Spring 2017

UAS-Collected Multispectral Imagery For The Identification Of Rangeland Vegetation In A Southern Mixed-Grass Prairie

Adam Rusk
Fort Hays State University, arrusk@mail.fhsu.edu

UAS-COLLECTED MULTISPECTRAL IMAGERY FOR THE IDENTIFICATION OF RANGELAND VEGETATION IN A SOUTHERN MIXED-GRASS PRAIRIE

Being

A Thesis Presented to the Graduate Faculty

of the Fort Hays State University, in

Partial Fulfilment of the Requirements for

the Degree of Master of Science

By

Adam Rusk

B.S., Fort Hays State University

M.S., Antioch University New England

Date: _____ Approved_____
Major Professor

Approved_____
Chair, Graduate Council

PREFACE

This thesis is written in the style of the Annals of the American Association of Geographers.

Keywords: unmanned aerial systems, supervised classification, rangeland, remote sensing

ABSTRACT

The interest in using unmanned aerial systems (UAS) for remote sensing of natural resources and ecology has grown rapidly in recent years and continues to develop. Recent improvements in the cost, size, and accessibility of consumer-grade UASs are now facilitating image collection on low-flying UAS platforms. Rangelands provide a unique opportunity to explore the uses of UAS remote sensing due to their large spatial extent and spatial and temporal heterogeneity. This study focuses on the use of UAS's to identify rangeland vegetation in a southern mixed-grass prairie in Kansas. Two primary questions were asked: 1) Do small UASs with low-cost sensors collect imagery useful for mapping rangeland plants? 2) Do different supervised classification techniques yield significant differences in their ability to classify rangeland vegetation? Data were collected over 100 acres of the Hadley Range in north-east Ellis County. Three separate modeling algorithms (Maximum Likelihood Classifiers, Random Forests, and Support Vector Machines) were compared to a random dataset to determine if imagery collected with UAS could identify land cover better than random assignment. While the classification algorithms did perform better than random in most regards, they did not perform sufficiently well as to replace, or even compare to field work. However, I expect that as technical, and technological improvement in spatial, spectral, and temporal resolution occur, UAS remote sensing's ability to aid in determining stocking rates for grazing, locating invasive or rare species, and estimating overall biodiversity will greatly increase.

ACKNOWLEDGEMENTS

I write this series of acknowledgements having now undergone the Master's Thesis endeavor more than once. Though my needs as a student have changed from my last thesis, the need for a support network and for guidance have not.

To Tom Schafer, my graduate advisor, for being supportive in my overambitious endeavors and for appreciating the interplay of my batting skills with his coaching skills.

To Rich Lisichenko for the expertise and the guidance for Python programming. Without his course and willingness to allow for expanded horizons, the amount of processing time required to conduct this work would have been impossible.

To Bill Stark for the equipment, the software, and the willingness to allow me to work with the UAVs here at FHSU. I started dreaming up a thesis before I knew it was possible, and you made it possible. For that, I am grateful.

To Rob Channell for being my sounding board when methods got weird and for being the catalyst of the whole mess that is my perspective on ecology. Thank you for introducing me to conservation, ecology, quantitative research (to a fault most likely), and for helping me through the tough analytical parts of this project.

To Dalton Hills and Sean Rogers. Two graduate students who I spent a lot of time working with and without whom, the work for this project would certainly not have been completed.

To Mitch Greer and the FHSU Range Management class of 2016 for working so diligently to collect the field data used in this thesis. A lot was asked of you and you delivered with flying colors. Thank you for making this process easier.

To my family and friends. In particular, my parents and siblings whom I neglected to mention previously but who likely deserve the most credit. They were with me when my interests in science began, and will be with me until the end. Also to my wife Bailey, who sees the daily struggle and is always supportive and understanding.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Sampling vegetation in an effort to quantify structure, function, and composition is a fundamental means of assessing ecological differences. Often, studies focused on the management of rangelands, which require our ability to quantify habitat, include a diversity of flora and fauna. Considerations of time and money tend to influence the quality and scope surveying vegetative distributions using traditional methods (Cruzan et al. 2016). However, were resources unlimited, the amount of disturbance that would occur when intensively sampling a rangeland negatively impacts the ecological system. Remote sensing adds value to ecosystem research in that time, money, and effort are greatly reduced. Unfortunately, traditional methods of remote sensing often lack sufficient resolutions when the ecosystem is heterogeneous. The use of UASs in heterogeneous ecosystems, such as rangelands, may be used to bridge this gap between insufficient remote sensing resolutions and practical inhibitors by collecting high spatial and temporal resolution data on difficult to survey ecosystems.

The interest in using unmanned aerial vehicles (UAS) for remote sensing of natural resources and ecology, such as vegetation cover, ecosystem health, and wildlife tracking, has grown rapidly in recent years and continues to increase. Colomina and Molina (2014) discuss an increase in the prevalence of unmanned aerial vehicle (UAS) and drone online research from 2005 to 2013. Traditionally, aerial photography collected using piloted aircraft provided data for rangeland monitoring and mapping. However, higher resolution imagery is often required for characterizing and quantifying finer spatial patterns and processes within a rangeland. Recent improvements in the cost, size, and accessibility of consumer-grade UASs are now facilitating image collection on low-

flying UAS platforms. These platforms are capable of collecting sub-decimeter resolution (as low as 1 cm) imagery at a lower cost, at an higher temporal frequency, on shorter notice than piloted aircraft, and are safer than manned aircraft (Laliberte and Rango 2011; Ishihama, Watabe, and Oguma 2012).

Rangelands provide a unique opportunity to explore the uses of UAS remote sensing. Rangelands are defined by the Society for Range Management as any land in which the climax vegetation is primarily grasses, forbs, or shrubs and is managed as a natural ecosystem, where fire, drought, and grazing inhibit the encroachment of succession to savannahs or forests. Ranges comprise nearly 40% of the United States and between 50 and 70% of the world (Rango et al. 2009), and often vary in expanse, ranging from fractions of a hectare, to thousands of hectares. They are spatially and temporally diverse, where an individual of a given species of vegetation may occupy no more than 10 cm^2, and species prevalent in the spring may not be prevalent in fall. A diverse assemblage of animal species rely on grassland or shrubland as habitat, from amphibians (green toad) and reptiles (longnose snake), to birds (greater prairie chicken) and hooved mammals (pronghorn). This includes livestock which humans manage for our own benefit. Researchers and land managers endeavor to quantify rangeland habitats to discern suitability for rangeland obligate species, monitor ecosystem health, and calculate stocking rate for livestock. The incorporation of low-altitude remote sensing using UASs may facilitate sampling in these vast ecosystems.

Prior work in the field of rangeland UAS remote sensing has focused on creating data collection workflows and collecting imagery using expensive equipment in

moderately uniform environments. This study aims to expand on prior research conducted primarily in New Mexico (e.g. Laliberte and Rango 2008) with intent to identify rangeland vegetative species in Kansas' southern-mixed grass prairie. Specifically, the following questions will be addressed: 1) Do small UASs with low-cost sensors collect imagery useful for mapping rangeland plants? and 2) Do different supervised classification techniques yield significant differences in their ability to classify rangeland vegetation?

A review of the literature provides context to this research and validates the research questions. Discussed below is a background of how remote sensing applies to ecological systems and vegetation, a history and description of UASs in rangeland research, and modes of image classification and how to assess their accuracy.

LITERATURE REVIEW

Context

Remote sensing is a valuable tool for collecting large amounts of data in a relatively short period of time. Satellites can capture images with spatial extents thousands of kilometers across (MODIS specifications 2330 km), though their spatial resolution, or size of a pixel, is poor. Conversely, near-Earth sensing platforms such as unmanned aerial vehicles can obtain images with a spatial extent the width and length of a hectare or smaller, with very fine resolution (decimeters to millimeters). Depending on the scale of a research project, both methods are valuable. Climate can successfully be modeled with a resolution of 1 km^2 and ecosystems, for example the difference between forests and grasslands, can often be separated using 30 m x 30 m (900 m^2) pixels. However, a 30 m x 30 m pixel is insufficient when looking at a rangeland for feeding cattle.

UAS remote sensing is being used worldwide to survey animals (Watts et al. 2010), identify invasive plants (Mitchell et al. 2012), locate and remove weeds (Xiang and Tian 2011; Gini et al. 2012; Lopez-Granados et al. 2015), and map fine-scale degradation of habitat (Mansour et al. 2015). Small unmanned systems or platforms are useful, because they are able to sample expansive and remote locales with minimal time commitment, and require minimal direct human contact with the ecosystem (Watts et al. 2010). These attributes are especially valuable when trying to monitor patches of vegetation across the large spatial extent of rangelands.

The identification of small objects, such as bunch grasses or individual forbs, requires the use of high spatial resolution imagery. Though high spatial resolution

imagery can be collected using satellites, with resolutions as low as meters or decimeters, on the ground or near-Earth sensors, such as those mounted on UASs are often of higher resolution and lower cost. While the use of fine spatial scale (meters to decimeters) remote sensing in precision agriculture dates back the 1980s, UAS-based remote sensing for rangelands remains relatively new (Mulla 2012). These ecosystems comprise 37% of the United States and up to 70% of the world's ecosystems (Lund 2007; Breckenridge and Dakins 2011). Understanding rangeland species assemblages and how those assemblages provide wildlife habitat, forage, and affect nutrient and hydrological cycles is essential for ecosystem research (Lauenroth, Burke, and Gutmann 1999), biodiversity/conservation (Patterson and Best 1996), and rangeland management (Ohlenbusch and Watson 1994). Work conducted at Jornada Experimental Range in New Mexico has yielded workflows for processing, orthorectification, and classification of arid rangeland ecosystems for assessment, monitoring, and management (Laliberte, Winters, and Rango 2008; Rango et al. 2009; Laliberte et al. 2010; Laliberte et al. 2011; Laliberte and Rango 2011; Laliberte, Winters, and Rango 2011).

Future research on identifying rangeland vegetation will require technological improvements, cost efficiency, and creativity in the field of remote sensing in order to effectively identify and map the distribution of vegetation. Knowledge of vegetative characteristics and phenology will be essential in determining what attributes of a plant can be measured and identified. The skill to use remote sensing software will be indispensable for classifying spectral characteristics of vegetation into something meaningful to data users.

Remote Sensing of Vegetation

Remote sensing of vegetation is possible because of the unique reflectance and absorbance properties of plant tissues. Vegetation reflectance is determined in part by foliar optical and biophysical properties (Asner 1998). Compounds, such as lignin and foliar nitrogen, can be measured using reflectance data and have been used classifying different tree species (Fourty et al. 1996; Martin and Aber 1997; Martin et al. 1998). Other studies have used remotely-sensed data to determine protein and polyphenol concentrations in trees and grasses as indicators of soil nutrient concentrations (Skidmore et al. 2010). Mutanga and Skidmore (2007) demonstrated that the shifting red edge, or the wavelength at which chlorophyll stops absorbing red light and starts reflecting infrared light, is a good indicator of nitrogen in grasses. Foliar nitrogen, a common compound in proteins and chlorophyll, is a primary determinant of the nutritional value of forage (Van Soest 1994).

While exact measurements of foliar nitrogen are valuable in remote sensing of vegetation, such measurements require more advanced sensors than are currently available for UASs. Given that high spectral resolution data is unavailable, researchers are forced to find innovative ways to use minimal information. Typically, when discussing remote sensing of flora, vegetation indices (VIs) are an important tool to consider and can be useful in developing measuring productivity of plants as a proxy for health or nitrogen content. However, VIs are beyond the scope of this review. Another creative method discussed in Laliberte and Rango (2008) involves incorporating texture,

intensity, hue, and saturation into species identification. These four measurements can be calculated using three-band imagery and improve rangeland species and other taxonomic identifications (Laliberte and Rango 2008). Innovations in spectral analysis methods, in addition to the high spatial resolution of UAS-collected imagery, provides managers with affordable tools to monitor and map rangelands.

UAS Flight and Data Acquisition

The recent increase in UAS popularity among civilians is approximately a decade old. As early as 2004, military UASs outnumbered civil and civilian UASs 50 to 1 (Rango et al. 2010). Since then the number of civilian-owned light UASs (under 150 kg) has increased, the use of UASs is expected in natural resource sciences (Rango et al. 2010). The price of current systems, including fixed-wings, helicopters, and multi-rotor copters, ranges from as little as $2,500 with minimal equipment to greater than $350,000 (Rango et al. 2006; Rango et al. 2009).

When determining the feasibility for UAS use in rangeland remote sensing, there are some important caveats to consider when compared to aerial photography via manned aircraft. While images collected on low-flying platforms have a higher spatial resolution than aircraft or satellite acquired data, UAS imagery often has low spectral and radiometric resolutions (Laliberte and Rango 2011). Additionally, UAS platforms are often less stable than larger aircraft, which makes the task of aligning the images correctly and integrating them into a single, cohesive image (registration and mosaicking) difficult; the large number of small images collected further complicates this challenge.

Similar to flying manned aircraft, the Federal Aviation Administration (FAA) requires a special certification (Part 107) to operate UASs commercially. These obstacles suggest that implementing UASs in rangeland research may not always be feasible. However, where possible, the use of UAS-mounted sensors continues to provide unique opportunities for rangeland data collection.

Additional considerations when using UASs include flight time and duration, equipment and logistics, and emergency planning. Flight time should range from two to six hours, and flights should occur between 1000 and 1500 hours to take advantage of direct sunlight (Rango et al. 2006). Ground control stations, such as ArduPlane, Horizon, and Norut Geo Viz, allow for the creation of waypoints to facilitate automated mission planning (Stodle, Borch, and Storvold 2013). These stations also allow user control over height of flights, image overlap, and speed of flights. Height determines the spatial extent and pixel resolution of the images, image overlap affects three-dimensional surface models and averaging of reflectance values in pixels, and speed of flights determines the amount of time spent in the air based on battery life. FAA regulations, payloads, sensors, and navigation are also important aspects of UAS flight to consider but are beyond the scope of this literature review (Colomina and Molina 2014).

