REMOTE SENSING TO ESTIMATE ABOVE-GROUND BIOMASS IN SOUTHEASTERN MONTANA GRASSLANDS

by

Shaelyn Joy Bauer

A professional paper submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Land Resources and Environmental Sciences

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2022

©COPYRIGHT

by

Shaelyn Joy Bauer

2022

All Rights Reserved

DEDICATION AND ACKNOWLEDGEMENTS

This work is dedicated to my family. Especially to my husband, Josh, who came into my life just as I was beginning this graduate school journey. Since then, we've married, moved, built a house in the country, moved again, and had two beautiful babies with another on the way within 8 weeks of this papers' submission. I wouldn't have been able to carry on toward this goal without his love, support, and encouragement. It's also dedicated to my three kids, Diehl Edmund, Truly Grace, and Baby Bauer #3, who have been the source of my motivation, distraction, and smiles.

I also want to dedicate this work to Chester and Rona Meyer, my parents and the owners of the study area ranch and my childhood home. They have been supportive throughout my education. Without their help with childcare while I collected samples and wrote this paper, I never would have been able to accomplish this project. Thank you, mom and dad, for the "ranch raising", for instilling in me a love of agriculture, the drive to be a good steward of the land and the desire help others.

A huge thank you to my advisor, Dr. Scott Powell, and his patience and guidance through my unconventional graduate school journey. I also want to thank all the LRES instructors, especially Nicholas Fox, for going above and beyond in helping me to learn the QGIS program which has been an enormous benefit to this project. I want to give special thanks to my "unofficial" advisor, Dr. Christopher Potter, who was the inspiration for this project and whose guidance was essential to its success.

I'd like to acknowledge and thank the Statistical Consulting and Research Service at Montana State University. Supervisor Mark Greenwood and graduate students Conner Demorest, Therese Lupariello and Braden Scherting helped me process and analyze my data and were an enormous help to me.

TABLE OF CONTENTS

1. INTRODUCTION 1
Background of Biomass Estimation.....1
Applications for remote sensing in grassland management.....2

2. INDICES FOR REMOTE SENSING ANALYSIS 4
Vegetation Indices 4
NDVI.....3
EVI.....5
Moisture Indices.....6
NDMI.....6

3. TECHNIQUES FOR REMOTE SENSING BIOMASS ANALYSIS.....6
Grassland Management Dynamics 8

4. PROJECT OBJECTIVES 9

5. MATERIALS AND METHODS.....10
Study Area 10
Sampling Methods 16
Satellite Image Acquisition 17
Data Analysis.....17

6. RESULTS 18

7. DISCUSSION.....22

8. LIMITATIONS AND FURTHER STUDY.....25

9. INCREASING ADOPTION OF REMOTE SENSING ABOVE GROUND BIOMASS ESTIMATION IN GRASSLAND MANAGEMENT 27

REFERENCES CITED.....30

LIST OF TABLES

Table Page

1. Landsat Spectral Indices products courtesy of the U.S. Geological Survey Earth Resources Observation and Science Center. 17

2. AICc Analysis Results 19

LIST OF FIGURES

Figure Page

1. Study Area in Carter County, Southeastern Montana....11

2. Rolling Rainfall of Study Year....11

3. North Section Soil Map. Each color denotes a distinct soil type as well as topographical variability.....12

4. South Section Soil Map. Each color denotes a distinct soil type as well as topographical variability.....13

5. Study Area....15

6. Clipping Biomass at Sampling Point....16

7. Validation Plot....20

8. Explanatory Variable Effect Plot....21

9. Southern Section Grazing Application....22

10. Northern Section Grazing Application....23

ABSTRACT

Grasslands provide numerous ecological services, including, but not limited to, carbon capture and storage, and wildlife habitat for numerous species including pollinators, domestic ungulates, wild herbivores, and their predators. Grasslands account for more than 40% of Earth's terrestrial surface and offer one of the most valuable land cover types, having the greatest diversity of grazing animals and predators on Earth. These grasslands serve as one of the largest reservoirs for terrestrial carbon, amounting to 10-30% of total world soil carbon (Hewins et al., 2018; Scurlock and Hall, 1998). Establishing carrying capacities of grasslands based on data, in both wildlife management and agricultural settings, is fundamental in grassland ecosystem management. However, ground observation of grassland biomass resources can be costly, time consuming, and is very limited in terms of scale, both spatially and temporally. Numerous remote-sensing applications have emerged to address the need for grassland biomass production data, with satellite-derived vegetation indices being the most widely implemented due to their wide variety of applications and the accessibility of free spectral imagery via federal agencies such as the U.S Geological Survey (USGS). Currently, there is no universal empirical model using statistical relationships between biomass and vegetation indices (VI), making custom-fitted models necessary for biomass estimation. Statistical analysis is limited by inadequate reference data and vulnerabilities to atmospheric factors. Results of multiple studies show strong correlations between field measured biomass and satellite-derived VI's. However, correlation strength depends on matching the best VI, or multiple VI's, to the land cover characteristics in question, coupled with validation tools and methodologies on the ground. Therefore, to explore the feasibility and accuracy of a site-specific biomass estimation model, I chose a privately owned ranch in eastern Montana for sampling and tested for combinations of indices that were best correlated with ground biomass measurements. I used several vegetation indices to compare the accuracy of univariate and multivariate modeling techniques for modeling above ground biomass. A multivariate model using the normalized difference vegetation index (NDVI), the enhanced vegetation index (EVI) and a native/introduced pasture classification variable arose as the model with the greatest predictive power for biomass estimation. This model accounted for 57% of the variation in response and was accurate to within 352.7 lbs/acre. My study lays out real-world applications for tailoring VIs to the application requirements of measuring biomass in an eastern Montana grassland coupled with ground validation. The methods in this study serve as a framework for fitting custom models to other grasslands under management.

1. INTRODUCTION

Background of Biomass Estimation

The major rangeland types of the world are grasslands, desert shrublands, savanna woodlands, forests, and tundra; each type has its own unique plant associations. Grasslands are marked by the dominance of plants within the Gramineae family and are relatively free of woody plants. These are the subset of rangelands focused on within my study.

Remotely sensed information of terrestrial plant growth is extremely useful for environmental management, biodiversity monitoring, conservation, agriculture, and forestry applications (Campbell, 2011). These types of data are applied to agricultural crop management, or precision agriculture, as it has been termed. This technology has only been 30 years in development (Xue and Su, 2017). Precision agriculture sometimes called “prescription farming” or “variable rate technology”, consists of applications and technologies aimed at increasing the efficiency of resource use. These technologies have not been widely applied to rangeland or, more specifically, grassland management, and those that have been developed for public use are not designed to remotely quantify rangeland/grassland resources such as above ground biomass production (ground validation is still needed). Above ground biomass represents current growing season organic matter resulting from plant photosynthetic activity. Global carbon cycling hinges on this photosynthetic activity (Zhou et al., 2021). The need for ongoing ground validation of above ground biomass greatly reduces the feasibility of adoption for grassland managers due to labor and monetary constraints.

For example, the Rangeland Analysis Platform (RAP) (Rangelands App, 2022) provides annual and 16-day above ground biomass estimates across the United States from 1986 to

present. Estimates are provided over broad regions using a process-based model for which traditional error metrics are not available. Alternatively, the estimates are compared to three other national scale datasets for accuracy. Due to the scale and validation method of the RAP, the results are not intended for the quantification of rangeland resources, determining thresholds, or to evaluate the effectiveness of management practices, all necessary features of PA. They are, however, useful for monitoring change and provide indicators of ecosystem trends over time.

Most of the present research and development in PA has been directed at arable crops rather than grasslands. This is likely due to the economic superiority per acre in production of commercial crops compared to milk and beef production in less desirable climate conditions, even though the latter is a substantial portion of agricultural land under management and contributes greatly to total agricultural capital and income in the U.S. (Jurgen et al., 2008). Precision focused technologies could provide major economic and environmental benefits through better grassland management decisions (Schellberg, 2007).

Applications for Remote Sensing in Agricultural Grassland Management

Practical applications of remotely sensed PA data could include variable rate grazing treatments, variable stocking rate and livestock nutritional supplementation recommendations. Grassland growth-rate monitoring and feed-on-offer projections (ratio of total above ground biomass that is available for livestock grazing) could also greatly increase decision making confidence for managers. Animal husbandry can benefit strongly from PA technology and is strongly linked to grassland utilization (Jurgen et al., 2008). Free access to visible and

multispectral data products, such as those from the Landsat satellite suite, removes cost prohibitive obstacles that prevent advancements in grassland management. Platforms such as Australia's Pastures from Space consortium, which can predict feed-on-offer to within 73% accuracy using Landsat sensors have greatly advanced grazing management practices in Australia and is free to use (Edirisinghe, 2004). Other online tools provide biomass forecasts (e.g., Grass-Cast, https://grasscast.unl.edu/ Peck et al., 2019), and, in a few select locations, standing herbaceous biomass across the growing season (e.g., RangeSAT, https://www.rangesat.org/). In other management uses, data are collected at the end of a grazing season and used retrospectively to assess management outcomes and inform decisions in future years or to track changes over time (Janson et al., 2022). More specifically, biomass residual vegetation data and utilization rates by livestock would be useful metrics gathered at the conclusion of a growing season.

