EXAMINATION OF HIGH RESOLUTION RAINFALL PRODUCTS AND SATELLITE GREENNESS INDICES FOR ESTIMATING PATCH AND LANDSCAPE FORAGE BIOMASS

A Dissertation

by

JAY PETER ANGERER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2008

Major Subject: Rangeland Ecology and Management

EXAMINATION OF HIGH RESOLUTION RAINFALL PRODUCTS AND SATELLITE GREENNESS INDICES FOR ESTIMATING PATCH AND LANDSCAPE FORAGE BIOMASS

A Dissertation

by

JAY PETER ANGERER

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	X. Ben Wu
Committee Members,	Bradford Wilcox
	William Grant
	Raghavan Srinivasan
Head of Department,	Steven Whisenant

May 2008

Major Subject: Rangeland Ecology and Management

ABSTRACT

Examination of High Resolution Rainfall Products and Satellite Greenness Indices for Estimating Patch and Landscape Forage Biomass. (May 2008)

Jay Peter Angerer, B.S., Texas Tech University;

M.S., Texas A&M University

Chair of Advisory Committee: Dr. X. Ben Wu

Assessment of vegetation productivity on rangelands is needed to assist in timely decision making with regard to management of the livestock enterprise as well as to protect the natural resource. Characterization of the vegetation resource over large landscapes can be time consuming, expensive and almost impossible to do on a near real-time basis. The overarching goal of this study was to examine available technologies for implementing near real-time systems to monitor forage biomass available to livestock on a given landscape. The primary objectives were to examine the ability of the Climate Prediction Center Morphing Product (CMORPH) and Next Generation Weather Radar (NEXRAD) rainfall products to detect and estimate rainfall at semi-arid sites in West Texas, to verify the ability of a simulation model (PHYGROW) to predict herbaceous biomass at selected sites (patches) in a semi-arid landscape using NEXRAD rainfall, and to examine the feasibility of using cokriging for integrating simulation model output and satellite greenness imagery (NDVI) for producing landscape maps of forage biomass in Mongolia's Gobi region.

The comparison of the NEXRAD and CMORPH rainfall products to gage collected rainfall revealed that NEXRAD outperformed the CMORPH rainfall with lower estimation bias, lower variability, and higher estimation efficiency. When NEXRAD was used as a driving variable in PHYGROW simulations that were calibrated using gage measured rainfall, model performance for estimating forage biomass was generally poor when compared to biomass measurements at the sites.

However, when model simulations were calibrated using NEXRAD rainfall, performance in estimating biomass was substantially better. A suggested reason for the improved performance was that calibration with NEXRAD adjusted the model for the general over or underestimation of rainfall by the NEXRAD product. In the Gobi region of Mongolia, the PHYGROW model performed well in predicting forage biomass except for overestimations in the Forest Steppe zone. Cross-validation revealed that cokriging of PHYGROW output with NDVI as a covariate performed well during the majority of the growing season. Cokriging of simulation model output and NDVI appears to hold promise for producing landscape maps of forage biomass as part of near real-time forage monitoring systems.

ACKNOWLEDGEMENTS

First of all, I would like to sincerely thank Dr. Ben Wu for serving as my committee chairman and advisor. Dr. Wu graciously took over as chairman after the passing of my original advisor, Dr. Jerry Stuth. His willingness to accept me as a student very late in my program, as well as his assistance and patience in completing the dissertation is greatly appreciated. I would also like to thank him for introducing me to geostatistics. The lectures and project in his Landscape Analysis class initiated my thoughts about how to combine simulation modeling and satellite imagery for mapping of forage biomass, and provided the basis for the study reported in Chapter IV of this dissertation.

I would also like to thank Dr. Bradford Wilcox, Dr. William Grant, and Dr. Raghavan Srinivasan for serving on my committee and their patience during the time it took me to complete my program. Dr. Wilcox, through his suggestion that I study Dr. Keith Beven's hydrological modeling research, provided me with a new perspective on simulation modeling. Dr. Grant's class was extremely helpful in understanding the simulation modeling process. Dr. Srinivasan's support and encouragement over the years is greatly appreciated.

I am very grateful for the willingness of the Kelleher and Hayne families to let us conduct research on the beautiful Catto-Gage Ranch and for allowing us to install two weather stations on their property. They also provided housing and vehicles during our visits to the ranch, and without their support, this research would not have been possible. I would also like to sincerely thank Don Keeling, senior ranch manager for the Catto-Gage, who was instrumental in helping us get the research started at the ranch and for his understanding of how this research could be beneficial for livestock production. Catto-Gage ranch manager, Brent Charlesworth, provided much needed assistance and logistical support, and made our stays at the ranch a pleasant experience. I would like to extend thanks to John Rizzo for coordinating access to the ranch with the family and for

his support during the past 4 years. The assistance of Frank Galvan, Kelly Guynes, and other ranch personnel is also appreciated.

The team at the Center for Natural Resources Information Technology is outstanding, and they provided valuable expertise and assistance in many aspects for these studies. The assistance and support from Dr. Richard Conner and Mr. Wayne Hamilton are greatly appreciated. Dr. Jimmy Wu built the scripts and platform for extracting the NEXRAD and CMORPH weather data. Jim Bucher assisted in transferring the plant and soil data into the PHYGROW modeling interface. Tim Brown has been a great friend over the years, helping with soil parameterization and data collection. Jason Jones, Bryce Thomas, and Stephen Prince assisted with data collection and weighing samples. Jennifer Jacobs and Dr. Balaji Narasimhan of the Spatial Sciences Laboratory provided advice and information on NEXRAD data. I appreciate each of these individuals' contributions in helping me complete this work.

The Gobi Forage team has been a pleasure to work with and was instrumental in the completion of the forage mapping study in Mongolia. Dr. Doug Tolleson has been a good friend and colleague, and made the trips to Mongolia even more enjoyable. Sean Granville-Ross' management of the project in Mongolia and the hospitality that he and his wife Charlie provided during my trips to Mongolia is greatly appreciated. Dr. Dennis Sheehy assisted with data collection and developing sampling strategies, and I thoroughly enjoyed our many conversations about Mongolia's rangelands. Dr. Tsogoo Damdin, Dr. Udval Gombosuren, Bolor-Erdene Lhamsuren, Urgamal Magsar, Narangerel Davaasuren, and Tsolmon Namkhainyam put in many long hours in both the office and the field to collect data for the PHYGROW modeling and landscape mapping. Without their help this work would not have been possible. Erka and Enkhbat did much of the driving to many of the remote sampling locations, and I appreciate their dedication and assistance. Steve Zimmerman, Tim Stewart, Peter Ormel, and many others from Mercy Corps Mongolia provided valuable assistance throughout the study.

I am very grateful to the Texas Water Resources Institute (TWRI) who provided funding to purchase the weather stations. I would also like to thank the The United States Agency for International Development (USAID) and the Global Livestock Collaborative Research Support Program (GL-CRSP) for their funding and support of the study in Mongolia.

I am also deeply indebted to Dr. Jerry Stuth who was instrumental in convincing me to pursue a Ph.D. and for giving me the opportunity to work on important rangeland issues throughout the world. Jerry had a limitless imagination and was always willing to listen to my ideas, regardless of how harebrained they might be. His untimely death in April 2006 brought sadness to all of us who knew him well, and his contributions and support will be missed for a very long time.

I greatly appreciate the support and encouragement that my parents, Jim and Vivian, have provided over the years. I am also grateful for the encouragement and support from my brothers (Chris, Doug and Nick), my sister (Barbara) and their families. I also appreciate the encouragement and kindness from my wife's family (Don, Martha, Gretcha, and Rob).

Lastly, I would like to express my gratitude and love for my beautiful wife SuZan. She has endured many long hours without me during my data collection trips to West Texas and Mongolia, as well as all the time I have spent studying and writing to complete my Ph.D. I will be forever grateful for her love, patience, understanding, encouragement and support that she has given me during this long process.

TABLE OF CONTENTS

CHAPTER Page

Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I
INTRODUCTION

The ability to characterize the productivity of vegetation over large landscapes can be an important component in the assessment of drought impacts, natural resource management options, environmental degradation, and economic impacts of changing technologies. However, the time and resources required to conduct accurate assessments of vegetation productivity over large landscapes are prohibitive. Another complicating factor is that decisions regarding livestock stocking/destocking may require near real-time information, especially in the face of drought. Vegetation productivity assessment is almost impossible to conduct over large land areas on a near real-time basis, thus the information needed for livestock decision making is not always available when it is needed most. The inability to make stock/destock decisions at critical times could lead to vegetation overuse, which in turn, could lead to thresholds being crossed that that move the vegetation and soil resources on a trajectory toward degradation (Trimble and Mendel 1995; Evans 1998).

Improvements in computing power and capacity, along with near real-time production of climate data and remote sensing imagery offer the opportunity to develop near real-time systems for monitoring vegetation on rangelands. In the past, much of the climate data available to users was from weather stations generally located in cities and towns. However, these data are not always reflective of the climate in the more remote, rangeland areas. The emergence of technology for estimating precipitation using techniques such as cold cloud temperatures (e.g., Herman et al. 1997; Xie and Arkin 1998) and Doppler radar (e.g., Whiton et al. 1998a,b) have made spatially explicit climate data available in these remote areas, thus increasing their potential for use in near real-time systems.

This dissertation follows the style and format of Landscape Ecology.

Improved computing power and capacity has also increased the use of simulation modeling for agriculture systems, including rangelands. The application of these models on rangelands includes simulations for hydrology, soil erosion, plant growth, or combinations of these (Bouraoui and Wolfe 1990). Models that have been used to predict plant biomass on rangelands include the Simulation of Production and Utilization of Rangelands (SPUR) (Wight and Skiles 1987; Carlson and Thurow 1992), Ekalaka Rangeland Hydrology and Yield Model (ERHYM-II) (Wight and Neff 1983), Water Erosion Prediction Project (WEPP) (Flanagan and Nearing 1995), Agricultural Land Management Alternatives with Numerical Assessment Criteria (ALMANAC) (Kiniry et al. 2002), Ecological Dynamics Simulation Model (EDYS) (Childress et al. 2002), and the Phytomass Growth Simulator Model (PHYGROW) (Stuth et al. 2003a). Of these models, only the PHYGROW model has been used for near real-time forage monitoring (Stuth et al. 2003b; Ryan 2005; Stuth et al. 2005). Little information is available on how these rangeland models perform using the currently available high resolution rainfall products and how these compare to performance using measured rainfall at a given site.

A limitation of many forage simulation models is that most provide simulation output for a specific point. Ideally, one would want to simulate as many points (or sites) as possible to represent a region or landscape, especially for the determination of biomass for livestock decision making. However, the amount of effort and cost for model parameterization on a large number of points can be prohibitive. An alternative approach is to conduct simulations for a number of points and then use geostatistical interpolation methods, such as kriging or cokriging, to create surface maps of simulation output for a region or landscape (Stuth et al. 2003b; Stuth et al. 2005). These surface maps can then be used represent spatially explicit vegetation production allowing users to monitor conditions and to determine livestock stocking rates.

As an interpolation method, ordinary kriging can provide estimates for unsampled points by using the weighted linear average of the available samples (Rossi et

al. 1994). Cokriging offers additional advantages over ordinary kriging in that it involves the use of a secondary variable (covariate) that is cross-correlated with the primary or sample variable of interest. The secondary variable is usually sampled more frequently and regularly (Isaaks and Srivastava 1989), thus allowing estimation of unsampled points using both variables. Remote sensing imagery provides a dense and exhaustive data set that can serve as a secondary variable for geostatistical interpolation given a correlation between the primary and secondary variable (Dungan 1998). Satellite derived vegetation indices, most notably the Normalized Difference Vegetation Index (NDVI), have been found to be correlated to vegetation productivity (Tucker et al. 1985; Tucker and Sellers 1986; Wylie et al. 1991; Sannier et al. 2002; Al-Bakri and Taylor 2003; Schino et al. 2003; Pineiro et al. 2006; Wessels et al. 2006), thus making these products suitable for use as a secondary variable in geostatistical analysis. On rangelands, NDVI has generally been used as a predictor variable for vegetation biomass (e.g., Tucker and Sellers 1986; Al-Bakri and Taylor 2003; Frank and Karn 2003), but has not been extensively used as a covariate in geostatistical interpolation of biomass. Vegetation indices produced through the National Oceanic and Atmospheric Administration Advanced Very High Resolution Radiometer (NOAA-AVHRR) and the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data streams have high temporal frequency (daily acquisition with 10 to 16 day compositing intervals) making them attractive for use in near real-time systems. The NOAA-AVHRR data has a relatively long historical record (1981 to present), global coverage, and a resolution of 1 km. This data set has been a major component of early warning systems for Africa (Hutchinson 1991; Rowland et al. 2005). The MODIS data collection is a more recent set of products (2000 to present) and is produced at multiple resolutions (250, 500 and 1000 m).

In many developing countries, infrastructure and funding is not available for characterization of vegetation conditions for livestock decision making. Since livestock

is a main component of wealth and livelihoods in many of these countries, shortages in forage supply brought about by drought and other climatic disasters can be devastating. The ability to characterize the vegetation resource over large land areas on a near real-time basis can improve lead time for decision making at local, regional, and national levels. Given that products such the NOAA-AVHRR and MODIS vegetation indices, along with several of the of the high resolution rainfall products, have global extent the ability to develop near real-time systems using these products increases the ability to more easily extend them to other areas, thus reducing costs and time of implementation.

The overall goal of this study was to examine available technologies for implementing a near real-time system for monitoring biomass available to livestock on a given landscape, thus allowing more precise monitoring of the forage resources to improve decision making about animal numbers and the vegetation resource. The specific objectives of this study were:

1. Examine the correspondence of two different near real-time rainfall products to that collected from fixed-location automated weather stations located in the Trans Pecos region of Texas (Chapter II). The rainfall products and resolutions were as follows:
    a. National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center Morphing Product (CMORPH) rainfall with 8 km resolution.
    b. National Weather Service (NWS), Next Generation Weather Radar (NEXRAD) New Precipitation Analysis with 4 km resolution.
2. Verify the ability of a simulation model (PHYGROW) to predict standing crop of herbaceous biomass at selected sites (patches) in the Trans Pecos region of Texas using the near real-time rainfall product having the best correspondence with measured rainfall (Chapter III).

3. Determine the feasibility of using the Normalized Difference Vegetation Index (NDVI) as a covariate in cokriging biomass output from a simulation model in order to produce landscape maps of herbaceous biomass on a near real-time basis in the Gobi region of Mongolia (Chapter IV).

CHAPTER II
A COMPARISON OF TWO HIGH-RESOLUTION RAINFALL PRODUCTS TO GAGE-MEASURED RAINFALL IN FAR WEST TEXAS

Introduction

During the past 30 years, efforts have increased to estimate the spatial distribution of rainfall to improve flood, drought, and water monitoring and management (Grimes et al. 1999; Legates 2000; Moon et al. 2004). Historically, rainfall has been measured in gages at point locations and this is generally viewed as the most accurate representation of precipitation amounts (Arkin and Meisner 1987; Schmidt et al. 2000). However, for regional remote area monitoring, the spatial distribution of rainfall over the region is needed. This becomes problematic where the number of gages is sparse (Grimes et al. 1999), but has been partially overcome by advances in estimation of precipitation from geostationary satellite imagery (e.g., Arkin and Meisner 1987; Herman et al. 1997; Joyce et al. 2004) and from radar (e.g., Fulton et al. 1998; Whiton et al. 1998b; Young et al. 2000) or combinations of these with traditional rain gage measurements (Grimes et al. 1999; Legates 2000; Seo and Breidenbach 2002; Moon et al. 2004).

Satellite estimation of precipitation generally involves the use of algorithms that estimate rainfall based on thermal infrared imagery collected by geostationary satellites (Arkin and Meisner 1987; Grimes et al. 1999). These algorithms calculate the temperatures of cloud tops based on the thermal infrared measurements and estimate rainfall amounts based on the temperature of the cloud and its duration over a given area (Herman et al. 1997). For example, Arkin and Meisner (1987) developed the GOES Precipitation Index which estimates precipitation based on the duration that cold clouds occupy a region. Their algorithm predicts 3 mm of precipitation for each hour that cloud top temperatures stay below 235º K within a 2.5 º x 2.5 º moving window. A limitation to this methodology is that it estimates precipitation indirectly and therefore is not

calibrated against any known precipitation measurements (Grimes et al. 1999). Herman et al. (1997) developed techniques to overcome this by doing a near real-time bias correction between gage collected rainfall and the GOES Precipitation Index. Grimes et al. (1999) notes that the gage correction works well in areas of dense gage networks, but it is limited in areas with a low number of gages. They developed techniques for improving satellite rainfall estimates by using historical gage data and an interpolation technique that does weighted averaging of the satellite and rainfall data.

A more recent satellite rainfall estimation technique is the Climate Prediction Center Morphing Technique (CMORPH) (Joyce et al. 2004). This method combines thermal infrared imaging for cold clouds along with passive microwave data to estimate precipitation around the globe every 30 minutes at relatively high resolution (8 km pixels). The algorithm estimates precipitation from passive microwave observations and propagates these in time and space using the thermal infrared imagery from geostationary satellites. Joyce et al. (2004) state that the passive microwave data provide better estimates of rainfall but their deployment on polar orbiting satellite platforms limits their spatial and temporal sampling characteristics. Conversely, the spatial and temporal sampling characteristics are quite good for thermal infrared measurements from geostationary satellites, but rainfall estimated from this imagery can be poor, especially in areas where the rainfall is orographic rather than convective. For the CMORPH algorithm, the passive microwave and the thermal infrared imagery have been combined to take advantage of the overall strengths of each. Joyce et al. (2004) reported that the CMORPH method generally outperformed other passive microwave and thermal infrared products and had similar performance to radar for estimating precipitation. One shortcoming of the method noted by the authors was the over-prediction of precipitation in areas of ice and snow cover.

Radar estimation of precipitation is a more direct method of precipitation measurement when compared to satellite methods in that the information gathered by the

radar is related to droplet size rather than some byproduct of precipitation such as cloud temperature (Arkin and Meisner 1987). Radar estimation of precipitation generally involves the use of Doppler radars. Since 1991, the US Government has installed a series of these radars across the US under the program known as Next Generation Weather Radar system (NEXRAD) (Fulton et al. 1998). The purpose of this program is to gather real-time data on precipitation to improve weather forecasting, flood monitoring, and hydrological monitoring (Fulton et al. 1998; Young et al. 1999).

The algorithm for estimating precipitation used by the NEXRAD system is a power law relationship between the radar reflectivity and the precipitation amount. The relationship is as follows:

$$R = aZ^b \qquad [2.1]$$

where R is rainfall rate (mm/h), a and b are parameters that are adjustable for different regions, and Z is the radar reflectivity factor (mm^6/m^3) (Fulton et al. 1998; Young et al. 1999). The Z factor is estimated from the back scatter power measured by the radar and is related to the amount of water droplets in the atmosphere (Young et al. 1999). Radar measurement has a dual strategy of estimating rainfall in a 360 degree pattern (horizontal) while also measuring the atmosphere at 0.5 to 20 degrees from the horizon (vertical). This hybrid scan attempts to measure rainfall as near to ground level as possible while also trying to minimize contamination of precipitation estimates caused by ground returns (Young et al. 1999).

Several sources of errors and issues with NEXRAD data have been identified that can lead to over or under estimation of precipitation. Legates (2000) described these as 1) errors associated with reflectivity measurements, 2) errors associated with the reflectivity/rainfall (Z-R) relationship, and 3) errors associated with below the radar beam effects. Errors associated with reflectivity measurements include ground clutter contamination and reflections from atmospheric disturbance which can lead to overestimation. Rain falling on the radar dome and beam widening with increasing

distance from the radar can lead to underestimation of precipitation. Young et al. (1999), in a study of NEXRAD in complex terrain, noted underestimation of precipitation by NEXRAD and attributed this to the variability in the terrain within the radar sampling area, distance from the radar, and ground clutter that resulted in poor reception or distortion of the returns to the radar.

With regard to errors associated with the Z-R relationship, variability in the dropsize distributions can lead to erratic effects, whereas presence of hail, sleet, and snow can lead to overestimates. Errors associated with below the radar beam effects include overestimation of precipitation because of rainfall that evaporates before reaching the ground, and over and underestimation resulting from strong horizontal winds that cause the precipitation to fall in a location other than where it was measured (Legates 2000).

In a semiarid region of New Mexico, Xie et al (2006) found seasonal differences in NEXRAD Stage III precipitation estimates when compared to gage data. The NEXRAD system generally overestimated rainfall during the monsoon season (June 1 to September 30) and underestimated rainfall during the non-monsoon season (October 1 to May 31). These differences were mainly attributed to how the NEXRAD Stage III algorithm software truncates the estimated precipitation during small events (i.e., makes events < 0.01 mm/h equal to 0). Rainfall events tended to be smaller during the non-monsoon season when compared to the monsoon period, thus leading to the underestimation observed in the non-monsoon seasons. However, on a rainfall accumulation basis, the NEXRAD Stage III overestimated total amount of rainfall by 11 to 88 percent in monsoon season and underestimated rainfall by 18 to 89 percent in the non-monsoon season.

Although these high resolution products have errors that, at times, reduce their accuracy in predicting rainfall, these high resolution rainfall products may be the only

way to access spatially explicit rainfall data, especially in rangelands where the number of rainfall gages is sparse. Also, many of these products are produced on a near real-time basis, thus making them useful for drought early warning and biophysical modeling to assess plant growth and production. In this study, two different near real-time rainfall products (CMORPH and NEXRAD) were compared to that of fixed-location automated weather stations located near Marathon in the Trans Pecos region of Texas. The overall goal of this study was to determine which of these products would be most suitable for use in biophysical modeling to estimate livestock forage. The objectives were to 1) compare the frequency statistics to assess how well the products predicted the presence or absence of precipitation, and 2) examine bias, error, and estimation efficiency between the rainfall products and gage measured rainfall in order to assess how these rainfall products perform in remote rangeland areas where recording rain gages are sparse.

Methods

Study Area

The study was conducted on the Catto-Gage Ranch, approximately 13 km west of Marathon, TX (30°12'23.90"N, 103°14'47.26"W; Figure 2.1) in Brewster County. The Catto-Gage is a 172,609 acre working ranch that has been grazed by livestock since the mid 1880's. Historically, the ranch was primarily a cow calf operation. However after an extended drought during 1999 to 2002, the operation was changed from a cow-calf operation to a yearling cattle stocker operation to provide greater flexibility in the livestock and grazing management (Don Keeling, personal communication). The study area is currently grazed all months except October and November at stocking rates determined by the ranch management.

The study area is situated with the Glass Mountains in the northern part and the Del Norte Mountains on the western side, with the majority of the land area falling within the greater Marathon Basin (Figure 2.2). The area consists of "high plateaus, rugged peaks and sierras, and broad, shallow intermontane valleys" (Smith 2001). Elevation at the study area follows a general northwest to southeast gradient with the highest elevations in the northwestern portion and the lowest in the southeastern portion (Figure 2.2). Elevation at the site ranges from 988 m to 1940 m.

The climate of the area is semiarid with cool, dry winters and hot summers. Temperatures range from an average low of 4º C during the winter months (December to March) to an average high of 33º C during the summer months (June to September) (Figure 2.3). Average monthly temperature across all months is 17º C (NCDC 2006). Precipitation averages 369 mm with the highest amount occurring during the summer months (average of 58 mm/month; Figure 2.3). On average, the lowest amount of precipitation occurs during March (6 mm; Figure 2.3). Precipitation during the summer months is generally from thunderstorms (Powell 1998) that can be of high intensity. Snowfall can occur during the winter months, but this form of precipitation generally averages less than 2 cm during the winter months (NCDC 2006).

High Resolution Rainfall Products

Two rainfall products were compared to examine their correspondence to fixed-location automated rain gages. The first product was the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center Morphing Product (CMORPH) rainfall (Joyce et al. 2004) (referred to hereafter as "CMORPH product"). This product is produced by NOAA each 24-hour period and represents the accumulated rainfall that occurs between 0:00 and 24:00 Greenwich Mean Time (GMT) (24:00 GMT is equivalent to 6:00 pm CST). The CMORPH product was acquired automatically from the NOAA servers via internet and downloaded to servers at the Center for Natural

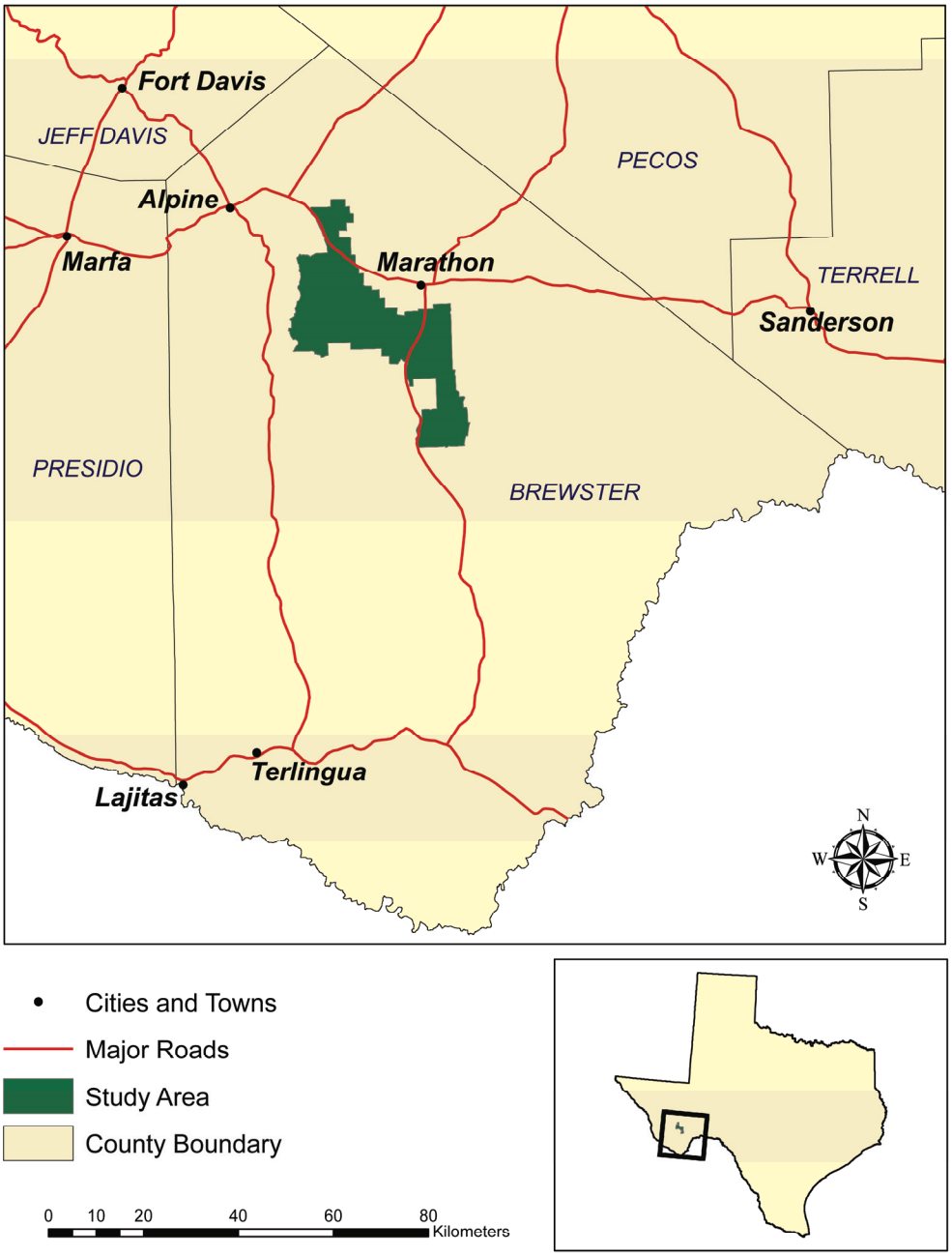

Figure 2.1. Location of the study area in Brewster County, TX.

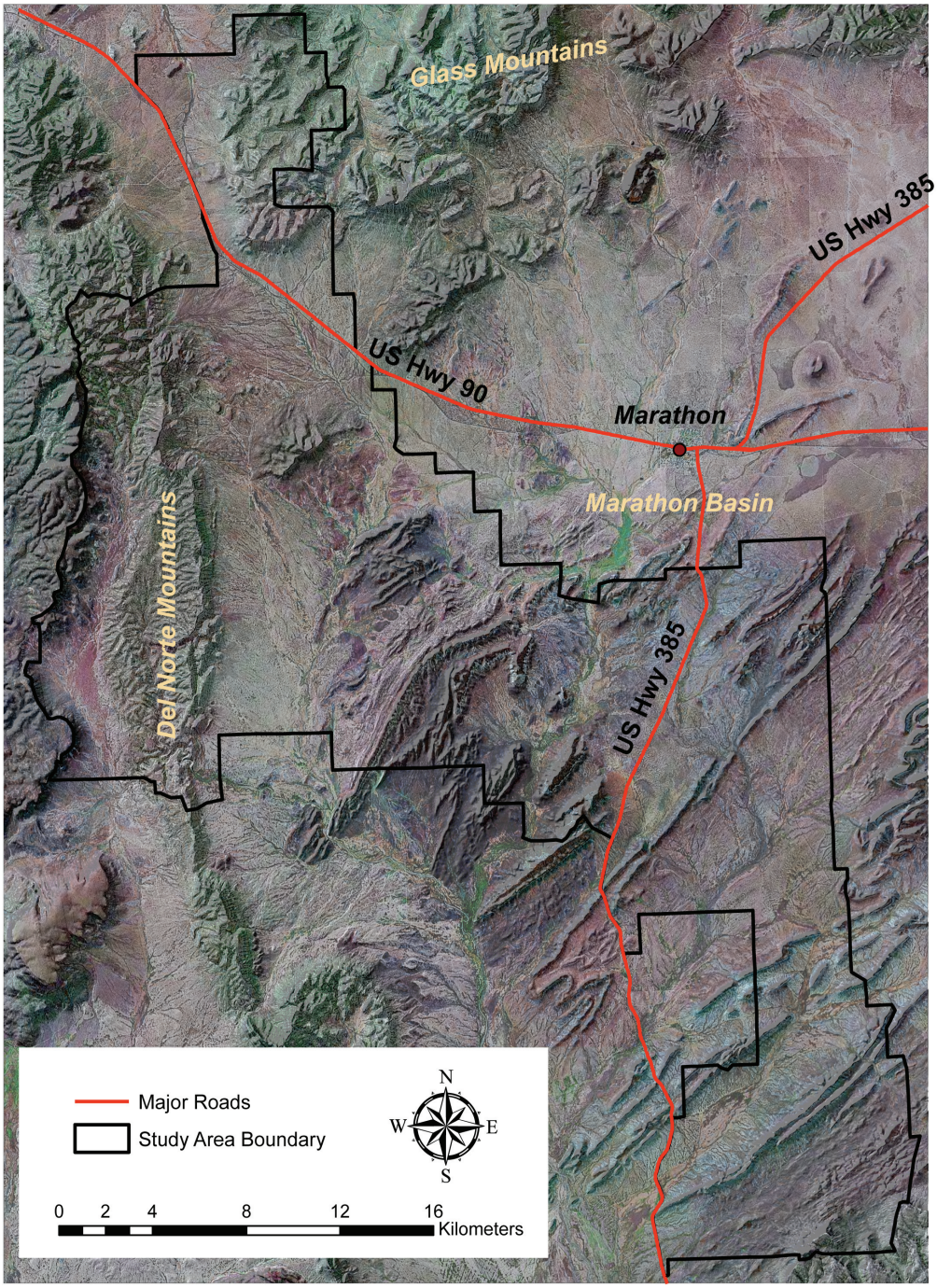

Figure 2.2. Landsat Enhanced Thematic Mapper Geocover mosaic image (MDA Federal 2004) draped over a hillshade representation of the 30-m Digital Elevation Model (USGS 1999) to provide an aerial view of the changing topography at the study area (outlined in black) near Marathon, TX.