Once data have been collected the imagery needs to be processed. Tie points, points/pixels common on multiple images, are used to georeference images with useful coordinate systems using software, such as Autopano (Kolor; Francin, France) using structure from motion techniques (Laliberte et al. 2011; Turner, Lucieer, and Watson 2012). Georeferencing places all images in appropriate orientation to one another, and in

geographical space. With multispectral imagery, radiometric correction may also be necessary (Laliberte and Rango 2011; Kelcey and Lucieer 2012). Once images have been appropriately aligned in space they are merged together (mosaicked) and corrected for topography (orthorectified) using software such as PreSync, Pix4D, or Agisoft Photoscan (Laliberte, Winters, and Rango 2008). After images have gone through alignment, georeferencing, mosaicking, and orthorectification, classification can begin.

Classification

After processing high-resolution imagery to generate a single, cohesive image, one common analytical goal is discerning vegetative cover or structure from the data. Classification is the act of categorizing all pixels in an image into groups or classes (Lillesand, Kiefer, and Chipman 2008). For the purpose of this review, only pixel-based supervised and unsupervised methods, and object-oriented classification methods are considered.

Pixel-based classification uses within-pixel spectral patterns to classify pixels. Pixels with similar digital number values, or combinations of values, are placed into categories. User input guides supervised pixel-based classification by specifying land cover classes of interest. The user delineates cover classes of interest on the image, often based on field data, and the program generates a key using attributes of each class. After building the key, attributes of other features are classified based on the categories in the key. Unsupervised pixel-based classification does not use training areas and instead classifies all pixels into a preselected number of clusters, which the analyst must then

combine into meaningful groups (Lillesand, Keifer, and Chipman 2008). Mitchell et al. (2012) demonstrated that different types of pixel-based classification provide adequate identification of vegetation.

Though historically pixel-based classification was dominant, object-based classification methods are becoming more common due to the increasing availability of high-resolution imagery (Hay et al. 2005). Object-based image analysis (OBIA) builds on older segmentation, edge detection, feature extraction, and classification concepts (Blaschke 2010). OBIA requires segmentation, defining of attributes, and the ability to connect objects in space and time (Hay and Castilla 2006). Segmentation is the process of partitioning an image into segments or objects. A group of pixels may represent a single object on the ground, and OBIA is used to classify that single object as the same category rather than each pixel separately. Using color, shape, size, texture, pattern, and context, OBIA overcomes intra-class spectral variability issues with high-resolution imagery by classifying after segmenting (Yu et al. 2006; Lillesand, Keifer, and Chipman 2008). As an example, with high resolution imagery, a single yucca (*Yucca glauca*) plant may be represented by hundreds of pixels. Some pixels may represent shadow, others may represent leaves or spaces between leaves. Using traditional pixel-based methods, each pixel would be classified differently. Segmentation would group all pixels of the yucca into a single object, then OBIA would classify the object as a whole. Many studies led by Laliberte, Rango, and Yu suggest that OBIA outperforms pixel-based classification methods.

Assessment is an essential component in determining the accuracy of classification. An error matrix is an effective tool for determining user, producer, and overall accuracy (Congalton and Green 2009). Producer's accuracy is a measure of omission errors and indicates how well an individual area can be classified. User's accuracy is a measure of commission errors and indicates the probability that a classified pixel on a map is representative of the material on the ground. Overall accuracy is a ratio of correctly identified pixels to the total number of pixels, and is reported as percent correct (Congalton and Green 2009). These are a few examples of many classification matrix metrics. Approximately fifty samples per category should be used to construct an efficient error matrix for accuracy assessment (Congalton 1991).

Summary

To be effective, or widely used, UAS applications for rangeland rely heavily on remote sensing techniques of vegetation, skills in UAS mission planning and deployment, imagery processing, and classification. However, the ability to identify vegetative species on hundred-hectare to tens-of-thousands-hectare properties to inform conservation and management is valuable. Identification of rangeland plants using UASs allow scientists and managers to identify rich soil locations, manage habitat for biodiversity, and calculate forage for stocking rates at a greater temporal frequency, finer spatial scale, and larger spatial extent than previously possible.

METHODS

Site Description

Hadley Range is a 1214-hectare (3000-acre) grassland in north-central Ellis county (Fig 1). The site has been actively grazed since prior to 1900 and is grazed by upwards of 500 cattle annually. In the 1940's oil was discovered on the property, and there are approximately 50 wells in production. The property is topographically diverse, comprised of lowlands, riparian areas, bluffs, and highlands. Within Hadley Range, a 40-hectare (100-acre) area was delineated as the study area for this project (Fig. 2). Vegetative communities in this study area consisted of ragweed-tall dropseed-sagewort lowlands (*Ambrosia psilostachya*, *Sporobolus compositus*, *Artemisia ludoviciana*, respectively), bluestem-yucca-grama grass uplands (*Schizachyrium//Andropogon spp.*, *Yucca glauca*, and *Bouteloua spp.*, respectively), and woody ravine areas. Other species common to the study area are Maximilian sunflowers (*Helianthus maximilianii*), Pitcher's sage (*Salvia azurea*), fragrant and smooth sumac (*Rhus aromatic* and *Rhus glaborous*), and rigid goldenrod (*Solidago rigida*).

Unmanned Aircraft and Imagery Collection

Aerial photographs were collected using a 3D Robotics (3DR) Iris+ (3D Robotics; Berkeley, CA) (Fig. 3). The Iris+ is a light-weight quadcopter weighing 1.2 kg including the lithium polymer (LiPo) battery and is 50 cm from motor to motor. A Pixhawk (3D Robotics; Berkeley, CA) autopilot system ensures the quadcopter can be launched from any flat surface and fly a predetermined path autonomously.

Flight plans were generated using Mission Planner (Version 1.3.40, DIY Drones; Berkeley, CA), an open-source ground station software. Waypoints, or coordinates for directing flight paths, were chosen using a map interface to generate flight lines that cover the entire study area (Fig. 4). Missions were planned to generate images with 80 percent front and side-lap where 80 percent of an image is also represented in the four images to the front, back, and sides of that image. Overlap is used to account for the effects of bidirectional reflectance, or the change in perceived reflectance based on the angle of image capture. Due to limited battery life and therefore flight time, the study area was divided into 10 equal sections. Each section was sampled on the same day between 10:00 am and 3:00 pm. The UAS carried a single sensor, a modified Sony RX 100 (Sony Corporation; Minato, Tokyo, Japan). Specific wavelength filters modified the sensor to collect imagery in the blue (approx. 400-490 nm), green (approx. 490-540 nm), and near-infrared (approx. 700-750 nm) spectral bands. Flying 6 m/s and 100 m above ground surface, the spatial resolution of the RX 100 was 9 cm^2.

Two series of flights were conducted in early September. Data collected during the first flight were considered the "full data set" with the specifications discussed above (Fig. 5). After deploying Daubenmire quadrats for field sampling, a second set of flights, known as the "reference data set", was conducted to allow for easier digitization of vegetation within the quadrats (Fig. 6). This data set was collected with reduced image overlap, which decreased collection time and a reduced the number of files to accommodate time restraints. Long periods of time between data collection events would

increase the chance of encountering inclement weather or cattle disturbance to the quadrats.

Field Measurements

I used a modified Daubenmire quadrat method to assess vegetation composition and arrangement (Daubenmire 1959). One hundred and ten, 1 m^2 quadrats were systematically placed across the study area using ArcGIS 10.4.1 (Esri; Redlands, CA) (Fig. 7) prior to the second data collection flights. Eighteen students of the Fort Hays State University Fall 2016 class on Rangeland Management collected the vegetative data by estimating percent cover of species based on Daubenmire's cover classes, drawing the distribution of plants within the quadrat on a standard form (Fig. 8), and taking a photo of the quadrat (Fig. 9).

Image Processing

Images captured in UAS flights had to be processed through a series of steps to provide the necessary quantitative information for classification. The workflow for this processing consisted of georeferencing, orthorectification, mosaicking, and calculation of vegetation indices. Images collected in the Sony RX 100 contained no spatial reference. Mission Planner was used to assign each image a geographic coordinate using time and telemetry data. Based on the time on the camera and the flight log, which recorded the each time the camera was triggered and the GPS location of the quadcopter, Mission Planner assigns a coordinate to each photograph. After images were geotagged, or assigned a coordinate, they were imported into Agisoft Photoscan (Agisoft LLC; St.

Petersburg, Russia) and aligned. Aligning places all images in correct position with respect to orientation, other images, and location on Earth's surface. A preliminary texture, or artificial merging of the images, was generated to allow for extended georeferencing. With Agisoft, five georeference points were used to further georectify the images based on field markers (Fig. 10). Once the images were realigned using the additional georeferenced data, they were combined into an orthomosaic, or a single image of the study area corrected for topography. This process was repeated for both the final data mosaic and the reference mosaic. The orthomosaics for both data sets were exported to ArcMap, and a final georectification was performed using a spline technique to ensure the images aligned.

Vegetation indices were calculated in ArcMap. Given the limited number of spectral bands available, only the green normalized difference vegetation index was calculated (GNDVI; Gitelson, Kaufman, and Merzylak 1996). Since classifications are performed on a single, multiband raster, GNDVI was stacked with the other three spectral bands for analysis.

Vegetation Digitization

The purpose of digitizing vegetation patches is to provide known species for the supervised classification. Field drawings and photographs provided the species identification and the proportion of quadrats occupied by those species. The vegetation in each quadrat was digitized by comparing the reference mosaic raster, field data forms, and field photographs. Polygons were used to delineate vegetation regions of interest, or

areas on the image where the species of vegetation is known, and were drawn in ArcGIS Pro 1.3. Laliberte and Rango (2011) found that digitizing vegetation patches directly on an image removed GPS inaccuracy, allowed for larger sample sizes, and reduced time and effort. Summaries of the field collected data indicated that approximately ten species were more abundant than all others. These ten species, along with classes for trees, easily identifiable shrubs, bare ground, litter, rock, and water were used in drawing and categorizing each polygon. After digitizing all quadrats on the reference image, 17 cover classes were represented by 286 polygons.

Classification and Analysis

The final data orthomosaic was segmented using ArcMap prior to classification (Fig. 11). Segmentation used mean, standard deviation, compactness, rectangularity, color, and count to generate "objects". Objects are groups of pixels that are clumped based on the above criteria. In OBIA, objects are classified rather than each pixel within. Spectral classification techniques are often insufficient with high-resolution imagery due to intra-object variation (Blaschke 2010). For example, shadows between leaves or leaves at different angles will yield slightly different spectral signatures. This error is reduced by segmenting pixels into objects before classification.

I iteratively executed three separate classification algorithms using random subsets of the vegetation polygons in Python 2.7 (Appendix 1). Eighty percent of the polygons were used to construct each model. Twenty percent of the polygons were reserved to test each classification's accuracy. It is important to note that, while the

polygons were digitized using the reference data set, the data used for classification was attained by extracting information from the segmented final data set. One hundred iterations of Random Forests (Breiman 2001), maximum likelihood, and Support Vector Machine (Cortes and Vapnik 1995) classifications were executed using different training subsets and the segmented final data set. Examples of each classification can be seen in Figure 11. Classifications from the classified image were extracted to the testing polygons and the attributes were exported as a comma separated values (CSV) file.

To evaluate the performance of the three classification algorithms listed above, I executed a random algorithm. The classifications of all 286 vegetation polygons were iteratively randomized with replacement. This processes randomly assigns a predicted class to each polygon which is then compared to the true, known class. Similar to above, randomization was executed 100 times.

The files containing the true and predicted values of each vegetation type were imported into R (R Development Core Team 2008). Confusion matrices and performance metrics were calculated using package "rminer" (Appendix 2) for each model (Cortez et al. 2010, Cortez 2016). The metrics were as follows: overall accuracy, classification error, balanced error rate, kappa index (Cohen 1960), Cramer's V (Cramer 1946), macro-sensitivity (the average of sensitivity), macro-specificity (the average of specificity), macro-precision (the average of precision), macro-F1 (the average of F1) (Powers 2011), and macro-MCC (the average of Matthews correlation coefficients) (Matthews 1975).

Due to collinearity of performance metrics, the final set of response variables included overall accuracy, Cramer's V, macro-sensitivity, and macro-specificity. The

data were evaluated for the assumptions of a parametric multivariate analysis of variance (MANOVA). A lack of multivariate normality and apparent heteroscedasticity of the data necessitated a non-parametric approach. A permutational analysis of variance on distance matrices was conducted using the adonis function in the vegan package of R to identify if a difference between model type and performance metrics existed (Anderson 2001, Oksanen et al. 2017). While the adonis method may produce significant results due to within group variation rather than between group variation, it appears to be more robust than other nonparametric MANOVA alternatives such as multiple response permutations procedure (MRPP) (Anderson 2001; R package vegan).

RESULTS

Of the 56 different cover types identified through direct observation in the field, only 17 were sufficiently identifiable from the aerial imagery to use in classification (Table 1). The resulting 286 cover polygons were iteratively divided into groups of 228 for model training and 57 for model testing. Completion time of the classifications varied in duration based on the model. Maximum Likelihood Classifiers and Random Forests both averaged 40 minutes per iteration while Support Vector Machines took approximately 16 hours to execute each iteration.

Evaluation of confusion matrices indicated that Random Forests and Support Vector Machines successfully removed categories with insufficient sample sizes, while Maximum Likelihood Classification maintained and predicted across all categories of cover (Table 2-4). Of the 5700 (100 iterations with 57 testing samples each) testing samples in the Maximum Likelihood Classifications, 2448 were misclassified as smooth sumac. Only buffalo grass was classified correctly more often than misclassified when maximum likelihood was the classifier (Table 2). The Random Forests classification algorithm most often correctly classified little bluestem, tall dropseed, Louisiana sagewort, bare ground, buffalo grass, rock, water, and trees. Sideoats grama was most often misclassified as buffalo grass, litter and yucca were most often misclassified as little bluestem, and all others were most often misclassified as a tree (Table 3). When comparing the classifications of the Support Vector Machine algorithm, little bluestem, tall dropseed, western ragweed, Louisiana sagewort, bare ground, buffalo grass, rock, water, and trees were most often correctly classified. Yucca, smooth sumac, and litter were most often misclassified as little bluestem. All others were most often misclassified

as a tree (Table 4). The random model misclassified nearly all classes as western ragweed more often than all other species.

