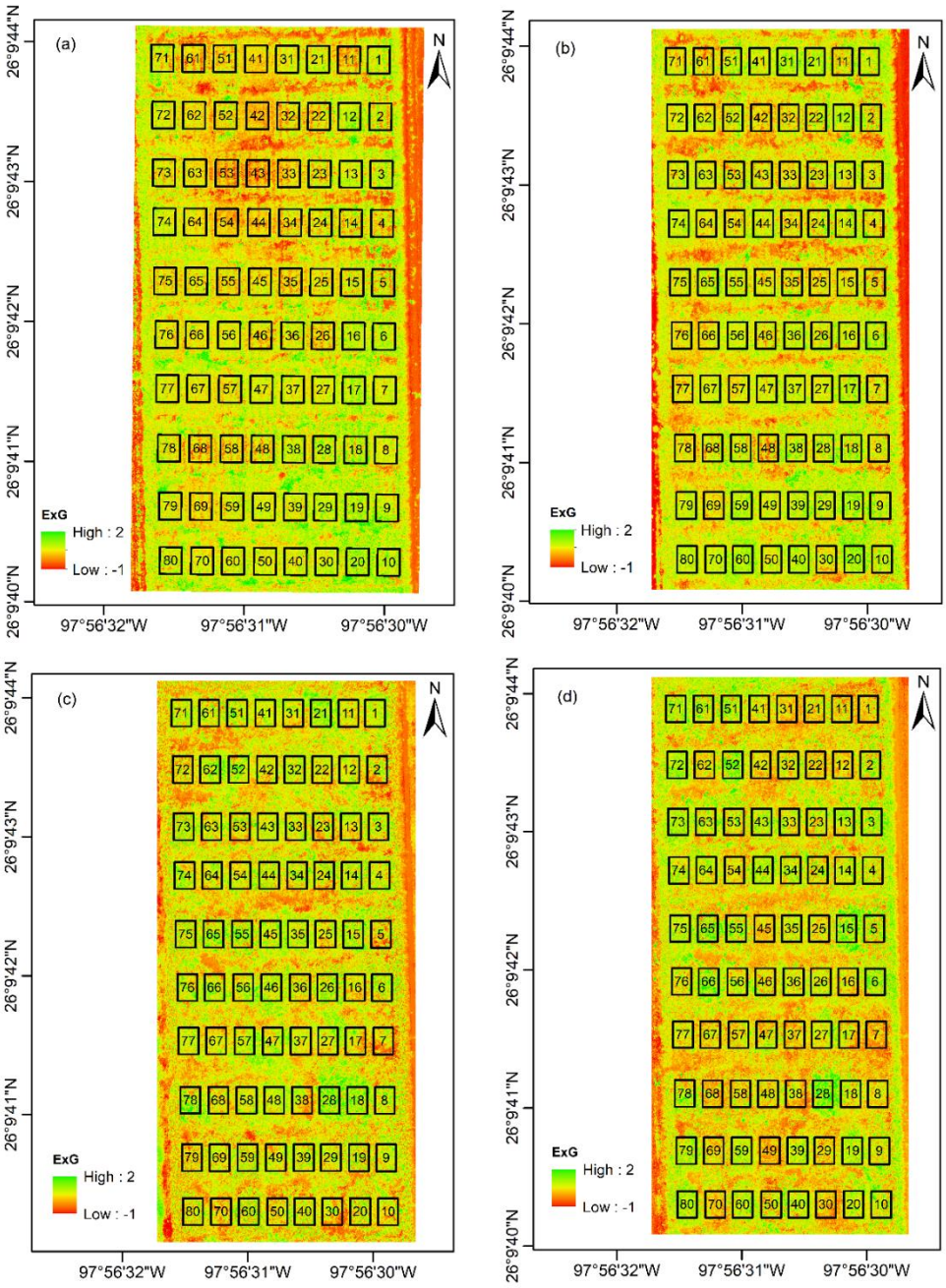

Figure 13. ExG maps for the first (a), second (b), third (c), and fourth (d) flights.