

PREDICTING COTTON ABOVE GROUND BIOMASS USING GEOAI AND
MULTISPECTRAL IMAGERY FROM UNCREWED AERIAL SYSTEMS

by

KYLE STEEN

(Under the Direction of Marguerite Madden)

ABSTRACT

Monitoring the biomass of cotton allows agriculturalists to modify their management practices to optimize yield and address issues such as damage from storms and impacts of climate change. Across the 2018 and 2019 growing seasons, remote sensing images were acquired by the U.S. Department of Agriculture (USDA) Agricultural Research Service (ARS) via Uncrewed Aerial Systems (UAS). These images provided high spatial and temporal resolution, multispectral image data for the analysis of cotton at the Ashburn Cooperator Farm in Ashburn, Georgia and the Ty Ty Cooperator Farm in Ty Ty, Georgia. Ground measurements of cotton biomass collected by the USDA-ARS harvesting plants in representative plots were used to scale up to the field level by using geographic artificial intelligence (GeoAI) machine learning models, Random Forest and XGBoost. Results displayed models using the raw bands predicted cotton biomass within the designated range and outperformed models using vegetation indices.

INDEX WORDS: Biomass, UAS, Cotton, Random Forest, XGBoost

PREDICTING COTTON ABOVE GROUND BIOMASS USING GEOAI AND
MULTISPECTRAL IMAGERY FROM UNCREWED AERIAL SYSTEMS

by

KYLE STEEN

B.S., University of Georgia, 2021

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment
of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2025

© 2025

Kyle Steen

All Rights Reserved

PREDICTING COTTON ABOVE GROUND BIOMASS USING GEOAI AND
MULTISPECTRAL IMAGERY FROM UNCREWED AERIAL SYSTEMS

by

KYLE STEEN

Major Professor:
Committee:

Marguerite Madden
Alisa Coffin
Lynne Seymour

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2025

DEDICATION

I dedicate this work to my family, for without them, I would not be where I am today.

ACKNOWLEDGEMENTS

I would like to express my gratitude to the USDA-ARS, SEWRL, and LTAR Network for financially supporting this project and providing data, support, and guidance throughout the course of this research. This project was supported by the Gulf Atlantic Long-Term Agroecosystem Research (LTAR) site (cooperative agreement with Marguerite Madden and Lynne Seymour grant contracts # 58-6048-7-036 and #58-6048-2-012).

I would also like to thank my advisor, Dr. Marguerite Madden, who guided me throughout my graduate studies, and my committee members, Dr. Alisa Coffin and Dr. Lynne Seymour, who provided critical insight into this project. I would also like to thank Dr. Sergio Bernardes, who provided geospatial wisdom whenever needed.

Additionally, I would also like to thank my friends and colleagues who supported me during my academic career including: Amanda Aragon, Vanessa Bailey, Justin Batchelor, Ben Bennett, Daisi Brand, Julie Brisson, Kayla Brown, Ellen Delgado Florian, Hassan Dogar, Seth Goss, Carl Griffiths, Eddie Griffiths, Justin Hong, Ryan Landry, Daniel Lane, Melina Le, Tyler Lynn, Shirin Malek, Kendrick Nelson, Ryan Palmer, Sarah Payne, Maya Rao, Megan Rich, Kelli Roberts, Shakirah Rogers, Kohta Schultz, Rajneesh Sharma, Coby Smith, Kayleigh Taillefer, Nathan Tesfayi, Nancee Uniyal, Hexiang Wang, Hailey Warren, Sydney Whilden, and Liz Wiggins.

I would like to give special thanks to the following individuals for their contributions to this study. Their support and assistance were instrumental in the successful completion of this research project.

- Dr. Marguerite Madden: University of Georgia - Department of Geography
- Dr. Lynne Seymour: University of Georgia - Department of Statistics
- Dr. Alisa Coffin: United States Department of Agriculture - Agricultural Research Service, Tifton, GA
- Dr. Sergio Bernardes: University of Georgia - Department of Geography
- Dr. Joseph Powell: University of Georgia - Department of Statistics
- Ivy Collins: University of Georgia – Department of Statistics
- Marley Holder: University of Georgia - Department of Statistics
- Kate Markham: University of Georgia - Department of Geography
- Tyler Lynn: University of Georgia - Department of Geography

Data provided by Dr. Alisa Coffin, Research Leader, of the U.S. Department of Agriculture-Agricultural Research Service, Southeast Watershed Research Laboratory, Tifton, GA, are being shared as part of the UGA NACA agreement (#58-6048-7-036). They are currently under embargo and may not be shared or published outside of use for thesis work. Public release of the data used in this thesis is forthcoming.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
CHAPTER	
1 INTRODUCTION	1
Study Area	4
2 LITERATURE REVIEW	8
Objectives	13
Data Sources	15
3 METHODS	20
Methods for Objective 1	20
Methods for Objective 2	23
Methods for Objective 3	26
4 RESULTS	27
Results for Objective 1.....	27
Results for Objective 2.....	28
Results for Objective 3.....	44
5 DISCUSSION AND CONCLUSIONS	48
REFERENCES	60

LIST OF APPENDICES

APPENDIX A.....	74
APPENDIX B.....	78

CHAPTER 1

INTRODUCTION

As the human population continues to grow exponentially, the demand to produce agricultural products also increases. Due to socioeconomic and population growth, it is estimated that the current agricultural production will need to double by the year 2050 to sustain humanity (Niu et al., 2019). Remotely sensed imagery and precision agriculture techniques enable the precise geolocation of within-field variations of crop health to predict Above Ground Biomass (AGB). Geographic technologies and analyses can, therefore, help mediate strains on current food and fiber production systems caused by changes in precipitation, temperature, and extreme weather events. Numerous studies in a variety of domains have geographically analyzed AGB at broad and fine scales including the prediction of global AGB of mangrove forests (Hu et al., 2020), and the prediction of AGB within an administrative zone of the Amazon Rainforest (D'Oliveira et al., 2020). Additionally, the pursuit of predicting AGB for agricultural areas has gained momentum over recent years (Bazzo et al., 2023).

Geographic Artificial Intelligence (GeoAI) is the use of artificial intelligence to solve geographic problems, and GeoAI has made significant advances in solving geographical problems over the past decades (Scheider & Richter, 2023). In particular, GeoAI uses remote sensing techniques to predict AGB at varying scales.

The use of Uncrewed Aerial Systems (UASs) promotes efficiency and innovation within modern agrarian practices, including the prediction of AGB. Numerous GeoAI studies have utilized imagery acquired from sensors onboard UASs and machine learning techniques to observe

vegetation and predict crop yield in relation to total AGB. For example, maize AGB was predicted utilizing Vegetation Indices (VIs) and a back-propagation neural network analysis for agricultural areas in North Dakota, U.S. (Panda et al., 2010). Predicting AGB is of critical importance to various groups including the agriculturalists, consumers, and government, for individuals from these groups can produce an accurate budget, prepare contingency plans, and anticipate market demands for specific crops (Deb et al., 2021).

The State of Georgia, U.S., is an agricultural hub dominated by cotton, peanuts, corn, and various other agricultural products (USDA, 2024a and 2024c). The southern region of Georgia, in the Coastal Plain of southeastern U.S., is where most of the state's agricultural products are grown. Cotton, (*Gossypium hirsutum* L.), is one of the most abundant crops within this region which generated over 868 million dollars for the state of Georgia in 2023, and this statistic places Georgia as the second highest cotton producing state in terms of sales (USDA 2024c). The U. S. Department of Agriculture - Agricultural Research Service (USDA-ARS) Southeast Watershed Research Laboratory (SEWRL) in Tifton, Georgia, has routinely monitored cotton agricultural areas in South Georgia, specifically within the Little River Experimental Watershed (LREW) (Coffin et al., 2022; USDA, 2024c).

The SEWRL and the LREW were created to evaluate the region's agricultural management practices and to further the understanding of hydrologic processes in the Coastal Plain region (Bosch et al., 2022). Throughout the history of the SEWRL, over 1000 journal articles have been published by SEWRL scientists from varying research fields including assessments of the hydrology in the watershed, characteristics of the riparian buffer system, impacts of land use changes, remote sensing, and water quality trends (Bosch et al. 2021). Utilizing UASs, based on its long history of agricultural research, the SEWRL manages the Gulf Atlantic Coastal Plain

(GACP) site within the Long-Term Agroecosystem Research (LTAR) network, a network operated by the USDA-ARS dedicated to researching agricultural solutions (*Figure 1.1*). Researchers at the GACP seek a better understanding of spatio-temporal patterns of crop growth and correlations with factors such as precipitation to scale crop production at multiple scales within and across LTAR regions.

The purpose for this research is to create a UAS and GeoAI framework and workflow which accurately predicts cotton AGB at two study sites within the LTAR-GACP for the 2018 and 2019 growing seasons. Scaling the prediction of AGB from harvested plants to plot and field levels was compared using numerous Vegetation Indices (VIs), remote sensing measurements used to assess landcover in addition to vegetation, derived from UAS imagery versus simply using the UAS raw bands, the latter being standard outputs from popular photogrammetric software, Pix4D

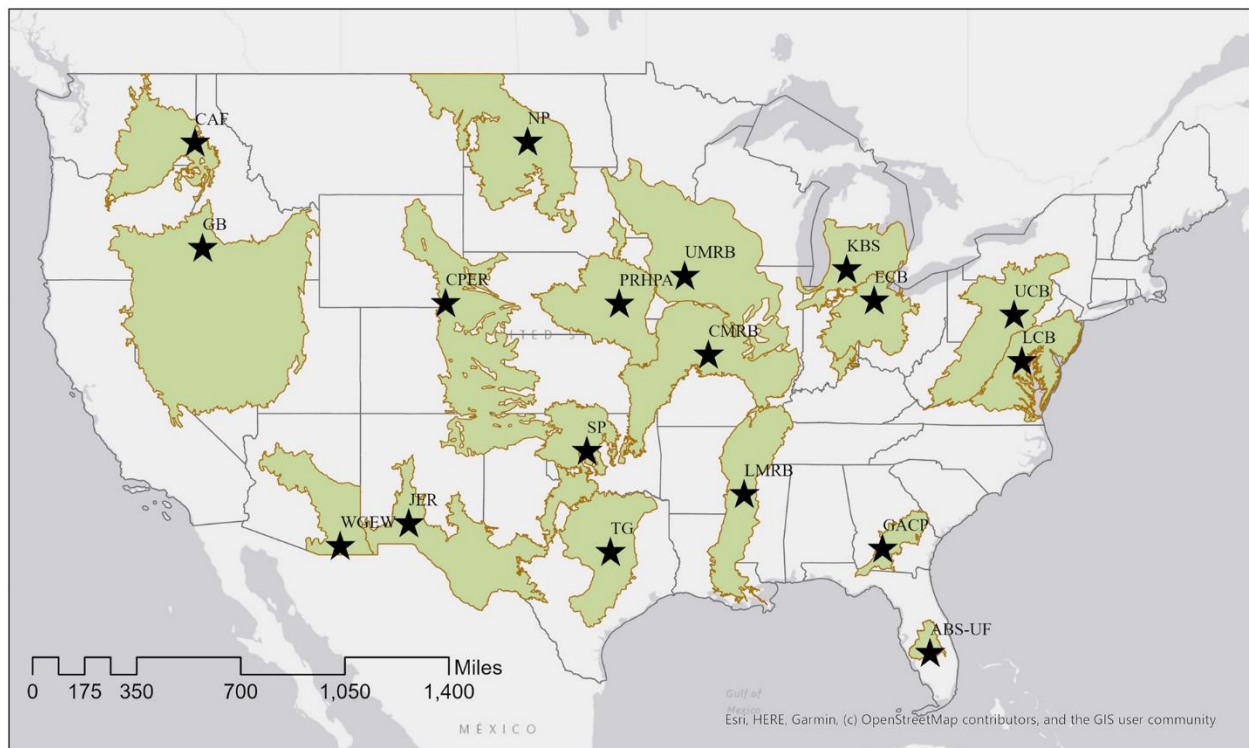


Figure 1.1: A map that displays all 18 LTAR agroecoregions across the United States. Each region consists of major agroecosystems such as cropland, graze/rangelands, and integrated systems, respectively (Bean et al., 2021).

(Ayanlade, 2017). The importance of scaling is critical to this project, as the stakeholders are extremely interested in scaling from plot and field level to regional levels. Additionally, a time-series of Normalized Difference Vegetation Index (NDVI) imagery acquired over the 2018 and 2019 growing seasons was used to determine phenological changes related to rainfall.

STUDY AREA

The overarching study area is the Gulf Atlantic Coastal Plain (GACP) LTAR site, containing 49 counties from Georgia and three from Florida, covering an area of 61,418 km² (Figure 1.2, Coffin et al., 2024). The GACP is located within the Southeastern Coastal Plain physiographic

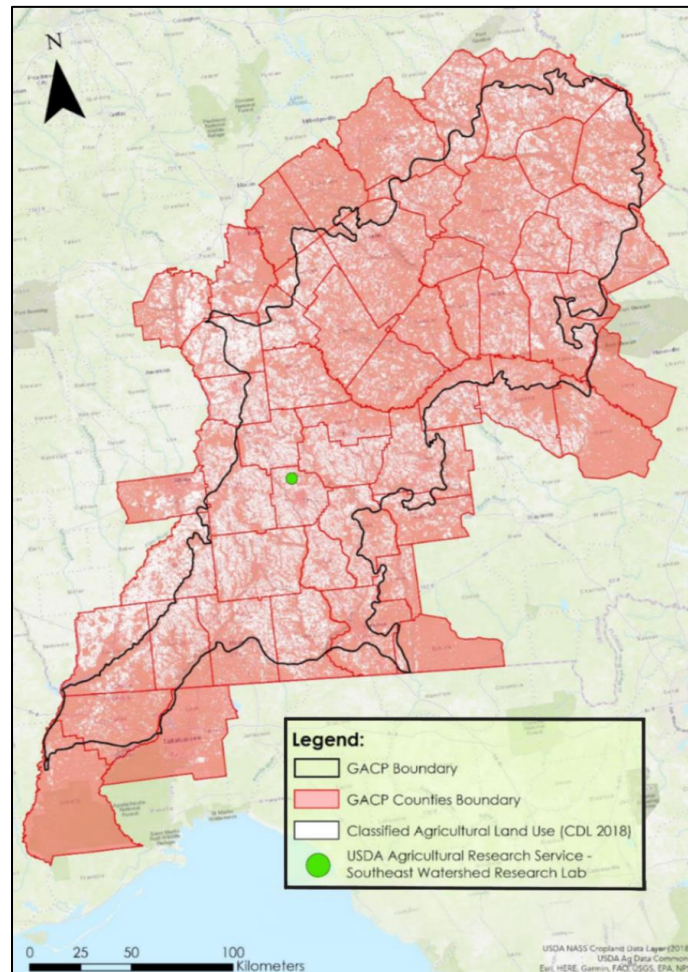


Figure 1.2: Map of the Gulf Atlantic Coastal Plain (GACP) boundary (outlined in black) intersected with county boundaries (indicated in red) as it spans from northern Florida through Georgia's Coastal Plain. The areas highlighted as white are classified agricultural land uses derived from the 2018 [Cropland Data Layer](#) (Stone, 2023).

region, and encompasses a majority of southern Georgia stretching from the northern portion of the Florida panhandle to the South Carolina border. It includes various agricultural, climatic, environmental, and social characteristics of the southeast U.S. (USDA, 2024b). The GACP has a topography of mild relief, mainly because it is entirely situated within the relatively flat Southeastern Coastal Plain. Elevation varies across the GACP ranging from 12 to 163 m above sea level (masl), with the average elevation being approximately 75 masl (Coffin et al., 2024a). Due to landscape consistency, numerous features are homogenous throughout the region including farming practices, geology, land use, and soil composition (Strickland, 2016b). The GACP has relatively high average annual rainfall at 1200mm, a mean annual temperature of 19°C, the coolest month is January at 11°C, and the warmest month is July at 27°C (Coffin et al., 2024a). Models predict extreme rainfall events to increase while moderate rainfall events decrease which heightens the potential for severe drought and flood risk (Coffin et al., 2024a).

Watering practices vary dramatically from systems of irrigation to natural precipitation. As drought conditions persist, supplying croplands with irrigation will become vital as seasonal precipitation does not perfectly align with crop production (Coffin et al., 2024a). The Fifth National Climate Assessment report of the U.S. Global Change Research Program (USGCRP) and the National Centers for Environmental Information of the National Oceanographic and Atmospheric Administration (NOAA) detail that agricultural systems within the region are subjected to concern due to increasing climatic trends of extreme heat and severe weather (NCEI 2023, USGCRP 2023). According to statistics derived from the USDA Cropland Data Layer, cotton overwhelmingly dominates the agricultural landscape within the GACP region (*Figure 1.3*) (USDA, 2024). The home institution for the GACP region of the LTAR network is the Southeast Watershed Research Laboratory (SEWRL) located in Tifton, Georgia (see *Figure 1.2*). One focus

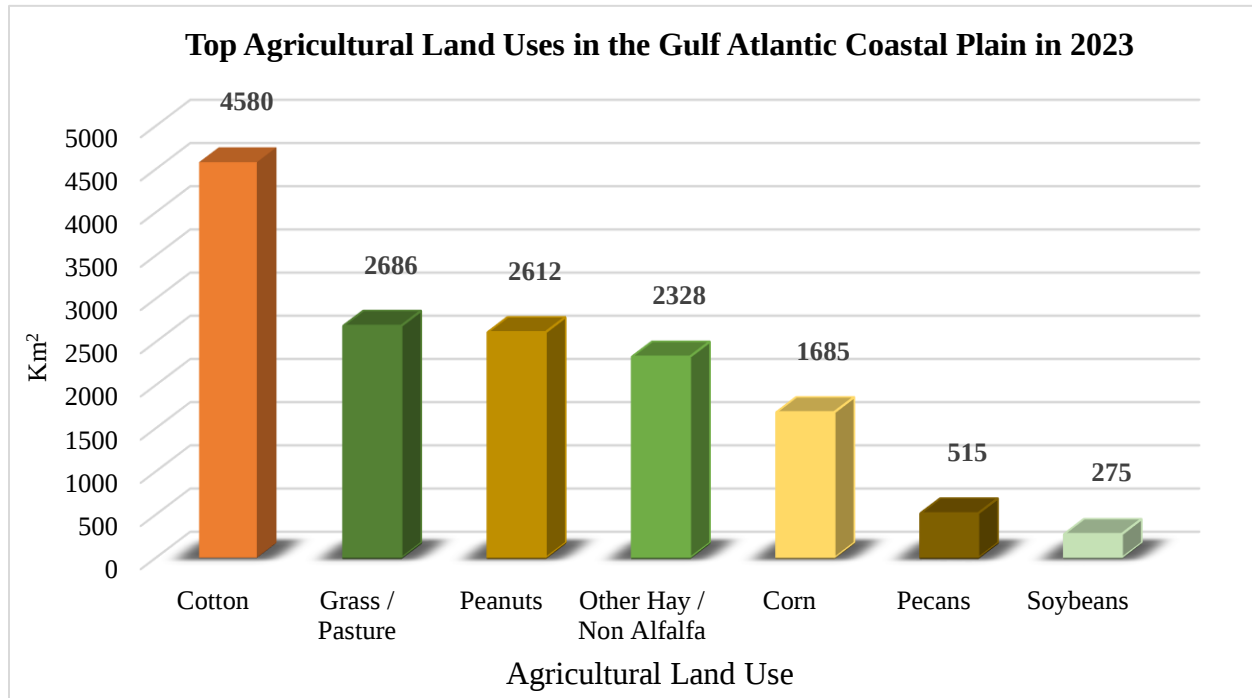


Figure 1.3: The top agricultural land uses in 2023 by km² as defined by the [Cropland Data Layer](#). Cotton is the overwhelming leader in agricultural land use in the GACP with 4580 km² of dedicated land followed by grass / pasture, peanuts, other hay/non-alfalfa, corn, pecans, and soybeans (USDA, 2024c).

of research by SEWRL is LTAR activities in collaboration with scientists from universities, other federal agencies and private landowners (USDA, 2024b). Researchers are committed to promoting sustainability within agricultural production by developing biofuels crops, incorporating production on marginal lands, integrating crop-livestock production, minimizing the importation of animal feeds, reducing agricultural water demand, and continuing to understand human impacts on agricultural landscapes (USDA, 2024b). The LREW, a 334 km² sub-watershed within the GACP region, is one of 12 national benchmark watersheds of key interest to the USDA-ARS (Sullivan et al., 2007). As a major U.S. interstate basin, the LREW is the headwaters of the Suwannee River Basin which begins in Georgia and empties into the Gulf of Mexico (Bosch et al., 2007). Within the LREW are two privately owned farms which have cooperative agreements with the SEWRL. They have been designated as the Ashburn Cooperator Farm (ACF), a non-irrigated

farm located in Ashburn, Georgia (*Figure 1.4a*), and the Ty Ty Cooperator Farm (TCF), an irrigated farm located in Ty Ty, Georgia (*Figure 1.4b*). These farms participate in the LTAR Network monitoring research for GACP and SEWRL researchers have been collecting AGB, soils, and biological data on the cropping systems here since 2014 (Coffin et al. 2024a).

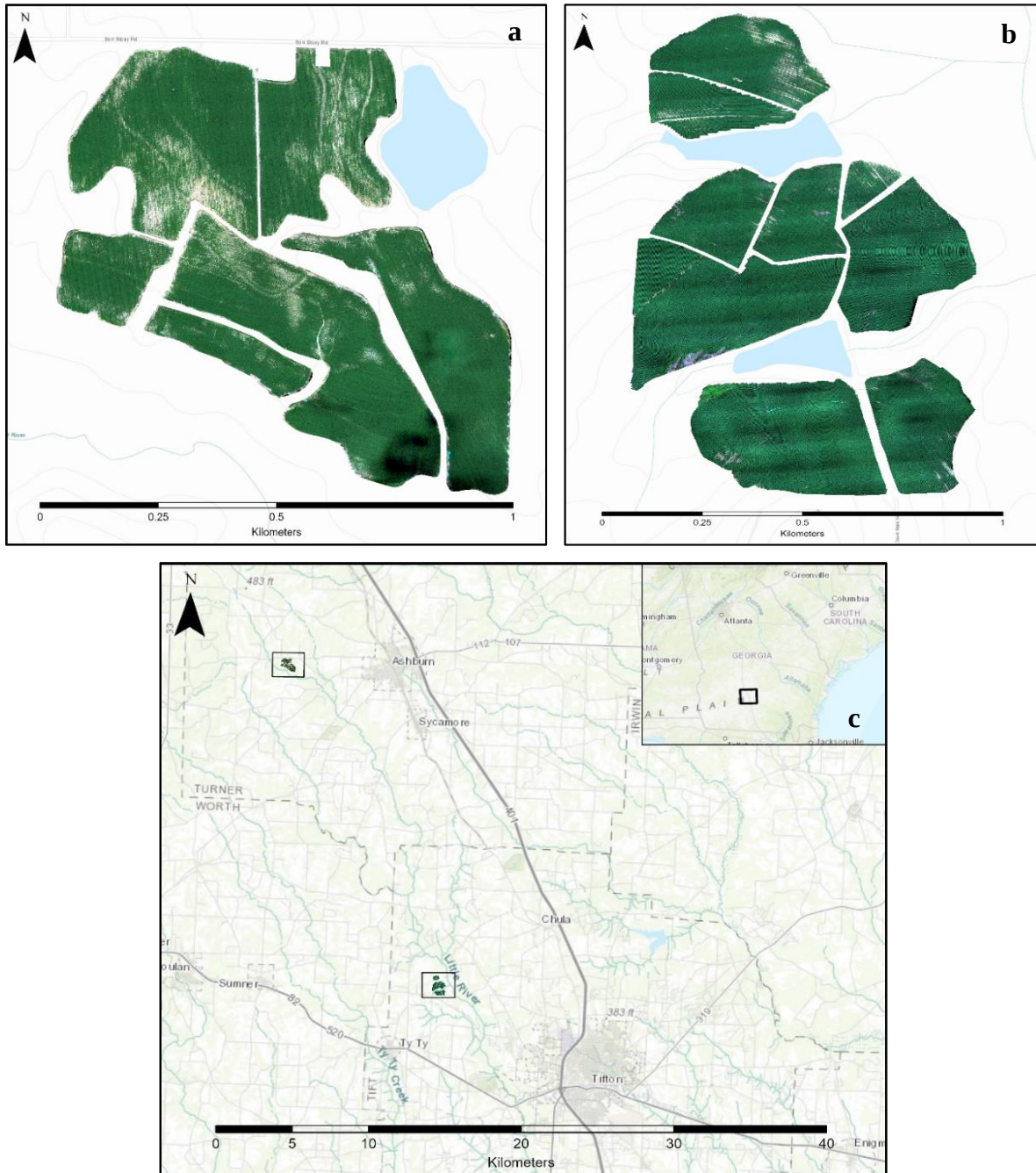


Figure 1.4: The Ashburn Cooperator Farm (a) and the Ty Ty Cooperator Farm (b) study site locations monitored by the LTAR GACP network, and (c) the surrounding contextual area.

CHAPTER 2

LITERATURE REVIEW

Above Ground Biomass (AGB) is an important agronomic parameter monitored in field studies, and it is frequently used to assess crop growth, crop health, the effectiveness of agricultural practices, and the carbon sequestration capacity of crops (Bendig et al., 2015). The accurate estimation of biomass is therefore key to understanding the current conditions impacting crops, and it is critical to predict biomass at varying scales. Predicting biomass can be accomplished using remotely sensed imagery and machine learning methods that fall within a Geographic Artificial Intelligence (GeoAI) framework, which is the application of AI techniques in geography and Earth sciences (Janowicz et al., 2019). Measuring biomass traditionally requires destructive measurements via harvesting, weighing, and recording crops, and this process is laborious, time-consuming, and extremely difficult to apply at broad scales for long-term measurements (Yang et al., 2018).

Non-destructive broad-scale biomass estimations have been previously pursued when attempting to measure AGB over vast areas of vegetation from remotely sensed images in relation to derived VIs computed from band ratios. One such study, pursued by Perry et al. (2014), utilized MODIS imagery to predict dryland wheat AGB across 37 paddocks in north-western Victoria, Australia. Results from this study yielded satisfactory metrics with a $R^2 = 0.81$ and a RMSE value of 164 kg ha^{-1} , but the spatial resolution of MODIS at 250 m, 500 m, or 1000 m depending on the band, is often too coarse to implement this framework for smaller farms. MODIS does have a tradeoff of high return time, but for this study the resolution is coarse. At a 30-m resolution,

Landsat imagery provides enhanced visualization for AGB prediction. Using imagery acquired by Landsat-7 Enhanced Thematic Mapper Plus and Landsat-8 Operational Land Imager sensors, winter wheat AGB was predicted for two growing seasons across 22 fields in southwestern Uruguay using a Simple Regression Model (SRM) with results indicating a RMSE value of 1532 and 966 kg ha⁻¹ for 2013 and 2014, respectively (Gasó et al., 2019). At a 10-m resolution with 13 bands ranging from the ultra-blue to the short-wave portion of the electromagnetic spectrum, Sentinel satellite imagery is often utilized for numerous agricultural applications. One study predicted sorghum AGB using Sentinel-2A and Sentinel-2B imagery across 42 fields near Bologna, Italy where an eXtreme Gradient Boosting – xgbtree (GBT) model performed consistently well ($r \geq 0.70$) in comparison to 13 other machine and deep learning models (Habyarimana et al., 2019). While satellite remote sensing is useful for monitoring biomass at broad scales, freely available satellite data are often constrained by spatial or radiometric resolution limitations which are unsuitable for precision agriculture, and satellite data are also subject to atmospheric conditions which may produce cloudy or unusable data (Wang et al., 2021).

Multispectral sensors mounted in crewed aircraft can be used to image a large area and spatial range, but they are typically expensive to acquire. Flying higher to capture a larger area within each frame and reduce costs may result in image data of relatively high spatial resolution, but the high cost is suitable for this project (Sofonia et al., 2019). UASs, on the other hand, have dramatically altered the ability to acquire imagery at a lower cost for precision agriculture. The utilization of UASs provides an optimized approach to farming tasks including biomass prediction, chemical spraying, field mapping, inventory counting, plant stress detection, and weed management. Additionally, UASs can fly visible, multispectral, hyperspectral and LiDAR data in different trajectories to photogrammetrically produce 3D models and maps (Hassler et al., 2019).

UASs provide a nondestructive approach to obtain spectral, structural, and texture features of agricultural areas across differing spatiotemporal scales (Jiang et al., 2019). Data from sensors mounted on UASs are suitable for high resolution, spatially explicit estimations of AGB, and they have great application potential for precision agriculture (Poley et al., 2020).

One application of utilizing UASs for precision agriculture was proposed by Gil-Docampo et al. (2020), and the authors detailed a methodology to predict herbaceous crop biomass using optical cameras and Structure from Motion (SfM) photogrammetry. They computed the difference between a digital surface model (DSM) and a bare Earth digital terrain model (DTM) to create a canopy height model (CHM). This method can be used to determine AGB by using an in-situ calibration density factor that relates crop volume and biomass, and the proposed method was intended to be generalized for the extrapolation of AGB in croplands by creating a CHM (Gil-Docampo et al., 2020). The comparison of SfM-based CHMs and LiDAR-based dense point clouds for the estimation of AGB for pasture grass has been analyzed by Walter et al. (2018). For clarification, it must be noted that DSM data can be produced from filtering the first returns of LiDAR point clouds or from photogrammetric SfM with multiple overlapping aerial images. No significant difference was observed between the AGB estimation from dense point clouds and CHMs; however, when the objective of the UAS survey is to precisely monitor crops, the dense point cloud data is preferred. If the purpose is simply to predict field biomass with little requirement of details, the SfM-based DSM data provide the values with comparable accuracy and enhanced efficiency (Walter et al., 2018). Additionally, Walter et al. (2018) identified that point cloud data derived from a UAS can accurately predict wheat volume with a strong correlation to AGB; and canopy height and harvest index values derived from the point cloud data correlated well to in-situ measurements.

Vegetation Indices (VIs) are derived from two or more spectral bands of remotely sensed imagery to assess and depict vegetation characteristics (Jensen, 2007). VIs are extremely useful for monitoring plant processes including AGB estimation, groundwater withdraw, net carbon fixation, phenology, primary productivity, and rainfall use efficiency (Glenn et al., 2008). One commonly used VI is the Normalized Difference Vegetation Index (NDVI) that is a ratio of the difference divided by the sum of the surface reflectance of red and near infrared (NIR) bands (Rouse et al. 1974, Kriegler et al. 1969). Glenn and Tabb (2018) note NDVI is used to assess green biomass based on the relatively strong absorption of red light and reflectance of NIR by chlorophyll in the leaves of healthy plants. They also noted NDVI was recognized as a measure of vegetation health because a reduction in chlorophyll production due to stressors such as disease, insect damage, lack of nutrients, salinity or drought, would result in an increase in red reflection and subsequent decrease in NDVI. Variations on NDVI soon emerged. For example, AGB estimation for rice crops in Portugal was pursued by Gerardo & de Lima (2023) using VIs derived from UAS RGB imagery. Results detailed that RGB derived indices Visible Atmospherically Resistant Index (VARI) and Triangular Greenness Index (TGI) provide a cost-effective for rice crop monitoring and management (Gerardo & de Lima., 2023).

