DISCRIMINATION OF MAIZE GENOTYPES THROUGH MULTI-TEMPORAL OBJECT-BASED REMOTE-SENSING CLASSIFICATION OF UNMANNED AIRCRAFT SYSTEM IMAGES

A Thesis

by

ANDREW CROCKETT WALTER

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

MASTER OF SCIENCE

Chair of Committee,	Anthony M. Filippi
Committee Members,	Muthukumar Bagavathiannan
	Michael Bishop
	Seth Murray
Head of Department,	David Cairns

May 2020

Major Subject: Geography

ABSTRACT

Remote-sensing classification in agriculture provides advanced techniques for expanding the capacity of high-throughput phenotyping in plant breeding and optimizing field management in precision agriculture. Methods applied in this study seek to advance remote sensing of plant breeding and precision agriculture through the classification of maize genotypes from unmanned aircraft system (UAS) images. Random forests (RF) and stochastic gradient boosting (SGB) algorithms were applied for classification of 12 maize genotypes at a row-scale analysis. Classification was achieved through the combined use of object-based image analysis (OBIA) and multi-temporal image analysis. The classification utilizes "object properties," or variables (image layer statistics, object texture, object geometry, Structure from Motion (SfM) derived height measurements, and time-series measurements), to discriminate among genotypes. Object variables are evaluated for discriminative capacity in a maize-genotype context. Classification results with an accuracy of 81.25% supported the discrimination of 12 maize genotypes using only RGB images. This study further supports the use of SGB over RF in small class number classification, but RF in assessments of a larger number of classes. Multi-temporal dimensionality proved beneficial for classification; image and variable stack SGB classification produced results with an increase in accuracy of 30% and greater over single-date classification. Finally, study results identified the most discriminative object variables, such as mean canopy height model (Mean-CHM) and FRAGSTATS Cohesion metric, for classification of maize genotypes with the objective that future research will make use of derived variables and their association with plant characteristics for selection of optimal varieties of maize.

ACKNOWLEDGEMENTS

I would like to express my appreciation to my committee chair, Dr. Anthony M. Filippi for all his support over the last two years. I am thankful for his guidance throughout this research and the many ideas he contributed to the work conducted. Thank you for always letting me drop by your office unannounced. Thank you also for your patience as “short meetings” that usually turned into conference calls.

I would also like to thank my committee members Dr. Bishop, Dr. Murray, and Dr. Bagavathiannan for their assistance. Thank you for your willingness to always answer questions, no matter how silly they might be, and to provide support from your own research and experience.

Thank you to Dr. Zhaohui Chi, Nathália Cruzato, Da Huo, and Dr. Steven Anderson for helping me understand the data analyzed in this study. Much of the research conducted in this thesis built upon data and hard work you contributed. I very much appreciate each of your help as this research could not have been done without your support. I thoroughly enjoyed working with each of you and learned so much through our collaboration.

To Billy Hales, thank you for your assistance in automating portions of the data analysis performed in this study. Your kindness helped to save countless hours of tedious clicking.

Xiao Lee, Xunwei Xu, and Cesar Castillo thank you for your friendship and always letting me ask questions, bounce ideas off you, verbally process research, and learn how to be a graduate student.

Dr. Lyle, in many ways, I feel like I have received two graduate degrees from everything I have learned from working with you the last three years. I am very thankful to have worked with you as a graduate teaching assistant. Thank you for taking time to invest in people. Many of the

photogrammetric and geodesy concepts I have learned from you have indirectly influenced the research of this thesis, but even more so, I have no doubt they will have an even greater impact on the research and career I explore in the future.

I would like to honor my parents Mike and Cathy and also my two sisters Cate and Rebecca. Thank you for all your encouragement and your confidence in me. I am convinced you believed in me and the dreams that fueled this research more than I have, thank you. Thank you for your love and understanding, for your listening ear and caring. This research would have been so hard without you. I am so grateful to have each you.

Kathryn Edmunds, I never anticipated the depth of friendship that could come from the adventure of pursuing a graduate degree. I am so glad that meeting you was one part of this journey. Thank you for your constant affirmation and for cheering me on this last year.

Finally, I would like to honor God. I acknowledge that every thought of inspiration, will to press through, and ability to work in excellence is through his Spirit and constant measures of grace. Thank you, God, for letting me come to Texas A&M and for planting in my heart the seeds that produced this research. May God be glorified through this thesis.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Associate Professor Anthony Filippi and Professor Michael Bishop of the Department of Geography, and Associate Professor Muthukumar Bagavathiannan and Professor Seth Murray of the Department of Soil and Crop Sciences.

The data analyzed in Chapter 3, including image mosaics; plot boundaries; and fieldnotes, were provided by Professor Seth Murray. Additional data analyzed in Chapter 3, wavelet data, was provided by Professor Michael Bishop.

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

Funding Sources

A portion of this work was funded by the Iowa Corn Promotion Board.

NOMENCLATURE

CHM	Canopy Height Model
DSM	Digital Surface Model
ExG	Excess Green
ExR	Excess Red
FAO	Food and Agriculture Organization of the United States
GEOBIA	Geographic Object-Based Image Analysis
LAI	Leaf Area Index
OBIA	Object-Based Image Analysis
RF	Random forests
SGB	Stochastic gradient boosting
RGB	Red Green Blue
SfM	Structure from Motion
SSWM	Site Specific Weed Management
SVM	Support Vector Machines
UAS	Unmanned Aircraft Systems
VEG	Vegetativen
V-Group	Variable Group
VI	Vegetation Index
YI	Yellow Index

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

CHAPTER 1

INTRODUCTION

1.1 Background

Maize (Zea mays L.) is the highest produced cereal crop by production in the world and is a valuable staple crop for many countries. The United States is the leading global producer of maize with maize being the United States' most widely produced crop (FAO 2009; NASS/USDA 2019). As world population is expected to increase to 8.3 billion people by 2050, agricultural production must rise to meet this demand. Estimates predict that crop production must increase by 20% to 75% to meet global demands (FAO 2009; Hunter et al. 2017). In addition, dependence on maize is remarkably high in many developing countries throughout Africa and the Americas, e.g. maize is estimated as ~36% of dietary energy for the country of Guatemala (Ray et al. 2013). According to Hyman et al. (2008), "one-third of all malnourished children are found in systems where maize is among the top three crops." Development of more productive maize varieties and farming practices is necessary to meet both current demand among developing countries and future population growth.

Remote-sensing platforms are increasingly being integrated into plant breeding research to help scientist identify genotypes with valuable traits such as higher yield, drought tolerance, lodging resistance, etc. High-throughput phenotyping is being used as means to quickly select genotypes by identifying physiological features of plants associated with beneficial traits or plant characteristics. Development of unmanned aerial systems (UAS) are also providing new opportunities for plant breeders to collect measurements with increased spatial and temporal resolution than previous methods, e.g. manual in field measurements or satellite platforms.

Application of high-throughput phenotyping and UAS image analysis can theoretically increase the speed, precision, and capacity at which plant breeders can identify improved hybrid maize breeds, if analytical tools are available and reliable. Plant breeders are using phenotypic features such as plant height, leaf area index (LAI), and plant health linked to desired traits to help quickly select unique genotypes for further assessment (Haghighattalab et al. 2016; Anderson et al. 2019). In addition to measuring plant phenotypic properties, remote-sensing classification is also being used to identify unique genotypes (Tu et al. 2018). Classification uses machine-learning algorithms to provide plant breeders with a reduced list of most discriminative genotypes for further evaluation. Classification also provides information as to which image variables generated and used as inputs are most valuable for describing a genotype. Variables derived from image analysis can further be assessed for descriptive information of genotypes and correlation to valuable plant traits.

1.2 Objectives

This thesis presents methodology for multi-temporal object-based high-spatial-resolution remote-sensing classification of agriculture. Different from previous remote-sensing studies performed on agriculture, a concept for remote-sensing classification was proposed, which provided multi-temporal object properties for classification of image objects at a subplot-scale. This thesis also differs from previous studies as genotype classification was performed exclusively using RGB data and not hyperspectral. Object properties (e.g., spectral and canopy height model (CHM) statistics, texture, geometry, and wavelet analysis results) derived from 18 image-acquisition dates and image stacking provided descriptive information as to the changing biophysical characteristics of a plant’s canopy throughout a growing season. This assisted classification by providing a larger feature space on which each algorithm was trained while also providing variables and classification which took into consideration plant development. Temporal information along with innovative use of CHM, wavelet, and Yellow Index assisted in classification in the absence of higher spectral resolution data. Classification of crop genotypes and the application of a semi-automated object-based image analysis (OBIA) workflow developed by this study was proposed to increase plant breeding and precision-agriculture efficiency by introducing improved methods for monitoring and selecting optimal maize genotypes. In addition, measures identified through the proposed remote-sensing classification provide a target for further research. The following research objectives sought to provide additional information for plant breeding of maize by achieving the following goals:

1. Identify which machine-learning algorithm provided the most accurate genotype classifications.

2. Identify the earliest point in the growing season that accurate maize genotype classification was possible.

3. Identify the most important variables for discriminating among maize genotypes.

CHAPTER 2

LITERATURE REVIEW

Remote-sensing classification has been used on a variety of different applications within agriculture from the discrimination of crops and weeds to tea tree cultivars (Pérez-Ortiz et al. 2016; Tu et al. 2018). Assessment of phenotypes (physiological properties of plants) via remote-sensing techniques has been conducted without classification in a large number of studies and using a wide range of sensors including RGB, multi-and hyperspectral, thermal, and LiDAR (Jay et al. 2015; Zaman-Allah et al. 2015; Haghighattalab et al. 2016; Vergara-Díaz et al. 2016; Malambo et al. 2018). However, studies focused specifically on the classification of crop cultivars are limited and methods applied very similar. Classification within such studies have primarily been limited to spectroradiometer data and to a lesser degree hyperspectral images with low to medium resolution derived from aerial or satellite platforms (Rao, Garg and Ghosh 2007; Shahi et al. 2007; Basso et al. 2010; Zhao et al. 2010; Das et al. 2018). Within studies using hyperspectral information for genotype classification, the number of classes between which discrimination occurred was limited to a range of 3 to 14 genotypes (Rao, Garg and Ghosh 2007; Das et al. 2018). Results with accuracy of 95-98% suggest classification of genotypes through hyperspectral analysis is possible and classification at the genotype level should be explored using alternative methods (Das et al. 2018; Tu et al. 2018). Similarly, spectroradiometer analysis and feature extraction of genotype characteristics was performed on sugarcane and wheat fields of 87 to 384 genotypes (a sample size better representing a full plant breeding scenario), but within these studies, classification was not performed which could have helped identify highly discriminative genotypes and the variables acting as the greatest contributor to distinguishing between genotypes (Zhao et al. 2012; Garriga et al. 2017). Within the last year, only one study used a UAS platform for hyperspectral classification of cultivars; however, this study focused on tea tree cultivar (*Camellia sinensis L.*) discrimination as opposed to annual crops. Eight cultivars were classified using support vector machines (SVM) and achieved an overall accuracy of 96.2% (Tu et al. 2018). Also similar to research conducted at the genotype level, classification of mangrove species was studied using UAS derived hyperspectral and 3-demensional digital surface model (DSM) data. Within this study, object-based image analysis (OBIA) of UAS images was performed to extract additional measures from image regions beyond spectral and height information inferred from individual pixels (Cao et al. 2018).

Previous studies in the literature have been concentrated on genotypic classification using spectroradiometer and low spatial resolution hyperspectral images. A significant gap in the literature remains for the classification of genotypes using UAS remote-sensing, RGB images, and height information. Studies have assessed maize genotypes using UAS, variables extracted from RGB bands, and height information from Structure from Motion (SfM) point clouds as are proposed in this study; however, though genotypes were selected in these studies (often through quantitative phenotypes), classification was not used to discriminate between them (Geipel, Link, and Claupein 2014; Zaman-Allah et al. 2015; Li et al. 2016; Malambo et al. 2018). Furthermore, multi-temporal OBIA has yet to be used for classification of annual crops using UAS-derived images. Limited work has applied OBIA for analyzing crops in a plant-breeding assessment (Hall et al. 2018). Three gaps identified in this literature review include 1) classification of annual crops using multi-temporal OBIA variables derived from UAS images; 2) classification of genotypes using images with only common red, green and blue (RGB) reflectance bands, rather than hyperspectral; and 3) use of SfM derived 3D information for classification of annual crops.

Object-based image analysis (OBIA), also referred to as geographic object-based image analysis (GEOBIA), works by grouping adjacent pixels within an image into meaningful image regions known as ‘image objects’ based on the homogeneity of pixel values in a multidimensional feature space (Liu et al. 2006; Blaschke 2010; Blaschke et al. 2014). OBIA is especially beneficial for very high-resolution image analysis as image objects provide the minimum unit for classification and are able to account for spectral variability and shadowing within plants (Torres-Sánchez, López-Granados, and Peña 2015,). Image objects provide additional information such as texture, geometry, and spatial context for classification beyond pixel-based analysis which is needed for complex vegetation classes. OBIA operates as a workflow of operations with segmentation being a core but not the only step for creating meaningful image objects (Hussain et al. 2013). Segments or image regions created by segmentation are optimized by merging, labeling, and feature extraction to produce image objects that more comprehensively describe targets within an image (ex. rows of maize, plants, leaves, etc.) (Blaschke 2010; Benz et al. 2014; Blaschke et al. 2014).

The multidimensional feature space used to create image objects can be comprised of multiple images layers or bands for a single point in time or from multiple points in time; therefore, image objects can, but do not have to, describe homogeneous regions within a single image or point in time but can be multi-temporal in nature (Hussain et al. 2013; Song et al. 2019). This suggest that image objects can describe a multi-temporal phenomenon and OBIA feature extraction can produce multi-temporal variables and summary statistics from multiple layers, images, and dates. After reviewing the literature, we were not able to identify any studies utilizing temporal object-based properties or variables for sub-field classification of agricultural UAS images. Multi-date assessment of maize has been conducted using UAS images, but inclusion of temporal information for classification was not used (Torres-Sánchez et al. 2014; Vega et al. 2015; Malambo et al. 2018). Multi-temporal classification up to this time has occurred almost entirely at the field and multi-field scales in agriculture due to the limited spatial resolution of satellite and aerial images (Petitjean et al. 2012; Moody et al. 2017; Peña, Liao, and Brenning 2017; Xiao et al. 2018). Classification of satellite images has made use of image stack techniques, multi-temporal vegetation indices, and OBIA methods as were adapted in this study (Petitjean et al. 2012; Osmólska and Hawrylo 2018; Song et al. 2019). In the area of very high-resolution image analysis, much work has been conducted recently on crop species and crop-weed classification with UAS platforms, but a clear void remains in the multi-temporal classification of agriculture at the sub-field scale (López-Granados 2011; Peña et al. 2013; Torres-Sánchez, López-Granados, and Peña 2015; David, Ballado, and Ieee 2016; Lopez-Granados et al. 2016; Perez-Ortiz et al. 2016; de Castro et al. 2018).

Most comparable to the methods proposed for this study is the object-based classification of crop and weeds using UAS images. A significant volume of literature has sought to address the classification of crop and weed species for site specific weed management (SSWM) (Peña et al. 2013; Torres-Sánchez, López-Granados, and Peña 2015; Lopez-Granados et al. 2016; de Castro et al. 2018). Studies have performed OBIA classification on the limited information of RGB and multispectral images; such research has produced promising results for between specie discrimination, specifically in the context of crop and weeds (Pérez-Ortiz et al. 2016). Literature on crop-weed classification have demonstrated the benefit of OBIA methods for high resolution and plant-level classification of agriculture with RGB images (Huang et al. 2018). Yet, some object-based methods which have proven very valuable for crop-weed classification, such as plant location in relation to crop-row objects, do not provide beneficial information for discrimination of genotypes. The location of an image object representing a plant within or outside of a crop row, a powerful variable used to identify weeds that occur between crop rows and highly contributing to classification accuracy in such studies, is not helpful in a genotype scenario as voluntary plant growth is purposefully eliminated (Perez-Ortiz et al. 2016). However, even within the broader field of crop-weed classification, no attempt has been made to use multi-temporal object properties for classification.

In addition to crop-weed classification, OBIA classification of UAS images was applied for mangrove species discrimination (Cao et al. 2018); however, application of OBIA for phenotype and genotype classification in a non-multi-temporal context is still under explored. Finally, we acknowledge mathematical morphology and object-based classification have been applied to RGB images of potato and banana cultivars, however in such articles classification was not performed on images of the plant (Przybyl et al. 2015; Dittakan et al. 2017). Classification in these studies was performed on harvested fruit and within a laboratory setting.

As of to date, no literature has been found on the classification of genotypes using solely optical or RGB UAS images. Multiple studies have used UAS to assess genotypes of cereal crops and derive phenotypic measurements such as plant height, leaf area index, and plant health from RGB images (Zaman-Allah et al. 2015, Haghighattalab et al. 2016; Vergara-Díaz et al. 2016; Anderson et al. 2019). Research as in Geipel et al. (2014) and Anderson et al. (2019) have assessed the viability of UAS for collecting reliable information for use in deriving phenotypic parameters of maize genotypes and predicting individual genotype yield. Whereas these studies make use of UAS for remote sensing of maize genotypes, no classification for the discrimination of genotypes has been conducted using only RGB or even multispectral sensors.

Remote-sensing classification of annual agriculture has largely been conducted using measurements of spectra, texture, and object geometry. Apart from work by Wu et al. (2017) and de Castro et al. (2018), the use of SfM and LiDAR derived 3D information for agricultural remote sensing has primarily been limited to plant analyses for annual crops with very little research devoted to classification. In a similar context, research has made use of SfM for classification of trees (Torres-Sánchez et al. 2015; Nevalainen et al. 2017; Cao et al. 2018). LiDAR has also been used extensively for the purpose of tree species and plant-foliage classification, and many techniques have been developed to assess the structure and characteristics of forests and individual trees (Hung, Bryson, and Sukkarieh 2006; Pipkins et al. 2014). Recent advancements in UAS technology have promoted increased use in LiDAR and stereotypic SfM for assessment and measuring of annual agriculture and crops with short height. Over approximately the past decade, multiple studies have assessed the use of SfM point clouds and digital surface models (DSMs) for predicting yield and identifying phenotypic traits of plants such as height and leaf area index (LAI) (Jay et al. 2015; Li et al. 2016). Though SfM measurements derived from passive sensors remain less robust than active remote sensing of LiDAR, research on LiDAR classification suggest 3D properties of agriculture may prove valuable information for classification. Research utilizing UASs for high-throughput phenotyping and plant monitoring further supports the use of 3D measurements for plant scale agricultural classification as previous studies were able to capture 10-40% increase in genetic variation (Aasen et al. 2015; Anderson et al. 2019).

Studies have made use of UAS derived 3D information for classification, but within the setting of annual agriculture, work has been limited (Wu et al. 2017; de Castro et al. 2018). Within past studies, height information was gathered at the plot-level. Research still needs to be conducted on the integration and benefit of using height information derived from SfM point clouds for classification. One area still under explored is the creation 3D derived object properties of annual agriculture from UAS which describe a portion of an individual plant's structure. Tree crown and height, for example, has been used for classification of tree species and to characterize plant structure (Hung, Bryson, and Sukkarieh 2006; Popescu and Zhou 2008). Future research using height information and spatial context could provide means for higher detailed feature extraction and characterization of plant architecture such as on stalks, tassels, or ears.

Methods developed in this study will contribute to improving the efficiency of plant breeding through extraction of increased descriptive information for genotype selection. Need for faster, standardized assessment of genotypes to increase capacity and efficiency of plant breeding is a driving factor for this research. Future work will further utilize variables derived and evaluated by this study for yield predictions of maize. Multi-temporal OBIA classification of UAS images will further contribute to the broader application of remote-sensing in precision agriculture (e.g. crop-weed management, pest detection, etc.)

CHAPTER 3

RESEARCH CONCEPTS AND METHODOLOGIES

3.1 Introduction

Methodologies conducted within this thesis are discussed in the following sections. A brief description is provided for the plant breeding study of hybrid maize genotypes on which analysis was performed. Preprocessing for creation of OBIA inputs and image quality corrections for multi-temporal analysis are provided. OBIA was performed at the row scale for maize within this study; a discussion of object scale at the row versus the plant level is presented. A proposed workflow for OBIA of single and multi-temporal creation of maize row image objects is provided. In addition to the full list of variables used in this study, software specifications for feature extraction of RGB derived information are provided. Methods for multi-temporal grouping of variables for description of individual maize rows is presented. Along with feature selection and algorithms for classification of genotypes, Figure 1 provides an outline of methodologies discussed in this chapter.

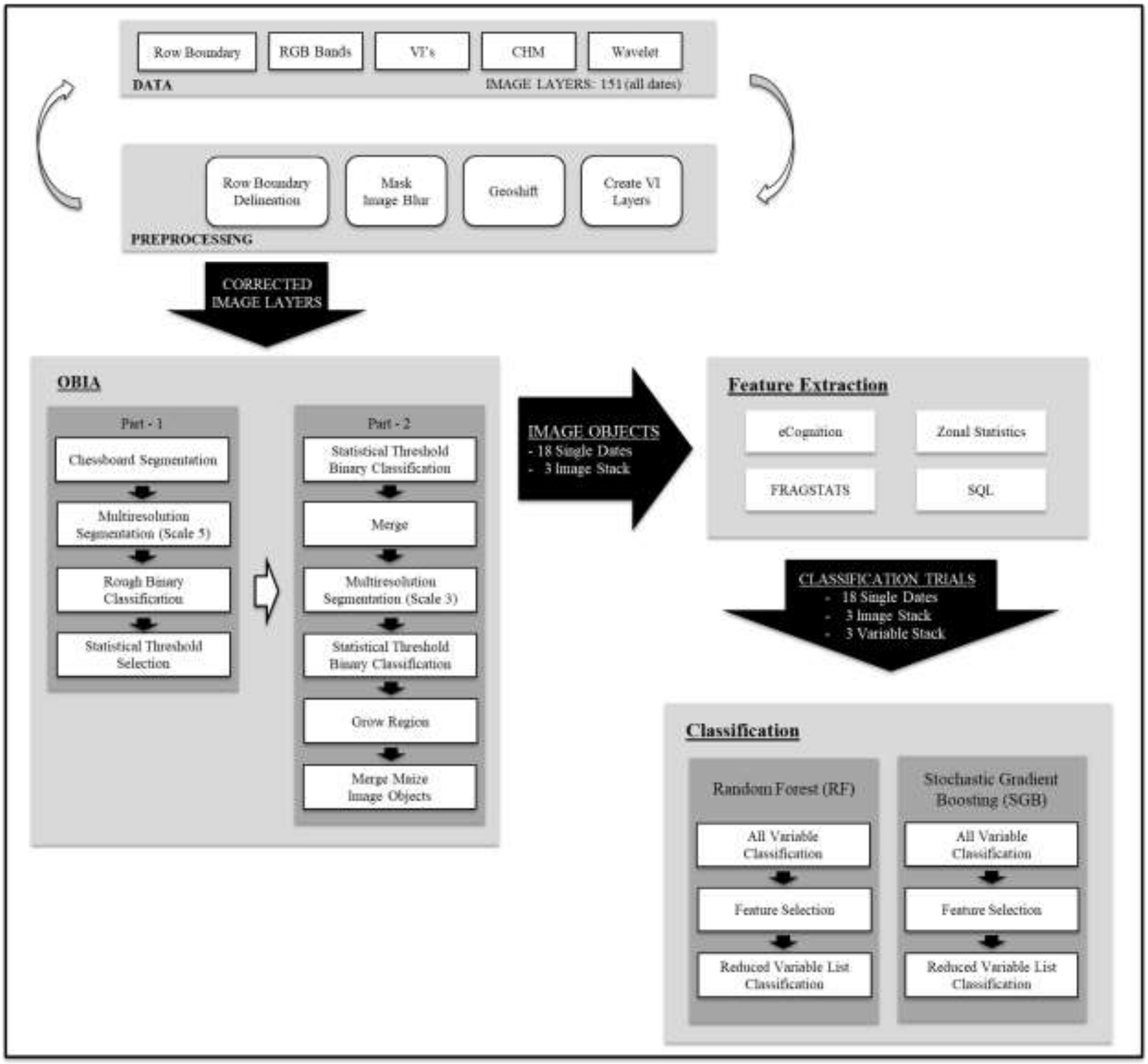


Figure 1:Outline of methods applied for multi-temporal object-based classification of maize.

3.2 Study Site

The study site was located at Texas A&M University AgriLife Research Farm (30.55 N, 96.43 W) south of College Station, Texas near the Brazos River. Field plots were designed and managed by the corn breeding program from the Department of Soil and Crop Sciences at Texas A&M University and used material from the national Genomes to Fields (G2F;

genomes2fields.org) program. Experimental hybrid maize varieties were planted in three field management conditions: field (G2FE) was irrigated/fertilized for optimal conditions, field (DG2F) was dryland (non-irrigated) and had reduced fertilizer to increase stress conditions, and field (G2LA) was late-planted but irrigated to increase heat and moisture stress. Analysis for this study is focused on the experimental condition G2FE (optimal). Within the G2FE field, 280 genotypes were planted. The field was organized as plots; each plot consisted of two paired rows for each genotype. With the exception of six (6) genotypes, all genotypes had a second replicate plot present in the field. Therefore, all genotypes with the exception of genotype pedigree GEMS-0264/3IIH6 had a minimum of 4 rows located in the field. Refer to Appendix 1 for the list of genotypes with more than two plots. Plots were organized randomly within the field, within replications, and within field management conditions. In other words, a split-split-split plot design was used where the first split was field management, the second split was replication, and the third was genotypes randomized within the replications. The field condition G2FE and an example of a two-row plot can be seen in Figure 2 for the date May 12 (05/12).

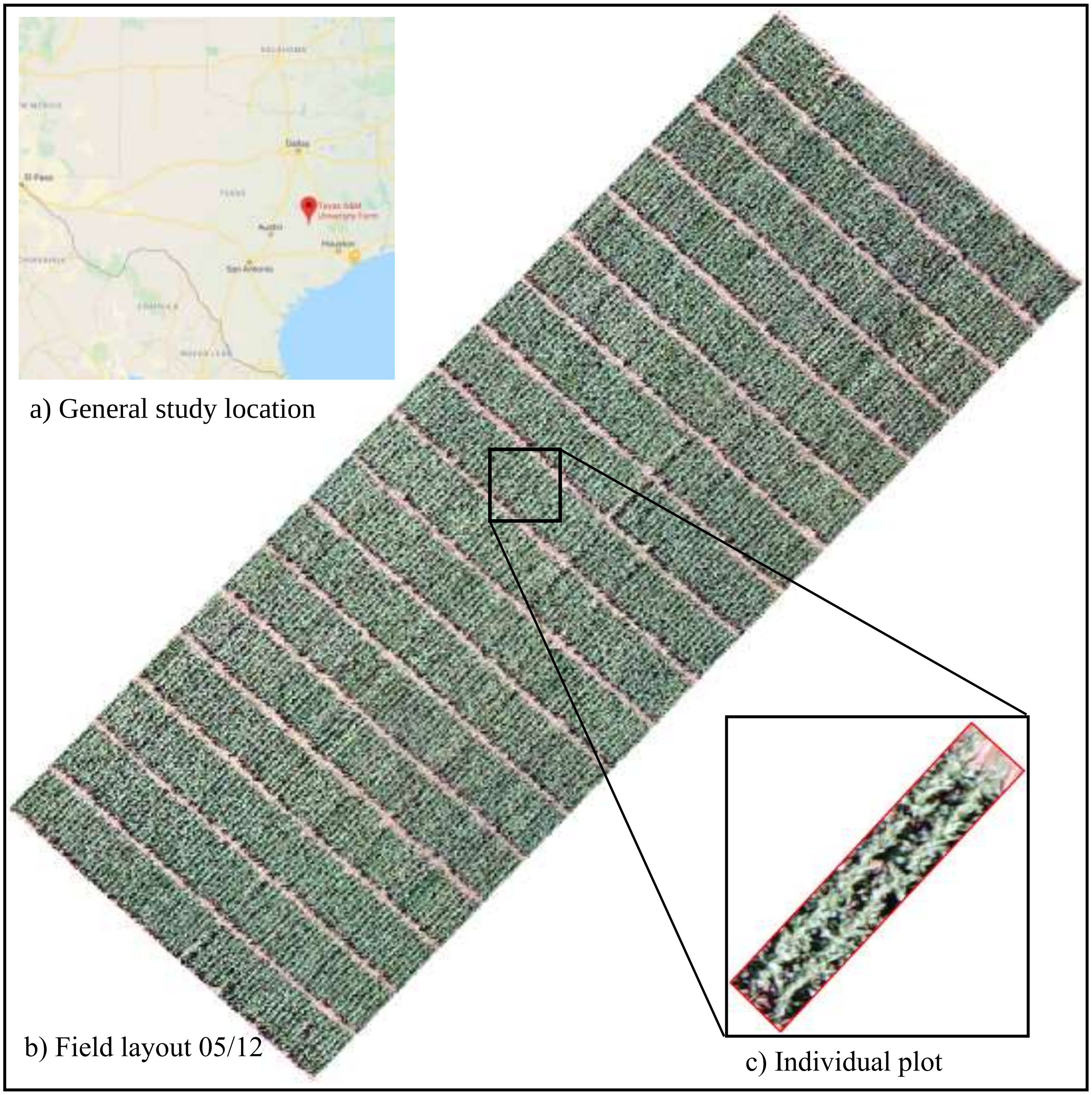


Figure 2: Study area. a) inset of general study site shown in Google Maps®, b) image mosaic of 05/12 for field condition G2FE, c) inset of mosaicked image plot.

3.3 Data Acquisition

UAS images were collected by Dr. Dale Cope in the Department of Mechanical Engineering at Texas A&M University for the maize growing season of 2017 between the dates of March 21 and July 13. Over the 2017 growing season, images were collected on 18 separate dates capturing a range of plant development from emergence to maturity. In Figure 3, three of the eighteen dates where image collection occurred are provided as an example of phenotypic change seen in maize canopy over the timeseries of this study. In an effort to maximize collected data during the flowering period, flights occurred biweekly, as weather permitted, between May 2 to June 16. Data was collected on the following dates: 03/21, 04/07, 04/20, 05/02, 05/05, 05/09, 05/12, 05/15, 05/24, 05/30, 06/02, 06/06, 06/09, 06/12, 06/16, 06/23, 06/29, 07/13. Refer to Appendix 2 for a full list of image collection dates. Images were collected with a Tuffwing UAV Mapper (http://www.tuffwing.com), a fixed-wing UAS. The UAS was mounted with a Sony ILCE-6000 (16mm) camera. The camera has the following specifications: 6000 x 4000 resolution, 16mm focal length, and 4 x 4 µm pixel size. Images were collected at an altitude of 120 meters and had a spatial resolution of 2.66-2.83cm. Spatial resolution varied slightly between images. Orthorectification and mosaicking performed in Agisoft PhotoScan Professional (Agisoft, St. Petersburg, Russia) combined overlapping photos into single images for each date; image overlap was 80%. In addition to orthomosaic images, point clouds for each date were also produced via Structure from Motion (SfM) photogrammetry using Agisoft PhotoScan Professional. This data was processed by the Texas A&M AgriLife Research Flight and Data Analysis Services

Figure 3: An example of phenological change in maize canopy over the 2017 growing season as shown by the dates 04/07, 05/02, and 06/06 listed left to right.

eCognition Developer 9.0 (Trimble Germany GmbH, Munich, Germany) provided the primary platform for object-based image analysis performed in this study (Benz et al. 2004). Image mosaics, SfM point clouds, and a vector layer defining plot boundaries were provided by Dr. Murray's research group and the Texas A&M AgriLife Research Flight and Data Analysis Services and are available at CyVerse (Murray et al. 2019). Eight image layers and one thematic layer were produced for analysis in eCognition. Image layers included bands: Red, Green, and Blue; the vegetation indices: Excess Green (ExG), Vegetativen (VEG), Excess Red (ExR), and Yellow Index; and canopy height model (CHM). Image layers described above were created for all dates with the exception of a CHM layer for the date March 21 (03/21). The CHM layer for 03/21 was omitted from this study as SfM point clouds had difficulty characterizing plant height at the juvenile stage of maize development present for this date. A single thematic layer composed of vectors defining row boundaries for each row of maize was also provided for analysis across all dates. Additional image layers derived from wavelet analysis were included as inputs for feature extraction and genotype classification. Wavelet image layers were provided by Dr. Bishop's research team in the Department of Geography at Texas A&M University. A continuous wavelet transform was used to measure the periodicity of a signal over multiple spatial scales. Wavelet image layers used in this study were created at a scale of 0.2 to 5.0 (0.35-8.87cm) at increments of 0.2 in scale for the dates of May 2 through June 29. Wavelet analysis was included in this study to provide additional information as to spatial patterns in image reflectance such as break points, trends, self-similarity, and higher-derivative discontinuities (Bishop et al. 2018).

3.4 Data Preprocessing

Preprocessing of data through quality assessment and multi-temporal alignment were fundamental to the application of multi-temporal object-based methods performed in this study. Prior to image segmentation, spatial and spectral corrections were applied to point cloud and orthomosaic images. The following paragraphs provide information regarding data preparation such as: conversion of CHM to 2-diminsional image layers, creation and comparison of vegetation indices, evaluation of image quality and masking of blurry image regions, and geospatial correction of image layers. These steps included the preprocessing of row boundary delineation, mask image blur, geoshift, and the creation of VI layers shown in Figure 1 on page 12.

3.4.1 Canopy Height Model Conversion

To reduce computational demand for object-based operations employed within eCognition, CHM data -digital terrain models specifically referring to measurements of crop height- were converted from point clouds to 2.5-diminsional raster images. Point clouds were converted to raster using the tool "LAS Dataset To Raster" in ESRI ArcMap. The interpolation type applied was Triangulation and the interpolation method was Linear. Selection method was set to Maximum and resolution was set to 0.06. Sampling type was Cell Size and set to a value of 0.1. Figure 4 provides an illustration of 2.5-dimensional CHM used in this study.

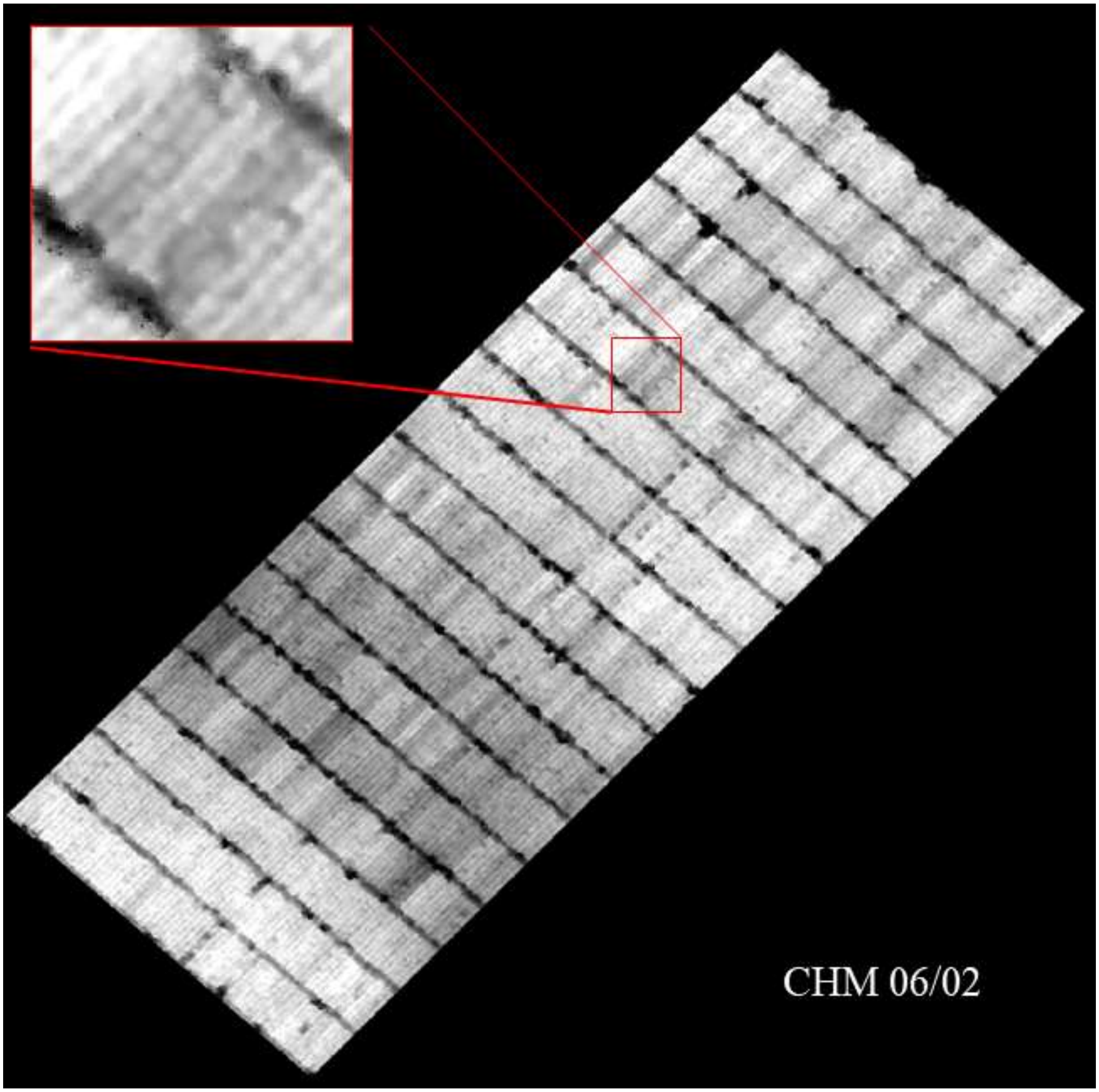


Figure 4: CHM for the date 06/02.

3.4.2 Optical Vegetation Indices

True color vegetation indices (VIs) provided a valuable component of our analysis as inputs for segmentation, thresholds for binary classification of maize and non-maize, and additional information for feature extraction. To emphasize different spectral or physiological characteristics of plants and to maximize the discriminative capacity of RGB bands, a comparison of twenty true color vegetation indices was conducted. Only true color VI's were generated as multispectral (red edge) band was at a different resolution and hyperspectral bands were not available. VI raster layers were generated using Raster Calculator in ESRI ArcMap and were evaluated through visual assessment for clarity and descriptive ability. The following indices were compared: CIVE, COM, COM(2), ExG, ExGR, ExR, GLI, GR, GRRI, NDI, NGBDI, NGRDI, VARI, VDVI, VEG, VIg, RGB Yellow, WI, and YI (YI was a custom index designed for this study) Table 1 provides a list of VI names, formulas, citations and related information. (David, Ballado, and Ieee 2016). All VIs produced for this study were derived from the bands of single date; multi-temporal true color VIs were not integrated in this study. Normalization, as shown in Table 2, was applied to select VIs where normalization was included in the original study.