Other potential applications include tracking changes in productivity over time. Satellite observation data are available dating to the 1970s when the normalized difference vegetation index (NDVI) was developed. Adjusted for rainfall, these data may provide insight into overall ecosystem sustainability trends over the last 40+ years and inform land managers of the long-term effects of their management practices. Photosynthetic activity driving above ground biomass is a sensitive indicator of ecological function, resilience and environmental shifts (Segoli et al., 2008). It is also an indicator of hydrologic function, resistance to erosion degradation, and carbon cycling.

2. INDICES FOR REMOTE SENSING ANALYSIS

Vegetation Indices

A vegetation index (VI) is defined as the arithmetic combination of two or more spectral bands related to vegetation characteristics (Xue & Su et al., 2017). VI's are used for vegetation classification, phenologic monitoring and the biophysical derivation of radiometric and structural vegetation parameters.

NDVI

The normalized difference vegetation index is the most widely used VI; it was developed by Rouse Jr. et al. (1974), and can be expressed as:

$$NDVI = (NIR - R) / (NIR + R) \quad \text{(Equation 1)}$$

NDVI represents the normalized ratio between the red (R) and near infrared bands (NIR) (Rouse, 1974). NDVI is sensitive to green vegetation and photosynthetic activity, even at low canopy cover, and can be used to measure canopy growth or vigor and leaf area. NDVI is known to saturate in dense biomass stands, is sensitive to soil brightness and color factors, and like all VI's, is sensitive to cloud cover and shadow. Garonna et al. (2009) found that atmospheric contaminants primarily affected NDVI estimates in wet climates and seasons and had less impact in dry climates and seasons. The result of soil and atmospheric interactions is for the reduction of one to often increase the other (a negative correlation between NDVI and soil brightness)

(Garonna et al., 2009). Because of these relationships, NDVI is more suited to biomass estimation in semi-arid regions. Indices like the Soil-Adjusted Vegetation Index (SAVI) have been proposed to correct for these soil and atmospheric effects where characteristics warrant. Alternatively, a feedback mechanism has been introduced that contains a parameter which corrects for soil and atmospheric effects simultaneously; this parameter is the Enhanced Vegetation Index (EVI) (Xue & Su, 2017).

EVI

EVI, like NDVI, is correlated with the "greenness" of the vegetation canopy which is an indicator of vegetative activity. However, EVI is more sensitive in areas with dense vegetation and corrects for some atmospheric conditions. EVI can be expressed as follows:

$$EVI = G * ((NIR - R) / (NIR + C1 * R – C2 * B + L)) \text{ (Equation 2)}$$

which includes near infra-red (NIR), red (R) and blue (B) band values corrected by the atmosphere. "L" represents a soil adjustment parameter which corrects for canopy background. "C" values are coefficients for atmospheric resistance (Landsat Missions, 2022).

The soil adjustment factor, "L" in the EVI can make it sensitive to topographic conditions. Matsushita et al. (2007) recommended that topographic effect, meaning the change in radiance that comes with a change in orientation from a horizontal to an inclined surface, be removed in the reflectance data before EVI is calculated over rough terrain. It is unknown how influential

this topographic effect is when estimating biomass via EVI and its effect beyond the scope of this research but could be a topic of further study. Additionally, the effectiveness of EVI in varying climates for biomass estimation seems to be less documented than NDVI.

Moisture Indices

NDMI

The Normalized Difference Moisture Index (NDMI) describes the water stress level, or water content, of vegetation and is calculated as follows:

(NIR - SWIR) / (NIR + SWIR) (Equation 3)

which uses near infrared (NIR) and short-wave infrared (SWIR) bands (Landsat Missions, 2022). A model combining NDMI and NDVI has previously been used to monitor bare soil expansion and soil wetness change under drought conditions (Rahman, 2019). It has also used for measuring live fuel moisture content in wildfire danger monitoring (Costa-Saura et al., 2021). NDMI has also been used in conjunction with NDVI and EVI in monitoring post-disaster recovery in mangrove areas (Tickman et al., 2021). Additionally, NDMI and NDVI patterns have been used in land surface temperature studies (Sarif et al., 2020).

3. TECHNIQUES FOR REMOTE SENSING BIOMASS ANALYSIS

Models for the estimation of above ground biomass via satellite remote sensing data are generally empirical in nature. Empirical models use statistical relationships between biomass and vegetation indices (Tucker, 1997). Statistical analysis based on vegetation indices is the most common method for biomass estimation, but is limited by inadequate reference data and vulnerability to atmospheric factors, such as cloud cover (Chao et al., 2019).

Machine learning algorithm models have also emerged as a very accurate alternative to more traditional statistical modeling techniques. One study used high accuracy surface modeling (HASM), which is a machine learning approach for above ground biomass modeling and found that this technique reproduced the spatial distribution characteristics of the biomass well and was able to predict above ground biomass to within 85% accuracy; performing better than other algorithm-based, machine learning models such as the support vector machine or random forest machine learning models (Zhou et al., 2021). Grassland biomass production is influenced by many factors such as terrain, climate factors, and soil properties and HASM has been successfully applied to incorporate these parameters to surface simulations. Zhou et al. (2011) incorporated climate (temperature, precipitation and solar radiation), terrain, and vegetation index as model input parameters and effectively reproduced the spatial distribution characteristics of above ground biomass .

The moisture factor is a critical component of above ground biomass production, but models combining vegetation indices with moisture index values for biomass estimation purposes are not well studied. NDVI has been used in studies as a factor of vegetations effect on

land surface temperature for post-disaster recovery and drought assessments (Ticman et al., 2021; Karnieli et al., 2010). Inversely, moisture indices such as NDMI are not commonly used in models to predict biomass. There may be opportunities in modeling biomass production more accurately by accounting for vegetation moisture content, as this is strongly correlated with productivity (within reasonable parameters).

Considerations in Grassland ABG Model Development

The heterogeneity of grasslands, within and among fields, is considered a significant constraint to precision grassland agriculture development and application. For example, the application of remote sensing techniques remains difficult on native grassland that is diverse in species composition, plant spacing, color and morphology compared to a single crop which exhibits distinct and consistent canopy characteristics. Understanding this heterogeneity would better enable the implementation of PA where competitive and cost efficient (Jurgen et al., 2008).

When using hyperspectral remote sensing data for biomass estimation, the issue of high data dimensionality arises, making carefully designed research and modelling very important. Many observations are required to achieve adequate predictability in a highly dimensional space, such as species rich grasslands in agriculture production. This can lead to poor understanding of the correlation among individual variables and how changes in one variable are associated with changes in other variables (Clevers et al., 2007). This can result in poor prediction results compared to models with fewer explanatory variables. Thus, selecting models with a suitable

number of explanatory variables to appropriately account for grassland heterogeneity is important.

Contributing to the difficulty of understanding grassland above ground biomass estimation, contrasting the defoliation of the swath for silage and hay, grazing animals create spatial patterns of grazed biomass that continuously change throughout the grazing year. Palatability of forage, season of use, vicinity to water, fencing infrastructure and ruminant species, age, behavior, and stage of production all have an effect on how a grassland will be used. These dynamics have considerable effects on spatial heterogeneity of grasslands which adds a layer of complexity when attempting to model and predict the effects of these dynamics (Rook et al., 2004). These inherent constraints of accurate modeling and predictions for PA applications within agricultural grasslands also mark the need for more precise management of this valuable resource. Where there is variability, precision is necessary.

To explore the feasibility and accuracy of a site-specific biomass estimation model, I studied a privately owned ranch in eastern Montana for sampling and testing the combination of remotely sensed vegetation indices that are best correlated with ground measurements. I am familiar with the many livestock/grassland management dynamics at play across the study area which enabled me to custom fit the most appropriate input parameters to account for the heterogeneity inherent within this grassland agricultural operation.

4. PROJECT OBJECTIVES

The objectives of my study are to: 1) collect above ground biomass reference data on a ranch in eastern Montana; 2) compare satellite-derived vegetation indices and moisture indices and combinations thereof with respect to their correlation to the reference data; 3) explore statistical approaches for estimating above ground biomass based on satellite-derived imagery; and 4) review approaches for increasing the adoption of remote sensing technology by rangeland managers for above ground biomass estimation.