While edible crops are of obvious importance to agriculturalists, cotton is an extremely valuable fiber crop typically used for textile purposes. Cotton breeding programs and management purposes are of vital importance, so it is critical for agriculturalists to monitor AGB and yield for precision agriculture (Siegfried 2021). In one study, VIs derived from UAS imagery were compared to cotton yield over numerous years, and NDVI produced more accurate predictions of yield than the Normalized Difference Red Edge Index (NDRE), which indicates that NDVI is more suitable for quantifying cotton yield (Siegfried 2021). As a temporal feature, NDVI was included

along with non-temporal and qualitative data in a machine learning framework to predict cotton yield, and results of an artificial neural network predicted yield with reasonable accuracy ($R^2 = 0.72$) even at 70 days after planting. Deep Learning methods have also been explored for predicting cotton AGB and yield. Li et al. (2022) employed a modified DCNN model using pixel-level segmentation to predict cotton yield in Xinjiang, China, and results detail an average error in yield estimates as low as 6.2% (Li et al., 2022). The relationship between cotton AGB and yield is imperative since the cotton fiber represents 90% of the total economic value (Singh et al., 2023). Contrary to other crops cotton has a low harvest index, the ratio of harvested biomass to yield, of approximately 15-20% (Constable & Bange, 2015). In this research, the primary focus was predicting cotton AGB, but the relationship between AGB and yield was explored.

Stone (2023) pursued a methodology to predict cotton AGB using a Random Forest regression with ArcGIS Pro for the 2018 growing season at the Ashburn Cooperator Farm in Ashburn, Georgia. A total of 12 VIs were calculated to predict cotton AGB. Pursuing the Random Forest regression within ArcGIS Pro required a specific number of training points to be extracted from biomass sample plots, in this case 6,000, which was approximately 20% of the pixels within each biomass sample plot. To pursue agricultural regionalization, the models were then scaled up across space to the field level.

This research builds upon Stone's (2023) workflow and results. The comparisons of predicted AGB based on ground-based, plant harvest measurements scaled up to plot and field levels were performed using an additional machine learning algorithm, time-series UAS imagery, and derived VIs for two different farms over two growing seasons. An additional 10 VIs, totaling 21, were calculated to observe if more VIs increased model accuracy. Random Forest and XGBoost regressions were pursued using the python library scikit-learn to compare predicted

cotton AGB across the 2018 and 2019 growing seasons at the Ashburn Cooperator Farm which relies on rainfall and the irrigated Ty Ty Cooperator Farm. Random Forest models were first detailed by Chen (2016), and these models can increase the tree-based size to have higher accuracy for training and unseen data. XGBoost models were first detailed by Tin (1995), and these models use decision trees where new trees correct the errors made by previous trees. Working with python allows for enhanced user flexibility during hyperparameter turning. With this methodology, all pixels are utilized within the biomass sample plots, which provides a more robust model for cotton AGB prediction.

OBJECTIVES

Utilizing a GeoAI framework, this study utilizes in-situ biomass data and UAS imagery to predict cotton AGB at plot and field level scales for the Ashburn and Ty Ty Cooperator Farms across the 2018 and 2019 growing seasons.

Objective 1: Reprocess UAS Imagery and Compute Weighted Biomass

The first objective of this study is to reprocess UAS imagery acquired by the USDA for two farms within the LREW watershed that represent the LTAR GACP region. The reprocessed UAS imagery acquired on multiple dates during the 2018 and 2019 growing seasons ensures correct radiometric outputs and improved computation of weighted biomass within USDA field sample plots. Original processing contained UAS flight lines to and from the launch station between battery swaps, so sections of the field can look irregularly illuminated due to changing light conditions. Additionally, for a few UAS flight dates, a radiometric calibration error existed for a single image which produced incorrect radiometric outputs. After manually identifying and removing the corrupt images within the image properties editor, the data were reprocessed using the 'Camera and Sun Irradiance' parameter, and the correct outputs were prepared to compute the

Weighted Wet and Dry Biomass, as described below, for the sample plots from both the Ashburn and Ty Ty Cooperator Farms.

Hypothesis: Reprocessed UAS imagery will enhance analysis of cotton AGB to produce higher accuracy for total end of season AGB. This will produce higher confidence in the end of season AGB predictions.

Objective 2: Predict AGB for Cotton Fields at Ashburn and Ty Ty farms

The second objective examines the use of 21 VIs to predict cotton AGB at plot and field level scales for the end of season harvest dates at the Ashburn and Ty Ty Cooperator Farms for the 2018 and 2019 growing seasons. After running a correlation matrix, the most significant VIs are identified and implemented for predicting biomass for the remainder of the 2018 and 2019 biomass dates. Models using only the raw UAS bands plus NDVI are also compared against models using the best VIs for the end of season harvest dates. The comparison of results between the two model types is significant because the raw bands plus NDVI are output during standard USDA photogrammetric processing, so it is noteworthy to observe if the effort of computing and using numerous VIs increases model accuracy. This analysis predicts cotton AGB grown in irrigated vs. non-irrigated fields (Ashburn and Ty Ty farms, respectively) during a year of relatively normal precipitation in 2018 and the 2019 drought in Georgia.

Hypothesis: Models using only the raw bands and NDVI will have adequate results for predicting cotton AGB, but the models incorporating the best of 12 VIs will increase model accuracy.

Objective 3: Analyze Cotton Growth and Development in Relation to Precipitation

The third objective compares cotton growth and development over the entirety of the two growing seasons for both study sites based on UAS-derived time-series of computed NDVI.

Precipitation trends are examined in comparison to the NDVI time-series and the predicted biomass values to examine any existing relationships among rainfall/irrigation, cotton crop health indicated by NDVI values and predicted AGB.

Hypothesis: Cotton AGB predictions will be heavily correlated with the results from the NDVI time series and the precipitation analysis. Study sites with higher NDVI and precipitation values will yield higher cotton AGB predictions.

DATA SOURCES

Remote Sensing images of the Ashburn Cooperator Farm and the Ty Ty Cooperator Farm were acquired by the USDA-ARS from a MicaSense RedEdge-3 sensor mounted to a DJI Matrice 100 (Coffin et al., 2023, Coffin et al., 2024b, *Table 2.2*). The sensor produced output images at a 9-cm spatial resolution and contained a total of five spectral bands including: blue, green, red, red-edge, and near-infrared (*Table 2.1*). The data were reprocessed as a part of Objective 1 and reflectance maps of the raw bands were output, in addition to NDVI maps for use in this research (*Figure 2.1*). Precipitation analysis of rainfall trends were conducted for the 2018 & 2019 growing

Table 2.1: The center wavelength and bandwidth of the five MicaSense multispectral sensor bands.

MicaSense Band	Wavelength of Band Center (nm)	Bandwidth (nm)
Blue	475	20
Green	560	20
Red	668	10
RedEdge	717	10
Near Infrared	840	40

seasons using public access data from STEWARDS (STEWARDS, 2025). Datasets from STEWARDS recorded total daily precipitation in mm, and the data for each study site and time frame were manually reorganized to display cumulative precipitation in mm beginning at planting and ending on October 14, which is the latest defoliation date across study sites and years. As cotton is defoliated, the plant essentially dies, so no precipitation data should be considered for unproductive plants. Two of the four study sites and years did not contain defoliation data, so the latest defoliation date of October 14th was used as the precipitation cutoff for all study sites and years. Researchers at the USDA-ARS SEWRL collected field measurements of harvested cotton throughout the growing season in 2018 and 2019. Measurements included plant height, density, and water content, along with wet and dry biomass from defined plot boundaries distributed within the farm fields and were published in the National Agricultural Library Ag Data Commons repository (<https://agdatacommons.nal.usda.gov/>; Coffin et al., 2023; Coffin et al., 2024b). To acquire the biomass measurements, researchers followed a systematic approach that allowed them to harvest the crop and measure biomass while also preserving the integrity of the surrounding plants (*Figure 2.2*). Cotton plants were sampled from within a 30-m x 10-crop row area, which are hereafter referred to as the sample plots. In 2019, a total of 9 biomass sample plots were used at the Ashburn Cooperator farm, and 6 biomass plots were used in 2018. For 2019, the Ty Ty Cooperator Farm had 3 biomass sample plots, but for 2018, only 2 biomass sample plots were used. A diagram of the collection protocol for 1 x 1 m harvest areas, hereafter called sub-plots, within the 30-m x 10-row sample plots is shown in *Figure 2.2* (Coffin et al., 2024b). Management practices were considered for this study. The planting date was May 25th for the ACF in 2018, May 24th for the TCF in 2018, May 16th for the ACF in 2019, and May 15th for the TCF in 2019. The harvest date was November 23rd for the ACF in 2018, October 30th for the TCF in 2018, October

28th for the ACF in 2019, and November 4th for the TCF in 2019. The TCF was irrigated when the agriculturalist deemed it necessary. The ACF is unirrigated, so no additional water was provided.

Table 2.2: UAS flight dates used in this study which correlate with in-situ biomass collections.

UAS Flight Date (MM/DD/YYYY)	Ashburn Cooperator Farm	Ty Ty Cooperator Farm
10/24/2019	✓	
10/07/2019		✓
09/19/2019		✓
09/18/2019	✓	
08/26/2019		✓
08/23/2019	✓	
07/31/2019	✓	
07/30/2019		✓
07/08/2019		✓
07/05/2019	✓	
06/13/2019	✓	
10/17/2018	✓	
10/15/2018		✓
09/21/2018	✓	
09/20/2018		✓
08/31/2018	✓	
08/30/2018		✓
08/07/2018		✓
08/03/2018	✓	
07/13/2018		✓
07/11/2018	✓	
06/19/2018		✓
06/18/2018	✓	

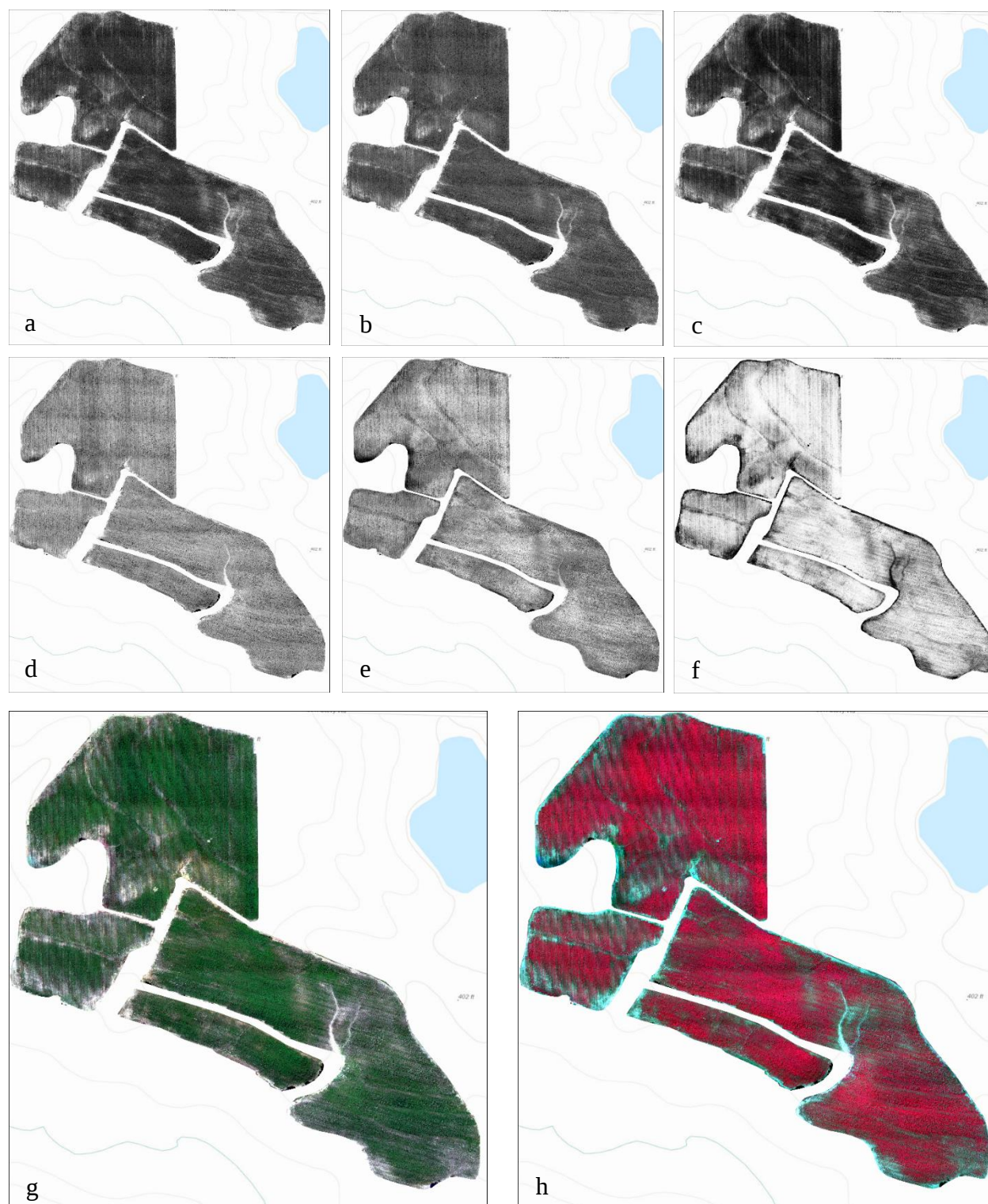


Figure 2.1: The Ashburn Cooperator Farm imagery with bands, indices, and color composite with the blue band (a), green band, (b), red band (c), rededge band, (d), near-infrared band (e), NDVI (f), RGB composite (g), and Color Infrared composite (h).

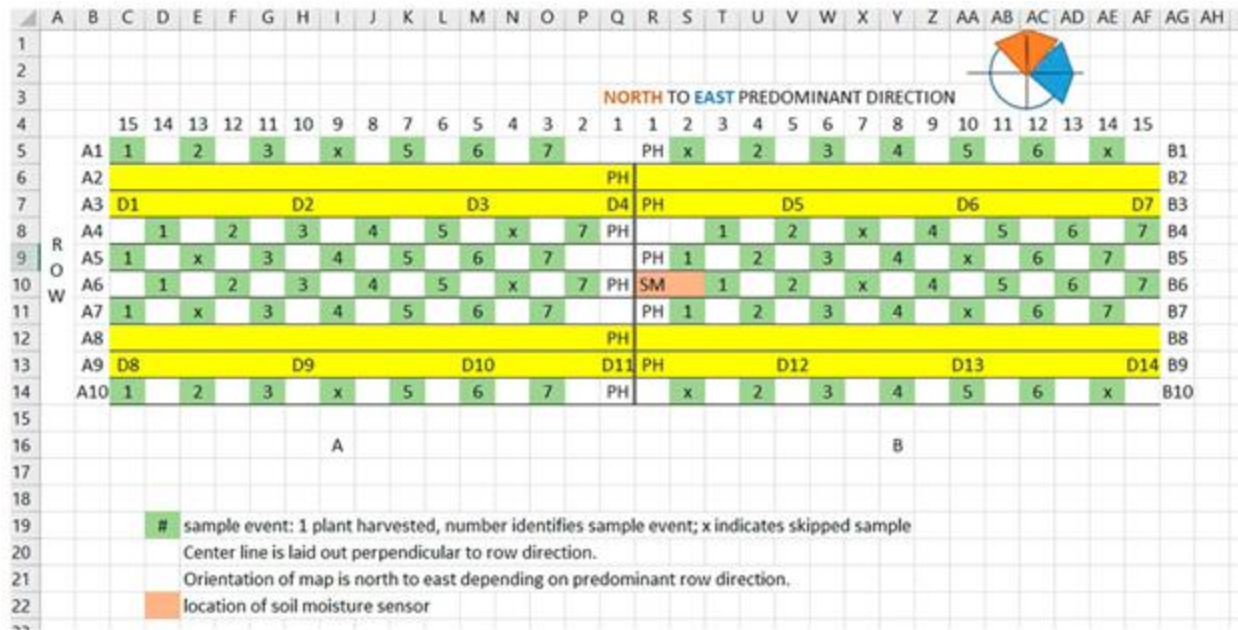


Figure 2.2: Sample site map showing data collection plan. Numbers across the top indicate meter length subsections (15-1, 1-15) of the sample site. Legend: A, left side of sample site; B, right side of sample site; SM, soil moisture measurement location (orange); PH, plant height measurement location; D1 through D14 were internal flag locations; yellow colored rows indicate areas where no destructive sampling occurred; green numbered blocks indicate biomass sampling locations by sortie number (x indicates no sample) (Coffin et al., 2024b, figure used with permission).

Chapter 3

METHODS

Methods for Objective 1: Reprocess UAS Imagery and Compute Weighted Biomass

UAS imagery of the Ashburn and Ty Ty Cooperator Farms was originally processed by SEWRL personnel using Pix4D Mapper by Pix4D S.A. to generate reflectance orthomosaic mosaics of the study areas during 2018 and 2019. However, due to sensor errors and the presence of nonparallel flight lines, the data had to be reprocessed in this research to ensure correct radiometric outputs. Images acquired within the departure and return flight lines from the UAS launch location were manually excluded to ensure changing light conditions did not disrupt the appropriate reflectance values across the study sites. Additionally, there were few flight dates where an unexplained radiometric calibration error existed for a single image which produced incorrect orthomosaics, so a handful of flight dates contained corrupted images which would produce irregular reflectance values if not excluded from the processing (*Figure 3.1*). Upon removing the problematic images, the index and reflectance maps were reprocessed. The process for generating the index and reflectance map uses Pix4D to perform SfM, create a DSM and photogrammetrically orthorectify each of the images per flight. The images are then stitched together to create a seamless multispectral orthomosaic of 9-cm spatial resolution. Raw UAS image bands were initially clipped to the boundary of the crop fields, and 21 VIs were then calculated at field level. To further examine vegetation characteristics including cotton growth and development and precipitation responses, 9 more VIs were utilized in addition to Stone's (2023) original VIs derived from the UAS imagery. In total, 21 VIs were used to explore any advantages

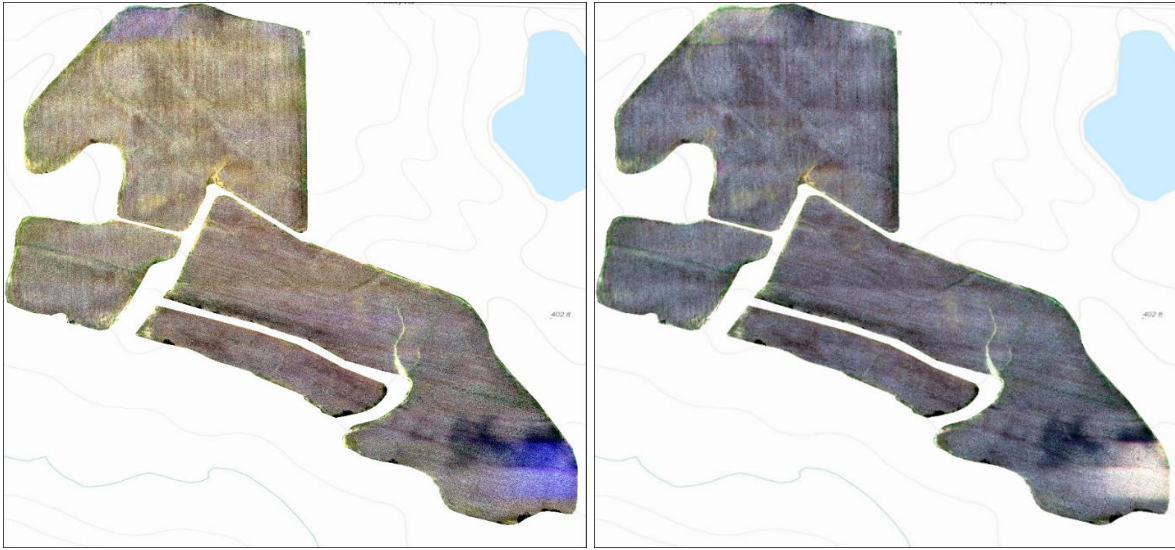


Figure 3.1: The UAS orthomosaic of the Ashburn Cooperator Farm before reprocessing (a). Note the blue discoloration in the corners of the image. The UAS orthomosaic after reprocessing (a) has corrected radiometric calibration.

of using additional VIs that are variations of NDVI (e.g., TDVI), use the RedEdge band (e.g., RRI1 and RRI2) and use only visible bands (e.g., GDVI) (*Table 3.1*). Raw bands and VIs were clipped to the boundary of each 30-m x 10-row sample plot. The raw bands and VIs were then resampled using a bilinear interpolation method to ensure all data possess identical dimensions which is critical for running the machine learning models. The methodology by Stone (2023) was replicated in this study where NDVI values within each 30-m x 10-row sample plot were averaged, and on a per pixel basis, NDVI was divided by the average NDVI value for the plot to produce the Weighted NDVI Layer which depicts the distance of each pixel's NDVI value from the mean NDVI within the given plot. The Weighted NDVI Layer was then multiplied by the biomass per pixel variable which is computed by multiplying the average wet weight of harvested biomass for the plot in grams, the plant density per square meter, and the unit conversion coefficient. The unit conversion coefficient is calculated by dividing the area of one UAS pixel by the area of one meter squared, which results in $(9\text{cm} \times 9\text{cm} / 100\text{cm} \times 100\text{cm}) = (81\text{cm}^2 / 10,000\text{cm}^2) = 0.0081$. This coefficient parses the average wet weight biomass times the plant density of the plot, to each individual UAS

Table 3.1: The 21 Vegetation Indices used for initially predicting cotton AGB plus NDVI, along with the citation for the source of the index equation. The indices which are bold and underlined are the new indices which Stone (2023) did not pursue in his methodology.

Vegetation Index Name	Abbr.	Equation	Reference
Chlorophyll Green Index	CGI	$\left(\frac{NIR}{GREEN}\right) - 1$	Gitelson and Merzlyak, 1994
<u>Chlorophyll Index RedEdge</u>	<u>CRededge</u>	$\frac{(NIR)}{(REDEGE)} - 1$	Gitelson et al., 2003
Chlorophyll Vegetation Index	CVI	$\frac{(NIR * RED)}{(GREEN^2)}$	Vincini et al., 2008
<u>Enhanced Vegetation Index</u>	<u>EVI</u>	$2.5 \frac{(NIR - RED)}{(NIR + 6 * RED - 7.5 * BLUE) + 1}$	Huete et al., 2002
Enhanced Vegetation Index (2-band)	EV12	$2.5 * \frac{(NIR - RED)}{(NIR + 2.4 * RED + 1)}$	Jiang et al., 2008
<u>Green Difference Vegetation Index</u>	<u>GDVI</u>	$NIR - GREEN$	Tucker, 1979
Green Leaf Index	GLI	$\frac{(2.0 * GREEN - RED - BLUE)}{(2.0 * GREEN + RED + BLUE)}$	Gobron et al. 2000
Green Normalized Difference Vegetation Index	GNDVI	$\frac{(NIR - GREEN)}{(NIR + GREEN)}$	Gitelson et al., 1996
Modified Soil Adjusted Vegetation Index	MSAVI	$\frac{(2 * NIR + 1 - (((2 * NIR + 1)^2 - 8 * (NIR - RED))^{0.5}))}{2}$	Qi et al., 1994
Normalized Difference Red Edge	NDRE	$\frac{(NIR - REDEGE)}{(NIR + REDEGE)}$	Gitelson and Merzlyak, 1994
Normalized Difference Vegetation Index	NDVI	$\frac{(NIR - RED)}{(NIR + RED)}$	Rouse, 1974
Normalized Difference Water Index	NDWI	$\frac{(GREEN - NIR)}{(GREEN + NIR)}$	McFeeters, 1996
Optimized Soil Adjusted Vegetation Index	OSAVI	$\frac{(NIR - RED)}{(NIR + RED + 0.16)}$	Rondeaux et al., 1996
<u>Normalized Difference Red/Green Redness Index</u>	<u>RI</u>	$\frac{(RED - GREEN)}{(RED + GREEN)}$	Escadafal and Huete, 1991
<u>RedEdge Ratio Index 1</u>	<u>RRI1</u>	$\frac{NIR}{REDEGE}$	Ehammer et al., 2010
<u>RedEdge Ratio Index 2</u>	<u>RRI2</u>	$\frac{REDEGE}{RED}$	Ehammer et al., 2010
Soil Adjusted Vegetation Index	SAVI	$(1 + L) * \frac{(NIR - RED)}{(NIR + RED + L)}$	Huete, 1988
<u>Simplified Canopy Chlorophyll Content Index</u>	<u>SCCCI</u>	$\frac{NDRE}{NDVI}$	Raper and Varco, 2015
<u>Simple Ratio</u>	<u>SR</u>	$\frac{NIR}{RED}$	Jordan, 1969
<u>Transformed Difference Vegetation Index</u>	<u>TDVI</u>	$1.5 * \left(\frac{NIR - RED}{\sqrt{NIR^2 + RED + 0.5}}\right)$	Bannari et al., 2002
<u>Triangular Vegetation Index</u>	<u>TVI</u>	$0.5(120(NIR - GREEN) - 200(RED - GREEN))$	Broge and Leblanc., 2001
Visible Atmospherically Resistant Index	VARI	$\frac{(GREEN - RED)}{(GREEN + RED - BLUE)}$	Gitelson et al., 2004

pixel. The AGB protocols are extensively documented by the LTAR (Wilke et al., 2024). Additionally, note the weights in grams are the sum of all 10 plants harvested for measuring biomass, so dividing by an additional factor of 10 gives the average biomass weight in grams for each specific harvest date. Inputting the division by a factor of 10 to the unit conversion coefficient provides a simple multiplication of three numbers to acquire the biomass in grams per pixel, namely, the average fresh weight, referred from now on as the average wet biomass weight, the plant density, and the unit conversion coefficient of 0.00081. The biomass per pixel variable is multiplied by the Weighted NDVI Layer which produces the Weighted Biomass Layer, and this layer details the calculated biomass on a per pixel basis within each sample plot for the corresponding dates (*Figure 3.2 and Figure 3.3*).

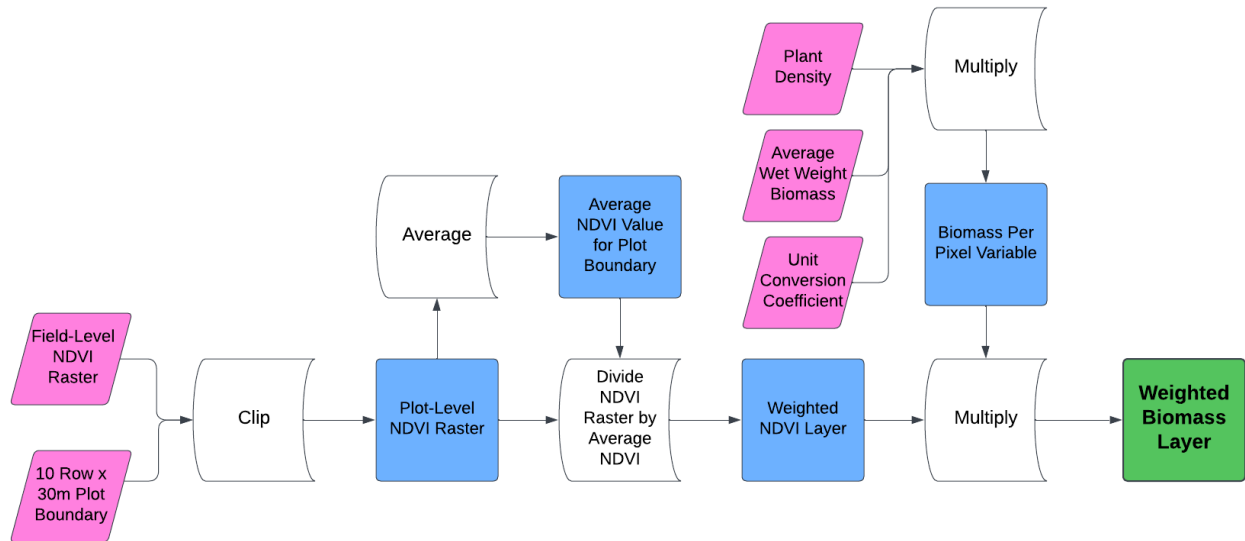


Figure 3.2: Workflow for producing a Weighted Biomass Layer for a singular biomass layer.

Methods for Objective 2: Predict AGB for Cotton Fields at the Ashburn and Ty Ty Farms

After acquiring the Weighted Biomass Layer for each corresponding plot and date, as the ground-measured AGB, these data are used in machine learning models to predict cotton AGB first at the plot level and then at the field level for the Ashburn and Ty Ty farms for 2018 and 2019.

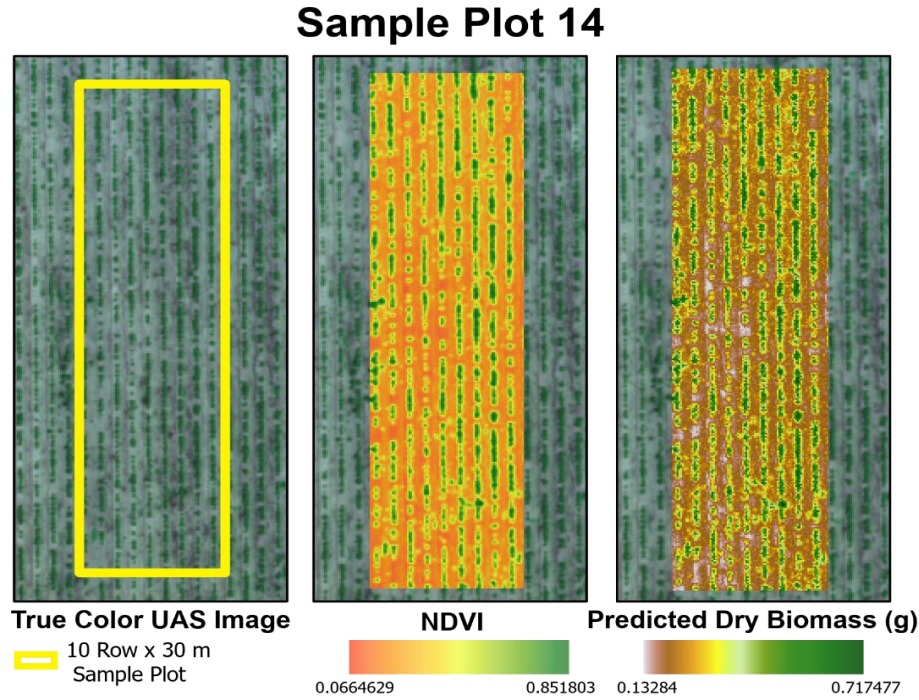


Figure 3.3: True color, NDVI, and predicted dry biomass images of sample plot 14 at the Ashburn Cooperator Farm (Stone, 2023).