Table 1: Vegetation indices tested for use in segmentation and genotype classification.

Vegetation Index	Abbreviation	Formula	Citation
Color index of vegetation extraction	CIVE	$0.441r - 0.881g + -0.385b + 18.78745$	Kataoka et al. 2003
Combination	COM	$0.25ExG + 0.3ExGR + 0.33CIVE + 0.12VEG$	Guijarro et al. 2011
Combination 2	COM(2)	$0.36ExG + 0.47CIVE + 0.17VEG$	Guerrero et al. 2012
Excess Green	ExG	$2g-r-b$	Woebbecke et al. 1995
Excess Green - Excess Red	ExGR	$3g-2.4r-b$	Meyer and Neto 2008
Excess Red	ExR	$1.4r-g$	Meyer, Hindman, & Laksmi. 1999
Ground Level Image	GLI	$(2g-b-r)/(2g+b+r)$	Louhaichi, Borman, and Johnson 2001
Green Ratio	GR	$Green / (Red + Green + Blue)$	Ide and Oguma 2013
Green-red Ratio Index	GRRI	g / r	Gamon and Surfus 1999
Normalized Difference Index	NDI	$(g-r) / (g+r)$	Woebbecke et al. 1993
Normalized Green-Blue Difference Index	NGBDI	$(Green-Red) / (Green + Blue)$	Xue and Su 2017
Normalized Green-Red Difference Index	NGRDI	$(Green-Red) / (Green + Red)$	Xue and Su 2017
Visible Atmospherically Resistant Index$_{green}$	VARI	$(Green-Red) / (Green+Red-Blue)$	Gitelson et al. 2002
Visible-Band Difference Vegetation Index	VDVI	$(2 * g - r - b) / (2 * g + r + b)$	Xue and Su 2017
Vegetativen	VEG	$g / (r^a b^{(1-a)})$	Hague, Tillett, & Wheeler. 2006
Vegetation Index$_{green}$	VIg	$(Green-Red) / (Green+Red)$	Gitelson et al. 2002
RGB Yellow	Yellow	$R + G$	Doi 2012
Woebbecke Index	WI	$(g-b) / (r-g)$	Woebbecke et al. 1995
Yellow Index	YI	$3g-2r$	Original

Table 2: Normalization applied to some vegetation indices.

Bands
r = R / (R+G+B)
g = G / (R+G+B)
b = B / (R+G+B)

After visual assessment, we chose the following VIs as we felt these provided valuable and unique information for our analysis: ExG, ExR, VEG, and YI. ExG is a well-known algorithm supported in the literature which provides information as to the greenness of vegetation and has been shown to perform better than other true color VIs in multiple studies (Woebbecke et al. 1995). Based on visual assessment, ExG provides noticeable contrast between green vegetation and background. In addition, ExG appeared less influenced by shadowing and glare effects or brightness within images as did other VIs tested such as the index VEG. ExG provided the primary threshold for our binary labeling of maize and non-maize image objects described in greater detail later in this study. VEG was chosen to provide supplemental information as to the greenness of vegetation. In previous studies, VEG was used for vegetation assessment (Hague, Tillett, & Wheeler. 2006). VEG was chosen as a second descriptor of greenness which provided additional information during feature extraction. As VEG image layers appeared less pixelated than those of ExG, VEG was included to improve segmentation. The index ExR provided information as to the redness of plant canopy (Meyer, Hindman, & Laksmi. 1999). As plants matured, ExR was used to measure browning foliage and as additional threshold for labeling maize and non-maize image objects. In addition to ExR, we felt an index designed to highlight yellow regions within an image would be especially descriptive of maize as plants began to tassel and later dry out and brown. Two vegetation indices (RGB yellow and Green/Blue) were evaluated for the purpose of highlighting yellow image regions; however, in our evaluation both indices were unable to isolate or provide meaning yellowness values for image regions solely using RGB bands (Doi. 2012). The vegetation index Yellow Index (YI) was a custom algorithm created to identify yellow image regions and provide measurements as to the yellowness of image objects. YI proved more effective at describing yellow image regions and distinguishing between yellow and other image colors than the indices RGB yellow and Green/Blue. YI was not included as an input for segmentation or as a threshold for binary labeling but was included as an additional input for classification. VIs chosen for this study are shown for the date 06/02 in Figure 5. As shown in Figure 5, results for ExR and YI are very similar; however, ExR appears to assign slightly higher intensity to non-vegetation than YI. This suggest, based on visual assessment, YI may be better suited to analyze yellow or dying vegetation than ExR.

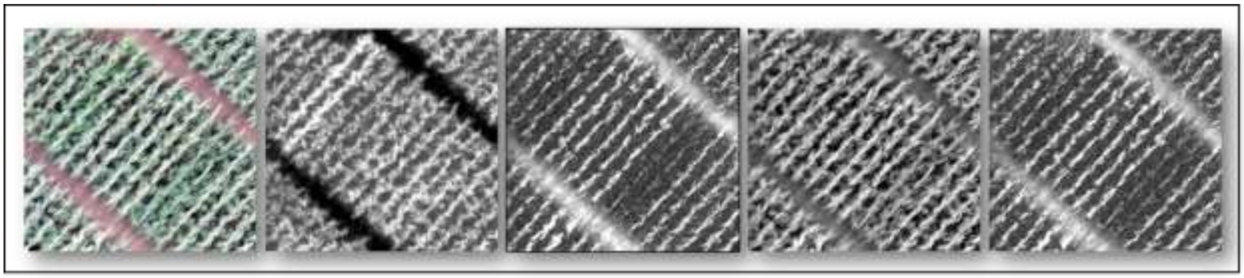


Figure 5: RGB image and Vegetation indices ExG, ExR, VEG, and YI for the date 06/02 as listed from left to right.

3.4.3 Geospatial Correction

During orthorectification, mosaic images were geospatially corrected using ground control points; however, further corrections were necessary for multi-temporal alignment of maize rows across all dates. Though photos were tied to ground control points during orthorectification and mosaicking, image distortion and spatial inaccuracy were noticeable in image mosaics. Geospatial correction applied at this step did not correct images based on RMS error produced during mosaicking and orthorectification, but images were corrected by aligning rows of maize within a standard geospatial unit. An ESRI shapefile of row boundaries provided by Anderson et al. (2019) was used as a constant geospatial unit for which images were corrected to. Each date was visually evaluated as to how well maize rows were centered within row boundary vectors. Figure 6 provides an example of a row canopy being split by the row boundary vector. To correct dates, the minimum x and y value of image layers were adjusted so that vegetation for a genotype shifted to the center of row boundary vectors. This geolocation shift was applied in the software ENVI version 4.8 (Exelis Visual Information Solutions, Boulder, Colorado). using the function “Edit ENVI Header”. Dates with geospatial corrections are provided in Table 3. Geospatial correction during this step was necessary to prevent segmentation from producing image objects from the spectra and height information of two different genotypes. This was especially important for image stack segmentation where image layers from multiple dates were stacked for a single segmentation.

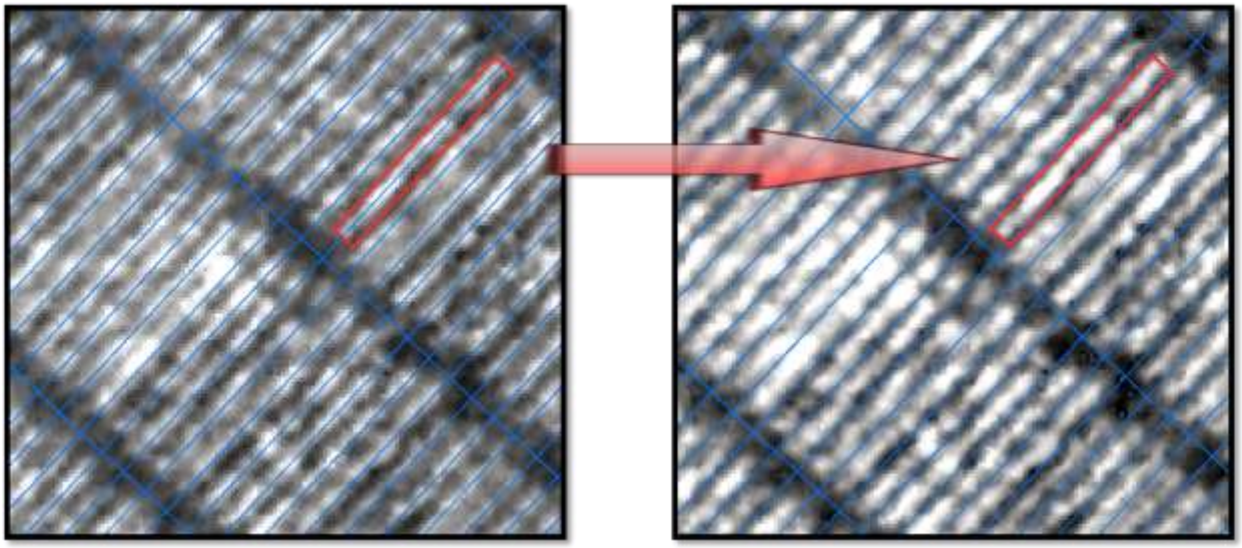


Figure 6: Geoshift applied to spatially align images of all dates to a row boundary polygon.

Table 3: Image dates and corrections applied.

Date	Geoshift	Mask Image Blur
March 21		
April 07	x	x
April 20		x
May 2		x
May 5	x	x
May 9		x
May 12	x	
May 15	x	x
May 24	x	x
May 30	x	x
June 02	x	
June 06	x	x
June 09	x	x
June 12		x
June 16	x	x
June 23		
June 29	x	x
July 13		x

3.4.4 Mask Image Blur

Blurring within images noted after mosaicking and orthorectification posed a potential inhibitor for genotype classification. Image blur or the smearing of portions of a mosaicked image was present in patches of varying levels of intensities across most dates. Image blur reduced the sharpness and clarity of image features creating spectrally and spatially distorted descriptive information for genotype classification. Boundaries in image clarity could be seen at right angles and along straight lines across orthomosaics. Figure 7 provides an illustration of image blur observed and the extensiveness of masking applied for the date 04/07.

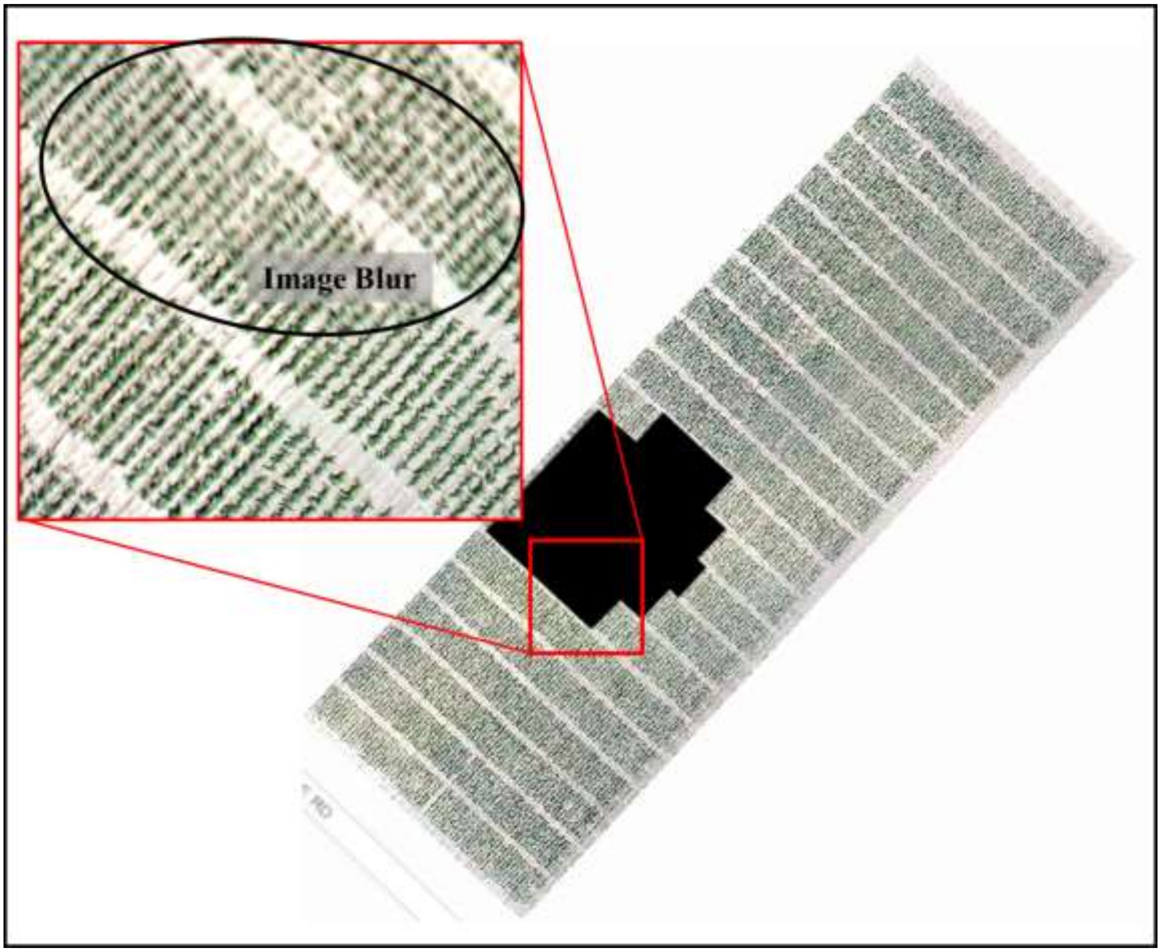


Figure 7: Example of extent and intensity of image blur for date 04/07.

As the focus of our study was not to address the optimization of image clarity, we decided to take a qualitative approach to compensate for blur. Mosaics of each date were manually assessed for “significant image blur”. This method was subjective to user evaluation and should be automated as a quantitative detection in future research. Blur did occur at varying levels within mosaics and also varying levels between dates. Increased blur was noticed at later dates. Based on user discretion, regions of “significant image blur” were identified and masked. Masking of image blur was carried out along row boundary vectors. Rows with partially clear and blurred portions did occur throughout mosaics. Rows with vegetation having greater than ~80% image blur present were masked. It should be noted that since smearing occurred at varying levels not all image blur was removed from a mosaic. It should also be noted that different levels of distortion were deemed appropriate to keep based on the date, overall quality of the mosaic or prevalence of blur over the entire mosaic. For example, since intensity of image blurring was high and very prevalent in later dates but clear spectral contrast was noticed between genotypes, regions with higher blur than was clipped in earlier dates was deemed acceptable and remained unmasked. For dates where the majority of the mosaic had image blur present, only regions with very high distortion were removed. Masking was applied to all dates within the time series except 03/21, 05/12, 06/02, and 06/23. Table 3 on page 24 list all dates on which quality corrections were applied.

3.5 Selecting Scale for Multi-Temporal OBIA

3.5.1 Scale of Objects

The first step in the application of object-based image analysis was to decide what level of maize architecture in a field image objects would represent. As noted previously, image objects rather than pixels provided the base unit for classification in OBIA; therefore, it was necessary to determine at what scale within a field image objects could or should be created. This was a critical point in our analysis as two scales of canopy architecture were considered: 1) plant or individual stalk-level image objects and 2) row-level image objects. Based on the scale of analysis chosen, segmentation parameters were adjusted accordingly to create image objects best representing individual plants of a genotype or individual rows of a genotype. Challenges to creating image objects representative of the canopy architecture proposed were 1)the overlapping of plant leaves between stalks within a row and 2)the overlap of leaves between rows.

As leaves overlapped between plants, occlusion occurred making visual identification of individual plants difficult or not feasible at the current spatial resolution of 2.66-2.83cm. At the same time, spectral differences were captured within individual stalks of maize; for example, spectral differences could be seen between portions of a plant such as leaves, tassels, etc. Apart from date 03/21, a few weeks after planting, overlap in leaf canopy was present for all dates. After performing preliminary segmentation analysis, it was noticed that spectral similarity was in fact greater between portions of different plants than within a single plant. Thus, when testing segmentation at a plant-level analysis, segmentation produced image objects describing plant center and outer overlapping leaves of more than one plant. Plant-level segmentation produced results where image objects for multiple plants, portions of a single plant (ex. plant core or tassel region), and combined portions of multiple plants (ex. overlapping leaves) but had difficulty separating each individual plant as a unique image object.

Similar to plant-level analysis, row-level segmentation was also challenged by overlapping canopy between rows. Where overlap between row canopies occurred, greater spectral similarity was present between the outer leaves of two rows and within the center of rows than within an entire individual row. Testing demonstrated that segmentation at the row scale for image dates with canopy overlap present produced image objects covering portions of two rows or portions of only a single row but rarely a single full row. An inverse relationship was created when adjusting the scale parameter in segmentation. As the scale parameter was increased, image objects were more likely to cover portions of multiple rows. As the scale parameter was lowered, segmentation produces fewer results which covered two rows but was unable to capture an entire row. At the same time, segmentation preferred to create image objects for outer overlapping leaves along each row and for the darker or tasseled center of a row. Within row scale segmentation, results produced images objects encompassing an entire row, a portion of a row, portions of two rows, and multiple rows.

Results from preliminary analysis suggested that a single multi-resolution segmentation operation in eCognition was not adequate alone for creating representative image objects for entire and only one maize plant or row across an entire field regardless of the scale parameter chosen. The proposed OBIA workflow of this study provided a solution to the unrepresentative segmentation noticed in preliminary analysis of plant- and row-level analysis. The use of a row boundary vector is suggested for row-level object-based analysis. As this study is performed on a maize field designed for plant breeding with exact specification at planting, rows have precise spacing, length, and orientation. A tool recently developed by Anderson et al. (2019) was applied to delineate plot boundaries based on spacing, orientation, length and starting position of maize rows within a field. From plot boundary vectors, rectangular boundaries for each row can further be defined as seen in Figure 8. An identification number was assigned to each row boundary vector. The inclusion of a thematic layer of row boundary vectors as rectangles provided the necessary information needed to limit row-level segmentation to a single row. Row boundary vectors also provided a constant geospatial extent for which image objects of maize, varying in size and location over a growing season, could be connected for multi-temporal analysis. Because of the inclusion of row boundary vectors, we were able to create clear image objects for rows of maize across all dates. We therefore chose a row-level scale of analysis for this study.

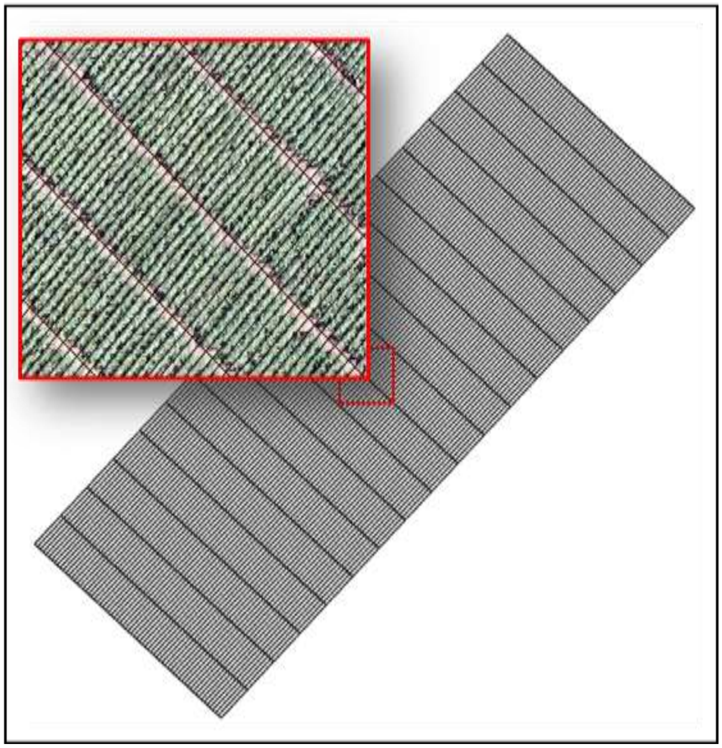


Figure 8: Row boundary polygons for field condition G2FE. Subset shows row boundaries overlaying image 05/12.

3.5.2 Scale Parameter for Multi-Temporal Segmentation

A second challenge to implementing OBIA on a multi-temporal dataset was the selection of appropriate scale parameters for eCognition multiresolution segmentation. Scale parameter determined the level of spectral heterogeneity within image objects and thus directly influenced the size of image objects produced by segmentation (Benz et al. 2004; Drăguţ et al. 2014). The scale parameter established at what threshold pixels were no longer combined into a particle image object and at what point the next image object should be formed. In previous literature, the appropriate scale parameter was often determined based on user expertise through empirical trial and error analysis (Blaschke 2010). The necessary scale parameter was influenced not only by the size but spectral uniformity of a phenomena being captured in a given image. As canopy size and architecture of maize changed during the growing season as plants matured, evolving size and spectral makeup of objects required the selection of a new scale parameter for each date when using conventional segmentation analysis. Manual selection of the scale parameter through “trial and error” experimentation did not provide a practical option for the segmentation of 18 individual dates and three image stacks conducted in this study. The multi-temporal nature of this study emphasized the need for a more automated way to select segmentation parameters than through repeated “trial and error” assessment. To counter the issue of selecting a unique scale parameter for each date in our timeseries, we proposed a new object-based workflow using row boundary vectors and constant segmentation scale parameters across all dates. This method functioned through a series of sub-plant multiresolution segmentation, Grow Region, and Merge Region operations (segmentation and segmentation reshaping algorithms provided in eCognition) within a thematic layer of row boundary polygons to create image objects characteristic of maize vegetation for a row (eCognition Developer 2014).. During this workflow, two scale parameters were used for all dates; a second set of scale parameters were chosen for image stack segmentation. Trial and error assessment was still applied to select the first and second scale parameters used; however, within our OBIA workflow, a single segmentation was not responsible for creating image objects describing an entire row but only portions of a row. Selection of a segmentation scale parameter using our proposed method was more flexible as less emphasis was placed on identifying the perfect scale for delineating row objects. The method proposed only required that the scale parameter be set to produce image objects smaller than the target object to be captured, in this case smaller than a full row of maize. Our method thus significantly reduced the tedious and subjective use of trial and error selection of the scale parameter within multiresolution segmentation.

3.6 Multi-Temporal Analysis

In this section, the multi-temporal components of analysis applied in this study is described in length. As previously referenced, a new OBIA workflow was applied to a timeseries of 18 orthomosaic images expanding the length of the maize growing season. Multi-temporal analysis within this study occurred in three ways: as image stack segmentation which provided multi-temporal variables for classification, as variable stack classification using variables extracted from OBIA results of individual dates, and a multi-temporal comparison of classification of single dates. This study extends beyond multi-temporal comparisons of separate classifications by applying techniques that captured temporal dimensionality in the growing season, which was then included as input for classification. A clear distinction between the application of variable stack classification and image stack segmentation are made in the following sections.

3.6.1 Image Stack for Segmentation

In our analysis, we wanted to perform a series of independent segmentations for each date of data acquisition. We also wanted to conduct three multi-temporal segmentations where image layers of multiple dates were grouped together for a single segmentation. To allow for independent segmentation of each date to be performed in eCognition, maps were created using the image layers of a single date. Three additional maps were created for image stack segmentation; all image layers from all dates were added to these maps for various levels of inclusion in segmentation and feature extraction. The first image stack map was created for a three-date segmentation. All images after the date 03/21 without image blur masking were selected for this segmentation. Image date 03/21 was not used in this segmentation due to very limited canopy extent. Image layers used in segmentation of this three-date image stack map included 05/12, 06/02, and 06/23; these dates were selected as they were the only dates available without image blur masking. Within this map, no portion of the field was influenced by void or null image regions created during image blur masking. The second image stack map was used for a six-date segmentation which included the six dates with least image blur masking applied. Dates used for six-date image stack segmentation were 05/02, 05/12, 05/15, 06/02, 06/16, and 06/23. A third image stack map was created for a full multi-temporal segmentation of the timeseries and used all image layers between the dates 04/07 to 06/23; however, dates 03/21, 06/29, and 07/13 were still omitted from segmentation in this map due to large levels of masking (06/29, and 07/13) or limited canopy size (03/21). Dates used in each image stack segmentation are shown in Table 4. All image layers were added to image stack maps; however, only dates previously listed were used as inputs for image stack segmentation. Feature extraction applied to image stack segmentation results however made use of all image layers, both layers included in segmentation and additional layers from across the growing season which were not used to generate image stack objects. This provided the ability to extract valuable multi-temporal information from all dates from a multi-temporal image object.

Table 4: Dates used as inputs in each image stack segmentation.

Date	3-Date	6-Date	15-Date
March 21			
April 07			x
April 20			x
May 2		x	x
May 5			x
May 9			x
May 12	x	x	x
May 15		x	x
May 24			x
May 30			x
June 02	x	x	x
June 06			x
June 09			x
June 12			x

Date	3-Date	6-Date	15-Date
June 16		x	x
June 23	x	x	x
June 29			
July 13			

Image stack segmentation, not to be confused with variable stack classification, creates a multi-temporal image object produced by the spectral homogeneity of pixels across all image layers included as inputs (Hussain et al. 2013; Osmólska and Hawrylo 2018; Song et al. 2019). In our case, image layers were provided from multiple dates. Image objects therefore produced by this operation did not describe the extent of maximum or minimum vegetative canopy for a row of maize but can better be defined as the core region of a maize row over a growing season given a set of dates. Figure 9 provides an example of the different segmentation results produced by three single date segmentations and a stack segmentation combining the same three dates. Stack segmentation produces new image objects with geometric information describing a multi-temporal phenomenon across a set of dates and image layers provided. Geometrical object properties described a new multi-temporal image object that was characteristic of not just one image, but all images included in segmentation. Spectral and textural properties were not extracted from new multi-temporal image layers but were summary statistics of original bands calculated within a new multi-temporal spatial extent or image object. Simply put, spectral and textural measurements were calculated from the same image layers as individual date feature extraction but over a different geospatial area.

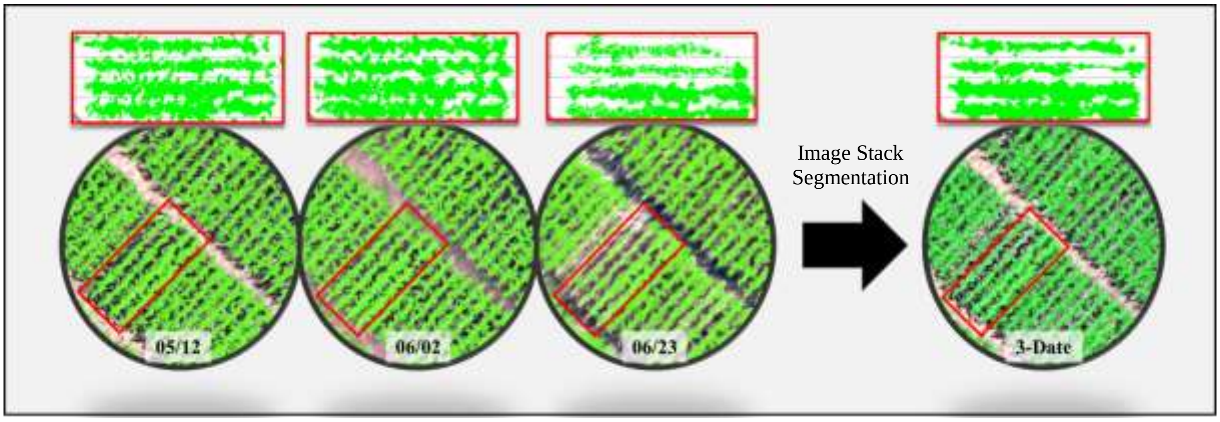


Figure 9: Image objects produced from the image layers 05/12, 06/02, and 06/23 and the combination of all three images as a 3-date image stack segmentation overlaying the image 05/12.

3.6.2 Variable Stack for Classification

Beyond performing classification on data produced by a single OBIA workflow or series of segmentations, we also joined separate segmentation results for a multi-temporal classification which we refer to as variable stack classification. Variables included in this analysis and how each was derived is provided in the following section *Feature Extraction* and in Table 7 on page 52. Variables extracted from multiple dates during feature extraction were combined for three multi-temporal variable stack classifications. Variables from different dates were combined to form variable stack datasets. This multi-temporal component of our research, though not producing new multi-temporally descriptive variables, was used to train classification over a multi-temporal feature-space as it combines information from multiple, independent OBIA results for each date in the timeseries. Our approach is innovative as we propose a manner to stack segmentation results (image objects) verses pixels from multiple dates. This provided a solution to generating object properties descriptive of an entire and only one maize row in the complex scenario where 1) one or multiple image objects could comprise a single row of maize, 2) the number of image objects changed between dates for the same maize row, and 3) the spatial extent and position of image objects changed between dates. Identifying corresponding image objects representative of a target unit or phenomenon (a single row of maize in our analysis) that are produced by separate segmentations has been cited in object-based change detection literature as a difficult challenge faced when conducting multi-temporal OBIA (Hussain et al. 2013). To perform variable stack classification, a dataset was generated based on the following criteria: 1) Image objects had to be constricted to only describe a single maize row. This allowed classification to be trained at the row object level. 2) Segmentation often produced multiple image objects for a single maize row. To produce variables descriptive of an entire maize row, summary statistics were generated from all separate image objects making up a row. 3) Unlike pixels which are consistent in spatial extent and location, a constant spatial entity between separate segmentations had to be generated to connect row image objects. Using a row boundary thematic layer as shown in Figure 10 and described previously in the section *Selecting Scale for Multi-Temporal OBIA*, we were able to group the object property information of all image objects comprising a single maize row at a single date while also connecting the image objects describing that same row from all dates/images in the timeseries for use in a single classification. This is different than work described by Petitjean et al. (2012) as vector derived features or properties in our analysis are not assigned to pixels but to a row boundary image object or thematic layer. To our knowledge, multi-temporal object-based variable stack classification of UAS images has not been performed according to our assessment of the literature.

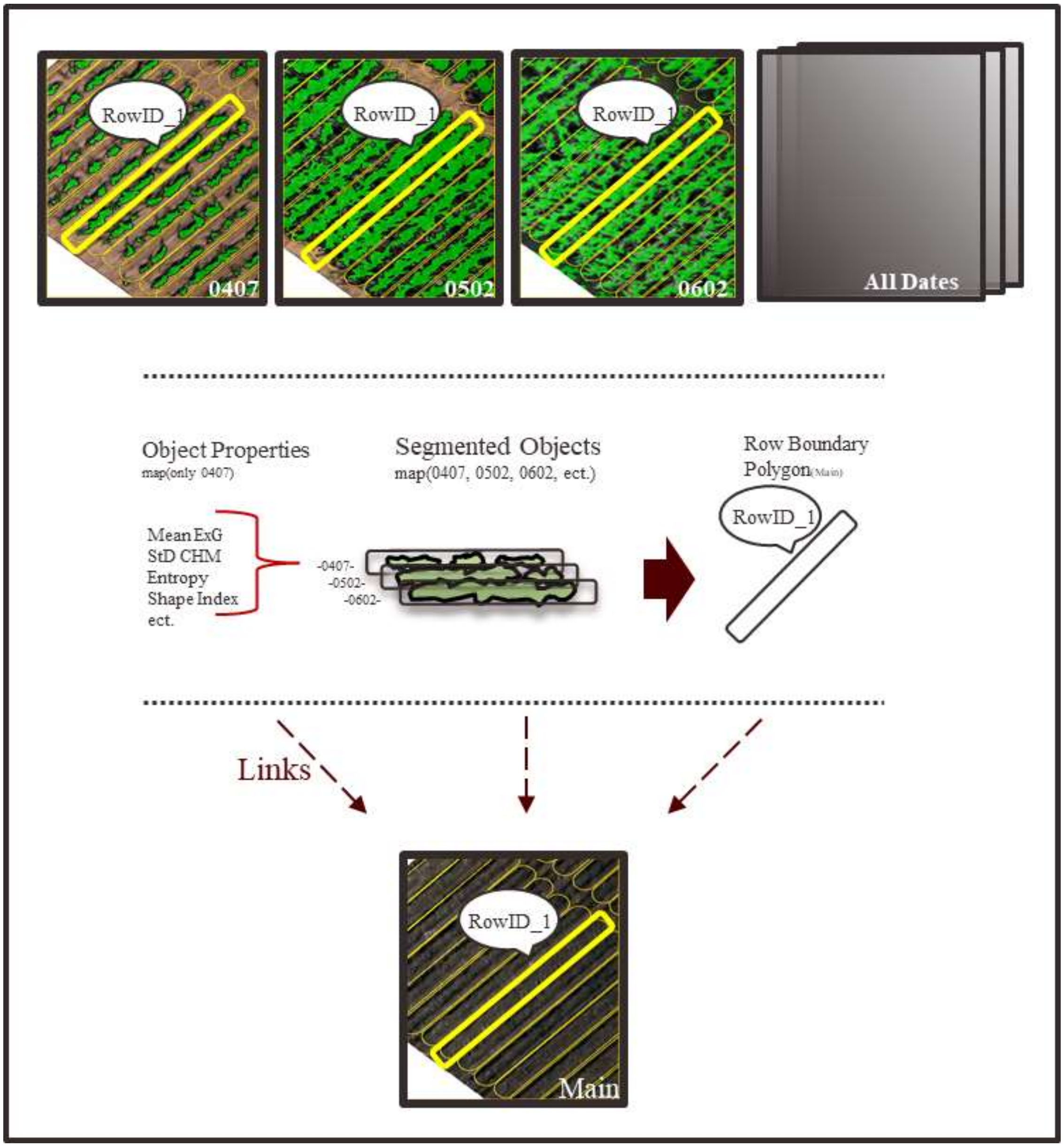


Figure 10: Connecting image objects from multiple dates using row identification numbers from row boundary polygons.

3.7 Proposed Object-based Image Analysis Methods

3.7.1 eCognition Workspace

Image layers (spectral bands, VI's, & CHM) and the thematic layer of row boundary vectors were added as inputs to an eCognition workspace. Components of the eCognition workspace included all the image layers to be used for segmentation and feature extraction, parameters defining null values within a project, and a process tree where a list of algorithms used to perform OBIA operations were defined. Within the workspace, individual maps were created for each date of image acquisition and image stack segmentation. eCognition operates in a hierarchy structure where multiple segmentations can be performed in a single map and at multiple image levels within a map. Segmentations performed on image levels within a map are not independent of one another. Image object boundaries produced by segmentation on levels higher in the image layer hierarchy influence and limit the spatial extent of image objects of subsequent levels. To create image objects characteristic of a maize row for a single date of image acquisition or image stack segmentation, operations for each date and image stack were conducted in separate maps. Use of separate eCognition maps has not been described in previous OBIA studies of UAS agriculture as this step is not necessary for segmentation focused on creating image objects for only one date or image stack. In our analysis, it was necessary to create image objects for each date and image stack where OBIA results were not influenced by the segmentation of another date. This was maintained by performing separate, independent OBIA workflows in a unique map for each date or image stack. This step was a significant contribution of the analysis performed in this study as it allowed for two types of multi-temporal analysis or classification. 1) The first being image stack segmentation and classification where image layers (bands red, green, blue, CHM, etc.) were stacked for a single OBIA segmentation and then summary statistics extracted from multiple image layers from more than one date. 2) The second was a variable stack classification using object properties generated from feature extraction applied to multiple independently generated image object layers produced by separate OBIA workflows but describing a single spatial entity across all images in a timeseries. Each image object layer was unique to the OBIA result of a single date and then combined for classification.

3.7.2 Two-Part OBIA Workflow

In order to generate object properties for multi-temporal image stack and variable stack classification described previously, a new object-based workflow was proposed in order to create image objects descriptive of a single row of maize and also to reduce tedium and subjectivity in trial and error selection of appropriate scale parameter in segmentation. Like methods used by Pérez-Ortiz et al. (2016) and de Castro et al. (2018), our object-based analysis performed a sequential segmentation and object labeling operation for input in a vegetation versus vegetation classification. In this study, object labeling was a binary classification between maize and combined soil and weed image regions, and vegetation versus vegetation classification was the discrimination of maize genotypes. Our proposed two-part OBIA workflow was conducted as a series of segmentation and binary classifications aided by an additional spatial autocorrelation component. This workflow was used to derive object properties descriptive of only one maize row and for the entirety of a row. Our proposed methodology further differs from previous work in object-based agricultural remote-sensing by including a thematic layer of row boundary vectors as input for segmentation.

Within the two-part workflow proposed, Part 1 identified binary classification thresholds through an initial simple, rough classification of maize and statistical assessment of the generated rough maize class. Statistical thresholds are then used for labeling of maize and non-maize in a series of second OBIA operations. Part 2 performed a series of segmentation, binary classification, and segmentation reshaping operations using classification thresholds derived in Part 1. In the following paragraphs (and Figure 11), we provide greater detail as to the steps conducted in the proposed two-step object-based image analysis. The workflow presented here was repeated for each single date and image stack map. For an illustrated example of results produced by OBIA workflow described in the following sections, refer to Figure 12.

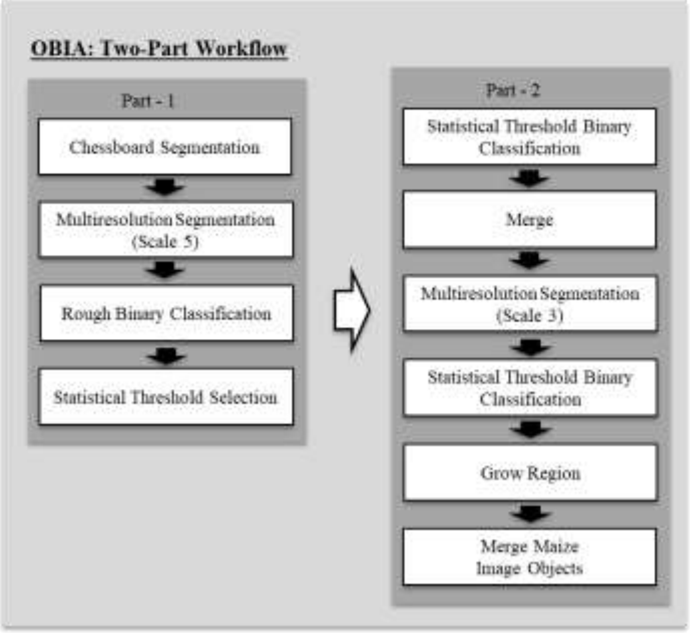


Figure 11: A two-part OBIA workflow for generating row image objects of maize.

