

MODELING SOIL WATER CONTENT
FOR PRECISION RANGE MANAGEMENT

by

Joel Brown Sankey

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

of

Master of Science

in

Land Rehabilitation

MONTANA STATE UNIVERSITY
Bozeman, MT

September 2005

© COPYRIGHT

by

Joel Brown Sankey

2005

All Rights Reserved

APPROVAL

of a thesis submitted by

Joel Brown Sankey

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Rick Lawrence

Approved for the Department of Land Resources and Environmental Sciences

Jon Wraith

Approved for the College of Graduate Studies

Joseph J. Fedock

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or parts may be granted only by the copyright holder.

Joel Brown Sankey

September 15, 2005

ACKNOWLEDGEMENTS

I would like to thank my advisor Rick Lawrence for setting an example of thoughtful, thorough, and efficient work. Some of his approach has rubbed off on me, hopefully. I would like to thank committee members David Brown, Gordon Decker, and Jon Wraith for their good ideas and willingness to consider all my questions. All the folks in David Brown's lab group – including Melissa Borino, Kevin Dalton, and Austin Allen – made a fun crew to work with and were instrumental in helping me navigate my way around the lab. Gordon Decker and Wes Henthorne were great hosts at their ranches and helped make field work very enjoyable. I would also like to thank the Upper Midwest Aerospace Consortium, especially MSU UMAC members Christine Sommers-Austin and Jerry Nielsen, for funding and supporting much of my work. Finally, I'd like to thank my family for their unwavering support during the ups and downs. My wife Teki, especially, has made the last two years a treat.

TABLE OF CONTENTS

1. INTRODUCTION	1
References Cited.....	5
2. LITERATURE REVIEW	6
Soil Water Storage.....	7
Evapotranspiration and Grazing	10
Remote Sensing of Evapotranspiration in Rangeland Systems	11
Terrain and Soil Controls on Soil Water Distribution	14
References Cited.....	21
3. AN EMPRICAL APPROACH TO MODELING AND MAPPING SPRING SOIL WATER DISTRIBUTION AT A RANCH SCALE	27
Introduction.....	27
Remote Sensing Approaches to Modeling Soil Water Content.....	28
Digital Terrain Modeling of Soil Water Content.....	30
Soil Attributes Used as Predictors of Soil Water.....	30
Objectives	31
Study Sites	33
Methods	35
Field Data Collection.....	36
GIS Data Set Development.....	37
Satellite Imagery	37
DEM-Derived Terrain Elements.....	37
SSURGO-Derived Soil Attribute Maps.....	38
Data Analysis.....	40
Results.....	43
Multiple Regression Analysis.....	43
Regression Tree Analysis	48
Spring Soil Water Content Maps.....	53
Discussion.....	54
Multiple Regression Models.....	57
Regression Tree Models	60
Conclusions.....	62
References Cited.....	64
4. EVALUATING THE SUITABILITY OF SOIL SURVEY AS A SOURCE OF SITE-SPECIFIC SOILS DATA FOR PREDICTIVE SOIL WATER CONTENT MODELING	68

TABLE OF CONTENTS -CONTINUED

Introduction.....	68
Soil Characterization With Diffuse Reflectance Spectroscopy	69
Objective.....	71
Study Sites	71
Methods	73
Field Data Collection	75
Lab Characterization.....	76
Carbon Analysis.....	77
Particle Size Analysis	78
Spectral Modeling of Soil Characteristics	78
Calculation of AWC and PAW.....	82
Soil Survey-Derived Characterization Data	83
GIS Data Set Development.....	84
Satellite Imagery	84
DEM-Derived Terrain Elements.....	84
Data Analysis.....	85
Results.....	87
Individual Predictors.....	87
Multiple Regression Models.....	89
Regression Tree Models	92
Discussion.....	97
Conclusion	101
References Cited.....	103
5. CONCLUSIONS.....	106

LIST OF TABLES

Table	Page
1. Average soil profile (100cm) mass water content (profile) models that performed best for calibration and were selected for validation for the BBar and Decker/Bales (D/B) ranches. Band numbers represent Landsat TM 5 reflectance bands. Slope is percent topographic slope. Aspect is cosine transformation of topographic aspect in degrees from North. Clay is entire profile weighted average percent clay of soil survey map unit major component. All models are significant at p -value < 0.05 and all variables are significant at p -value < 0.05 unless noted with [#] . Adjusted R ² values are presented.	44
2. Model validation results for models in Table 1. Models were independently validated with half the data set at each ranch (50 validation samples for Decker/Bales (D/B) and 41 for BBar ranch). Levene and Mann-Whitney statistics are p -values.....	44
3. Average soil profile (100cm) mass water content models constructed with decreased sample sizes at the Decker/Bales ranch. Band numbers represent Landsat TM 5 reflectance bands. Slope is percent topographic slope. Aspect is cosine transformation of topographic aspect in degrees from North. Clay is entire profile weighted average percent clay of soil survey map unit major component. OM is entire profile weighted average percent organic matter of soil survey map unit major component. PAW is entire profile weighted average plant available water (cm) of soil survey map unit major component. All models are significant at p -value < 0.05 and all variables are significant at p -value < 0.05 unless noted with [#] . Adjusted R ² values are presented.	47
4. Average soil profile (100cm) mass water content models constructed with decreased sample sizes for the BBar ranch. Band numbers represent Landsat TM 5 reflectance bands. Slope is percent topographic slope. Clay is entire profile weighted average percent clay of soil survey map unit major component. All models are significant at p -value < 0.05 and all variables are significant at p -value < 0.05 . Adjusted R ² values are presented.....	47

LIST OF TABLES -CONTINUED

Table	Page
5. Validation results for decreased sample size models constructed for the Decker/Bales ranch and presented in Table 3. Models were independently validated with the full validation data set (50 validation samples for Decker/Bales). Mann-Whitney statistic is p -value.	48
6. Validation results for decreased sample size models constructed for the BBar ranch and presented in Table 4. Models were independently validated with the full validation data set (41 validation samples for BBar). Mann-Whitney statistic is p -value.	48
7. Validation results for all regression tree models constructed for the Decker/Bales ranch. Results for the regression tree constructed with 50 samples refers to validation of the Decker/Bales tree presented in Figure 5. Subsequent results are for trees constructed with decreased sample sizes. Tree models were independently validated with the full validation data set (50 validation samples for Decker/Bales). Mann-Whitney statistic is p -value.	51
8. Validation results for all regression tree models constructed for the BBar ranch. Results for the regression tree constructed with 41 samples refers to validation of the BBar tree presented in Figure 6. Subsequent results are for trees constructed with decreased sample sizes. Tree models were independently validated with the full validation data set (41 validation samples for BBar). Mann-Whitney statistic is p -value.....	52
9. Site-specific validation results for clay DRS models. Weights refer to relative weighting of samples from NRCS and ranch data sets in boosted regression tree models. Root mean square deviation (RMSD) is average prediction error for clay models. Models were constructed with 10 iterations of a 1/10 holdout using data from both ranches. The 10 held out subsets were stratified by ranch, so validation was per ranch by 1/5 holdout.....	81

LIST OF TABLES - CONTINUED

Table	Page
10. Site-specific validation results for SOC DRS models. Weights refer to relative weighting of samples from NRCS and ranch data sets in boosted regression tree models. Root mean square error (RMSD) is average prediction error for SOC models, respectively. Models were constructed with 10 iterations of a 1/10 holdout using data from both ranches. The 10 held out subsets were stratified by ranch, so validation was per ranch by 1/5 holdout.....	81
11. Soil variables as individual predictors of average profile (100cm) mass water content. Soil survey variables were derived from soil survey maps and attribute data. Lab characterization variables were derived from characterization with diffuse reflectance spectroscopy and field characterization.....	88
12. Average soil profile (100cm) mass water content (profile) models using lab characterization data that performed well for calibration and were selected for validation for the BBar and Decker/Bales (D/B) ranches. Band numbers represent Landsat TM 5 reflectance bands. Slope is percent topographic slope. Clay lab is average of percent clay content for characterized profile samples. Field depth is depth to root restrictive layer within 100cm. All models are significant at p -value < 0.05 and all variables are significant at p -value < 0.05 unless noted with #. Adjusted R^2 values are presented.....	90
13. Average soil profile (100cm) mass water content (profile) models using soil survey data that performed well for calibration and were selected for validation for the BBar and Decker/Bales (D/B) ranches. Band numbers represent Landsat TM 5 reflectance bands. Slope is percent topographic slope. Aspect is cosine transformation of topographic aspect in degrees from North. Clay is entire profile weighted average percent clay of soil survey map unit major component. All models are significant at p -value < 0.05 and all variables are significant at p -value < 0.05 unless noted with #. Adjusted R^2 values are presented.....	90

LIST OF TABLES - CONTINUED

Table	Page
14. Validation statistics for models in Tables 12 and 13. Models were independently validated with half the data set at each ranch (50 validation samples for Decker/Bales (D/B) and 41 for BBar ranch). Levene and Mann-Whitney statistics are <i>p</i> -values.....	91
15. Hypothesis test <i>p</i> -values comparing the differences between predicted and observed water content values for multiple regression models from the two data sources for each ranch.....	92
16. Validation statistics for regression tree models constructed with soil survey soil variables and lab characterization soil variables.....	93
17. Mann-Whitney paired comparison of differences between predicted and observed water content values for regression tree models from the two data sources for each ranch.	94

LIST OF FIGURES

Figure	Page
1. Study site locations in Montana. BBar ranch is located in Sweet Grass County and Decker/Bales ranch is in Powder River County.	34
2. Diagram outlining the general procedure for model development followed in Chapter 3.	35
3. Semivariograms of residuals from regression models presented in Table 1: (A) BBar multiple regression model, (B) Decker/Bales multiple regression model.	44
4. Predicted versus observed water content plots for: (A) BBar multiple regression model (Table 1), and (B) Decker/Bales multiple regression model (Table 1). Solid line represents least squares regression of independent validation water content as a function of predicted water content for specific model. Dashed line represents 1:1 line ($y = 0 + 1x$).	46
5. Decker/Bales regression tree constructed with full calibration sample size ($n = 50$).	50
6. BBar regression tree constructed with full calibration sample size ($n = 41$).	51
7. Predicted versus observed water content graphs for: (A) BBar full calibration sample regression tree model, and (B) Decker/Bales full calibration sample regression tree model. Solid line represents least squares regression for independent validation water content as a function of predicted water content. Dashed line represents 1:1 line ($y = 0 + 1x$).	52
8. Example mass water content maps constructed with: (A) the Decker/Bales (D/B) multiple regression model developed with 20 calibration samples, (B) the D/B multiple regression model developed with the full calibration sample size for the Decker/Bales ranch, (C) the Decker/Bales regression tree model developed with 30 samples, (D) the Decker/Bales regression tree model developed with the full calibration sample size, (E) the BBar multiple regression model developed with 20 samples, and (F) the BBar multiple regression model developed with the full calibration sample size.	54

LIST OF FIGURES - CONTINUED

Figure	Page
9. Study site locations in Montana. BBar ranch is located in Sweet Grass county and Decker/Bales ranch is in Powder River county.	73
10. Diagram outlining the general procedure, from development of the water content response variable to testing of the study hypothesis, followed in Chapter 4.	74
11. Predicted versus observed water content graphs for: (A) BBar soil survey variable regression model, (B) BBar lab and field characterization regression model, (C) Decker/Bales soil survey variable regression model, (D) Decker/Bales lab and field characterization regression model. Solid line represents least squares regression of independent validation water content as a function of predicted water content for specific model. Dashed line represents 1:1 line ($y = 0 + 1x$).	91
12. Predicted versus observed water content graphs for: (A) BBar soil survey variable regression tree model, (B) BBar lab and field characterization regression tree model, (C) Decker/Bales soil survey variable regression tree model, (D) Decker/Bales lab and field characterization regression tree model. Solid line represents simple linear regression of observed water content as a function of predicted water content for specific model. Dashed line represents $y = 0 + x$	94
13. BBar lab characterization regression tree. Terminal node values are mass water content (g_{water}/g_{soil}).	95
14. BBar soil survey regression tree. Terminal node values are mass water content (g_{water}/g_{soil}).	95
15. Decker/Bales lab characterization regression tree. Terminal node values are mass water content (g_{water}/g_{soil}).	96
16. Decker/Bales soil survey regression tree. Terminal node values are mass water content (g_{water}/g_{soil}).	96

ABSTRACT

I developed site-specific empirical models to predict spring soil water content for two Montana ranches. The models used publicly available Landsat TM 5, USGS DEM, and soil survey-derived data as predictor variables. The goal of the project was to test whether ranchers could collect a limited size soil water content data set, build site-specific regression models based on the data set, and construct soil water content maps based on the models. The response variable for models consisted of 100 and 82 average soil profile mass water content samples for each ranch, respectively. Half the samples were used for model calibration and half for model validation. Multiple regression models had calibration R^2 of 0.64 and 0.43 for each ranch, respectively. Validation showed that the multiple regression models predicted the validation data sets with average error (RMSD) within 0.04 mass water content and regression tree models predicted within 0.055 mass water content. The majority of the validation MSD for all models was accounted for by a lack of correlation between predicted and observed values along a 1:1 line. Models were then constructed with decreased sample sizes. Regression tree models and multiple regression models constructed with 20 to 30 samples predicted soil water content with similar, though still limited, accuracy and precision to full sample models. Site-specific field and lab soil characterization data developed with diffuse reflectance spectroscopy modeling was used to assess the suitability of the soil survey based predictor variables. The multiple regression models with the site-specific data predicted soil water content with average prediction errors (RMSD) of 0.035 and 0.036 mass water content for the two ranches, respectively. Soil survey model predictions were statistically significantly different than site-specific model predictions for one ranch but not the other. Especially dry conditions were a factor contributing to the difficulty in accurately modeling and predicting soil water content encountered at both study sites. Landsat imagery from the peak of the previous growing season, DEM-derived slope and aspect variables, and soil survey attribute data each showed promise as significant predictors of spring soil water content, particularly considering the dry conditions of the data collection period.

CHAPTER 1

INTRODUCTION

More than 700,000 pastured livestock operations exist in the U.S. (Kellog, 2002). These farms and ranches have combined yearly livestock sales exceeding \$17 billion (Kellog, 2002). Some states are more heavily invested in agriculture and livestock production than others. This is as much a function of a state's landscape as it is of the local economy.

Montana accounts for approximately 6% of the farm and ranch acreage in the United States, and livestock sales bring the state \$1.1 billion annually (Montana Agricultural Statistics Service, 2005). Much of the state's federally managed land, especially east of the Continental Divide, has operational grazing leases. Over 50% of Montana's non-federal lands are presently managed as rangeland (Natural Resource Inventory, 1992). More than 75% percent of non-federal lands are managed as rangeland in portions of south-central and most of southeastern Montana (Natural Resource Inventory, 1992). In short, much of Montana's livelihood and landscape are based on livestock and rangelands.

Montana ranchers make a difficult economic decision every spring. They set stocking rates according to the number of animal units they anticipate that their pastures can support in the coming growing season. Ranchers must predict the amount of forage their pastures will produce in order to set stocking rates (Holocek, 1988). These forage production estimates are largely based on a combination of guess work and expert

knowledge that might often be heavily influenced by the successes or failures of the previous growing season.

Annual forage production has been significantly correlated with two factors in non-irrigated, semi-arid rangelands like those in south-central and southeastern Montana. These are the amount of water stored in the soil preceding the growing season and the amount of precipitation that falls during the growing season (Neff and Wright, 1977, Rogler and Haas, 1947). It is not feasible to predict the amount of rain that will fall in an upcoming summer. It might be possible, however, to model and map the spatial distribution of spring, pre-growing season soil water content. Ranchers could use such estimates of pre-growing season soil water content to help estimate the upcoming season's forage production and to set stocking rates for their pastures.

The advent and growth of precision agriculture has created an interest in using geospatial tools to aid traditional agricultural practices. Though precision range management is not at the stage of application of precision farming, some ranchers are technologically savvy and use GPS, GIS, and remotely sensed imagery for ranch management and inventory purposes. The development of a spring soil water content modeling and mapping methodology that implemented these geospatial tools would both aid ranchers interested in using precision agricultural techniques for management tasks, such as setting stocking rates, and contribute to the advancement of precision agriculture in ranching culture.

There are several remote sensing and GIS-based land inventory products that are both publicly available and potentially useful in the spatially explicit estimation of pre-

growing season soil water content. Three publicly available data sources are digitized soil maps with associated attribute data, digital elevation models, and multispectral remotely sensed imagery. This project proposed to combine these three data sources with field samples to develop spatially explicit predictions of spring-time soil water contents on two ranches in southeastern and south-central Montana. The success of the project was evaluated based on whether soil water content models could be parameterized to a specific ranch based on field and GIS work that could be completed by a precision agriculture-minded livestock producer.

The main goal of this project was to create and test a spatially explicit model that combined GIS, remote sensing, and field measurements to estimate pre-growing season soil water contents for each ranch in the study. A secondary goal was to evaluate the suitability of using soil survey data in such a model. These goals were accomplished through the following objectives:

- 1) create a model for each ranch that predicts a spatially explicit set of pre-growing season soil water content measurements using the predictor variables of Landsat imagery from peak production during the previous growing season, DEM-derived terrain attributes of slope and aspect, and soil survey-derived soil attribute maps;
- 2) validate the model with the reserved portion of the data set and then test the model with decreased sizes of the data set;
- 3) evaluate the suitability of soil survey data for the models developed in the first objective by comparing soil survey-derived predictor variables to the same set of soil attributes derived from field and lab characterization.

The remainder of this thesis is organized in four chapters. Chapter 2 reviews the pertinent literature. Chapter 3 pertains to the research associated with Objectives 1 and 2. Chapter 4 pertains to the research associated with Objective 3. Chapter 5 summarizes and concludes the entire project.

References Cited

- Holocek, J.L. (1988). An approach for setting the stocking rate. *Rangelands*. 10:10-14.
- Kellog, R.L. (2002). Profile of farms with livestock in the United States: a statistical summary. <http://www.nrcs.usda.gov/technical/land/pubs/livestockfarm.html>. Natural Resources Conservation Service, USDA. Washington, D.C.
- Montana Agricultural Statistics Service. (2005). <http://www.nass.usda.gov/mt/>. National Agricultural Statistics Service, USDA. Helena, Montana.
- Natural Resource Inventory. (1992). Natural Resources Conservation Service, USDA. Washington, D.C.
- Neff, E.L., Wight, J.R. (1977). Overwinter soil water recharge and herbage production as influenced by contour furrowing on Eastern Montana rangelands. *Journal of Range Management*. 30: 193-195.
- Rogler, G., Haas, H.J. (1946). Range production as related to soil moisture and precipitation on the Northern Great Plains. *Journal of the American Society of Agronomy*. 378-389.

CHAPTER 2

LITERATURE REVIEW

Water is the most limiting factor in forage production in non-irrigated semi-arid rangelands (Rogler and Haas, 1947). The semi-arid rangelands in southeastern Montana are on the boundary between the Ustic and Aridic soil moisture regimes (Soil Survey Staff, 1971). The semi-arid rangelands of south-central Montana are more accurately characterized by the Ustic moisture regime than the Aridic moisture regime (Soil Survey Staff, 2004). The upper soil profile is often relatively dry for much of the year in the Ustic moisture regime, but the yearly peak of plant available water usually coincides with the early summer growing season (Soil Survey Division Staff, 1993). The Aridic moisture regime trends towards drier and more water limiting conditions throughout the year (Soil Survey Division Staff, 1993).

Range plant production has been significantly correlated with precipitation and soil moisture in the Northern Great Plains (Rogler and Haas, 1947; Neff and Wight, 1977; Cannon and Nielsen, 1984). Annual forage production has been highly correlated to spring (pre-growing season) plant available water in Montana's semi-arid rangelands (Neff and Wight, 1977). Forage production in Montana's non-irrigated semi-arid rangelands is a function of the amount of water stored in the soil prior to the growing season and the amount of rain that falls during the growing season.

The spatial distribution of soil water in such rainfall-driven systems should reflect the spatial distribution of the components of the basic hydrologic budget. The basic

hydrologic budget states that in non-irrigated systems: Inputs (precipitation) = Outputs (evapotranspiration) + Change in Storage (soil water content) (Fetter, 1996). The basic hydrologic budget is used as a context for modeling the spatial distribution of soil water in rangelands (Grayson et al., 1997; Western et al., 1999; Pachepsky et al., 2001; Salve and Allen-Diaz, 2001; Chamran et al., 2002) and for relating soil water conditions to range vegetation production and productivity (Rogler and Haas, 1947; Neff and Wight, 1977; Nouvellon et al., 2001). The ability to model and predict the spatial distribution of pre-growing season plant available water might be a useful management tool in the range systems of south-central and southeastern Montana, where much of the yearly precipitation falls during the beginning of the growing season. It might provide ranchers with the capacity to make predictions about forage production, make management decisions according to these predictions, and monitor the outcome of the predictions with each precipitation event. Setting yearly stocking rates is one key management decision for which reliable predictions of forage production are necessary (Holochek, 1988).

Soil Water Storage

Soil water content is related to the pore space in a given volume of soil. Pore space can be occupied by liquid or vapor. Soil water content can be represented in several ways, which include (Marshall et al., 1996):

1. volume of water / volume of total pore space
2. volume of water / (volume of air + volume of water)
3. volume of water / volume of total unit of soil
4. mass of water / mass of total unit of soil.

It is often most useful to think of water content in terms of volume water content (3) or mass water content (4). Conversion between the volume and mass units can be made using the bulk density of the unit of soil in question and the density of water. The porosity and bulk density of a given unit of soil are functions of the soil texture, which is the relative composition of sand, silt, and clay sized particles (Marshall et al., 1996). Porosity and bulk density are further influenced by coarse fragment content, organic matter content, and plant roots (Marshall et al., 1996). Soils are considered to develop from five factors: geologic parent material, climate, biology, topography, and time (Jenny, 1941). Porosity and bulk density are functions of soil development and result indirectly from the five factors. A sixth factor of soil development is management practices (Troeh et al., 1999). Management practices intended to increase soil water storage and water use efficiency are common, as are practices that negatively influence soil water storage by decreasing porosity and increasing bulk density (Troeh et al, 1999; Hatfield et al., 2001).

Equivalent depth of plant available water is a common and useful measurement of soil water content in range management (Soil Survey Division Staff, 1993). Plant available water is estimated as the difference between soil water content at field capacity and permanent wilting point (Marshall et al., 1996). Depending on measurement type, the water content often must be converted from a mass water content to a volume water content using the bulk density of the soil (Marshall et al., 1996). Equivalent depth of plant available water (inches or centimeters) is calculated by multiplying the volume

water content of plant available water by rooting depth or depth to root restrictive layer (Soil Survey Division Staff, 1993). This produces a linear value of plant available water.

Permanent wilting point refers to the water content at which plants can no longer remove water from the soil (Marshall et al., 1996). A certain amount of water will be retained by the soil depending on its texture and will not be available to plants regardless of how dry conditions become for plants. This amount of water unavailable to plants is loosely defined as permanent wilting point and varies according to soil particle size distribution, coarse fragment content, mineralogy, and organic matter content (Marshall et al., 1996).

On the wet end of the spectrum, there is a difference between saturation and the more temporally stable state of field capacity (Marshall et al., 1996). Field capacity is defined as the amount of water in a soil at field conditions following a saturation event after the water that is readily available to freely drain to underlying layers has done so (Marshall et al., 1996). Some give the amount of time necessary for this to occur as a specific value (e.g., 2 days), but it is more important to recognize the potential variability in this value based on soil development, soil characteristics (texture, porosity, drainage potential), and soil management (Marshall et al., 1996).

Available water capacity is the amount of plant available water a given unit of soil can hold at field capacity. It is calculated as the difference between water content at field capacity and water content at permanent wilting point. It can be expressed as a volume or a depth, as well as a mass ratio (Soil Survey Division Staff, 1993). Estimates of available water capacity for a given pedon of soil are generally adjusted for depth to root

limiting layer (Soil Survey Division Staff, 1993) and can be adjusted for rooting depth as well.

Evapotranspiration and Grazing

Evapotranspiration plays an important role in depleting stored soil water in water limiting environments. Evapotranspiration is the combined result of evaporation and plant transpiration. It has been shown that both peak evapotranspiration rates, and evapotranspiration rates in general, follow precipitation events in semi-arid rangelands (Frank, 2003). These peaks correspond to times when evapotranspiration is equal to or near potential evapotranspiration, after soil wetting precipitation events.