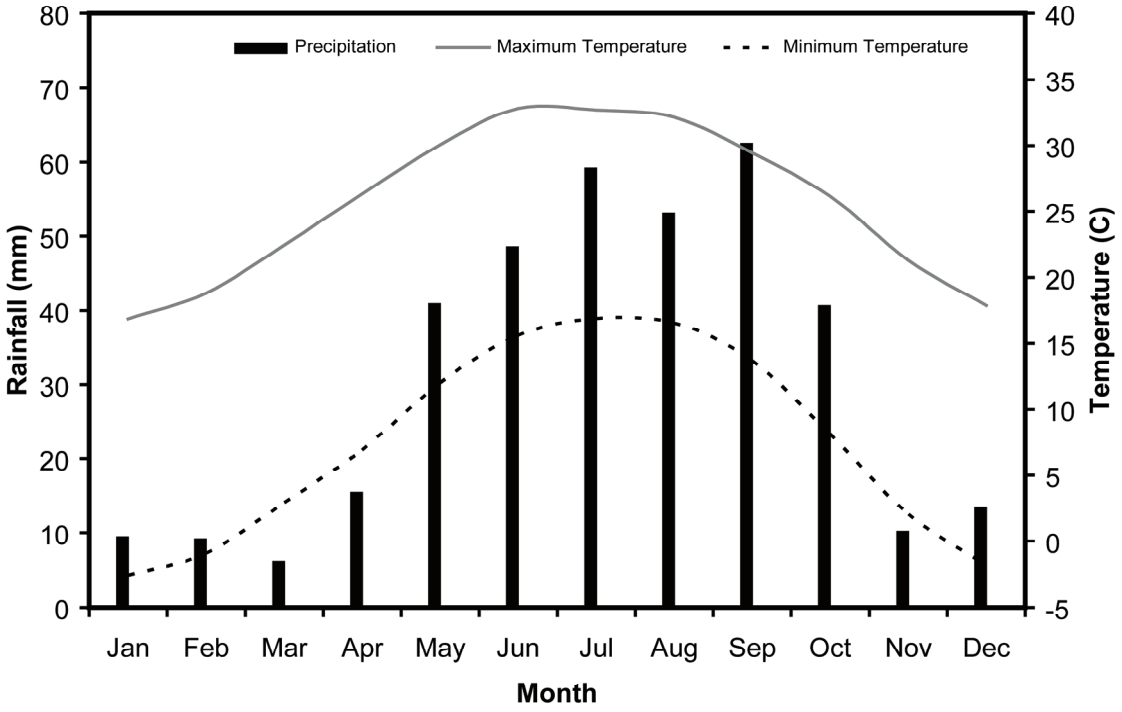

Figure 2.3. Average monthly rainfall (mm), average monthly maximum temperature (ºC) and average monthly minimum temperature (ºC) during the period from 1970 to 2000 in Marathon, TX (NCDC 2006).

Resource Information Technology (CNRIT), Texas A&M University. The rainfall product was delivered as a gridded image that had a geographic range of 130.0º to 65.0º West longitude and 22.0º to 50.0º North latitude, and covered the contiguous 48 states. Grid cell spacing of the image was 0.07276º in the longitudinal direction and 0.07277º in the latitudinal direction (approximately 8 km at the equator).

The second product examined was produced by NOAA National Weather Service (NWS) and was called the "New Precipitation Analysis" (http://www.srh.noaa.gov/rfcshare/precip_analysis_new.php) (referred to hereafter as the "NEXRAD product"). The precipitation for this product is estimated using a multi-sensor approach (NOAA 2007) where the WSR-88D NEXRAD radar precipitation estimates are compared to reported ground station rainfall measurements and a bias correction is calculated to correct the NWS NEXRAD precipitation estimates (P1 protocol as

described in Young et al. 2000). In areas where radar coverage is sparse or non-existent, satellite precipitation estimates are used in place of the radar estimates and corrected using the ground station data. The NEXRAD data are processed daily by NWS with the accumulation period starting at 12:00 GMT (6 am CST) and ending 24 hours later. The data were made available in shapefile format. The shapefile contains the NWS Hydrologic Rainfall Analysis Project (HRAP) grid cell where the radar has detected rainfall and the amount. Each HRAP grid cell is approximately 4 by 4 km (Reed and Maidment 1999). For the study area, the NEXRAD radar located in Midland, TX covered the majority of the site (Figure 2.4).

The NEXRAD product was downloaded from the NWS servers on a daily basis and stored in the CNRIT weather database. The daily data were joined to a master table that contained all the HRAP cells for the continental US. Since the NWS data contains only those HRAP cells where rainfall was estimated to occur, any cells not having a precipitation estimate were given a zero (0) to indicate that no rainfall occurred in the area represented by the HRAP grid cell.

Automated Weather Stations

Two automated weather stations were installed in grazing exclosures on the study area. The exclosures were built in 2003 and had been previously subjected to continuous grazing by cattle. One exclosure was located in the north-central part of the study area and allowed the automated station to be positioned in both a unique HRAP and CMORPH grid cell (Figure 2.5; West Point station). The second exclosure was located in the south-central portion of the study area and also allowed placement of the station in unique HRAP and CMORPH grid cells (Figure 2.5; Twin China station). The distance between the two stations was approximately 20 km.

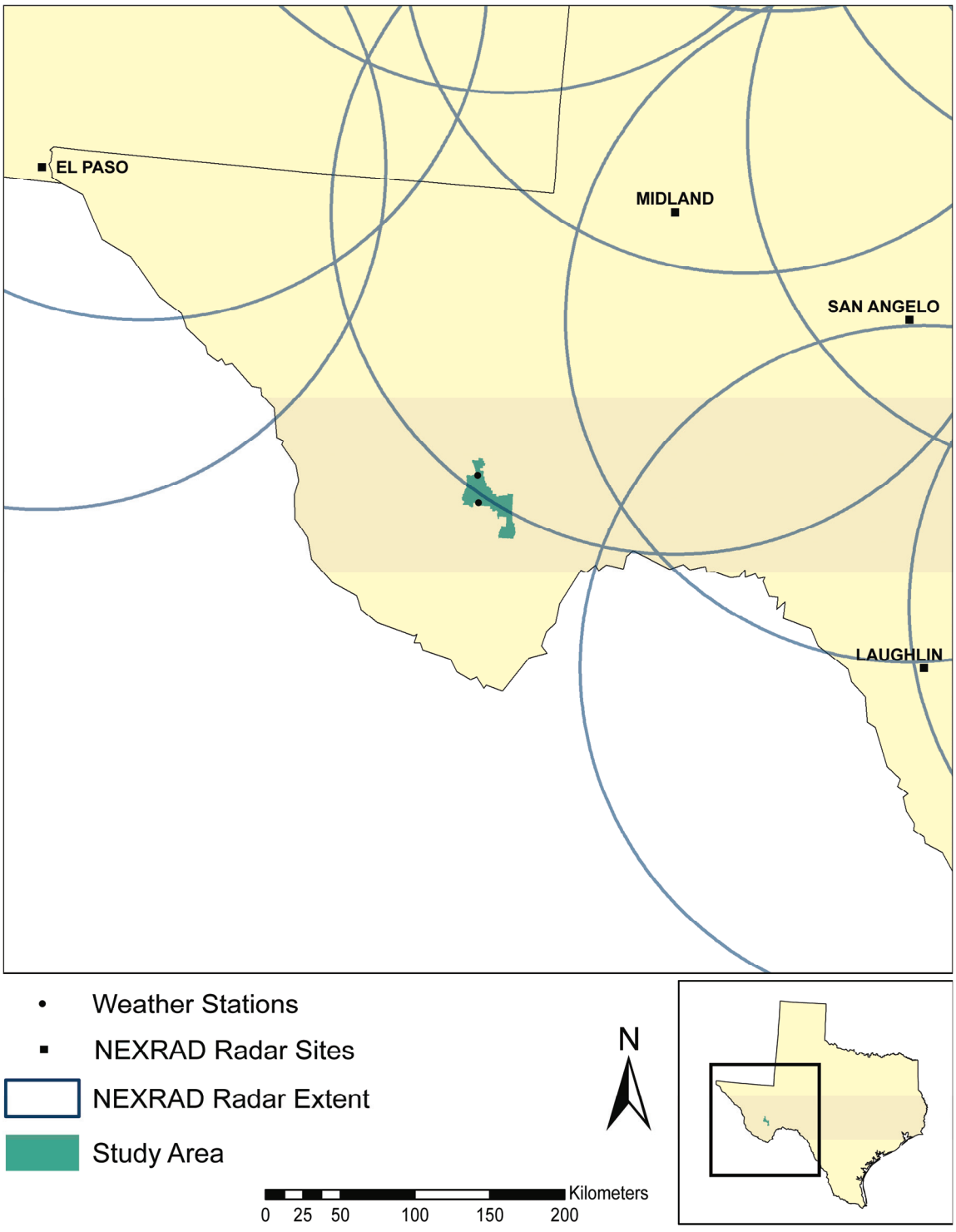

Figure 2.4. Location of the study area and automated weather stations in West Texas in relation to NEXRAD radar locations.

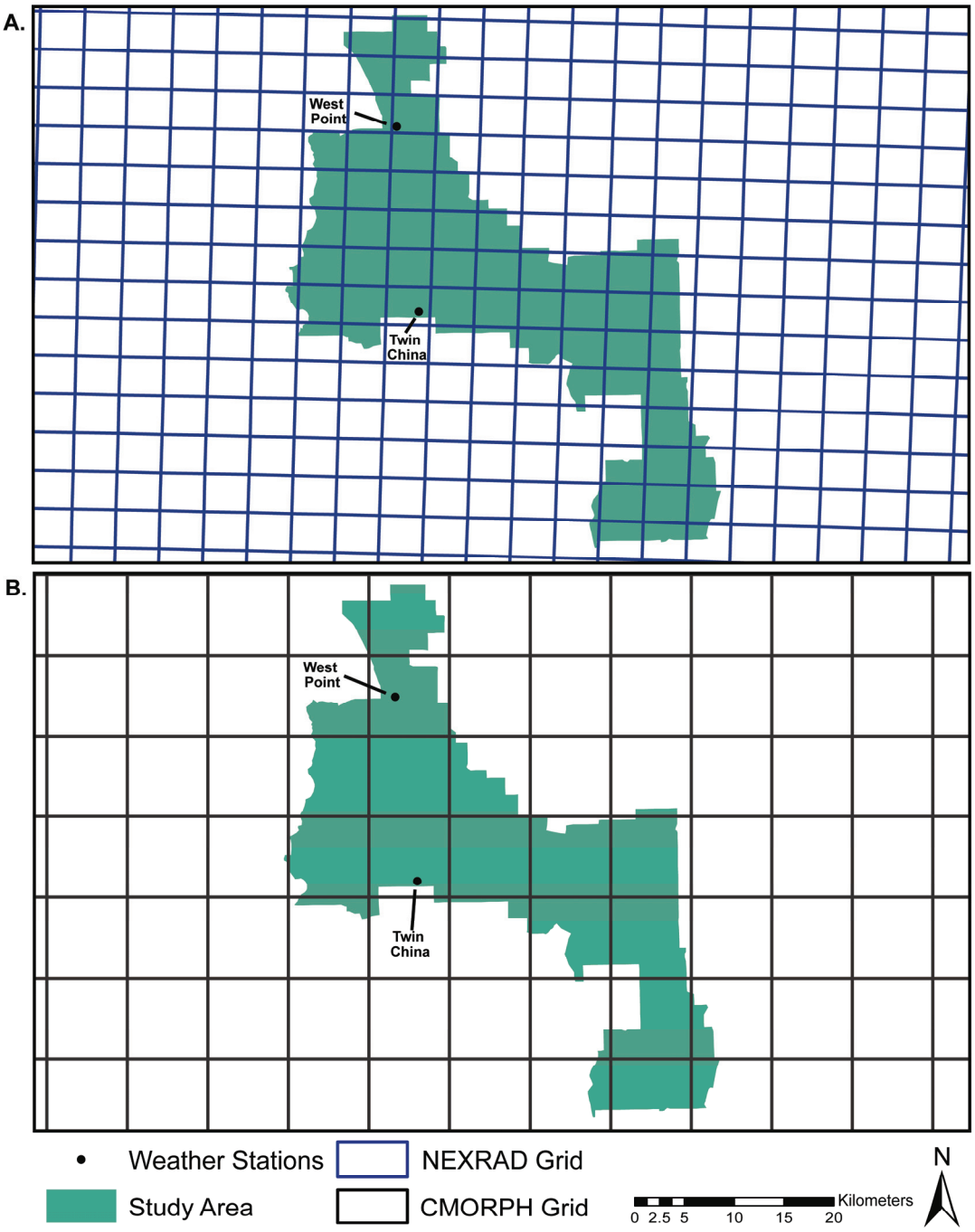

Figure 2.5. Spatial resolution (grid size) of the A) CMORPH and B) NEXRAD rainfall products in relation to the weather station locations at the study site.

Each automated weather station was equipped with an 8-inch orifice tipping bucket rain gage (Texas Electronics TE525WS, 2.54 mm/tip), a combination temperature and relative humidity probe enclosed in a gill radiation shield (Campbell Scientific CS215), a pyranometer for solar radiation measurement (LiCor 200X), and a wind anemometer (RM Young Wind Sentry 3101) for wind speed and direction measurement. Each of these instruments was wired to a Campbell CR10X datalogger that allowed continuous measurement of these climate variables. The CR10X was programmed to take measurements each minute. The minutely readings were averaged (or summed for rainfall) by the CR10X software and hourly measurements for each climate variable were stored by the datalogger. At the end of each 24-hour period (12:00 am to 11:59 pm CST), the hourly measurement of climate variables were averaged or summed by the CR10X software to produce 24-hour measurements. At periodic intervals, the hourly and 24-hour measurements were downloaded from the CR10X and stored in a database at CNRIT for subsequent comparison to the high resolution rainfall products.

Scripting tools were developed to extract both the CMORPH and NEXRAD precipitation data for a given latitude and longitude. The CMORPH and NEXRAD data were extracted for the location of each automated weather station in the study area. To allow comparisons of gage rainfall to that of the NEXRAD and CMORPH rainfall products, the hourly rain gage data from the weather stations was aggregated to match the accumulation periods of both the NEXRAD and CMORPH rainfall products. For example, to match the CMORPH rainfall on August 2, the hourly station data was summed from 6:01 pm on August 1 to 6:00 pm on August 2. For the NEXRAD data, the station rainfall measurements for August 2 represented the sum of the hourly station data from 6:01 am on August 1 to 6:00 am on August 2.

The CMORPH rainfall product that provided coverage of the study area did not become available for use until February 2004. Therefore, for the analyses in this study,

the time period for all data sources (CMORPH, NEXRAD, and station gages) begins February 1, 2004 and continues to March 31, 2007.

Statistical Analysis

A two-way contingency table with a rain/no rain contingency was used to compare rainfall detection for each rainfall product. Analyses were conducted by product (CMORPH vs. NEXRAD), product and location (West Point station vs. Twin China station; Figure 2.5), and product and season (monsoon and non-monsoon; June 1 to September 30 and October 1 to May 31, respectively). For the two-way contingency analyses, the following variables were defined:

- if rainfall was measured at the station and also was detected by the rainfall product, this was defined as a "hit";
- if rainfall occurred at the station, but was not detected by the rainfall product, this was a "miss";
- if rainfall was detected by the rainfall product, but none was measured at the station, this was defined as a "false alarm";
- if no rainfall was measured at the station and none was detected by the rainfall product, this was defined as a "correct negative".

Using the above variables from the contingency tables, a set of statistics (Stanski et al. 1989; Johnson and Olsen 1998) were calculated to examine the ability of the rainfall products to detect rainfall at the study area. These were as follows:

$$Accuracy = \frac{hits + correct\ negatives}{total\ days} \quad [2.2]$$

$$Bias\ Score = \frac{hits + false\ alarms}{hits + misses} \quad [2.3]$$

$$Probability\ of\ Detection = \frac{hits}{hits + misses} \quad [2.4]$$

$$False\ Alarm\ Ratio = \frac{false\ alarms}{hits + false\ alarms} \quad [2.5]$$

$$Probability\ of\ False\ Detection = \frac{false\ alarms}{correct\ negatives + false\ alarms} \quad [2.6]$$

$$Critical\ Success\ Index = \frac{hits}{hits + misses + false\ alarms} \quad [2.7]$$

$$Equitable\ threat\ score = \frac{hits - hits_{random}}{hits + misses + false\ alarms - hits_{random`}} \quad [2.8]$$

$$\text{where } hits_{random} = \frac{(hits + misses) \times (hits + false\ alarms)}{total\ days} \quad [2.9]$$

For the above statistics, accuracy provides an indication of the fraction of the days that the rainfall product correctly estimated both the presence and absence of rainfall as observed at the weather stations. Values range between 0 and 1 with a 1 being perfect detection of the presence and absence of rainfall. Bias score is a ratio of the frequency of rainfall days detected by the rainfall product versus that of days in which rainfall was observed at the station. For bias scores less than 1, the rainfall product would have a tendency to detect less days of rainfall than the number of days in which rainfall was observed at the station. Bias scores greater than 1 indicate the opposite. Probability of detection (POD) is the fraction of days where rainfall was correctly detected by the rainfall product when rain was observed at the station. The POD can range between 0 and 1 with 1 being perfect detection of rainfall by the product. The false alarm ratio (FAR) provides an indication of the fraction of days in which the rainfall product detected rainfall when none was observed at the station. Values can range between 0 and 1 with a value of 0 indicating no false alarms. Probability of false detection (POFD) indicates the fraction of days that the product detected rainfall compared to total number of days when no rain was observed at the station. Scores can range between 0 and 1 with a value of 0 indicating perfect correspondence between the product and station. The critical success index (CSI) provides an indication of the number of days that rainfall were correctly detected by the rainfall product compared to

total number of days where rainfall was observed at the station (hits and misses) or incorrectly detected by the rainfall product (false alarms). It is similar to accuracy, but does not include days where no rainfall was detected by both the station and the product (correct negatives). Values of 1 indicate perfect detection of rainfall by the product. Lastly, the equitable threat score (ETS) provides an indication of how well the rainfall product detected rainfall correctly compared to random chance. ETS scores near 0 would indicate that the product had no skill in detecting rainfall. Scores near 1 would indicate near perfect detection of rainfall by the product.

To assess the ability of the rainfall products to estimate rainfall amounts, the time series data for each product, station, and location comparisons were examined for total difference in rainfall amount, estimation bias, estimation efficiency, slope, and root mean square difference. Total difference (TD) in rainfall is simply the subtraction of the sum of the total station rainfall during the time series from the sum of the total product rainfall estimated during the same time series. Estimation bias (BIAS) reflects the normalized difference between the precipitation product estimate and fixed weather station data and is equated as follows (Jayakrishnan et al. 2004; Moon et al. 2004):

$$BIAS\ (\%) = \frac{Rainfall\ Product\ Total - Station\ Total}{Station\ Total} \times 100 \qquad [2.10]$$

Positive estimation bias values indicate the overestimation by the rainfall product compared to the station gages whereas negative values indicate the opposite. Estimation efficiency (EE) is a measure of the deviation from a 1:1 line between station precipitation and the precipitation product estimate and is calculated as follows (Nash and Sutcliffe 1970; Jayakrishnan et al. 2004; Moon et al. 2004):

$$EE = 1.0 - \frac{\sum_{i=1}^{n}\left(R_i - W_i\right)^2}{\sum_{i=1}^{n}\left(R_i - R_m\right)^2} \qquad [2.11]$$

Where R_i is the station gage precipitation for day i, W_i is the precipitation product rainfall total in the grid cell where the station is located for day i, R_m is the average station gage precipitation over all days, and n is the total number of days. A value of 1

would reflect a perfect correspondence between the station data and the rainfall product. Values greater than 0 would indicate that a positive relationship exists between the station data and rainfall product and that the rainfall product data could be used as a good estimate for the station location. Values less than 0 indicate a low correspondence between the station data and rainfall product and that an average of the station data would be a better predictor of the rainfall at the station location than the rainfall product estimates (Moon et al. 2004). Root mean square difference (RMSD) is a measure of the average magnitude of the difference between the rainfall product and the station gage data. It is calculated as follows:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n}\left(R_i - W_i\right)^2}{n}} \qquad [2.12]$$

Linear regression with a zero intercept was conducted to examine slope for the rainfall products versus the station data. The linear regression and estimation efficiency analyses were performed on a conditional dataset where days when no rainfall was detected by both the rainfall product and the station (i.e., correct negatives as defined above) were excluded from the analysis. Inclusion of days where no rainfall was detected by both the rainfall product and the station inflates the statistics in favor of the rainfall product, thus reducing the quality assessment (Jayakrishnan et al. 2004).

Results

Rainfall Detection Ability

The rainfall products differed in their ability to detect rainfall at the study site. An examination of the rainfall product and station location contingencies had very similar results to that of the product only (both station locations combined) contingency analysis, therefore the rainfall product results without regard to location are presented. The NEXRAD product was slightly more accurate at detecting both rain and no rain at the study site compared to CMORPH (0.88 versus 0.82 respectively; Table 2.1). Both

products had a tendency to detect rainfall on a greater number of days than what was observed at the weather stations as indicated by their frequency bias scores with the CMORPH product detecting a greater number of days of rainfall than the NEXRAD product (Table 2.1). Both products had similar ability to detect rainfall when rainfall was observed at the stations, with the probability of detection being 0.78 for CMORPH versus 0.80 for the NEXRAD product (Table 2.1). However, the CMORPH product had the tendency to detect a higher number of days of rainfall than what was detected at the stations (i.e., more false alarms) when compared to the NEXRAD product (Table 2.1). This resulted in a higher probability of false detection for the CMORPH product when compared to the NEXRAD product (Table 2.1).

The CSI provides an indication of the correspondence between the number of events where rainfall was detected by the product versus that observed at the weather stations. Index values near 1 indicate that the rainfall products had high success in

Table 2.1. Two-way contingency tables (rain or no rain) and contingency statistics for a comparison of the ability of two high resolution rainfall products (CMORPH and NWS) to detect rainfall measured at weather stations near Marathon, TX.

		Station				Station	
		Rain	No Rain			Rain	No Rain
CMORPH	Rain	326	307	**NEXRAD**	Rain	334	180
	No Rain	92	1547		No Rain	86	1676
Total Days			2272				2276

Contingency Statistic	**CMORPH**	**NEXRAD**
Accuracy	0.82	0.88
Frequency Bias	1.51	1.22
Probability of Detection	0.78	0.80
False Alarm Ratio	0.48	0.35
Probability of False Detection	0.17	0.10
Critical Success Index	0.45	0.56
Equitable Threat Score	0.34	0.47

detecting rainfall and a value of zero indicates no success. Both the NEXRAD and CMORPH products had moderate success in detecting rainfall at the station sites with the NEXRAD product having a higher success rate (0.56 and 0.45, respectively; Table 2.1).

In contrast to the CSI, the equitable threat score provides an indication of how well the rainfall products correctly detected rainfall at the station site accounting for rain events that could have been detected due to random chance. Values near 1 indicate perfect skill in detecting rainfall and values less than 0.33 indicate low predictability. The CMORPH rainfall product had relatively low predictability (0.34) once random chance was accounted for, whereas the NEXRAD product had moderate predictability (0.47). However, both products equitable threat scores were lower than the critical success index indicating that, in this instance, accounting for random chance was important for assessing skill of the products to detect rainfall at these sites.

The contingency analysis was conducted for the rainfall products on a seasonal basis (monsoon and non-monsoon). In general, the seasonal contingency statistics (Table 2.2) for the product-season comparison followed the same general trend of the statistics of the product-only comparisons (Table 2.1) with the NEXRAD product having relatively better contingency statistics than CMORPH. However, some seasonal differences existed among the rainfall products. Both the CMORPH and NEXRAD products had higher accuracy in detecting both the presence and absence of rainfall during the non-monsoon season although this may be inflated by the greater number of days during the non-monsoon season. Accuracy, as calculated in the contingency analysis, is strongly influenced by the category with the greatest frequency of events (Stanski et al. 1989). In this case, the days where both the product and station detected no rainfall (correct negatives) was much greater during the non-monsoon season for both products. The probability of detection may be a more appropriate statistic for the seasonal comparison as it is influenced only by “hits” and “misses” which have

comparable totals for the product-season comparisons (Table 2.2). For both the CMORPH and NEXRAD products, the probability of detection was higher during the monsoon season compared to the non-monsoon season.

Individually, the products differed seasonally. The NEXRAD product during the non-monsoon season has slightly higher accuracy, lower bias, and lower probability of false detection when compared to the monsoon season (Table 2.2). However, during the monsoon season, a greater percentage of the rainfall events were correctly detected (95%) compared to the non-monsoon season (65%) (Table 2.2). The CMORPH product exhibited similar trends with higher accuracy, slightly lower bias, and lower probability of false detection during the non-monsoon season (Table 2.2). Probability of detection was also lower in the non-monsoon season compared to the monsoon season (Table 2.2).

Table 2.2. Two-way contingency tables (rain or no rain) and contingency statistics for a comparison of the ability of two high resolution rainfall products (CMORPH and NEXRAD) to detect rainfall measured during different rainfall seasons (monsoon and non-monsoon) at weather stations near Marathon, TX.

			Station				**Station**	
			Rain	No Rain			Rain	No Rain
Monsoon	**CMORPH**	Rain	165	138	**NEX-RAD**	Rain	174	97
		No Rain	31	338		No Rain	10	391
	Total Days			672				672
			Station				**Station**	
			Rain	No Rain			Rain	No Rain
Non-Monsoon	**CMORPH**	Rain	161	169	**NEX-RAD**	Rain	160	83
		No Rain	61	1209		No Rain	76	1285
	Total Days			1600				1604

	CMORPH			**NEXRAD**	
Contingency Statistic	**Monsoon**	**Non-Monsoon**		**Monsoon**	**Non-Monsoon**
Accuracy	0.75	0.86		0.84	0.90
Frequency Bias	1.55	1.49		1.47	1.03
Probability of Detection	0.84	0.73		0.95	0.68
False Alarm Ratio	0.46	0.51		0.36	0.34
Probability of False Detection	0.29	0.12		0.20	0.06
Critical Success Index	0.49	0.41		0.62	0.50
Equitable Threat Score	0.31	0.33		0.48	0.44

Rainfall Estimation Ability

The ability of the rainfall products to correctly estimate the amount of rainfall was highly contrasting. For both station locations, the CMORPH product estimated substantially more precipitation than what was measured at the station gages (Figure 2.6). An examination of estimation biases indicated that the CMORPH product overestimated rainfall amounts by 50% (765 mm) at the West Point location and by 41% (613 mm) at the Twin China location for the entire time series (Figure 2.6). In contrast, the NEXRAD product at the West Point location overestimated rainfall by 7% (102 mm) whereas rainfall was underestimated at the Twin China location by 9% (-141 mm) (Figure 2.7).

As expected, linear regression on the conditional data with respect to zero rain indicated a positive relationship between the station rainfall and that estimated by the rainfall product, with each product and location having slopes greater than 0.65 (Figure 2.6 and Figure 2.7). The proportion of the variability explained by the linear regression model was greatest for the NEXRAD product at both station locations when compared to the CMORPH product (r^2 values of 0.60 to 0.67 for the NEXRAD product and 0.40 to 0.47 for the CMORPH product; Figure 2.6 and Figure 2.7). Root mean square differences were lower for the NEXRAD product at both locations (6.7 to 7.1 mm) when compared to the CMORPH product (9.4 to 9.5 mm) indicating an almost 40% greater amount of variability in the CMORPH rainfall (Figure 2.6 and Figure 2.7) compared to the NEXRAD rainfall.

Estimation efficiency was very poor for the CMORPH product at both locations with the West Point site having a negative estimation efficiency (-0.17) and the Twin China site exhibiting a value near 0 (Figure 2.6). Estimation efficiency was much greater for the NEXRAD product which exhibited a value 0.50 for the Twin China

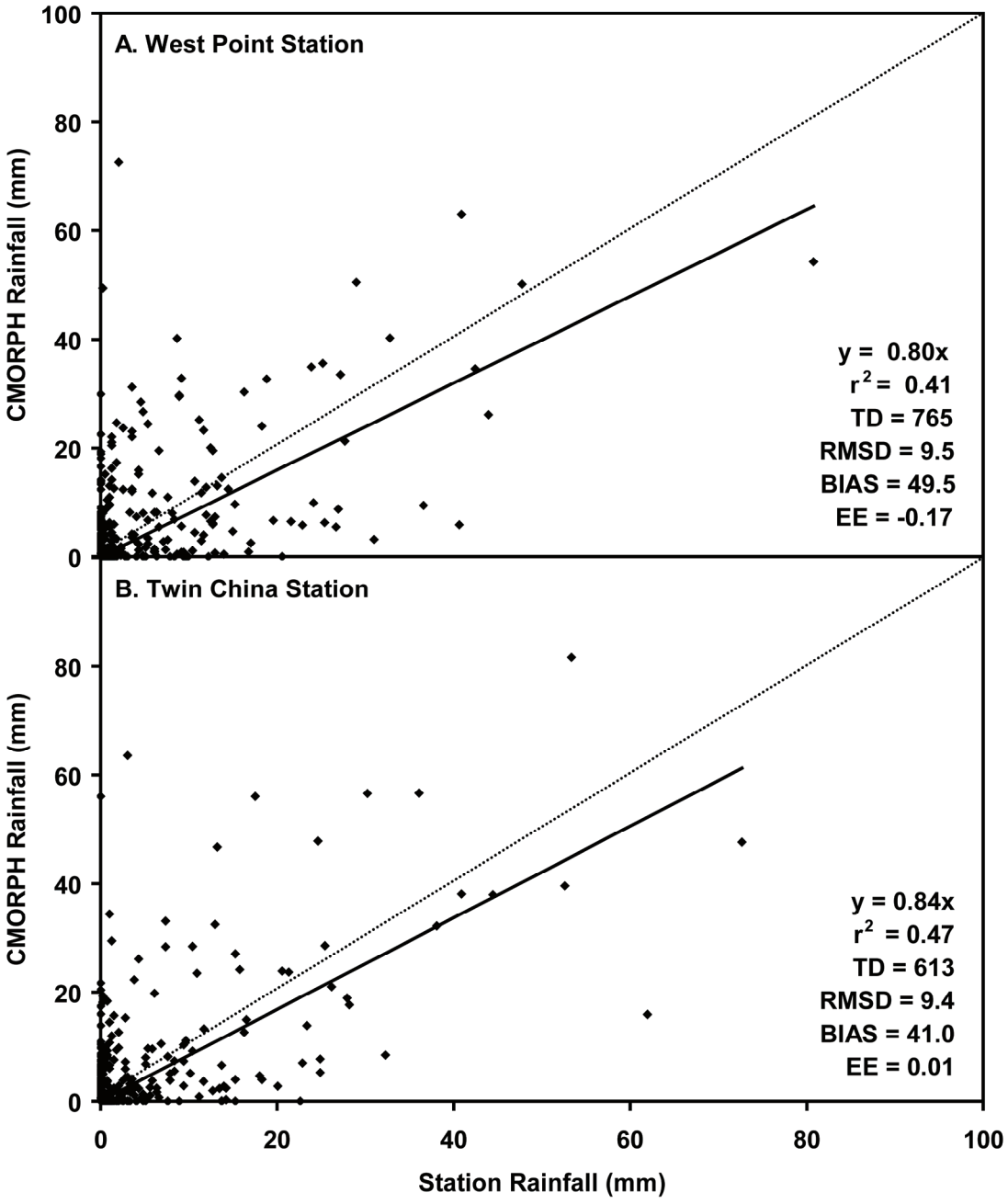

Figure 2.6. Comparison of daily station rainfall (mm) versus the CMORPH rainfall product estimate for the A) West Point station and B) Twin China station located at the study site near Marathon, TX. Data used were conditional with respect to zero rain (i.e., all station-product pairs with zeros on the same day were dropped from the analysis). Comparison statistics include no-intercept regression slope and r^2, total difference in absolute amounts of rainfall (TD; mm), root mean squared difference in (RMSD; mm), percent bias (BIAS) and estimation efficiency (EE). Dotted line represents 1:1 line.

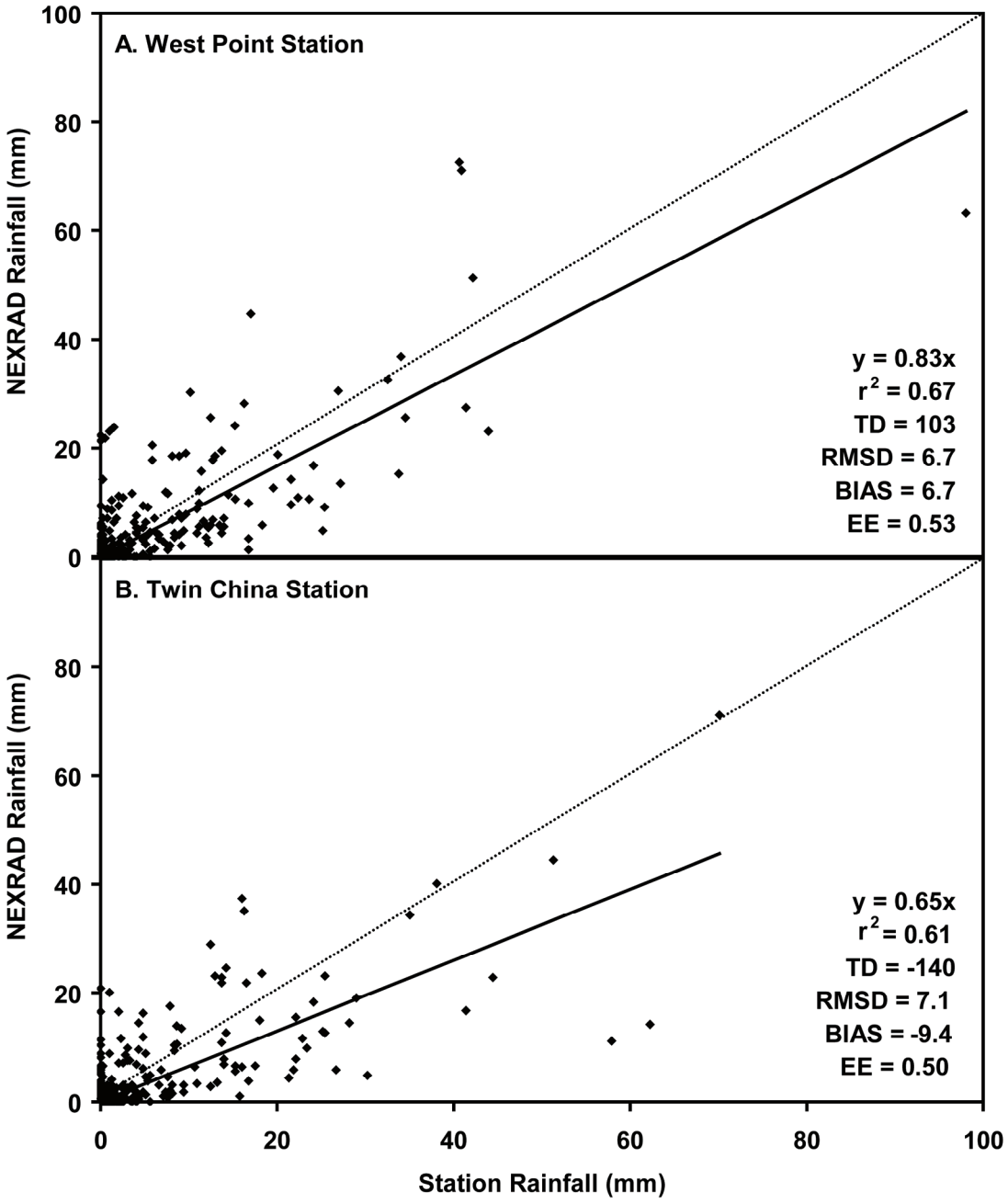

Figure 2.7. Comparison of daily station rainfall (mm) versus the NEXRAD rainfall product estimate for the A) West Point station and B) Twin China station located at the study site near Marathon, TX. Data used were conditional with respect to zero rain (i.e., all station-product pairs with zeros on the same day were dropped from the analysis). Comparison statistics include no-intercept regression slope and r^2, total difference in absolute amounts of rainfall (TD; mm), root mean squared difference (RMSD; mm), percent bias (BIAS) and estimation efficiency (EE). Dotted line represents 1:1 line.

station and 0.53 for the West Point Station locations (Figure 2.7). The low estimation efficiency for the CMORPH was likely the result of several extreme over or under predictions of rainfall by the CMORPH product. There were several instances where the rainfall predicted by CMORPH and that measured at the station differed by 45 to 70 mm. The NEXRAD product exhibited several instances of extreme over or under prediction of rainfall; however, the NEXRAD product had smaller absolute differences that ranged from 30 to 48 mm.

