Utah State University

DigitalCommons@USU

All Graduate Theses and Dissertations

Graduate Studies

12-2012

Using Biophysical Geospatial and Remotely Sensed Data to Classify Ecological Sites and States

Carson A. Stam
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Other Life Sciences Commons

Recommended Citation

Utah State University
MERRILL-CAZIER LIBRARY

USING BIOPHYSICAL GEOSPATIAL AND REMOTELY SENSED DATA TO

CLASSIFY ECOLOGICAL SITES

AND STATES

by

Carson A. Stam

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Range Science

Approved:

_____   _____
R. Douglas Ramsey   Janis L. Boettinger
Major Professor   Committee Member

_____   _____
Eugene W. Schupp   Mark R. McLellan
Committee Member   Vice President for Research and
                                          Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2012

ABSTRACT

Using Biophysical Geospatial and Remotely Sensed Data to Classify

Ecological Sites and States

by

Carson A. Stam, Master of Science

Utah State University, 2012

Major Professor: Dr. R. Douglas Ramsey
Department: Wildland Resources

Monitoring and identifying the state of rangelands on a landscape scale can be a time consuming process. In this thesis, remote sensing imagery has been used to show how the process of classifying different ecological sites and states can be done on a per pixel basis for a large landscape.

Twenty-seven years' worth of remotely sensed imagery was collected, atmospherically corrected, and radiometrically normalized. Several vegetation indices were extracted from the imagery along with derivatives from a digital elevation model. Dominant vegetation components from five major ecological sites in Rich County, Utah, were chosen for study. The vegetation components were Aspen, Douglas-fir, Utah juniper, mountain big sagebrush, and Wyoming big sagebrush. Training sites were extracted from within map units with a majority of one of the five ecological sites.

A Random Forests decision tree model was developed using an attribute table populated with spectral biophysical variables derived from the training sites. The overall

out-of-bag accuracy for the Random Forests model was 97.2%. The model was then applied to the predictor spectral and biophysical variables to spatially map the five major vegetation components for all of Rich County. Each vegetation class had greater than 90% accuracies except for Utah juniper at 81%. This process is further explained in chapter 2.

As a follow-on effort, we attempted to classify vegetation ecological states within a single ecological site (Wyoming big sagebrush). This was done using field data collected by previous studies as training data for all five ecological states documented for our chosen ecological site. A Maximum Likelihood classifier was applied to four years of Landsat 5 Thematic Mapper imagery to map each ecological state to pixels coincident to the map units correlated to the Wyoming big sagebrush ecological site. We used the Mahalanobis distance metric as an indicator of pixel membership to the Wyoming big sagebrush ecological site. Overall classification accuracy for the different ecological states was 64.7% for pixels with low Mahalanobis distance and less than 25% for higher distances.

(96 pages)

PUBLIC ABSTRACT

Using Biophysical Geospatial and Remotely Sensed Data to Classify

Ecological Sites and States

Within the Intermountain West, vast expanses of big sagebrush shrubland and steppe are considered emblems of the western range. Currently, there are approximately 60 million hectares of big sagebrush within the 11 western states, four million of which are in the state of Utah. However, the historic distribution of sagebrush has been impacted by conversion to other types of land cover through juniper encroachment, urbanization, invasive weeds, and agricultural expansion. In Utah alone, big sagebrush communities have been reduced to approximately 55% of their historic extent. A primary and current example of the cumulative impact of big sagebrush loss is the eminent listing of the Sage Grouse as an endangered species. This potential listing will force land management agencies to impose strict guidelines for future development of sagebrush-dominated landscapes. These growing pressures have led to a need to accurately estimate the actual and potential spatial distribution of sagebrush shrubland and steppe and their current ecological condition.

The Utah State University Remote Sensing and Geographic Information Systems laboratory proposed a two-year study to develop and demonstrate methods of ecological assessment using satellite and aerial imagery. This project will show how common remote sensing tools can help in the identification of unique ecological sites across an entire landscape. Ecological site descriptions describe the historic plant communities and

soils that existed on an ecological site (ES). Therefore, classifying ESs will allow land

managers to understand the potential vegetation communities that can exist at a site.

Because much of the historic vegetation in the Intermountain West has changed to

alternative land cover types, it is also important to assess the current vegetation condition

of the landscape. A remote sensing based classification was used to identify the

ecological state of Wyoming big sagebrush communities. A method of calculating the

probability of an area belonging to the Wyoming big sagebrush ES will also be explained.

The methodology described in this research will be easily replicated by those with

minimal training in remote sensing techniques. It is expected that these methods will

benefit both public and private land managers as they seek to produce sustainable policies.

Carson Stam

DEDICATION

To my beautiful and loving wife Marissa, our son Greyson, and our other

children yet to come.

ACKNOWLEDGMENTS

I first want to thank and recognize my major professor, Douglas Ramsey. I began meeting with Doug well before I was accepted to graduate school. He has been helpful and provided much appreciated guidance and direction on my thesis. He also employed me at the RS/GIS lab during my time in Logan. There I learned valuable skills that will help me in my future employment. I also want to express my gratitude to my other committee members, Janis Boettinger and Eugene "Geno" Schupp. One of my first courses at USU was an exceptional soils class taught by Janis. She has also offered important comments to my thesis project ideas and papers. I thank Geno for being willing to be on my committee and for his valuable comments and criticisms of my work.

If there was an award for "Honorary Committee Member" Alex Hernandez would win it. Alex shared his extensive knowledge of ecology and remote sensing with me whenever I needed help. Without Alex's help on my graduate work, my thesis would never have been completed. I would also like to thank Nate Payne. Nate helped me with my field work in Rich County, UT. I couldn't have asked for a better research assistant.

Thank you to everyone at the RS/GIS lab. Chris McGinty helped me with several details of my project and helped me get ready for my field work. He also included me on several projects for the lab. Chris Garrard taught one of my favorite classes I have ever taken. The things I learned in that programming class have and will continue to help me. She was also willing to help me with questions whenever I had them.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In the recent history of land management, ground-based techniques have been
used to monitor and assess the condition of ecosystems. The information gathered using
ground-based techniques is often extrapolated to larger landscapes. The application of
these techniques is usually sparse in time and space, leading to a mischaracterization of
the landscape (Pringle et al. 2006). This problem presents a challenge to natural resource
managers responsible for assessing and taking action to improve or maintain the
ecological condition of landscapes. Forbis et al. (2007) stated that one of the main, initial
issues facing resource managers is to quantitatively assess the ecological condition of
landscapes using limited financial resources which translates into limited field sampling
efforts. The subject of our research has been to investigate remote sensing methods and
strategies that can identify the potential and current ecological condition of a landscape.
Remote sensing is recognized as a cost-effective method for identifying ecological
conditions across large landscapes (Mumby et al. 1999). In fact, remote sensing is now
critical to the successful modeling of many natural resource processes (Jensen 2000). A
major consideration when using remote sensing to monitor ecological condition is the
contextual framework within which spectral data is interpreted. For this work, we have
opted to use a landscape level framework developed by the USDA Natural Resources
Conservation Service (NRCS).

The NRCS has been systematically classifying rangelands into ecological sites
(ES) that link soil characteristics to the defined historic plant community occupying that
soil. Ecological site descriptions (ESD) describe areas of specific biophysical properties

and associated plant communities that may be found at a given ES. These sites differ from other sites in their ability to produce a distinct kind and amount of vegetation. Areas of the same ES, but separated by geography, are also unique in that they are assumed to "respond similarly to management actions and natural disturbances" (U.S. Department of Agriculture, NRCS 2011). Ecological sites are primarily determined on the basis of soil characteristics and the resulting differences in plant species composition and production that occur on those soils.

Currently, ESs are only identified on a landscape as components within map units (MU). An MU is a spatially defined area that defines the soil characteristics at a location. A given MU can contain one or more different soil types that are termed components. Components are contiguous groupings of different soils whose extents are equal to or smaller than the MU. Map unit polygons therefore have a one-to-many relationship with ESs (Arid Land Research Programs 2010). The spatial and tabular data for MUs are stored in individual soil surveys and can be obtained from the NRCS SSURGO database (U.S. Department of Agriculture, NRCS-SSURGO 2012). Up to four different ecological site components (one per soil type) are combined into one MU and the SSURGO tabular database details the percentage of area each component occupies within a given MU; however, the database does not define the spatial location of a particular ES component within the MU.

Vegetation communities exist across their geographic distribution in various ecological states. These states can be viewed as nuances in community structure due to local environmental factors, or they can represent alterations forced by management actions or changes in climate. Information about the different ecological states that

communities can occupy, as well as the forces that promote the transitions between states, can be enumerated in state-and-transition models (STMs) (Westoby et al. 1989). These transitions can take place due to soil erosion, fire regimes, weather variability, and management (Briske et al. 2005). Westoby et al. (1989) suggested that the purpose of an STM are to 1. Define the states possible within a system, 2. Catalogue management action and other forces that drive transitions from one state to another, and 3. List the actions that could produce favorable transitions as well as the hazards of inaction that could result in unfavorable transitions. A state is defined as a recognizable, resistant, and resilient complex of soil base and vegetation structure (Stringham et al. 2003). The original STM framework did not indicate a need to identify a reference state. However, STMs adopted by the NRCS have been joined with the traditional range model so that STMs developed by the NRCS include a reference state that characterizes the historic plant community (Briske et al. 2005).

This thesis is composed of two substantive chapters bounded by this introduction and overall conclusion chapters. In chapter 2, we test whether a multi-temporal dataset of Landsat 5 Thematic Mapper (TM) imagery can be used in conjunction with a decision tree classifier to map the vegetation components of ESs within map units. Landsat 5 TM imagery was collected for a 26 year span. Each image was atmospherically corrected and normalized using an image-based method (Chavez 1996). Several remote sensing variables and topographic variables were explored for their ability to separate ES vegetation components. A cluster analysis was conducted to determine whether there was natural structure in the data that would allow for discrimination between vegetation types. A Random Forests model was developed and applied to a set of image and

topographic predictor variables to map the spatial distribution of ESs on a pixel basis. This ability to predict ecological sites on a pixel basis has been suggested as the next step in remote sensing applications to rangeland conservation (Hernandez 2011). With the knowledge of where these ecological sites can occur, resource managers are then able to understand the distribution of resources and the ecological potential of sites. This information will lead to better-informed management decision making.

In chapter 3, we explored whether different ecological states could be classified within Wyoming big sagebrush ecological sites. Field data collected by Peterson (2009) and the Utah Division of Wildlife Resources (2006) were used to train the classifier to map the different ecological states. A Maximum Likelihood classifier was used to classify a temporal image stack of TM imagery spanning four continuous years (2005-2008) into different ecological states. A Mahalanobis distance metric was calculated to estimate the probability of a pixel belonging to a specific ecological state. Field work was done to 1) assess the accuracy of our ecological state classification and 2) determine whether the Mahalanobis distance was a suitable indicator of membership in a Wyoming big sagebrush ecological site. The implications of the classification accuracies as well as the suitability of using Mahalanobis distance as a similarity metric are discussed.

LITERATURE CITED

Arid Land Research Programs. 2010. *Jornada Experimental Range*. Available at: http://jornada.nmsu.edu/esd/intro. Accessed 10 May 2012.

Briske, D. D., S.D. Fuhlendorf, and F. E. Smeins. 2005. State-and-transition models, thresholds, and rangeland health: a synthesis of ecological concepts and perspectives. *Rangeland Ecology & Management* 58:1-10.

Chavez, P. S. 1996. Image-based atmospheric corrections revisited and improved. *Photogrammetric Engineering and Remote Sensing* 62:1025-1036.

Forbis, T. A., L. Provencher, L. Turner, G. Medlyn, J. Thompson, and G. Jones. 2007. A method for landscape-scale vegetation assessment: application to Great Basin rangeland ecosystems. *Rangeland Ecology & Management* 60:209-217.

Hernandez, A. 2011. Spatiotemporal modeling of threats to big sagebrush ecological sites in northern Utah [dissertation]. Logan, UT: Utah State University. 163 p.

Jensen, J. R. 2000. Remote sensing of the environment: an earth resources perspective. Upper Saddle River, NJ: Prentice-Hall. 592 p.

Mumby P.J., E. P. Green, A. J. Edwards, and C. D. Clark. 1999. The cost-effectiveness of remote sensing for tropical coastal resources assessment and management. *Journal of Environmental Management* 55:157-166.

Peterson, K. A. 2009. Modeling potential native plant species distributions in Rich County, Utah [thesis]. Logan, UT: Utah State University. 107 p.

Pringle, H. J., I. W. Watson, and K. L. Tinley. 2006. Landscape improvement, or ongoing degradation – reconciling apparent contradictions from the arid rangelands of Western Australia. *Landscape Ecology* 21:1267-1279.

Stringham, T. K., W. C. Krueger, and P. L. Shaver. 2003. State and transition modeling: an ecological process approach. *Journal of Range Management* 56:106-113.

U.S. Department of Agriculture, NRCS. 2011. National soil survey handbook. Available at: http://soils.usda.gov/technical/handbook/ contents/part622.html#07. Accessed 5 January 2012.

U.S. Department of Agriculture, NRCS-SSURGO. 2012. Soil data mart. Available at: http://soildatamart.nrcs.usda.gov/. Accessed 10 October 2011.

Utah Division of Wildlife Resources. 2006. Range trend studies. Available at: http://wildlife.utah.gov/range/wmu4.htm. Accessed 10 April 2012.

Westoby, M., B. Walker, and I. Noy-Meir. 1989. Opportunistic management for rangelands not at equilibrium. *Journal of Range Management* 42:266-274.

CHAPTER 2

MAPPING VEGETATION COMPONENTS OF ECOLOGICAL SITES:

A REMOTE SENSING APROACH

INTRODUCTION

Ecological Site Descriptions (ESD) as defined by the Natural Resources

Conservation Service (NRCS), characterize sites of specific biophysical properties and

plant communities. These sites differ from other kinds of land in their ability to produce

a distinctive kind and amount of vegetation. Areas of the same ecological site (ES) are

also unique in that they will "respond similarly to management actions and natural

disturbances" (U.S. Department of Agriculture, NRCS 2012a). Ecological sites are

correlated on the basis of soils, geomorphology, hydrology, and the resulting differences

in plant species composition that occur on those soils. Because ESDs are based on the

plant community that existed at the time of European settlement (U.S. Department of

Agriculture, NRCS 2011), ESDs represent reference states for State and Transition

Models (STM).

Each complete ESD has an associated STM. The purposes of an STM are to 1.

Define the alternative stable states possible within a system 2. Catalogue the transitions

from one state to another including the conditions which induce the transitions and 3. List

the management actions that could produce favorable transitions as well as the hazards of

inaction that could produce unfavorable transitions (Westoby et al. 1989). A state is

defined as a recognizable, resistant, and resilient complex of soil base and vegetation

structure (Stringham et al. 2003). The original STM framework does not indicate the

need to identify a reference state. However, STMs developed by the Natural Resources

Conservation Service (NRCS) have been joined with the traditional range model so that

these STMs include a reference state that refers to the historic (pre-Columbian) plant

community (Briske et al. 2005). An ESD, therefore, is an important component of an

STM because it defines the reference state. Briske et al. also stated that ESDs are a

"critical feature of state-and-transition models because the descriptions provide the

interpretive information associated with these models" (p. 5).

Currently, ESs are spatially identified as components within map units (MU). An

MU is a spatially defined area that defines the soil characteristics at that location. A

given MU can contain one or more different soil types that are termed components.

Components are contiguous groupings of different soils whose extents are smaller than

the minimum mapping unit of the MU. The percentage of area each component occupies

within an MU is documented; however, the spatial location of a specific component

within an MU is not defined.

Bestelmeyer et al. (2009) formulated an approach to develop and apply ecological

sites along with STMs. They suggested a spatial hierarchy system for sampling which

used imagery to identify vegetation distribution. These mapped vegetation areas could

then infer possible ecological sites and states. They suggested that Southwest Regional

Gap (SWGAP) (Prior-Magee 2007) or Landsat imagery could be used for this purpose.

Maynard et al. (2007) found that there was a high correlation between field measures of

productivity and exposed soil when compared to the tasseled cap brightness component

extracted from Landsat Thematic Mapper (TM) imagery. The tasseled cap

transformation converts reflectance values obtained through remote sensing into a set of

composite values consisting of scene brightness, greenness and wetness. The brightness component represents the general intensity of reflectance per pixel across all spectral bands in a Landsat 5 TM scene. Differences in brightness have been shown to discriminate between deciduous shrubs (or harvested forest stands) and closed canopy forests (Dymond et al. 2002).

The Normalized Difference Vegetation Index (NDVI) (Rouse et al. 1974) quantifies the amount of live green vegetation found in a remotely sensed image. Gamon et al. (1995) discussed the usefulness of the NDVI as an indicator of photosynthetic activity as well as canopy structure, and plant nitrogen content. Jensen (2000) showed that NDVI was sensitive to canopy variations including soil visible through canopy openings" (p. 386). While the sensitivity to soil background has typically been seen as a disadvantage of NDVI for vegetation assessment, it could prove useful for studying ESs because areas of the same ES may have a similar amount and type of bare soil. Since NDVI is sensitive to these differences, it should be a good index for distinguishing different ESs. The NDVI values within the polygon of a soil mapping unit and the variation in the NDVI has also been used to distinguish between cover types (Pickup and Foran 1987).

Accurately classifying and identifying the spatial extent of ESs on a landscape level is a very time consuming process. At this point in time, only extensive field work can map the spatial distribution of ESs across a landscape due to the need to properly identify soils. While remotely sensed data cannot yet be used to obtain detailed data about soils, it can be used to identify the unique vegetation components of ESs. Being able to accurately identify the vegetation component of ESs should provide a means by

which soil field sample locations can be identified more efficiently. We postulate that using satellite derived NDVI and brightness, coupled with biophysical geospatial data (elevation, slope, and aspect) should allow areas of the same ES vegetation component to be mapped. If remote sensing indices allow for separation between ES vegetation components, then that process could help with accurately classifying the landscape into individual ESs and subsequently help with the formulation of STMs. Our objective, therefore, is to use NDVI, brightness, and biophysical geospatial data to determine whether we can accurately identify areas of the same ES vegetation component across a large landscape. This process of identifying sites using spectral and biophysical data could provide a way to identify and understand the various states an ES could occupy.

METHODS

Study Area

Our research was conducted in Rich County, Utah, located in the northeastern corner of the state (long 111°30'38.5'' – long 111°2'42.2'' West and lat 42°0'0''- lat 42°08'24.3'' North). The sites we sampled were from two Major Land Resource Areas (MLRA) including the Wasatch and Uinta Mountains (47) and Cool Central Desertic Basins and Plateaus (34A). MLRAs are classified by physiography, geology, climate, water, soils, biological resources, and land use (U.S. Department of Agriculture, NRCS 2005). The western portion of the study area is characterized by high elevations with vegetation consisting of aspen forests, subalpine conifer forests, and scattered mountain sagebrush steppe. Moving east, the elevation decreases, and the mountain sagebrush steppe becomes dominant. Both the mountain and foothills sections of the county are in

MLRA 47. Central and eastern Rich County is made up of relatively lower elevations

with vegetation consisting of basin big sagebrush steppe and shrubland, subalpine

grasslands, and agriculture. These sections of the county are in MLRA 34A.

The average elevation is 2093 m. The highest point is Bridger Peak at 2821 m

and the lowest point is about 1800 m. The climate is variable and is affected by the

changing topography of the county. The soil temperature regime is frigid and the soil

moisture regime is xeric for most of the county. North facing slopes in the higher

elevations have cryic soil temperature regimes. Higher elevations also transition to an

ustic soil moisture regime. The parent material is primarily derived from sandstone and

limestone. The large variations in elevation, slope, and climate make a detailed account

of all soils present in Rich County difficult in this document. For a detailed description

of the soils present in Rich County, visit the online NRCS Soil Survery Geographic

(SSURGO) Database (U.S. Department of Agriculture, NRCS-SSURGO 2012).