Evapotranspiration diverges more from potential evapotranspiration as the soil water supply decreases. More of evapotranspiration is accounted for by plant transpiration as soil water supply decreases, as well (Marshall et al., 1996).

Grazing has been shown to influence both evaporation and transpiration (Bremer, 2001; Frank, 2003). Grazing has the potential to increase evaporation of soil water by decreasing plant cover, allowing more solar radiation to the soil surface, and influencing soil temperatures and energy fluxes at the soil-air interface (Frank, 2003). Transpiration by an individual plant can decrease immediately following grazing due to reduction of leaf area, but the total seasonal transpiration by grazed versus ungrazed sites might be dependent on when grazing occurs during the growing season (Wraith et al. 1987; Bremer, 2001).

Grazing has been shown to decrease average growing season evapotranspiration by 6-8% (Bremer, 2001; Frank, 2003). This appears to be dependent on timing of grazing. Plants grazed in the spring can produce secondary growth that creates a peak of evapotranspiration later in the growing season in comparison to a comparable ungrazed grassland system (Bremer, 2001). Grazing should affect soil water storage if it affects evapotranspiration. Grazing has been shown to conserve soil water, and grazing has also been shown to deplete soil water (Bremer, 2001). The timing of grazing relative to precipitation, evaporative demand, and plant phenology is the most likely cause of the disparity in these conclusions (Bremer, 2001).

Remote Sensing of Evapotranspiration in Rangeland Systems

The use of multispectral satellite imagery has been investigated for sensing surface soil water content on bare agricultural fields (Moran et al., 2002; Peng et al., 2003). Much of the work involving the use of remote sensing tools in soil water content monitoring, however, has focused on the thermal and microwave range of the spectrum (Schmugge, 2002; Hunt et al., 2003). A review of the applications of remote sensing to rangeland management stated that the use of thermal and radar data for interpreting rangeland hydrological functioning is in the experimental versus operational stage (Hunt et al., 2003). Radar data is only useful for estimating water content of the surface soil (Schmugge, 2002), for example, the upper 10 cm for one rangeland study (Starks, 2002). The relative abilities of radar and Landsat thematic mapper (TM) imagery were compared for sensing surface soil moisture content in agricultural fields (Moran et al., 2002). Radar

data was suggested to be most useful when combined with the optical TM imagery. The combination of radar data with Normalized Difference Vegetation Index (NDVI) derived from Landsat TM imagery was used to estimate the surface soil water content in semi-arid rangelands (Wang et al., 2004). Surface roughness and vegetation cover posed major impediments to accurate soil water content estimates. The soil water content estimates were not field validated, but had limited correlation with precipitation records.

Soil color, texture, and water content have been shown to affect the spectral response of the land surface in more arid and less densely vegetated environments (Escadafal and Couralt, 1989; Mathieu et al., 1998). The short wave infrared portion of the spectrum is suggested to be better suited than the visible and near infrared portions of the spectrum to sensing soil water content (Lobell and Asner, 2002), particularly lower water contents (Weidong et al., 2002). The AVIRIS hyperspectral sensor was used to detect differences in spectral response based on soil surface organic matter and iron content (Palacios-Orueta and Ustin, 1998).

Apart from the direct measurement of soil characteristics, there is the potential for using remote sensing to characterize and estimate variables important to the soil water hydrologic budget. One example is evapotranspiration. The SEBAL model is a mechanistic approach to modeling land surface heat fluxes and relative evapotranspiration rates (Bastiaansen et al., 1998). It involves the use of satellite imagery and ground calibration of short wave atmospheric transmittance, surface temperature, and vegetation height measurements (Bastiaansen et al., 1998). The model's approach has been used most successfully to estimate evaporation rates from surface water bodies and

has been applied to estimating evapotranspiration rates from agricultural systems (Ayenew, 2003; Bastiaansen and Ali, 2003; Bastiaansen and Chandrapala, 2003; Hemakumara et al., 2003). This indicates some potential for the use of soil and vegetation reflectance in satellite imagery to account for evapotranspiration in empirical models as well.

From an empirical hydrologic budget modeling standpoint, there might be potential in using the capacity of satellite imagery to estimate relative plant productivity. Relative productivity might be expected to be a function not only of precipitation inputs in water limiting rangelands, but also of the outputs of water use and evapotranspiration – all important components to the spatial distribution of soil water (Hatfield et al., 2002; Obrist et al., 2003). Estimation of peak biomass production in rangelands was one of the first applications of research in the use of satellite imagery (Hunt et al., 2003). The use of such estimates has been more applicable in site-specific studies than in attempts to develop robust models for differing landscapes and scenes of imagery (Hunt et al., 2003). Some site-specific rangeland examples have used: (1) NDVI from AVHRR imagery to predict approximately 60% of the variation in live versus dead biomass (Thoma et al., 2002), (2) Landsat ETM+ and MODIS imagery to derive leaf area index with approximately 80% accuracy for ETM+ and less accuracy for the lower spatial resolution of MODIS (Cohen et al., 2003), and (3) Landsat imagery and a mechanistic ecosystem model to make 10 years of daily biomass, LAI, and soil water content predictions (Nouvellon et al., 2001). A more robust approach for assessing rangeland biomass with

Landsat imagery has been suggested to involve the use of bandwise regression versus vegetation indices like NDVI and band ratios (Maynard, 2004).

Multispectral satellite imagery might be used to account for the empirical relationship between evapotranspiration and the spatial distribution of soil water.

Landsat imagery has been used to estimate accurately leaf area (Qi et al., 2000; Wylie et al., 2002), which in turn should be highly correlated to evapotranspiration (Obrist et al., 2003). This is different than many recent studies that directly measure these parameters and develop mechanistic models for their prediction with satellite imagery (Bastiaansen et al., 1998; Boegh et al., 2002; Ayenew, 2003; Bastiaansen and Ali, 2003; Bastiaansen and Chandrapala, 2003; Hemakumara et al., 2003).

Terrain and Soil Controls on Soil Water Content

Soil landscape models are often used by field soil scientists to describe soil distribution and relative soil development at medium to coarse scales. These models are built upon qualitative observational or expert data. The U.S. National Cooperative Soil Survey system relies on such an approach to produce soil maps (Hudson, 1992). In their approach, sample points are located across the range of landform positions on a landscape. Point observation profile descriptions are then interpolated to a continuum of soil types. The continuum is built upon traditional soil and landscape concepts such as the five factors of soil formation, the catena, and three dimensional landform-soil models.

Topography is considered one of five factors most important to soil formation (Jenny, 1941). A process-based conceptual framework considers topography to be one of

several factors that dictates the specific combination and relative balance of biological, chemical, and physical processes that drive soil development at a specific location (Simonson, 1959). The catena was originally defined as a “topographic complex of soils” (Milne, 1936). In the catena concept, soil development is associated with hillslope position, and specifically hillslope-related exposure and translocation of soil parent materials (Milne, 1936; Bushnell, 1942). The hillslope soil landscape model was later extended to three-dimensional hillslope morphology at a drainage basin scale (Ruhe and Walker, 1968; Walker and Ruhe, 1968). Each of these traditional concepts is rooted in the relationship between topographic landform position (or the terrain and biophysical components that comprise landform position) and soil development (Wysocki et al., 2000). Subsequently, much of the current understanding of the distribution and formation of soil at a landscape scale is based on this topography-soil relationship as well.

The ability to analyze and model the influence of topography on soil characteristics and hydrology with digital elevation models has great potential. A distinction is made in DEM-based studies between primary topographic variables and compound topographic indices (Moore et al., 1991). Both are important to the analysis of relationships of soil and hydrology with topography. Slope, aspect, upslope catchment area, flow path length, profile curvature, and plan curvature are primary topographic variables important to hydrologic studies (Moore et al., 1991). Compound topographic indices reflect relationships between primary topographic variables.

Wetness index and radiation indices are examples of hydrologically important compound topographic indices (Moore et al., 1991). Wetness index is calculated as: $Wi = \ln(\text{upslope catchment area} / \tan(\text{slope}))$ (Western et al., 2002). Slope and upslope catchment area are the primary topographic variables important to wetness index.

Wetness index has been used to predict spatial distribution of relative amounts of soil water (Western et al., 2002). The radiation indices used in hydrologic studies are more involved calculations than wetness index, but include the topographic variables of slope and aspect (Moore et al., 1991). Radiation indices are used to account for soil water evapotranspiration. Slope and aspect can also be used as primary variables to account for relative amounts of evapotranspiration across the landscape (Western et al., 2002).

The influence of primary topographic variables and compound topographic indices on measured and modeled soil water distribution has been studied (Landon, 1995; Nyberg, 1996; Grayson et al., 1997; Western et al. 1999; Pachepsky et al., 2001; Chamran et al., 2002; Kozar, 2002). Wetness index and upslope catchment area were found to be significantly and substantially correlated to soil water content in a covered catchment area with controlled precipitation inputs (Nyberg, 1996). Evapotranspiration was not accounted for, and much of the unexplained variability in soil water content was concluded to be associated with losses due to evapotranspiration. A difference between relative dry soil states and wet soil states was cited as a factor in whether the wetness index explained more of the variability in soil water distribution than potential solar radiation index and its component primary variables (Grayson et al., 1997). The drier states were characterized by vertical movement of soil water, and the wetter states were

characterized by lateral subsurface and surface flow of water for areas with long-term seasonal differences in precipitation and evapotranspiration. Furthermore, the transition between dry and wet states, though difficult to characterize in terms of topography, were very small in comparison to the more stable dry and wet states. The temporal stability similarly was found to be greatest with increasing aridity of the soil water state in an agricultural system in Spain (Martinez-Fernandez and Ceballos, 2003). Periods of recharge showed the least temporal stability. Topography might directly influence soil water distribution in semi-arid environments because soil profiles might rarely be wetted to the point that lateral flow becomes possible (Landon, 1995; Kozar, 2002).

Upslope catchment area and wetness index correlated best to soil water distribution during wet conditions and potential radiation index correlated best to soil water distribution during dry conditions in a comparison of terrain indices in a temperate region of Australia (Western et al., 1999). The topographic indices explained considerably more of the variability in soil water distribution in wet conditions compared to dry conditions where hydrologic connectivity of soils across the landscape would have been less substantial. Subsurface flow, as related to slope, highly influenced stored soil water content distribution on a semi-arid hill slope for a catena of soils in California (Chamran et al., 2002). Three times the average rainfall was noted to have fallen the previous year. Soil water retention at the drier end of the moisture continuum was found to be significantly related to the terrain subdivisions of similar slope, slope shape (plan and profile curvature), and slope length (Chamran et al., 2002).

Landform classification is one method for representing hydrologically and pedologically meaningful terrain subdivisions. Digital landform classification methods based upon a heuristic approach to digital terrain modeling have been developed over the past three decades (Pennock et al., 1987; Burrough et al., 2000; Macmillan et al., 2000). These classification systems provide a methodologically explicit and reproducible tool for delineating landform components fundamental to the spatial relationships of soil formation and development. An issue of importance is whether a specific classification procedure uses the parameters of calculated flow paths and upslope contributing area. Single direction algorithms produce considerably different results than multidirection algorithms, and algorithms that allow for dispersion of flow create different results than those that do not (Tarboton, 1997). A flow routing algorithm that allows for dispersion could accurately reflect water movement in saturated conditions yet misrepresent the lack of hydrologic connectivity in dry conditions.

The complexity of the approach, if any, used for dealing with DEM artifacts is another important issue in landform classification and terrain modeling in general. Uncorrected DEMs can contain numerous depressions that confound flow routing applications (Macmillan, in review). The approaches used in dealing with these depressions can range from preprocessing the DEM with the assumption that all depressions are artifacts to complex systems of recognizing non-artifact depressions and preserving their characteristics (Macmillan, 2003; Macmillan, in review). The importance of this issue is also impacted by landscape characteristics. Many studies implementing heuristic landform classifications based on terrain modeling have focused

on hummocky landscapes of the continental, glaciated plains of North America (Pennock et al., 1987; Burrough et al., 2000; Macmillan et al., 2000). These environments are characterized by closed depressional drainage systems. Nonglaciated environments characterized by fluvially dissected, sedimentary bedrock controlled hills, such as in southeastern Montana, might not be as impacted by the ambiguity between actual and artifact DEM depressions.

Soil water distribution should be more closely related to hydrologically important soil characteristics like texture than to topographic variables, particularly in environments where potential evapotranspiration generally exceeds precipitation (Grayson et al., 1997; Western et al., 1999; Ridofi et al., 2003). Soil characteristics including texture, organic matter content, and A horizon thickness have been correlated to and predicted with DEM-derived terrain variables (Moore et al., 1993). Soil texture was significantly related to topographic variables when the relationship between texture and topographic variables was studied for soil modeling purposes (Pachepsky et al., 2001). Available water capacity was proposed to be more highly correlated to topographic variables than was texture, because the strongest relationship between soil water distribution and topographic variables was in the water content range of field capacity (Pachepsky et al., 2001). A wide range of soil variables made available through survey data has been used to predict soil water retention and saturated hydraulic conductivity with pedotransfer functions (Wosten et al., 2001). Such pedotransfer functions have increasingly focused on including landscape position as a grouping variable in model development (Wosten et al., 2001). Region-partitioning methods intended to group locations of similar soil and

landscape position resulted in limited correlations between soil water content and topographic regions in semi-arid Montana farm fields (Kozar, 2002). The addition of site-specific soils data to soil water management zone predictive models that use topographic and soil survey data was recommended for future research in semi-arid environments (Kozar, 2002). These findings and recommendations all emphasize some potential for using digital soil map and attribute data or lab characterized soil data in conjunction with terrain variables to predict the distribution of soil water.

References Cited

- Ayenew, T. (2003). Evapotranspiration estimation using thematic mapper spectral satellite data in the Ethiopian rift and adjacent highlands. *Journal of Hydrology*. 279:83-93.
- Bastiaanssen, W.G.M., Ali, S. (2003). A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan. *Agriculture, Ecosystems and Environment*. 94:321-340.
- Bastiaanssen, W.G.M., Chandrapala, L. (2003). Water balance variability across Sri Lanka for assessing agricultural and environmental water use. *Agricultural Water Management*. 58:171-192.
- Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., Holtslag, A.A.M. (1998). A remote sensing surface energy balance algorithm for land (SEBAL) 1. Formulation. *Journal of Hydrology*. 212-213:198-212.
- Boegh, E., Soegaard, H., Thomsen, A. (2002). Evaluating evapotranspiration rates and surface conditions using Landsat TM to estimate atmospheric resistance and surface resistance. *Remote Sensing of Environment*. 79:329-343.
- Bremer, D.J., Auen, L.A., Ham, J.M., Owensby, C.E. (2001). Evapotranspiration in a prairie ecosystem: effects of grazing by cattle. *Agronomy Journal*. 93:338-348.
- Burrough, P.A., van Gaans, P.F.M., Macmillan, R.A. (2000). High-resolution landform classification using fuzzy k-means. *Fuzzy Sets and Systems*. 113: 37-52.
- Bushnell, T.M. (1942). Some aspects of the soil catena concept. *Soil Science Society of America Proceedings*. 7:466-476.
- Cannon, M.E., Nielsen, G.A. (1984). Estimating production of range vegetation from easily measured soil characteristics. *Soil Science Society of America Journal*. 48:1393-1397.
- Chamran, F., Gessler, P.E., Chadwick, O.A. (2002). Spatially explicit treatment of soil-water dynamics along a semi-arid catena. *Soil Science Society of America Journal*. 66:1571-1583.
- Cohen, W.B., Maiersperger, T.K., Yang, Z., Gower, S.T., Turner, D.P., Ritts, W.D., Berterretche, M., Running, S.W. (2003). Comparisons of land cover and LAI estimates derived from ETM+ and MODIS for four sites in North America: a quality

- assessment of 2000/2001 provisional MODIS products. *Remote Sensing of Environment*. 88:233-255.
- Escadafal, R., Courault, M.G.D. (1989). Munsell soil color and soil reflectance in the visible spectral bands of Landsat MSS and TM data. *Remote Sensing of Environment*. 27:37-46.
- Fetter, C.W. (1994). *Applied Hydrogeology*. Prentice-Hall: Upper Saddle River, New Jersey.
- Frank, A.B. (2003). Evapotranspiration from northern semi-arid grasslands. *Agronomy Journal*. 95:1504-1509.
- Grayson, R.B., Western, A.W., Chiew, F.H.S. (1997). Preferred states in spatial soil moisture: local and nonlocal controls. *Water Resources Research*. 33: 2897-2908.
- Hatfield, J.L., Sauer, T.J., Prueger, J.H. (2001). Managing soils to achieve greater water use efficiency: a review. *Agronomy Journal*. 93:271-280.
- Hemakumara, H.M., Chandrapala, L., Moene, A.F. (2003). Evapotranspiration fluxes over mixed vegetation areas measured from large aperture scintillometer. *Agricultural Water Management*. 58:109-122.
- Holocek, J.L. (1988). An approach for setting the stocking rate. *Rangelands*. 10:10-14.
- Hudson, Berman. (1992). The soil survey as a paradigm based science. *Soil Science Society of America Journal*. 56: 836-841.
- Hunt, E.R., Everitt, J.H., Ritchie, J.C., Moran, M.S., Booth, D.T., Anderson, G.L., Clark, P.E., Seyfried, M.S. (2003). Applications and research using remote sensing for rangeland management. *Photogrammetric Engineering and Remote Sensing*. 69:675-693.
- Jenny, H. (1941). *Factors of soil formation: a system of quantitative pedology*. (originally published by McGraw-Hill; Mineola, N.Y.: Dover).
- Kozar, B.J. (2002). Predicting soil water distribution using topographic models within four Montana farm fields. M.S. thesis. Montana State University: Bozeman, Montana.
- Landon, M.A. (1995). Soil and terrain attributes for evaluation of leaching in a Montana farm field. M.S. thesis. Montana State University: Bozeman, Montana.

- Lobell, D.B., Asner, G.P. (2002). Moisture effects on soil reflectance. *Soil Science Society of America Journal*. 66:722-727.
- Macmillan, R.A., Pettapiece, W.W., Nolan, S.C., Goddard, T.W. (2000). A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. *Fuzzy Set and Systems*. 113: 81-109.
- Macmillan, R.A. (2003). *LandmapR Software toolkit- C++ Version: Users manual*. Landmapper Environmental Solutions Inc., Edmonton, AB. 110 pp.
- Macmillan, R.A. (in review). Extending current models of surface water flow to include explicit recognition of depressions and of connectivity between depressions.
- Marshall, T.J., Holmes, J.W., Rose, C.W. (1996). *Soil Physics*. Press Syndicate of the University of Cambridge: New York, New York.
- Martinez-Fernandez, J., Ceballos, A. (2003). Temporal stability of soil moisture in a large-field experiment in Spain. *Soil Science Society of America Journal*. 67:1647-1656.
- Mathieu, R., Pouget, M., Cervelle, B. Escadafal, R. (1998). Relationships between satellite-based radiometric indices simulated using laboratory reflectance data and typic soil color of an aridic environment. *Remote Sensing of Environment*. 66:17-28.
- Maynard, C.L. (2004). *Characterizing rangeland using multispectral remotely sensed data and multi-scale ecological units*. PhD. dissertation. Montana State University: Bozeman, Montana.
- Milne, G. (1936). Normal erosion as a factor in soil profile development. *Nature*. 138:548-549.
- Moran, M.S., Hymer, D.C., Qi, J., Kerr, Y. (2002). Comparison of ERS-2 SAR and Landsat TM imagery for monitoring agricultural crop and soil conditions. *Remote Sensing of Environment*. 79:243-252.
- Moore, I.D., Grayson, R.B., Ladson, A.R. (1991). Digital terrain modeling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes*. 5:3-30.
- Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A. (1993). Soil attribute prediction using terrain analysis. *Soil Science Society of America*. 57:443-452.

- Neff, E.L., Wight, J.R. (1977). Overwinter soil water recharge and herbage production as influenced by contour furrowing on Eastern Montana rangelands. *Journal of Range Management*. 30:193-195.
- Nouvellon, Y., Moran, M.S., Lo Seen, D., Bryant, R., Rambal, S., Ni, W., Beque, A., Chehbouni, A., Emmerich, W.E., Heilman, P., Qi, J. (2001). Coupling a grassland ecosystem model with Landsat imagery for a 10-year simulation of carbon and water budgets. *Remote Sensing of Environment*. 78:131-149.
- Nyberg, L. (1994). Spatial variability of soil water content in the covered catchment at Gardsjon, Sweden. *Hydrological Processes*. 10:89-103.
- Obrist, D., Verburg, P.S., Young, M.H., Coleman, J.S., Schorran, D.E., Arnone, J.A. (2003). Quantifying the effects of phenology on ecosystem evapotranspiration in planted grassland mesocosms using EcoCELL technology. *Agricultural and Forest Meteorology*. 118:173-183.
- Pachepsky, Y.A., Timlin, D.J., Rawls, W.J. (2001). Soil water retention as related to topographic variables. *Soil Science Society of America Journal*. 65:1787-1795.
- Palacios-Oureta, A., Ustin, S.L. (1998). Remote sensing of soil properties in the Santa Monica mountains i. spectral analysis. *Remote Sensing of Environment*. 65:170-183.
- Peng, W., Wheeler, D.B., Bell, J.C., Krusemark, M.G. (2003). Delineating patterns of soil drainage class on bare soils using remote sensing analyses. *Remote Sensing of Environment*. 115:261-279.
- Pennock, D.J., Zerbath, B.J., De Jong, E. (1987). Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. *Geoderma*. 40: 297-315.
- Qi, J., Kerr, Y.H., Moran, M.S., Weltz, M., Huete, A.R., Sorooshian, S., Bryant, R. (2000). Leaf area index estimates using remotely sensed data and BRDF models in a semi-arid region. *Remote Sensing of Environment*. 73:18-30.
- Ridolfi, L., D'Odorico, P., Porporato, A., Rodriguez-Iturbe, I. (2003). Stochastic soil moisture dynamics along a hillslope. *Journal of Hydrology*. 272:264-275.
- Rogler, G., Haas, H.J. (1946). Range production as related to soil moisture and precipitation on the Northern Great Plains. *Journal of the American Society of Agronomy*. 378-389.
- Ruhe, R.V., Walker, P.H. (1968). Hillslope models and soil formation. I. Open systems. *Transactions of the 9th International Congress of Soil Science*. 4:551-560.

- Salve, R., Allen-Diaz, B. (2001). Variations in soil moisture content in rangeland catchment. *Journal of Range Management*. 54:44-51.
- Schmugge, T.J., Kustas, W.P., Ritchie, J.C., Jackson, T.J., Rango, A. (2002). Remote sensing in hydrology. *Advances in Water Resources*. 25:1367-1385.
- Simonsen, R.W. (1959). Outline of a generalized theory of soil genesis. *Soil Science Society of America Proceedings*. 23:152-156.
- Soil Survey Division Staff. (1993). *Soil Survey Manual*. United States Department of Agriculture, Natural Resources Conservation Service: Washington D.C.
- Soil Survey Staff. (2004). *Soil Survey (SSURGO) database for Sweet Grass County Area, Montana*. United States Department of Agriculture, Natural Resources Conservation Service: Fort Worth, Texas. <http://soildatamart.nrcs.usda.gov>.
- Soil Survey Staff. (1971). *Soil Survey Powder River Area Montana*. United States Department of Agriculture, Soil Conservation Service: Washington, D.C.
- Starks, P.J. Jackson, T.J. (2002). Estimating soil water content in tallgrass prairie using remote sensing. *Journal of Range Management*. 55:474-481.
- Tarboton, D.G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*. 33: 309-319.
- Thoma, D.P., Bailey, D.W., Long, D.S., Nielsen, G.A., Henry, M.P., Breneman, M.C., Montagne, C. (2002). Short-term monitoring of rangeland forage conditions with AVHRR imagery. *Journal of Range Management*. 55: 383-389.
- Troeh, F.R., Hobbs, J.A., Donahue, R.L (1999). *Soil and Water Conservation – Productivity and Environmental Protection*. Prentice Hall: Upper Saddle River, New Jersey.
- Walker, P.H., Ruhe, R.V. (1968). Hillslope models and soil formation. II. Closed systems. *Transactions of the 9th International Congress of Soil Science*. 4:561-568.
- Wang, C., Qi, J., Moran, S., Marsett, R. (2004). Soil moisture estimation in a semi-arid rangeland using ERS-2 and TM imagery. *Remote Sensing of Environment*. 90: 178-189.
- Weidong, L., Baret, F., Xingfa, G., Quingxi, T., Lanfen, Z., Bing, Z. (2002). Relating soil surface moisture to reflectance. *Remote Sensing of Environment*. 81: 238-246.