For the NEXRAD product, there were strong location differences in estimation of rainfall. At the West Point station, in the northern portion of the study area, the NEXRAD product overestimated rainfall amounts by 6.7% compared to a 9.4% underestimation of rainfall at the Twin China station (Figure 2.7). The variability in rainfall was also slightly lower at the West Point location with a root mean square difference of 6.7 mm compared to 7.1 mm at the Twin China station. A possible reason for these differences is that the Twin China station is just beyond the extent boundary for the NEXRAD radar coverage (Figure 2.4) which could have led to poor signal returns from this site.

Seasonal differences were also apparent for the high resolution rainfall products. For both locations combined, CMORPH overestimated rainfall by 59% during the monsoon season compared to 29% during the non-monsoon season (Figure 2.8) as indicated by the estimation bias. For the NEXRAD product, rainfall was overestimated by 1% in the monsoon season whereas it was underestimated by 4% during the non-monsoon season (Figure 2.9). Like the differences described above for location, CMORPH overestimated precipitation at substantially higher amounts than the NEXRAD product for both seasons.

The CMORPH product during the monsoon season had the greatest degree of variability in estimating precipitation at the two stations as indicated by the low r^2 (0.38)

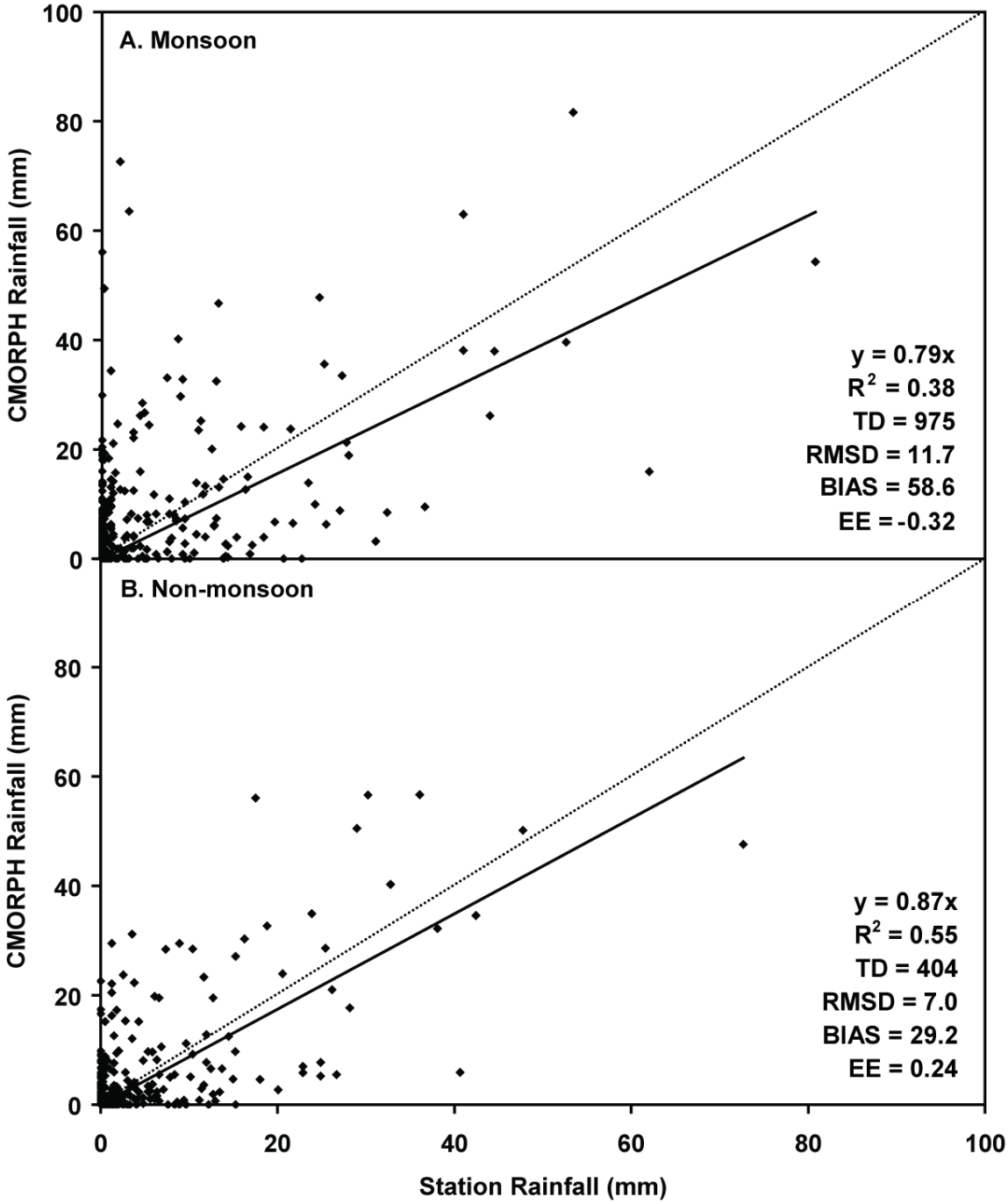

Figure 2.8. Comparison of daily station rainfall (mm) versus the CMORPH rainfall product estimate for the A) monsoon (June 1 to September 30) and B) non-monsoon season (October 1 to May 31). Data used were conditional with respect to zero rain (i.e., all station-product pairs with zeros on the same day were dropped from the analysis). Comparison statistics include no-intercept regression slope and r^2, total difference in absolute amounts of rainfall (TD; mm), root mean squared difference (RMSD; mm), percent bias (BIAS) and estimation efficiency (EE). Dotted line represents 1:1 line.

and the relatively high root mean squared difference (11.7) (Figure 2.8). For the non-monsoon season, the CMORPH product exhibited lower variability than during the monsoon season (r^2 =0.55 and RMSD =7.0 mm; Figure 2.8). This was only slightly higher than the amount of variability exhibited by the NEXRAD product during the non-monsoon season (r^2 = 0.60 and RMSD=56.1 mm; Figure 2.9).

Estimation efficiency was lowest for the CMORPH product during the monsoon season (Figure 2.8) with a value of -0.32 indicating that CMORPH's skill of estimating rainfall was extremely poor during this season. Estimation efficiency was greater for CMORPH during the non-monsoon season (Figure 2.8). However this was less than that of the NEXRAD product regardless of season (Figure 2.9). The NEXRAD product's estimation efficiency was better in the monsoon season compared to the non-monsoon.

Discussion

Rainfall Detection

The ability to detect the presence of rainfall at the study site varied among the high resolution rainfall products examined in this study. In general, the NEXRAD radar product performed better than the CMORPH product with higher accuracy in detecting events, less over-prediction of the number of rainfall events, less false alarms, and higher skill scores (CSI and ETS) than CMORPH (Table 2.1). In the initial validation of the CMORPH product across the entire United States, Joyce et al. (2004) noted that CMORPH had a higher average rates of false alarms than radar. However, in contrast to this study, they observed an overall higher probability of detection for CMORPH when compared to radar, but state that inclusion of mountainous areas in the western US where radar coverage suffers from terrain blockages may have factored into the lower detection for radar. Joyce et al. (2004) also report equitable threat scores ranging from 0.37 to 0.43 for the CMORPH validation which were higher than that observed in this study

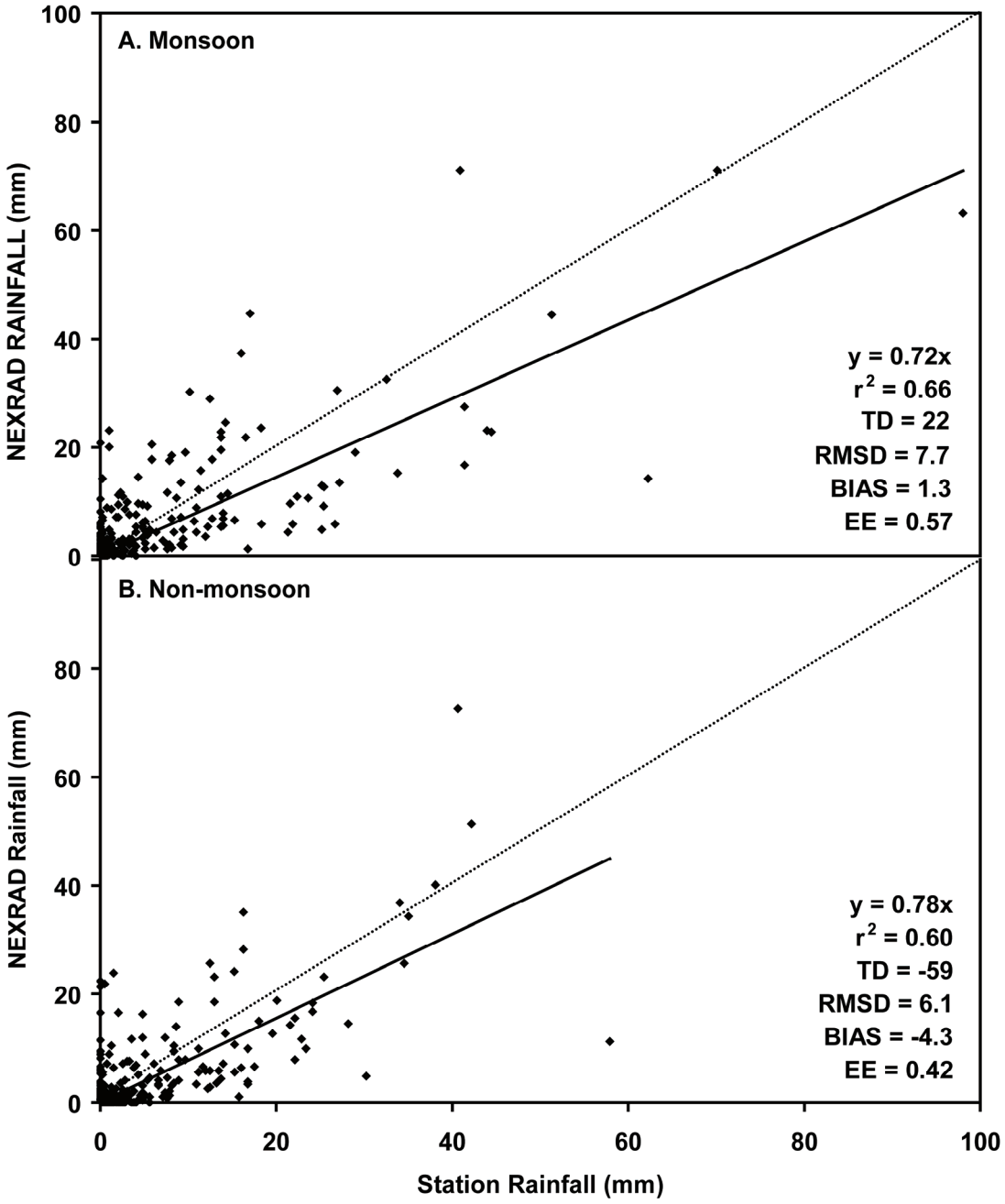

Figure 2.9. Comparison of daily station rainfall (mm) versus the NEXRAD rainfall product estimate for the A) monsoon (June 1 to September 30) and B) non-monsoon season (October 1 to May 31). Data used were conditional with respect to zero rain (i.e., all station-product pairs with zeros on the same day were dropped from the analysis). Comparison statistics include no-intercept regression slope and r^2, total difference in absolute amounts of rainfall (TD; mm), root mean squared difference (RMSD; in mm), percent bias (BIAS) and estimation efficiency (EE). Dotted line represents 1:1 line.

(0.34; Table 2.1). Possible reasons for these differences are that the CMORPH validation included over 7,000 stations across the US and the time period for the validation was much shorter (June to November 2003) than this study.

Seasonal differences in the ability to detect rainfall by the products were apparent for this study site. Both products exhibited greater probability of detection in the monsoon season with the NEXRAD product have the highest probability of detection. Joyce et al. (2004) reported similar trends of higher probability of detection for both radar and CMORPH in summer months (June to August; comparable to monsoon period in this study) than in the fall period (September to November; early part of non-monsoon period in this study). At mid-latitudes in Australia, Ebert et al. (2007) reported that CMORPH had greater probability of detection in the summer period (December – February) than in the winter period (June to August). They attributed this difference to the lower number of convective rainfall events during the winter period that are not as easily detected by the infrared and passive microwave sensors on satellites used for CMORPH data collection. In a study of NEXRAD Stage III radar (a precursor to the NEXRAD product used in this study that employs an alternative methodology for gage correcting the radar data [Young et al. 2000]) in central New Mexico, Xie et al. (2006) noted a similar pattern with higher probability of detection in the monsoon compared to the non-monsoon season as was found in this study. McCollum et al. (2002), in another study using NEXRAD Stage III radar data, reported a decline in the detection capability of the radar data from warm summer months to the colder winter months. They attributed this to the seasonal differences in the type of rainfall events with less convection type events and more shallow, stratiform type events during the cooler winter months. The NEXRAD radar is less capable of detecting stratiform type of events due to beam overshoot.

Rainfall Estimation

An understanding of the estimation ability of the high resolution rainfall products is needed to assess the quality of these products for use as driving variables in biophysical models. The use of a product with consistent over or under estimation of rainfall could lead to erroneous predictions by the model. In this study, the NEXRAD product generally performed better than the CMORPH product in its ability to estimate rainfall. The NEXRAD product had lower estimation bias and lower variability by station location (Figures 3.3 and 3.4) and season (Figure 2.8 and 3.6). CMORPH in this study overestimated rainfall by 58% (2.9 mm/event overestimate on average) in the monsoon season and by 29% (1 mm/event overestimate on average) in the non-monsoon season (Figure 2.8) compared to the NEXRAD radar product that had overestimates of 1.3 percent in monsoon months (0.08 mm/event overestimate on average) and 4.3 percent underestimate of rainfall in the non-monsoon season (0.19 mm/event underestimate on average) (Figure 2.9). In slight contrast to the results at this study site, Joyce et al. (2004) reported that in a comparison of CMORPH and radar estimates to gage data, NEXRAD radar generally underestimated rainfall across the continental US in most months and CMORPH generally overestimated rainfall in most months, especially during the summer (June to August). They also found that root mean square differences for CMORPH and radar across the continental US generally tracked each other with values of 6 to 10 mm for the summer months and 3 to 7 mm in the fall (September to November). In the study reported here, the NEXRAD radar product had much lower root mean square difference in the monsoon season (summer) than the CMORPH product (7.0 mm vs., 11.7 mm, respectively); however, the non-monsoon season had more comparable differences (6.1 mm vs. 7.0 mm, respectively). In a study comparing the performance of satellite rainfall products, rainfall prediction models, and radar for the continental US, Ebert et al. (2007) found that NEXRAD radar, when compared to gage data, generally underestimated rainfall during the winter months (January to

March) except in the central plains region of the US, and that infrared/passive microwave (IR/PMW) products (CMORPH included) also underestimated rainfall except in the mountainous regions in the northwestern US. In the summer months (June to August), both radar and IR/PMW products overestimated rainfall by 2-3 mm, on average, throughout the central and south-central US. They attributed this to rain gages possibly missing the short-lived convective storms that occur in the summertime in this region that are better detected by the radar and the IR/PMW products.

On a location basis, the NEXRAD product had variable estimation statistics at the study site even though the stations were only 20 km apart. The NEXRAD radar product overestimated rainfall by 6.7% at the West Point station and underestimated rainfall at by 9.4% the Twin China station (Figure 2.7). Estimation efficiency and root mean square differences were similar between the two station locations. For other studies on the performance of NEXRAD radar, the results are mixed with regard to overestimation or underestimation of rainfall and appear to be related to multiple issues including ground features in relation to radar location/elevation (Smith et al. 1996; Young et al. 1999), stage of NEXRAD processing (Young et al. 1999; Young et al. 2000; Jayakrishnan et al. 2004; Wang et al. 2008), storm type (Xie et al. 2006; Wang et al. 2008) and spatial resolution differences between the radar grid and the rain gage (Ciach and Krajewski 1999; Wang et al. 2008). For example, Young et al. (1999), in a study of NEXRAD Stage II performance in mountainous terrain, found consistent underestimation of precipitation by radar and attributed these differences to radar beam blockage, ground returns, and non-detection of rainfall by the radar. Jayakrishnan et al. (2004), in a study of NEXRAD Stage III in the Texas-Gulf Basin, found that the radar product generally underestimated rainfall in the first 3 years of the study (1995 to 1997) but gradually began overestimating rainfall at a larger number of sites in the last 2 years of the study (1998 to 1999). They attributed this to changes in the NEXRAD processing algorithm over time and stated that although there was this increased tendency of overestimation, the algorithms had improved during the study period with an increased

percentage of the stations having estimation biases of within ±20%. In a recent study in the Texas Hill County using the same NEXRAD product as used in this study, Wang et al. (2008) reported overall underestimates of 7% but noted that estimations were variable between the uniform and non-uniform (more spatially heterogeneous) events.

At this study site, beam blockage and radar range problems could certainly have affected the estimates. The radar coverage at the study site is from the Midland, TX radar (Figure 2.4) and the Glass Mountains are in the path between the radar and study site. With regard to radar range, the West Point station falls within the ring of coverage of the Midland radar (Figure 2.4). However, the Twin China station is just beyond the outer general range of coverage, therefore the rainfall may have been estimated from gage corrected satellite data (NOAA 2007). The differences in the statistics for these two NEXRAD grids with a relatively short distance between them provides an indication that significant local variation in precipitation estimates is an issue that may need to be considered when using this product in biophysical modeling where gage and radar coverage are sparse.

The CMORPH product was more consistent statistically between the two station locations examined in this study (Figure 2.6). However, the large overestimation of rainfall by this product would make it unsuitable for use in forage biophysical modeling at this site. The low estimation efficiency by this product is partially due to several extreme overestimates between the CMORPH estimate and that measured at the station. Overestimation of rainfall by CMORPH and other satellite-derived products has been attributed to detection of rainfall by the satellite, but the rainfall evaporates before reaching the ground surface (McCollum et al. 2000; McCollum et al. 2002; Janowiak 2005). This is most evident in summer months in the arid and semiarid portions of the US when convective storms are more prevalent. McCollum et al. (2000) state that cloud bases for convective clouds in drier areas are generally higher than those formed in

moist environments which can lead to greater evaporation of the falling rain, thus leading to less precipitation reaching the soil surface.

Although CMORPH consistently overestimated rainfall throughout the year at this study site, there was a substantial difference in monsoon versus non-monsoon rainfall estimation which could be attributed to rainfall evaporation during convective storms. During the monsoon season when convective storms were prevalent, CMORPH overestimated rainfall by 59% compared to 29% during the non-monsoon season (Figure 2.8). Variability in the rainfall estimates was also higher in the monsoon season (Figure 2.8). The consistent overestimation of rainfall by CMORPH at this study site would make this product unsuitable for use in biophysical modeling of livestock forage. Of particular concern would be the large overestimation of rainfall in summer months during the growing season for most of the forage plant species which could lead to overestimation of forage amounts. To make this product more useful for biophysical modeling, bias corrections using the rain gage network like that done for NEXRAD products could reduce overestimation.

CHAPTER III
USE OF HIGH RESOLUTION NEXRAD RAINFALL IN BIOPHYSICAL MODELING OF FORAGE BIOMASS ON RANGELANDS: AN EVALUATION OF MODEL PERFORMANCE

Introduction

In order to properly assess the number of grazing animals that can be allocated to a given area for a given amount of time (i.e. stocking rate), an assessment of the vegetation resource is needed to determine the amount of plant biomass available for grazing (Holechek et al. 1995). In order to maximize the number of grazing animals, while at the same time minimizing the impacts to the vegetation resource, one would need to conduct periodic assessments of the vegetation resource so that animal numbers could be modified given the condition of the resource (e.g., reduce numbers during drought). However, the process of determining carrying capacity and stocking rate is time consuming, and for large areas, can be very costly and labor intensive. To overcome some of these issues, rangeland scientists have developed methods to predict plant biomass using models. These include relatively simple correlative models, more complex biophysical models, remote sensing techniques, or various combinations of these.

Correlative models are generally simple models that use one or more variables to predict biomass using regression methods. For example, O'Connor et al. (2001), using a stepwise regression model, found that precipitation and species composition accounted for 66% of the variation in plant biomass in a semi-arid grassland in South Africa. In the California annual grasslands, Duncan and Woodmansee (1975) reported that rainfall in October, December and May had the highest predictability when used in a linear regression to predict grass biomass. Drawbacks of these types of correlative models include lack of confidence in the predictions when the bounds of the original data are

exceeded (Ott and Longnecker 2001) and the inability to delineate the processes underlying the prediction (Grant et al. 1997). These models also have limitations when applying them at sites other than where they were originally developed (Bouraoui and Wolfe 1990).

With the increased capacity and accessibility to computers and programming languages in recent years, the development of more complex models for simulating biomass production on rangelands has occurred. These simulation models have differing levels of complexity and many are designed to not only simulate biomass production, but to examine other aspects such as hydrology, erosion, livestock production, and/or economics in an integrated, interacting framework. This framework allows users to examine ecosystem processes and management alternatives, and to predict response to differing alternatives (Wight and Skiles 1987; Bouraoui and Wolfe 1990; Carlson and Thurow 1996).

Examples of these multipurpose rangeland simulation models include the Ekalaka Rangeland Hydrology and Yield Model (ERHYM-II) (Wight and Neff 1983) , the Simulation of Production and Utilization of Rangelands (SPUR) model (Wight and Skiles 1987), and the Phytomass Growth Simulator (PHYGROW) model (Stuth et al. 2003a). The ERHYM-II model simulates biomass production using a relationship between actual-to-potential evapotranspiration and potential biomass yield where biomass yield is maximized when actual-to-potential evapotranspiration is 1 (water non-limiting conditions) (Wight and Neff 1983). SPUR is a physically-based model having integrated climate, hydrology, plant, animal, and economic modules and had been modified over time to improve functionality (Carlson and Thurow 1992; Hanson et al. 1992; Foy et al. 1999; Pierson et al. 2001). It has been evaluated in several different locations to ascertain its ability to predict biomass production in both pasture (Corson et al. 2006) and rangeland plant communities (Pierson et al. 2001; Teague and Foy 2004). The PHYGROW model simulates biomass production, selective grazing by multiple

kinds/classes of livestock, and changes in stocking rates brought about by changing forage conditions. PHYGROW has been used as part of bioeconomic studies for climate change (Butt et al. 2005), forage forecasting (Alhamad 2002; Alhamad et al. 2007) and is the foundation model for the regional drought early warning system on grazinglands in East Africa (Stuth et al. 2003b; Ryan 2005; Stuth et al. 2005).

With the increased availability of satellite remote sensing data, models have been developed that use remote sensing products to predict biomass on rangelands. As with the biophysical simulation models, the level of complexity varies among the different models with some using a strictly correlative approach, whereas others combine biophysical modeling with remote sensing products as inputs. Correlative models generally involve prediction of biomass using regression relationships between the remote sensing product and biomass collected from ground measurements (e.g., Dungan 1998; Thoma et al. 2002; Al-Bakri and Taylor 2003; Frank and Karn 2003; Kogan et al. 2004). For example, Tucker et al. (1983) used both a linear and logarithmic regression between the Normalized Difference Vegetation Index (NDVI) and ground collected biomass data to predict biomass in the Sahel region of Senegal. The more complex models have involved the combination of biophysical models and various remote sensing inputs. For example, Reeves et al. (2001) used fraction of absorbed photosynthetically active radiation (fAPAR) and leaf area index (LAI) products from the MODIS system with a light use efficiency model to estimate aboveground net primary productivity (ANPP) on rangelands in the northwestern US. Hunt and Miyake (2006) used a similar light use efficiency model to estimate biomass and available forage in order to predict stocking rates within 1 km^2 cells for the entire state of Wyoming.

The increased availability of remote sensing products for use as driving variables in biophysical simulation modeling offers many opportunities for monitoring and decision support on rangelands. Advances in remote sensing products that predict rainfall on a spatially explicit basis are especially attractive since reporting rain gauges

in rangeland areas are sparse and rainfall is generally one of the major factors limiting forage productivity for livestock. In the previous chapter (Chapter II), two remotely sensed rainfall products were examined for their ability to detect and estimate rainfall, and to assess their suitability for use in biophysical modeling on rangelands. The Next Generation Weather Radar system (NEXRAD) (Fulton et al. 1998) product was determined to be the most suitable because of lower estimation bias, higher detection ability, and higher estimation efficiency when compared to the CMORPH (Joyce et al. 2004) rainfall product. From a modeling standpoint, the NEXRAD product has been used mostly for hydrological studies and monitoring (e.g., Young et al. 2000; Ajami et al. 2004; Moon et al. 2004; Kalin and Hantush 2006). To date, few studies have been conducted to evaluate the use of NEXRAD data as a driving variable for biophysical modeling of forage biomass on rangelands.

For this study, the overall goal was to evaluate the ability of a biophysical simulation model (PHYGROW) to accurately predict herbaceous biomass on a heterogeneous semiarid landscape in the Trans-Pecos ecoregion of Texas using the NEXRAD rainfall product as a driving variable. The objectives of the study were to: 1) calibrate the PHYGROW model and evaluate its performance for predicting herbaceous biomass at two ungrazed sites (patches) using rainfall measured from automated rain gages located at the sites; 2) evaluate the calibrated PHYGROW model's performance when NEXRAD rainfall was substituted for that measured at the sites and compare to simulation results for gage-measured rainfall; and 3) evaluate the model's performance using NEXRAD data at multiple grazed locations representing the dominant plant communities across the study area landscape.

Methods

Study Area

The study was conducted on the Catto-Gage Ranch, approximately 13 km west of Marathon, TX (30°12'23.90"N, 103°14'47.26"W; Figure 2.1) in Brewster county. The Catto-Gage is a 69852 ha (172,609 acre) working ranch that has been grazed by livestock since the mid 1880's. Historically, the ranch was primarily a cow calf operation. However after an extended drought during 1999 to 2002, the operation was changed from a cow-calf operation to a yearling cattle stocker operation to provide greater flexibility in the livestock and grazing management (Don Keeling, personal communication). The study area is currently grazed all months except October and November at stocking rates determined by the ranch management.

The study area is situated with the Glass Mountains in the northern part and the Del Norte Mountains on the western side, with the majority of the land area falling within the greater Marathon Basin (Figure 3.1). The area consists of "high plateaus, rugged peaks and sierras, and broad, shallow intermontane valleys" (Smith 2001). Elevation at the study area follows a general northwest to southeast gradient with the highest elevations in the northwestern portion and the lowest elevations in the southeastern portion (Figure 3.1). Elevation at the site ranges from 988 m to 1940 m.

The climate of the area is semiarid with cool, dry winters and hot summers. Temperatures range from an average low of 4º C during the winter months (December to March) to an average high of 33º C during the summer months (June to September) (Figure 2.3). Average monthly temperature across all months is 17º C (NCDC 2006). Precipitation averages 369 mm with the highest amount occurring during the summer months (average of 58 mm/month) (Figure 3.2). On average, the lowest amount of precipitation occurs during March (6 mm) (Figure 3.2). Precipitation during the summer

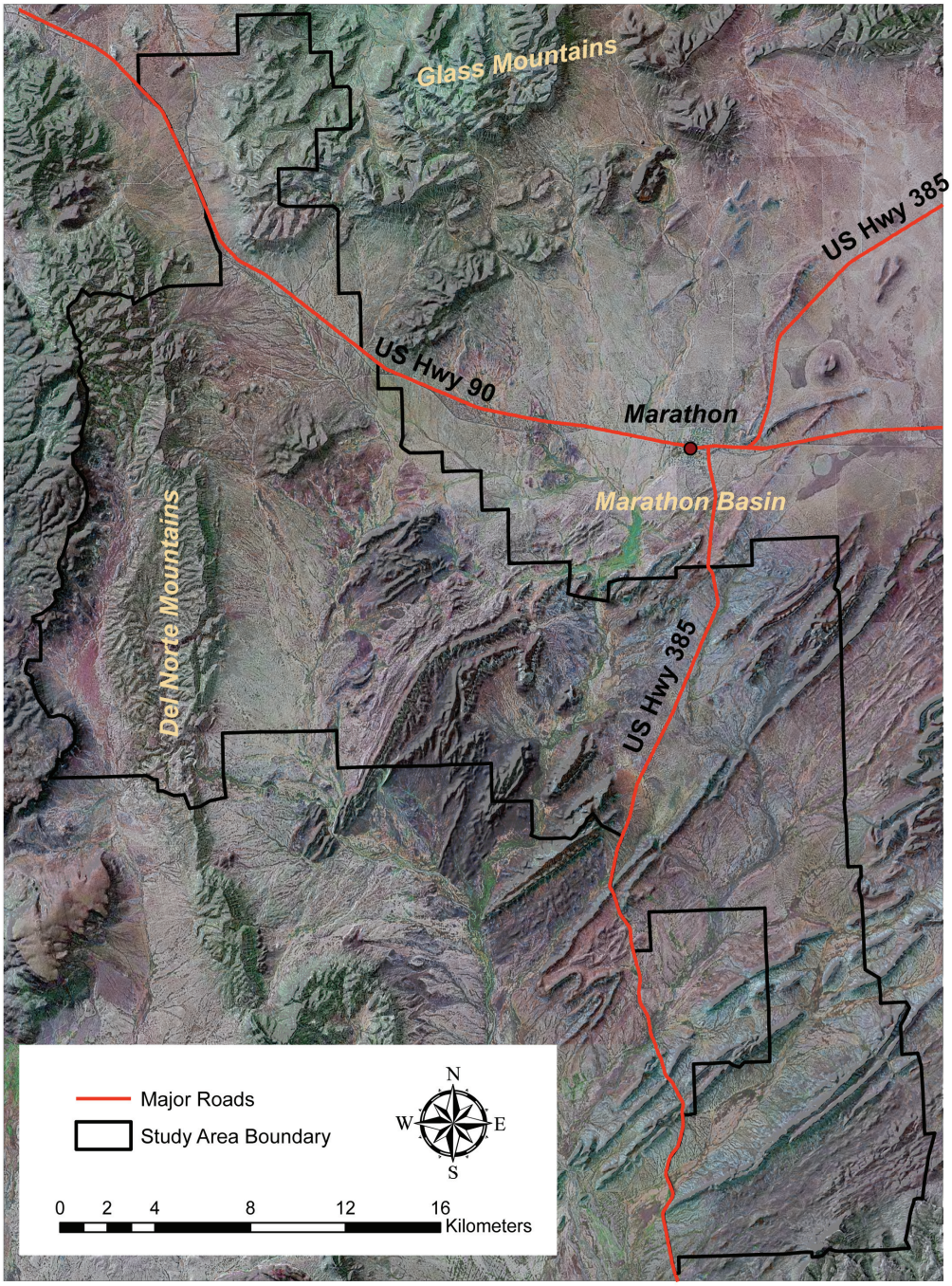

Figure 3.1. A Landsat Enhanced Thematic Mapper Geocover mosaic image (MDA Federal 2004) draped over a hillshade representation of the 30-m Digital Elevation Model (USGS 1999) to depict the terrain and changing elevation in the study area and surrounding environment.