The majority of the land is in private ownership at 58.8%. The federal

government is the next largest landowner with 33.6% with land split between the Bureau

of Land Management and the U.S. Forest Service. The state of Utah owns only 7.6% of

the land area which is mostly composed of State Trust Lands (Utah Office of Tourism

2009). Disturbances that have affected the area include agriculture, grazing, logging, and

burning.

Biophysical Geospatial Datasets

A series of Landsat 5 TM images (Path 38/ Row 31) for each year between 1984-

2011 with Julian date as close to 207 (July 26th) as possible was collected from the U.S.

Geological Survey Global Visualization Viewer (GLOVIS). The Julian date of 207 was

chosen by averaging the date for each year that displayed the greatest variance in NDVI

between different land cover types. The dates were obtained by examining line graphs of

mean NDVI values collected by AVHRR of evergreen forests, shrubs, and deciduous

forests. These graphs can be obtained through GLOVIS using a tool called "NDVI

graph" (U.S. Geological Survey 2011). Figure 2-1 is an example of one of these graphs

from 2009. Images with minimal cloud cover and collection dates closest to the Julian

date 207 were selected. Of the 28 years' images, 18 were within 20 days of 207, 5 more

were within 30 days of 207, and 3 more were within forty days of 207. The cloud free

scene closest to Julian date 207 from 1987 had a Julian date of 153 and was 54 days off.

The year 2001 was the only year that a late spring or summer image was not available

due to cloud cover.

All images were rectified and resampled to UTM Zone 12 NAD 1983 map

projection. Each image's raw digital numbers were converted to reflectance values using

an image-based atmospheric correction (Chavez 1996) and the calibration coefficients for

Landsat 5 TM (Chander et al. 2009). Following image standardization, we calculated

NDVI using the formula (NIR - RED) / (NIR + RED). We then used a 5 x 5 pixel

(22,500 m^2 ground area) focal window to calculate the standard deviation in NDVI for

each pixel. A 5x5 focal window was not used in calculating NDVI because it was not

necessary and doing so would only decrease the spatial accuracy of the NDVI values.

The brightness component was calculated using the published transformation coefficients

for the Landsat 5 TM imagery for each year (Crist and Cicone 1984). These variables

were collected for multiple years based on literature indicating that longer time series of

remotely sensed data were necessary to adequately characterize different ecological states

due to inherent year-to-year variance (Hernandez 2011).

A 30 m digital elevation model (DEM) was obtained from the Automated

Geographic Reference Center (2011) for Rich County. Slope and aspect were then

calculated using Spatial Analyst in ArcMap^{TM}. Elevation, slope, and aspect have been

shown to determine the microclimate and therefore the spatial distribution and patterns of

vegetation (Jin et al. 2008).

Ecological Sites

For this study, five ES vegetation components were selected. They included

Wyoming big sagebrush (*Artemisia tridentata* ssp. *wyomingensis*), mountain big

sagebrush (*Artemisia tridentata* ssp. *vaseyana*), Utah juniper (*Juniperus osteosperma*),

Douglas-fir (*Pseudotsuga menziesii*), and aspen (*Populus tremuloides*). With the

exception of Utah juniper, these vegetation components were selected because of their

prevalence in the county. Wyoming big sagebrush accounts for much of the vegetation in

MLRA 34A, and MRLA 47 is mostly comprised of aspen, Douglas-fir, and mountain big

sagebrush. Utah juniper is not prevalent in either MRLA; however, we thought it an

important vegetation component to classify due to its potential encroachment into

sagebrush steppe communities (Miller and Rose 1999). Together, these vegetation

components represent approximately 71% of the county by area.

Map unit (MU) spatial and tabular data were obtained from the NRCS SSURGO

database. For the purposes of this study, we selected MU's which were predominantly

made up of one of our targeted components (70% areal composition). This was done to

help ensure that sites chosen for sampling would have low spatial soil and land cover variability. Land cover data from the SWGAP analysis were used to identify MU's that represented the defined ES. For instance, an MU was selected that consisted of a >=70% component Wyoming big sagebrush. If the SWGAP analysis land cover also identified the area as containing a big sagebrush land cover class, then that MU was used for this study. Twenty polygons were digitized for each ES vegetation component of interest using the intersection of the SSURGO and SWGAP data and the visible boundaries of the vegetation component as photointerpreted from the 2009 National Agricultural Imagery Project (NAIP) 1m resolution orthoimagery. In total, one-hundred polygons were created (20 for each ES vegetation component).

Cluster Analysis and Dataset Preparation

We applied a cluster analysis to determine if the spectral and biophysical characteristics of our 100 training polygons would allow us to separate each vegetation type from the others. Cluster analysis was conducted to determine if there was natural structure in the data that would allow separation between dissimilar ES vegetation types. Cluster analysis is suited for this task because it does not take into account any training data. Clusters are created based solely on the distance, in n-dimensional space, of one cluster to another. For our purposes, we used the agglomerative hierarchical clustering (AHC) method. AHC starts with n clusters where each initial observation is its own single observation cluster. On the first iteration, the two closest observations are merged into a composite cluster so that there then exists $n - 1$ clusters. This process continues until there is one cluster that contains all observations. The distance between clusters can

be defined multiple ways in AHC. The most common are single-link, complete-link, and average-link clustering. Single-link clustering measures the distance of the two most similar observations within a cluster. Complete-link clustering measures the distance of the two most dissimilar observations within a cluster. Average-link clustering measures the distance between each observation in a cluster and all the observations in another cluster. The two clusters with the lowest average distance are combined to form a new cluster. There are drawbacks to each method. The single-link method is sensitive to noise and outliers. The complete-link method is not sensitive to noise and outliers, but can break large clusters into smaller clusters. The average-link method is a compromise between the two (Kotsiantis and Pintelas 2004). We chose to use the average-link method because of this compromise.

Polygons were intersected with the topographic data layers, yearly NDVI imagery, and yearly brightness component images. For each polygon, the mean values of topographic and brightness variables were extracted along with the mean and standard deviation of each NDVI image. Instead of including the brightness component, NDVI, and standard deviation of NDVI for each year for each polygon in our data matrix, we created 5-year averages for these variables. This was done to minimize the effects of interannual climate variability and clouds. Interannual climate variability has been shown to affect some plant species productivity (Goulden et al. 1996; Arain et al. 2002) and ecologic processes (Westerling and Swetnam 2003). The resulting data matrix was therefore composed of the ES vegetation component name followed by three columns for the DEM derivatives, five sets of 5-year averages for the remotely sensed variables, and one set of 3-year (2009 – 2011) averages for the remotely sensed variables. Because our

variables contained different units of measurement (degrees, meters, and vegetation indices), we normalized each variable by subtracting the mean of that variable from the actual value and dividing by the standard deviation (Sakrejda-Leavitt 2009). To perform cluster analysis, we used R code written by Everitt and Hothorn (2010).

Random Forests

The purpose of running cluster analysis on the data matrix was to determine if there was enough structure in the data to spatially map ES vegetation components using these variables. If we determined that there was structure to the data, it was then our goal to develop a decision tree model utilizing these data to map the distribution of our selected ESs across the study area. We chose Random Forests (Breiman 2001) for its high accuracy in ecological applications (Cutler et al. 2007), automatic variable selection, and generation of an internal unbiased estimate of the generalization error. We also wanted the ability to interpret what variables were most important in deriving the decision tree model. Random Forests is well suited to this task because of its easy to produce variable importance plots. Random Forests uses a bootstrap sample of the dataset to "fit" several classification trees. Observations not included in the bootstrap sample are called out-of-bag observations. Each fitted classification tree is then used to predict the out-of-bag observations. The out-of-bag accuracy (cross-validation) is calculated for each observation using the out-of-bag predictions (Cutler et al. 2007). This process is repeated hundreds of times until a final classification and cross-validation accuracy is produced.

Vegetation type (i.e. aspen, Douglas-fir, mountain big sagebrush, Wyoming big sagebrush, and Utah juniper) was used as the class variable to be predicted while the remotely sensed and topographic variables were used as predictors. We used the default 500 iterations as our bootstrap.

Image Classification and Validation

After a Random Forests decision tree model was created, we applied it to a geospatial data stack of Rich County using the image imputation package in R (Crookston and Finley 2008). This geospatial data stack contained the normalized variables (see the Cluster Analysis and Data Set Preparation section above) used to develop the model including the multiple year averages, NDVI layers, the matching spatial variance layers, matching brightness layers, as well as the topographic layers. The output of the image imputation package (Fig. 2-2) was assessed for accuracy by generating random points within each class. Fifty random points were generated within the classified areas for each vegetation type. Each point was validated using NAIP 1 m resolution imagery.

RESULTS

The 100 polygons representing the five different vegetation components (20 each) varied in size. The smallest polygon was approximately 4 acres and the largest was 124 acres. The reason for this range of area is that some ESs had larger areas of contiguous coverage (e.g. Wyoming big sagebrush) while others had smaller areas of contiguous coverage (e.g. Utah juniper). Area did not vary as much within a given ES. Table 2-1

contains the averaged spectral values and topographic data as well as polygon size collected for each polygon.

An exploratory analysis was conducted to determine whether vegetation components had unique NDVI and brightness values. A series of graphs plotted each observation (polygon) against different variables. Figure 2-3 shows the 28-year mean of the average NDVI value for each polygon plotted against the 28-year mean of the standard deviation of NDVI for each polygon. This analysis showed that our selected vegetation components occupied unique NDVI mean and spatial variance regions. Some overlap occurred between Wyoming big sagebrush and Utah juniper and between Douglas-fir and aspen ESs. We then tested whether brightness could also help separate the five vegetation types. This was done by plotting each observation on a graph continuing to use the 28-year mean NDVI on the x-axis and 28-year mean brightness on the y-axis (Fig. 2-4). The brightness component was able to cleanly separate Aspen polygons from the Douglas-fir polygons. However, brightness provided little separation between Utah juniper and Wyoming big sagebrush.

We plotted each polygon against elevation and slope (Fig. 2-5) and also against elevation and the cosine of aspect (Fig. 2-6). Topographic variables alone were able to somewhat separate vegetation components along an elevation gradient (as expected). Slope seemed to be a good variable to separate Utah Juniper from Wyoming big sagebrush and Douglas-fir from Aspen. Aspect was not useful for distinguishing between any vegetation types. Every vegetation component had observations with wide ranges of aspect that overlapped dissimilar vegetation component observations. Because

aspect did not seem to separate any vegetation components, it was omitted from the data matrix when performing cluster analysis.

The areas in spectral space that the 20 samples from each vegetation component occupied (Figs 2-2 and 2-3) were where we anticipated they would be. The Wyoming big sagebrush polygons had low greenness and low spatial variation in greenness. Utah juniper sites had similarly low average greenness, but due to high contrast between green juniper trees and a relatively larger amount of bare ground, these sites had higher spatial variation in greenness. Mountain big sagebrush had higher average NDVI values. This was expected since mountain big sagebrush occurs in higher elevations that receive more precipitation than either Wyoming big sagebrush or Utah juniper and therefore is associated with higher plant production. Aspen polygons tended to have higher NDVI values than Douglas-fir polygons with both ES vegetation components having a similar, relatively large distribution of spatial variance.

Cluster Analysis

Figure 2-7 shows a graphical representation of the cluster analysis for the one-hundred vegetation component polygons using the average-link method. Each time large clusters were created, the data was closely examined to determine whether observations with like vegetation components were being agglomerated. Most of the aspen observations were in one cluster that contained 16 of the 20 aspen observations. All 20 Douglas-fir observations were present in one cluster. Seventeen of the 20 mountain big sagebrush sites were in one cluster. Eighteen of the 20 Utah juniper polygons were present in one cluster that also contained 2 Wyoming big sagebrush polygons. The last

large cluster contained 18 of the 20 Wyoming big sagebrush polygons. Besides these

large clusters, two smaller clusters were also formed that contained four observations

each. One of these small clusters contained one Utah juniper polygon and three mountain

big sagebrush polygons. This small cluster was appended to the cluster formed by the

large Wyoming big sagebrush and Utah juniper clusters. The other small cluster

contained four aspen observations. This small cluster was appended to the large

mountain big sagebrush cluster. One lone Utah juniper observation was also appended to

the large mountain big sagebrush cluster.

As seen in Figure 2-7, the linkages between the large Utah juniper and Wyoming

big sagebrush clusters, the large aspen, Douglas-fir, and mountain big sagebrush clusters

were the last agglomeration to occur. This means that these ES vegetation components

were the most distant from each other in terms of spectral and biophysical space. This is

not surprising due to the difference in elevation and precipitation between these groups.

That break also loosely represents the division between the two MRLAs present in Rich

County.

Cluster Analysis Validation

The validation of the cluster analysis was done to 1) Make sure that each polygon

accurately represented the vegetation component that we were classifying them as, and 2)

Determine why some sites (two Utah juniper, four aspen, three mountain big sagebrush,

and two Wyoming big sagebrush) were not clustered with the rest of their respective

observations. A site being classified as a different vegetation type meant that the site was

more similar to a vegetation component of different type than to its own. Validation of

vegetation type was performed using high resolution NAIP imagery for all ESs. Close examination of the DEM derivatives for all observations was also done to explain the incorrect clustering of certain observations.

It was found that the clustering of two Utah juniper sites with mountain big sagebrush clusters was caused by a low brightness component values. These two sites are mostly on west facing slopes that would have been shaded during image acquisition. The rest of the Utah juniper sites were characterized by relatively higher brightness values compared to mountain big sagebrush sites due to the high amount of bare soil typical of juniper sites. The Utah juniper site that was clustered with the large mountain big sagebrush cluster had a higher standard deviation in NDVI than the rest of the Utah juniper sites. This juniper site straddles a ridge so that it has both north and south facing slopes. The multiple topographic aspects within this polygon caused the high spatial variance in NDVI. Along with brightness value, the high standard deviation in NDVI made it more similar to the large cluster of mountain big sagebrush observations. The other Utah juniper site was clustered with three mountain big sagebrush sites that together were agglomerated to the combination of the large Wyoming big sagebrush and Utah Juniper clusters. These three mountain big sagebrush sites had low standard deviations in NDVI which were more typical of Wyoming big sagebrush and Utah Juniper sites as seen in Figure 2-3. The low standard deviations were a product of low variability in vegetation cover, whereas the other mountain big sagebrush sites had large percentages of bare ground cover which increased the standard deviation in NDVI for those sites.

The two Wyoming big sagebrush sites that were clustered with the large Utah Juniper cluster were a product of having high slopes and slightly higher standard deviations in NDVI. Several other sites had similarly high standard deviations or high slopes, but no other Wyoming big sagebrush sites had both of these conditions.

Four aspen sites were clustered together and then added to the large cluster of mountain big sagebrush observations. These sites had relatively high standard deviations when compared to the majority of aspen sites. These sites also had slightly lower NDVI values. Three of these sites appeared to have lower aspen canopy cover. The other site contained a mix of immature aspen trees and shrubs which caused high standard deviation values.

Random Forests

Due to the relatively clean separation of types as shown by the simple cluster analysis, the resulting cross-validation accuracy of our decision tree model derived from Random Forests was approximately 97.2%. We note that because Random Forests uses an iterative process that employs a random sub-sample of the training data to fit multiple classification trees, cross-validation accuracies change slightly with each Random Forests analysis. We therefore have reported the average cross-validation accuracy produced from 20 independent runs of Random Forests. The standard deviation of the cross-validation accuracies from these 20 runs was 0.616. The model with the most conservative estimate of cross-validation accuracy was 96% accurate. This model resulted in three Utah juniper polygons incorrectly classified as Wyoming big sagebrush and one Wyoming big sagebrush polygon incorrectly classified as Utah juniper. These

incorrect classifications were not surprising given the results of the cluster analysis and the visible overlap in mean and spatial variance in NDVI and brightness values for these two vegetation components (Figs. 1-2 and 1-3).

We were also interested in which variables were most important in the development of the decision tree model. To determine variable importance, random values are substituted in place of the original values for a specific variable for each out-of-bag observation. The difference between the misclassification rate for the modified and original out-of-bag data, divided by the standard error, is the measure of variable importance (Cutler et al. 2007). Because of the way these values are computed, they can be thought of as z-scores. Variable importance was calculated for each variable in our model and the results plotted on a variable importance plot (Fig. 1-7). This graph ranks the variables (top to bottom on vertical axis) according to the "mean decrease in accuracy" caused by the substitution of that variable with random numbers. Of the 21 predictor variables (six 5-year averages each for NDVI, standard deviation in NDVI, and brightness, as well as elevation, slope, and aspect), NDVI variables were generally ranked highest in importance (occupying the 1-5 and 7 rank values), the standard deviation in NDVI variables were ranked 12, 14, 15, 17, 18, and 20 and the brightness variables were ranked 6, 9-11, 13, and 16. Slope was ranked 8, elevation was ranked 19, and aspect was ranked 21. It was unsurprising that aspect had the lowest variable importance since it also showed the least visual separation between our sampled vegetation components (Fig. 1-5).

We calculated a correlation matrix for all variable pairs and determined that all combinations of like variables (e.g. comparing each NDVI variable to each other) were

significantly correlated. Most of the correlations produced Pearson coefficients greater than or equal to 0.92. These high correlations suggested that only one 5-year group is needed for accurate classification. To confirm this, we fit several Random Forests classifications with random combinations of only one variable per NDVI, standard deviation in NDVI, and the brightness component. Each of the out-of-bag accuracies for these Random Forests classifications was equally accurate with our initial Random Forests model using several multi-year variables.

Image Classification Accuracies

The results of the accuracy assessment are summarized in Table 2-2. Because there were a few vegetation component classes that we did not account for in our model (e.g. black sagebrush, mountain mahogany, shadscale) that were present in Rich County, we expected many errors of commission (i.e. identifying a pixel as belonging to a vegetation type that does not belong to that vegetation type). We tried to limit these by only performing the accuracy assessment within the MUs that were predominantly made up of one of our five ESs. However, because there were still minority components within virtually every MU, these other vegetation components not accounted for in our model still occurred in our accuracy assessment. These errors are summarized in the "Other" row of Table 2-2. Since these errors were due to vegetation components not accounted for in our model, and therefore, not due to the inability of our model to discriminate between these types, we did not use these errors in the calculation of the percent correctly classified for each of our target vegetation types.

Those vegetation components not accounted for in our model were primarily classified as Utah juniper. Utah juniper also had the lowest percent correctly classified (81%) due to confusion with Wyoming big sagebrush. This was not surprising given the obvious overlap with certain variables. The percent correctly classified for Mountain big sagebrush was 95%. Wyoming big sagebrush had the highest percent correctly classified (98%). Only 2% of pixels classified as Wyoming big sagebrush belonged to the "Other" category. Douglas-fir and Aspen had similarly high percent correctly classified measures with 96% and 94%, respectively. When not omitting the error introduced by vegetation types not accounted for in our model, the user's accuracies decreased. The user's accuracies for each ES were as follows: Utah Juniper 44%, mountain big sagebrush 84%, aspen 92%, Wyoming big sagebrush 96%, and Douglas-fir 96%.

DISCUSSION

Identifying ES components of MUs on a landscape scale can be very time consuming. Remote sensing offers a cost-efficient alternative and has been found to be effective in evaluating the spatial dynamics of large landscapes (Brandon et al. 2003; Hunt et al. 2003; Washington-Allen et al. 2006). We have shown that using variables derived from remotely sensed images as well as biophysical geospatial data, ES vegetation components can be discriminated on a per-pixel basis.

Our initial cluster analysis showed that 89% of all observations were first grouped with the observations of their respective vegetation components before being combined with other clusters. Those observations not clustered with observations of the same ES

vegetation component were shown to have topographic or plant community properties not typical of the sampled ES vegetation communities for that type.

Decision tree-based algorithms (such as Random Forests) differ from cluster analysis in that they identify thresholds in each variable that best reduce the deviance in a response variable (Breiman et al. 1984). Cluster analysis does not produce a response variable and is thus incapable of doing this. Creating thresholds allows classifiers such as Random Forests to adjust the point at which classes are separated until the most accurate result is produced. An examination of the distribution of the observations in Figures 2-3 – 2-5 shows that drawing thresholds for different variables, instead of relying on distance from a centroid can produce cleaner results. This is particularly evident in Figure 2-5 when separating Wyoming big sagebrush from Utah juniper and mountain big sagebrush using slope and elevation. To a lesser extent, the advantage of thresholds can also be seen in Figures 2-2 and 2-3 when separating Utah juniper from mountain big sagebrush using NDVI.