5. MATERIALS AND METHODS

Study Area

The study area is a 9,500 acre, privately owned ranch located west of the town of Ekalaka in Carter County, Montana (Figure 1). The ecological site description encompassing the study area is descriptive of dense clay sites receiving 10-14 inches of annual precipitation (websoilsurvey.org). The rolling rainfall prior to sampling, conducted in July of 2021, was only 6.56 inches from July 2020 to June 2021 according to nearby AgACIS weather stations labeled, MIZPAH 4 NNW, MT and Ekalaka, MT (AgACIS.org). This was approximately 46% of average rainfall for the area compared to long-term averages (14.79 inches) (Figure 2).

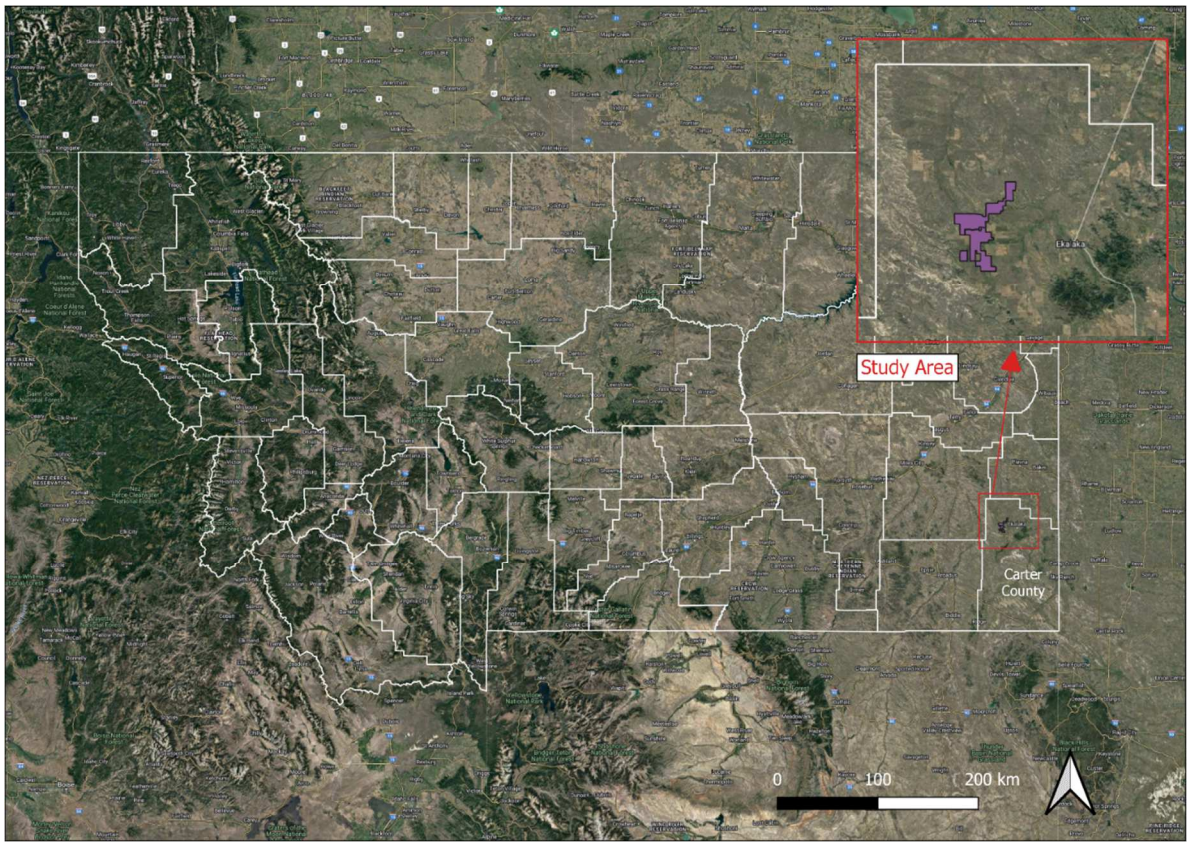


Figure 1. Study area in Carter County, southeastern Montana.

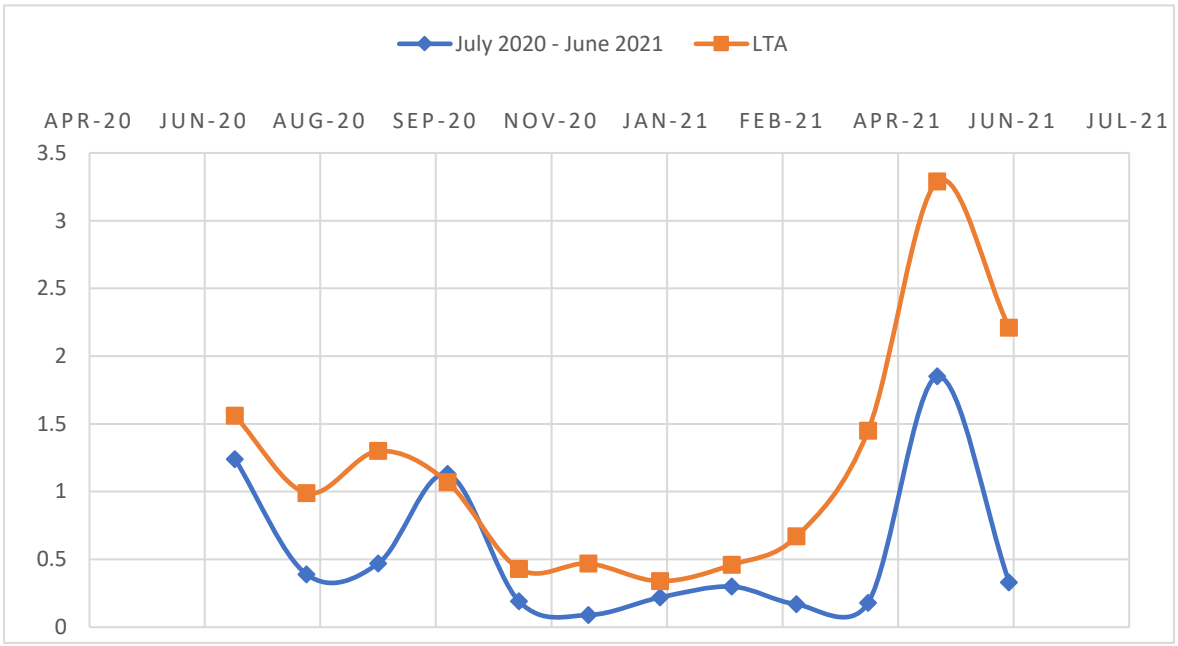


Figure 2. Rolling rainfall of study year. Long-term average based on 30-year dataset.

The topography and soil types across the study area vary, with a distinct difference between the northern and southern portions of the study area. The northern half is predominantly comprised of unfarmed native range with diverse topography and species composition and ash tree lined drainages (Figure 3). Primary species composition includes sand dropseed (*Sporobolus cryptandrus*), blue grama (*Bouteloua gracilis*), needle and thread (*Hesperostipa comata*), western wheatgrass (*Elymus smithii*), buffalograss (*Bouteloua dactyloides*), prairie junegrass