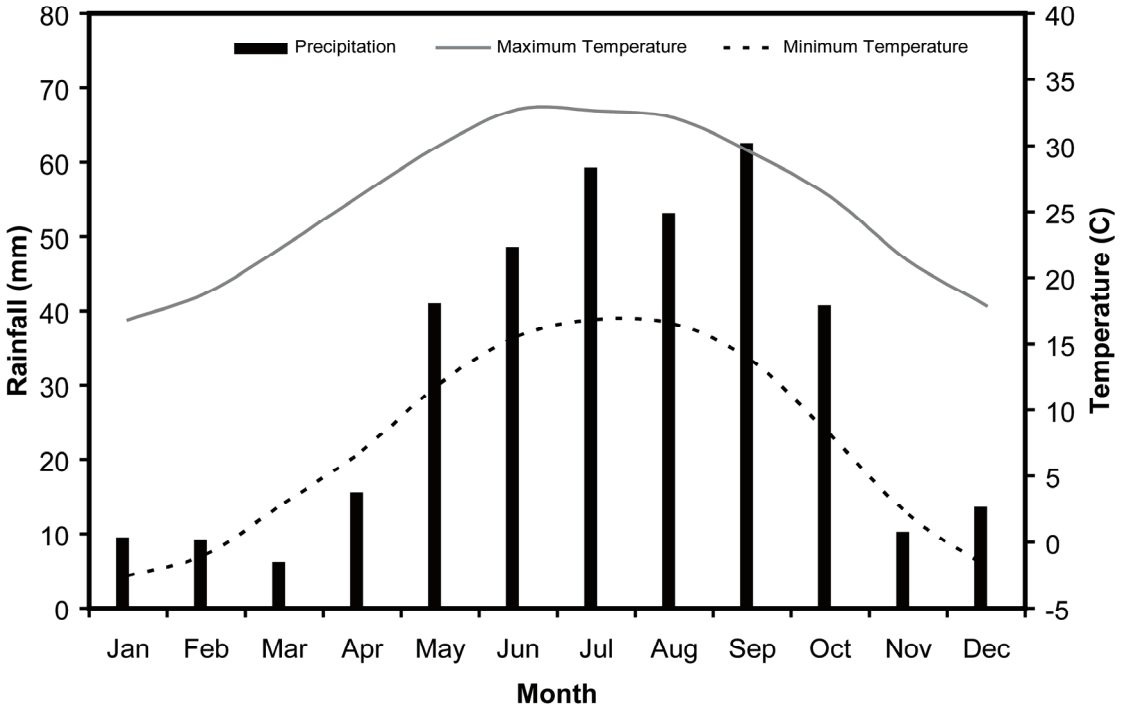

Figure 3.2. Average monthly rainfall (mm), average maximum temperatures (ºC), and average minimum temperatures (ºC), during the period from 1970 to 2000 at the official recording station closest to the study area (Marathon, TX) (NCDC 2006).

months is generally from thunderstorms (Powell 1998) that can be of high intensity. Snowfall can occur during the winter months, but this form of precipitation generally averages less than 2 cm during the winter months (NCDC 2006).

Native vegetation in the study area is diverse due to the strong elevational gradient (Figure 3.1) and differing soil parent materials. Powell (1998) describes 5 distinct vegetation types that can occur in the Trans-Pecos region of Texas: 1) Oak-Pinyon-Juniper Woodlands, 2) Grasslands, 3) Chihuahuan Desert Scrub, 4) Conifer Forest, and 5) Riparian Communities. Of these, the first three can be found in the study area. In the Glass and Del Norte Mountains (northern and northwestern portions of the study area), the Oak-Pinyon-Juniper woodlands predominate. In elevations ranging from 1200 to 1500 m, pinyon pine (*Pinus ponderosa*), Mohr oak (*Quercus mohriana*) and redberry juniper (*Juniperus coahuilensis*) are common tree and shrub species. Black

grama (*Bouteloua eriopoda*), bush muhly (*Muhlenbergia porteri*), hairy grama (*Bouteloua hirsuta*), and sideoats grama (*Bouteloua curtipendula*) are common grasses. The dominant soil in these areas is the Altuda series (Soil Survey Staff 2007).

The valleys, plains, and basins in the study area can be divided into two distinct regions that represent the Grassland and the Chihuahuan Desert Scrub vegetation types. In the north central and central portion of the study area, the Grassland type is present as these areas are mostly open grasslands interspersed with small shrubs. In the south central and southern part of the study area, the Chihuahuan Desert Scrub vegetation type is present with the valley and plains vegetation being primarily shrublands, with some open grasslands in basins. These distinct vegetation types are primarily caused by the increased aridity along the elevational gradient in the study area.

In the Grassland vegetation type found in the central portion of the study area, the dominant grass species are bush muhly (*Muhlenbergia porteri*) cane bluestem (*Bothriochloa barbinodis*), burrograss (*Scleropogon brevifolius*), hairy grama (*Bouteloua hirsuta*), black grama (*Bouteloua eriopoda*) and sideoats grama (*Bouteloua curtipendula*). Dominant forbs include gray globemallow (*Sphaeralcea incana*), desert eveningprimrose (*Oenothera primiveris*) and dogweed species (*Dyssodia spp.*). Common shrubs include gregg dalea (*Dalea greggii*), feather dalea (*Dalea formosa*), javelinabush (*Condalia ericoides*), broom snakeweed (*Gutierrezia sarothrae*) and agarito (*Mahonia trifoliolata*). Dominant soils include the Crossen, Cienega, Stovall, Boracho, Paisano, Espy and Musquiz series (Soil Survey Staff 2007).

In the southern and south-central region of the study area where the Chihuahuan Desert Scrub is the dominant type, shrub species include creosotebush (*Larrea tridentata*) tarbush (*Flourensia cernua*), catclaw acacia (*Acacia greggii*), mesquite (*Prosopis glandulosa*), whitebrush (*Aloysia gratissima*), mariola (*Parthenium incanum*) and range ratany (*Krameria erecta*). Grasses include tanglehead (*Scleropogon*

brevifolius), hairy grama (*Bouteloua hirsuta*), black grama (*Bouteloua eriopoda*) sideoats grama (*Bouteloua curtipendula*), low woollygrass (*Dasyochloa pulchella*) and purple threeawn (*Aristida purpurea*). Common soils in this area include the Bullis, Catto, Crossen, Paisano, and Stovall series (Soil Survey Staff 2007).

Simulation Model

The Phytomass Growth Simulation Model (PHYGROW) (Stuth et al. 2003a) was selected for evaluation to predict herbaceous biomass at the study site. PHYGROW is a point model that contains 4 integrated submodels: climate, soil, plant growth and grazing. The model simulates a soil water balance, multi species/functional group plant growth, and livestock grazing on a daily time step. PHYGROW, at its core, is a light use efficiency model (Montieth 1972; Montieth 1977) that simulates plant growth under water non-limiting (optimal conditions). The model then discounts plant growth based on the amount of water stress, temperature stress, and livestock grazing demand based on the input climate variables and model parameters.

The model contains parameters for soil surface and layer information, plant species and community data, livestock grazing management and stocking rates, and is driven by daily climate data (Stuth et al. 2003b). The soil subcomponent of the model has 13 unique parameters that include soil depth, bulk density, infiltration, and water holding capacity variables. The plant subcomponent can be parameterized for individual species or functional groups. Plant community composition parameters include initial standing crop, percent basal cover for grasses, frequency of forbs, and canopy cover of shrubs and trees. For each individual plant species/functional group in the model, there are 27 parameters including minimum, optimal and maximum temperatures for growth, radiation use efficiency, leaf area index, leaf and wood turnover, leaf and wood decomposition, and canopy water movement. The grazing subcomponent of the model has 19 variables related to each kind/class of grazing animal including forage intake,

stocking rate, and grazing preference class for each plant species parameterized in the model. Lastly, the climate subcomponent has 6 variables which include year, day, maximum and minimum temperature, rainfall and solar radiation.

Model Parameterization and Evaluation at Weather Stations

To gather the necessary plant community parameters for the PHYGROW model, a permanent 100-m vegetation transect was established within each exclosures and near (< 100 m) the automated weather stations (Figure 3.3). Along each transect, a modified point-frame method (Ryan 2005) was used to collect percent basal cover of grasses, frequency of forbs, and shrub canopy cover. This was done by placing the modified point frame on the soil surface at 1-m increments along the transect. Each point on the frame was examined to determine if the point intersected the basal area of a grass species, plant litter, bare ground, or rock. If a basal area of a grass species was encountered, this was recorded as a "hit". Within a 5 x 5-cm quadrat around each point, each presence of a unique forb species was defined as a "hit". If a shrub or tree canopy intersected in an upward, perpendicular line from the point, the shrub or tree species was recorded as a "hit". Along each transect, a total of 500 points were sampled. The "hits" of grass, forbs, and shrub/tree species were divided by the total possible hits (500) and these values were entered as the plant community composition variable in the PHYGROW model (Table 3.1).

Herbaceous biomass at each transect was measured at the time of transect establishment and approximately every 3 to 4 months thereafter during the period from March 2004 to January 2007. A 0.25 or 0.50-m^2 quadrat was placed at 10-m increments along the 100-m transect (n=10 sample size per transect). Within each quadrat, the herbaceous biomass (grass and forbs, but not shrubs or trees) was clipped to a 1-cm stubble height. The clipped biomass was placed in paper bags and taken back to the

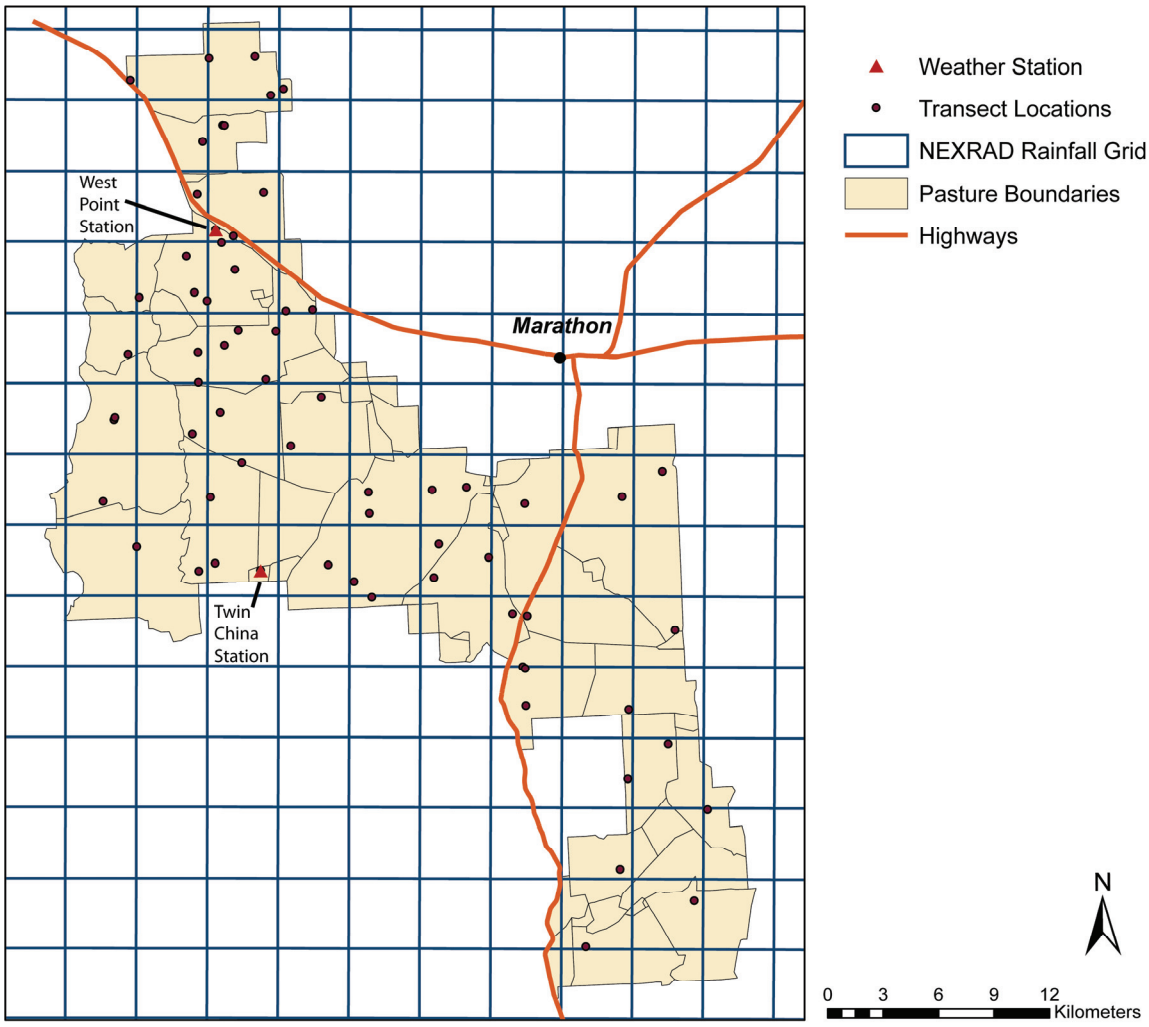

Figure 3.3. Location of automated weather stations and transects used for evaluation of the PHYGROW simulation model for predicting herbaceous biomass on rangeland near Marathon, TX.

Table 3.1. The percent plant community composition measured on transects located at the West Point and Twin China weather stations near Marathon, TX.

Species/ Functional Group	Scientific name	Growth Habit	Community Composition (%)
West Point Station			
desert baileya	*Baileya multiradiata*	Forb	0.58
cane bluestem	*Bothriochloa barbinodis*	Grass	0.19
sideoats grama	*Bouteloua curtipendula*	Grass	0.39
black grama	*Bouteloua eriopoda*	Grass	2.72
hairy grama	*Bouteloua hirsuta*	Grass	1.36
cool season forb		Forb	2.14
low woollygrass	*Dasyochloa pulchella*	Grass	0.19
broom snakeweed	*Gutierrezia sarothrae*	Shrub	1.94
curlymesquite	*Hilaria belangeri*	Grass	1.17
bush muhly	*Muhlenbergia porteri*	Grass	1.17
desert eveningprimrose	*Oenothera primiveris*	Forb	2.14
prickly pear species	*Opuntia*	Shrub	1.75
wooly plantain	*Plantago patagonica*	Forb	0.19
slim tridens	*Tridens muticus*	Grass	0.19
warm season forb		Grass	0.19
yucca species	*Yucca*	Shrub	0.19
Twin China Station			
catclaw acacia	*Acacia greggii*	Shrub	5.25
whitebrush	*Aloysia gratissima*	Shrub	3.00
purple threeawn	*Aristida purpurea*	Grass	0.38
fourwing saltbush	*Atriplex canescens*.	Shrub	0.38
sideoats grama	*Bouteloua curtipendula*.	Grass	0.38
slender grama	*Bouteloua repens*	Grass	0.19
cool season forb		Forb	0.19
low woollygrass	*Dasyochloa pulchella*	Grass	0.19
vine ephedra	*Ephedra pedunculata*	Shrub	1.50
tarbush	*Flourensia cernua*	Shrub	7.88
elbowbush	*Forestiera pubescens*	Shrub	0.38
range ratany	*Krameria erecta*	Shrub	1.13
creosotebush	*Larrea tridentate*	Shrub	6.57
bush muhly	*Muhlenbergia porteri*	Grass	0.38
tasajillo	*Opuntia leptocaulis*	Shrub	0.94
halls panicum	*Panicum hallii*	Grass	0.19
mesquite	*Prosopis glandulosa*	Shrub	3.19

laboratory and dried in a forced air oven at 60º C for 48 hours. After drying, the samples were weighed with a digital scale. The sample weights were then multiplied by the appropriate plot factor in relation to the quadrat size to convert the biomass to kg/ha units. The 10 samples were averaged and the mean was used for comparison to the simulation model output for each sampling date.

To parameterize the soil components in PHYGROW for each transect location, the soil series was identified using the digital version of the Brewster County Soil Survey (Soil Survey Staff 2007) for the latitude and longitude of each transect. Parameters needed for PHYGROW were extracted from the soil survey database for each soil series. When a needed parameter was missing from the soil survey data, the Map Unit Use File software (MUUF) (Baumer et al. 1987) was used to estimate the parameter.

The hourly weather data collected at the weather stations was processed to produce a daily climate dataset for each transect site. Daily climate variables included minimum temperature, maximum temperature, rainfall, and solar radiation. The hourly rainfall data for each station were summed to match the production schedule for the NEXRAD rainfall product (see Chapter II for details).

The calibration procedure for PHYGROW involved running the model with the climate data and comparing the modeled herbaceous biomass output to that measured during the first 2-3 biomass clipping dates in time sequence. If the model output fell within ± 1 standard error of the mean for the herbaceous biomass measured on the transect, the model was considered calibrated. If the model output fell outside ± 1 standard error of the measured data, parameters were adjusted to in an attempt to move the modeled biomass estimate to within the standard error. This process was repeated for each time period data was collected until the model was considered calibrated. Parameter adjustments were generally limited to species maximum rooting depths, green

and dead leaf turnover rates, and surface soil layer thickness (influences depth of soil water evaporation). After the model was considered calibrated, the parameters were no longer adjusted and the data were used to evaluate model performance during subsequent herbaceous biomass clipping events (model verification).

For the weather station sites, two different modeling scenarios were evaluated. The first was to conduct simulations using the actual rainfall measured at the sites to evaluate the model performance in two very different plant communities (Table 3.1). In this scenario, the PHYGROW model was calibrated and the model output for herbaceous biomass was compared to that clipped, over time, in the transects adjacent to the weather station for the period from March 2004 to January 2007. The second scenario was designed to assess how well the calibrated model performed at each site using NEXRAD rainfall (extracted from the appropriate NEXRAD grid cell) in place of the weather station rainfall (for information on the NEXRAD product and procedures for acquiring and extracting the data, see Chapter II). All other parameters and climate data were kept the same as that used in the first scenario. Each scenario was evaluated using the statistics described below.

The means and standard deviations of the simulated and observed herbaceous biomass were calculated and linear regression was used to examine model predictive strength (Carlson and Thurow 1996). Difference statistics were calculated to examine bias and variability between the simulated and observed data. These statistics included percent estimation bias (BIAS), mean bias error (MBE), mean absolute error (MAE), and root mean square difference (RMSD). Estimation bias reflects the normalized difference between the simulation model output and the observed data and is expressed as follows:

$$BIAS(\%) = \frac{\overline{P} - \overline{O}}{\overline{O}} \times 100 \qquad [3.1]$$

where $\overline{P}$ is the mean of the simulation model predictions and $\overline{O}$ is the mean for the observed predictions. Positive estimation bias values indicate the overestimation of biomass by the simulation model whereas negative values indicate the opposite. Mean bias error provides an indication of the average magnitude of the over-prediction or under-prediction by the simulation model in the units of the biomass (kg/ha) (Andales et al. 2005). It is calculated as:

$$MBE = \frac{\sum_{i=1}^{n}(P_i - O_i)}{n} \quad [3.2]$$

where P_i is the i^{th} predicted value, O_i is the i^{th} for observed value and n is the number of simulated and observed data pairs. Mean absolute error provides an indication of the average absolute difference between the simulated and observed values in the series of data pairs being evaluated and is calculated as (Legates and McCabe Jr. 1999):

$$MAE = \frac{\sum_{i=1}^{n}|P_i - O_i|}{n} \quad [3.3]$$

Root mean square difference (RMSD) is a measure of the average magnitude of the difference between the simulation and observed biomass data in the units of the data (kg/ha). RMSD is similar to MAE error, however it is more sensitive to extreme differences between the simulation and observed data (Willmott 1982). It is generally greater than MAE and the degree of difference is related to the number of outliers in the data (Legates and McCabe Jr. 1999). RMSD is calculated as follows:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{n}} \quad [3.4]$$

Relative error measures (goodness-of-fit measures) were also used to evaluate performance of the PHYGROW simulation model at the weather station sites. These included estimation efficiency (EE) (Nash and Sutcliffe 1970; Legates and McCabe Jr. 1999) and the index of agreement (*d*) (Willmott. et al. 1985; Legates and McCabe Jr. 1999). Estimation efficiency is a measure of the deviation from a 1:1 line between simulation model output and the observed data and is calculated as:

$$EE = 1.0 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2} \quad [3.5]$$

An EE value of 1 would reflect a perfect correspondence between the simulated output and the measured data. Values greater than 0 would indicate that a positive relationship exists between the simulation output and the observed data and that the simulation data could be used as a good estimate for the location where the observed data was collected. Values less than 0 indicate a low correspondence between the simulation output and the observed data suggesting that the mean of the observed data would serve as a better predictor than the simulation model output (Legates and McCabe Jr. 1999; Moon et al. 2004). The index of agreement is measure of the tightness between the simulation predictions and the observed data to a 1:1 line (Willmott. et al. 1985; Andales et al. 2005) and is expressed as follows:

$$d = 1.0 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}\left(\left|P_i - \overline{O}\right| + \left|O_i - \overline{O}\right|\right)^2} \quad [3.6]$$

Values of *d* can range from 0 to 1 with a 1 indicating perfect agreement between the simulation output and the observed data.

Model Evaluation of Grazed Location Simulations Using NEXRAD

An additional 60 sites were selected across the study area (Figure 3.3) to evaluate the ability of the PHYGROW model to predict herbaceous biomass at the patch scale using the NEXRAD rainfall product. These sites were dispersed across the study area and located in the major plant communities. The NRCS ecological site map (Soil Survey Staff 2007) was used, along with Landsat satellite imagery, to identify possible sample locations prior to going to the field. Once in the field, the areas identified from the soil and Landsat maps were visited and a determination was made in the field as to whether the site was suitable for sampling. Once a site was identified, a GPS was used to record the start and end of a 100-m transect. Plant species composition and cover, along with

herbaceous biomass data, were collected from each transect using the modified point frame procedure described above for the weather station locations. These data were used to parameterize the plant communities for the PHYGROW simulations.

After transects were established at each of the 60 sites, the majority of the site were revisited at least once during the period from March 2004 to March 2007 to collect additional herbaceous biomass measurements for model evaluation. The methodology for collecting herbaceous biomass was the same as described above for the weather station locations.

Soil data for parameterization of the PHYGROW model simulations at the 60 sites was extracted from the Brewster County, TX soil survey in the same manner as that described above for the weather station locations. Since these sites were grazed by cattle, the grazing module of PHYGROW was parameterized using the pasture stocking rate information provided by the Catto-Gage Ranch management.

The simulations for the grazed sites were conducted using NEXRAD data as the rainfall source. The same procedure for calibration described above for the weather station locations was used for each site. Because the number of sample dates for model evaluation at each site was small (n=2 to 5), the data pairs for simulated and observed herbaceous biomass were pooled. The performance of the PHYGROW model was then assessed on the pooled data set using the statistics described above for the weather stations.

Results

Simulation with Measured Rainfall

At the West Point weather station location (Figure 3.3), the plant community can be described as a desert grassland with black grama (*Bouteloua eriopoda*), bush muhly (*Muhlenbergia porteri*), and curlymesquite (*Hilaria belangeri*) as the dominant grass species (Table 3.1). The simulation for this plant community, using rainfall collected from the weather station at the site, required the first three sampling dates to calibrate the model (Figure 3.4A). After this, model parameters were no longer adjusted. For the entire time series, the average herbaceous biomass measured at the site was 1054 kg/ha whereas the average predicted by the PHYGROW model was 964 kg/ha. This resulted in an overall 8.57% under-prediction in herbaceous biomass by the model (Table 3.2). The standard deviation for the simulated herbaceous biomass (sd_s) was slightly higher than that for the observed biomass (sd_o) showing there was more variability in the model estimates for these observation dates (Table 3.2). The root mean square difference (RMSD) between the simulated and observed data was 246 kg/ha, which was only slightly higher than the mean absolute error (MAE) (213 kg/ha; Table 3.2), providing an indication that there were few extreme differences between the model output and the herbaceous biomass at this site. However, the model did have a tendency to under-predict biomass during the spring (May) in 2004 and during mid-winter (January) in both 2006 and 2007 (Figure 3.4A).

Linear regression indicated a reasonable correspondence between the simulated biomass and that measured at the site with the 72% of the variability in the measured biomass explained by the simulation model results (Table 3.2). Model estimation efficiency (EE) and the index of agreement (d) also provided evidence that the simulation model had moderate to good skill in predicting herbaceous biomass at the site (EE = 0.58 and d = 0.91; Table 3.2).

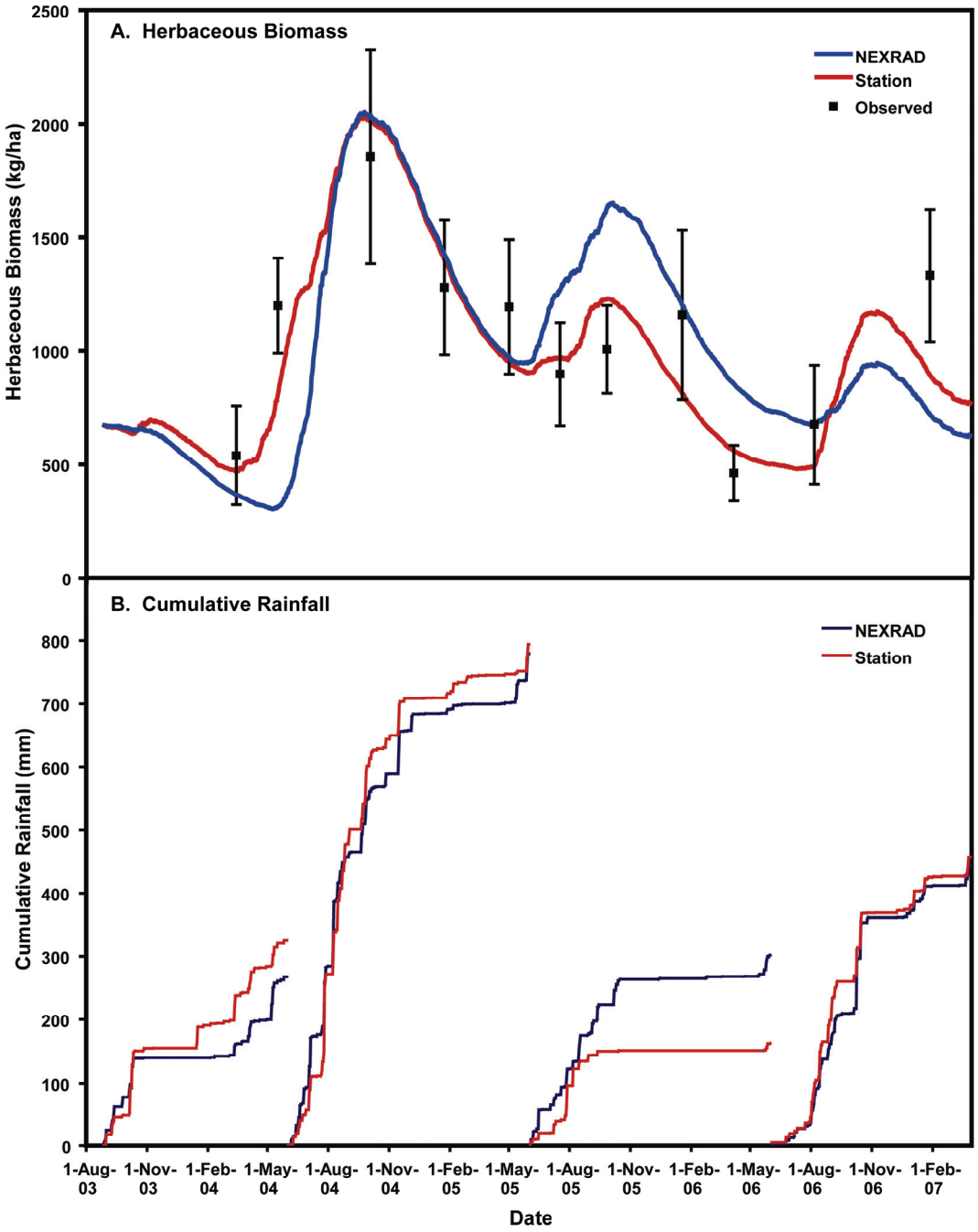

Figure 3.4. Comparison of A) observed mean herbaceous biomass (kg/ha) measurements to the herbaceous biomass predicted by the PHYGROW simulation model and the corresponding B) cumulative rainfall (mm) for two different modeling scenarios at the West Point weather station location. Scenarios evaluated were 1) simulation using rainfall collected at the study site (Station) and 2) simulation using NEXRAD rainfall for the study site location (NEXRAD). Bars on observed values represent the standard error of the mean.

Table 3.2. Statistics for evaluation of the PHYGROW model's ability to simulate herbaceous biomass production (kg/ha) at the West Point study site near Marathon, TX. Two different modeling scenarios were evaluated: 1) simulation using rainfall collected at the study site (station simulation) and 2) simulation using NEXRAD rainfall for the study site location (NEXRAD simulation).

	Scenario 1	**Scenario 2**
Statistic	**Station Simulation vs. Observed**	**NEXRAD Simulation vs. Observed**
Observed Mean (kg/ha)	1054	1054
Simulated Mean (kg/ha)	964	1041
[1]sd_o (kg/ha)	401	401
sd_s (kg/ha)	452	532
Bias (%)	-9	-1
MBE (kg/ha)	-90	-13
MAE (kg/ha)	213	332
RMSD (kg/ha)	246	424
r^2	0.72	0.33
EE	0.58	-0.23
d	0.91	0.75
n	11	11

[1]sd_o = standard deviation for observed; sd_s = standard deviation for simulation; MBE = Mean Bias Error; MAE = Mean Absolute Error; RMSD = Root Mean Square Difference; r^2 = coefficient of determination; EE = estimation efficiency; *d* = index of agreement; *n* = number of samples

At the Twin China site (Figure 3.3), the plant community can be characterized as a creosotebush (*Larrea tridentata*) and tarbush (*Flourensia cernua*) shrubland interspersed with grasses such as sideoats grama (*Bouteloua curtipendulata*) and bush muhly (*Muhlenbergia porteri*) (Table 3.1). The observed herbaceous biomass data during the study period was highly variable with standard errors for the sampling dates ranging from 70 to 300 kg/ha. This was due to the large amount of bareground between the herbaceous plant species growing at the site. Over the time series, the observed herbaceous biomass averaged 260 kg/ha (Table 3.3).

The simulation for the Twin China site, using rainfall data collected from the weather station, also required the first three collection dates to calibrate the model (Figure 3.5A). The herbaceous biomass predicted by the simulation model averaged 254 kg/ha exhibiting a slight negative bias of approximately 2% (Table 3.3). The RMSD during the time series was 74 kg/ha (Table 3.3). The greatest differences between the

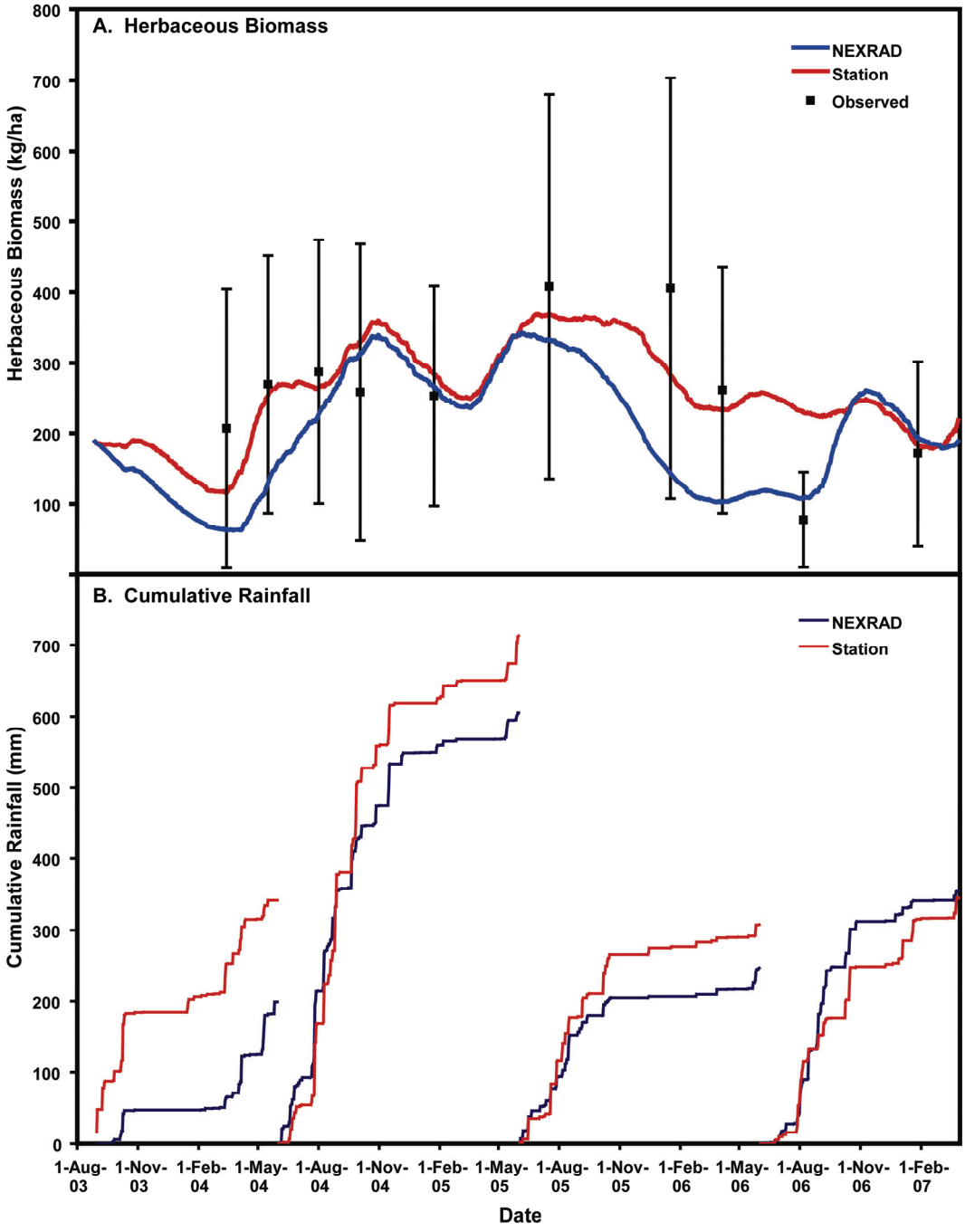

Figure 3.5. Comparison of A) observed mean herbaceous biomass (kg/ha) measurements versus the herbaceous biomass predicted by the PHYGROW simulation model and B) the corresponding cumulative rainfall (mm) for two different modeling scenarios at the Twin China weather station location. Scenarios evaluated were 1) simulation using rainfall collected at the study site (Station) and 2) simulation using NEXRAD rainfall for the study site location (NEXRAD). Bars on observed values represent the standard error of the mean.