- Western, A.W., Grayson, R.B., Bloschl, G. (2002). Scaling of soil moisture: a hydrologic perspective. *Annual Review of Earth and Planetary Sciences*. 30:149-180.
- Western, A.W., Grayson, R.B., Bloschl, G., Willgoose, G.R., McMahon, T.A. (1999). Observed spatial organization of soil moisture and its relation to terrain indices. *Water Resources Research*. 35: 797-810.
- Wosten, J.H., Pachepsky, Y.A., Rawls, W.J. (2001). Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*. 251:123-150.
- Wraith, J.M., Johnson, D.A., Hanks, R.J., Sisson, D.V. (1987). Soil and plant water relations in a crested wheatgrass pasture: response to spring grazing by cattle. *Oecologia*. 73: 573-578.
- Wylie, B.K., Meyer, D.J., Tieszen, L.L., Mannel, S. (2002). Satellite mapping of surface biophysical parameters at the biome scale over the North American grasslands a case study. *Remote Sensing of Environment*. 79:266-278.
- Wysocki, D.A., Schoenberger, P.J., LaGarry, H.E. (2000). Geomorphology of soil landscapes. In: Sumner, M.E. (ed.) *Handbook of Soil Science*. CRC Press Inc., Boca Raton, FL.

CHAPTER 3

AN EMPIRICAL APPROACH TO MODELING AND MAPPING SPRING SOIL
WATER DISTRIBUTION AT A RANCH SCALEIntroduction

Ranchers across the Northern Great Plains of North America make difficult economic decisions each spring. They predict the number of animal units their pastures will support during the coming growing season. These predictions are based on expert knowledge and best guesses as to the amount of forage their pastures will produce.

Annual forage production has been significantly correlated with two factors in the semi-arid rangelands of the Northern Great Plains. These are the amount of water stored in the soil prior to the growing season and the amount of precipitation that falls during the growing season (Rogler and Haas, 1947; Neff and Wight, 1977). It is difficult to predict the amount of rain that will fall in a growing season. It might be possible, however, to model and map the spatial distribution of spring, pre-growing season soil water content with field data, GIS, and remote sensing tools. Ranchers could use such estimates of pre-growing season soil water status to help predict the upcoming season's forage production and set stocking rates for their pastures.

Some ranchers have become technologically savvy through the advent and growth of precision agriculture. They use GPS, GIS, and remotely sensed imagery for ranch management and inventory purposes. This group of ranchers might be interested in

models developed to run based on both publicly available geospatial data and easily collected, small sized field data sets. There are several remote sensing and GIS-based land inventory products that are both publicly available and potentially useful in the spatially explicit estimation of pre-growing season soil water content. Three publicly available data sources are Landsat imagery made available to producers through the Digital Northern Great Plains (<http://digital-ngp.aero.und.edu/>), 30-m resolution DEMs available from the USGS (<http://gisdata.usgs.net/website/Seamless>), and digitized soil surveys with associated attribute data available from the National Cooperative Soil Survey (<http://soildatamart.nrcs.usda.gov>).

Remote Sensing Approaches to Modeling Soil Water Content

Much of the work involving the use of remote sensing tools in soil water content monitoring has focused on the thermal and microwave range of the spectrum (Hunt et al., 2003). The use of thermal and radar data for assessing rangeland water resources is in the experimental versus operational stage according to a review of the applications of remote sensing to rangeland management (Hunt et al., 2003). Radar data is only useful for estimating water content of the surface soil, for example, the upper 10 cm for one rangeland study (Starks, 2002). The combination of radar data with Normalized Difference Vegetation Index (NDVI) derived from Landsat thematic mapper (TM) imagery was used to estimate the surface soil water content in semi-arid rangelands (Wang et al., 2004). Surface roughness and vegetation cover posed major impediments to accurate soil water content estimates and no field validation was completed.

There is potential for using satellite image remote sensing in the physical modeling of variables important to the soil water hydrologic budget, apart from the direct measurement of soil water content. Recent studies have used mechanistic approaches involving satellite imagery, ground measurements, and calibration to model land surface heat fluxes (Bastiaansen et al., 1998; Ayenew, 2003; Bastiaansen and Ali, 2003; Bastiaansen and Chandrapala, 2003; Hemakumara et al., 2003). This approach has been applied to estimating evapotranspiration rates from agricultural systems (Ayenew, 2003; Bastiaansen and Ali, 2003; Bastiaansen and Chandrapala, 2003; Hemakumara et al., 2003).

Multispectral satellite imagery also might be used to account for the empirical relationship between evapotranspiration and the spatial distribution of soil water. Landsat imagery has been accurately used to estimate leaf area (Qi et al., 2000), which in turn should be highly correlated to evapotranspiration (Obrist et al., 2003). This is potentially an alternative approach to recent studies that directly measure these parameters and develop mechanistic models for their prediction with satellite imagery (Bastiaansen et al., 1998; Ayenew, 2003; Bastiaansen and Ali, 2003; Bastiaansen and Chandrapala, 2003; Hemakumara et al., 2003). Empirical relationships developed between evapotranspiration and soil water content are site and date specific, but are considerably easier to develop than mechanistic approaches. Such empirical models avoid the radiometric correction and universal calibration issues that mechanistic models must confront.

Digital Terrain Modeling of Soil Water Content

Slope and aspect are two primary topographic variables important to hydrologic studies (Moore et al., 1991). These topographic variables can be used to account for relative amounts of evapotranspiration across a landscape, both as primary variables and in compound indices (Western et al., 2002). Terrain has been shown to be a better predictor of soil water content in wet versus dry conditions (Western et al., 1999; Kozar, 2002). Wetter periods in an environment can be characterized by vertical flow and lateral subsurface and surface flow of water, whereas drier periods are mostly expected to be characterized by vertical flow (Grayson et al., 1997). Relationships between water content and topographic factors can exist in dry conditions, but soil water fluxes are expected to be more difficult to model. Soil water retention in drier conditions has been significantly correlated to terrain subdivisions of similar slope gradient, aspect, length, and curvature (Chamran et al., 2002). Soil water content in semi-arid Montana environments, however, has been found to have limited correlation with terrain subdivisions and topographic indices (Landon, 1995; Kozar, 2002).

Soil Attributes Used as Predictors of Soil Water Content

Soil water distribution might be more closely related to hydrologically important soil characteristics, such as texture, than to topographic variables in semi-arid Montana rangelands. Topography is considered one of the important components of soil genesis (Jenny, 1941; Ruhe and Walker, 1968; Walker and Ruhe, 1968). Soil texture and water retention have been statistically significantly related to topographic variables in soil

modeling exercises (Pachepsky et al., 2001). This emphasizes the potential for using spatially explicit soil attribute data in conjunction with terrain variables to predict the distribution of soil water.

Soil surveys provide one source of spatially explicit soil attribute data. The U.S. National Cooperative Soil Survey relies on a qualitative, expert system approach to building soil surveys (Hudson, 1992). Point observation profile descriptions are interpolated to a continuum of soil types in this approach. The continuum is built upon the soil and landscape concepts of Jenny's factors of soil formation, Simonson's process model, Milne's catena, and Ruhe and Walker's three dimensional landform-soil models (Milne, 1936; Jenny, 1941; Simonson, 1959; Ruhe and Walker, 1968; Walker and Ruhe, 1968). Each of these concepts is rooted in the relationship between landform position (or the terrain and biophysical components that comprise landform position) and soil development.

Soil surveys are limited as sources of spatially explicit soil attribute data. Their point accuracy has been estimated at 45% - 65% for the 1:63,360 scale and 65% - 85% for the 1:25,000 scale (Burrough et al., 1971). Attribute data is often interpolated and/or extrapolated from a handful of lab characterized pedons for an entire survey area (Hudson, 1992). Soil surveys, however, provide the most geographically extensive geospatial soils database within the United States.

Objectives

The overall objective of this study was to develop a method for mapping rangeland soil water content that would be easily developed on a site- and date-specific

basis as a readily implemented management tool, rather than try to develop a universal mechanistic model that might be impractical for a rancher to implement. The specific goal was to develop and test an empirical approach to mapping soil water content using two Montana ranch study sites. The models were developed based on pre-growing season soil samples collected at each study site. The predictor variables used in the models were Landsat TM imagery from August of the previous growing season, DEM-derived slope and aspect layers, and soil survey-derived attribute layers of texture, soil depth, available water capacity, and plant available water capacity. The developed models were site and date specific and would have to be parameterized before being used at another location.

The models were independently validated with half of the soil water content data set from each ranch. New models were constructed with decreased soil water content sample sizes and then validated. Validation of the decreased sample size models was intended to determine whether the parameterization process could be quite simple and based on data collected during one day of work. These site and date specific models have an advantage over more global mechanistic approaches in that imagery issues of radiometric correction among scenes and dates are avoided and complex ground measurement and calibration procedures are avoided. The disadvantage is that a new empirical model based on a new set of soil samples is required to predict soil water content for a different site or date. The approach could be very practical for soil water content modeling, however, if the sampling and modeling are easily accomplished. The specific hypotheses tested at each study site were:

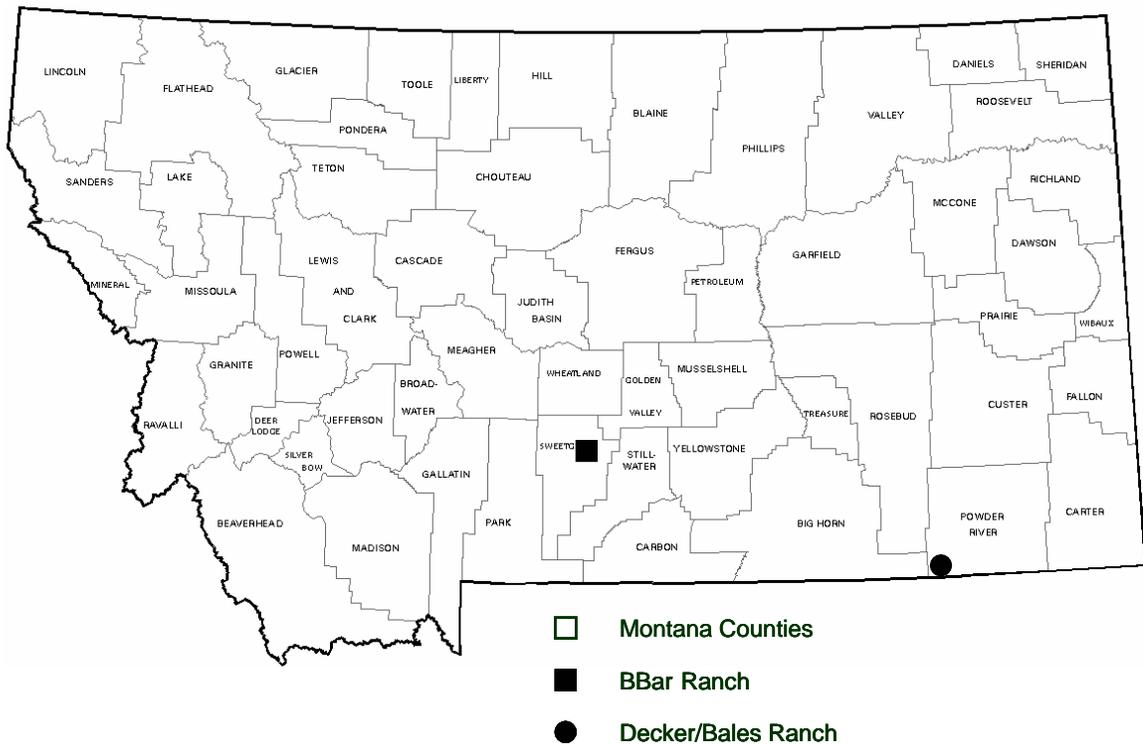
1. A model with the three data sources (imagery, topography, and soils) statistically significantly predicts the spatially explicit soil water content data set.
2. There is no significant difference in the independent validation data between measured and predicted soil water content.
3. There is no significant difference in each reduced size model validation data set between measured and predicted soil water content.

Study Sites

The Decker/Bales ranch and the BBar ranch served as the two study sites (Figure 1). The Decker/Bales ranch is approximately 100 km² and is located in southwestern Powder River County in southeastern Montana. It is mostly in the Tongue River watershed, but includes an area of the divide between the Tongue and Powder Rivers. The landscape is part of Montana's non-glaciated plains and is characterized by dissected sedimentary layers that form a low relief, fluviially incised landscape. Range vegetation consists of grassland communities of western wheatgrass (*Agropyron smithii* Rydb.), needle and thread (*Stipa comata* Trin. & Rupr.), and blue grama (*Bouteloua gracilis* Willd. ex Kunth), with a varying presence of big sagebrush (*Artemisia tridentata* Nutt.) (Montagne et al., 1982). Soils include loamy, calcareous Ustorthents formed in siltstones, clayey, calcareous Ustorthents formed in shales, fine to coarse-loamy Haplustalfs formed in slope alluvium, loamy-skeletal Haplustalfs formed in scoria beds, and fine Natrustalfs that are often associated with prairie dog communities (Veseth and Montagne, 1980; Montagne et al., 1982). The area receives approximately 30 cm of

mean annual precipitation, the soil temperature regime is on the boundary between Mesic and Frigid, and the soil moisture regime is on the boundary between Ustic and Aridic (Soil Survey Staff, 1971).

Figure 1: Study site locations in Montana. BBar ranch is located in Sweet Grass County and Decker/Bales ranch is in Powder River County.



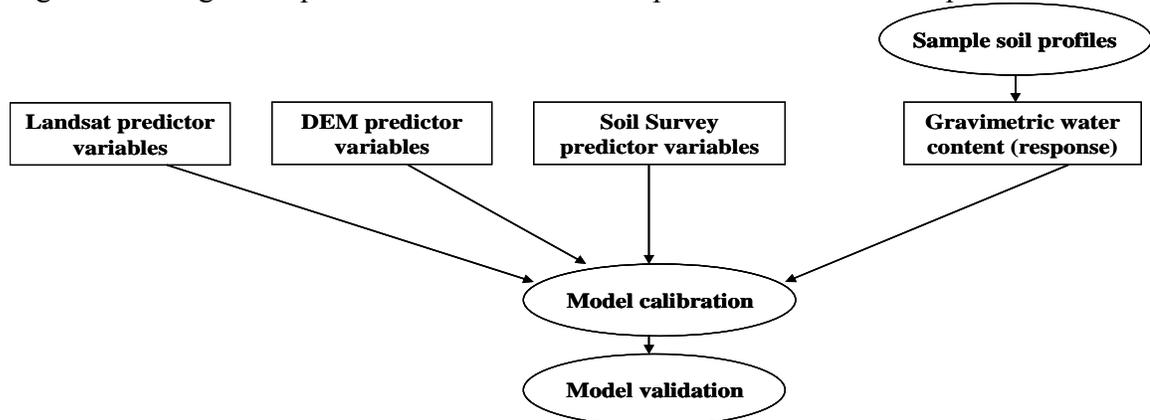
The BBar ranch is approximately 30 km² and is located in northern Sweet Grass County in south-central Montana. It lies in a valley at the Rocky Mountain front in the westernmost extent of Montana's non-glaciated plains. The landscape consists of rolling, sedimentary bedrock-controlled hills vegetated with grassland communities of western wheatgrass, little bluestem (*Andropogon scoparius* Michx.), needle and thread, and blue grama (Montagne et al., 1982). There are isolated surfaces of alluvium and outwash

associated with the Crazy Mountains. Soils that form in these parent materials range from fine Argiustolls on backslopes, footslopes and toeslopes, to loamy-skeletal Ustorthents on summit and shoulder positions, as well as fine Natrustalfs on toeslopes and valley floor positions, and fine and fine-loamy Torrifluvents in drainageways (Veseth and Montagne, 1980; Montagne et al., 1982). The area receives approximately 35 cm of mean annual precipitation, the soil temperature regime is Frigid, and the moisture regime is Ustic (Soil Survey Staff, 2004).

Methods

Methods implemented at both study sites included field soil sampling and gravimetric calculation of mass water content for soil samples (Figure 2). Models were developed to predict water content based on a set of predictor variables derived from Landsat, DEM, and soil survey data sources. Models were independently validated with a reserved soil water content data set. Models constructed with reduced calibration sample size to test the third hypothesis of the study were developed and validated in a similar fashion to models based on the full calibration sample size.

Figure 2: The general procedure for model development followed in Chapter 3.



Field Data Collection

A digitized representation of each ranch's boundary was considered the extent of each study area. A statistical power test (Pfaffenberger and Patterson, 1981) was performed on a small preliminary data set of depth of moist soil measurements from the Decker/Bales ranch (n=11) with a mean of 74 cm of moist soil and a standard deviation of 9.6 cm of moist soil. The test assumes normal distribution, constant variance in the data set, and an alpha level of 0.05. The power test showed that 41 points were necessary to be able to detect a significant difference of 7.6 cm of moist soil. At least 41 points for model development and 41 points for validation were targeted for each study area during sampling.

Sample locations were stratified based on the soil survey of the county in which each ranch resides. Soil survey maps were used to account for both the variability in soil type as well as the variability in slope, aspect, landform, and landform position at each ranch. The spatial data layer of each digitized survey was clipped to the extent of each ranch's digitized boundary. Random points were selected within each soil survey map unit, with at least one location for each named map unit.

Each sample point was identified with an X,Y location in UTM coordinates. Navigation to the points was accomplished with map and a GPS receiver with an accuracy of < 1m. The location of each sample point was logged in the GPS receiver as a waypoint and coordinates were recorded on the datasheet as well.

A hand auger was used to collect soil samples in eight 10 cm increments, from the soil surface to 100 cm depth at each sample location. It was assumed that variability in

soil characteristics would decrease with depth. Samples from 60 to 70 cm and 80 to 90 cm were not collected for logistical and efficiency purposes.

A total of 100 locations were sampled at the Decker/Bales ranch and 82 locations were sampled at the BBar ranch. Sampling was completed during the first week of May 2004 at the Decker/ Bales ranch and during the second week of May 2004 at the BBar ranch. The samples were collected in wax lined paper bags and were transported to the lab at the end of field work. Samples were weighed at field moist state, oven dried at 105° C, and then weighed again to gravimetrically calculate mass water content. Mass water content was averaged for the entire sampled profile at each sample location. Average soil profile mass water content served as the response variable for water content modeling at both study sites.

GIS Data Set Development

Satellite Imagery. A Landsat 5 TM scene was selected from the previous growing season for each study site. Scenes were selected by proximity to the peak of growing season biomass production and cloud free quality. A scene from 1 August 2003 was selected for the BBar site and was clipped to the extent of the digitized ranch boundary. A scene from 3 August 2003 was selected and clipped for the Decker/Bales ranch.

DEM-Derived Terrain Elements. A seamless, 30-m DEM was downloaded from the USGS Seamless Data Distribution Center (<http://gisdata.usgs.net/website/Seamless>) for each ranch. Percent slope and aspect layers were created in ARCGIS using the spatial analyst surface function. Aspect was transformed to the cosine of aspect from degrees

from north. Northerly aspects were positive values from 0 to 1 and southerly aspects were negative values from 0 to -1.

SSURGO-Derived Soil Attribute Maps. SSURGO digitized soil maps and associated attribute data were downloaded for each study site from the National Cooperative Soil Survey Soil Data Mart Distribution Center (<http://soildatamart.nrcs.usda.gov>). The soil survey maps were clipped to the extent of the respective ranch boundary. Soil maps were developed from the soil survey and attribute data in ARCGIS for percent clay content, soil depth to root restrictive layer, percent soil organic matter, bulk density, and water content at field capacity. Clay content, organic matter, bulk density, and water content at field capacity were calculated as weighted average values for the entire profile of the major component for each soil map unit. Soil depth to root restrictive layer was considered the depth to lithic or paralithic material or the maximum recorded soil depth for the major component of each soil map unit. Soil characteristic values of each major component were assigned to all pixels within representative soil map unit boundaries, and the resulting map was stored in raster format.

Maps of mass water content (θ_m) at permanent wilting point (PWP, -15 bar equivalent) were created for each ranch using the organic matter (OM) and clay layers in the ARCGIS raster calculator function and the following equation (Decker, 1972):

$$\text{PWP} = (0.29 * \% \text{Clay}) + (0.58 * \% \text{OM}) + 2.1$$

where PWP is the θ_m at -15 bar matric potential. The equation for water content at PWP was originally developed for use in Montana based on 930 samples ($R^2 = 0.84$) from 186 Montana pedons dominated by Mollisols and Entisols (Decker, 1972).

The permanent wilting point layer was subtracted from the field capacity (FC) layer and the result was multiplied by the respective bulk density (D_b) layer to create a thematic layer of plant available water holding capacity (AWC) by volume (θ_v) for each ranch. This was completed in the ARCGIS raster calculator and the following equation (Marshall et al., 1996):

$$AWC = (FC - PWP) * (D_b / D_w)$$

where FC is θ_m at -1/3 bar matric potential, PWP is θ_m at -15 bar matric potential, and D_b and D_w (density of water) are in units of g/cm^3 . This map was then adjusted to an equivalent depth (cm) of plant available water (PAW) map by multiplying it by the depth to root restrictive layer (cm) for each ranch with the following equation (Marshall et al., 1996):

$$PAW = AWC * Depth$$

where AWC is in θ_v units and depth is in cm. The final maps used in model development for the first objective were the percent clay content, percent organic matter, bulk density,

depth to root restrictive layer, available water capacity, and equivalent depth of plant available water layers.

Data Analysis

The average soil profile (100 cm) mass water content data set for each ranch was split randomly into two equally sized data sets prior to analysis ($n = 50$ Decker/Bales and $n = 41$ BBar). One data set for each ranch was used for model calibration and the other for independent validation. Multiple regression models were constructed in S-Plus 6.2 using stepwise regression (forward and backward stepping) in the first analysis step. To address the first hypothesis of the study, overall model significance and the combination of predictor variables used were considered. A model that included a predictor variable from each of the three data sources and was significant at the 0.05 confidence level failed to reject the first hypothesis that a model with predictor variables from each of the three data sources significantly explains variability in the soil water content data set.

The second hypothesis of this study tested whether the models developed in the first analysis step could be independently validated. Models that performed well in calibration were selected for each ranch and subsequently validated with the reserved data set. The ordinary least squares multiple regression model required the assumption that errors in the model were independent. Semivariograms were constructed to check for spatial autocorrelation in the residuals of the models selected for validation.

Independent model validation consisted of predicting water content for the reserved data set. A least squares regression of the validation water content as a function of predicted values was constructed, and a scatterplot of the relationship containing

points, regression line, and a 1:1 line (slope = 1, intercept = 0) was examined. A mean squared deviation (MSD) and root mean square deviation (RMSD) were calculated for the predicted versus observed values and the MSD was broken into components of standard bias (SB), non-unity (NU), and lack of correlation (LC) with the following equations (Gauch et al., 2003):

$$\text{MSD} = \sum_n (\text{Predicted } \theta_{m,n} - \text{Validation } \theta_{m,n})^2 / N$$

$$\text{SB} = (\mu(\text{Predicted } \theta_m) - \mu(\text{Validation } \theta_m))^2$$

$$\text{NU} = (1 - b)^2 * \sum_n (\text{Predicted } \theta_{m,n} - \mu(\text{Predicted } \theta_m))^2 / N$$

$$\text{LC} = (1 - r^2) * \sum_n (\text{Validation } \theta_{m,n} - \mu(\text{Validation } \theta_m))^2 / N$$

$$\text{MSD} = \text{SB} + \text{NU} + \text{LC}$$

$$\text{RMSD} = \sqrt{\text{MSD}}$$

$$\text{Bias} = \sqrt{\text{SB}}$$

where b refers to the slope of the least squares regression line of validation θ_m as a function of predicted θ_m , and r^2 is the square of the correlation. SB quantifies the proportion of the MSD related to the deviance of the least squares fit from a 1:1 relationship in the y direction (intercept). NU quantifies the proportion of the MSD related to the deviance of the least squares fit from a 1:1 relationship in the slope of the fitted line. LC quantifies the proportion of the MSD related to the scatter of the points in relation to the 1:1 line.