When comparing the final set of classification matrix metrics, Support Vector Machines had a higher average overall accuracy, Cramer's V, and macro-sensitivity than the other scenarios. The random model showed a higher mean with regards to macro-specificity (Table 5).

The test methods, which were compared using a permutational multivariate analysis of variance using distance matrices (adonis), showed a significant difference in model performance ($F_{3, 398} = 103.9$, $p = 0.001$). Due to the significant result of the multivariate test, individual Kruskal-Wallis tests were performed for each response variable to evaluate significant differences and significant results were followed by multiple comparison tests. Overall accuracy differed significantly between the algorithms ($X^2 = 249.0$, $df = 3$, $p < 0.001$). Cramer's V also differed significantly between the classification algorithms ($X^2 = 240.1$, $df = 3$, $p < 0.001$). The Kruskal-Wallis test of macro-sensitivity also indicated that the four different models differed significantly ($X^2 = 227.6$, $df = 3$, $p < 0.001$). Finally, macro-specificity indicated a significant difference between the classification types ($X^2 = 142.4$, $df = 3$, $p < 0.001$).

Given that each Kruskal-Wallis test indicated a significant difference between at least two of the modeling algorithms, nonparametric Tukey's Honestly Significant Difference (Tukey's HSD) tests were performed. Support Vector Machines and Random Forests had significantly different overall accuracies from Maximum Likelihood or random models but were not different from each other (Table 6; Fig. 12). Maximum

likelihood, Support Vector Machines, and Random Forests cannot be statistically differentiated with regards to Cramer's V, but were all significantly different than the random model (Table 6; Fig. 12). The random models showed significantly different macro-sensitivity than the other models, and Support Vector Machines and Random Forests performed significantly different than Maximum Likelihood (Table 6; Fig. 12). Finally, Support Vector Machines and Random Forests performed as well as each other and significantly different from Maximum Likelihood with regard to macro-specificity. The random models performed significantly different from all other models when measured by the same metric (Table 6; Fig. 12).

Due to the nonparametric nature and complex series of steps in these analyses, determination of which initial predictor variable was the most important in identifying vegetation or cover is extremely difficult. Since the output of the initial classifications did not provide variable performances, the importance of each variable cannot be discussed.

DISCUSSION

This study expands on a body of research that uses low-flying UAS to differentiate among types of ground cover (Laliberte, Winters, and Rango 2008; Rango et al. 2009; Laliberte et al. 2010; Laliberte et al. 2011; Laliberte and Rango 2011; Laliberte, Winters, and Rango 2011). In particular, little effort has been made to identify vegetative species in a southern mixed-grass prairie. The cost, quality, and availability of equipment and the heterogeneity of the ecosystem studied provide unique and complex challenges to this research. Below, I evaluate each performance metric with regards to all model scenarios before summarizing common patterns and overall trends.

Overall Accuracy

Accuracy as calculated here is the total number of correct classifications divided by the total number of classifications. This includes true positive and true negative rates. All models should, theoretically, outperform a random scenario in overall accuracy because they are constructed with parameters used to supervise classifications. Maximum Likelihoods inability to perform better than the random classifications suggests poor data or an insufficient modeling algorithm. As previously mentioned, Maximum Likelihood was unable to vet, or identify and remove, underrepresented categories and therefore showed a high number of misclassifications, such that the overall accuracy was worse than random. However, this was not the case with Random Forests and Support Vector Machines, which were able to reduce the number of predicted cover types to those most likely and those best represented.

Random Forests and Support Vector Machines had significantly higher overall accuracies than both the Maximum Likelihood and random scenarios, though they were not statistically different from each other. Both algorithms created rules for splitting cases into a predefined group, and the method for doing such is similar in both algorithms. Therefore, it is not surprising that the results were similar and fine adjustment of each model's parameters in a more in depth study of model classification ability might be necessary to identify differences in Random Forest and Support Vector Machine performance.

Cramer's V

Based on the multiple comparison test there is no statistical difference among the performance Maximum Likelihood, Support Vector Machines, or Random Forests regarding Cramer's V. Though it can be argued that Support Vector Machines outperformed the Maximum Likelihood Classifier, there is no difference between Support Vector Machines and Random Forests. Given that Random Forests did not perform significantly different than Maximum Likelihood, the three model algorithms cannot be statistically differentiated. However, all three classifiers outperformed the random model.

It is also important to note the variability in model performances. Unlike the other performance metrics where the random scenario appears to have the only significantly different variance, there is marked difference between variances of all classifiers with regards to Cramer's V. Maximum likelihood appears to be the least consistent, followed

by Random Forests, and Support Vector Machines. The distribution of Cramer's V for the random scenario is relatively small. This phenomenon will be further discussed in the random models section.

Macro-Sensitivity

Random Forests and Support Vector Machines were statistically equivalent when evaluated by macro-sensitivity. Maximum likelihood had variation to the other models but overall performed poorly. The random model had the lowest performance in macro-sensitivity, suggesting that the true positive rate of random classifiers was almost never above ten percent.

Overall, the best models when comparing sensitivity showed a true positive rate of 50% while the average rates were typically below 20%. This suggests all models were poor at correctly identifying cover type. This is contrasted by each model's ability to identify what a cover type should not be classified as, or macro-specificity.

Macro-Specificity

Macro-specificity is the only performance metric in which the random model outperformed the other classification algorithms. As with accuracy and macro-sensitivity, Support Vector Machines and Random Forests were not statistically different from one another, though both algorithms outperformed the Maximum Likelihood Classifier. The average true negative rate for all scenarios was greater than 92% suggesting that all models were sufficiently able to identify what a given cover type was not. As an

example, all models could classify a portion of the study area as not being big bluestem when it was not big bluestem. A benefit to high specificity is the reduction in type I error rate, though the type II error rates might be relatively high.

The difference in mean specificity between all models was less than two percent. Significance was most likely the result of the relatively large variation in the three modeling algorithms compared to the small variation in the random scenario.

Random Models

The construction of the classification matrices for the random models involved iteratively randomizing the entire training data set. Rather than calculating metrics based on 100 sets of 57 values as was the case in the training subsets of the other three models, performance metrics were calculated based on 100 sets of 258 classifications. Though the random model was bootstrapped, or randomized with replacement, the overall sample sizes likely reduce the variation in performance metrics. This is a confounding factor when comparing model performances. Had the data been randomly subsetted for the random model as it was for all other models, the performance of the random model would likely have decreased.

Quantity of vegetation polygons, per species, used to build the model also confounded classification performance. Of the 286 polygons used in classification, 238 were represented by six cover types. Two classes, western ragweed and little bluestem, account for 42% of all samples. When randomly subsetting data, these six classes were the most likely to be chosen, with few polygons representing the other classes. In this

case, the type II error rate was greatly inflated because incorrect classifications become more prevalent. However, this also increases the number of instances where, for each cover type, the true condition of any cover type was not classified as that specific cover type. For example, for every polygon that was not little bluestem, and was classified as not being little bluestem (regardless of how it was classified), the specificity improves. An instance where sensitivity and specificity were both high would suggest a low error rate and high accuracy, though this was not the case for the random data. For this reason, despite having the highest specificity, the random data model was not better than any other model.

Overall Trends

When looking at all four performance metrics, the random scenario outperformed all other models with regards to only one, though its specificity was only slightly higher than the other models. Maximum Likelihood Classification was generally worse than the other two classification algorithms, except in the case of the Cramer's V statistic. Random forests and Support Vector Machines were never statistically different but outperformed the other two methods half the time. Based on these four performance metrics, these two algorithms were the best methods for classifying cover in this study.

As mentioned earlier, execution of the Maximum Likelihood Classifiers and the Random Forests classifiers took significantly less time than the Support Vector Machines. The nature of machine learning algorithms requires longer processing time (Huang et al. 2002, Nitze et al. 2012). Interestingly, this increase in processing time was

not apparent in the Random Forest algorithm. Contrary to Nitze et al. (2012), Random Forests were computationally less costly than Support Vector Machines and yielded similar results. Based on an OBIA demonstration by Esri professionals (2016 User Conference), I expected Support Vector Machines to outperform all other algorithms. This was not the case. A combination of performance quality and execution time suggest that Random Forest is the preferable classification algorithm for conducting object based image analysis with high resolution, heterogeneous data.

CONCLUSION

Conclusions for this Study

The purpose of this study was two-fold: to determine if low-cost sensors and platforms collect useful data for mapping rangeland vegetation and to identify differences in model performance when attempting to classify rangeland vegetation. In this instance, low-cost sensors yielded 3 cm (9 cm^2) resolution imagery in only three spectral bands. When classifying this imagery based on intensive field methods, cover was accurately classified, on average, only 30% of the time. The models with the best performance exhibited a 53% accuracy. While this is significantly better than random, these results do not suggest that low-cost platforms and sensors are better than rigorous field work. While some time and money might be saved in the long run, this study conducted more field work than the average study of a similar magnitude, and yielded less-than-satisfactory classification results.

Despite these models' inability to adequately map rangeland vegetation accurately, I was able to detect differences in model performance. Random forests and Support Vector Machines were indistinguishable statistically, but Random Forest models executed in a fraction of the time compared to Support Vector Machines. For this reason, given the structure and limits of this study, Random Forests are the best modeling algorithm of the three tested for classifying rangeland vegetation.

Methodological Considerations

In reviewing sampling and analysis methods, there are several aspects of this study that could be improved without expanding or drastically changing the methods. When conducting field sampling, prior knowledge of vegetative communities would facilitate a stratified random sampling strategy. Stratifying the quadrats in the field would yield more even count of samples of each cover class and hopefully improve classification accuracy. This could also be achieved by removing all but the most dominant cover types rather than using all cover types discernable from the aerial imagery.

Similar to the issue stated above, the sample size of the random scenario yielded some confounding results. By reducing the sample size to the same as the other models, comparisons of the four scenarios might yield more realistic results. Similarly, an exploration of R packages, and a workflow facilitating the different stages of Object Based Image Analysis using available R packages might allow for execution of classifications in the R environment, which would allow for more flexibility and control than is available in ArcGIS.

Finally, this study was initially intended to have been conducted using a fixed wing platform. Due to technical limitations, a quadcopter was used instead. Use of a fixed wing would have allowed for faster data collection which would have reduced the effect of sun movement throughout the data collection period. Though all data were collected in an acceptable time span, between 10 AM and 3 PM some spectral streaking was noticed along flight paths. Though the streaking did not appear to affect the

classifications, a reduction in time between the first and last flight should improve the quality of the imagery.

Future Work

The best ways to improve the results of this study, despite those covered in the methodological considerations, involve improvement in three of remote sensing's resolutions: spatial resolution, spectral resolution, and temporal resolution. The short duration of the available field season and the available technology limited the above resolutions. However, each can be improved in future studies.

Vegetation phenology and varying appearance through the year could improve identification of rangeland vegetation. Nonnative, cool season, and warm season grasses all green-up and senesce at different times, as do forbs and woody plants. Presence of flowers, if large enough, could also be beneficial in differentiating between similarly structured or colored species. By conducting multiple flights at various stages of the growing season, I expect that the ability to differentiate between species improves.

Another consideration is the size of most rangeland vegetation. While clusters or mats of grasses can occupy square meters of space, individual grass or forbs can occupy areas smaller than 9 cm^2. By mounting a sensor with a higher spatial resolution on a lower, faster flying platform, the increase in spatial resolution could be sufficient to identify smaller grasses or forbs. This might allow for the inclusion of small flowers to be used in identifying plants such as Pitcher's sage.

Finally, improvements in spectral resolution pose a unique challenge. Current modified cameras only allow for the recording of three spectral ranges. In order to collect data in more than three bands, more sensors are needed. While these sensors exist, they are generally used in agricultural applications where fine resolution is less important. Thus, an increase in spectral resolution would require a forced decrease in spatial resolution. As new technology is developed for non-agricultural applications, the use of more spectral bands in classifying rangeland vegetation will become more feasible. Another plausible method to address the need for spectral resolution would be to use multiple modified cameras on the same platform. However, the need to co-register images from the two cameras becomes increasingly difficult, especially when the proximity of the two sensors might change with every flight.

Implications for Ecology and Management

Unmanned aerial vehicles provide the opportunity to increase the frequency and spatial resolution that remotely sensed data is collected. These considerations are especially important when studying fine-spatial-scale and phenologically diverse patterns such as those found in the mixed-grass communities of western Kansas. The use of light-weight, low-cost UASs and sensors in the field of rangeland management can facilitate the identification of rangeland vegetation better than random assignment using three spectral bands, though aforementioned considerations should be accounted for. It is important to accomplish these tasks on a limited budget, thereby stressing the importance of low-cost platforms, sensors, and open-source software. As technology improves,

including the production of high-quality, low-cost sensors and faster computers, the use of UASs for determining stocking rates for grazing, locating invasive or rare species, and estimating overall biodiversity will become more feasible. Until then, we must work with the available technology and work to better understand how increases in spatial resolution and modeling paradigms affect our ability, as ecologists and managers, to understand the systems with which we work.

TABLES

Table 1: List of all cover types in the study area including common name, scientific name, and code. Highlighted rows indicate the final 17 covers included in analyses.