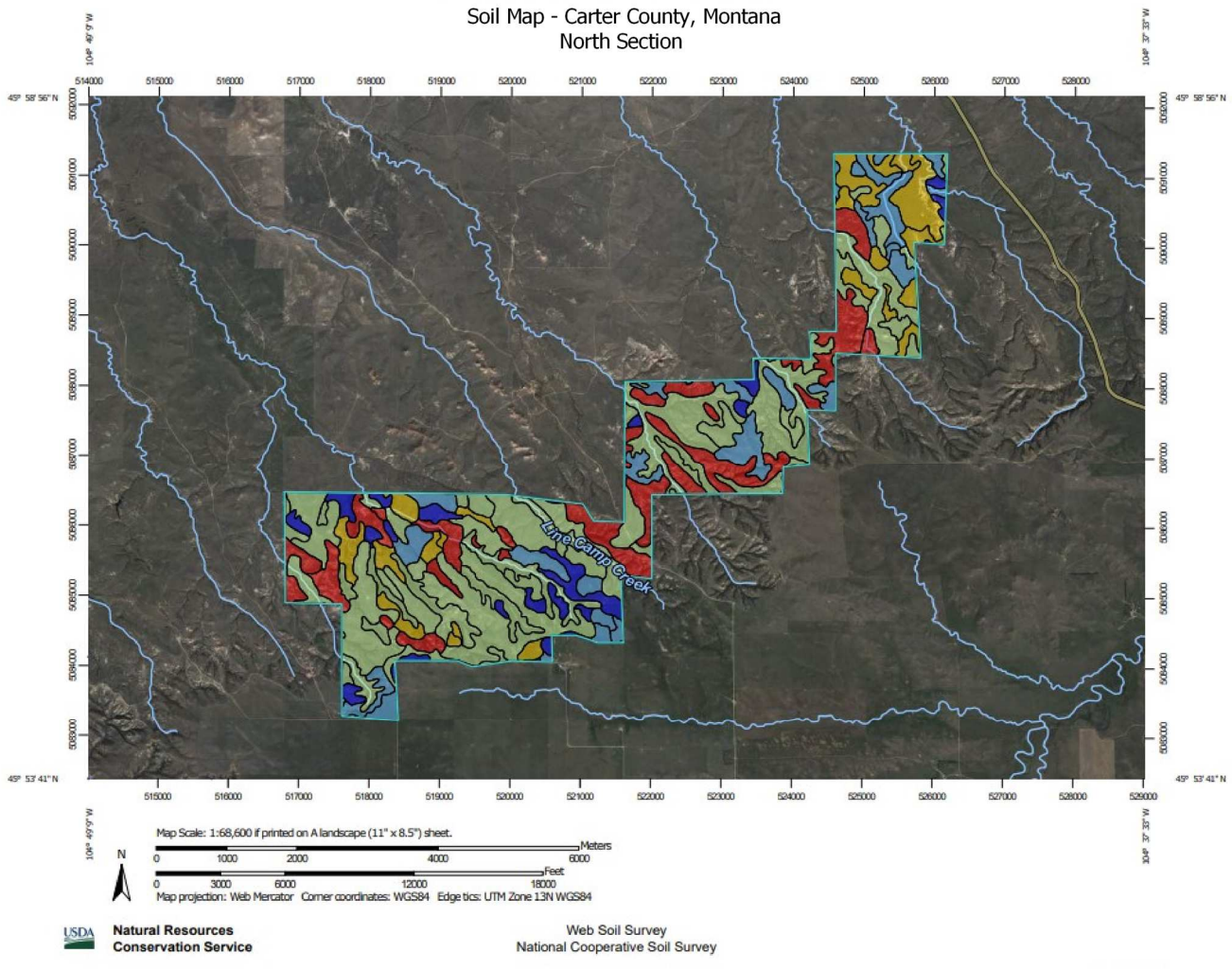


Figure 3. North section soil map. Each color denotes a distinct soil type as well as topographical variability.

(*Koeleria macrantha*), sandberg bluegrass (*Poa sandbergii*) and fringe sagebrush (*Artemisia frigida*). The northern half has much greater variability in soil type, depth, and productivity than the southern portion of the study area (Figure 4).

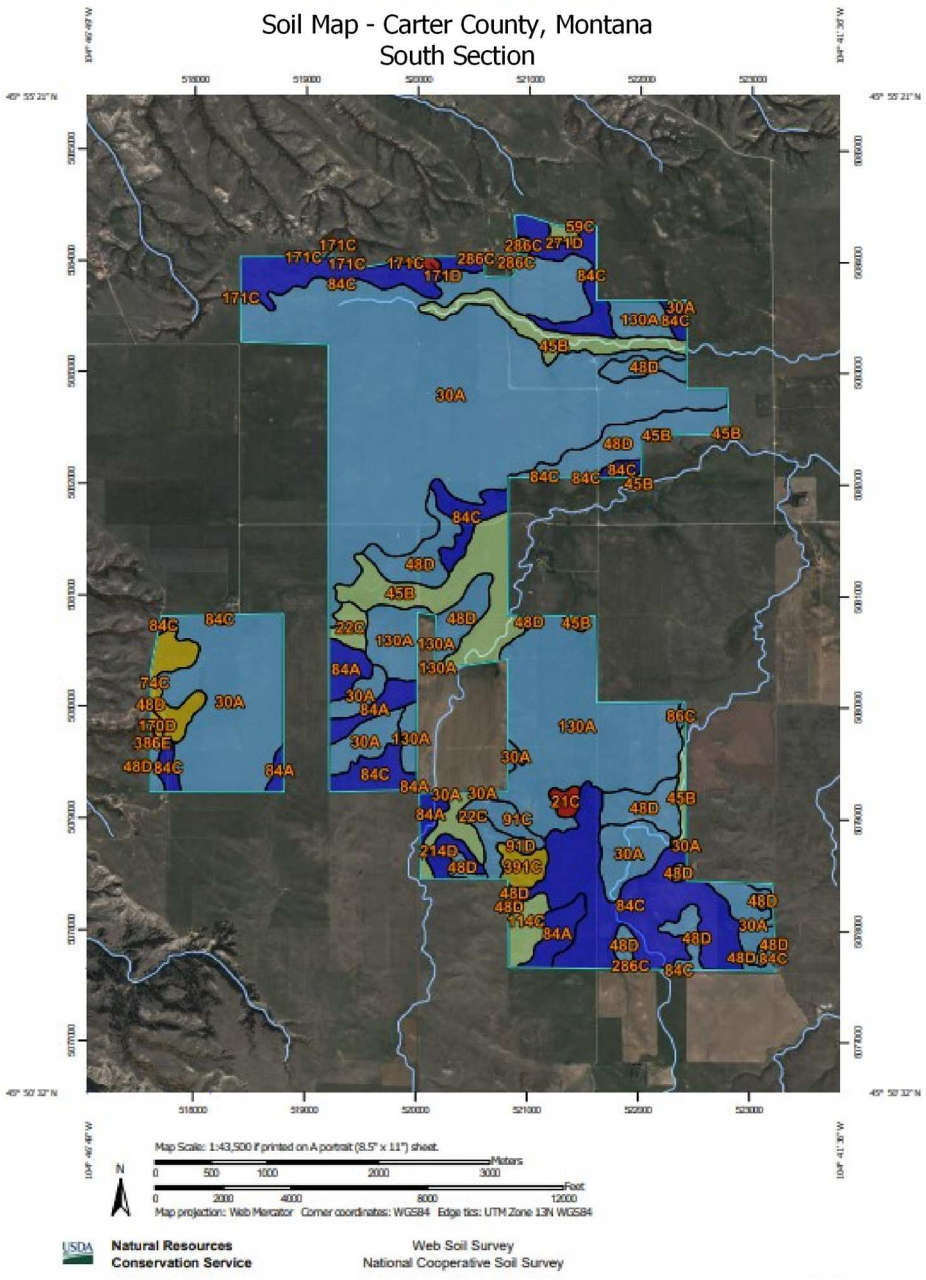


Figure 4. South section soil map. Each color denotes a distinct soil type as well as topographical variability.

The southern portion of the study area was historically cultivated (tilled) and farmed for monoculture grain crops but has since been converted back to non-native grass pasture. With less species diversity than the northern section, this southern section is predominately comprised of just three species, pubescent wheatgrass (*Agropryon trichophorum*), western wheatgrass (*Pascopyrum smithii*), alfalfa (*Medicago sativa*), and Kentucky bluegrass (*Poa pretensis*). This southern, previously farmed, non-native pasture is referred to as "tilled" in this study. The southern pastures are mostly flat, with a more consistent soil type and depth with less variation in biomass productivity within pixels. There are also few trees that could affect vegetation index outputs.

Land use also varies between the two sections (Figure 5). For example, fenced pastures or paddocks in the northern section are larger, between 200 to 600 acres, and grazed less intensively, and for longer durations than pastures in the southern half (0.9 to 2.75 standard animal units per acre). The pastures in the southern section have been fenced into long laneways that are then cross-fenced into 3-5-acre paddocks with electric polywire fencing and very intensively grazed at high stock densities and for short durations (100-160 standard animal units per acre). This creates a grazing pattern that is more uniform and, likewise, a more uniform regrowth pattern as well. This difference in grazing dynamic may influence the accuracy of biomass estimation.