The second hypothesis of the study stated that there is no significant difference between predicted and observed validation soil water content and was tested for each validated model. A t-test and F-test can be used to test for the equality of means and

variances, respectively, of predicted and observed samples (Wosten et al., 2001, Feng et al., 2005). The Levene's test is an alternative to the F-test for equal variances and the Mann-Whitney test is an alternative to the paired t-test. Both tests are resistant to departures from normality, making them appropriate for the long-tailed distributions of the validation water content samples for both ranches. The Levene's test was used to test whether predicted and observed sample populations had significantly different variances (Feng et al., 2005). The Mann-Whitney test of the paired predicted and validation samples was used to test whether the mean of the differences between the samples was statistically significantly different than zero (Feng et al., 2005). For Levene's test, an insignificant p -value at a specified level of confidence showed no evidence that the populations had unequal variances. An insignificant Mann-Whitney p -value at a specified confidence level failed to reject the null hypothesis of the test that the mean of the differences between predicted and observed water content is zero. Model predictions had to produce insignificant results for both the variance and mean hypothesis tests to fail to reject the second hypothesis of the study that there is no significant difference between predicted and observed validation soil water content

The third hypothesis tested whether predicted and observed soil water content was significantly different for models constructed with reduced sample sizes. The reduced sample sizes were 40, 30, 20, and 10 samples for the Decker/Bales site and 30, 20, and 10 samples for the BBar site. Models were constructed in the same manner as with the full data sets. Soil water content was predicted for each reduced sample size model using the same validation data set for each site. Models were independently validated in the same

manner as the regression models constructed with the full sample size. Levene's test and the Mann-Whitney test were once again used to test for statistically significant differences between predicted and validation data sets.

All three hypotheses were also tested using regression tree analysis. Regression trees were built in S-Plus with all bands of imagery, aspect, slope, and all soil layers as potential predictors. Cross validation pruning was used to determine the number of nodes for constructed trees (Breiman et al., 1984). Regression tree models were validated as with multiple regression models.

Results

Multiple Regression Analysis

Multiple regression models built with the stepwise procedure were evaluated based on significance of predictor variables at the 0.05 level, adjusted R^2 values, overall model significance at the 0.05 level, and model interpretability. The most parsimonious models that explained a great amount of variability in the calibration data relative to other models constructed were selected for validation. Models selected for validation (Table 1) contained a variable from each of the three data sources and statistically significantly explained variability in the soil water content data set. The multiple regression analysis, therefore, failed to reject the first hypothesis of the study at both sites. The models, however, failed to explain a substantial amount of the variability in soil water content data sets, in particular for the Decker/Bales site.

Table 1: Average soil profile (100cm) mass water content (profile) models that performed best for calibration and were selected for validation for the BBar and Decker/Bales (D/B) ranches. Band numbers represent Landsat TM 5 reflectance bands. Slope is percent topographic slope. Aspect is cosine transformation of topographic aspect in degrees from North. Clay is entire profile weighted average percent clay of soil survey map unit major component. All models are significant at p -value < 0.05 and all variables are significant at p -value < 0.05 unless noted with $\#$. Adjusted R^2 values are presented.

Model ID	Model	R^2
BBar	$\text{profile} = 0.7840 - 0.0146(\text{band3})^\# - 0.0040(\text{band4}) - 0.0376(\text{slope}) + 0.0012(\text{clay}) + 0.0001(\text{band3}^2)^\# - 0.0004(\text{slope}*\text{band3}) + 0.0009(\text{slope}*\text{band4})$	0.64
D/B	$\text{profile} = 0.7080 - 0.0296(\text{band3}) + 0.0084(\text{band4}) + 0.0120(\text{aspect}) + 0.0007(\text{clay}) + 0.0004(\text{band3}^2) - 0.000003(\text{band4}*\text{band3}^2) - 0.0022(\text{band6})^\#$	0.43

The two models that performed best in calibration were validated (Table 2).

Semivariograms of the residuals from the two models were constructed (Figure 3). No pattern was evident in any of the semivariograms, thus the assumption that residuals were spatially independent was acceptable.

Figure 3: Semivariograms of residuals from regression models presented in Table 1: (A) BBar multiple regression model, (B) Decker/Bales multiple regression model.

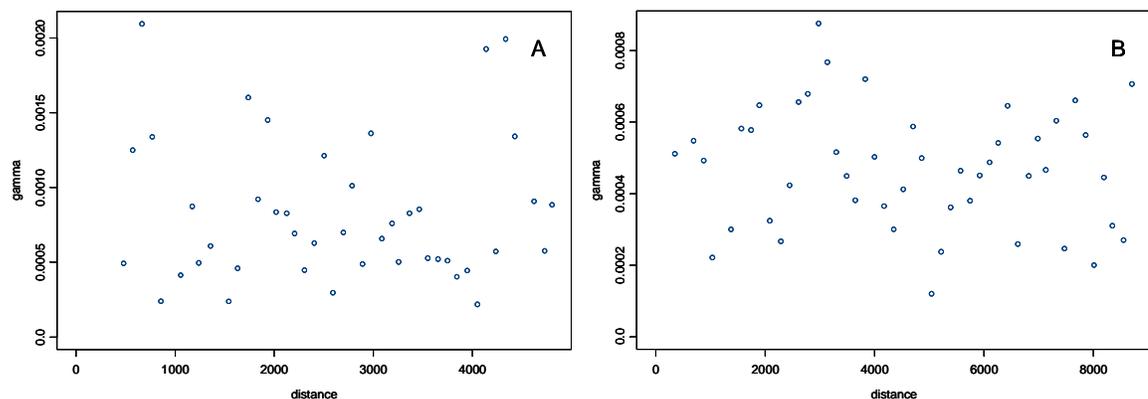


Table 2: Model validation results for models in Table 1. Models were independently validated with half the data set at each ranch (50 validation samples for Decker/Bales (D/B) and 41 for BBar ranch). Levene and Mann-Whitney statistics are p -values.

Model	RMSD	Bias	MSD	SB	NU	LC	r^2	Levene	Mann-Whitney
BBar	0.039	0.005	0.002	0.0000	0.0000	0.0015	0.54	0.81	0.35
D/B	0.040	0.006	0.002	0.0000	0.0005	0.0011	0.00	0.00	0.14

RMSD values calculated for validation (Table 2) of the two best models represent an unbiased estimate of the average error in predicted water content when compared to observed values in the independent validation data. The models for each site predicted mass water content within 0.04 (Table 2). Insignificant Levene test p -values at the 0.05 level failed to suggest that predicted and observed water content populations had unequal variances for the BBar model but not the Decker/Bales model (Table 2). Insignificant p -values at the 0.05 level failed to reject the null hypothesis of the Mann-Whitney paired sample test (Table 2) that the mean of the differences of the predicted and observed values is zero for both models. No statistically significant difference between predicted and observed water content in the validation data was found for the BBar model. The Decker/Bales model, however, rejected this hypothesis. The models from both ranches failed to explain the variability in soil water content at a suitable level of precision for practical purposes. This is illustrated by the scatterplots of predicted versus observed water content for the models from each ranch (Figure 4). The MSD components (Table 2) suggested that the major source of validation error in the models was in the lack of correlation between predicted and validation values in terms of a 1:1 relationship. This

was corroborated by low r^2 values that suggested the BBar model predictions explained a low proportion of variability in the validation data set, and that the Decker/Bales model predictions explained no variability in the respective validation data set.

Models constructed with decreased sample size (Tables 3 and 4) had different combinations of predictor variables than the models developed with the full calibration data set at each site. Reduced sample size models were validated with the full validation sample size for each ranch (Tables 5 and 6). Levene's test p -values suggested significant differences between the variances of predicted and observed sample populations for the 3 largest reduced sample models at the Decker/Bales ranch (Table 5). This rejected the third hypothesis of the study that there is no significant difference between observed and predicted water content in each reduced size calibration data set. Hypothesis test p -values failed to reject the third hypothesis of the study for the model constructed with 10 samples at the Decker/Bales ranch. The 10 sample model predictions, however, failed to explain a suitable amount of variability in the validation data set for this ranch.

Figure 4: Predicted versus observed water content plots for: (A) BBar multiple regression model (Table 1), and (B) Decker/Bales multiple regression model (Table 1). Solid line represents least squares regression of predicted water content versus validation water content for specific model. Dashed line represents 1:1 line ($y = 0 + 1x$).

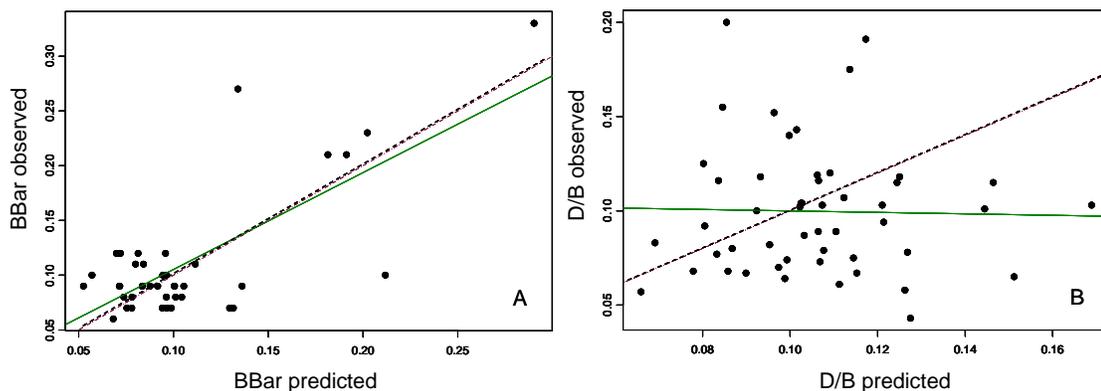


Table 3: Average soil profile (100cm) mass water content models constructed with decreased sample sizes at the Decker/Bales ranch. Band numbers represent Landsat TM 5 reflectance bands. Slope is percent topographic slope. Aspect is cosine transformation of topographic aspect in degrees from North. Clay is entire profile weighted average percent clay of soil survey map unit major component. OM is entire profile weighted average percent organic matter of soil survey map unit major component. PAW is entire profile weighted average plant available water (cm) of soil survey map unit major component. All models are significant at p -value < 0.05 and all variables are significant at p -value < 0.05 unless noted with $\#$. Adjusted R^2 values are presented.

Calibration n	Model	R^2
40	profile = 1.0112 - 0.0388(band 3) + 0.0088(band 4) - 0.0021(slope) + 0.0006(band3 ²) - 0.0027(band 6) - 0.000003(band 4* band3 ²)	0.52
30	profile = 0.2121 - 0.0026(band 4) + 0.0490(aspect) - 0.0007(paw) + 0.0230(om) [#] + 0.00002(band3 ²) - 0.0513(aspect*om)	0.41
20	profile = 0.0904 - 0.0348(band 3) + 0.0149(band 4) + 0.0006(band3 ²) - 0.000004(band 4*band3 ²)	0.61
10	profile = 1.3751 - 0.0436(band 3) + 0.0004(band3 ²) - 0.0017(clay)	0.85

Table 4: Average soil profile (100cm) mass water content models constructed with decreased sample sizes for the BBar ranch. Band numbers represent Landsat TM 5 reflectance bands. Slope is percent topographic slope. Clay is entire profile weighted average percent clay of soil survey map unit major component. All models are significant at p -value < 0.05 and all variables are significant at p -value < 0.05 . Adjusted R^2 values are presented.

Calibration n	Model	R^2
30	profile = 0.4101 - 0.0035(band 3) - 0.0024(band 4) - 0.0470(slope) + 0.0016(clay) + 0.0007(slope*band 4)	0.61
20	profile = 0.7798 - 0.0057(band 3) - 0.0053(band 4) - 0.0664(slope) + 0.0010(slope*band 4)	0.79
10	profile = 0.6261 - 0.0225(band 2) + 0.0057(band 7)	0.65

Table 5: Validation results for decreased sample size models constructed for the Decker/Bales ranch and presented in Table 3. Models were independently validated with the full validation data set (50 validation samples for Decker/Bales). Mann-Whitney statistic is p-value.

n	RMSD	Bias	MSD	SB	NU	LC	r ²	Levene	Mann-Whitney
40	0.040	0.006	0.002	0.0000	0.0004	0.0011	0.00	0.00	0.17
30	0.034	-0.002	0.001	0.0000	0.0001	0.0011	0.06	0.00	0.93
20	0.043	0.008	0.002	0.0001	0.0007	0.0011	0.01	0.00	0.11
10	0.062	0.008	0.004	0.0001	0.0026	0.0011	0.05	0.62	0.66

Table 6: Validation results for decreased sample size models constructed for the BBar ranch and presented in Table 4. Models were independently validated with the full validation data set (41 validation samples for BBar). Mann-Whitney statistic is p-value.

n	RMSD	Bias	MSD	SB	NU	LC	r ²	Levene	Mann-Whitney
30	0.042	-0.010	0.002	0.0001	0.0000	0.0016	0.49	0.67	0.23
20	0.044	-0.010	0.002	0.0001	0.0002	0.0016	0.51	0.49	0.09
10	0.100	-0.075	0.010	0.0057	0.0015	0.0029	0.10	0.23	0.00

Hypothesis test results for the BBar ranch were insignificant (p -value > 0.05) for models developed with 30 and 20 samples, but were significant (p -value < 0.05) for the model developed with 10 samples (Table 6). This suggested a significant difference between predicted and observed water content for the smallest sample size ($n=10$) model, and therefore rejected the third hypothesis of the study for the BBar model at this sample size. This failed to find a significant difference between predicted and observed water

content for the models developed with 30 and 20 samples, respectively, at this site.

These two reduced sample models had similar MSD values (Table 6) to the full sample model at the BBar ranch, and the model predictions appeared to explain a similar amount of the variability in the validation data set as the full sample model.

Regression Tree Analysis

Regression tree models did not provide a continuous predicted response variable. One important step in addressing the first hypothesis of the study for these models was to review the range of soil water content values predicted. The range should fit within the range of soil water content values observed in the field. The Decker/Bales regression tree predicted discrete mass water content values between 0.08 and 0.16 (Figure 5). The BBar regression tree predicted water content values between 0.08 and 0.25 (Figure 6). The regression trees used different sets of predictor variables than the multiple regression models. Landsat bands such as band 1 and soil variables including bulk density and organic matter were important predictors in the regression tree models. The regression tree model for the Decker/Bales site failed to reject the first hypothesis of the study because it significantly predicted soil water content and contained a variable from each of the data sources. The BBar tree significantly predicted soil water content but did not use a DEM-derived slope or aspect variable, and therefore rejected the first hypothesis of the study.

Validation of the regression tree constructed with the full calibration sample size for the Decker/Bales ranch failed to reject the second hypothesis of the study that there is no significant difference in the validation data between predicted and observed water

content (Table 7). The hypothesis was rejected by validation of the BBar full sample regression tree (Table 8), for which the Mann-Whitney p -value suggested significant differences between the predicted and validation samples. The regression tree models predicted soil water content with a substantially larger average error (RMSD) than the regression model for the BBar site but not the Decker/Bales site, respectively. The regression tree models explained even less of the variability in soil water content than the multiple regression models at both sites, and showed similar lack of correlation between predicted and validation samples in terms of a 1:1 relationship (Figure 7).

Figure 5: Decker/Bales regression tree constructed with full calibration sample size ($n = 50$).

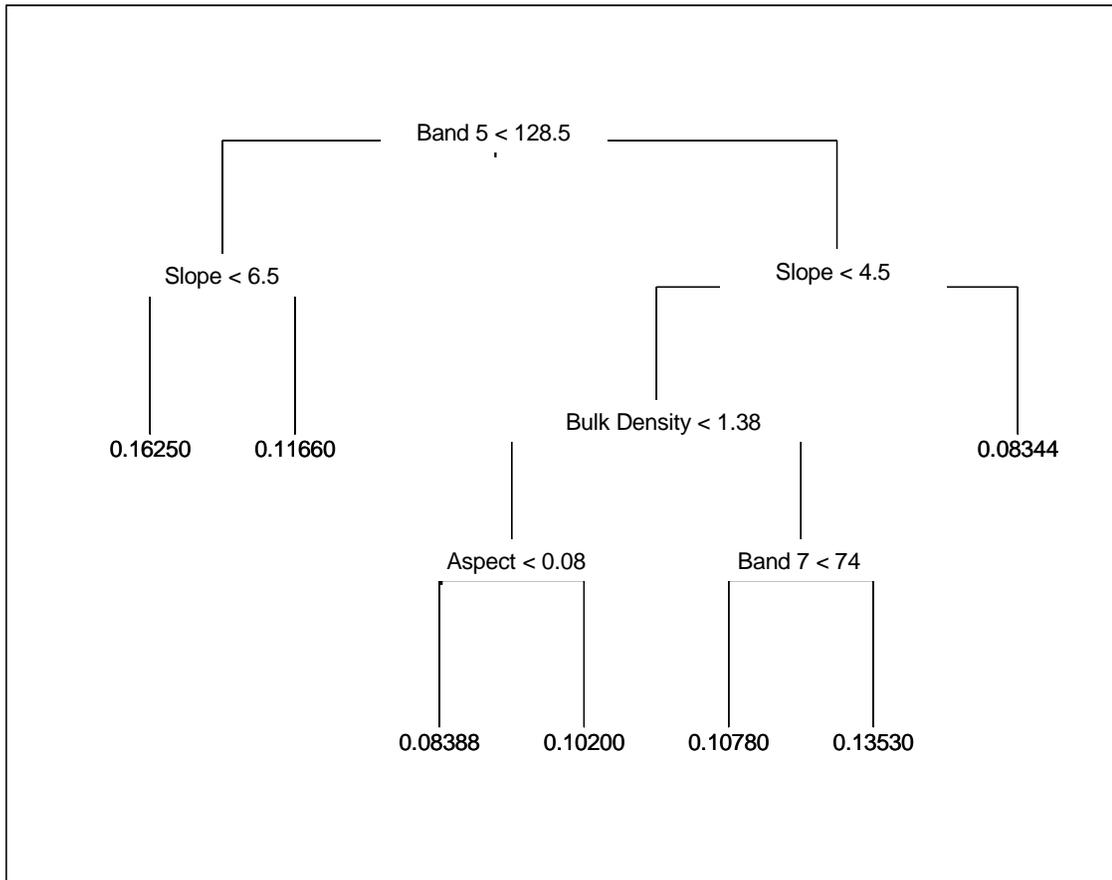


Figure 6: BBar regression tree constructed with full calibration sample size (n = 41).

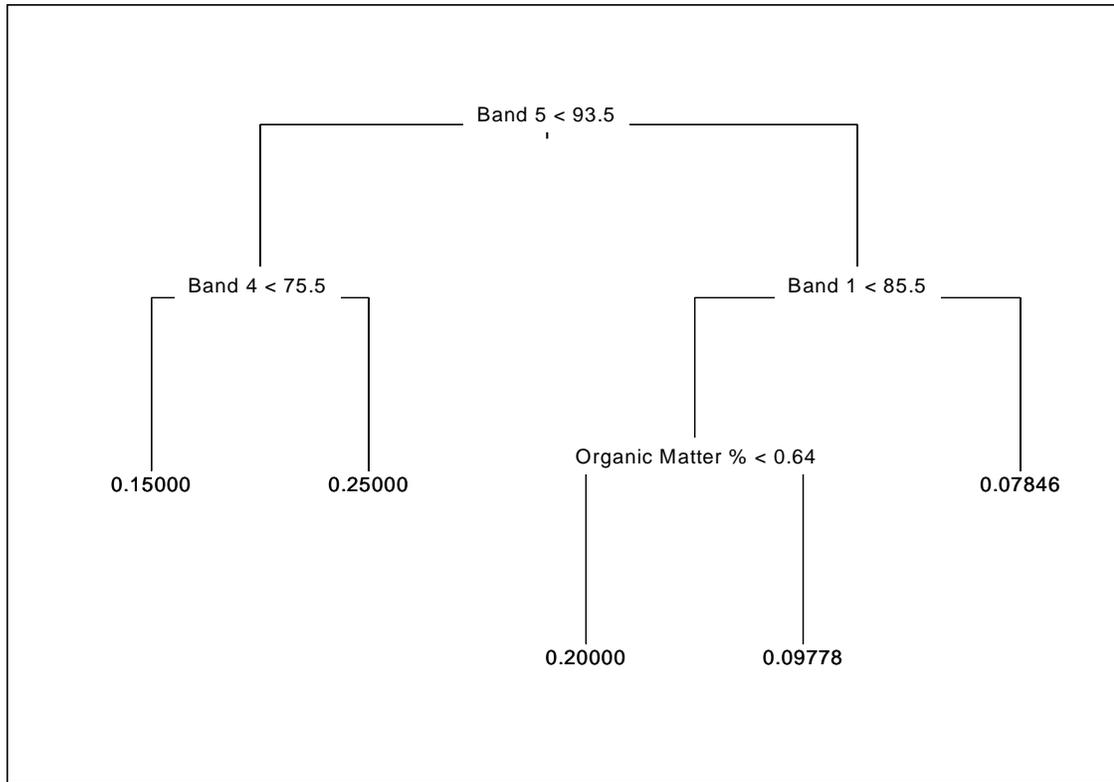


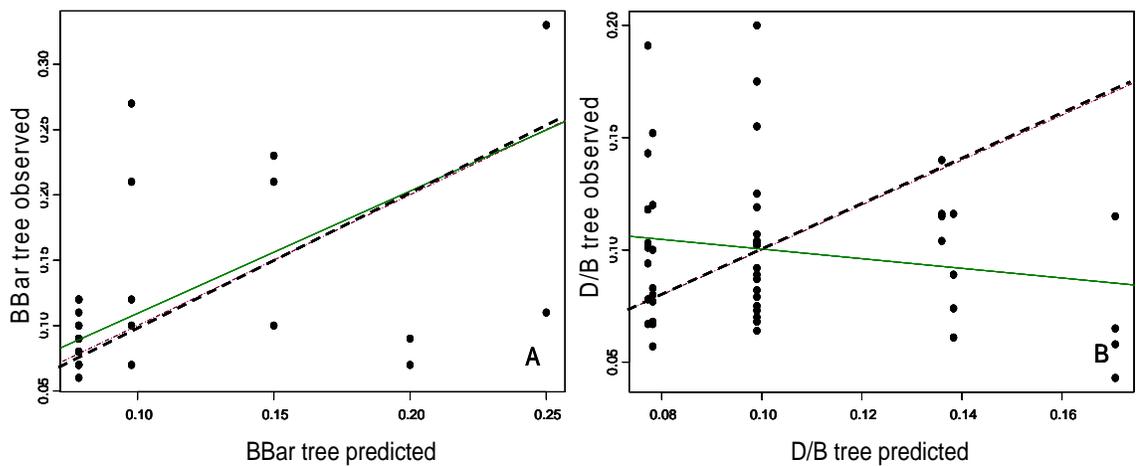
Table 7: Validation results for all regression tree models constructed for the Decker/Bales ranch. 50 sample regression tree results refer to validation of the Decker/Bales tree presented in Figure 5. Subsequent results are for trees constructed with decreased sample sizes. Tree models were independently validated with the full validation data set (50 validation samples for Decker/Bales). Mann-Whitney and Levene statistics are p -values.

n	RMSD	Bias	MSD	SB	NU	LC	r^2	Levene	Mann-Whitney
50	0.048	0.003	0.002	0.0000	0.0012	0.0011	0.03	0.13	0.52
40	0.045	0.001	0.002	0.0000	0.0009	0.0011	0.00	0.49	0.65
30	0.037	0.000	0.001	0.0000	0.0000	0.0000	0.00	0.00	0.70
20	0.067	0.030	0.004	0.0009	0.0024	0.0011	0.00	0.03	0.00
10	0.062	0.024	0.004	0.0006	0.0021	0.0011	0.02	0.03	0.03

Table 8: Validation results for all regression tree models constructed for the BBar ranch. Results for the regression tree constructed with 41 samples refers to validation of the BBar tree presented in Figure 6. Subsequent results are for trees constructed with decreased sample sizes. Tree models were independently validated with the full validation data set (41 validation samples for BBar). Mann-Whitney and Levene statistics are p -values.

n	RMSD	Bias	MSD	SB	NU	LC	r^2	Levene	Mann-Whitney
41	0.055	-0.010	0.003	0.0001	0.0004	0.0025	0.22	0.42	0.01
30	0.047	-0.008	0.002	0.0001	0.0002	0.0020	0.38	0.46	0.19
20	0.051	-0.003	0.003	0.0000	0.0006	0.0020	0.39	0.88	0.10
10	0.063	-0.022	0.004	0.0005	0.0003	0.0032	0.02	0.15	0.04

Figure 7: Predicted versus observed water content graphs for: (A) BBar full calibration sample regression tree model, and (B) Decker/Bales full calibration sample regression tree model. Solid line represents least squares regression for independent validation water content as a function of predicted water content. Dashed line represents 1:1 line ($y = 0 + 1x$).

