Spring 2023

Modeling Growth and Stress Factors for Converted Silvopasture Systems in the Missouri Ozarks

Bailee N. Suedmeyer

MODELING GROWTH AND STRESS FACTORS FOR CONVERTED SILVOPASTURE SYSTEMS IN THE MISSOURI OZARKS

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Plant Science

By

Bailee Suedmeyer

May 2023

MODELING GROWTH AND STRESS FACTORS FOR CONVERTED SILVOPASTURE SYSTEMS IN THE MISSOURI OZARKS

Environmental Plant Science and Natural Resources

Missouri State University, May 2023

Master of Science

Bailee Suedmeyer

ABSTRACT

Silvopasture systems are becoming increasingly popular among sustainable agriculture ranchers, due to the increase in knowledge of benefits to the cattle and ability to grow cool season grasses beneath the canopy. This project focuses on the forest crop aspect of silvopasture systems from monitoring of the health of the trees over time to recommendations for thinning management to keep it functioning as viable silvopasture. The study site consists of five acres of upland hardwood forest area in Southern Missouri with 18 monumented fixed-area plots. Arial and ground data was collected at each plot throughout the growing season, along with data from a weather station located in the stand. Multispectral indices and cover percentages were extracted from aerial data through Metashape and ArcGIS processing. The data was used to develop prediction models for chlorophyll and water potential as well as canopy cover percentages. An additional analysis was conducted, comparing data from our drone operations with publicly accessible data through Landsat and local weather stations. It was determined that with multiple years of data collected from the same month, strong prediction models are achievable for all three of the variables in question. While ground metrics do improve the models, they can be removed when modeling over multiple years, with little decrease in model predictive power. Climate metrics had little influence on any of the models and were therefore not used for most final predictions. Similarly, Landsat multispectral indices are also able to predict strong models without the inclusion of weather station or ground metrics.

KEYWORDS: silvopasture, ArcGIS, Landsat, prediction models, multispectral, R Studio

MODELING GROWTH AND STRESS FACTORS FOR CONVERTED SILVOPASTURE SYSTEMS IN THE MISSOURI OZARKS

By

Bailee Suedmeyer

A Master's Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Plant Science

May 2023

Approved:

Michael Goerndt, Ph.D., Thesis Committee Chair

Melissa Bledsoe, Ph.D., Committee Member

Will McClain, Ph.D., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

ACKNOWLEDGEMENTS

I would like to say thank you to many people that have helped me out over the course of my graduate studies at Missouri State University. They have helped me in growing in my academics as well as a person and I will forever be grateful for their support. First, I want to thank my professor, Dr. Michael Goerndt, for allowing me to come to Missouri State University and trusting me with this project. He has helped me throughout this whole experience, being there for me when I needed it, but also letting me gain independence and confidence in my work. I will forever be grateful for the small jokes and experiences Dr. Goerndt has opened the door for. Next, I want to thank Dr. Toby Dogwiler for the many questions answered regarding the GIS aspect of this project. Dr. Dogwiler not only helped me with this project, but also has helped me gain a love for this line of work and the problem solving that comes with it. I would also like to thank Dr. McClain and Dr. Bledsoe for all their help with the field data collection aspect of this project. I am very thankful for the great team we were able to have at all the field collection days. Carsen Goodman and Stewart McCollum have also helped me throughout this project. I want to thank Carsen for all the help throughout my first year at Missouri State with data collection and maintenance. Finally, I want to thank my family and everyone that I have met through this process and throughout my time at Missouri State University. I could not have done it without the overwhelming support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Increases in degradation of land by intensive agriculture have caused a worldwide movement to discover more sustainable ways of reaching producer goals. For the cattle industry, this has meant implementing practices such as rotational grazing and the incorporation of silvopasture systems. Silvopasture is a form of sustainable agroforestry and has positive impacts both for livestock and the land. Research has focused primarily on the livestock aspect of silvopasture and studies have shown there have been many positive impacts of rotating cattle through savannah like stands. Benefits to livestock have included increased weight gain, decreased heat stress, and continued growth of nutrient dense, cool season grasses through hot summer months.

While the health of trees have not been studied extensively for silvopasture, studies in forested areas to identify stressors related to climate change and other factors are common. Remote sensing has been used to positively link stress indicators such as water potential with multispectral bands and indices through statistical analysis. Increased fluctuations and seasonal changes from year to year also can play a major role in stress of the trees. Weather stations can help to capture these changes and potentially be used as variables to predict stress factors.

Canopy cover changes are another factor that has not been heavily studied in silvopasture systems. Natural savannahs however, have been studied in this realm. Since this study uses a converted forest stand in the central hardwoods region, it will likely not show exact growth rates of a natural savannah and will require more frequent work to maintain the ideal canopy cover due to the absence of disturbance. It is important to ensure that canopy

cover stays under 40-60% to ensure adequate light can reach ground level of the understory. This will allow growth of cool season grasses to provide forage for grazing cattle.

The objective of this study was to further the knowledge of tree health in silvopasture systems. This was done through the creation of models for predicting stand stress and growth for the converted silvopasture system at Journagan Ranch (Missouri State University, Mountain Grove, MO). Many different models were used to compare predicting potential of the covariates. This included models with only one year of data, two years of data, only one month of data per year, inclusion of cover percentages, and a comparison between drone and satellite data. From these models, statistics was done in R Studio to determine relationships between x and y variables and analyze the ability of the model to create predictions based on the variables and datasets. An additional study was conducted to determine how often thinning needs to take place to keep the stand functioning as a silvopasture.

LITERATURE REVIEW

Silvopasture

For generations, ranchers have looked at forested land as having little value. However, as a natural resource, silvopasture is beneficial for carbon sequestration and soil fertility (Soares et al., 2018). Through research, the value of actively managed, savannah like, forest systems has gained popularity in their benefits to both livestock and the environment (USDA, 2008).

Livestock benefit from silvopasture stands in many ways. Shaded understory has been positively linked with increased weight gain in hot, summer months. Understory can be 11-25 °F cooler allowing for protection from heat and less weight loss due to sweat (Klopfenstein et al., 1997). Forage is also improved in the understory due to the ability to grow cool season grasses. In a typical system, lower quality warm season grasses will begin to grow as the summer gets hotter and less water is available. Growth of cool season grasses throughout the growing season gives cattle access to more nutritious, dense forage (Klopfenstein et al., 1997). Maintaining 40-60% light penetration into the understory will allow cool season grasses to grow and survive in the understory (USDA, 2008). Silvopasture also retains more water than a standard system due to protection of soil by the canopy (USDA, 2008). Protection of soil and retention of water could help keep forage from experiencing drought-like conditions.

Remote Sensing in Forestry

Remote sensing of forested areas has gained interest and popularity to study effects of disturbances and changes in the environment caused naturally or by human involvement (Lausch et al., 2016). Multispectral data are commonly used for forest composition and classification with advances in technology continuing to improve accuracy of data (Ustin et al., 1998).

Data from remote sensing allows for visualization and statistical analysis of what is happening to those trees under specific conditions. The results help to define what is a healthy stand and what are ideal conditions for keeping the stand functioning at its best (Lopez-Sanchez and Roig, 2017). When a stand is under stress, whether disturbance or a resource limitation, stress indicators will be apparent in statistical analysis. Agencies at federal and state levels have developed programs for monitoring these changes and identifying patterns and trends to better understand what is happening as well as what this may mean for the future of forests (Lausch et al., 2016).

Historically, forest health has had many definitions depending on what is being studied. At tree level, forest health is commonly determined by visual characteristics such as: disease presence, damage, crown status, or leaf characteristics (Lausch et al., 2016). Remote sensing offers many other forest health indicators. Without multispectral indices, observations are limited to wavelengths of the visible light spectrum (Sykas, 2020).

Remote sensing also helps bridge data gaps locally, and across the world. Online data clouds allow for mass sharing and availability of resources from around the globe (Lausch et al.,

2016). Historical archived data can also be used for various future projects to effectively analyze trends over time without spending years collecting data before results can be obtained.

Weather Stations

Traditionally, weather stations have not accounted for forest stands. Guidelines set out by the World Meteorological Organization state that sensors should be installed far away from all trees (De Frenne et al., 2016). Lack of forest weather station data puts a large gap in data regarding climate change and indicators because 27% of land surface is under forest cover. Forests are also home to two-thirds of the biodiversity on Earth that are being affected by different climatic factors (De Frenne et al., 2016).

Canopy Cover Modeling

In forestry, relationships between aerial photography and multispectral imagery have been used to examine health of trees. However, recent studies have examined canopy cover changes with normalized difference vegetation index changes (Soares et al., 2018). Monitoring change in forests using aerial data has been shown to be effective in seeing changes over time and in reducing time used for data collection (Ustin et al., 1998). Arial data with GPS pinpoints allow for exact comparisons of stands overtime by directly overlaying images. Lining data up accurately is important when pinpointing factors leading to mortality such as infections, pollution, natural die off, and drought (Ustin et al., 1998). Distinguishing between trees and herbaceous understory can help to understand changes in ecology and composition of a site (Carreiras et al., 2006). Due to changes in interception of rainfall and radiation, separating the

canopy can give us better measure of water and chlorophyll fluxuation within trees leading to stronger prediction for the future (Carreiras et al., 2006).

Landsat

Landsat was created in the 1960s to capture Earths dynamics and changes in the land overtime (Loveland and Dwyer, 2012). Since it's original launch, there have been six successful missions. These missions have improved capabilities and clarity of satellite imagery. Landsat data used in this project is from Landsat 9. Vandenberg Space Force Based launched Landsat 9 from California in September of 2021 (NASA, 2020). NASA states that the satellite orbits Earth 14 times a day and takes over 700 images per day. Landsat 9 is equipped with more sensors than the 8th version. It now can extract nine spectral bands and two thermal inferred bands (Lulla et al., 2021). Spatial resolution of the Landsat 9 satellite has also increased to 15m, 30m or 100m, depending on the band. Lulla et al. (2021) states that testing has been done to check scientific usage of different Landsat Satellite based location accuracy. It was determined that the satellite images are accurately calibrated to use as long-term comparisons. The ability to compare data long-term has made Landsat Satellites the backbone of Earth observations overtime (Lulla et al., 2021). All 50+ years of data collected from various Landsat satellites can be used to look at changes over time without issues in lining up images.

Chlorophyll and Water Potential

Studies have shown there is a relationship between chlorophyll and water potential contents in leaves of vegetation with image spectrometer data. Seasonal changes and drought

have been detected and can be predicted with spectral data (Ustin et al., 1998). Chlorophyll content in leaves plays a key role in analyzing the health of many crops (Zhou et al., 2020). Chlorophyll plays a crucial role in photosynthesis as it absorbs light. Changes in abundance and presence of chlorophyll can impact survival of many plants. Chlorophyll can be analyzed using remote sensing techniques will not damage the plant itself (Zhou et al., 2020). One major cause of stress on trees starts with lack of water. Drought is a dangerous, but natural, occurrence in the world and monitoring helps understand effects of climate change (Abid et al., 2018).

Multispectral Predictors

Light reflectance is used commonly for identification of objects based on the visible light spectrum. Using multispectral sensors, images can be manipulated to enhance or alter coloring of certain factors to help distinguish between objects (Horning, 2004). Multispectral images have been shown to be effective in identifying healthy and unhealthy vegetation. There are five bands, all consisting of different wavelengths, that are commonly used and combined to emphasize different factors of a landscape. These bands are red, blue, green, red edge, and near inferred (NIR) (Sykas, 2020). The human eye is only able to distinguish between blue, green, and red, but plant stress is most visible at the red edge band and near inferred wavelengths (Micasense, 2018). The red color band helps distinguish vegetation out of an image. Vegetation absorbs all red light making it useful when determining vegetation from soil or other features. The red band can be referred to as the chlorophyll absorption band (Horning, 2004) The blue color band is commonly used for distinguishing between plant life or soil and water and the green band (Horning, 2004). Red edge shows the greatest difference in

reflectance between health and stressed vegetation will be seen while NIR is where stressed plants reflectance will exceed healthy plants reflectance as seen in Figure 1 (Micasense, 2018).

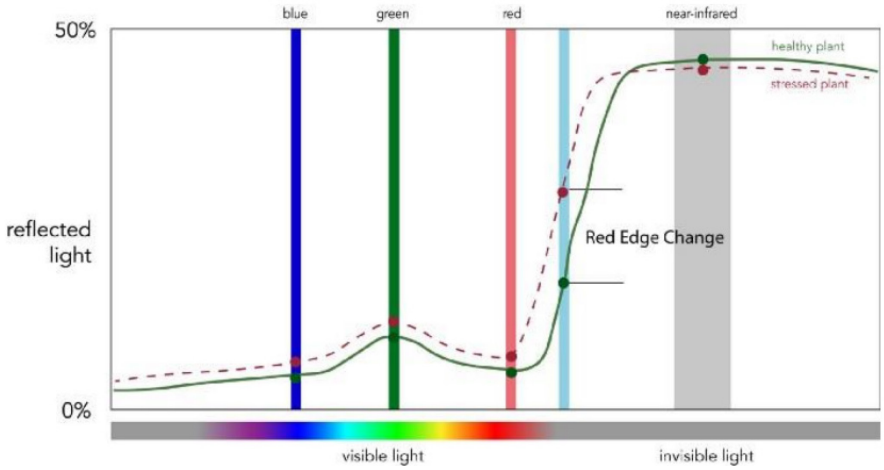


Figure 1. Micasense Red Edge-M reflected band chart showing health of vegetation in relation to light reflectance (Micasense, 2018)

With five color bands, different combinations can be created to help create a greater emphasis on target factors of vegetation such as chlorophyll and water presence in the leaves. The Normalized Difference Vegetation Index (NDVI) is a combination of the NIR and red bands. NDVI combination highlights NIR reflectance indicating healthy vegetation and has also been referred to as the green-ness index. A lower value for NDVI would indicate vegetation that is weak or stressed (Sykas, 2020). NDVI is commonly used to map vegetation cover and characteristics. NDVI changes throughout the season allowing for visualization of what is happening to the canopy throughout the growing period. NDVI is also used commonly for chlorophyll predictions due to having a strong relationship between the NDVI value and chlorophyll presence in vegetation (Ustin et al., 1998). Green Normalized Difference Vegetation Index (GDNVI) uses green and near inferred bands to estimate photosynthetic

activity. GNDVI combination is used to visualize nitrogen and water movement into the canopy (Evangelides and Nobajas, 2020). Red Edge Normalized Vegetation Index (RENDVI) is similar to NDVI but focuses on a smaller waveband for chlorophyll absorption and is more related to chlorophyll absorption than NDVI. Due to the narrow band used, sensitivities are also more prevalent in the RENDVI combination (Evangelides and Nobajas, 2020).