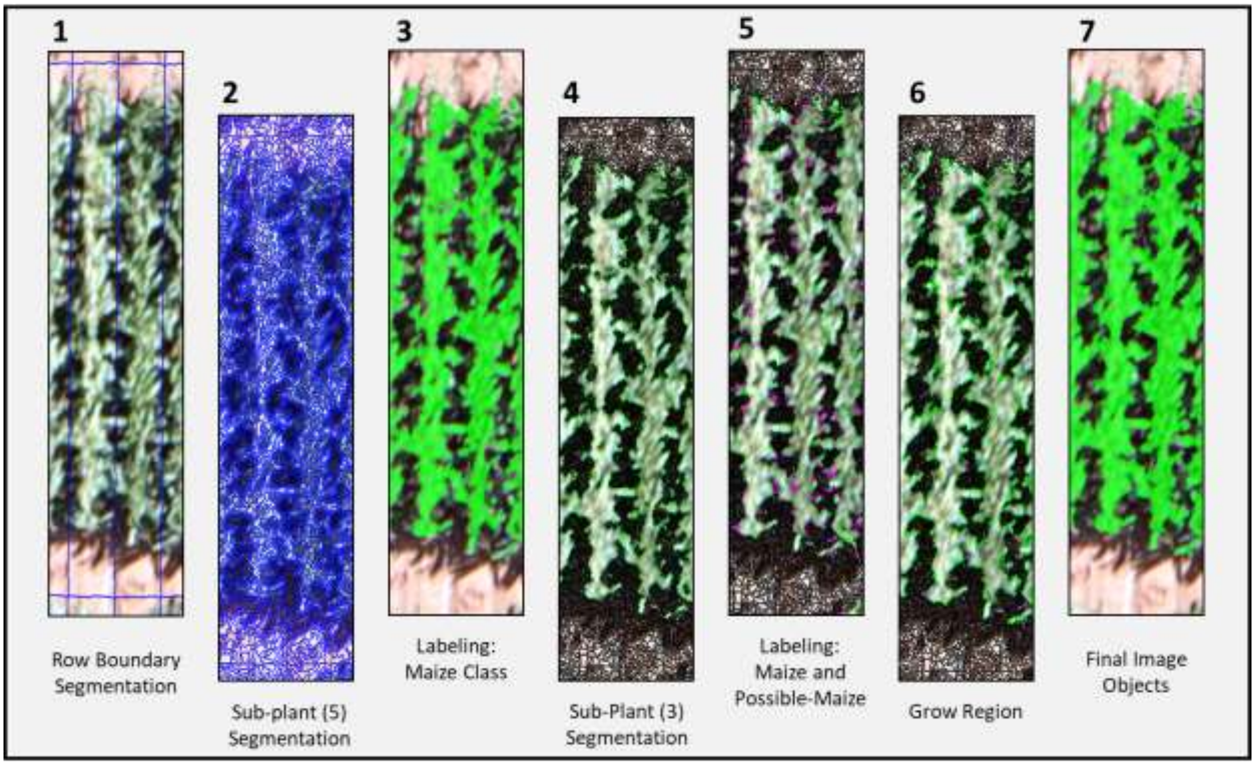


Figure 12: Illustrated OBIA methods performed to create maize row image objects.

3.7.3 Two-Part OBIA Workflow: Part 1

A chessboard segmentation using the thematic layer of row boundary vectors was conducted as the initial operation in Part 1 of the OBIA workflow. Chessboard segmentation was the initial process in the segmentation-classification series. This step was essentially used to convert row boundary information from an eCognition thematic layer (an ESRI shapefile) into image objects within a new map layer for use in segmentation of image objects in following steps. Row boundary image objects within the first layer of the image layer hierarchy provided the spatial extent of objects produced by following segmentation and merging operations. In other words, this step enforced that image objects produced in subsequent steps only describe one row. For chessboard segmentation, the eCognition setting of ‘Object Size’ was set to 999,999 and was performed on a pixel level domain. Object Size of 999,999 was arbitrary number set very high to insure that a single image object was produced for each row boundary vector. Chessboard segmentation was chosen for this step due to algorithm speed and rectangular shape of row boundary polygons.

For the next step, a multiresolution segmentation was performed on an image object level domain produced by the initial chessboard segmentation. A segmentation scale of 5 was selected in order to create sub-plant image objects. Additional inputs included in multiresolution segmentation were the image layers red, green, blue, ExG, ExR, and VEG of a single date or selected dates defined by an image stack. All image layers were given a weight of 1. Default values for parameters of Shape and Compactness (0.1 and 0.5) were used. As this workflow was repeated for eighteen dates and three image stacks, trial and error selection of parameter weights were not practical to evaluate. Based on preliminary trials and supporting evidence by Drăguţ et al. (2014), parameters of shape and compactness had minimal impact on segmentation. CHM was not included as an input for segmentation as preliminary trials using CHM image layers did not produce improved results. Tests for further optimizations of weight, shape, and compactness parameters could be conducted in future research but was not included in this analysis.

Multiresolution segmentation in our workflow produced sub-plant scale image objects. The use of sub-plant image objects allowed for semi-pixel/object-based binary classification of maize and non-maize. This was possible as image objects produced at a scale of 5 were not representative of an entire plant or discrete portion of a plant’s architecture, suggesting a pure object-based approach, but were also not evenly distributed or equally proportional entities as might represent the nature of pixel-based analysis. One could think of our method as a semi-pixel/object-based binary classification of maize and non-maize, as image objects were not representative of an entire plant or discrete portion of a plant's architecture but where, at the same time, image objects were not evenly distributed or equally proportionate which might represent the nature of a pixel-based analysis. Sub-plant image objects provided the ability to select a constant scale parameter for all dates within a first and second segmentation. An additional scale parameter was selected for all image stack segmentation. Use of sub-plant image objects and a series of OBIA operations meant scale parameter did not have to be as refined an input as seen in OBIA methods where a single segmentation on a pixel-based domain is used. Since scale parameter was constant for all single date analysis after initial trial and error selection, selection of appropriate thresholds for binary classification became the crux or flexible component of our analysis.

A rough classification was then conducted on the sub-plant image objects produced by multiresolution segmentation at a scale of 5. A threshold of mean ExG or combination of mean ExG and mean ExR was chosen for each date. We refer to our initial threshold for maize as a rough maize classification, because the threshold was non-exhaustively chosen and intentionally set low so as to include all possible image objects which could describe maize canopy. This threshold was chosen to separate maize from background soil and weeds. ExG was used as the primary threshold for separating maize from non-maize within our study. The VI ExG has proven useful for labeling of vegetation image objects in previous studies of crop versus weed classification (Pérez-Ortiz et al. 2016). We also included ExR and CHM for classification of later dates where weeds and variation in maize leaf color occurred. A statistical assessment was then conducted on the extracted ExG and ExR value of each image object classified as the rough maize class. After calculating standard deviation at percentiles of 10, 16, 20 and 30 for each date, the 16th and 30th percentiles were chosen as binary classification thresholds for labeling of image objects in Part 2. The selected percentiles were not used for all dates due to complexity in maize canopy as plants yellowed with tassels and other emergences and browned with maturity and senescence in later dates; however, selected percentile values were used for most dates and provided a baseline for manual threshold selection. Table 5 provides rough maize threshold values, percentiles, and selected thresholds for each date and image stack. Statistical assessment of rough binary classification was conducted to help reduce subjectivity in selection of thresholds and provide standardized percentiles for all dates for quick threshold selection.

Table 5: Thresholds used for labeling of maize and possible-maize image objects.

Date	Rough Threshold	Maize Threshold	Pos.-Maize Threshold	Percentile	Description
March 21	ExG ≥ 4	ExG ≥ 4.22	2.64 ≤ ExG ≤ 4.22	10th, 16th	
April 07	ExG ≥ 13	ExG ≥ 34.66 and CHM ≥ -10	21.51 ≤ ExG ≤ 34.66 and CHM ≥ -10	30th, 16th	Generic CHM threshold set low to remove non-vegetation.
April 20	ExG ≥ 10	ExG ≥ 42.69 and CHM ≥ -10	26.61 ≤ ExG ≤ 42.69 and CHM ≥ -10	30th, 16th	Generic CHM threshold set low to remove non-vegetation.
May 2	ExG ≥ 15	ExG ≥ 38.58 and CHM ≥ -10	26.81 ≤ ExG ≤ 38.58 and CHM ≥ -10	30th, 16th	Generic CHM threshold set low to remove non-vegetation.
May 5	ExG ≥ 7	ExG ≥ 39.09 and CHM ≥ 0.3	21.28 ≤ ExG ≤ 39.09 and CHM ≥ 0.3	30th, 16th	CHM used to remove soil image objects with high ExG
May 9	ExG ≥ 10	ExG ≥ 35.18 and CHM ≥ -10	22.23 ≤ ExG ≤ 35.18 and CHM ≥ -10	30th, 16th	Generic CHM threshold set low to remove non-vegetation.
May 12	ExG ≥ 10	ExG ≥ 26.93 and CHM ≥ -10	17.03 ≤ ExG ≤ 26.93 and CHM ≥ -10	30th, 16th	Generic CHM threshold set low to remove non-vegetation.
May 15	ExG ≥ 8	ExG ≥ 25.78 and CHM ≥ 0.5	16.50 ≤ ExG ≤ 25.78 and CHM ≥ 0.5	30th, 16th	CHM used to limit presence of weeds included in row image object
May 24	ExG ≥ 8	ExG ≥ 26.88 and CHM ≥ 0.5	16.35 ≤ ExG ≤ 26.88 and CHM ≥ 0.5	30th, 16th	CHM used to limit labeling weeds as maize
May 30	Maize = ExG ≥ 10 and CHM ≥ 0.75	ExG ≥ 60 and CHM ≥ 0.75	NULL	Trial/Error	CHM used to limit labeling weeds and glare on soil as maize. Possible-maize not generated as T/E labeling produced good results.
	Removed = ExG ≤ 10 and ExR > 50	ExG ≤ 60 and ExR ≥ 50 and CHM ≥ 0.75	NULL	Trial/Error	CHM used to limit labeling weeds and glare on soil as maize. Possible-maize not generated as T/E labeling produced good results.
June 02	ExG ≥ 10	ExG ≥ 30.54 and CHM ≥ 0.5	19.61 ≤ ExG ≤ 30.54 and CHM 0.5	30th, 16th	Threshold set to capture green maize canopy.
		ExG ≥ 5.03 and ExR ≥ 28.17 and CHM ≥ 0.5	NULL	5th and 30th	Threshold set to capture yellow or white maize canopy.
June 06	ExG ≥ 7	ExG ≥ 33.38 and CHM ≥ 0.8	20.24 ≤ ExG ≤ 33.38 and CHM ≥ 0.8	30th, 16th	

Table 5 Continued

Date	Rough Threshold	Maize Threshold	Pos.-Maize Threshold	Percentile	Description
June 09	ExG ≥ 5	ExG ≥ 36.17 and CHM ≥ 0.8	21.14 ≤ ExG ≤ 36.17 and CHM ≥ 0.8	30th, 16th	
June 12	Maize = ExG ≥ 10	ExG ≥ 45 and CHM ≥ 0.9	NULL	Trial/Error	Possible-maize not generated as T/E labeling produced good results.
	Removed = ExG ≤ 45 and ExR ≥ 50	10 ≤ ExG < 45 and ExR ≤ 50 and CHM ≤ 0.9	NULL	Trial/Error	Possible-maize not generated as T/E labeling produced good results.
June 16	ExG ≥ 10	ExG ≥ 45.19 and CHM ≥ 0.8	25.73 ≤ ExG ≤ 45.19 and CHM ≥ 0.8	30th, 16th	Threshold set to capture green maize canopy.
		ExG ≥ 13.94 and ExR ≥ 100 and CHM ≥ 1	ExG ≥ -0.22 and ExR ≥ 100 and CHM ≥ 1	5th and Trial/Error	Threshold set to capture browning maize canopy.
June 23	ExG ≥ 20 and CHM ≥ 0.8	ExG ≥ 43.95 and CHM ≥ 0.8	ExG ≥ 32.42 and CHM ≥ 0.8	30th, 16th	Threshold set to capture green maize canopy.
	ExR ≥ 100 and CHM ≥ 1 and ExG ≥ -6 and ExG ≤ 20	0 ≤ ExG ≤ 20 and ExR ≥ 100 and CHM ≥ 1	NULL	Trial/Error	Threshold set to capture browning maize canopy.
June 29	ExG ≥ 2 and CHM (05/15) ≥ 1 and CHM (06/12) ≥ 1	ExG ≥ 14.70 and CHM (05/15) ≥ 0.8 and CHM (06/12) ≥ 0.8	ExG ≥ 6.77 and CHM (05/15) ≥ 1 and CHM (06/12) ≥ 1	30th, 16th	Threshold set to capture green maize canopy.
	ExR ≥ 50 and CHM (05/15) ≥ 1 and CHM (06/12) ≥ 1	ExR ≥ 62.96 and CHM (05/15) ≥ 0.8 and CHM (06/12) ≥ 0.8	ExR ≥ 51.14 and CHM (05/15) ≥ 0.8 and CHM (06/12) ≥ 0.8	30th, 16th	Threshold set to capture browning maize canopy.
July 13	ExR ≥ 65 and CHM (05/15) ≥ 1 and CHM (06/12) ≥ 1	ExR ≥ 85.85 and CHM (05/15) ≥ 0.8 and CHM (06/12) ≥ 0.8	ExR ≥ 75.24 and CHM (05/15) ≥ 0.8 and CHM (06/12) ≥ 0.8	30th, 16th	
3-Date ImStk	NULL	Sum Mean ExG ≥ 75	NULL	Trial/Error	Sum of ExG threshold from dates used in segmentation.
6-Date ImStk	NULL	Sum Mean ExG ≥ 350	NULL	Trial/Error	Sum of ExG threshold from dates used in segmentation.
15-Date ImStk	NULL	Sum Mean ExG ≥ 350	NULL	Trial/Error	Sum of ExG threshold from dates used in segmentation.

The 30th percentile was chosen as a threshold for the class maize. We felt confident that maize could be labeled by a threshold of ExG at the 30th percentile. The 16th percentile of ExG was used to label image objects as a possible-maize class. Possible-maize was not a fuzzy classification but a classification of maize with less confidence. This class provided information for the eCognition algorithm Grow Region which reclassifies objects based on principles of spatial autocorrelation. We discuss this operation in greater detail later in this paper. The class of possible-maize included image objects we considered too low for inclusion in the maize class with confidence but had a high ExG value. In addition, a second threshold of CHM was included in maize and non-maize labeling of all dates except 03/21, CHM was not available for the date 03/21. The CHM threshold was constant for all dates where the statistical threshold selection was applied. It was chosen by a single trial and error analysis and set very low as to only remove soil and avoid removing any potential vegetation.

Labeling maize image objects within multiple images across a growing season presented a complex situation for selecting appropriate classification thresholds. Multiple factors influenced the selection of appropriate thresholds for maize and possible-maize classes across a growing season. Image brightness, glare, shadowing, and presence of weeds varied between dates and influenced optimal thresholds. Physiological properties of maize, such as leaf color, changed across the growing season and varied within the same image for different genotypes. This added complexity in labeling image objects which differed from similar studies such as those in weed and maize classification. In maize-weed classification, maize was classified at a single stage of maturity and early in the growing season when it had very green phenotypic properties. Only one threshold had to be selected (Pérez-Ortiz et al. 2016). For select dates in our study (05/30, 06/12, 06/16, 06/23 and image stack maps), statistically derived thresholds produced by the rough maize class did not prove adequate for separating maize and non-maize image objects. To compensate for additional complexity in images due to presences of weeds, yellowing of tassels, and browning as plants matured, multiple thresholds were used to label maize for these dates. Thresholds of ExR and CHM played an important role here in labeling image objects. Refer to Table 5 for dates and thresholds chosen through trial and error selection.

To classify objects produced by image stack segmentation as maize, a custom feature or property was created in eCognition to combine the mean ExG for each date included in an image stack segmentation. This custom feature was created by calculating the sum of mean ExG for all dates used in segmentation of an image stack map. Image objects with high summed ExG from observed dates were classified as core maize objects or image stack maize; however, this was challenging for rows where masking occurred. Some rows were masked in multiple images and thus had a lower combined value of mean ExG. Date 05/24 was removed from the combined property to create a more meaningful combined value for labeling the image stack maize class. Challenge in selecting a valuable image stack maize class prompted the creation of 3 image stack segmentations where different number of dates selected based on level of image masking were chosen.

3.7.4 Two-Part OBIA Workflow: Part 2

Sub-plant image objects produced by segmentation with a scale of 5 were classified as maize, possible maize, and non-maize using thresholds derived from 30th and 16th percentiles extracted in rough classification statistical assessment above. Image objects produced in Part 1 of the workflow were reclassified after performing the rough classification assessment. Using a threshold for ExG at the 16th percentile and greater, we classified image objects as possible-maize. Maize image objects were then classified using the 30th percentile and greater for ExG. All image objects below the possible-maize class threshold were left unclassified or as non-maize. After labels were applied, all image objects of the maize class were merged within a row boundary. Since the row boundary image object was present in the image layer hierarchy above the image objects created by the multiresolution segmentation with a scale of 5, only maize image objects within a row boundary were able to be merged. This created separate objects for each row. At this stage, all image objects not labeled as maize were also merged.

A third segmentation was then applied to image objects not labeled in the previous step. A multiresolution segmentation was conducted at a scale of 3 on the merged image objects. All segmentation parameters were the same as in the previous segmentation. This step was used to capture potentially isolated maize pixels such as outlying canopy, shadowed leaves, etc. that were missed in the previous segmentation and classification. We performed this step as an extra measure to capture any map regions that were missed. Only image objects not labeled as maize were segmented in this step. Labeling of maize, possible-maize, and non-maize were conducted on image objects created by this segmentation. 16th and 30th percentiles of ExG were again used as thresholds for labeling image objects in this step.

The eCognition algorithm Grow Region was then used to label all image objects of the possible-maize class which were adjacent to the class maize as maize image objects (eCognition Developer 2014). The premise behind the combined process of possible-maize classification and then Grow Region was based on the concept of spatial-autocorrelation. Dense maize canopy, which we had high confidence in our ability to classify, was labeled first using a high threshold of ExG, 30th percentile of rough maize class. We then created the possible-maize class using the 16th percentile of rough maize. Considering the concept of spatial-autocorrelation, we felt image objects of the possible-maize class near or adjacent to the maize class were more likely apart of a maize row than image objects that were isolated from the maize class. In this step, we hoped to eliminate image objects with higher ExG values but not maize (such as glare on soil) while also including maize canopy with lower ExG (shadowed leaves, mixed pixels, etc.). The number of Grow Region cycles was set to one. A second set of Merge Region operations was then applied to all maize image objects. After this, all possible-maize objects remaining were removed from the map and now considered non-maize. Figure 13 provides an example of Grow Region results.

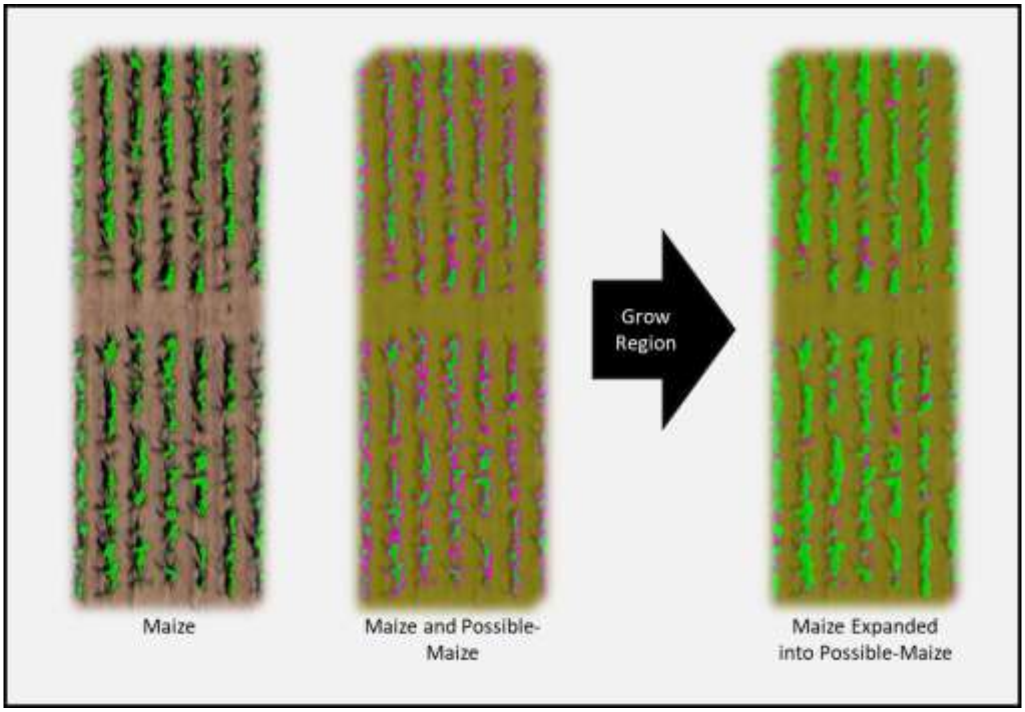


Figure 13: Illustration of the Grow Region operation adding spatially-autocorrelated image objects of the possible-maize class to the maize class.

3.7.5 Feature Extraction

The OBIA workflow described in previous sections provided the geospatial objects from which variables describing each individual row were created for training of remote-sensing classification of maize genotypes. Image objects produced by segmentation and binary classification of each date and image stack described entire or portions of a single row of maize. Multiple image objects describing a single row of maize, as shown in Figure 13, were later combined to produce a single geospatial object or multipart polygon for each row. All image objects within the spatial extent of a row boundary vector were assigned a unique row identification number from that row boundary image object. The object properties generated in eCognition and derived from multiple image objects comprising a row were combined using this row identification number. eCognition feature extraction produces summary statistics of an image object's spatial and spectral characteristics from the image layer values of each pixel within the spatial extent of an object. In this manner, we refer to object properties as the descriptive information created in eCognition that describes a single image object or "patch" as referred to in FRAGSTATS (McGarigal, Cushman, and Ene 2012). Similar terms also include metrics, descriptors, and variables; definitions for each are provided in Table 6. We use the term variable to refer to the descriptive information of combined image objects describing an entire row of maize. The use of the term object property therefore is only used to describe information derived from a single image object and not always an entire row. Variable is used to describe the information derived from all objects or a multipart polygon which comprised an entire row. The terms object property and variable can refer to the same data when a row is comprised of only one image object but when multiple image objects make up a row of maize their definitions differ. Prior to combining image objects into multipart polygons within row boundaries, object properties were created for each date and image stack through eCognition feature extraction. Types of object properties created in eCognition for this analysis included: texture and geometry. Refer to Table 7 for a full list of variables created within this study both within and outside of the software eCognition.

Table 6: Definitions for image units and features derived during OBIA.

Image Units	Class (FRAGSTATS)	FRAGSTATS term which refers to all the patches of a single type; all the image objects labeled as maize that comprise a single row of maize
	Image Object	An image region or cluster of pixels defined during segmentation
	Landscape	FRAGSTATS term which refers to all the patches and classes that make up a row boundary vector in our analysis; this includes both patches of the maize and non-maize classes
	Multipart Polygon	ArcGIS term which refers to a polygon that has more than one parts or has a hole
	Patch	FRAGSTATS term which refers to image objects or individual segmented regions of a map

Image Features	Descriptor	Term used in classification algorithms to refer to variables
	Metric	FRAGSTATS term which refers to the geometrical information describing a single image region at the patch, class, or landscape level
	Object Property	Spatial, textural, and geometrical information derived from the pixels comprising a segmented single image object
	Variable	Spatial, textural, and geometrical information derived from the pixels within image objects comprising an entire row of maize

Table 7: Variables derived from feature extraction and used for classification of maize genotypes.

Variable Group	Variable	Variations	Software Used
CHM			
	Mean	CHM	ArcMap: Zonal Statistics
	Standard Deviation	CHM	ArcMap: Zonal Statistics
Spectral			
	Mean	red, green, blue	ArcMap: Zonal Statistics
	Standard Deviation	red, green, blue	ArcMap: Zonal Statistics
Vegetation Indices			
	Mean	ExG, ExR, VEG, YI	ArcMap: Zonal Statistics
	Standard Deviation	ExG, ExR, VEG, YI	ArcMap: Zonal Statistics
Wavelet			
	Mean	Scale 0.2-5 at 0.2 increments	ArcMap: Zonal Statistics
	Standard Deviation	Scale 0.2-5 at 0.2 increments	ArcMap: Zonal Statistics
Texture			
	GLCM Contrast (45˚)	red, green, blue, ExG, ExR, VEG, YI, CHM, all	eCognition
	GLCM Contrast	All Directions	eCognition
	GLCM Entropy (45˚)	red, green, blue, ExG, ExR, VEG, YI, CHM, all	eCognition
	GLCM Entropy	All Directions	eCognition
	GLCM Mean (45˚)	red, green, blue, ExG, ExR, VEG, YI, CHM, all	eCognition
	GLCM Mean	All Directions	eCognition
Geometrical			
	Area / Number of objects		SQL
	Asymmetry		eCognition
	Border Index		eCognition
	Border Length - Length of longest edge		eCognition
	CA - Total Class Area		FRAGSTATS
	Cohesion - Patch Cohesion Index	mean, area-weighted mean, standard deviation	FRAGSTATS
	CONTIG - Contig Index Distribution	mean, area-weighted mean, range	FRAGSTATS
	Curvature / Length (of main)		SQL
	Degree of skeleton branching		eCognition
	ED - Edge Density		FRAGSTATS
	FRAC - Fractal Dimension Index	mean, area-weighted mean, standard deviation	FRAGSTATS

Table 7 Continued

Variable Group	Variable	Variations	Software Used
	Length		eCognition
	Length / Width		SQL
	Length of longest edge		eCognition
	Length of main line		eCognition
	NDCA -Number of Disjunct Core Areas		FRAGSTATS
	Number of inner objects	Number of image objects within an image objects	eCognition
	Number of objects	Number of image objects comprising the multi-part polygon of a row	eCognition
	PARA - perimeter-area ratio	mean, area-weighted mean, range	FRAGSTATS
	PLAND - Percent of Landscape		FRAGSTATS
	SHAPE	mean, area-weighted mean, standard deviation	FRAGSTATS
	TE - Total Edge		FRAGSTATS
	Width		eCognition

Unlike related studies, we did not create image layer (ex. mean CHM, mean red band, etc.) object properties within eCognition. Since multiple image objects were created for most rows, often rows having a main image object describing the body of the row and smaller image objects which captured smaller row components, image layer statistics were created outside of eCognition on a multipart polygon of combined image objects of a row to prevent producing summary statistics of summary statistics. Using ESRI shapefiles exported from eCognition, the zonal statistics tool in ArcMap (ESRI, Redlands California) was used on multipart polygons. Zonal statistics created image layer variables for each spectral band, vegetation index, CHM, and wavelet analysis scale. Object properties of texture and geometry were exported from eCognition as text files and combined using groupby and join queries in Sequel Query Language. Additional variables, listed in Table 7, were also created outside of eCognition using SQL. Segmentation and binary classification results were further exported from eCognition as raster and ESRI shapefiles.

Additional spatial properties were produced using patch and landscape analysis in the software FRAGSTATS (University of Massachusetts, Amherst). FRAGSTATS produces metrics that describe image objects, referred to as patches, and describe patch shape and relationship to other patches, classes, and the entire landscape. We chose to use class metrics as our goal was to create variables which described all the patches of the maize class within a row boundary polygon. Raster images displaying the binary classification of maize and non-maize were used as the input for FRAGSTATS analysis. FRAGSTATS produced metrics for both classes, maize and non-maize. In order to create separate metrics for individual rows, a user provided tile was created from row boundary vectors and included as an algorithm parameter. Using the row boundary tile, each row was then able to be assessed as an individual landscape. Since only two classes were present, maize and non-maize, patch and landscape metrics were not included in our analysis; only class metrics were used. To create additional core metrics, the edge depth was set to 133.38. This was the equivalent of five pixels after adjusting raster images for FRAGSTATS analysis. Since FRAGSTATS was designed for landscape analysis over large areas, metrics are provided in the measurement unit of hectares. This was problematic to our study as image objects or patches could be as small as 3cm and FRAGSTATS only provides results in hectares and without a sufficient number of decimals. Raster images of binary classification were adjusted in the software ENVI prior to use in FRAGSTATS. Pixel size was multiplied by 100 for all dates, image stacks, and the row boundary raster.

All variables generated for a date or image stack were joined using SQL. Additional groupings of variables were created by joining all the variables from multiple dates together: three dates with no image blur masking, six dates with least amount of masking, and all dates regardless masking level. The three variable groups created in this step were used as datasets for variable stack classification. Variables produced by slight pixelization error when rasterizing shapefiles in eCognition were removed for masked rows by applying the query “delete PLAND < 1” and “delete PLAND is Null”. PLAND refers to percent of the landscape or in our case the percent of the row boundary area. This query removed all rows which had a row object area of less than 1% the size of the row boundary area. This step was applied to all dates with image blur masking. Additional cleanup of data included the removal of ‘gap rows’ or rows with significant gaps in canopy observed at planting. Rows were not removed for genotypes where gaps in canopy appeared during or later in the growing season. Rows with gaps appearing after early dates were included in our dataset as we felt the loss of canopy or lodging of plants could be a characteristic descriptive of some genotypes.

3.8 Genotype Classification

3.8.1 Classification Algorithms

We performed classification of genotypes and evaluated the discriminative value of variables using two algorithms: random forests (RF) and stochastic gradient boosting (SGB). Random forests is a common classification algorithm which has demonstrated high accuracy when compared to other supervised classifiers in object-based studies (Breiman 2001; Ma et al. 2017). Stochastic gradient boosting was selected for its strength in working with small sample sizes and large variable dimensionality (Friedman 2002; Güneralp, Filippi, and Randall 2014). Unlike RF, SGB is not a common algorithm found in the literature on UAS agricultural image classification. RF was chosen as a standard algorithm for comparing genotype classification to related research in the literature and SGB provided a novel method for hyperspatial classification of agriculture. In addition to classification, both algorithms provided ranked variable lists based on variable contribution to classification (Freeman et al. 2015).

Classification was performed using a sample size of four rows per class as only four rows were present within the field experimental condition for each genotype. Training and testing classifiers on a sample size of four points is not ideal for remote-sensing classification but was the data available for this study at the row-level image object classification. We recognize that the limited sample size of this study presents a difficult and non-ideal scenario for remote-sensing classification, as classification is typically performed with sample sizes of 10s or 100s of data points and on fewer number of classes. Classification of 12, 46, 125 and 249 classes using a sample size of 4 as was performed in this study is extreme. Algorithm parameters were therefore chosen to best compensate for the unconventional nature of this dataset; many parameters chosen for our analysis differed from default or selected values found in the literature (Güneralp, Filippi, and Randall 2014; Salford Systems Modeler 2018).

To compensate for small sample size within the data, bootstrapping without replacement was chosen for RF (Güneralp, Filippi, and Randall 2014). Parameters for the number of trees was set to 800, and the number of variables was set to two times the square root of the total number of variables used in the classification (2*sqrt # of variables). Parent node minimum was set to 2. For SGB, a 4-fold cross-validation was used for training and testing classification (Güneralp, Filippi, and Randall 2014). Cross Entropy was selected as criterion determining number of trees optimal for logistic model. Results were recorded as classification accuracy with the default probability cutoff chosen, “assign class if probability exceeds 50%”. Subsample fraction was set to 1. Number of trees to build was set to 400 and learn rate to 0.1. We decided on 400 trees as classification accuracy began to plane off in preliminary trials around 300 trees; 200 trees are the default setting and a suggested starting point in the literature (Salford Systems Modeler 2018). Max nodes per tree was left as the default of 6. Modification of the terminal node in trial and error tests did not improve classification results. Terminal node minimum was set to 1 to compensate for each class only being comprised of four samples. Influence trimming was set to zero. This prevented outliers from being removed from our dataset during classification, but we felt it was necessary to keep all data points due to the small sample size of 4 points which was available for each class in our study. "Vary tree sizes randomly (As Poisson)" was selected to add randomness to the model as partially performed in RF. Both RF and SGB parameters were selected empirically.

3.8.2 Classification Experimental Trials

Three classification experiments were conducted using both RF and SGB algorithms. Experiment 1 compared classification accuracy when the number of classes was reduced from 249 to 12. Classification was assessed on four class sizes: 12, 46, 125, and 249. Experiment 2 was performed to evaluate accuracy and variable importance within variable groups or category. Experiment 3 assessed classification accuracy and variable importance between each single date, image stack, and variable stack. Feature selection was used in Experiments 1 and 3 to improve classification accuracy. Prior to classification, a brief test was conducted to evaluate two feature selection methods.

Prior to classification in Experiments 1-3, the benefit of using feature selection was tested in preliminary trials to evaluate whether classification accuracy improved when using a reduced number of variables. Feature selection methods were tested on a 3-date variable stack using both RF and SGB. Two in-house methods, variable group and in-trial feature selection, were developed and tested for this study. Variable group feature selection selected highest ranked variables from each variable group classified in Experiment 2. In-trial feature selection used the thirty to fifty (30-

50) highest ranked variables selected from an initial all variable classification for each date, image stack, and variable stack. The variable group method was further modified by including an in-trial feature selection on variable groups with more than thirty (30) variables (landscape, texture, and wavelet). We refer to this modified variable group method as an optimized variable group feature selection. For this variable group feature selection, the number of variables from each category was determined by user expertise rather an automatic selection of top 10 variables in each group. In both variable group and in-trial methods, variables were selected based on a variable ranking derived during an initial classification trial. A list of variables ranked by a variable importance score was generated as a result for both RF and SGB algorithms. Variable ranking was influenced by which classes were present, the number of classes, and what algorithm was used for classification.

For this study, we chose to use in-trial feature selection for feature selection in subsequent experiments. Experiments 1 and 3 were conducted using this method. A reduced variable list was thus selected by evaluating relative variable importance scores in an initial classification of each trial and selecting thirty variables for that trial with the highest scores (Dube et al. 2014). Classification for each trial was then conducted a second time but only using the selected 30-50 most valuable variables identified in the first classification. After trial and error assessment on the optimal number of variables to include in the reduced list of variables, the top thirty variables were chosen for all trials with less than 2,000 original variables. Fifty was selected as the ideal number of variables for the reduced variable list when classifying the variable stack trial all-dates in experiment 3 which had over 2,000 original variables. To ensure feature selection was optimal for each trial, this reduced variable list was generated separately for each date, image stack, and variable stack classification.

Experiment 1 evaluated classification accuracy based on the number of classes present within a trial. This step was conducted to see how classification performed on a smaller number of classes (a number of classes comparable to other remote-sensing studies), but also how classification performed on a full field or actual plant breeding scenario. We therefore conducted four trials for this experiment using the following number of classes: 12, 46, 125, and 249. A variable stack or combined group of variables from the three least masked mosaic dates was used as our test sample for trials comparing number of classes. Due to computational limitations of the stochastic gradient boosting software used, a single classification of 280 genotypes was not possible. The largest number of classes which could be handled was 249; therefore, 249 classes were chosen to represent a full field classification scenario. We did not feel this reduction in the number of genotypes was unrepresentative of a true plant breeding scenario as field conditions of non-irrigated and late-planted, also available for the year 2017 but not included in this study, only consisted of 250 genotypes rather than the 280 present in the field condition of an optimal irrigated trial from which we selected our sample data.

A reduced number of genotypes (12, 46, and 125) was produced by evaluating the sample size of each genotype over the growing season after masking of image blur was applied. Masking was applied to 14 of 18 image mosaic dates in this study. Masking lowered the number of rows of maize available in an image to train classification for individual date trials and created null variable information for rows in trials of combined dates. The number of rows for each genotype that were present for an image after masking were summed for all dates across the growing season. The sum of available rows for a genotype across the temporal timeline provided a manner in which to rank genotypes based on presence of sample data. We refer to this step as Mask Ranking. Appendix 3 provides totals for available rows after masking for each date and the Mask Rank value of each genotype. Genotypes were then selected based on natural breaks in the Mask Rank score. Natural breaks occurred at 12, 46, and 125; we also then selected the highest ranked or least masked 249 genotypes. We acknowledge that selected genotypes are not necessarily the most optimal genotype list for all single date analysis as this method provided lower sample sizes for some dates rather than choosing the highest mask ranked genotypes for any particular date; however, multi-temporal mask ranking calculation did allow for the same and genotypes with the best sample data to be selected for comparison of classification for all dates. It also provided genotypes with the largest number of available variables for combined date classification included in experiment 3 trials.

To prevent genotypes with significantly large sample size from creating bias in classification accuracy, five genotypes were removed from this analysis. Except for the genotypes 2, 14, 45, 81, and 238 and genotype class number 59 which only had 1 plot, all other genotypes available for this study had two plots for a total of four rows present in field at planting. Preliminary classification trials tested weighted versus non-weighted sample classes. Balanced and unit were tested for the parameter class weight. Tests also assessed accuracy when genotypes with sample size greater than four rows were present or absent. Individual class and overall accuracy in classification were assessed in preliminary trials. When genotypes with large sample size were included, non-weighted training sample classification produced higher raw accuracy than weighted; however, results of raw accuracy did not reflect ability to classify all classes. Genotypes listed in Table 8 were removed prior to classification.

Table 8: Genotypes removed from sample data prior to classification.

Class Number	Pedigree	**# of Rows**	**Mask Rank**
2	2396/LH123HT	60	928
14	B73/PHZ51	8	114
45	DKB 64-69	6	92
81	LH195/PHZ51	8	126
238	PHW52/PHN82	8	120

The second experiment of our study performed classification using both RF and SGB on variable groupings. Variable group classification was conducted to evaluate the importance each variable type or category contributed toward classification. Variables were grouped into 6 categories: CHM, Landscape/Geometric, Spectral, Texture, Vegetation Indices, and Wavelet. Table 7 on page 52 provides a list of the variables assigned to each category. Variable group classification was performed using the 3-date variable stack of dates 05/12, 06/02, and 06/23; this was the same sample data used in trials assessing class number. 3-date variable stack was chosen to test variable grouping and class number classifications as it provided variables from multiple dates and sample data without image mask reductions. Variable group classification trials in this experiment were conducted both with and without in-trial feature selection.

Trials performed in Experiment 3 produced classification accuracy and variable importance for each individual date of the growing season, for three image stack segmentations, and for three variable stacks or date groupings. All classification was performed on datasets of twelve (12) genotypes. Each trial was conducted using the same list of genotypes for all dates, image stacks, and variable stacks. Classification was conducted on a reduced variable list of top 30 most important variables (top 50 variables for the all-date variable stack trial) derived using our in-trial feature selection method. A reduced variable list was produced for all single date classification as described in previous sections; however, for image stack and variable stack trials a 3-step classification was used to reduce most important variables from a full list to 75 highest ranked variables and then to a final 30 highest ranked variables. On trials where a large variable list was present, this additional variable selection step produced better results. Final classification in each trial produced classification accuracy and ranking of variable importance on the selected variable list.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Feature Selection Results

Preliminary trial and error experimentation demonstrated a reduced variable list consistently produced higher results than a full variable list. Table 9 shows the improved accuracy for the 3-date image stack classification using in-trial feature selection. Feature selection using the in-trial method produced higher accuracy than variable group feature selection where the top ten variables from each variable group classification trial were chosen. Though SGB accuracy was similar between both methods (in-trial accuracy being 2.1% higher than variable group method), RF classification produced results with the in-trail method at an 8.4% difference in accuracy. Table 10 provides results for both in-trial and variable group feature selection methods. In-trial feature selection was also more efficient and automated than the variable group method. The in-trial method required a single dataset for both feature selection and final classification; whereas, variable group required seven datasets for input in classification: 6 variable group datasets and a full dataset with all variables for final classification. In-trial feature selection was conducted via two classification trials. The optimized variable group feature selection did improve variable group

Table 9: Improved accuracy when performing a 12-genotype classification on a reduced variable list.