Our out-of-bag accuracy for Random Forests of 97.2% demonstrated that we could accurately classify our observations. Some may suggest that this high level of accuracy is a product of over-fitting our classification to the data. However, out-of-bag accuracies are considered to be unbiased estimates of error (Breiman et al. 1984). Furthermore, over-fitting is not likely to occur in Random Forests (Prasad et al. 2006). Our Random Forests accuracy was also validated by applying the tree model to a geospatial data stack and randomly testing the output. This resulted in an overall accuracy estimate of 94%. The 3.2% reduction in accuracy when compared to the out-of-

bag accuracy may be attributed to the vegetation components present on the landscape, but that were not captured in our sampling.

We acknowledge that we have only shown the ability to accurately identify five vegetation communities out of several in Rich County. Twenty-nine percent of the ES vegetation components by land area were not considered. The inclusion of these other vegetation types would undoubtedly decrease our accuracy. The accuracy of our methodology is dependent on the spectral and ecological separability of vegetation types.

Even though we focused on only five of the ES vegetation types in Rich County, we have demonstrated that we can also identify vegetation components within an MU that did not belong to the majority vegetation component. An example of this is the mapping of mountain big sagebrush communities in MUs that were predominantly composed of aspen and Douglas-fir and did not identify mountain big sagebrush as a component. These results could help direct future soil mapping and also derive finer resolution MUs.

For areas of Wyoming big sagebrush and mountain big sagebrush, our error increased at intermediate elevations where these varieties intermix and create hybrids known as Bonneville big sagebrush (U.S. Department of Agriculture, NRCS 2012b). This intermixing presents obvious difficulties in identifying distinct ecological sites in transition areas. Currently, a precise identification of these types in intermediate elevations will require field-work.

Our variable importance plots produced by Random Forests showed that 5-year averages of NDVI were typically the most important remotely sensed variables, followed by the brightness components and then by the spatial variance in NDVI variables. We

also concluded that only one set of 5-year groups is needed to accurately map our

vegetation components. This conclusion does not go against those of other papers who

suggested that multi-temporal datasets were important for remote sensing classifications.

One 5-year average variable is still a product of multiple years' worth of remote sensing

imagery. We tested our conclusion that 5-year averages were necessary by creating a

Random Forests model using just one year for each remotely sensed variable. The out-

of-bag accuracy for this model was significantly lower that the accuracy from our model

with 5-year averages. Other multi-temporal datasets such as multi-seasonal remote

sensing data could be useful for ecological site classification and have been proven to be

effective in land cover classification (Andres et al. 1994; Kasischke and French 1995).

It was somewhat surprising that elevation was ranked relatively low in variable

importance. We concluded that this was due to the fact that almost all of these ESs

overlap on an elevation gradient. Additionally, even when elevation is assigned random

values during variable importance calculations, NDVI acts as somewhat of a proxy for

elevation because of increased precipitation in higher elevations leading to higher NDVI

values.

IMPLICATIONS

Prediction of the spatial distribution of ESs on a pixel basis has been suggested as

the next step in remote sensing applications to rangeland conservation (Hernandez 2011).

We have described and implemented a methodological approach to identify ES

vegetation components within individual MUs. We stress that we have not developed a

remote sensing solution for identifying complete ecological sites. To accomplish this we

need to accurately identify soil characteristics in addition to a more detailed description of the vegetation component. Our method accurately identifies and discriminates between vegetation components that are unique to certain ESs. The product from our method identifies where vegetation components occur spatially within MUs that previously only contained vegetation data on a percent composition level. This information can be used by those responsible for delineating ESs on a landscape scale to identify areas that have a high probability of ownership to a certain ES. Field work, particularly soil identification, can then be done to validate ES locations.

We have found that there are a few variables that were used in our analysis that only marginally improved our predictive ability (e.g aspect) and there are a few variables not used in this study that should be considered. We suggest that multi-seasonal imagery could be used as an independent variable. Another variable that should be considered is the map unit name. This variable could help in limiting the area that a certain ES can be mapped. If the training data used to build the classifier for a particular ES do not fall within the boundaries of certain MUs, then it will be unlikely that the particular ES will be mapped in those MUs. This will not help to differentiate between vegetation types that occupy the same MUs; however, it will help discriminate between vegetation components that may have similar remote sensing index values but do not occur on the same MUs. In our study, Utah juniper was significantly over estimated across the landscape. This problem could likely be solved by using the map unit name as a categorical variable in our Random Forests classifier since Utah juniper only occurs on specific map units. There are also several topographic variables derived from DEMs such as topographic wetness index, curvature, hillshade, and others. Certain

combinations of Landsat ETM bands have also been used to estimate soil composition

(Nield et al. 2007). Exploratory analyses, including scatterplots and cluster analysis,

should be conducted to determine what variables will be essential for accurate

classification of ES vegetation components. However, it is possible for variables to show

little added separation during cluster analysis and still be useful in Random Forests

classification.

LITERATURE CITED

Andres, L., W. A. Sala, and D. Skole. 1994. Fourier analysis of multi-temporal AVHRR data applied to a land cover classification. *International Journal of Remote Sensing* 15:1115-1121.

Arain, M. A., T. A. Black, A. G. Barr, P. G. Jarvis, J. M. Massheder, D. L. Verseghy, and Z. Nesic. 2002. Effects of seasonal and interannual climate variability on net ecosystem productivity of boreal deciduous and conifer forests. *Canadian Journal of Forest Research* 32:878-891.

Automated Geographic Reference Center. 2011. Utah GIS portal. Available at http://gis.utah.gov/elevation. Accessed 10 October 2011.

Bestelmeyer, B. T., A. J. Tugel, G. L. Peacock, D. G. Robinett, P. L. Shaver, J. R. Brown, J. E. Herrick, H. Sanchez, and K. M. Havstad. 2009. State-and-transition models for heterogeneous landscapes: a strategy for development and application. *Rangeland Ecology & Management* 52:1-15.

Brandon, T. B., J. R. Brown, K. M. Havstad, R. Alexander, G. Chavez, and J. E. Herrick. 2003. Development and use of state and transition models for rangelands. *Journal of Range Management* 56:114-126.

Breiman, L. 2001. Random forests. *Machine Learning* 45:5-32.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. Classification and Regression Trees. Belmont, CA: Wadsworth International Group. 358 p.

Briske, D. D., S.D. Fuhlendorf, and F. E. Smeins. 2005. State-and-transition models, thresholds, and rangleland health: a synthesis of ecological concepts and perspectives. *Rangeland Ecology & Management* 58:1-10.

Chander, G., B. L. Markham, and D. L. Helder. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment* 113:893-903.

Chavez, P. S. 1996. Image-based atmospheric corrections revisited and improved. *Photogrammetric Engineering and Remote Sensing* 62:1025-1036.

Crist, E. P., and R. C. Cicone. 1984. A physically-based transformation of Thematic Mapper data – the TM Tasseled Cap. *IEEE Transactions on Geosciences and Remote Sensing* 22:256-263.

Crookston, N. L., and A. O. Finley. 2008. yaImpute: an R package for KNN imputation. *Journal of Statistical Software* 23:1-16.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. *Ecology* 88:2783-2792.

Dymond, C. C., D. J. Mladenoff, and V. C. Radeloff. 2002. Phenological differences in tasseled cap indices improve deciduous forest classification. *Remotes Sensing of Environment* 80:460-472.

Everitt, B. S., and T. Hothorn. 2010. A handbook of statistical analyses using R. Boca Raton, FL: Taylor & Francis Group. 207 p.

Gamon, J. A., C. B. Field, M. L. Goulden, K. L. Griffin, and A. E. Hartley. 1995. Relationships between NDVI, canopy structure, and photosynthesis in three Californian vegetation types. *Ecological Applications* 5:28-41.

Goulden, M. L., J. W. Munger, S. Fan, B. C. Daube, and S. C. Wofsy. 1996. Exchange of carbon dioxide by a deciduous forest: response to interannual climate variability. *Science* 271:1576-1578.

Hernandez, A. 2011. Spatiotemporal modeling of threats to big sagebrush ecological sites in northern Utah [dissertation]. Logan, UT: Utah State University. 163 p.

Hunt, E. R., J. H. Everitt, J. C. Ritchie, M. S. Moran, D. T. Booth, G. L. Anderson, P. E. Clark, and M. S. Seyfried. 2003. Applications and research using remote sensing for rangeland management. *Photogrammetric Engineering and Remote Sensing* 69:675-693.

Jensen, J. R. 2000. Remote sensing of the environment: an earth resources perspective. Upper Saddle River, NJ: Prentice-Hall. 592 p.

Jin, X. M., Y. K. Zhang, M. E. Schaepman, J. G. Clevers, and Z. Su. 2008. Impact of elevation and aspect on the spatial distribution of vegetation in the Qilian mountain area with remote sensing data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 37:1385-1390.

Kasischke, E. S., and N. H. French. 1995. Locating and estimating the areal extent of wildfires in alaskan boreal forests using multiple-season AVHRR NDVI composite data. *Remote Sensing of Environment* 51:263-275.

Kotsiantis, S. B., and P. E. Pintelas. 2004. Recent advances in clustering: a brief survey. *WSEAS Transactions on Information Science & Applications* 1:73-81.

Maynard, C. L., R. L. Lawrence, G. D. Nielsen, and G. Decker. 2007. Ecological site descriptions and remotely sensed imagery as a tool for rangeland evaluation. *Canadian Journal of Remote Sensing* 33:109-115.

Miller, R. F., and J. A. Rose. 1999. Fire history and western juniper encroachment in sagebrush steppe. *Journal of Range Management* 52:550-559.

Nield, J. L., J. L. Boettinger, and R. D. Ramsey. 2007. Digitally mapping gypsic and nitric soil areas using Landsat ETM data. *Soil Science Society of America Journal* 71:245-252.

Pickup, G., and B. D. Foran. 1987. The use of spectral and spatial variability to monitor cover change on inert landscapes. *Remote Sensing of Environment* 23:351-363.

Prasad, A. M., L. R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems* 9:181-199.

Prior-Magee, J. S. 2007. Chapter 1. Moscow, ID: U.S. Geological Survey, Final Report for Southwest Regional Gap Analysis. 13 p.

Rouse, J. W., R. H. Haas, J. A. Schell, and D. W. Deering. 1974. Monitoring vegetation systems in the great plains with ERTS. Proceedings of the 3rd Earth Resources Technology Satellite-1 Symposium, 10-14 December 1973; Greenbelt, MD: National Aeronautics and Space Administration. p. 3010-2017.

Sakrejda-Leavitt, K. 2009. Fast function for centering and standardizing variables. Available at: https://stat.ethz.ch/pipermail/r-help/2009-June/200055.html. Accessed 6 January 2012.

Stringham, T. K., W. C. Krueger, and P. L. Shaver. 2003. State and transition modeling: an ecological process approach. *Journal of Range Management* 56:106-113.

U.S. Department of Agriculture, NRCS. 2005. MRLA definitions. Available at: http://soils.usda.gov/survey/geography/mlra/mlra_definitions.html. Accessed 10 January 2012.

U.S. Department of Agriculture, NRCS. 2011. National soil survey handbook. Available at: http://soils.usda.gov/technical/handbook/ contents/part622.html#07. Accessed 5 January 2012.

U.S. Department of Agriculture, NRCS. 2012a. Ecological Sites. Available at: http://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/landuse/rangepasture/?cid=STELPRDB1043235. Accessed 17 October 2011.

U.S. Department of Agriculture, NRCS. 2012b. Plant guide: big sagebrush. Available at: http://plants.usda.gov/plantguide/pdf/pg_artrw8.pdf. Accessed 25 April 2012.

U.S. Department of Agriculture, NRCS-SSURGO. 2012. Soil data mart. Available at: http://soildatamart.nrcs.usda.gov/. Accessed 10 October 2011.

U.S. Geological Survey. 2011. *NDVI Graph.* Available at: http://glovis.usgs.gov/. Accessed 10 January 2012.

Utah Office of Tourism. 2009. Rich County tourism profile. Available at: http://www.travel.utah.gov/research_and_planning/documents/RichCounty_005.pdf. Accessed 5 January 2012.

Washington-Allen, R., N. E. West, R. D. Ramsey, and R. A. Efroymson. 2006. A protocol for retrospective remote sensing-based ecological monitoring of rangelands. *Rangeland Ecology & Management* 59:19-29.

Westerling, A. L., and T. W. Swetnam. 2003. Interannual to decadal drought and wildfire in the western United States. *Eos, Transactions, American Geophysical Union* 84:545-560.

Westoby, M., B. Walker, and I. Noy-Meir. 1989. Opportunistic management for rangelands not at equilibrium. *Journal of Range Management* 42:266-274.

Table 2-1. Site-by-attribute table for each training site. There are 20 training sites for each ecological site vegetation component. Table continues on next two pages.

name	NDVI mean	NDVI SD	brightness	elevation	slope	aspect	area
				---m---	-----degrees-----		---acres---
aspen1	180.81	4.82	130.48	2565	9.85	276.06	58
aspen2	183.07	4.99	145.52	2404	6.84	107.39	56
aspen3	181.94	7.59	146.05	2383	9.32	105.35	43
aspen4	182.26	3.23	131.38	2417	8.8	89.51	24
aspen5	181.65	3.82	154.58	2224	12.3	87.6	33
aspen6	185.28	4.45	141.71	2151	13.01	312.51	50
aspen7	186.04	3.92	149.12	2085	13.93	114.6	14
aspen8	185.21	3.53	133.05	2174	7.38	97.07	16
aspen9	179.6	6.87	144.07	2359	12.48	42.23	87
aspen10	176.97	8.04	148.83	2563	9.43	135.91	48
aspen11	179.74	4.17	132.36	2427	8.73	93.21	91
aspen12	174.26	9.71	129.08	2567	6.64	137.47	28
aspen13	178.43	5.6	118.45	2450	10.92	48.38	21
aspen14	183.92	5.66	135.55	2515	10.78	114.67	18
aspen15	183.11	3.25	131.39	2570	8.8	120.86	13
aspen16	182.25	6.13	134.81	2483	19.42	224.6	8
aspen17	186.88	3.69	144.75	2321	15.26	111.59	13
aspen18	182.92	5.36	144.63	2431	6.11	86.37	18
aspen19	183.44	5.19	138.26	2643	9.7	139.43	12
aspen20	183.47	4.62	128.31	2566	5.71	167.71	14
df1	174.62	5.46	85.87	2364	22.99	268.83	25
df2	174.95	6.67	82.33	2365	23.56	252.16	18
df3	180.04	8.68	95.81	2121	25.14	28.88	15
df4	181.71	4.62	93.43	2365	23.56	252.16	31
df5	174.95	6.67	82.33	2198	18.9	162.02	23
df6	172.09	4.84	99.73	2496	13.96	81.52	23
df7	182.07	3.81	85.19	2152	27.38	87.64	19
df8	182.56	3.82	85.54	2171	19.8	285.01	25
df9	177.69	3.19	76.75	2403	14.49	84.14	61
df10	177.09	5.04	81.06	2478	9.35	259.95	24
df11	177.08	4.53	91.57	2402	9.82	221.03	13
df12	179.36	4.23	82.74	2372	9.65	264.54	44
df13	176.86	5.13	74.07	2411	28.71	294.94	22
df14	175.29	7.84	83.39	2535	21.32	256.53	12
df15	176.27	4.39	88.35	2312	27.82	298.69	27
df16	172.71	6.46	82.27	2266	28.5	338.85	4
df17	177.66	5.71	88.63	2455	17.82	271.13	16
df18	178.38	4.68	86.97	2604	13.71	254.92	53
df19	171.41	7.01	90.01	2298	26.24	291.75	18
df20	178.55	3.4	87.36	2282	8.09	64.69	18

Table 2-1. Continued.

name	NDVI mean	NDVI SD	brightness	elevation	slope	aspect	area
				---m---	-----degrees-----		---acres---
juniper1	129.44	2.3	180.58	2136	15.6	92.12	48
juniper2	129.47	3.29	175.92	2122	11.65	93.89	19
juniper3	131.49	2.13	170.74	2118	11.68	133.3	24
juniper4	125.91	3.78	191.16	2091	22.41	161.51	13
juniper5	132.8	2.45	169.93	2149	12.46	124.96	9
juniper6	129.98	4.51	175.67	2130	18.64	171.16	11
juniper7	130.46	3.42	171.1	2220	22.8	180.05	32
juniper8	131.79	3.05	173.88	2188	13.71	151.03	39
juniper9	130.64	2.35	173.94	1970	16.71	138.11	28
juniper10	134.78	4.36	160.52	1968	16.79	165.69	85
juniper11	126.79	4.96	197.95	1998	16.65	162.37	45
juniper12	130.02	3.7	167.98	2103	11.8	126.82	28
juniper13	131.09	2.74	168.51	2115	7.93	95.84	124
juniper14	130.48	2.44	172.65	2117	12.28	100.61	28
juniper15	133.58	3.11	130.64	1990	21.99	268.54	21
juniper16	126.44	3.5	189.65	1982	16.17	164.27	4
juniper17	129.64	3.11	157.36	1965	14.4	215.3	16
juniper18	127.01	3.72	196.82	2038	17.89	168.63	27
juniper19	127.74	3.42	177.6	2155	10.15	189.96	12
juniper20	130.54	5.77	141.47	2140	25.97	269.37	15
mbs1	146.3	5.61	142.17	2248	6.08	82.87	89
mbs2	148.53	6.24	145.71	2220	1.67	222.07	26
mbs3	151.88	3.55	139.16	2195	6.62	180.92	23
mbs4	142.02	4.39	148.32	2085	9.18	121.2	25
mbs5	155.52	7.03	135.58	2138	17.91	230.99	46
mbs6	141.93	3.45	154.1	2186	4.79	143.38	22
mbs7	159.67	7.34	145.65	2368	9.83	159.92	16
mbs8	154.25	5.11	146.23	2322	11.04	130.38	47
mbs9	151.17	8.09	137.99	2314	6.74	224.13	56
mbs10	150.09	6.72	143.46	2363	8.34	102.96	23

Table 2-1. Continued.

name	NDVI mean	NDVI SD	brightness	elevation	slope	aspect	area
				---m---	-----degrees-----		---acres---
mbs11	149.19	8.9	161.07	2382	13.71	155.94	95
mbs12	148.98	7.73	143.59	2402	6.13	244.23	60
mbs13	150.4	7.16	146.06	2177	9.76	97.21	59
mbs14	150.34	6.64	150.74	2177	10.83	90.08	28
mbs15	154.46	6.6	136.2	2149	9.68	65.04	13
mbs16	138.22	7.4	163.37	2194	9.43	163.16	31
mbs17	161.91	7.77	140.39	2323	8.59	145.72	12
mbs18	145.79	8.05	143.48	2096	12.74	71.09	23
mbs19	146.82	5.18	141.17	2119	13.44	178.5	35
mbs20	162.6	7.7	149.62	2432	13.6	170.91	19
wbs1	124.12	1.24	183.76	1936	0.36	127.76	119
wbs2	125.4	1.59	181.34	1956	1.27	117.84	68
wbs3	124.6	1.45	198.44	1946	1.57	205.6	51
wbs4	129.57	2.03	161.31	2054	0.96	136.15	55
wbs5	129.45	1.66	157.93	2045	1.62	102.82	52
wbs6	127.2	1.74	166.36	2038	2.73	233.43	25
wbs7	129.28	1.74	155.4	2034	6.16	84.56	20
wbs8	125.77	1.66	173.57	1924	1.07	50.46	12
wbs9	125.86	1.6	167.51	2048	1.6	101.69	22
wbs10	125.08	1.78	189.98	1975	1.37	313.34	23
wbs11	132.33	1.88	152.67	2028	6.92	75.17	16
wbs12	126.94	1.7	170.09	2018	1.96	116.38	124
wbs13	123.79	1.47	192.89	1959	1.41	112.06	23
wbs14	129.13	2.77	159.15	1984	2.79	117.3	42
wbs15	126.76	1.66	163.34	2041	2.56	160.63	30
wbs16	128.37	2.33	174.61	2066	5.38	181.92	8
wbs17	127.07	1.61	180.02	1943	2.63	65.85	9
wbs18	126.16	1.42	172.56	2025	2.43	90.24	45
wbs19	128.72	1.94	161.01	2017	4.79	314.33	29
wbs20	131.87	2.76	162.75	2047	8.35	35.03	18