Leadplant	Amorpha canescens	AMCA	Switchgrass	Panicum virgatum	PAVI
Annual Broomweed	Amphiachyris dracunculoide	AMDR	Scurf pea	Pediomelum tenuiflorum	PETE
Ragweed	Ambrosia psilostachya	AMPS	Clammyweed	Polanisia dodecandra	PODO
Big Bluestem	Andropogon gerardii	ANGE	Rose	Prunus	Prunus
Lousiana Sagewort	Artemisia ludoviciana	ARLU	Fragrant sumac	Rhus aromatica	RHAR
Purple Threeawn	Aristida purpurea	ARPU	Smooth sumac	Rhus glaborous	RHGL
Bare ground		BG	Rock/Gravel		ROCK
Sideoats Gramma	Bouteloua curtipendula	BOCU	Blackberry	Rubus ostryifolius	RUOS
Buffalograss	Bouteloua dactloides	BODA	Pitcher's sage	Salvia azurea	SAAZ
Blue Gramma	Bouteloua gracilis	BOGR	Little Bluestem	Schizachyrium scoparium	SCSC
Hairy Gramma	Bouteloua hirsuta	BOHI	Compass plant	Silphium laciniatum	SILA
Silver Bluestem	Bothriochloa laguroides	BOLA	Indiangrass	Sorghastrum nutans	SONU
Downy Brome	Bromus tectorum	BRTE	Rigid goldenrod	Solidago rigida	SORI
Purple Poppy-Mallow	Callirhoe involucrata	CAIN	Tall dropseed	Sporobolus compositus	SPCO
Ditchweed	Cannabis sp.	CASP	Sand Dropseed	Sporobolus cryptandrus	SPCR
Marestail	Conyza canadensis	COCA	Prairie Cordgrass	Spartina pectinata	SPPE
Coroton	Croton texensis	CRTE	Stenosiphon (False Gaura	Stenosiphon linifolius	STLI
Black Sampson	Echinacea angustifolia	ECAN	Heath aster	Symphyotrichum ericoides	SYER
Snow On The Mountain	Euphorbia marginata	EUMA	Aromatic Aster	Symphyotrichum oblongifoliur	SYOB
Curlycup Gumweed	Grindelia squarrosa	GRSQ	Buckbrush	Symphoricarpos orbiculatus	SYOR
Broom snakeweed	Gutierezzia sarothrae	GUSA	Intermediate wheatgrass	Thinopyrum intermedium	THIN
Maximillian sunflower	Helianthus maximilianii	HEMA	Tree (unknown)		TREE
Ashy sunflower	Helianthus mollis	HEMO	Baldwins ironweed	Vervonia baldwinii	VEBA
Eastern Red Cedar	Juniperus virginiana	JUVI	Riverbank grape	Vitis riparia	VIRI
Dotted Gayfeather	Liatris punctata	LIPU	Water		WATER
Litter		LITTER	Yucca	Yucca glauca	YUGL

Table 2: Classification matrix of the Maximum Likelihood Classifier. Cells highlighted in grey indicate the highest number of true/predicted instances.

		Predicted																
		SCSC	**SPCO**	**AMPS**	**ANGE**	**ARLU**	**BG**	**BOCU**	**BODA**	**LITTER**	**ROCK**	**WATER**	**JUVI**	**RHAR**	**RHGL**	**TREE**	**VIRI**	**YUGL**
TRUE	**SCSC**	71	58	75	37	9		7	73	22		2	25	200	379	21	6	97
	SPCO	4	92	20	1	5		18	136	12		3	8	10	224			35
	AMPS	28	65	443	29	7	5	2	101	153	3	2	11	1	533	1		83
	ANGE				4	7		4							45	19	5	5
	ARLU	3	10	5	8	69	26	4		17	2				372			27
	BG	4	4	1	13	36	41	6	21	5	54	1	1		410			34
	BOCU		16		1	3	1	12	45	2		6			34			30
	BODA		12	6	2	2		10	231	16	1	44			167			69
	LITTER	2	1	13	6	11		3	1					3	54			4
	ROCK		2	3	3	5	7				12		21		115			2
	WATER	1			1				20			37				14		30
	JUVI												1	4	29			9
	RHAR	6	20			2		1					16	8	33			25
	RHGL	19											3					
	TREE															52		2
	VIRI				3										12			3
	YUGL	1	2												41			4

Table 3: Classification matrix of the Random Forest Classifier. Cells highlighted in grey indicate the highest number of true/predicted instances.

		Predicted												
		SCSC	**SPCO**	**AMPS**	**ARLU**	**BG**	**BOCU**	**BODA**	**LITTER**	**ROCK**	**WATER**	**JUVI**	**RHAR**	**TREE**
TRUE	**SCSC**	554	116	50	1	2		46				31	6	253
	SPCO	131	303	35	11	16		81	5			21		1
	AMPS	319	406	361	18	28		121	4			68		85
	ANGE	2		6	10	7		1				4		35
	ARLU	23	47	14	274	124		9		27		9		25
	BG	81	29	27	101	177	1	58		155				17
	BOCU	53	20			8		62	20					
	BODA	94	158	12	2	8		328	1			2		1
	LITTER	58	17	2		8		14						2
	ROCK				21	36				81		13		4
	WATER	1				13		4		7	64			1
	JUVI			6								17	1	25
	RHAR	26											2	85
	RHGL													20
	TREE													60
	VIRI	1										2		21
	YUGL	40												

Table 4: Classification matrix of the Support Vector Machine Classifier. Cells highlighted in grey indicate the highest number of true/predicted instances.

		Predicted												
		SCSC	SPCO	AMPS	ANGE	ARLU	BG	BOCU	BODA	LITTER	ROCK	WATER	JUVI	RHAR	TREE
TRUE	SCSC	612	79	59	1	10	22		31				15	5	211
	SPCO	104	385	38		9	5		51				31		
	AMPS	223	377	498	1	29	47		142	4			32		37
	ANGE	15		11		4	2		2						47
	ARLU	74	38	27		277	99		5		26		26		9
	BG	71	21	11		151	156	7	37	4	171	2			11
	BOCU	44	19	8					76	2					
	BODA	72	194	37	1	1	24		284						
	LITTER	57	18	1			24		1						2
	ROCK					14	47			1	77		7		13
	WATER											91			
	JUVI	2											10	6	25
	RHAR	60													48
	RHGL														20
	TREE		1										2		56
	VIRI														16
	YUGL	26											3		5

Table 5: Comparison of classification metrics between the model scenarios. Bolded terms indicate the highest value for that metric.

TEST	Overall Accuracy	Cramer's V	Macro-Sensitivity	Macro-Specificity
MLC	13.7 (8.9)	0.48 (0.1)	11.5 (8.0)	92.3 (1.3)
RAND	13.7 (2.0)	0.23 (0.03)	5.8 (1.4)	**94.1 (0.1)**
RF	28.4 (7.7)	0.53 (0.06)	21.6 (10.1)	93.2 (0.89)
SVM	**32.4 (7.3)**	**0.55 (0.05)**	**22.3 (9.6)**	93.6 (0.8)

Table 6: Results of nonparametric Tukey's HSD tests. Pairwise model comparisons are indicated in the first column of each section. The last column of each section indicates that there is a significant difference between the two scenarios.

Overall Accuracy

Comparison	Observed Difference	Critical Difference	Significant Difference
MLC-RAND	3.385	43.13638045	FALSE
MLC-RF	160.55	43.13638045	TRUE
MLC-SVM	197.055	43.13638045	TRUE
RAND-RF	163.935	43.13638045	TRUE
RAND-SVM	200.44	43.13638045	TRUE
RF-SVM	36.505	43.13638045	FALSE

Macro-Sensitivity

Comparison	Observed Difference	Critical Difference	Significant Difference
MLC-RAND	85.68	43.13638045	TRUE
MLC-RF	115.645	43.13638045	TRUE
MLC-SVM	125.515	43.13638045	TRUE
RAND-RF	201.325	43.13638045	TRUE
RAND-SVM	211.195	43.13638045	TRUE
RF-SVM	9.87	43.13638045	FALSE

Cramer's V

Comparison	Observed Difference	Critical Difference	Significant Difference
MLC-RAND	163.64	43.13638045	TRUE
MLC-RF	41.25	43.13638045	FALSE
MLC-SVM	66.35	43.13638045	TRUE
RAND-RF	204.89	43.13638045	TRUE
RAND-SVM	229.99	43.13638045	TRUE
RF-SVM	25.1	43.13638045	FALSE

Macro-Specificity

Comparison	Observed Difference	Critical Difference	Significant Difference
MLC-RAND	190.56	43.13638045	TRUE
MLC-RF	70.51	43.13638045	TRUE
MLC-SVM	112.17	43.13638045	TRUE
RAND-RF	120.05	43.13638045	TRUE
RAND-SVM	78.39	43.13638045	TRUE
RF-SVM	41.66	43.13638045	FALSE

FIGURES

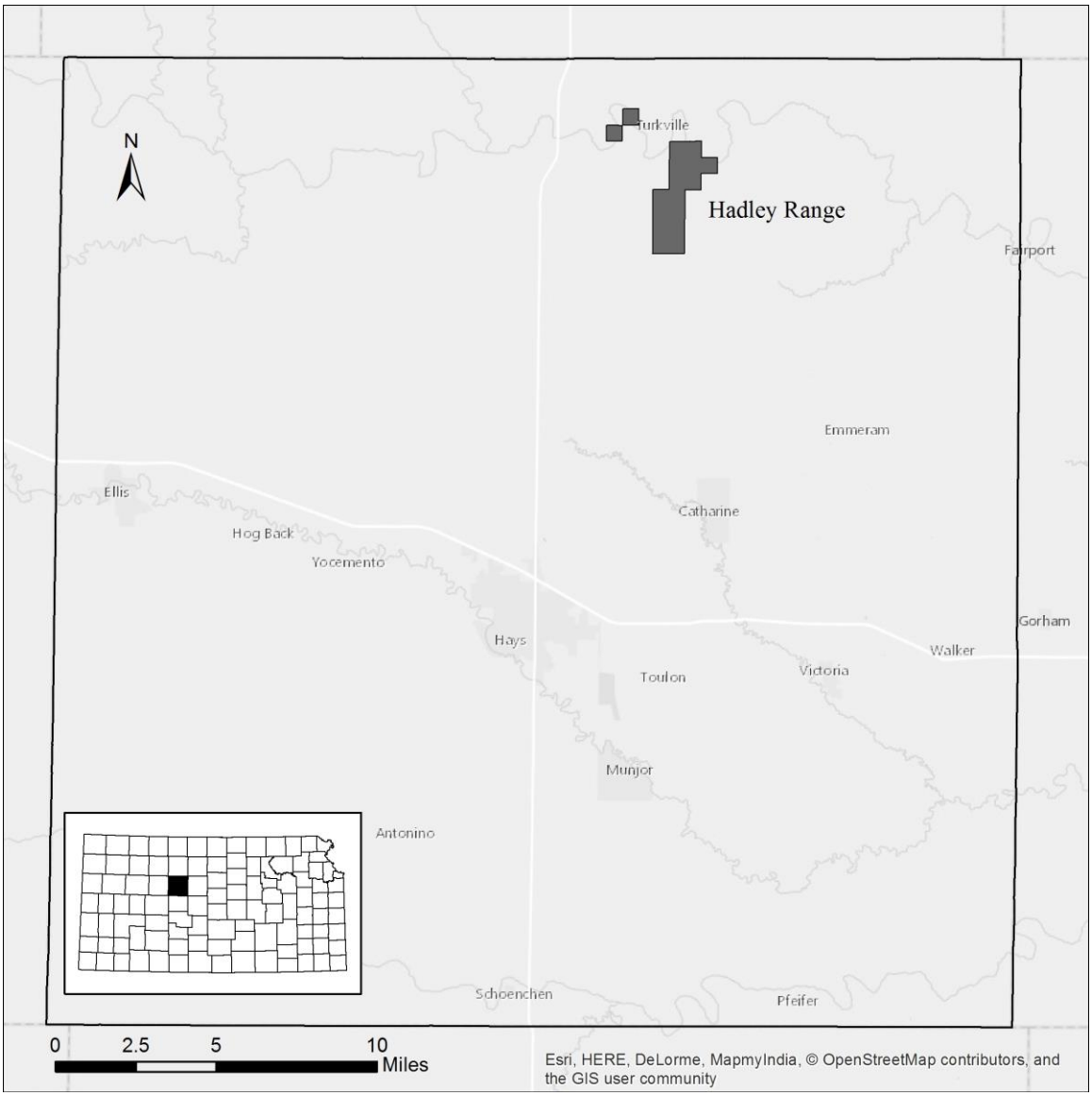

Figure 1: Locus map of Hadley Range within Ellis county. Hadley is the black polygon in the upper right. Ellis county is highlighted on the KS map.

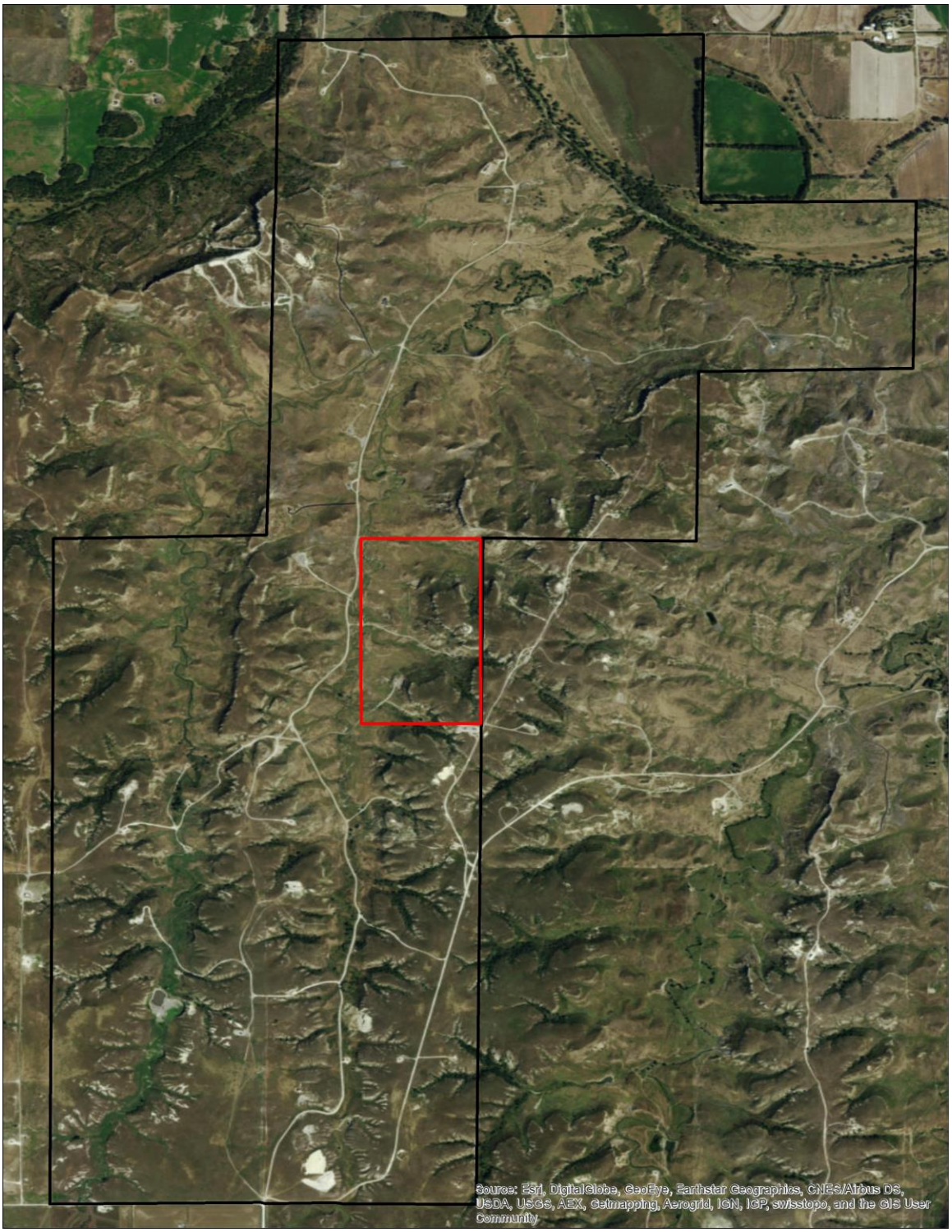

Figure 2: Map of the 40-hectare study area within Hadley Range. Hadley is outlined in black, the study area is outlined in red.