Statistical Modeling

Regression is used to find relationships between independent and dependent variables in many statistical analyses. Variables used for predicting stress can be seen in Table 1. Linear regression models are commonly used in accessing relationships between aerial and LiDAR (Light Detection and Ranging) data (Soares et al., 2018). Once determined to have a linear relationship, models can then be used to estimate different factors based on already collected data. Statistical modeling involves examining the relationship and significance of all of independent variables in determining dependent variable. Data can be split into training and testing sets to then check predictability of the final model through cross-validation (Zach, 2019).

Table 1. List of covariates for the models including multispectral bands, multispectral indices, ground measurements, and weather station measurements

Description	Variables
Multispectral Bands	Red
	Blue
	Green
	Red Edge
	NIR
Multispectral Indices	NDVI
	GDNVI
	RENDVI
	Cover
Ground Measurements	Temperature
	Moisture
Weather Station Measurements	Pressure Monthly (pm)
	Pressure Weekly (pw
	Temperature Monthly (tm)
	Temperature Weekly (tw)
	Moisture Monthly (mm)
	Moisture Weekly (mw)

METHODS

Study Site

The data for this study was collected from a converted silvopasture stand at Journagan Ranch in Mountain Grove, Missouri. Journagan was donated to Missouri State University in 2010. The 3300-acre ranch is in the heart of the Ozarks and includes many habitats such as grasslands, woodlands, and riparian areas. The study area is in a valley that consists of open pastures, dense woodland, and glades as seen in Figure 2. The five-acre stand was selected in 2019 and is on a NW facing slope. Soil in the converted stand is very rocky, increasing towards the southeast corner. Species composition is diverse, including oaks (*Quercus*), hickories (*Carya*), walnuts (*Juglans*), ash (*Fraxinus*), dogwood (*Cornus*), redbud (*Cercis*), sassafras (*Sassafrass*), spicebush (*Lindera*), and buckbrush (*Ceanothus*).

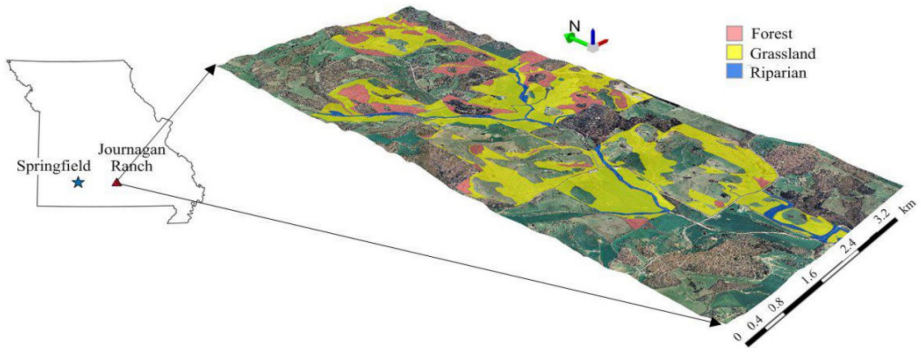


Figure 2. Journagan Ranch map (McCollum, 2021)

The original thinning of the converted stand took place in 2019. Prior to thinning, an inventory was done to determine the stocking level. Trees were selected to be removed and a

bulldozer was used to extract trees, including roots. After thinning, a secondary inventory was done. With around 50% of the stand removed and 30% stocking remaining, the stand went from fully stocked to understocked, according to the Gingrich Stocking Chart (Gingrich, 1967). Once removed, the ground was smoothed and later seeded for fescue in 2020. A timeline can be found in the Appendix (A-1).

Ground Data

Ground data was taken around three times per growing season, depending on weather conditions. These were July, August, and September. Data was taken systematically with five marked points for all 18 (.1 acre) plots in the stand as seen in Figure 3. At these points, light intensity, soil moisture, and soil temperature were recorded. At the beginning and end of each row, full light intensity was recorded to have a baseline measurement. An Apogee Instruments® MQ-306 Line Quantum Photosynthetically Active Radiation sensor was used to measure light intensity, a Campbell Scientific® HydroSense II handheld soil moisture meter measured soil moisture, and a SpotOn® digital temperature probe measure soil temperature at all 90 points within the stand. Leaves in the upper canopy of marked trees in all 18 plots were extracted by shotgun and put into pressure chambers to record chlorophyll and water potential. Chlorophyll was measured using an Apogee Instruments MC-100 chlorophyll meter and was given a relative value based on greenness of the leaves. Water potential in the leaves were measured using a PMS Instrument Company Model 600 pressure chamber. Water potential and chlorophyll data were entered into a datasheet for statistical analysis. Annual inventories consisting of diameter

at breast height (DBH), basal area (BA), crown width, and crown density were also collected for annual growth comparisons.

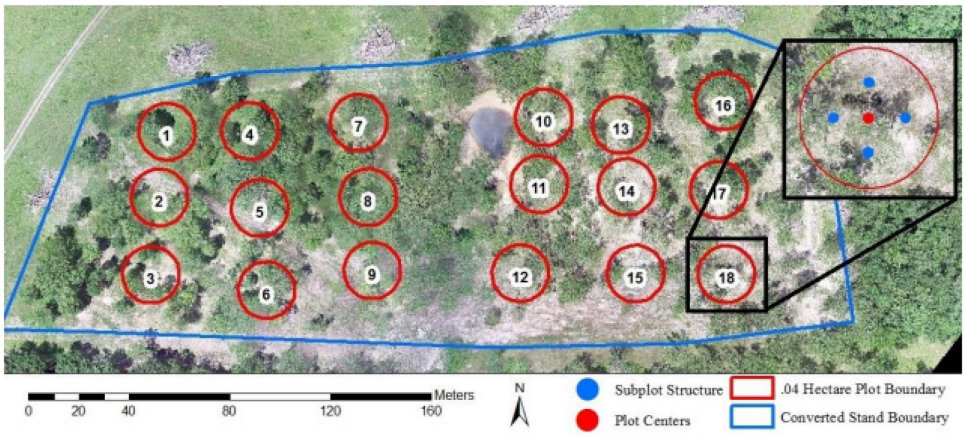


Figure 3. Layout of the 18 - .1-acre plots in the stand. The stand boundary is outlined in blue, and individual plots are outlined in red. Numbers represent the plot identification number. Each plot has five marked subplots, including the plot center, where ground data was taken (McCollum, 2021)

Aerial Data

Aerial data was collected using a Phantom 4 Professional Quadcopter Unmanned Aerial Vehicle (UAV). The drone was equipped with multispectral and remote sensing cameras including Micasence Red Edge-M, allowing for five bands: red, green, blue, red edge, and near inferred (NIR). Arial data was recorded during full sun hours of the day (11am-1pm) on low wind, low cloud cover days. The drone flew a predetermined pattern at 85 meters above ground with 80% forward and side overlap, resulting in around 25 images for every point on the ground. Resolution for remote sensing images was 2.3 cm and multispectral images where 5.9 cm. Markers were set in specific areas to use during later processing to ensure accuracy during modeling. Figure 4 shows the product of a drone flight, post processing.

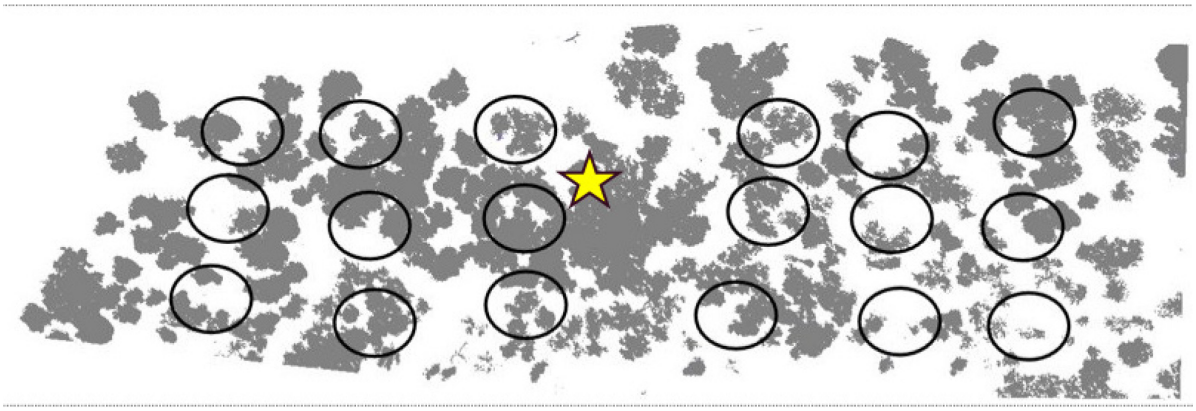


Figure 4. Pictured is the canopy cover model for 2022 with .1-acre plots circled and a star representing approximate location of the weather station in the stand.

Weather Station Data

A weather station was positioned in the middle of the converted stand at Journagan Ranch. The weather station collected data on temperature above and below ground, soil moisture, humidity, pressure, precipitation, and wind direction and speed. Due to issues with data collection and extraction from the Journagan weather station during certain time periods, data was also collected from a nearby weather station in Mountain Grove, about nine miles North of the stand location. The Mountain Grove weather station is owned by the University of Missouri and is at the Fruit Experiment Station.

Stand Model

The stand model was created using both remote sensing and multispectral images extracted from the drone. Sensors on the drone capture height of the canopy and ground layers by shooting lasers straight down and capturing time it takes to return to the drone. Differences in time are caused by canopy interactions with the laser and can then be used to

make canopy height or cover models. The same interactions are done to record values for multispectral bands by determining what is being absorbed and reflected by the object the laser is interacting with. Images were then put into Agisoft Metashape PhotoScan where they were put through many processes to create the full stand model. Metashape workflow, USGS Unmanned Aircraft Systems Data Post-Processing, was used to create the models, with minor alterations to account for noise reduction in the canopy (USGS, 2017). Model creation in Metashape was done for remote sensing and multispectral images separately. Products of each model were a digital elevation model (DEM), digital surface model (DSM), and an orthomosaic photo. A processing workflow can be seen in Appendix B-1.

Canopy Height Model (CHM)

The canopy height model was created using DEM and DSM products from Agisoft Metashape. These images were opened in ARCGIS Pro and clipped to the stand boundary. CHM was produced by subtracting DSM from DEM as seen in Figure 5. CHM calculation was done with a raster calculator and resulted in a model of just the canopy. From the CHM product, canopy cover percentages for each plot were extracted. The reclassify tool in ARCGIS was used to change all pixel values of the canopy image as either “1” or “NO DATA”. Reclassification of the CHM created a clipped polygon of the crown itself. Plot boundaries were then added to the image and the sum of pixels in every plot was extracted using zonal statistics. To convert summed pixel values into percentages, reclassification was set to make all values equal “1”. The number of pixels contained in every plot were recorded and became the baseline for percentages of crown cover per plot. The baseline value for number of pixels per

plot was re-calculated for every data collection date due to slight changes in values between flights. Changes in values between flights was likely due to variations in height of the drone over the plot from one flight to the next. A workflow of the CHM process in ArcGIS Pro can be seen in Appendix B-2.

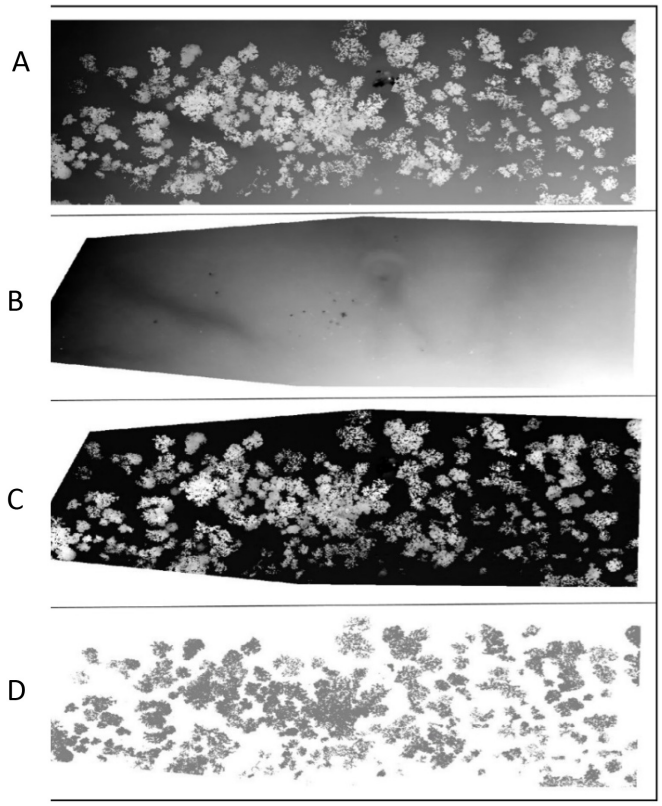


Figure 5. Canopy Height Model (CHM) creation images. Image A, DSM (highpoints), minus B, DEM (Ground points), creates C, canopy height model. Image D is the reclassified polygon.

Multispectral Indices

The multispectral indices were extracted using the orthophoto derived from the MS stand model in Metashape. The orthophoto was then put into ArcGIS Pro and red, blue, green, NIR, and red edge bands were extracted separately. Raster calculations shown in Figure 6 were used to extract the remaining indices (NDVI (1), GNDVI (2), RENDVI (3)). Once these were all

separate images, plot boundaries were included into the images. Spatial analyst zonal statistics were performed on all images and averages were recorded for each plot. Full workflow for creating and extracting multispectral indices in ArcGIS Pro can be seen in Appendix B-3.

$$\frac{NIR - Red}{NIR + Red} = NDVI \quad (1)$$

$$\frac{NIR - Green}{NIR + Green} = GNDVI \quad (2)$$

$$\frac{NIR - RedEdge}{NIR + RedEdge} = RENDVI \quad (3)$$

Figure 6. Equations for multispectral indices (Iizuka et al., 2019). Bands were respective of program band numbering used for the drone data and Landsat data.

Statistical Analysis

Statistical analysis of data was completed using R Studio Statistical Software. Our variables of interest for predicting were chlorophyll (CH), water potential (WP), and cover percentage (Cover). A stand level summary of ground data measurements for each data collection day can be viewed in Table 2.

Seven sets of data (Table 3) were run with six types of variable combinations for each set. Variable combinations included: WP with ground data included (WGW), WP without ground data (NGW), WP with only multispectral data (MSW), and respectively for CH predictions (WGC, NGC, MSC). The July data set was also run to predict cover percentage and had three additional sets for predicting cover (WGCover, NGCover, MSCover).