3-Date Image Stack	RF	SGB	Number of Variables
Full Variable List	12.5%	22.9%	1020
Reduced Variable list	50.0%	68.8%	30

accuracy over in-trial results. As depicted in Figure 14, optimized variable group feature selection produced results with higher accuracy for both algorithms; however, the optimized variable group method was more subjective and less automated than both in-trial and the original variable group feature selection methods. Due to improved automation and similar accuracy, in-trial feature selection was chosen for following experiments over variable and optimized variable group feature selection.

Table 10: Feature selections methods tested using RF and SGB algorithms on the 3-date variable stack.

Feature Selection Methods	RF	SGB
Original Variable Group	33.3%	50.0%
Optimized Variable Group	43.8%	54.2%
In-Trial Method	41.7%	52.1%

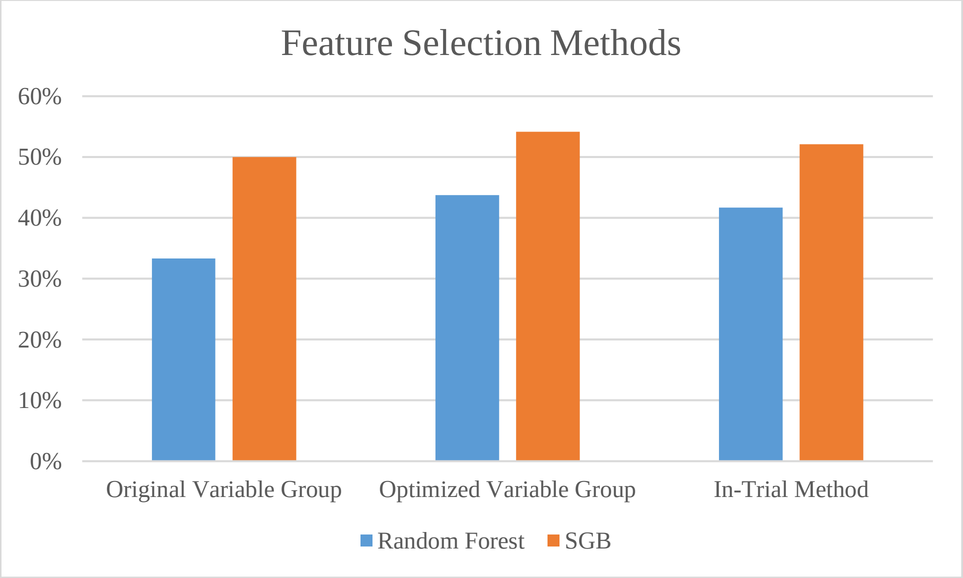


Figure 14: Feature selections methods tested using RF and SGB algorithms on the 3-date variable stack.

4.2 Experiment 1: Class Number Comparison

As demonstrated in Table 11 and Figure 15, an inverse relationship occurs in the number of classes present in a trial and classification accuracy. As the number of classes was reduced from a moderate plant breeding scenario of 249 genotypes down to 12, percent classification accuracy dramatically increased for both algorithms. For both algorithms, classification accuracy was highest for 12 genotype trials and decreased in accuracy for trials of 46, 125, and 249 in that order as class number increased. SGB classification of a 12-genotype trial produced the best accuracy at 52.1%. SGB accuracy diminished more rapidly than RF as class number increased. Accuracy for SGB decreased by 37.1% between 12-gentoype and 249-genotype trials. Whereas RF accuracy for a 12-genotype trial was lower than SGB at 41.7%, classification of 249 genotypes only saw a 12% reduction in accuracy. For all trials except 12-genotype, RF outperformed SGB; therefore, SGB produced both the highest and lowest accuracy within this experiment. Results suggest that RF can better handle classification with a greater number of classes than SGB, but when the number of classes is small, SGB noticeably outperforms RF.

Table 11: RF and SGB algorithm performance on trials of different class size performed on the 3-date variable stack.

Number of Genotypes	RF	SGB
12	41.7%	52.1%
46	35.9%	31%
125	32.0%	16.2%
249	29.7%	15%

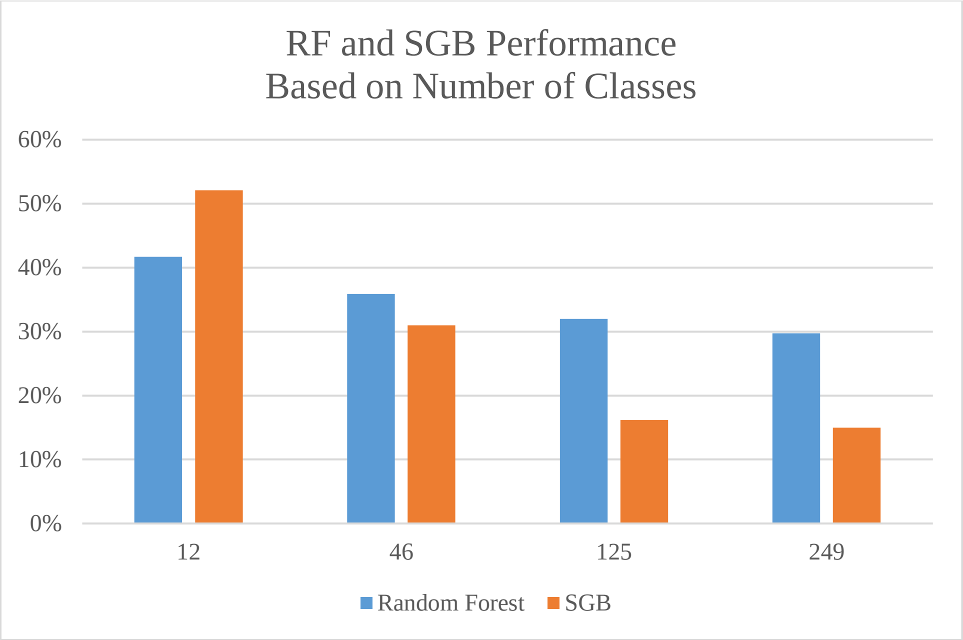


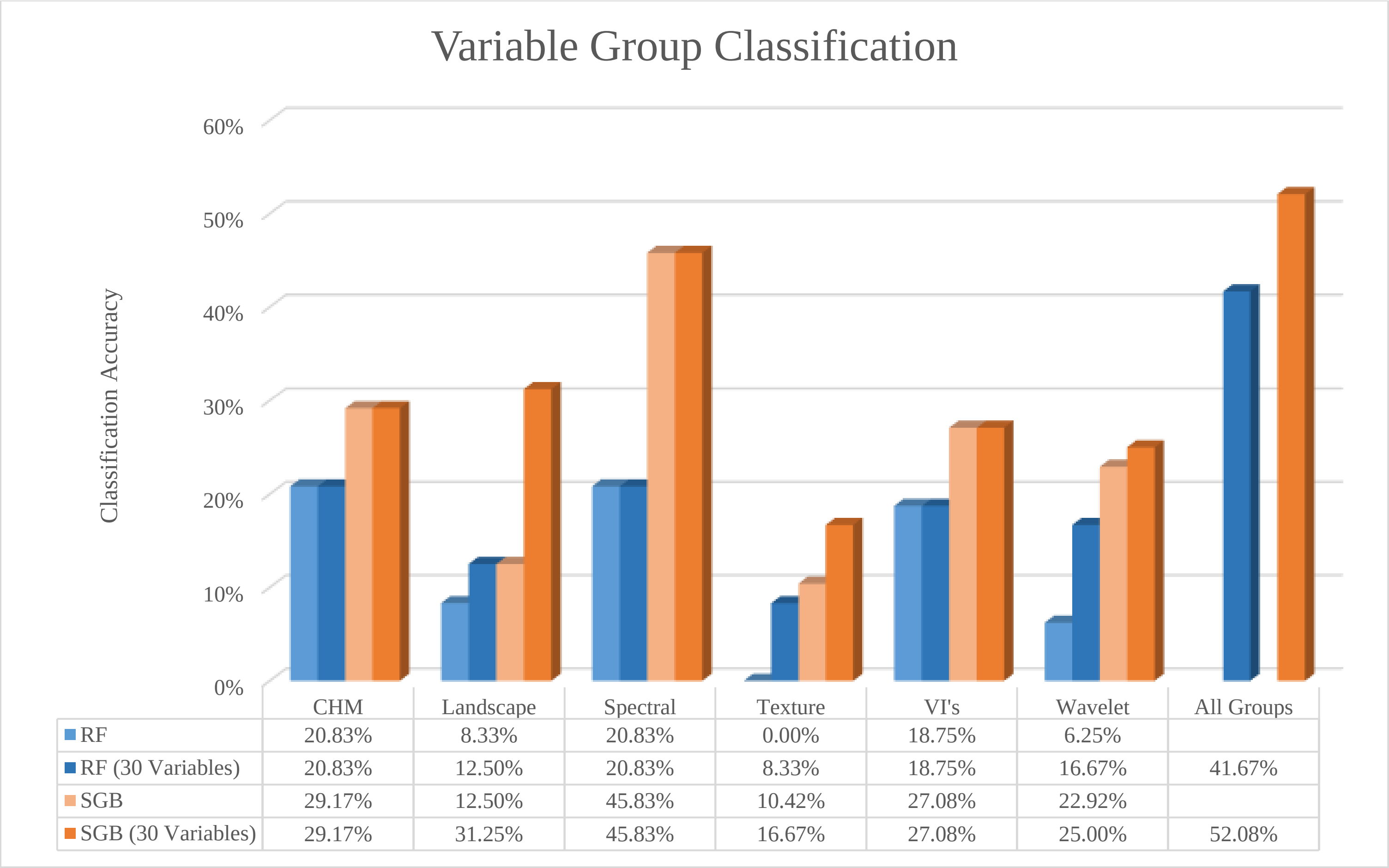
Figure 15: RF and SGB algorithm performance on trials of different class size performed on the 3-date variable stack.

4.3 Experiment 2: Variable Group Classification

Classification using only variables within each variable category known as variable groups (V-Groups) produced lower accuracy than combined variables of all categories applied to a 12-genotype trial in Experiment 1. See Table 12 for variable categories, variable counts per category, and SGB and RF results. The V-Group Spectral produced the highest results at 45.8% using SGB followed by Landscape in one trial (31.3%) and CHM (29.2%); however, accuracy for the Spectral category was 6.3% lower than SGB classification using all V-Groups as test in Experiments 1 and 3. A full variable list of all groups combined performed with In-Trial feature selection had an accuracy of 52.1%. SGB produced higher results for all V-Groups over RF. Figure 16 provides an illustration of SGB and RF accuracy for all V-Groups and a V-Group classification.

Table 12: Variable group classification results and variable counts on the 3-date variable stack.

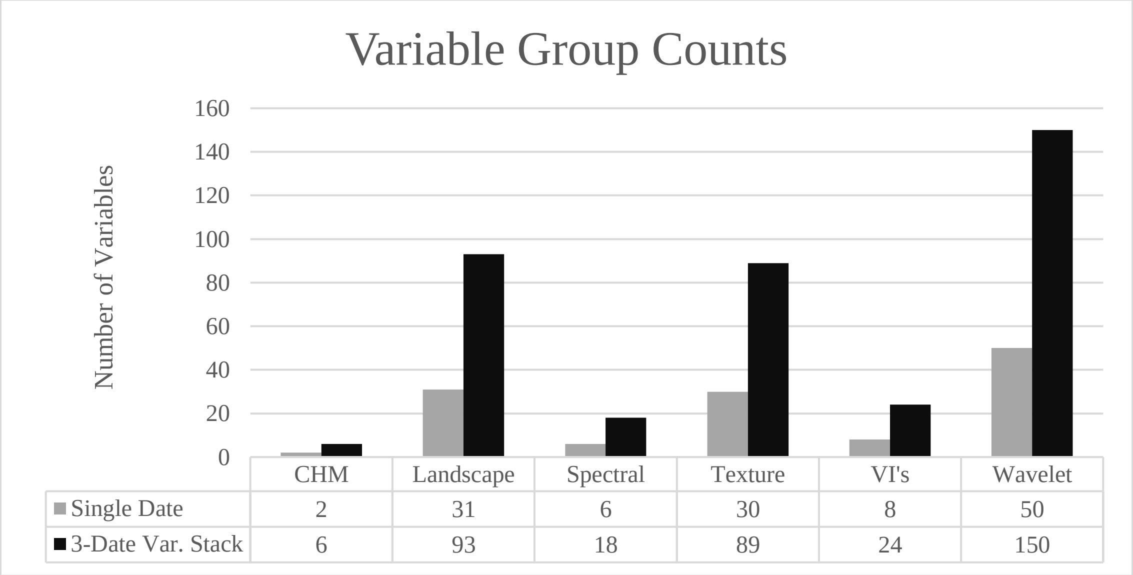
Variable Groupings	Number of Vars.	RF	RF (30 Vars.)	SGB	SGB (30 Vars.)	RF Increase	SGB Increase	Feature Selection
CHM	6	20.8%	**20.8%**	29.2%	**29.2%**	0.0%	0.0%	
Landscape	93	8.3%	**12.5%**	12.5%	**31.3%**	4.2%	18.8%	x
Spectral	18	20.8%	**20.8%**	45.8%	**45.8%**	0.0%	0.0%	
Texture	89	0.0%	**8.3%**	10.4%	**16.7%**	8.3%	6.3%	x
VI's	24	18.8%	**18.8%**	27.1%	**27.1%**	0.0%	0.0%	
Wavelet	150	6.3%	**16.7%**	22.9%	**25.0%**	10.4%	2.1%	x



	CHM	Landscape	Spectral	Texture	VI's	Wavelet	All Groups
■ RF	20.83%	8.33%	20.83%	0.00%	18.75%	6.25%	
■ RF (30 Variables)	20.83%	12.50%	20.83%	8.33%	18.75%	16.67%	41.67%
■ SGB	29.17%	12.50%	45.83%	10.42%	27.08%	22.92%	
■ SGB (30 Variables)	29.17%	31.25%	45.83%	16.67%	27.08%	25.00%	52.08%

Figure 16: Variable Group and all variable classification results on the 3-date variable stack.

Within an initial classification using all variables within a V-Group, V-Groups with a large number of variables performed lower than smaller V-Groups. The V-Groups Landscape, Texture, and Wavelet had over three times the number of variables as did CHM, Spectral, and VI's; see Figure 17. In-Trial feature selection was performed on V-Groups with over 30 variables (Landscape, Texture, and Wavelet). Figure 16 illustrates the improved accuracy of reducing the number of variables in each category. Increase in accuracy after feature selection suggest that variable importance of some V-Groups could be lost due to large variable numbers and potential variable correlation within categories. All V-Groups on which feature selection was applied improved in accuracy. Landscape jumped from the second lowest V-Group to the second highest after feature selection. This was possibly due to the Landscape group having the largest number of variables to begin with. The V-Group Texture, though improved by feature selection, produced the lowest results before and after feature selection was applied.



	CHM	Landscape	Spectral	Texture	VI's	Wavelet
■ Single Date	2	31	6	30	8	50
■ 3-Date Var. Stack	6	93	18	89	24	150

Figure 17: Variable counts for each variable group on a single date and 3-date variable stack trial.

Results from V-Group classification suggested that Spectral and CHM categories were significant contributors to classification in all variable trials as they produced higher accuracy than other V-Groups, but also did so on a smaller number of variables. Results from variable group classification suggested that texture was the smallest contributor to classifying genotypes at the row level. This is not to say that texture variables are not valuable for classification since variable ranking results in Experiment 3 suggest that individual texture variables score higher than variables of other categories in some trials and prove specifically beneficial in image stack classification.

In 3-date variable stack used for V-Group classification, each variable was present three times as a descriptor for training classification, once for each date (e.g. Mean_CHM_05/12, Mean_CHM_06/02, Mean_CHM_06/23, etc.). Variable rank results for V-Groups Landscape and CHM on a 3-date variable stack classification are provided in Tables 13 and 14 as an example indicate that variable importance changes based on algorithm. For example, the variable "Length by Width" had a very different ranking produced by SGB for each date (1st, 26th and 38th out of ninety-three (93) landscape variables). RF ranked the Landscape V-Group differently and ordered Length by Width as the 11th, 54th, and 84th most important variable contributing to classification. In both algorithms Length by Width for the date 06/23 ranked the highest; however, Length by Width for the date 05/12 was ranked as more important than 06/02 in SGB versus RF. However in the case of the CHM V-Group shown in Table 15, mean CHM for all dates from a 3-date variable stack had a higher importance score than standard deviation of CHM using SGB. These results suggest that variable ranking is dependent on the date described and somewhat on algorithm when discriminative value is low. Reference to Appendix D for a full list of variable names and function.

Table 13: SGB variable rankings for landscape variable group.

Rank		Rank		Rank		Rank	
1	LG_BY_WD_0623	13	ED_0512	25	WIDTH_0512	37	PARA_AM_0602
2	CURV_BY_LGT_0623	14	LENGTH_LONG_EDGE_0512	26	LG_BY_WD_0512	38	LG_BY_WD_0602
3	SHAPE_AM_0602	15	#_INNER_OBJS_0602	27	PARA_AM_0512	39	CONTIG_MN_0602
4	SHAPE_SD_0512	16	PARA_MN_0512	28	LGT_MAIN_0623	40	SHAPE_MN_0512
5	LENGTH_LONG_EDGE_0623	17	CURV_BY_LGT_0602	29	LGT_MAIN_0602	41	#_OF_OBJS_0602
6	ASYMMETRY_0512	18	CONTIG_AM_0602	30	FRAC_SD_0512	42	FRAC_MN_0623
7	AREA_BY#_OBJ_0623	19	LENGTH_LONG_EDGE_0602	31	PARA_MN_0602	43	TE_0512
8	#_OF_OBJS_0623	20	NDCA_0623	32	AREA_BY#_OBJ_0602	44	SHAPE_MN_0602
9	BORDER_INDEX_0602	21	NDCA_0512	33	FRAC_AM_0602	45	WIDTH_0623
10	BORDER_INDEX_0512	22	BL_SUB_LLED_0512	34	BL_SUB_LLED_0602	46	CONTIG_AM_0512
11	FRAC_MN_0512	23	ASYMMETRY_0602	35	COHESION_0602	47	FRAC_MN_0602
12	PARA_RA_0623	24	LENGTH_0623	36	LGT_MAIN_0512	48	PARA_AM_0623

Table 13 Continued

Rank		**Rank**		**Rank**		**Rank**	
49	LENGTH_0602	61	CONTIG_AM_0623	73	#_INNER_OBJS_0623	85	PARA_RA_0512
50	#_INNER_OBJS_0512	62	DEG_SKL_BRCH_0602	74	WIDTH_0602	86	CA_0602
51	BL_SUB_LLED_0623	63	LENGTH_0512	75	FRAC_AM_0623	87	TE_0623
52	SHAPE_AM_0623	64	NDCA_0602	76	CONTIG_MN_0512	88	COHESION_0623
53	ED_0602	65	FRAC_SD_0623	77	CA_0623	89	CONTIG_RA_0602
54	SHAPE_AM_0512	66	FRAC_AM_0512	78	ED_0623	90	PLAND_0512
55	BORDER_INDEX_0623	67	#_OF_OBJS_0512	79	CONTIG_RA_0512		
56	CONTIG_RA_0623	68	TE_0602	80	SHAPE_SD_0602		
57	SHAPE_MN_0623	69	AREA_BY#_OBJ_0512	81	PARA_RA_0602		
58	SHAPE_SD_0623	70	FRAC_SD_0602	82	CONTIG_MN_0623		
59	COHESION_0512	71	PLAND_0623	83	PLAND_0602		
60	ASYMMETRY_0623	72	CURV_BY_LGT_0512	84	PARA_MN_0623		

Table 14: RF variable rankings for landscape variable group.

Rank		Rank		Rank		Rank	
1	SHAPE_AM_0602	13	#_OF_OBJS_0623	25	TE_0623	37	CONTIG_AM_0623
2	FRAC_AM_0602	14	PARA_RA_0623	26	PARA_MN_0623	38	CONTIG_MN_0602
3	ASYMMETRY_0512	15	BL_SUB_LLED_0623	27	SHAPE_MN_0623	39	ASYMMETRY_0602
4	FRAC_MN_0512	16	SHAPE_SD_0512	28	BL_SUB_LLED_0602	40	AREA_BY#_OBJ_0602
5	CURV_BY_LGT_0623	17	#_INNER_OBJS_0602	29	#_OF_OBJS_0602	41	BL_SUB_LLED_0512
6	ED_0623	18	ED_0602	30	CONTIG_MN_0623	42	SHAPE_AM_0623
7	AREA_BY#_OBJ_0623	19	TE_0602	31	FRAC_MN_0602	43	FRAC_AM_0623
8	LGT_MAIN_0602	20	LGT_MAIN_0512	32	PARA_AM_0623	44	PARA_AM_0512
9	LENGTH_0623	21	CONTIG_RA_0623	33	BORDER_INDEX_0623	45	PARA_RA_0602
10	LENGTH_LONG_EDGE_0623	22	LENGTH_LONG_EDGE_0512	34	CONTIG_AM_0602	46	SHAPE_SD_0602
11	LG_BY_WD_0623	23	COHESION_0623	35	FRAC_SD_0623	47	ED_0512
12	BORDER_INDEX_0602	24	SHAPE_SD_0623	36	COHESION_0602	48	LENGTH_LONG_EDGE_0602

Table 14 Continued

Rank		Rank		Rank		Rank	
49	LGT_MAIN_0623	61	LENGTH_0602	73	FRAC_AM_0512	85	WIDTH_0512
50	SHAPE_MN_0512	62	ASYMMETRY_0623	74	CA_0623	86	PLAND_0512
51	FRAC_SD_0512	63	AREA_BY#_OBJ_0512	75	WIDTH_0623	87	WIDTH_0602
52	CONTIG_RA_0602	64	NDCA_0623	76	#_OF_OBJS_0512	88	PARA_RA_0512
53	PARA_MN_0602	65	PLAND_0602	77	CURV_BY_LGT_0512	89	#_INNER_OBJS_0512
54	LG_BY_WD_0602	66	NDCA_0512	78	PLAND_0623	90	COHESION_0512
55	FRAC_MN_0623	67	TE_0512	79	CA_0512	91	DEG_SKL_BRCH_0623
56	PARA_MN_0512	68	SHAPE_MN_0602	80	CA_0602	92	DEG_SKL_BRCH_0512
57	PARA_AM_0602	69	NDCA_0602	81	SHAPE_AM_0512	93	DEG_SKL_BRCH_0602
58	CONTIG_AM_0512	70	#_INNER_OBJS_0623	82	FRAC_SD_0602		
59	BORDER_INDEX_0512	71	CONTIG_RA_0512	83	LENGTH_0512		
60	CURV_BY_LGT_0602	72	CONTIG_MN_0512	84	LG_BY_WD_0512		

Table 15: SGB and RF variable ranking for CHM variable group.

Rank	SGB - CHM Variable Group
1	MEAN_CHM_0623
2	MEAN_CHM_0512
3	MEAN_CHM_0602
4	STD_CHM_0512
5	STD_CHM_0602
6	STD_CHM_0623

Rank	RF - CHM Variable Group
1	MEAN_CHM_0623
2	MEAN_CHM_0602
3	STD_CHM_0512
4	MEAN_CHM_0512
5	STD_CHM_0623
6	STD_CHM_0602

4.4 Experiment 3: Single Date, Image Stack, and Variable Stack Classification

SGB yielded better performance over RF in all single date, image stack, and variable stack classifications with the exception of one single date trial, 06/06, as shown in Table 16. SGB had an average 13.5% improvement in accuracy over RF across all trials. In the single occasion where RF outperformed SGB, RF demonstrated less than 2% advantage in accuracy. The best results for RF were produced on the trial all-date variable stack with an accuracy of 60.4%, 6.3% lower than SGB classification of the same trial. SGB classification of 6-date image stack produced the highest accuracy of all trials at 79.2%. This trial was further modified to include model randomization to improve accuracy to 81.3% (parameters predictors per node and predictors per tree were both set to 15). This allowed for only 15 variables of the 30 available variables of the reduced feature list to be available for training each decision tree and each node in a decision tree. Model randomization parameters of predictors per node and predictors per tree were not used in any other trial as preliminary analysis did not demonstrate consistent benefits in results across multiple classifications. Results with and without model randomization are provided for 6-date stack classification as this trial demonstrated highest accuracy. When comparing variable importance, results are assessed for 6-date image stack with model randomization included.

Table 16: Correct Classification of 12 maize genotypes over a growing season using RF and SGB on single date, image stack, and variable stack datasets. Highest accuracy for each algorithm are shown in bold.

Variable Groups	RF	SGB	Initial Num. of Variables	Final Num. of Variables	Wavelet Present
Single Dates					x = present
03/21	8.3%	22.9%	72	30	
04/07	12.5%	22.9%	77	30	
04/20	18.8%	18.8%	77	30	
05/02	18.8 %	27.1%	127	30	x
05/05	20.8%	33.3%	127	30	x
05/09	25.0%	31.3%	127	30	x
05/12	20.8%	41.7%	127	30	x
05/15	16.7%	29.2%	127	30	x
05/24	18.8%	33.3%	127	30	x
05/30	10.4%	31.3%	127	30	x
06/02	20.8%	43.8%	127	30	x
06/06	**39.3%**	37.5%	127	30	x
06/09	34.2%	**50.0%**	127	30	x
06/12	12.5%	35.4%	127	30	x
06/16	20.8%	29.2%	127	30	x
06/23	14.6%	29.2%	127	30	x
06/29	15.9%	20.8%	127	30	x
07/13	23.1%	**50.0%**	77	30	
Image Stack					
3 Date Image Stack	50.0%	68.8%	1480	30	x
6 Date Image Stack	56.3%	79.2%	1480	30	x
6 Date Image Stack (MR)	N/A	**81.3%**	1480	30	x - Model Randomization
15 Date Image Stack	50.0%	66.7%	1480	30	x
Variable Stack					
3 Date Variable Stack	41.7%	52.1%	381	30	x
6 Date Variable Stack	35.4%	50.0%	762	30	x
All Date Variable Stack	60.4%	**66.7%**	2081	50	x

SGB classification of all-date variable stack yielded similar accuracy of 66.7 and 68.8% to 3-date image stack and 15-date image stack trials. Higher accuracy in image stack over variable stack trials suggested that a smaller number of multi-temporal descriptive landscape variables are more beneficial for classification than landscape variables derived from each single date segmentation, but landscape variables in image stack trials were ranked significantly lower than those in variable stack trials. Within image stack trials of both RF and SGB, texture variables ranked much higher than in variable stack or single date trials. Variables in image stack trials can describe regions of maize canopy, core regions of a maize canopy, and possible non-maize areas for a given date. Description of core canopy regions of maize and non-maize regions in some images may be a contributor of improved accuracy.

Results clearly revealed multi-temporal dimensionality **significantly** contributed to higher accuracy as depicted in Figure 18. Improved or equal accuracy in 10 out of 12 image stack and variable stack trials support the use of multi-temporal dimensionality for classification of maize genotypes. For multi-temporal trials of SGB, average accuracy was 64.2%, 14% higher than the highest single date trial of either algorithm. The 6-date variable stack produced the lowest result for multi-temporal SGB trials but was equal to 06/09 and 07/13 which yielded the highest accuracy at 50% of any single date trials. Average accuracy for multi-temporal SGB trials was more than 30% higher than average single date accuracy at 32.6%. Multi-temporal RF trials also yielded results with an average accuracy of 30% greater than single date RF. Performance in multi-temporal RF trials 3-date variable stack and 6-date variable stack were lower than three single date SGB trials (06/02, 06/09, and 07/13); however, only the RF trial 06/06 outperformed the RF multi-temporal trial of 6-date variable stack. Mean accuracy across all single date trials was 19.6% for

RF and 32.6% SGB. Single date classification did not achieve accuracy above 50%. Whereas accuracy of 60% and greater was achieved for five multi-temporal trials, one RF and three SGB.

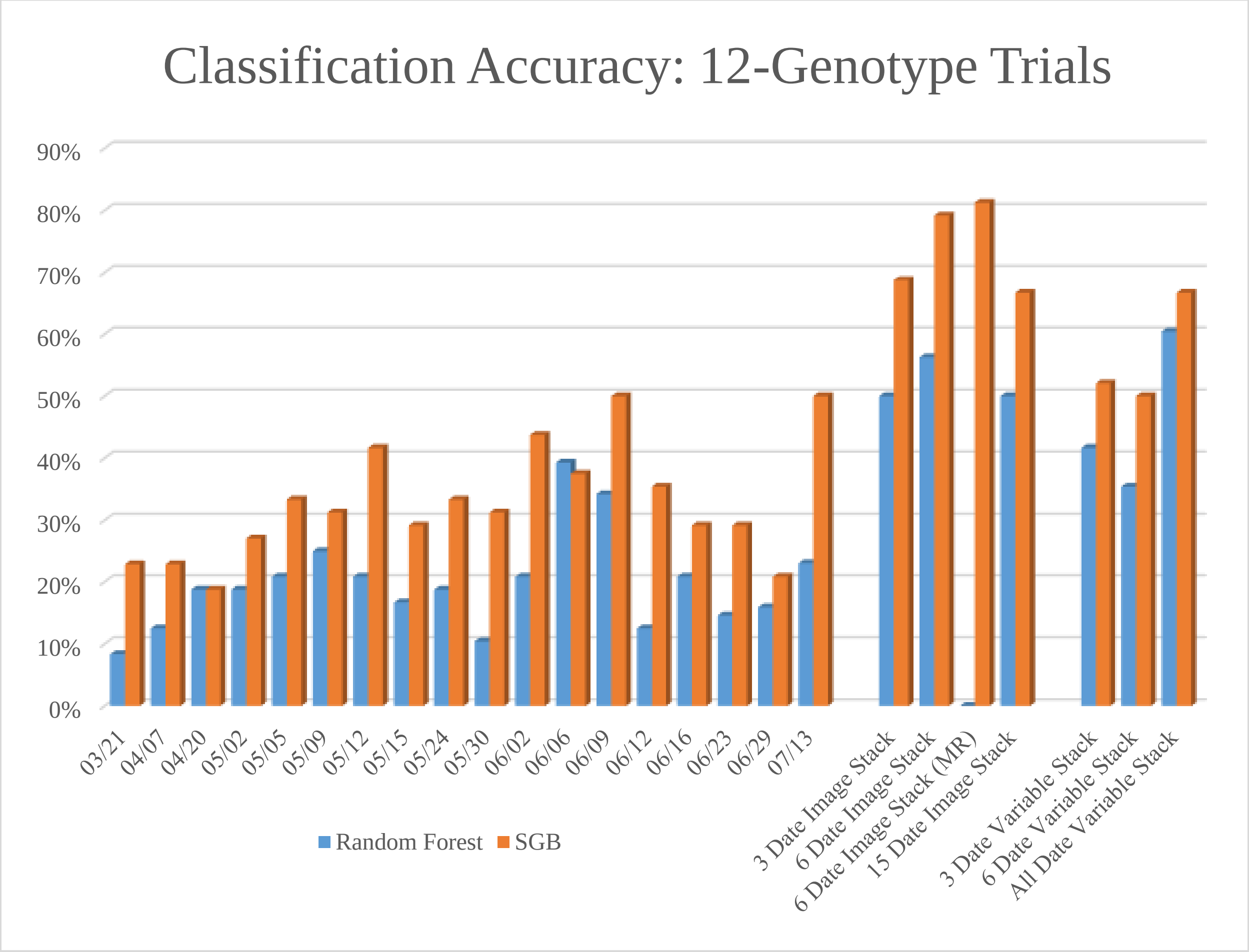


Figure 18: Classification accuracy shown for single date, image stack, and variable stack classification of 12 genotype trials.

When assessing both algorithms, single date accuracy was lower early in the growing season. The earliest date with accuracy above 40% was achieved on May 12. Results for single date trials at the given spatial resolution between 2.66-2.83cm and temporal resolution of this study suggest accurate classification of genotypes using RGB images of a single date is not yet possible.

A general bell-shaped curve in accuracy shown in Figure 18 suggest that mid to later season images provide the most beneficial information for classification. This is due to substantially more differentiation between mature plants that can be visually observed in the field, potentially due to differences accumulating in plants growth and response cumulatively up to that point and likely also due to more structure in the plant that can be used to discriminate in adult plants (more nodes, more leaves, tassels) than in juvenile plants. Accuracy within single date trials of 40-50% suggested variables from single date classification can provide valuable information for future plant breeding analysis and should be assessed further.

4.4 Variable Evaluation of Experiment 3

In this section, variables were evaluated in two ways: 1) by importance ranking and 2) variable occurrence within and across trials. Variable rank was evaluated based on relative score for each single date, image stack, and variable stack, specifically within the most accurate multi-temporal trials (6-date image stack and all-date variable stack). Variable occurrence was measured by 1) calculating the sum of each variable's presence within the top 10 variables of all trials by algorithm, 2) variable occurrence within the combined reduced variable lists of both RF and SGB algorithms for the all-date variable stack, and 3) examining the top 10 variables of SGB and RF on all-date variable stack.

When assessing variable importance by variable ranking and relative score, individual variable and variable group contribution to classification differed noticeably between datasets and algorithms. The order of variables ranked in RF trials differed from SGB. Variable rank results for the top ten (10) variables from each date and multi-temporal classification shown in Tables 17 and 19 indicated that variable importance changed based on date and algorithm. Some variables received a high score in both trials and in some cases the same rank such as in trials 05/05, 05/30, and 06/02 where mean CHM was ranked as the most important descriptor. In fact, CHM variables (mean and standard deviation) were consistently ranked high across all trials and for both algorithms. Mean CHM was the highest ranked variable in 21 out of 48 trials between RF and SGB algorithms. Within single date trials, the spectral variable 'mean blue band' also showed high ranking in several trials. Overall however, importance of individual variables varied widely between trials. Though variables of the landscape category occurred among the highest ranked variables in multiple trials, the same variable or variables did not show up as a consistently important contributor as noticed in variables mean CHM and mean blue. Variable ranking of both RF and SGB suggest that all variable categories provided valuable information for classification. As a trend, texture variables were ranked lower than all other variable groups except in image stack trials. Wavelet variables were generally ranked high in most trials where available. The wavelet variables 'mean - scale of 1.4' and 'standard deviation - scale of 1.2' showed repeated use across several trials. Though wavelet variables showed importance across many trials, highest ranked variables of this category differed between multiple scales. The number of wavelet scales included in a feature list should be reduced in future work to evaluate the importance of individual wavelet scales. This could also further improve classification by reducing the number of correlated variables being used in a trial. VI's provided valuable information in all trials with ExG variables showing prominence over ExR, VEG, and YI. In single date trials where ExR and YI were high, ExG tended to have a lower score than trials where these variables were less important. Reference to Appendix D for a full list of variable names and function.

Table 17: Top 10 variables from each SGB classification of 12 genotype trial.