Table 2-2. Confusion matrix for random forests classification of Rich County, UT. The "Other" row displays the number of accuracy assessment sites that were classified as each class but in reality were part of an ecological site vegetation component not accounted for in our classification. MBS, mountain big sagebrush; WBS, Wyoming big sagebrush.
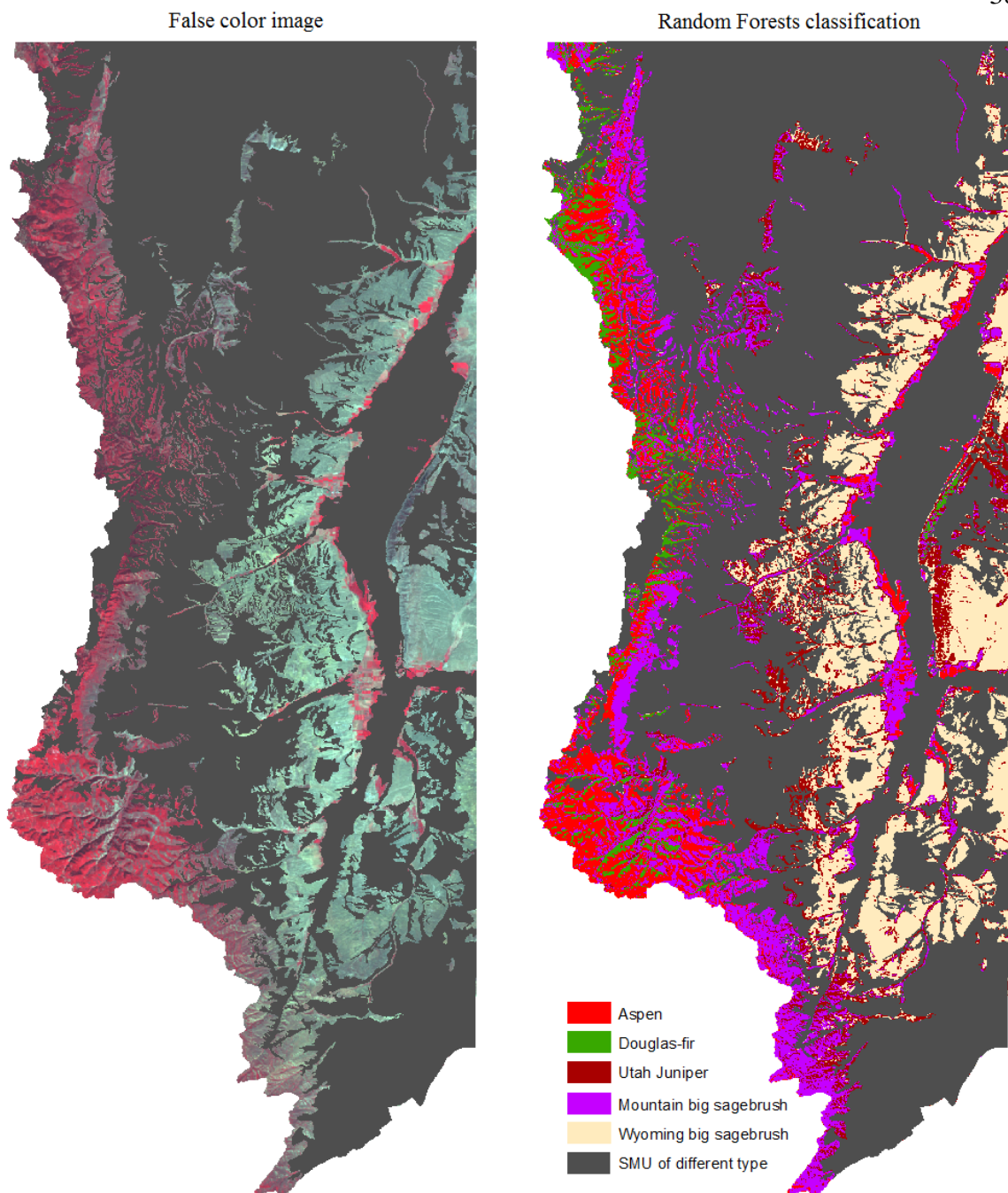
		Predicted				
		Aspen	Douglas-fir	Utah Juniper	MBS	WBS
Actual	Aspen	**46**	1	0	0	0
	Douglas-fir	1	**48**	1	0	0
	Utah Juniper	0	1	**22**	2	1
	MBS	2	0	1	**42**	0
	WBS	0	0	3	0	**48**
	Other	1	0	23	6	1

Table 2-3. Remote sensing and topographic variables used in the cluster analysis and Random Forests model.

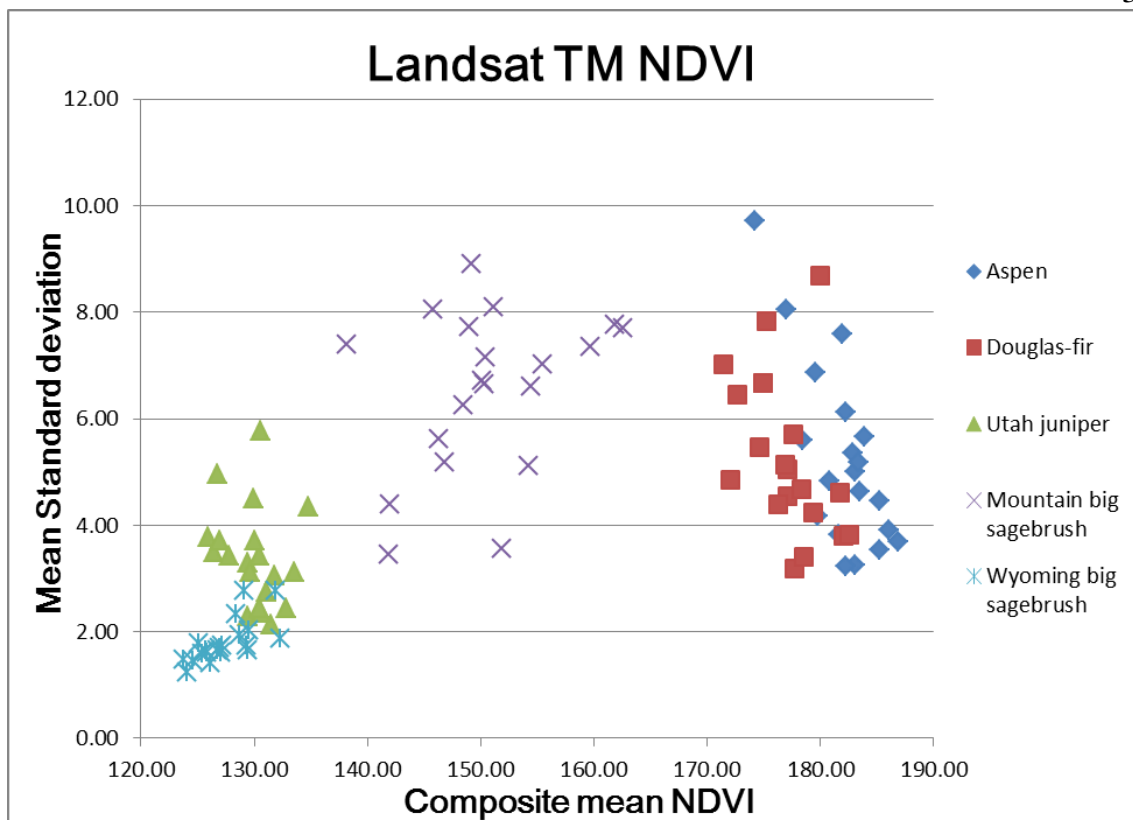| Variables			
5 year averages NDVI	5 year averages standard deviation of NDVI	5 year averages Brightness component	Topographic
NDVI84_88	SD84_88	BRIGHT84_88	slope30m
NDVI89_93	SD89_93	BRIGHT89_93	elevation30m
NDVI94_98	SD94_98	BRIGHT94_98	aspect30m
NDVI99_03	SD99_03	BRIGHT99_03	
NDVI04_08	SD04_08	BRIGHT04_08	
NDVI09_11	SD09_11	BRIGHT09_11	
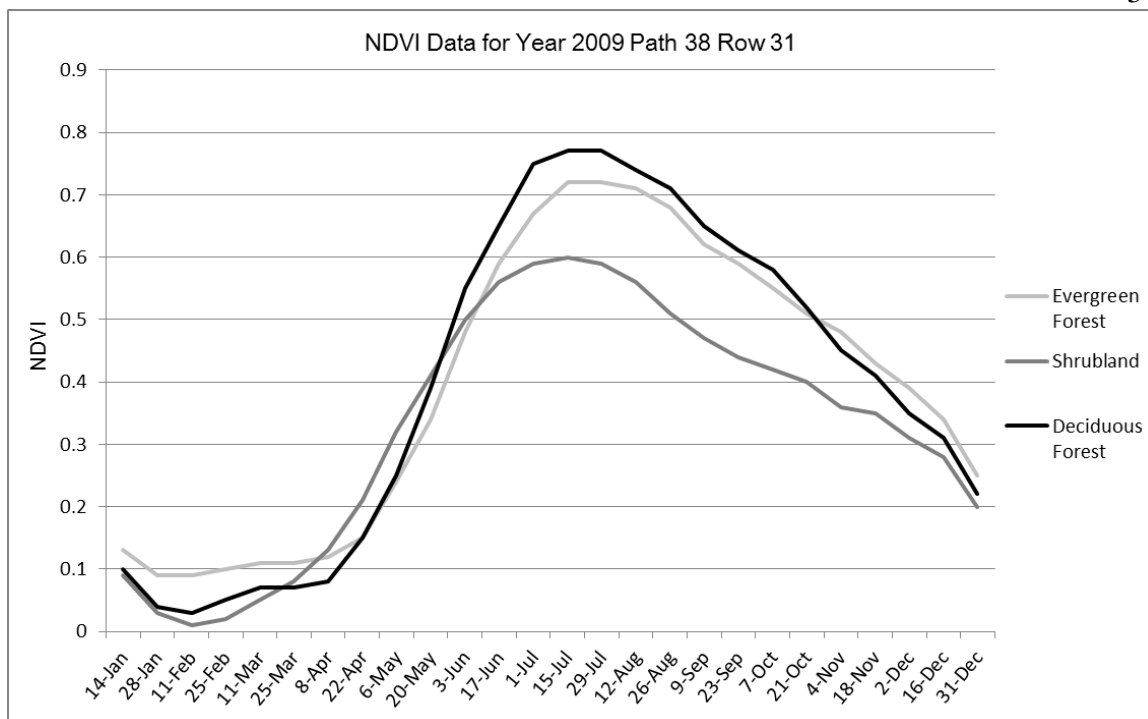
Figure 2-1. Line graph of annual fluctuations in NDVI for evergreen forests, shrublands, and deciduous forests. The largest differences in NDVI can be seen in mid-summer. Similar graphs can be obtained from the USGS GLOVIS Visualization Viewer at http://glovis.usgs.gov/. NDVI, Normalized Difference Vegetation Index.

False color image

Random Forests classification

Legend:
- Aspen
- Douglas-fir
- Utah Juniper
- Mountain big sagebrush
- Wyoming big sagebrush
- SMU of different type

Figure 2-2. Random Forests classification of Rich County, UT. Black areas represent map units (MU) whose majority component was not Aspen, Douglas-fir, Utah Juniper, Mountain big sagebrush, or Wyoming big sagebrush.
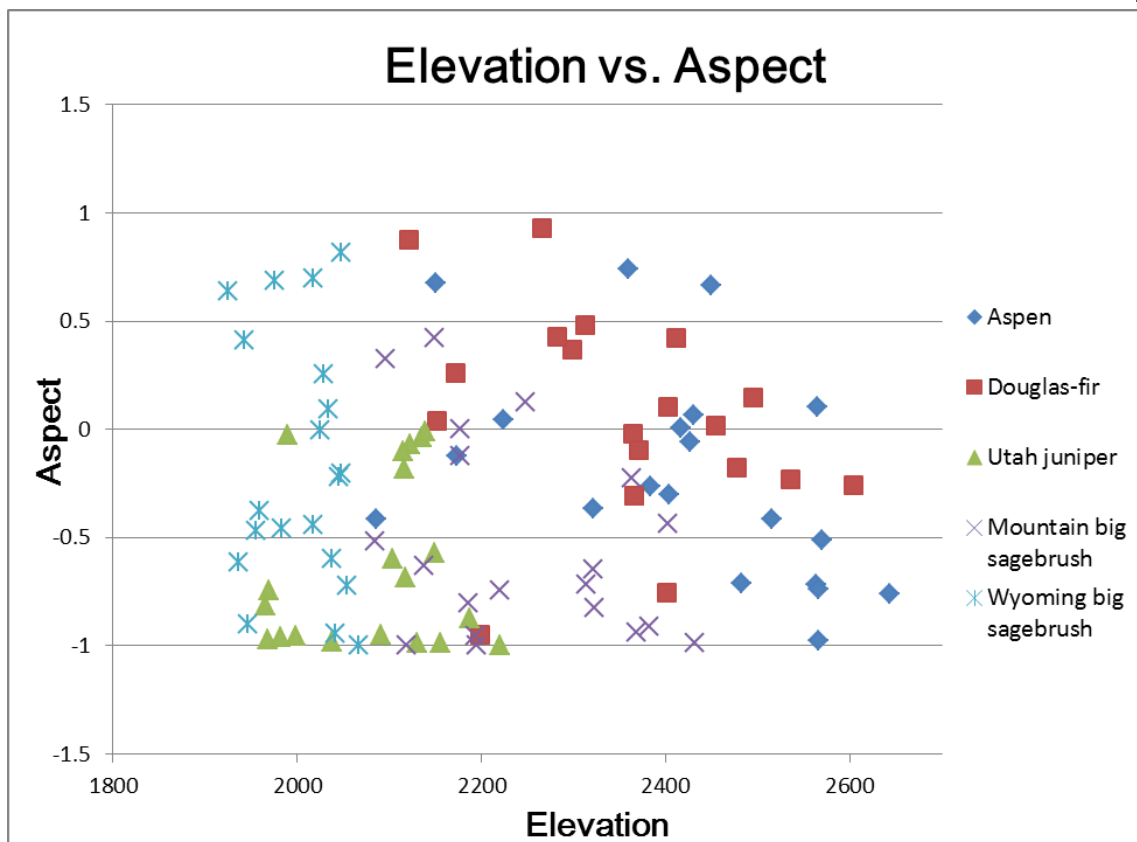
Figure 2-3. Scatter-plot showing the distribution of each ecological site vegetation component in our study with average NDVI value on the x-axis and the average standard deviation in NDVI on the y-axis. Both of these variables provide some separation between vegetation classes. TM, Thematic Mapper; NDVI, Normalized Difference Vegetation Index.
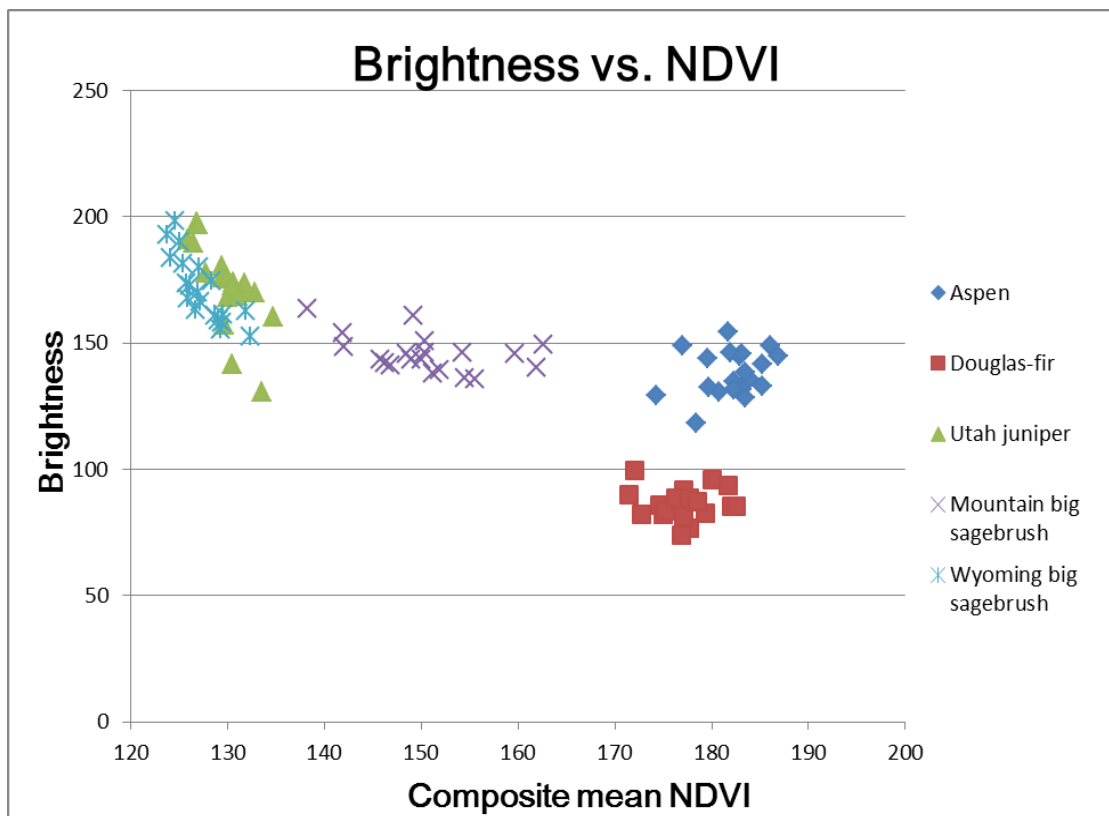
Figure 2-4. Scatter-plot showing the distribution of each ecological site vegetation component in our study with average NDVI value on the x-axis and the average brightness component (obtained from the tasseled cap transformation) value on the y-axis. This graph shows that brightness provides added separation between ecological site vegetation classes. NDVI, Normalized Difference Vegetation Index.
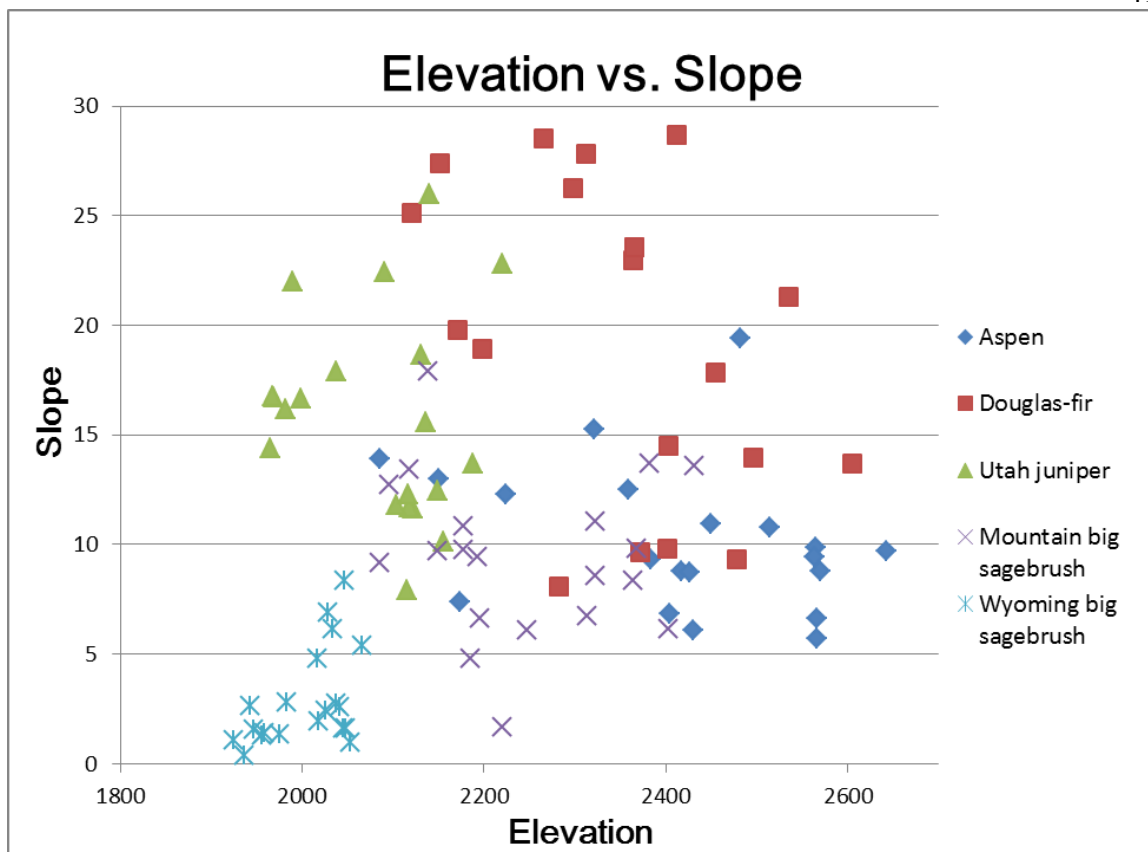
Figure 2-5. Scatter-plot showing the distribution of each ecological site vegetation component in our study with elevation on the x-axis and slope on the y-axis. Most vegetation classes overlap one another. However, slope does help with separating Wyoming big sagebrush from Utah juniper.
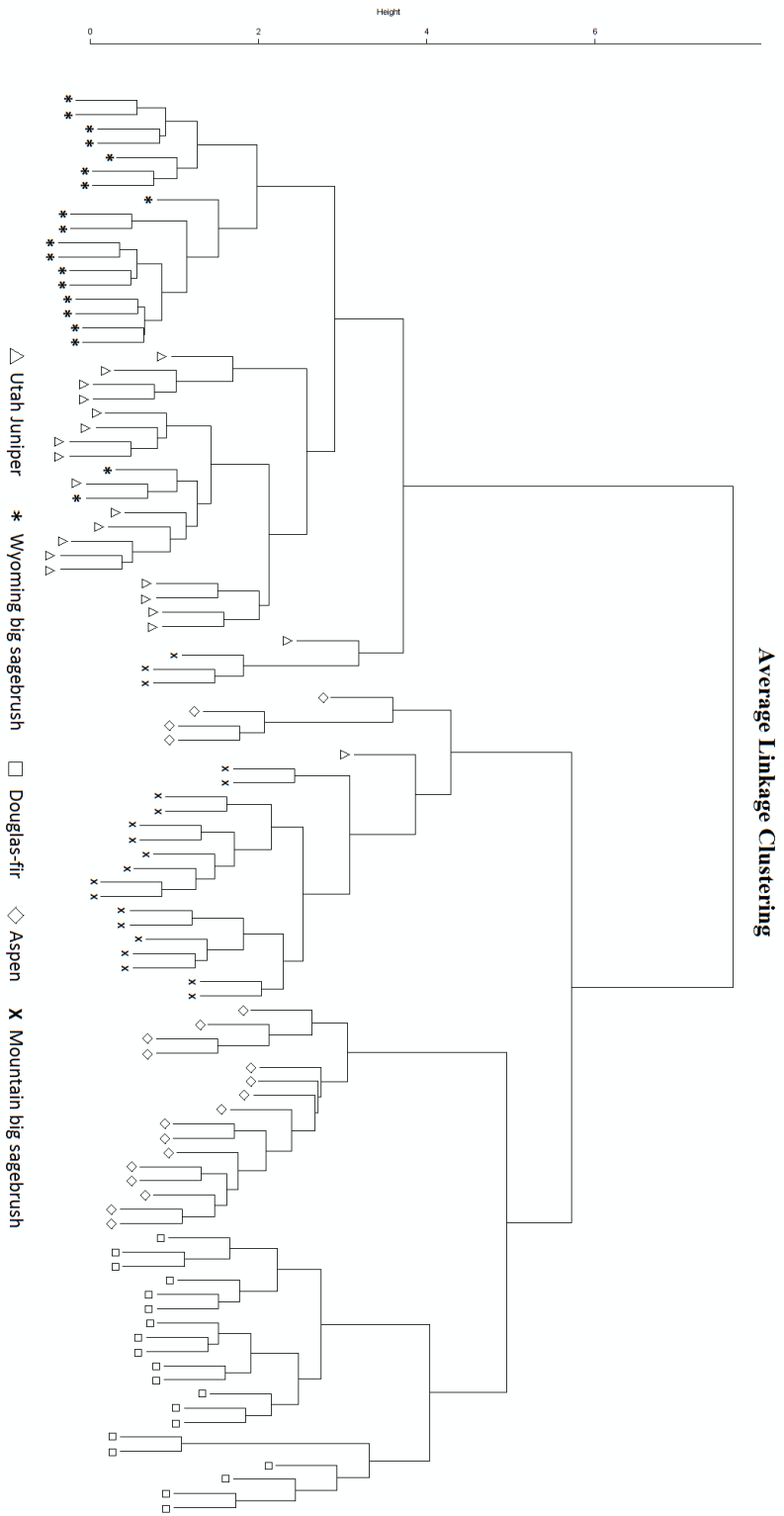
**Figure 2-6.** Scatter-plot showing the distribution of each ecological site vegetation component in our study with elevation on the x-axis and aspect on the y-axis. Aspect does not appear to separate any vegetation classes.