As seen in Table 3, gray shaded boxes were not considered for specific models. August 2020 data was corrupted and was therefore not included for the 2020 or 2020-2022 datasets.

Cover was only used for July with cover dataset. Landsat 9 satellite does not include sensors for Red Edge which also impacts availability of RENDVI data which was also not included in the model. While cover was available for the Landsat dates based on our image collection, it was not publicly available data and was not considered in the Landsat model.

Table 2. Summary Table including estimates of mean, min, max, and standard deviation for stand level ground data (CH, WP, Soil Moisture, Soil Temperature) for each collection date. (n=18).

		Chlorophyll			
Year	Month	mean	min	max	stDev
2020	7	13.82	7.00	27.20	4.32
	9	14.49	6.30	21.90	5.08
2021	7	13.53	8.20	20.00	3.11
	8	14.26	4.60	24.90	4.64
	9	13.71	7.20	24.20	4.73
2022	7	12.74	6.60	20.80	4.43

		Water Potential (bars)			
Year	Month	mean	min	max	stDev
2020	7	25.89	11.00	35.00	6.39
	9	27.72	21.00	39.00	4.75
2021	7	24.83	21.00	31.00	2.83
	8	32.14	25.00	40.00	4.83
	9	32.06	22.00	40.00	5.18
2022	7	26.39	23.00	37.50	3.69

		Soil Moisture (%)			
Year	Month	mean	min	max	stDev
2020	7	17.39	10.90	28.12	4.68
	9	4.24	1.28	11.02	2.44
2021	7	20.22	13.58	28.72	3.83
	8	5.34	2.16	10.00	2.30
	9	4.48	2.50	7.60	1.59
2022	7	2.83	0.90	6.32	1.41

		Soil Temperature (fahrenheit)			
Year	Month	mean	min	max	stDev
2020	7	79.94	76.64	84.38	2.52
	9	65.78	62.10	69.26	2.18
2021	7	73.52	71.58	75.98	1.32
	8	81.00	77.32	85.18	1.98
	9	74.13	71.38	75.92	1.19
2022	7	79.78	75.10	84.70	2.56

The first analysis step on individual, month-specific data sets was stepwise selection (Marhuenda et al. 2014). Weather data was not considered in any month-specific analysis due

Table 3. The Table shows variables included prior to stepwise selection for all six data sets. Gray boxes symbolize factors that were not included in models for the specific data sets.

Data Set	Months			Years			Y Variables		X Variables																	
	7	8	9	20	21	22	WP	CH	Red	Blue	Green	Red Edge	NIR	NDVI	GNDVI	RENDVI	Moisture	Temperature	Cover	pm	pw	tm	tw	mm	mw
2020	x		x	x			x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x
2020-2022	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x
No Cover	x			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x
With Cover	x			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Landsat Comparison	x			x	x	x	x	x	x	x	x		x	x	x		x	x		x	x	x	x	x	x
Landsat	x			x	x	x	x	x	x	x	x		x	x	x		x	x		x	x	x	x	x	x

to inability to run linear models on data with duplicated values. Stepwise selection tests significance of independent variables in the model and how they relate to the dependent variable (Marhuenda et al., 2014). The remaining variables were put into the full model for the next step. Once significant variables were identified, month-specific linear models were run using the lm() function from the LMER package in R Studio (Bates et al., 2015). Results of the lm() function included p-values and adjusted-r^2 values which were assessed for model strength. Workflow for multiple linear regression for month-specific data can be seen in Appendix B-4.

Stepwise selection was also run first on yearly data sets from Table 3. Once variables of interest were determined, mixed linear regression models were run from the LMER package using the lmer() function (Bates et al., 2015). For mixed regression analysis, date was considered a random time variable. Weather data was included prior to stepwise selection, but not significant enough to be included for most mixed models. Once mixed models were complete, summary data including p-values were obtained. Residual plots were then created and extracted from R Studio. In these models, p-values under .05 indicate significance for

individual variables. P-values between .05 and .1 indicate marginal significance. Workflow for mixed regression models can be seen in Appendix B-5.

Leave one out cross validation was then run on the models to determine their ability to predict values not included in observations. Cross validation removes observation specific effects that could cause bias in the model. The result is a percentage error showing precision of the model predictions for a set of data not used in formulation of the model versus actual measure values for that same dataset. This was done by randomly splitting the dataset into training and testing models. Predictions from the testing dataset were then preformed. Glm() function was used, from the Boot package, and cross validation error was extracted from the results (Davison and Hinkley, 1997). Cross validation errors for all models were compared to look for predictability based on the model.

Landowner Modeling System

A landowner modeling system was developed using information that is publicly available for landowners to freely access. Information for the landowner model was extracted from Landsat 9 and a University of Missouri weather station. Landsat 9 is not equipped to capture Red Edge band, which excludes RENDVI and Red Edge raster values from the model. Canopy cover is available through Landsat Cover Analysis but is not calculated annually and was therefore not used for the landowner model. ArcGIS was used for the model as well to duplicate the process of the rest of the models on this project. The landowner model was designed for individuals who do not have the ability to use a drone or weather station can do to capture and predict canopy closure and stress levels on their stands. Data used for these

models were taken on dates close to the July 2020-2022 dates which allowed for comparisons between drone and satellite data. Using these similar images, it can be determined how accurately plant stress metrics can be predicted with publicly available information. For a landowner to predict stress in their own stands, ground measures of CH and WP would be needed for model training. The statistical software, R Studio, is available to the public for free downloading and usage.

Study Limitations

One limitation stated previously is inability to use climate metrics that were collected from the weather station directly in the stand. This was due to missing dates. While the weather station that was used had similar values, they were all higher and could have been why climate metrics are not significantly used in any models.

Potential causes of error with the landowner model included changes in raster cell sizes from drone data to satellite data. The drone has a resolution of 6cm while the satellite's lowest band resolution is 15m. Inability to extract RedEdge and RENDVI values from Landsat data was a limitation, but these covariates were excluded for the July comparison model. However, with future Landsat Missions, these indices may be available. Another limitation with the Landowner model is inability to get images from the exact same day and time as drone collected data. Satellite images were chosen based on date proximity and clarity.

The cover models and cover prediction models were also limited on accuracy of the cover models due to movement of the canopy creating noise in the dataset and inability to capture every leaf on the trees. Days were selected to reduce noise by wind and cloud cover,

and the same processing was performed when removing noise from the models to ensure adequate detection of growth changes. The thinning model was also limited to using plots experiencing increases in crown cover. This reduced sample size when developing the growth prediction model and did not account for natural mortality in the canopy.

RESULTS

Climate Stress Models

2020. From the July month-specific multiple regression model, majority of p-values for water potential were significant and adjusted-r^2 was relatively high. The chlorophyll model had low significance of variables with only one variable having a p-value below .05 for both the model including and excluding ground data. Adjusted-r^2 was also low for the chlorophyll model. The September 2020 model for water potential had no p-values under .05 and had a low adjusted-r^2. The chlorophyll model for September did improve with an adjusted-r^2 of 47% without ground data and 52% with ground data included. P-values were only significant in the model excluding ground data.

The full 2020 data was put into three mixed models for predicting WP and CH: all variables (WGW and WGC), no ground data (NGW and NGC), and MS data only (MSW and MSC). When looking at p-values, there was overall low significance for covariates among all models when compared to later models as seen in C-1 through C-6 of the Appendices. The model with highest significance, relative to p-values, for water potential was the WGW model. The chlorophyll model with highest significance was the MSC model. According to Table 4, the NGC model was the same as MSC due to including the same variables, excluding weather data. Variable selection among chlorophyll models stayed consistent, using NIR, RedEdge, and NDVI in every model. For WG and NG models, weather station data was included, but was eliminated in the stepwise selection process for all models except the model including ground data for CH. The CH model only used month-long pressure average, and it was deemed to have marginal

significance indicating that it does not have a strong influence in prediction of CH. In models including ground data, moisture and temperature were used every time. P-values for these models can be seen in Table 4.

Table 4. Mixed regression model p-values for the variables used to predict water potential and chlorophyll from the lmer() function of the LMER package in R Studio (Bates et al. 2015) using the 2020 data. Predictions models were created using ground data (WGW and WGC), excluding ground data (NGW and NGC) and multispectral data only (MSW and MSC). P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

	Water Potential			Chlorophyll		
Coefficiant	WGW	NGW	MSW	WGC	NGC	MSC
intercept	0.7140*	0.7443	0.0001**	0.0563*	0.0000**	0.0000**
Red		0.1137				
Blue					0.1073	0.1073
Green		0.0561*				
NIR	0.0530*			0.0011**	0.0003**	0.0003**
RedEdge	0.0510*		0.0909*	0.0021**	0.0005**	0.0005**
NDVI	0.0447**	0.1060		0.1601	0.1583	0.1583
GNDVI	0.1207	0.0546*				
RENDVI			0.0879*		0.0005**	0.0005**
Moisture	0.0005**			0.0022**		
Temperature	0.0014**			0.1474		
Pressure M				0.0606*		
Pressure W						
Temperature M						
Temperature W						
Moisture M						
Moisture W						

The residual plots for 2020 models can be seen in Figure 7. When looking at residuals, we can assume that the data is normal due to lack of heteroscedasticity or bias in the plots. While margins of residual values on the y axis are not large, residual points within the plot are not as tightly clustered about the zero-line, indicating that predictions were not as strong as they could be. When comparing models that included ground data (A and D) excluded ground data (B and E) and multispectral only (C and F), there is no obvious change in predictability as they all have similar shapes and scatter of residuals along the x and y axes.

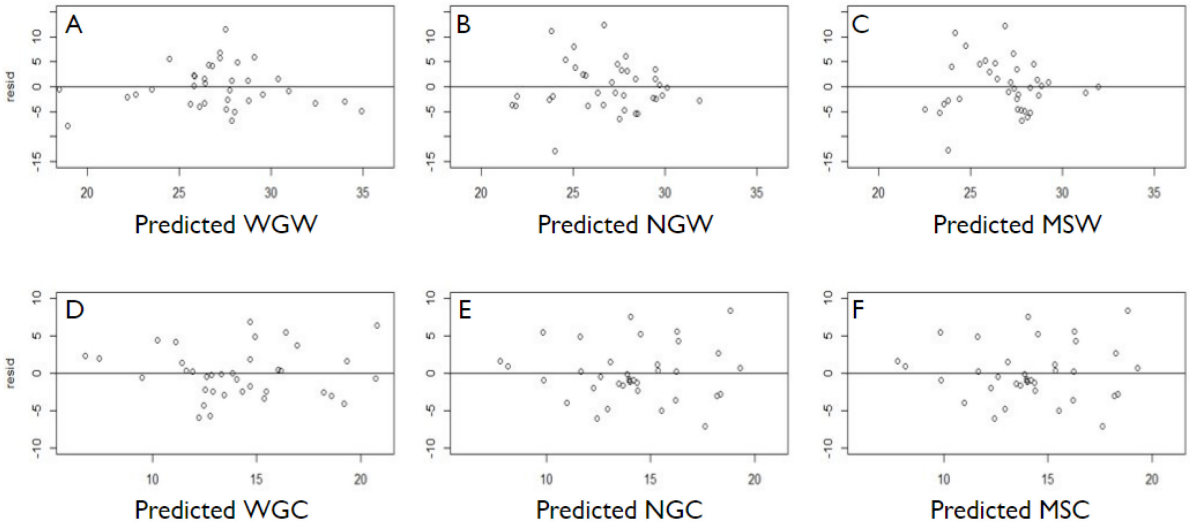


Figure 7. These are the residual plots for the 2020 model. Plots A, B, and C are predicting water potential in bars and plots D, E, and F are predicting chlorophyll. Plots A and D include ground data, B and E do not include ground data, and C and F include only multispectral data.

2020-2022. For the multiple linear regression model, individually looking at each month, adjusted-r^2 WP for most months was under 50% and CH percentages did not exceed 70%. For month-specific models, 60% of variables were considered at least marginally significant in individual models. 40% of variables that remained in the model after stepwise selection were not significant enough to contribute greatly to the model. Covariates selected to stay in the

model most often were red and green bands. Ground variables, moisture and temperature, were only included half of the time.

When comparing models based on significance of p-values alone from the mixed model, there was high significance for most variables across all models as seen Table 5. For multispectral variable selection when predicting water potential, red band, NIR and NDVI were used for all models. When looking at chlorophyll, red band, blue band, green band, NIR, NDVI, and GNDVI were included in every model.

The residual plots for mixed models of predictions based on the 2020-2022 models (Figure 8) show relatively similar patterns and margins, when compared to the 2020 residual plots in Figure 7. With inclusion of more points, but no obvious change in residual margins, it can be inferred that the model is becoming stronger with inclusion of more data. This is because additional residual points are clustered about the zero line tighter, which means predictions were closer to the actual value. An increase in residual margins would indicate prediction potential is not changing or getting worse with more data. Predictions for WGW are more spread out across the x axis, however, remain close to the zero-line.

Leave one out cross-validation is a test of the models' ability to predict values from a dataset that were not included in model creation. Error is determined by percent change from the actual, measured, value to the predicted value based on the model equation. When comparing cross-validation error percentages (Figure 9) for the models including only one year (2020) and three years (2020-2022), it was evident that three years had less error in the prediction calculations for both chlorophyll and water potential. For water potential

predictions, models including ground data (WG) had a lower percentage error. However, for chlorophyll predictions, ground data increased percentage error.