Var. Groups	03/21	04/07	04/20	05/02	05/05	05/09
CHM	BORDER_INDEX	MEAN_BLUE	MEAN_BLUE	ASYMMETRY	MEAN_CHM	MEAN_CHM
Landscape	PARA_RA	STD_EXG	LENGTH_LONG_EDGE	WV_SCALE__5STD	BORDER_INDEX	WV_SCALE__1_6MEAN
Spectral	SHAPE_AM	SHAPE_SD	MEAN_RED	STD_BLUE	GLCM_MEAN_ALL_DIR	WV_SCALE__2_2STD
Texture	DEG_SKL_BRCH	GLCM_MEAN_B_45˚	CA	WV_SCALE__0_6STD	MEAN_RED	ASYMMETRY
VI's	MEAN_EXR	LENGTH	NDCA	GLCM_CONTRAST_45˚	GLCM_CONTRAST_ALL_DIR	MEAN_BLUE
Wavelength	MEAN_EXG	STD_CHM	MEAN_CHM	BORDER_INDEX	WV_SCALE__3_2STD	TE
	LGT_MAIN	GLCM_CONTRAST_EXR_45˚	STD_EXR	MEAN_EXG	MEAN_BLUE	MEAN_GREEN
	GLCM_CONTRAST_ALL_DIR	BL_SUB_LLED	MEAN_EXG	AREA_BY_#OBJ	WV_SCALE__3_4STD	CURV_BY_LGT
	FRAC_AM	MEAN_CHM	LENGTH	WV_SCALE__4_6MEAN	WV_SCALE__1_6STD	SHAPE_AM
	CURV_BY_LGT	GLCM_MEAN_CHM_45˚	GLCM_MEAN_B_45˚	MEAN_CHM	WV_SCALE__1_4STD	PARA_RA

05/12	05/15	05/24	05/30	06/02	06/06
FRAC_MN	MEAN_CHM	MEAN_BLUE	MEAN_CHM	MEAN_CHM	STD_CHM
STD_GREEN	MEAN_YEL	MEAN_CHM	WV_SCALE__3_4MEAN	MEAN_BLUE	MEAN_CHM
MEAN_GREEN	SHAPE_SD	SHAPE_SD	WV_SCALE__0_8STD	WV_SCALE__1_4MEAN	STD_BLUE
MEAN_CHM	GLCM_MEAN_G_45˚	MEAN_RED	MEAN_GREEN	MEAN_EXG	WV_SCALE__1_8STD
MEAN_RED	MEAN_RED	WV_SCALE__1_2STD	GLCM_CONTRAST_EXG_45˚	WV_SCALE__0_2STD	MEAN_GREEN
STD_BLUE	MEAN_EXR	COHESION	LGT_MAIN	LGT_MAIN	WV_SCALE__5STD
MEAN_EXG	FRAC_SD	STD_GREEN	WV_SCALE__1STD	SHAPE_AM	GLCM_MEAN_EXG_45˚
WV_SCALE__0_2MEAN	WV_SCALE__1_2STD	LENGTH_LONG_EDGE	FRAC_MN	WV_SCALE__1_2STD	GLCM_MEAN_YEL_45˚
STD_RED	STD_CHM	CA	WIDTH	WV_SCALE__3_4STD	WV_SCALE__1_2MEAN
GLCM_CONTRAST_EXG_45˚	GLCM_CONTRAST_ALL_DIR	NDCA	WV_SCALE__1_4STD	LENGTH_LONG_EDGE	WV_SCALE__2_4MEAN

Table 17 Continued

Var. Groups	06/09	06/12	06/16	06/23	06/29	07/13
CHM	MEAN_BLUE	MEAN_CHM	STD_EXR	MEAN_CHM	PARA_AM	STD_EXG
Landscape	MEAN_CHM	ASYMMETRY	MEAN_CHM	SHAPE_AM	STD_CHM	PARA_AM
Spectral	STD_EXG	MEAN_EXG	LGT_MAIN	MEAN_BLUE	ASYMMETRY	STD_GREEN
Texture	WV_SCALE__2_8STD	GLCM_MEAN_CHM_45˚	MEAN_VEG	LG_BY_WD	WV_SCALE__1_2STD	CURV_BY_LGT
VI's	LENGTH_LONG_EDGE	SHAPE_SD	GLCM_MEAN_YEL_45˚	GLCM_CONTRAST_EXR_45˚	MEAN_CHM	LGT_MAIN
Wavelength	COHESION	GLCM_MEAN_B_45˚	STD_YEL	AREA_BY_#OBJ	LENGTH	FRAC_MN
	WV_SCALE__4_8STD	PARA_RA	WV_SCALE__4_8STD	PARA_RA	GLCM_CONTRAST_45˚	MEAN_CHM
	PARA_AM	NDCA	WV_SCALE__2_2STD	#_of_Objs	CA	PARA_RA
	CURV_BY_LGT	WV_SCALE__0_6STD	WV_SCALE__0_2MEAN	CURV_BY_LGT	AREA_BY_#OBJ	LG_BY_WD
	MEAN_GREEN	MEAN_EXR	STD_RED	WV_SCALE__5MEAN	FRAC_SD	SHAPE_MN

3-Date Image Stack	6-Date Image Stack	15-Date Image Stack	3-Date Variable Stack	6-Date Variable Stack	All-Date Variable Stack
MEAN_EXR_0524	MEAN_CHM_0515	GLCM_CONTRAST_B_0623_45˚	MEAN_CHM_0623	FRAC_MN_0602	COHESION_0524
STDEV_EXR_0407	MEAN_EXR_0524	GLCM_ENTROPY_B_0505_45˚	WV0602_SCALE__1_4MEAN	WV0602_SCALE__1_4MEAN	STD_BLUE_0505
MEAN_CHM_0515	WV0509SC1_4_MEAN	GLCM_ENTROPY_YEL_0505_45˚	LG_BY_WD_0623	STD_YEL_0616	WV0602_SCALE__1_2STD
GLCM_MEAN_CHM_0623_45˚	STDEV_YEL_0616	STDEV_CHM_0512	MEAN_RED_0512	MEAN_YEL_0515	BORDER_INDEX_0321
GLCM_MEAN_EXR_0321_45˚	WV0524SC1_4_STD	GLCM_MEAN_CHM_0713_45˚	LGT_MAIN_0602	WV0616_SCALE__2_2STD	MEAN_GREEN_0713
WV0502SC0_2_3D_STD	MEAN_CHM_0616	GLCM_ENTROPY_EXR_0509_45˚	WV0602_SCALE__0_2STD	WV0602_SCALE__0_2STD	GLCM_MEAN_CHM_0612_45˚
MEAN_CHM_0713	GLCM_CONTRAST_45˚_IMSTK_6D	WV0515SC1_8_MJ_STD	MEAN_BLUE_0602	GLCM_MEAN_G_0515_45˚	WV0509_SCALE__4_4STD
MEAN_EXG_0515	GLCM_CONTRAST_B_0629_45˚	MEAN_EXR_0524	SHAPE_AM_0602	WV0602_SCALE__1_2STD	MEAN_CHM_0606
MEAN_CHM_0609	WV0515SC1_8_STD	MEAN_CHM_0609	WV0602_SCALE__1_8MEAN	GLCM_CONTRAST_EXR_0616_45˚	MEAN_EXG_0713
GLCM_ENTROPY_B_0713_45˚	MEAN_G_0509	MEAN_CHM_0612	GLCM_MEAN_ALL_DIR_0602	MEAN_CHM_0515	MEAN_RED_0616

Table 18: Top 10 variables from each RF classification of 12 genotype trial.

Var. Groups	03/21	04/07	04/20	05/02	05/05	05/09
CHM	DEG_SKL_BRCH	STD_CHM	MEAN_BLUE	MEAN_CHM	MEAN_CHM	LENGTH
Landscape	BORDER_INDEX	STD_EXG	MEAN_CHM	GLCM_CONTRAST_EXR_45˚	MEAN_RED	MEAN_CHM
Spectral	SHAPE_AM	MEAN_CHM	STD_EXG	ASYMMETRY	MEAN_BLUE	WV_SCALE__1STD
Texture	FRAC_AM	MEAN_EXR	LENGTH_LONG_EDGE	WV_SCALE__0_6STD	GLCM_MEAN_ALL_DIR	MEAN_GREEN
VI's	LENGTH_LONG_EDGE	GLCM_CONTRAST_EXR_45˚	FRAC_SD	AREA_BY_#OBJ	GLCM_CONTRAST_B_45˚	ED
Wavelength	PARA_RA	MEAN_BLUE	MEAN_EXG	SHAPE_MN	WV_SCALE__0_2MEAN	WV_SCALE__1_6MEAN
	MEAN_EXG	FRAC_AM	STD_GREEN	WV_SCALE__0_8STD	GLCM_CONTRAST_45˚	PLAND
	SHAPE_MN	FRAC_SD	MEAN_RED	WV_SCALE__1_4STD	MEAN_VEG	CONTIG_RA
	SHAPE_SD	CA	LENGTH	WV_SCALE__1STD	GLCM_CONTRAST_R_45˚	MEAN_EXG
	CURV_BY_LGT	GLCM_CONTRAST_YEL_45˚	STD_VEG	GLCM_MEAN_CHM_45˚	MEAN_GREEN	MEAN_BLUE

05/12	05/15	05/24	05/30	06/02	06/06
MEAN_CHM	MEAN_EXR	MEAN_RED	MEAN_CHM	MEAN_CHM	STD_CHM
MEAN_GREEN	MEAN_YEL	MEAN_CHM	LGT_MAIN	MEAN_BLUE	MEAN_CHM
MEAN_VEG	MEAN_CHM	MEAN_EXR	WV_SCALE__1STD	SHAPE_AM	MEAN_BLUE
STD_VEG	MEAN_BLUE	MEAN_YEL	MEAN_RED	STD_BLUE	LENGTH_LONG_EDGE
STD_GREEN	FRAC_SD	MEAN_BLUE	GLCM_CONTRAST_EXG_45˚	WV_SCALE__1_4MEAN	GLCM_MEAN_EXG_45˚
FRAC_MN	CURV_BY_LGT	LENGTH_LONG_EDGE	MEAN_VEG	WV_SCALE__1_6MEAN	WV_SCALE__1_8STD
STD_CHM	WV_SCALE__5MEAN	WV_SCALE__1_2STD	MEAN_GREEN	GLCM_MEAN_R_45˚	MEAN_GREEN
MEAN_RED	MEAN_RED	COHESION	MEAN_EXG	WV_SCALE__2MEAN	STD_RED
STD_RED	GLCM_MEAN_G_45˚	LENGTH	STD_CHM	WV_SCALE__1_8MEAN	GLCM_MEAN_YEL_45˚
ASYMMETRY	WV_SCALE__4_8MEAN	STD_CHM	SHAPE_SD	WV_SCALE__0_2STD	MEAN_VEG

Table 18 Continued

Var. Groups	06/09	06/12	06/16	06/23	06/29	07/13
CHM	MEAN_CHM	MEAN_CHM	MEAN_CHM	MEAN_CHM	MEAN_YEL	STD_EXG
Landscape	MEAN_BLUE	ASYMMETRY	STD_YEL	STD_EXG	CA	CONTIG_AM
Spectral	GLCM_ENTROPY_CHM_45˚	LENGTH_LONG_EDGE	STD_EXR	MEAN_VEG	PLAND	CURV_BY_LGT
Texture	MEAN_GREEN	MEAN_EXG	GLCM_CONTRAST_EXG_45˚	STD_BLUE	MEAN_EXR	STD_GREEN
VI's	MEAN_VEG	SHAPE_SD	LGT_MAIN	STD_GREEN	STD_CHM	PLAND
Wavelength	WV_SCALE__1_8MEAN	WV_SCALE__2_2STD	GLCM_CONTRAST_YEL_45˚	LENGTH_LONG_EDGE	MEAN_CHM	GLCM_MEAN_EXG_45˚
	WV_SCALE__1_4MEAN	STD_CHM	GLCM_MEAN_YEL_45˚	CURV_BY_LGT	ASYMMETRY	MEAN_CHM
	WV_SCALE__2_4MEAN	ED	MEAN_VEG	MEAN_GREEN	MEAN_EXG	LGT_MAIN
	STD_EXG	WV_SCALE__2STD	STD_RED	LENGTH	LENGTH	PARA_AM
	GLCM_MEAN_CHM_45˚	WV_SCALE__2_4STD	GLCM_CONTRAST_B_45˚	AREA_BY_#OBJ	NDCA	CA

3-Date Image Stack	6-Date Image Stack	15-Date Image Stack	3-Date Variable Stack	6-Date Variable Stack	All-Date Variable Stack
MEAN_EXG_0509	GLCM_CONTRAST_B_0629_45˚	MEAN_CHM_0609	MEAN_CHM_0623	MEAN_CHM_0616	STD_GREEN_0524
MEAN_EXR_0524	MEAN_CHM_0616	GLCM_CONTRAST_B_0629_45˚	WV0602_SCALE__1_6MEAN	STD_YEL_0616	MEAN_CHM_0606
MEAN_CHM_0609	MEAN_CHM_0515	GLCM_CONTRAST_B_0623_45˚	SHAPE_AM_0602	MEAN_EXR_0515	GLCM_MEAN_CHM_0612_45˚
MEAN_YEL_0524	MEAN_R_0524	GLCM_ENTROPY_EXG_0602_45˚	WV0602_SCALE__1_4MEAN	STD_EXR_0616	MEAN_EXG_0713
MEAN_EXG_0530	STDEV_YEL_0616	GLCM_CONTRAST_R_0623_45˚	FRAC_AM_0602	SHAPE_AM_0602	WV0602_SCALE__1_2STD
STDEV_EXR_0407	MEAN_CHM_0609	MEAN_CHM_0505	MEAN_CHM_0602	MEAN_YEL_0515	STD_BLUE_0616
MEAN_CHM_0515	MEAN_B_0524	STDEV_YEL_0616	WV0602_SCALE__1_8MEAN	WV0602_SCALE__1_8MEAN	BL_SUB_LLED_0629
MEAN_CHM_0616	MEAN_YEL_0524	GLCM_CONTRAST_EXR_0509_45˚	STD_BLUE_0602	MEAN_CHM_0623	WV0602_SCALE__1_4STD
GLCM_MEAN_VEG_0602_45˚	GLCM_ENTROPY_EXG_0509_45˚	MEAN_YEL_0524	MEAN_VEG_0512	WV0602_SCALE__1_4MEAN	SHAPE_MN_0602
STDEV_YEL_0616	MEAN_EXR_0524	GLCM_MEAN_B_0515_45˚	STD_VEG_0512	STD_BLUE_0602	COHESION_0524

Variables were also evaluated based on repeated occurrence in the reduced variable list of each trial. Tables 19 provides a sum of the number times each variable by algorithm occurred within the top 10 highest ranked variable list of each date. For both RF and SGB, mean CHM, mean blue, and mean ExG were the top three most commonly occurring variables within the two summary lists of top ten variable rankings. Mean CHM occurred 30 times in the top ten ranking of RF trials and 27 times in SGB trials. Mean blue occurred 10 times in RF trials and 9 times in SGB. ExG shared a similar occurrence as mean blue. As can be seen, variable occurrence decreased sharply after mean CHM and varied for following variables between trials. For the spectral category of variables, mean demonstrated greater importance over standard deviation based on occurrence. Variable types mean and contrast for the Texture V-Group occurred more often and with a higher rank than the variable type entropy. Tables 20 and 21 provide further summary statistics as to the number variables present in each V-Group.

Table 19: Occurrence of variable by algorithm within the top 10 ranked variables of all dates, image stack and variable stack trials.

SGB – Variables	Count	V-Groups	RF - Variables	Count
MEAN_CHM	27	CHM	MEAN_CHM	30
MEAN_BLUE	9	Landscape	MEAN_BLUE	10
MEAN_EXG	7	Spectral	MEAN_EXG	9
MEAN_RED	7	Texture	MEAN_GREEN	8
MEAN_GREEN	7	VI's	MEAN_VEG	8
LGT_MAIN	6	Wavelength	STD_GREEN	8
MEAN_EXR	6		MEAN_EXR	7
WV_SCALE__1_2STD	6		MEAN_YEL	7
GLCM_MEAN_CHM_45˚	5		STD_CHM	7
SHAPE_AM	5		STD_EXG	7
STD_CHM	5		LENGTH_LONG_EDGE	6
CURV_BY_LGT	5		MEAN_RED	6
PARA_RA	5		GLCM_CONTRAST_B_45˚	5
ASYMMETRY	4		LENGTH	5
BORDER_INDEX	4		STD_BLUE	5
FRAC_MN	4		STD_YEL	5
LENGTH_LONG_EDGE	4		ASYMMETRY	4

Table 19 Continued

SGB – Variables	Count	V-Groups	RF - Variables	Count
SHAPE_SD	4	CHM	CURV_BY_LGT	4
STD_BLUE	4	Landscape	FRAC_AM	4
WV_SCALE__0_2STD	4	Spectral	SHAPE_AM	4
WV_SCALE__1_4MEAN	4	Texture	WV_SCALE__1_4MEAN	4
NDCA	3	VI's	WV_SCALE__1_8MEAN	4
STD_EXG	3	Wavelength	CA	3
AREA_BY_#OBJ	3		FRAC_SD	3
CA	3		GLCM_CONTRAST_EXR_45°	3
COHESION	3		PLAND	3
GLCM_CONTRAST_ALL_DIR	3		SHAPE_MN	3
GLCM_MEAN_B_45°	3		SHAPE_SD	3
LENGTH	3		STD_VEG	3
LG_BY_WD	3		WV_SCALE__1_6MEAN	3
PARA_AM	3		WV_SCALE__1STD	3
STD_GREEN	3		AREA_BY_#OBJ	2
WV_SCALE__2_2STD	3		COHESION	2
GLCM_CONTRAST_45°	3		ED	2
STD_EXR	3		GLCM_CONTRAST_EXG_45°	2
STD_YEL	3		GLCM_CONTRAST_R_45°	2
WV_SCALE__1_4STD	3		GLCM_CONTRAST_YEL_45°	2
WV_SCALE__1_8STD	3		GLCM_ENTROPY_EXG_45°	2
STD_RED	2		GLCM_MEAN_CHM_45°	2
GLCM_CONTRAST_EXR_45°	2		GLCM_MEAN_EXG_45°	2
GLCM_MEAN_G_45°	2		GLCM_MEAN_YEL_45°	2
FRAC_SD	2		LGT_MAIN	2
GLCM_CONTRAST_B_45°	2		WV_SCALE__1_4STD	2
GLCM_CONTRAST_EXG_45°	2		BL_SUB_LLED_0629	1
GLCM_ENTROPY_B_45°	2		BORDER_INDEX	1
GLCM_MEAN_ALL_DIR	2		CONTIG_AM	1
GLCM_MEAN_EXR_45°	2		CONTIG_RA	1
GLCM_MEAN_YEL_45°	2		DEG_SKL_BRCH	1
MEAN_YEL	2		GLCM_CONTRAST_45°	1
WV_SCALE__0_2MEAN	2		GLCM_ENTROPY_CHM_45°	1
WV_SCALE__0_6STD	2		GLCM_MEAN_ALL_DIR	1
WV_SCALE__3_4STD	2		GLCM_MEAN_B_45°	1
WV_SCALE__4_8STD	2		GLCM_MEAN_CHM_0612_45°	1
#_of_OBJS	1		GLCM_MEAN_G_45°	1
BL_SUB_LLED	1		GLCM_MEAN_R_45°	1
DEG_SKL_BRCH	1		GLCM_MEAN_VEG_45°	1
FRAC_AM	1		NDCA	1
GLCM_ENTROPY_EXR_45°	1		PARA_AM	1
GLCM_ENTROPY_YEL_45°	1		PARA_RA	1
GLCM_MEAN_EXG_45°	1		STDEV_EXR_0407	1
MEAN_VEG	1		WV_SCALE__0_2MEAN	1
TE	1		WV_SCALE__0_2STD	1

Table 19 Continued

SGB – Variables…continued	Count	V-Groups	RF - Variables…continued	Count
WIDTH	1	CHM	WV_SCALE__0_6STD	1
WV_SCALE__0_8STD	1	Landscape	WV_SCALE__0_8STD	1
WV_SCALE__1_2MEAN	1	Spectral	WV_SCALE__1_2STD	1
WV_SCALE__1_6MEAN	1	Texture	WV_SCALE__1_8STD	1
WV_SCALE__1_6STD	1	VI's	WV_SCALE__2_2STD	1
WV_SCALE__1_8MEAN	1	Wavelength	WV_SCALE__2_4MEAN	1
WV_SCALE__1STD	1		WV_SCALE__2_4STD	1
WV_SCALE__2_4MEAN	1		WV_SCALE__2MEAN	1
WV_SCALE__2_8STD	1		WV_SCALE__2STD	1
WV_SCALE__3_2STD	1		WV_SCALE__4_8MEAN	1
WV_SCALE__3_4MEAN	1		WV_SCALE__5MEAN	1
WV_SCALE__4_4STD	1			
WV_SCALE__4_6MEAN	1			
WV_SCALE__5MEAN	1			
WV_SCALE__5STD	1			
SHAPE_MN	1			

Table 20: Variables of each variable group present in the Top 10 ranked variables from each SGB trial.

Variable Group	Sum top 10 variables (all trials)	Sum all variables (all trials)	Percent of Top 10 out of all variables
CHM	32	116	27.59%
Landscape	71	2975	2.39%
Spectral	32	576	5.56%
Texture	33	2340	1.41%
VI's	25	624	4.01%
Wavelet	45	3100	1.45%

Table 21: Variables of each variable group present in the Top 10 ranked variables from each RF trial.

Variable Group	Sum top 10 variables (all trials)	Sum all variables (all trials)	Percent of Top 10 out of all variables
CHM	37	116	31.90%
Landscape	58	2975	1.95%
Spectral	29	576	5.03%
Texture	30	2340	1.28%
VI's	47	624	7.53%
Wavelet	29	3100	0.94%

One issue with assessing most commonly occurring variables from an all trial comparison is that trials with low accuracy have potential to skew variable importance. Within early dates where classification produced very low accuracy, such low classification accuracy suggests none or few of the variables of this trial are very significant for classification regardless of ranking. This could prove problematic when identifying the most valuable variables based on occurrence across and between all trials. The assessment of variables within trials yielding more accurate results could provide more valuable information regarding variable importance. A combined RF and SGB variable assessment on the trial all-date variable stack was chosen to shed more light as to the most important variables based on occurrence and ranked order. All-date variable stack produced the highest accuracy when looking at both RF and SGB trials and did so on a reduced list of 50 rather than 30 variables. This trial was ideal for assessing variable importance based on variable occurrence as this trial took into consideration all variables from all dates. Within each trial (RF and SGB), multiple instances of the same variable but from a different date could occur in the variable ranking. Image stack trials, though demonstrating higher accuracy for SGB than the all-date variable stack, were not ideal for assessing variable importance across our timeseries as image stack trials work with a reduced number of landscape variables. This is due to image stack trials having only one OBIA result from which variables are extracted while variable stack trials extract variables from multiple image object layers. Using a variable stack trial allows for multiple occurrences of landscape variables that image stack trials cannot provide; therefore, when assessing variables in a variable stack trial, landscape variables are not underrepresented when compared to other V-Groups. Assessment of image stack variables would have been problematic as classification in these trials was performed with a smaller list of multi-temporal landscape variables. Variable importance was also considered in this assessment by comparing variable rank based on date. A particular date could therefore stand out as a greater contributor to classification.

Table 22: Occurrence of variables in combined SGB and RF reduced variable list of all-date variable stack trials.

Variables	Count
STD_BLUE	10
BORDER_INDEX	5
COHESION	5
MEAN_CHM	5
MEAN_EXG	4
MEAN_GREEN	4
WV_SCALE__5STD	4
MEAN_RED	3
MEAN_YEL	3
SHAPE_AM	3
DEG_SKL_BRCH	2
GLCM_MEAN_CHM_45°	2
GLCM_MEAN_VEG_45°	2
MEAN_EXR	2
MEAN_VEG	2
PARA_MN	2
SHAPE_MN	2
STD_EXG	2
STD_EXR	2
STD_GREEN	2
WV_SCALE__1_2STD	2
WV_SCALE__1_4STD	2
WV_SCALE__1_6STD	2
WV_SCALE__3_4MEAN	2
#_of_OBJS	1
continued…Variables	**Count**
AREA_BY_#OBJ	1
ASYMMETRY	1
BL_SUB_LLED	1
CA	1
CONTIG_MN	1
CONTIG_RA	1
CURV_BY_LGT	1
ED	1
FRAC_MN	1
GLCM_CONTRAST_B_45°	1
GLCM_ENTROPY_CHM_45°	1
GLCM_ENTROPY_YEL_45°	1
GLCM_MEAN_B_45°	1
LENGTH_LONG_EDGE	1
NDCA	1
SHAPE_SD	1
STD_RED	1
STD_VEG	1
WV_SCALE__0_2MEAN	1
WV_SCALE__0_8STD	1
WV_SCALE__1_2MEAN	1
WV_SCALE__2_2MEAN	1
WV_SCALE__2_2STD	1
WV_SCALE__3_6STD	1
WV_SCALE__4_4STD	1

When assessing variable importance based on variable occurrence in 100 combined variables of 50 RF and 50 SGB all-date variable stack rankings, mean CHM and mean ExG were in the top 5 most common variables as shown in Table 22. Unlike the sum of top 10 variables from all trials, mean CHM and mean ExG however were not the highest occurring variables. The highest ranked variables by occurrence in these combined trials were standard deviation of the blue band at a frequency of 10, border index at 5, cohesion at 5, and mean CHM at 5 followed by mean ExG at 4. The only variable group not represented within the top 10 most occurring variables was texture; however, the texture variable ‘GLCM mean CHM 06/12’ did occur twice within the top highest ranked variables by relative importance score in both RF and SGB all-date variable stack trials; see Table 23. There were also eight occurrences of texture variables within the 100 variable occurrence assessment of the combined RF and SGB all-date variable stacks; refer to Tables 22 and 24. The repeated occurrence of the landscape variable ‘cohesion’ and its rank based on relative score (1 in SGB and 10 in RF out of 2081 variables available prior to feature selection) suggest future research should investigate use of the variable cohesion for high-throughput plant breeding. The cohesion variable is a measure of how connected image objects labeled as maize are to one another in a row boundary. This variable could help provide plant breeders with information as to how clumped maize canopy is across a row. The custom VI of Yellow Index was present in the top 10 of 100 variables of the combined RF and SGB variable occurrence assessment. Yellow Index for date 06/16 was ranked 20th by relative importance in the SGB all-date variable stack trial. When comparing the top 10 variables in all-date variable stack trials of RF and SGB as shown in Table 23, 5 variables are correlated between trials (cohesion, standard deviation of wavelet scale 1.2, GLCM mean CHM, mean CHM, and mean ExG). Results from variable assessment support that all variable groups provided beneficial information for classification. One interesting thing to note is that variable importance in all-date variable stack trials did not always correlate to variable ranking in single date trials. For example, the highest ranked variable for the SGB all-date variable stack trial cohesion extracted from date 05/24 was ranked 6th in the single date trial of 05/24. The variable standard deviation of the blue band from date 05/05 was ranked second in all-date variable stack but twenty-first in single date trial 05/05. Assessing variable importance based on rank and presence in RF and SGB all-date variable stack reduced variable list is meaningful as these trials consider all variables from all dates. If a particular variable is considerably more valuable for classification, it should occur more often in a reduced feature list of selected variables. In addition, variables in these trials must compete against all other variables for relative importance. Greater competition in multi-temporal trials supports the use of identified variables in this assessment.

Table 23: Top 10 variables by importance for both SGB and RF all-date variable stack trials.

SGB: Top 10 Variables	RF: Top 10 Variables
COHESION_0524	STD_GREEN_0524
STD_BLUE_0505	MEAN_CHM_0606
WV0602_SCALE__1_2STD	GLCM_MEAN_CHM_0612_45˚
BORDER_INDEX_0321	MEAN_EXG_0713
MEAN_GREEN_0713	WV0602_SCALE__1_2STD
GLCM_MEAN_CHM_0612_45˚	STD_BLUE_0616
WV0509_SCALE__4_4STD	BL_SUB_LLED_0629
MEAN_CHM_0606	WV0602_SCALE__1_4STD
MEAN_EXG_0713	SHAPE_MN_0602
MEAN_RED_0616	COHESION_0524

Table 24: Number of variables by V-Group for combined top 50 SGB and RF reduced feature lists.

Variable Group	Total out of 100 (50SGB + 50RF)	Number of Variables in V-Group	Percent of Variable Group
CHM	5	34	14.71%
Landscape	32	1116	2.87%
Spectral	18	216	8.33%
Texture	8	1080	0.74%
VI's	16	288	5.56%
Wavelet	19	1500	1.27%
Total all available Variables (SGB + RF trials)	100	4234	

4.5 Discussion of Potential Error

Prior to concluding the discussion of results provided in this chapter, we acknowledge potential sources of error and limitations of this study such as small sample of data available for training and running classification algorithms and variation in image properties or quality due to multi-temporal acquisition and mosaicking. Within this study, classification was performed on a sample size of four rows or multipart polygons per genotype. Conventional remote-sensing classification is performed using training samples ranging from 10s to 1000s of sample points. To compensate for the limited data available for this study, RF classification was trained using bootstrapping without replacement, and SGB was performed using a 4-fold cross validation. We acknowledge that classification accuracy may be influenced by overfitting due to the limited pool of training data available for classification. As training data was not separate from the dataset classified because bootstrapping and cross validation sampling methods were used, this could also introduce error. It is, however, our assessment based on this study that genotype classification is possible using only RGB-derived data and that methods proposed in this study present valuable information which should be applied for future genotype classification.

Potential error could also be introduced due to varying spatial resolution of images collected across the timeseries of this study. Pixel resolution ranging between 2.66-2.83cm could affect accuracy of segmentation and value of object properties used in classification. In addition, no radiometric corrections were applied as all data was provided as RGB images with digital numbers. Contrast in pixel brightness and image blur were also noticed after mosaicking and orthorectification. Large areas of significant image blur were manually removed prior analysis; however, blur did occur at lesser intensity in various portions of many images. Finally, we acknowledge that potential error could also have been introduced by changing solar geometry during data acquisition as time of day and month changed.

CHAPTER 5

CONCLUSION

In this study, a multi-temporal object-based image analysis assessment of UAS images for classification of maize genotypes was introduced. A semi-automated OBIA workflow was designed for separate segmentation and labeling of 18 images and 3 multi-temporal image stacks. In addition, this study presents methodology for image object variable stacking, a complex phenomenon as changing maize canopy produces varying number and located image objects over the multi-temporal scenario of a maize growing season. Machine-learning-algorithms random forests (RF) and stochastic gradient boosting (SGB) were tested for classification of maize genotypes on a dataset of very small sample size, four row-scale image objects per class. Viability of methods and algorithms employed by this study were tested on three classification experiments: four trials with different number of classes, variable group trials, and 12 genotype classification of each individual date, image stack, and variable stack.

Results show that classification of maize genotypes on a study size of 12 classes can achieve accurate results as high as 81% using SGB; however, an inverse relationship between the number of classes and classification accuracy was demonstrated as the number of genotypes included in a trial increased. This suggests that object-based multi-temporal classification of RGB UAS images can be applied for accurate SGB classification of maize genotypes on studies with small number of classes. Results from this study did not achieve high accuracy for classification of 46, 125, and 249 genotypes, which better represent a full plant breeding scenario, but high accuracy in 12 genotype classification using only RGB-derived data, improved accuracy when applying feature selection, and demonstrated benefit in multi-temporal dimensionality suggest analysis performed in this study have potential for future research on a larger number of classes. Whereas SGB consistently outperformed RF on trials of 12 genotypes, results indicate RF classification performs better than SGB as the number of classes included in a trial increase. SGB machine learning performed better than RF in all 12-genotype trials but one for each single date, image stack, and variable stack assessments. This study supports the use of SGB over RF for small class number classification of maize genotypes.

Multi-temporal dimensionality of image stack and variable stack classification in this study demonstrate multi-temporal analysis improves classification. Results for multi-temporal trials demonstrated an average accuracy of 30% greater than single date classification. Accurate single date classification was not achieved by this study. The earliest point classification was achieved in this study with 40% and greater accuracy was on the date May 12th. Accuracy increased up to 50% by the date June 9th. Highest accuracy for both single date and multi-temporal trials was achieved using SGB. Classification of image objects produced through image stack segmentation provided the highest results at 81.8% followed by all-date variable stack classification. Low accuracy in single date classification and significantly higher accuracy in multi-temporal classification demonstrates that temporal information can be extracted and is valuable for describing genotypes.

Results demonstrate classification using a reduced variable list provide superior performance over trials conducted with all available variables. This reduces the variable dimensionality provided to classification algorithms, and thus greater dependence is place on the most valuable variables. In-trial feature selection of ranked variables from an initial classification provided a more automated method and equivalent accuracy to variable group feature selection tested in this study. Variables were assessed by variable group, relative importance ranking, and occurrence within top ten variables of each date and top 50 variables of all-date variable stack.

Through variable group classification, spectral and landscape categories closely followed by CHM demonstrated the greatest value for classification. Variable ranking and occurrence assessments support the use of CHM for assessing maize genotypes as CHM was consistently ranked high across single date and multi-temporal variable rankings and was the most frequently occurring variable across top 10 ranked variables in the timeseries. The vegetation index mean-ExG and landscape variable Cohesion also demonstrated value for classification and should be further assessed for benefit in genotype discrimination and high-throughput phenotyping. All variable types demonstrated value for discriminating between genotypes. Both wavelet and Yellow Index variables, newly tested variables, showed promise for classification. Future classification should seek to better identify appropriate wavelet scale/s to reduce correlation between wavelet variables. Texture and landscape variables also demonstrated value for classification in this study but should be further refined to maximize variable type contribution.

Future work should most importantly validate these variables on other datasets, such as different genotypes and environments to determine they are robust. If validated, further exploration as to what theses variable do or mean biologically should be undertaken with other tools. Future work should assess automated ways to detect and correct for image blur in UAS images. Potential use of target detection, Hough Transform as conducted in Perez-Ortiz et al. (2016), or other methods for defining a row centroid could provide a more robust alternative to Anderson's plot boundary delineation tool which provided row boundaries in this study (Anderson et al. 2019). Development of such tools could provide broader application for methods derived in this study beyond the setting of plant breeding where row width, length, and orientation are fixed. Future research should apply remote-sensing classification of maize on a more refined variable list. In this study, algorithms were trained on a feature space of 72-127 variables for single date classification and as many as 2081 variables for the multi-temporal trial all-date variable stack. Better selection of variables prior to feature selection could improve results by reducing the number of correlated variables and feature space size. Classification should also be tested in future studies on a larger sample size and dataset. This could help prevent errors caused overfitting and testing on training dataset. In addition, conducting SGB classification on a dataset with larger sample size would allow inputs of subsample fraction and model randomization to be applied, components beneficial to random forests. The limited sample size of data used in this study prevented the use of these but inclusion of such parameters in future research could potentially lead to increased SGB performance.

REFERENCES

Aasen, H., A. Burkart, A. Bolten, and G. Bareth. 2015. "Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance." *ISPRS Journal of Photogrammetry and Remote Sensing* 108:245-259. doi: 10.1016/j.isprsjprs.2015.08.002.

Anderson, Steven L., Seth C. Murray, Lonesome Malambo, Colby Ratcliff, Sorin Popescu, Dale Cope, Anjin Chang, Jinha Jung, and J. Alex Thomasson. 2019. "Prediction of maize grain yield before maturity using improved temporal height estimates of unmanned aerial systems." *The Plant Phenome Journal* 2 (1).

Basso, Bruno, Davide Cammarano, Giovanni Cafiero, Stefano Marino, and Arturo Alvino. 2010. "Cultivar discrimination at different site elevations with remotely sensed vegetation indices." *Italian Journal of Agronomy* 6 (1):1.

Benz, Ursula C., Peter Hofmann, Gregor Willhauck, Iris Lingenfelder, and Markus Heynen. 2004. "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information." *ISPRS Journal of photogrammetry and remote sensing* 58 (3-4):239-258.

Bishop, Michael P., M. V. Bagavathiannan, D.A. Cope, D. Huo, S.C. Murray, J.A. Olsenholler, W.L. Rooney, J.A. Thomasson, J. Valasek, B.W. Young, A.M. Filippi, D.B. Hays, L. Malambo, S.C. Popescu, N. Rajan, V.P. Singh, B. McCutchen, B. Avant, M. Vidrine. 2018. "High-resolution UAS imagery in agricultural research concepts, issues, and research directions." In High Spatial Resolution Remote Sensing: Data, Analysis, and Applications, edited by Qihao Weng Yuhong He, 405. Boca Raton: CRC Press.

Blaschke, Thomas. 2010. "Object based image analysis for remote sensing." *ISPRS journal of photogrammetry and remote sensing* 65 (1):2-16.

Blaschke, Thomas, Geoffrey J. Hay, Maggi Kelly, Stefan Lang, Peter Hofmann, Elisabeth Addink, Raul Queiroz Feitosa, Freek Van der Meer, Harald Van der Werff, and Frieke Van Coillie. 2014. "Geographic object-based image analysis–towards a new paradigm." *ISPRS journal of photogrammetry and remote sensing* 87:180-191.

Breiman, Leo. 2001. "Random forests." *Machine learning* 45 (1):5-32.

Cao, J. J., W. C. Leng, K. Liu, L. Liu, Z. He, and Y. H. Zhu. 2018. "Object-based mangrove dpecies classification using unmanned aerial vehicle hyperspectral images and digital surface models." *Remote Sensing* 10 (1). doi: 10.3390/rs10010089.

Das, Bappa, R.N. Sahoo, Ankur Biswas, Sourabh Pargal, Gopal Krishna, Rakesh Verma, Viswanathan Chinnusamy, Vinay K Sehgal, and Vinod K Gupta. 2018. "Discrimination of rice genotypes using field spectroradiometry." *Geocarto International* (just-accepted):1-28.

David, L. C. G., A. H. Ballado, and Ieee. 2016. "Vegetation indices and textures in object-based weed detection from UAV imagery." *2016 6th Ieee International Conference on Control System, Computing and Engineering*.

de Castro, Ana I., Jorge Torres-Sanchez, Jose M. Pena, Francisco M. Jimenez-Brenes, Ovidiu Csillik, and Francisca Lopez-Granados. 2018. "An automatic random forest-OBIA algorithm for early weed mapping between and within crop rows using UAV imagery." *Remote Sensing* 10 (2). doi: 10.3390/rs10020285.

Dittakan, Kwankamon, Nawanol Theera-Ampornpunt, Waraphon Witthayarat, Sararat Hinnoy, Supawit Klaiwan, and Thunyatorn Pratheep. 2017. "Banana cultivar classification using scale invariant shape analysis." Information Technology (INCIT), 2017 2nd International Conference on.

Doi, Ryoichi. 2012. "Simple luminosity normalization of greenness, yellowness and redness/greenness for comparison of leaf spectral profiles in multi-temporally acquired remote sensing images." *Journal of biosciences* 37 (4):723-730.

Drăguţ, L., O. Csillik, C. Eisank, and D. Tiede. 2014. "Automated parameterisation for multi-scale image segmentation on multiple layers." *ISPRS Journal of Photogrammetry and Remote Sensing* 88:119-127.

Dube, Timothy, Onisimo Mutanga, Adam Elhadi, and Riyad Ismail. 2014. "Intra-and-inter species biomass prediction in a plantation forest: testing the utility of high spatial resolution spaceborne multispectral rapideye sensor and advanced machine learning algorithms." *Sensors* 14 (8):15348-15370.

eCognition Developer, Trimble. 2014. "9.0 User Guide." *Trimble Germany GmbH: Munich, Germany*.

FAO, Food and Agriculture Organization of the United States. 2009. "FAO’s director-general on how to feed the world in 2050." *Population and Development Review* 35:837-839.

Freeman, Elizabeth A., Gretchen G. Moisen, John W. Coulston, and Barry T. Wilson. 2015. "Random forests and stochastic gradient boosting for predicting tree canopy cover: comparing tuning processes and model performance." *Canadian Journal of Forest Research* 46 (3):323-339.

Friedman, Jerome H. 2002. "Stochastic gradient boosting." *Computational statistics & data analysis* 38 (4):367-378.

Garriga, Miguel, Sebastián Romero-Bravo, Félix Estrada, Alejandro Escobar, Iván A Matus, Alejandro del Pozo, Cesar A Astudillo, and Gustavo A Lobos. 2017. "Assessing Wheat

Traits by Spectral Reflectance: Do We Really Need to Focus on Predicted Trait-values or Directly Identify the Elite Genotypes Group?" *Frontiers in plant science* 8:280.

Gamon, J.A., and J.S. Surfus. 1999. "Assessing leaf pigment content and activity with a reflectometer." *The New Phytologist* 143 (1):105-117.

Geipel, Jakob, Johanna Link, and Wilhelm Claupein. 2014. "Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system." *Remote Sensing* 6 (11):10335-10355.

Gitelson, Anatoly A., Yoram J. Kaufman, Robert Stark, and Don Rundquist. 2002. "Novel algorithms for remote estimation of vegetation fraction." *Remote sensing of Environment* 80 (1):76-87.

Guerrero, José Miguel, Gonzalo Pajares, Martín Montalvo, Juan Romeo, and María Guijarro. 2012. "Support vector machines for crop/weeds identification in maize fields." *Expert Systems with Applications* 39 (12):11149-11155.

Guijarro, Marıa, Gonzalo Pajares, Isabel Riomoros, P.J. Herrera, X.P. Burgos-Artizzu, and Angela Ribeiro. 2011. "Automatic segmentation of relevant textures in agricultural images." *Computers and Electronics in Agriculture* 75 (1):75-83.

Güneralp, İnci, Anthony M. Filippi, and Jarom Randall. 2014. "Estimation of floodplain aboveground biomass using multispectral remote sensing and nonparametric modeling." *International Journal of Applied Earth Observation and Geoinformation* 33:119-126.