Predicted AGB at the plot level are optimized by comparing predictions made with the five raw bands and NDVI vs. predictions using all or some of the additional 21 VIs computed from the UAS imagery. Explanatory variables included the raw bands and the selected vegetation indices which were used for biomass prediction. Comparatively, predicted cotton AGB is the response variable. Once the best models are determined, optimized input variables were reshaped into a 2D array using the NumPy reshape function which gives a new shape to the input array without changing the data. The reshaped data were input to the machine learning algorithms to predict AGB for the entire field. Specific methods to first predict AGB at the plot level are depicted in the workflow shown in *Figure 3.4*. This workflow was repeated with the substitution of the XGBoost machine learning algorithm for the Random Forest Regressor. The train/test split for this project was 80/20 respectively. The ground-measured AGB data per UAS pixel for the sample plots were imported into python along with the field-level UAS surface reflectance for the five UAS bands

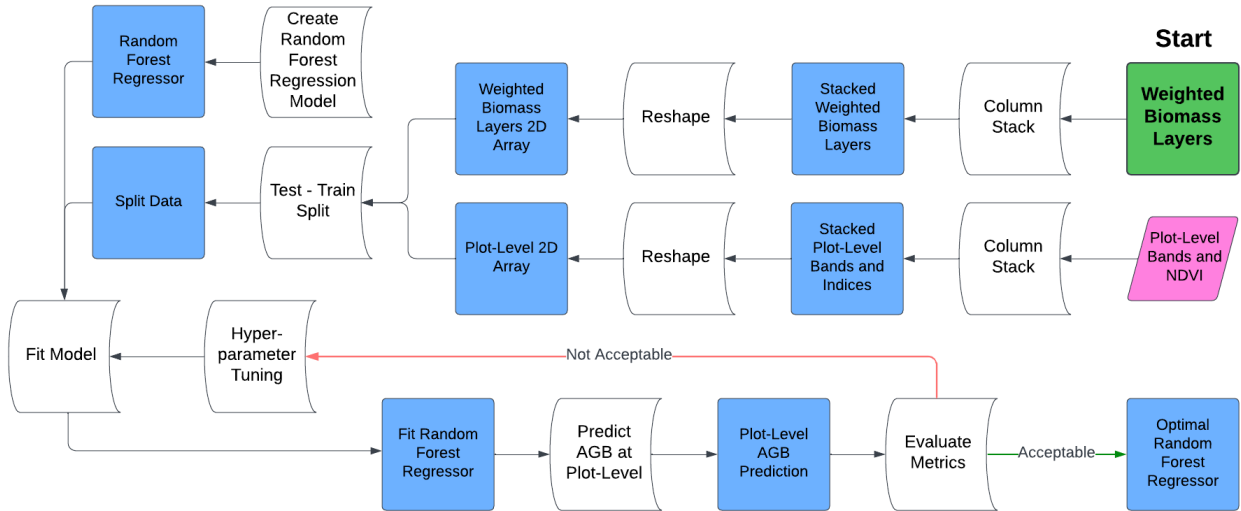


Figure 3.4: Workflow, reading from right to left, for predicting cotton AGB at plot-level.

and 22 computed VIs. Using scikit-learn, a Python machine learning library, AGB, was first predicted at the 30-m x 10-row plot level using Random Forest and XGBoost models. Detailed in Table 3.1 above, 21 VIs were utilized to predict AGB of cotton. These 21 VIs were derived from each individual UAS flight date, and the additional 10 were selected because they are well-known VIs utilized for vegetation analysis, many of which are slight variations of more common indices. After achieving appropriate results, the AGB was predicted for the end-of-season dates for both study sites in 2018 and 2019.

Within the Random Forest and XGBoost models, a test-train split of 80% / 20% was applied. All other model parameters were left as default for further exploration of model optimization. A correlation matrix was produced to determine the most significant, i.e., uncorrelated, indices. and these were used to predict AGB for the remainder of the biomass UAS

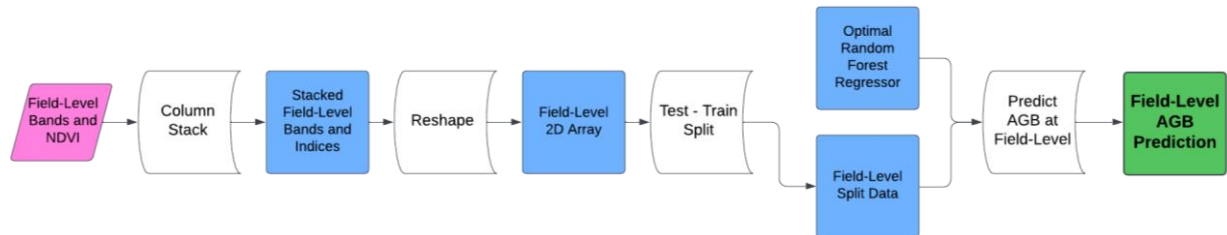


Figure 3.5: Workflow for predicting cotton AGB at field-level.

flight dates. A root mean squared error metric and mean absolute error metric was performed to evaluate model performance and error. Optimal models were then extended from the plot to the field levels using field-level bands and indices to predict AGB of cotton in fields of irrigated vs. non-irrigated farms and 2018 vs. 2019 trends in AGB and natural rainfall (*Figure 3.5*).

Methods for Objective 3: Analyze Cotton Growth and Development in Relation to Precipitation

Precipitation datasets were acquired from public access data available from STEWARDS (STEWARDS, 2025). Precipitation analysis to compare rainfall trends was conducted by calculating and comparing the cumulative precipitation throughout the 2018 and 2019 growing seasons for both study sites. The time series of rainfall from 2018 with a typical rainfall trend in terms of amount, timing and duration, was compared to rainfall trends during the drought year of 2019. Trends in cotton vegetation health or greenness were compared to rainfall trends by computing NDVI from multiple dates of UAS imagery acquired for the Ashburn and Ty Ty farm fields. NDVI analysis was conducted by first clipping the multi-date output NDVI raster datasets to the extent of the corresponding fields. Field-level NDVI was calculated by averaging all values to produce a single median NDVI value, used for removing the skew from outliers, for the corresponding UAS flight date. NDVI values for each flight date were plotted over time to analyze the cotton growth and development in relation to precipitation.

Chapter 4

RESULTS

Results for Objective 1: *Reprocess UAS Imagery and Compute Weighted Biomass*

Multispectral UAS datasets across the 2018 & 2019 growing seasons at the Ashburn and Ty Ty Cooperator Farms were reprocessed due to issues with the multispectral sensor and lighting conditions. Thermal and visible imagery datasets were not reprocessed during this study. A total of 55 flight dates were reprocessed totaling in excess of 1.35 TBs of data. Only 23 of the UAS flight dates were utilized for the biomass prediction aspect of this research because they correspond with the dates of in-situ data collection of AGB (see *Table 2.2*). The NDVI and precipitation analysis required 41 of the UAS datasets. All reprocessed datasets had the nonparallel flight lines where the UAS departed from and returned to the launch site removed to ensure changing light conditions did not produce radiometrically incorrect orthomosaics. The remaining reprocessed datasets were from different locations not analyzed in this study.

Additionally, 4 datasets contained corrupt data, where out of the thousands of images used for photogrammetric processing, a single image lacked radiometric correction, and no sun irradiance information was provided (*Figure 4.1*). Within Pix4D Mapper, the processing software used to produce the orthomosaics, the corrupt images had to be manually removed within the Image Properties Editor before the Correction Type option ‘Camera and Sun Irradiance’ could be selected within the Index Calculator Pane (*Figure 4.1*). Located entirely at the Ashburn Cooperator Farm, the datasets with corrupt images are entitled ‘L_2018020’ (11 July 2018), ‘L_2018025’ (3 August 2018), ‘L_2018038’ (17 October 2018), and ‘L_2019008’ (30 April 2019).

Upon reprocessing the imagery, output bands were then able to produce field level VIs which were then clipped down to sample plot level for the computation of the Weighted Biomass Layers. A total of 234 Weighted Biomass Layers were computed for this research with half of the layers representing wet biomass and the remaining half representing dry biomass.

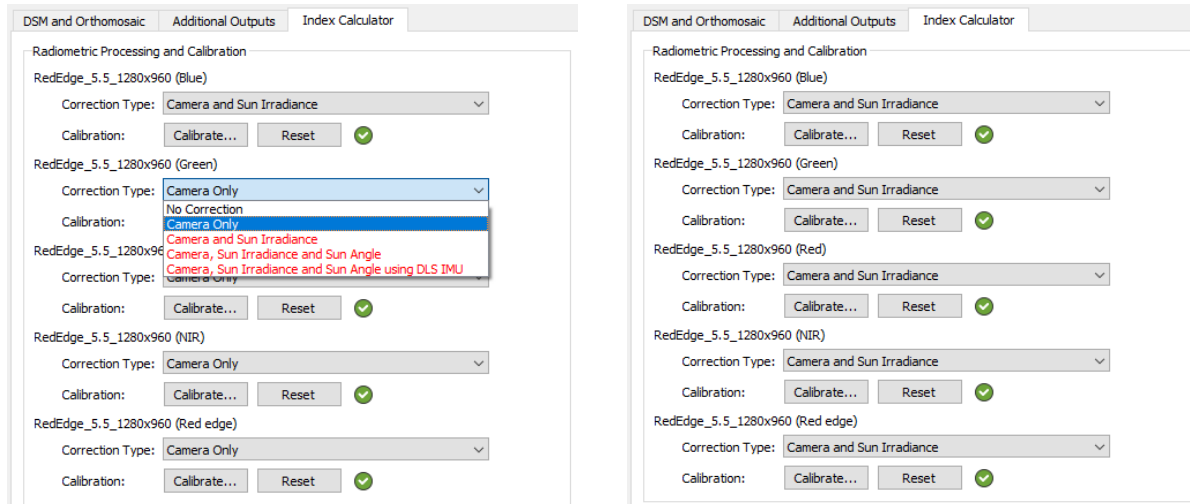


Figure 4.1: The correction type for the raw bands before (left) and after (right) removing corrupt images.

Results for Objective 2: Predict AGB for Cotton Fields at the Ashburn and Ty Ty Farms

A total of 21 Vegetation Indices (VIs) were computed from the UAS multispectral imagery using band ratios and formulas described in Table 3.1 as potential predictors of AGB. Collinearity analysis was performed to assess associations among the predictor variables and remove redundant variables from the model. In this way, only the most important VIs contributing to predicted AGB were identified. Collinearity matrices were produced for 21 VIs, wet and dry biomass and wet and dry biomass per pixel for each year of UAS imagery, i.e., 2018 and 2019, for both the ACF and TCF farms (Appendix A).

Variables listed on the X axis can be paired with variables listed on the Y axis and the color-coded value in the intersecting cell indicates the strength of the correlation. A bright red cell with values approaching 1.00 indicates a pair of variables with strong positive collinearity, i.e., as

measurements of the predictor variable increase measurements of the paired variable also increase. Bright blue cells with collinearity values approaching -1.00 conversely indicate variables that are negatively correlated. Ramped colors and values ranging between 1.00 and -1.00 indicate varying degrees of collinearity.

Figure 4.2 depicts the collinearity matrix for the ACF in 2019 and represents the findings from all 4 matrices included in Appendix A. Results indicate the VIs were generally highly correlated and the 21 original VIs could be reduced to 5 optimal VIs: Normalized Difference Water Index (NDWI), Optimized Soil Adjusted Vegetation Index (OSAVI), RedEdge Ratio Index 2 (RRI2), Triangular Vegetation Index (TV1), and Visible Atmospherically Resistant Index (VARI). The Southeast Watershed Research Laboratory, USDA-ARS harvested, dried and weighed cotton in the ACF. A separate dataset details the end of season, ground-sampled dry AGB values for the ACF in 2018 with a range from 2,500-5,500 kg/ha for seed cotton and about 5,000 – 16,000 kg/ha for total biomass (Dr. Tim Strickland, personal communication, June 1, 2023). Since it is known that the end of season dry biomass range is within 5,000 – 16,000 kg/ha, the raw bands and VIs were used to first predict cotton AGB for the end of season biomass in kg/ha at the ACF in 2018. With this known range, the models can be assessed to determine if the models using VIs (*Table 4.1*) or Raw Bands (*Table 4.2*) have higher performance to ensure the models are accurately predicting to field level scales. Validating the predicted end of season biomass within the range of measured AGB in kg/ha for the ACF in 2018 is critical to assess model accuracy. It was determined that although model metrics from the raw bands are comparable to the model metrics from the VIs, the range in predicted AGB in kg/ha from Random Forest to XGBoost for the models using the raw bands are substantially lower. To visually show the end of season biomass predictions, the

results are graphically displayed showing the models using raw bands (*Figure 4.3*) versus the models using VIs (*Figure 4.4*).

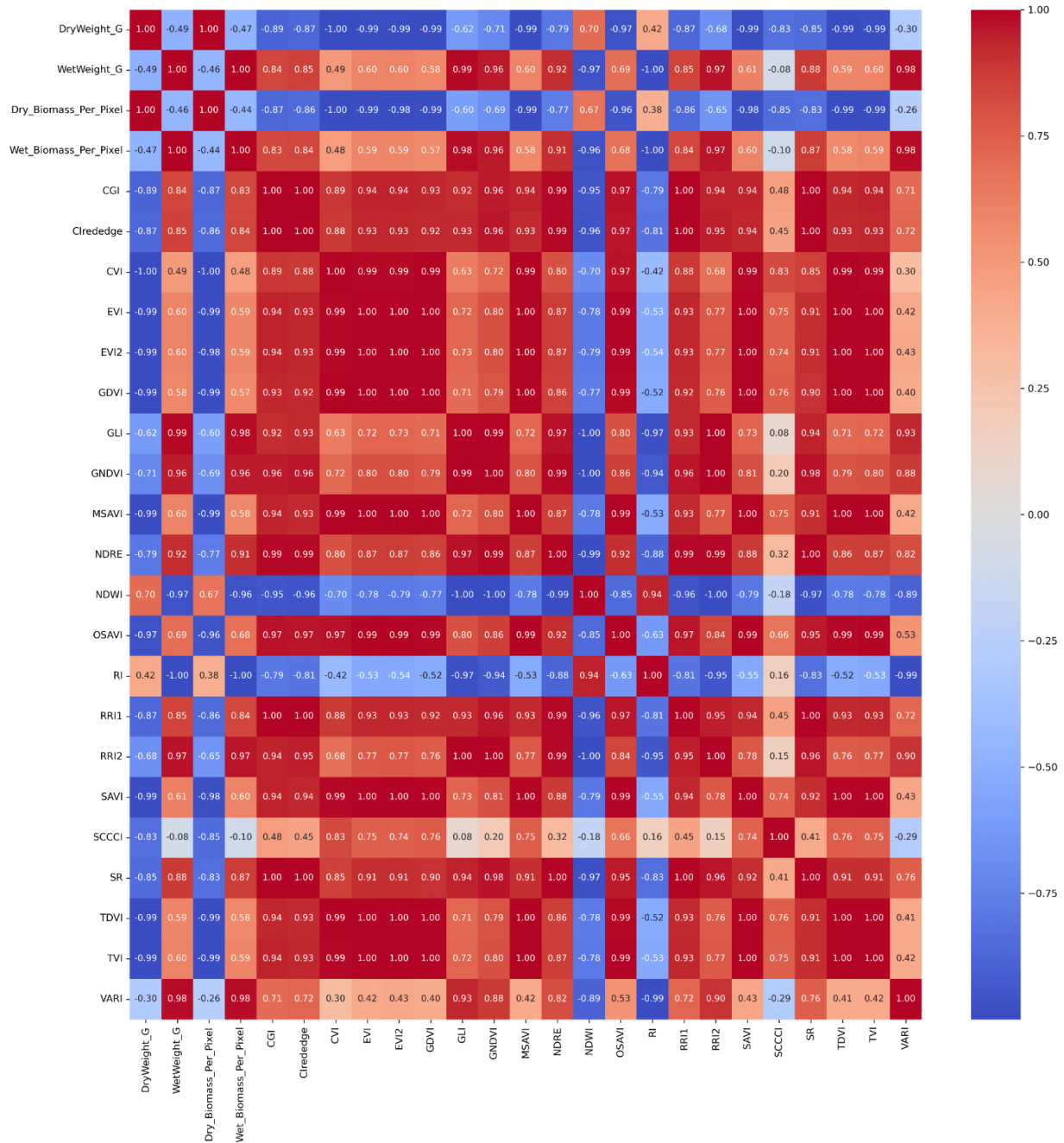


Figure 4.2: Correlation matrix of the 21 VIs used for predicting cotton AGB at the ACF for 2019.

Table 4.1: Model metrics and predicted total end of season above ground biomass (AGB) using 5 optimal VIs and two Machine Learning algorithms. Model metrics include accuracy of training and testing data, root mean square error (RMSE) and mean absolute error (MAE).

		Predicted AGB Using Vegetation Indices									
		Random Forest					XGBoost				
		Predicted AGB (kg/ha)	Train	Test	RMSE	MAE	Predicted AGB (kg/ha)	Train	Test	RMSE	MAE
2018 ACF	Wet Biomass	30,161	0.9894	0.9239	1.4882	0.9204	44,477	0.9395	0.9255	1.4721	0.9424
	Dry Biomass	16,938	0.9935	0.9538	0.5922	0.3822	23,305	0.9625	0.9547	0.5862	0.3892
2018 TCF	Wet Biomass	45,567	0.9988	0.9917	0.8146	0.4401	55,706	0.9934	0.9917	0.8122	0.4577
	Dry Biomass	20,398	0.9993	0.9949	0.2889	0.1568	26,515	0.9959	0.9949	0.2869	0.1620
2019 ACF	Wet Biomass	36,315	0.9836	0.8835	2.4603	1.1809	47,834	0.8969	0.8809	2.4879	1.2287
	Dry Biomass	20,247	0.9867	0.9056	1.2654	0.6985	25,866	0.9160	0.9040	1.2764	0.7181
2019 TCF	Wet Biomass	26,384	0.9989	0.9925	0.7771	0.2406	40,349	0.9814	0.9788	1.3118	0.5950
	Dry Biomass	14,203	0.9990	0.9933	0.3942	0.1213	22,410	0.9833	0.9810	0.6644	0.2970

Table 4.2: Model metrics and predicted total end of season above ground biomass (AGB) using Raw Bands and two Machine Learning algorithms. Model metrics include accuracy of training and testing data, root mean square error (RMSE) and mean absolute error (MAE).

		Predicted AGB Using Raw Bands									
		Random Forest					XGBoost				
		Predicted AGB (kg/ha)	Train	Test	RMSE	MAE	Predicted AGB (kg/ha)	Train	Test	RMSE	MAE
2018 ACF	Wet Biomass	13,976	0.9892	0.9221	1.5160	0.9481	14,062	0.9209	0.9132	1.6003	1.0146
	Dry Biomass	7,302	0.9934	0.9528	0.6040	0.3937	7,341	0.9499	0.9458	0.6474	0.4193
2018 TCF	Wet Biomass	20,794	0.9999	0.9999	0.0598	0.0251	20,712	0.9986	0.9975	0.4347	0.0733
	Dry Biomass	9,559	0.9999	0.9999	0.0173	0.0093	9,528	0.9985	0.9974	0.2007	0.0263
2019 ACF	Wet Biomass	18,336	0.9827	0.8772	2.5338	1.2459	18,182	0.8831	0.8722	2.5844	1.3051
	Dry Biomass	10,689	0.9859	0.9004	1.3030	0.7281	10,598	0.9038	0.8946	1.3406	0.7631
2019 TCF	Wet Biomass	26,424	0.9987	0.9914	0.8320	0.4292	26,315	0.9904	0.9895	0.9232	0.4429
	Dry Biomass	14,146	0.9994	0.9961	0.3015	0.1559	14,167	0.9942	0.9938	0.3781	0.1647

Histograms depicting the area per pixel value and maps depicting the field-level biomass on a per pixel basis are displayed for the following end of season models: 2018 ACF Wet Biomass Random Forest model (*Figure 4.5, Figure 4.6*), 2018 ACF Wet Biomass XGBoost model (*Figure 4.7, Figure 4.8*), 2018 ACF Dry Biomass Random Forest model (*Figure 4.9, Figure 4.10*), 2018 ACF Dry Biomass XGBoost model (*Figure 4.11, Figure 4.12*), 2018 TCF Dry Biomass Random Forest model (*Figure 4.13, Figure 4.14*), 2018 TCF Dry Biomass XGBoost model (*Figure 4.15, Figure 4.16*), 2019 ACF Dry Biomass Random Forest model (*Figure 4.17, Figure 4.18*), 2019 ACF Dry Biomass XGBoost model (*Figure 4.19, Figure 4.20*), 2019 TCF Dry Biomass Random Forest model (*Figure 4.21, Figure 4.22*), 2019 TCF Dry Biomass XGBoost model (*Figure 4.23, Figure 4.24*). Figure axes are not standardized, and this was done to ensure maximum vertical distribution for each graph. It may appear that some values are zero in the histograms, but no matter how small the value on the Y axis, each value range on the X axis has an associated value for the Y axis. Enlarged predicted biomass maps are included in Appendix B.

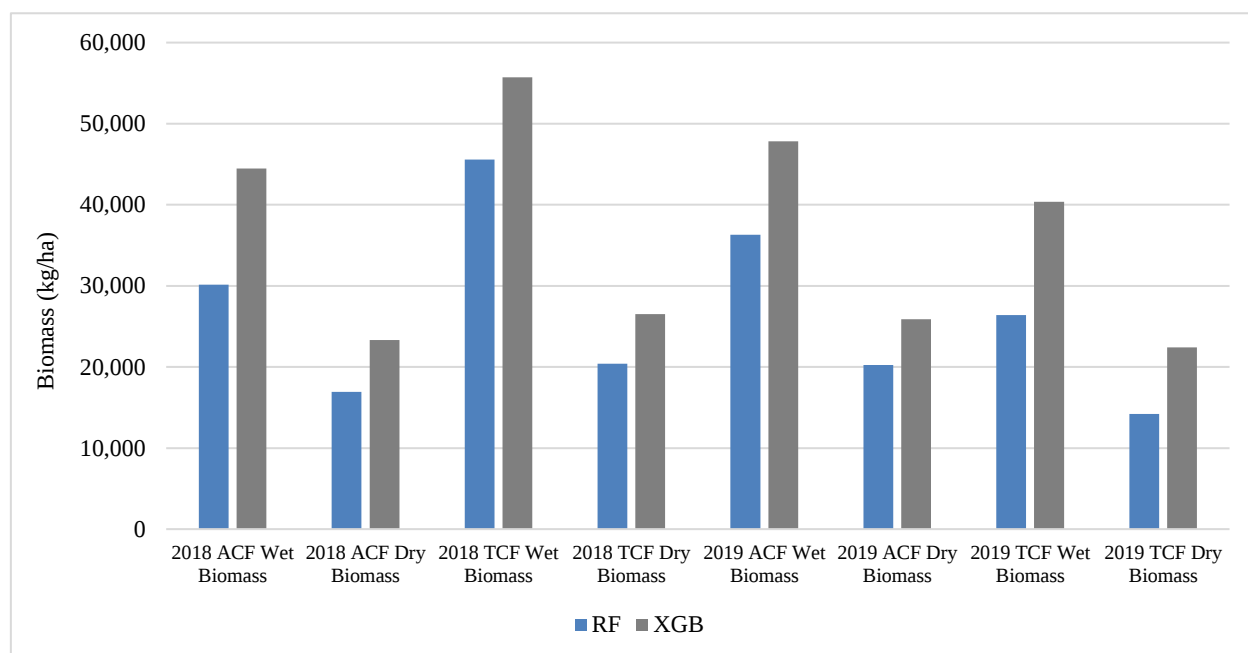


Figure 4.3: Total end of season predicted biomass as kg/ha using the 5 VIs.

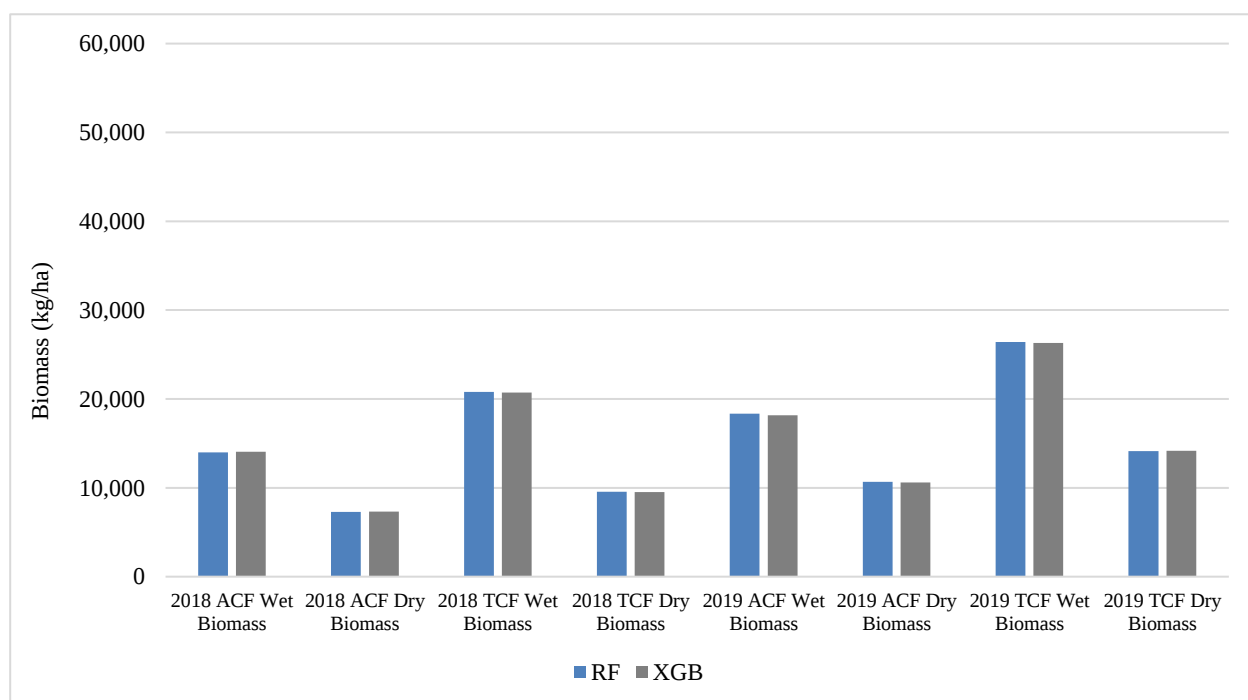


Figure 4.4: Total end of season predicted biomass as kg/ha using the Raw Bands.

Figure 4.5: Histogram displaying the area per pixel value for the 2018 ACF End of Season Random Field-Level Forest Predicted Wet Biomass model.

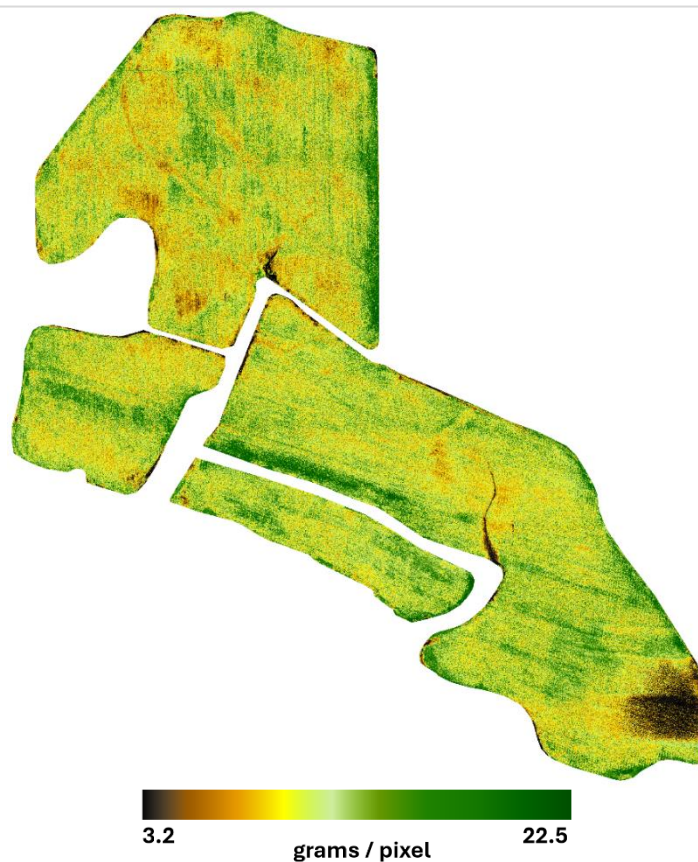
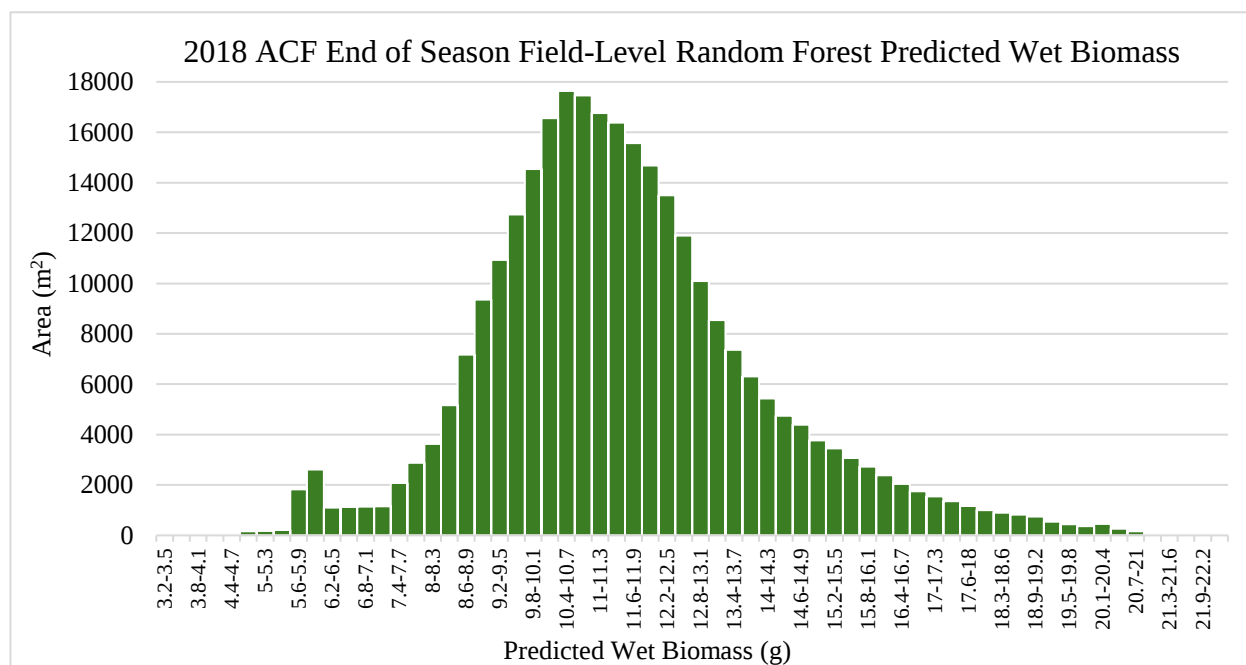


Figure 4.6: Map of field-level predicted biomass for the 2018 ACF End of Season Random Forest Field-Level Forest Predicted Wet Biomass model.

Figure 4.7: Histogram displaying the area per pixel value for the 2018 ACF End of Season Field-Level XGBoost Predicted Wet Biomass model.