Table 5. Mixed regression model p-values for the variables used to predict water potential and chlorophyll from the lmer() function of the LMER package in R Studio (Bates et al. 2015) using the 2020-2022 data. Predictions models were created using ground data (WGW and WGC), excluding ground data (NGW and NGC) and multispectral data only (MSW and MSC). P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

	Water Potential			Chlorophyll		
Coefficiant	WGW	NGW	MSW	WGC	NGC	MSC
intercept	0.01917**	0.0006**	0.1375	0.0741*	0.0051**	0.0051**
Red	0.0050**	0.0053**	0.0062**	0.0000**	0.0004**	0.0004**
Blue				0.0188**	0.0732	0.0732
Green				0.0000**	0.0000**	0.0000**
NIR	0.0085**	0.0106**	0.0117**	0.0000**	0.0000**	0.0000**
RedEdge						
NDVI	0.0267**	0.0430**	0.2470	0.0001**	0.0001**	0.0001**
GNDVI			0.0649	0.0000**	0.0000**	0.0000**
RENDVI						
Moisture	0.0000**					
Temperature	0.0000**			0.0159**		
Pressure M		0.0001**				
Pressure W		0.0331**		0.0396**		
Temperature M						
Temperature W						
Moisture M	0.0000**	0.0258**				
Moisture W						

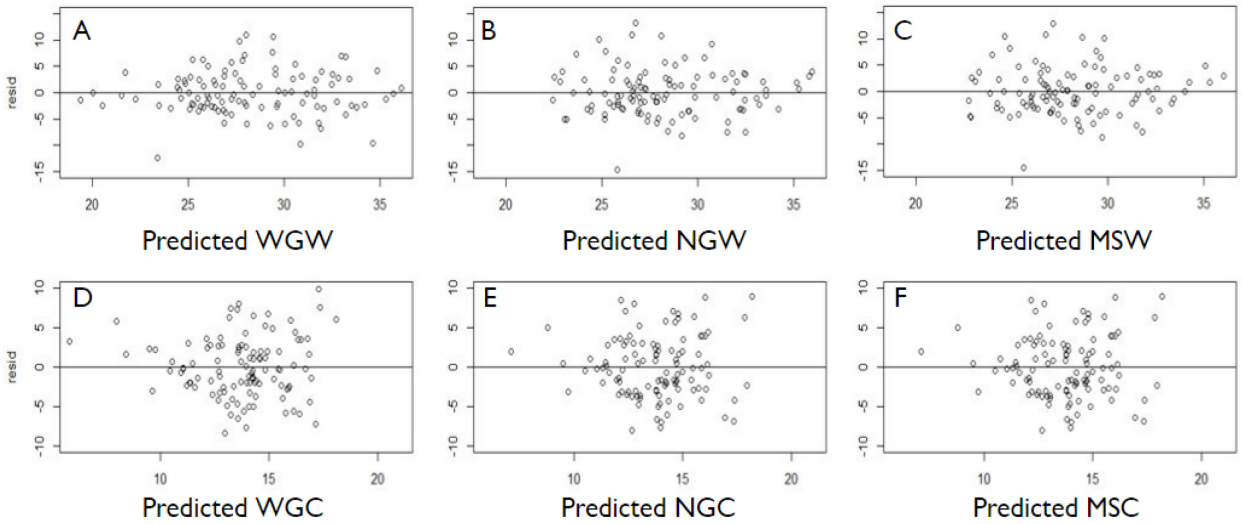


Figure 8. These are the residual plots for the full 2020-2022 model. Plots A, B, and C are predicting water potential in bars and plots D, E, and F are predicting chlorophyll. Plots A and D include ground data, B and E do not include ground data, and C and F include only multispectral data.

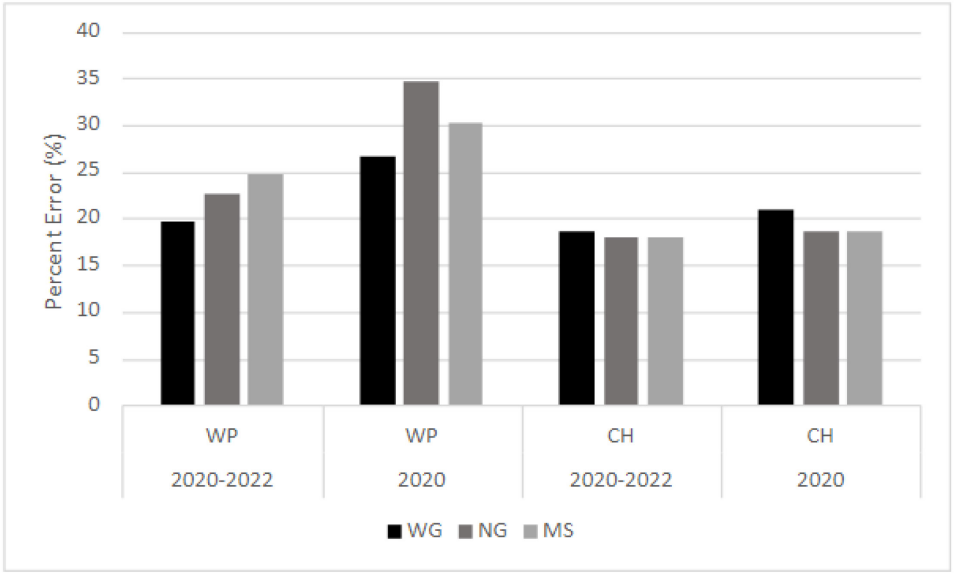


Figure 9. Cross Validation Error Percentage when predicting water potential (WP) and chlorophyll (CH) for 2020 and 2020-2022 data. WG = all variables, NG = without ground measurements, MS = multispectral indices only.

July No Cover. When looking at p-values (Table 6) from mixed regression of the July dataset, not including cover, all models showed high significance for individual variables. Every variable included in the models after stepwise selection had a p-value of less than .05. Across all models for July No Cover models, the most used covariates after stepwise selection were the green band, GNDVI, and RENDVI. These indices were kept in the final models after stepwise selection 67% of the time. Temperature was included in every model that it was available.

The residuals for these models (Figure 10) continue to show similar patterns as previous residual plots. Margins for the residual points for both water potential and chlorophyll are smaller, as the points are clustered tighter about the zero-line. These plots continue to support assumptions that the model does not have any issues with bias or non-linearity. Like the 2020-2022 residual plots in Figure 9, the WGW plot continues to be stretched further on the x axis.

The comparison of percent error from cross validation of the 2020-2022 models and the July no cover models in Figure 11, July no cover had a lower percentage error in making predictions for water potential and chlorophyll. The pattern of ground data increasing the ability for the model to predict water potential accurately and decreasing accuracy for chlorophyll continues with July no cover model.

Table 6. Mixed regression model p-values for the variables used to predict water potential and chlorophyll from the lmer() function of the LMER package in R Studio (Bates et al. 2015) using the July no cover data. Predictions models were created using ground data (WGW and WGC), excluding ground data (NGW and NGC) and multispectral data only (MSW and MSC). P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

	Water Potential			Chlorophyll		
Coefficiant	WGW	NGW	MSW	WGC	NGC	MSC
intercept	0.0005**	0.0287**	0.0287**	0.0374**	0.0062**	0.0062**
Red				0.0117**		
Blue		0.0028**	0.0028**			
Green	0.0072**			0.0045**	0.0054**	0.0054**
NIR	0.0073**			0.0161**	0.0048**	0.0048**
RedEdge		0.0004**	0.0004**			
NDVI				0.0074**		
GNDVI	0.0049**	0.0024**	0.0024**	0.0084**		
RENDVI		0.0034**	0.0034**		0.0423**	0.0423**
Moisture	0.0032**					
Temperature	0.0001**			0.0189**		
Pressure M	0.0000**					
Pressure W	0.0012**			0.0381**		
Temperature M						
Temperature W						
Moisture M						
Moisture W						

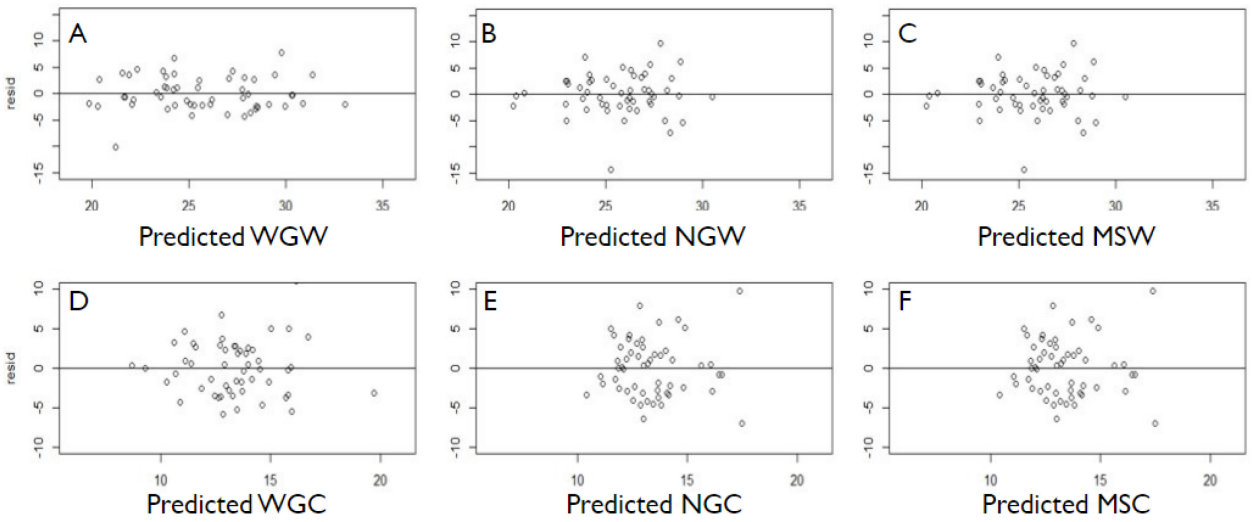


Figure 10. These are the residual plots for the July 2020-2022 model not including cover. Plots A, B, and C are predicting water potential in bars and plots D, E, and F are predicting chlorophyll. Plots A and D include ground data, B and E do not include ground data, C and F.

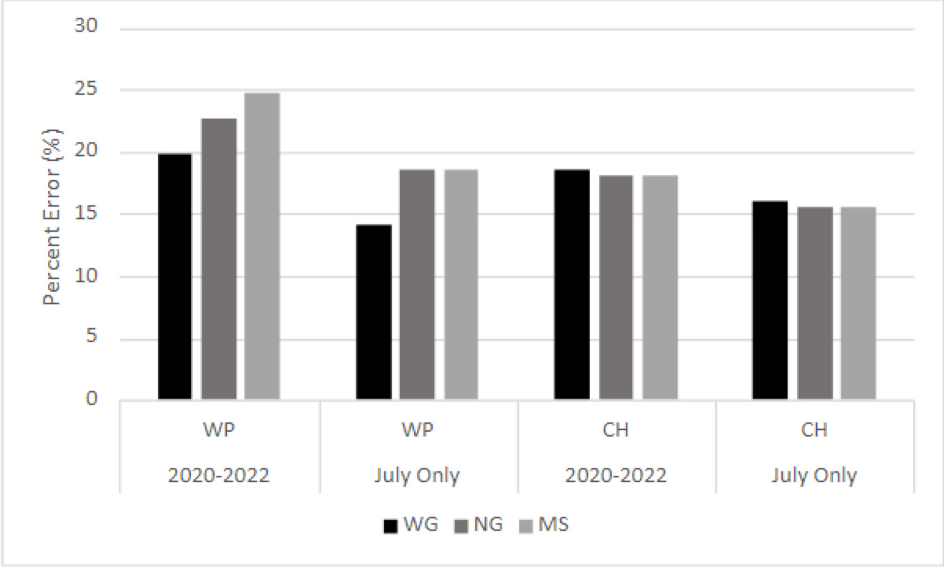


Figure 11. Cross Validation Error Percentage for all data sets from 2020-2022 and only July data sets from 2020-2022 for water potential (WP) and chlorophyll (CH) predictions. WG = all variables, NG = without ground measurements, MS = multispectral indices only.

July with Cover. The mixed regression model for the July data from 2020-2022 including cover had p-values that were all less than .05 as seen in Table 7. Percentage canopy cover was

added in these models to assess the influence on cover percentage on prediction capability of the models. Cover was eliminated during stepwise selection in all WP models, but kept in CH models. Water potential models remained the same as the July without cover models (Table 6). Cover was also used as a dependent variable and did have significant covariates as seen in Table 8. For the July model including cover, variables kept in the full models most often were NIR and GNDVI. In models including temperature and moisture, temperature was used in both the WGW and WGC models and moisture was used in only the water potential model.

The residuals for the models (Figure 12) for water potential remain the same as previously viewed residuals for no cover data (Figure 10). For chlorophyll models, residual plots had very little change with inclusion of cover. The residual points are slightly more spread out when cover is included. In Figure 13, cover prediction residuals appear fairly scattered with large margins. This shows that cover is not being accurately predicted in this model.

As previously mentioned, the water potential models between the July models including and excluding cover used the same variables and therefore have the same cross validation percent error (Figure 14). When looking at chlorophyll models, use of cover in the models did slightly help reduce error for this set of models. The same pattern does continue with ground data hindering the ability of chlorophyll to be predicted. As expected, cover prediction cross validation error in Figure 15 is extremely high.

Table 7. Mixed regression model p-values for the variables used to predict water potential and chlorophyll from the lmer() function of the LMER package in R Studio (Bates et al. 2015) using the July with cover data. Predictions models were created using ground data (WGW and WGC), excluding ground data (NGW and NGC) and multispectral data only (MSW and MSC). P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

	Water Potential			Chlorophyll		
Coefficiant	WGW	NGW	MSW	WGC	NGC	MSC
intercept	0.0005**	0.0286**	0.0286**	0.0362**	0.010**	0.0080**
Red				0.0531*		
Blue		0.0028**	0.0028**			
Green	0.0072**			0.00953**	0.0016**	
NIR	0.0073**			0.0150**	0.0014**	
RedEdge		0.0004**	0.0004**			
NDVI				0.0203**		
GNDVI	0.0049**	0.0024**	0.0024**	0.0205**		0.0028**
RENDVI		0.0034**	0.0034**		0.1061	0.09729*
Cover				0.0131**	0.0133**	0.0151**
Moisture	0.0005**					
Temperature	0.0032**			0.0124**		
Pressure M	0.0000**					
Pressure W	0.0012**			0.0363**		
Temperature M						
Temperature W						
Moisture M						
Moisture W						

Table 8. Mixed regression model p-values for the variables used to predict cover percentage from the lmer() function of the LMER package in R Studio (Bates et al. 2015) using the July 2020-2022 data. Predictions models were created using ground data (WG), excluding ground data (NG) and multispectral data only (MS). P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

	Cover		
Coefficiant	WG	NG	MS
intercept	0.0521*	0.0521*	0.3606
Red	0.0005**	0.0005**	0.0457**
Blue			
Green			0.0918*
NIR	0.0169**	0.0169**	
RedEdge			
NDVI			0.1210
GNDVI			0.0588*
RENDVI			
Moisture			
Temperature			
Pressure M	.0505**	.0505**	
Pressure W			
Temperature M			
Temperature W			
Moisture M			
Moisture W			

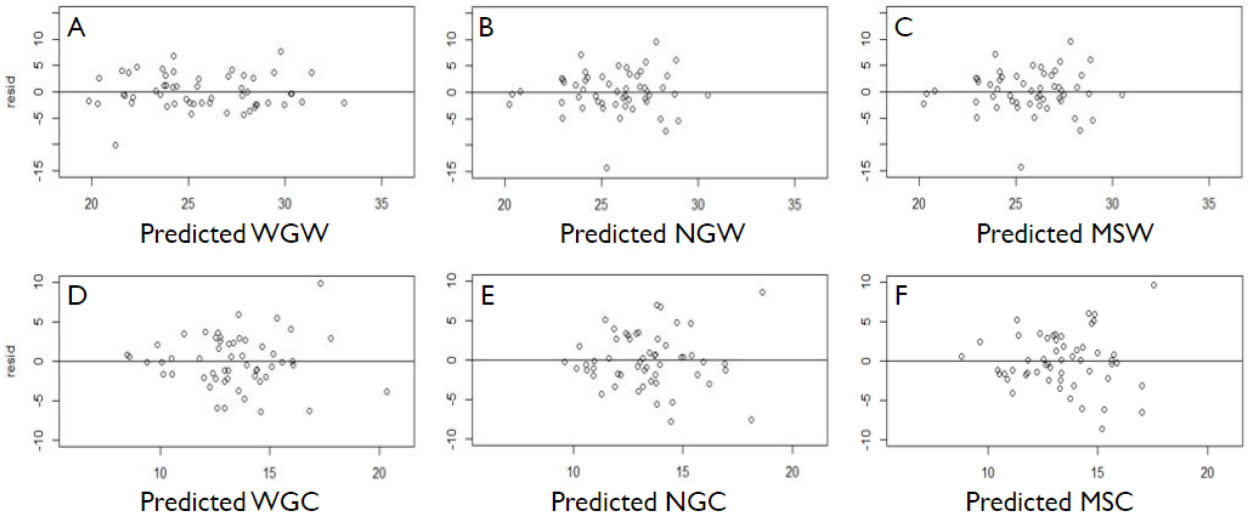


Figure 12. These are the residual plots for the July 2020-2022 model including cover. Plots A, B, and C are predicting water potential in bars and plots D, E, and F are predicting chlorophyll. Plots A and D include ground data, B and E do not include ground data, and C and F include only multispectral data.