Haghighattalab, Atena, Lorena González Pérez, Suchismita Mondal, Daljit Singh, Dale Schinstock, Jessica Rutkoski, Ivan Ortiz-Monasterio, Ravi Prakash Singh, Douglas Goodin, and Jesse Poland. 2016. "Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries." *Plant Methods* 12 (1):35.

Hague, T., N.D. Tillett, and H. Wheeler. 2006. "Automated crop and weed monitoring in widely spaced cereals." *Precision Agriculture* 7 (1):21-32.

Hall, Ola, Sigrun Dahlin, Håkan Marstorp, Maria Archila Bustos, Ingrid Öborn, and Magnus Jirström. 2018. "Classification of maize in complex smallholder farming systems using UAV imagery." *Drones* 2 (3):22.

Huang, Huasheng, Jizhong Deng, Yubin Lan, Aqing Yang, Xiaoling Deng, and Lei Zhang. 2018. "A fully convolutional network for weed mapping of unmanned aerial vehicle (UAV) imagery." *Plos One* 13 (4). doi: 10.1371/journal.pone.0196302.

Hung, C., M. Bryson, and S. Sukkarieh. 2011. "Vision-based shadow-aided tree crown detection and classification algorithm using imagery from an unmanned airborne vehicle." *In 34th International Symposium for Remote Sensing of the Environment.*

Hunter, Mitchell C, Richard G Smith, Meagan E Schipanski, Lesley W Atwood, and David A Mortensen. 2017. "Agriculture in 2050: recalibrating targets for sustainable intensification." *Bioscience* 67 (4):386-391.

Hussain, Masroor, Dongmei Chen, Angela Cheng, Hui Wei, and David Stanley. 2013. "Change detection from remotely sensed images: From pixel-based to object-based approaches." *ISPRS Journal of photogrammetry and remote sensing* 80:91-106.

Hyman, Glenn, Sam Fujisaka, Peter Jones, Stanley Wood, M Carmen De Vicente, and John Dixon. 2008. "Strategic approaches to targeting technology generation: Assessing the coincidence of poverty and drought-prone crop production." *Agricultural Systems* 98 (1):50-61.

Ide, Reiko, and Hiroyuki Oguma. 2013. "A cost-effective monitoring method using digital time-lapse cameras for detecting temporal and spatial variations of snowmelt and vegetation phenology in alpine ecosystems." *Ecological Informatics* 16:25-34.

Jay, Sylvain, Gilles Rabatel, Xavier Hadoux, Daniel Moura, and Nathalie Gorretta. 2015. "In-field crop row phenotyping from 3D modeling performed using Structure from Motion." *Computers and Electronics in Agriculture* 110:70-77.

Kataoka, Takashi, Toshihiro Kaneko, Hiroshi Okamoto, and S. Hata. 2003. "Crop growth estimation system using machine vision." *Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics* (AIM 2003).

Li, Wang, Zheng Niu, Hanyue Chen, Dong Li, Mingquan Wu, and Wei Zhao. 2016. "Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system." *Ecological indicators* 67:637-648.

Liu, Yongxue, Manchun Li, Liang Mao, Feifei Xu, and Shuo Huang. 2006. "Review of remotely sensed imagery classification patterns based on object-oriented image analysis." *Chinese Geographical Science* 16 (3):282-288.

Lopez-Granados, Francisca, Jorge Torres-Sanchez, Angelica Serrano-Perez, Ana I. de Castro, Fco-Javier Mesas-Carrascosa, and Jose-Manuel Pena. 2016. "Early season weed mapping in sunflower using UAV technology: variability of herbicide treatment maps against weed thresholds." *Precision Agriculture* 17 (2):183-199. doi: 10.1007/s11119-015-9415-8.

López-Granados, Francisca. 2011. "Weed detection for site-specific weed management: mapping and real-time approaches." *Weed Research* 51 (1):1-11.

Louhaichi, Mounir, Michael M. Borman, and Douglas E. Johnson. 2001. "Spatially located platform and aerial photography for documentation of grazing impacts on wheat." *Geocarto International* 16 (1):65-70.

Ma, Lei, Manchun Li, Xiaoxue Ma, Liang Cheng, Peijun Du, and Yongxue Liu. 2017. "A review of supervised object-based land-cover image classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 130:277-293.

Malambo, Lonesome, Sorin C Popescu, Seth C. Murray, E. Putman, N. Ace Pugh, David W. Horne, G. Richardson, R. Sheridan, William L. Rooney, and R. Avant. 2018. "Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery." *International journal of applied earth observation and geoinformation* 64:31-42.

McGarigal, K., S.A. Cushman, and E. Ene. 2012. "FRAGSTATS v4: Spatial pattern analysis program for categorical and continuous maps." Computer software program produced by the authors at the University of Massachusetts, Amherst. Available at the following web site: http://www.umass.edu/landeco/research/fragstats/fragstats.html

McGarigal, Kevin. 2015. "FRAGSTATS help." University of Massachusetts: Amherst, MA, USA.

Meyer, George E., Timothy W. Hindman, and Koppolu Laksmi. 1999. "Machine vision detection parameters for plant species identification." Precision Agriculture and Biological Quality.

Meyer, George E., and João Camargo Neto. 2008. "Verification of color vegetation indices for automated crop imaging applications." *Computers and electronics in agriculture* 63 (2):282-293.

Moody, Daniela I., Steven P. Brumby, Rick Chartrand, Ryan Keisler, Nathan Longbotham, Carly Mertes, Samuel W. Skillman, and Michael S. Warren. 2017. "Crop classification using temporal stacks of multispectral satellite imagery." In Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery Xxiii, edited by M. VelezReyes and D. W. Messinger.

Murray, S.C., L. Malambo, S. Popescu, D. Cope, SL Anderson, A. Chang, J. Jung, N. Cruzato, S. Wilde, and R.L. Walls. 2019. "G2F Maize UAV Data, College Station, Texas 2017." *CyVerse Data Commons. doi* 10.

Neto, Joao Camargo. 2004. *A combined statistical-soft computing approach for classification and mapping weed species in minimum-tillage systems*: The University of Nebraska-Lincoln.

Nevalainen, Olli, Eija Honkavaara, Sakari Tuominen, Niko Viljanen, Teemu Hakala, Xiaowei Yu, Juha Hyyppa, Heikki Saari, Ilkka Polonen, Nilton N. Imai, and Antonio M. G.

Tommaselli. 2017. "Individual tree detection and classification with UAV-based photogrammetric point clouds and hyperspectral imaging." *Remote Sensing* 9 (3). doi: 10.3390/rs9030185.

Osmolska, Anna, and Pawel Hawrylo. 2018. "Using a GEOBIA framework for integrating different data sources and classification methods in context of land use/land cover mapping." *Geodesy and Cartography* 67 (1):99-116. doi: 10.24425/118699.

Peña, J. M., J. Torres-Sánchez, A. I. de Castro, M. Kelly, and F. López-Granados. 2013. "Weed mapping in early-season maize fields using object-based analysis of unmanned aerial behicle (UAV) images." *PLoS ONE* 8 (10). doi: 10.1371/journal.pone.0077151.

Peña, M. A., R. Liao, and A. Brenning. 2017. "Using spectrotemporal indices to improve the fruit-tree crop classification accuracy." *ISPRS Journal of Photogrammetry and Remote Sensing* 128:158-169. doi: 10.1016/j.isprsjprs.2017.03.019.

Perez-Ortiz, Maria, Jose Manuel Pena, Pedro Antonio Gutierrez, Jorge Torres-Sanchez, Cesar Hervas-Martinez, and Francisca Lopez-Granados. 2016. "Selecting patterns and features for between- and within-crop-row weed mapping using UAV-imagery." *Expert Systems with Applications* 47:85-94. doi: 10.1016/j.eswa.2015.10.043.

Petitjean, François, Camille Kurtz, Nicolas Passat, and Pierre Gançarski. 2012. "Spatio-temporal reasoning for the classification of satellite image time series." *Pattern Recognition Letters* 33 (13):1805-1815.

Pipkins, Kyle, Michael Foerster, Anne Clasen, Tobias Schmidt, and Birgit Kleinschmit. 2014. "A Comparison of feature selection methods for multitemporal tree species classification." In *Earth Resources and Environmental Remote Sensing/Gis Applications V*, edited by U. Michel, K. Schulz, M. Ehlers, K. G. Nikolakopoulos and D. L. Civco.

Popescu, Sorin C, and Kaiguang Zhao. 2008. "A voxel-based lidar method for estimating crown base height for deciduous and pine trees." Remote sensing of environment 112 (3):767-781.

Przybyl, K., K. Gorna, D. Wojcieszak, W. Czekala, A. Ludwiczak, A. Przybylak, P. Boniecki, K. Koszela, M. Zaborowicz, D. Janczak, and A. Lewicki. 2015. "The recognition of potato varieties using neural image analysis method." In *Seventh International Conference on Digital Image Processing*, edited by C. M. Falco and X. Jiang.

Ranum, Peter, Juan Pablo Peña-Rosas, and Maria Nieves Garcia-Casal. 2014. "Global maize production, utilization, and consumption." *Annals of the New York Academy of Sciences* 1312 (1):105-112.

Rao, N. Rama, P. K. Garg, and S. K. Ghosh. 2007. "Development of an agricultural crops spectral library and classification of crops at cultivar level using hyperspectral data." *Precision Agriculture* 8 (4-5):173-185. doi: 10.1007/s11119-007-9037-x.

Ray, Deepak K., Nathaniel D. Mueller, Paul C. West, and Jonathan A. Foley. 2013. "Yield trends are insufficient to double global crop production by 2050." *PloS one* 8 (6):e66428.

Salford Systems Modeler. 2018. "Introducing TreeNet gradient boosting machine" https://www.salford-systems.com/products/spm/userguide#treenet%C2%AE-stochastic-gradient-boosting

Shahi, U. P., S. Kumar, N. P. Singh, A. K. Chaubey, and Y. Kumar. 2007. "Potato varietal discrimination using ground based multiband radiometer." *Journal of the Indian Society of Remote Sensing* 35 (1):53-65. doi: 10.1007/BF02991833.

Song, Qian, Mingtao Xiang, Ciara Hovis, Qingbo Zhou, Miao Lu, Huajun Tang, and Wenbin Wu. 2019. "Object-based feature selection for crop classification using multi-temporal high-resolution imagery." *International Journal of Remote Sensing* 40 (5-6):2053-2068.

Torres-Sánchez, J., F. López-Granados, and J. M. Peña. 2015. "An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops." *Computers and Electronics in Agriculture* 114:43-52. doi: 10.1016/j.compag.2015.03.019.

Torres-Sánchez, Jorge, Francisca López-Granados, Nicolás Serrano, Octavio Arquero, and José M Peña. 2015. "High-throughput 3-D monitoring of agricultural-tree plantations with unmanned aerial vehicle (UAV) technology." *PloS one* 10 (6):e0130479.

Torres-Sánchez, Jorge, José Manuel Peña, Ana Isabel de Castro, and Fransisca López-Granados. 2014. "Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV." *Computers and Electronics in Agriculture* 103:104-113.

Trimble. 2019. "eCognition Developer 9.5 Reference Book." Trimble Germany GmbH: Munich, Germany.

Tu, Y., M. Bian, Y. Wan, and T. Fei. 2018. "Tea cultivar classification and biochemical parameter estimation from hyperspectral imagery obtained by UAV." *PeerJ* 2018 (5). doi: 10.7717/peerj.4858.

NASS/USDA. “Acreage”. National Agricultural Statistics Service, U.S. Department of Agriculture, ISSN: 1949-1522, June 2019. https://downloads.usda.library.cornell.edu/usda-esmis/files/j098zb09z/0k225n39n/jw827p632/acrg0619.pdf

Vega, Francisco Agüera, Fernando Carvajal Ramirez, Monica Perez Saiz, and Francisco Orgaz Rosua. 2015. "Multi-temporal imaging using an unmanned aerial vehicle for monitoring a sunflower crop." *Biosystems Engineering* 132:19-27.

Vergara-Díaz, Omar, Mainassara A. Zaman-Allah, Benhildah Masuka, Alberto Hornero, Pablo Zarco-Tejada, Boddupalli M Prasanna, Jill E. Cairns, and José L. Araus. 2016. "A novel remote sensing approach for prediction of maize yield under different conditions of nitrogen fertilization." *Frontiers in plant science* 7:666.

Woebbecke, David M., George E. Meyer, K. Von Bargen, and D.A. Mortensen. 1995. "Color indices for weed identification under various soil, residue, and lighting conditions." *Transactions of the ASAE* 38 (1):259-269.

Woebbecke, David M., George E. Meyer, Kenneth Von Bargen, and David A. Mortensen. 1993. "Plant species identification, size, and enumeration using machine vision techniques on near-binary images." Optics in Agriculture and Forestry.

Wu, M. Q., C. H. Yang, X. Y. Song, W. C. Hoffmann, W. J. Huang, Z. Niu, C. Y. Wang, and W. Li. 2017. "Evaluation of orthomosics and digital surface models derived from aerial imagery for crop type mapping." *Remote Sensing* 9 (3). doi: 10.3390/rs9030239.

Xiao, J., H. Wu, C. Wang, and H. Xia. 2018. "Land cover classification using features generated from annual time-series Landsat data." *IEEE Geoscience and Remote Sensing Letters* 15 (5):739-743. doi: 10.1109/LGRS.2018.2805715.

Xue, Jinru, and Baofeng Su. 2017. "Significant remote sensing vegetation indices: A review of developments and applications." *Journal of Sensors* 2017.

Yu, Zhenghong, Zhiguo Cao, Xi Wu, Xiaodong Bai, Yueming Qin, Wen Zhuo, Yang Xiao, Xuefen Zhang, and Hongxi Xue. 2013. "Automatic image-based detection technology for two critical growth stages of maize: Emergence and three-leaf stage." *Agricultural and Forest Meteorology* 174:65-84.

Zaman-Allah, Mainassara, O. Vergara, J.L. Araus, A. Tarekegne, C. Magorokosho, P.J. Zarco-Tejada, A. Hornero, A. Hernández Albà, B. Das, and P. Craufurd. 2015. "Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize." *Plant methods* 11 (1):35.

Zhao, Chunjiang, Jihua Wang, Wenjiang Huang, and Qifa Zhou. 2010. "Spectral indices sensitively discriminating wheat genotypes of different canopy architectures." *Precision Agriculture* 11 (5):557-567.

Zhao, Duli, Neil C. Glynn, Barry Glaz, Jack C. Comstock, and Richard M Johnson. 2012. "Development of leaf spectral models for evaluating large numbers of sugarcane genotypes." *Crop science* 52 (4):1837-1847.

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

APPENDIX A

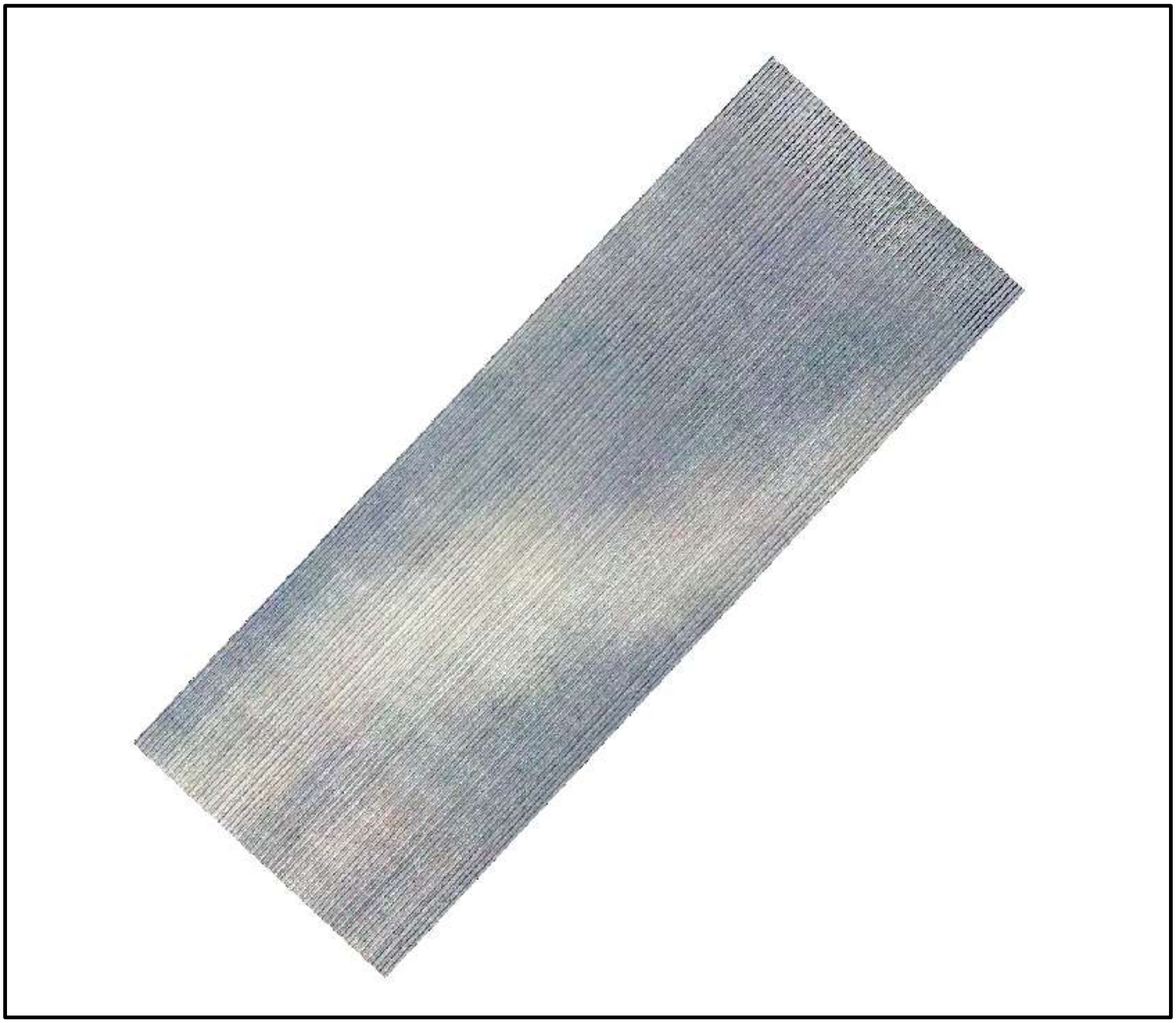
Class Number	Pedigree	**# of Rows**	**Mask Rank**
2	2396/LH123HT	60	928
14	B73/PHZ51	8	114
45	DKB 64-69	6	92
81	LH195/PHZ51	8	126
238	PHW52/PHN82	8	120

Genotypes with more than 4 rows (2 plots) in field condition G2FE.

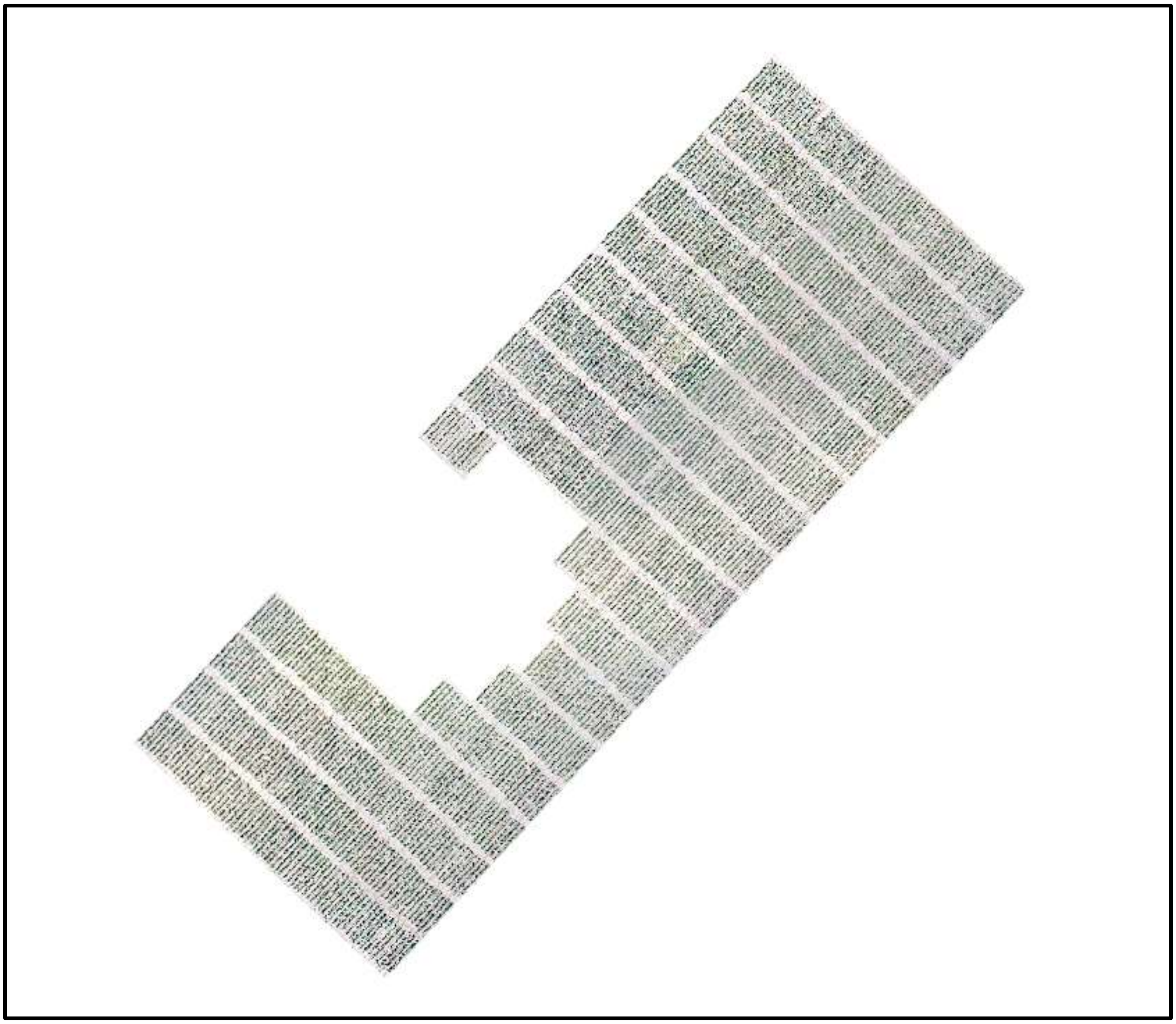
Date	**Mask Image Blur**	**Geoshift**
March 21		
April 07	x	x
April 20	x	
May 2	x	
May 5	x	x
May 9	x	
May 12		x
May 15	x	x
May 24	x	x
May 30	x	x
June 02		x
June 06	x	x
June 09	x	x
June 12	x	
June 16	x	x
June 23		
June 29	x	x
July 13	x	

Image dates and corrections applied.

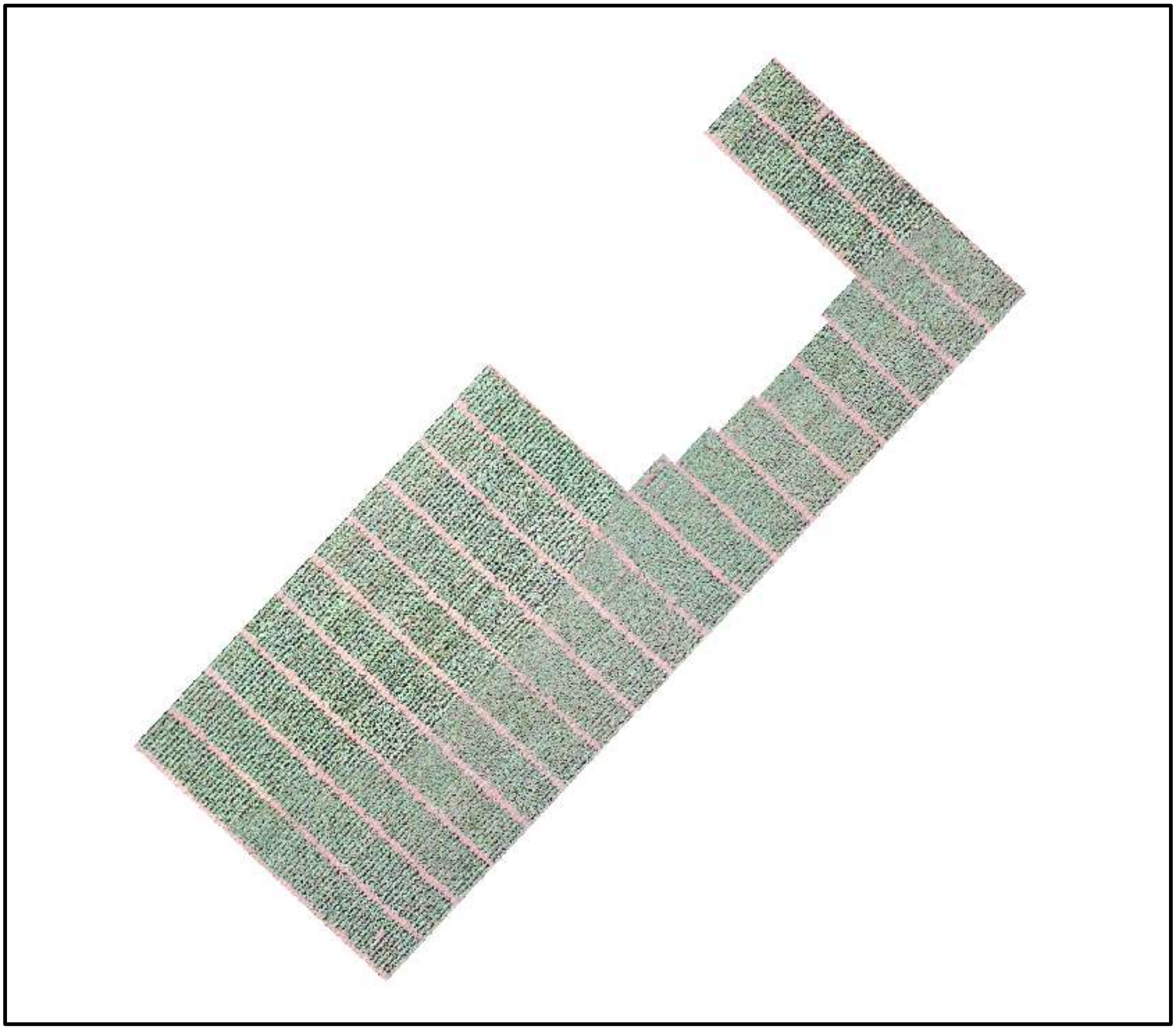
APPENDIX B



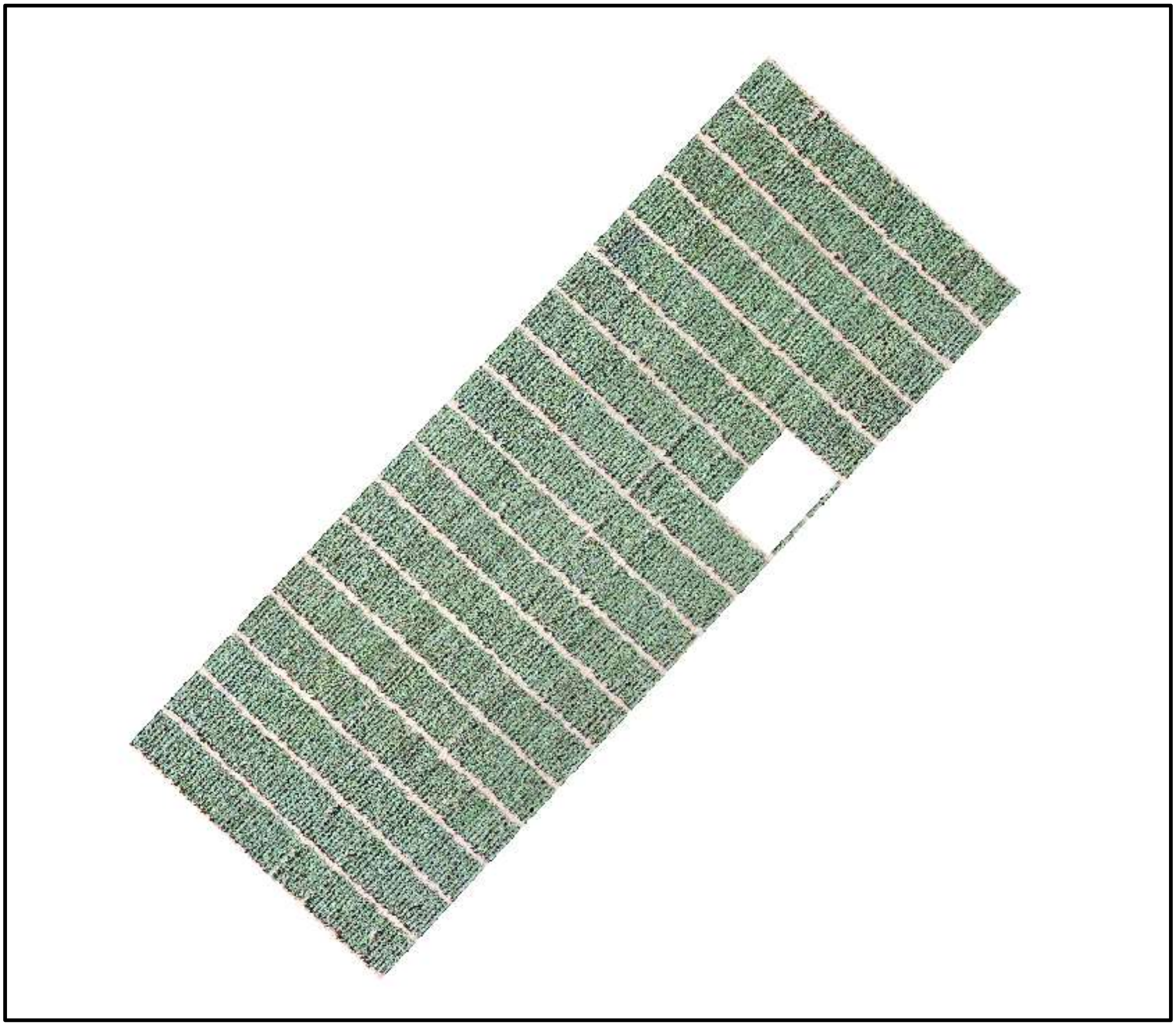
RGB mosaic for date 03/21 after quality corrections were applied.



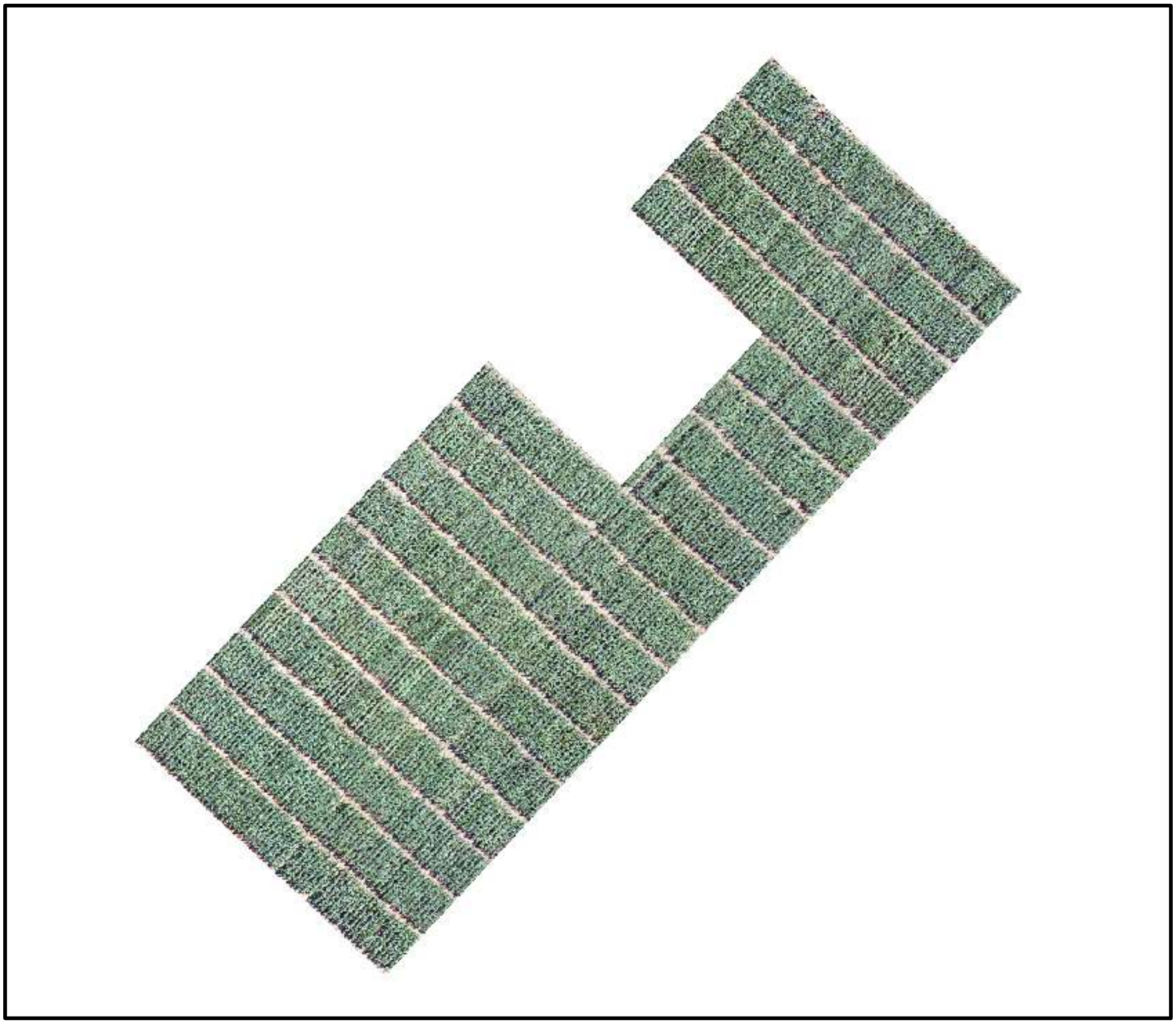
RGB mosaic for date 04/07 after quality corrections were applied.



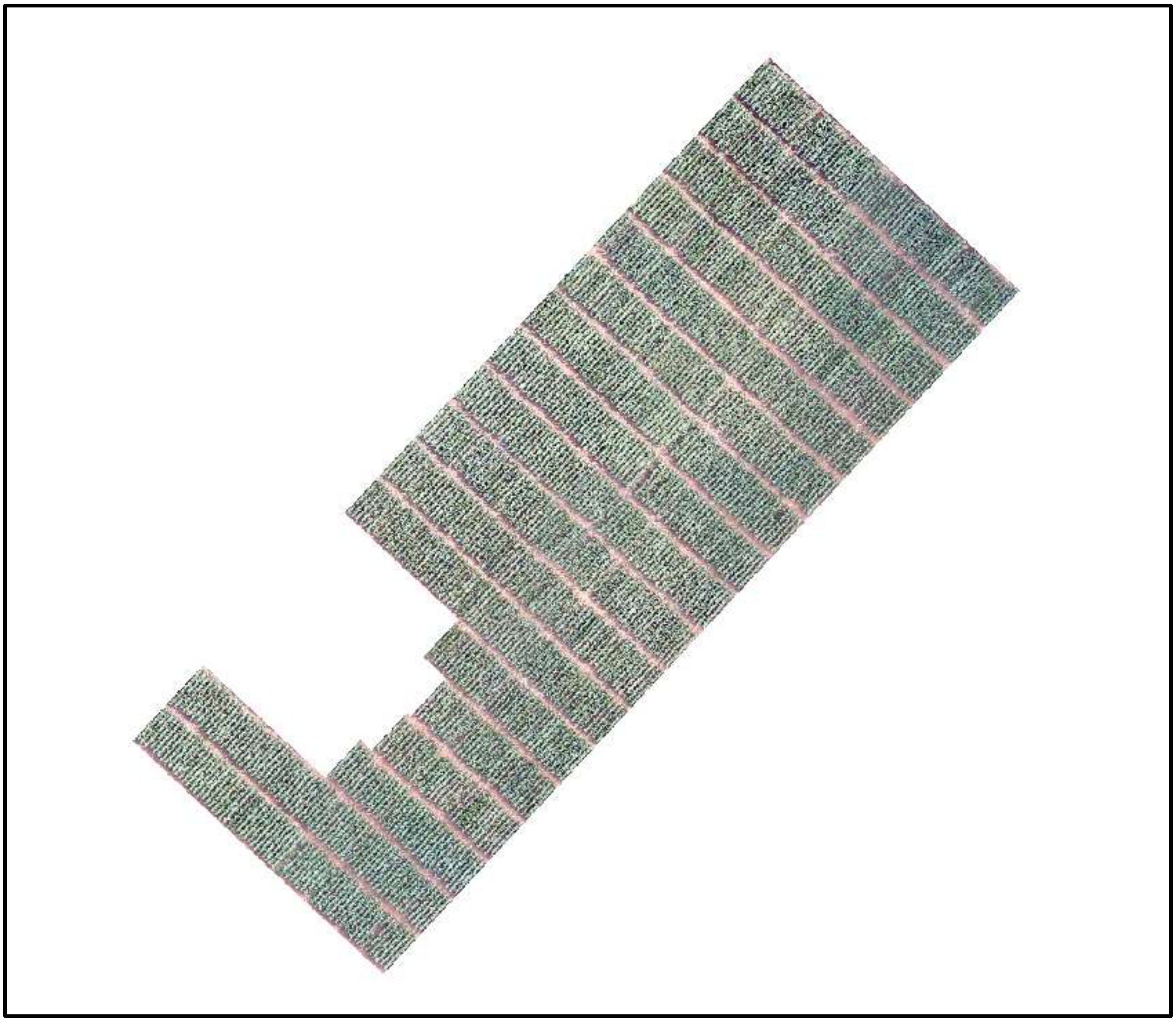
RGB mosaic for date 04/20 after quality corrections were applied.