Figure 2-7. Dendogram of Average Linkage Clustering. This figure shows the grouping of ecological site descriptions (ESD) into their own clusters. There are few ESDs that grouped with other ESDs.

Variable Importance for Random Forests
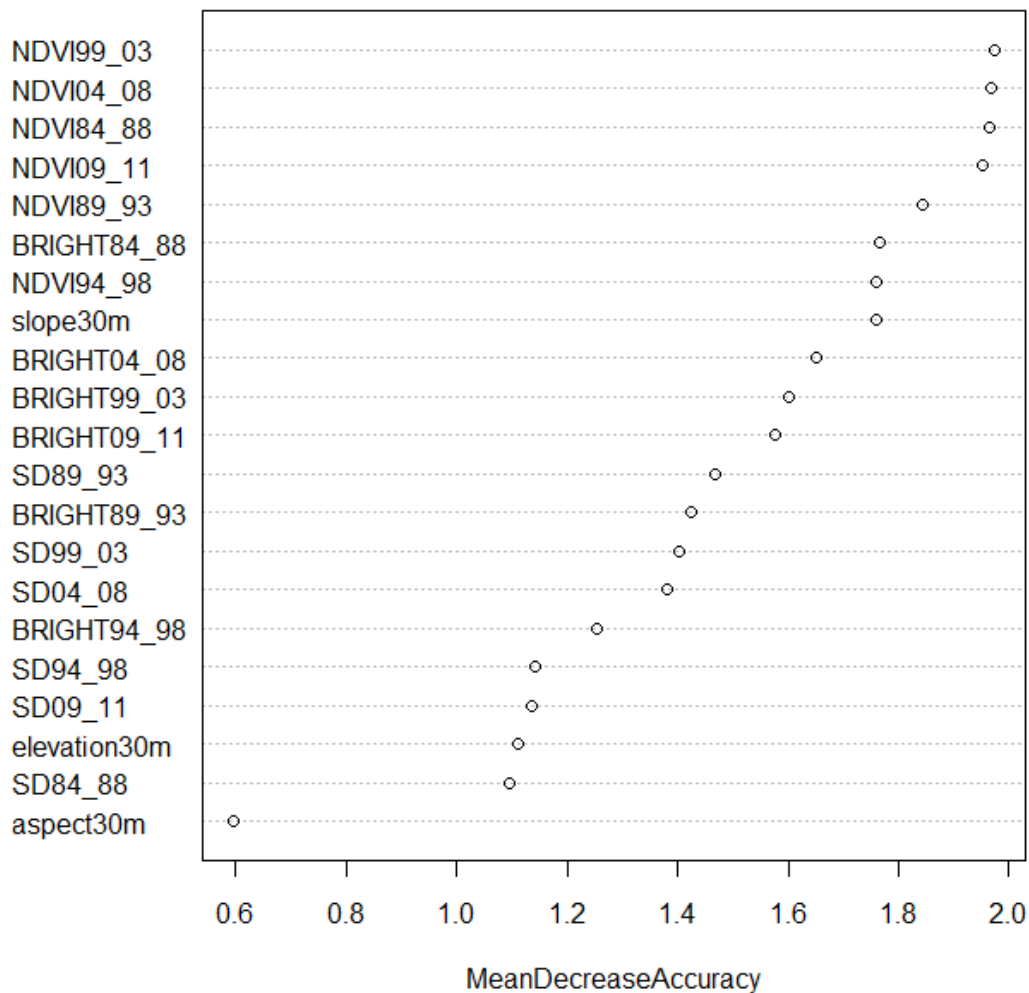
Figure 2-8. Variable importance plot produced by random forests model. Variables with higher mean decrease in accuracy values provided more separation between classes.

CHAPTER 3

ECOLOGICAL SITE AND STATE CLASSIFICATION OF WYOMING BIG

SAGEBRUSH IN RICH COUNTY,

UTAH

INTRODUCTION

Within the Intermountain West, vast expanses of big sagebrush shrubland and
steppe are considered emblems of the western range. Currently, there are approximately
60 million hectares of big sagebrush within the 11 western states (Beetle 1960), four
million of which are in the state of Utah (Lowry et al. 2007). However, the historic
distribution of sagebrush has been impacted by conversion to other types of land cover
(e.g., encroachment by Juniper and invasion by annual weeds) (Miller and Rose 1999),
and anthropogenic land use (agriculture and urbanization). In Utah alone, Big Sagebrush
communities have been reduced to approximately 55% of their historic extent (Landfire –
EVT 2008; Landfire – BPS 2008). Changes to alternative land cover types have been
facilitated by an alteration of disturbance regimes, namely fire return intervals, grazing,
mechanical treatments, and urbanization (Knick et al. 2003). A primary and current
example of the cumulative impact of big sagebrush loss is the eminent listing of the Sage
Grouse as a threatened and endangered species (Connelly et al. 2004). This potential
listing will force land management agencies to impose strict guidelines for future
development of sagebrush-dominated landscapes. These growing pressures have led to a
need to accurately estimate the current spatial distribution of sagebrush shrubland and
steppe and their current ecological condition.

Big Sagebrush communities, as well as other semiarid vegetation communities, exist across their geographic distribution in various ecological states. These states can be viewed as nuances in community structure due to local environmental factors, or they can represent alterations forced by management actions or changes in climate. Information about the different ecological states that sagebrush communities can occupy, as well as the forces that promote the transitions between states, can be enumerated in state-and-transition models (STMs) (Westoby et al. 1989). These transitions can take place due to soil erosion, fire regimes, weather variability, and management (Briske et al. 2005). Westoby et al. suggested that the purposes of STMs are to 1. Define the states possible within a system 2. Catalogue management action and other forces that drive transitions from one state to another and 3. List the actions that could produce favorable transitions as well as the hazards of inaction that could result in unfavorable transitions. A state is defined as a recognizable, resistant, and resilient complex of soil base and vegetation structure (Stringham et al. 2003). The original STM framework does not indicate a need to identify a reference state. However, STMs adopted by the USDA Natural Resources Conservation Service (NRCS) have been joined with the traditional range model so that STMs developed by the NRCS include a reference state that characterizes the historic plant community (Briske et al. 2005).

The NRCS has been systematically classifying rangelands into ecological sites (ES) that link soil characteristics to the defined historic plant community occupying that soil. Ecological site descriptions (ESDs) describe areas of specific biophysical properties and associated plant communities that may be found at a given site. These sites differ from other sites in their ability to produce a distinct kind and amount of vegetation.

Areas of the same ES, but separated by geography, are also unique in that they are assumed to "respond similarly to management actions and natural disturbances" (U.S. Department of Agriculture 2011). Ecological sites are primarily determined on the basis of soil characteristics and the resulting differences in plant species composition and production that occur on those soils. Because ESDs are based on the plant community that existed at the time of European settlement (U.S. Department of Agriculture, NRCS 2011), ESDs represent reference states in STMs.

Currently, ESs are identified on a landscape as components within map units (MU). An MU is a spatially defined area that enumerates the soil characteristics at that location. A given MU can contain one or more different soil types that are termed components. Components are contiguous groupings of different soils whose extents are equal to or smaller than the MU. Map unit polygons therefore have a one-to-many relationship with ESs (Arid Land Research Programs 2010). The spatial and tabular data for MUs are stored in individual soil surveys and can be obtained from the NRCS SSURGO database (U.S. Department of Agriculture, NRCS-SSURGO 2012). Up to four different ecological site components (one per soil type) are combined into one MU and the SSURGO tabular database details the percentage of area each component occupies within a given MU; however, the database does not define the spatial location of a particular ES component within the MU. It will be the goal of this research to create and use a remote sensing based similarity index to map the spatial distribution of an ES component and its states across a landscape.

Similarity indices are not new to the ES process. The NRCS adopted a similarity index in an effort to standardize definitions and quantify ecological states. This effort

followed an initial lack of universally accepted definitions of STMs that subsequently led

to confusion and criticism (Iglesias and Kothmann 1997). The NRCS's similarity index

provides a way to compare vegetation states to one another. This is done by comparing

the present state of vegetation on a site to the kinds, proportions, and amounts of

vegetation that existed in the reference/historic climax plant community state (U.S.

Department of Agriculture, NRCS 2006). The similarity index indicates the percent of

the plant community present during the reference state that is still present today. Before

the similarity index for a site can be calculated, a field inventory is carried out to estimate

the annual productivity for each species present at the site. Like all field work, this

process takes a great deal of time and is therefore costly.

Hernandez (2011) postulated a method for creating a similarity index, referred to

as "ecodistance," using remotely sensed imagery. This was done by comparing the mean

and standard deviations in the soil adjusted vegetation index (SAVI) for a given location

to identical metrics of undesirable alternative states (e.g., cheatgrass and juniper

encroachment). These alternative undesirable states served as benchmarks from which to

compare all other sites with similar ESs (West 1991). Similarity was quantified by using

a Euclidean distance metric, measured in standardized units of mean and standard

deviations in SAVI, between a given geographic location and the alternative state

benchmarks. Sites with low distance were considered very similar to the conditions of

the benchmark.

Other studies have used remotely sensed data to help classify and discriminate

between different ESs and the different ecological states possible within an ES. Maynard

et al. (2007) found that the tasseled cap components were correlated with variations in

ground measurements of biomass and exposed soil when sites were stratified by

ecological site. Gamon et al. (1995) discussed the usefulness of the NDVI as an indicator

of photosynthetic activity as well as canopy structure, and plant nitrogen content. Jensen

(2000) showed that NDVI was sensitive to canopy variations including soil visible

through canopy openings" (p. 386). While the sensitivity to soil background has

typically been seen as a disadvantage of NDVI for vegetation assessment (Huete et al.

2002), it could prove useful for studying states within an ES since areas of the same

ecological state will have a similar amount and type of bare soil. Since NDVI is sensitive

to these differences, we feel that it would be a suitable index for distinguishing between

states and approximating distance to states. The NDVI values within the polygon of a

soil mapping unit and the variation in the NDVI has also been used to distinguish

between states (Hernandez 2011).

Because ESs are not explicitly mapped, it is not surprising that ecological states

within a given ESs STM have also not been mapped. We were only able to find one

study that attempted to map ecological states. Steele et al. (2012) used a manual mapping

approach that combined aerial photo interpretation supplemented with field data to map

ecological states in New Mexico. We wish to build upon Hernandez's work by first

classifying each pixel in the ES R034AY2ggUT (Semi-desert Loam: Wyoming big

sagebrush/Caespitose bluebunch wheatgrass) in Rich County, UT, to one of the states

identified in the STM. We will then calculate a similarity index represented by the

Mahalanobis distance for each image pixel to the most probable state identified by the

corresponding STM. We have chosen to work with the Rich County, Utah, soil survey

area (NRCS soil survey UT604) where there are 679 individual MUs whose largest

component (40 - 95% of the area) is R034AY2ggUT. By applying the similarity index developed here to every remotely sensed pixel within a given MU, pixels that have large distances to any one of our predefined benchmark states should either be inclusions (not R034AYggUT) or states not previously considered for R034AYggUT. Doing this will create a cost efficient and standardized way to map the spatial extent of ESs and their respective ecological states. We expect this work to be valuable to those responsible for identifying and defining ESs as well as those responsible for creating and updating MUs and STMs.

METHODS

Study Area

Our research was conducted in Rich County, Utah, located in the northeastern corner of the state (long 111°30'38.5'' – long 111°2'42.2'' West and lat 42°0'0'' – lat 42°08'24.3'' North). Rich County is made up of two Major Land Resource Areas (MLRA) including the Wasatch and Uinta Mountains (47) and Cool Central Desertic Basins and Plateaus (34A). MLRAs are generalized areas similar to ecoregions that are classified by physiography, geology, climate, water, soils, biological resources, and land use (U.S. Department of Agriculture, NRCS 2005). The western portion of Rich County is characterized by high elevations with vegetation consisting of aspen forests, subalpine conifer forests, and scattered mountain sagebrush steppe. Moving east, the elevation decreases, and the mountain sagebrush steppe becomes dominant. Both the mountain and foothills sections of the county are in MLRA 47. The ES that we were interested (R034AY2ggUT) is in MLRA 34A which is primarily located in central and eastern Rich

County. This MRLA is made up of relatively lower elevations with vegetation consisting

of big sagebrush steppe and shrubland, subalpine grasslands, and agriculture.

The average elevation in Rich County for areas dominated by R034AY2ggUT is

1990 m. The highest elevation is 2300 m and the lowest point is about 1891 m. The soil

temperature regime is frigid and the soil moisture regime is xeric for most of the county.

The parent material is primarily derived from sandstone and limestone. The source of the

parent material is alluvium. Plants in R034AY2ggUT occur on xeric soils that are

shallower than those occupied by other sagebrush species such as basin and mountain big

sagebrush. R034AY2ggUT soils typically contain a large amount of clay or sometimes

silt. Wyoming big sagebrush does not do well on coarse textured soils (Frisina and

Wambolt 2004). For a detailed description of the soils present in the study area, read the

Soil Survey of Rich County Utah (U.S. Department of Agriculture, NRCS-SSURGO

1982).

A slight majority of the land occupied by R034AY2ggUT is in private ownership

at 52.8%. The federal government is the next largest landowner with 40.2% which is

managed by the Bureau of Land Management. The state of Utah owns only 7% of the

land area which is mostly composed of State Trust Lands (Utah Office of Tourism 2009).

Much of the private land (22%) is owned by Deseret Land and Livestock. Disturbances

that have affected the area include agriculture, grazing, and burning.

Ecological Site

We chose the ES R034AY2ggUT since it is a preferred plant community of

wintering sage-grouse (Welch et al. 1991) and its large distribution across the

Intermountain West. In Rich County, three other ESs are identified as having a dominant

component of Wyoming big sagebrush. R034AY2ggUT was chosen because it is the

most commonly occurring of the four ESs. The reference vegetation component

(historical plant community) for the ES R034AY2ggUT is Wyoming big sagebrush

(*Artemisia tridentata* Nutt. ssp. *wyomingensis*) with varying amounts of bluebunch

wheatgrass (*Pseduoroegneria spicata* [Pursh] Á. Löve), yellow rabbitbrush

(*Chrysothamnus viscidiflorus* [Hook.] Nutt.), and other native perennial bunchgrasses

(Fig. 3-1). While a general estimation of the historic pre-Columbian plant community

can be made, a confident quantitative estimate is not possible for this ES due to a lack of

direct historical documentation preceding European settlement. The first reports of

dominant plant species were made in the late 19th century from a cadastral survey

conducted by the General Land Office (Galatowitsch 1990). Human management in this

area was introduced well before European settlement by Shoshone Indians who grazed

horses and set fires to alter the vegetation for their needs (Parson 1996).