Table 3.3. Statistics for evaluation of the PHYGROW model's ability to simulate herbaceous biomass production (kg/ha) at the Twin China study site near Marathon, TX. Two different modeling comparisons were evaluated: 1) simulation using rainfall collected at the study site (station simulation) and 2) simulation using NEXRAD rainfall for the study site location (NEXRAD simulation).

Statistic	Scenario 1 Station Simulation vs. Observed	Scenario 2 NEXRAD Simulation vs. Observed
Observed Mean (kg/ha)	260	260
Simulated Mean (kg/ha)	254	188
[1] sd_o (kg/ha)	99	99
sd_s (kg/ha)	71	93
Bias (%)	-2	-28
MBE (kg/ha)	-5	-72
MAE (kg/ha)	58	96
RMSD (kg/ha)	74	121
r^2	0.38	0.18
EE	0.37	-0.68
d	0.74	0.60
n	10	10

[1] sd_o = standard deviation for observed; sd_s = standard deviation for simulation; MBE = Mean Bias Error; MAE = Mean Absolute Error; RMSD = Root Mean Square Difference; r^2 = coefficient of determination; EE = estimation efficiency; *d* = index of agreement; *n* = number of samples

model predictions and the observed biomass occurred during January 2006 when the model under-predicted biomass by 122 kg/ha and during August 2006 when the model over-predicted biomass by 151 kg/ha (Figure 3.5A).

Linear regression indicated a weak correspondence between the simulated and observed herbaceous biomass with only 38% of the variability in the observed biomass being explained by the simulation model output (Table 3.3). EE was also low (0.37); however, it does show that the simulation model was a slightly better predictor than the overall mean of the observed herbaceous biomass. The *d* index was 0.74 indicating a moderate correspondence of the simulation model predictions and the observed data to a 1:1 line.

Simulation with NEXRAD Rainfall

At the West Point location, replacement of the station rainfall with NEXRAD rainfall into the calibrated model resulted in a different pattern of herbaceous biomass production response when compared to both the observed data and the station simulation (Figure 3.4A). Although the average biomass between the NEXRAD simulation and the observed data were similar (1041 vs. 1054 kg/ha, respectively), the variability in the NEXRAD simulation predictions was almost 33% greater than the observed data as indicated by the standard deviations (Table 3.2). The variability between the simulated biomass and that measured at the site was high with a RMSD of 424 kg/ha. This was an almost 75% increase in that seen for the simulation using rainfall collected at the site (Table 3.2). Goodness-of-fit statistics were low with a linear regression r^2 of 0.33 and an EE of -0.23 (Table 3.2). The negative EE indicates that the variability in the NEXRAD simulation was so large, that the mean of the observed herbaceous biomass would be a better predictor of biomass than the NEXRAD simulation. The *d* index indicated a moderate correspondence (0.75) between the NEXRAD simulation and observed data, and was lower than that observed for the station simulation (Table 3.2).

The differences in herbaceous biomass between the simulations using NEXRAD rainfall and the rainfall collected at the site appear to be due to differences in the timing and amounts of precipitation in the NEXRAD product since all other model parameters were unchanged. An examination of the rainfall amounts during the 30 days prior to when a large divergence occurred between the model outputs for NEXRAD and the station data (e.g., October 2003, June 2005, and September 2006) reveals that the NEXRAD and the station rainfall differed by 25 to 40 mm of rainfall in each instance. However, when the NEXRAD rainfall tracked the station rainfall more closely, such as it did during the first 9 months of the 2004 to 2005 monsoon cycle (June 1, 2004 to May 31, 2005), the amount of biomass predicted by the model using NEXRAD data was almost identical to the biomass predicted by the station rainfall model. This provides an indication of the sensitivity of the model to the timing and amount of rainfall and that a

series of over- or under-predictions of rainfall by the NEXRAD product can lead to substantially different predictions in biomass.

At the Twin China site, the replacement of the station rainfall with the NEXRAD rainfall also led to a different and more variable pattern of herbaceous biomass prediction. Biomass predictions from the NEXRAD simulation were generally lower than the station rainfall biomass predictions for almost the entire time series (Figure 3.5A) with an overall estimation bias of -28% (Table 3.3). The herbaceous biomass in the NEXRAD simulation average 188 kg/ha for the time series whereas that measured on site was 260 kg/ha. The MAE and the RMSD were 96 and 121 kg/ha respectively, which represented and almost 63% increase in variability when compared to the station simulation (Table 3.3). Like that observed at the West Point site, the use of the NEXRAD rainfall in the simulation reduced the goodness-of-fit statistics. The EE statistic was most affected with a value of -0.68 for the NEXRAD simulation compared to 0.37 for the station simulation (Table 3.3). Since the EE statistic is sensitive to large outliers, the relatively large under-prediction of herbaceous biomass by the NEXRAD simulation during most of 2004 and 2006 led to this reduced EE value.

As was observed at the West Point site, the timing and amount of rainfall strongly influenced the differences seen between the NEXRAD and station rainfall simulations. During September through October 2003, the NEXRAD rainfall underestimated rainfall at the site by almost 140 mm (Figure 3.5B). However, this large difference in rainfall resulted in only a 50 kg/ha difference in the herbaceous biomass when compared to the station simulations (Figure 3.5A). Because this site is dominated by shrub species, and grasses and forbs comprise only a small proportion of the community composition (Table 3.1), the grasses and forbs have a limited potential to respond to the additional water in the model.

During the period from June 2005 to August 2006, the NEXRAD simulation also had a large deviation in herbaceous biomass when compared to both the measured biomass and the station rainfall simulation. (Figure 3.5A). NEXRAD underestimated rainfall by 26 mm during May 2005 and this appears to be what triggered the trajectory change for the NEXRAD simulation.

Simulation Using NEXRAD Rainfall on Grazed Sites

The PHYGROW model was parameterized for 60 sites (Figure 3.3) across the study area that represented the major plant communities accessible to grazing by livestock. NEXRAD rainfall data were used to drive the PHYGROW model. Since there were not enough observations from each site to examine model performance individually, the PHYGROW simulated and observed data pairs were pooled across sites for evaluation.

Approximately half of the sites required two observation dates to calibrate the model. Approximately a third of the sites required no calibration after the first data collection and 10 percent required three observation dates to calibrate. Across all sites and collection dates in the calibration data set, the observed herbaceous biomass averaged 917 kg/ha. The average herbaceous biomass predicted by the simulation model was 851 kg/ha, which resulted in a negative bias of 7% for the calibration (Table 3.4). The variability in the data for both the observed and simulated biomass was high, but was similar (sd_o = 886 and sd_s = 832; Table 3.4). The RMSD was about 63% greater than the MAE (Table 3.4), indicating that several extreme outliers were increasing the variability between the observed and simulated biomass (Figure 3.6A). These extreme outliers were mostly associated with large under-predictions of biomass (700 to 900 kg/ha) by the simulation model during the early portion of the growing season in May 2004. This same trend was also observed at the West Point site (Figure 3.4A).

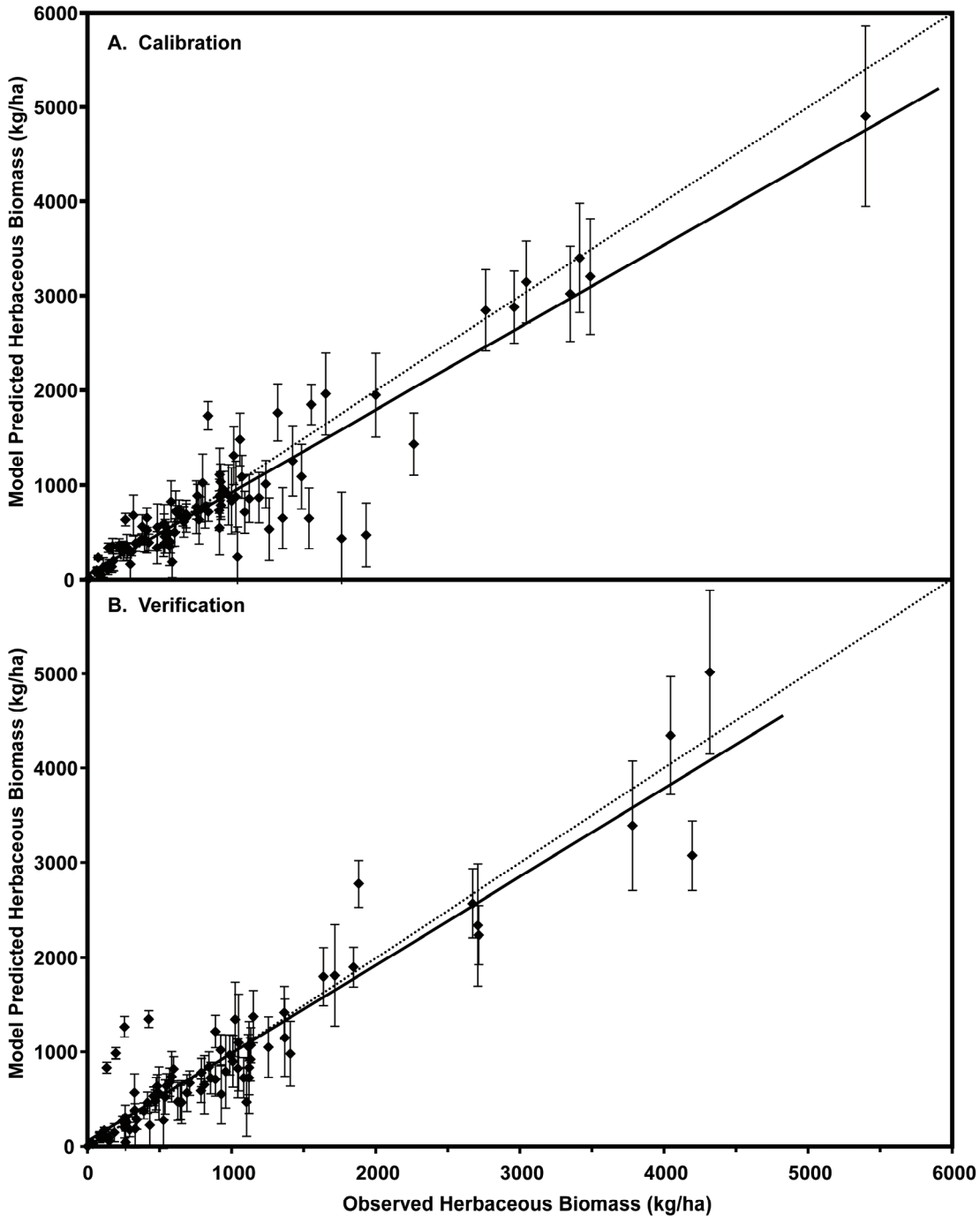

Figure 3.6. A comparison of observed mean herbaceous biomass measurements (kg/ha) to those predicted by the PHYGROW model during A) calibration and B) model verification using NEXRAD rainfall. Bars on observation mean indicate standard error.

Table 3.4. Statistics for calibration and validation performance on the ability of the PHYGROW model, using NEXRAD rainfall, to predict herbaceous biomass at multiple sites across the study area near Marathon, TX.

Statistic	Calibration	Verification
Observed Mean (kg/ha)	917	917
Simulated Mean (kg/ha)	851	907
[1]sd_o (kg/ha)	886	851
sd_s (kg/ha)	832	774
Bias (%)	-7	-1
MBE (kg/ha)	-66	-9
MAE (kg/ha)	206	208
RMSD (kg/ha)	333	320
r^2	0.86	0.88
EE	0.86	0.88
d	0.96	0.97
n	98	82

[1]sd_o = standard deviation for observed; sd_s = standard deviation for simulation; MBE = Mean Bias Error; MAE = Mean Absolute Error; RMSD = Root Mean Square Difference; r^2 = coefficient of determination; EE = estimation efficiency; *d* = index of agreement; *n* = number of samples

The relative error statistics for the model calibration across the grazed sites were good. Linear regression analysis revealed a good correspondence between the observed and simulated biomass with 86% of the variability in observed biomass explained by the simulation predictions (Table 3.4). The EE and *d* index were also quite high (0.86 and 0.97, respectively) indicating that the observed and simulated biomass pairs generally conformed to the 1:1 line.

Model verification statistics were quite similar to the calibration statistics across all grazed sites (Table 3.4). The observed herbaceous biomass was 917 kg/ha which, surprisingly, was the same as the observed mean for the calibration set. For the verification, PHYGROW continued to under-predict herbaceous biomass; however, the percentage was less than in the calibration (1% vs. 7%, respectively; Table 3.4). The variability between observed and simulated biomass in the verification was less than the calibration (RMSD of 320 kg/ha vs. 333 kg/ha, respectively). This reduced variability

could also be seen in the percent difference between the RMSD and the MAE which was 55% for the model verification (Table 3.4). The EE and r^2 statistics were both slightly higher for the model verification compared to the calibration. The *d* index did not change (Table 3.4). The model verification statistics indicate that across all sites the calibrated PHYGROW model did quite well in predicting herbaceous biomass.

Several obvious outliers can be seen in the verification data set (Figure 3.6B). One cluster of 4 data pairs was examined where PHYGROW over-predicted the biomass by more than 600 kg/ha (Figure 3.6B, lower left quadrant). Three of the 4 sites represented data collected during August 2006 and the other was collected in January 2007. Notations on the data sheets for these sites indicated moderate to heavy utilization of the herbaceous biomass by livestock. Since the grazing algorithm in PHYGROW is parameterized using stocking rates for the entire pasture, the model is not able to capture localized overgrazing that may occur on the site being modeled; therefore, this could lead to large outliers on verification sites.

Discussion

The PHYGROW simulation model's performance, using rainfall data collected at the site, was moderately low to good depending on the performance measure (EE or *d*) and the site examined (Table 3.2 and 4.3). Both the West Point and Twin China sites had sampling dates where the model had relatively large over or under-predictions of the herbaceous biomass which reduced the performance measures. Assessing the source of these differences becomes problematic since the differences were not consistent in the time series or across sites. The source of these differences may be associated with inadequacies in the model algorithms, parameterization, sampling methodology for the observed data, or combinations of these. In some cases, the model appears to lag observed conditions by 20 to 30 days (e.g., May 2004 and August 2006 at the West Point site; Fig 4.2A). This may reflect problems with the model algorithm in not being able to

respond as quickly to conditions as plants do in nature. A higher temporal frequency in biomass sampling coupled with soil moisture monitoring may be an approach to gathering the data needed for model improvement.

The response of the PHYGROW model to the replacement of the site measured rainfall with NEXRAD was surprising. The NEXRAD simulations resulted in very different temporal curves in biomass change when compared to the station simulations. It appears that differences of greater than 25 mm between the NEXRAD and station rainfall during a 30 day time period resulted in PHYGROW predicting very different trajectories of biomass growth (Figure 3.4A and 3.5A). These trajectory changes occurred during October 2003, May 2005, and October 2006 at both sites and seem to correspond to the start of the growing season or with growth in the latter part of the growing season prior to winter. In each case where NEXRAD simulation changed trajectory from the station simulation, it later caught up and the biomass estimates between the two simulations were comparable.

Bias trends in the comparison statistics for each site follow the results of the location comparison of NEXRAD rainfall to station rainfall described in Chapter II. In that study, NEXRAD rainfall had a tendency to overestimate rainfall at the West Point site by 7% and underestimate rainfall at the Twin China site by 9% (Figure 2.7). A comparison of the change in bias for the station simulation versus the NEXRAD simulation at the West Point site shows that the bias increased from -9.0% to -1.0% (Table 3.2) confirming the overestimation of rainfall by NEXRAD translated into increased biomass production. For the Twin China Site, the bias decreased from -2.08% to -27.67% (Table 3.3) indicating the underestimation of NEXRAD reduced biomass.

The seasonal differences in rainfall bias for the NEXRAD product described in Chapter II are not as easily to discern with the simulation results. The seasonal statistics for NEXRAD indicate that it overestimated rainfall by 1.3% in the monsoon season

(June 1 to September 30) and underestimated rainfall by 4.3% (Figure 2.9). The differences in biomass predictions between NEXRAD and station rainfall seemed to be influenced most by the rainfall in a 30 day window near the end of the monsoon period (May) or the start of the monsoon period (October) (Figure 3.4 and 3.5) rather than the differences in the total amount of rainfall that occurred during the season.

The performance measures were reduced for the NEXRAD simulation when compared to the station rainfall (Table 3.2 and 3.3). The variability in the NEXRAD simulation increased substantially at both sites with RMSD increasing by 72% at the West Point site and by 63% at the Twin China site. The increased variability reduced the goodness-of-fit measures, especially EE which became < 0 for both sites indicating that the variability in the NEXRAD simulation data was much greater than the observed biomass data (Table 3.2 and 3.3). Based on the EE statistic alone, one could conclude that the use of NEXRAD rainfall in PHYGROW simulations for predicting biomass does not provide any additional skill and using the mean of the observed biomass would be a more appropriate predictor (Wilcox et al. 1990; Legates and McCabe Jr. 1999). The index of agreement statistic (*d*) was reduced for the NEXRAD simulation at both sites, but not as drastically as the EE (Table 3.2 and 3.3). It continued to show a reasonable agreement between the NEXRAD rainfall simulation and the observed herbaceous biomass.

The performance of the PHYGROW simulations on grazed locations using NEXRAD rainfall is somewhat in conflict with the performance at the weather station locations. EE and *d* for both calibration and validation were greater than that observed for the NEXRAD simulations at the West Point site and the Twin China (Table 3.2, 3.3, and 3.4) Possible explanations for this include calibration procedures and sensitivity of the performance statistics to the range of biomass values in the data pairs. With regard to the calibration procedure, all of the grazed sites were calibrated using NEXRAD weather data. For sites where NEXRAD might consistently over or underestimate

rainfall, calibrating the model using the NEXRAD data could overcome this bias because variables in the model are adjusted accordingly to match the first data collections points. A potential pitfall of calibration with NEXRAD, especially in biomass prediction for near real-time drought early warning or stocking rate assessment, would be when the NEXRAD rainfall estimation is inconsistent. Uncertainty increases in not knowing how the model would respond, especially during those critical windows when differences in the actual rainfall and the NEXRAD estimate can cause the biomass to take a different trajectory in the model simulation as seen at the West Point and Twin China sites.

A second reason for the differences in model performance statistics between the grazed and ungrazed sites is likely related the greater range of biomass values and sample size for the grazed sites versus the ungrazed station locations. The EE and *d* statistics are sensitive to extreme values (Legates and McCabe Jr. 1999) and increased data pairs reduce the effect of individual outliers. For example, if the data pairs for the grazed site verification data are split along the mean value of 917 kg/ha (Table 3.4), and the EE and d values are recalculated separately for the data pairs on either side of the mean, the overall performance statistics drop. For those data pairs with biomass less than the 917 kg/ha average, the EE was reduced to 0.62 and the *d* reduced to 0.88. These values become comparable to the performance measures for the simulations using station rainfall at the West Point site (Table 3.2). For the data pairs above 917 kg/ha, the EE was 0.83 and the *d* index was 0.95, indicating good correspondence between the NEXRAD simulations and the observed biomass at the sites with higher productivity.

The results for the grazed sites that were calibrated using the NEXRAD rainfall look promising for prediction of herbaceous biomass at the patch scale on semiarid landscapes. However, the stark differences in biomass predictions that were seen for the NEXRAD simulations at the weather station sites (Figure 3.4 and 4.3) suggest the need for additional research to assess performance over a longer period of time. More

frequent temporal sampling at selected sites would be recommended to address the reliability of the NEXRAD simulations especially during and after the time periods where different trajectories in biomass production were detected (start of monsoon and end of monsoon period). This information will be needed to fully understand the ramifications for using PHYGROW with NEXRAD rainfall in a drought early warning system or near real-time stocking rate assessment. Overestimations by the model could lead to erroneous recommendations for increasing animal numbers, which in turn could lead to degradation of the resource due to overgrazing. In the case of drought, overestimation could lead to keeping animals longer than the forage can support (again leading to degradation) and underestimation could lead to ranchers selling animals when they did not need to, thus increasing their operating costs. Additional study could assist in reducing these uncertainties.

CHAPTER IV
COKRIGING OF BIOPHYISCAL MODEL OUTPUT AND A SATELLITE GREENNESS INDEX TO PREDICT FORAGE BIOMASS IN THE GOBI REGION OF MONGOLIA

Introduction

The ability to characterize the vegetation productivity over large landscapes can be an important component in the assessment of drought impacts, natural resource management options, environmental degradation, and economic impacts of changing technologies. For pastoralists, an understanding of the vegetation productivity in the surrounding landscape can assist in determining whether to move, buy or sell animals, and assess the level of risk for decision making. However, the time and resources required to conduct accurate assessments of vegetation productivity over large landscapes are prohibitive, and in many developing countries such as Mongolia, the infrastructure and funding do not exist for large-scale characterization. Another complicating factor is that decisions regarding livestock movement and stocking/de-stocking may require near real-time information, especially in the face of drought. Vegetation productivity assessment is almost impossible to conduct over large land areas on a near real-time basis, thus the information needed for livestock related decisions is not always available when it is needed most. The inability to make decisions at critical times could lead to vegetation overuse, which in turn, could lead to rangeland degradation (Weber et al. 2000).

Improvements in computing power and capacity, along with near real-time production of climate data and remote sensing imagery offer the opportunity to develop near real-time systems for monitoring vegetation on rangelands. Improved computing power and availability of climate data has increased the use of simulation modeling for near real-time monitoring in agriculture systems, including rangelands (e.g., Nain et al.

2002; Stuth et al. 2005). A limitation of many rangeland simulation models is that most provide simulation output for a specific point. Ideally, one would want to simulate as many points (or sites) as possible to represent a region or landscape, especially for the determination of vegetation productivity across the landscape. However, the amount of effort and cost for gathering the data essential for model parameterization on a large number of monitoring points can be prohibitive. Geostatistical interpolation methods such as kriging and cokriging provide the opportunity to extend data collected or simulated for a given set of points to unsampled areas by taking advantage of spatial correlations in the data (Isaaks and Srivastava 1989; Rossi et al. 1994).

As an interpolation method, kriging can provide estimates for unsampled points by using the weighted linear average of the available samples (Rossi et al. 1994). Ordinary kriging is often described as the Best Linear Unbiased Estimator (B.L.U.E.; Isaaks and Srivastava 1989). It is "best" because the variance of the errors is minimized, linear because the estimates are weighted linear combinations of the sample data, and unbiased in that the average error is equal to zero. Goovaerts (1998) states that one of the primary advantages of kriging over other interpolation techniques such as inverse distance weighting is that kriging accounts for the pattern of spatial variability (both range and direction) through semivariogram modeling.

Cokriging involves the use of a secondary variable (covariate) that is cross-correlated with the primary or sample variable of interest and offers additional advantages over ordinary kriging. The secondary variable is usually sampled more frequently and/or regularly, thus allowing estimation at unsampled points using both variables (Isaaks and Srivastava 1989; Goovaerts 1998). Generally, the greater the degree of spatial and cross-correlation that exists between the primary and secondary variables being analyzed, the greater the benefit of using cokriging over kriging (Goovaerts 1998). This can aid in minimizing the error variance of the estimation (Isaaks and Srivastava 1989).

Kriging and cokriging have been used for a variety of applications including mapping of ore bodies in mining (Journel and Huijbregts 1978), mapping and estimating soil physical and chemical properties(Gloaguen et al. 2001; Bekele et al. 2003; Ersahin 2003; Mueller and Pierce 2003), and soil erosion monitoring (Wang et al. 2003). Kriging and cokriging methods have also been employed to estimate plant biomass and other plant parameters on croplands (Atkinson et al. 1994; Dobermann and Ping 2004; Chokmani et al. 2005), forested lands (King et al. 2003; Nanos et al. 2004), and grazinglands (Mutanga and Rugege 2006).

With regard to cokriging in these instances, data from remotely sensed images were used as a covariate in the analysis. Remote sensing imagery provides a dense and exhaustive data set that can serve as a secondary variable for geostatistical interpolation given a correlation between the primary and secondary variable (Dungan 1998). Satellite derived vegetation indices (i.e., greenness indices), most notably the Normalized Difference Vegetation Index (NDVI), have been found to be correlated to vegetation productivity (Tucker et al. 1985; Tucker and Sellers 1986; Wylie et al. 1991; Sannier et al. 2002; Al-Bakri and Taylor 2003; Schino et al. 2003; Pineiro et al. 2006; Wessels et al. 2006), thus making these products suitable for use as a secondary variable in geostatistical analysis. On rangelands, NDVI has generally been used as a predictor variable for vegetation biomass (e.g., Tucker and Sellers 1986; Al-Bakri and Taylor 2003; Frank and Karn 2003), but has not been extensively used as a covariate in geostatistical interpolation of biomass. Vegetation indices produced through the National Oceanic and Atmospheric Administration Advanced Very High Resolution Radiometer (NOAA-AVHRR) and the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data streams have high temporal frequency (daily acquisition with 10 to 16 day compositing intervals) making them attractive for use in near real-time systems. The NOAA-AVHRR data has a relatively long historical record (1981 to present), global coverage, and a resolution of 1 km. This data set has been a major

component of drought and famine early warning systems for Africa (Hutchinson 1991; Rowland et al. 2005).

The assessment of vegetation productivity on a near real-time basis is especially important in Mongolia where drought and winter disasters (dzud) that deplete vegetation resources represent a major risk confronting nomadic livestock producers. During the period from 1999 to 2001, as much as 35% of the nation's livestock was lost to drought and winter disasters. In the Gobi region of the country, livestock mortality reached 50%, with many households losing entire herds (Siurua and Swift 2002). Since the majority of the livestock producers are semi-nomadic (Bedunah and Schmidt 2004), knowledge of the surrounding forage conditions is critical for making decisions about livestock, especially during drought (Kogan et al. 2004). Currently, the majority of herders respond to drought by moving animals to another location, but the movement is not always coordinated due to the lack of information about vegetation condition, thus leading to increased animal numbers in non-drought affected areas. In 2004, a study was implemented in the Gobi region of Mongolia to examine the feasibility of developing a forage monitoring system that would provide near real-time spatial and temporal assessment of livestock forage conditions. The objectives of this study were to 1) assess the ability of the PHYGROW forage simulation model to accurately predict forage biomass at selected sites across the landscape using a near real-time, high resolution rainfall product, and 2) determine the feasibility of using the geostatistical tool of cokriging to integrate simulation model output with satellite greenness indices (NDVI) to produce landscape level maps of forage production across the region.

Methods

Study Area

Mongolia is a landlocked country having a land area of over 1.5 million square kilometers of which of which over 90% is rangelands. Livestock producers are generally semi-nomadic herders who extensively graze their animals in surrounding regions during the spring, summer, and fall, then return to protected camps for the winter months (Bedunah and Schmidt 2004). Sheep and goats are the predominant kinds of livestock, followed by cattle, horses, yaks and camels.

Mongolia's climate is continental with extremely cold, dry winters and warm summers. Precipitation generally occurs in the form of rainfall during the summer months (June – August) which coincides with the general growing season for most plants. The country-wide average temperature is 20º C during the summer months and -24 º C during the winter months. Precipitation is most abundant in the northern regions of the country averaging 200 to 350 mm per year and least abundant in the southern regions which average 100 to 200 mm. A large portion of the country is prone to extreme winter disasters (dzuds) which are periods of intensely cold temperatures (<-40 º C) accompanied by snow and/or ice. They usually follow periods of summer drought which can lead to large losses of livestock because animals are in poor condition and cannot survive the extreme temperatures. The most recent large-scale occurrence of dzud in Mongolia was during 1999 to 2001 (Siurua & Swift 2002).

This study was conducted in the Gobi region of Mongolia (Figure 4.1). The study area included the administrative aimags (provinces) of Gobi Altai, Bayankhongor, Ovorkhangai, Omnogobi, Dundgobi, Dornogobi, Gobi Sumber, and Tov (Figure 4.1). The area can be classified into 5 natural zones (Yunatov et al. 1979) that generally follow the north to south elevational gradient and include the High Mountain, Mountain

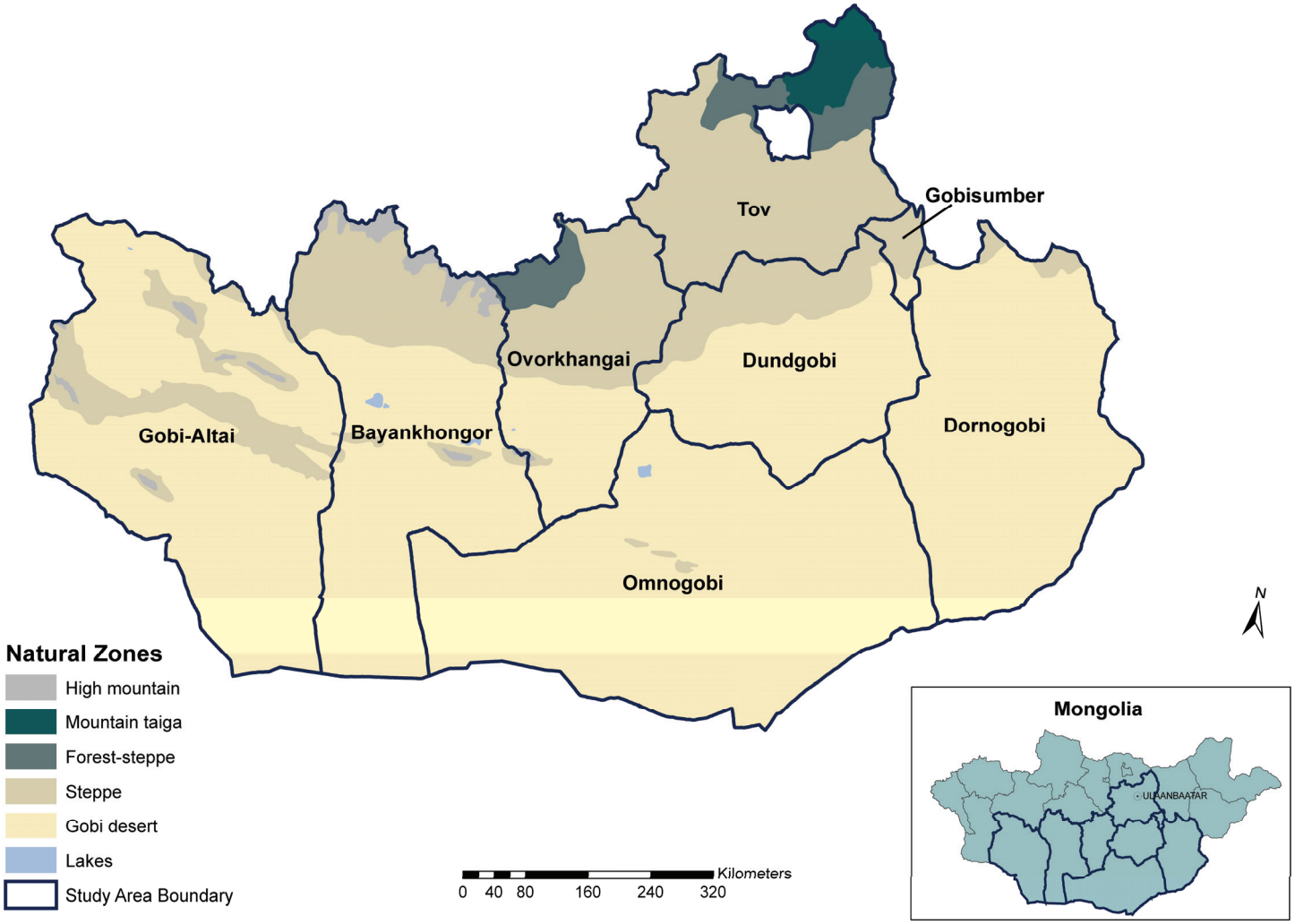

Figure 4.1. Aimag (province) boundaries and natural zones within the study area in Mongolia. The five natural zones generally follow the north to south gradient of elevation within the study area.