Figure 3: 3DR Iris+ quadcopter.

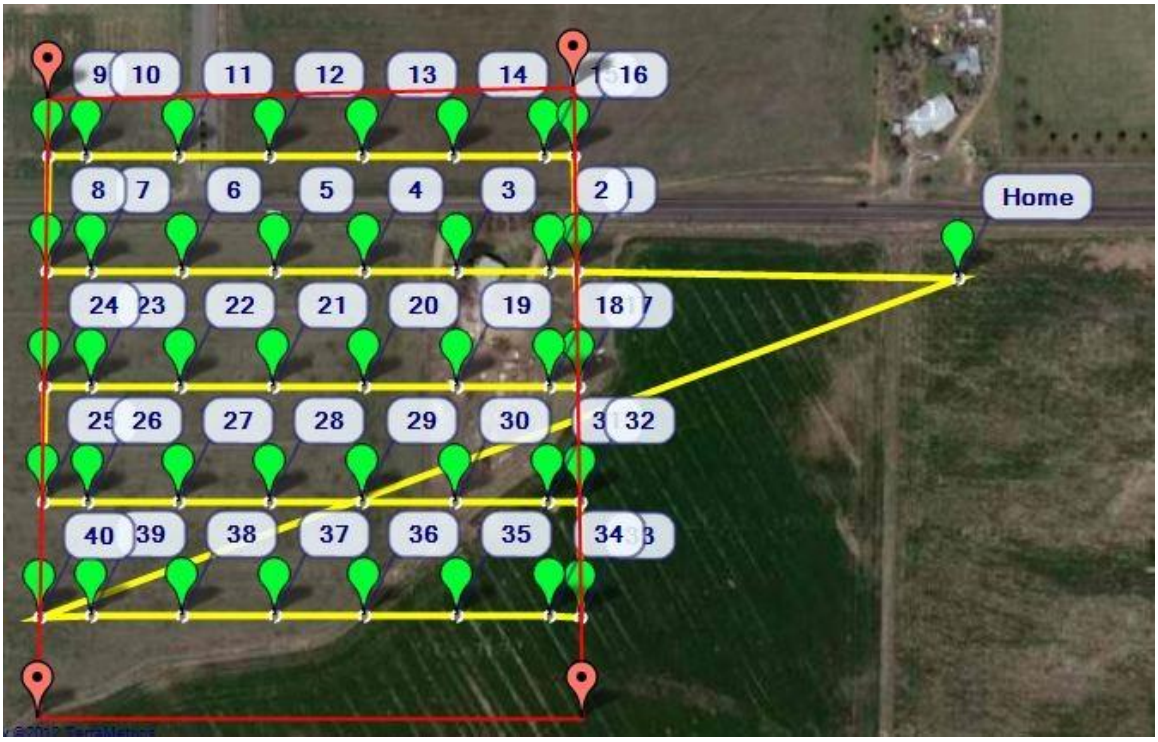

Figure 4: Mission Planner displaying waypoints used to calculate flight paths. The distance between paths is determined by sideways overlap whereas the frequency of images is determined by forward overlap.

Figure 5: Flight paths from the full data collection. Ten total flights with 80% overlap. Takeoff locations are marked with a black X.

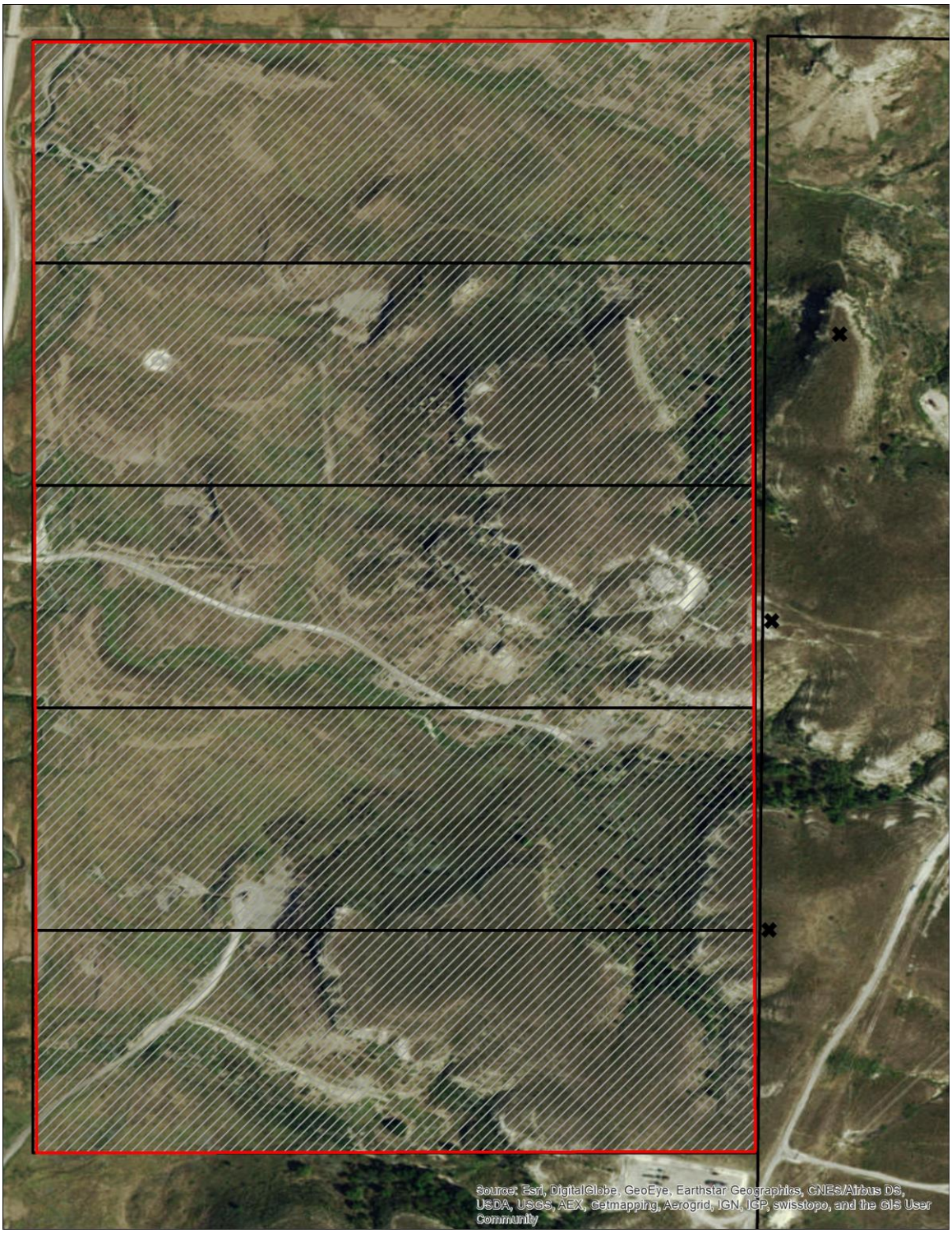

Figure 6: Flight paths from the reference data collection. Five total flights with reduced overlap. Takeoff locations are marked with a black X.

Figure 7: Locations of 110 systematically placed, one square meter Daubenmire quadrats.

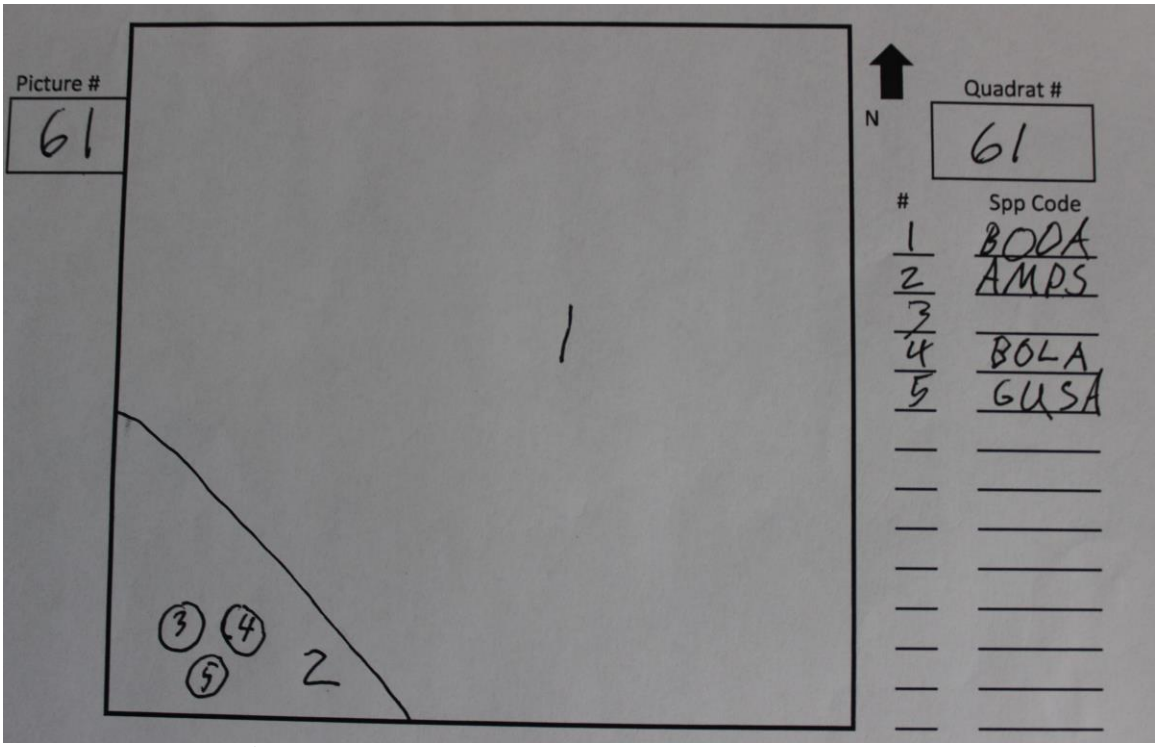

Figure 8: Data collection form where the square represents the Daubenmire quadrat. Polygons drawn inside were used in digitization when compared to Figure 9 and aerial imagery.

Figure 9: Image taken from beside the quadrat. Used when comparing Figure 8 to aerial imagery for digitization of cover types.

Figure 10: Image of one on-ground control points used in georectification of imagery.

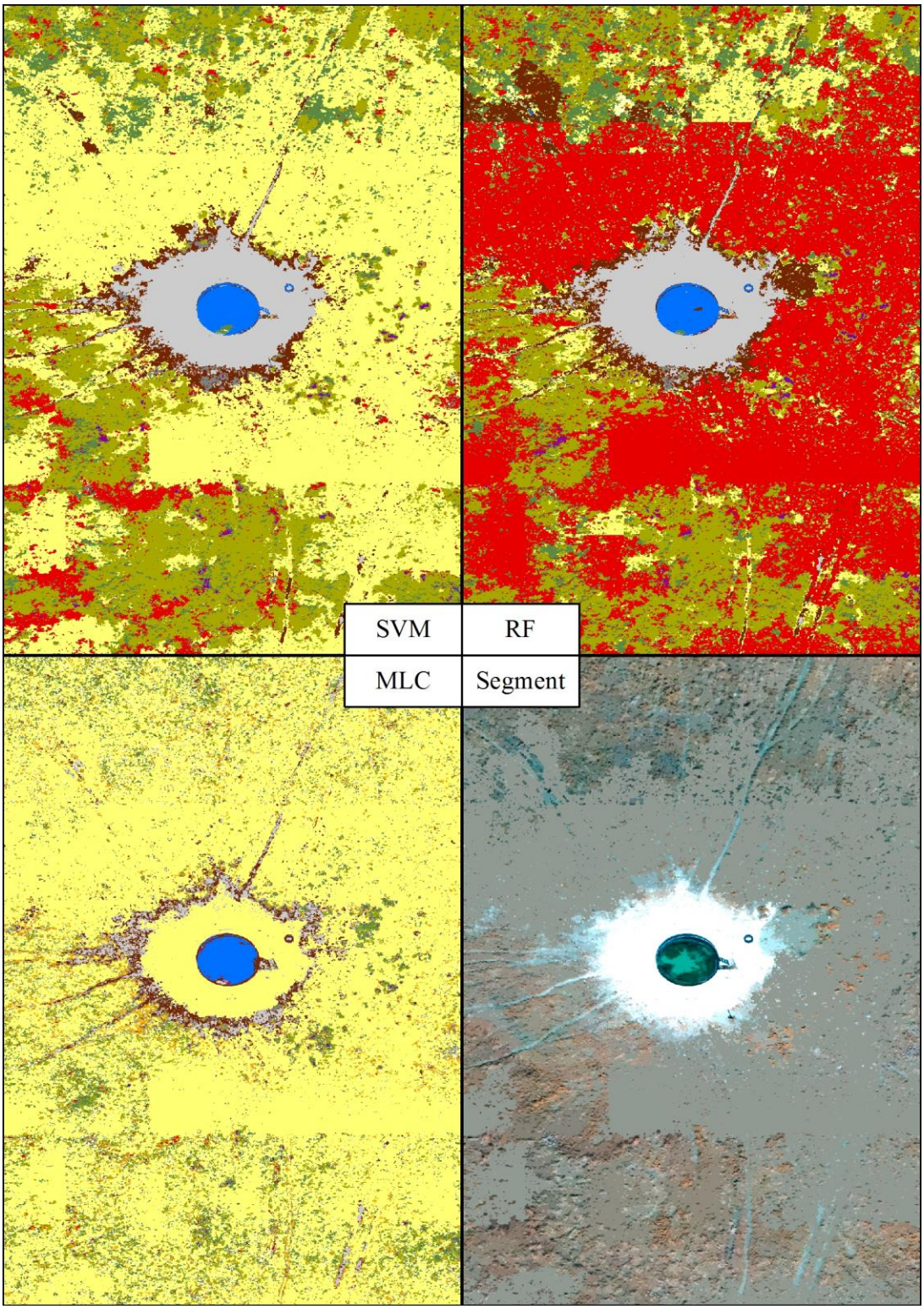

Figure 11: Results of a single classification execution from each modeling algorithm. The colors for each classification are uniform. Therefore, what is red in one panel should be red in all three if the models were equal. The bottom right corner is the segmented image, present for comparison.