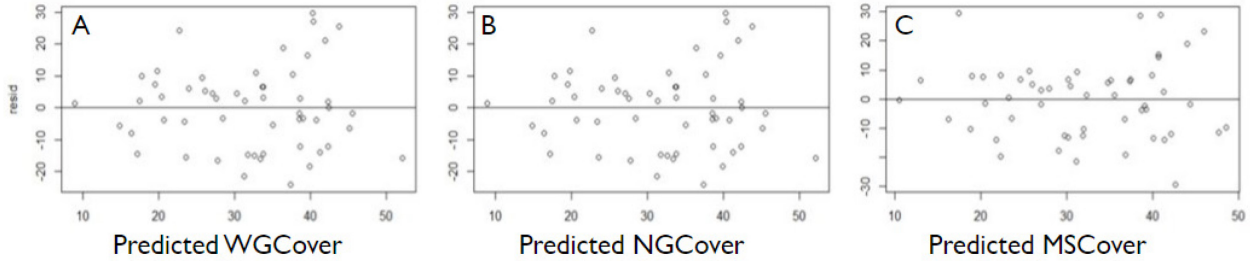


Figure 13. These are the residual plots for the July 2020-2022 model predicting cover. Plots A includes ground data, B does not include ground data, and C includes only multispectral data.

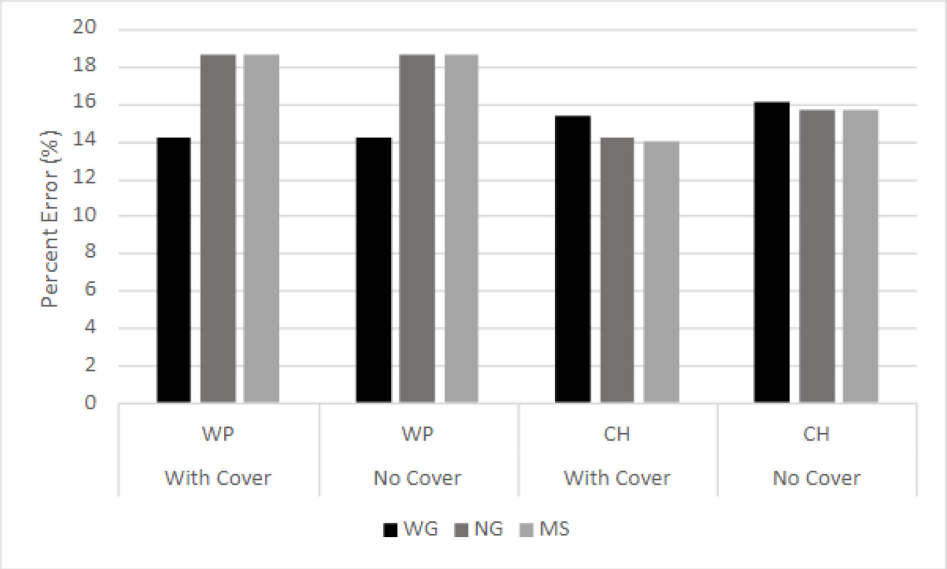


Figure 14. Cross Validation Error Percentage for July data sets from 2020-2022 with and without the inclusion of cover for water potential (WP) and chlorophyll (CH) predictions.. WG = all variables, NG = without ground measurements, MS = multispectral indices only.

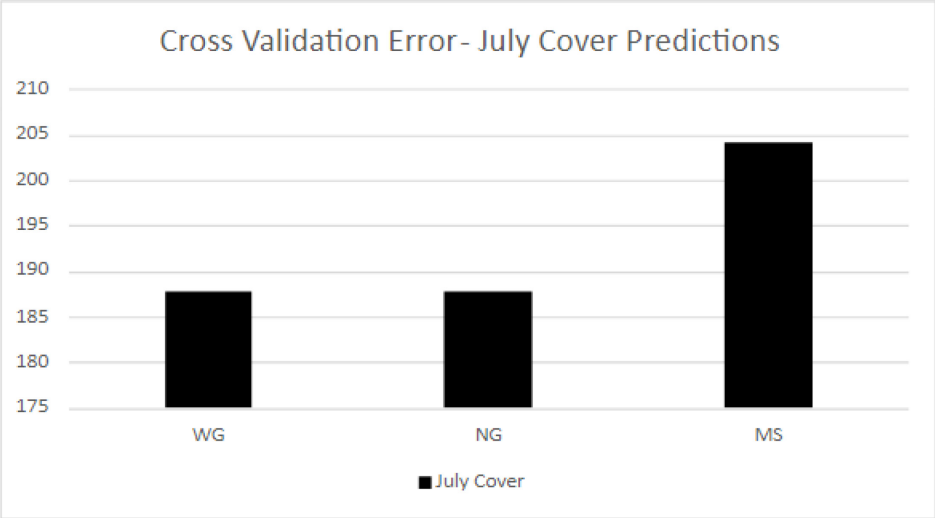


Figure 15. Cross Validation Error Percentage for cover predictions for July 2020-2022 data. WG = all variables, NG = without ground measurements, MS = multispectral indices only.

July Landsat Comparison. The July dataset was used for these models, excluding cover, red edge, and RENDVI to allow for a direct comparison with Landsat satellite models. P-values

for variables used to predict water potential were all significant. The chlorophyll models only using multispectral data did have less significance. For both predictions, WG models had higher significance as seen in Table 9. Variables kept in the equation after stepwise selection most often were NIR and GNDVI for the July Landsat comparison models, both being used in all models. Temperature was used every time that it was available, but moisture was only used for WP prediction modeling.

Table 9. Mixed regression model p-values for the variables used to predict cover percentage from the lmer() function of the LMER package in R Studio (Bates et al. 2015) using the July Landsat comparison data. Predictions models were created using ground data (WG), excluding ground data (NG) and multispectral data only (MS). P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

	Water Potential			Chlorophyll		
Coefficiant	WGW	NGW	MSW	WGC	NGC	MSC
intercept	0.0005**	0.2013	0.2013	0.0374**	0.7913	0.7913
Red				0.0117**	0.0586*	0.0586*
Blue		0.0358**	0.0358**			
Green	0.0072**			0.0045**	0.0278**	0.0278**
NIR	0.0073**	0.0052**	0.0052**	0.0161**	0.0885*	0.0885*
NDVI				0.0074**	0.0705*	0.0705*
GNDVI	0.0049**	0.0422**	0.0422**	0.0084**	0.0733*	0.0733*
Moisture	0.0032**					
Temperature	0.0001**			0.0189**		
Pressure M	0.0000**					
Pressure W	0.0012**			0.0381**		
Temperature M						
Temperature W						
Moisture M						
Moisture W						

The residual plots remained similar to previous July model residual plots. The WGW plot continues to be stretched further across the x axis showing a larger range of predicted values (Figure 16). Chlorophyll models all remain similar with the ground data causing a slight wider spread of residuals.

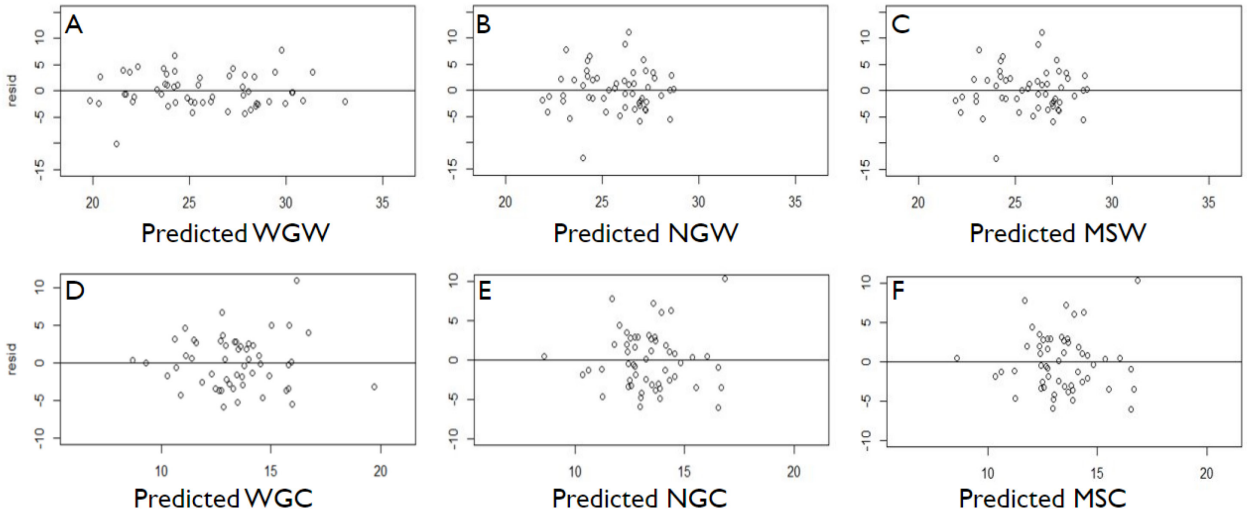


Figure 16. These are the residual plots for the July 2020-2022 model not including cover, red edge, or RENDVI. Plots A, B, and C are predicting water potential in bars and plots D, E, and F are predicting chlorophyll. Plots A and D include ground data, B and E do not include ground data, and C and F include only multispectral data.

LandSat. The variables used after stepwise selection to create predictions for the satellite captured Landsat data were very limited and showed little consistency in variable selection as seen in Table 10. Variables that were kept in the equation for water potential predictions showed high significance. The chlorophyll prediction model including ground (WGC) also showed high significance, but the multispectral model had only marginal significance. During stepwise selection for WP models, all Landsat satellite data was excluded when ground variables were available. In models excluding ground data, only NDVI showed significance in

predicting water potential. All three chlorophyll models were consistent in using blue and green bands.

Table 10. Mixed regression model p-values for the variables used to predict cover percentage from the lmer() function of the LMER package in R Studio (Bates et al. 2015) using the Landsat data. Predictions models were created using ground data (WG), excluding ground data (NG) and multispectral data only (MS). P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

	Water Potential			Chlorophyll		
Coefficiant	WGW	NGW	MSW	WGC	NGC	MSC
intercept	0.0002**	0.0000**	0.0000**	0.5139	0.4081	0.4081
Red				0.0161**		
Blue				0.0236**	0.0626*	0.0626*
Green				0.0036**	0.4660	0.4660
NIR				0.0163**		
NDVI		0.0054**	0.0054**			
GNDVI						
Moisture	0.0063**			0.0310**		
Temperature	0.0001**					
Pressure M	0.0003**					
Pressure W	0.0001**					
Temperature M						
Temperature W						
Moisture M						
Moisture W						

Although there were a limited number of variables used in prediction models for water potential and chlorophyll, the residual plots are relatively strong and show no issues within the data itself such as outliers or sways. When compared to July drone derived Landsat comparison

residual plots in Figure 17, there are very similar patterns, and the satellite residuals are tighter about the zero-line.

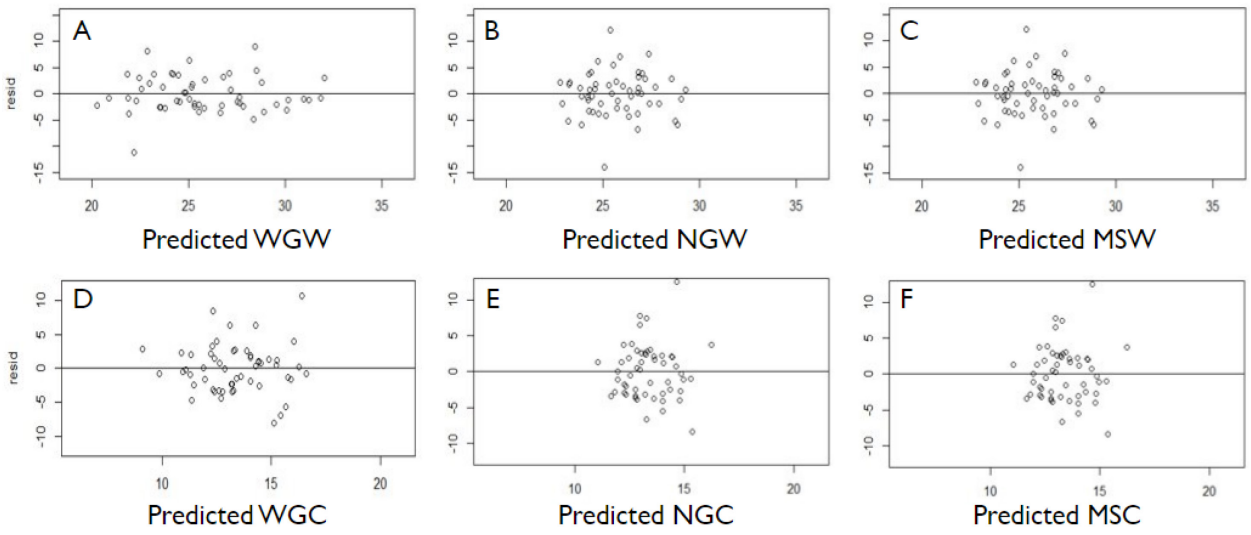


Figure 17. These are the residual plots for the Landsat 2020-2022 model. Plots A, B, and C are predicting water potential in bars and plots D, E, and F are predicting chlorophyll. Plots A and D include ground data, B and E do not include ground data, and C and F include only multispectral data.