Since then, several other and more frequent disturbances have occurred that have

caused transitions from the defined reference state to alternative states.   These changes

are modeled in Figure 3-1 (U.S. Department of Agriculture, NRCS 2012). This first

transition is from the reference state to an alternative state (State 2) that is very close to

the approximation of the reference state. State 2 is identical to the reference state with

the exception of a small component of introduced non-natives into the plant community.

The second alternative state (State 3) is a Wyoming big sagebrush super-dominance state

which is caused by heavy, year-round grazing by cattle, sheep, and horses. From this

state, three different transitions can occur that can move the ES into one of three

additional states. State 4 is an increased invasives state caused by prescribed grazing,

unusually wet climate, soil anoxia, insects, and/or wildfire. State 5 is a crested

wheatgrass state that can be transitioned to from either State 3 or 4 by brush management.

State 6 is a Wyoming big sagebrush and native grass state that can be transitioned from

either State 3 or 4 by means of prescribed grazing.

Datasets

Because we wanted to calculate the similarity of all areas within the

R034AY2ggUT ES to the state of most probable membership, we needed to have a

representative sample of each state defined in the STM. The reference state (State 1) is

not represented because it is assumed that this state no longer exists. Training sites were

acquired from fieldwork conducted by Peterson (2009) and the Utah Division of Wildlife

Resources (2006). Both datasets included the geographic location of the site along with

the percent cover of each species present. From this information, we created polygons

that represented the area sampled for each site. We also assigned a state number to each

site if it appeared to be in one of the states present in the STM. These assignments were

based on the percent cover for each species at each site. The minimum number of sites

that were assigned to a single state was three. It was important that each state have the

same number of training sites so that none would be over or underestimated. This led us

to use only three training sites for each state. If a state had more than three sites, three of

them were randomly selected for use in our classification and distance computations.

Remotely sensed imagery provided the data used to calculate our NDVI metrics.

Four Landsat 5 Thematic Mapper (TM) images (Path 38/ Row 31) for years 2005-2008

with Julian date as close to 207 (July 26th) as possible were downloaded from the U.S.

Geological Survey Global Visualization Viewer (GLOVIS). The Julian date of 207 was

chosen by averaging the date for each year that displayed the greatest variance in NDVI

between different land cover types. The dates were obtained by examining line graphs of

mean NDVI values collected by AVHRR of grasslands, shrubs, and deciduous forests.

These graphs can be obtained through GLOVIS using a tool called "NDVI graph" (U.S.

Geological Survey 2011). Figure 3-2 shows an example of one of these graphs from

2008. Landsat 5 TM images with minimal cloud cover and collection dates closest to 207

were selected. All images were rectified and resampled to the UTM Zone 12 NAD 1983

map projection. Each image was converted to percent reflectance values using an image-

based atmospheric correction (Chavez 1996) and the calibration coefficients for Landsat

5 TM (Chander et al. 2009). Following image standardization, we calculated NDVI

using the formula (NIR-RED)/ (NIR+RED). We also calculated the standard deviation in

NDVI using a 5x5 (22500 m^2 ground area) focal window that produced a standard

deviation in NDVI value for each pixel.

Classification and Similarity

To calculate the similarity of all MUs containing R034AY2ggUT to the state of

most probable membership, we first classified the area encompassed by these MUs into

the five alternative states using our training data (Fig. 3-3). The variables that we used

included the NDVI and standard deviation of NDVI calculated from the Landsat 5 TM

scenes. The classifier we used was a maximum likelihood classifier which is a form of

linear discriminant analysis. Maximum likelihood classification is one of the most

widely used supervised classification algorithms (McIver and Friedl 2002; Wu and Shao 2002).

All pixels in the MU with R034AY2ggUT as the largest component ES were classified into one of the five R034AY2ggUT states even though many of the pixels represented areas of much different vegetation type (e.g., agriculture, juniper, riparian zones). When our maximum likelihood classifier was executed, a Mahalanobis distance image was also produced (Fig. 3-4) as a standard output of the classification process. The pixels in this image were enumerated with the Mahalanobis distance to whichever state the pixel had been assigned to in the classification. Mahalanobis distance calculates the similarity of an observation with n-variables to a group of observations (training sets in our case) with n-variables (Mahalanobis 1936). Mahalanobis differs from Euclidean distance measures in that it takes into account the correlations of variables within the data set and it is scale invariant. Because Mahalanobis distance accounts for unequal variances and correlations between variables, it is able to assign different weights to the variables. Only when variables are uncorrelated will the Mahalanobis distance be equal to the Euclidean distance (Xian et al. 2008).

Field Work

Following the calculation of a Mahalanobis distance for each pixel, we verified that the distance metric corresponded with conditions in the field. Our assumption was that pixels with the largest Mahalanobis distances represented pixels that were less likely to be associated with any of the five different alternative states defined by the R034AY2ggUT STM model. These pixels were either another ES or they represented previously unconsidered states for the R034AY2ggUT ES. Conversely, pixels with low

distances represented vegetation cover conditions similar to one of the five states and pixels with moderate distances were somewhat similar. It was our belief that ecological state classifications would also be more accurate for sites with smaller Mahalanobis distances.

A stratified random sample of the Mahalanobis distance image's values was used to select field sites to validate. Because Mahalanobis distances are unitless, thresholding distances into similar, somewhat similar, and dissimilar can be subjective. To do this as objectively as possible we used the distance image's histogram (Fig. 3-5) to select these thresholds. The distribution of the distances was skewed to the right. The pixel value with the maximum occurrence in the image was 12. At the Mahalanobis distance of 52 a point of inflection occurred. Previous studies have used the maximum value and inflection points to identify similar threshold values such as dark object values and phenological stages (Chavez 1988; Sakamoto et al. 2005). With these thresholds we described distances of 0 - 12 as being similar, 12 - 52 as somewhat similar, and > 52 as dissimilar. Conceptually, the threshold at the distance of 12 represented the point at which every following interval of Mahalanobis distance had a lower pixel frequency. The threshold at 52 represented the point at which every following interval of distance had a much lower decrease in pixel frequency. While these thresholds did not necessarily relate to the ecological conditions of the areas represented by the pixels, they did serve as a starting point for identifying actual ecological breaks.

With our data stratified into three groups, we randomly selected twenty sites in each group for field validation. Areas identified as agriculture by the Southwest Regional Gap (GAP) (Prior-Magee 2007) were not included in the potential sample area.

Validation sites were visited during the summer of 2012. Two 60 m transects, with their

center point being one random sample location were used to apply the Daubenmire field

method. The Daubenmire method was chosen for its utility in estimating percent cover,

its simplicity and rapid application (U.S. Department of Agriculture, NRCS 1999). The

two transects were run in north-south and east-west directions. Square 1 m quadrats were

placed every 5 m along each transect and percent canopy cover was recorded for each

plot. When all plot canopy covers were collected, a site percent canopy cover was

calculated by averaging the plot percent canopy estimates. The percent canopy cover

data for each point was examined to determine whether there were plant species present

that were not indicative of one of the states described in the STM or that were correlated

with other ESs. This was a binary approach of recording whether each point had non-

R034AY2ggUT plant species present or not. The area sampled at each site was equal to

four Landsat 5 TM pixels. Of the 60 sites that were randomly generated, we were able to

access 56. Access to private property was the largest factor in not being able to sample

all points. Two of these points were in our similar class and two were in our dissimilar

class.

RESULTS

Ecological State Classification Accuracy

Each of our 56 validation points was assigned a state by comparing the percent

canopy cover collected during the field work to the plant communities described for each

state in the STM. Our *a priori* knowledge that many of our points in the dissimilar and

somewhat similar classes would not be correctly classified due to the fact that our MUs

included areas of completely different ESs (and therefore states) led us to construct three separate confusion matrices. A confusion matrix was built for each distance class (similar, somewhat similar, and dissimilar distances) (Tables 3-1 — 3-3). Because there were states present in our area that were not considered, an additional column was added to represent when a pixel was classified as a state from the R034AY2ggUT STM but in reality the pixel belonged to a state not identified within the STM. The percent correctly classified (PCC) for the points with Mahalanobis distance 0 - 12 was 64.7%. The Kappa value for these pixels was 0.50. Points with Mahalanobis distance 12 - 52 had a PCC of 17.7% and had a Kappa value of 0.03. The PCC for points with a Mahalanobis distance > 52 was 25.0% and had a Kappa value of 0.14.

States 2 and 4 (Fig. 3-1) had the highest PCC at 71.4% and 80.0% respectively when looking at points in the similar class. Using only points from the similar class, State 6 had the lowest PCC at 33.3% and no points were classified as belonging to State 3. When only using points from the somewhat similar class, the highest PCC was for State 2 at 28.6% and the lowest PCC was for State 4 and 6 at 0.0%. All of the State PCCs for the dissimilar points were 0.0% accurate except for State 2 which had a PCC of 100%. All distance classes contained points that were misclassified as belonging to a state from the R034AY2ggUT STM.

Ecological Site Similarity Assessment

Of the 56 sampled areas, at least a portion of 18 of them were in an ecological state not identified for that particular ES. Of the 18 points whose Mahalanobis pixels were classified as being similar (0 - 12), only one had plant species present that were not

associated with R034AY2ggUT states. Three out of the 20 points that were classified as

being somewhat similar (12 – 52) had plant species present that were not associated with

R034AY2ggUT states. Of the 18 points that had Mahalanobis pixel values classified as

dissimilar (> 52) only four had exclusively R034AY2ggUT plant species present. These

results are summarized in Figure 3-6.

Because species data were recorded for the area within 30 m of each random point

location, we also examined the spectral data by averaging the four nearest pixels' values

to the point (60 meter buffer). We calculated the differences between the point pixel

Mahalanobis distance values and the area-averaged Mahalanobis distance values.

Overall, the Mahalanobis values differed by less than 0.5; however, some of the

differences were quite large with one sample location showing a 2000% difference in

Mahalanobis distance between the averaged value and the point specific value. Using the

averaged values, 14 validation sites had Mahalanobis distance values below the first

threshold (< 12), 26 sites occurred in the somewhat similar class (12 – 52), and 16 sites

were found in the dissimilar class (> 52). Only one of the sites with Mahalanobis

distance less than 12 had plant species present that were not linked with the

R034AY2ggUT ES. Of the sites with distances between 12 and 52, three had plant

species that were not associated with our specific ES. Fourteen of the sites with distances

greater than 52 had plant species present that were associated with dissimilar ESs. These

results (Fig. 3-7) are very similar to those summarized in Figure 3-6.

After obtaining the percentages of areas that had plant species present from other

ESs for each class (0 - 12, 12 - 52, and > 52) we desired to see if the trend of increasing

percentages of non-R034AY2ggUT plant species could be seen within these classes (Fig.

3-8). We created six classes from the three by adding a threshold at the halfway point within each class. Because no halfway point existed for the dissimilar class (the class represented values of 52 to infinite), we instead created two classes which each held half of the samples. These six new classes were separated at thresholds of 6, 12, 32, 52, and 120. Both areas with Mahalanobis distance between 0 and 6 had only R034AY2ggUT plant species present. Eleven of 12 areas in the distance class of 6-12 were exclusively made up of R034AY2ggUT plant species. Fifteen of 16 areas in the class from 12 - 32 was made up of areas with only R034AY2ggUT plant species. Eight of 10 areas in the distance range of 32-52 contained only R034AY2ggUT plant species. The distance range of 52-120 had only two of its eight areas exclusively made up of R034AY2ggUT plant species. The last distance class, greater than 120, had no points out of eight that were exclusively made up of R034AY2ggUT plant species.

<div align="center">DISCUSSION</div>

Ecological State Classification

Implementations of STM concepts are increasing in the Western United States for field-level assessments of vegetation and soil condition at discrete locations (Steele et al. 2012). These field-level assessments cannot be used for comprehensive management of large landscapes (Fuhlendorf et al. 2006; Briske et al. 2008). With an increasing desire to incorporate detailed ecological data for landscape scale decision-making , a repeatable and dependable method of mapping ecological states across a large landscape is necessary(Karl and Sadowski 2005; Forbis et al. 2007; Ludwig et al. 2007; Steele et al. 2012). We have demonstrated that remote sensing can aid in this process. We calculated

a PCC of 64.7% for all pixels with Mahalanobis distances less than 12. These pixels

comprised about 26% of all non-irrigated areas within the R034AY2ggUT ES. Percent

Correctly Classified dropped significantly for pixels with higher Mahalanobis distances

showing that the Mahalanobis distance is an appropriate metric to identify areas that were

either correctly or incorrectly classified. Land managers can use the Mahalanobis

distance to identify areas where the automated state classification product will be helpful

in creating ecological state maps.

The difficulty with accurately classifying states within the R034AY2ggUT ES lies

in the fact that the differences in plant species composition in each state do not provide a

sufficient spectral discrimination. For example, the differences between ecological states

2 and 4 are functionally very small. These states are nearly identical with the exception

of an increase in invasives such as mustards and cheatgrass in State 4. The dominant

plant species, Wyoming big sagebrush, is constant throughout both states. Therefore, we

find that this method of pixel-based classification to map ecological states is appropriate

for those states that are distinct from each other, but not for states that have subtle

difference. These findings in part confirm the conclusions of Steele et al. (2012) that the

accurate mapping of ecological states using common classification algorithms is difficult.

However, ecological state classification can have a significant utility as a supportive,

ancillary dataset to assist land managers in the process of drawing new MU boundaries to

more closely match specific ecological sites.

Ecological Site Inclusions

Several types of inclusions and one ecological state that were not accounted for in the STM for R034AY2ggUT were identified through this process Approximately 80% of the validation sites located in the dissimilar class contained plant species that were either associated with other ESs or were not accounted for in the STM. Of these points, 29% contained black sagebrush (*Artemisia nova* A. Nelson) which occurs on shallow, stony soils (Zamora and Tueller 1973). Fourteen percent of the points contained basin big sagebrush which is generally found on deep, well-drained soils in valley bottoms. Another 29% contained plant communities that are typical of another ES which is a mixture of basin big sagebrush (*Artemisia tridentata* Nutt. ssp. *tridentata*), basin wildrye (*Leymus cinereus* [Scribn. & Merr.] Á. Löve), and thickspike wheatgrass (*Elymus lanceolatus* [Scribn. & J.G. Sm.] Gould ssp. *lanceolatus*). The remaining points (21%) contained greasewood (*Sarcobatus* Nees) which is part of another ES occurring on finely textured, highly saline soils. All of these areas were considered inclusions (which are defined as minority ESs within an MU) because ESs existed whose plant profile matched the plant communities at these sites. These plants' ESs frequently occur within the same MU as the R034AY2ggUT ES. We have demonstrated that it is possible to map these inclusions within MUs through the use of the Mahalanobis distance.

Only one of these sites could be considered an alternative ecological state of the R034AY2ggUT ES but was not accounted for in the associated STM. This site contained a high amount of Utah juniper (*Juniperus osteosperma* [Torr.] Little) at 26.7% canopy cover as well as plants that were typical of R034AY2ggUT such as Wyoming big sagebrush (4.5%), rabbitbrush (4%), and Kentucky bluegrass (5.6%). However, there is

no state within the R034AY2ggUT STM that details any encroachment of Utah juniper. Additionally, there is another ES, R034AY2rrUT (Semi-desert Shallow Breaks (Caespitose Bluebunch Wheatgrass/ Utah Juniper)), that has a similar described plant community to what we found at the site. A decision must be made as to whether an update to the R034AY2ggUT STM needs to be made or whether this site is a completely different ES. This decision would be based on the soil characteristics at the site.

Four field sites containing plant species not attributed to the R034AY2ggUT ES were found in the similar and somewhat similar distance classes. One of these sites, located within the similar distance pixels, was largely made up of black sagebrush (22% canopy cover). We have no explanation as to why this site was classified as being similar to the R034AY2ggUT ES. The mean and standard deviation of NDVI values at this site were similar to those of our training data. The other three sites that had different plant species present were found in the somewhat similar class' pixels. These sites contained different combinations of basin big sagebrush, black sagebrush, and Utah Serviceberry (*Amelanchier utahensis* Koehne). Finding a few sites with plant species typical of other ESs was expected for points in the somewhat similar class. Likewise, we also expected to have a few sites that were part of the R034AY2ggUT ES in the dissimilar class.

A few factors may have contributed to the inability to identify a Mahalanobis distance value that cleanly separated pixels that represented areas of different ES. One of these issues could have been the standard deviation in NDVI variable that was used. This variable allowed us to separate areas in our ES of interest, which have low spectral variance, from other ESs that have higher variance such as riparian areas. However, this variable also expanded the estimated area of dissimilar ES around each area of higher

variance. This was caused by the way that the standard deviation was calculated. The only way to calculate the standard deviation for an area is to consider the pixels surrounding the pixel of interest. We used a 5 x 5 pixel focal window in each calculation of standard deviation. This means that the standard deviation of a pixel that was in reality an R034AY2ggUT pixel could potentially be mischaracterized by an area of high variance up to 60 m away.

Another issue was heterogeneity among states within the R034AY2ggUT ES. Sometimes the MU containing our specific ES would have two or more states in close proximity. If these states had contrasting NDVI values, then this caused the standard deviation in NDVI to increase above normal levels and an exaggerated Mahalanobis distance would be obtained.

There are other limitations to this methodology. Remote sensing cannot be used to obtain detailed information about soils. Our methodology makes the assumption that since plants from other ESs (and plants not detailed in the associated STM) were present at a site, that at least some of the area was part of a different ES. We did not attempt to verify this assumption through soil work. Only soil sampling can positively identify the extent of ESs. The distance image we produced with its probabilities of ES membership could provide a way to effectively choose sample sites for soil field work.

IMPLICATIONS

Our research has shown that a pixel-based classification shows promise as a means of separating distinct ecological states, but has difficulty separating states that are compositionally similar. Therefore, this method should be used in combination with

other techniques to identify ecological states within a specific ES. This technique can assist and supplement manual delineations of ecological states as described by Steele et al. (2012). Areas with small Mahalanobis distance had a much higher classification accuracy and could therefore be used as a basis for where states occurred.

A similarity index like the Mahalanobis distance can be applied at the landscape scale to locate areas of similarity to a specific ecological state. The method described here can help define where ecological states of a given ES occur on a landscape. Furthermore, the similarity index can be used in its original pixel value, categorized into discrete similarity categories, or converted into probability of ES membership through field work and used as a predictor variable in a more advanced classification algorithm such as random forests or an object-oriented classification tool. This would be helpful when classifying multiple states from a variety of possible ESs across a large area.

This method can be easily replicated by land managers for multiple ecological sites and states. Existing field data is available from a variety of government and educational organizations that could be used to both classify ecological states and calculate Mahalanobis distances. However, a posteriori field work will need to be done similar to our study to validate at what Mahalanobis distances the probability of ES membership decreases dramatically. After this data is created, it could be distributed with the tabular soil data in the NRCS SSURGO database (U.S. Department of Agriculture, NRCS-SSURGO 2012).

LITERATURE CITED

Arid Land Research Programs. 2010. Jornada Experimental Range. Available at: http://jornada.nmsu.edu/esd/intro. Accessed 10 May 2012.

Beetle, A. A. 1960. A study of sagebrush—the section Tridentatae of *Artemisia*. Laramie, WY: Wyoming Agricultural Experiment Station Bulletin 368. 83 p.

Briske, D. D., S.D. Fuhlendorf, and F. E. Smeins. 2005. State-and-transition models, thresholds, and rangleland health: a synthesis of ecological concepts and perspectives. *Rangeland Ecology & Management* 58:1-10.

Briske, D. D., B. T. Bestelmeyer, T. K. Stringham, and P. L. Shaver. 2008. Recommendations for development of resilience-based state-and-transition models. *Rangeland Ecology and Management* 61:359-367.

Chander, G., B. L. Markham, and D. L. Helder. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment* 113:893-903.

Chavez, P. S. 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment* 24: 459-479.

Chavez, P. S. 1996. Image-based atmospheric corrections revisited and improved. *Photogrammetric Engineering and Remote Sensing* 62:1025-1036.