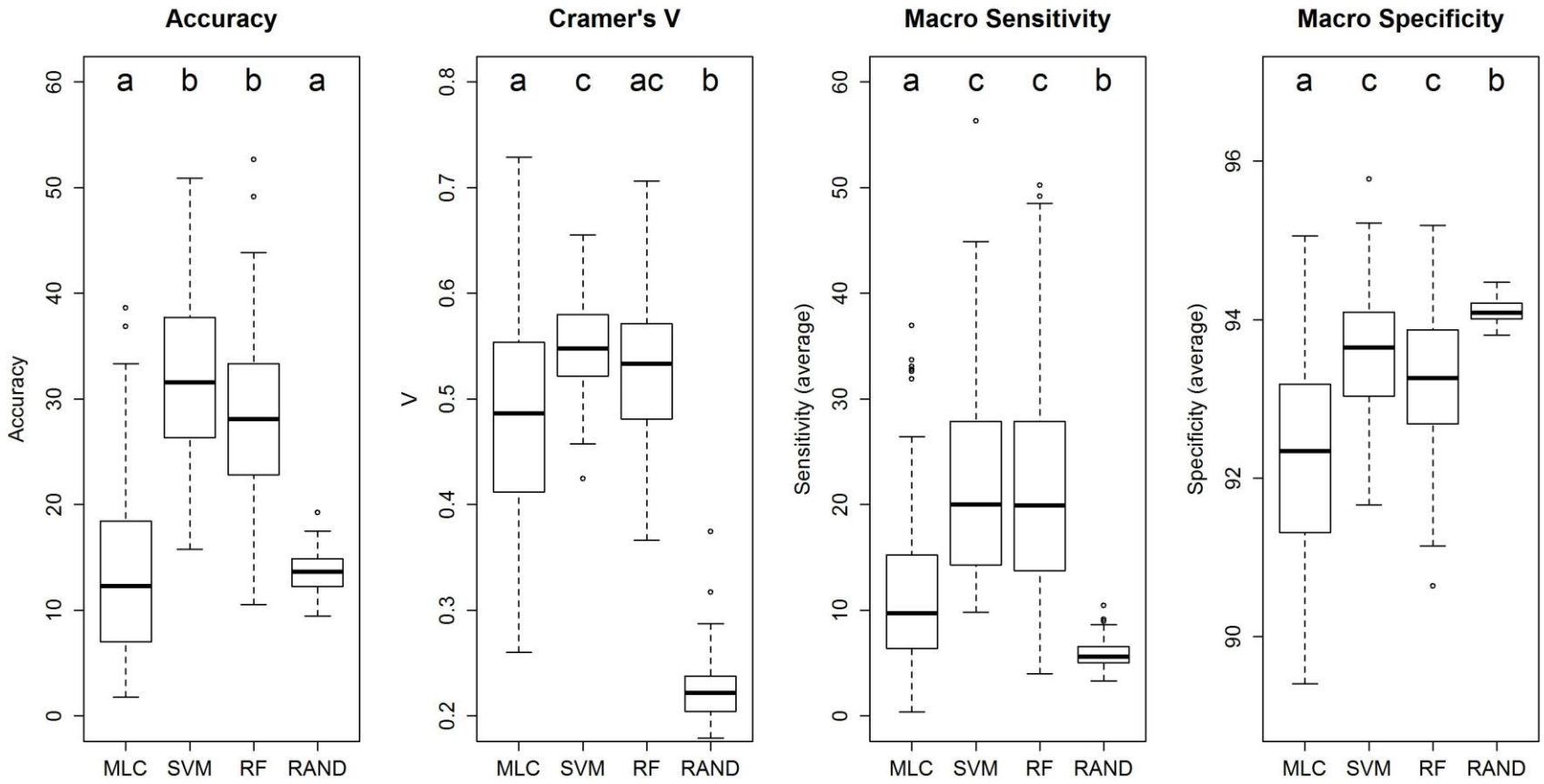

Figure 12: Boxplots indicating the median and variation of each performance metric within each model type. Letters above the bars indicate models which are grouped based on Tukey's HSD tests. Different letters indicate statistical difference. MLC = Maximum Likelihood Classifier, SVM = Support Vector Machine, RF = Random Forest, RAND = random model.

APPENDICES

Appendix 1: Python Code

```
1.  import arcpy
2.  from arcpy import sa
3.  from arcpy import env
4.  from arcpy.sa import *
5.  arcpy.env.overwriteOutput = True
6.
7.  #Execute Maximum Likelihood
8.  roi = "I:/Runs/Shape.gdb/Final"
9.  trainwksp = "I:/Runs/Training.gdb/"
10. testwksp = "I:/Runs/Testing.gdb/"
11. ecdwksp = "I:/Runs/ECD/"
12. mtrxwksp = "I:/Runs/Matrix.gdb/"
13. clsfdwksp = "I:/Runs/Classified.gdb/"
14. accasswksp = "I:/Runs/Accass.gdb/"
15. lookup = "I:/Runs/Lookup.gdb/"
16. updaccasswksp = "I:/Runs/Updaccass.gdb/"
17. csvwksp = "I:/Runs/CSV/MLC"
18. image = "I:/Runs/final.tif"
19. segimage = "I:/Runs/Segment.gdb/Segmented"
20. attributes = "COLOR;MEAN;STD;COUNT;COMPACTNESS;RECTANGULARITY"
21.
22. count = 1
23. while count <= 100 :
24.     arcpy.SubsetFeatures_ga(roi, trainwksp + "MLCroitrain" + str(count), testwksp + "MLCroitest" + str(count
    ), 80, "PERCENTAGE_OF_INPUT")
25.     train = trainwksp + "MLCroitrain" + str(count)
26.     test = testwksp + "MLCroitest" + str(count)
27.     ecd = TrainMaximumLikelihoodClassifier(segimage, train, ecdwksp + "MLCclassifier" + str(count) +".ecd
    ", image, attributes)
28.     classied = ClassifyRaster(image, ecd, segimage)
29.     classied.save(clsfdwksp + "MLclassified" + str(count))
30.     arcpy.FeatureToPoint_management(test, accasswksp + "accasspt" + str(count), "CENTROID")
31.     accass = accasswksp + "accasspt" + str(count)
32.     Look = Lookup(classied, "Classvalue")
33.     Look.save(lookup + "MLClookup" + str(count))
34.     lookuprast = lookup + "MLClookup" + str(count)
35.     ExtractValuesToPoints(accass, lookuprast, updaccasswksp + "updaccasspt" + str(count), "NONE", "ALL")

36.     uaa = updaccasswksp + "updaccasspt" + str(count)
37.     arcpy.ExportXYv_stats(uaa, "OBJECTID;Classvalue;RASTERVALU", "COMMA", csvwksp + str(count)
     + ".csv", "ADD_FIELD_NAMES")
38.     count = count + 1
39.
40. #Execute Random Forests
41. roi = "I:/Runs/Shape.gdb/Final"
42. trainwksp = "I:/Runs/Training.gdb/"
43. testwksp = "I:/Runs/Testing.gdb/"
44. ecdwksp = "I:/Runs/ECD/"
45. mtrxwksp = "I:/Runs/Matrix.gdb/"
46. clsfdwksp = "I:/Runs/Classified.gdb/"
47. accasswksp = "I:/Runs/Accass.gdb/"
48. lookup = "I:/Runs/Lookup.gdb/"
```

```
updaccasswksp = "I:/Runs/Updaccass.gdb/"
csvwksp = "I:/Runs/CSV/RF"
image = "I:/Runs/final.tif"
segimage = "I:/Runs/Segment.gdb/Segmented"
attributes = "COLOR;MEAN;STD;COUNT;COMPACTNESS;RECTANGULARITY"

count = 1
while count <= 100 :
    arcpy.SubsetFeatures_ga(roi, trainwksp + "RFroitrain" + str(count), testwksp + "RFroitest" + str(count), 80, "PERCENTAGE_OF_INPUT")
    train = trainwksp + "RFroitrain" + str(count)
    test = testwksp + "RFroitest" + str(count)
    ecd = TrainRandomTreesClassifier(segimage, train, ecdwksp + "RFclassifier" + str(count) +".ecd", image, "", "", "0",attributes)
    classied = ClassifyRaster(image, ecd, segimage)
    classied.save(clsfdwksp + "RFclassified" + str(count))
    arcpy.FeatureToPoint_management(test, accasswksp + "accasspt" + str(count), "CENTROID")
    accass = accasswksp + "accasspt" + str(count)
    Look = Lookup(classied, "Classvalue")
    Look.save(lookup + "RFlookup" + str(count))
    lookuprast = lookup + "RFlookup" + str(count)
    ExtractValuesToPoints(accass, lookuprast, updaccasswksp + "updaccasspt" + str(count), "NONE", "ALL")

    uaa = updaccasswksp + "updaccasspt" + str(count)
    arcpy.ExportXYv_stats(uaa, "OBJECTID;Classvalue;RASTERVALU", "COMMA", csvwksp + str(count) + ".csv", "ADD_FIELD_NAMES")
    count = count + 1

#Execute Support Vector Machine
roi = "I:/Runs/Shape.gdb/Final"
trainwksp = "I:/Runs/Training.gdb/"
testwksp = "I:/Runs/Testing.gdb/"
ecdwksp = "I:/Runs/ECD/"
mtrxwksp = "I:/Runs/Matrix.gdb/"
clsfdwksp = "I:/Runs/Classified.gdb/"
accasswksp = "I:/Runs/Accass.gdb/"
lookup = "I:/Runs/Lookup.gdb/"
updaccasswksp = "I:/Runs/Updaccass.gdb/"
csvwksp = "I:/Runs/CSV/SVM"
image = "I:/Runs/final.tif"
segimage = "I:/Runs/Segment.gdb/Segmented"
attributes = "COLOR;MEAN;STD;COUNT;COMPACTNESS;RECTANGULARITY"

count = 1
while count <= 100 :
    arcpy.SubsetFeatures_ga(roi, trainwksp + "SVMroitrain" + str(count), testwksp + "SVMroitest" + str(count), 80, "PERCENTAGE_OF_INPUT")
    train = trainwksp + "SVMroitrain" + str(count)
    test = testwksp + "SVMroitest" + str(count)
    ecd = TrainSupportVectorMachineClassifier(segimage, train, ecdwksp + "SVMclassifier" + str(count) +".ecd", image, "0", attributes)
    classied = ClassifyRaster(image, ecd, segimage)
    classied.save(clsfdwksp + "SVclassified" + str(count))
    arcpy.FeatureToPoint_management(test, accasswksp + "accasspt" + str(count), "CENTROID")
    accass = accasswksp + "accasspt" + str(count)
    Look = Lookup(classied, "Classvalue")
    Look.save(lookup + "SVMlookup" + str(count))
    lookuprast = lookup + "SVMlookup" + str(count)
```

```
101.    ExtractValuesToPoints(accass, lookuprast, updaccasswksp + "updaccasspt" + str(count), "NONE", "ALL")

102.    uaa = updaccasswksp + "updaccasspt" + str(count)
103.    arcpy.ExportXYv_stats(uaa, "OBJECTID;Classvalue;RASTERVALU", "COMMA", csvwksp + str(count)
        + ".csv", "ADD_FIELD_NAMES")
104.    count = count + 1
```