The landowner prediction model created through Landsat images resulted in a very similar, strong model when compared to the July LS model with the same available variables, before stepwise selection (Figure 18). The Landsat model included less variables in model equations but had a slightly lower cross validation error for both CH and WP predictions. The decrease in error was most likely due to consistency of satellite imagery across all three growing seasons.

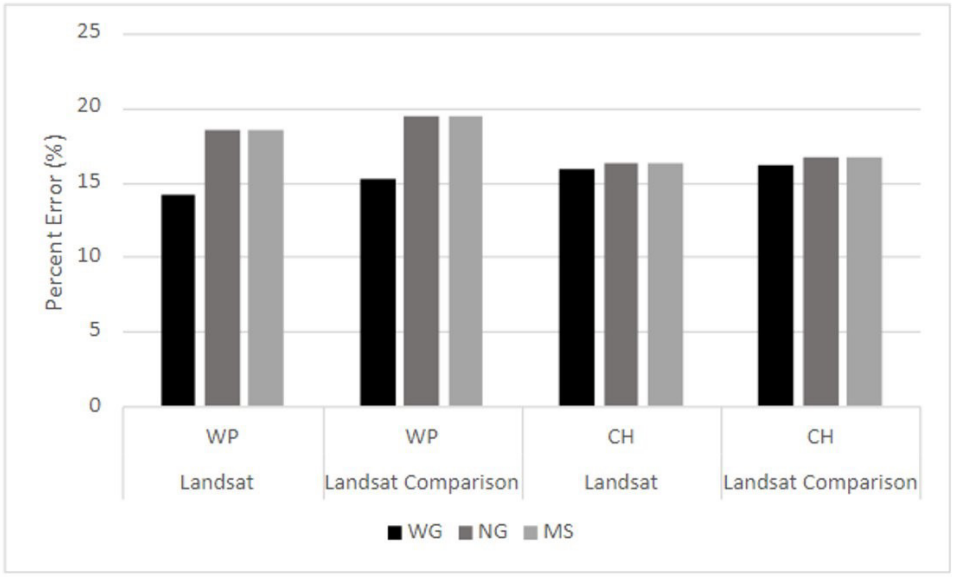


Figure 18. Cross Validation Error Percentage for drone derived July datasets from 2020-2022 and Landsat derived July datasets from 2020-2022 for water potential (WP) and chlorophyll (CH) predictions. WG = all variables, NG = without ground measurements, MS = multispectral indices Only.

Crown Cover Modeling

From year one to year two, 13 of the 18 plots experienced positive growth with an average of 9% growth. From year two to year three, 14 plots experienced an average of 23% growth. These changes across the stand can be seen in Figure 19. With an average yearly growth of 16%, an equation was formed to predict the years till 60% and 40% canopy cover within each plot for each year. The average predicted time for each plot to get to 60% cover from their year one canopy cover percentage was 9.5 years, year two was 9.2 years, and year three was 7.5 years. With an average of 16% growth for each year, it was expected that it would take an average of 7 years from initial thin to reach 60% cover based on Figure 12. For 40% cover, the same process was done. Average predicted time from the first year of data to 40% cover was 5.4 years, the second year of data was 5.2 years, and the third year was 4.7

years, with an average of five years from initial thin to 40%. Growth predictions based on 16% yearly growth can be seen on the graph in Figure 20.

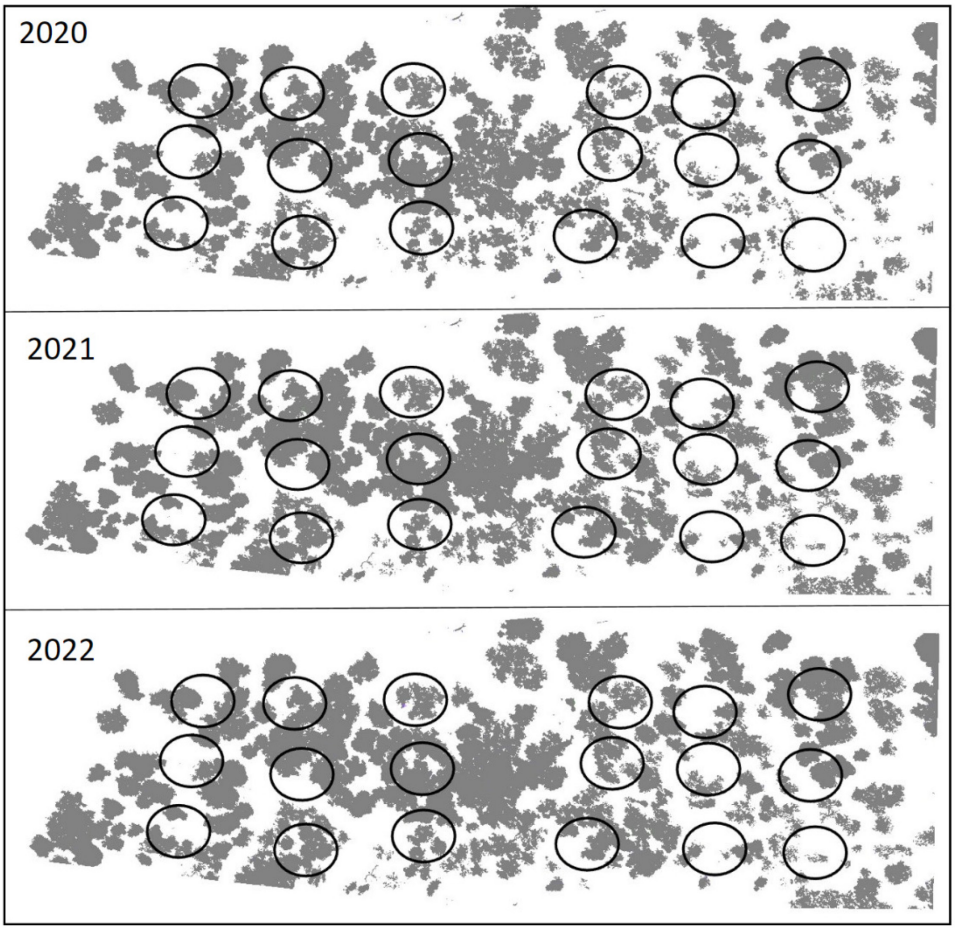


Figure 19. Changes in canopy cover from year one to year three across all 18 plots in the converted stand.

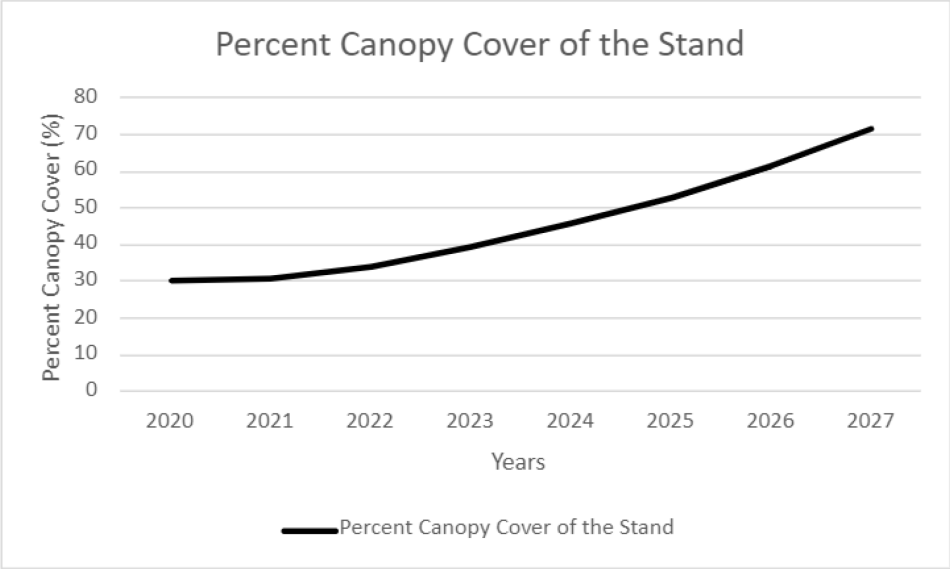


Figure 20. Average canopy cover prediction for the stand from 2020 to 2027 for an average growth of 16% per year.

DISCUSSION

Climate Stress Models

2020. The 2020 model was calculated using only July and September data for the year. Upon viewing, the August multispectral data was corrupted. After recalculation, it was determined that errors were likely due to a defect in images for the August collection date. As expected, accuracy of models using only July and September of 2020 was low when compared to other models.

Looking specifically at results of the July and September multiple regression models, it could be inferred that water potential was not properly captured for the month of September. Limitations within our data collecting tools could be the reason for lack of significant variables and low adjusted-r^2. Leading up to the September flight, the stand experienced drought conditions. It is possible that water potential exceeded what was able to be detected with the equipment used in this study. The CH model, however, did show more significance within the individual variables and a higher adjusted-r^2.

For the 2020 mixed regression water potential models, covariate significance improved when ground data was added into the model (Table 4). This indicates that with only two months of data, soil moisture and temperature are required to produce useful prediction models. Chlorophyll prediction models did continue to show significance in covariates even when ground data was not present, showing that the relationship between chlorophyll and multispectral bands such as NIR and RedEdge, have a strong relationship.

Even with predictively weak water potential model for September, residual plots for these models did show predictability potential (Figure 7). While it is not tight in relation to residuals for other models, random scatter and even points above and below the zero-line for both water potential and chlorophyll residual plots give indication that with more data, strength could be improved. The WGW and MSC plots that were stronger models for p-values did prove to have the most linearity for residuals as well. While p-values and residuals for only one year of data did show significance and ability to predict, when looking at Appendix C, these models did have the highest percentage error across all models for this study.

2020-2022. The full 2020-22 model was developed using all six of the collection dates from July of 2020 to July of 2022, excluding the 7th date, August of 2020 due to errors in values as previously stated.

For mixed regression models, stepwise selection was performed to identify which variables would create the strongest prediction model. Multispectral variables included in all models, red band, NIR, and NDVI, reflect what would be expected to have a relationship with water potential and chlorophyll. When given the option to include soil moisture for chlorophyll prediction, it was removed during stepwise selection. This indicated that soil moisture and chlorophyll did not have a strong enough relationship to create predictions.

When looking at p-values for 2020-2022 models (Table 5), most values for individual covariates were highly significant. For water potential models, p-values remained significant when ground data was included. MSW values showed less significance in the model, but when compared to 2020 water potential models, there was an improvement in covariate significance without the need for soil moisture and temperature. Soil moisture was consistently removed

during stepwise selection for all chlorophyll prediction models. This indicates that while soil moisture and chlorophyll do have a relationship in the process of photosynthesis, the relationship is not strong enough to use in predictions.

When viewing residual plots, all three CH models are similar and remain evenly distributed about the zero-line indicating no bias issues (Figure 8). It can also be assumed that there is homogeneous error as there is no significant increase to the spread of residuals with an increase or decrease of prediction values. The observed grouping of residuals is expected, as CH is a variable with a narrow range of recorded values. WP models show less grouping and more linearity of residuals, with the strongest being models excluding ground data. WGW had the widest range of predictions but remains close to the zero-line indicating strong predictions.

From the cross-validation error figure for 2020 and 2020-2022 data (Figure 9), it is obvious that inclusion of more years of data reduce error for water potential predictions. When predicting chlorophyll, based on this Figure, error decreases with inclusion of ground data, while water potential prediction error decreases with exclusion of ground data.

July No Cover. The July no cover model was created to look at a direct comparison between using only one month of data for every year, versus multiple months of data for mixed regression models.

After stepwise selection, many variables were included in the model (Table 6). For most models, green and NIR bands, along with MS indices GNDVI and RENDVI were included in the predictions. This is continues to show the relationship of light reflectance and absorbance in water potential and chlorophyll prediction. The strong relationship for stress identification in

trees with these indices have been found in other studies looking at drought response and chlorophyll content (Raddi et al., 2021).

When looking at the significance of the covariates of mixed models, p-values for WP, the MSW model had the highest percentage of significant variables. The prediction equation using all variables for CH (WGC) had the most significant correlation of covariates for chlorophyll models. When comparing the full 2020-2022 models and July only models, p-values for covariates improved using only one date per year. This is likely due to reducing seasonal stress changes over the growing period by focusing on one specific month rather than a whole growing season.

Looking at residual plots in Figure 10, the WGW model has the strongest plot for predicted water potential with no obvious grouping or swaying of residuals and shows the ability to predict a wider range with added ground data while keeping error the same across the points. The CH models do have slight grouping of residuals, regardless of the inclusion of ground data, and prediction potential remains consistent throughout the three model types.

The cross-validation figure, comparing the July only and 2020-2022 models show an increase in error percentage with an increase in collection dates for one year (Figure 11). From this figure, it can be inferred that while ground data inclusion helps with water potential prediction, accuracy of the prediction model is still higher with just one day of multispectral data a year. For chlorophyll predictions of these models, percent error was lowest using only multispectral data for the July only models. From this study, it can be inferred that just one day of multispectral data for both water potential and chlorophyll will give strong predictions.

July With Cover. The July 2020-22 models including cover were used to show effects on the model when cover is available as a predictor variable and how well multispectral and ground data can predict cover. Only July collection dates were included for the model due to inability to see changes in canopy cover percentages from month to month both due to noise and lack of visible change within the growing season. Leaves moving in the canopy during drone data collection create noise in the image, which was later removed through processing steps. Therefore, different growing season changes are more likely to have a higher degree of accuracy and change. Even with a year's difference, there are many factors that can affect the percentage that we see, such as minute changes in height the drone is flew, movements in leaves, and general noise in the images.