Connelly, J. W., S. T. Knick, M. A. Schroeder, and S. J. Stiver. 2004. Conservation assessment of greater sage-grouse and sagebrush habitats. Cheyenne, WY: Western Association of Fish and Wildlife Agencies, Unpublished Report. 610 p.

Forbis, T. A., L. Provencher, L. Turner, G. Medlyn, J. Thompson, and G. Jones. 2007. A method for landscape-scale vegetation assessment: application to Great Basin rangeland ecosystems. *Rangeland Ecology & Management* 60:209-217.

Frisina, M. R., and C. L. Wambolt. 2004. A guide to identifying the four subspecies of big sagebrush. *Rangelands* 26:12-16.

Fuhlendorf, S. D., W. C. Harrell, D. M. Engle, R. G. Hamilton, C. A. Davis, and D. M. Leslie. 2006 Should heterogeneity be the basis for conservation? Grassland bird response to fire and grazing. *Ecological Applications* 16:1706-1716.

Galatowitsch, S.M. 1990. Using the original land survey notes to reconstruct pre-settlement landscapes in the American West. *Great Basin Naturalist* 50:181-191.

Gamon, J. A., C. B. Field, M. L. Goulden, K. L. Griffin, and A. E. Hartley. 1995. Relationships between NDVI, canopy structure, and photosynthesis in three Californian vegetation types. *Ecological Applications* 5:28-41.

Hernandez, A. 2011. Spatiotemporal modeling of threats to big sagebrush ecological sites in northern Utah [dissertation]. Logan, UT: Utah State University. 163 p.

Huete, A. R., K. Didan, T. Miura, E. P. Rodriguez, X. Goa, and G. Gerreira. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83:195-213.

Iglesias, R. R., and M. K. Kothmann. 1997. Structure and causes of vegetation change in state and transition model applications. *Journal of Range Management* 50:399-408.

Jensen, J. R. 2000. Remote sensing of the environment: an earth resources perspective. Upper Saddle River, NJ: Prentice-Hall. 592 p.

Karl, M., and J. Sadowski. 2005. Assessing big sagebrush at multiple spatial scales: an example in southeast Oregon. Denver, CO, USA: US Department of the Interior, Bureau of Land Management, Technical note 417. 41 p.

Knick, S. T., A. L. Holmes, J. T. Rotenberry, M. A. Schroeder, W. M. Vander Haegen, and C. van Ripper III. 2003. Teetering on the edge or too late? Conservation and research issues for avifauna of sagebrush habitats. *Condor* 105:611-634.

Landfire Biophysical Settings (BPS) Map. 2008. Landfire. Available at: http://www.landfire.gov/index.php. Accessed 12 March 2012.

Landfire Existing Vegetation Type (EVT) Map. 2008. Landfire. Available at: http://www.landfire.gov/index.php. Accessed 12 March 2012.

Lowry, J.L. Jr., R. D. Ramsey, K. Boykin, D. Bradford, P. Comer, S. Falzarano, W. Kepner, J. Kirby, L. Langs, J. Prior-Magee, G. Manis, L. O'Brien, K. Pohs, W. Rieth, T. Sajwaj, W. Schrader, K. A. Thomas, D. Schrupp, K. Schulz, B. Thompson, C. Wallace, C. Velasquez, E. Waller, and B. Wolk. 2007. Mapping Meso-scale land cover over very large geographic areas within a collaborative framework: a case study of the Southwest Regional Gap Analysis Project (SWReGAP). *Remote Sensing of Environment* 108:59-73.

Ludwig, J. A., G. N. Bastin, J. F. Wallace, and T. M. McVicar. 2007. Assessing landscape health by scaling with remote sensing: when is it not enough? *Landscape Ecology* 22:163-169.

Mahalanobis, P. C. 1936. On the generalized distance in statistics. Proceedings of the National Institute of Sciences of India; 16 April 1936; India: National Institute of Sciences of India. p. 49-55.

Maynard, C. L., R. L. Lawrence, G. A. Nielsen, and G. Decker. 2007. Modeling vegetation amount using bandwise regression and ecological site descriptions as an alternative to vegetation indices. *GIScience & Remote Sensing* 44:68-81.

McIver, D. K., and M. A. Friedl. 2002. Using prior probabilities in decision-tree classification of remotely sensed data. *Remote Sensing of* Environment 81:253-261.

Miller, R. F., and J. A. Rose. 1999. Fire history and western juniper encroachment in sagebrush steppe. *Journal of Range Management* 52:550-559.

Parson, R. E. 1996. A History of Rich County. Salt Lake City, UT, USA: Utah State Historical Society, County Commission, Rich County, Utah.

Peterson, K. A. 2009. Modeling potential native plant species distributions in Rich County, Utah [thesis]. Logan, UT: Utah State University. 107 p.

Prior-Magee, J. S. 2007. Chapter 1. Moscow, ID: U.S. Geological Survey, Final Report for Southwest Regional Gap Analysis. 13 p.

Sakamoto, T., M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka, and H. Ohno. 2005. A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment* 96:366-374.

Stringham, T. K., W. C. Krueger, and P. L. Shaver. 2003. State and transition modeling: an ecological process approach. *Journal of Range Management* 56:106-113.

Steele, C. M., B. T. Bestelmeyer, L. M. Burkett, P. L. Smith, and S. Yanoff. 2012. Spatially explicit representation of state-and-transition models. *Rangeland Ecology &* Management 65:213-222.

U.S. Department of Agriculture. 2011. Ecological Sites. Available at: http://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/landuse/rangepasture/?cid=STELPRDB1043235. Accessed 17 October 2011.

U.S. Department of Agriculture, NRCS. 1999. Sampling Vegetation Attributes. Available at: http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1044175.pdf. Accessed 10 May 2012.

U.S. Department of Agriculture, NRCS. 2005. MRLA definitions. Available at: http://soils.usda.gov/survey/geography/mlra/mlra_definitions.html. Accessed 10 January 2012.

U.S. Department of Agriculture, NRCS. 2006. Inventorying and monitoring grazing land resources. *In:* U.S. Department of Agriculture, NRCS [EDS.]. National range and pasture handbook. Washington, DC, USA: United States Department of Agriculture.

U.S. Department of Agriculture, NRCS. 2011. National soil survey handbook. Available at: http://soils.usda.gov/technical/handbook/contents/part622.html#07. Accessed 5 January 2012.

U.S. Department of Agriculture, NRCS, 2012. Ecological Site Description. Available at: http://esis.sc.egov.usda.gov/Welcome/pgESDWelcome.aspx. Accessed 10 January 2012.

U.S. Department of Agriculture, NRCS-SSURGO. 1982. Soil Survey of Rich County Utah. Available at: http://soildatamart.nrcs.usda.gov/manuscripts/UT604/0/ rich.pdf. Accessed 8 May 2012.

U.S. Department of Agriculture, NRCS-SSURGO. 2012. Soil data mart. Available at: http://soildatamart.nrcs.usda.gov/. Accessed 10 October 2011.

U.S. Geological Survey. 2011. *NDVI Graph.* Available at: http://glovis.usgs.gov/. Accessed 10 January 2012.

Utah Division of Wildlife Resources. 2006. Range trend studies. Available at: http://wildlife.utah.gov/range/wmu4.htm. Accessed 10 April 2012.

Utah Office of Tourism. 2009. Rich County tourism profile. Available at: http://www.travel.utah.gov/research_and_planning/documents/ RichCounty_005.pdf. Accessed 5 January 2012.

Welch, B. L., F. J. Wagstaff, and J. A. Roberson. 1991. Preference of wintering sage grouse for big sagebrush. *Journal of Range Management* 44:462-465.

West, N. E. 1991. Benchmarks for rangeland management and environmental quality. *In:* L. F. James, J. O. Evans, M. H. Ralphs, and R. D. Child [eds.]. Noxious range weeds. Boulder, CO: Westview Press. p. 31-44.

Westoby, M., B. Walker, and I. Noy-Meir. 1989. Opportunistic management for rangelands not at equilibrium. *Journal of Range Management* 42:266-274.

Wu, W., and G. Shao. 2002. Optimal combinations of data, classifiers, and sampling methods for accurate characterizations of deforestation. *Canadian Journal of Remote Sensing* 28:601-609.

Xian, S., F. Nie, and C. Zhang. 2008. Learning a Mahalanobis distance metric for data clustering and classification. *Pattern Recognition* 41:3600-3612.

Zamora, B., and P. T. Tueller. 1973. *Artemisia arbuscula, A. longiloba,* and *A. nova* habitat types in northern Nevada. *Great Basin Naturalist* 33:225-242.

Table 3-1. Confusion matrix for the similar field sites. Similarity classes were based on the distribution of Mahalanobis distances for each pixel classified by the Natural Resources Conservation service as being part of a map unit with a majority R034AY2ggUT ecological site (ES). A column was added and labeled "Other ES" to represent when a pixel was classified as being one of the five states but in reality was in a different ES altogether. ES, ecological site.

		Similar field sites				
		Predicted				
	State	2	3	4	5	6
Actual	2	**5**		1		2
	3				1	
	4			**4**		
	5	2			**1**	
	6					**1**
	Other ES			1		

Table 3-2. Confusion matrix for the somewhat similar field sites. Similarity classes were based on the distribution of Mahalanobis distances for each pixel classified by the Natural Resources Conservation service as being part of a map unit with a majority R034AY2ggUT ecological site (ES). A column was added and labeled "Other ES" to represent when a pixel was classified as being one of the five states but in reality was in a different ES altogether. ES, ecological site.

|		Somewhat similar field sites				
		Predicted				
	State	2	3	4	5	6
Actual	2	2				
	3	3	1			1
	4	1	2			2
	5	1	2	1		1
	6					
	Other ES		1			2

Table 3-3. Confusion matrix for the dissimilar field sites. Similarity classes were based on the distribution of Mahalanobis distances for each pixel classified by the Natural Resources Conservation service as being part of a map unit with a majority R034AY2ggUT ecological site (ES). A column was added and labeled "Other ES" to represent when a pixel was classified as being one of the five states but in reality was in a different ES altogether. ES, ecological.

		Dissimilar field sites				
		Predicted				
	State	2	3	4	5	6
Actual	2	1				
	3					
	4		1			
	5			1		
	6		1			
	Other ES	3	8		1	2

**R034AY2ggUT: Semi-desert Loam
(Wyoming Big Sagebrush/ Caespitose Bluebunch Wheatgrass)**
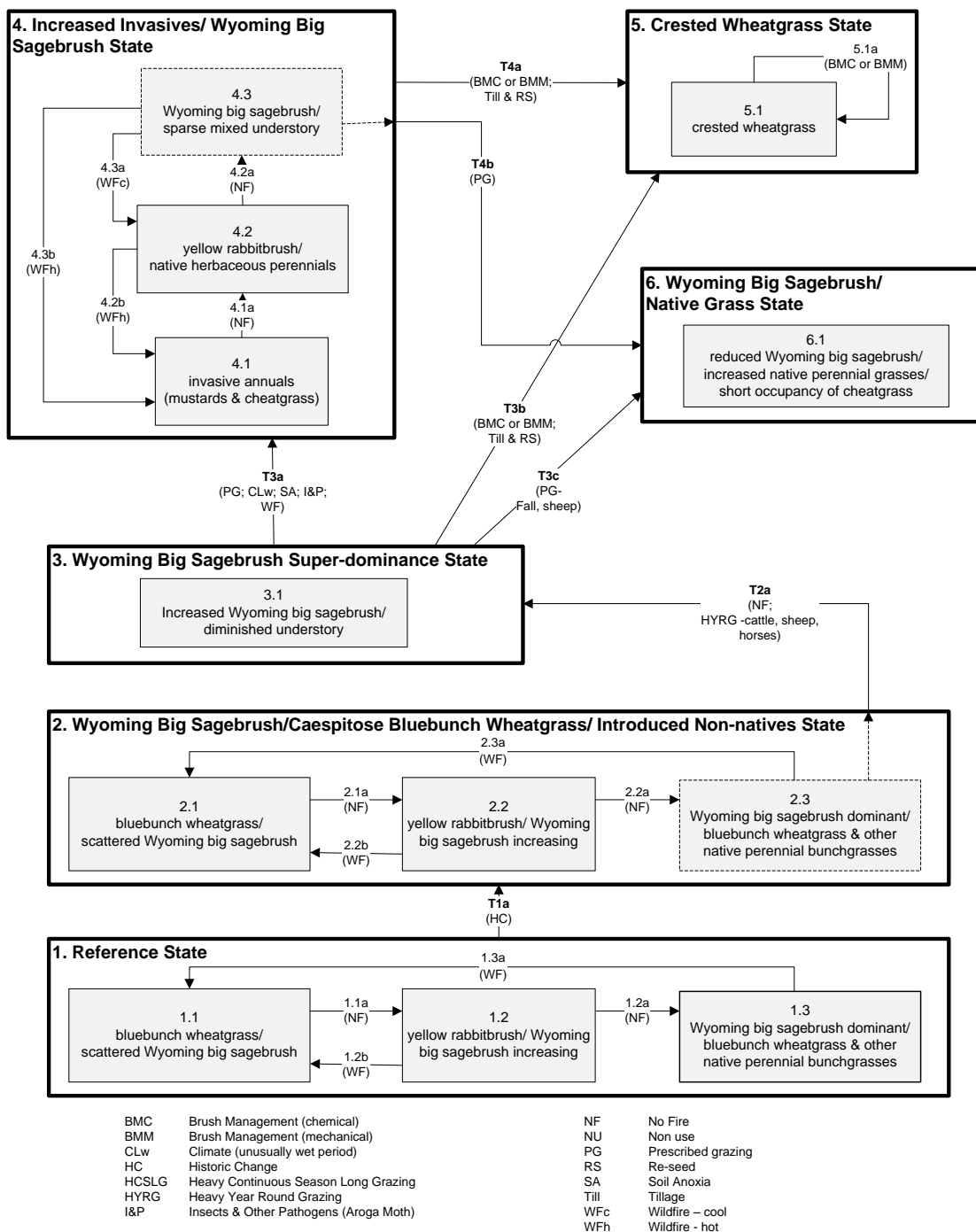
Figure 3-1. State-and-transition model for the R034AY2ggUT ecological site (U.S. Department of Agriculture, NRCS 2012). Each numbered box represents a state. Boxes with decimal numbers represent phases within a state.
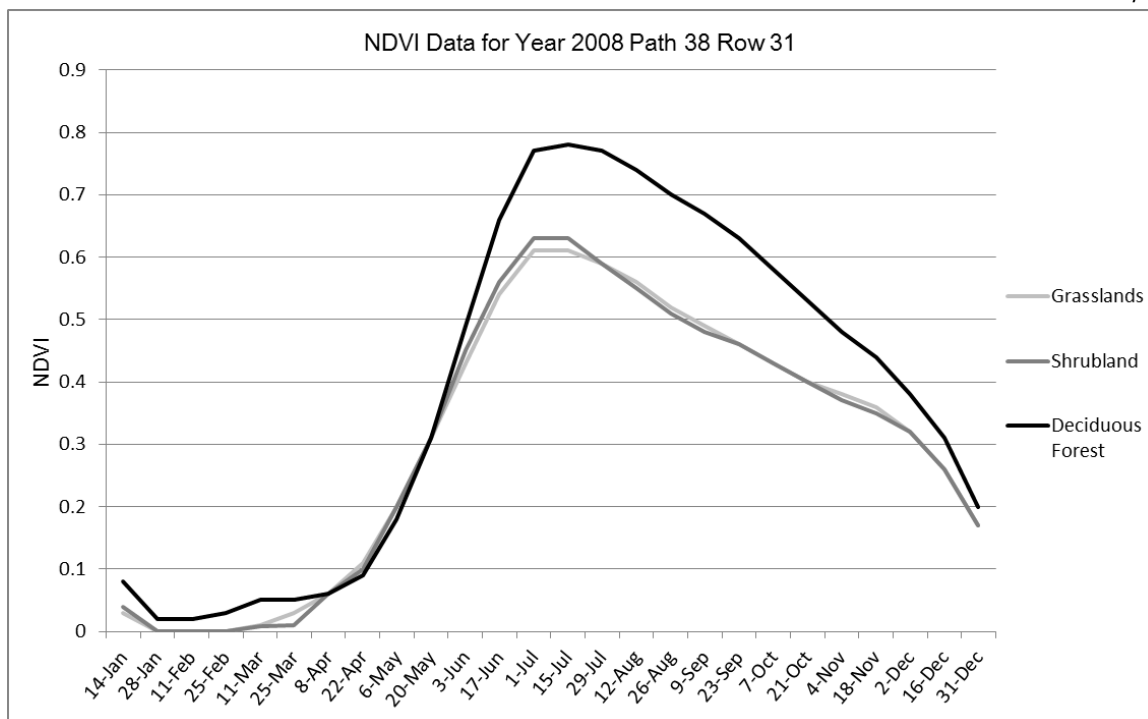
Figure 3-2. Line graph of annual fluctuations in NDVI for grasslands, shrublands, and deciduous forests. The largest differences in NDVI can be seen in mid-summer. Similar graphs can be obtained from the USGS GLOVIS Visualization Viewer at http://glovis.usgs.gov/. NDVI, Normalized Difference Vegetation Index.
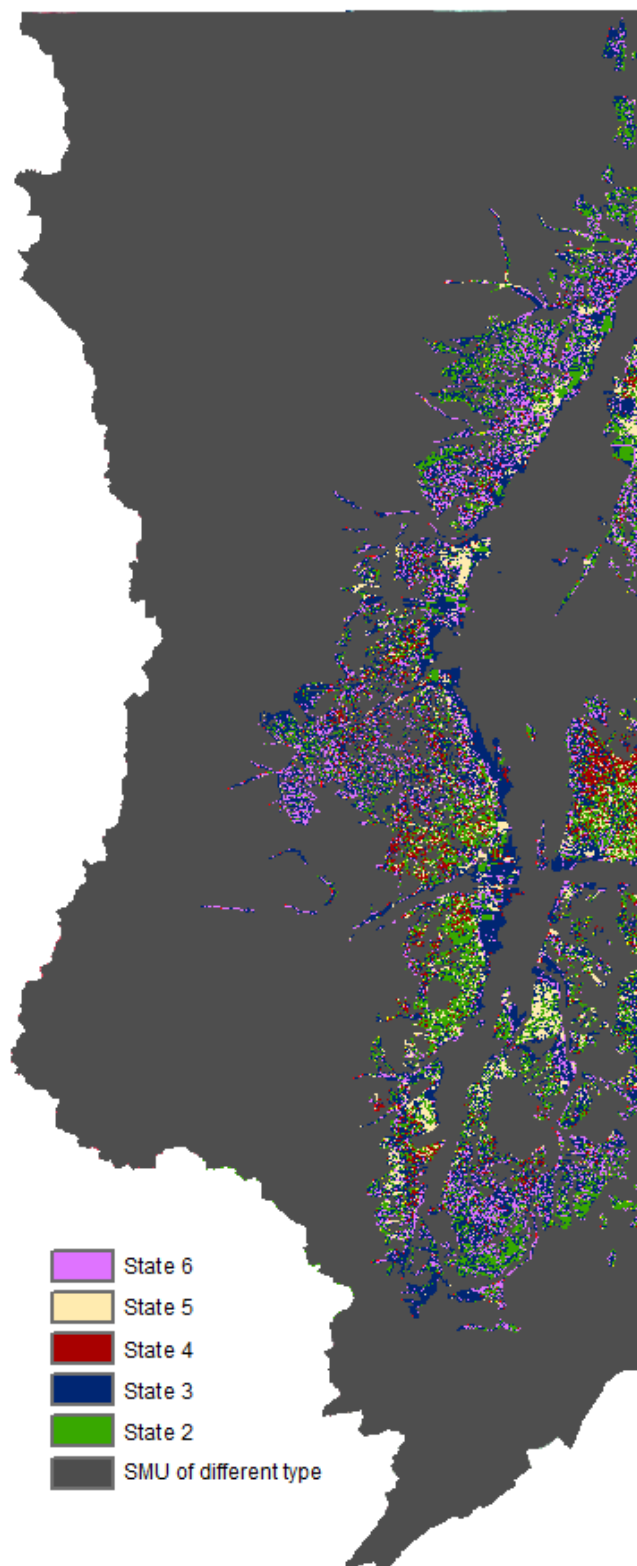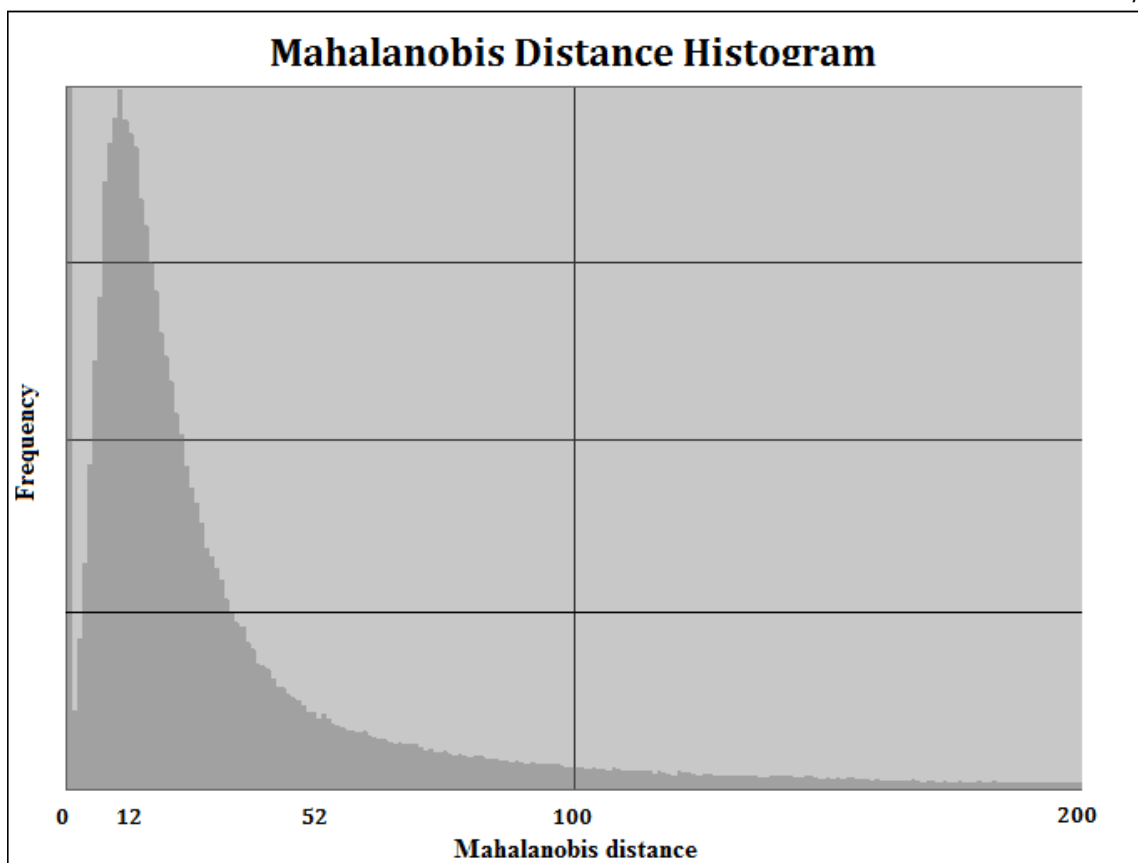
Figure 3-3. State Classification map of all areas in the R034AY2ggUT ecological site in Rich County, UT.
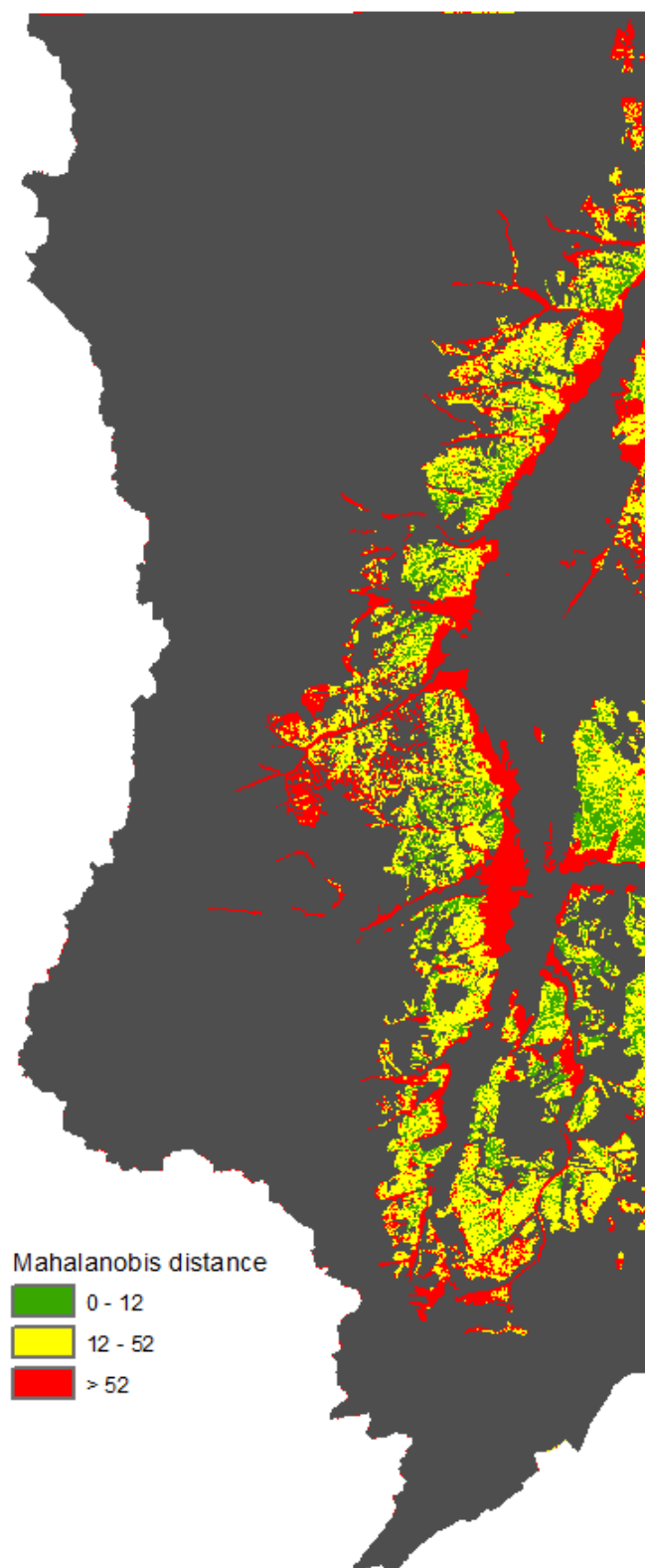
Figure 3-4. Mahalanobis distances for all areas in the R034AY2ggUT ecological site in Rich County, UT.