Appendix 2: R Code

```
1.  setwd("B:\\FHSU_Thesis\\CSV_Data\\MLC\\") #set the path name to the CSVs
2.  file.names <- dir(pattern ="*.csv")
3.  finaldata <- as.data.frame(matrix(0, ncol= 11, nrow=1))
4.  names <-
     c("TEST", "ACC", "CE", "BER", "KAPPA", "CRAMERV", "MACRO_SENS", "MACRO_SPEC", "MACR
    O_PREC", "MACRO_F1", "MACRO_MCC")
5.  colnames(finaldata) <- names
6.  library(rminer)
7.  library(plyr)
8.
9.  for(i in 1:length(file.names)){
10.   mat <- read.csv(file.names[[i]])
11.   conf <- mmetric(factor(mat$CLASSVALUE), factor(mat$RASTERVALU), metric="ALL")
12.   confmat <- as.data.frame(conf)
13.   title <- 1
14.   ACC <- confmat[1,]
15.   CE <- confmat[2,]
16.   BER <- confmat[3,]
17.   KAPPA <- confmat[4,]
18.   CRAMERV <- confmat[5,]
19.   MACRO_SENS <- mean(subset(confmat, grepl("TPR", rownames(confmat)), drop=TRUE))
20.   MACRO_SPEC <- mean(subset(confmat, grepl("TNR", rownames(confmat)), drop=TRUE))
21.   MACRO_PREC <- mean(subset(confmat, grepl("PREC", rownames(confmat)), drop=TRUE))
22.   MACRO_F1 <- mean(subset(confmat, grepl("F1", rownames(confmat)), drop=TRUE))
23.   MACRO_MCC <- mean(subset(confmat, grepl("MCC", rownames(confmat)), drop=TRUE))
24.   run <-
     c(title, ACC, CE, BER, KAPPA, CRAMERV, MACRO_SENS, MACRO_SPEC, MACRO_PREC, MACR
    O_F1, MACRO_MCC)
25.   finaldata <- rbind(finaldata, run)
26. }
27.
28. setwd("B:\\FHSU_Thesis\\CSV_Data\\SVM\\") #set the path name to the CSVs
29. file.names <- dir(pattern ="*.csv")
30. for(i in 1:length(file.names)){
31.   mat <- read.csv(file.names[[i]])
32.   conf <- mmetric(factor(mat$CLASSVALUE), factor(mat$RASTERVALU), metric="ALL")
33.   confmat <- as.data.frame(conf)
34.   title <- 3
35.   ACC <- confmat[1,]
36.   CE <- confmat[2,]
37.   BER <- confmat[3,]
38.   KAPPA <- confmat[4,]
39.   CRAMERV <- confmat[5,]
40.   MACRO_SENS <- mean(subset(confmat, grepl("TPR", rownames(confmat)), drop=TRUE))
41.   MACRO_SPEC <- mean(subset(confmat, grepl("TNR", rownames(confmat)), drop=TRUE))
42.   MACRO_PREC <- mean(subset(confmat, grepl("PREC", rownames(confmat)), drop=TRUE))
43.   MACRO_F1 <- mean(subset(confmat, grepl("F1", rownames(confmat)), drop=TRUE))
```

```
  MACRO_MCC <- mean(subset(confmat, grepl("MCC", rownames(confmat)), drop=TRUE))
  run <-
 c(title, ACC, CE, BER, KAPPA, CRAMERV, MACRO_SENS, MACRO_SPEC, MACRO_PREC, MACR
O_F1, MACRO_MCC)
  finaldata <- rbind(finaldata, run)
}

setwd("B:\\FHSU_Thesis\\CSV_Data\\RF\\") #set the path name to the CSVs
file.names <- dir(pattern ="*.csv")
for(i in 1:length(file.names)){
  mat <- read.csv(file.names[[i]])
  conf <- mmetric(factor(mat$CLASSVALUE), factor(mat$RASTERVALU), metric="ALL")
  confmat <- as.data.frame(conf)
  title <- 2
  ACC <- confmat[1,]
  CE <- confmat[2,]
  BER <- confmat[3,]
  KAPPA <- confmat[4,]
  CRAMERV <- confmat[5,]
  MACRO_SENS <- mean(subset(confmat, grepl("TPR", rownames(confmat)), drop=TRUE))
  MACRO_SPEC <- mean(subset(confmat, grepl("TNR", rownames(confmat)), drop=TRUE))
  MACRO_PREC <- mean(subset(confmat, grepl("PREC", rownames(confmat)), drop=TRUE))
  MACRO_F1 <- mean(subset(confmat, grepl("F1", rownames(confmat)), drop=TRUE))
  MACRO_MCC <- mean(subset(confmat, grepl("MCC", rownames(confmat)), drop=TRUE))
  run <-
 c(title, ACC, CE, BER, KAPPA, CRAMERV, MACRO_SENS, MACRO_SPEC, MACRO_PREC, MACR
O_F1, MACRO_MCC)
  finaldata <- rbind(finaldata, run)
}

finaldata <- finaldata[-c(1),]

random <- read.csv("B:/FHSU_Thesis/CSV_Data/random.csv")
random <- random[,1:2]

count = 3

while(count < 103){
  random[,count] <- sample(random$CLASSVALUE, replace=TRUE)
  count = count + 1
}

true <- random$CLASSVALUE
pred <- random[,3:length(random)]

for(i in names(pred)){
  y <- pred[i]
  y <- unlist(y)
  conf <- mmetric(factor(true), factor(y), metric="ALL")
  confmat <- as.data.frame(conf)
  title <- 4
  ACC <- confmat[1,]
  CE <- confmat[2,]
  BER <- confmat[3,]
  KAPPA <- confmat[4,]
  CRAMERV <- confmat[5,]
  MACRO_SENS <- mean(subset(confmat, grepl("TPR", rownames(confmat)), drop=TRUE))
  MACRO_SPEC <- mean(subset(confmat, grepl("TNR", rownames(confmat)), drop=TRUE))
  MACRO_PREC <- mean(subset(confmat, grepl("PREC", rownames(confmat)), drop=TRUE))
```

```
  MACRO_F1 <- mean(subset(confmat, grepl("F1", rownames(confmat)), drop=TRUE))
  MACRO_MCC <- mean(subset(confmat, grepl("MCC", rownames(confmat)), drop=TRUE))
  run <-
   c(title, ACC, CE, BER, KAPPA, CRAMERV, MACRO_SENS, MACRO_SPEC, MACRO_PREC, MACRO_F1, MACRO_MCC)
  finaldata <- rbind(finaldata, run)
}

### Testing Assumptions ###
final <- finaldata

library(mvoutlier)

alg_m <- final[final$TEST=="MLC",]
alg_rf <- final[final$TEST=="RF",]
alg_s <- final[final$TEST=="SVM",]
alg_rnd <- final[final$TEST=="RAND",]

outliers <-
 aq.plot(final[c("ACC", "BER", "CE", "CRAMERV", "KAPPA", "MACRO_SENS", "MACRO_SPEC", "MACRO_PREC")])

x <- finaldata[,2:length(finaldata)]
View(cor(x, method="spearman"))

par(mfrow=c(2,4))
qqnorm((finaldata$ACC)^0.66, xlab="ACC")
qqline((finaldata$ACC)^0.66, col="red")
qqnorm((finaldata$CE)^2.3, xlab="CE")
qqline((finaldata$CE)^2.3, col="red")
qqnorm((finaldata$BER)^5.8, xlab = "BER")
qqline((finaldata$BER)^5.8, col="red")
qqnorm(finaldata$KAPPA, xlab="KAPPA")
qqline(finaldata$KAPPA, col="red")
qqnorm(finaldata$CRAMERV, xlab="Cramer V")
qqline(finaldata$CRAMERV, col="red")
qqnorm(sqrt(finaldata$MACRO_SENS), xlab="SENS")
qqline(sqrt(finaldata$MACRO_SENS), col = "red")
qqnorm(finaldata$MACRO_SPEC, xlab="Spec")
qqline(finaldata$MACRO_SPEC, col="red")
qqnorm(finaldata$MACRO_PREC, xlab="Prec")
qqline(finaldata$MACRO_PREC, col="red")

library(vegan)
library(pgirmess)
library(multcompView)

finaldata2 <- finaldata[ -c(3,4,5, 9, 10, 11)] #remove collinear variables

d <- vegdist(finaldata2[2:5], method="manhattan")
x <- betadisper(d, finaldata2$TEST, type = c("median", "centroid"))
anova(x)
plot(x)
boxplot(x, ylab= "Distance to centroid")
TukeyHSD(x) #Variances are unequal

###Proceed with nonparametric tests###

attach(finaldata)
```

```
adonis.final <-
    adonis(ACC + CRAMERV + MACRO_SENS + MACRO_SPEC + MACRO_PREC~TEST, data = finaldata
2, method = "manhattan")
adonis.final
#This shows a significant result; let this be OK because while variances are unequal,
#there are "large" samples and sample sizes are equal

ag <- aggregate(.~TEST, finaldata2, function(x) c(mean=mean(x), sd=sd(x)))

#Following up on a significant adonis with multiple Kruskal-Wallis tests
kruskal.test(ACC~TEST, data=finaldata2)
kruskal.test(CRAMERV~TEST, data=finaldata2)
kruskal.test(MACRO_SENS~TEST, data=finaldata2)
kruskal.test(MACRO_SPEC~TEST, data=finaldata2)

kruskalmc(ACC~TEST, data=finaldata2)
kruskalmc(CRAMERV~TEST, data=finaldata2)
kruskalmc(MACRO_SENS~TEST, data=finaldata2)
kruskalmc(MACRO_SPEC~TEST, data=finaldata2)

mlcacc <- finaldata$ACC[finaldata$TEST=="MLC"]
mlcce <- finaldata$CE[finaldata$TEST=="MLC"]
mlcber<- finaldata$BER[finaldata$TEST=="MLC"]
mlckap<- finaldata$KAPPA[finaldata$TEST=="MLC"]
mlccrv <- finaldata$CRAMERV[finaldata$TEST=="MLC"]
mlcmacsen <- finaldata$MACRO_SENS[finaldata$TEST=="MLC"]
mlcmacspec<- finaldata$MACRO_SPEC[finaldata$TEST=="MLC"]
mlcmacprec<- finaldata$MACRO_PREC[finaldata$TEST=="MLC"]
mlcmac51<- finaldata$MACRO_F1[finaldata$TEST=="MLC"]
mlcmacmcc<- finaldata$MACRO_MCC[finaldata$TEST=="MLC"]


svmacc <- finaldata$ACC[finaldata$TEST=="SVM"]
svmce <- finaldata$CE[finaldata$TEST=="SVM"]
svmber<- finaldata$BER[finaldata$TEST=="SVM"]
svmkap<- finaldata$KAPPA[finaldata$TEST=="SVM"]
svmcrv <- finaldata$CRAMERV[finaldata$TEST=="SVM"]
svmmacsen <- finaldata$MACRO_SENS[finaldata$TEST=="SVM"]
svmmacspec<- finaldata$MACRO_SPEC[finaldata$TEST=="SVM"]
svmmacprec<- finaldata$MACRO_PREC[finaldata$TEST=="SVM"]
svmmac51<- finaldata$MACRO_F1[finaldata$TEST=="SVM"]
svmmacmcc<- finaldata$MACRO_MCC[finaldata$TEST=="SVM"]

rfacc <- finaldata$ACC[finaldata$TEST=="RF"]
rfce <- finaldata$CE[finaldata$TEST=="RF"]
rfber<- finaldata$BER[finaldata$TEST=="RF"]
rfkap<- finaldata$KAPPA[finaldata$TEST=="RF"]
rfcrv <- finaldata$CRAMERV[finaldata$TEST=="RF"]
rfmacsen <- finaldata$MACRO_SENS[finaldata$TEST=="RF"]
rfmacspec<- finaldata$MACRO_SPEC[finaldata$TEST=="RF"]
rfmacprec<- finaldata$MACRO_PREC[finaldata$TEST=="RF"]
rfmac51<- finaldata$MACRO_F1[finaldata$TEST=="RF"]
rfmacmcc<- finaldata$MACRO_MCC[finaldata$TEST=="RF"]

randacc <- finaldata$ACC[finaldata$TEST=="RAND"]
randce <- finaldata$CE[finaldata$TEST=="RAND"]
randber<- finaldata$BER[finaldata$TEST=="RAND"]
randkap<- finaldata$KAPPA[finaldata$TEST=="RAND"]
randcrv <- finaldata$CRAMERV[finaldata$TEST=="RAND"]
```

```
randmacsen <- finaldata$MACRO_SENS[finaldata$TEST=="RAND"]
randmacspec<- finaldata$MACRO_SPEC[finaldata$TEST=="RAND"]
randmacprec<- finaldata$MACRO_PREC[finaldata$TEST=="RAND"]
randmac51<- finaldata$MACRO_F1[finaldata$TEST=="RAND"]
randmacmcc<- finaldata$MACRO_MCC[finaldata$TEST=="RAND"]

###boxplots with letters###
par(mfrow=c(1,4))
mct_acc <- kruskalmc(ACC~TEST, data=finaldata2)
test_acc <- mct_acc$dif.com$difference
names(test_acc) <- row.names(mct_acc$dif.com)
let_acc <- multcompLetters(test_acc, compare = "<", threshold = 0.05)
let_acc
boxplot(mlcacc, svmacc, rfacc, randacc, main="Accuracy", ylab="Accuracy", ylim=c(0,60), names=c("MLC
", "SVM", "RF", "RAND"))
mtext(c("a", "b", "b", "a"), at=1:4, line = -2)

mct_crv <- kruskalmc(CRAMERV~TEST, data=finaldata2)
test_crv <- mct_crv$dif.com$difference
names(test_crv) <- row.names(mct_crv$dif.com)
let_crv <- multcompLetters(test_crv, compare = "<", threshold = 0.05)
let_crv
boxplot(mlccrv, svmcrv, rfcrv, randcrv, main = "Cramer's V", ylab = "V", ylim=c(0.2,0.8), names=c("MLC",
"SVM", "RF", "RAND"))
mtext(c("a", "c", "ac", "b"), at=1:4, line = -2)

mct_sens <- kruskalmc(MACRO_SENS~TEST, data=finaldata2)
test_sens <- mct_sens$dif.com$difference
names(test_sens) <- row.names(mct_sens$dif.com)
let_sens <- multcompLetters(test_sens, compare = "<", threshold = 0.05)
let_sens
boxplot(mlcmacsen, svmmacsen, rfmacsen, randmacsen, main="Macro Sensitivity", ylab = "Sensitivity (aver
age)", ylim=c(0,60),names=c("MLC", "SVM", "RF", "RAND"))
mtext(c("a", "c", "c", "b"), at=1:4, line =-2)

mct_spec <- kruskalmc(MACRO_SPEC~TEST, data=finaldata2)
test_spec <- mct_spec$dif.com$difference
names(test_spec) <- row.names(mct_spec$dif.com)
let_spec <- multcompLetters(test_spec, compare = "<", threshold = 0.05)
let_spec
boxplot(mlcmacspec, svmmacspec, rfmacspec, randmacspec, main="Macro Specificity", ylab = "Specificity (
average)", ylim = c(89,97), names=c("MLC", "SVM", "RF", "RAND"))
mtext(c("a", "c", "c", "b"), at=1:4, line = -2)

###export csvs ###
setwd("K:/Thesis/FHSU/Excel/MCT")
write.csv(mct_acc, "MCT_ACC.csv")
write.csv(mct_crv, "MCT_CRV.csv")
write.csv(mct_sens, "MCT_SENS.csv")
write.csv(mct_spec, "MCT_SPEC.csv")
```

LITERATURE CITED

Anderson, M.J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26: 32–46.