For this data set, cover was included as a potential variable for every model prior to stepwise selection. After stepwise selection, it was only kept in models predicting chlorophyll (Table 7). Since cover was not used for water potential, water potential will be left out of this discussion section. The chlorophyll cover relationship is likely because more leaves in the canopy would indicate more chlorophyll present. For these models, GNDVI was included in almost every model. For the cover prediction models created with this dataset, red band and NIR were used most often (Table 8).

When looking at p-values for chlorophyll, and cover, almost all covariates included in the models showed significance, including cover. When looking at residual plots in Figure 12 the CH plots do have grouping similar to other models, but show very little bias. This inclusion of cover or ground data did not change overall shape or prediction potential according to residual plots. Residuals for cover prediction in Figure 13 all have wide margins and no

grouping. This is likely due to variation in percentages of canopy cover in each plot and would require more years of data to strengthen the model. Based on these plots, there was no change with inclusion of ground data.

The cross-validation error percentages for chlorophyll predictions slightly increase with inclusion of cover data (Figure 14). This could have been caused by variation in percent cover from plot-to-plot related to factors outside of this study such as death in the canopy, size of the trees, and presence of trees. This model may be more useful in an even-aged, pure plantation stand that would have less variation from plot to plot. Cover prediction cross validation was very high and accounts for the many outside factors mentioned above (Figure 15).

July Landsat Comparison. The July Landsat Comparison models were created for a direct comparison with publicly accessible Landsat and weather station data. This dataset is a duplicate of the July no cover dataset but excludes RedEdge and RENDVI from the model prior to stepwise selection.

The multispectral variables included in the model after stepwise selection remained the same throughout all three chlorophyll models and included red and green bands, NIR, NDVI, and GNDVI (Table 9). The WGC model also included temperature and one weather station variable. Water potential models all included NIR and GNDVI in the models. These are similar to all previous models.

The p-values for water potential models were all significant and chlorophyll models were all at least marginally significant. All water potential models seem to be even in their strength based on number of significant variables in relation to available variables prior to

stepwise selection. The WGC model, however, is clearly strongest for chlorophyll predictions based on individual covariates.

July's Landsat comparison residual plots (Figure 16) for water potential using all three model types appear evenly distributed about the zero-line showing absence of bias and linearity issues within the model. Inclusion of ground variables for water potential (WGW) created a more linear pattern while maintaining the same assumption of constant error variance. This indicates that ground data allows for a wider range of predictions than with only multispectral data. This would likely result in slightly better predictions. For chlorophyll plots, general grouping seen in all previous CH models can be seen and residual points are even on top and bottom indicating a strong prediction model.

Landsat. Landowner models included data only for July from 2020-2022. Dates were selected to be as close to the actual drone collected dates as possible. Data extracted from Landsat was used and compared to drone collected data. Images went through the same processing as drone collected data in ArcGIS. Weather information was derived from the weather station in Mountain Grove that previous datasets have included. Ground measurements, temperature and moisture did come from the drone datasets to use as a control for the model in the all variables models. Model training predictor variables, WP and CH, were used from the July months datasets.

From a side-by-side comparison for individual multispectral indices, Landsat data did have consistently higher values. Similarity of the values resulted in similar prediction models to the drone collected data.

During stepwise selection process, many of the variables were excluded from the final models. Chlorophyll prediction models continually kept the blue band and green band in the models, and water potential models excluded all multispectral data for the WGW model and only kept NDVI for the NGW and MSW models.

While this could mean that data from the Landsat satellite covariates were not strongly related to the chlorophyll and water potential variables, the high number of significant variables, of variables left in the models, indicate otherwise (Table 10). When looking only at p-values it could be inferred that NG and MS only models for both water potential and chlorophyll would be unable to make predictions based on the lack of significant variables in the models.

When looking at residuals (Figure 17) only, it appears to be normal and very similar to the previous residual models including ground data being able to extend the range of predictions for water potential and chlorophyll being relatively tightly grouped. Both water potential and chlorophyll residual points are closer together than previously mentioned residual models (Figure 16) of the comparison model indicating less variability in predicted values.

The cross-validation percentage for the Landsat and Landsat comparison model (Figure 18) also helps to re-enforce the idea that lack of variables can still create strong models for prediction. This figure shows that percentage error for cross-validation for both drone derived models and Landsat models are very similar. Landsat cross validation error percentages are slightly lower for all of the models. While this could be due solely to consistency of the satellite, this could also be due to inability for Landsat to capture at the same level of precision as the

drone, causing each plot to have less variability in multispectral bands and indices compared to the drone model. There have been other studies conducted using Landsat-based multispectral imagery to extract bands, looking at the ability to quickly assess plant health with satellite data that have concluded that this is one step in the right direction, but there is still room for improvements and advancements (Gillani et al., 2023).

Model Type Comparison

For our three model types, with ground, no ground, and multispectral only, the with ground variable predictions for WP (WGW) were slightly stronger than no ground (NGW) and multispectral only (MSW) for most models. The CH predictions decreased in model strength with the inclusion of ground data.. Using only multispectral variables will give relatively low error predictions for both predictions of WP and chlorophyll. Previous studies have been successful in predicting water stress in orchards using only multispectral imaging and looking at narrow bands using a photochemical reflective index (Suarez et al., 2011). By evaluating p-values for all models (Appendices C-1 through C-6) and cross-validation error for all models (Appendix D-1) it can be determined that using multispectral variables only will result in relatively strong models with multiple years of data. This will result in faster data collection and less field days.

Thinning Prediction

The canopy heights were only compared for July dates due to inability to see measurable month to month changes. In post processing of the images, noise was deleted

from the images. Noise commonly includes leaves in the canopy if they move during the drone flight. Even with mild filtering, allowing for noise, and low wind days, many of these points are deleted reducing the number of useable data points. While canopy was not as full in the images as it was in real life, growth predictions can still be done when comparing images that all went through the same filtering processes.

Growth predictions were created by a simple model of average percent change by plot from year to year. Some plots did have negative growth, or a decrease in canopy cover. Negative values could be due to changing conditions during the flight and noise in the images. For some plots, there were verifiable decreases due to death, such as ash trees in plots affected by the Emerald Ash Borer (*Agrilus planipennis*) and some trees injured during the thinning operation in 2020. For the purposes of the model, negative percent change values were excluded.

Model predictions indicate that the crown cover of the stand will increase by an average of 16% per year. This means that when the stand reaches a minimum of five years and a maximum of seven years after the initial thin, a thinning operation needs to take place to allow for 40% to 60% light into the understory (Figure 20). This coincides directly with the USDA Forest Services Working Trees Silvopasture brochure which recommends thinning every 5-7 years for cool season grasses (USDA, 2008). This will facilitate continued growth of cool season grasses in the understory which will maintain rich forage for cattle to graze in the silvopasture system. This calculation is based on average growth of the stand and does not account for death or decreases in the plots. This calculation also does not account for other factors contributing to rate of canopy closure. Slowed expansion due to decrease of space in the

canopy will likely not be a factor for most plots if percent cover remains under 60%. The trees should not reach a point where they have limited room for canopy growth.

CONCLUSION

Multispectral and remote sensing data can be used to create prediction models for both stress and growth of the canopy in forested stands. Stress in trees is best seen at wavelengths human eyes cannot see, making it difficult to detect health of a tree. Multispectral imaging allows for manipulation of images to extract band level data that can then be used to formulate calculations of stress based on measurable responses of vegetation such water potential and chlorophyll. This allows for prediction of those responses based on imaging, which allows for plot and stand level data for large areas much quicker than taking water potential and chlorophyll measurements of every tree in a forest. Looking at changes in canopy cover over time in response to thinning operations also allows for a quicker estimation of when a canopy may need to be thinned. In a silvopasture, this is important in maintaining growth of cool season grasses throughout the entire growing season. Waiting too long to thin a stand will result in a decrease of light into the understory and forage will not be able to grow.

This study shows that multispectral data is able to predict both chlorophyll and water potential. When predicting water potential, model error decreases with inclusion of soil moisture and temperature. The error of the models also decreases with inclusion of multiple years of data. It was found that error is lowest when only looking at one month for every year versus one to three months for every year. This means that rather than going into a stand multiple times a year to collect ground data, it is better to pick one day a year and collect that same day every year. When cover was added there was no significant impact on the model's ability to predict chlorophyll or water potential. While ground data does help increase model

strength for water potential, multispectral-only models had very similar results based on residual plots and cross validation. With more years of data, multispectral only data will likely increase in strength for water potential prediction models. When predicting chlorophyll, based on this study, using only multispectral data is sufficient on its own and decreases with inclusion of ground metrics. Landsat satellite data proved to have low error percentages when predicting both chlorophyll and water potential and was very similar to the drone data, however, due to the limited variables that were used in the model and lack of precision due to cell size of Landsat images, it cannot be concluded that Landsat is more sufficient in making predictions. Landsat likely would not be able to detect small changes due to lack of variation in the dataset. More years of data would be needed to make further conclusions on the use of Landsat data in comparison to drone data.

For canopy cover predictions in this study, there were many factors that went into percent change both for yearly changes and when comparing plots. Some plots experienced mortality and were not able to be used to predict overall growth of the stand as a whole. Average growth among the plots that did not decrease was 16% per year. In order to maintain a silvopasture, the stand needs to be kept between 40-60% cover to ensure growth of cool season forage. For the converted stand at Journagan Ranch, which was thinned to 30% in 2020, that will likely be around 5-7 years post thin. Maintaining this stand will lead to an increase in nutrient uptake for cattle, or grazing livestock, which increases weight gain. Shaded canopy will also serve as a shelter from heat, helping maintain weight gain in cattle through the hottest summer months.

REFERENCES

Abid, M., Ali, S., Qi, L. K., Zahoor, R., Tian, Z., Jiang, D., Snider, J. L., & Dai, T., 2018. Physiological and biochemical changes during drought and recovery periods at tillering and jointing stages in wheat (Triticum aestivum L.). Scientific reports, 8(1), 4615. https://doi.org/10.1038/s41598-018-21441-7

Bates, D., Mächler, M., Bolker, B., & Walker, S., 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Carreiras, J. M. B., Pereira, J. M. C., & Pereira, J. S., 2006. Estimation of tree canopy cover in Evergreen Oak Woodlands using remote sensing. *Forest Ecology and Management*, *223*(1-3), 45–53. https://doi.org/10.1016/j.foreco.2005.10.056

De Frenne, P., & Verheyen, K., 2016. Weather stations lack forest data. *Science*, *351*(6270), 234. http://www.jstor.org/sTable/24741426

Evangelides, Christos & Nobajas, Alexandre., 2020. Red-Edge Normalized Difference Vegetation Index (NDVI705) from Sentinel-2 imagery to assess post-fire regeneration. Remote Sensing Applications Society and Environment. 17. 100283. 10.1016/j.rsase.2019.100283.

Gillani, S. S., Tahir, M. N., Anwar, A., Haq, S. I., Awais, M., Iqbal, M., Iqbal, J., Malik, H. A., Naqvi, S. M., Ullah, R., & Khan, M. A., 2023. Real time estimation of wheat chlorophyll content retrieve from landsat 8 imagery under rainfed condition. *Sarhad Journal of Agriculture*, *39*(1). https://doi.org/10.17582/journal.sja/2023.39.1.147.155

Gingrich SF, 1967. Measuring and evaluating stocking and stand density in upland hardwood forests in the central states. Forest Science 13:38-52

Davison AC, Hinkley DV., 1997. *Bootstrap Methods and Their Applications*. Cambridge University Press, Cambridge. ISBN 0-521-57391-2, http://statwww.epfl.ch/davison/BMA/.

Horning, N., 2004. Selecting the appropriate band combination for an RGB image using Landsat imagery Version 1.0. American Museum of Natural History, Center for Biodiversity and Conservation. Available from http://biodiversityinformatics.amnh.org. (accessed on date)

Iizuka, K., Kato, T., Silsigia, S., Soufiningrum, A. Y., & Kozan, O., 2019. Estimating and examining the sensitivity of different vegetation indices to fractions of vegetation cover at different scaling grids for early stage acacia plantation forests using a fixed-wing UAS. *Remote Sensing*, *11*(15), 1816. https://doi.org/10.3390/rs11151816

Klopfenstein NB, Rietveld WJ, Carman RC, Clason TR., 1997. Silvopasture: An Agroforestry Practice. Agroforestry notes (USDA-NAC)

Lausch, A., Erasmi, S., King, D., Magdon, P., & Heurich, M., 2016. Understanding Forest Health with remote sensing -part I—a review of spectral traits, processes and remote-sensing characteristics. *Remote Sensing*, *8*(12), 1029. https://doi.org/10.3390/rs8121029

Loveland, T. R., & Dwyer, J. L., 2012. Landsat: Building a strong future. *Remote Sensing of Environment*, *122*, 22–29. https://doi.org/10.1016/j.rse.2011.09.022

Lulla, K., Nellis, M. D., Rundquist, B., Srivastava, P. K., & Szabo, S., 2021. Mission to earth: Landsat 9 will continue to view the world. *Geocarto International*, *36*(20), 2261–2263. https://doi.org/10.1080/10106049.2021.1991634

Lopez-Sanchez, A., Dirzo, R., & Roig, S., 2017. Changes in livestock footprint and tree layer coverage in Mediterranean Dehesas: A Six-decade study based on Remote Sensing. *International Journal of Remote Sensing*, *39*(14), 4727–4743. https://doi.org/10.1080/01431161.2017.1365391

McCollum, S. M. C. J., 2021. *Modeling the Growth and Establishment of Plantation and Converted Silvopasture Systems in the Missouri Ozarks Region* (thesis). Missouri State University

Micasense, 2018. Why Narrow Bands Matter. In: Micasense. https://micasense.com/whynarrow-bands-matter/ Accessed 12 April 2021

Marhuenda, Y., Morales, D. and Pardo, M.C., 2014. Information criteria for Fay-Herriot model selection. Computational Statistics and Data Analysis 70, 268-280

National Aeronautics and Space Administration, 2020. *Landsat 9 Mission Brochure*.