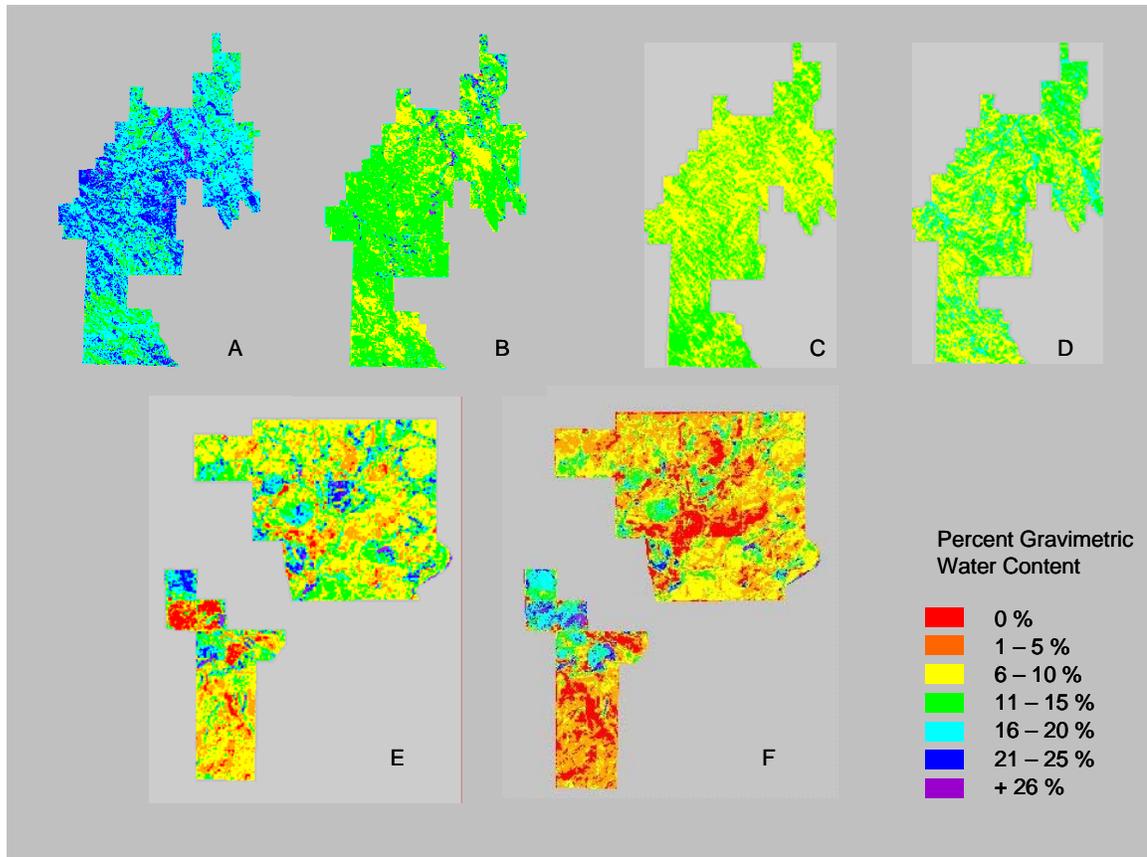


Regression trees constructed with 10 samples at both sites and with 20 and 30 samples at the Decker/Bales site rejected the third hypothesis of the study that there is no significant difference between predicted and observed water content for the decreased sample size models (Tables 7 and 8). Regression trees constructed with 40 samples for the Decker/Bales site failed to show a significant difference between predicted and observed water content (Table 7). Regression trees constructed with 30 and 20 samples for the BBar site failed to show a significant difference between predicted and observed water content (Table 8). The decreased sample size regression tree models for which predicted and validation water content were not shown to be significantly different, explained a similar amount of variability in soil water content and had similar error components as their respective full sample tree. Interestingly, the regression tree models constructed with 20 and 30 samples for the BBar ranch appeared to produce predictions with a slightly stronger correlation to observed water contents than the full sample size tree.

Spring Soil Water Content Maps

Mass water content maps were developed from the models constructed and validated in this study. The multiple regression models predicted a continuous water content response variable and the regression tree models predicted a set of discrete water contents. Maps constructed with both types of models for full and reduced sample sizes were categorized into classes of 0.05 mass water content (Figure 8).

Figure 8: Example mass water content maps constructed with: (A) the Decker/Bales (D/B) multiple regression model developed with 20 calibration samples, (B) the D/B multiple regression model developed with the full calibration sample size for the Decker/Bales ranch, (C) the Decker/Bales regression tree model developed with 30 samples, (D) the Decker/Bales regression tree model developed with the full calibration sample size, (E) the BBar multiple regression model developed with 20 samples, and (F) the BBar multiple regression model developed with the full calibration sample size.



Discussion

Spatially explicit, spring soil water content data sets were predicted with statistically significant models derived from Landsat TM imagery from the previous growing season, DEM-derived slope and aspect layers, and soil thematic maps derived from publicly available soil surveys. The models, however, explained a limited amount

of variability in calibration soil water content samples. The multiple regression models used a variable from each of the data sources, but the regression tree for the BBar site did not use a DEM-derived variable.

Independent validation of both types of models produced problematic results. The model validation results suggested that a wide range of validation statistics should be carefully considered when using the decreased sample size modeling approach and resultant maps. An average error (RMSD) within 0.05 mass water content found for validation of many of the models in this study might seem an acceptable level of accuracy to a rancher interested in using soil water content maps for forage production estimates and other management decisions traditionally based on the rancher's expert knowledge. The average prediction error of 0.05 mass water content appeared substantial, however, considering that the range of observed soil water contents was largely between 0.05 and 0.15 mass water content at both ranches. The predicted and validation samples were not found to be significantly different by the variance and mean difference hypothesis tests for many models. The regression of predicted vs. observed values and the MSD components, however, pointed to a low level of precision for all model predictions. Specifically, the low r^2 values and large LC values, for all models, pointed to a low correlation between predicted and observed values and a great deal of scatter about the 1:1 line, highlighting the imprecise relationship. Validation results were nonetheless different for the two ranches.

Model predictions for the Decker/Bales site explained almost none of the variability in the validation samples. RMSD, LC, and r^2 statistics suggested that model

predictions for the BBar site explained variability in the validation samples, but did so with limiting average error, low precision, and unacceptable individual error between certain prediction – validation pairs that approached 0.10 mass water content. Validation results for models constructed with decreased sample sizes suggested ranchers might parameterize soil water content models for their particular ranches with as few as 20 samples, but these results varied by study site and model. This should be tested for models that perform better in validation in the future.

Maps of spring soil water content can be developed from these models to provide a tool for visualizing model predictions. Maps created with multiple regression models constructed with 20 samples appeared to predict wetter conditions compared to maps of the full sample size regression models at both ranches (Figure 8). There were some obvious exceptions to this generalization in the BBar site maps (Figure 8 E and F), in particular several center pivot irrigation circles in the full sample size map were predicted to be drier in the smaller sample size map. With these notable exceptions, locations mapped as drier with the larger sample model for the BBar site appeared to be mapped as relatively drier locations by the smaller sample size model. This was similarly true for relatively wetter locations in the two maps. This suggested that both maps for the BBar site might portray spring water content with some level of accuracy, though the model validation results certainly pointed to poor precision, corroborating the interpretation of the validation statistics for these BBar models.

Though the models performed poorly for the Decker/Bales site, it is interesting to compare maps developed with the full sample models to maps constructed with the

reduced sample models. The range of measured soil water content at the Decker/Bales site was between 0.04 and 0.25 mass water content, with an average of 0.10 and a standard deviation of 0.03. Both the full sample multiple regression model (Figure 8 B) and the 20 sample multiple regression model (Figure 8 A) appeared to produce maps that over-predicted water content. The map for the reduced sample multiple regression model, in particular, predicted high mass water contents of greater than 0.11 for almost the entire ranch and greater than 0.21 for large areas. There was more between-class variability for wetter locations in the map of the reduced sample multiple regression model (Figure 8 A), when compared to the more homogenous map of the full sample multiple regression model (Figure 8 B). Maps developed with regression tree models for the Decker/Bales ranch appeared more similar between full and reduced sample sizes (Figure 8 C and D) than maps for the multiple regression models.

It is also interesting to consider how the models function in terms of soil, water, and plant relationships with these results in mind. These relationships will be discussed for both study sites with the realization that the poor validation results limit the extent to which model-landscape interpretations can be considered realistically. These relationships will be discussed first for the best multiple regression models developed with the full calibration sample size and then for the regression tree models developed with the full calibration sample size.

Multiple Regression Models

Landsat TM 5 imagery from the peak of the previous growing season served as a useful predictor of spatially explicit soil water contents at both study sites (Table 1).

Bands 3 and 4 were the most common imagery predictor variables in the multiple regression models developed for this study. The bands were useful both as individual predictors and in interaction terms, suggesting they were dependent on one and other, and on factors such as topographic slope. Band 3 had a negative coefficient in all models as an individual predictor (Table 1). Band 4 had a negative coefficient in the BBar model and a positive coefficient in the Decker/Bales model (Table 1). Reflectance values that are high for band 4 and low for band 3 can result from land surfaces covered with healthy green vegetation (Jensen, 1996). A positive band 4 coefficient and a negative band 3 coefficient, such as in the Decker/Bales model, might suggest that locations with more growing season green biomass had higher spring soil water contents. This might suggest that water collecting landscape positions, or positions that held more water due to soil characteristics like texture and depth, might have supported both higher plant productivity in the growing season and higher soil water content in the spring. This is the opposite of an evapotranspiration driven interpretation where areas of lower leaf area with resulting limited evapotranspiration might have been expected to conserve soil water for later seasons, such as in a fallow agricultural field (Hatfield et al., 2001).

The negative band 3 and 4 coefficients for the BBar site (Table 1) suggest that spring soil water content decreased with an increase in both red and NIR reflectance during the previous growing season. TM band 3 and 4 reflectance have been shown to decrease with increased surface soil water content on fallow agricultural surfaces (Moran et al., 2002). Reflectance for both bands has been suggested to be high for both increased cover of senescent litter (Moran et al., 2002) and exposed bare soil (Asner et al., 2000,

Hill and Schutt, 2000). The negative coefficients for bands 3 and 4 in the BBar models might suggest that spring soil water content was more related to surface water content, senescent vegetation cover, and/or bare soil than abundance of healthy green biomass. Both study sites had non-irrigated pastures with relatively high amounts of exposed bare soil surfaces. Some of the BBar site was irrigated with both flood irrigation and center pivots. Timing and amount of irrigation might have influenced both surface water content and stage of vegetation growth or senescence during the time of image acquisition. Timing and intensity of grazing, similarly, might have influenced vegetative cover and relative amounts of exposed bare soil at both sites. Both timing and intensity of plant defoliation by grazing have been shown to influence soil water storage (Bremer, 2001).

The TM thermal band (band 6) was a useful predictor of spring soil water content at the Decker/Bales site when bands 3 and 4 were in the model (Table 1). Emittance measured by the thermal band might be influenced by ground surface temperature and water content (Jensen, 1996). Remote sensing in the thermal range has been used to link ground surface temperature with evapotranspiration rates (Schmugge et al., 2002). High band 6 emittance at the peak of the previous growing season might have suggested higher surface temperatures, which might have indicated water-limited areas or areas with high exposed bare soil and low vegetation cover. Higher evaporation rates can lead to greater soil water depletion and lower soil water content in later seasons (Hatfield et al., 2001).

Aspect was only used as a predictor variable for the model developed for the Decker/Bales site (Table 1). The positive coefficient for aspect in the Decker/Bales

model suggests that southerly aspects had lower soil water contents and northerly aspects had relatively higher soil water contents. This is an expected relationship in the Northern Hemisphere where southerly aspects of hill slopes receive higher potential solar radiation than northerly sloping hill sides. Another expected relationship exists between topographic slope and soil water content. The sign for the coefficient for the percent slope variable suggests that water content was lower on steeper slopes at the BBar site (Table 1). Water content is generally expected to be lower on steeper slopes due to surface and sub-surface flow (Grayson et al., 1997; Western et al., 1999). The redistribution of soil water by sub-surface flow, however, is probably not substantial in semi-arid environments where soil water content might not be highly influenced by terrain (Landon, 1995; Kozar, 2002). The water content and slope relationship might have been mitigated by soil characteristics like texture, organic matter content, and depth, as well as vegetation characteristics (Pachepsky et al., 2001; Chamran et al., 2002).

Soils with higher clay content tend to have a higher amount of small pores compared to soils comprised of larger particle sizes (Brady, 1990). The clay predictor variable had positive coefficients for models at both study sites (Table 1). This suggests that locations with higher average clay content in the upper 100 cm of the soil profile had higher spring soil water content.

Regression Tree Models

Regression tree analysis offered several advantages to multiple regression analysis, though the average soil water content prediction error was generally greater. The trees took advantage of variables that the multiple regression models did not utilize.

The trees were also somewhat easier to interpret. For example, trees for both sites began with a split between higher band 5 values and lower band 5 values (Figures 4 and 5). The Landsat TM 5 band 5 is a middle infrared sensor and is sensitive to plant and soil surface water content (Jensen, 1996). Trees for both sites gave the highest spring soil water content for locations with low growing season MIR reflectance.

The Decker/Bales regression tree (Figure 5) highlighted some expected soil water and soil landscape relationships. Southerly slopes were drier than more northerly slopes. Soil water content was as much a function of soil characteristics of bulk density and organic matter as it was of slope gradient at drier locations, an expected relationship in drier states (Grayson et al., 1997; Western et al., 1999; Chamran et al., 2002; Pachepsky et al., 2001). Steeper slopes supported drier spring soil profiles at lower middle infrared values, which were generally wetter locations. This relationship is generally expected in wetter soil water states, but not in the semi-arid environment studied (Landon, 1995; Grayson et al., 1997; Western et al., 1999; Kozar, 2002).

The relationships between soil characteristics and spring soil water content were somewhat counterintuitive for both sites. Soils with higher organic matter and lower bulk density are generally expected to have higher capacity for water storage (Brady, 1990). It remains unclear why there was higher spring water content in soils with lower organic matter at the BBar site, possibly a texture or depth difference existed for lower organic matter versus higher organic matter soils (Figure 6). There might be an explanation for why soils with higher bulk density had higher water content at the Decker/Bales site (Figure 5). A branch of the Decker/Bales tree (Figure 5) predicted that locations with

low slope (< 4.5%) and high bulk density were some of the wetter locations. Perhaps the low slope/high bulk density locations were wetter due to higher infiltration capacity compared to more finely textured, low bulk density soils in similar gently sloping, water collecting positions on the landscape.

Conclusions

Statistically significant models of spring soil water content were developed with the publicly available data of Landsat imagery from the previous growing season, USGS DEM-derived variables, and digitized soil survey attribute layers. Models developed in this study explained a limited amount of variability in spring soil water content, however, and when validated showed limited accuracy and low precision. These limitations are probably not surprising when one considers the relatively coarse scale of the satellite imagery (30m), the interpolated nature of the DEMs (USGS digitized contours), and the qualitative landscape-scale soil characterization of the soil surveys.

Ranchers might be able to predict spring soil water content within acceptable error using a small field collected data set of 20 samples, publicly available GIS layers, and empirical multiple regression models. The models developed in this study were site and date specific and would have to be parameterized to a particular ranch with a local soil water content data set. The site-specific parameterization approach might be successful if more accurate and precise soil water content models could be developed. Data sets of 20 samples are small enough to collect in one day and were shown in this study to validate similarly to larger sample models in some cases. Future work should

focus on testing the site-specific parameterization approach if better performing empirical soil water content models can be developed. Ranchers in the Northern Great Plains would benefit from maps developed from such models when predicting forage production based on established relationships between production and spring soil water content. Such maps would compliment the guess work and expert knowledge used traditionally for setting stocking rates.

References Cited

- Asner, G.P., Wessman, C.A., Bateson, C.A., Privette, J.L. (2000). Impact of tissue, canopy, and landscape factors on the hyperspectral reflectance variability of arid ecosystems. *Remote Sensing of Environment*. 74:69-84.
- Ayenew, T. (2003). Evapotranspiration estimation using thematic mapper spectral satellite data in the Ethiopian rift and adjacent highlands. *Journal of Hydrology*. 279:83-93.
- Bastiaanssen, W.G.M., Ali, S. (2003). A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan. *Agriculture, Ecosystems and Environment*. 94:321-340.
- Bastiaanssen, W.G.M., Chandrapala, L. (2003). Water balance variability across Sri Lanka for assessing agricultural and environmental water use. *Agricultural Water Management*. 58:171-192.
- Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., Holtslag, A.A.M. (1998). A remote sensing surface energy balance algorithm for land (SEBAL) 1. Formulation. *Journal of Hydrology*. 212:198-212.
- Brady, N. (1990). *The Nature and Properties of Soils*. Macmillan: New York.
- Bremer, D.J., Auen, L.A., Ham, J.M., Owensby, C.E. (2001). Evapotranspiration in a prairie ecosystem: effects of grazing by cattle. *Agronomy Journal*. 93:338-348.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984). *Classification and Regression Trees*. Wadsworth International Group, Belmont, CA.
- Burrough, P.A., Beckett, P.H.T., Jarvis, H.G. (1971). The relation between cost and utility in soil survey (I-III). *Soil Science Society of America Journal*. 22:368-381.
- Chamran, F., Gessler, P.E., Chadwick, O.A. (2002). Spatially explicit treatment of soil-water dynamics along a semi-arid catena. *Soil Science Society of America Journal*. 66:1571-1583.
- Decker, G.L. (1972). *Automatic Retrieval and Analysis of Soil Characterization Data*. PhD. dissertation. Montana State University. Bozeman, MT.

- Feng, C-X. J., Yu, Z-G. S., Wang, J-H. J. (2005). Validation and data splitting in predictive regression modeling of honing surface roughness data. *International journal of production research*. 43:1555-1571.
- Gauch, H.G., Hwang, J.T., Fivk, G.W. (2003). Model evaluation by comparison of model-based predictions and measured values. *Agronomy Journal*. 95:1442-1446.
- Grayson, R.B., Western, A.W., Chiew, F.H.S. (1997). Preferred states in spatial soil moisture: local and nonlocal controls. *Water Resources Research*. 33:2897-2908.
- Hatfield, J.L, Sauer, T.J., Prueger, J.H. (2001). Managing soils to achieve greater water use efficiency: a review. *Agronomy Journal*. 93:271-280.
- Hemakumara, H.M., Chandrapala, L., Moene, A.F. (2003). Evapotranspiration fluxes over mixed vegetation areas measured from large aperture scintillometer. *Agricultural Water Management*. 58:109-122.
- Hill, J., Schutt, B. (2000). Mapping complex patterns of erosion and stability in dry mediterranean ecosystems. *Remote Sensing of Environment*. 74:557-569.
- <http://digital-ngp.aero.und.edu>. Digital Northern Great Plains. Upper Midwest aerospace Consortium, University of North Dakota: Grand Forks, ND.
- <http://gisdata.usgs.net/website/Seamless>. Earth Resources Observation and Science Seamless Data Distribution Center. United States Department of the Interior, U.S. Geological Survey: Washington, D.C.
- <http://soildatamart.nrcs.usda.gov>. Soil Data Mart. United States Department of Agriculture, Natural Resources Conservation Service: Fort Worth, Texas.
- Hudson, Berman. (1992). The soil survey as a paradigm based science. *Soil Science Society of America Journal*. 56:836-841.
- Hunt, E.R., Everitt, J.H., Ritchie, J.C., Moran, M.S., Booth, D.T., Anderson, G.L., Clark, P.E., Seyfried, M.S. (2003). Applications and research using remote sensing for rangeland management. *Photogrammetric Engineering and Remote Sensing*. 69:675-693.
- Jenny, H. (1941). *Factors of soil formation: a system of quantitative pedology*. (originally published by McGraw-Hill; Mineola, N.Y.: Dover).
- Jensen, J.R. (1996). *Introductory digital image processing: a remote sensing perspective*. Prentice Hall, Englewood Cliffs, New Jersey.

- Kozar, B.J. (2002). Predicting soil water distribution using topographic models within four Montana farm fields. M.S. thesis. Montana State University: Bozeman, Montana.
- Landon, M.A. (1995). Soil and terrain attributes for evaluation of leaching in a Montana farm field. M.S. thesis. Montana State University: Bozeman, Montana.
- Marshall, T.J., Holmes, J.W., Rose, C.W. (1996). Soil Physics. Press Syndicate of the University of Cambridge: New York, New York.
- Milne, G. (1936). Normal erosion as a factor in soil profile development. *Nature*. 138:548-549.
- Montagne, C., Munn, L.C. Nielsen, G.A., Rogers, J.W., Hunter, H.E. (1982). Soils of Montana. Montana Agricultural Experiment Station, Montana State University. Bozeman, MT.
- Moore, I.D., Grayson, R.B., Ladson, A.R. (1991). Digital terrain modeling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes*. 5:3-30.
- Moran, S.M., Hymer, D.C., Qi, J., Kerr, K. (2002). Comparison of ERS-2 SAR and Landsat TM imagery for monitoring agricultural crop and soil conditions. *Remote Sensing of Environment*. 79:243-252.
- Neff, E.L., Wight, J.R. (1977). Overwinter soil water recharge and herbage production as influenced by contour furrowing on Eastern Montana rangelands. *Journal of Range Management*. 30:193-195.
- Obrist, D., Verburg, P.S., Young, M.H., Coleman, J.S., Schorran, D.E., Arnone, J.A. (2003). Quantifying the effects of phenology on ecosystem evapotranspiration in planted grassland mesocosms using EcoCELL technology. *Agricultural and Forest Meteorology*. 118:173-183.
- Pachepsky, Y.A., Timlin, D.J., Rawls, W.J. (2001). Soil water retention as related to topographic variables. *Soil Science Society of America Journal*. 65:1787-1795.
- Pfaffenberger, R.C., Patterson, J.H. (1981). Statistical Methods. Richard D. Irwin, Inc.: Homewood, Illinois.
- Qi, J., Kerr, Y.H., Moran, M.S., Wetz, M., Huete, A.R., Sorooshian, S., Bryant, R. (2000). Leaf area index estimates using remotely sensed data and BRDF models in a semi-arid region. *Remote Sensing of Environment*. 73:18-30.

- Rogler, G., Haas, H.J. (1946). Range production as related to soil moisture and precipitation on the Northern Great Plains. *Journal of the American Society of Agronomy*. 378-389.
- Ruhe, R.V., Walker, P.H. (1968). Hillslope models and soil formation. I. Open systems. *Transactions of the 9th International Congress of Soil Science*. 4:551-560.
- Schmugge, T.J., Kustas, W.P., Ritchie, J.C., Jackson, T.J., Rango, A. (2002). Remote sensing in hydrology. *Advances in Water Resources*. 25:1367-1385.
- Simonson, R.W. (1959). Outline of a generalized theory of soil genesis. *Soil Science Society of America Proceedings*. 23:152-156.
- Soil Survey Staff. (1971). Soil Survey Powder River Area Montana. United States Department of Agriculture, Soil Conservation Service: Washington, D.C.
- Soil Survey Staff. (2004). Soil Survey (SSURGO) database for Sweet Grass County Area, Montana. United States Department of Agriculture, Natural Resources Conservation Service: Fort Worth, Texas. <http://soildatamart.nrcs.usda.gov>.
- Starks, P.J., Jackson, T.J. (2002). Estimating soil water content in tallgrass prairie using remote sensing. *Journal of Range Management*. 55:474-481.
- Veseth, R., Montagne, C. (1980). Geologic Parent Materials of Montana Soils. Bulletin 721. Montana Agricultural Experiment Station and United States Department of Agriculture, Soil Conservation Service: Bozeman, MT.
- Walker, P.H., Ruhe, R.V. (1968). Hillslope models and soil formation. II. Closed systems. *Transactions of the 9th International Congress of Soil Science*. 4:561-568.
- Wang, C., Qi, J., Moran, S., Marsett, R. (2004). Soil moisture estimation in a semi-arid rangeland using ERS-2 and TM imagery. *Remote Sensing of Environment*. 90:178-189.
- Western, A.W., Grayson, R.B., Bloschl, G. (2002). Scaling of soil moisture: a hydrologic perspective. *Annual Review of Earth and Planetary Sciences*. 30:149-180.
- Western, A.W., Grayson, R.B., Bloschl, G., Willgoose, G.R., McMahon, T.A. (1999). Observed spatial organization of soil moisture and its relation to terrain indices. *Water Resources Research*. 35:797-810.
- Wosten, J.H., Pachepsky, Y.A., Rawls, W.J. (2001). Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*. 251:123-150.