Asner, G. P. 1998. Biophysical and Biochemical Sources of Variability in Canopy Reflectance. *Remote Sensing of Environment* 64 (3): 234–53. doi:10.1016/S0034-4257(98)00014-5

Blaschke, T. 2010. Object Based Image Analysis for Remote Sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (1): 2–16. doi:10.1016/j.isprsjprs.2009.06.004.

Breckenridge, R. P., and M. E. Dakins. 2011. Evluation of Bare Ground on Rangelands Using Unmanned Aerial Vehicles: A Case Study. *GIScience & Remote Sensing* 48 (1): 74–85. doi:10.2747/1548-1603.48.1.74.

Breiman, L. 2001. Random Forests. *Machine Learning*, *45*(1): 5-32.

Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement, 20*(1): 37-46.

Colomina, I., and P. Molina. 2014. Unmanned Aerial Systems for Photogrammetry and Remote Sensing: A Review. *ISPRS Journal of Photogrammetry and Remote Sensing* 92 (June): 79–97. doi:10.1016/j.isprsjprs.2014.02.013.

Congalton, R. G. 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing of Environment* 37 (1): 35–46. doi:10.1016/0034-4257(91)90048-B.

Congalton, R.G., and K. Green. 2009. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices.* 2nd. Ed. Boca Raton, FL: CRC Press/Taylor & Francis.

Cortes, C., and V. Vapnik. 1995. Support-Vector Networks. *Machine learning*, *20*(3), 273-297.

Cortez, P. 2016. Rminer: Data Mining Classification and Regression Methods. R package version 1.4.2. https://CRAN.R-project.org/package=rminer

Cortez. P., A. Correia, P. Sousa, M. Rocha, M. Rio. 2010. Data Mining with Neural Networks and Support Vector Machines Using the R/rminer Tool. *Advances in Data Mining: Applications and Theoretical Aspects, Lecture Notes in Computer Science, 6171* 479-489.

Cramer, H. 1946. Mathematical Methods of Statistics. *Princeton: Princeton University Press.* p. 282

Cruzan, M.B., B.G. Weinstein, M.R. Grasty, B.F. Kohrn, E.C. Hendrickson, T.M. Arredondo, P.G. Thompson. 2016. Small Unmanned Aerial Vehicles (Micro-UASs, Drones) in Plant Ecology. *Applications in Plant Sciences, 4*(9):1600039

Daubenmire, R.F. 1959. A Canopy-Cover Method of Vegetation Analysis. *Northwest Science* 33:43-46.

Fourty, T., F. Baret, S. Jacquemoud, G. Schmuck, and J.Verdebout. 1996. Leaf Optical Properties with Explicit Description of Its Biochemical Composition: Direct and Inverse Problems. *Remote Sensing of Environment* 56: 104-117. Doi: 10.1016/0034-4257(95)00234-0.

Gini, R., D. Passoni, L. Pinto, and G. Sona. 2012. Aerial Images from an UAS System: 3d Modeling and Tree Species Classification in a Park Area. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 39 (B1): 361–66. doi:10.5194/isprsarchives-XXXIX-B1-361-2012.

Gitelson, A. A., Y. J. Kaufman, and M. N. Merzylak. (1996). Use of a Green Channel in Remote Sensing of Global Vegetation from EOS-MODIS. *Remote Sensing of Environment* 58 (3): 289-298. doi:10.1016/S0034-4257(96)00072-7.

Hay, G. J., G. Castilla, M. A. Wulder, and J. R. Ruiz. 2005. An Automated Object-Based Approach for the Multiscale Image Segmentation of Forest Scenes. *International Journal of Applied Earth Observation and Geoinformation* 7 (4): 339–59. doi:10.1016/j.jag.2005.06.005.

Hay, G. J., and G. Castilla. 2006. Object-Based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT). In *Proceedings of the 1st International Conference on Object-Based Image Analysis*, 4–5. http://www.isprs.org/proceedings/xxxvi/4-c42/Papers/OBIA2006_Hay_Castilla.pdf.

Huang, C., L.C. Davis, and J.R. Townshend. 2002. An Assessment of Support Vector Machines for Land Cover Classification. *International Journal of Remote Sensing*. 23:725-749.

Ishihama, F., Y. Watabe, and H. Oguma. 2012. Validation of a High-resolution, Remotely Operated Aerial Remote-sensing System for the Identification of Herbaceous Plant Species. Edited by Aaron Moody. *Applied Vegetation Science* 15 (3): 383–89. doi:10.1111/j.1654-109X.2012.01184.x.

Kelcey, J., and A. Lucieer. 2012. Sensor Correction of a 6-Band Multispectral Imaging Sensor for UAS Remote Sensing. *Remote Sensing* 4 (12): 1462–93. doi:10.3390/rs4051462.

Laliberte, A. S., M. A. Goforth, C. M. Steele, and A. Rango. 2011. Multispectral Remote Sensing from Unmanned Aircraft: Image Processing Workflows and Applications for Rangeland Environments. *Remote Sensing* 3 (12): 2529–51. doi:10.3390/rs3112529.

Laliberte, A. S., J. E. Herrick, A. Rango, and C. Winters. 2010. Acquisition, Orthorectification, and Object-based Classification of Unmanned Aerial Vehicle (UAS) Imagery for Rangeland Monitoring. *Photogrammetric Engineering & Remote Sensing* 76 (6): 661–72. doi:10.14358/PERS.76.6.661.

Laliberte, A. S., and A. Rango. 2008. Incorporation of Texture, Intensity, Hue, and Saturation for Rangeland Monitoring with Unmanned Aircraft Imagery. *The International Archives of the Photogrammetry, Remote Sensing, and Spatial Information Sciences*. http://www.isprs.org/proceedings/xxxviii/4-c1/sessions/Session4/6555_Laliberte_Proc.pdf.

———. 2011. Image Processing and Classification Procedures for Analysis of Sub-Decimeter Imagery Acquired with an Unmanned Aircraft over Arid Rangelands. *GIScience & Remote Sensing* 48 (1): 4–23. doi:10.2747/1548-1603.48.1.4.

Laliberte, A. S., C. Winters, and A. Rango. 2008. A Procedure for Orthorectification of Sub-Decimeter Resolution Imagery Obtained with an Unmanned Aerial Vehicle (UAS). In *Proceedings from the American Society of Photogrammetry and Remote Sensing Annual Conference*, 08–047. http://jornada-data.nmsu.edu/bibliography/08-047.pdf.

———. 2011. UAS Remote Sensing Missions for Rangeland Applications. *Geocarto International* 26 (2): 141–56. doi:10.1080/10106049.2010.534557.

Lauenroth, W. K., I.C. Burke, and M.P. Gutmann. 1999. The Structure and Function of Ecosystems in the Central North American Grassland Region. *Great Plains Research*, 223-259.

Lillesand, T. M., R. W. Kiefer, and J. W. Chipman. Digital Image Interpretation and Analysis. In *Remote Sensing and Image Interpretation*. 6th ed. Hoboken, NJ: John Wiley & Sons, 2008.

López-Granados, F., J. Torres-Sánchez, A. Serrano-Pérez, A. I. de Castro, F. Mesas-Carrascosa, and J. Peña. 2015. Early Season Weed Mapping in Sunflower Using

UAS Technology: Variability of Herbicide Treatment Maps Against Weed Thresholds. *Precision Agriculture*, August. doi:10.1007/s11119-015-9415-8.

Lund, H. G. 2007. Accounting for the World's Rangelands. *Rangelands* 29 (1): 3-10. doi:10.2111/1551-501X(2007)29[3:AFTWR]2.0.CO;2.

Mansour, K., O. Mutanga, E. Adam, and E. M. Abdel-Rahman. 2015. Multispectral Remote Sensing for Mapping Grassland Degradation Using the Key Indicators of Grass Species and Edaphic Factors. *Geocarto International*, July, 1–15. doi:10.1080/10106049.2015.1059898.

Martin, M. E., J. D. Aber. 1997. High Spectral Resolution Remote Sensing of Forest Canopy Lignin, Nitrogen, and Ecosystem Process. *Ecological Applications* 7 (2): 431-443. doi:10.1890/1051-0761(1997)007[0431:HSRRSO]2.0.CO;2.

Martin, M. E., S. D. Newman, J. D. Aber, and R. G. Congalton. 1998. Determining Forest Species Composition Using High Spectral Resolution Remote Sensing Data. *Remote Sensing of Environment* 65 (3): 249–54. doi: 10.1016/S0034-4257(98)00035-2

Matthews, B. W. 1975. Comparison of the Predicted and Observed Secondary Structure of T4 Phage Lysozyme". *Biochimica et Biophysica Acta (BBA) - Protein Structure*. *405* (2): 442–451.

Mitchell, J. J., N. F. Glenn, M. O. Anderson, R. C. Hruska, A. Halford, C. Baun, and N. Nydegger. 2012. Unmanned Aerial Vehicle (UAS) Hyperspectral Remote Sensing for Dryland Vegetation Monitoring. In *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2012 4th Workshop On*, 1–10. IEEE. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6874315.

Mulla, D. J. 2012. Twenty Five Years of Remote Sensing in Precision Agriculture: Key Advances and Remaining Knowledge Gaps. *Biosystems Engineering* 114 (4): 358–71. doi:10.1016/j.biosystemseng.2012.08.009.

Mutanga, O., and A. K. Skidmore. 2007. Red Edge Shift and Biochemical Content in Grass Canopies. *ISPRS Journal of Photogrammetry and Remote Sensing* 62 (1): 34–42. doi:10.1016/j.isprsjprs.2007.02.001.

Nitze, I., U. Schulthess, & H. Asche. (2012). Comparison of Machine Learning Algorithms Random Forest, Artificial Neural Network and Support Vector Machine to Maximum Likelihood for Supervised Crop Type Classification. Proc. of the 4th GEOBIA, 7-9.

Ohlenbusch, P.D. and S.L. Watson. 1994. Stocking Rate and Grazing Management. *Kansas State University*.

Oksanen, J., F.G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P.R. Minchin, R.B. O'Hara, G.L. Simpson, P. Solymos, M. Henry, H. Stevens, E. Szoecs, and H. Wagner. 2017. vegan: Community Ecology Package. R package version 2.4-2. https://CRAN.R-project.org/package=vegan

Patterson, M. P. and L.B. Best. 1996. Bird abundance and nesting success in Iowa CRP fields: the importance of vegetation structure and composition. *American Midland Naturalist*, 153-167.

Powers, D.M.W. 2011. Evaluation: From Precision, Recall, and F-Measure to ROC, Informedness, Markedness and Correlation. *Journal of Machine Learning Technologies, 2*(1): 37-63.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rango, A., A. Laliberte, K. Havstad, C. Winters, C. Steele, and D. Browning. 2010. Rangeland Resource Assessment, Monitoring, and Management Using Unmanned Aerial Vehicle-Based Remote Sensing. In *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, 608–11. IEEE. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5651659.

Rango, A., A. Laliberte, J. E. Herrick, C. Winters, K. Havstad, C. Steele, and D. Browning. 2009. Unmanned Aerial Vehicle-based Remote Sensing for Rangeland Assessment, Monitoring, and Management. *Journal of Applied Remote Sensing* 3 (1): 033542–033542. doi:10.1371/journal.pone.0109881.

Rango, A., A. Laliberte, C. Steele, J. E. Herrick, B. Bestelmeyer, T. Schmugge, A. Roanhorse, and V. Jenkins. 2006. Using Unmanned Aerial Vehicles for Rangelands: Current Applications and Future Potentials. *Environmental Practice* 8 (03): 159–68. doi: 10.1017/S1466046606060224.

Skidmore, A. K., J. G. Ferwerda, O. Mutanga, S. E. Van Wieren, M. Peel, R. C. Grant, H. H. T. Prins, F. B. Balcik, and V. Venus. 2010. Forage Quality of Savannas — Simultaneously Mapping Foliar Protein and Polyphenols for Trees and Grass Using Hyperspectral Imagery. *Remote Sensing of Environment* 114 (1): 64–72. doi:10.1016/j.rse.2009.08.010.

Stødle, P., N. Borch, and R. Storvold. 2013. High-Performance Visualization of UAS Sensor and Image Data with Raster Maps and Topography in 3D. *ISPRS-Int. Arch. Photogram. Rem. Sens. Spat. Inform. Sci., XL-1 W* 2: 275–80.

Turner, D., A. Lucieer, and C. Watson. 2012. An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAS) Imagery, Based on Structure from Motion (SfM) Point Clouds. *Remote Sensing* 4 (12): 1392–1410. doi:10.3390/rs4051392.

Van Soest, P. J. *Nutritional Ecology of the Ruminant*. 2nd ed. Ithaca: Comstock Pub., 1994.

Watts, A. C., J. H. Perry, S. E. Smith, M. A. Burgess, B. E. Wilkinson, Z. Szantoi, P. G. Ifju, and H. F. Percival. 2010. Small Unmanned Aircraft Systems for Low-Altitude Aerial Surveys. *Journal of Wildlife Management* 74 (7): 1614–19. doi:10.2193/2009-425.

Xiang, H., and L. Tian. 2011. Development of a Low-cost Agricultural Remote Sensing System Based on an Autonomous Unmanned Aerial Vehicle (UAS). *Biosystems Engineering* 108 (2): 174–90. doi:10.1016/j.biosystemseng.2010.11.010.

Yu, Q., P. Gong, N. Clinton, G. Biging, M. Kelly, and D. Schirokauer. 2006. Object-Based Detailed Vegetation Classification with Airborne High Spatial Resolution Remote Sensing Imagery. *Photogrammetric Engineering & Remote Sensing* 72 (7): 799–811. doi: 10.14358/PERS.72.7.799.