Figure 3-5. Distribution of pixels based on Mahalanobis distance from whichever state the pixel was classified as. Pixels with larger distances are more probable to be a different ecological site. Thresholds to stratify the data were placed at the Mahalanobis distances of 12 and 52.
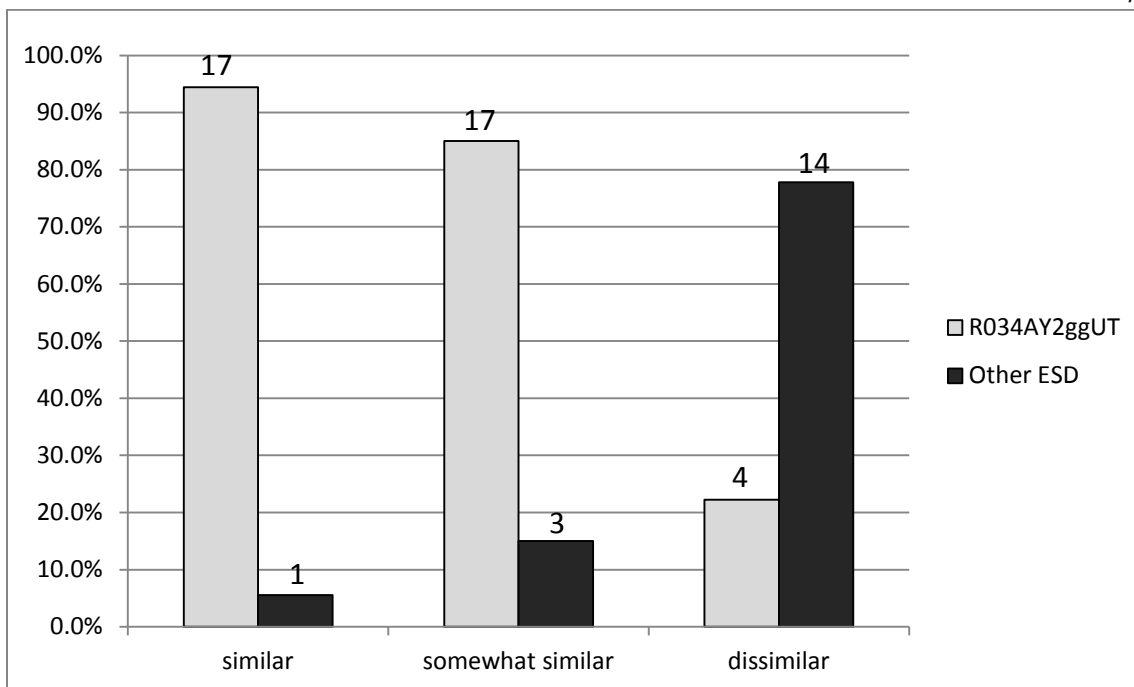
Figure 3-6. Bar graph showing the percentages of each similarity class that actually were within the R034AY2ggUT ecological site. Numbers at the top of the bars represent the total number of field sites in each category.
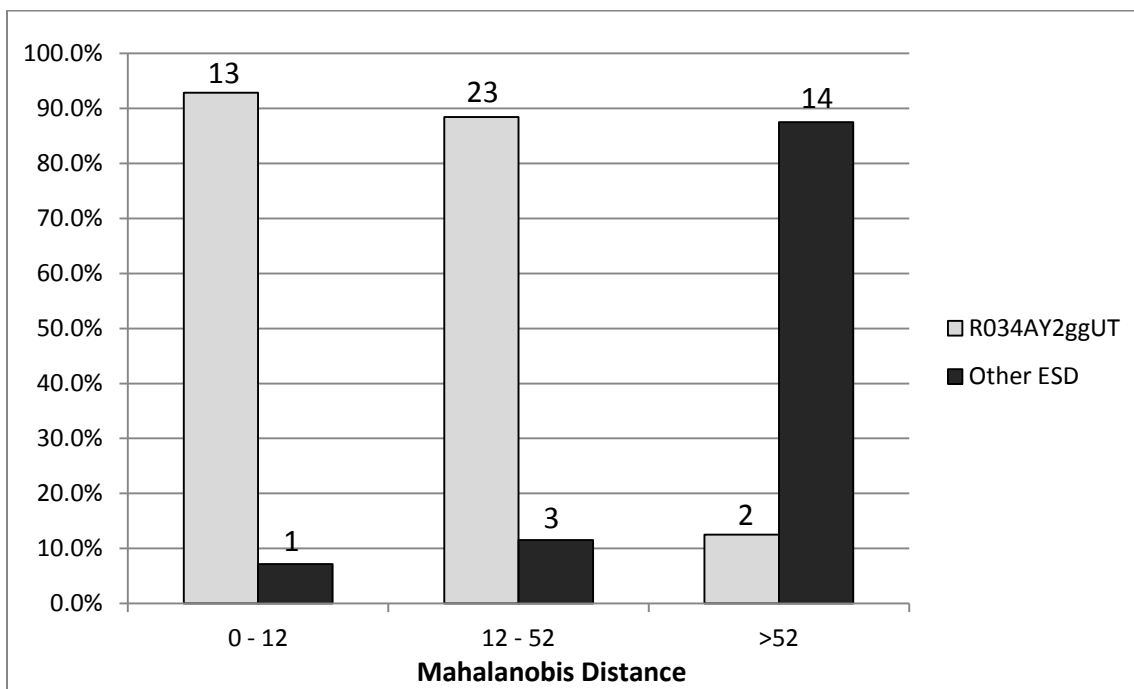
Figure 3-7. Bar graph showing the percentages of areas within each Mahalanobis distance range that actually were within the R034AY2ggUT ecological site. Numbers at the top of the bars represent the total number of field sites in each category.
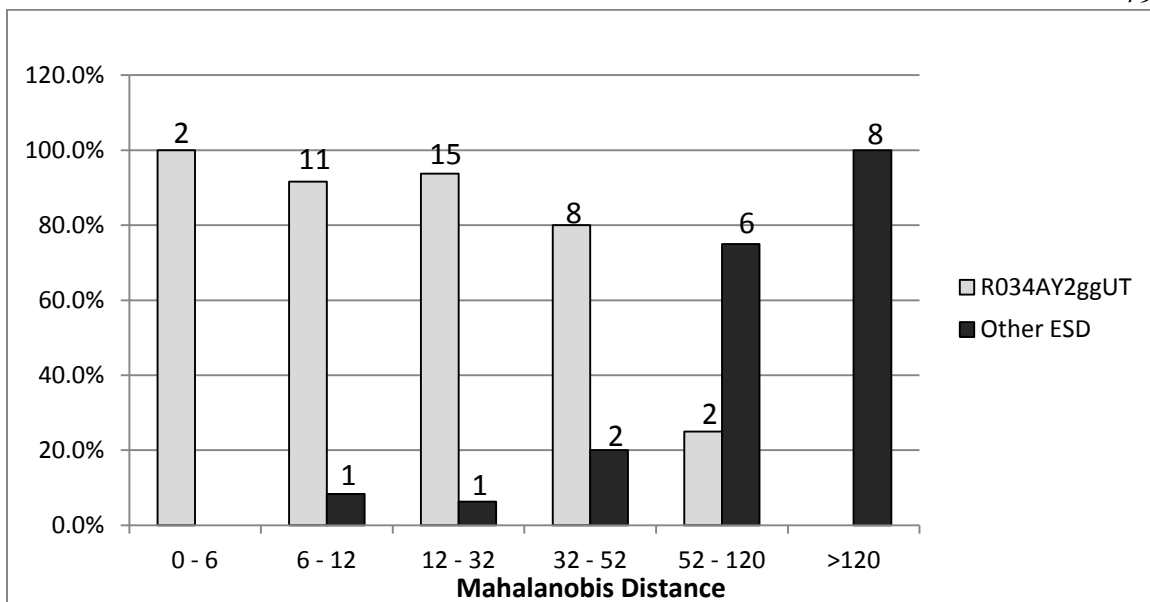
Figure 3-8. Bar graph showing the percentages of areas within each Mahalanobis distance range that actually were within the R034AY2ggUT ecological site. Extra ranges added by including additional thresholds at the midpoint of each range. Numbers at the top of the bars represent the total number of field sites in each category.

CHAPTER 4

CONCLUSION

The conceptual framework of ecological site descriptions (ESD) and state and transition models (STM) (Westoby et al. 1989) provides a way to record the historic plant communities as well as the current soil and plant properties at a given location. An ESD represents unique soil characteristics and the resulting plant species composition that occur on those soils. Ecological sites differ from one another in their ability to produce a distinct kind and amount of vegetation. Areas of the same ES, but separated by geography, are also unique in that they are assumed to "respond similarly to management actions and natural disturbances" (U.S. Department of Agriculture 2011). Each ESD has an associated STM that describes the different ecological states that can occur within an ES. STMs also describe how transitions to different states occur. Because of the information contained in ESDs and their associated STMs, they are a valuable decision support system that land managers can use in fragile ecosystems (Hernandez 2011).

The issue with the current state of ESs that we have identified in this thesis is that they are not explicitly delineated. Currently, ESs are identified on a landscape as components within map units (MU) with no specific spatial extent. In order for ESDs to be more useful to land managers, the spatial extent of ESs must be identified. Once ESs are mapped, their utility should be improved (Steele et al. 2012). The main goals of this research were to utilize common remote sensing techniques to 1) identify the spatial distribution of ecological sites and 2) identify the spatial distribution of states within ecological sites.

In Chapter 2 we addressed the first goal by identifying vegetation indices as well as biophysical variables that allowed us to discriminate between the vegetation components of selected ESs. The normalized difference vegetation index (NDVI) (Rouse et al. 1974) provided the most separation between vegetation components followed by the brightness component and then by the spatial variance of NDVI. A cluster analysis showed that the natural structure in the data would allow for separation between classes. We then applied the Random Forests decision tree algorithm (Breiman 2001) to our data resulting in an out-of-bag accuracy (cross-validation) of 97.2%. Our Random Forests model was then applied to all of Rich County, UT. Most of the vegetation components in our selected ESs were classified at greater than 90% accuracy. Our method accurately identified and discriminated between vegetation components that are unique to specific ESs. The resulting classified image from this process mapped the specific boundaries of vegetation components within MUs.

Chapter 3 utilized field work collected by Peterson (2009) and the Utah Division of Wildlife Resources (2006) to address both objectives using a similarity index rather than a decision tree model. Field sites were assigned an ecological state outlined by the STM for the Semi-desert Loam: Wyoming big sagebrush ES. A representative sample of each state was used to train a Maximum Likelihood classifier and subsequently classify each pixel identified by the U.S. Department of Agriculture: Natural Resources Conservation Service (NRCS) as being within our specific ES. A per-pixel Mahalanobis distance metric was produced during the image classification. The classification accuracy for pixels with low Mahalanobis distances was 64.7%. Classification accuracies were very low (<25%) for pixels with higher Mahalanobis distances (low

similarity). We found that the Mahalanobis distance metric is a suitable indicator of pixel membership to various ecological states of the Semi-desert Loam: Wyoming big sagebrush ES. We propose that Mahalanobis distances can be converted to probabilities of ecological site membership by performing field work. These results could help land managers delineate ecological sites and lead to a better understanding of landscape potential.

The work presented in Chapters 2 and 3 has demonstrated how common remote sensing techniques can help in the classification of ecological sites and ecological states. If implemented by land management agencies, these techniques will help clarify the vegetation potential of landscapes and help in policy-making decisions. The techniques in both chapters have implemented multi-temporal remotely sensed data sets. These were needed to average yearly changes in vegetation production due to climate variability.

Multi-temporal imagery coupled with field reconnaissance can be used to better delineate ecological sites and to some degree map different ecological states. Improved knowledge of the spatial distribution and extent of ES vegetation components can lead to improved delineation of soils as well as a better understanding of the different ecological state-and-transition forces occurring on these landscapes.

LITERATURE CITED

Breiman, L. 2001. Random forests. *Machine Learning* 45:5-32.

Hernandez, A. 2011. Spatiotemporal modeling of threats to big sagebrush ecological sites in northern Utah [dissertation]. Logan, UT: Utah State University. 163 p.

Peterson, K. A. 2009. Modeling potential native plant species distributions in Rich County, Utah [thesis]. Logan, UT: Utah State University. 107 p.

Rouse, J. W., R. H. Haas, J. A. Schell, and D. W. Deering. 1974. Monitoring vegetation systems in the great plains with ERTS. Proceedings of the 3rd Earth Resources Technology Satellite-1 Symposium, 10-14 December 1973; Greenbelt, MD: National Aeronautics and Space Administration. p. 3010-2017.

Steele, C. M., B. T. Bestelmeyer, L. M. Burkett, P. L. Smith, and S. Yanoff. 2012. Spatially explicit representation of state-and-transition models. *Rangeland Ecology &* Management 65:213-222.

U.S. Department of Agriculture, NRCS. 2011. National soil survey handbook. Available at: http://soils.usda.gov/technical/handbook/ contents/part622.html#07. Accessed 5 January 2012.

Utah Division of Wildlife Resources. 2006. Range trend studies. Available at: http://wildlife.utah.gov/range/wmu4.htm. Accessed 10 April 2012.

Westoby, M., B. Walker, and I. Noy-Meir. 1989. Opportunistic management for rangelands not at equilibrium. *Journal of Range Management* 42:266-274.