CHAPTER 4

EVALUATING THE SUITABILITY OF SOIL SURVEY AS A SOURCE OF SITE-SPECIFIC SOILS DATA FOR PREDICTIVE SOIL WATER CONTENT MODELING

Introduction

Rangeland forage production in the semi-arid Northern Great Plains is a function of spring soil water storage and growing season precipitation (Rogler and Haas, 1946; Neff and Wight, 1977). Soil water content models and predictive spring water content maps might provide ranchers with useful information for range management decisions. Spatially explicit soil water content was empirically modeled with combinations of predictor variables derived from: Landsat imagery from the previous growing season, DEM-derived terrain attributes, and soil survey-derived soil attributes (Chapter 3). The models explained a limited amount of variability in spring water content and with independent validation were shown to predict with little precision (Chapter 3). There is a question of how accurate the soil survey data source is, however, and whether this affected modeling efforts.

The addition of site-specific soils data in terrain and soil survey based soil water content modeling has been recommended for future research in semi-arid agricultural systems (Kozar, 2002). Publically available soil survey maps and attribute data produced by the National Cooperative Soil Survey provide the most spatially contiguous soil data source in the U.S. The suitability of traditional soil survey attribute data for site-specific

applications is questionable, however, because the data is often interpolated and/or extrapolated from a handful of lab characterized pedons for an entire survey area (Hudson, 1992). This study uses site-specific characterization data to evaluate and assess the limitations of soil survey-derived attributes as predictor variables in spring soil water content models. The site-specific characterization data used in this study is derived from diffuse reflectance spectroscopy (DRS), which has been acknowledged as a potential tool for rapid and economically efficient characterization of soil samples (Dunn et al., 2002; Shepherd and Walsh, 2002). The evaluation of the soil survey data source is accomplished through a comparison of the relative predictive abilities of models developed with the soil survey versus DRS characterization data sources. Models developed with both soil data sources include the Landsat and DEM-derived variables previously shown to be useful predictors of spring soil water content (Chapter 3).

Soil Characterization with Diffuse Reflectance Spectroscopy

The process commonly termed visible (VIS) and near-infrared (NIR) DRS by the laboratory spectroscopy community senses soil reflectance from 350 to 2500 nm in the electromagnetic spectrum. Specific mineral classes, such as certain iron oxides, carbonates, and clay minerals, have distinct bands of absorption in this range of the spectrum (Clark, 1999). Soils and their parent materials are assemblages of such minerals and, therefore, have spectral signatures that are often muted combinations of the signatures of the specific minerals present (Hunt, 1989). The presence of hydroxyl groups, water, carbonate, sulfate, and phosphate can lead to inherent vibrational overtones that can act individually or in combination to produce distinct bands of

absorption in this portion of the spectrum (Hunt, 1989). DRS characterization of soils developed from sedimentary rock, as many of the soils of the Northern Great Plains are, might be expected to rely heavily on relative composition of carbonates, clay mineralogy, and iron oxides. It might also be expected to rely on the distribution of organic matter in which carbon – hydrogen bonds have vibrational overtone and combination absorption bands in the short wave infrared portion of the spectrum (Clark, 1999). Color is yet another diagnostic soil characteristic that has recognizable spectral features in the visible portion of the spectrum and is related to clay content and organic matter, among other properties (Ben-Dor et al., 1999; Clark, 1999).

DRS has been used to predict a range of soil characteristics in agricultural systems from cation exchange capacity, to particle size fractions, total, organic, and inorganic carbon, as well as concentrations of exchangeable cations including calcium and potassium (Dunn et al., 2002; McCarty et al., 2002; Shepherd and Walsh, 2002; Brown et al., in press a; Brown et al., in press b). Clay content and soil organic carbon (SOC) predictions, specifically, have been validated within a substantial range of average error. Validation root mean square deviations (RMSD) from 75 g/kg to 95 g/kg have been reported for clay predictions with VIS and NIR first derivative reflectance (Shepherd and Walsh, 2002; Brown et al., in press b). RMSD from 1.26 g/kg SOC in local field calibration studies (Brown et al., in press a) to 3.1 g/kg to 9.0 g/kg SOC in regional to global calibration studies have been reported (McCarty et al., 2002; Shepherd and Walsh, 2002; Brown et al., in press b). The disparity in average prediction error has been suggested to be related to lab procedures (specifically for SOC) on which validation

reference samples are based (McCarty et al., 2002), the proportion of calibration to validation samples (Shepherd and Walsh, 2002), validation approach used (e.g., cross validation versus random holdout), and relevance of validation samples to model geographic extent (Brown et al., in press a). An untested approach of using weighted samples from a global spectral library and a small set of local calibration samples is employed in this study to model clay content and SOC from first derivative VIS and NIR spectral reflectance.

Objective

This study examines the effect of the expected relative imprecision of soil survey data compared to sample point specific characterized soil properties in predictive models based on data collected for this study. This study addresses this issue in the context of empirical soil water content prediction models. The relative soil water content predictive abilities of a soil survey-derived data set and a field and lab characterization data set modeled from DRS data were compared for two Montana ranches. The specific hypothesis of the study is that there is no statistically significant difference between soil water content predictions by models constructed with soil survey predictor variables and similarly constructed models with field and lab characterization predictor variables.

Study Sites

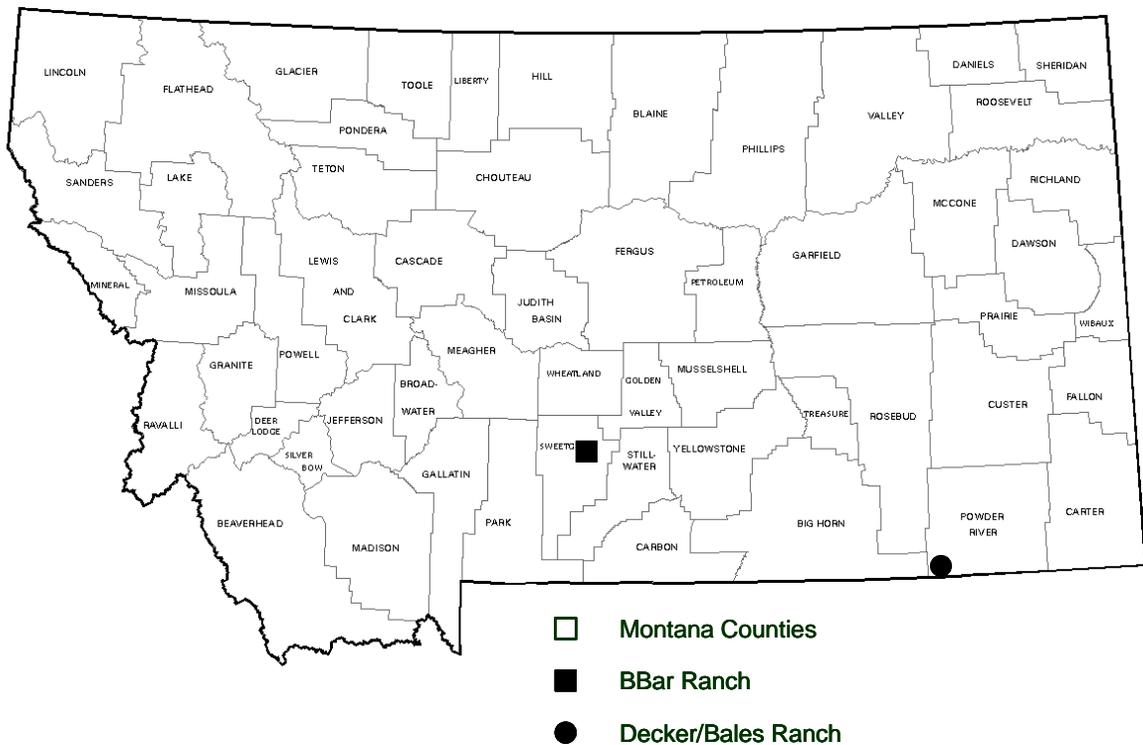
The Decker/Bales ranch and the BBar ranch served as the two study sites (Figure 9). The Decker/Bales ranch is approximately 100 km² and is located in southwestern

Powder River County in southeastern Montana. It is mostly in the Tongue River watershed, but includes an area of the divide between the Tongue and Powder rivers. The landscape is part of Montana's non-glaciated plains and is characterized by dissected sedimentary layers that form a low relief, fluvially incised landscape. Range vegetation consists of grassland communities of western wheatgrass (*Agropyron smithii* Rydb.), needle and thread (*Stipa comata* Trin. & Rupr.), and blue grama (*Bouteloua gracilis* Willd. ex Kunth), with a varying presence of big sagebrush (*Artemisia tridentata* Nutt.) (Montagne et al., 1982). Soils include loamy, calcareous Ustorthents formed in siltstones, clayey, calcareous Ustorthents formed in shales, fine to coarse-loamy Haplustalfs formed in slope alluvium, loamy-skeletal Haplustalfs formed in scoria beds, and fine Natrustalfs that are often associated with prairie dog communities (Veseth and Montagne, 1980; Montagne et al., 1982). The area receives approximately 30 cm of mean annual precipitation, the soil temperature regime is on the boundary between Mesic and Frigid, and the soil moisture regime is on the boundary between Ustic and Aridic (Soil Survey Staff, 1971).

The BBar ranch is approximately 30 km² and is located in northern Sweet Grass County in south-central Montana (Figure 9). It lies in a valley at the Rocky Mountain front in the westernmost extent of Montana's non-glaciated plains. The landscape consists of rolling, sedimentary bedrock-controlled hills vegetated with grassland communities of western wheatgrass, little bluestem (*Andropogon scoparius* Michx.), needle and thread, and blue grama (Montagne et al., 1982). There are isolated surfaces of alluvium and outwash associated with the Crazy Mountains. Soils that form in these

parent materials range from fine Argiustolls on backslopes, footslopes, and toeslopes, to loamy-skeletal Ustorthents on summit and shoulder positions, as well as fine Natrustalfs on toeslopes and valley floor positions, and fine and fine-loamy Torrifuvents in drainageways (Veseth and Montagne, 1980; Montagne et al., 1982). The area receives approximately 35 cm of mean annual precipitation, the soil temperature regime is Frigid, and the moisture regime is Ustic (Soil Survey Staff, 2004).

Figure 9: Study site locations in Montana. BBar ranch is located in Sweet Grass county and Decker/Bales ranch is in Powder River county.

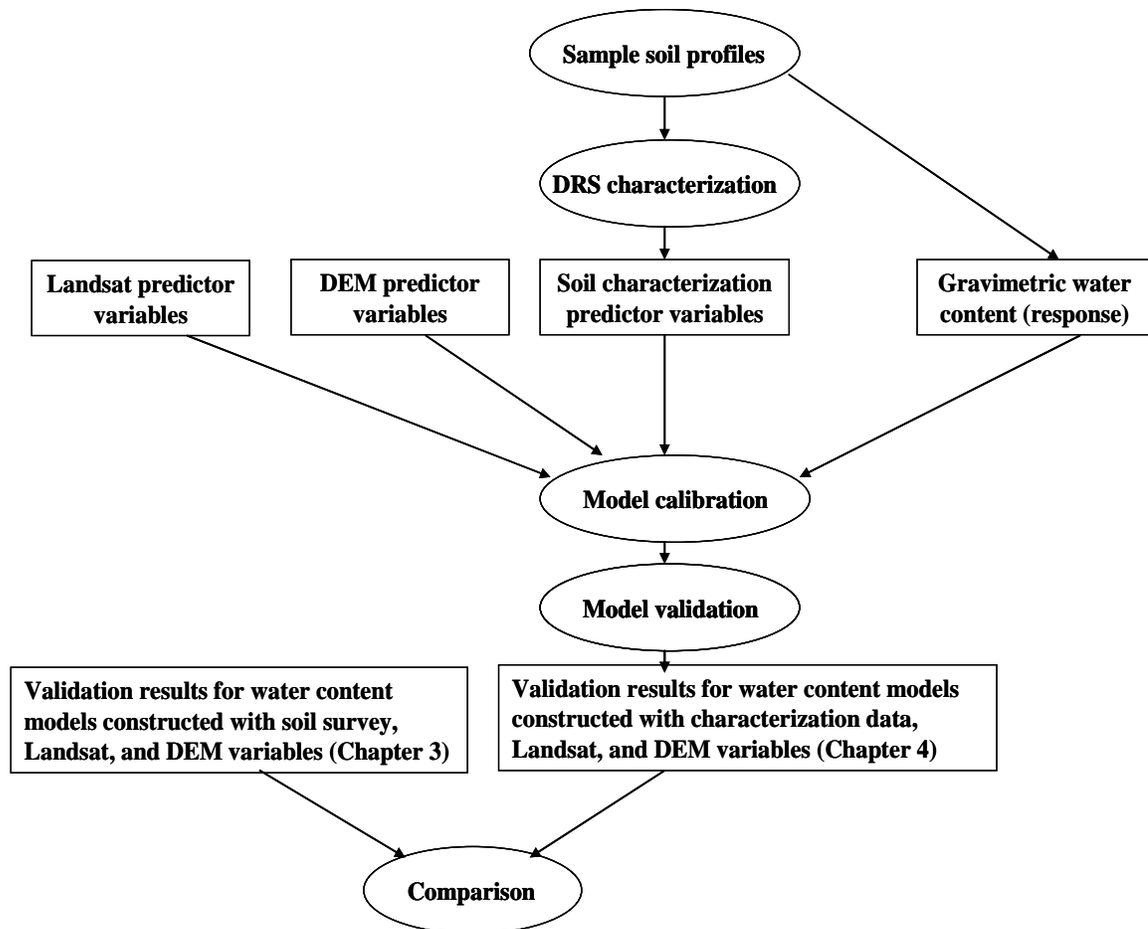


Methods

Methods implemented for both study sites in this study included field soil sampling, measurement of mass water content for samples, and laboratory

characterization of samples with DRS (Figure 10). Models were developed to predict soil water content with predictor variables derived from Landsat, DEM, and soil characterization data sources. Models were independently validated with a reserved soil water content data set. Model validation results were compared with validation results for models constructed similarly, but with a soil survey predictor variable data source instead of soil characterization data source in Chapter 3 (Figure 10).

Figure 10: Diagram outlining the general procedure, from development of the water content response variable to testing of the study hypothesis, followed in Chapter 4.



Field Data Collection

A digitized representation of each ranch's boundary was considered the extent of each study area. A statistical power test (Pfaffenberger and Patterson, 1981) was performed on a small preliminary data set of depth of moist soil measurements from the Decker/Bales ranch (n=11) with a mean of 74 cm of moist soil and a standard deviation of 9.6 cm of moist soil. The test assumes normal distribution, constant variance in the data set, and an alpha level of 0.05. The power test showed that 41 points were necessary to be able to detect a significant difference of 7.6 cm of moist soil. At least 41 points for model development and 41 points for validation were targeted for each study area during sampling.

Sample locations were stratified based on the soil survey of the county in which each ranch resides. Soil survey maps were used to account for both the variability in soil type as well as the variability in slope, aspect, landform, and landform position at each ranch. The spatial data layer of each digitized survey was clipped to the extent of each ranch's digitized boundary. Random points were selected within each soil survey map unit, with at least one location for each named map unit.

Each sample point was identified with an X,Y location in UTM coordinates. Navigation to the points was accomplished with map and a GPS receiver with an accuracy of < 1m. The location of each sample point was logged in the GPS receiver as a waypoint and coordinates were recorded on the datasheet as well.

A hand auger was used to collect soil samples in eight 10 cm increments, from the soil surface to 100 cm depth at each sample location. It was assumed that variability in

soil characteristics would decrease with depth. Samples from 60 to 70 cm and 80 to 90 cm were not collected for logistical and efficiency purposes. Depth to root restrictive layer was recorded for each profile if observed within 100 cm.

A total of 100 locations were sampled at the Decker/Bales ranch and 82 locations were sampled at the BBar ranch. Sampling was completed during the first week of May 2004 at the Decker/ Bales ranch and during the second week of May 2004 at the BBar ranch. The samples were collected in wax lined paper bags and were transported to the lab at the end of field work. Samples were weighed at field moist state, oven dried at 105° C, and then weighed again to calculate mass water content. Mass water content was averaged for the entire sampled profile at each sample location. Average soil profile mass water content served as the response variable for water content modeling at both study sites.

Lab Characterization

All samples from both ranches from the 0-10 cm, 30-40 cm, and 70-80 cm depths were selected as a subset for lab characterization. The fine earth fraction (< 2mm) was separated from coarse fragments by grinding and sieving. The fine earth fraction for the approximately 540 samples was scanned with an ASD “Fieldspec Pro FR” spectroradiometer (Analytical Spectral Devices, Boulder, CO). The spectroradiometer has a spectral range of 350-2500 nm, a 2 nm sampling resolution, a spectral resolution of 3 nm at 700 nm, and a spectral resolution of 10 nm at 1400 and 2100 nm. The spectroradiometer was set up to record a composite reflectance signature of 10 internally averaged scans between 350 and 2500 nm. The samples were scanned in an optical

quality glass petri dish. Two scans were collected for each sample with a 90 degree rotation between scans. Replicate scan spectra were compared and samples were rescanned when possible errors were detected in reflectance and 1st derivatives. Replicate spectra were averaged for each sample, smoothed, and 1st derivative values were extracted in 10 nm segments from 360 to 2490 nm.

A subset of samples was selected from the combined set of scanned samples for total carbon, inorganic carbon, and particle size analysis. Reflectance data was used to predict clay content, inorganic carbon, total carbon, and clay mineralogy classes of Vermiculite, Montmorillonite, and Kaolonite for all the scanned samples. This was done with existing reflectance models for each soil characteristic developed with 3794 globally collected Natural Resources Conservation Service (NRCS) samples (Brown et al., in press b). The samples with maximum and minimum values for each modeled characteristic were selected as members of the subset for lab characterization. The remaining samples from those particular profiles were selected as well so that the subset would contain entire characterized profiles. Several remaining entire profiles of samples were then selected at random in order to fill out the subset to approximately 100 samples, a logistically feasible size for lab characterization. The final subset contained 106 samples representing 37 soil profiles, with 44 samples from the BBar ranch and 62 samples for the Decker/Bales ranch.

Carbon Analysis. Total carbon and total nitrogen were measured for each sample of the subset with a LECO C/N/S 2000 analyzer (LECO Corporation, St Joseph, MI, USA). The LECO machine measures C and N by dry combustion of 1 gram of milled,

fine earth fraction of sample. Inorganic carbon was measured by a modified pressure calcimeter method (Sherrod et al., 2002). HCl was added to 1 gram of milled, fine earth fraction of each sample in a stoppered and capped vial. Pressure resultant from HCl reaction with CaCO_3 was measured after a two hour reaction period. Pressure was measured via a pressure transducer connected by rubber tubing to a hypodermic needle used to puncture the stoppered vial.

Soil organic carbon was calculated as the difference between total carbon and inorganic carbon for each sample. The ratio of percent organic carbon to percent nitrogen was reviewed for each sample. Those samples for which the ratio was outside the range of 3:1 and 15:1 were rerun for inorganic carbon analysis (Robertson et al., 1997; Brown et al., in press a).

Particle Size Analysis. Particle size analysis was performed by the pipette method (Gee and Bauder, 1986, Soil Survey Staff, 1996). 10 grams of milled, fine earth fraction of each sample was treated with HCl to remove carbonates. NaOCl was added to samples for organic matter removal. Samples were dispersed with NaHMP and shaken overnight prior to sieving for sand separation and sedimentation in a graduated cylinder for clay separation by pipette method. Sand, silt, and clay fractions for each sample were determined as percent by weight basis.

Spectral Modeling of Soil Characteristics. TreeNet® software was used to model percent clay content (clay) and soil organic carbon (SOC) with boosted regression tree models. A maximum of 1000 trees was specified, with minimum and maximum number of nodes per tree, 10 and 12, respectively – parameters arrived at heuristically in a

previous study (Brown et al., in press b). The 106 lab characterized samples from this study were pooled with 1,566 NRCS samples that had previously been characterized for particle size and carbon analysis and scanned with a spectroradiometer. This set of samples served as training data for boosted regression tree development. Ten iterations of boosted regression tree calibration were performed with a 1/10 holdout of the 106 lab samples. The 10 calibration and validation subsets were stratified by study site and soil profile. 5 subsets contained sampled profiles from the BBar site and 5 from the Decker/Bales site. This allowed for site-specific model validation.

The majority of the samples used in model development were from characterized profiles in the NRCS archives (Brown et al., in press b), making the applied models relatively global in nature. The importance of local calibration by geographic weighting was tested by sequentially applying relative weights to the NRCS and site-specific samples in several iterations of the boosted regression tree modeling procedure. Local samples from the two study sites were always given a full weight of 1, and NRCS samples were given weights of 0.01, 0.25, 0.50, 0.75, and 1 for five iterations of model development. Models from each iteration were validated with the 1/10 local sample holdout. The 10 holdout subsets were stratified by study site and soil profile, and model validation was assessed individually for the two study sites.

Predicted and measured values were compared for validation by calculating a mean squared deviation (MSD) and root mean square deviation (RMSD). The MSD was broken into components of standard bias (SB), non-unity (NU), and lack of correlation (LC) with the following equations (Gauch et al., 2003):

$$\text{MSD} = \sum_n (\text{Predicted}_n - \text{Validation}_n)^2 / N$$

$$\text{SB} = (\mu(\text{Predicted}) - \mu(\text{Validation}))^2$$

$$\text{NU} = (1 - b)^2 * \sum_n (\text{Predicted}_n - \mu(\text{Predicted}))^2 / N$$

$$\text{LC} = (1 - r^2) * \sum_n (\text{Validation}_n - \mu(\text{Validation}))^2 / N$$

$$\text{MSD} = \text{SB} + \text{NU} + \text{LC}$$

$$\text{RMSD} = \sqrt{\text{MSD}}$$

$$\text{Bias} = \sqrt{\text{SB}}$$

where b refers to the slope of the least squares regression line through the plot of measured values as a function of predicted values, and r^2 is the square of the correlation. SB quantifies the proportion of the MSD related to the deviance of the least squares fit from a 1:1 relationship in the y direction (intercept). NU quantifies the proportion of the MSD related to the deviance of the least squares fit from a 1:1 relationship in the slope of the fitted line. LC quantifies the proportion of the MSD related to the scatter of the points in relation to the 1:1 line.

Geographic weighting did not appear to improve clay prediction (Table 9) but did appear to improve SOC prediction (Table 10). The appropriate weight for SOC prediction appeared to be site-specific. A single weight was required for the two ranches because models were constructed with samples from both ranches. A weight of 0.50, specifically, appeared to produce the best SOC model validation results for the BBar ranch and acceptable results for the Decker/Bales ranch (Table 10).

Table 9: Site-specific validation results for clay DRS models. Weights refer to relative weighting of samples from NRCS and ranch data sets in boosted regression tree models. Root mean square deviation (RMSD) is average prediction error for clay models. Models were constructed with 10 iterations of a 1/10 holdout using data from both ranches. The 10 held out subsets were stratified by ranch, so validation was per ranch by 1/5 holdout.

Site	NRCS Data Weight	Local Data Weight	RMSD (% clay)	MSD	SB	NU	LC	r^2
BBar	0.01	1	9.6	92	0.2	4.0	87.6	0.19
	0.25	1	9.7	95	0.0	5.0	90.0	0.17
	0.50	1	9.9	97	0.1	5.3	91.9	0.15
	0.75	1	9.6	93	0.0	4.4	88.2	0.19
	1	1	9.4	88	0.0	1.7	85.8	0.21
D/B	0.01	1	11.0	121	4.5	1.5	115.3	0.39
	0.25	1	10.4	108	3.7	2.0	101.9	0.46
	0.50	1	10.1	102	4.4	1.0	96.3	0.49
	0.75	1	10.2	103	3.8	1.3	98.1	0.48
	1	1	10.1	101	4.4	0.7	96.1	0.49

Table 10: Site-specific validation results for SOC DRS models. Weights refer to relative weighting of samples from NRCS and ranch data sets in boosted regression tree models. Root mean square error (RMSD) is average prediction error for SOC models, respectively. Models were constructed with 10 iterations of a 1/10 holdout using data from both ranches. The 10 held out subsets were stratified by ranch, so validation was per ranch by 1/5 holdout.