Taiga (Forest), Forest Steppe, Steppe, and Gobi Desert zones. The High Mountain zone represents areas above the tree line and consists mainly of tundra vegetation. The Mountain Taiga zone is dominated by forest species, mainly Siberian larch (*Larix sibirica*) and Siberian pine (*Pinus sibirica*). The Forest Steppe zone represents a transition between the Mountain Taiga and Steppe zones and consists of grasslands interspersed with forested areas. Trees such as Siberian larch (*Larix sibirica*) and Siberian pine (*Pinus sibirica*) can be found on north slopes and *Stipa* and *Festuca* grasses on southern slopes. The Steppe zone consists of grasslands dominated by *Stipa*, and *Cliestogenes* grass species and *Artemisia* forbs and have the largest concentration of livestock production within the study area. The Gobi Desert zone is the most arid zone (<200 mm of precipitation) and with the dominant plants consisting of *Stipa* and *Allium* species and sub-shrubs such as *Caragayna* and *Amygdalus* species.

Simulation Model

The Phytomass Growth Simulation Model (PHYGROW) (Stuth et al. 2003a) was used for the prediction of forage biomass for monitoring sites within the study region. PHYGROW is a point model that contains 4 integrated submodels: climate, soil, plant growth and grazing. The model simulates a soil water balance, multi species/functional group plant growth, and livestock grazing on a daily time step. PHYGROW is based on the light use efficiency model concept (Montieth 1972; Montieth 1977) that simulates plant growth under water non-limiting (optimal conditions). The model then discounts plant growth based on the amount of water stress (calculated from the water balance), temperature stress (based on species temperature tolerances for growth), and livestock grazing demand.

The model contains parameters for soil surface and layer information, plant species and community data, livestock grazing management and stocking rates, and is driven by daily climate data (Stuth et al. 2003a). The soil subcomponent of the model

has 13 unique parameters that include soil depth, bulk density, infiltration, and water holding capacity variables. The plant subcomponent can be parameterized for individual species or functional groups. Plant community composition parameters include initial standing crop, percent basal cover for grasses, frequency of forbs, and canopy cover of shrubs and trees. For each individual plant species/functional group in the model, there are 27 parameters including minimum, optimal and maximum temperatures for growth, radiation use efficiency, leaf area index, leaf and wood turnover, leaf and wood decomposition, and canopy water movement. The grazing subcomponent of the model has 19 variables related to each kind/class of grazing animal including forage intake, stocking rate, and grazing preference class for each plant species parameterized in the model. Lastly, the climate subcomponent has 6 variables which include year, day, maximum and minimum temperature, rainfall and solar radiation.

Site Selection and Model Parameterization

A series of monitoring sites were established across the study area. Sites were chosen randomly from a grid representing the resolution of the CMORPH rainfall data. To insure that sites would be accessible, grids were stratified by selecting those that were within 30 km of roads. From the stratified grids, a subset of grids was randomly selected within each aimag (Figure 4.1) with the number of grids proportional to the land area of the aimag. Within each randomly selected grid, the dominant plant community was identified through field reconnaissance and a permanent vegetation transect was established. Due to the large geographic area, the transects were installed in phases with the first phase occurring in the Gobi Altai, Bayankhongor, and Ovorkhangai aimags during 2004 (Figure 4.2). In 2005, transects were established in Omnogobi, Dundgobi, Gobisumber, and Dornogobi aimags. Transects in the Tov aimag were established in 2006. A total of 243 monitoring sites were installed across the region (Figure 4.2).

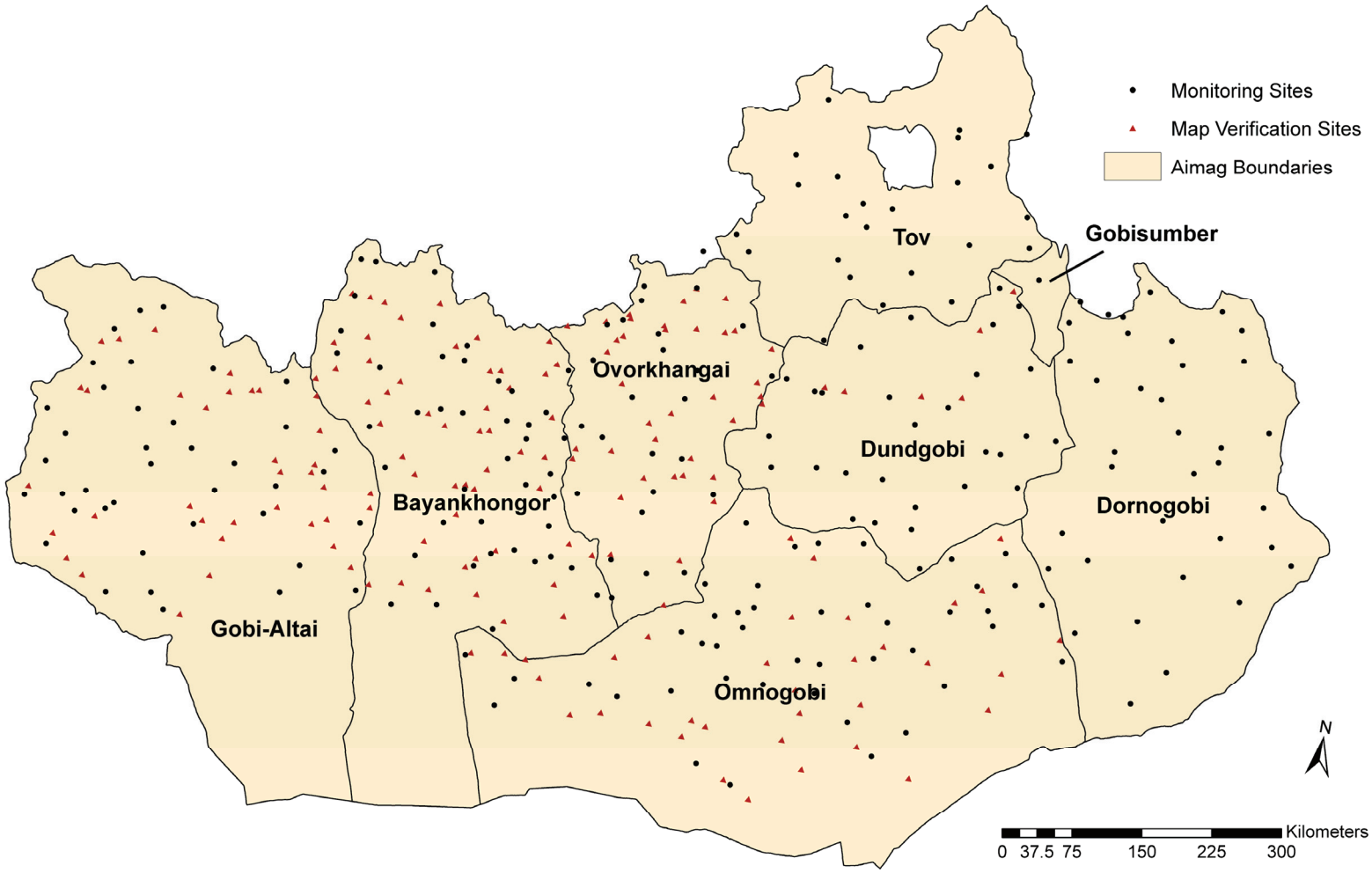

Figure 4.2. Location of monitoring sites (black dots) and independent map verification sites (red triangles) within the study area in the Gobi region of Mongolia.

To gather the necessary plant community parameters for the PHYGROW model at each monitoring site, a modified point-frame method (Ryan 2005) was used to collect percent basal cover of grasses, frequency of forbs, and shrub canopy cover along each permanent transect. Transect lengths ranged from 100 to 500m with the lengths varying based on vegetation cover and plant spacing at the sites. Sites having sparse vegetation and low plant cover had longer transects.

Along each transect, the modified point frame was placed on the soil surface and each point on the frame was examined to determine if it intersected the basal area of a grass species, plant litter, bare ground, or rock. If a basal area of a grass species was encountered, this was recorded as a "hit". Within a 5 x 5-cm quadrat around each point, each presence of a unique forb species was defined as a "hit". If a shrub or tree canopy intersected an upward, perpendicular line from the point, the shrub or tree species was recorded as a "hit". A total of 250 to 500 points were sampled with the number varying based on conditions the vegetation cover and plant spacing. The "hits" of grass, forbs, and shrub/tree species were divided by the total possible hits and these values were entered as the plant community composition variable in the PHYGROW model.

Herbaceous biomass at each transect was measured at the time of transect establishment and at least once more during the period from March 2004 to October 2007. A 0.25 or 0.50-m^2 quadrat was placed at equal increments along the transect (n=10 sample size per transect) and the herbaceous biomass (grass and forbs) was clipped to a 1-cm stubble height. If shrubs were located within the quadrat and they were palatable to livestock, the current year's growth was clipped from the plant. The clipped biomass was placed in paper bags and taken back to the laboratory and dried in a forced air oven at 60º C for 48 hours. After drying, the samples were weighed with a digital scale. The sample weights were then multiplied by the appropriate plot factor in

relation to the quadrat size to convert the biomass to kg/ha units. The 10 samples were averaged and the mean was used for comparison to the simulation model output for each sampling date.

Plant species and functional group parameters for the species encountered during transect establishment were acquired from published literature and online databases such as EcoCrop (FAO 1994) and the Global Leaf Area Index Database (Scurlock et al. 2001). When no information could be found for a species, an expert judgment was made based on the plant genus, functional group, and information on growth characteristics gathered from plant experts in Mongolia.

To parameterize the soil components in PHYGROW for each monitoring site, soil information was acquired from the Mongolia National Soil Laboratory and through consultations with the national soil scientists. When parameters for soil were incomplete, they were estimated from texture using a soil parameter estimation tool (Saxton et al. 1986).

Stocking rate information was calculated from soum (district) censuses of livestock that were conducted during each year of the study. The total number of each kind of livestock was divided by the land area of the soum and this number was used as the stocking rate parameter in PHYGROW. Seasonal dry matter intake for each kind of livestock was determined through consultation with ruminant nutrition scientists with the Mongolian Agriculture University Research Institute for Animal Husbandry (RIAH).

Climate Data Sources

The National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center Morphing Product (CMORPH) rainfall (Joyce et al. 2004) (referred to hereafter as the "CMORPH product") was chosen for use as a driving variable in the forage

simulation modeling. This product is produced by NOAA each 24-hour period and represents the accumulated rainfall that occurs between 0:00 and 24:00 Greenwich Mean Time (GMT) (24:00 GMT is equivalent to 6:00 pm CST). The CMORPH product was acquired automatically from the NOAA servers via internet and downloaded to servers at the Center for Natural Resource Information Technology (CNRIT), Texas A&M University. The rainfall product was delivered as a gridded image that had a geographic range of 80.0º to 120.0º East longitude and 40.0º to 55.0º North latitude, covering the entire country of Mongolia and portions of northern China and southern Russia. Grid cell spacing of the image was 0.07276º in the longitudinal direction and 0.07277º in the latitudinal direction (approximately 8 km at the equator). During the initial comparisons of CMORPH rainfall estimates to station rainfall collected in Mongolia, it was discovered that the product was overestimating rainfall in many locations within the study area, especially in the Steppe and Forest Steppe zones. Large overestimations occurred during the summer months (peak rainfall) and may have been related to the known problem with CMORPH and other satellite rainfall products where rainfall is detected but none reaches the surface because of evaporation (Janowiak 2005) (see discussion on CMORPH in Chapter II). A daily bias correction was calculated and applied to the product using rainfall data collected from approximately 200 weather stations within the Mongolia CMORPH domain. The station data were acquired on a near real-time basis from NOAA as part of the Global Telecommunications System (GTS) data feed. GTS is a world-wide network of climate monitoring stations that provide data to the World Meteorological Organization (WMO) as part of the World Weather Watch system. The bias-adjusted CMORPH data were used for PHYGROW simulation modeling.

Temperature data for the model was acquired from the NOAA Global Data Assimilation System (GDAS) which produces daily maximum and minimum temperature surfaces for the entire globe. Resolution of the data is 1 degree at the equator (approximately 110 km).

Model Calibration and Evaluation

The calibration procedure for PHYGROW involved running the model with the climate data and comparing the simulated forage biomass output to that measured during the transect establishment and subsequent biomass clipping at later dates. If the model output fell within ± 1 standard error of the mean for the herbaceous biomass measured on the transect, the model was considered calibrated for that data collection period. If the model output fell outside ± 1 standard error of the measured data, parameters were adjusted to in an attempt to move the modeled biomass estimate to within the standard error. This process was repeated for each time period data was collected until the model was considered calibrated. Model parameter adjustments for calibration were generally limited to species maximum rooting depths, green and dead leaf turnover rates, and soil layer thickness at the surface (influences depth of soil water evaporation). After the model was considered calibrated, the model parameters for a site were no longer adjusted and the data were used to evaluate model performance during subsequent forage biomass clipping events (model verification).

Geostatistical Interpolation

The bimonthly average forage biomass estimated by the PHYGROW model for each of the monitoring sites was subjected to the geostatistical methods of ordinary kriging and cokriging to determine the feasibility of mapping herbaceous biomass at the landscape scale. Ordinary kriging is an interpolation procedure that predicts the values at unsampled points through a weighted linear averaging of surrounding sampled points (Rossi et al. 1994). The ordinary kriging estimator can be expressed as:

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \qquad [4.1]$$

where $Z^*(x_0)$ represents the value that is to be estimated at the unsampled point x_0, $Z(x_i)$ are the values at sampled points, and λ_i are the weights for the sampled points (Isaaks and Srivastava 1989; Rossi et al. 1994). The weights for the linear averaging are influenced by the degree of spatial correlation (continuity) between points. Spatial continuity can be modeled using semivariance (or variogram) modeling. Semivariance modeling is based on random function theory and allows an examination of sample variability in both direction and distance (Rossi et al. 1994). Semivariance can be expressed as:

$$\gamma^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2 \qquad [4.2]$$

where $\gamma^*(h)$ is the estimated semivariance for separation (lag) distance h, $N(h)$ is the number of pairs of sample points that are separated by distance h, $z(x)$ is the value of the sample at location x, and $z(x+h)$ is the value of another sample at some direction and distance (h) away from $z(x)$ (Rossi et al. 1994). The semivariance can be plotted against distance classes (lags) in the form of a variogram to examine the spatial structure. In modeling the spatial structure, the functional form of the model must be positive definitive to insure that only one stable solution exists (Isaaks and Srivastava 1989). Only a few of these positive definite models are commonly used and they include the 1) nugget effect model, 2) spherical model, 3) exponential model, 4) Gaussian model, and 5) linear model. Linear combinations of these models, which are also positive definite, can be used to model more complex variograms.

Cokriging is a kriging method that involves the use of secondary variables (or covariates) that are spatially cross-correlated with the primary variable that is being estimated (Isaaks and Srivastava 1989). In cokriging, the semivariance analysis is conducted on both the primary and secondary variables in the same manner as for kriging. However, to capture the cross-correlation between the primary and secondary variables, the cross-semivariance is computed in the following functional form:

$$\gamma_{12}^{*}(h) = \frac{1}{2N(h)} \sum_{k=1}^{N(h)} \left[Z_1(x_k) - Z_1(x_k + h) \right] \times \left[Z_2(x_k) - Z_2(x_k + h) \right] \quad [4.3]$$

where $\gamma_{12}^{*}(h)$ is the estimated cross semivariance between the primary variable and the secondary variable, $N(h)$ is the number of pairs of sample points that are separated by lag distance h, Z_1 is the value of the primary variable at locations x and $x + \mathrm{h}$, Z_2 is the value of the secondary variable at these same locations (Hudak et al. 2002; Ersahin 2003).

For a primary variable Z_1 and a secondary variable Z_2, the cokriging model estimates Z^* for location x in the following functional form:

$$Z^*(x_0) = \sum_{k_1=1}^{n_1} \lambda_{k_1} Z_1(x_{k_1}) + \sum_{k_2=1}^{n_2} \lambda_{k_2} Z_2(x_{k_2}) \quad [4.4]$$

where $Z^*(x_0)$ is the cokriging estimated value for the primary variable, λ_{k_1} and λ_{k_2} are the weights for the n_1 primary and n_2 secondary data, respectively, and x_{k_1} and x_{k_2} are the locations of the primary and secondary variables, respectively. The size of n_1 and n_2 are defined when the search neighborhoods for the primary and secondary variable are set in the cokriging analysis (Hudak et al. 2002; Bekele et al. 2003).

The forage biomass from the PHYGROW simulation model for the monitoring sites was used as the primary variable in the kriging and cokriging analyses. For the secondary variable, the NDVIg product of the Global Inventory Modeling and Mapping Studies (Tucker et al. 2005), was acquired from the National Atmospheric and Space Administration (NASA). The NDVIg has a spatial resolution of 8 km (at the equator) and is a global product produced twice per month. The NDVI values for each half month represent a composite of maximum daily NDVI value that occurred in each individual pixel during the period. The images were acquired and stored on the CNRIT server and ArcGIS software (ESRI 2005a) was used to extract the NDVI data from all pixels within the study area.

For each bimonthly period of the NDVI images during the growing season in 2005 and 2006, the forage biomass predicted by the PHYGROW model was averaged for each of the monitoring sites and collocated with the NDVI. Pearson's correlation coefficients were calculated to assess the degree of the linear relationship between the forage biomass and NDVI. Statistical significance of the correlations was assessed using the CRH modified t-test (Clifford et al. 1989) option in the PASSaGE software (Rosenberg 2000). The CRH modified t-test adjusts the degrees of freedom for the degree of autocorrelation in the data since the presence of autocorrelation violates the independence assumption. The modified degrees of freedom are then used in the significance test for the correlations (Rosenberg 2000).

Semivariance modeling and subsequent ordinary kriging and cokriging were conducted using the Geostatistical Analyst extension in the ArcGIS 9 software (ESRI 2005b). Semivariance modeling included selection of an appropriate positive-definite model (exponential, spherical, or Gaussian) that best matched the spatial structure for the forage biomass and the NDVI. Because of the north to south gradient of elevation in the study area, anisotropy (i.e. different spatial structure in different directions) was examined. Kriging and cokriging were conducted with the selected variogram models and a landscape map of forage production was produced for each bimonthly period during the growing season (June to September) in 2005 and 2006. Cross validation (Isaaks and Srivastava 1989) was conducted for each of the chosen kriging and cokriging semivariance models to assess performance and accuracy of the interpolation procedure for the landscape maps of forage biomass.

Independent Map Verification

A set of independent sites were established within the study area for an independent verification of the interpolated maps. The sites were chosen randomly using the same

methodology as described above for the monitoring sites. A total of 164 map verification sites were established during 2005 and 2006 in 6 of the 8 aimags (Figure 4.2). Tov and Dornogobi aimags were not sampled due to logistical constraints. At each of the sites, a transect was established using the same procedures as for the monitoring sites. The forage biomass samples (n=10) collected at each site were taken back to the laboratory for oven-drying and weighing. The forage biomass samples weights were averaged and then paired with the forage estimate from the interpolated maps for the time period when the forage biomass was collected. The observed forage biomass was then compared statistically to the interpolated values to assess how well the interpolated maps performed in predicting biomass in unsampled areas.

Statistical Measures of Performance

The PHYGROW model calibration and validation data, the interpolated map cross-validation, and the independent map verification data sets were each subjected to a series of calculations to assess error and performance. Means and standard deviations for predicted and observed forage biomass for each data set were calculated and linear regression was used to examine model predictive strength (r^2) (Carlson and Thurow 1996). Difference statistics were calculated to examine bias and variability between the prediction and observed data. These statistics included percent estimation bias (BIAS), mean bias error (MBE), mean absolute error (MAE), and root mean square difference (RMSD). Estimation bias reflects the normalized difference between the simulation model output and the observed data and is expressed as follows:

$$BIAS(\%) = \frac{\overline{P} - \overline{O}}{\overline{O}} \times 100 \qquad [4.5]$$

where $\overline{P}$ is the mean of the predictions and $\overline{O}$ is the mean for the observed data. Positive estimation bias values indicate the overestimation of biomass by the simulation model whereas negative values indicate the opposite. Mean bias error provides an indication of the average magnitude of the over-prediction or under-prediction by the

simulation model in the units of the biomass (kg/ha) (Andales et al. 2005). It is calculated as:

$$MBE = \frac{\sum_{i=1}^{n}(P_i - O_i)}{n} \quad [4.6]$$

where P_i is the i^{th} predicted value, O_i is the i^{th} for observed value and n is the number of predicted and observed data pairs. Mean absolute error provides an indication of the average absolute difference between the predicted and observed values in the series of data pairs being evaluated and is calculated as (Legates and McCabe Jr. 1999):

$$MAE = \frac{\sum_{i=1}^{n}|P_i - O_i|}{n} \quad [4.7]$$

Root mean square difference (RMSD) is a measure of the average magnitude of the difference between the predicted and observed biomass data in the units of the data (kg/ha). RMSD is similar to MAE error, however it is more sensitive to extreme differences between the simulation and observed data (Willmott 1982). It is generally greater than MAE and the degree of difference is related to the number of outliers in the data (Legates and McCabe Jr. 1999). RMSD is calculated as follows:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{n}} \quad [4.8]$$

Relative error measures (goodness-of-fit measures) were also used to evaluate performance of the model calibration, model validation, cross-validation of the interpolated maps, and independent verification of the interpolated maps. Relative error measures included estimation efficiency (EE) (Nash and Sutcliffe 1970; Legates and McCabe Jr. 1999) and the index of agreement (*d*) (Willmott. et al. 1985; Legates and McCabe Jr. 1999). Estimation efficiency is a measure of the deviation from a 1:1 line between predicted and the observed data and is calculated as:

$$EE = 1.0 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2} \quad [4.9]$$

An EE value of 1 would reflect a perfect correspondence between the predicted biomass and the measured data. Values greater than 0 would indicate that a positive relationship exists between the predicted and the observed data and that the predicted data is a good estimate of the observed data. Values less than 0 indicate a low correspondence between the predicted and observed data suggesting that the mean of the observed data would serve as a better predictor than the method used to predict biomass (Legates and McCabe Jr. 1999; Moon et al. 2004). The index of agreement is measure of the tightness between the predicted and observed data to a 1:1 line (Willmott. et al. 1985; Andales et al. 2005) and is expressed as follows:

$$d = 1.0 - \frac{\sum_{i=1}^{n}\left(O_i - P_i\right)^2}{\sum_{i=1}^{n}\left(\left|P_i - \overline{O}\right| + \left|O_i - \overline{O}\right|\right)^2} \qquad [4.10]$$

Values of *d* can range from 0 to 1 with a 1 indicating perfect agreement between the simulation output and the observed data.

To determine whether cokriging with NDVI improves the prediction of forage biomass across the landscape compared to kriging, relative improvement was assessed. Relative improvement (RI) assesses the improvement in precision of evaluated methods compared to a reference method (Bekele et al. 2003; Dobermann and Ping 2004) and is calculated as:

$$RI = \left(\frac{RMSE_R - RMSE_E}{RMSE_R}\right) \times 100 \qquad [4.11]$$

where $RMSE_R$ and $RMSE_E$ represent the root mean square errors for the reference method (kriging in this case) and the evaluation method (cokriging with NDVI), respectively.

Results

Simulation Model Performance

For calibration of the PHYGROW model, the majority of the sites in the region were accepted as calibrated after the second biomass sampling. Approximately 10% of the sites required three sampling events for calibration. Across all sites and collection dates, the PHYGROW model predictions of forage biomass averaged 157 kg/ha which was 1 kg/ha greater than the average biomass measured across all sites (Table 4.1). The variability across sites and sampling dates was high for both the simulation model predictions and the observed biomass. The standard deviation for the observed forage biomass was of equal magnitude to the observed mean, but the simulation model estimates were more variable with the standard deviation slightly larger than the mean (Table 4.1). Mean absolute error (MAE) and root mean squared difference (RMSD) were 61 and 93 kg/ha, respectively. The RMSD was 52% greater than the MAE indicating the presence of large outliers in the calibration data set. Several of the larger

Table 4.1. Statistics for performance assessment of the PHYGROW model to predict forage biomass at monitoring sites established across the Gobi region of Mongolia under model calibration and verification.

Statistic	**Calibration**	**Verification**
Observed Mean (kg/ha)	156	133
Simulated Mean (kg/ha)	157	115
[1]sd_o (kg/ha)	157	181
sd_s (kg/ha)	171	185
Bias (%)	0.6	-14
MBE (kg/ha)	1	-18
MAE (kg/ha)	61	59
RMSD (kg/ha)	93	94
r^2	0.71	0.76
EE	0.70	0.74
d	0.91	0.93
n	459	117

[1]sd_o = standard deviation for observed; sd_s = standard deviation for simulation; MBE = Mean Bias Error; MAE = Mean Absolute Error; RMSD = Root Mean Square Difference; r^2 = coefficient of determination; EE = estimation efficiency; *d* = index of agreement; *n* = number of samples

outliers (Figure 4.3) were associated with sites located in the Forest Steppe zones (Figure 4.1) where the model had a tendency to overpredict biomass compared to the observed.

Performance statistics indicated that the PHYGROW model did a reasonably good job of estimating forage biomass under calibration. Linear regression analysis revealed a good correspondence between the observed and simulated biomass (Figure 4.3A) with 71 % of the variability in observed biomass explained by the simulation predictions (Table 4.1). Estimation efficiency (EE) was 0.70 and the index of agreement (*d*) was 0.91 indicating general conformance of the data pairs to the 1:1 line (Table 4.1 and Figure 4.3) and that the model had reasonably good skill in predicting forage biomass under calibration.

For model verification, the number of predicted and observed data pairs across all sites was much reduced compared to calibration due to short length of the study. For model verification, PHYGROW had a tendency to underestimate forage biomass across sites by 14% with an overall mean bias error (MBE) of -18 kg/ha (Table 4.1). The variability in both the observed and PHYGROW predicted biomass, as indicated by their standard deviations, was high and was much greater than for calibration. However, the MAE and RMSD for verification and calibration were very similar (Table 4.1). Several large outliers existed in the verification data pairs (Figure 4.3B) and these were sites located in the Forest Steppe and Steppe areas of the Ovorkhangai aimag (Figure 4.1).

Performance measures indicated that the model performed reasonably well under verification. Linear regression analysis indicated a good correspondence between the PHYGROW predicted biomass and the observed data ($r^2 = 0.76$; Figure 4.3B). EE and *d* statistics also indicated good correspondence and were slightly higher than that observed under calibration.

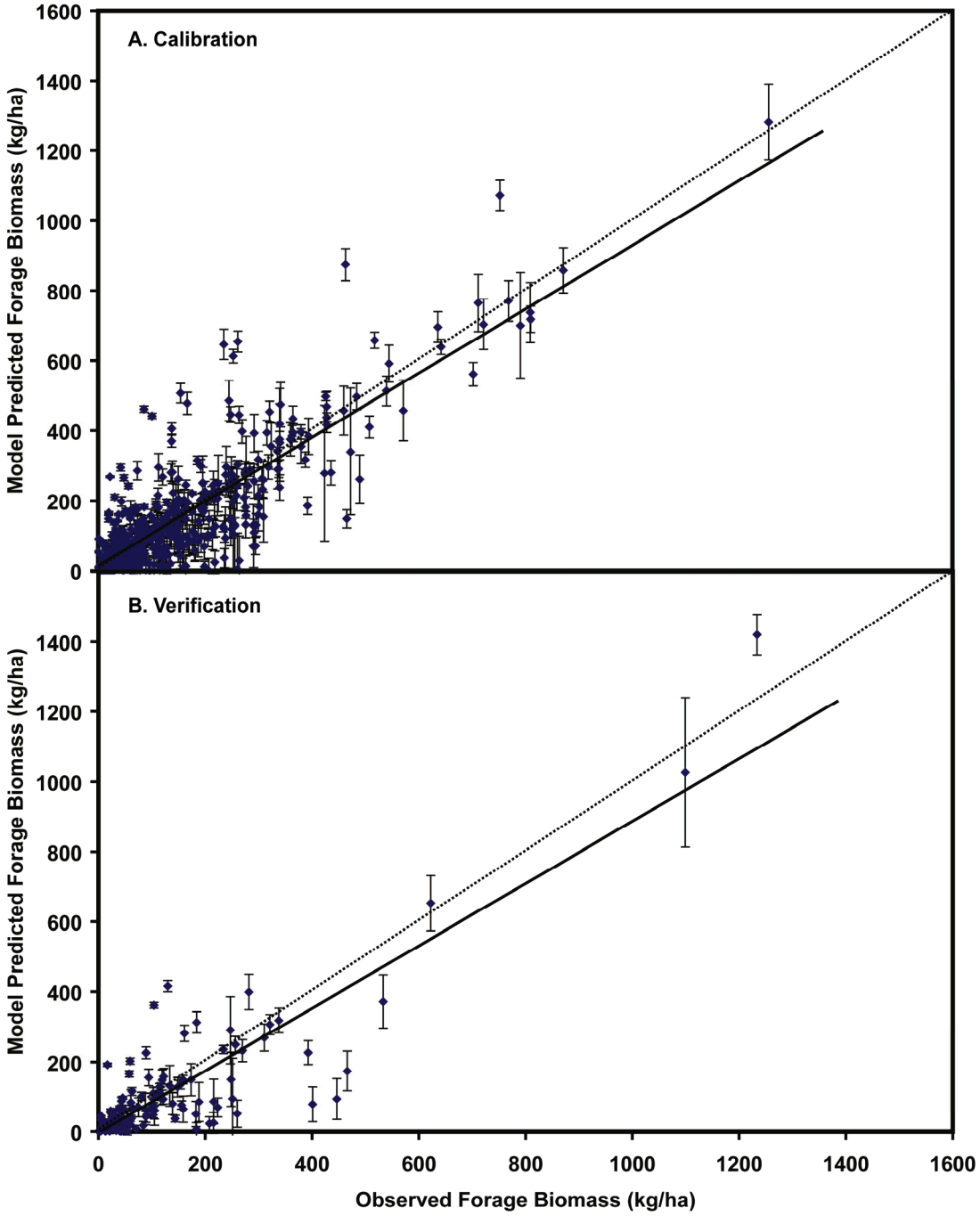

Figure 4.3. Relationship between observed forage biomass (kg/ha ± standard error bars) and PHYGROW model predicted forage biomass for monitoring sites that were A) calibrated and B) verified in the Gobi Region of Mongolia. Dotted line represents 1:1 line.

Cokriging of Forage Biomass

For the 2005 and 2006 growing seasons, average forage biomass during each bimonthly period (corresponding to NDVI image delivery) increased each period until it peaked in late August/early September and then began do decline (Table 4.2). Standard deviations also increased; however, the coefficient of variability indicated greater variability in predicted biomass across sites at the beginning and end of the growing season than during the middle (Table 4.2). Mean biomass and variability were also greater in 2006 for most of the time periods when compared to 2005 (Table 4.2).

Mean NDVI , across the study area during both 2005 and 2006, was lowest during the start of the growing season in June, increased during June and July, peaked in August and then began to decline (Table 4.3). The peaks in NDVI were slightly earlier than that of the PHYGROW predicted forage biomass (Tables 4.2 and 4.3). Like forage,

Table 4.2. Bimonthly statistics for forage biomass predictions from the PHYGROW model at monitoring sites in the Gobi region of Mongolia during the growing season in 2005 and 2006.

	2005 Statistics							
Statistic	**June 1-15**	**June 16-30**	**July 1-15**	**July 16-31**	**August 1-15**	**August 16-31**	**Sept 1-15**	**Sept 16-30**
Mean	48	85	134	170	217	243	244	220
[1]*sd*	58	97	145	185	237	272	276	263
Maximum	368	665	1182	1563	1748	1826	1772	1670
Minimum	0	0	0	0	0	0	0	0
CV (%)	122	114	108	108	109	112	113	120
n	243	243	243	243	243	243	243	243
	2006 Statistics							
	June 1-15	**June 16-30**	**July 1-15**	**July 16-31**	**August 1-15**	**August 16-31**	**Sept 1-15**	**Sept 16-30**
Mean	35	68	116	180	231	245	235	214
sd	64	97	149	213	268	293	292	276
Maximum	497	846	1184	1552	1726	1775	1705	1680
Minimum	0	0	0	0	0	0	0	0
CV (%)	180	144	128	118	116	119	124	129
n	243	243	243	243	243	243	243	243

[1]*sd* = standard deviation; CV = coefficient of variation; *n= sample size*

Table 4.3. Bimonthly statistics for Normalized Difference Vegetation Index (NDVI) in 8 x 8 km grid resolution across the Gobi region of Mongolia during the growing season in 2005 and 2006.

	2005 Statistics							
Statistic	**June 1-15**	**June 16-30**	**July 1-15**	**July 16-31**	**August 1-15**	**August 16-31**	**Sept 1-15**	**Sept 16-30**
Mean	0.128	0.141	0.157	0.155	0.177	0.190	0.172	0.148
[1]*sd*	0.125	0.179	0.202	0.205	0.212	0.195	0.161	0.113
Maximum	0.780	1.000	1.000	1.000	1.000	0.921	0.842	0.623
Minimum	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CV (%)	98	126	129	133	120	102	94	76
n	0.128	0.141	0.157	0.155	0.177	0.190	0.172	0.148
	2006 Statistics							
	June 1-15	**June 16-30**	**July 1-15**	**July 16-31**	**August 1-15**	**August 16-31**	**Sept 1-15**	**Sept 16-30**
Mean	0.126	0.131	0.167	0.195	0.214	0.200	0.183	0.139
sd	0.124	0.154	0.193	0.211	0.217	0.187	0.158	0.105
Maximum	0.942	0.805	0.883	1.000	1.000	0.900	0.790	0.690
Minimum	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CV (%)	98	118	116	108	102	94	87	76
n	0.126	0.131	0.167	0.195	0.214	0.200	0.183	0.139

[1]*sd* = standard deviation; CV = coefficient of variation; *n= sample size*

the NDVI was quite variable over time. However variability was greater during late June and July for both years. Similar to forage biomass, NDVI was generally higher in 2006 when compared to 2005. However, in contrast to the forage, the NDVI variability was generally greater in 2005 than in 2006.