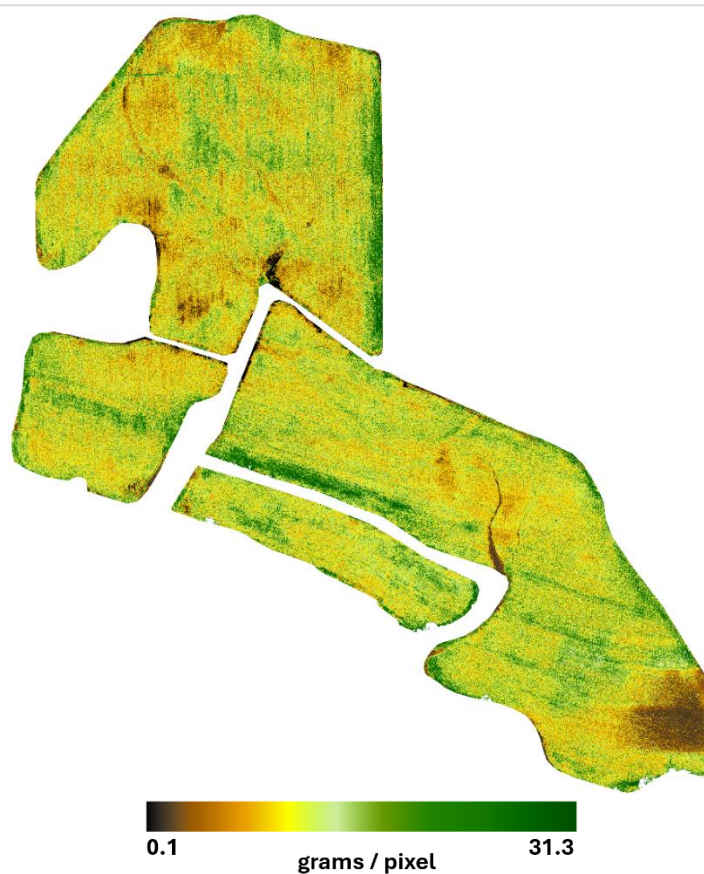
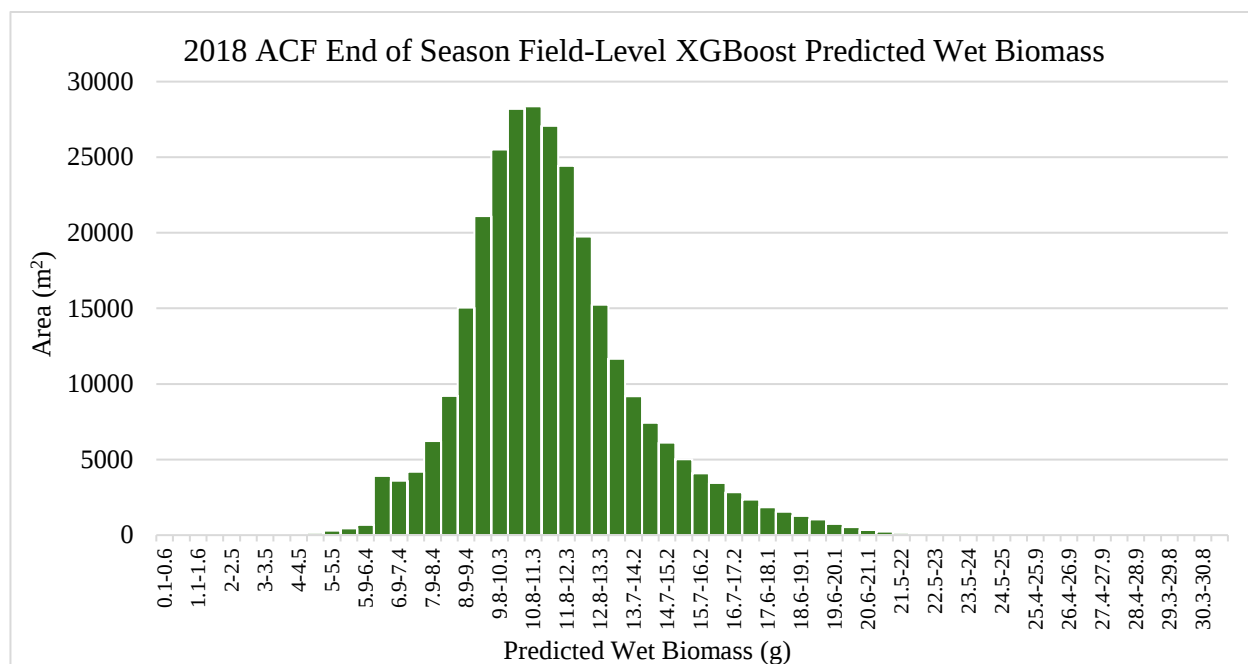


Figure 4.8: Map of field-level predicted biomass for the 2018 ACF End of Season Field-Level XGBoost Predicted Wet Biomass model.

Figure 4.9: Histogram displaying the area per pixel value for the 2018 ACF End of Season Field-Level Random Forest Predicted Dry Biomass model.

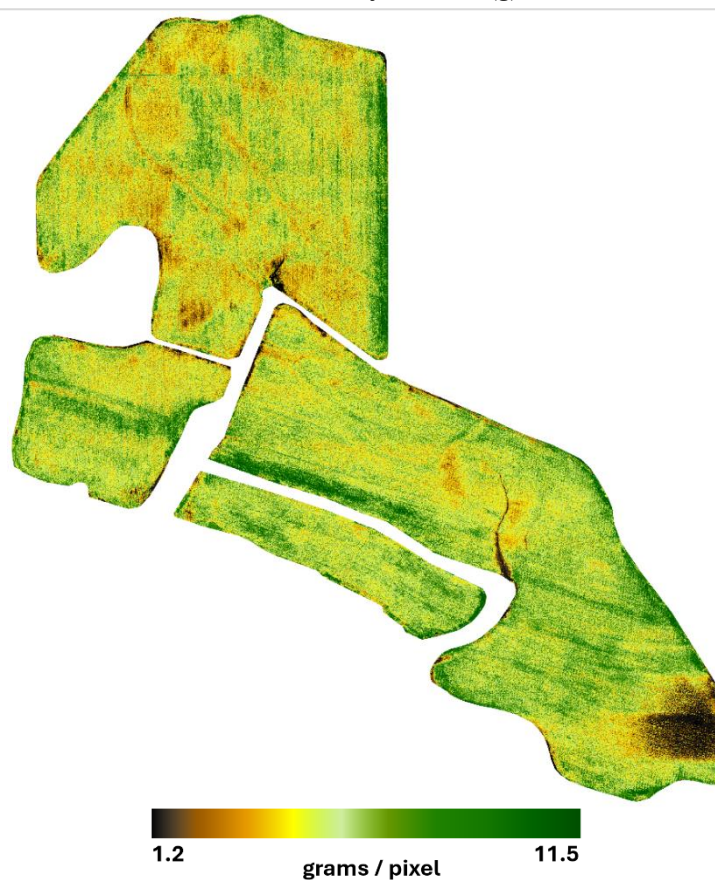
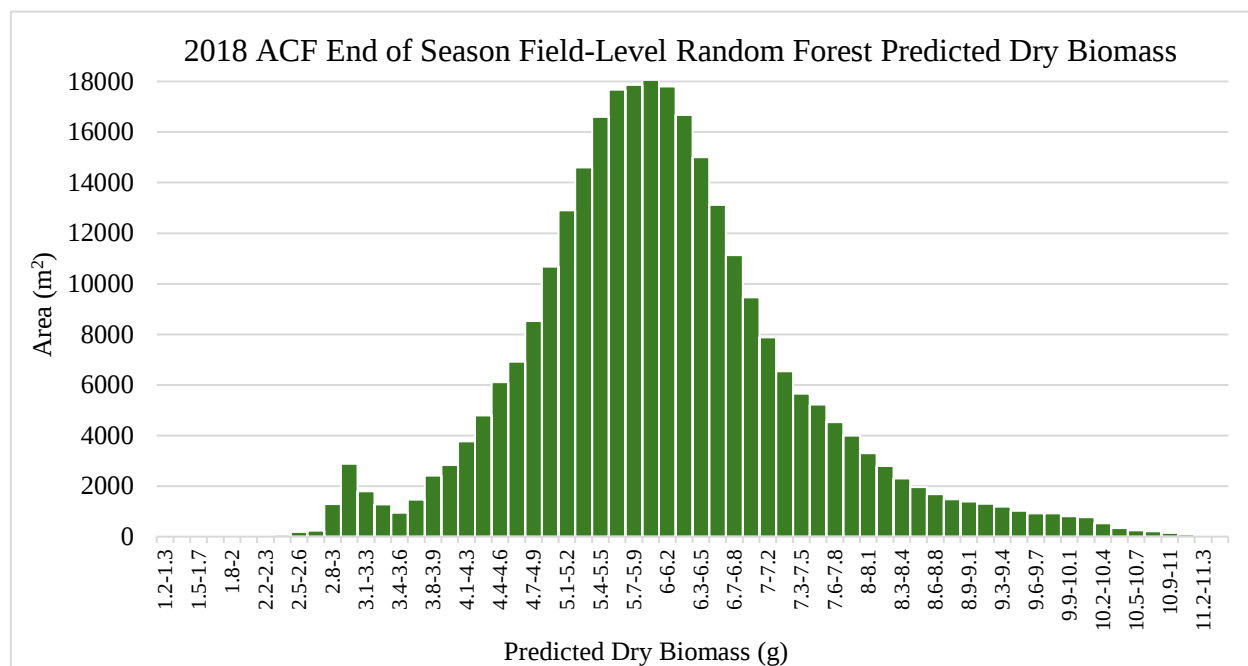


Figure 4.10: Map of field-level predicted biomass for 2018 ACF End of Season Field-Level Random Forest Predicted Dry Biomass model.

Figure 4.11: Histogram displaying the area per pixel value for the 2018 ACF End of Season Field-Level XGBoost Predicted Dry Biomass model.

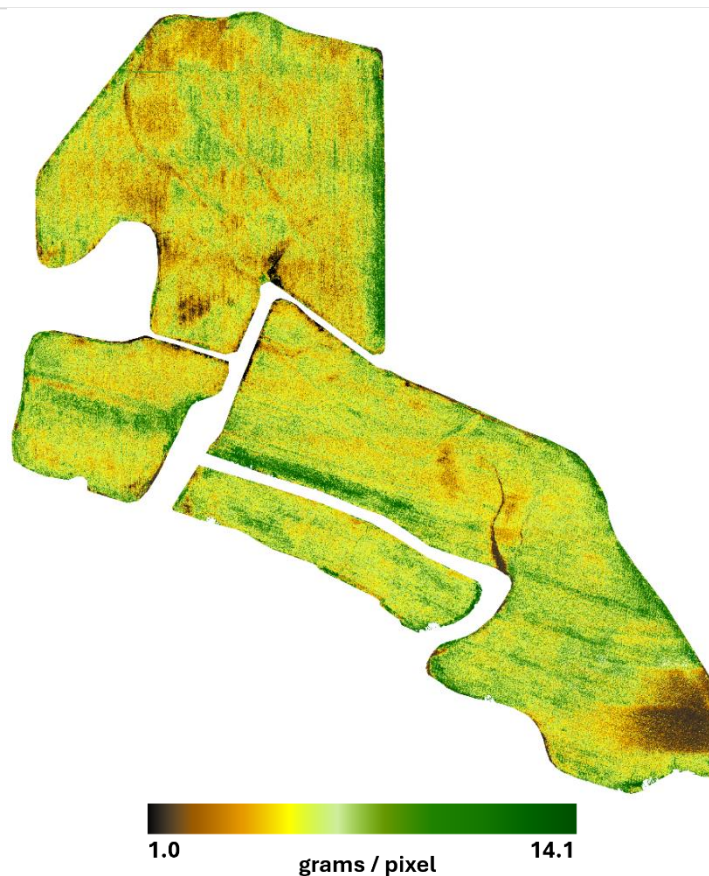
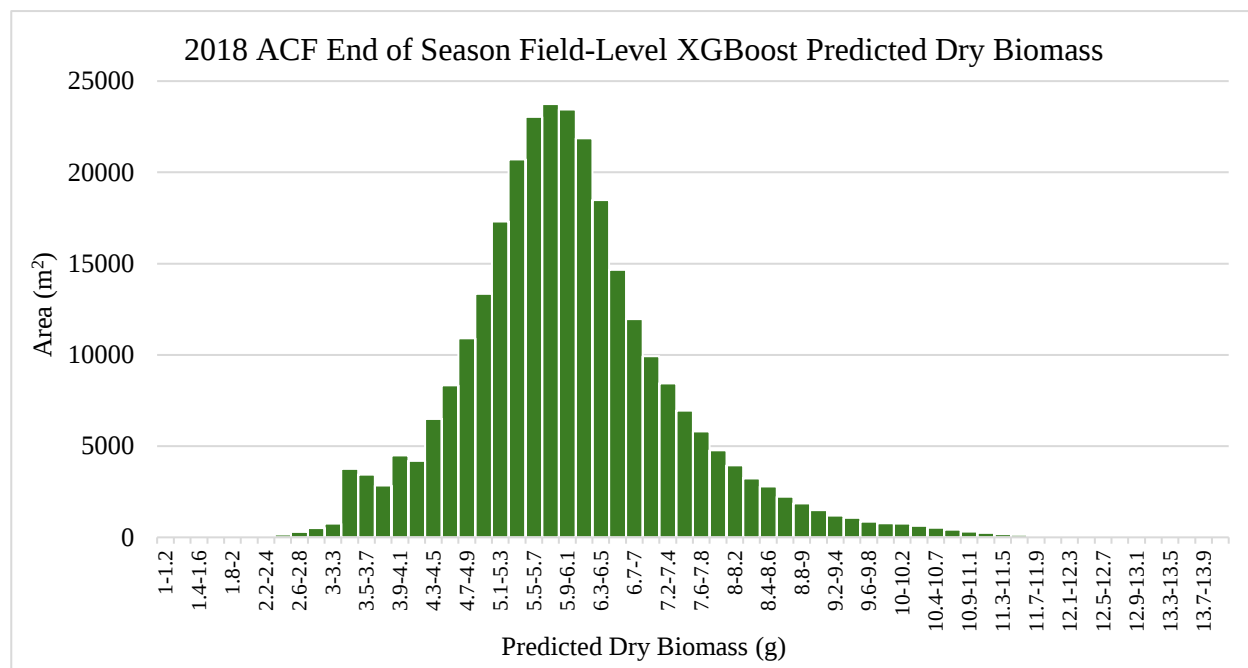


Figure 4.12: Map of field-level predicted biomass for 2018 ACF End of Season Field-Level XGBoost Predicted Dry Biomass model.

Figure 4.13: Histogram displaying the area per pixel value for the 2018 TCF End of Season Field-Level Random Forest Predicted Dry Biomass model.

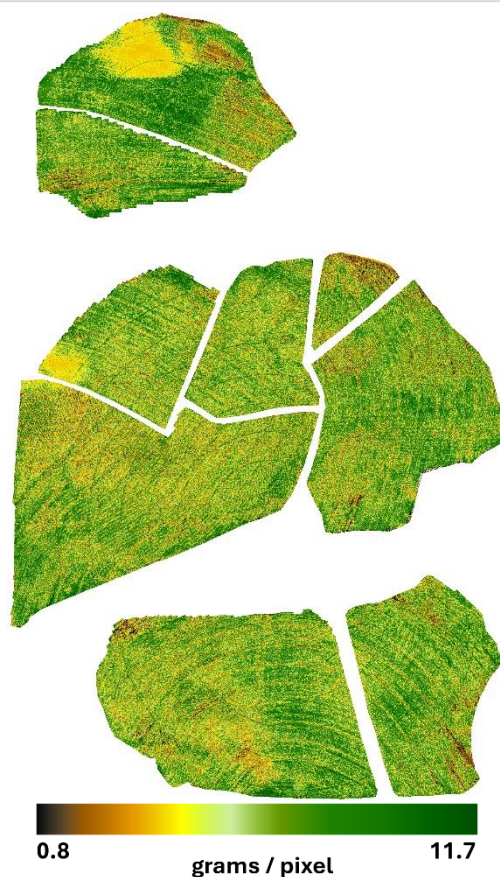
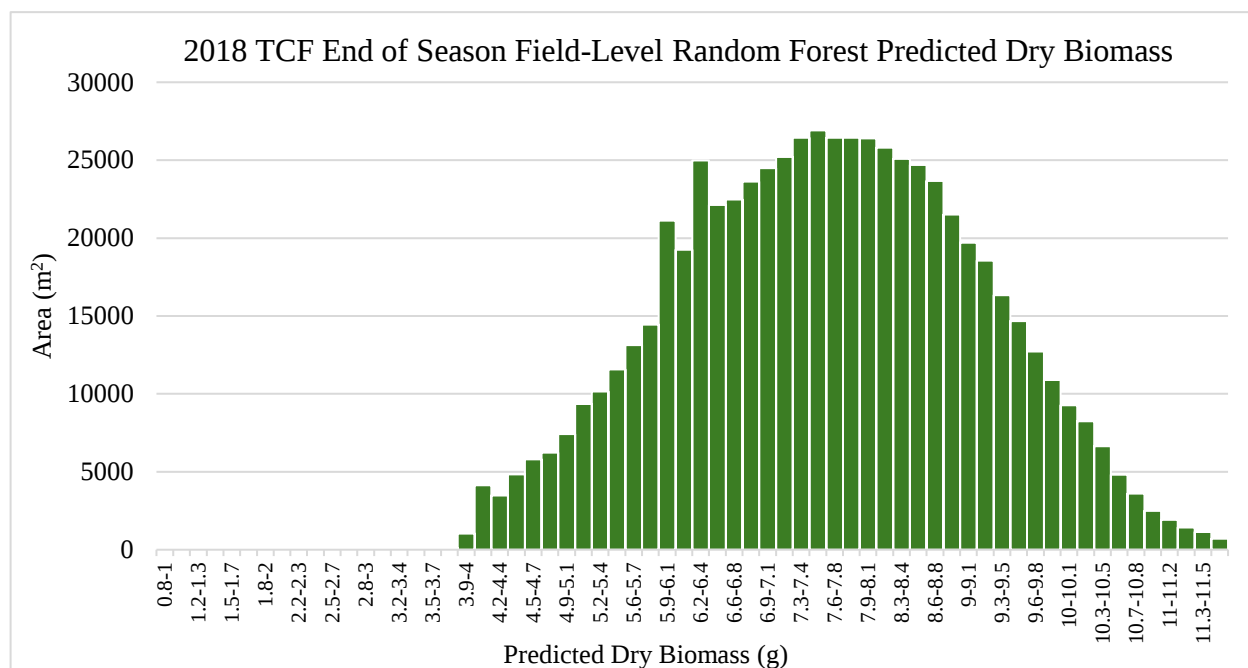


Figure 4.14: Map of field-level predicted biomass for the 2018 TCF End of Season Field-Level Random Forest Predicted Dry Biomass model.

Figure 4.15: Histogram displaying the area per pixel value for the 2018 TCF End of Season Field-Level XGBoost Predicted Dry Biomass model.

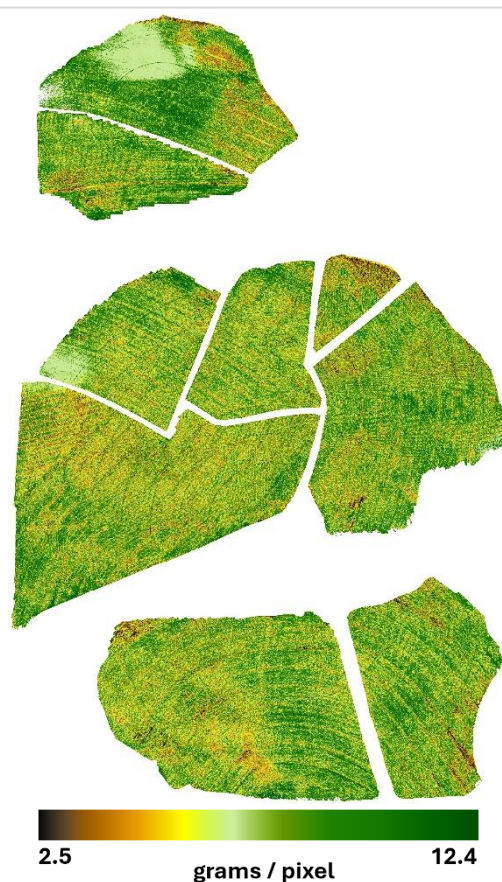
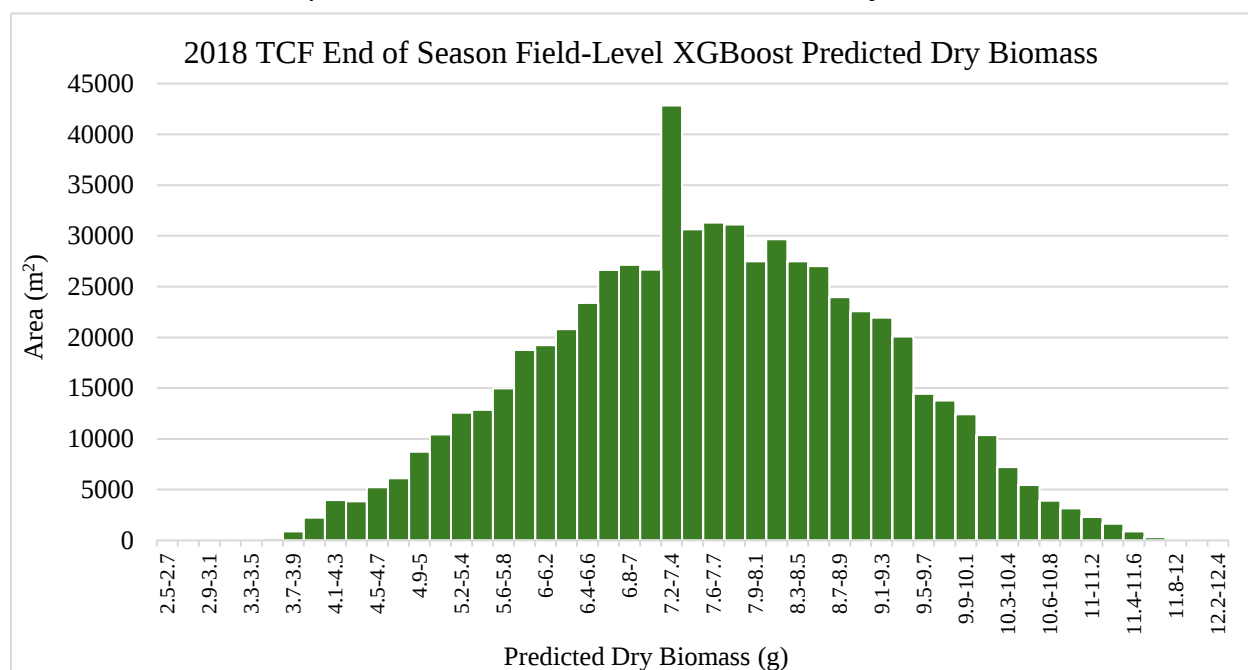


Figure 4.16: Map of field-level predicted biomass for the 2018 TCF End of Season Field-Level XGBoost Predicted Dry Biomass model.

Figure 4.17: Histogram displaying the area per pixel value for the 2019 ACF End of Season Field-Level Random Forest Predicted Dry Biomass model.

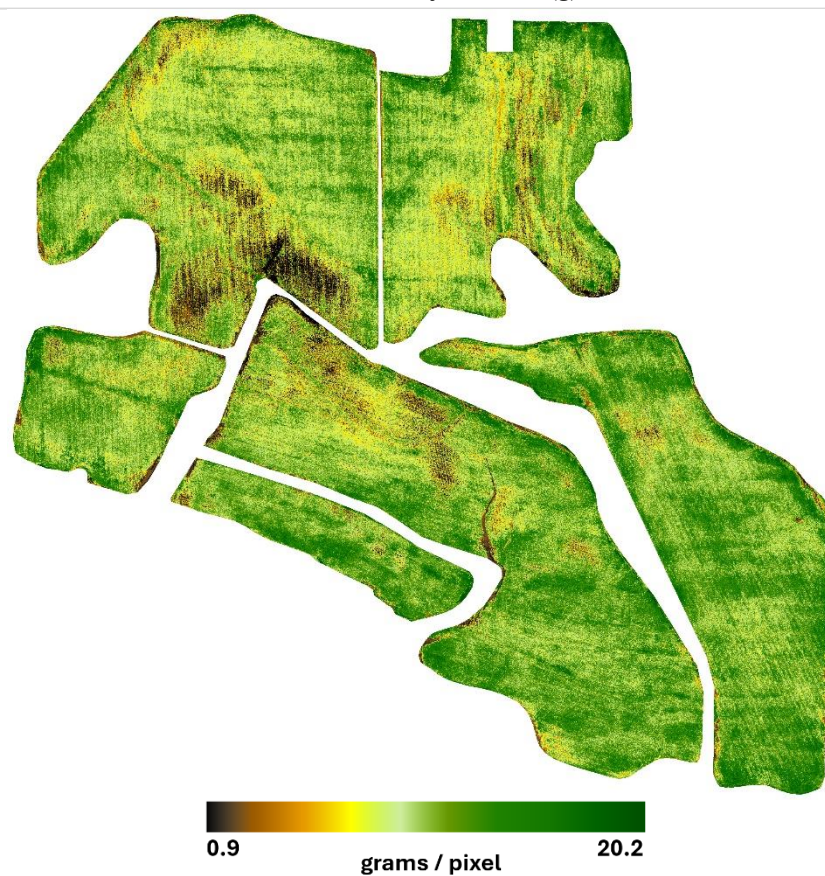
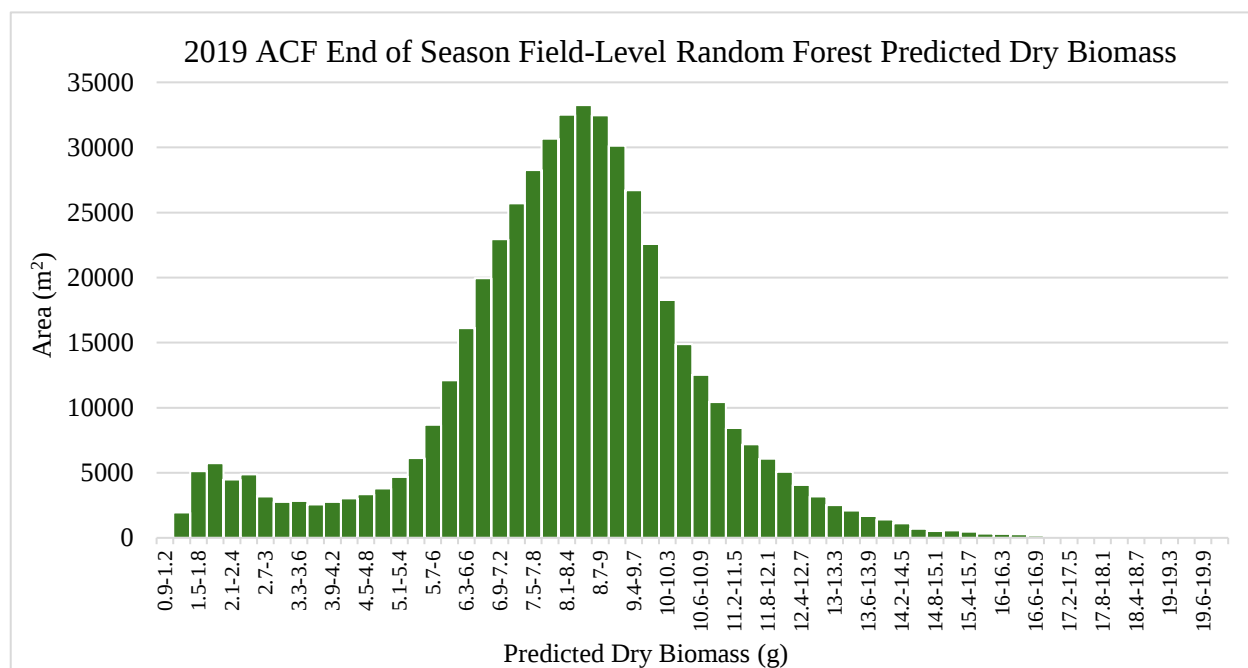


Figure 4.18: Map of field-level predicted biomass for 2019 ACF End of Season Field-Level Random Forest Predicted Dry Biomass model.

Figure 4.19: Histogram displaying the area per pixel value for the 2019 ACF End of Season Field-Level XGBoost Predicted Dry Biomass model.

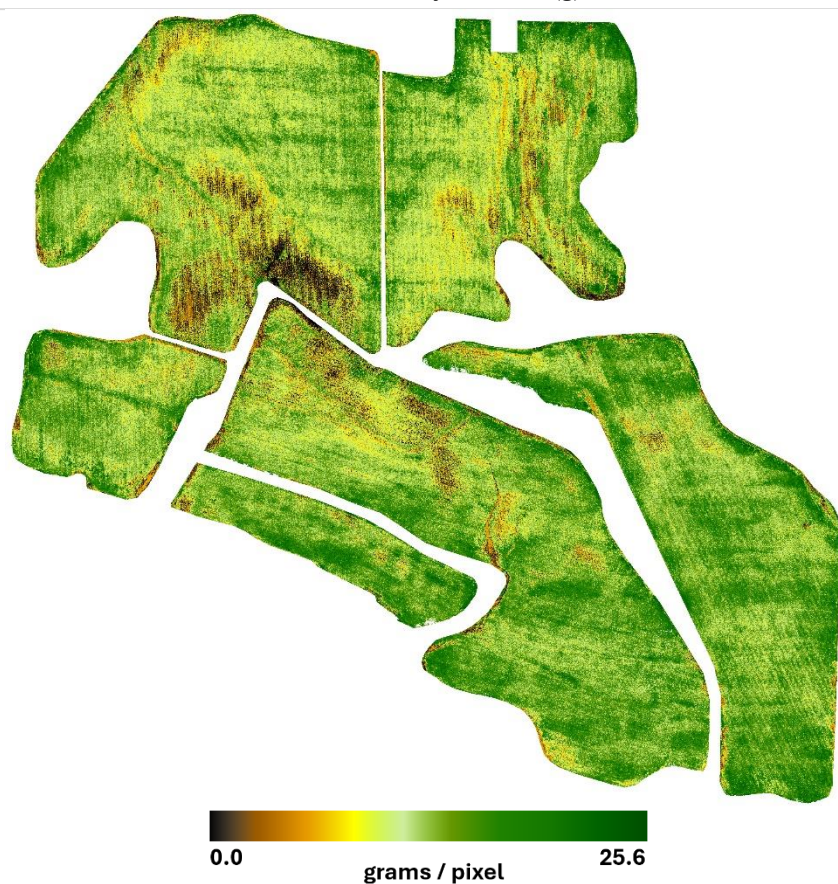
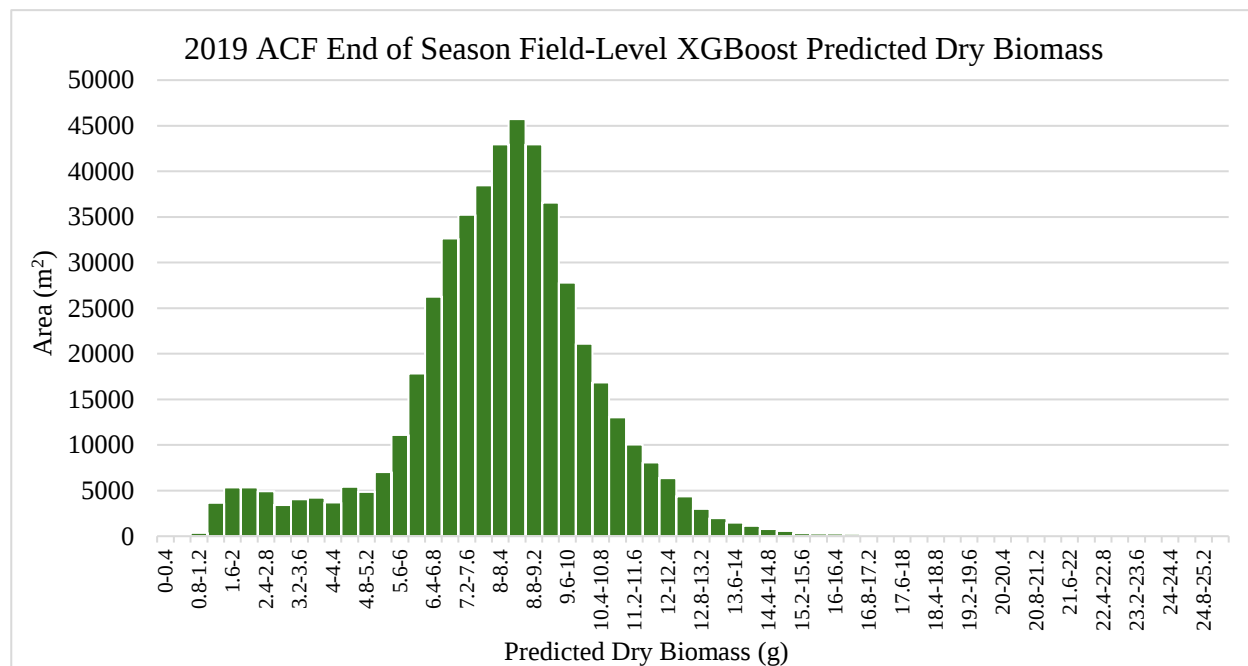


Figure 4.20: Map of field-level predicted biomass for 2019 ACF End of Season Field-Level XGBoost Predicted Dry Biomass model.