Site	NRCS Data Weight	Local Data Weight	RMSD (gSOC/100gsoil)	MSD	SB	NU	LC	r^2
BBar	0.01	1	0.71	0.50	0.000	0.016	0.483	0.86
	0.25	1	0.41	0.17	0.000	0.008	0.159	0.95
	0.50	1	0.34	0.12	0.003	0.000	0.113	0.97
	0.75	1	0.36	0.13	0.002	0.003	0.126	0.96
	1	1	0.37	0.14	0.003	0.000	0.134	0.96
D/B	0.01	1	0.64	0.42	0.000	0.012	0.404	0.72
	0.25	1	0.71	0.50	0.001	0.017	0.485	0.66
	0.50	1	0.75	0.56	0.001	0.005	0.554	0.61
	0.75	1	0.77	0.59	0.000	0.008	0.580	0.59
	1	1	0.76	0.58	0.000	0.006	0.571	0.60

Clay and SOC were predicted for the approximately 430 samples from the two study sites that had been scanned but not characterized. The clay prediction model used did not weight NRCS and local samples differently. The SOC model used the 0.50 weight for NRCS samples and 1.0 for local samples. SOC was adjusted to percent organic matter (OM) by multiplying by 1.72, to allow for a comparison between the lab characterization and soil survey data sources (Sparks, 1995). The soil survey organic carbon data used in this study was presented in percent organic matter.

Calculation of AWC and PAW. Mass water content (θ_m) at permanent wilting point (PWP) was estimated using average predicted OM and clay values for each sampled profile with the following equation (Decker, 1972):

$$\text{PWP} = (0.29 * \% \text{Clay}) + (0.58 * \% \text{OM}) + 2.1$$

where PWP is the θ_m at -15 bar matric potential. Permanent wilting point estimates were subtracted from estimated mass water content at field capacity taken from the soil survey. The results were multiplied by estimated bulk density (D_b) from the soil survey to estimate available water capacity (AWC) as a volume basis (θ_v) for each sampled profile. This followed the equation (Marshall et al., 1996):

$$\text{AWC} = (\text{FC} - \text{PWP}) * (D_b / D_w)$$

where FC is the θ_m at -1/3 bar matric potential, PWP is the θ_m at -15 bar matric potential, and D_b and D_w (density of water) are in g/cm^3 .

AWC estimates were adjusted to an equivalent depth (cm) of plant available water (PAW) by multiplying by the depth to root restrictive layer measured in the field, as in the following equation (Marshall et al., 1996):

$$\text{PAW} = \text{AWC} * \text{Depth}$$

where AWC is in θ_v units, and depth is in cm. The final field and lab characterization values used to develop water content models were percent clay content, percent organic matter, depth to root restrictive layer, AWC, and PAW. This data set will be referred to as the lab characterization data set to avoid confusion with the soil data set derived solely from soil survey attribute data, though the AWC and PAW variables were calculated with soil survey estimates of bulk density and water content at field capacity,.

Soil Survey-Derived Characterization Data

SSURGO digitized soil maps and associated attribute data were downloaded for each study site from the National Cooperative Soil Survey Soil Data Mart Distribution Center (<http://soildatamart.nrcs.usda.gov>). The soil survey maps were clipped to the extent of the respective ranch boundary. Soil maps were developed from the soil survey and attribute data in ARCGIS for percent clay content, soil depth to root restrictive layer, percent soil organic matter, bulk density, and water content at field capacity. Clay content, organic matter, bulk density, and water content at field capacity were calculated as weighted average values for the entire profile of the major component for each soil map unit. Soil depth to root restrictive layer was considered the depth to lithic or paralithic material or the maximum recorded soil depth for the major component of each

soil map unit. Soil characteristic values of each major component were assigned to all pixels within representative soil map unit boundaries, and the resulting map was stored in raster format.

Maps of PWP, AWC, and PAW were constructed with the same calculations used for the field and lab characterization data set. Soil survey variables were used solely in these calculations, however, and the calculations were performed with the ARCGIS raster calculator so that the calculation inputs and outputs were raster data layers. The final maps used in model development with soil survey variables were percent clay content, percent organic matter, depth to root restrictive layer, AWC, and equivalent depth PAW layers.

GIS Data Set Development

Satellite Imagery. Landsat 5 TM scene was selected from the previous growing season for each study site. Scenes were selected by proximity to the peak of growing season biomass production and cloud free quality. A scene from 1 August 2003 was selected for the BBar site and was clipped to the extent of the digitized ranch boundary. A scene from 3 August 2003 was selected and clipped for the Decker/Bales ranch.

DEM-Derived Terrain Elements. A seamless, 30-m DEM was downloaded from the USGS Seamless Data Distribution Center (<http://gisdata.usgs.net/website/Seamless>) for each ranch. Slope and aspect layers were created in ARCGIS using the spatial analyst surface function. Slope was coded in percent. The aspect layers were coded with a cosine transformation of aspect in degrees from North. Northerly aspects were positive values from 0 to 1 and southerly aspects were negative values from 0 to -1.

Data Analysis

The individual variables in the soil survey and lab characterization data sets were used as individual predictors of mass water content. This analysis did not directly relate to the main objective of the study, but did provide some insight into the relative abilities of variables from the two data sources as predictors of soil water contents at the two ranches. To address the study objective, multiple regression models were constructed to predict mass water content with Landsat bands, DEM-derived slope and aspect variables, and the lab characterization soil data set of depth, clay, OM, AWC, and PAW. Models were constructed with half of the sampled data set at each ranch. Models were independently validated with the remaining half of the data set at each ranch.

Independent model validation consisted of predicting values for the reserved data set. A least squares regression of the validation water content as a function of predicted values was constructed, and a scatterplot of the relationship containing points, regression line, and a 1:1 line (slope = 1, intercept = 0) was examined. A mean squared deviation (MSD) and root mean square deviation (RMSD) were calculated for the predicted versus observed values and the MSD was broken into components of standard bias (SB), non-unity (NU), and lack of correlation (LC) with the approach previously described for DRS model validation.

Hypothesis tests were used to test for significant differences of the means and variances of predicted and observed water content samples (Wosten et al., 2001; Feng et al., 2005). The Levene's test was used to test whether predicted and observed sample populations had significantly different variances (Feng et al., 2005). The Mann-Whitney

test of the paired predicted and validation samples was used to test whether the mean of the differences between the samples was statistically significantly different than zero (Feng et al., 2005). The Levene's test is an alternative to the F-test for equal variances and the Mann-Whitney test is an alternative to the paired t-test. Both tests are resistant to departures from normality, making them appropriate for the long-tailed distributions of the validation water content samples for both ranches. For Levene's test, an insignificant p -value at a specified level of confidence gave no evidence that the populations had unequal variances. An insignificant Mann-Whitney p -value at a specified confidence level failed to reject the null hypothesis of the test that the mean of the differences between predicted and observed water content is zero.

Models constructed with lab characterization data were compared with models previously developed with Landsat bands, DEM-derived slope and aspect, and the soil survey data set for each study site. Models for each ranch developed with the two soil data sources were compared based on adjusted R^2 values and validation results. The Levene's test for equality of variances and Mann-Whitney test of mean difference were used to test the hypothesis of the study. The differences between predicted and observed water content were compared for models developed with the two data sources. The differences between predicted and validation samples had long-tailed distributions like the validation data set for each ranch. A significant p -value for either test would have rejected the hypothesis of the study that there is no significant difference between soil water content predictions by models with soil survey variables compared to models with lab characterization variables.

Regression tree models were constructed with the same set of imagery, DEM, and lab characterization soil variables as for multiple regression models. Cross validation pruning was used to determine the number of nodes for constructed trees (Breiman et al., 1984). Models were compared with previously developed regression trees built with the soil survey data set and the same set of imagery and DEM variables. Bulk density was found to be a useful predictor in these previously constructed trees. It was not feasible to characterize field sampled soils for bulk density in this study, so the same bulk density variable from the soil survey data set was included in the lab characterization data set for regression tree analysis. Validation was completed in the same fashion as for multiple regression models, and the study hypothesis was tested similarly as well.

Results

Individual Predictors

Individual predictor variables from the lab characterization data set generally explained more of the variability in mass water content than the individual predictors from the soil survey data set at each ranch (Table 11). Only depth at the BBar ranch was significant at the 0.05 level of the soil survey variables at both ranches. The best individual predictor at the 95% confidence level appeared to be depth to root restrictive layer measured in the field. The best individual predictor characterized with DRS modeling appeared to be clay at both ranches. Clay might not have been much better of an individual predictor than OM and PAW at the Decker/Bales ranch, however. A variable's relative ability as an individual predictor is not necessarily indicative of its ability to explain variability in the response with other predictors in a model.

Table 11: Soil variables as individual predictors of average profile (100cm) mass water content. Soil survey variables were derived from soil survey maps and attribute data. Lab characterization variables were derived from characterization with diffuse reflectance spectroscopy and field characterization.

Data Source	Model	<i>p</i> -value	<i>r</i> ²
D/B Soil Survey	profile = 0.0918 + 0.0004(clay)	0.37	0.02
	profile = 0.0928 + 0.0001(depth)	0.54	0.01
	profile = 0.1324 + -0.1278(AWC)	0.22	0.03
	profile = 0.1049 + -0.0001(PAW)	0.90	0.00
	profile = .0869 + .0218(OM)	0.21	0.03
D/B field and lab	profile = .0635 + .0013(lab clay)	0.02	0.11
	profile = 0.1882 + -0.0009(field depth)	0.00	0.20
	profile = 0.1212 + -0.0386(lab AWC)	0.08	0.06
	profile = 0.1243 + -0.0005(lab PAW)	0.02	0.10
	profile = 0.0766 + 0.0137(lab OM)	0.03	0.09
BBar Soil Survey	profile = 0.0572 + 0.0015(clay)	0.11	0.06
	profile = 0.0519 + 0.0004(depth)	0.04	0.10
	profile = 0.1126 + -0.0871(AWC)	0.70	0.00
	profile = 0.0695 + 0.0011(PAW)	0.09	0.07
	profile = 0.1175 + -0.0179(OM)	0.22	0.04
BBar field and lab	profile = -0.0075 + 0.0037(lab clay)	0.01	0.16
	profile = 0.0365 + 0.0008(field depth)	0.00	0.20
	profile = 0.0905 + 0.0525(lab AWC)	0.56	0.00
	profile = 0.0754 + 0.0016(lab PAW)	0.08	0.08
	profile = 0.0942 + 0.0032(lab OM)	0.61	0.01

Multiple Regression Models

Multiple regression models were constructed with the field and lab soil data set, Landsat bands, and slope and aspect variables. Model performance was considered based on parsimonious nature, overall significance at 95% confidence, significance of individual predictor variables at 95% confidence, and adjusted R^2 values. The best model developed for each ranch with lab characterization soil data (Table 12) was selected for validation (Table 14). Each of these models contained at least one soil variable as well as a variable from the Landsat and DEM data sources. Hypothesis test p -values (Table 14) were insignificant at the 0.05 level, which failed to find significant differences between predicted and observed water contents for the lab characterization models.

The lab characterization models (Table 12) were compared to previously developed models constructed with Landsat bands, slope and aspect, and the soil survey data set (Table 13) using calibration adjusted R^2 values and validation statistics (Table 14). The best model for each ranch constructed with lab characterization data had a higher adjusted R^2 value than the best model developed with the soil variables derived solely from soil survey. The adjusted R^2 calculation is presumably not affected by the number of predictor variables in the model. Nonetheless, there was a model constructed with lab characterization data for each ranch with one fewer predictor variables that explained more of the variability in its respective calibration water content response variable than the best models constructed with soil survey variables (Tables 12 and 13). The lab characterization model predicted water content with lower average error (RMSE) than the best model constructed with soil survey variables for each ranch (Table 14).

Comparison of plots of predicted vs. validation water content (Figure 11) and MSD components (Table 14) for models from both data sources suggested that while the lab characterization data improved model predictions of water content, there was still a considerable lack of precision in the predictions. Several unacceptable maximum individual prediction errors approaching 0.10 mass water content were also observed.

Table 12: Average soil profile mass water content (profile) models using lab characterization data that were validated for the BBar and Decker/Bales (D/B) ranches. Band numbers represent Landsat TM 5 reflectance bands. Slope is percent topographic slope. Clay lab is average of percent clay content for characterized profile samples. Field depth is depth to root restrictive layer within 100cm. All models are significant at p -value < 0.05 and all variables are significant at p -value < 0.05 unless noted with [#]. Adjusted R^2 values are presented.

Model ID	Model (Lab Characterization Data)	R^2
BBar lab1	profile = .5193 + -.0040(band3) + -.0040(band4) + -.0545(slope) + .0010(field depth) + .0009(slope*band4) + -.0001(field depth*slope)	0.67
D/B lab1	profile = .5406 + -.0093(band3) + -.0011(band5) + .0001(band3 ²) + -.0032(slope) + .0018(clay lab) + -.0007(field depth)	0.64

Table 13: Average soil profile (100cm) mass water content (profile) models using soil survey data that were selected for validation for the BBar and Decker/Bales (D/B) ranches. Band numbers represent Landsat TM 5 reflectance bands. Slope is percent topographic slope. Aspect is cosine transformation of topographic aspect in degrees from North. Clay is entire profile weighted average percent clay of soil survey map unit major component. All models are significant at p -value < 0.05 and all variables are significant at p -value < 0.05 unless noted with [#]. Adjusted R^2 values are presented.

Model ID	Model (Soil Survey Data)	R^2
BBar SS 1	profile = .7840 + -.0146(band3) [#] + -.0040(band4) + -.0376(slope) + .0012(clay) + .0001(band3 ²) [#] + -.0004(slope*band3) + .0009(slope*band4)	0.64
D/B SS 1	profile = .7080 + -.0296(band3) + .0084(band4) + .0120(aspect) + .0007(clay) + .0004(band3 ²) + -.000003(band4*band3 ²) + -.0022(band6) [#]	0.43

Figure 11: Predicted versus observed water content graphs for: (A) BBar soil survey variable regression model, (B) BBar lab and field characterization regression model, (C) Decker/Bales soil survey variable regression model, (D) Decker/Bales lab and field characterization regression model. Solid line represents least squares regression of predicted versus observed water content. Dashed line represents 1:1 line ($y = 0 + 1x$).

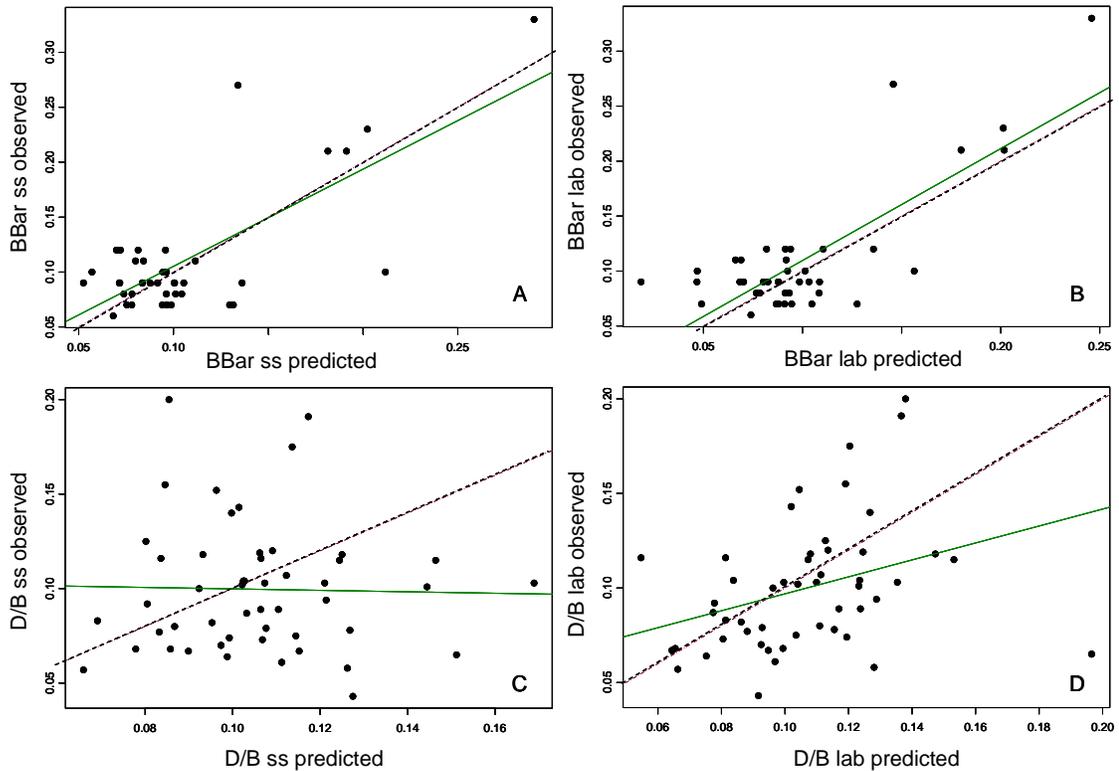


Table 14: Validation statistics for models in Tables 12 and 13. Models were validated with half the data set at each ranch (50 validation samples for Decker/Bales (D/B) and 41 for BBar ranch). Levene and Mann-Whitney statistics are p -values.

Model	RMSD	Bias	MSD	SB	NU	LC	r^2	Levene	Mann-Whitney
BBar lab 1	0.036	0.010	0.001	0.0001	0.0000	0.0012	0.62	0.77	0.14
BBar ss 1	0.039	0.005	0.002	0.0000	0.0000	0.0015	0.54	0.81	0.35
D/B lab 1	0.035	0.007	0.001	0.0000	0.0002	0.0010	0.11	0.11	0.18
D/B ss 1	0.040	0.006	0.002	0.0000	0.0005	0.0011	0.00	0.00	0.14

The mean difference between predicted and observed water content differences for multiple regression models constructed with the two soil data sources was suggested to be significantly different than zero at 95% confidence for the BBar ranch but not the Decker/Bales ranch with the Mann-Whitney test (Table 15). This rejected the hypothesis of the study for the BBar ranch and failed to reject it for the Decker/Bales site, suggesting there was a significant difference between water content predictions by multiple regression models constructed with the two data sources at BBar but not Decker/Bales.

Table 15: Hypothesis test p -values comparing the differences between predicted and observed water content values for multiple regression models from the two data sources for each ranch.

Comparison	Levene	Mann-Whitney
BBar SS (pred. – observ.) and BBar lab (pred. – observ.)	0.46	0.000
D/B SS (pred. – observ.) and D/B lab (pred. – observ.)	0.46	0.942

Regression Tree Models

Validation average prediction errors (RMSD) and the MSD components (Table 16 and Figure 12) showed that the regression tree model constructed with characterization soil variables for the BBar ranch (Figure 13) was not substantially better than the tree constructed with soil survey variables (Figure 14) for the same site. Interestingly, the Mann-Whitney test suggested the mean difference between predicted and observed water contents could be zero for the lab characterization tree but not for the soil survey tree at this site at 95% confidence. The average prediction error (Table 16) appeared to be slightly smaller for the Decker/Bales regression tree constructed with lab characterization soil variables (Figure 15) compared to the soil survey regression tree (Figure 16) for that

site. The Levene's test suggested, however, that the variances of predicted and observed water content were unequal for the Decker/Bales lab characterization tree.

The predicted and observed water content differences for regression tree models constructed with the two soil data sources were suggested to be significantly different for the BBar ranch but not the Decker/Bales ranch by the Levene and Mann-Whitney tests (Table 17). This rejected the null hypothesis of the study for the BBar ranch and failed to reject it for the Decker/Bales site, suggesting there were significant differences between water content predictions by regression tree models constructed with the two data sources at BBar but not Decker/Bales. Plots of predicted water content versus validation water content suggested, however, that tree models with both types of soil variables failed to predict water content with much accuracy or precision at both ranches (Figure 12).

Table 16: Validation statistics for regression tree models constructed with soil survey soil variables and lab characterization soil variables.

Model	RMSD	Bias	MSD	SB	NU	LC	r^2	Levene	Mann-Whitney
BBar lab tree	0.054	-0.004	0.003	0.0000	0.0003	0.0025	0.21	0.53	0.73
BBar ss tree	0.055	-0.010	0.003	0.0001	0.0004	0.0025	0.22	0.42	0.01
D/B lab tree	0.034	0.003	0.001	0.0000	0.0001	0.0011	0.08	0.01	0.24
D/B ss tree	0.048	0.003	0.002	0.0000	0.0012	0.0011	0.03	0.13	0.52

Figure 12: Predicted versus observed water content graphs for: (A) BBar soil survey variable regression tree model, (B) BBar lab and field characterization regression tree model, (C) Decker/Bales soil survey variable regression tree model, (D) Decker/Bales lab and field characterization regression tree model. Solid line represents simple linear regression of observed water content as a function of predicted water content for specific model. Dashed line represents $y = 0 + 1x$.

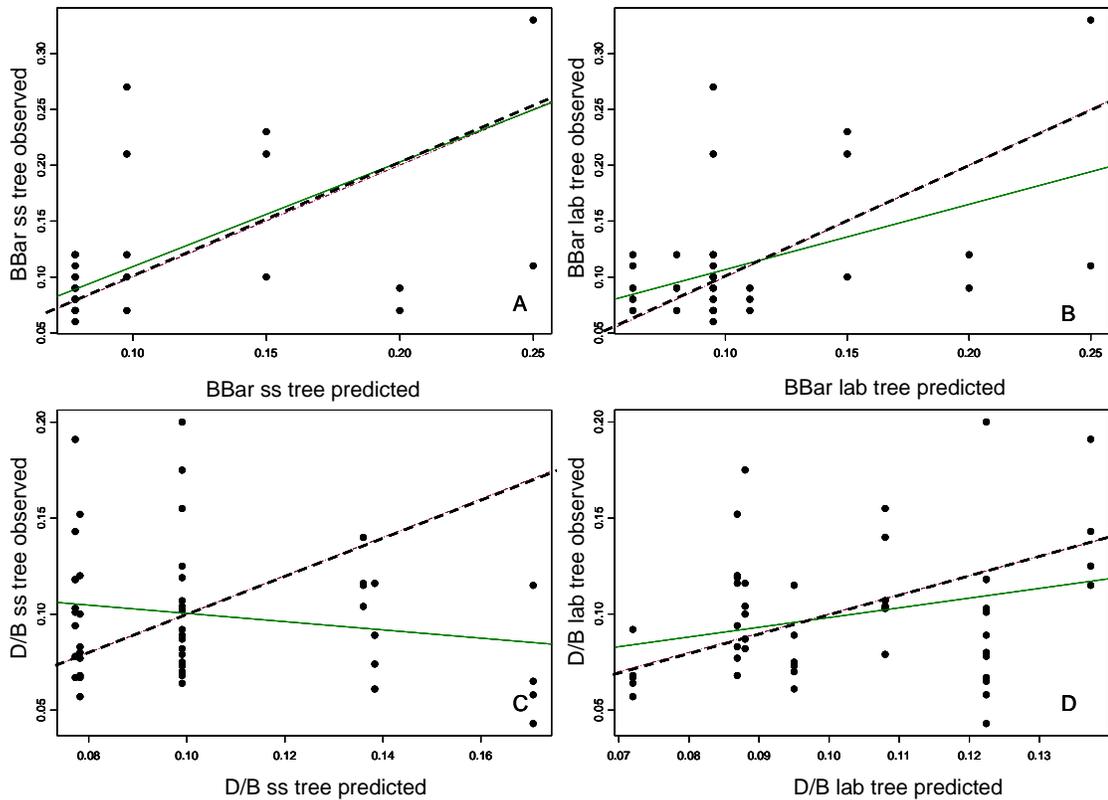


Table 17: Mann-Whitney paired comparison of differences between predicted and observed water content values for regression tree models from the two data sources for each ranch.

Comparison	Levene	Mann-Whitney
BBar SS tree (pred. – observ.) and BBar lab tree (pred. – observ.)	0.04	0.005
D/B SS tree (pred. – observ.) and D/B lab tree (pred. – observ.)	0.08	0.954

Figure 13: BBar lab characterization regression tree. Terminal node values are mass water content ($\text{g}_{\text{water}}/\text{g}_{\text{soil}}$).