Pearson's correlation analysis generally indicated a moderately high correlation between simulated forage biomass and NDVI for most of the growing season. Correlations were lowest (r =0.45) at the start of the growing season during both years (Table 4.4). Correlation increased between the two variables as the season progressed with correlations ranging from 0.73 to 0.76 during late July to early September. Correlation was greatest during late July in 2005 and in early September during 2006. (Table 4.4). Correlations between biomass and NDVI were slightly greater in 2006 when compared to 2005.

To examine spatial structure of the forage biomass and NDVI, semivariance analysis was conducted. For both forage and NDVI, a spherical model best represented the structure in the isotropic empirical semivariogram. Although anisotropy was detected for both forage biomass and NDVI (Figure 4.4, semivariance surfaces), model performance statistics generally did not indicate any better fit for anisotropic models than for the isotropic form. Therefore, the results presented here are for the isotropic model.

For forage biomass, the range of spatial structure increased as the growing season progressed (Table 4.5). During the beginning of the growing season, the effective range of spatial dependence was approximately 500 km during both years (Table 4.5). As the amount of biomass increased and peaked, the range of spatial dependence increased to approximately 650 km in 2005, and 690 km in 2006. Figure 4.5A provides an example of a fitted variogram model for forage biomass during the September 1 to 15, 2006 time period.

Table 4.4. Pearsons correlation coefficients (r) between PHYGROW simulated forage biomass and Normalized Difference Vegetation Index values during the 2005 and 2006 growing season for monitoring sites in the Gobi region of Mongolia. Time periods represent production periods for NDVI images. All correlations were significant ($p<0.05$) after adjustment of the degrees of freedom for autocorrelation in the data using the CRH modified t-test (Clifford et al. 1989).

	Year	
Time period	**2005**	**2006**
June 1 – 15	0.45	0.45
June 16 – 30	0.60	0.57
July 1 – 15	0.68	0.71
July 16 – 31	0.74	0.75
August 1 – 15	0.73	0.74
August 16 – 31	0.73	0.74
September 1 – 15	0.71	0.76
September 16-30	0.67	0.68

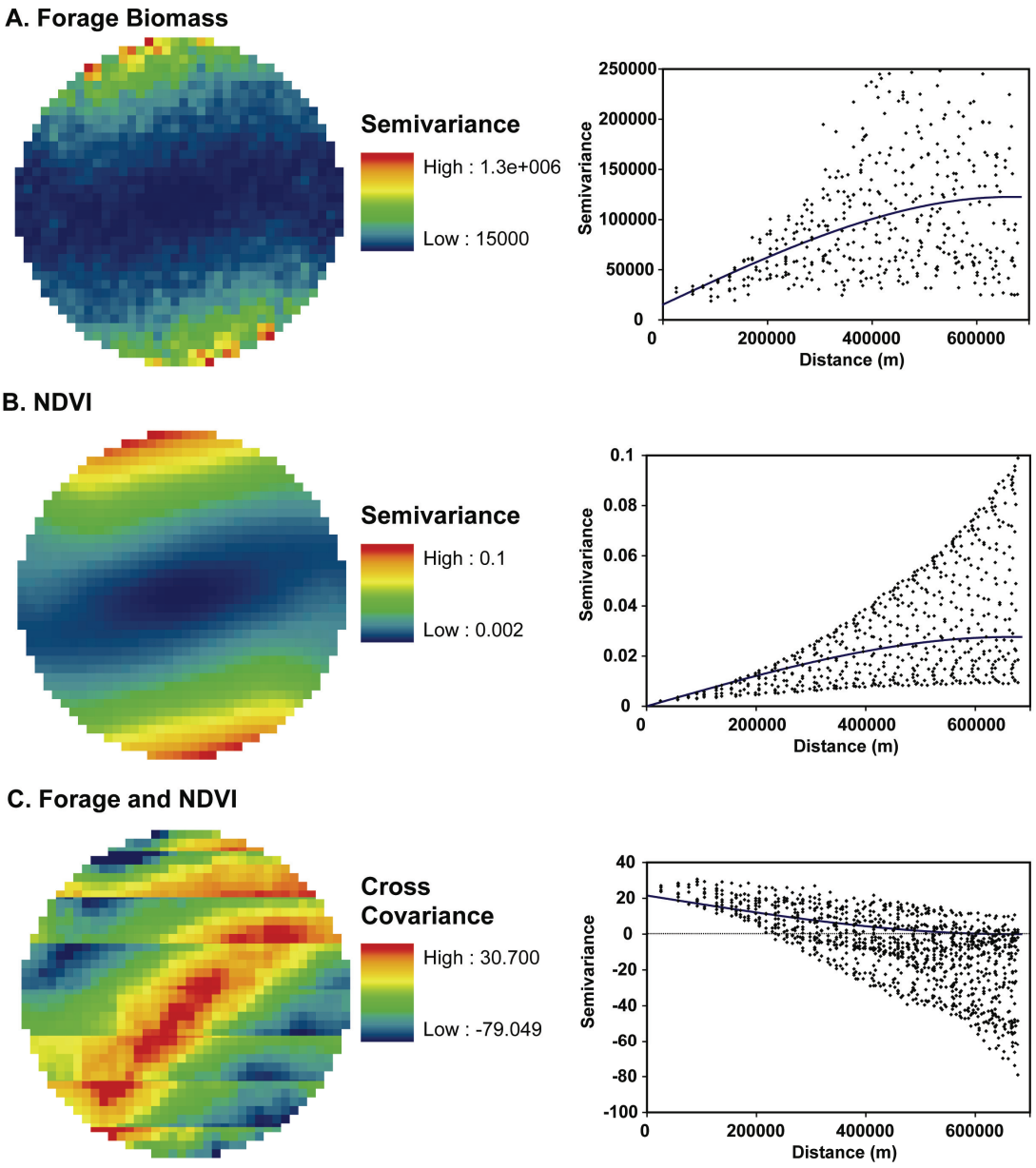

Figure 4.4. Semivariance/cross covariance surfaces and the associated empirical variogram/covariogram and fitted spherical models for the A) forage biomass, B) NDVI, and the C) cross-covariance between biomass and NDVI for the period of September 1 to September 15, 2006.

Table 4.5. Parameters for semivariance models used to examine spatial structure in PHYGROW simulated forage biomass, Normalized Difference Vegetation Index (NDVI), and the cross covariance between the forage biomass and NDVI at monitoring sites during bimonthly periods in 2005 and 2006. Each empirical variogram was fit using a spherical model with lag sizes for 36000 m (36 km).

	Forage Biomass Variogram Model			**NDVI Variogram Model**			**Cross Covariance Model**	
Time Period	**[1]C_o**	**[2]C_o+C**	**[3]C/ (C_o+C)**	**C_o**	**C_o+C**	**C/ (C_o+C)**	**[4]C**	**[5]A**
2005								
June 1-15	1644	4040	0.59	0	0.012	1	2.7	502068
June 16-30	2860	10922	0.74	0	0.027	1	8.4	502068
July 1-15	4543	27649	0.84	0	0.045	1	11.8	644490
July 16-31	6344	44264	0.86	0	0.046	1	17.4	611470
August 1-15	9320	78947	0.88	0	0.050	1	24.0	640350
August 16-31	13334	104658	0.87	0	0.042	1	26.9	628595
September 1-15	15122	109574	0.86	0	0.027	1	21.7	642253
September 16-30	14737	99277	0.85	0	0.014	1	13.8	649663
2006								
June 1-15	996	4543	0.78	0	0.013	1	3.5	502068
June 16-30	2304	10362	0.78	0	0.020	1	6.9	502068
July 1-15	3825	26746	0.86	0	0.037	1	14.0	574310
July 16-31	5625	64587	0.91	0	0.054	1	19.7	693339
August 1-15	10518	103227	0.90	0	0.055	1	27.5	677355
August 16-31	13922	124334	0.89	0	0.040	1	25.4	671308
September 1-15	15313	122586	0.88	0	0.028	1	21.5	663446
September 16-30	13630	110719	0.88	0	0.012	1	12.5	667412

[1]C_o = nugget variance
[2]C_o+C = Sill or overall variance
[3]C/(Co+C) = proportion of sill variance explained by spatial structure
[4] C = Partial sill or variance explained by spatial structure
[5]A = Range of spatial structure (m)

The relative contribution of the spatial structure to the overall variance in forage biomass also increased as the growing season progressed (Table 4.5). For example, during the time period of June 1 to 15, 2005, the range of spatial dependence accounted for 59% of the total variance. After this date, the proportion of total variance explained by spatial structure increased 74 to 88% in 2005, and 78 to 90% in 2006, coinciding with the peaks in biomass (Table 4.5). The overall variance in biomass during these same time periods also increased indicating an increase in spatial variability as the growing season progressed (Table 4.5).

The variance not accounted for by the spatial structure is referred to as the nugget variance and provides an indication of the microscale variation and/or measurement error in the data (Isaaks and Srivastava 1989). For the early portion of the growing season in both years, the nugget variance made up a larger proportion of the overall variance than during other sampling dates. This implies a greater degree of variability in the forage biomass at scales less than the first lag (36 km) used for the semivariance modeling of forage biomass.

Like forage biomass, the overall variance in NDVI was lowest at the beginning of the growing season and progressively increased until peak NDVI in August of both years implying that the spatial variability in NDVI was greatest during this time period. The fitted variogram models for NDVI resulted in nugget variances of 0 for all time periods. Thus, the spatial structure accounted for 100% of the overall variance in NDVI (Table 4.5). Figure 4.4B provides an example fitted model for the empirical variogram for NDVI and its associated semivariance surface for the September 1 to 15, 2006 time period.

The cross-covariance between forage biomass and NDVI displayed a similar pattern of change as that for biomass and NDVI individually. Cross covariance was lowest at the beginning of the growing season and increased until early August and then declined in both years (Table 4.5).

Using the fitted variogram models, kriging and cokriging were conducted for each time period and cross-validations were conducted to evaluate performance of the interpolations. Performance statistics between the kriging and cokriging interpolations were very similar. However, an examination of the Relative Improvement (RI) statistic for assessing the improvement in RMSD with cokriging generally indicated that cokriging was slightly better than kriging with a 1 to 5% reduction in RMSD (Table 4.6).

Therefore, only the results for cokriging will be presented. The performance statistics for kriging are provided in Appendix A.1.

Cross validation analysis indicated that cokriging generally resulted in a slight underprediction (1-4%) of forage biomass over time (Table 4.6). RMSD and MAE were lowest at the beginning of the growing season each year and increased throughout the growing season until peak biomass in August (Table 4.6). However, the proportion of RMSD and MAE in relation to the mean forage biomass was higher (0.82 to 1.4) at the start and end of the growing season than at peak forage biomass (0.7 to 0.8) indicating higher variability in the biomass estimates during these periods. Although the RMSD for cokriging cross-validations for July to September were higher than that observed for the PHYGROW model calibration and verification (RMSD = 93 and 94 kg/ha, respectively) the proportion of error relative to the means were similar (Table 4.1 and 4.6).

During the month of June in both years, the goodness-of-fit statistics for the cokriging cross-validation were low (Table 4.6). The EE statistics were negative, implying that the overall mean for the PHYGROW simulated biomass was a better predictor of biomass than the cokriging model. The low performance during the month of June generally corresponded to when the nugget variance made up a greater proportion of the overall variance compared to the other time periods (Table 4.5). Apparently, the variability in adjacent samples was quite high during the early part of the growing season and reduced the prediction capability of the cokriging model. Goodness-of-fit statistics were improved for the July, August, and September months in both years, but slightly more so during 2006. Linear regression r^2 and the *d* index values approached that observed for PHYGROW model calibration and validation (Table 4.1 and 4.6). However, the EE statistics were much lower for the cokriging (Table 4.6) when compared to the PHYGROW model calibration and validation (Table 4.1). The

Table 4.6. Cross-validation analysis statistics for cokriging of PHYGROW simulation model and Normalized Difference Vegetation Index (NDVI) data to estimate forage standing crop across the Gobi region of Mongolia during the growing season (June to September) in 2005 and 2006. Time periods noted below represent production periods for the NDVI product.

	2005 Cross-validation							
Statistic	**June 1-15**	**June 16-30**	**July 1-15**	**July 16-31**	**August 1-15**	**August 16-31**	**Sept 1-15**	**Sept 16-30**
Simulation Mean (kg/ha)	48	85	134	170	217	243	244	220
Cokriged Map Mean (kg/ha)	47	82	132	169	213	238	240	216
[1]sd_o (kg/ha)	58	97	145	185	237	272	276	263
sd_s (kg/ha)	30	65	107	144	192	217	220	207
Bias (%)	-2	-3	-2	-1	-2	-2	-1	-2
MBE (kg/ha)	-1	-2	-2	-1	-3	-4	-4	-4
MAE (kg/ha)	36	50	64	77	93	104	105	102
RMSD (kg/ha)	50	71	96	117	145	170	180	170
R^2	0.26	0.46	0.56	0.60	0.63	0.61	0.58	0.58
EE	-1.75	-0.20	0.19	0.34	0.43	0.38	0.33	0.32
d	0.63	0.79	0.84	0.86	0.88	0.87	0.85	0.85
n	243	243	243	243	243	243	243	243
RI (%)	1.04	4.32	2.89	3.23	1.64	0.86	-1.20	0.23

	2006 Cross-validation							
	June 1-15	**June 16-30**	**July 1-15**	**July 16-31**	**August 1-15**	**August 16-31**	**Sept 1-15**	**Sept 16-30**
Simulation Mean (kg/ha)	35	68	116	180	231	245	235	214
Cokriged Map Mean (kg/ha)	34	65	113	175	226	241	233	211
[1]sd_o (kg/ha)	64	97	149	213	268	293	292	276
sd_s (kg/ha)	41	63	112	170	216	237	239	225
Bias (%)	-4	-3	-3	-2	-2	-2	-1	-2
MBE (kg/ha)	-1	-2	-3	-4	-5	-4	-3	-4
MAE (kg/ha)	29	45	61	79	102	108	106	100
RMSD (kg/ha)	49	73	94	123	158	169	165	158
R^2	0.40	0.43	0.60	0.67	0.65	0.67	0.68	0.67
EE	-0.43	-0.34	0.28	0.47	0.46	0.49	0.52	0.51
d	0.75	0.76	0.86	0.89	0.88	0.89	0.90	0.89
n	243	243	243	243	243	243	243	243
RI (%)	-1.97	0.26	4.52	3.36	1.48	2.10	2.18	1.24

[1]sd_o = standard deviation for model simulated biomass; sd_s = standard deviation for cokriged biomass; MBE = Mean Bias Error; MAE = Mean Absolute Error; RMSD = Root Mean Square Difference; r^2 = coefficient of determination; EE = estimation efficiency; *d* = index of agreement; *n* = number of samples; RI = Relative improvement over kriging

EE statistics is sensitive to extreme outliers (Legates and McCabe Jr. 1999) and this leads to an overall reduction in this statistic. An examination of the monitoring site locations having the greatest differences in predicted and observed values in the cokriging cross-validation revealed extreme differences for 8-9 locations in the Forest Steppe and Steppe regions in the Tov and Ovorkhangai aimags (Figure 4.2). These included several of the monitoring sites that were found to be outliers in the PHYGROW model calibration and validation.

Overall, the cross-validation results indicated that cokriging had moderate to good utility in interpolating biomass during the months of July, August and September, and low skill during June, the early portion of the growing season. Examples of the interpolated maps for each of the bimonthly time periods are presented in Figure 4.5 (2005) and Figure 4.6 (2006).

Independent Map Verification

A comparison of the biomass estimates from the cokriging interpolated maps to that of the biomass measured at 167 independent monitoring sites (Figure 4.2) revealed that interpolation of the PHYGROW model forage biomass resulted in an overall 14% underestimation of forage biomass (Table 4.7). RMSD and MAE were larger than that observed for PHYGROW model calibration/verification and cokriging cross-validation in terms of both the absolute amount and the proportion of error relative to the observed mean (Tables 4.1, 4.6, and 4.7). Goodness-of-fit statistics showed an overall low performance for cokriging of PHYGROW model biomass with a linear regression r^2 of 0.37, EE of 0.34, and a *d* index of 0.74. There were several sets of extreme deviations from the 1:1 line that reduced the overall performance of the map verification (Figure 4.7). The majority of these sites were situated in the northern regions of the study area in the High Mountain and Forest Steppe zones in the Bayankhongor and Ovorkhangai aimags (Figure 4.2).

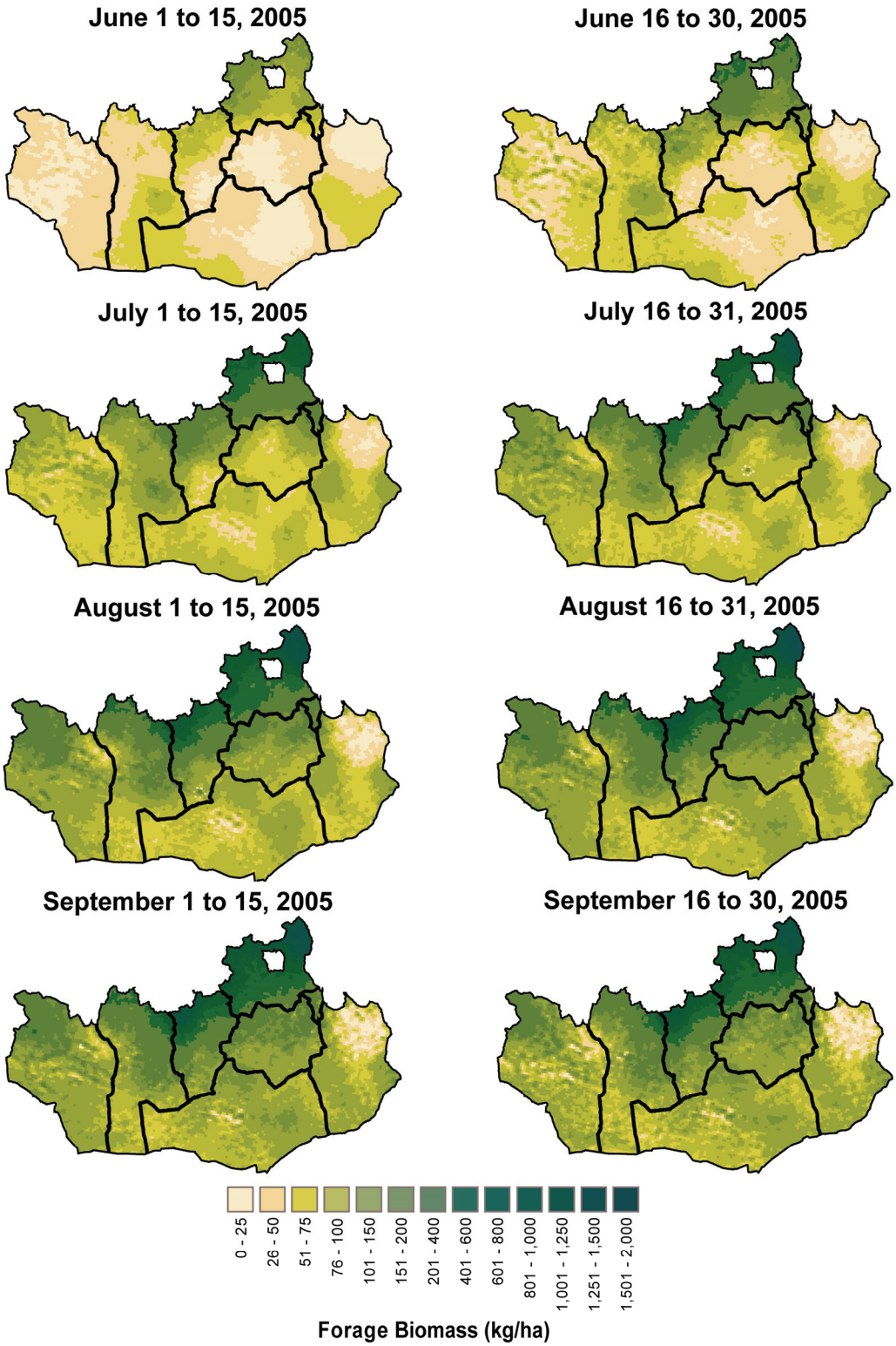

Figure 4.5. Bimonthly cokriged maps of forage biomass (kg/ha) during the 2005 growing season for the Gobi region in Mongolia. Cross validation statistics for each map can be found in Table 4.6.

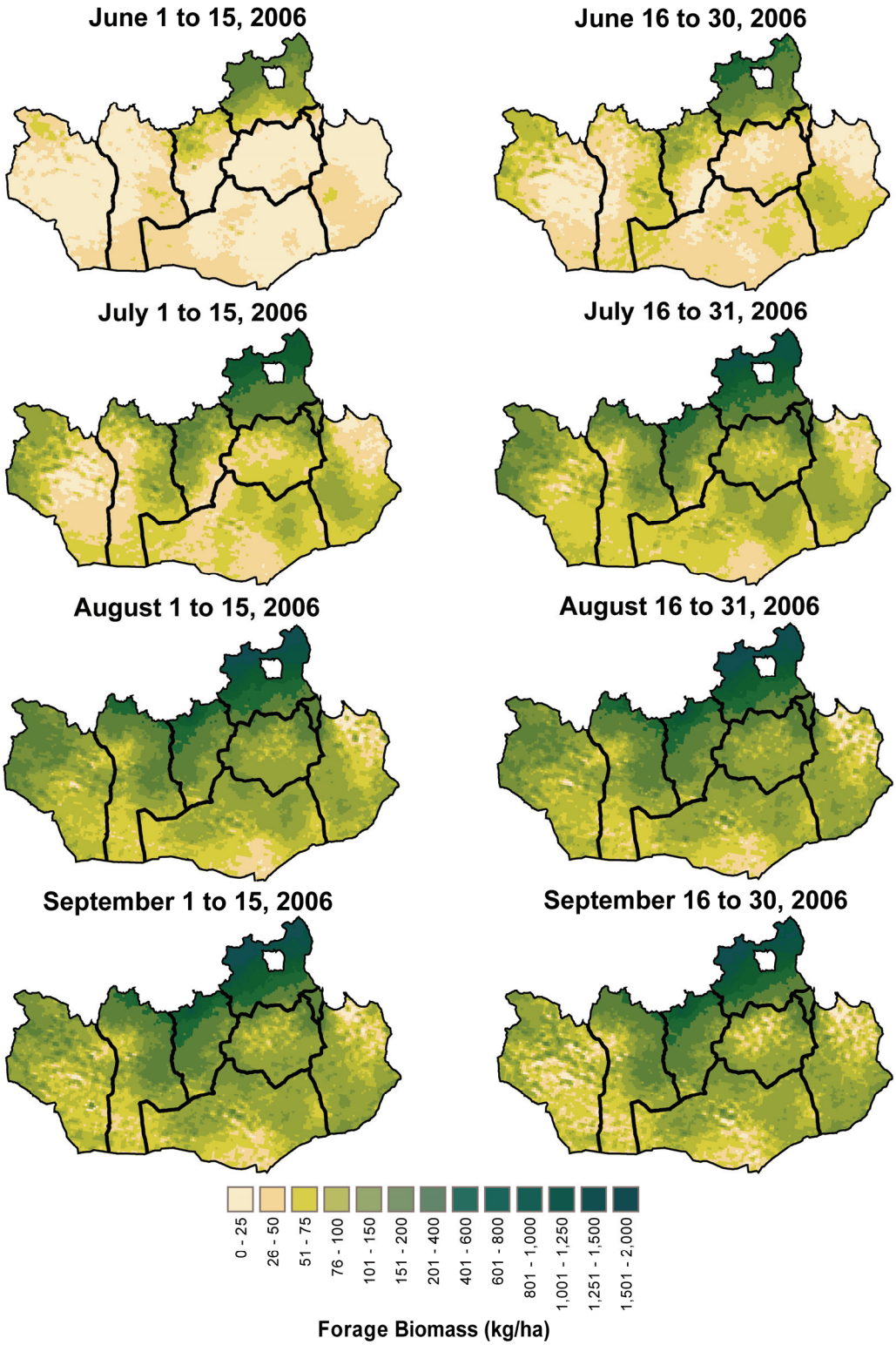

Figure 4.6. Bimonthly cokriged maps of forage biomass (kg/ha) during the 2006 growing season for the Gobi region in Mongolia. Cross validation statistics for each map can be found in Table 4.6.

Table 4.7. Statistics for evaluating the performance of cokriging interpolation of PHYGROW derived biomass at independent map verification sites established across the Gobi region of Mongolia.

Statistic	Map Verification
Observed Mean (kg/ha)	189
Cokriged Map Mean (kg/ha)	163
[1]sd_o (kg/ha)	220
sd_s (kg/ha)	162
Bias (%)	-14
MBE (kg/ha)	-26
MAE (kg/ha)	115
RMSD (kg/ha)	178
r^2	0.37
EE	0.34
d	0.74
n	164

[1]sd_o = standard deviation for simulation output; sd_s = standard deviation for cokriged map data ; MBE = Mean Bias Error; MAE = Mean Absolute Error; RMSD = Root Mean Square Difference; r^2 = coefficient of determination; EE = estimation efficiency; *d* = index of agreement; *n* = number of samples

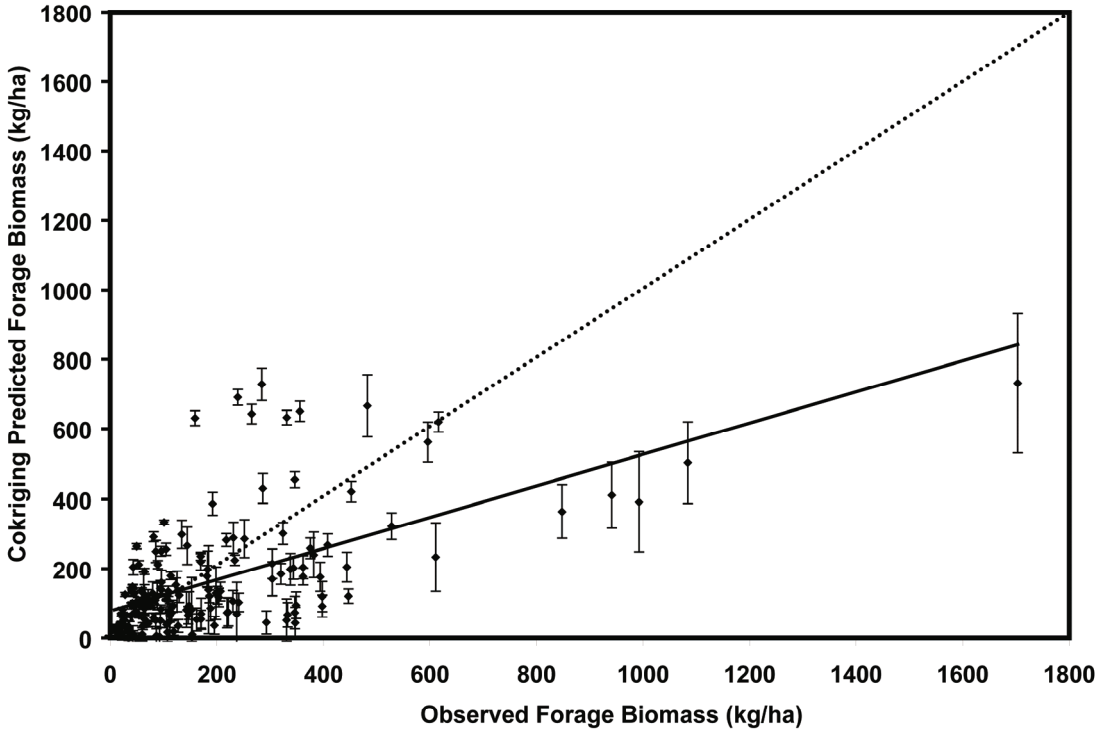

Figure 4.7. Relationship between observed forage biomass (kg/ha ± standard error bars) and cokriging predicted forage biomass for independent map verification sites in the Gobi region of Mongolia. Dotted line represents 1:1 line.

Discussion

The PHYGROW model performed reasonably well in predicting forage biomass at the majority of the monitoring sites (Table 4.1, Figure 4.3). Although the verification data was not extensive due to the short time frame of the study and the large geographic area covered, initial results showed good correspondence between the model predicted biomass and that measured in the field. An examination of outliers in both the calibration and verification datasets indicated that monitoring sites located in the in the northern portions of the study area had erratic correspondence between modeled biomass and that measured in the field. These sites were generally located in the Forest Steppe zone and its transition into the Steppe zone (Figure 4.1). Although these sites were considered calibrated, it appears that additional and more frequent data needs to be collected to improve model predictions. The model generally overestimated biomass at these sites. Possible reasons for overestimation include misparameterization of soils at the site and problems with the CMORPH estimation of rainfall. With regard to misparameterization of soils, the soils at the sites were identified using a national soil map with 1:1,000,000 scale. Within this Forest Steppe zone, the topography, geology, and aspect are highly variable, so the soil chosen for use in the model may not have been appropriate for the site. The PHYGROW model is also sensitive to the depth of bedrock or indurated layer. Information from the soil survey descriptions on depth to bedrock may not have been appropriate. Future monitoring at these sites should include a more complete soil characterization to rule soil out as a possible problem in the model.

Discrepancies between the CMORPH predicted rainfall and what was actually received at the site could lead to large differences between the model predicted and observed biomass at the site. The PHYGROW model predictions are very sensitive to the timing and amounts of precipitation (see discussion in Chapter III). CMORPH has a general tendency to overestimate rainfall (see discussion in Chapter II), and although the CMORPH rainfall used in this study had a bias correction, it may not have adequately

corrected the overestimate, thus leading to higher biomass estimates by the PHYGROW model. A daily examination of the CMORPH rainfall images for anomalous rainfall cells might prove helpful to identify areas where the rainfall may not be adequately corrected by the bias correction. Because there is a lack of rain gauges in the Forest Steppe areas, future monitoring may need to include collection of rainfall at these sites to eliminate CMORPH as a potential reason for poor model performance.

Cross-validation indicated that for cokriging of the PHYGROW output with NDVI as a covariate performed well during the months of July, August, and September with performance measures slightly less than the PHYGROW model calibration and verification performance measures (Tables 4.1 and 4.6). The exception to this was the EE statistic which was lower due to its sensitivity to extreme outliers. The outliers identified in the cross-validation analysis were sites located in the Forest Steppe and its transition into the Steppe zone in the northern portions of the Tov and Ovorkhangai aimags. Cokriging underestimated forage at these sites biomass by 300 to 500 kg/ha. The majority of these outliers were sites identified as outliers in the PHYGROW model validation. Future improvements to the model calibration at the sites identified as outliers should improve the overall cokriging performance. Since the mountainous terrain in the Forest Steppe is highly variable with grassland meadow vegetation on southern exposures and Siberian larch (*Larix sibirica*)/Siberian pine (*Pinus sibirica*) forests on the northern exposures (Gunin et al. 1999), increasing the number of monitoring sites in these areas would also improve the cokriging estimates.

The poor performance of the cokriging interpolation at the start of the growing season during both years (Table 4.6) appears to be related to a greater amount of variability in the PHYGROW model estimates during that period. The nugget variances in the semivariance modeling of forage biomass were proportionally larger during the start of the growing season when compared to the other time periods indicating a larger degree of spatial variability in the forage biomass at distances less than the first lag in

the semivariogram model (36 km). As the nugget variance increases, the weights for the weighted averaging in the cokriging become more equitable leading to estimates for points during interpolation to be more like simple averages (Isaaks and Srivastava 1989). Therefore as the nugget variance increases relative to the overall variance, the effect of spatial continuity in the cokriging estimates are much reduced.

The higher degree of variability in biomass estimates during the early growing season appears to be an issue with differential timing for plants to start growth in the PHYGROW model. Since the majority of the monitoring sites had biomass amounts near zero for forbs and grasses after the winter months, biomass production generally does not begin until rainfall is received. Since each monitoring site was established in a separate CMORPH rainfall grid, and given the sensitivity of the PHYGROW model to rainfall at the beginning of the growing season (see discussion in Chapter III), biomass production in the model could be quite different at sites that are short distances apart. As the growing season progresses and rainfall increases, the productivity between sites close together may become more synchronized with each other. Additional field sampling with a higher frequency of forage biomass collection, coupled with rainfall measurement at selected groups of sites could assist in better defining the source of variability and improve model calibrations.