Figure 4.21: Histogram displaying the area per pixel value for the 2019 TCF End of Season Field-Level Random Forest Predicted Dry Biomass model.

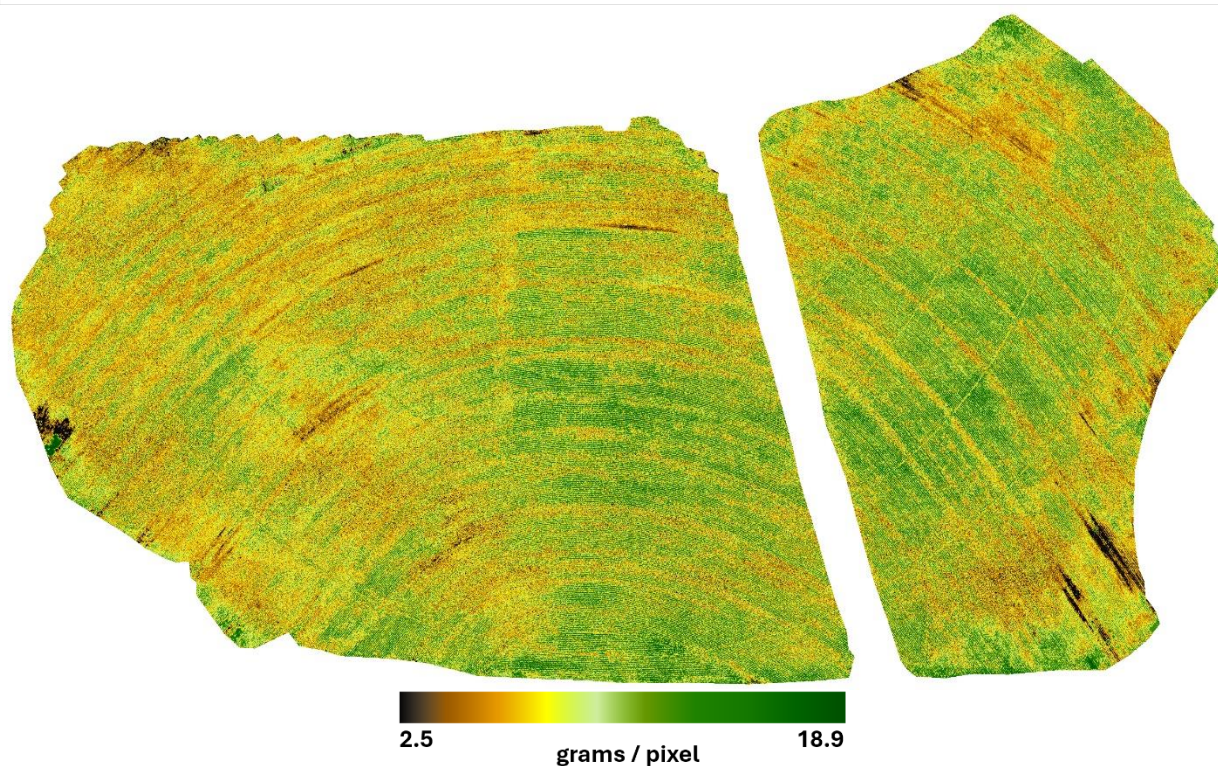
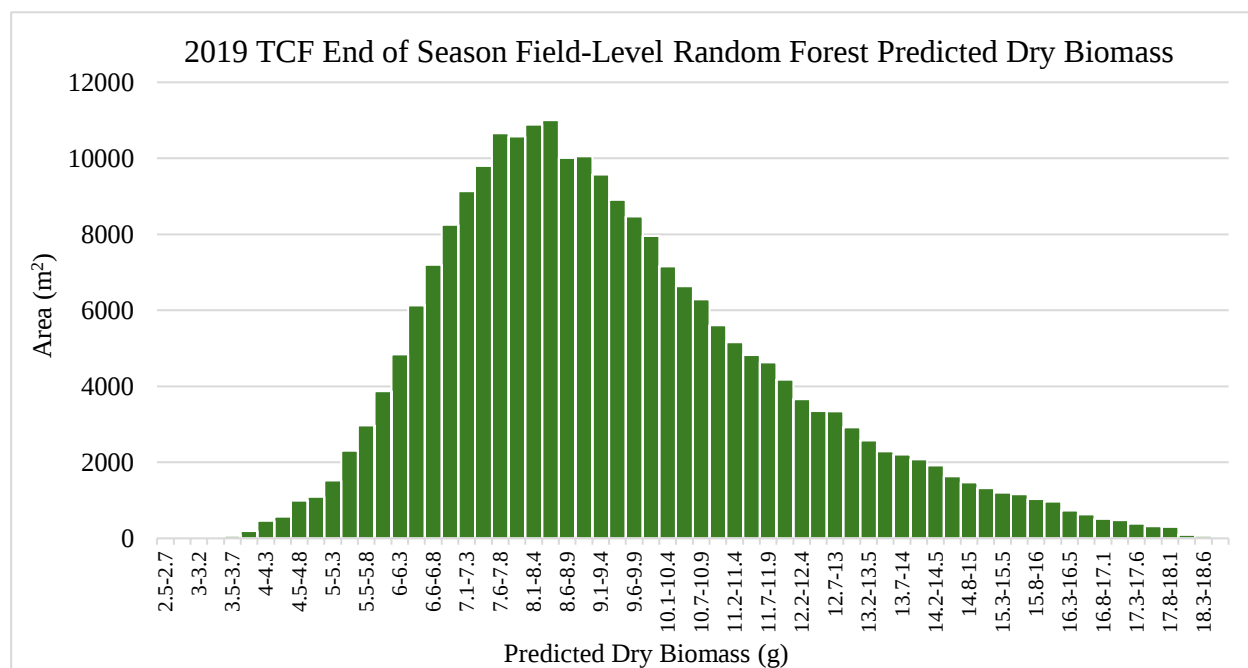


Figure 4.22: Map of field-level predicted biomass for the 2019 TCF End of Season Field-Level Random Forest Predicted Dry Biomass model.

Figure 4.23: Histogram displaying the area per pixel value for the 2019 TCF End of Season Field-Level XGBoost Predicted Dry Biomass model.

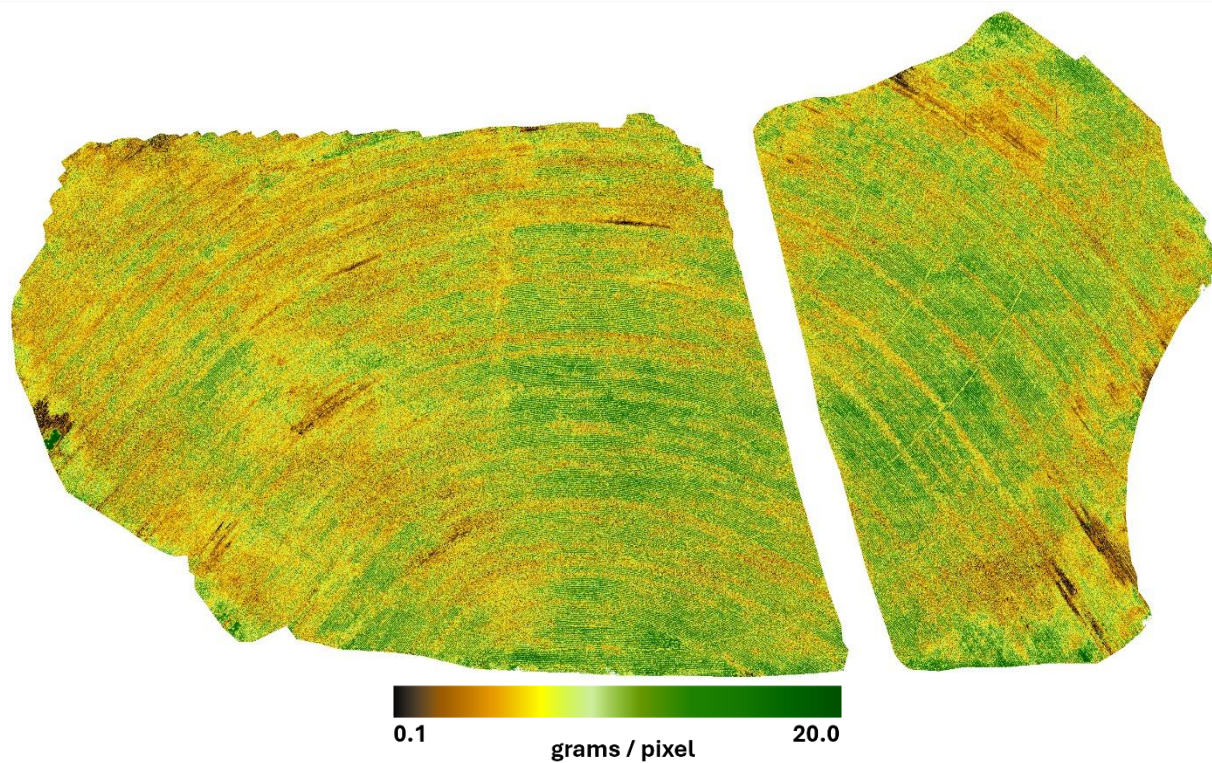
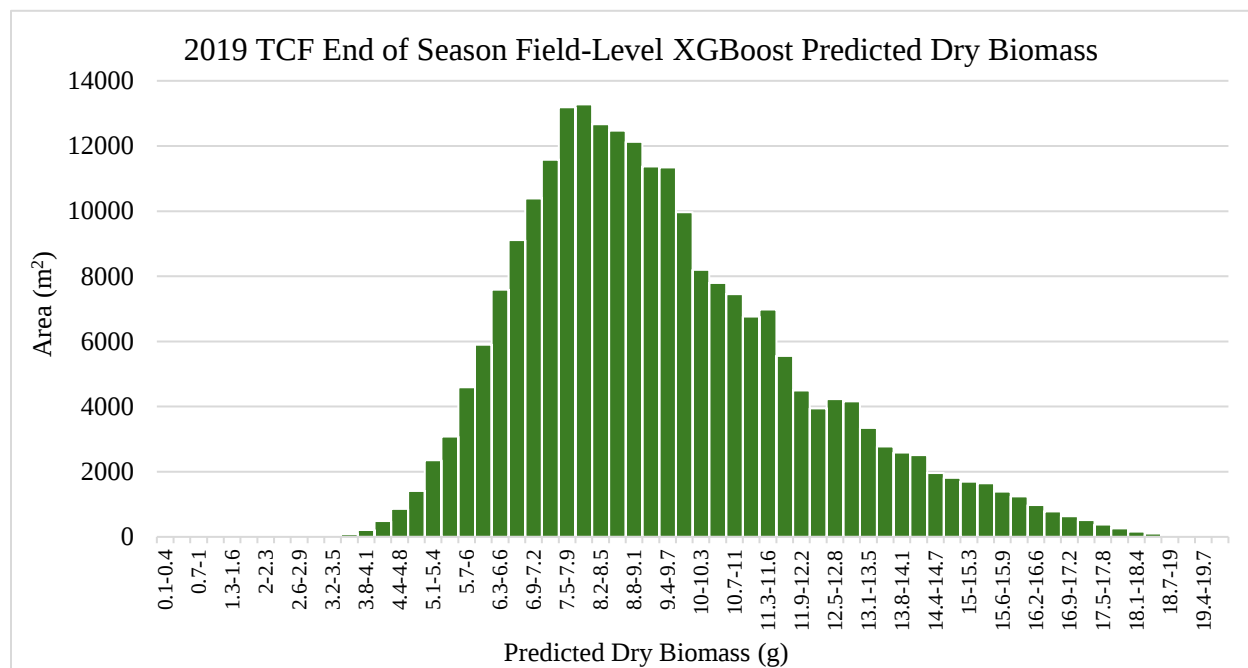


Figure 4.24: Map of field-level predicted biomass for the 2019 TCF End of Season Field-Level XGBoost Predicted Dry Biomass model.

Results for Objective 3: Analyze Cotton Growth and Development in Relation to Precipitation

Cotton growth and development was analyzed against cumulative precipitation using the NDVI time series acquired from the UAS multispectral imagery. Cumulative precipitation are rainfall totals that do not include irrigation. NDVI values for both study sites for the 2018 and 2019 growing seasons years are plotted in green against cumulative precipitation values shown in blue (Figure 4.25). The NDVI time series and cumulative precipitation values are plotted against each other individually for the ACF (Figure 4.26) and the TCF (Figure 4.27).

Cotton AGB was analyzed against cumulative precipitation using the total field-level biomass predictions acquired from the machine learning models. AGB (kg/ha) values for both study sites for the 2018 and 2019 growing seasons years are plotted against cumulative precipitation values (Figure 4.28). The AGB values and cumulative precipitation values are plotted against each other individually for the ACF (Figure 4.29) and the TCF (Figure 4.30).

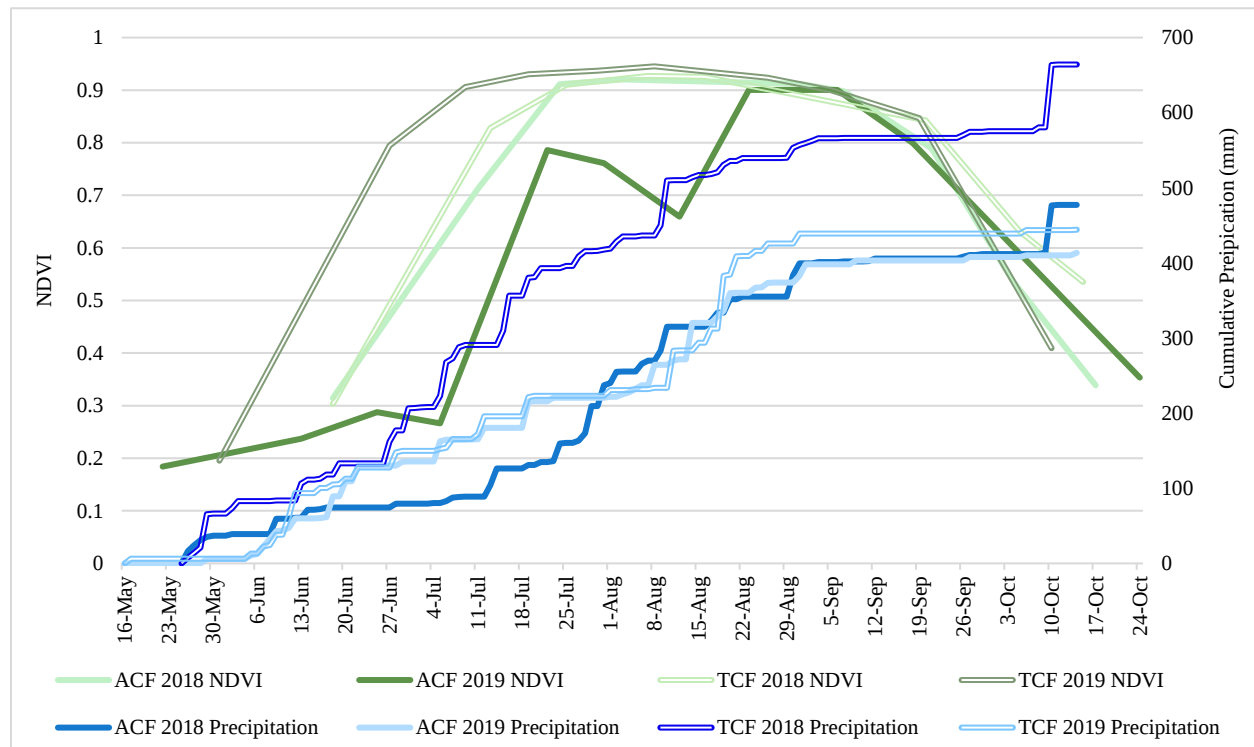


Figure 4.25: NDVI and cumulative precipitation values for the ACF and TCF throughout the 2018 and 2019 growing seasons.

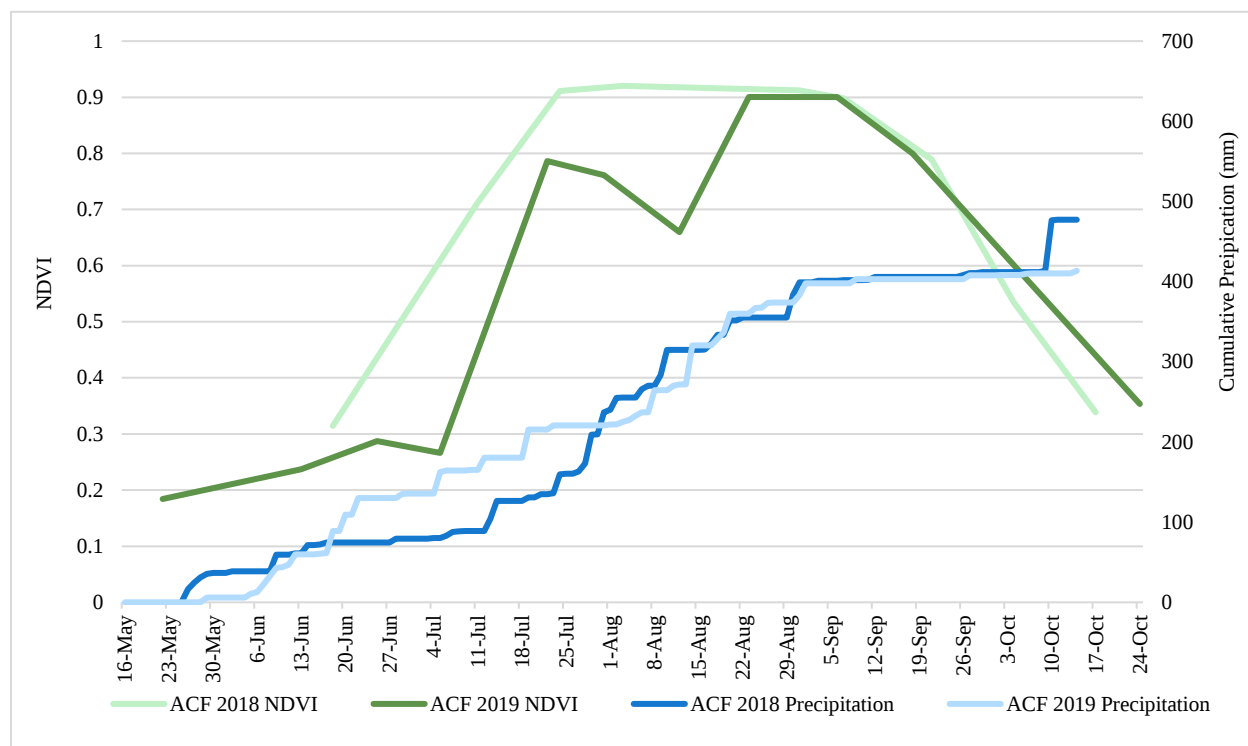


Figure 4.26: NDVI and cumulative precipitation values for the ACF throughout the 2018 and 2019 growing seasons.

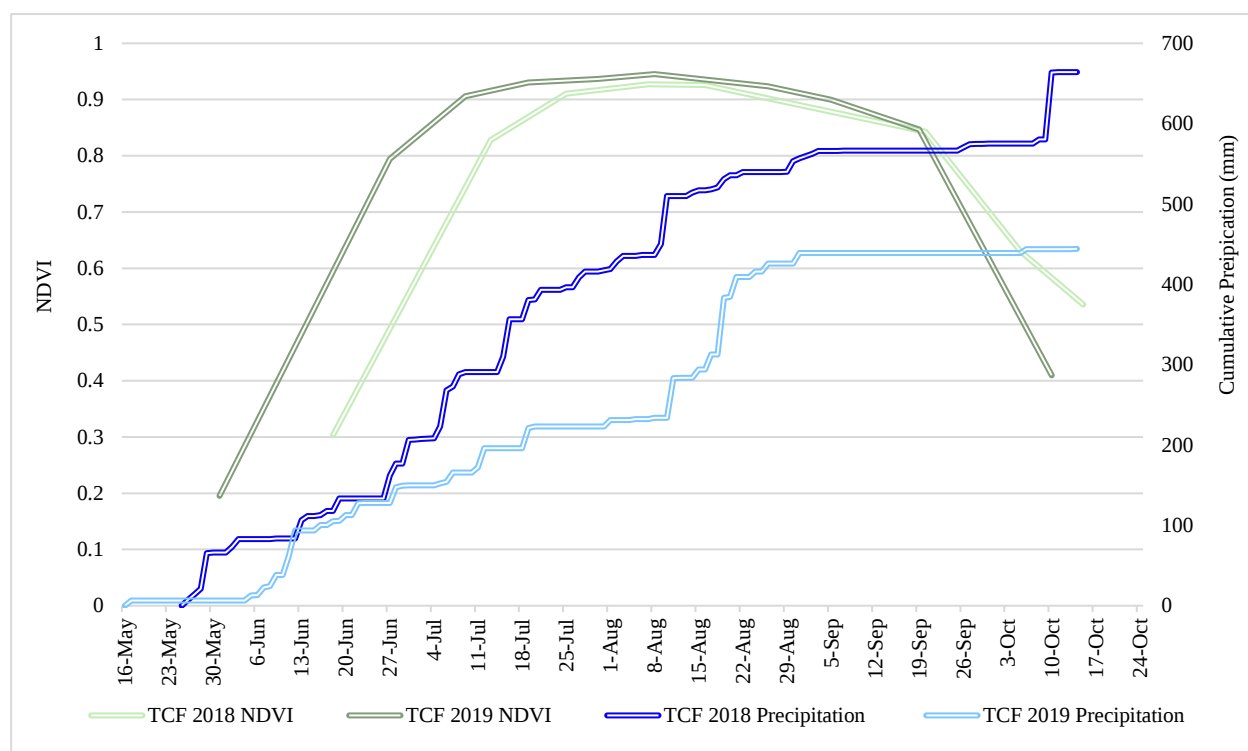


Figure 4.27: NDVI and cumulative precipitation values for the TCF throughout the 2018 and 2019 growing seasons.

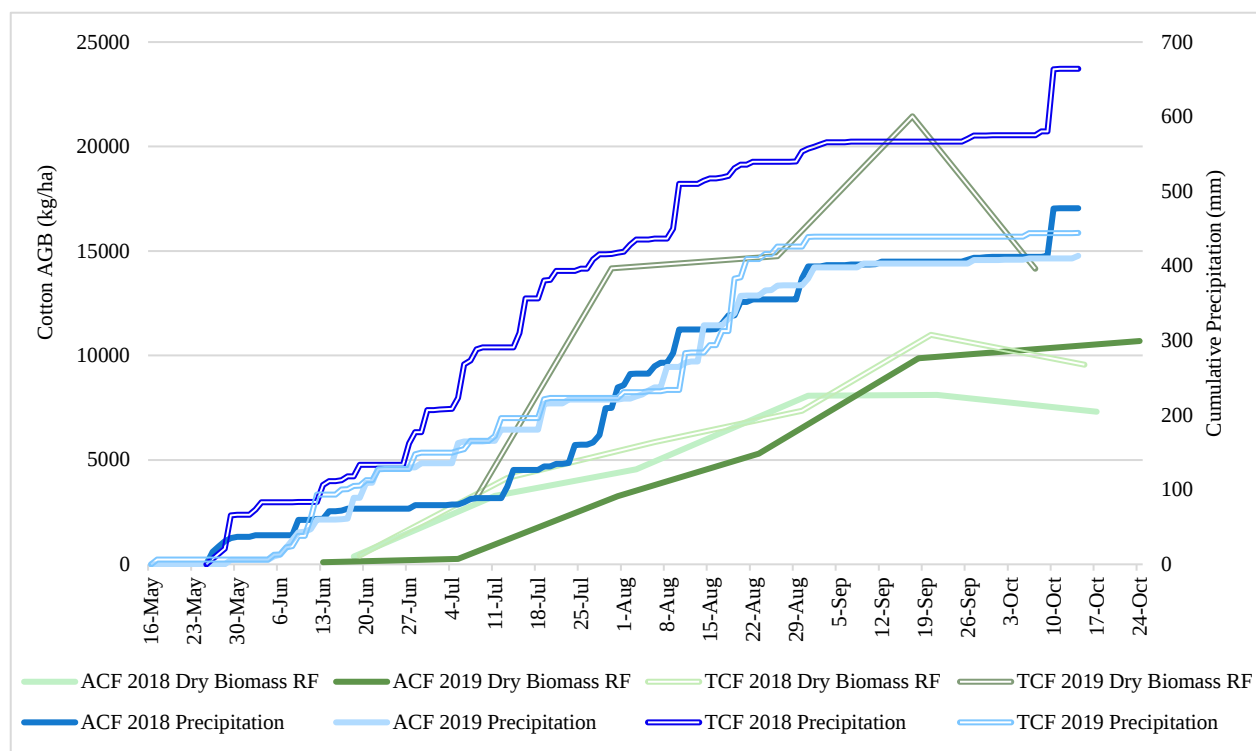


Figure 4.28: Cotton AGB and cumulative precipitation values for the ACF and TCF throughout the 2018 and 2019 growing seasons.

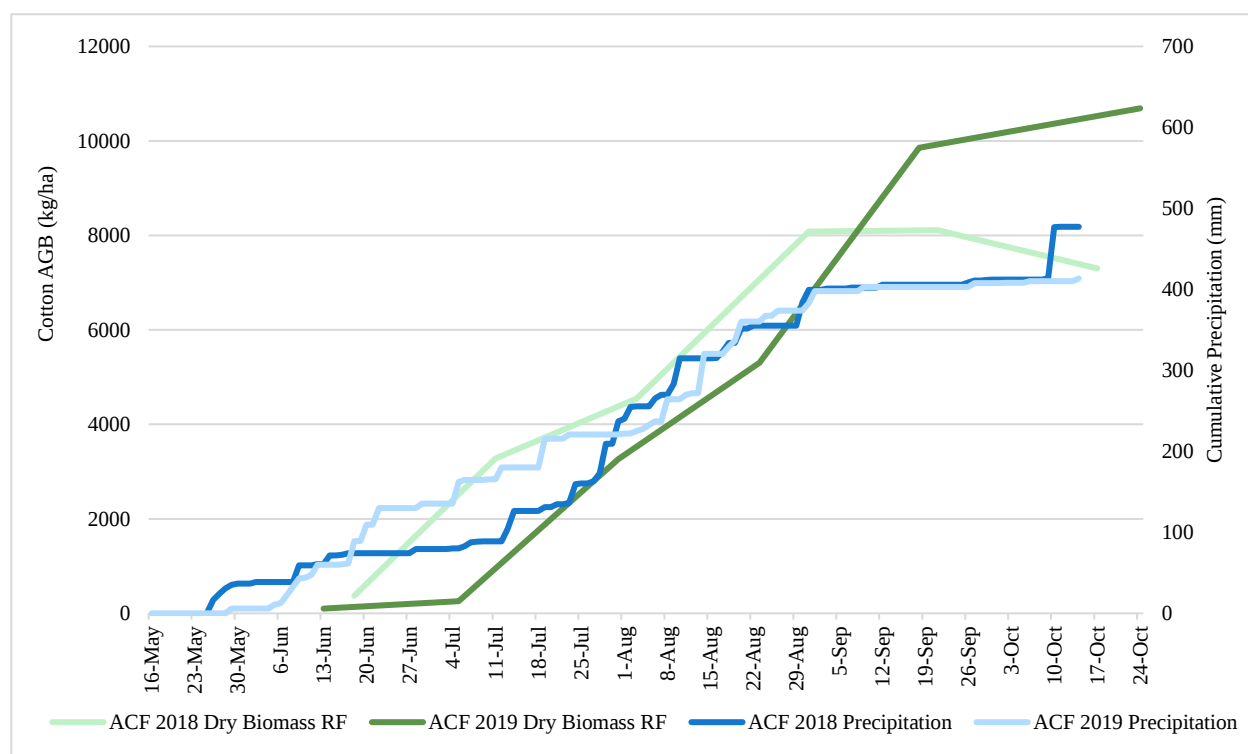


Figure 4.29: Cotton AGB and cumulative precipitation values for the ACF throughout the 2018 and 2019 growing seasons.

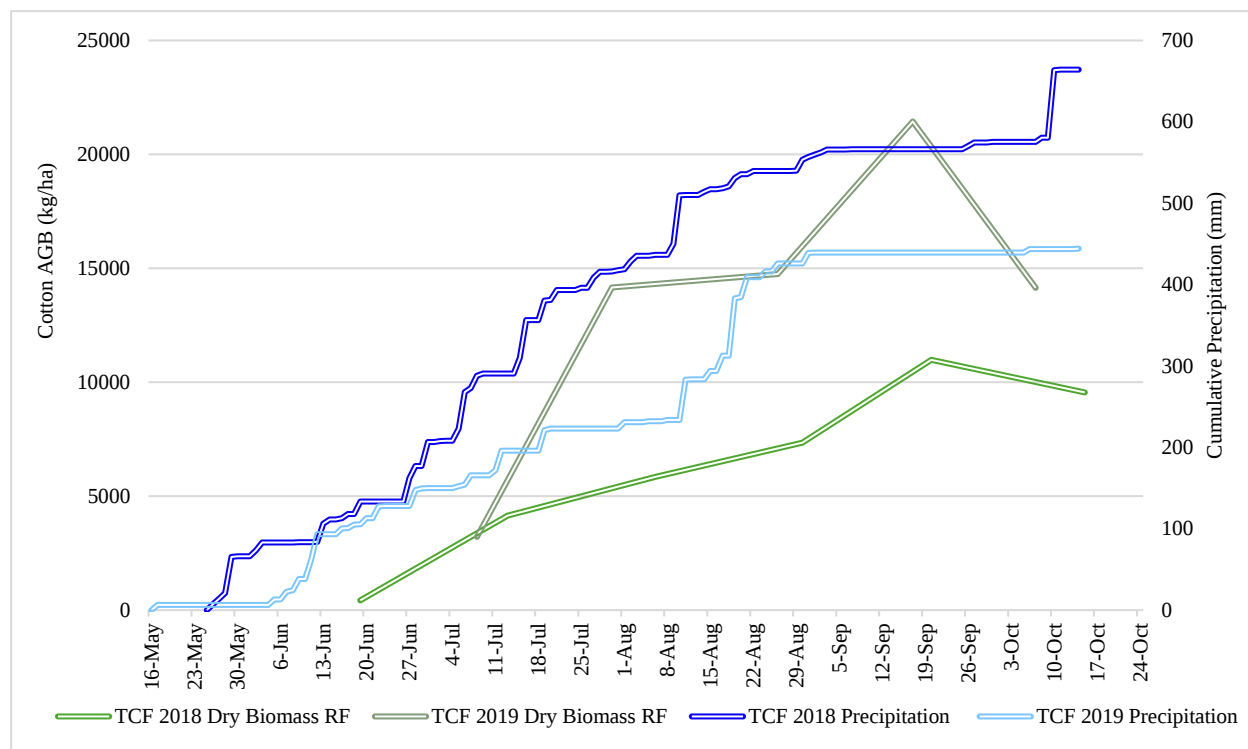


Figure 4.30: Cotton AGB and cumulative precipitation values for the TCF throughout the 2018 and 2019 growing seasons.