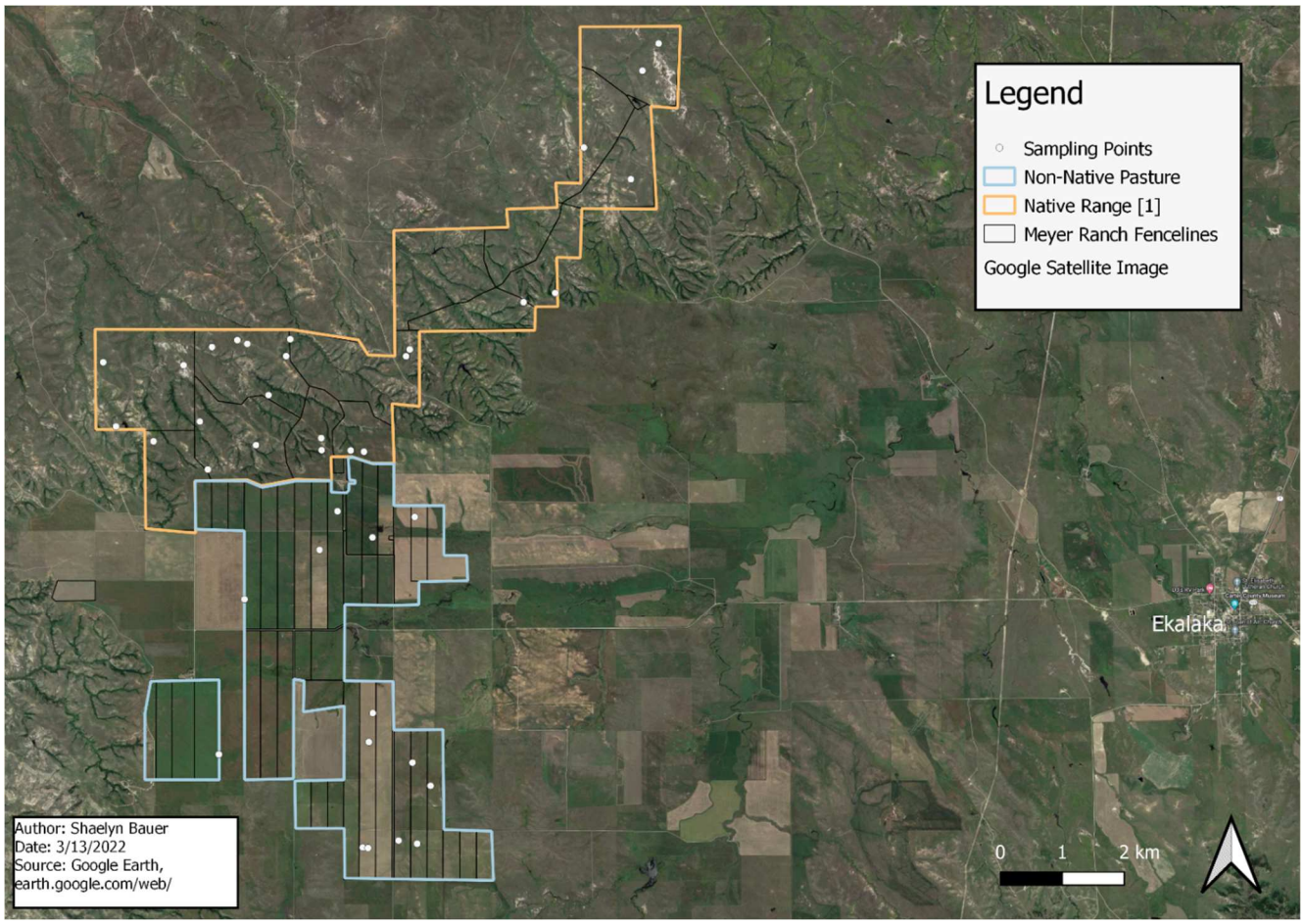


Figure 5. Study area.

Sampling Methods

I randomly selected 40 samples from within the study area boundary using a random point generator in the global information software QGIS 3.12 (QGIS Development Team, 2022). I sampled biomass from July 15-17, 2021. I clipped all standing biomass from within a 4.8 ft² cable hoop placed at each sampling point. I located points with a standard, civilian GPS unit, accurate to within ten feet. After placing the hoop at each sampling site, I clipped, bagged and weighed each sample with a gram scale which had been zeroed out to account for the weight of the bag. I dried the samples in a convection oven until dry and brittle to the touch and reweighed to obtain a dry weight.

4.80 ft^2
Radius = 1.24 feet
Hoop Circumference = 7.77 feet
Square Plot Dimensions = 2.19 X 2.19 ft
Conversion Factor = Grams X 20 = lbs/ac
= Grams X 22.42 = kg/ha

Equation 4. Sampling Hoop Specs and Conversion Factor (NRCS.gov)



Figure 6. Clipping biomass at sampling point.

I applied a conversion factor of twenty to the dried clipping weights to estimate the dry biomass weight measured in pounds (lbs.) per acre.

Satellite Image Acquisition.

Landsat Analysis-ready data , encompassing the study area, for NDVI, NDMI and EVI indices were then ordered for the closest date to the sampling period with the best quality images. Images taken on 07/20/2021 were pulled from the Earth Resources Observation and Science (EROS) Science Processing Architecture (ESPA) on Demand Interface of the USGS database (https://espa.cr.usgs.gov/). All images obtained were offered as "analysis-ready," meaning they were pre-processed for corrections to surface reflectance. I then obtained pixel values at sampling points, utilizing QGIS software, by extracting pixel values for each of the indices. Extracted pixel values and sampling point details were exported to an Excel spreadsheet for initial analysis and to identify key variables likely to be highly correlated.

Data Analysis.

Several of the original 40 sample points were left out of the analysis for differing reasons. One discrepancy was found between the randomly selected sampling point and the actual coordinates sampled, likely due to nearby tree interference, and was eliminated (point #35). Other high outlier points with NDVI values greater than 3,000 (point #33 & 17) were eliminated

Table 1. Landsat spectral indices products courtesy of the U.S. Geological Survey Earth Resources Observation and Science Center.

Section	Number of Points	Mean Biomass/ Acre (lbs.)	Biomass/ Acre Range	Mean NDVI	NDVI Range	Mean NDMI	NDMI Range	Mean EVI	EVI Range
North	24	587.5	160 - 1840	2430.5	2079-2898	-2242	-2652 to -1747	1482	1360 - 1809
South	13	1207.7	520-2100	2407	2027-2788	-2238	-2414 to -2030	1559	1247-2286

from the sample set due to the lack of supporting data points to justify biomass prediction at that level of greenness. This left a remaining 37 data points (Table 1).

Multiple linear regression model selection was performed using R 4.1.0 (R Core Team, 2021). Predictor variables included in the model selection were NDVI, EVI, NDMI, and the land class (Tilled/Native) with the response variable being predicted biomass per acre. For ideal model selection, the Akaike Information Criterion, with a correction for small sample size, (AICc) was used in conjunction with "data dredging", which considers all possible combinations of explanatory variables and compares the fit of all possible regression models, finding the best fit model to explain the greatest amount of variation in the data while penalizing models that use excessive parameters (Barton, 2022 & Greenwood, 2021). An AICc is an estimate of the distance (or discrepancy or divergence) between a candidate model and the true model, on a log-scale. The models that are closer (have a smaller distance or divergence value) to the truth are more accurate (Greenwood, 2021).

Leave-one-out cross validation was then performed using the cross-validation for generalized linear models (cv.glm) function in the boot package to characterize predictive performance (Davison & Hinkley, 1997). This function removes one observation from the data

set, fits the model to the other 36 observations, and estimates the biomass of the removed observation point based on that observation's predictors. This procedure is repeated for each observation in the data set. The errors in these predictions are squared and then averaged to give the overall cross-validation error. It should be noted that in using AICc's for model selection, p-values do not carry the same interpretation and will not be reported in these findings (Greenwood, 2021).

6. RESULTS

The AICc analysis generated a list of the top 8 models tested. The predictor variables NDMI, NDVI, EVI, and land class (Tilled) revealed the model with the greatest predictive capability to be NDVI, EVI and land class (Table 2). The tilled variable is to be assigned a value of "0", with native range receiving a value of "1" to assign numerical value to these categorical predictors. Results are in pounds of dry biomass per acre (lbs/acre). The model with the greatest predictive capability was:

$$\text{Lbs/acre} \sim (-4.291)*\text{EVI} + (2.282)*\text{NDVI} + (491.145)*\text{Tilled} \quad \text{(Equation 4)}$$

The second highest rated model included NDMI as well as the EVI, NDVI, and Tilled factors. This model has an AICc score of 547.5; 0.7 units larger, and less accurate, than the simpler model. As stated before, simpler models are given preference under AICc analysis.

Table 2. AICc analysis results. When a numeric predictor is included in a model, the coefficient estimate of that parameter is displayed. When a categorical predictor is included, the table displays a "+", however the coefficient estimates for these predictors are not displayed. When a second "+" is displayed, this represents a two-level categorical predictor, NDV:Tll for example, that also has an un-displayed single coefficient.