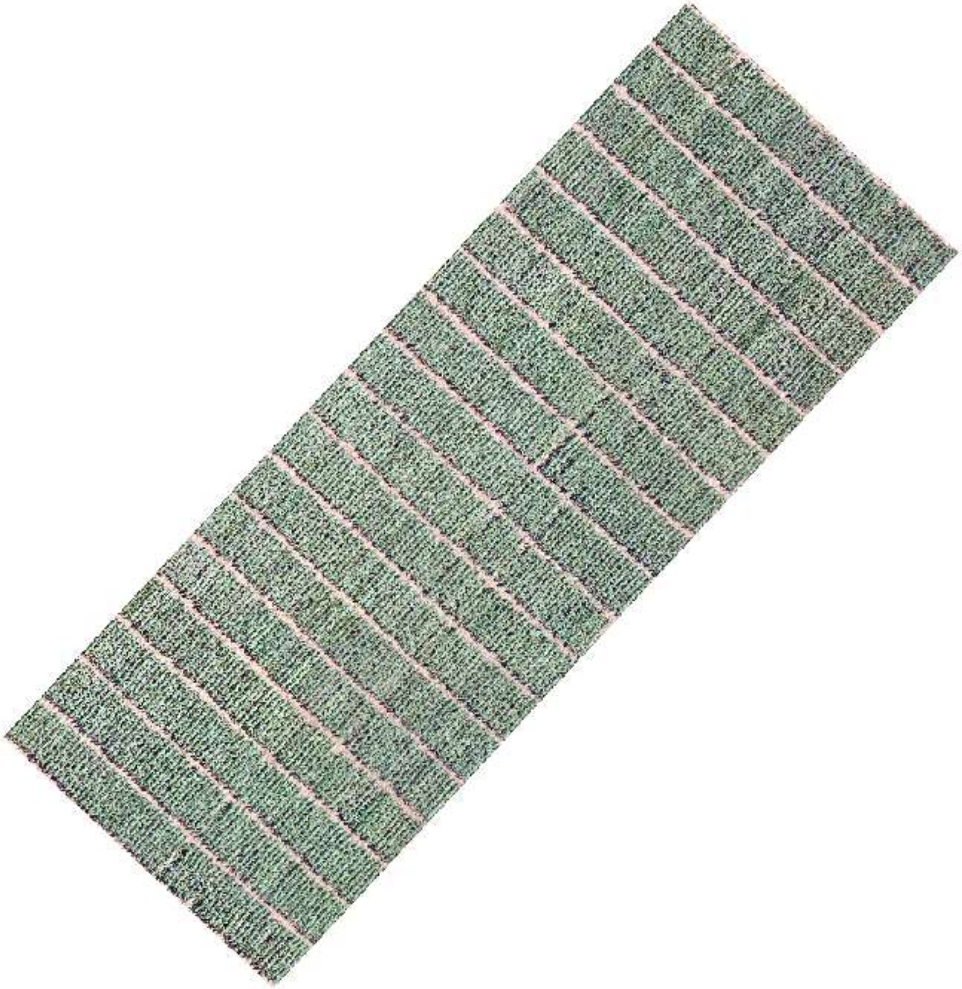
RGB mosaic for date 05/02 after quality corrections were applied.



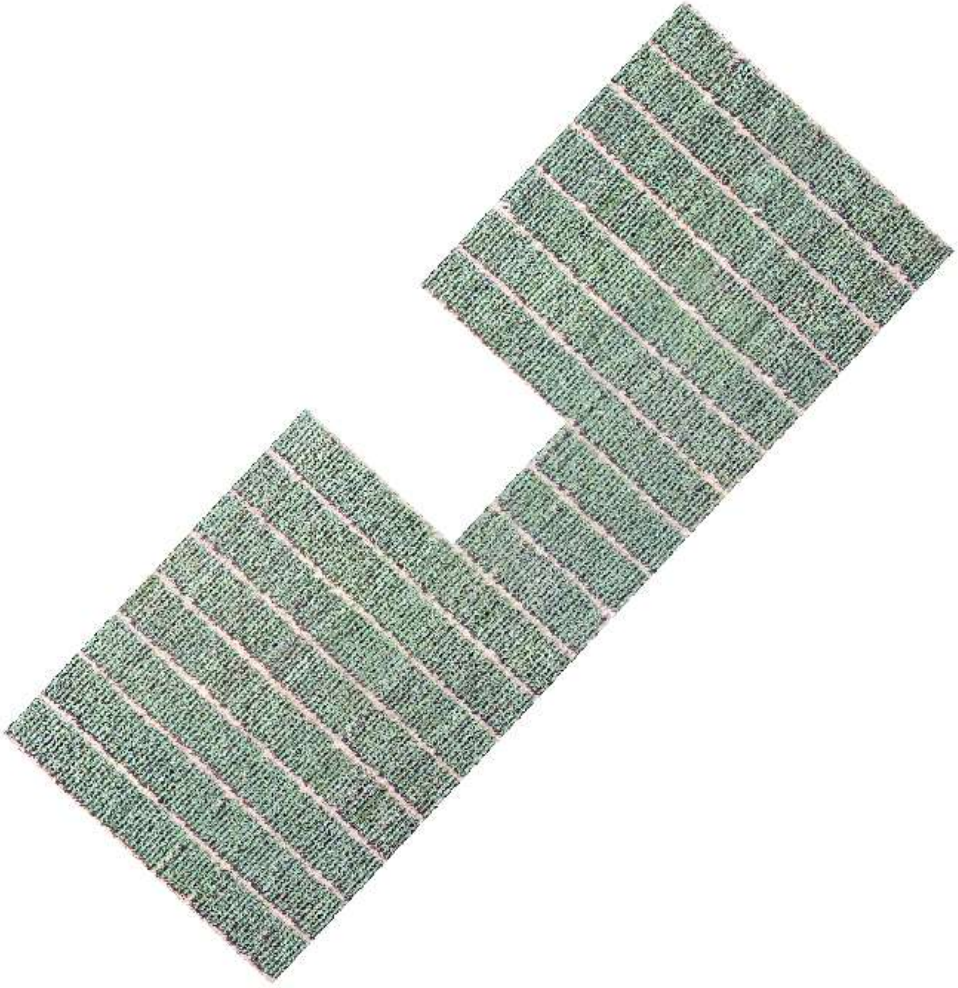
RGB mosaic for date 05/05 after quality corrections were applied.



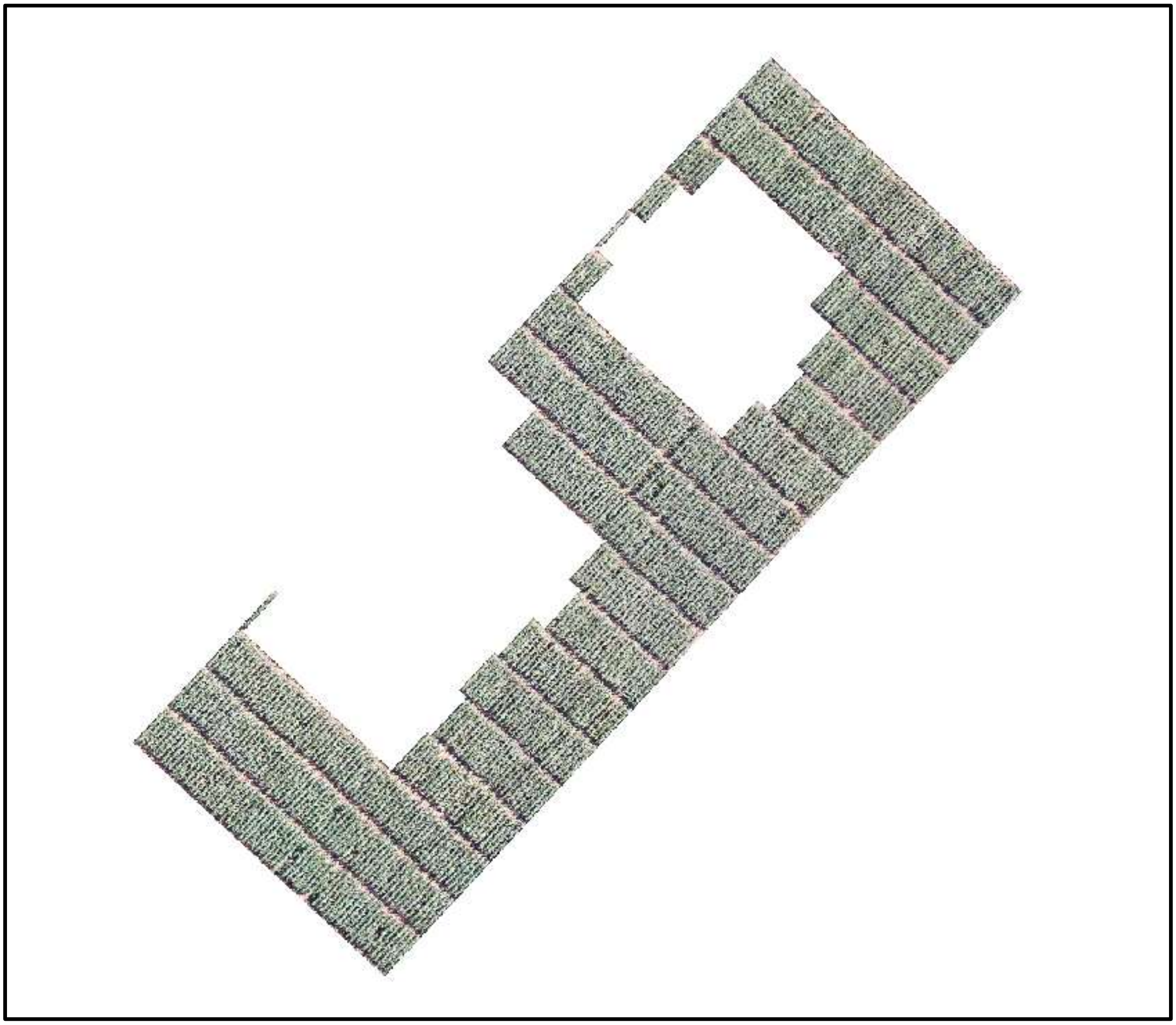
RGB mosaic for date 05/09 after quality corrections were applied.



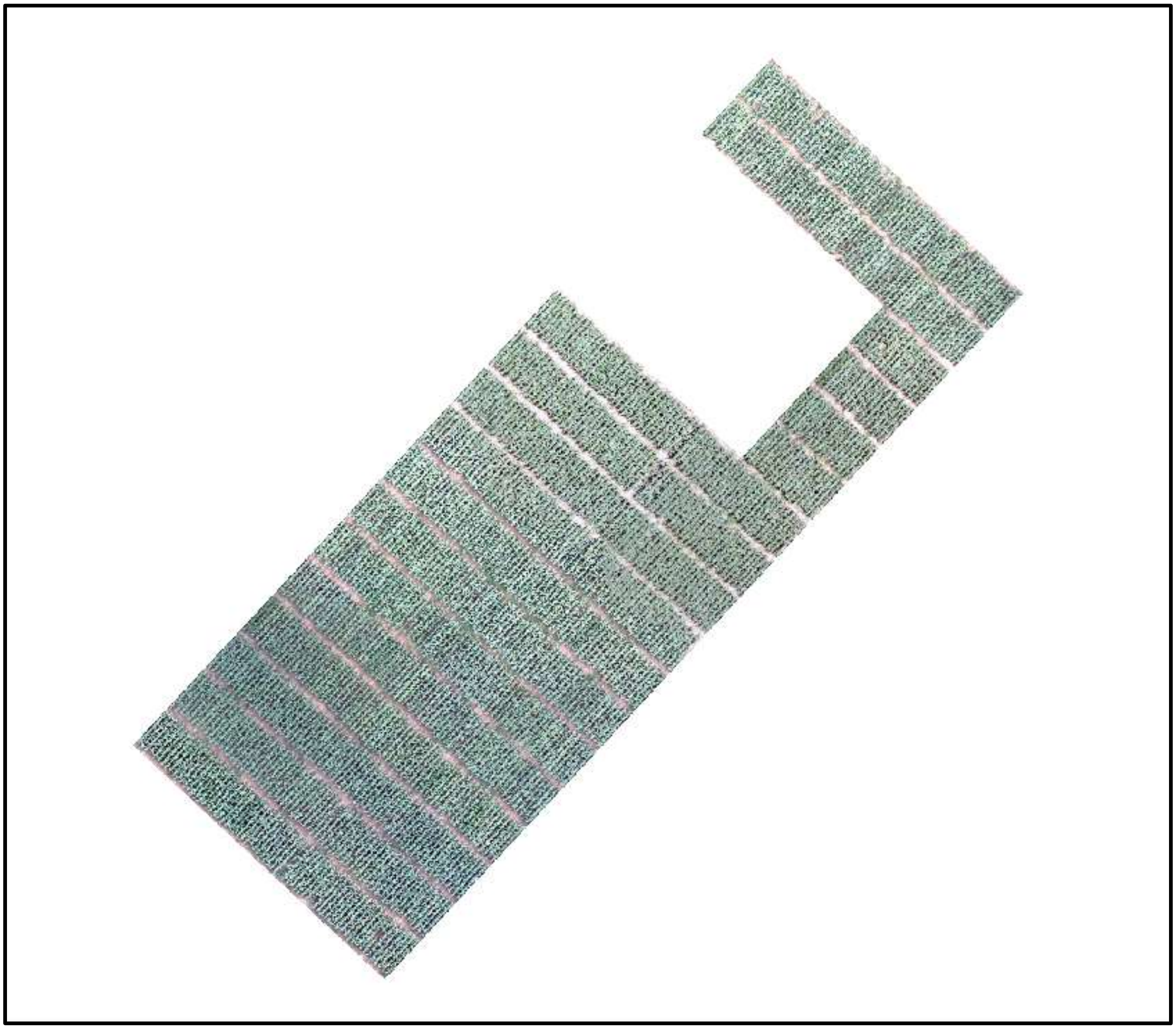
RGB mosaic for date 05/12 after quality corrections were applied.



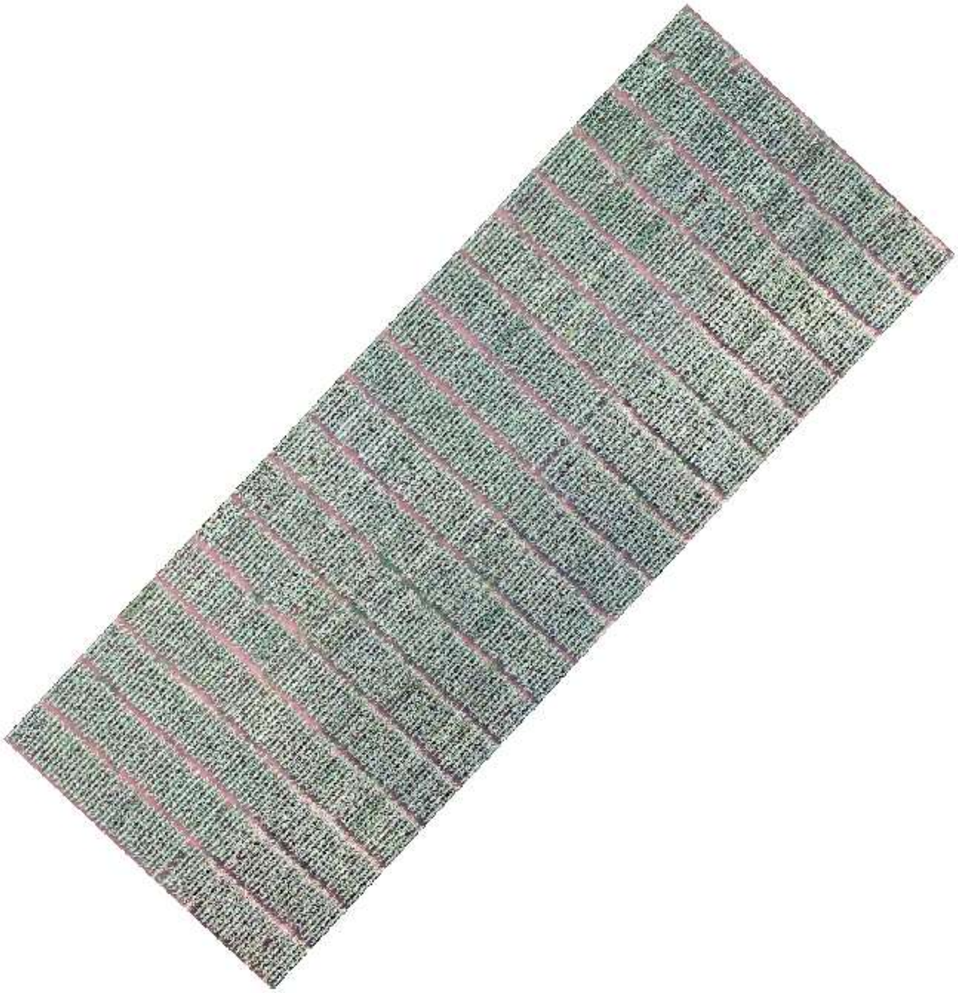
RGB mosaic for date 05/15 after quality corrections were applied.



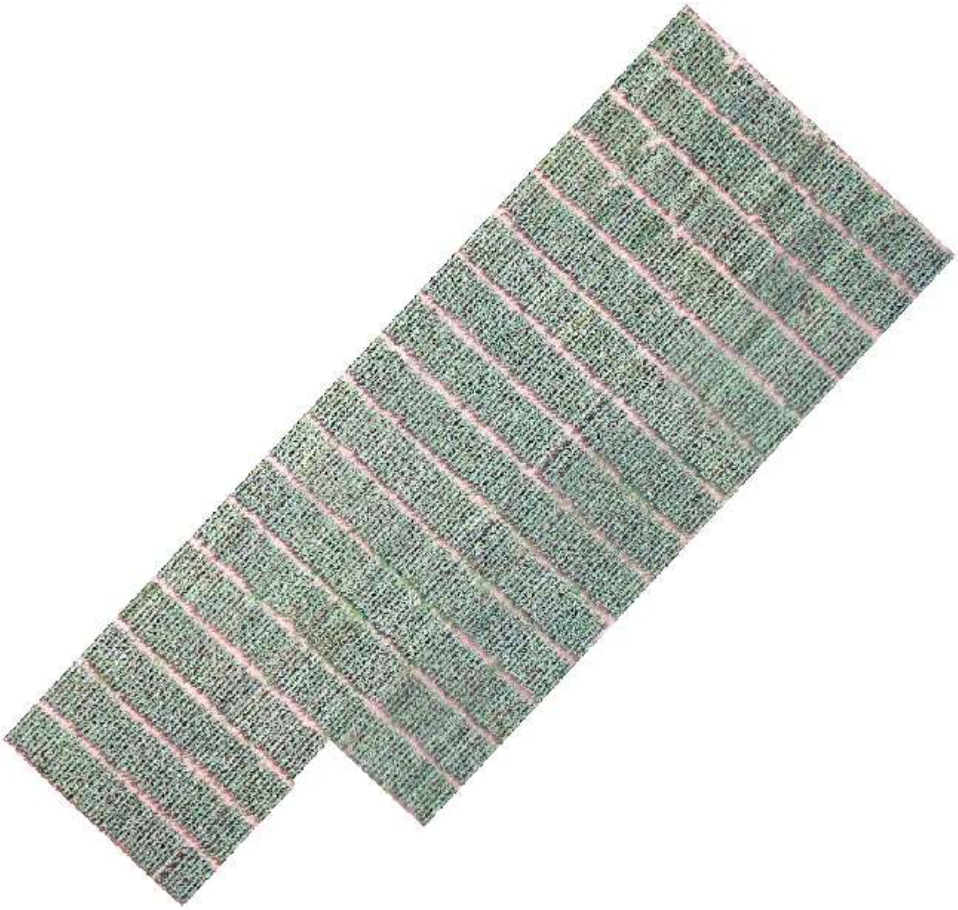
RGB mosaic for date 05/24 after quality corrections were applied.



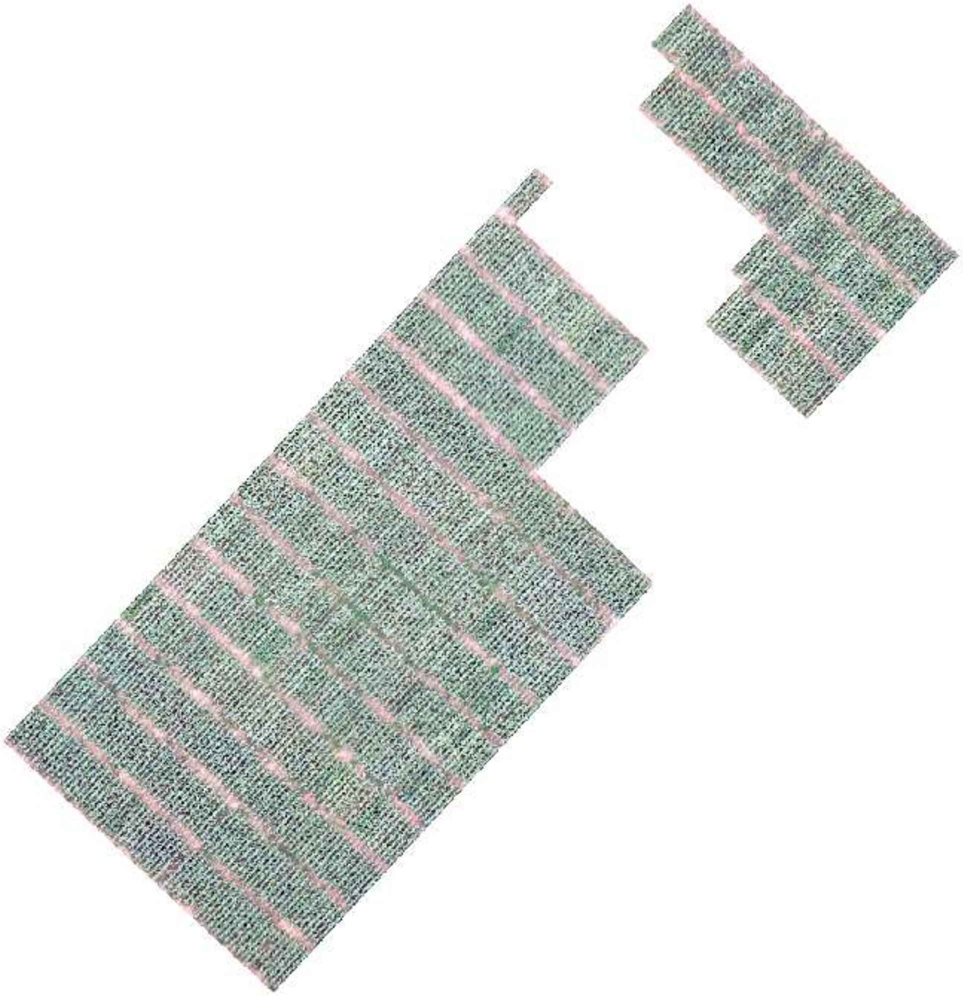
RGB mosaic for date 05/30 after quality corrections were applied.



RGB mosaic for date 06/02 after quality corrections were applied.



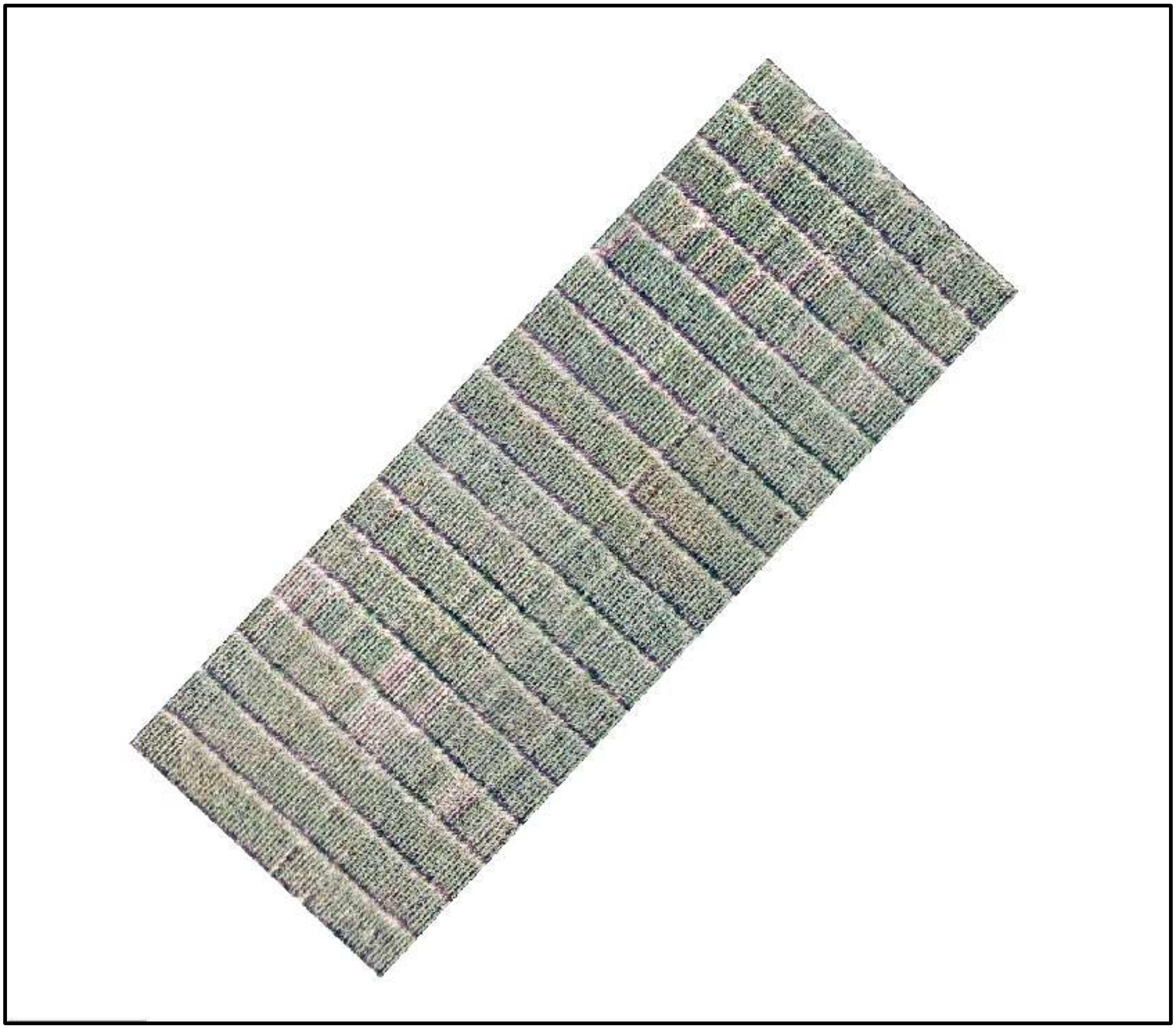
RGB mosaic for date 06/06 after quality corrections were applied.



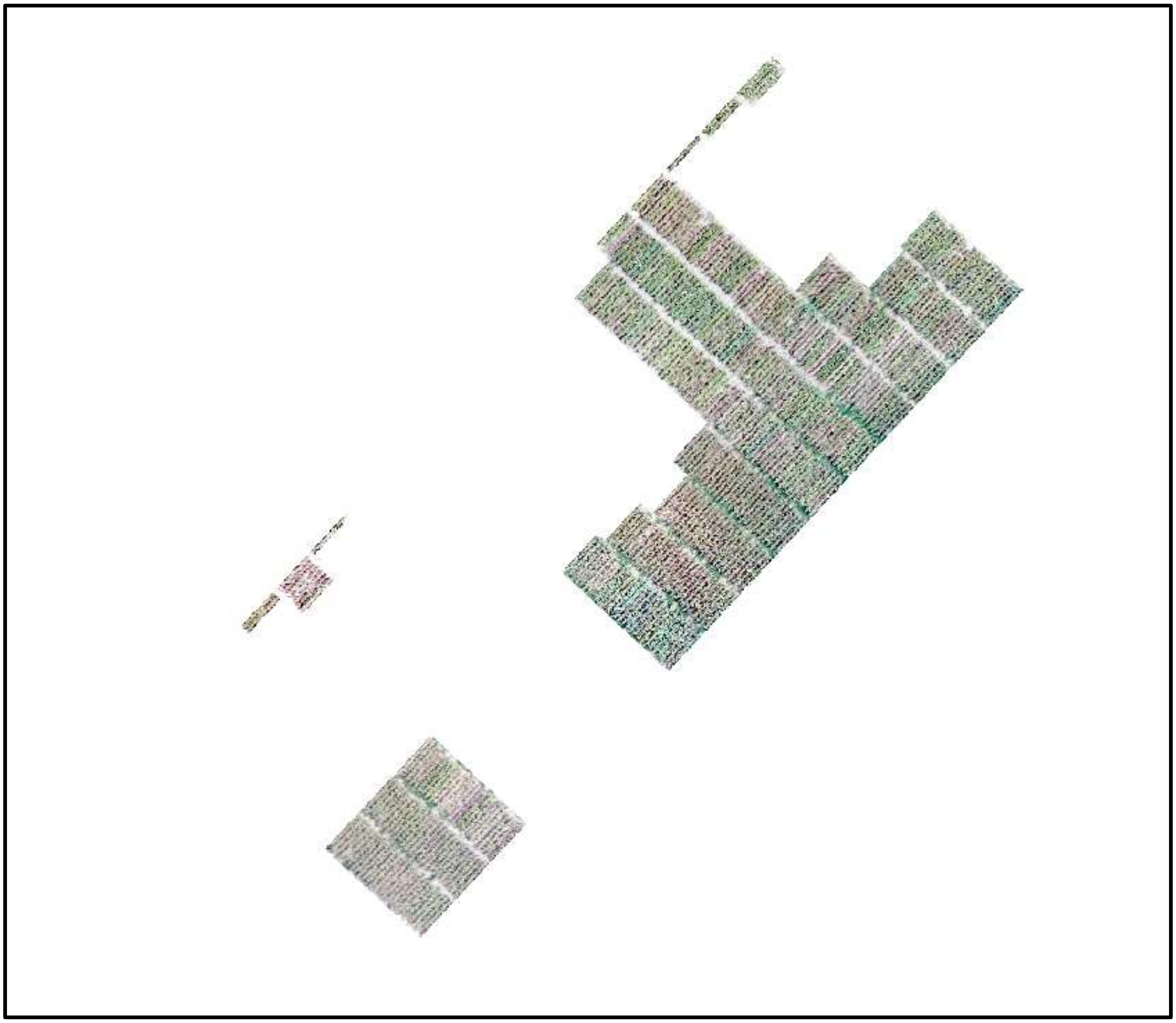
RGB mosaic for date 06/09 after quality corrections were applied.

RGB mosaic for date 06/12 after quality corrections were applied.

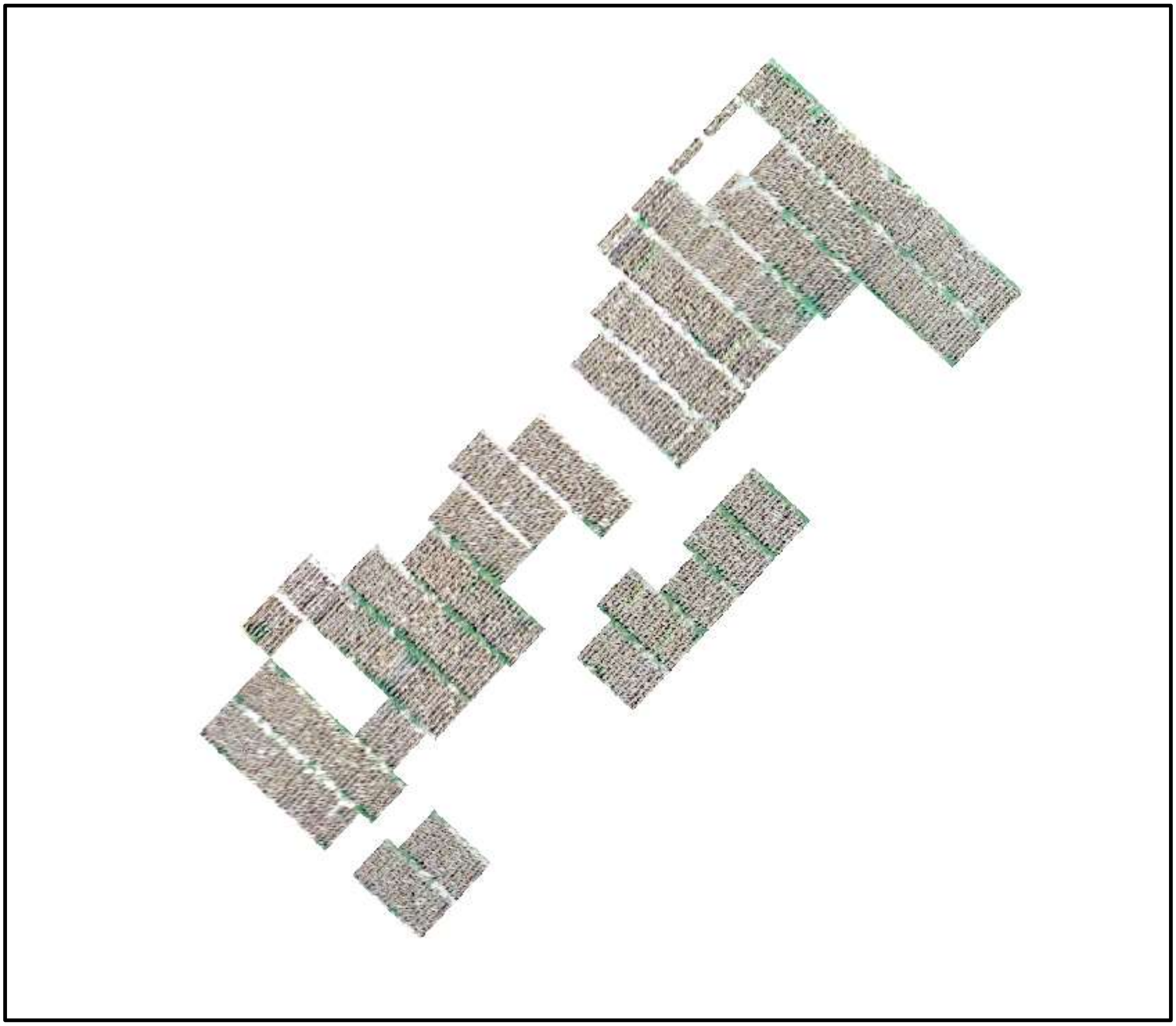
RGB mosaic for date 06/16 after quality corrections were applied.



RGB mosaic for date 06/23 after quality corrections were applied.



RGB mosaic for date 06/29 after quality corrections were applied.



RGB mosaic for date 07/13 after quality corrections were applied.

APPENDIX C

Mask Ranking based on the number of rows available for each genotype after image blur masking.