The relative improvement in error reduction in using cokriging of biomass with NDVI compared to kriging of biomass alone (Table 4.6) was less than expected given the moderately high correlation between NDVI and forage. Dungan (1998), in a study using a synthetic dataset to examine biomass estimation methods, noted that vegetation quantity predictability with cokriging increased as the correlation between the primary and secondary variable increases. Mutanga and Rugege (2006) in a study comparing kriging and cokriging with NDVI to estimate biomass at Kruger National Park, South Africa, found a relative improvement of 21% with cokriging. The fact that NDVI did improve predictions somewhat is encouraging. Improvement in PHYGROW model

predictions of biomass in the Forest Steppe zones would likely improve the correlations between NDVI and forage biomass, thus leading to further reduction in error compared to kriging only.

The performance of cokriging as assessed through independent verification was low (Table 4.7); however, the performance results did show that some skill existed in the cokriging of forage biomass in the region (e.g., the positive EE statistic). The problems associated with uncertainty in the biomass predictions by the PHYGROW model at sites in the northern portion of the study area carried through to the cokriging, resulting in poor estimates of biomass. Another issue may be related to the scale of the transects (100 to 500 m) versus the scale of the rainfall and NDVI grids (8 km). Additional data may need to be collected in the future to address whether the transects are representative of the large grids, especially in the northern regions of the study area. An additional positive benefit of the independent map verification is that is allowed and identification of areas where the greatest deviations in biomass predictions occurred, thus assisting in identifying new areas where monitoring sites could be installed in the future to improve model and cokriging predictions.

The overall methodology of geostatistically integrating the PHYGROW model output with NDVI to produce landscape maps of forage biomass shows promise for implementing as a near real-time system for drought monitoring in Mongolia. Improvements in the calibration of the PHYGROW model for sites in the northern regions, as well as more rigorous validation at selected sites across should help reduce the uncertainty in PHYGROW model predictions, and therefore improve accuracy of the interpolated maps. Cokriging by natural zone is another option that could be examined to strengthen the predictions within the individual zones across the region

CHAPTER V
SUMMARY

Assessment of vegetation productivity on rangelands is needed to assist in timely decision making with regard to management of the livestock enterprise as well as to protect the natural resource. Characterization of the vegetation resource over large landscapes can be time consuming, expensive and almost impossible to do on a near real-time basis. Recent advances in computer capacity, remote sensing, and climate data availability, provide the necessary tools and framework to develop systems for monitoring vegetation on a near real-time basis. Given that many of the remote sensing and climate products are available globally, developing near real-time systems using these products increases the ability to more easily extend them to other areas, thus reducing costs and time of implementation.

This study was implemented with the overarching goal of examining available technologies for implementing a near real-time system to monitoring biomass available to livestock on a given landscape, thus allowing improved monitoring of the forage resources to assist in decision making. This study had three objectives 1) examine the ability of the Climate Prediction Center Morphing Product (CMORPH) and Next Generation Weather Radar (NEXRAD) rainfall products to detect and estimate rainfall at a semi-arid site in West Texas, 2) verify the ability of a simulation model (PHYGROW) to predict herbaceous biomass at selected sites (patches) in a semi-arid landscape using NEXRAD rainfall and 3) examine the feasibility of using cokriging for integrating simulation model (PHYGROW) output and satellite greenness imagery (NDVI) to predict herbaceous biomass across the landscape in the Gobi region of Mongolia. The overall results and conclusions regarding each of these objectives are summarized below.

Ability of NEXRAD and CMORPH to Detect and Estimate Rainfall

The rainfall from the NEXRAD and CMORPH products were compared to rainfall collected at two automated weather stations in West Texas to assess the products ability for detection and estimation of rainfall at the study site and their suitability for use in biophysical modeling. Frequency statistics, bias, error, and estimation efficiency were used to for the comparisons.

At the West Texas study site, the NEXRAD rainfall product outperformed the CMORPH rainfall in terms of both rainfall detection and estimation. NEXRAD had higher accuracy in detecting events, less over-prediction of the number of rainfall events, less false alarms, and higher skill scores than CMORPH. From a rainfall estimation standpoint, NEXRAD had lower estimation bias, lower variability, higher temporal correlations, and higher estimation efficiency than CMORPH. These traits make the NEXRAD product more suitable for use in biophysical modeling compared to CMORPH.

Seasonal differences were apparent in the rainfall detection and estimates for both products. Each had higher probability of detection and overestimation of rainfall in the monsoon season (June 1 to September 30) compared to the non-monsoon season (October 1 to May 31). CMORPH overestimated rainfall in both seasons but had greater overestimation in the monsoon period. Overestimation by CMORPH in the monsoon season may be related to detection of rainfall from convective events where the rainfall evaporates before it reaches the soil surface.

Location effects were apparent with the NEXRAD product with the two stations having opposite trends in estimation even though the sites were 20 km apart. This may be related to radar beam blockage and radar range effects. Although this study used a very limited number of rain gages for comparison to the rainfall products, it does

highlight the local variation that can exist in these rainfall products. Before using these products in biophysical modeling, it may be useful to conduct local validation in order to understand the variability, especially in areas where reporting rain gage networks are sparse.

PHYGROW Simulation Model Performance Using NEXRAD Rainfall

The PHYGROW biophysical simulation model was evaluated to assess its performance in accurately predicting herbaceous biomass at selected sites in West Texas using the NEXRAD rainfall product as a driving variable. The model was first evaluated at the location of two weather stations and calibrated using the rainfall collected from the station gages. The model was then evaluated after substituting the gage collected rainfall with NEXRAD rainfall. Lastly the model was evaluated using NEXRAD data at multiple grazed locations representing the dominant plant communities across the study area.

The PHYGROW model's performance, using rainfall data collected at the weather stations, was moderate to good depending on the performance measure evaluated (estimation efficiency or index of agreement) and the location examined. At both station locations, the PHYGROW model generally tracked the biomass measured at the site, but during several periods lagged the observed data by 20 to 30 days. Additional study is recommended to address these issues to improve model performance.

The replacement of station collected rainfall with NEXRAD rainfall in the calibrated PHYGROW model resulted in poor model performance when compared to the observed biomass data at the two weather station sites. The variability in the predictions increased and the goodness-of-fit statistics dropped, especially for the estimation efficiency statistic. A comparison of the biomass predictions for the NEXRAD and the station simulations indicated that where the NEXRAD and station simulations diverged,

rainfall differences in the 30 days prior to divergence were 25 mm or more. The PHYGROW model appeared to be most sensitive to these differences at the beginning or the end of the monsoon period.

The results of the PHYGROW simulations on the multiple grazed sites using NEXRAD precipitation indicated good correspondence between simulated biomass and that measured at 60 sites across the study area. These results were much better than those observed with the NEXRAD rainfall at the weather station sites. Possible reasons for these differences could be that the multiple grazed site simulations were calibrated with the NEXRAD rainfall, whereas the simulations with NEXRAD at the station sites used models calibrated with rainfall from the station. Parameter adjustment during calibration for the NEXRAD simulations on grazed sites may have adjusted for consistent over or underestimation of rainfall by the NEXRAD product. A second reason is that the sample size and range of the data pairs were greater for the evaluation at the grazed sites, thus reducing the impact of single outliers in the overall evaluation of model performance on the grazed sites.

Results of PHYGROW simulations calibrated with and using NEXRAD corresponded well with the herbaceous biomass collected at multiple sites across a heterogeneous semi-arid landscape and the methodology looks promising for predicting biomass at the patch scale on a near real-time basis. However, additional research is needed to better understand the uncertainties in forage predictions associated with the patterns of over and underestimation of rainfall by the NEXRAD product.

Cokriging to Predict Forage Biomass in the Gobi Region of Mongolia

This study was implemented in the Gobi region of Mongolia to examine the feasibility of developing a forage monitoring system that could provide near real-time spatial and temporal assessment of livestock forage conditions. As part of this assessment the PHYGROW simulation model was evaluated to determine its ability to predict forage

biomass at selected sites across the landscape using a near real-time, high resolution rainfall (CMORPH). A second objective was to evaluate methodology for using the geostatistical technique of cokriging to integrate PHYGROW model output with NDVI to produce landscape maps of forage biomass that could be produced on a near real-time basis.

The PHYGROW model performed reasonably well in predicting forage biomass at the majority of the monitoring sites across the Gobi region. Site that had reduced performance generally overpredicted biomass and were located in the Forest Steppe zones in the northern portions of the study area. Poor performance at these sites was attributed to misparameterization of soils at the site and problems with the CMORPH estimation of rainfall. Additional and more frequent monitoring is needed at these sites to assist in improving model calibration.

Cross-validation indicated that cokriging of PHYGROW output with NDVI as a covariate performed well during the months of July, August, and September with performance measures slightly less than the PHYGROW model calibration and verification performance measures. The exception to this was the estimation efficiency statistic which is sensitive to extreme outliers. The majority of the outliers were located in the Forest Steppe zone and several were identified previously as outliers in the PHYGROW model verification.

The performance of cokriging as assessed through an independent verification was low with extreme outliers in the data again occurring in the Forest Steppe zone. The problems associated with uncertainty in the biomass predictions by the PHYGROW model at sites in the northern portion of the study area carried through to the cokriging, resulting in poor estimates of biomass in the northern regions of the study area.

The overall methodology of geostatistically integrating the PHYGROW model output with NDVI to produce landscape maps of forage biomass shows promise for implementing as a near real-time system for drought monitoring in Mongolia. Improvements in the calibration of the PHYGROW model for sites in the northern regions, as well as more rigorous validation at selected sites across should help reduce the uncertainty in PHYGROW model predictions, and therefore improve accuracy of the interpolated maps.

REFERENCES

Ajami NK, Gupta H, Wagener T, Sorooshian S (2004) Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. Journal of Hydrology 298:112-135

Al-Bakri JT, Taylor JC (2003) Application of NOAA AVHRR for monitoring vegetation conditions and biomass in Jordan. Journal of Arid Environments 54:579-593

Alhamad MN (2002) Integration of point biophysical modeling and NDVI data to improve forecasting of near term forage conditions in Texas. PhD Dissertation, Texas A&M University

Alhamad MN, Stuth J, Vannucci M (2007) Biophysical modelling and NDVI time series to project near-term forage supply: spectral analysis aided by wavelet denoising and ARIMA modelling. International Journal of Remote Sensing 28:2513-2548

Andales AA, Derner JD, Bartling PNS, Ahuja LR, Dunn GH, Hart RH, Hanson JD (2005) Evaluation of GPFAR for simulation of forage production and cow-calf weights. Rangeland Ecology and Management 58:247-255

Arkin PA, Meisner BN (1987) The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. Monthly Weather Review 115:51-74

Atkinson PM, Webster R, Curran PJ (1994) Cokriging with airbone MSS imagery. Remote Sensing of Environment 50:335-345

Baumer OW, Wenberg RD, Rice JW (1987) The use of soil water retention curves in DRAINMOD. Technical Paper, USDA, Soil Conservation Service, Mid-West National Technical Center, Lincoln, NE

Bedunah DJ, Schmidt SM (2004) Pastoralism and protected area management in Mongolia's Gobi Gurvansaikhan National Park. Development and Change 35:167-191

Bekele A, Downer RG, Wolcott MC, Hudnall WH, Moore SH (2003) Comparative evaluation of spatial prediction methods in a field experiment for mapping soil potassium. Soil Science 168:15-28

Bouraoui F, Wolfe ML (1990) Application of hydrologic models to rangelands. Journal of Hydrology 121:173-191

Butt TA, McCarl BA, Angerer J, Dyke PT, Stuth JW (2005) The economic and food security implications of climate change in Mali. Climatic Change 68:355-378

Carlson DH, Thurow TL (1992) SPUR-91:Workbook and user guide. Texas A&M University, Department of Rangeland Ecology and Management in cooperation with USDA Soil Conservation Service. College Station, TX, Pub. No. MP-1743. p. 259

Carlson DH, Thurow TL (1996) Comprehensive evaluation of the improved SPUR model (SPUR-91). Ecological Modelling 85:229-240

Childress WM, Coldren CL, McLendon T (2002) Applying a complex, general ecosystem model (EDYS) in large-scale land management. Ecological Modelling 153:97-108

Chokmani K, Viau AA, Bourgeois G (2005) Regionalization of outputs of two crop protection models using geostatistical tools and NOAA-AVHRR images. Agronomy for Sustainable Development 25:79-92

Ciach GJ, Krajewski WF (1999) On the estimation of radar rainfall error variance. Advances in Water Resources 22:585-595

Clifford P, Richardson S, Hemon D (1989) Assessing the significance of the correlation between two spatial processes. Biometrics 45:123-134

Corson MS, Skinner RH, Rotz CA (2006) Modification of the SPUR rangeland model to simulate species composition and pasture productivity in humid temperate regions. Agricultural Systems 87:169-191

Dobermann A, Ping JL (2004) Geostatistical integration of yield monitor data and remote sensing improves yield maps. Agronomy Journal 96:285-297

Duncan D, Woodmansee R (1975) Forecasting forage yield from precipitation in California's annual rangeland. Journal of Range Management 28:327-329

Dungan J (1998) Spatial prediction of vegetation quantities using ground and image data. International Journal of Remote Sensing 19:267-285

Ebert EE, Janowiak JE, Kidd C (2007) Comparison of near-real-time precipitation estimates from satellite observations and numerical models. Bulletin of the American Meteorological Society 88:47-64

Ersahin S (2003) Comparing ordinary kriging and cokriging to estimate infiltration rate. Soil Science Society of America Journal 67:1848-1855

ESRI (2005a) ArcGIS 9: Using ArcGIS Geostatistical Analyst. ESRI, Redlands, CA

ESRI (2005b) ArcGIS 9: Using ArcMap. ESRI, Redlands, CA

Evans R (1998) Progress reports: The erosional impacts of grazing animals. Progress in Physical Geography 22:251-268

FAO [Food and Agriculture Organization] (1994) EcoCrop Version 1.1. Food and Agriculture Organization. Rome, Italy

Flanagan DC, Nearing MA (1995) WEPP technical documentation. National Soil Erosion Research, Report No. 10. West Lafayette, IN

Foy JK, Teague WR, Hanson JD (1999) Evaluation of the upgraded SPUR model (SPUR2.4). Ecological Modelling 118:149-165

Frank AB, Karn JF (2003) Vegetation indices, CO2 flux, and biomass for Northern Plains Grasslands. Journal of Range Management 56:382-387

Fulton RA, Breidenbach JP, Seo DJ, Miller DA, O'Bannon T (1998) The WSR-88D rainfall algorithm. Weather and Forecasting 13:377-395

Gloaguen E, Chouteau M, Marcotte D, Chapuis R (2001) Estimation of hydraulic conductivity of an unconfined aquifer using cokriging of GPR and hydrostratigraphic data. Journal of Applied Geophysics 47:135-152

Goovaerts P (1998) Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. Biology and Fertility of Soils 27:315-334

Grant WE, Pedersen EK, Marín SL (1997) Ecology and natural resource management: systems analysis and simulation. Wiley, NY

Grimes DIF, Pardo-Iguzquiza E, Bonifacio R (1999) Optimal areal rainfall estimation using rain gauges and satellite data. Journal of Hydrology 222:93-108

Gunin PD, Vostokova EA, Dorofeyuk NI, Tarasov PE, Black CC (1999) Vegetation dynamics of Mongolia. Kluwer Academic Publishers, Boston, MA

Hanson JD, Baker BB, Bourdon RM (1992) SPUR2 documentation and user guide. GPSR Technical Report No. 1. USDA, Fort Collins, CO

Herman A, Kumar VB, Arkin PA, Kousky JV (1997) Objectively determined 10-day African rainfall estimates created for famine early warning systems. International Journal of Remote Sensing 18:2147-2159

Holechek J, Pieper RD, Herbel CH (1995) Range management: principles and practices. Prentice Hall, Englewood Cliffs, NJ

Hudak AT, Lefsky MA, Cohen WB, Berterretche M (2002) Integration of lidar and Landsat ETM plus data for estimating and mapping forest canopy height. Remote Sensing of Environment 82:397-416

Hunt ER, Miyake BA (2006) Comparison of stocking rates from remote sensing and geospatial data. Rangeland Ecology and Management 59:11-18

Hutchinson CF (1991) Uses of satellite data for famine early warning in sub-Saharan Africa. International Journal of Remote Sensing 12:1405-1421

Isaaks EH, Srivastava RM (1989) Applied geostatistics. Oxford University Press, NY

Janowiak JE, V. E. Kousky, and R. J. Joyce (2005) Diurnal cycle of precipitation determined from the CMORPH high spatial and temporal resolution global precipitation analyses. Journal of Geophysical Research 110:D23105

Jayakrishnan R, Srinivasan R, Arnold JG (2004) Comparison of raingage and WSR-88D Stage III precipitation data over the Texas-Gulf basin. Journal of Hydrology 292:135-152

Johnson LE, Olsen BG (1998) Assessment of quantitative precipitation forecasts. Weather and Forecasting 13:75-83.

Journel AG, Huijbregts CJ (1978) Mining geostatistics. Academic Press, NY.

Joyce RJ, Janowiak JE, Arkin PA, Xie PP (2004) CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. Journal of Hydrometeorology 5:487-503

Kalin L, Hantush MM (2006) Hydrologic modeling of an eastern Pennsylvania watershed with NEXRAD and rain gauge data. Journal of Hydrologic Engineering 11:555-569

King SL, Lister AJ, Hoppus ML (2003) Estimating and mapping five forest attributes with satellite ancillary data. In: G.J. Arthaud and T.M. Barrett (eds) Systems Analysis in Forest Resources - Proceedings of the Eighth Symposium. Snowmass Village, CO, September 27-30, 2000. Kluwer Academic Publishing, Boston, MA p 340

Kiniry JR, Sanchez H, Greenwade J, Seidensticker E, Bell JR, Pringle F, Peacock G, Rives J (2002) Simulating grass productivity on diverse range sites in Texas. Journal of Soil and Water Conservation 57:144-150

Kogan F, Stark R, Gitleson A, Jargalsaikhan C, Dugrajav C, Tsooj S (2004) Derivation of pasture biomass in Mongolia from AVHRR-based vegetation health indices. International Journal of Remote Sensing 25:2889-2896

Legates DR, McCabe Jr. GJ (1999) Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resources Research 35:233–241

Legates DR (2000) Real-time calibration of radar precipitation estimates. Professional Geographer 52:235-246

McCollum JR, Gruber A, Ba MB (2000) Discrepancy between gauges and satellite estimates of rainfall in equatorial Africa. Journal of Applied Meteorology 39:666-679

McCollum JR, Krajewski WF, Ferraro RR, Ba MB (2002) Evaluation of biases of satellite rainfall estimation algorithms over the continental United States. Journal of Applied Meteorology 41:1065-1080

MDA Federal (2004) Landsat GeoCover ETM+ 2000 Edition Mosaics Tiles N-13-25 and N-13-30. ETM-EarthSat-MrSID, 1.0. USGS, Sioux Falls, SD. Available from http://glcf.umiacs.umd.edu/data/mosaic (accessed on 9/17/2007)

Montieth JL (1972) Solar radiation and productivity in tropical ecosystems. Applied Ecology 9:747-766

Montieth JL (1977) Climate and efficiency of crop production in Britain. Philosophical Transactions of the Royal Society of London, Series B 281:277-294

Moon J, Srinivasan R, Jacobs JH (2004) Stream flow estimation using spatially distributed rainfall in the Trinity River Basin, Texas. Transactions of the ASAE 47:1445-1451

Mueller TG, Pierce FJ (2003) Soil carbon maps: Enhancing spatial estimates with simple terrain attributes at multiple scales. Soil Science Society of America Journal 67:258-267

Mutanga O, Rugege D (2006) Integrating remote sensing and spatial statistics to model herbaceous biomass distribution in a tropical savanna. International Journal of Remote Sensing 27:3499-3514

Nain AS, Dadhwal VK, Singh TP (2002) Real time wheat yield assessment using technology trend and crop simulation model with minimal data set. Current Science 82:1255-1258

Nanos N, Calama R, Montero G, Gil L (2004) Geostatistical prediction of height/diameter models. Forest Ecology and Management 195:221-235

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I - A discussion of principles. Journal of Hydrology 10:282-290

NCDC [National Climatic Data Center] (2006) Monthly Station Climate Summaries, Climatography of the United States No.20 1971-2000. National Climatic Data Center, Asheville, NC. Available from http://cdo.ncdc.noaa.gov/climatenormals/clim20/tx/415579.pdf (accessed 4/15/2007)

NOAA [National Oceanic and Atmospheric Administration] (2007) National Weather Service New Precipitation Analysis. National Oceanic and Atmospheric Administration. Available from http://www.srh.noaa.gov/rfcshare/precip_about.php (accessed July 1, 2007)

O'Connor TG, Haines LM, Snyman HA (2001) Influence of precipitation and species composition on phytomass of a semi-arid African grassland. Journal of Ecology 89:850-860

Ott L, Longnecker M (2001) An introduction to statistical methods and data analysis. Duxbury, Pacific Grove, CA

Pierson FB, Carlson DH, Spaeth KE (2001) A process-based hydrology submodel dynamically linked to the plant component of the simulation of production and utilization on rangelands SPUR model. Ecological Modelling 141:241-260

Pineiro G, Oesterheld M, Paruelo JM (2006) Seasonal variation in aboveground production and radiation-use efficiency of temperate rangelands estimated through remote sensing. Ecosystems 9:357-373

Powell AM (1998) Trees and shrubs of the Trans-Pecos and adjacent areas. University of Texas Press, Austin, TX

Reed SM, Maidment DR (1999) Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling. Journal of Hydrologic Engineering 4:174-182

Reeves MC, Winslow JC, Running SW (2001) Mapping weekly rangeland vegetation productivity using MODIS algorithms. Journal of Range Management 54:90-105

Rosenberg MS (2000) PASSAGE. Pattern analysis, spatial statistics, and geographic exegesis. Version 1.0. Department of Biology, Arizona State University, Tempe, AZ

Rossi RE, Dungan JL, Beck LR (1994) Kriging in the shadows - geostatistical interpolation for remote-sensing. Remote Sensing of Environment 49:32-40

Rowland J, Verdin J, Adoum A, Senay G (2005) Drought monitoring techniques for famine early warning systems in Africa. In: Boken VK, Cracknell AP and Heathcote RL (eds), Monitoring and predicting agricultural drought: a global study. Oxford University Press, NY, pp. 252-263

Ryan Z (2005) Establishment and evaluation of a livestock early warning system for Laikipia, Kenya. Thesis, Texas A&M University

Sannier CAD, Taylor JC, Du Plessis W (2002) Real-time monitoring of vegetation biomass with NOAA-AVHRR in Etosha National Park, Namibia, for fire risk assessment. International Journal of Remote Sensing 23:71-89

Saxton KE, Rawls WJ, Romberger JS, Papendick RI (1986) Estimating generalized soil-water characteristics from texture. Soil Science Society of America Journal 50:1031-1036

Schino G, Borfecchia F, De Cecco L, Dibari C, Iannetta M, Martini S, Pedrotti F (2003) Satellite estimate of grass biomass in a mountainous range in central Italy. Agroforestry Systems 59:157-162

Schmidt J, Lawrence B, Olsen B (2000) A comparison of operational precipitation processing methodologies. National Oceanic and Atmospheric Administration, Fort Worth, TX. Available from http://www.srh.noaa.gov/abrfc/p1vol.html (accessed June 1, 2007)

Scurlock JMO, Asner GP, Gower ST (2001). Global leaf area index data from field measurements, 1932-2000. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, TN. Available from http://www.daac.ornl.gov (accessed 11/15/2007)

Seo D-J, Breidenbach JP (2002) Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. Journal of Hydrometeorology 3:93-111

Siurua H, Swift J (2002) Drought and dzud but no famine (yet) in the Mongolian herding economy. Institute of Development Studies Bulletin 33:88-97

Smith JA, Seo DJ, Baeck ML, Hudlow MD (1996) An intercomparison study of NEXRAD precipitation estimates. Water Resources Research 32:2035-2046

Smith R (2001) Hydrogeology of the Marathon Basin, Brewster County, Texas. In: Mace RE, Mullican WF and Angle ES (eds), Aquifers of West Texas. Texas Water Development Board, Austin, TX, pp. 190-206.

Soil Survey Staff (2007) Soil Survey geographic (SSURGO) database for Brewster County, Texas. Natural Resources Conservation Service, United States Department of Agriculture. Available from http://soildatamart.nrcs.usda.gov (accessed 12/1/2007).

Stanski HR, Wilson LJ, Burrows W (1989) Survey of common verification methods in meteorology. World Weather Watch Report No. 8. TD No. 358. World Meteorological Organization, Geneva, Switzerland

Stuth JW, Angerer J, Kaitho R, Zander K, Jama A, Heath C, Bucher J, Hamilton W, Conner R, Inbody D (2003b) The Livestock Early Warning System (LEWS): Blending technology and the human dimension to support grazing decisions. Arid Lands Newsletter, University of Arizona. Available from http://cals.arizona.edu/OALS/ALN/aln53/stuth.html (accessed 7/1/2007)

Stuth JW, Schmitt D, Rowan RC, Angerer JP, Zander K (2003a) PHYGROW Users Guide and Technical Documentation. Texas A&M University. Available from http://cnrit.tamu.edu/phygrow/PHYGROW_userguide.pdf (accessed June 1, 2007)

Stuth JW, Angerer J, Kaitho R, Jama A, Marambii R (2005) Livestock early warning system for Africa rangelands. In: Boken VK, Cracknell AP and Heathcote RL (eds), Monitoring and predicting agricultural drought: a global study. Oxford University Press, NY, pp. 283-294

Teague WR, Foy JK (2004) Can the SPUR rangeland simulation model enhance understanding of field experiments? Arid Land Research and Management 18:217-228

Thoma DP, Bailey DW, Long DS, Nielsen GA, Henry MP, Breneman MC, Montagne C (2002) Short-term monitoring of rangeland forage conditions with AVHRR imagery. Journal of Range Management 55:383-389

Trimble SW, Mendel AC (1995) The cow as a geomorphic agent - A critical review. Geomorphology 13:233-253

Tucker CJ, Vanpraet C, Boerwinkel E, Gaston A (1983) Satellite remote-sensing of total dry-matter production in the Senegalese Sahel. Remote Sensing of Environment 13: 461-474

Tucker CJ, Vanpraet CL, Sharman MJ, Vanittersum G (1985) Satellite remote-sensing of total herbaceous biomass production in the Senegalese Sahel - 1980-1984. Remote Sensing of Environment 17:233-249.

Tucker CJ, Sellers PJ (1986) Satellite remote-sensing of primary production. International Journal of Remote Sensing 7:1395-1416

Tucker CJ, Pinzon JE, Brown ME, Slayback DA, Pak EW, Mahoney R, Vermote EF, El Saleous N (2005) An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. International Journal of Remote Sensing 26:4485-4498

USGS [United States Geological Survey] (1999). USGS national elevation dataset. EROS Data Center, Sioux Falls, SD. Available from http://ned.usgs.gov/ (accessed 12/15/2007)

Wang GX, Gertner G, Fang SF, Anderson AB (2003) Mapping multiple variables for predicting soil loss by geostatistical methods with TM images and a slope map. Photogrammetric Engineering and Remote Sensing 69:889-898

Wang X, Xie H, Sharif H, Zeitler J (2008) Validating NEXRAD MPE and Stage III precipitation products for uniform rainfall on the Upper Guadalupe River Basin of the Texas Hill Country. Journal of Hydrology 348:73-86

Weber GE, Moloney K, Jeltsch F (2000) Simulated long-term vegetation response to alternative stocking strategies in savanna rangelands. Plant Ecology 150:77-96.

Wessels KJ, Prince SD, Zambatis N, Macfadyen S, Frost PE, Van Zyl D (2006) Relationship between herbaceous biomass and 1-km^2 Advanced Very High Resolution Radiometer (AVHRR) NDVI in Kruger National Park, South Africa. International Journal of Remote Sensing 27:951-973

Whiton RC, Smith PL, Bigler SG, Wilk KE, Harbuck AC (1998a) History of operational use of weather radar by U.S. weather services. Part I: The pre-NEXRAD era. Weather and Forecasting 13:219-243

Whiton RC, Smith PL, Bigler SG, Wilk KE, Harbuck AC (1998b) History of operational use of weather radar by U.S. weather services. Part II: Development of operational Doppler weather radars. Weather and Forecasting 13:244-252.

Wight JR, Neff EL (1983) Soil-vegetation-hydrology studies; volume 2, a user manual for ERHYM: The Ekalaka Rangeland Hydrology and Yield Model. Agriculture Research Results, ARR-W-29. U.S. Department of Agriculture, Agricultural Research Service. Washington, DC, p 38

Wight JR, Skiles JW (1987) SPUR: Simulation and production and utilization of rangelands, documentation and user's guide. ARS Publication 63, United States Department of Agriculture, Agricultural Research Service, USDA, Fort Collins, CO, p 272

Wilcox BP, Rawls WJ, Brakensiek DL, Wight JR (1990) Predicting runoff from rangeland catchments: a comparison of two models. Water Resources Research 26:2401-2410

Willmott CJ (1982) Some Comments on the Evaluation of Model Performance. Bulletin of the American Meteorological Society 63:1309-1313

Willmott. CJ, Ackleson SG, Davis RE, Feddema JJ, Klink KM, Legates DR, O'Donnell J, Rowe CM (1985) Statistics for the evaluation and comparison of models. Journal of Geophysical Research 90:8995–9005

Wylie BK, Harrington JA, Prince SD, Denda I (1991) Satellite and ground-based pasture production assessment in Niger: 1986-1988. International Journal of Remote Sensing 12:1281-1300

Xie HJ, Zhou XB, Hendrickx JMH, Vivoni ER, Guan HD, Tian YQ, Small EE (2006) Evaluation of NEXRAD Stage III precipitation data over a semiarid region. Journal of the American Water Resources Association 42:237-256

Xie P, Arkin PA (1998) Global monthly precipitation estimates from satellite-observed outgoing longwave radiation. Journal of Climate 11:137-164

Young CB, Nelson BR, Bradley AA, Smith JA, Peters-Lidard CD, Kruger A, Baeck ML (1999) An evaluation of NEXRAD precipitation estimates in complex terrain. Journal of Geophysical Research-Atmospheres 104:19691-19703

Young CB, Bradley AA, Krajewski WF, Kruger A, Morrissey ML (2000) Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. Journal of Hydrometeorology 1:241-254

Yunatov AA, Dashnima B, Gerbikh AA (1979) Vegetation Map of the Mongolian People's Republic. Naukia. Moscow, Russia

APPENDIX A

Appendix A.1. Cross-validation analysis statistics for kriging of PHYGROW simulation model output to estimate forage standing crop across the Gobi region of Mongolia during the growing seasons (June to September) in 2005 and 2006.

	2005 Cross-validation							
Statistic	**June 1-15**	**June 16-30**	**July 1-15**	**July 16-31**	**August 1-15**	**August 16-31**	**Sept 1-15**	**Sept 16-30**
Simulation Mean (kg/ha)	48	85	134	170	217	243	244	220
Cokriged Map Mean (kg/ha)	47	84	133	169	215	241	242	218
[1]sd_o (kg/ha)	58	97	145	185	237	272	276	263
sd_s (kg/ha)	30	63	106	140	187	215	214	202
Bias (%)	-1	-1	-1	-1	-1	-1	-1	-1
MBE (kg/ha)	0	-1	-1	-1	-2	-1	-2	-2
MAE (kg/ha)	36	52	65	78	93	103	105	99
RMSD (kg/ha)	51	74	99	121	147	172	177	170
r^2	0.24	0.41	0.53	0.57	0.61	0.60	0.59	0.58
EE	-1.84	-0.39	0.13	0.25	0.38	0.36	0.31	0.29
d	0.62	0.76	0.83	0.85	0.87	0.86	0.85	0.85
n	243	243	243	243	243	243	243	243

	2006 Cross-validation							
	June 1-15	**June 16-30**	**July 1-15**	**July 16-31**	**August 1-15**	**August 16-31**	**Sept 1-15**	**Sept 16-30**
Simulation Mean (kg/ha)	35	68	116	180	231	245	235	214
Cokriged Map Mean (kg/ha)	34	66	114	178	229	244	234	213
[1]sd_o (kg/ha)	64	97	149	213	268	293	292	276
sd_s (kg/ha)	41	64	109	167	210	233	236	223
Bias (%)	-3	-2	-2	-1	-1	-1	-1	-1
MBE (kg/ha)	-1	-2	-2	-2	-2	-1	-1	-1
MAE (kg/ha)	29	45	63	81	103	109	106	98
RMSD (kg/ha)	48	74	99	127	160	172	169	160
r^2	0.42	0.42	0.56	0.64	0.64	0.65	0.67	0.67
EE	-0.41	-0.35	0.18	0.42	0.41	0.45	0.49	0.49
d	0.76	0.76	0.84	0.88	0.88	0.88	0.89	0.89
n	243	243	243	243	243	243	243	243

[1]sd_o = standard deviation for model simulated biomass; sd_s = standard deviation for cokriged biomass; MBE = Mean Bias Error; MAE = Mean Absolute Error; RMSD = Root Mean Square Difference; r^2 = coefficient of determination; EE = estimation efficiency; *d* = index of agreement; *n* = number of samples

VITA

Name: Jay Peter Angerer

Address: Department of Ecosystem Science and Management, Texas A&M University, 2138 TAMU, College Station, TX 77843-2138

Email Address: jangerer@tamu.edu

Education: B.S., Range Management, Texas Tech University, 1986
M.S., Range Science, Texas A&M University, 1991
Ph.D. Rangeland Ecology and Management, Texas A&M University, 2008