Number	Intercept	EVI	NDMI	NDVI	Tilled	EVI:Tilled	NDMI:Tilled	NDV:Tll	Degrees of Freedom	AICc
14	1485	-4.29		2.282	+				5	546.8
16	3488	-4.447	0.5246	2.038	+				6	547.5
78	1672	-4.194		2.145	+			+	6	549.3
30	1405	-4.219		2.271	+	+			6	549.6
80	3732	-4.343	0.534	1.882	+			+	7	550.2
48	3484	-4.474	0.553	2.083	+		+		7	550.4
32	3422	-4.396	0.5222	2.031	+	+			7	550.5

The residual standard error of the top model was 352.7 with a multiple r squared of 0.57, which means that 57% of the variation in response can be explained and we could expect to estimate biomass production to within 352.7 lbs of dry biomass per acre, our Root Mean Squared Error (RMSE), while a mean-only model, or the model with the least predictive capability, gave a RMSE of 523.9 lbs/acre. The AICc score of this top-rated model is 546.8.

The validation plot below shows actual the prediction error for each data point by plotting actual values vs predicted values (Figure 7). Idealy, points are centered around the x=y line with no curvature or bias above or below.

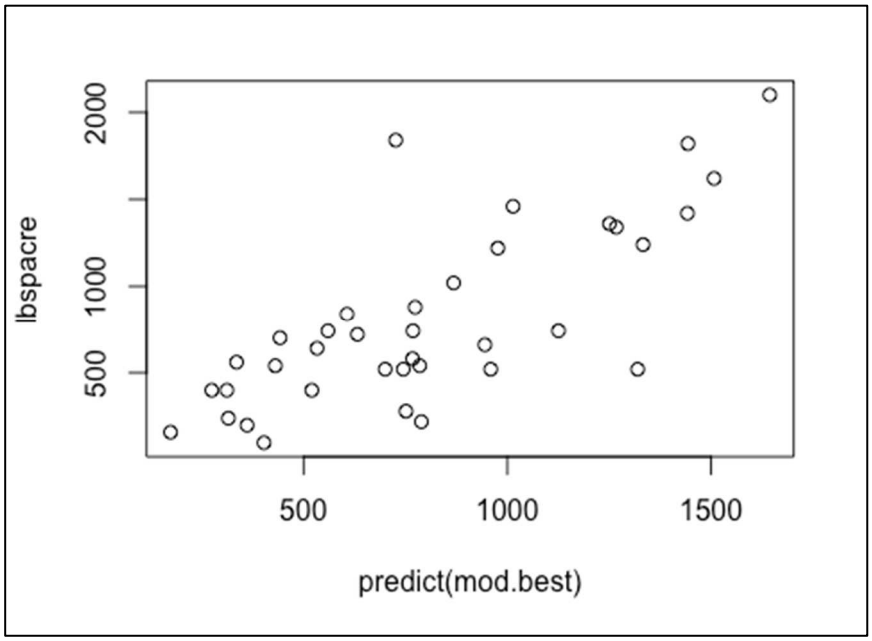


Figure 7. Validation plot.

Below (Figure 8) are effect plots depicting the predicted outcome values of each input variable of the top model tested (*lbspacre ~ (-4.291)*EVI + (2.282)*NDVI + (491.145)*Tilled)* with the other variables held constant. These plots show how correlated each input variable is with measured values. We can see from these effect plots that the predictions of our vegetation indices, the EVI and NDVI explanatory variables, are highly correlated, with some error in measured values compared to predicted values. We also see that the tilled and introduced grassland has higher biomass production than the native grassland.

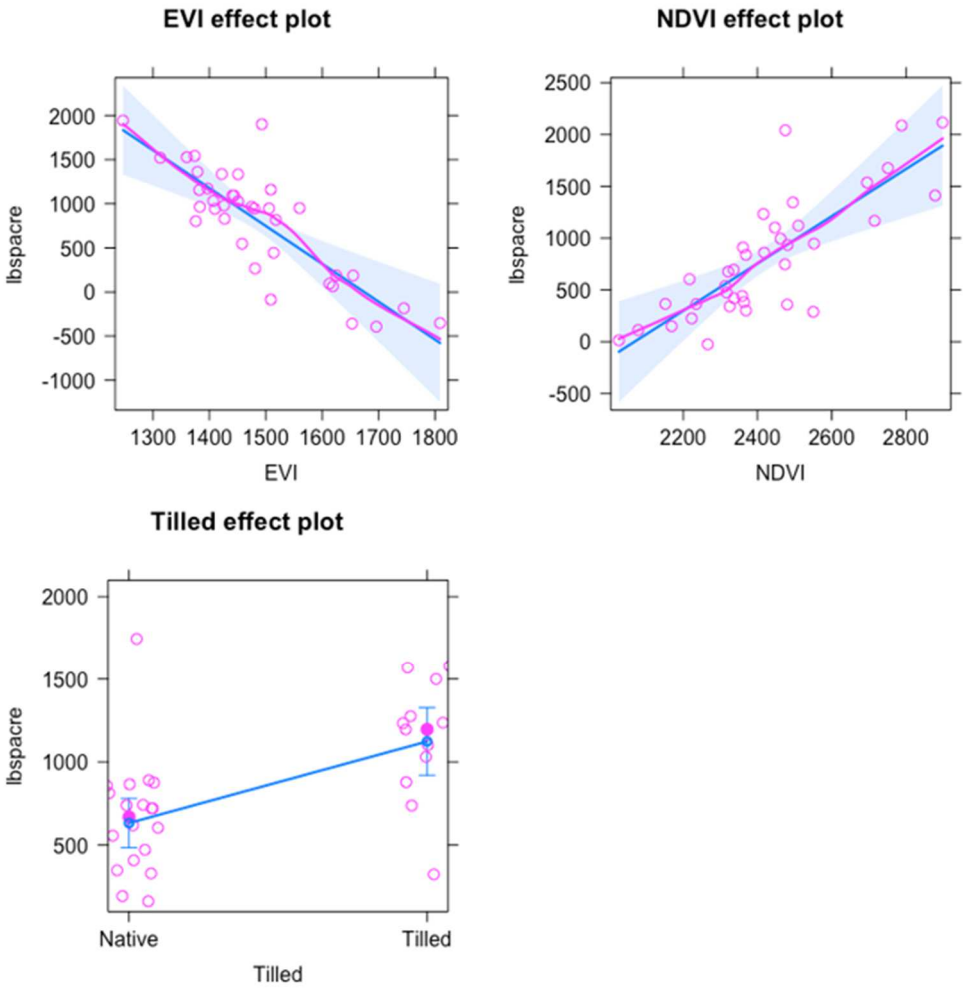


Figure 8. Explanatory variable effect plot (Fox & Weisberg, 2018)

7. DISCUSSION

My study shows promise in biomass estimation using a multivariate approach including NDVI and EVI and land use characteristics. However, more sampling needs to be done over multiple years to potentially improve the accuracy of this model before it can be used for any precision management applications.

There were some confounding variables not directly accounted for within this model. For example, as mentioned, there is a distinct difference in grazing regime between northern and southern sections of the study area. There were also differences in average production, with the southern section production more biomass than the northern section. The tilled/untilled variable attempts to adjust for this difference in livestock grazing management as well as the difference in forage type between the northern (native) and southern (introduced) pasture types, however because grazing occurs year-round, and at a variety of densities and durations, the complexity of



Figure 9. Southern section grazing application.

this variable is difficult to effectively incorporate into a model. For visual comparison, Figure 9 is representative of the grazing intensity regularly applied on the southern portion, while Figure 10 is representative of grazing density routinely applied in the northern section. We can see from these two figures that animal density, biomass utilization and trampling, and the greenness of

remaining forage are quite different despite the difference in vantage point. The duration of grazing events between the two areas will also be very different, with grazes in the southern section only being approximately 12 hours long at the grazing density pictured and up to 2 weeks long at the density pictured in the northern section of the ranch.

This study was conducted July 15-17 when, in an average rainfall year, greenness and peak forage production would be peaking. In 2021, however, with the drought conditions, the period of peak greenness had passed and senescence was occurring across most of the study



Figure 10. Northern section grazing application.

area. Overall production was also extremely low; approximately 50% to 65% of average production according to local knowledge and comparisons of sampled values to what had been observed in more average precipitation years. It is unknown just how the drought variable has

affected the outcome of this study, but repetitive analysis across future growing seasons, under variable rainfall parameters, would begin to answer this question.