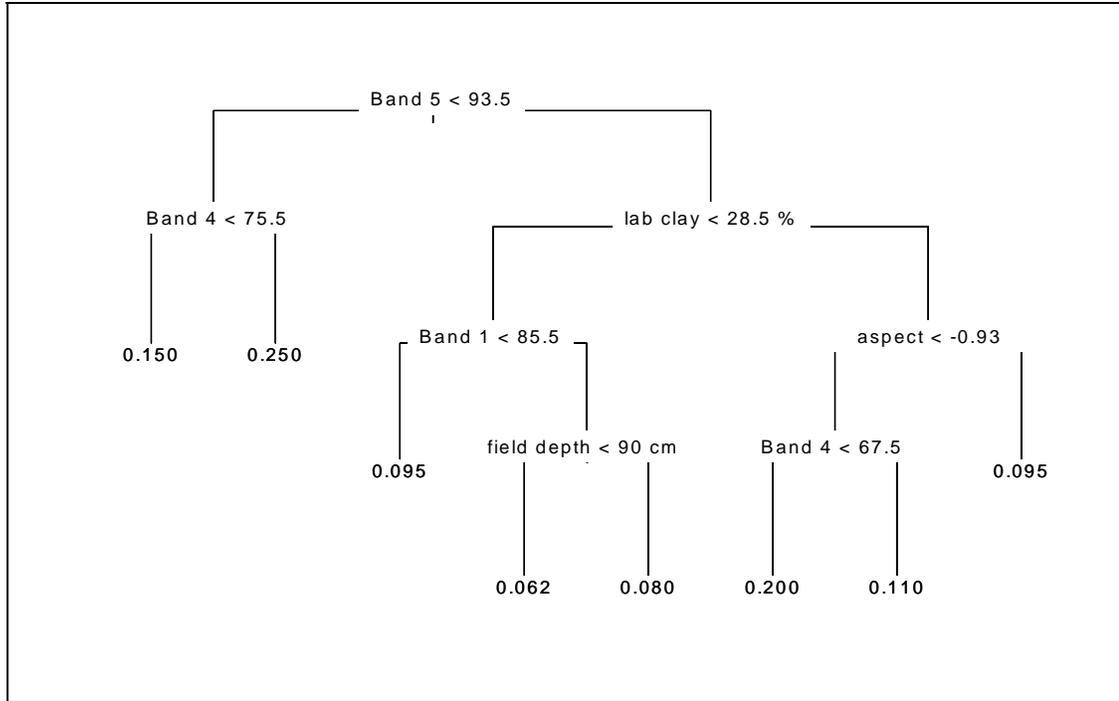


Figure 14: BBar soil survey regression tree. Terminal node values are mass water content ($\text{g}_{\text{water}}/\text{g}_{\text{soil}}$).

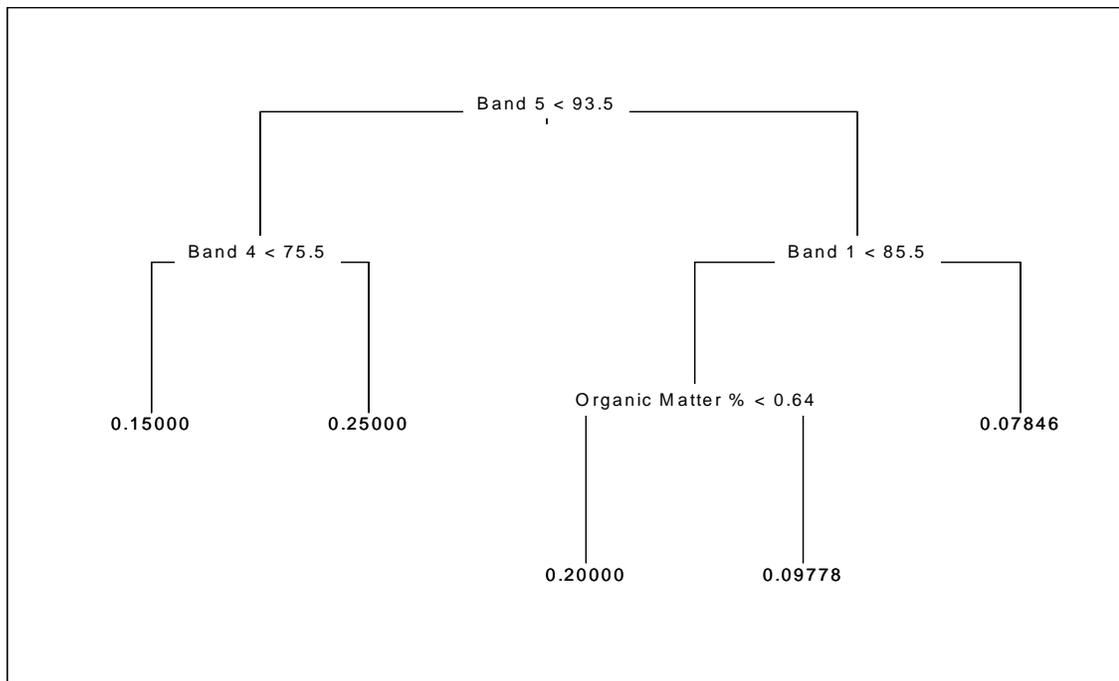


Figure 15: Decker/Bales lab characterization regression tree. Terminal node values are mass water content ($g_{\text{water}}/g_{\text{soil}}$).

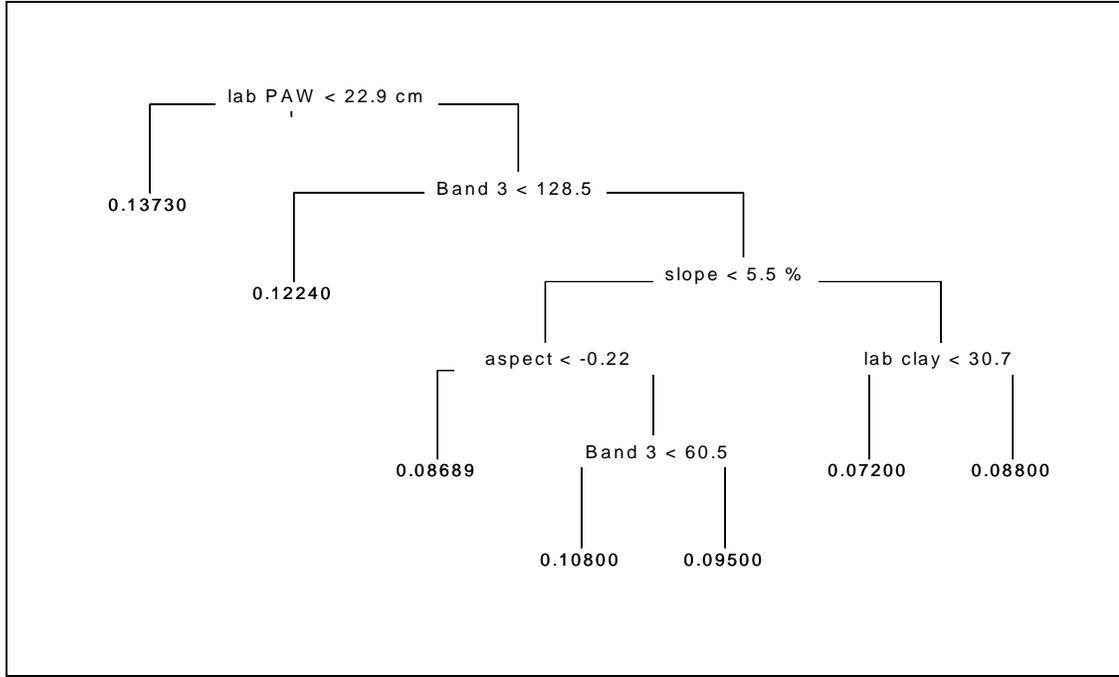
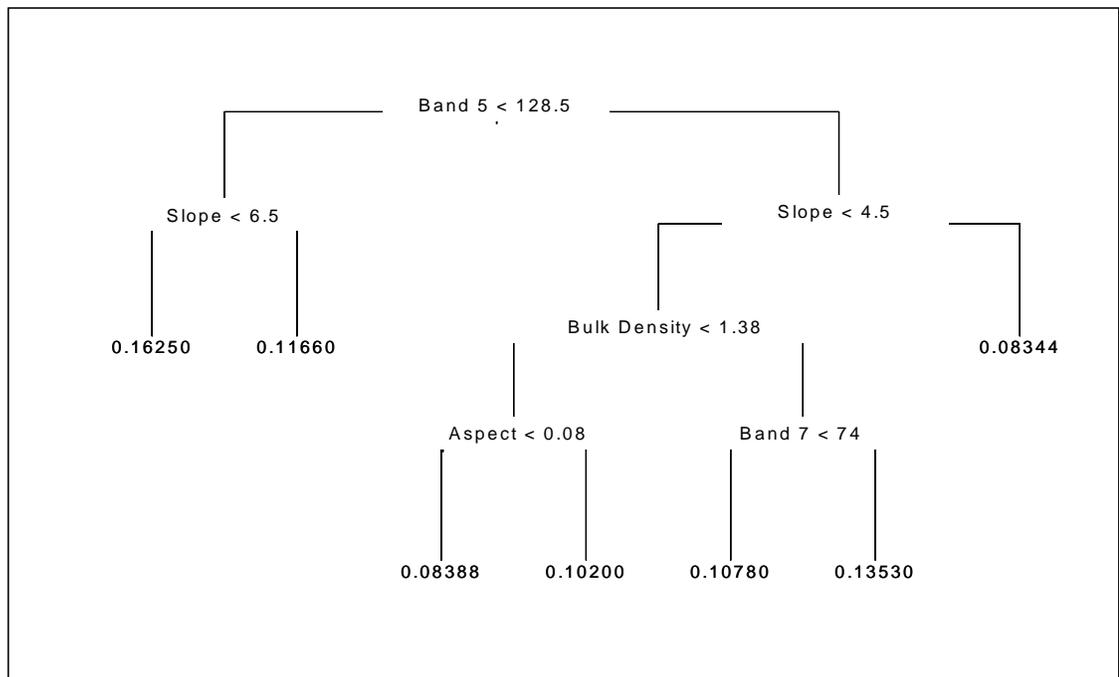


Figure 16: Decker/Bales soil survey regression tree. Terminal node values are mass water content ($g_{\text{water}}/g_{\text{soil}}$).



Discussion

Statistically significant differences between soil water content predictions for models with predictors derived from Order 2 soil surveys compared to site-specific characterization data were detected for one of two study sites. Differences, when detected, were slight. The use of soil predictor variables derived from field and lab characterization in empirical soil water content models appeared to slightly reduce average prediction error (RMSD) compared to similarly constructed models with soil survey-derived variables. This might suggest that the use of site-specific soils data from field and lab characterization improved the general prediction accuracy of the soil water content models. This is supported by larger r^2 values for the predicted vs. observed relationship for models developed with site-specific soils data compared to soil survey data, at each respective ranch. Statistically significant differences between soil water content predictions by models constructed with soil survey predictor variables and similarly constructed models with field and lab characterization predictor variables were found for both multiple regression models and regression tree models at the BBar ranch but not the Decker/Bales ranch. These hypothesis test results in conjunction with average prediction error and MSD component results support the conclusion that model predictions were significantly better for the models developed with site-specific data than the models developed with soil survey data at the BBar ranch. No such conclusion can be made for the Decker/Bales ranch. Models constructed with both sources of soil variables were found, however, to predict spring soil water content with a lack of

precision. This was made evident for all models by the large proportion of the MSD due to the lack of correlation (increased scatter) of observed water content as a function of predicted water content along a 1:1 relationship.

The absence of a large disparity between predictive accuracy of models derived from the two soil data sources was unexpected. Soils data, in general, were expected to provide predictive ability for soil water content in addition to that provided by the Landsat and DEM data sources. The soil survey data was expected to provide less precise estimates of soil characteristics at the soil water content sample locations than the characterization data. It was in turn expected that the soil survey data would contribute to less precise predictions of soil water contents than the sample location specific characterization data. Predictions by models developed with the two soils data sources at both ranches, however, appeared to predict soil water contents with a similar lack of precision.

The dry conditions encountered during the sampling period might help explain some of the limitations of both soil data sources as soil water content predictors. The mass water content measured at both ranches was predominantly in the range of 0.05 to 0.15, with a handful of wetter samples at each site. Soil water content is difficult to model at a landscape scale, and difficulties are generally confounded with increasing aridity (Grayson et al., 1997; Western et al., 1999). This is in part because the spatial distribution of soil water at a landscape scale is generally increasingly random and decreasingly influenced by topography as conditions become drier (Western, 1999, Ridofli et al., 2003). Topography might not be expected to substantially influence the

distribution of soil water even during average or wetter years in semi-arid, Montana environments (Landon, 1995; Kozar, 2002). Soil characteristics such as texture, however, also influence the distribution of soil water across a landscape and are expected to act as larger scale, local controls on soil water content in more arid states and systems (Grayson et al., 1997). Multiple regression model interaction terms and regression tree nodes in this study suggested that water content predictions by soil variables such as texture and depth tended to be influenced by topography, a finding as well in other studies (Pachepsky et al., 2001, Ridofli et al., 2003). Soil water content prediction by soil characterization variables, therefore, might be expected to become more limited as the relationship between soil water distribution and topography becomes more stochastic with increasing aridity.

The most common soil predictor variable in models that used field and lab characterization data in this study was soil depth. This was the easiest soil characteristic to measure, being simply the depth at which root restrictive lithic or paralithic material was found, if at all, within 100 cm of the soil surface. The BBar site regression tree constructed with field and lab characterization data suggested that shallower soils were slightly drier (Figure 13). This was corroborated by the positive coefficient for the field depth predictor in the BBar lab characterization multiple regression model, which suggested that average profile water content increased with an increase in depth (Table 12). Interestingly, the coefficient for the field depth predictor in the Decker/Bales lab characterization multiple regression model suggested an opposite relationship (Table 12).

Different relationships between depth and soil texture might have existed between the two study sites.

Soils with higher clay content tend to have a higher proportion of smaller pores, and therefore tend to hold water much more tightly (Brady, 1990). Surface area might be a more important control on soil water content than macroporosity in drier conditions, and therefore clay content might be a better predictor of soil water content than soil variables such as sand content and bulk density (Wosten et al., 2001). Regression tree nodes and multiple regression model coefficients for the lab characterization clay variable at both ranches suggested that soils with more clay were generally wetter. The clay variable modeled with DRS appeared to be a better individual predictor of soil water content at both ranches than the respective soil survey clay variable (Table 11). This might have been expected considering that the characterization clay variable was an estimate of clay content more specific to the soil water content sample locations than was the soil survey clay variable.

It is difficult to say, however, how different model predictions with the soil survey clay variable were to model predictions with the DRS characterization clay variable at either ranch. The DRS characterization model used to predict clay content with first derivative soil reflectance for this project was shown with validation to predict with an average error of 9 to 10% clay content and to suffer from a lack of correlation between predicted and measured values that limited prediction precision (Table 9). The inclusion of relatively easily measured auxiliary predictor variables such as sand content and pH, has been suggested to improve DRS first derivative model predictions (Brown et

al., in press b). More involved DRS modeling and characterization methods such as the use of auxiliary predictors could be used to produce a more accurate characterization data set. Such a data set could be used in a comparison with soil survey-derived variables in the future. This might provide a more accurate reference data with which to assess the suitability of soil survey-derived variables as a site-specific data source for soil water content modeling.

Conclusion

This study found that publicly available Order 2 soil surveys provided predictive ability for modeling soil water content at a landscape scale that was significantly different than more sample point specific field and lab characterization data at one of the two study sites. Differences, when detected however, were slight and probably not substantial for practical purposes. Soil water content was problematic to model with both sources of soils data, even with additional Landsat imagery and DEM-derived predictor variables. Models explained a limited amount of variability in calibration water content data sets and had almost no predictive ability for one study site and limited accuracy and poor precision at the other site. Very dry soil water states during the sampling periods possibly contributed to the limited explanatory ability of soil variables. Future work on assessing the suitability of soil survey data as a site-specific data source for predictive modeling might benefit from comparison with a more accurate clay content characterization predictor variable. Such a variable might be developed with auxiliary predictor variables in DRS modeling. DRS characterized clay content and easily

measured soil depth to root restrictive layer, particularly in conjunction with topographic slope, hinted at important soil water content prediction ability. These relationships appear especially substantial considering the dry conditions of the study period.

References Cited

- Ben-Dor, E., Irons, J.R., Epema, G.F. (1999). Soil reflectance. In: N. Rencz (Editor), Remote sensing for the earth sciences: manual of remote sensing. John Wiley and Sons: New York.
- Brady, N. (1990). The Nature and Properties of Soils. Macmillan: New York.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984). Classification and Regression Trees. Wadsworth International Group, Belmont, CA.
- Brown, D.J., Bricklemeyer, R.S., Miller, P.R. (in press a). Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. Geoderma.
- Brown, D.J., Shepherd, K.D., Walsh, M.G., Mays, D.M., Reinsch, T.G. (in press b). Global soil characterization with a VNIR diffuse reflectance library. Geoderma.
- Clark, R.N. (1999). Spectroscopy of rocks and minerals, and principles of spectroscopy. In: N. Rencz (Editor), Remote sensing for the earth sciences: manual of remote sensing. John Wiley and Sons: New York.
- Decker, G.L. (1972). Automatic Retrieval and Analysis of Soil Characterization Data. PhD. dissertation. Montana State University. Bozeman, MT.
- Dunn, B.W., Beecher, H.G., Batten, G.D., Ciavarella, S. (2002). The potential of near-infrared reflectance spectroscopy for soil analysis – a case study from the Riverine Plain of south-eastern Australia. Australian Journal of Experimental Agriculture. 42:607-614.
- Feng, C-X. J., Yu, Z-G. S., Wang, J-H. J. (2005). Validation and data splitting in predictive regression modeling of honing surface roughness data. International journal of production research. 43:1555-1571.
- Gauch, H.G., Hwang, J.T., Fivk, G.W. (2003). Model evaluation by comparison of model-based predictions and measured values. Agronomy Journal. 95:1442-1446.
- Gee, G.W., Bauder, J.W. (1986). Particle Size Analysis. In: Methods of Soil Analysis, Part 1. Physical and Mineralogical Methods. Soil Science Society of America: Madison, Wisconsin.

- Grayson, R.B., Western, A.W., Chiew, F.H.S. (1997). Preferred states in spatial soil moisture: local and nonlocal controls. *Water Resources Research*. 33: 2897-2908.
- Hudson, Berman. (1992). The soil survey as a paradigm based science. *Soil Science Society of America Journal*. 56:836-841.
- Hunt, G.R. (1989). Spectroscopic properties of rocks and minerals. In: R.S. Carmichael (Editor.), *Practical handbook of physical properties of rocks and minerals*. CRC Press: Boca Raton, Florida.
- Kozar, B.J. (2002). Predicting soil water distribution using topographic models within four Montana farm fields. (thesis). Montana State University: Bozeman, Montana.
- Landon, M.A. (1995). Soil and terrain attributes for evaluation of leaching in a Montana farm field. (thesis). Montana State University: Bozeman, Montana.
- Marshall, T.J., Holmes, J.W., Rose, C.W. (1996). *Soil Physics*. Press Syndicate of the University of Cambridge: New York, New York.
- McCarty, G.W., Reeves III, J.B., Reeves, V.B., Follett, R.F., Kimble, J.M. (2002). Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Science Society of America Journal*. 66:640-646.
- Montagne, C., Munn, L.C. Nielsen, G.A., Rogers, J.W., Hunter, H.E. (1982). *Soils of Montana*. Montana Agricultural Experiment Station, Montana State University. Bozeman, MT.
- Neff, E.L., Wight, J.R. (1977). Overwinter soil water recharge and herbage production as influenced by contour furrowing on Eastern Montana rangelands. *Journal of Range Management*. 30:193-195.
- Pachepsky, Y.A., Timlin, D.J., Rawls, W.J. (2001). Soil water retention as related to topographic variables. *Soil Science Society of America Journal*. 65:1787-1795.
- Pfaffenberger, R.C., Patterson, J.H. (1981). *Statistical Methods*. Richard D. Irwin, Inc.: Homewood, Illinois.
- Ridolfi, L., D'Odorico, P., Porporato, A., Rodriguez-Iturbe, I. (2003). Stochastic soil moisture dynamics along a hillslope. *Journal of Hydrology*. 272:264-275.
- Robertson, G. P., Klingensmith, K.M., Klug, M.J., Paul, E.A., Crum, J.R., Ellis, B.G. (1997). Soil resources, microbial activity, and primary production across an agricultural system. *Ecological Applications*. 7:158-170.

- Rogler, G., Haas, H.J. (1946). Range production as related to soil moisture and precipitation on the Northern Great Plains. *Journal of the American Society of Agronomy*. 378-389.
- Shepherd, K.D., Walsh, M.G. (2002). Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of America Journal*. 66:988-998.
- Sherrod, L.A. (2002). Inorganic carbon analysis by modified pressure calcimeter method. *Soil Science Society of America Journal*. 66:299-305.
- Soil Survey Staff. (2004). Soil Survey (SSURGO) database for Sweet Grass County Area, Montana. United States Department of Agriculture, Natural Resources Conservation Service: Fort Worth, Texas. <http://soildatamart.nrcs.usda.gov>.
- Soil Survey Staff. (1971). Soil Survey Powder River Area Montana. United States Department of Agriculture, Soil Conservation Service: Washington, D.C.
- Soil Survey Staff. (1996). Soil Survey Laboratory Methods – Soil Survey Investigations Report Number 42. United States Department of Agriculture, Natural Resources Conservation Service, National Soil Survey Center: Washington D.C.
- Sparks, D. (1995). *Environmental Soil Chemistry*. Academic Press: San Diego, CA.
- Veseth, R., Montagne, C. (1980). Geologic Parent Materials of Montana Soils. Bulletin 721. Montana Agricultural Experiment Station and United States Department of Agriculture, Soil Conservation Service: Bozeman, MT.
- Western, A.W., Grayson, R.B., Bloschl, G., Willgoose, G.R., McMahon, T.A. (1999). Observed spatial organization of soil moisture and its relation to terrain indices. *Water Resources Research*. 35:797-810.
- Wosten, J.H., Pachepsky, Y.A., Rawls, W.J. (2001). Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*. 251:123-150.

CHAPTER 5

CONCLUSIONS

The development of relatively simple, site-specific models to accurately predict spring soil water contents with publicly available GIS data sources proved a difficult and problematic task. The relationships between Landsat imagery, USGS DEM-derived topographic slope and aspect, National Cooperative Soil Survey data, and rangeland spring soil water contents might require more complex modeling approaches or larger sample sizes. The need for an increase in modeling complexity or sample sizes would defeat the overall goal of this project, which was to develop and validate a system for ranchers to collect a limited size soil water content data set, build site-specific regression models based on the data set, and construct soil water content maps based on the models.

The use of a small ($n = 20-30$ samples) field collected calibration data set to parameterize site-specific models did appear to hold some promise. This approach might side step the complex global calibration issues confronted when using remotely sensed imagery in modeling efforts. The site-specific modeling approach would only be useful, however, if better predictive empirical soil water content models could be developed. This might require alternatives to the data sources used in this project.

The results of this study did not conclusively suggest that the soil survey data source could be substantially improved upon by an alternative more site-specific soil data source. The soil survey data source contributed variables to models that predicted soil water content without a substantial difference in accuracy compared to models developed

with field and lab characterization data, in some cases. The unexpectedly low predictive ability of site-specific characterization data in this study might be attributed to the aridity of the study period and the semi-arid climate of the study sites. The predictive ability of the soil survey-derived variables might, presumably, have been similarly afflicted. The somewhat limited accuracy of the DRS characterization models developed for clay content prediction was another possible explanation of the relatively low predictive ability in the case of the site-specific characterization data.

There are other predictive tools and data sources that, though not available in an easily usable form today, will likely become available to ranchers in the future.

Remotely sensed radar data is an example of a data source that is complex to implement in a precision agriculture setting today. It might have future potential in surface soil water content prediction applications, however, and might complement several of the imagery, terrain, and soil data sources shown to have soil water content explanatory ability in this project. Similarly, this study did not use grazing as a predictor variable in soil water content models. Grazing is generally expected to influence soil water storage in rangelands, and spatial data sets could be developed to represent a rancher's description of the timing and intensity of grazing at a pasture scale.

Precision agriculture is a relatively young science and precision range management is even younger, with relatively few tested applications in use by ranchers today. This study constitutes another step in the direction of providing ranchers with methods for balancing traditional expert knowledge with the insights that geospatial tools can provide in making informed range management decisions.