Chapter 5

DISCUSSION AND CONCLUSIONS

This research is a continuation and enhancement of the work performed by Stone (2023) entitled, “A methodology for scaling agricultural biomass from ground to regional scales using remote sensing and machine learning analysis”. For his M.S. thesis research, Stone (2023) analyzed the ACF in 2018 using VIs and a Random Forest Model in ArcGIS Pro. Building on Stone’s findings, the study sites in this current research were the ACF & TCF for both 2018 & 2019. Additionally, Raw Bands and a total of 21 VIs were compared to determine the effectiveness of each in predicting cotton AGB scaled from the ground to field level with Random Forest and XGBoost models using the python scikit-learn machine learning library.

The first objective of this study involved reprocessing UAS imagery acquired by the USDA for the two farms on multiple dates during the 2018 and 2019 growing seasons. Checking and reprocessing the imagery is critical to ensuring correct radiometric outputs and improved computation of weighted biomass within USDA field sample plots for scaling AGB to field levels. For example, the end-of-season field-level biomass raster maps for the ACF in 2018 displayed a dark strip in the southeast corner. This flight date, October 17, 2018, was one of four reprocessed datasets where the radiometric calibration had to be reprocessed as detailed in the *Results* section. The other three reprocessed datasets include July 11, 2018, at the ACF, August 3, 2018, at the ACF, and April 30, 2018, at the ACF. The three flight dates in 2018 were aligned with biomass collection flight dates in 2018, and the 2019 flight date was the baseline bare earth flight date before planting. After reprocessing, the field-level raster was radiometrically correct, but extreme

illumination in this area resulted in machine learning models being unable to accurately predict AGB. Acknowledging the challenges of UAS image data collection in southwest Georgia during the humid afternoons at the peak of the growing season, these results emphasize the importance of selecting flight days that have consistent sun illumination. If possible, it is best to avoid: 1) flying on sunny days with spotty clouds moving in during the duration of the flight; and 2) early morning and late afternoon flights with low sun angles and long shadows.

A total of 21 Vegetation Indices (VIs) were computed from band ratios of the UAS multispectral imagery and assessed as potential predictors of AGB. Collinearity analysis was performed to remove redundant variables and improve the efficiency of the predictive models. Results for the ACF in 2019 were depicted as a collinearity matrix in *Figure 4.2* that represents the collinearity findings from both farms in 2018 and 2019 (see Appendix A). The 12 original VIs were highly correlated and reduced to 5 optimal VIs, namely, Normalized Difference Water Index (NDWI), Optimized Soil Adjusted Vegetation Index (OSAVI), RedEdge Ratio Index 2 (RRI2), Triangular Vegetation Index (TVI), and Visible Atmospherically Resistant Index (VARI). Although selecting the most important VIs for the correlation analysis was challenging, eliminating the most redundant VIs determined that NDWI, OSAVI, RRI2, TVI, and VARI were the most important VIs for field-level biomass prediction. While it was difficult to find a threshold to remove variables, and the five VIs did have some collinearity, the five VIs chosen predicted cotton AGB accurately across the two study sites and years while having less collinearity in comparison to the other VIs.

Analyzing the metrics for the models in Objective 2 also was critical to decide whether the five optimal Vegetation Indices (VIs) or Raw Bands are best suited for predicting AGB and creating the field-level biomass maps. While the training, testing, RMSE, and MAE scores were

comparable between the models using VIs and Raw Bands, the amount of predicted AGB in kg/ha values were different. Particularly, the kg/ha values for the 2018 ACF dry biomass were examined since field-measured dry AGB provided known end-of-season biomass values that ranged between 5,000 – 16,000 kg/ha (Dr. Tim Strickland, personal communication, June 1, 2023). Using the VIs, the 2018 Random Forest dry biomass predicted value was 16,938 kg/ha, and the XGBoost value was 23,305 kg/ha. For the Raw Bands, the 2018 Random Forest dry biomass value was 7,302 and the XGBoost value was 7,341 kg/ha. Both models using the VIs predicted AGB greater than the highest value in the known biomass range, and they are drastically different from each other with the XGBoost prediction being over 6,000 kg/ha higher than the Random Forest prediction. Conversely, the predicted AGB values for the two models using Raw Bands were within the range of the field-measured AGB amounts and had less than a 50 kg/ha difference.

These trends were consistent for the predicted AGB in the ACF in 2019 and in the TCF in 2018 and 2019 (*see Table 4.1, Table 4.2, Figure 4.3, and Figure 4.4*). The largest difference between Random Forest and XGBoost for the models using the VIs was 14,316 kg/ha for the 2018 ACF wet biomass. Alternatively, the largest difference between Random Forest and XGBoost for the models using the Raw Bands was 154 kg/ha for the 2019 ACF wet biomass. Due to the inflated and large differences between the kg/ha values for the models using the VIs, the models using the Raw Bands were utilized for scaling up AGB predictions and creating field-level maps. Utilizing the Raw Bands over the VIs is beneficial because it provides the ability to streamline the process of regionalization. Removing the calculations of VIs allows users to simply use the Raw Bands for biomass predictions, and this method can be scaled up to predict cotton at larger geographic ranges by using only the remote sensing data directly from the sensor.

Predicted AGB scaled to field-level was computed for each UAS pixel in the end of season flights (assumed to be the maximum growth of cotton) in 2018 and 2019 for the ACF and TCF farms. Histograms depicting the area per pixel value of predicted wet AGB using Random Forest and XGBoost (see *Figure 4.5* and *Figure 4.7*) and dry AGB (see *Figure 4.9* and *Figure 4.11*) for the 2018 ACF end-of-season biomass flight date show a normal distribution of AGB peaking at 9 to 10 g for wet and 5 to 6 g for dry AGB. There is a slight bimodal peak towards the lower values (approximately 3 to 5 g) of both wet and dry biomass predicted using both models. Based upon the maps of field-level predicted wet and dry biomass for the ACF (see *Figures 4.6, 4.8, 4.10* and *4.12*), it can be inferred that the small bimodal peaks are a result of the intense illumination in the southeast section of the fields resulting in low biomass predictions. The histograms for the Random Forest models have a distribution more akin to a normal distribution in comparison to the histograms for the XGBoost models. The tails on either end of the histograms are enhanced and have more extreme values for the XGBoost models in comparison to the histograms for the Random Forest models. Due to the distribution of biomass values, the field-level maps of predicted AGB from Random Forest models appear greener in comparison to the XGBoost models. Additionally, the maps displaying dry biomass also appear greener than the models displaying wet biomass. The distribution of biomass values in the maps could be due to the different ranges of predicted biomass and therefore different color ramps. Also, wet biomass contains more water content, so the NIR reflectance values would be reduced for wet biomass values.

The remainder of this discussion of results will focus on the field-level histograms and maps displaying predicted dry biomass for the TCF. The histogram depicting the area per pixel values of predicted AGB for the 2018 TCF end-of-season flight date using the Random Forest model (see *Figure 4.13*) has a wider normal distribution with peak values ranging from about 6.5

to 9 g and a long tail skewing the data to the lower AGB values of the histogram (i.e., less than 2 g). Otherwise, the data seems relatively normally distributed with the exception of two spikes at the 5.9 – 6.1 and 6.2 – 6.4 g biomass ranges. The histogram depicting the area per pixel value for the 2018 TCF end-of-season flight date using the XGBoost model (see *Figure 4.15*) also has a wide peak distribution ranging from approximately 7.5 to 8.5 g with two tails on either side of the normal distribution, and an extremely high spike in the 7.2 – 7.4 g biomass range. When looking at the field-level predicted biomass maps (see *Figure 4.14* and *Figure 4.16*), there are areas in the northern portion of the top field and the northwestern portion of the middle field with uniformly low biomass values. These areas are the result of cloud shadows entering the image and modifying the reflectance values for the bands, and it can be inferred that these areas result in the high spikes within the histograms. In general, the map for the Random Forest model appears slightly greener than the map for the XGBoost model.

The histogram depicting the area per pixel value for the 2019 ACF end-of-season flight date using the Random Forest model (see *Figure 4.17*) exhibits a narrower normal distribution peaking at approximately 7.5 to 9.5 g and a bimodal peak near the lower end of the biomass values at 1.5 to 2.5 g, and a long tail skewing the dataset towards the higher end of the biomass values out to 19.5 g. The histogram depicting the area per pixel value for the 2019 ACF end-of-season flight date using the XGBoost model (see *Figure 4.19*) is similar, but the tail is longer skewing the dataset towards the higher end of the biomass values out to 25 g. There is an area near the center of the fields with low biomass values visible in the field-level predicted AGB maps (see *Figure 4.18* and *Figure 4.20*). These areas of lower biomass values are a result of the low elevation within the microtopography of the ACF which appears to serve as a drainage and runoff area within the field. Additionally, along the edge of the fields, tree shadow is present which modifies the biomass

predictions. For the Random Forest model, these areas are depicted as low biomass values; however, for the XGBoost models, these areas were assigned No Data values, and removed from the image. It is difficult to distinguish, but the Random Forest model appears to be slightly greener.

The 2019 TCF maps only have the two most southerly fields in comparison to the 2018 TCF maps because cotton was only grown in these field during the 2019 growing season. The histogram depicting the area per pixel value for the 2019 TCF end-of-season flight date using the Random Forest model (see *Figure 4.21*) is normally distributed with a relatively wide peak (i.e., ranging from 7 to 9.5 g) and is skewed on either end of the peak with wide tails along the X axis. The lower values have a more abrupt cutoff at 3.5 g, and the higher values exhibit a steady decline from 9.5 to 18.5 g. The histogram depicting the area per pixel value for the 2019 TCF end-of-season flight date using the XGBoost model (see *Figure 4.23*) is similar to the Random Forest histogram peaking at 7.5 to 9.5 g, but the distribution is more compressed. It is difficult to distinguish which map has a greener appearance.

Random Forest and XGBoost models were utilized in this study because they are suitable for the given scenario of scaling cotton AGB from the ground to field level. A similar study where Kaur Dhaliwal et al. (2022) predicted cotton and found that tree-based models (Random Forest and XGBoost) outperformed four other models (Kaur Dhaliwal et al. 2022). These findings from Kaur Dhaliwal et al. (2022) support the notion that Random Forest and XGBoost models were suitable for this project. The model metrics tables, histograms, and maps were analyzed for the Raw Bands models. In every category, Random Forest out-performed XGBoost. These findings are similar to the results from Chen et al. (2024) where Random Forest Regression also outperformed a Bayesian Ridge Regression (Chen et al. 2024).

Random Forest models have also performed comparative to deep learning models in cotton AGB predictions. Ashapure et al. (2020) developed a machine learning-based framework to predict cotton and found that a Random Forest Regression produced a 0.846 R^2 value for the average of 10 runs whereas a deep learning Artificial Neural Network produced a 0.861 R^2 value for the average of 10 runs (Ashapure et al. 2020). While the Artificial Neural Network did outperform the Random Forest Model, the difference was not drastic.

One significant benefit for XGBoost is the decreased time it takes for the models to run. In this research, the Random Forest models typically required 20-30 minutes of processing time, but the XGBoost models typically required 5-10 minutes. Upon investigating the inner working for Random Forest and XGBoost, it is revealed that the difference in computational time is a result of how the models' algorithms are designed. Random Forest uses a technique called bagging to create full decision trees with random samples from the dataset, but XGBoost uses gradient boosting to aggregate numerous models into a strong model (Fatima et al., 2023). Due to the architecture, XGBoost models are scalable with high processing speeds in comparison to similar models. Still, the model metrics from this research demonstrate that Random Forest was more accurate. It must be noted, however, that if computational time is a strain on a similar project, then XGBoost may be favored due to its comparable performance.

In Objective 3 of this research, the median field-level NDVI values for the study areas were plotted against cumulative precipitation values, acquired from STEWARDS public access data (STEWARDS, 2025), to assess the impact of precipitation on cotton growth and development (see *Figure 4.25*). The 2018 TCF growing season had higher cumulative, and ended up having over 100 mm more precipitation at the end of the growing season. However, in 2019 for the TCF,

cumulative precipitation started out similar to 2018 in mid-June, was lower from early July to the end of August and then leveled out with nearly no rainfall in the month of September.

In 2018 for the ACF, rainfall was very slow, and began to pick up during July. At the end of the season it was the second highest precipitation dataset. For the ACF in 2019, rainfall was steady, but there was a stretch of minimal rainfall from the beginning of September all the way through October 14th. This difference in rainfall for the two years was reflected in lower NDVI values in 2019. In general, the NDVI values were very low during the beginning and end of the growing seasons, and plateaued around 0.9 for the middle portion of the growing seasons, from about mid-June to mid-September.

The predicted AGB values were plotted against cumulative precipitation values (see *Figure 4.28*). The non-irrigated ACF showed higher predicted dry AGB biomass in the wetter 2018 than the dryer 2019 growing season from June to early September, while the irrigated TCF showed higher predicted dry AGB biomass in 2019. Surprisingly, the 2019 growing season had less cumulative precipitation, yet higher end-of-season AGB values. For both years, the TCF had higher end-of-season AGB values. As discussed previously, there are areas in the 2018 maps for both the ACF and the TCF with extremely low values which skewed the data. Due to these factors, it could explain why the predicted amount of dry AGB biomass in kg/ha in 2018 is lower than 2019 even though the cumulative precipitation in 2018 was higher. Additionally, the agriculturalist managing these fields at the TCF prefers years of less precipitation because cotton does not grow as well in oversaturated conditions.

Cotton overwhelmingly dominates the global natural fiber industry where in 2021, cotton alone represented 76% of the global natural fiber industry, and the average cotton lint yield in the U.S. from 2013 – 2023 was 946 kg/ha (Singh et al., 2023). However, under irrigated conditions, a

potential lint yield of 3,500 kg/ha is obtainable (Constable & Bange, 2015). Exploring the relationship between cotton AGB and yield by analyzing the cotton AGB using a harvest index (HI), the ratio of harvested biomass to yield, is imperative to understanding the economic value from a harvest. In comparison to other crops, cotton has a low HI of approximately 15-20% (Constable & Bange, 2015). Using the Raw Band models and the 20% HI value, the predicted 2018 field-level ACF dry yield values are 1,460 kg/ha and 1,468 kg/ha for Random Forest and XGBoost, respectively (*Table 5.1*). Alternatively, using the VI models and the 20% HI value, the 2018 ACF dry yield values are 3,388 and 4,661 kg/ha for Random Forest and XGBoost, respectively (*Table 5.2*). While the yield results using the Raw Band models are above the 10-year average from 2013 – 2023, they are still well below the potential maximum of 3,500 kg/ha. The yield results for the VI models are extremely high with both the Random Forest and XGBoost predictions exceeding the potential 3,500 kg/ha maximum which is unexpected, especially considering the ACF is not irrigated.

Table 5.1: Predicted cotton yield from models using the VIs with a 20% Harvest Index value.

		Predicted Yield Values Using Vegetation Indices	
		Random Forest	XGBoost
		Predicted Yield (kg/ha)	Predicted Yield (kg/ha)
2018 ACF	Wet Yield	6,032	8,895
	Dry Yield	3,388	4,661
2018 TCF	Wet Yield	9,113	11,141
	Dry Yield	4,079	5,303
2019 ACF	Wet Yield	7,263	9,567
	Dry Yield	4,049	5,173
2019 TCF	Wet Yield	5,277	8,070
	Dry Yield	2,841	4,482

Table 5.2: Predicted cotton yield from models using the Raw Bands with a 20% Harvest Index value.

		Predicted Yield Values Using Raw Bands	
		Random Forest	XGBoost
		Predicted Yield (kg/ha)	Predicted Yield (kg/ha)
2018 ACF	Wet Yield	2,795	2,812
	Dry Yield	1,460	1,468
2018 TCF	Wet Yield	4,159	4,142
	Dry Yield	1,912	1,905
2019 ACF	Wet Yield	3,667	3,636
	Dry Yield	2,138	2,120
2019 TCF	Wet Yield	5,285	5,263
	Dry Yield	2,829	2,833

CONCLUSIONS

Regarding the hypothesis for Objective 1, the reprocessed UAS imagery enhanced the analysis of cotton AGB. By removing the nonparallel flight lines and the images lacking the sun irradiance information, the reprocessed, radiometrically correct orthomosaics provide higher confidence in the end-of-season cotton AGB predictions.

Regarding the hypothesis for Objective 2, the Raw Band models outperformed the models using the VIs. Additionally, 12 VIs still provided extremely high collinearity, so the number of VIs was reduced to five to minimize collinearity and to match the number variables in the Raw Band models. NDVI was not used in this analysis since it is used for creating the weighted biomass layers, which is what the models are attempting to predict to at plot-level, yet further exploration of using NDVI is suggested.

Regarding the hypothesis for Objective 3, in general, the NDVI values for the ACF & the TCF for 2018 & 2019 growing seasons were comparable. Precipitation for the 2018 growing

season was higher than 2019 for nearly the entirety of the growing seasons, but the end-of-season predicted cotton AGB was higher for the study sites in 2019 than in 2018.

In this research, it was found that models scaling up plot-level ground measurements of cotton AGB to predicted field levels utilizing the Raw Bands of multispectral UAS imagery of high spatial resolution outperformed the predictive models utilizing the VIs. Also, the Random Forest models outperformed the XGBoost models, but the XGBoost models had a shorter processing time. The end-of-season predicted cotton AGB was comparable for the models using the Raw Bands, but there were stark differences with the models using the VIs.

Other Remote Sensing imagery, including hyperspectral data from UASs or satellite imagery, can be used for scaling up ground measurements and predicting cotton AGB at broader scales. As the LTAR aims to scale up AGB predictions to regional level, this study acknowledged as a framework for scaling up to field-level, and the methodology in this paper can be utilized for further extrapolation to county, watershed, and even regional levels. The models using surface reflectance from 5-band multispectral imagery from the MicaSense RedEdge-3 sensor provided better results than the computed VIs, so it would be interesting to observe how effective the models would be with dozens of raw bands. Utilizing the microtopography of the field is also expected to provide insight into the field variations affecting plant growth and field-level predicted AGB. Pursuing a Principal Component analysis will allow for further understanding of collinearity between the various VIs and raw bands. Exploring the cost benefit of analysis for the irrigation at the TCF is important to determine if spending the extra funds to irrigate the crops is worthwhile for the increased crop yield. Lastly, exploring other machine learning and deep learning algorithms will improve the confidence of the current models including linear regression, lasso regression, ridge regression, support vector regression and artificial neural network.

This research will aid in the prediction of above ground biomass across the LTAR Gulf Atlantic Coastal Plain region by scaling ground measurements at the plot level to field levels using UAS imagery, best performing vegetation indices, raw bands, and optimized machine learning predictive models. Predicting biomass of crops at varying scales is important to farm owners because, in general, it allows agriculturalists to have an accurate prediction of the potential crop yield. Additionally, the prediction of biomass is also important to the government and consumers because it allows accurate market forecasts for the crops to be analyzed. While the biomass prediction of cotton for this project is constrained to localized areas, the overarching goal for the USDA is to extrapolate to watershed, county, and potentially even regional scales. Future work will leverage the methodology and results presented to predict cotton AGB at broader scales.

REFERENCES

- Ashapure, A., Jung, J., Yeom, J., Chang, A., Maeda, M., Maeda, A., & Landivar, J. (2019). A novel framework to detect conventional tillage and no-tillage cropping system effect on cotton growth and development using multi-temporal UAS data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 49–64.
<https://doi.org/10.1016/j.isprsjprs.2019.04.003>
- Ashapure, A., Jung, J., Chang, A., Oh, S., Yeom, J., Maeda, M., Maeda, A., Dube, N., Landivar, J., Hague, S., & Smith, W. (2020). Developing a machine learning based cotton yield estimation framework using multi-temporal UAS data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169, 180–194.
<https://doi.org/10.1016/j.isprsjprs.2020.09.015>
- Ayanlade, A. (2017). Remote sensing vegetation dynamics analytical methods: a review of vegetation indices techniques. *Geoinformatica Polonica*, 16, 7-17.
<https://doi.org/10.4467/21995923gp.17.001.7188>
- Bannari, A., Asalhi, H., & Teillet, P. M. (2002). Transformed difference vegetation index (TDVI) for vegetation cover mapping. In *IEEE International Geoscience and Remote Sensing Symposium* (Vol. 5, pp. 3053-3055). IEEE.
<https://doi.org/10.1109/IGARSS.2002.1026867>
- Bazzo, C. O. G., Kamali, B., Hütt, C., Bareth, G., & Gaiser, T. (2023). A Review of Estimation Methods for Aboveground Biomass in Grasslands Using UAV. *Remote Sensing*, 15(3), 639. <https://doi.org/10.3390/rs15030639>

- Bean, A. R., Coffin, A. W., Arthur, D. K., Baffaut, C., Holifield Collins, C., Goslee, S. C., Ponce-Campos, G. E., Sclater, V. L., Strickland, T. C., & Yasarer, L. M. (2021). Regional Frameworks for the USDA Long-Term Agroecosystem Research Network. *Frontiers in Sustainable Food Systems*, 4. <https://doi.org/10.3389/fsufs.2020.612785>
- Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., et al. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation*, 39, 79-87. <https://doi.org/10.1016/j.jag.2015.02.012>
- Bosch, D.D., Sheridan, J.M., and Marshall, L.K. (2007). Precipitation, soil moisture, and climate database, Little River Experimental Watershed, Georgia, United States. *Water Resources Research*, 43(9). <https://doi.org/10.1029/2006WR005834>
- Bosch, D.D., Coffin A.W., Sheridan J., Pisani O., Endale D.M., Strickland T.C. (2021). Little River Experimental Watershed, a keystone in understanding of Coastal Plain watersheds. *Hydrological Processes*, 35:e14334. <https://doi.org/10.1002/hyp.14334>.
- Bosch, D.D., Sheridan J.M., Davis F.M., Strickland T.C., Endale D.M., Coffin A.W., Pisani O. (2022). “Little River Experimental Watershed History”, in: J. S. Latimer, et al., *The Seventh Interagency Conference on Research in the Watersheds*, U.S. Department of Agriculture Forest Service, Southern Research Station. <https://doi.org/10.2737/SRS-GTR-264>
- Broge, N. H., & Leblanc, E. (2001, January 1). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, 76(2), 156–172. [https://doi.org/10.1016/S0034-4257\(00\)00197-8](https://doi.org/10.1016/S0034-4257(00)00197-8)

Chen, M., Yin, C., Lin, T., Liu, H., Wang, Z., Jiang, P., Ali, S., Tang, Q., & Jin, X. (2024).

Integration of unmanned aerial vehicle spectral and textural features for accurate above-ground biomass estimation in cotton. *Agronomy*, 14(6).

<https://doi.org/10.3390/agronomy14061313>

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In

Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794). <https://doi.org/10.1145/2939672.2939785>

Coffin A.W., Bosch D.D., Strickland T.C., Endale D.M., Pisani O., Lowrance R. (2022).

“Cropping patterns over two decades in the Little River Experimental Watershed, Georgia, USA”, in: J. S. Latimer, et al. (eds.), *The Seventh Interagency Conference on Research in the Watersheds*, U.S. Department of Agriculture Forest Service, Southern Research Station, Hosted virtually by the USDA-ARS and the University of Georgia, Tifton, GA. pp. 186-190.

https://www.srs.fs.usda.gov/pubs/gtr/gtr_srs264.pdf#page=202

Coffin, A. W., Cosh, M. H., Pisarello, K. (2023). Data from: Two Years of Cotton (*Gossypium hirsutum* L.) Data from the Gulf Atlantic Coastal Plain LTAR Network Site. Ag Data Commons. Dataset.

<https://www.sidalc.net/search/Record/dat-usda-us-article24857439/Description>

Coffin, A. W., Pisani, O., Pisarello, K., Porter, K., Bosch, D. D., & Strickland, T. C. (2024a).

The LTAR Croplands Common Experiment at the Gulf Atlantic Coastal Plain. *Journal of Environmental Quality*. <https://doi.org/10.1002/jeq2.20645>

- Coffin, A.W., Cosh, M.H. & Pisarello, K. (2024b). Two Years of Cotton (*Gossypium hirsutum* L.) Data from the Georgia Coastal Plain, USA. *Sci Data* 11, 1037.
<https://doi.org/10.1038/s41597-024-03716-z>
- Constable, G. A., & Bange, M. P. (2015). The yield potential of cotton (*Gossypium hirsutum* L.). *Field Crops Research*, 182, 98-106. <https://doi.org/10.1016/j.fcr.2015.07.017>
- Deb, D., Mandal, S., Deb, S., Choudhury, A., Hembram, S. (2021). Crop Production Estimation Using Remote Sensing. In: Mitran, T., Meena, R.S., Chakraborty, A. (eds) *Geospatial Technologies for Crops and Soils*. Springer, Singapore. https://doi.org/10.1007/978-981-15-6864-0_6
- D'Oliveira, M. V. N., Broadbent, E. N., Oliveira, L. C., Almeida, D. R. A., Papa, D. A., Ferreira, M. E., Zambrano, A. M. A., Silva, C. A., Avino, F. S., Prata, G. A., Mello, R. A., Figueiredo, E. O., Castro Jorge, L. A. de, Junior, L., Albuquerque, R. W., Brancalion, P. H. S., Wilkinson, B., & Oliveira-Da-Costa, M. (2020). Aboveground biomass estimation in Amazonian tropical forests: a comparison of aircraft- and GatorEye UAV-borne LiDAR data in the chico mendes extractive reserve in Acre, Brazil. *Remote Sensing*, 12(11). <https://doi.org/https://www.mdpi.com/2072-4292/12/11/1754>
- Ehammer, A., Fritsch, S., Conrad, C., Lamers, J., & Dech, S. (2010, October). Statistical derivation of fPAR and LAI for irrigated cotton and rice in arid Uzbekistan by combining multi-temporal RapidEye data and ground measurements. In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XII* (Vol. 7824, pp. 66-75). SPIE.
<https://doi.org/10.1117/12.864796>
- Escadafal, R., & Huete, A. (1991). Improvement in Remote-Sensing of Low Vegetation Cover in Arid Regions by Correcting Vegetation Indexes for Soil Noise. *Comptes Rendus De L*

Academie Des Sciences Serie Ii, 312(11), 1385–1391.

<https://www.osti.gov/etdeweb/biblio/22340487>

Fatima, S., Hussain, A., Amir, S. B., Ahmed, S. H., & Aslam, S. M. H. (2023). XGBoost and random forest algorithms: an in depth analysis. *Pakistan Journal of Scientific Research*, 3(1), 26-31. <https://doi.org/10.57041/pjosr.v3i1.946>

Feng AiJing, F. A., Zhang MeiNa, Z. M., Sudduth, K. A., Vories, E. D., & Zhou JianFeng, Z. J. (2019). Cotton yield estimation from UAV-based plant height. *Transactions of the ASABE*, 62(2), 393–404. <https://doi.org/10.13031/trans.13067>

Feng, A., Zhou, J., Vories, E. D., Sudduth, K. A., & Zhang, M. (2020). Yield estimation in cotton using UAV-based multi-sensor imagery. *Biosystems Engineering*, 193, 101–114. <https://doi.org/10.1016/j.biosystemseng.2020.02.014>

Gaso, D. V., Berger, A. G., & Ciganda, V. S. (2019). Predicting wheat grain yield and spatial variability at field scale using a simple regression or a crop model in conjunction with Landsat images. *Computers and Electronics in Agriculture*, 159, 75-83. <https://doi.org/10.1016/j.compag.2019.02.026>

Gerardo, R., de Lima, I., (2023). Applying RGB-Based Vegetation Indices Obtained from UAS Imagery for Monitoring the Rice Crop at the Field Scale: A Case Study in Portugal. *Agriculture*, 13(10), 1916. <https://doi.org/10.3390/agriculture13101916>

Gil-Docampo, M. D. L. L., Arza-García, M., Ortiz-Sanz, J., Martínez-Rodríguez, S., Marcos-Robles, J. L., & Sánchez-Sastre, L. F. (2020). Above-ground biomass estimation of arable crops using UAV-based SfM photogrammetry. *Geocarto International*, 35(7), 687-699. <https://doi.org/10.1080/10106049.2018.1552322>

- Gitelson, A., & Merzlyak, M. N. (1994). Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology*, 22(3), 247–252. DOI: [https://doi.org/10.1016/1011-1344\(93\)06963-4](https://doi.org/10.1016/1011-1344(93)06963-4)
- Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS- MODIS. *Remote Sensing of Environment*, 58(3), 289–298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7)
- Gitelson, A. A., Viña, A., Arkebauer, T. J., Rundquist, D. C., Keydan, G., & Leavitt, B. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophysical Research Letters*, 30(5), 1248. <https://doi.org/10.1029/2002GL016450>
- Gitelson, A. A. (2004). Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation. *Journal of Plant Physiology*, 161(2), 165–173. DOI: <https://doi.org/https://doi.org/10.1078/0176-1617-01176>
- Glenn, E. P., Huete, A. R., Nagler, P. L., & Nelson, S. G. (2008). Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors*, 8(4), 2136–2160. <https://doi.org/10.3390/s8042136>
- Glenn, D. M., & Tabb, A. (2018). Evaluation of Five Methods to Measure Normalized Difference Vegetation Index (NDVI) in Apple and Citrus. *International Journal of Fruit Science*, 19(2), 191–210. <https://doi.org/10.1080/15538362.2018.1502720>
- Gobron, N., Pinty, B., Verstraete, M. M., & Widlowski, J. L. (2000). Advanced vegetation indices optimized for up-coming sensors: Design, performance, and applications. *IEEE Transactions on Geoscience and Remote Sensing*, 38(6), 2489–2505.

<https://doi.org/10.1109/36.885197>

Habyarimana, E., Piccard I., Catellani, M., De Franceschi, P., Dall'Agata, M. (2019). Towards Predictive Modeling of Sorghum Biomass Yields Using Fraction of Absorbed Photosynthetically Active Radiation Derived from Sentinel-2 Satellite Imagery and Supervised Machine Learning Techniques. *Agronomy*. 9(4):203.

<https://doi.org/10.3390/agronomy9040203>

Hassler, S. C., & Baysal-Gurel, F. (2019). Unmanned Aircraft System (UAS) Technology and Applications in Agriculture. *Agronomy*, 9(10), 618.

<https://doi.org/10.3390/agronomy9100618>

Hu, T., Zhang, Y., Su, Y., Zheng, Y., Lin, G., & Guo, Q. (2020). Mapping the global mangrove forest aboveground biomass using multisource remote sensing data. *Remote Sensing*, 12(10). <https://doi.org/https://www.mdpi.com/2072-4292/12/10/1690>

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1-2), 195-213.

[https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)

Janowicz, K., Gao, S., McKenzie, G., Hu, Y., & Bhaduri, B. (2019). GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *International Journal of Geographical Information Science*, 34(4), 625–636.