As discussed, heterogeneity within and among natural grasslands is another constraint to the application of effectively modeling biomass production for the application of precision agriculture practices. These dynamics can make it difficult to account for the many variables that can affect modeling results, such as species composition, and land management variables. For example, the harvesting of grasslands with ruminant animals is never completely uniform like the swath of a harvester. Additionally, the harvest of grasslands can feasibly take place any day of the year, potentially complicating ground validation efforts and modeling. If grazing status (whether a grassland has been used by ruminants) affects model accuracy, accounting for this variable will take further research. I did test grazed vs ungrazed plots and their correlation to NDVI, EVI, and NDMI to see if grazing status seemed to affect modeling accuracy, and the results were unclear. I considered any plots located within pastures grazed within the current grazing season (April – October) as “grazed”. However, it is unknown what amount of regrowth had occurred since the last graze, to what degree the plants within the plots located in those pastures were defoliated, and how grazing pressure varied within pixels located within those pastures. Because of the distinct difference in grazing pressure, amount and uniformity of defoliation, and vegetation type between the North and South sections, I chose to include the tilled/native parameter in the data dredging process in hopes of accounting for some of these grazing dynamics at play within this specific study area between the native north section and the previously farmed southern section. Additionally, vegetation dynamics vary widely depending

on current weather conditions aside from precipitation which could, in theory, move any benchmarks set attempting to account for grassland/herbivory dynamics.

8. LIMITATIONS AND FURTHER STUDY

Limitations of remote sensing are related to pixel resolution (30 m per pixel for Landsat) and orbit period (16 d for Landsat). Other commercial satellite products provide higher resolution imagery with sensors more focused to particular ranges of the electromagnetic spectrum, providing features inaccessible on publicly available platforms, however these platforms can be cost prohibitive (Xue and Su, 2017). The cloud cover, which is impenetrable to passive sensors, also poses an obstacle to regular monitoring. Unmanned Aerial Vehicle (UAV) platforms can solve resolution and revisitation intervals; however, these technologies can also be cost prohibitive and cannot be implemented at broad scales due to limited flight times.

Being restricted both temporally and in scale, this study has a scope of inference that is limited to the study area but could reasonably be generalized to any location within the study area for NDVI measurements between 2,000 and 3,000. NDVI measurements over 3,000 were problematic because they were far from other values, were possible influential points, or because the points were within pixels containing trees.

These results cannot yet be applied to other time periods or areas outside of the study area and sampling period without additional validation. Additionally, the predictive accuracy of this current model is not well suited to precision agriculture applications due to the generous margin of error. However, there is certainly potential for other ecological monitoring applications.

This study template is a guideline for model development fitted to property specific needs and management variables. By repeating this study process and sampling random points within a property, collecting yield data and tuning a model based on that data, one could use this template to make trustworthy predictions and management decisions. The applicability of this specific model to other properties is unknown, but the potential is there for expanding the dataset to include a broader scale and to decrease prediction error.

The utility of short-term, site-specific monitoring data, such as what was collected in this study, to inform broad-scale management decisions, depends on spatial and temporal correlations. There is reason to believe that there will be some temporal correlation within the measurement sites. It would also be reasonable to expect correlation between this models' predictions and measurements taken from other locations within the same ecological site as this study area, provided that the land class variable is known. However, these assumptions would need further validation as the prediction error estimate is only for predictions within the study area.

As previously stated, this model was created under drought conditions which may have skewed the results. Further validation is needed to confirm correlations between these data and differing climate conditions such as an average and above average precipitation growing seasons. If accuracy improves when this model is applied in wetter years, one potential solution to the dry, brown biomass variable, would be substituting a VI better suited to measuring brown forage. For example, Jansen et al., (2018) successfully measured biomass during both wet and dry years by using a two-model approach, substituting the normalized tillage index (NTDI) for when plants

were brown, or the normalized burn ratio (NBR) for when plants were green. NDVI was used to guide which model (brown or green) was used.

Additionally, it is unknown whether the native grassland or introduced pasture predictor variable would improve predictions in another area containing the two land classes. It is unknown whether this predictor variable improved the model's accuracy due to the difference in grazing regime between the two land classes, or whether forage amount or type was the primary driver.

9. INCREASING ADOPTION OF REMOTE SENSING TO ESTIMATE ABOVE GROUND BIOMASS IN GRASSLAND MANAGEMENT

The primary drivers of the adoption and use of any innovative agricultural technology are its ability to address a problem, its accuracy or reliability, and accessibility. When a source of information or a technology is new to managers, it is important to build evidence based on empirical relationships to begin relating common in-field observations to remote sensing metrics. One bottleneck discussed previously is the lack of reference data, or ground observations, needed to calibrate custom models. The success of models built in conjunction with ground observations hinges on not only the accuracy of the reference data, but also the scale and abundance of reference data collected. Diverse grasslands and dynamic grazing patterns require many observations to adequately capture the inherent variability within the system. However, when

ground observation is costly and time consuming, there needs to be some creativity in how reference data are collected and compiled.

There is no optimal field-based method for every management unit. However certain intra-method calibration has been shown to improve the accuracy of biomass estimates when measurement-based methods (e.g., clipping and weighing) are done on a small number of subplots to calibrate observations over a larger area with less intensive ocular estimates (Janson, 2022). This is similar to double sampling but does not seek to establish a statistical relationship between the two methods, only to improve accuracy of the less intensive measurement technique (Herrick, 2009). This could present an acceptable trade-off between accuracy and scale when collecting data for building custom fit models. Even if there is a slight decline in accuracy of measurements, if there is a corresponding increase in reference data, the resulting model may benefit overall.

Intra-method calibration techniques also open the door to citizen-science participation. With training and the appropriate supplies, there is an opportunity for landowner participation in the collecting of reference data for the development of more custom fit models in collaboration with government agencies, universities, nonprofits or independent consultants. In these “citizen science” scenarios, when multiple observers are compiling datasets, training is crucial for maintaining consistency and accuracy.

Additionally, when adoption is the goal, associating relevant and traditional field vegetation data commonly used in grassland management to remotely sensed data, an important frame of reference is maintained that increases the usefulness and applicability of satellite data for management and monitoring purposes (Janson, 2021). For example, presenting data in

familiar metrics to what managers currently use is helpful in integrating these technologies into agricultural settings. This highlights a need for involvement of land managers and ranchers in future study design and implementation. The advancement of these remote sensing technologies also hinges on further education and technical training opportunities for range management professionals, both government and private, who work directly with land managers and ranchers and can facilitate the data collection and fitting of custom biomass estimation models.

REFERENCES CITED

AgACIS, NOAA Regional Climate Centers, http://agacis.rcc-acis.org/.

Ajay Kumar Taloor, Drinder Singh Manhas, Girish Chandra Kothyari, Retrieval of land surface temperature, normalized difference moisture index, normalized difference water index of the Ravi basin using Landsat data, Applied Computing and Geosciences, Volume 9, 2021, 100051, ISSN 2590-1974, https://doi.org/10.1016/j.acags.2020.100051.

A.J. Rook, B. Dumont, J. Isselstein, K. Osoro, M.F. WallisDeVries, G. Parente, J. Mills, Matching type of livestock to desired biodiversity outcomes in pastures – a review, Biological Conservation, Volume 119, Issue 2, 2004, Pages 137-150, ISSN 0006-3207, https://www.sciencedirect.com/science/article/pii/S0006320703004518

Campbell, J. B., & Wynne, R. H. (2011). Introduction to Remote Sensing (5th ed.). New York, NY: The Guilford Press.

Costa-Saura, J., Balaguer-Beser, Á., Ruiz, L. A., Pardo-Pascual, J., & Soriano-Sancho, J. (2021). Empirical models for spatio-temporal live fuel moisture content estimation in mixed mediterranean vegetation areas using sentinel-2 indices and meteorological data. *Remote Sensing, 13*(18), 3726. doi:http://dx.doi.org/10.3390/rs13183726

Clevers, J.G.P.W., et al. "Estimating Grassland Biomass Using SVM Band Shaving of Hyperspectral Data." *Photogrammetric Engineering & Remote Sensing*, vol. 73, no. 10, 2007, pp. 1141–1148., https://doi.org/10.14358/pers.73.10.1141.

Davison, A.C. & Hinkley, D.V. (1997) Functions and datasets for bootstrapping from the book "Bootstrap Methods and Their Application" originally written by Angelo Canty for S.

Edirisinghee, A & Donald, G. & Henry, D & Mata, G. & Gherardi, Stephen & Oldham, C. & Gittins, S. & Smith, R. (2004). Pastures From Space – Validating Remotely Sensed Estimates of Feed On-Offer.

Fabricius, C., M. Burger, and P. A. R. Hockey. 2010. "Comparing Biodiversity between Protected Areas and Adjacent Rangeland in Xeric Succulent Thicket, South Africa: Arthropods and Reptiles." *Journal of Applied Ecology* 40 (2): 392–403. doi:10.1046/j.1365-2664.2003.00793.x.

Garonna, I.; Fazey, I.; Brown, M. E.; Pettorelli, N. Rapid primary productivity changes in one of the last coastal rainforests: the case of Kahua, Solomon Islands. Environ. Conserv. 2009, 36, 253–260.