Raddi, S., Giannetti, F., Martini, S., Farinella, F., Chirici, G., Tani, A., Maltoni, A., & Mariotti, B., 2021. Monitoring drought response and chlorophyll content in Quercus by consumer-grade, near-infrared (NIR) camera: A comparison with reflectance spectroscopy. *New Forests*, *53*(2), 241–265. https://doi.org/10.1007/s11056-021-09848-z

Soares, C., Príncipe, A., Köbel, M., Nunes, A., Branquinho, C., & Pinho, P., 2018. Tracking tree canopy cover changes in space and time in high nature value farmland to prioritize reforestation efforts. *International Journal of Remote Sensing*, *39*(14), 4714–4726. https://doi.org/10.1080/01431161.2018.1475777

Suarez, L., Zarco-Tejada, P. J., Berni, J. A. J., González-Dugo, V., & Fereres, E., 2011. Orchard Water Stress Detection using high-resolution imagery. *Acta Horticulturae*, (922), 35–39. https://doi.org/10.17660/actahortic.2011.922.3

Sykas, D., 2020. *Spectral Indices with Multispectral Satellite Data*. GIS and Earth Observation University. Retrieved April 12, 2023, from https://www.geo.university/pages/blog?p=spectral-indices-with-multispectral-satellite-data

United States Department of Agriculture National Agroforestry Center, 2008. *Silvopasture an Agroforestry Practice*. Retrieved 2023, from https://www.fs.usda.gov/nac/assets/documents/workingtrees/brochures/wts.pdf.

United States Geological Survey, 2017. Unmanned Aircraft Systems Data Post-Processing Structure-from-Motion Photogrammetry. USGS National UAS Project Office

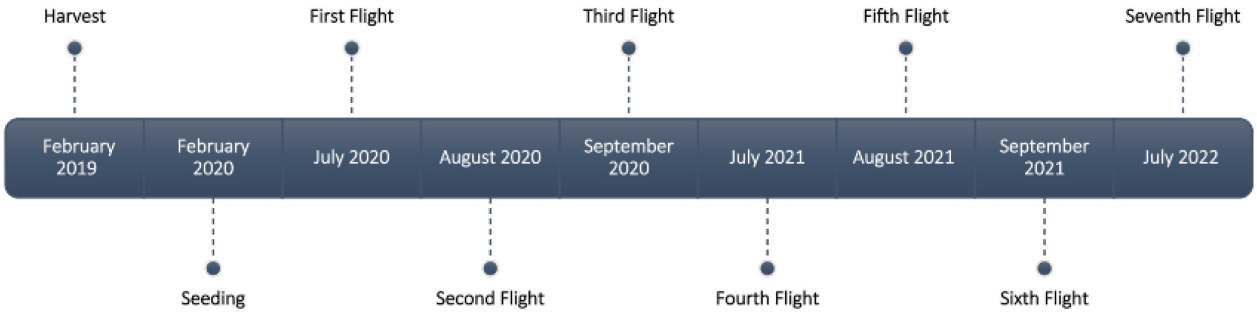
Ustin, S. L., Kasischke, E. S., Goetz, S., Hansen, M. C., Ozdogan, M., Rogan, J., Ustin, S. L., & Woodcock, C. E., 1998. Chapter 4: Temperate and Boreal Forests. In *Remote Sensing for Natural Resource Management and Environmental Monitoring* (3rd ed., Vol. 4, pp. 147–235). essay, John Wiley & Sons.

Zach., 2019. *How to perform cross validation for model performance in R*. Statology. Retrieved April 13, 2023, from https://www.statology.org/how-to-perform-cross-validation-for-model-performance-in-r/

Zhou, X., Zhang, J., Chen, D., Huang, Y., Kong, W., Yuan, L., Ye, H., & Huang, W., 2020. Assessment of leaf chlorophyll content models for winter wheat using landsat-8 multispectral remote sensing data. *Remote Sensing*, *12*(16), 2574. https://doi.org/10.3390/rs12162574

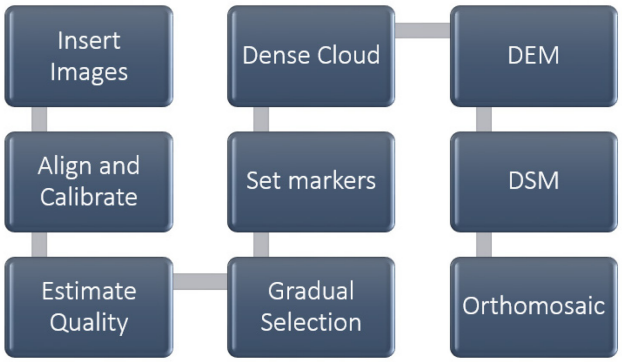
APPENDICES

Appendix A: Journagan ranch converted stand timeline

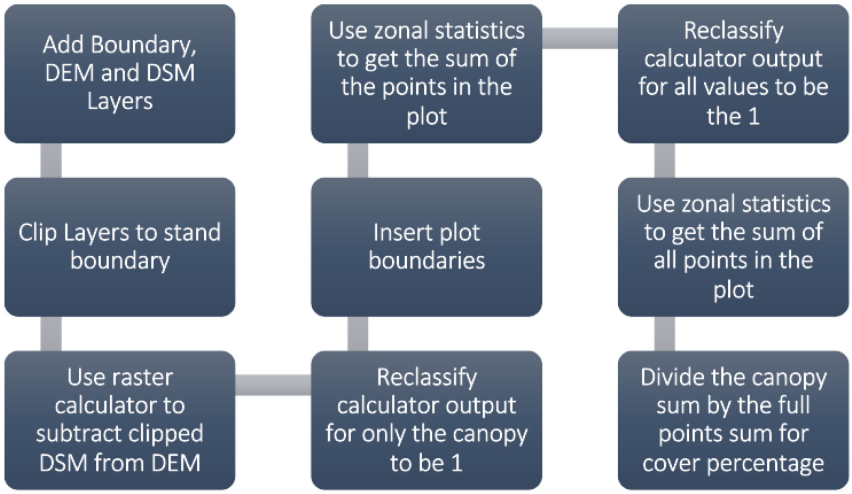


Appendix B: Processing Workflows

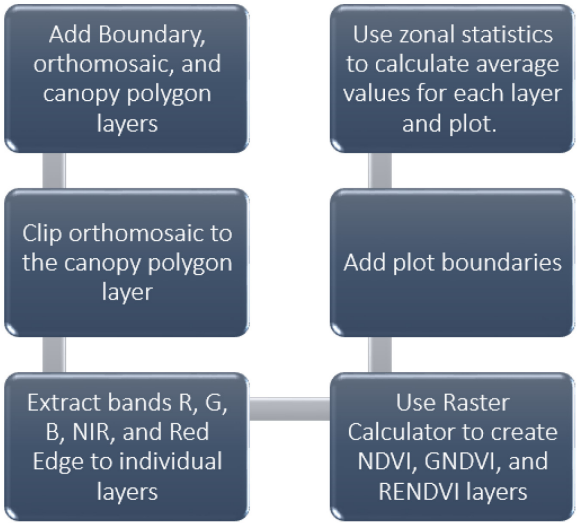
Appendix B-1: Metashape processing workflow for RGB and multispectral images.



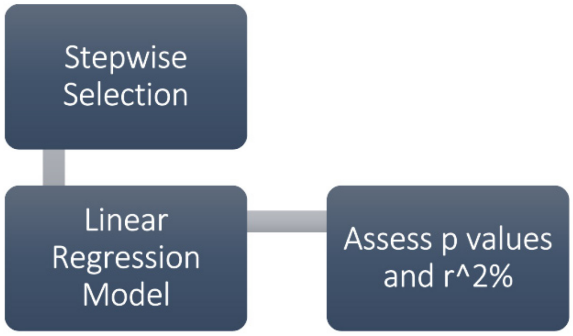
Appendix B-2: ArcGIS processing workflow for RGB Images to create canopy height models (CHM).



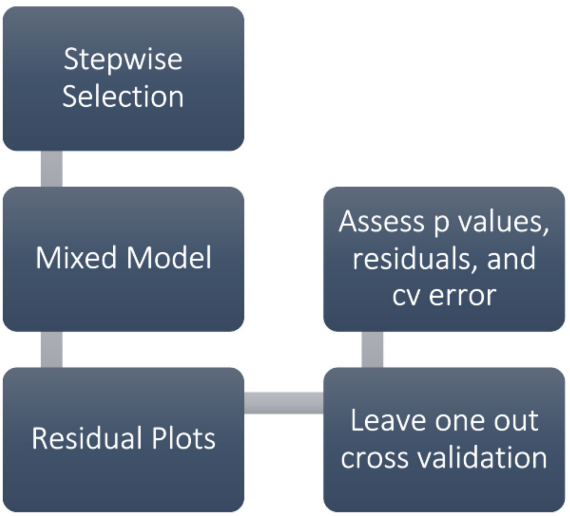
Appendix B-3: ArcGIS processing workflow for multispectral images to extract band specific data and create multispectral indices.



Appendix B-4: R Studio workflow to run month-specific multiple linear regression.



Appendix B-5: R Studio workflow to run mixed regression models on the full datasets.



Appendix C: P-values for all models

Appendix C-1: P-values for the Water Potential Predictions of all Datasets with all Available Variables Included. Some data sets have variable exclusions prior to the stepwise selection that can be seen in Table 3. P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

Coefficiant	2020	2020-2022	No Cover	With Cover	Landsat Comparison	Landsat
intercept	0.7140	0.0192**	0.0005**	0.0005**	0.0005**	0.0002**
Red		0.0049**				
Blue						
Green			0.0072**	0.0072**	0.0072**	
NIR	0.0530*	0.0085**	0.0073**	0.0073**	0.0073**	
RedEdge	0.0509*					
NDVI	0.0447**	0.0267**				
GNDVI	0.1207		0.0049*	0.0049*	0.0049*	
RENDVI						
Cover						
Moisture	0.0005**	0.0000**	0.0032*	0.0032*	0.0032*	0.0063**
Temperature	0.0014**	0.0000**	0.0001*	0.0001*	0.0001*	0.0001**
Pressure M			0.0000**	0.0000**	0.0000**	0.0003**
Pressure W			0.0012*	0.0012*	0.0012*	0.0001**
Temperature M						
Temperature W						
Moisture M		0.0000**				
Moisture W						

Appendix C-2: P-values for the Chlorophyll Predictions of all Datasets with all Available Variables Included. Some data sets have variable exclusions prior to the stepwise selection that can be seen in Table 3. P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

Coefficiant	2020	2020-2022	No Cover	With Cover	Landsat Comparison	Landsat
intercept	0.0563*	0.0741*	0.0374**	0.0362**	0.0374**	0.5139
Red		0.0000**	0.0117**	0.0531*	0.0117**	0.0161**
Blue		0.0188**				0.0236**
Green		0.0000**	0.0045**	0.0095**	0.0045**	0.0036**
NIR	0.0011**	0.0000**	0.0161**	0.015**	0.0161**	0.0163**
RedEdge	0.0021**					
NDVI	0.1601	0.0001**	0.0074**	0.0203**	0.0074**	
GNDVI		0.0000**	0.0084**	0.0205**	0.0084**	
RENDVI						
Cover				0.0131**		
Moisture	0.0022**					0.0301**
Temperature	0.1474	0.0159**	0.0189**	0.0124**	0.0189**	
Pressure M	0.0606*					
Pressure W		0.0396**	0.0381**	0.0363**	0.0381**	
Temperature M						
Temperature W						
Moisture M						
Moisture W						

Appendix C-3: P-values for the Water Potential Predictions of all Datasets with Ground Data Excluded. Some data sets have variable exclusions prior to the stepwise selection that can be seen in Table 3. P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

Coefficiant	2020	2020-2022	No Cover	With Cover	Landsat Comparison	Landsat
intercept	0.7443	0.0006**	0.0286**	0.0286**	0.2013	0.0000**
Red	0.1137	0.0053**				
Blue			0.0028**	0.0028**	0.0358**	
Green	0.0561*					
NIR		0.0106**			0.0052**	
RedEdge			0.0004**	0.0004**		
NDVI	0.1060	0.0430**				0.0054**
GNDVI	0.0546*		0.0024**	0.0024**	0.0422**	
RENDVI			0.0034**	0.0034**		
Cover						
Pressure M		0.0001**				
Pressure W		0.0331**				
Temperature M						
Temperature W						
Moisture M		0.0258**				
Moisture W						

Appendix C-4: P-values for the Chlorophyll Predictions of all Datasets with Ground Data Excluded. Some data sets have variable exclusions prior to the stepwise selection that can be seen in Table 3. P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

Coefficiant	2020	2020-2022	No Cover	With Cover	Landsat Comparison	Landsat
intercept	0.0000**	0.0051**	0.0062**	0.0102**	0.7913	0.4081
Red		0.0004**			0.0586*	
Blue	0.1073	0.0732*				0.0626*
Green		0.0000**	0.0054**	0.0016**	0.0278**	0.4660
NIR	0.0003**	0.0000**	0.0048**	0.0014**	0.0885*	
RedEdge	0.0005**					
NDVI	0.1583	0.0010**			0.0705*	
GNDVI		0.0000**			0.0733*	
RENDVI	0.0005**		0.0423**	0.1061		
Cover				0.0133**		
Pressure M						
Pressure W						
Temperature M						
Temperature W						
Moisture M						
Moisture W						

Appendix C-5: P-values for the Water Potential Predictions of all Datasets with Only Multispectral Data Included. Some data sets have variable exclusions prior to the stepwise selection that can be seen in Table 3. P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

Coefficiant	2020	2020-2022	No Cover	With Cover	Landsat Comparison	Landsat
intercept	0.0001**	0.1375	0.0286**	0.0286**	0.2013	0.0000**
Red		0.0062**				
Blue			0.0028**	0.0028**	0.0358**	
Green						
NIR		0.0117**			0.0052**	
RedEdge	0.0909*		0.0004**	0.0004**		
NDVI		0.2470				0.0054**
GNDVI		0.0649*	0.0024**	0.0024**	0.0422**	
RENDVI	0.0879*		0.0034**	0.0034**		
Cover						

Appendix C-6: P-values for the Chlorophyll Predictions of all Datasets with Only Multispectral Data Included. Some data sets have variable exclusions prior to the stepwise selection that can be seen in Table 3. P-values with two stars (**) indicate significance (p-value < .05). P-values with a single star (*) indicate marginal significance (p-value < .1).

Coefficiant	2020	2020-2022	No Cover	With Cover	Landsat Comparison	Landsat
intercept	0.0000**	0.0051**	0.0062**	0.0080**	0.7913	0.4081
Red		0.0004**			0.0586*	
Blue	0.1073	0.0732*				0.0626*
Green		0.0000**	0.0054**		0.0278**	0.4660
NIR	0.0003**	0.0000**	0.0048**		0.0885*	
RedEdge	0.0005**					
NDVI	0.1583	0.0010**			0.0705*	
GNDVI		0.0000**		0.0028**	0.0733*	
RENDVI	0.0005**		0.0423**	0.0973*		
Cover				0.0151**		

Appendix D: Cross validation error across all datasets and model types. WG = all variables, NG = without ground measurements, MS = multispectral indices Only