<https://doi.org/10.1080/13658816.2019.1684500>

- Jensen, J. R. (2007). *Remote Sensing of the Environment: An Earth Resource Perspective* (2nd ed.). Pearson Prentice Hall.
<https://archive.org/details/remotesensingofe0000jens/page/n1/mode/2up>
- Jiang, Q., Fang, S., Peng, Y., Gong, Y., Zhu, R., Ma, Y., Wu, X., Duan, B., Liu, J. (2019). UAV-Based Biomass Estimation for Rice-Combining Spectral, TIN-Based Structural and Meteorological Features. *Remote Sensing*, 11(7), 890. <https://doi.org/10.3390/rs11070890>
- Jiang, Z., Huete, A. R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112(10), 3833–3845. <https://doi.org/10.1016/j.rse.2008.06.006>
- Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, 50(4), 663-666. <https://doi.org/10.2307/1936256>
- Kaur Dhaliwal, J., Panday, D., Saha, D., Lee, J., Jagadamma, S., Schaeffer, S., & Mengistu, A. (2022). Predicting and interpreting cotton yield and its determinants under long-term conservation management practices using machine learning. *Computers and Electronics in Agriculture*, 199. <https://doi.org/10.1016/j.compag.2022.107107>
- Kriegler, F.J., W.A. Malila, R.F. Nalepka, and W. Richardson. (1969). Preprocessing transformations and their effects on multispectral recognition. Proceedings of the Sixth International Symposium on Remote Sensing of Environment. Ann Arbor, Michigan, USA: Remote Sensing of Environment 1, 97–131.
<https://cir.nii.ac.jp/crid/1572261549021269760>
- Li, F., Bai, J., Zhang, M., & Zhang, R. (2022). Yield estimation of high-density cotton fields using low-altitude UAV imaging and deep learning. *Plant Methods*, 18(1), 1–11.
<https://doi.org/10.1186/s13007-022-00881-3>

- McFeeters, S. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425–1432. <https://doi.org/10.1080/01431169608948714>
- NCEI. (2023). *Annual 2023 National Climate Report*. Annual 2023 National Climate Report | National Centers for Environmental Information (NCEI). <https://www.ncei.noaa.gov/access/monitoring/monthly-report/national/202313>
- Niu, Y., Zhang, L., Zhang, H., Han, W., and Peng, X. (2019). Estimating above-ground biomass of maize using features derived from UAV-based RGB imagery. *Remote Sensing*, 11(11), 1261. <https://doi.org/10.3390/rs11111261>
- Panda, S. S., Ames, D. P., & Panigrahi, S. (2010). Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques. *Remote Sensing*, 2(3), 673–696. <https://doi.org/10.3390/rs2030673>
- Perry, E. M., Morse-McNabb, E. M., Nuttall, J. G., O Leary, G. J., & Clark, R. (2014). Managing Wheat From Space: Linking MODIS NDVI and Crop Models for Predicting Australian Dryland Wheat Biomass. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing, 7(9), 3724–3731. <https://doi.org/10.1109/JSTARS.2014.2323705>
- Poley, L. G., & Mcdermid, G. J. (2020). A systematic review of the factors influencing the estimation of vegetation aboveground biomass using unmanned aerial systems. *Remote Sensing*, 12(7). <https://doi.org/https://www.mdpi.com/2072-4292/12/7/1052>

- Psiroukis, V., Papadopoulos, G., Kasimati, A., Tsoulas, N., & Fountas, S. (2023). Cotton Growth Modelling Using UAS-Derived DSM and RGB Imagery. *Remote Sensing*, 15(5), 1214. <https://doi.org/10.3390/rs15051214>
- Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil vegetation adjusted index. *Remote Sensing of Environment*, 48(2), 119–126. [https://doi.org/10.1016/0034-4257\(94\)90134-1](https://doi.org/10.1016/0034-4257(94)90134-1)
- Qiu, Z., Ma, F., Li, Z., Xu, X., & Du, C. (2022). Development of Prediction Models for Estimating Key Rice Growth Variables Using Visible and NIR Images from Unmanned Aerial Systems. *Remote Sensing*, 14(6), 1384. <https://doi.org/10.3390/rs14061384>
- Raper, T. B., & Varco, J. J. (2015). Canopy-scale wavelength and vegetative index sensitivities to cotton growth parameters and nitrogen status. *Precision Agriculture*, 16, 62–76. <https://doi.org/10.1007/s11119-014-9383-4>
- Rondeaux, G., Steven, M., and Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*. 55: 95– 107. DOI: [https://doi.org/10.1016/0034-4257\(95\)00186-7](https://doi.org/10.1016/0034-4257(95)00186-7)
- Rouse, J. W., Haas, R. H., Schell, J. A. & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ.* 351(1), 309. https://books.google.com/books?hl=en&lr=&id=cl42FB2_UEcC&oi=fnd&pg=PA309&dq=Monitoring+vegetation+systems+in+the+Great+Plains+with+ERTS&ots=vtbNPtbR4x&sig=sqi9udccs2FZxeTOVxom60x4ZDI
- Scheider, S., Richter, KF. *GeoAI. Künstl Intell* 37, 5–9 (2023). <https://doi.org/10.1007/s13218-022-00797-z>

- Siegfried, J. A. (2021). *Unmanned Aerial Remote Sensing for Estimating Cotton Yield*. [Doctoral dissertation, Texas A&M University]. OAKTrust. <https://hdl.handle.net/1969.1/196253>
- Singh, J., Gamble, A. V., Brown, S., Campbell, B. T., Jenkins, J., Koebernick, J., ... & Sanz-Saez, A. (2023). 65 years of cotton lint yield progress in the USA: Uncovering key influential yield components. *Field Crops Research*, 302, 109058.
<https://doi.org/10.1016/j.fcr.2023.109058>
- Sofonia, J., Shendryk, Y., Phinn, S., Roelfsema, C., Kendoul, F., and Skocaj, D. (2019). Monitoring sugarcane growth response to varying nitrogen application rates: a comparison of UAV SLAM LIDAR and photogrammetry. *International Journal of Applied Earth Observation and Geoinformation*, 82, 101878.
<https://doi.org/10.1016/j.jag.2019.05.011>
- Stone, A. (2023). *A Methodology for Scaling Agricultural Biomass from Ground to Regional Scales using Remote Sensing and Machine Learning Analysis*. (Order No. 30567251) [Master's thesis, University of Georgia]. ProQuest Dissertations and Theses Global.
<https://www.proquest.com/dissertations-theses/methodology-scaling-agricultural-biomass-ground/docview/2859490096/se-2?accountid=14537>
- Strickland, T., Bosch, D. D., Endale, D. M., & Potter, T. L. (2016a). Gulf Atlantic Coastal Plain Long Term Agroecosystem Research site, Tifton, GA. *Headwaters to Estuaries: Advances in Watershed Science and Management*, 265.
https://www.srs.fs.usda.gov/pubs/gtr/gtr_srs211.pdf
- Strickland, T., Bosch, D., Hubbard, R., Lowrance, R., Potter, T., Reza, M., & Savabi, H., Truman, C. (2016b). Southeastern Coastal Plain LTAR Application: Little River Experimental Watershed, Southeast Watershed Research Laboratory, Tifton, GA, 22 p.

<https://www.ars.usda.gov/ARSTUserFiles/np211/Tifton SEWRL LTAR Proposal.pdf>

STEWARDS. (2025). *Stewards v4.0: Access to ars ceap benchmark watershed data*.

STEWARDS Version 4.0. <https://www.nrrig.mwa.ars.usda.gov/stewards/stewards.html>

Sullivan, D. G., Batten, H. L., Bosch, D. D., Sheridan, J. M., & Strickland, T. C. (2007). Little River Experimental Watershed, Tifton, Georgia, United States; a geographic database.

Water Resources Research, 43(9). <https://doi.org/10.1029/2006WR005836>

Tianhai Wang, Yadong Liu, Minghui Wang, Qing Fan, Hongkun Tian, Xi Qiao, & Yanzhou Li.

(2021). Applications of UAS in Crop Biomass Monitoring: A Review. *Frontiers in Plant Science*, 12. <https://doi.org/10.3389/fpls.2021.616689>

Tin Kam Ho. (1995) "Random decision forests," Proceedings of 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 1995, pp. 278-282 vol.1.

<https://ieeexplore.ieee.org/abstract/document/598994>

Tucker, C. J., Elgin, J. H., Jr., McMurtrey, J. E., III, & Fan, C. J. (1979). Monitoring corn and soybean crop development with hand-held radiometer spectral data. *Remote Sensing of Environment*, 8(3), 237–248. [https://doi.org/10.1016/0034-4257\(79\)90004-X](https://doi.org/10.1016/0034-4257(79)90004-X)

USDA. (2020). Little River Watershed Public Access Site. Public Data: USDA ARS.

<https://www.ars.usda.gov/southeast-area/tifton-ga/southeast-watershed-research/research/public-data/>

USDA. (2024a). Georgia Agricultural Facts. United States Department of Agriculture National Agricultural Statistics Service.

https://www.nass.usda.gov/Statistics_by_State/Georgia/Publications/More_Features/GA_AgFacts2022.pdf

- USDA. (2024b). *Gulf Atlantic Coastal Plain*. USDA Agricultural Research Service LTAR.
<https://ltar.ars.usda.gov/sites/gacp/>
- USDA. (2024c). USDA Cropland Data Layer. United States Department of Agriculture National Agricultural Statistics Survey.
https://www.nass.usda.gov/Research_and_Science/Cropland/Release/index.php
- USGCRP. (2023). *Fifth National Climate Assessment*. Crimmins, A.R., Avery, C.W., Easterling, D.R., Kunkel, K.E., Stewart, B.C. and Maycock, T.K., Eds. U.S. Global Change Research Program, Washington, DC, USA. <https://doi.org/10.7930/NCA5.2023>
- Vincini, M., Frazzi, E., & D'Alessio, P. (2008). A broad-band leaf chlorophyll vegetation index at the canopy scale. *Precision Agriculture*, 9(5), 303–319. DOI:
<https://doi.org/10.1007/S11119-008-9075-Z/FIGURES/7>
- Walter, J., Edwards, J., McDonald, G., & Kuchel, H. (2018). Photogrammetry for the estimation of wheat biomass and harvest index. *Field Crops Research*, 216, 165–174.
<https://doi.org/https://www.sciencedirect.com/science/journal/03784290>
- Wilke, B., Abendroth, L., VanderWulp, S. 2024. USDA LTAR Common Experiment measurement: Aboveground biomass. protocols.io
<https://dx.doi.org/10.17504/protocols.io.bp2l62zmkgqe/v1>
- Xu, W., Chen, P., Zhan, Y., Chen, S., Zhang, L., & Lan, Y. (2021). Cotton yield estimation model based on machine learning using time series UAV remote sensing data. *International Journal of Applied Earth Observation and Geoinformation*, 104.
<https://doi.org/10.1016/j.jag.2021.102511>
- Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, 17 p.

<https://doi.org/10.1155/2017/1353691>

Yang, S., Feng, Q., Liang, T., Liu, B., Zhang, W., and Xie, H. (2018). Modeling grassland above-ground biomass based on artificial neural network and remote sensing in the Three-River Headwaters Region. *Remote Sensing of Environment*, 204, 448-455.

<https://doi.org/10.1016/j.rse.2017.10.011>

Yiru Ma, Lulu Ma, Qiang Zhang, Changping Huang, Xiang Yi, Xiangyu Chen, Tongyu Hou, Xin Lv, & Ze Zhang. (2022). Cotton Yield Estimation Based on Vegetation Indices and Texture Features Derived From RGB Image. *Frontiers in Plant Science*, 13.

<https://doi.org/10.3389/fpls.2022.925986>

Zhang, K., Sekiyama, A., Okazawa, H., Yamazaki, Y., Hayashi, K., Tsuji, O., & Akimoto, M. (2023). Comparison of crop surface models and 3D point clouds by UAV imagery on estimating plant height and biomass volume of pasture grass. *International Journal of Environmental & Rural Development*, 13(2). <https://doi.org/https://iserd.net/ijerd132/13-2-21.pdf>

APPENDIX A

Correlation Matrices Among the Varying Vegetation Indices

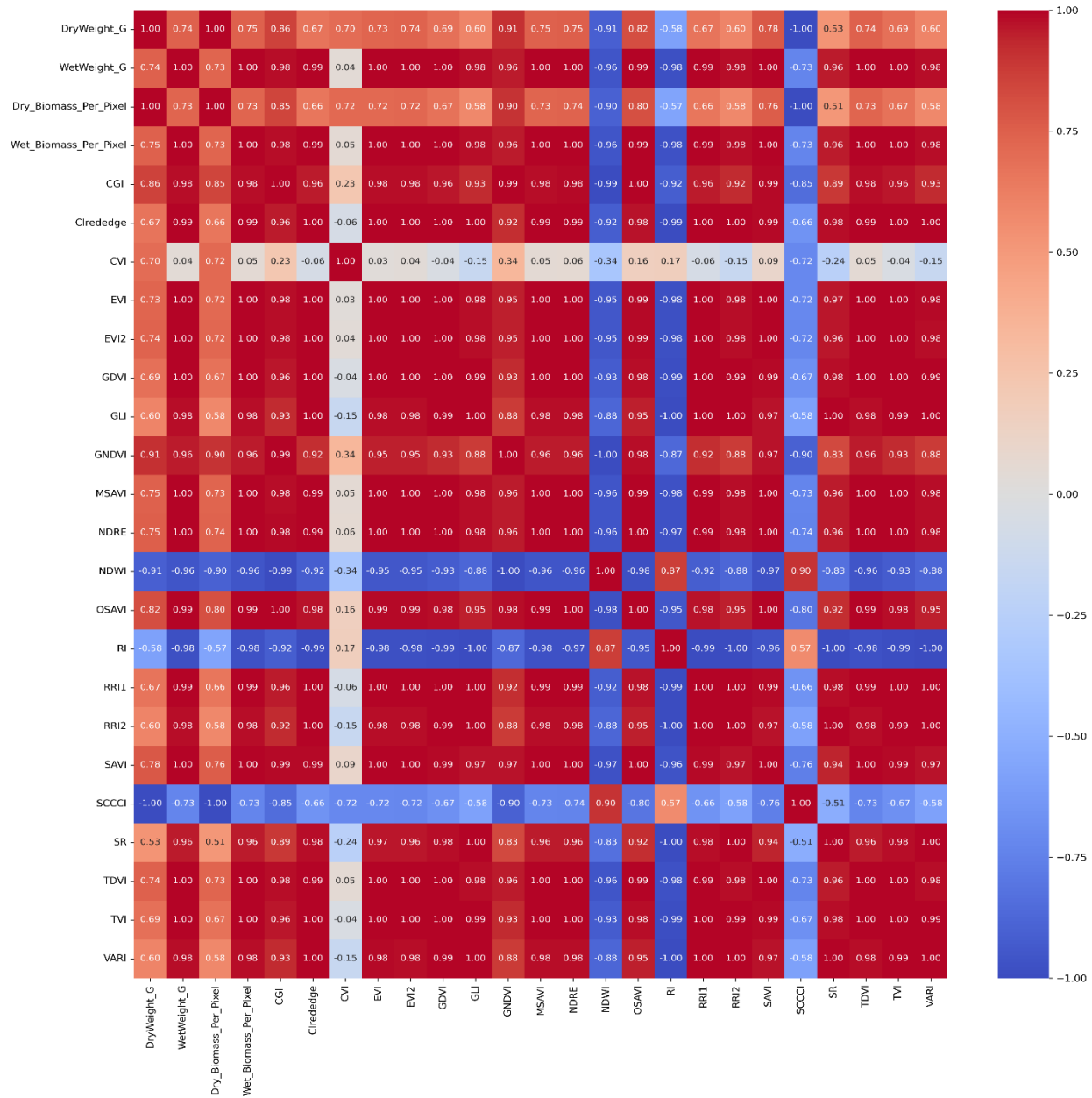


Figure A.1: Correlation matrix of the 21 VIs used for predicting cotton AGB at the ACF for 2018.

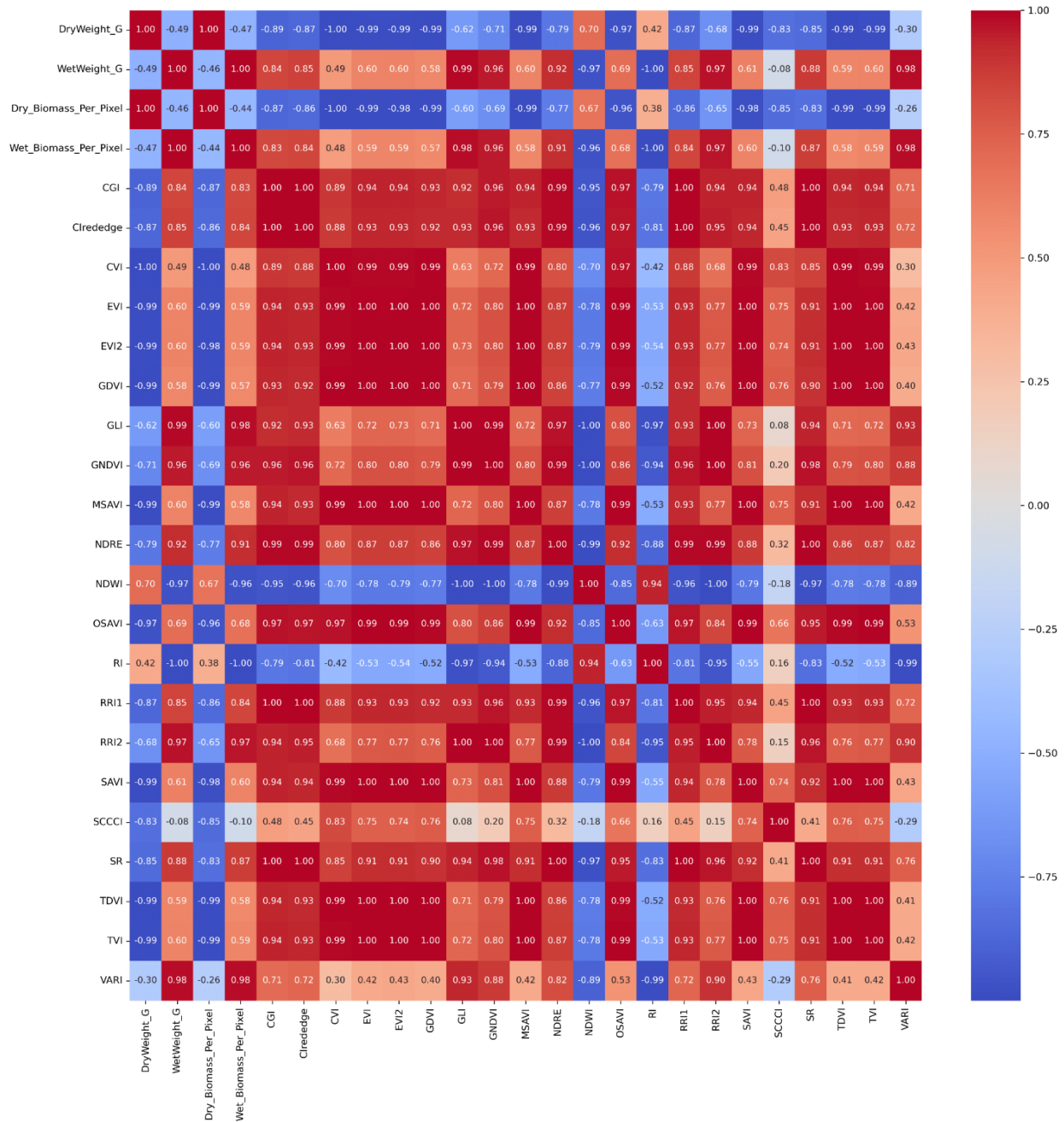


Figure A.2: Correlation matrix of the 21 VIs used for predicting cotton AGB at the ACF for 2019.

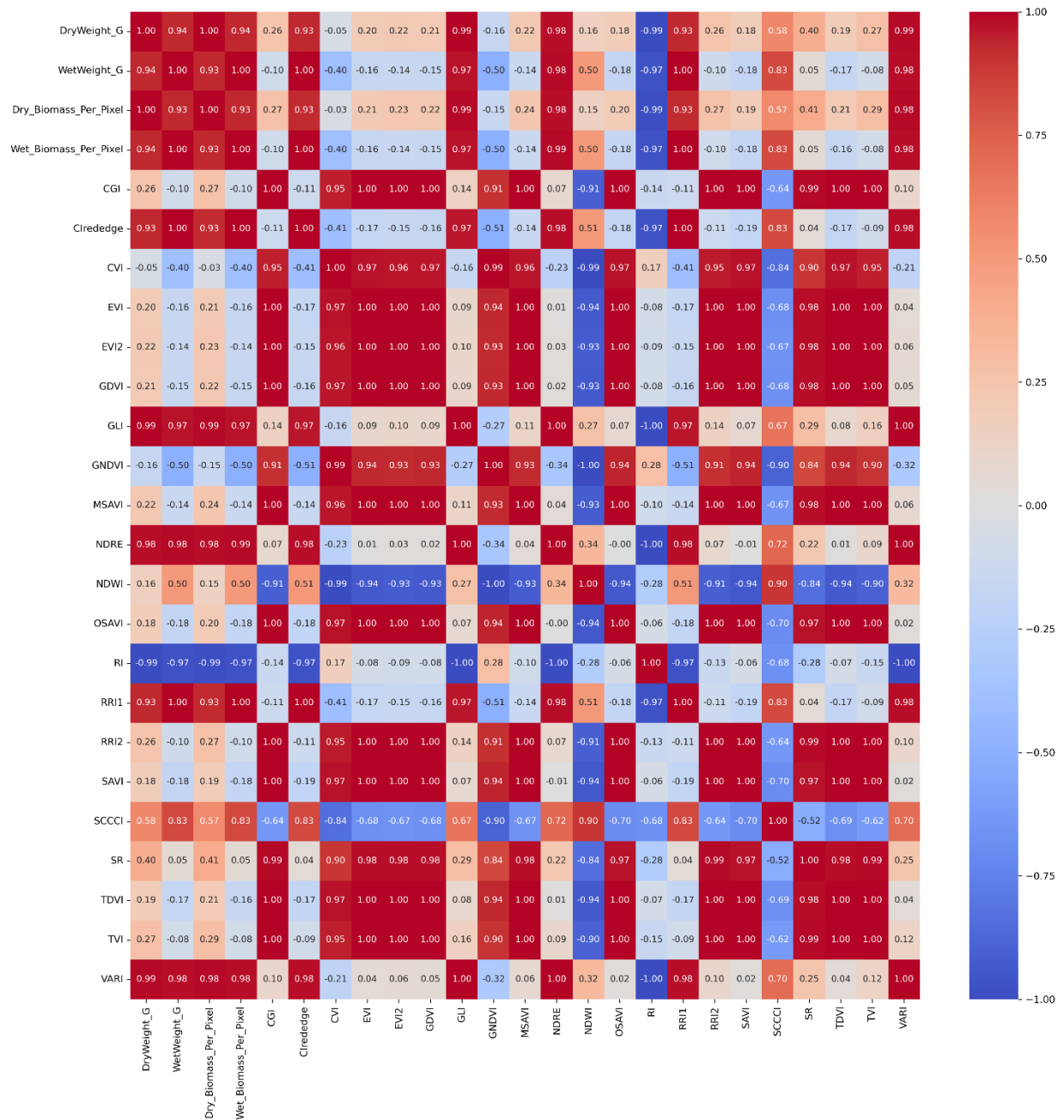


Figure A.3: Correlation matrix of the 21 VIs used for predicting cotton AGB at the TCF for 2018.

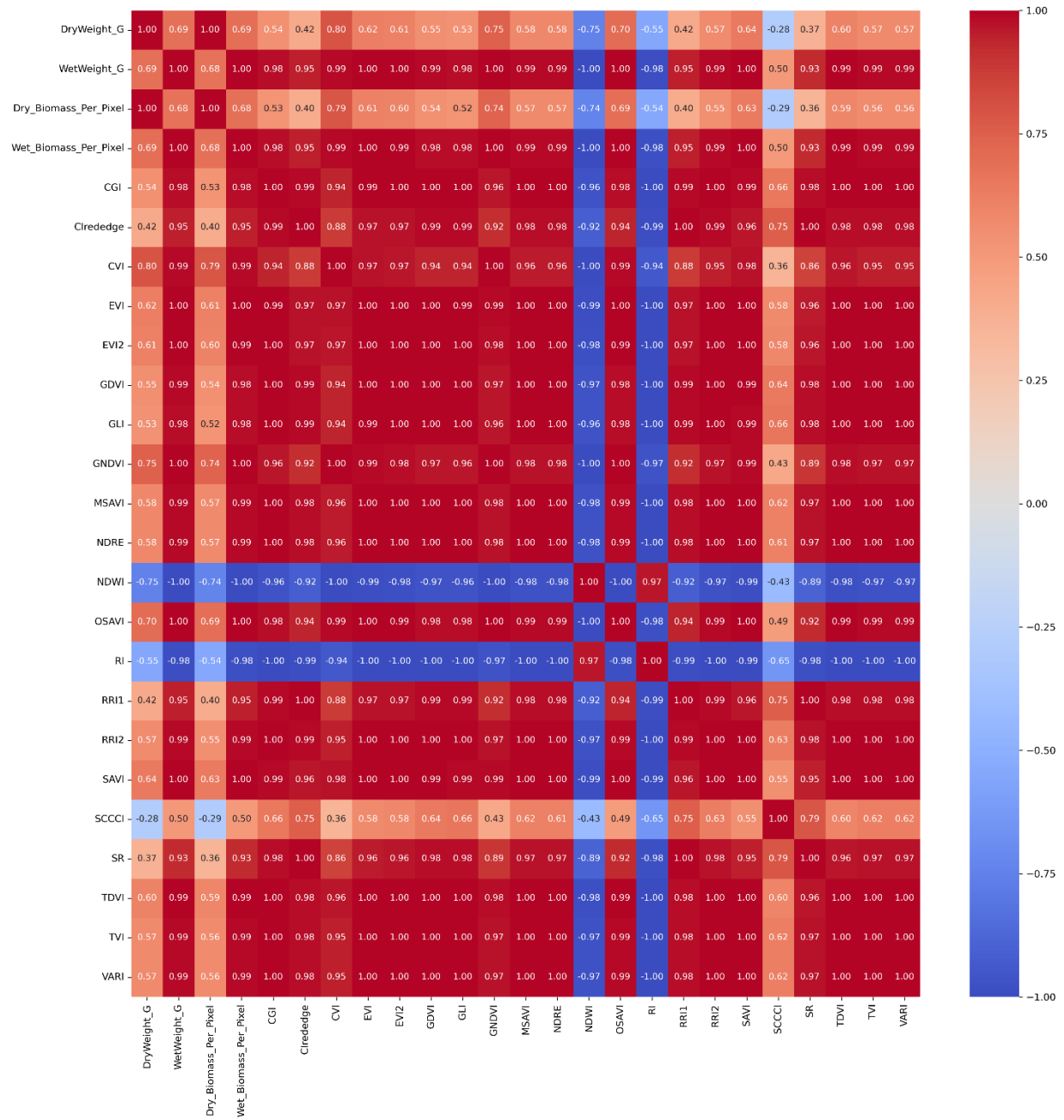


Figure A.4: Correlation matrix of the 21 VIs used for predicting cotton AGB at the TCF for 2019.

APPENDIX B

Predicted Biomass Raster Maps

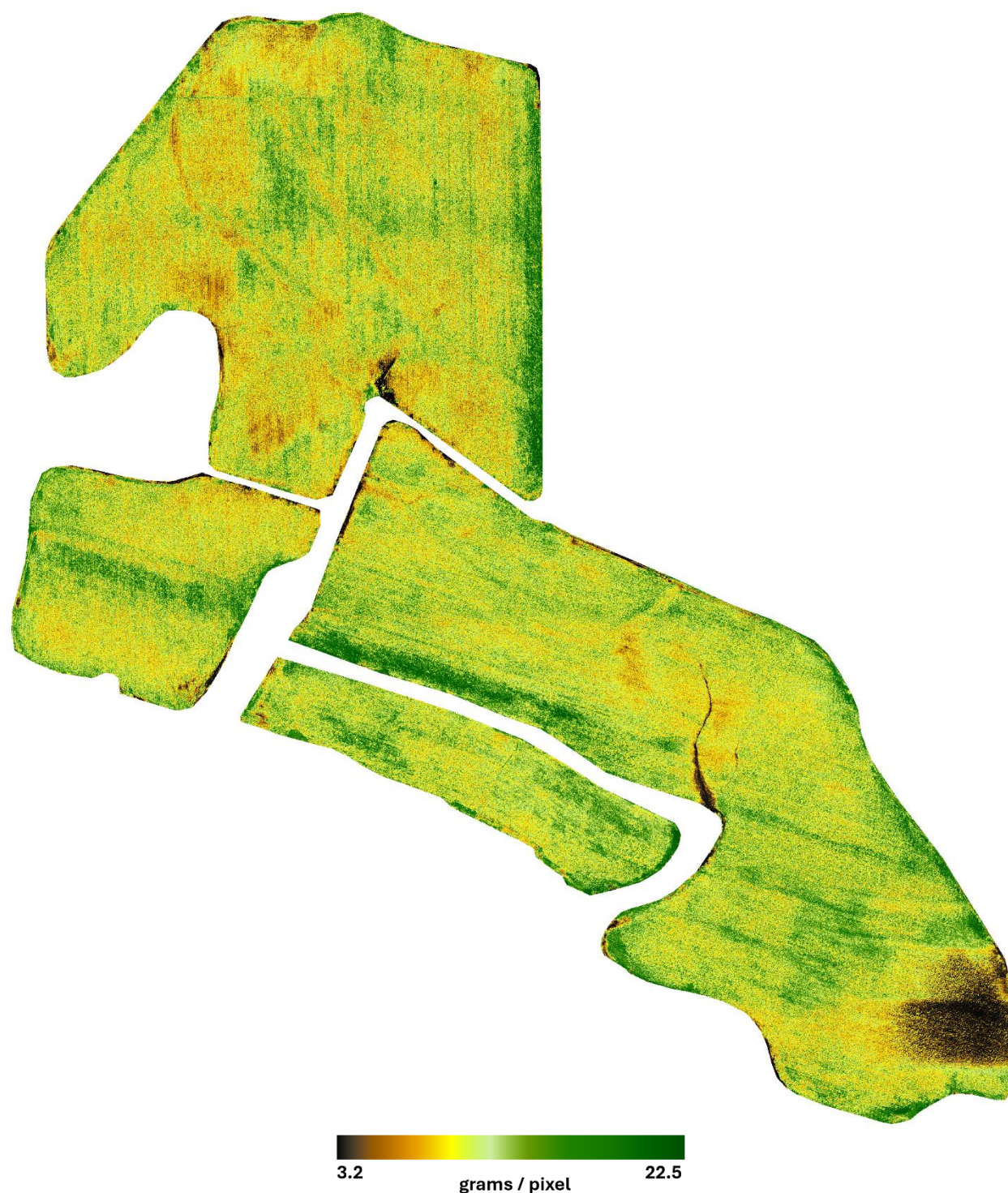


Figure B.1: Field-level predicted biomass for the 2018 ACF End of Season Random Field-Level Forest Predicted Wet Biomass model.