Greenwood, Mark C. *Intermediate Statistics with R*. 3rd ed., 2021.

Hewins, D. B., Lyseng, M. P., Schoderbek, D. F., Alexander, M., Willms, W. D., Carlyle, C. N., . . . Bork, E. W. (2018). Grazing and climate effects on soil organic carbon concentration and particle-size association in northern grasslands. *Scientific Reports (Nature Publisher Group), 8*, 1-9. doi:http://dx.doi.org/10.1038/s41598-018-19785-1

Huete, A.R.; Justice C. MODIS vegetation index (MOD13) algorithm theoretical basis document. Ver. 3, 1999.

Jansen VS, Kolden CA, Schmalz HJ. (2018). The Development of Near Real-Time Biomass and Cover Estimates for Adaptive Rangeland Management Using Landsat 7 and Landsat 8 Surface Reflectance Products. *Remote Sensing*, 10(7):1057. https://doi.org/10.3390/rs10071057

Jansen, V. S., Kolden, C. A., Schmalz, H. J., Karl, J. W., & Taylor, R. V. (2021). Using satellite-based vegetation data for short-term grazing monitoring to inform adaptive management. *Rangeland Ecology & Management*, *76*, 30–42. https://doi.org/10.1016/j.rama.2021.01.006

Jansen, V., Traynor, A. C. E., Karl, J. W., Lepak,, N., & Sprinkle, J. (2022). Monitoring grazing use: Strategies for leveraging technology and adapting to variability. *Rangelands*, *44*(1), 64–77. https://doi.org/10.1016/j.rala.2021.07.005

JE Herrick, JW Van Zee, KM Havstad, LM Burkett, WG. Whitford. (2009). The Monitoring Manual for Grassland, Shrubland and Savanna Ecosystems, *Volume 2,* Design, Supplementary Methods and Interpretation. https://www.sciencedirect.com/science/article/pii/S0190052821000687#bib0032

John Fox, Sanford Weisberg (2018). Visualizing Fit and Lack of Fit in Complex Regression Models with Predictor Effect Plots and Partial Residuals. Journal of Statistical Software, 87(9), 1-27. doi 10.18637/jss.v087.i09

Jürgen Schellberg, Michael J. Hill, Roland Gerhards, Matthias Rothmund, Matthias Braun, Precision agriculture on grassland: Applications, perspectives and constraints, European Journal of Agronomy, Volume 29, Issues 2–3, 2008, Pages 59-71, ISSN 1161-0301, https://www.sciencedirect.com/science/article/pii/S1161030108000646

J. W. Rouse Jr., R. Haas, J. Schell, and D. Deering, “Monitoring vegetation systems in the great plains with erts,” NASA Special Publication 351, 309, 1974

Kamil Barton (2022). MuMIn: Multi-Model Inference. R package version 1.46.0. https://CRAN.R-project.org/package=MuMIn

Karnieli, Arnon, Nurit Agam, Rachel T. Pinker, Martha Anderson, Marc L. Imhoff, Garik G. Gutman, Natalya Panov, and Alexander Goldberg. " Use of NDVI and Land Surface Temperature for Drought Assessment: Merits and Limitations". *Journal of Climate* 23.3 (2010): 618-633. < https://doi.org/10.1175/2009JCLI2900.1>. Web. 14 Apr. 2022.

Landsat Missions. (2022). *Landsat surface reflectance-derived spectral indices*. Landsat Surface Reflectance-derived Spectral Indices | U.S. Geological Survey. Retrieved April 26, 2022, from https://www.usgs.gov/landsat-missions/landsat-surface-reflectance-derived-spectral-indices

Masek, J.G., Vermote, E.F., Saleous, N., Wolfe, R., Hall, F.G., Huemmrich, F., Gao, F., Kutler, J., and Lim, T.K. (2006). A Landsat surface reflectance data set for North America, 1990-100, IEEE Geoscience and Remote Sensing Letters. 3:68-72.

Matsushita, B., Yang, W., Chen, J., Onda, Y., & Qiu, G. (2007). Sensitivity of the enhanced vegetation index (EVI) and Normalized Difference Vegetation Index (NDVI) to topographic effects: A case study in high-density Cypress Forest. *Sensors*, *7*(11), 2636–2651. https://doi.org/10.3390/s7112636

Matthew O. Jones, Nathaniel P. Robinson, David E. Naugle, Jeremy D. Maestas, Matthew C. Reeves, Robert W. Lankston, Brady W. Allred, Annual and 16-Day Rangeland Production Estimates for the Western United States, Rangeland Ecology & Management, Volume 77, 2021, Pages 112-117, ISSN 1550-7424, https://www.sciencedirect.com/science/article/pii/S1550742421000464

Ogle, Dan, and Brandan Brazee. “Estimating Initial Stocking Rates.” *USDA - Natural Resource Conservation Service Technical Note*, vol. 3, June 2009, pp. 4–5., https://doi.org/https://www.nrcs.usda.gov/Internet/FSE_PLANTMATERIALS/publications/idpmstn9390.pdf .

QGIS Development Team. “QGIS Software.” *QGIS*, 3.12, QGIS Organization, 2022, https://www.qgis.org/. Accessed 11 Feb. 2022.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rahman, S. (2019). Change vector analysis, tasseled cap, and NDVI-NDMI for measuring land Use/Cover changes caused by a sudden short-term severe drought: 2011 texas event. *Remote Sensing, 11*(19), 2217. doi:http://dx.doi.org/10.3390/rs11192217

“Rangelands App.” *Rangeland Analysis Platform Version 3.0*, https://rangelands.app/.

“Rangelands Gateway.” *Vegetation Types on Rangelands* , Rangelands Gateway, 2022, https://rangelandsgateway.org/topics/rangeland-ecology/vegetation-types-rangelands

Rook, A.J., Dumont, B., Isselstein, J., Osoro, K., WallisDeVries, M.F., Parente, G., Mills, J., 2004. Matching type of livestock to desire

Rouse, J.W., R.H. Haas, J.A. Schell, and D.W. Deering, 1974. Monitoring vegetation systems in the Great Plains with ERTS, In: S.C. Freden, E.P. Mercanti, and M. Becker (eds) Third Earth Resources Technology Satellite–1 Syposium. Volume I: Technical Presentations, NASA SP-351, NASA, Washington, D.C., pp. 309-317.

Sarif, Md Omar, Bhagawat Rimal, and Nigel E Stork. "Assessment of Changes in Land Use/Land Cover and Land Surface Temperatures and Their Impact on Surface Urban Heat Island Phenomena in the Kathmandu Valley (1988-2018)." *ISPRS International Journal of Geo-information* 9.12 (2020): 726. Web.

Scurlock, J. M. O, and D. O Hall. "The Global Carbon Sink: A Grassland Perspective." *Global Change Biology* 4.2 (1998): 229-33. Web.

Segoli, M., E. D. Ungar, and M. Shachak. 2008. “Shrubs Enhance Resilience of a Semi-Arid Ecosystem by Engineering and Regrowth.” *Ecohydrology* 1 (4): 330–339. doi:10.1002/eco.21.

Ticman, K. V., S. G. Salmo III, K. E. Cabello, M. Q. Germentil, D. M. Burgos, and A. C. Blanco. "MONITORING POST-DISASTER MANGROVE FOREST RECOVERIES IN LAWAAN-BALANGIGA, EASTERN SAMAR USING TIME SERIES ANALYSIS OF MOISTURE AND VEGETATION INDICES." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.* XLVI-4/W6-2021 (2021): 295-301. Web.

Vermote, E., Justice, C., Claverie, M., & Franch, B. (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. Remote Sensing of Environment, 185, 46-56.

Wei Zhou, Haoran Li, Lijuan Xie, Xuemin Nie, Zong Wang, Zhengping Du, Tianxiang Yue, Remote sensing inversion of grassland aboveground biomass based on high accuracy surface modeling, Ecological Indicators, Volume 121, 2021, 107215, ISSN 1470-160X, https://doi.org/10.1016/j.ecolind.2020.107215.

Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, *2017*, 1–17. https://doi.org/10.1155/2017/1353691

Zach. “How to Calculate AIC in R (Including Examples).” *Statology*, 13 Mar. 2022, https://www.statology.org/aic-in-r/#:~:text=The%20AIC%20is%20designed%20to,AIC%20value%20of%20each%20model

Zhenhua Chao, Ning Liu, Peidong Zhang, Tianyu Ying, Kaihui Song, Estimation methods developing with remote sensing information for energy crop biomass: A comparative review, Biomass and Bioenergy ,Volume 122, 2019, Pages 414-425, ISSN 0961-9534, https://doi.org/10.1016/j.biombioe.2019.02.002.