Genotype Class	03/ 21	04/ 07	04/ 20	05/ 02	05/ 05	05/ 09	05/ 12	05/ 15	05/ 24	05/ 30	06/ 02	06/ 06	06/ 09	06/ 12	06/ 16	06/ 23	06/ 29	07/ 13	All Date Sum
2	60	60	46	56	48	56	60	48	56	58	60	50	52	53	48	60	21	36	**928**
81	8	6	6	8	6	8	8	8	4	6	8	8	6	6	8	8	6	8	**126**
238	8	4	8	8	8	6	8	8	4	8	8	8	8	4	8	8	2	4	**120**
14	8	6	6	8	8	6	8	8	2	4	8	8	4	6	8	8	4	4	**114**
45	6	6	4	6	4	6	6	6	6	6	6	6	6	6	4	6		2	**92**
89	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	4	4	**70**
90	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	4	4	**70**
149	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	4	**70**
91	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	4	2	**68**
96	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	4	2	**68**
97	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	4	2	**68**
100	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	4	2	**68**
148	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	2	4	**68**
151	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	4	4	4	**68**
152	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	4	4	4	**68**
159	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	4	4	4	**68**
99	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	2	3	**67**
64	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
65	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
69	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
75	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
77	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
92	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	2	2	**66**
94	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	2	2	**66**
98	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	2	2	**66**
135	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	2	2	**66**
153	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	4	4	2	**66**
158	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	4	4	2	**66**
180	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
182	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
185	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
201	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
216	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
225	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**

Genotype Class	03/ 21	04/ 07	04/ 20	05/ 02	05/ 05	05/ 09	05/ 12	05/ 15	05/ 24	05/ 30	06/ 02	06/ 06	06/ 09	06/ 12	06/ 16	06/ 23	06/ 29	07/ 13	All Date Sum
234	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
236	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
243	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	2	2	**66**
101	4	4	4	4	4	4	4	2	4	4	4	4	4	4	4	4	1	3	**66**
104	4	4	4	4	4	4	4	2	4	4	4	4	4	4	4	4		4	**66**
105	4	4	4	4	4	4	4	2	4	4	4	4	4	4	4	4		4	**66**
107	4	4	4	4	4	4	4	2	4	4	4	4	4	4	4	4		4	**66**
108	4	4	4	4	4	4	4	2	4	4	4	4	4	4	4	4		4	**66**
110	4	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4		4	**66**
111	4	4	4	4	4	4	4	2	4	4	4	4	4	4	4	4		4	**66**
172	4	4	4	4	4	4	4	4	4	4	4	3	4	2	4	4	4		**65**
156	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	4	4	1	**65**
50	4	4	4	4	4	2	4	4	3	4	4	4	4	2	4	4	4	2	**65**
124	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	1	2	**65**
109	3	4	4	4	4	4	4	2	4	4	4	4	4	4	4	4		4	**65**
154	4	4	4	3	4	4	4	4	4	4	4	2	2	2	4	4	4	4	**65**
179	4	4	4	4	4	4	4	4	3	4	4	4	4	2	4	4		4	**65**
1	4	4	4	4	4	4	4	4	4	4	4	2	4	2	4	4	4		**64**
51	4	4	4	4	4	4	4	4	4	4	4	2	4	2	4	4	4		**64**
95	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	2		**64**
106	4	4	4	4	4	4	4	2	4	4	4	4	4	4	4	4	2		**64**
157	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	4	4		**64**
174	4	4	4	4	4	4	4	4	4	4	4	2	4	2	4	4	4		**64**
40	4	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	2	2	**64**
66	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**64**
68	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**64**
70	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**64**
103	4	4	4	4	4	4	4	2	4	4	4	2	4	4	4	4	2	2	**64**
165	4	4	3	4	4	4	4	4	4	4	4	4	2	4	4	4	1	2	**64**
191	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**64**
203	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**64**
214	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**64**
223	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**64**
232	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**64**
233	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**64**
242	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**64**
268	4	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	2	2	**64**
102	4	4	4	4	4	4	4	2	4	4	4	4	4	4	2	4		4	**64**
150	4	4	4	2	4	4	4	4	4	4	4	4	2	2	4	4	2	4	**64**

Genotype Class	03/21	04/07	04/20	05/02	05/05	05/09	05/12	05/15	05/24	05/30	06/02	06/06	06/09	06/12	06/16	06/23	06/29	07/13	All Date Sum
187	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4	2	4	**64**
189	4	2	4	4	4	2	4	4	2	4	4	4	4	4	4	4	2	4	**64**
193	4	2	4	4	4	4	4	4	4	4	4	4	4	2	4	4		4	**64**
194	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4	2	4	**64**
211	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4	2	4	**64**
221	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4	2	4	**64**
228	4	2	4	4	4	2	4	4	2	4	4	4	4	4	4	4	2	4	**64**
164	4	4	4	4	4	4	4	4	4	4	4	4	2	3	4	4	2		**63**
169	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	4	2	1	**63**
168	4	4	2	4	3	4	4	4	4	4	4	4	4	4	4	4		2	**63**
237	4	3	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**63**
269	4	4	4	3	4	4	4	4	4	4	4	4	2	2	4	4	2	2	**63**
155	4	4	4	2	4	4	4	4	4	4	4	2	2	2	4	4	4	3	**63**
230	4	2	4	4	4	4	4	4	4	4	4	4	4	2	4	4		3	**63**
39	4	4	2	4	3	4	4	4	2	2	4	4	4	4	4	4	2	4	**63**
197	4	2	4	4	4	2	4	4	2	2	4	4	4	4	4	4	3	4	**63**
206	3	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4	2	4	**63**
208	4	2	4	4	4	4	4	4	2	4	4	4	4	3	4	4		4	**63**
226	4	2	4	4	4	4	4	4	2	4	4	4	4	3	4	4		4	**63**
52	4	4	4	4	4	4	4	4	4	2	4	2	4	2	4	4	4		**62**
162	4	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	2		**62**
163	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	4	2		**62**
171	4	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4	2		**62**
204	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4			**62**
3	4	4	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2	2	**62**
33	4	4	2	4	2	4	4	4	4	4	4	4	4	4	4	4		2	**62**
130	4	4	4	4	4	2	4	4	4	4	4	4	2	4	4	4		2	**62**
133	4	4	4	4	4	2	4	4	4	4	4	4	2	4	4	4		2	**62**
167	4	4	4	2	4	4	4	4	4	4	4	4	2	2	4	4	2	2	**62**
227	4	2	4	4	4	4	4	4	4	4	4	4	4	2	4	4		2	**62**
245	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4	2	2	**62**
260	4	4	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2	2	**62**
264	4	4	4	4	4	4	4	4	2	2	4	4	2	4	4	4	2	2	**62**
4	4	4	2	4	2	4	4	4	2	2	4	4	2	4	4	4	4	4	**62**
15	4	4	2	4	2	4	4	4	2	2	4	4	2	4	4	4	4	4	**62**
16	4	4	2	4	2	4	4	4	2	2	4	4	2	4	4	4	4	4	**62**
62	4	4	2	4	2	4	4	4	2	2	4	4	2	4	4	4	4	4	**62**
74	4	2	4	4	4	4	4	4	2	4	4	4	2	2	4	4	2	4	**62**
76	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4		4	**62**

Genotype Class	03/21	04/07	04/20	05/02	05/05	05/09	05/12	05/15	05/24	05/30	06/02	06/06	06/09	06/12	06/16	06/23	06/29	07/13	All Date Sum
78	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4		4	**62**
192	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4		4	**62**
205	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4		4	**62**
218	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4		4	**62**
241	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4		4	**62**
252	4	4	2	4	2	4	4	4	2	2	4	4	2	4	4	4	4	4	**62**
255	4	4	2	4	2	4	4	4	2	2	4	4	2	4	4	4	4	4	**62**
256	4	4	2	4	2	4	4	4	2	2	4	4	2	4	4	4	4	4	**62**
47	4	4	2	4	4	4	4	4	2	2	4	4	2	4	4	4	3	2	**61**
71	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4	1	2	**61**
160	4	4	2	4	4	4	4	4	4	4	4	2	3	4	4	4		2	**61**
166	4	4	4	2	4	4	4	4	4	4	4	4	2	2	3	4	2	2	**61**
267	4	4	2	4	4	4	4	4	2	2	4	4	3	4	4	4	2	2	**61**
257	4	4	2	4	2	4	4	4	2	2	4	4	2	4	4	4	4	3	**61**
265	4	4	2	4	2	4	4	4	2	2	4	4	4	4	4	4	2	3	**61**
57	4	4	2	4	4	4	4	4	1	2	4	4	2	4	4	4	2	4	**61**
82	3	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2	4	**61**
85	4	3	2	4	4	2	4	4	2	2	4	4	4	4	4	4	2	4	**61**
67	4	4	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2		**60**
72	4	4	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2		**60**
73	4	4	4	4	4	4	4	4	4	4	4	4	2	2	4	4			**60**
112	4	4	4	4	4	4	4	2	2	4	4	2	4	4	4	4	2		**60**
176	4	4	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2		**60**
195	4	4	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2		**60**
196	4	4	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2		**60**
235	4	4	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2		**60**
277	4	4	4	4	4	4	4	4	4	2	4	4	2	2	4	4	2		**60**
161	4	4	4	3	4	4	4	4	4	4	4	2	2	2	4	4	2	1	**60**
7	4	4	2	4	4	4	4	4	2	2	4	2	2	4	4	4	4	2	**60**
8	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2	2	**60**
9	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2	2	**60**
12	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2	2	**60**
13	4	2	4	4	4	2	4	4	2	2	4	4	4	4	4	4	2	2	**60**
17	4	4	2	4	2	4	4	4	2	2	4	4	2	4	4	4	4	2	**60**
34	4	2	4	4	4	4	4	4	2	4	4	4	2	2	4	4	2	2	**60**
61	4	2	4	4	4	4	4	4	2	4	4	4	2	2	4	4	2	2	**60**
63	4	4	2	4	4	4	4	4	2	2	4	2	2	4	4	4	4	2	**60**
79	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4		2	**60**
84	4	4	2	4	4	4	4	4	2	2	4	2	2	4	4	4	4	2	**60**

Genotype Class	03/21	04/07	04/20	05/02	05/05	05/09	05/12	05/15	05/24	05/30	06/02	06/06	06/09	06/12	06/16	06/23	06/29	07/13	All Date Sum
113	4	4	2	4	2	4	4	4	4	4	4	4	4	4	2	4		2	**60**
117	4	4	2	4	2	4	4	4	4	4	4	4	4	4	2	4		2	**60**
121	4	4	2	4	2	4	4	4	4	4	4	4	4	4	2	4		2	**60**
123	4	4	2	4	2	4	4	4	4	4	4	4	4	4	2	4		2	**60**
136	4	4	4	4	4	4	4	2	2	4	4	4	4	4	2	4		2	**60**
137	4	4	4	4	4	4	4	2	2	4	4	4	4	4	2	4		2	**60**
138	4	4	4	4	4	4	4	2	2	4	4	4	4	4	2	4		2	**60**
144	4	4	4	4	4	4	4	2	2	4	4	4	4	4	2	4		2	**60**
170	4	4	2	4	2	4	4	4	4	4	4	2	4	4	4	4		2	**60**
173	4	4	4	4	4	2	4	4	2	2	4	4	4	2	4	4	2	2	**60**
183	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4		2	**60**
188	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4		2	**60**
198	4	4	2	4	4	4	4	4	2	2	4	2	2	4	4	4	4	2	**60**
219	4	2	4	4	4	4	4	4	2	4	4	4	4	2	4	4		2	**60**
248	4	4	2	4	4	4	4	4	2	2	4	2	2	4	4	4	4	2	**60**
249	4	4	2	4	2	4	4	4	2	2	4	4	2	4	4	4	4	2	**60**
251	4	4	2	4	4	4	4	4	4	2	4	2	2	2	4	4	4	2	**60**
261	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4	2	2	**60**
270	4	2	4	4	4	4	4	4	2	4	4	4	2	2	4	4	2	2	**60**
5	4	4	2	4	4	4	4	4	2	2	4	4	2	2	4	4	2	4	**60**
10	4	4	2	4	4	4	4	4		2	4	4	2	4	4	4	2	4	**60**
19	4	4	2	4	4	4	4	4		2	4	4	2	4	4	4	2	4	**60**
186	4	4	2	4	4	4	4	4	2	2	4	4	2	4	4	4		4	**60**
202	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4		4	**60**
215	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4		4	**60**
240	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4		4	**60**
244	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4		4	**60**
246	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4		4	**60**
247	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4		4	**60**
253	3	4	2	4	2	4	4	4	2	2	4	4	2	3	4	4	4	4	**60**
11	4	4	4	4	4	2	4	4	2	2	4	4	4	3	4	4	2		**59**
217	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4	1	2	**59**
28	4	4	4	4	4	4	4	4	2	2	4	4	2	2	4	4	2		**58**
30	4	2	4	4	4	4	4	4	2	4	4	4	2	2	4	4	2		**58**
38	4	2	4	4	4	4	4	4	2	4	4	4	2	2	4	4	2		**58**
56	4	2	4	4	4	4	4	4	2	4	4	4	2	2	4	4	2		**58**
114	4	4	2	4	2	4	4	4	4	4	4	4	4	4	2	4			**58**
116	4	4	2	4	2	4	4	4	4	4	4	4	4	4	2	4			**58**
279	4	4	4	4	4	4	4	2	2	4	4	4	2	2	4	4	2		**58**

Genotype Class	03/21	04/07	04/20	05/02	05/05	05/09	05/12	05/15	05/24	05/30	06/02	06/06	06/09	06/12	06/16	06/23	06/29	07/13	All Date Sum
281	4	2	4	3	4	4	4	4	2	4	4	4	2	2	4	4	2	1	**58**
18	4	4	2	4	2	4	4	4	2	2	4	4	2	2	4	4	4	2	**58**
120	4	4	2	4	2	4	4	4	4	4	4	4	4	2	2	4		2	**58**
122	4	4	2	4	2	4	4	4	4	4	4	4	4	2	2	4		2	**58**
125	4	4	2	4	2	2	4	2	4	4	4	4	4	4	4	4		2	**58**
127	4	4	2	4	2	2	4	2	4	4	4	4	4	4	4	4		2	**58**
128	4	4	2	4	2	2	4	2	4	4	4	4	4	4	4	4		2	**58**
129	4	4	2	4	2	2	4	2	4	4	4	4	4	4	4	4		2	**58**
131	4	4	2	4	2	2	4	2	4	4	4	4	4	4	4	4		2	**58**
132	3	4	2	4	3	2	4	2	4	4	4	4	4	4	4	4		2	**58**
134	4	4	2	4	2	2	4	2	4	4	4	4	4	4	4	4		2	**58**
139	4	4	2	4	2	4	4	2	4	4	4	4	4	4	2	4		2	**58**
147	4	4	2	4	2	4	4	2	4	4	4	4	4	4	2	4		2	**58**
177	4	4	2	4	4	4	4	4	2	2	4	2	2	2	4	4	4	2	**58**
178	4	2	4	4	4	2	4	4	2	4	4	4	4	2	4	4		2	**58**
184	4	4	2	4	4	4	4	4	2	2	4	4	2	4	4	4		2	**58**
207	4	4	2	4	4	4	4	4	2	2	4	4	2	4	4	4		2	**58**
220	4	4	2	4	4	4	4	4	2	2	4	4	2	4	4	4		2	**58**
231	4	4	2	4	4	4	4	4	2	2	4	4	2	4	4	4		2	**58**
250	4	4	2	4	2	4	4	4		2	4	4	2	4	4	4	4	2	**58**
258	4	2	4	4	4	2	4	4	2	2	4	4	4	2	4	4	2	2	**58**
263	4	2	4	4	4	2	4	4		2	4	4	4	4	4	4	2	2	**58**
22	4	2	2	4	2	4	4	4		2	4	4	4	4	4	4	2	4	**58**
32	4	4	2	4	2	4	4	2		2	4	4	4	4	4	4	2	4	**58**
37	4	2	2	4	4	4	4	4		2	4	4	2	4	4	4	2	4	**58**
42	4	2	2	4	4	2	4	4		2	4	4	2	4	4	4	4	4	**58**
49	4	4	2	4	4	2	4	4		2	4	4	2	4	4	4	2	4	**58**
199	4	2	2	4	4	2	4	4		2	4	4	2	4	4	4	4	4	**58**
262	4	2	2	4	4	2	4	4		2	4	4	2	4	4	4	4	4	**58**
271	4	2	2	4	4	4	4	4		2	4	4	2	4	4	4	2	4	**58**
119	4	4	2	4	2	4	4	4	4	4	4	4	4	3	2	4			**57**
41	4	2	4	4	4	4	4	2	2	4	4	4	2	2	4	4	2	1	**57**
145	4	4	2	4	2	2	4	2	4	4	4	4	4	4	4	4		1	**57**
273	4	2	4	2	4	4	4	4	2	4	4	4	2	2	4	4	2	1	**57**
115	3	4	2	4	2	4	4	4	4	4	4	4	4	2	2	4		2	**57**
259	4	4	2	4	4	2	4	4		2	4	4	3	4	4	4	2	2	**57**
266	4	2	4	2	4	4	4	2	3	4	4	4	2	2	4	4	2	2	**57**
6	4	4	2	4	4	3	4	4		2	4	4	2	4	4	4		4	**57**
26	3	4	2	4	2	4	4	2		2	4	4	4	4	4	4	2	4	**57**

Genotype Class	03/21	04/07	04/20	05/02	05/05	05/09	05/12	05/15	05/24	05/30	06/02	06/06	06/09	06/12	06/16	06/23	06/29	07/13	All Date Sum
54	4	3	2	4	2	4	4	2		2	4	4	4	4	4	4	2	4	**57**
222	4	2	2	4	4	4	4	4	1	2	4	4	2	4	4	4		4	**57**
278	4	2	2	4	3	4	4	4		2	4	4	2	4	4	4	2	4	**57**
29	4	4	4	2	4	4	4	2	2	4	4	4	2	2	4	4	2		**56**
31	4	4	4	2	4	4	4	2	2	4	4	4	2	2	4	4	2		**56**
53	4	2	4	2	4	4	4	4	2	4	4	4	2	2	4	4	2		**56**
118	4	4	2	4	2	4	4	4	4	4	4	4	4	2	2	4			**56**
140	4	4	2	4	2	4	4	2	4	4	4	4	4	4	2	4			**56**
141	4	4	2	4	2	4	4	2	4	4	4	4	4	4	2	4			**56**
142	4	4	2	4	2	2	4	2	4	4	4	4	4	4	4	4			**56**
143	4	4	2	4	2	2	4	2	4	4	4	4	4	4	4	4			**56**
146	4	4	2	4	2	2	4	2	4	4	4	4	4	4	4	4			**56**
209	4	4	2	4	4	4	4	4	4	2	4	4	2	2	4	4			**56**
213	4	4	2	4	4	4	4	4	4	2	4	4	2	2	4	4			**56**
25	4	2	4	4	4	4	4	2	2	2	4	4	2	2	4	4	2	2	**56**
60	4	2	2	4	2	4	4	4		2	4	4	4	4	4	4	2	2	**56**
88	4	4	2	4	2	4	4	2	2	2	4	4	4	4	4	4		2	**56**
175	4	4	2	4	4	2	4	4		2	4	4	2	4	4	4	2	2	**56**
254	4	4	2	4	2	4	4	4		2	4	4	2	4	4	4	2	2	**56**
272	4	2	2	4	2	4	4	4		2	4	4	4	4	4	4	2	2	**56**
274	4	2	2	4	4	4	4	4		2	4	4	2	4	4	4	2	2	**56**
276	4	2	2	4	4	4	4	4		2	4	4	2	4	4	4	2	2	**56**
280	4	2	2	4	2	4	4	4		2	4	4	4	4	4	4	2	2	**56**
58	4	2	2	4	2	4	4	2		2	4	4	4	4	4	4	2	4	**56**
83	4	2	2	4	4	2	4	4		2	4	4	2	4	4	4	2	4	**56**
210	4	2	2	4	4	4	4	4		2	4	4	2	4	4	4		4	**56**
229	4	2	2	4	4	4	4	4		2	4	4	2	4	4	4		4	**56**
212	4	4	2	4	4	4	4	4	3	2	4	4	2	2	4	4			**55**
23	4	3	2	4	2	4	4	2		2	4	4	4	4	4	4	2	2	**55**
24	3	2	2	4	2	4	4	4		2	4	4	4	4	4	4	2	2	**55**
35	3	2	4	2	4	4	4	2	2	4	4	4	2	2	4	4	2	2	**55**
36	4	2	2	4	2	4	4	2	1	2	4	4	4	4	4	4	2	2	**55**
55	3	2	2	4	2	4	4	4		2	4	4	4	4	4	4	2	2	**55**
46	3	2	2	4	4	2	4	4		2	4	4	2	4	4	4	2	4	**55**
86	4	2	2	4	2	2	4	4	1	2	4	4	4	2	4	4	2	4	**55**
224	4	2	4	4	4	2	4	4	2	4	4	4	2	2	4	4			**54**
20	4	2	2	4	2	4	4	2		2	4	4	4	4	4	4	2	2	**54**
21	4	2	2	4	2	4	4	2	2	2	4	4	4	4	4	4		2	**54**
48	4	4	2	4	4	2	4	4		2	4	4	2	4	4	4		2	**54**

Genotype Class	03/ 21	04/ 07	04/ 20	05/ 02	05/ 05	05/ 09	05/ 12	05/ 15	05/ 24	05/ 30	06/ 02	06/ 06	06/ 09	06/ 12	06/ 16	06/ 23	06/ 29	07/ 13	All Date Sum
87	4	4	2	4	4	2	4	4		2	4	4	2	2	4	4	2	2	**54**
275	4	2	2	4	2	4	4	2		2	4	4	4	4	4	4	2	2	**54**
43	4	2	2	4	2	2	4	4		2	4	4	4	2	4	4	2	4	**54**
44	4	2	2	4	4	2	4	4		2	4	4	2	4	4	4		4	**54**
27	2	4	3	4	4	4	4	2		2	4	4	2	3	4	4	2	1	**53**
181	4	2	2	4	4	2	4	4	2	2	4	4	2	2	4	4	1	2	**53**
200	4	2	2	4	4	2	4	4		2	4	4	2	4	4	4		2	**52**
80	4	2	2	4	4	2	4	4		2	4	4	2	2	4	4		2	**50**
93	3	3	3	3	3	3	3	3	3	3	3	2	3	3	3	3	2		**49**
190	3	3	3	3	3	3	3	3	3	3	3	3	3	2	3	3		1	**48**
126	3	3	3	3	3	1	3	3	3	3	3	3	3	3	3	3		2	**48**
59	2	2	2	2	2	2	2	2			2	2			2	2	2		**26**

APPENDIX D

Variable	**Abbreviation**	**Definition**
Image Layer Summary Statistics		
Mean-Blue	MEAN_BLUE	Combined mean of the blue image layer for all image objects over a row
Mean-CHM	MEAN_CHM	Combined mean of the CHM image layer for all image objects over a row
Mean-Green	MEAN_GREEN	Combined mean of the green image layer for all image objects over a row
Mean-ExG	MEAN_EXG	Combined mean of the ExG image layer for all image objects over a row
Mean-ExR	MEAN_EXR	Combined mean of the ExR image layer for all image objects over a row
Mean-Red	MEAN_RED	Combined mean of the RED image layer for all image objects over a row
Mean-VEG	MEAN_VEG	Combined mean of the VEG image layer for all image objects over a row
Mean-YI	MEAN_YI	Combined mean of the YI image layer for all image objects over a row
Mean-Wavelet Scale 0.2	WV_SCALE_0_2MEAN	Combined mean of the wavelet image layer at a scale of 0.2 for all image objects over a row
Mean-Wavelet Scale 0.4	WV_SCALE_0_4MEAN	Combined mean of the wavelet image layer at a scale of 0.4 for all image objects over a row
Mean-Wavelet Scale 0.6	WV_SCALE_0_6MEAN	Combined mean of the wavelet image layer at a scale of 0.6 for all image objects over a row
Mean-Wavelet Scale 0.8	WV_SCALE_0_8MEAN	Combined mean of the wavelet image layer at a scale of 0.8 for all image objects over a row

Mean-Wavelet Scale 1	WV_SCALE_1MEAN	Combined mean of the wavelet image layer at a scale of 1 for all image objects over a row
Mean-Wavelet Scale 1.2	WV_SCALE_1_2MEAN	Combined mean of the wavelet image layer at a scale of 1.2 for all image objects over a row
Mean-Wavelet Scale 1.4	WV_SCALE_1_4MEAN	Combined mean of the wavelet image layer at a scale of 1.4 for all image objects over a row
Mean-Wavelet Scale 1.6	WV_SCALE_1_6MEAN	Combined mean of the wavelet image layer at a scale of 1.6 for all image objects over a row
Mean-Wavelet Scale 1.8	WV_SCALE_1_8MEAN	Combined mean of the wavelet image layer at a scale of 1.8 for all image objects over a row
Mean-Wavelet Scale 2	WV_SCALE_2MEAN	Combined mean of the wavelet image layer at a scale of 2 for all image objects over a row
Mean-Wavelet Scale 2.2	WV_SCALE_2_2MEAN	Combined mean of the wavelet image layer at a scale of 2.2 for all image objects over a row
Mean-Wavelet Scale 2.4	WV_SCALE_2_4MEAN	Combined mean of the wavelet image layer at a scale of 2.4 for all image objects over a row
Mean-Wavelet Scale 2.6	WV_SCALE_2_6MEAN	Combined mean of the wavelet image layer at a scale of 2.6 for all image objects over a row
Mean-Wavelet Scale 2.8	WV_SCALE_2_8MEAN	Combined mean of the wavelet image layer at a scale of 2.8 for all image objects over a row
Mean-Wavelet Scale 3	WV_SCALE_3MEAN	Combined mean of the wavelet image layer at a scale of 3 for all image objects over a row
Mean-Wavelet Scale 3.2	WV_SCALE_3_2MEAN	Combined mean of the wavelet image layer at a scale of 3.2 for all image objects over a row

Mean-Wavelet Scale 3.4	WV_SCALE_3_4MEAN	Combined mean of the wavelet image layer at a scale of 3.4 for all image objects over a row
Mean-Wavelet Scale 3.6	WV_SCALE_3_6MEAN	Combined mean of the wavelet image layer at a scale of 3.6 for all image objects over a row
Mean-Wavelet Scale 3.8	WV_SCALE_3_8MEAN	Combined mean of the wavelet image layer at a scale of 3.8 for all image objects over a row
Mean-Wavelet Scale 4	WV_SCALE_4MEAN	Combined mean of the wavelet image layer at a scale of 4 for all image objects over a row
Mean-Wavelet Scale 4.2	WV_SCALE_4_2MEAN	Combined mean of the wavelet image layer at a scale of 4.2 for all image objects over a row
Mean-Wavelet Scale 4.4	WV_SCALE_4_4MEAN	Combined mean of the wavelet image layer at a scale of 4.4 for all image objects over a row
Mean-Wavelet Scale 4.6	WV_SCALE_4_6MEAN	Combined mean of the wavelet image layer at a scale of 4.6 for all image objects over a row
Mean-Wavelet Scale 4.8	WV_SCALE_4_8MEAN	Combined mean of the wavelet image layer at a scale of 4.8 for all image objects over a row
Mean-Wavelet Scale 5	WV_SCALE_5MEAN	Combined mean of the wavelet image layer at a scale of 5 for all image objects over a row
Standard Deviation-Blue	STD_BLUE	Combined standard deviation of the blue image layer for all image objects over a row
Standard Deviation-CHM	STD_CHM	Combined standard deviation of the CHM image layer for all image objects over a row
Standard Deviation-Green	STD_GREEN	Combined standard deviation of the green image layer for all image objects over a row

Standard Deviation-ExG	STD_ExG	Combined standard deviation of the ExG image layer for all image objects over a row
Standard Deviation-ExR	STD_ExR	Combined standard deviation of the ExR image layer for all image objects over a row
Standard Deviation-Red	STD_Red	Combined standard deviation of the RED image layer for all image objects over a row
Standard Deviation-VEG	STD_VEG	Combined standard deviation of the VEG image layer for all image objects over a row
Standard Deviation-YI	STD_YI	Combined standard deviation of the YI image layer for all image objects over a row
Standard Deviation-Wavelet Scale 0.2	WV_SCALE_0_2STD	Combined standard deviation of the wavelet image layer at a scale of 0.2 for all image objects over a row
Standard Deviation-Wavelet Scale 0.4	WV_SCALE_0_4STD	Combined standard deviation of the wavelet image layer at a scale of 0.4 for all image objects over a row
Standard Deviation-Wavelet Scale 0.6	WV_SCALE_0_6STD	Combined standard deviation of the wavelet image layer at a scale of 0.6 for all image objects over a row
Standard Deviation-Wavelet Scale 0.8	WV_SCALE_0_8STD	Combined standard deviation of the wavelet image layer at a scale of 0.8 for all image objects over a row
Standard Deviation-Wavelet Scale 1	WV_SCALE_1STD	Combined standard deviation of the wavelet image layer at a scale of 1 for all image objects over a row
Standard Deviation-Wavelet Scale 1.2	WV_SCALE_1_2STD	Combined standard deviation of the wavelet image layer at a scale of 1.2 for all image objects over a row

Standard Deviation-Wavelet Scale 1.4	WV_SCALE_1_4STD	Combined standard deviation of the wavelet image layer at a scale of 1.4 for all image objects over a row
Standard Deviation-Wavelet Scale 1.6	WV_SCALE_1_6STD	Combined standard deviation of the wavelet image layer at a scale of 1.6 for all image objects over a row
Standard Deviation-Wavelet Scale 1.8	WV_SCALE_1_8STD	Combined standard deviation of the wavelet image layer at a scale of 1.8 for all image objects over a row
Standard Deviation-Wavelet Scale 2	WV_SCALE_2STD	Combined standard deviation of the wavelet image layer at a scale of 2 for all image objects over a row
Standard Deviation-Wavelet Scale 2.2	WV_SCALE_2_2STD	Combined standard deviation of the wavelet image layer at a scale of 2.2 for all image objects over a row
Standard Deviation-Wavelet Scale 2.4	WV_SCALE_2_4STD	Combined standard deviation of the wavelet image layer at a scale of 2.4 for all image objects over a row
Standard Deviation-Wavelet Scale 2.6	WV_SCALE_2_6STD	Combined standard deviation of the wavelet image layer at a scale of 2.6 for all image objects over a row
Standard Deviation-Wavelet Scale 2.8	WV_SCALE_2_8STD	Combined standard deviation of the wavelet image layer at a scale of 2.8 for all image objects over a row
Standard Deviation-Wavelet Scale 3	WV_SCALE_3STD	Combined standard deviation of the wavelet image layer at a scale of 3 for all image objects over a row
Standard Deviation-Wavelet Scale 3.2	WV_SCALE_3_2STD	Combined standard deviation of the wavelet image layer at a scale of 3.2 for all image objects over a row
Standard Deviation-Wavelet Scale 3.4	WV_SCALE_3_4STD	Combined standard deviation of the wavelet image layer at a scale of 3.4 for all image objects over a row
Standard Deviation-Wavelet Scale 3.6	WV_SCALE_3_6STD	Combined standard deviation of the wavelet image layer at a scale of 3.6 for all image objects over a row

Standard Deviation-Wavelet Scale 3.8	WV_SCALE_3_8STD	Combined standard deviation of the wavelet image layer at a scale of 3.8 for all image objects over a row
Standard Deviation-Wavelet Scale 4	WV_SCALE_4STD	Combined standard deviation of the wavelet image layer at a scale of 4 for all image objects over a row
Standard Deviation-Wavelet Scale 4.2	WV_SCALE_4_2STD	Combined standard deviation of the wavelet image layer at a scale of 4.2 for all image objects over a row
Standard Deviation-Wavelet Scale 4.4	WV_SCALE_4_4STD	Combined standard deviation of the wavelet image layer at a scale of 4.4 for all image objects over a row
Standard Deviation-Wavelet Scale 4.6	WV_SCALE_4_6STD	Combined standard deviation of the wavelet image layer at a scale of 4.6 for all image objects over a row
Standard Deviation-Wavelet Scale 4.8	WV_SCALE_4_8STD	Combined standard deviation of the wavelet image layer at a scale of 4.8 for all image objects over a row
Standard Deviation-Wavelet Scale 5	WV_SCALE_5STD	Combined standard deviation of the wavelet image layer at a scale of 5 for all image objects over a row
Texture		Gray-level co-occurrence matrices (GLCM) is a tabulation of how often different combinations of pixel gray levels occur in a scene.
GLCM Contrast (45˚)	GLCM_CONTRAST_45˚	The amount of local variation in gray-level co-occurrence matrices in an image for all image layers at a 45˚ direction
GLCM Contrast (45˚)-CHM	GLCM_CONTRAST_CHM_45˚	The amount of local variation in gray-level co-occurrence matrices in an image for CHM image layer at a 45˚ direction
GLCM Contrast (45˚)-Blue	GLCM_CONTRAST_BLUE_45˚	The amount of local variation in gray-level co-occurrence matrices in an image for blue image layer at a 45˚ direction
GLCM Contrast (45˚)-Green	GLCM_CONTRAST_GREEN_45˚	The amount of local variation in gray-level co-occurrence matrices in an image for green image layer at a 45˚ direction

GLCM Contrast (45˚)-Red	GLCM_CONTRAST_RED_45˚	The amount of local variation in gray-level co-occurrence matrices in an image for red image layer at a 45˚ direction
GLCM Contrast (45˚)-ExG	GLCM_CONTRAST_EXG_45˚	The amount of local variation in gray-level co-occurrence matrices in an image for ExG image layer at a 45˚ direction
GLCM Contrast (45˚)-ExR	GLCM_CONTRAST_EXR_45˚	The amount of local variation in gray-level co-occurrence matrices in an image for ExR image layer at a 45˚ direction
GLCM Contrast (45˚)-VEG	GLCM_CONTRAST_VEG_45˚	The amount of local variation in gray-level co-occurrence matrices in an image for VEG image layer at a 45˚ direction
GLCM Contrast (45˚)-YI	GLCM_CONTRAST_YEL_45˚	The amount of local variation in gray-level co-occurrence matrices in an image for YI image layer at a 45˚ direction
GLCM Contrast All Directions	GLCM_CONTRAST_All_DIR	The amount of local variation in gray-level co-occurrence matrices of all image layers at 45˚, 90˚, 135˚, and 180˚ directions
GLCM Entropy (45˚)	GLCM_ENTROPY_45˚	A measurement of how equally distributed elements of the grey-level co-occurrence matrices are in an image for all image layers at a 45˚ direction
GLCM Entropy (45˚)-CHM	GLCM_ENTROPY_CHM_45˚	A measurement of how equally distributed elements of the grey-level co-occurrence matrices are in an image for CHM image layer at a 45˚ direction
GLCM Entropy (45˚)-Blue	GLCM_ENTROPY_BLUE_45˚	A measurement of how equally distributed elements of the grey-level co-occurrence matrices are in an image for blue image layer at a 45˚ direction
GLCM Entropy (45˚)-Green	GLCM_ENTROPY_GREEN_45˚	A measurement of how equally distributed elements of the grey-level co-occurrence matrices are in an image for green image layer at a 45˚ direction
GLCM Entropy (45˚)-Red	GLCM_ENTROPY_RED_45˚	A measurement of how equally distributed elements of the grey-level co-occurrence matrices are in an image for red image layer at a 45˚ direction

GLCM Entropy (45˚)-ExG	GLCM_ENTROPY_EXG_45˚	A measurement of how equally distributed elements of the grey-level co-occurrence matrices are in an image for ExG image layer at a 45˚ direction
GLCM Entropy (45˚)-ExR	GLCM_ENTROPY_EXR_45˚	A measurement of how equally distributed elements of the grey-level co-occurrence matrices are in an image for ExR image layer at a 45˚ direction
GLCM Entropy (45˚)-VEG	GLCM_ENTROPY_VEG_45˚	A measurement of how equally distributed elements of the grey-level co-occurrence matrices are in an image for VEG image layer at a 45˚ direction
GLCM Entropy (45˚)-YI	GLCM_ENTROPY_YEL_45˚	A measurement of how equally distributed elements of the grey-level co-occurrence matrices are in an image for YI image layer at a 45˚ direction
GLCM Entropy All Directions	GLCM_ENTROPY_All_DIR	A measurement of how equally distributed elements of the grey-level co-occurrence matrices are in all image layers at 45˚, 90˚, 135˚, and 180˚ directions
GLCM Mean (45˚)	GLCM_MEAN_45˚	The average of the grey-level co-occurrence matrices in an image for all image layers at a 45˚ direction
GLCM Mean (45˚)-CHM	GLCM_MEAN_CHM_45˚	The average of the grey-level co-occurrence matrices in an image for CHM image layer at a 45˚ direction
GLCM Mean (45˚)-Blue	GLCM_MEAN_BLUE_45˚	The average of the grey-level co-occurrence matrices in an image for blue image layer at a 45˚ direction
GLCM Mean (45˚)-Green	GLCM_MEAN_GREEN_45˚	The average of the grey-level co-occurrence matrices in an image for green image layer at a 45˚ direction

GLCM Mean (45˚)-Red	GLCM_MEAN_RED_45˚	The average of the grey-level co-occurrence matrices in an image for red image layer at a 45˚ direction
GLCM Mean (45˚)-ExG	GLCM_MEAN_EXG_45˚	The average of the grey-level co-occurrence matrices in an image for ExG image layer at a 45˚ direction
GLCM Mean (45˚)-ExR	GLCM_MEAN_EXR_45˚	The average of the grey-level co-occurrence matrices in an image for ExR image layer at a 45˚ direction
GLCM Mean (45˚)-VEG	GLCM_MEAN_VEG_45˚	The average of the grey-level co-occurrence matrices in an image for VEG image layer at a 45˚ direction
GLCM Mean (45˚)-YI	GLCM_MEAN_YEL_45˚	The average of the grey-level co-occurrence matrices in an image for YI image layer at a 45˚ direction
GLCM Mean All Directions	GLCM_MEAN_All_DIR	The average of the grey-level co-occurrence matrices in all image layers at 45˚, 90˚, 135˚, and 180˚ directions
Geometric		
Area / Number of objects	AREA_BY#_OBJ	summed area of all image objects comprising a row multipart polygon divided by the number of images
Asymmetry	ASYMMETRY	compares an image object with an approximated ellipse around the given image object. The calculation is based on the variance in x-direction and y-direction, which results in a feature value that increases with an increasing image object asymmetry
Border Index	BORDER_INDEX	uses a rectangular approximation to measure how jagged an image object is. It is calculated as the ratio between the length and width of the image object and the smallest rectangle enclosing rectangle.
Border Length - Length of longest edge	BL_SUB_LLED	sum of all image objects's edges minus the length of the longest edge

Contig Index Distribution mean	CONTIG_MN	mean of spatial connectedness, or contiguity, of cells within a grid-cell patch to provide an index on patch boundary configuration and thus patch shape for all patches of a class.
Contig Index Distribution area weighted mean	CONTIG_AM	area weighted mean of spatial connectedness, or contiguity, of cells within a grid-cell patch to provide an index on patch boundary configuration and thus patch shape for all patches of a class.
Contig Index Distribution range	CONTIG_RA	range of spatial connectedness, or contiguity, of cells within a grid-cell patch to provide an index on patch boundary configuration and thus patch shape for all patches of a class.
Curvature / Length (of main)	CURV_BY_LGT	length-to-curvature ratio of the main line of an object. The curvature is the sum of all changes in direction of the mainline. Changes in direction are expressed by the acute angle a in which sections of the mainline, built by the connection between the nodes, cross each other
Degree of skeleton branching	DEG_SKL_BRCH	The highest order of branching in an object.
Edge Density	ED	the sum of the lengths of all edge segments involving the corresponding patch type, divided by the total landscape area, multiplied by 2 10,000 (to convert to hectares).
Fractal Dimension Index mean	FRAC_MN	mean of 2 times the logarithm of patch perimeter divided by the logarithm of patch area for all patches of a class; the perimeter is adjusted to correct for the raster bias 2 in perimeter.
Fractal Dimension Index area weighted mean	FRAC_AM	area weighted mean of 2 times the logarithm of patch perimeter divided by the logarithm of patch area for all patches of a class; the perimeter is adjusted to correct for the raster bias 2 in perimeter.

Fractal Dimension Index standard deviation	FRAC_SD	standard deviation of 2 times the logarithm of patch perimeter divided by the logarithm of patch area for all patches of a class; the perimeter is adjusted to correct for the raster bias 2 in perimeter.
Length	LENGTH	length of the smallest enclosing rectangle able to fit around an image object
Length / Width	LG_BY_WD	length of the smallest enclosing rectangle able to fit around an image object divided by the width of the smallest enclosing rectangle able to fit around an image object
Length of longest edge	LENGTH-LONG-EDGE	length of longest image object edge or border
Length of main line	LGT_MAIN	The sum of all distances between the nodes of the main line of an object
Number of Disjunct Core Areas	NDCA	The sum of the number of disjunct core areas contained within each patch of the corresponding patch type; that is, the number of disjunct core areas contained within the landscape
Number of inner objects	#_INNER_OBJS	The number of inner polygons that are completely surrounded by a selected outer polygon.
Number of objects	#_OF_OBJS	Number of image objects within the row boundary polygon
Patch Cohesion Index mean	COHESION_MN	this value is the mean of all patches of a class and is created as 1 minus the sum of patch perimeter (in terms of number of cell surfaces) divided by the sum of patch perimeter times the square root of patch area (in terms of number of cells) for patches of the corresponding patch type, divided by 1 minus 1 over the square root of the total number of cells in the landscape, multiplied by 100 to convert to a percentage.

Patch Cohesion Index area weighted mean	COHESION_AM	this value is the area weighted mean of all patches of a class and is created as 1 minus the sum of patch perimeter (in terms of number of cell surfaces) divided by the sum of patch perimeter times the square root of patch area (in terms of number of cells) for patches of the corresponding patch type, divided by 1 minus 1 over the square root of the total number of cells in the landscape, multiplied by 100 to convert to a percentage.
Patch Cohesion Index standard deviation	COHESION_SD	this value is the standard deviation of all patches of a class and is created as 1 minus the sum of patch perimeter (in terms of number of cell surfaces) divided by the sum of patch perimeter times the square root of patch area (in terms of number of cells) for patches of the corresponding patch type, divided by 1 minus 1 over the square root of the total number of cells in the landscape, multiplied by 100 to convert to a percentage.
Perimeter-area ratio mean	PARA_MN	mean of the ratio of the patch perimeter to area for all patches of a class
Perimeter-area ratio area weighted mean	PARA_AM	area weighted mean of the ratio of the patch perimeter to area for all patches of a class
Perimeter-area ratio standard deviation	PARA_SD	standard deviation of the ratio of the patch perimeter to area for all patches of a class
Percent of Landscape	PLAND	the sum of the areas of all patches of the corresponding 2 patch type, divided by total landscape area, multiplied by 100 (to convert to a 2 percentage); in other words, PLAND equals the percentage the landscape comprised of the corresponding patch type.
Shape Index mean	SHAPE_MN	mean of the patch perimeter divided by the square root of patch area for all patches of a designated class, adjusted by a constant to adjust for a square standard

Shape Index area weighted mean	SHAPE_AM	area weighted mean of the patch perimeter divided by the square root of patch area for all patches of a designated class, adjusted by a constant to adjust for a square standard
Shape Index standard deviation	SHAPE_SD	the standard deviation patch perimeter divided by the square root of patch area for all patches of a designated class, adjusted by a constant to adjust for a square standard
Total Class Area	CA	equals the sum of the areas of all patches of the corresponding patch 2 type, divided by 10,000 (to convert to hectares)
Total Edge	TE	the sum of the lengths of all edge segments involving the corresponding patch type
Width	WIDTH	width of the smallest enclosing rectangle able to fit around an image object
		(McGarigal 2015) (Trimble 2019)