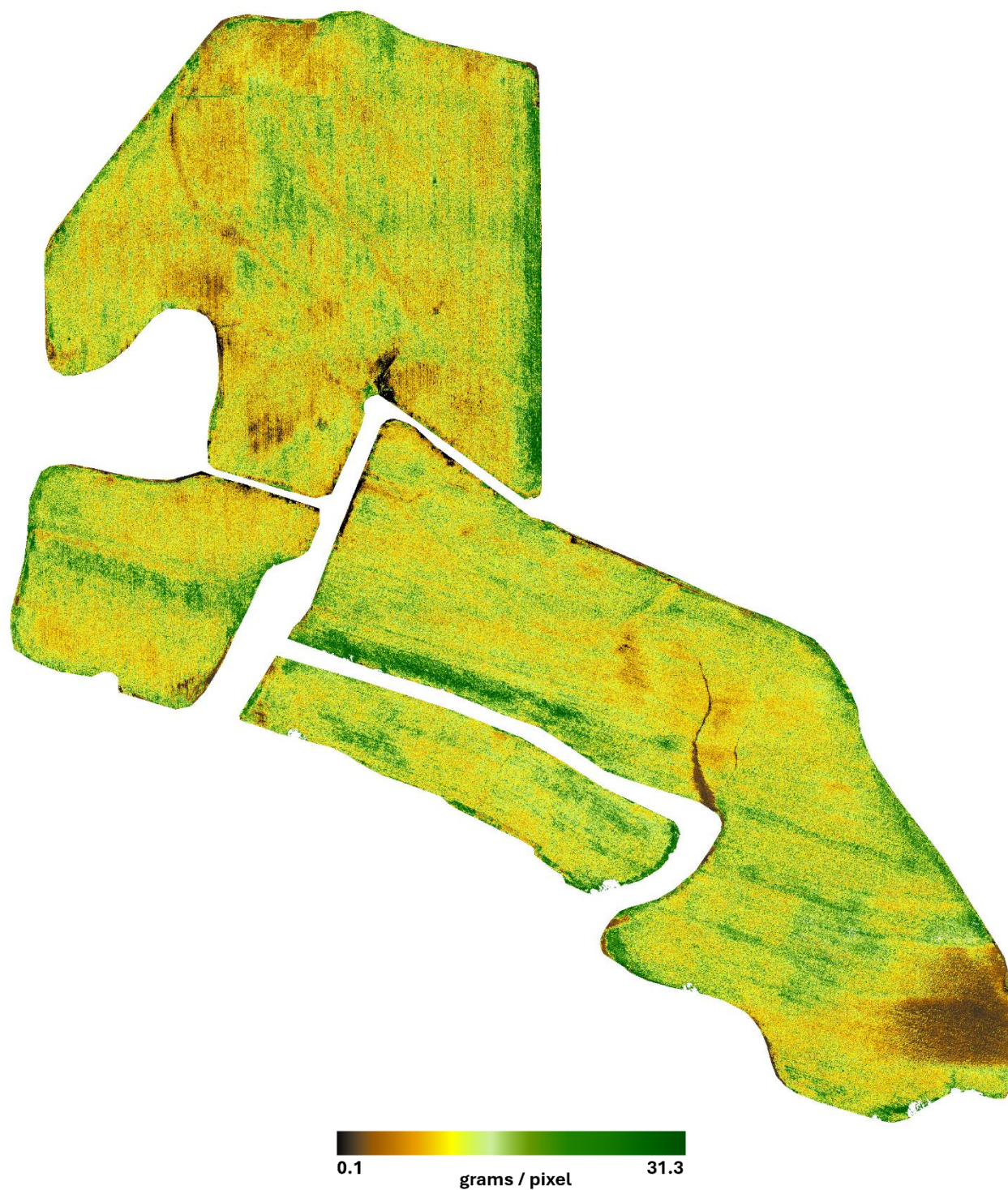


Figure B.2: Field-level predicted biomass for the 2018 ACF End of Season Field-Level XGBoost Predicted Wet Biomass model.

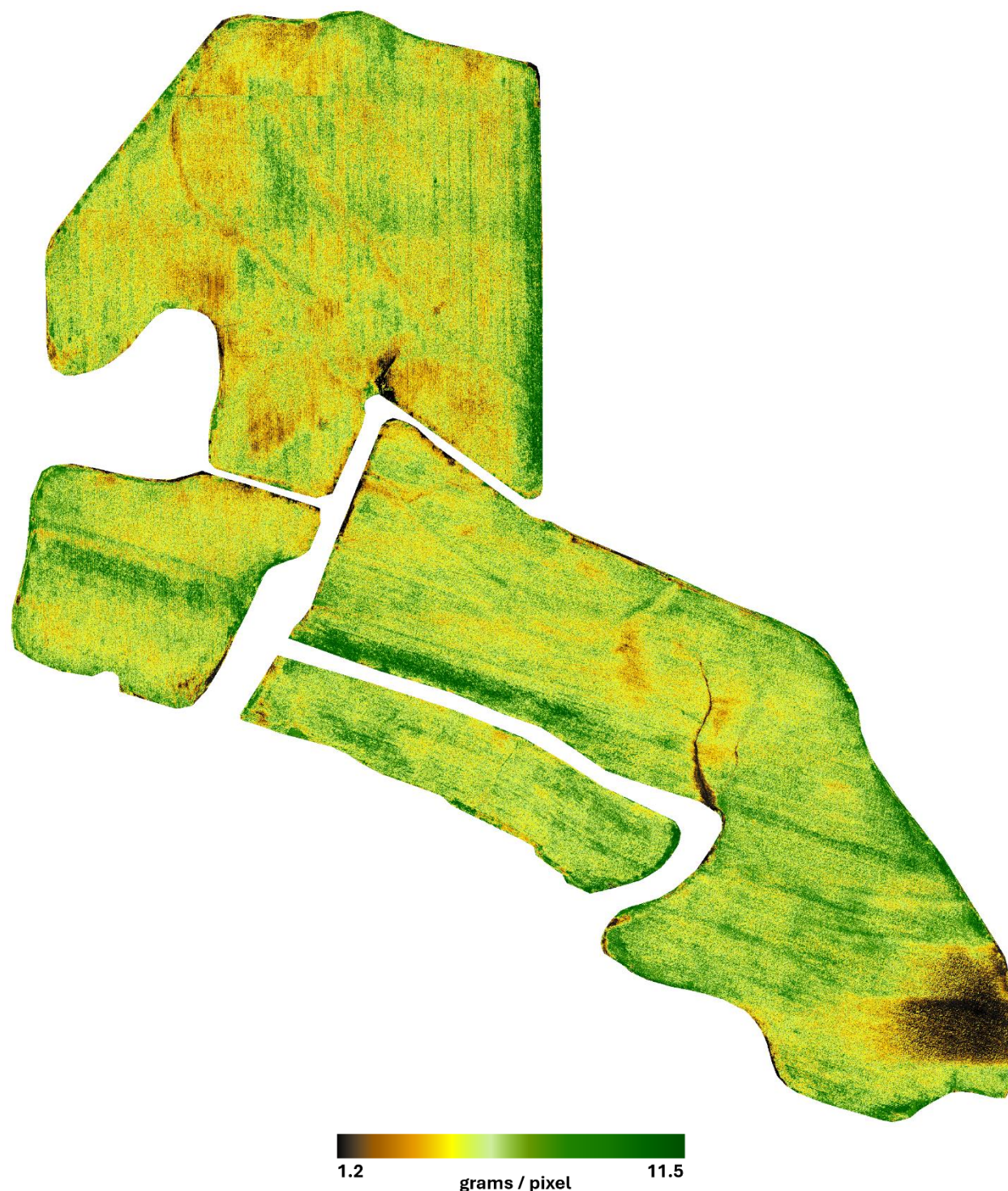


Figure B.3: Field-level predicted biomass for 2018 ACF End of Season Field-Level Random Forest Predicted Dry Biomass model.

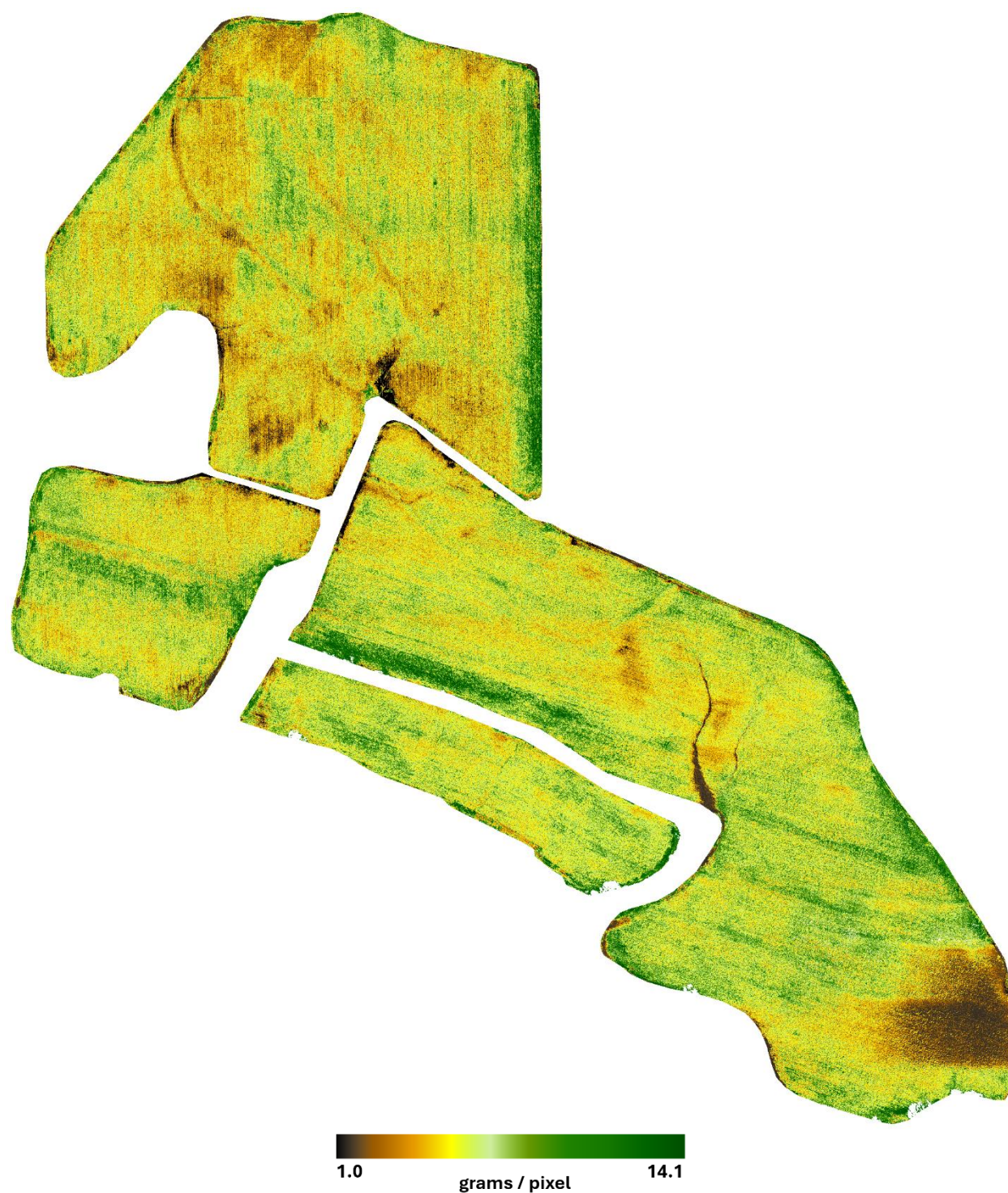


Figure B.4: Field-level predicted biomass for 2018 ACF End of Season Field-Level XGBoost Predicted Dry Biomass model.

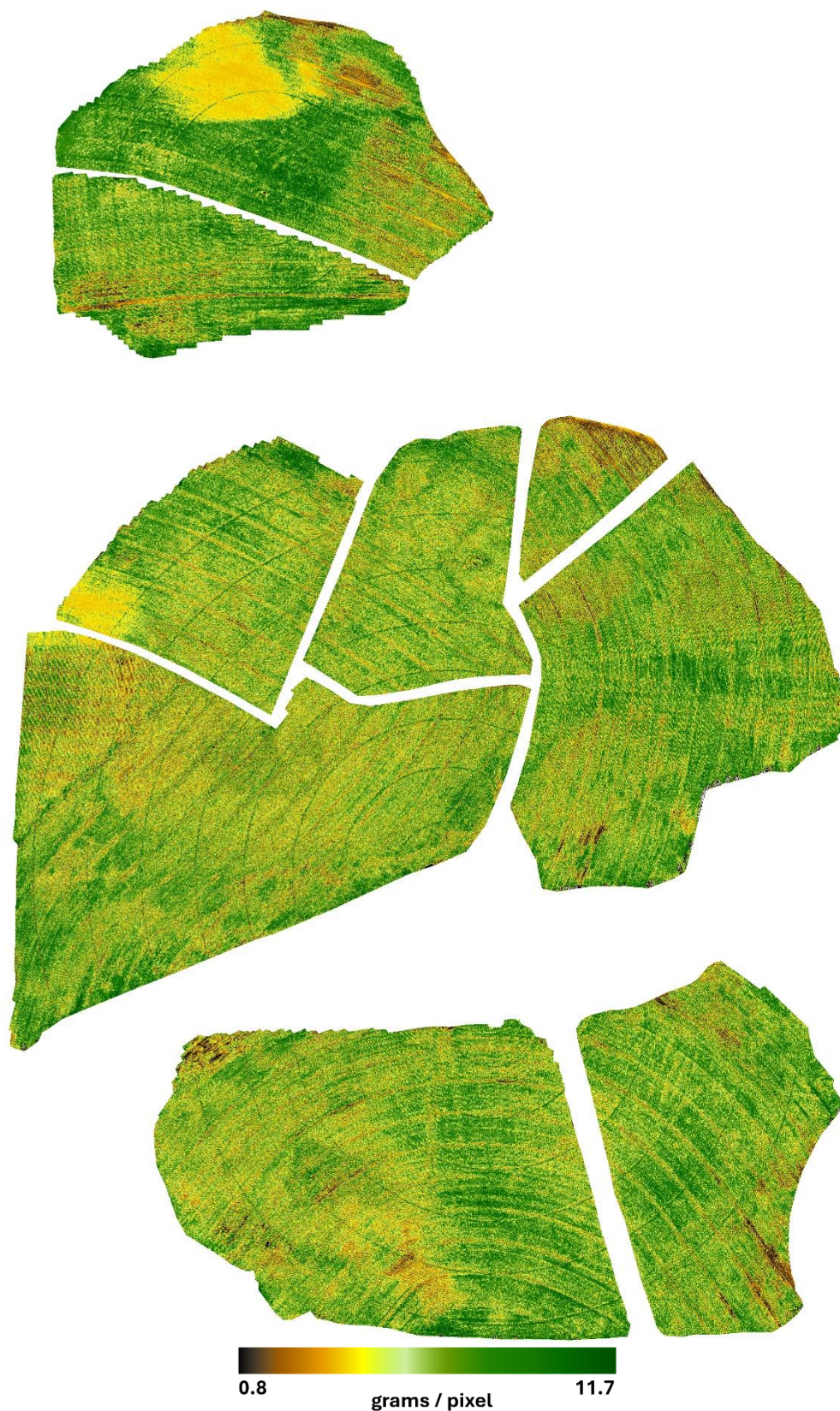


Figure B.5: Field-level predicted biomass for the 2018 TCF End of Season Field-Level Random Forest Predicted Dry Biomass model.

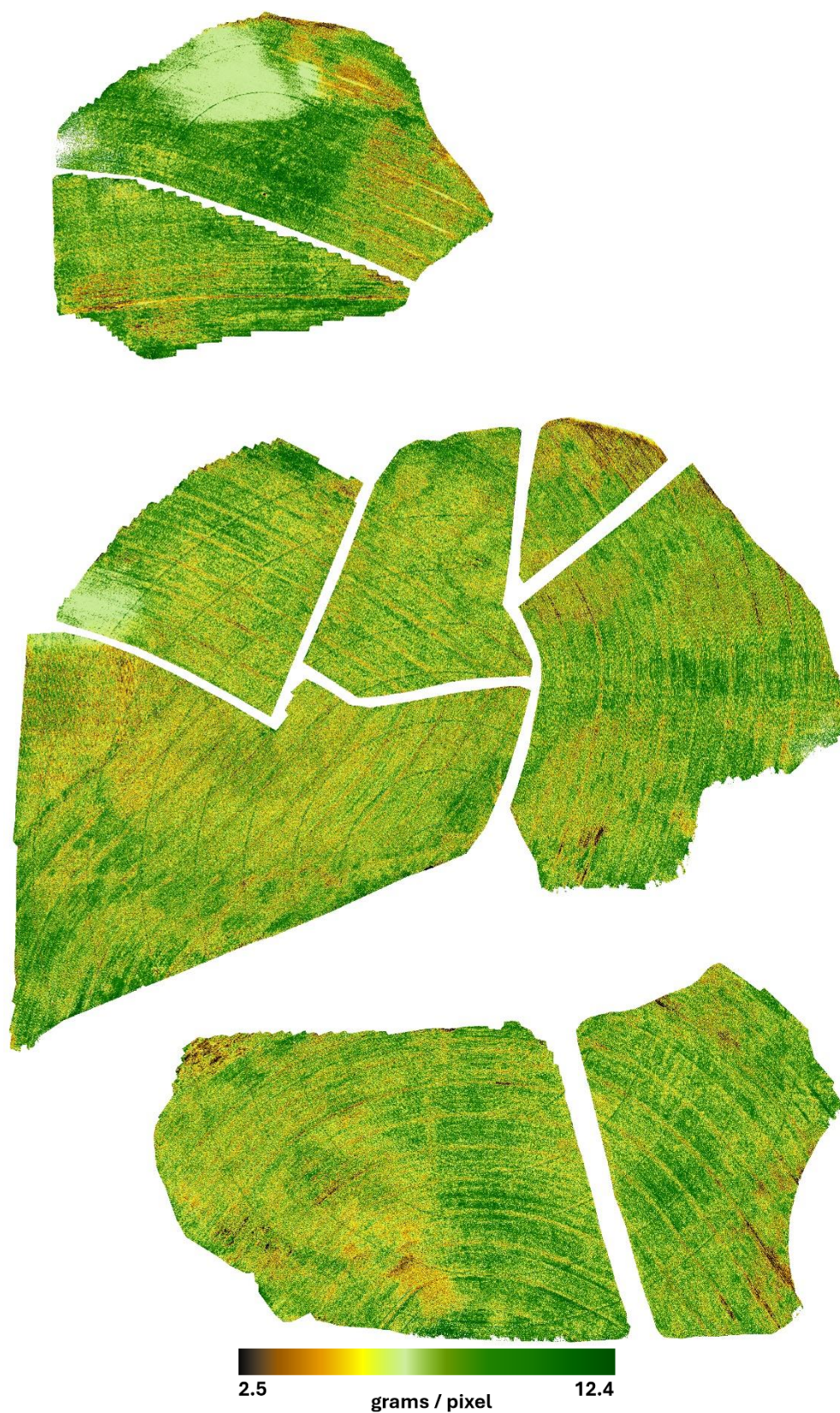


Figure B.6: Field-level predicted biomass for the 2018 TCF End of Season Field-Level XGBoost Predicted Dry Biomass model.

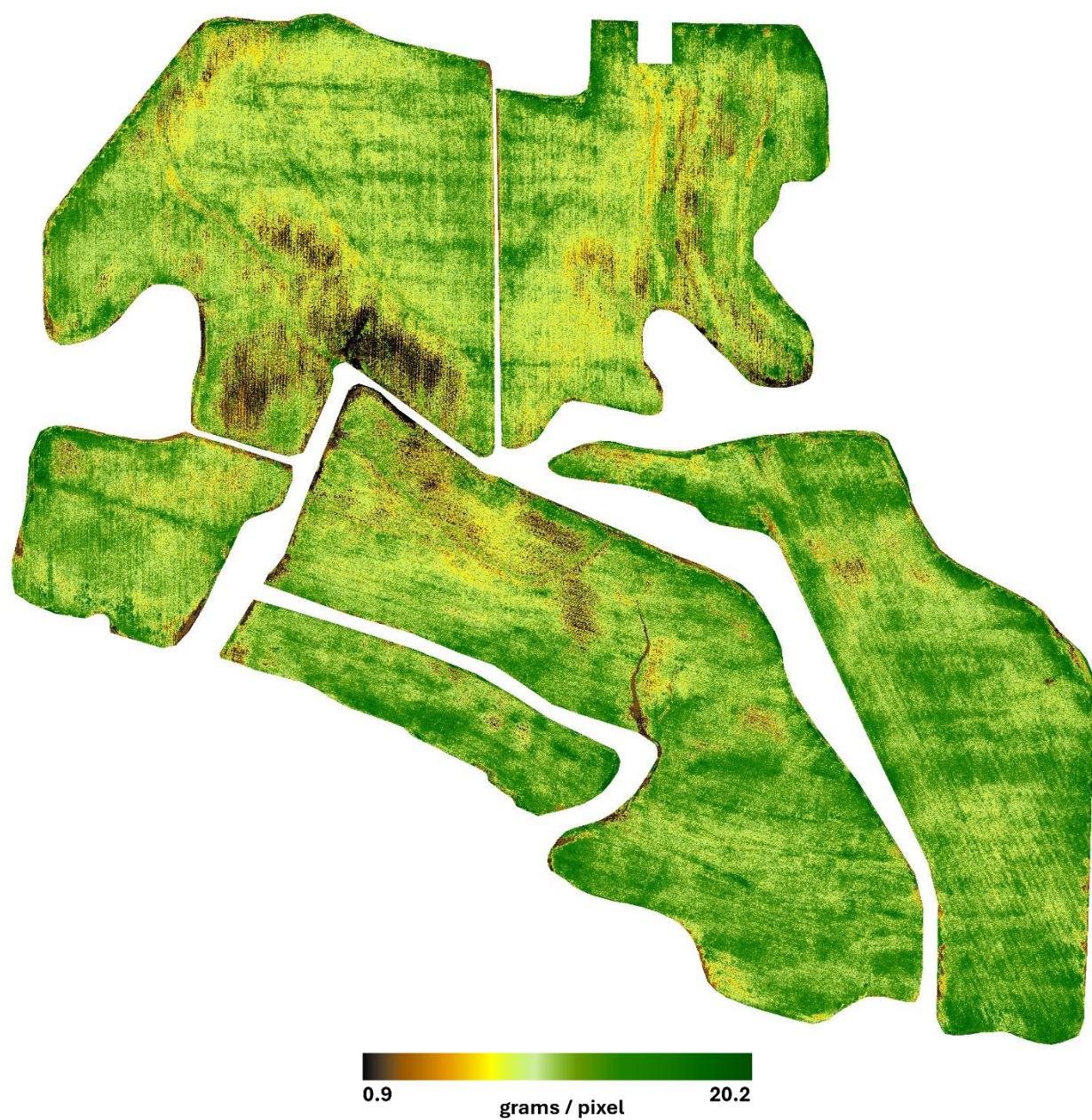


Figure B.7: Field-level predicted biomass for 2019 ACF End of Season Field-Level Random Forest Predicted Dry Biomass model.

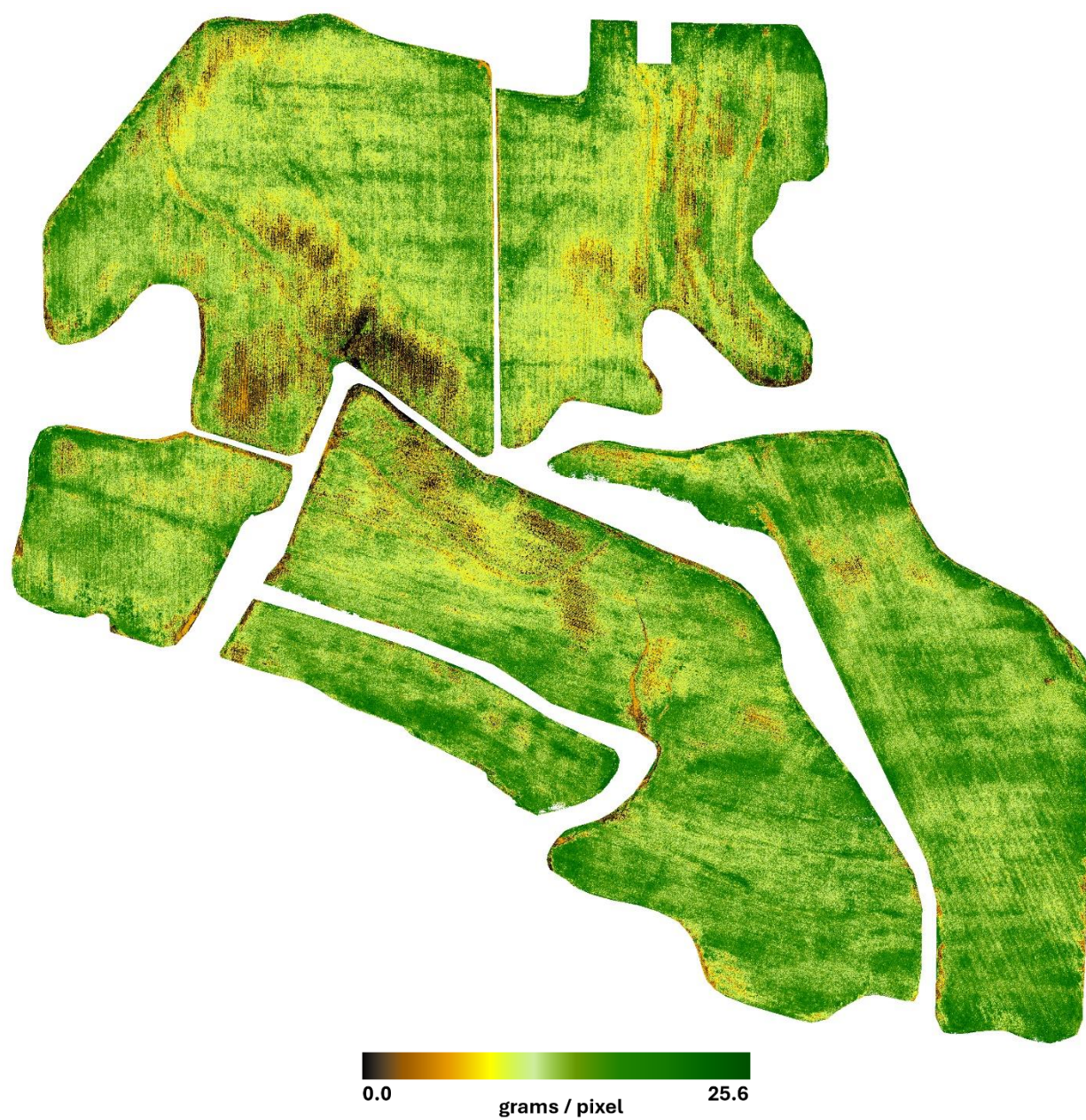


Figure B.8: Field-level predicted biomass for 2019 ACF End of Season Field-Level XGBoost Predicted Dry Biomass model.

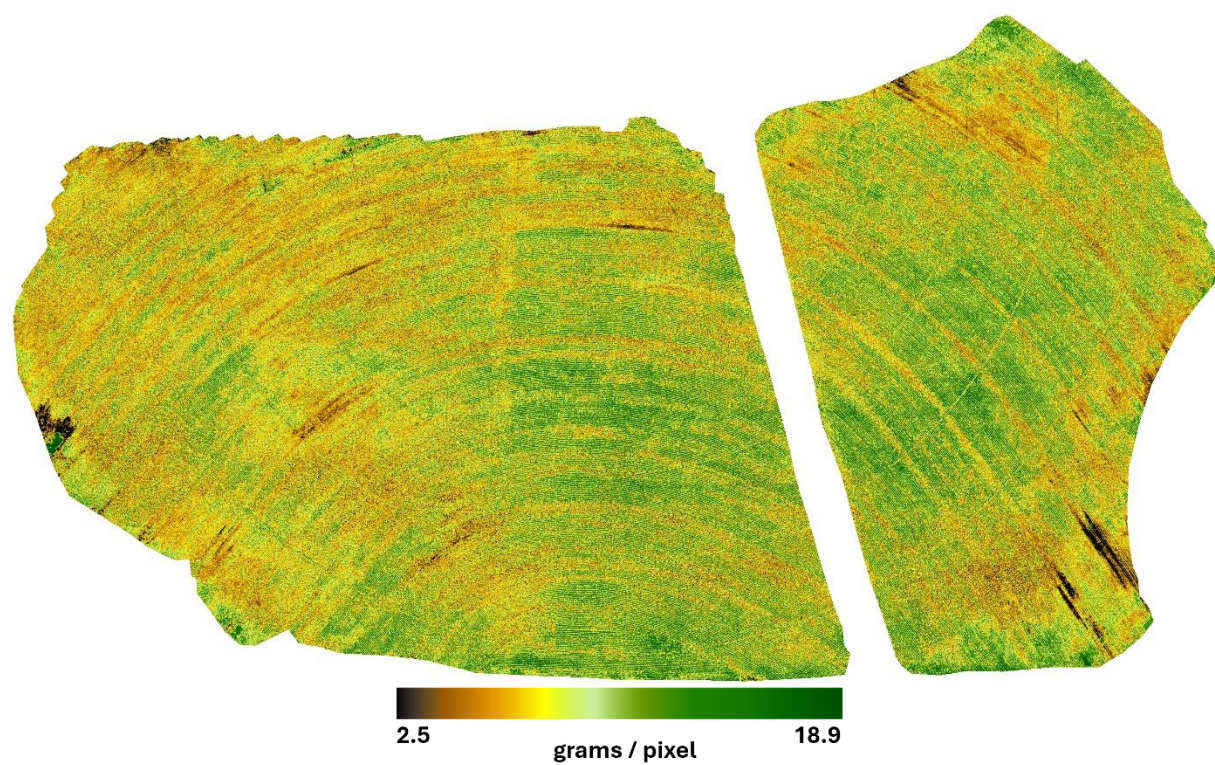


Figure B.9: Field-level predicted biomass for the 2019 TCF End of Season Field-Level Random Forest Predicted Dry Biomass model.

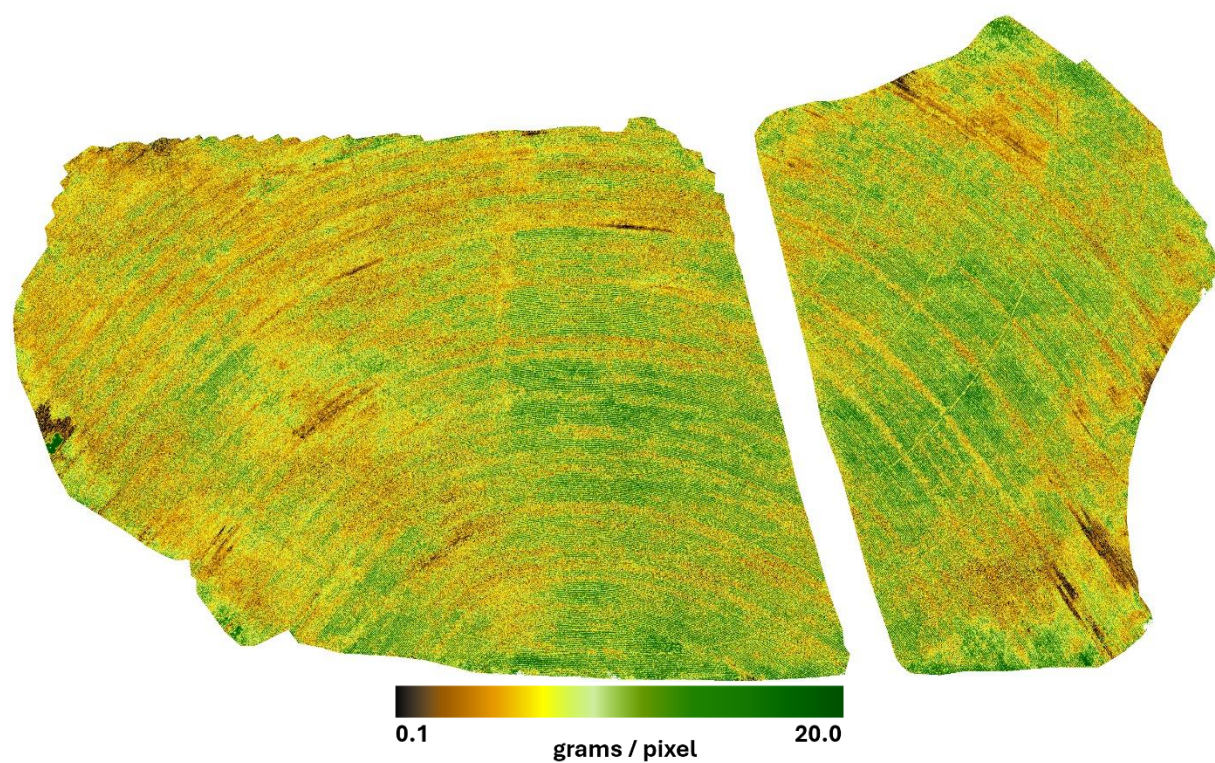


Figure B.10: Field-level predicted biomass for the 2019 TCF End of Season Field-Level XGBoost Predicted Dry Biomass model.