



Wheat yield estimates using multi-temporal AVHRR-NDVI satellite imagery
by Mari Patricia Henry

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Land Resources and Environmental Sciences

Montana State University

© Copyright by Mari Patricia Henry (1999)

Abstract:

Crop condition monitoring and early season yield estimates could provide information for assessing crop stress, fertilizer demand, and productivity and assist in farm management decisions. I examined the application of the Advanced Very High Resolution Radiometer (AVHRR) Normalized Difference Vegetation Index (NDVI) for making crop yield estimates at regional, county, and farm levels in Montana. Seasonal growth profiles were examined to determine if yield and/or protein concentration could be accurately estimated prior to harvest.

Results are presented for six regions, thirty-nine counties, and five farms in Montana. Biweekly NDVI values were extracted to produce yearly growth profiles representing April through mid-September for the years 1989 through 1997. Wheat yield data were supplied from Montana Agricultural Statistics Service and from cooperating farmers. Protein concentration - NDVI parameter relationships were developed for four farm sites. NDVI values were integrated and summed in various ways across the growth profile to find the best model for wheat yield estimation. Sixteen NDVI growth profile parameters were computed at the region and county level whereas twenty-nine were computed at the farm level. A multiple linear regression model was used to determine overall relationships between NDVI parameters and yield or protein concentration.

. Regions, counties, and farms were included as indicator variables in the model, along with interaction terms, when significant (p -value = 0.05).

Results show correlation between yield and several NDVI growth profile parameters at the region, county, and farm level. Regions showed strong relationships between wheat yields and integrated NDVI over the entire growing season (adj. $R^2 = 0.753$, $p = 0.0001$) and with late season NDVI parameters (summation period 7/6 - 8/30 adj. $R^2 = 0.686$; summation through August adj. $R^2 = 0.745$). Counties exhibited similar, yet weaker relationships between wheat yield and NDVI parameters. Farms revealed strong relationships between integrated NDVI over the entire growing season and spring wheat yields (adj. $R^2 = 0.628$) that improved when adjusted to the apparent growing season (AGS) (adj. $R^2 = 0.688$). Protein concentration was also strongly correlated with integrated NDVI when adjusted for AGS (adj. $R^2 = 0.789$). Early season NDVI parameters showed weak relationships with spring wheat yields and percent protein content and thus, could not provide accurate yield and protein concentration estimates.

Results indicate a need for region, county, and farm-specific calibration of NDVI growth profiles and refinement of integration and summation periods to improve wheat yield and protein concentration estimates. The use of AVHRR-NDVI growth profiles at the regional level provided the best yield estimates. At the farm scale, the spatial resolution (1km^2) limited the certainty for accurate portrayal of field locations. However, our models provide a basis for further examination of time-series NDVI data with satellite/sensor systems with higher spatial and spectral resolution to be launched in the near future.

**WHEAT YIELD ESTIMATES USING MULTI-TEMPORAL AVHRR-NDVI
SATELLITE IMAGERY**

by

Mari Patricia Henry

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

of

Master of Science

in

Land Resources and Environmental Sciences

MONTANA STATE UNIVERSITY-BOZEMAN
Bozeman, Montana

April 1999

N378
H3965

APPROVAL

of a thesis submitted by

Mari P. Henry

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

Gerald Nielsen Gerald Nielsen 4-14-99

Approved for the Department of Land Resources and Environmental Sciences

Jeff Jacobsen Jeff Jacobsen 4/19/99

Approved for the College of Graduate Studies

Bruce R. McLeod Bruce R. McLeod 4-24-99

STATEMENT OF PERMISSION TO USE

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

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

Signature Mari P. Henry

Date 4-13-99

ACKNOWLEDGMENTS

First, I would like to thank my major professor, Jerry Nielsen, for providing me an opportunity to earn my degree, and for believing in me and encouraging me along the way. Thanks also to my committee members, Rick Lawrence, Rick Engel, and Dan Long, for their advice, assistance, and support in this project. Special thanks to Linda Gillette and Bob Snyder for assisting me in getting this project off the ground and lending me data to use. Thanks to Dave Thoma, Shane Porter, Chris Erlien, Elizabeth Roberts, and Diana Cooksey for listening and providing advice and encouragement. Finally, I wish to thank my husband, Dale, for helping me make this possible, and my daughter Kayla, for reminding me what is really important.

Funding for this project was provided in part from NASA/UMAC grants #NCC5-310 and NAG5-3616, by the Upper Midwest Aerospace Consortium, Montana Space Grant Consortium, and the Montana Agricultural Experiment Station and the Montana State University Department of Land Resources and Environmental Sciences. I would like to thank these organizations for making this project possible.

TABLE OF CONTENTS

	Page
ABSTRACT	ix
CHAPTER 1 - GENERAL INTRODUCTION	1
Literature Review	3
CHAPTER 2 - REGIONAL AND COUNTY SCALE WHEAT YIELD ESTIMATION USING MULTI-TEMPORAL AVHRR-NDVI SATELLITE IMAGERY	10
Introduction	10
Methods and Materials	11
Montana physiography and climate	11
Montana non-irrigated agricultural lands	13
Montana county yield data	16
AVHRR-NDVI processing	16
NDVI growth profile analysis	18
Results	22
Wheat yield – Regional NDVI parameters relationships	22
Analysis of best model scatterplots	24
County Yield – NDVI relationships	27
Discussion	30
Regional yield – NDVI relationships	30
County yield – NDVI relationships	33
Sources of error	35
Conclusions	36
CHAPTER 3 - FARM-SCALE WHEAT YIELD AND PROTEIN CONCENTRATION ESTIMATION USING MULTI-TEMPORAL AVHRR-NDVI SATELLITE IMAGERY	37
Introduction	37
Methods and Materials	39
Study sites	39
Satellite data processing and analysis	42
Growth profile analysis	43

Results	46
Spring wheat yield – NDVI parameters relationships.....	46
Analysis of best model scatterplots.....	49
Protein concentration – NDVI parameters relationships	52
Analysis of best model scatterplots.....	55
Discussion	58
Spring wheat yield – NDVI parameter relationships	58
Protein concentration – NDVI parameters relationships	61
Sources of error	63
Potential for real-time crop monitoring	64
Conclusions	67
REFERENCES.....	68

LIST OF TABLES

	Page
Table 1. Counties and percentage of acres for each county included in study.....	15
Table 2. Wheat yield – NDVI parameter relationships for regional study sites.....	22
Table 3. Wheat yield – NDVI parameter relationships for county study sites.....	28
Table 4. Spring wheat yield – NDVI parameter relationships for farm study sites.....	47
Table 5. Protein concentration – NDVI parameter relationships for farm study sites.....	53

LIST OF FIGURES

		Page
Figure 1.	MAPS Atlas coverage of dryland agriculture in Montana.....	14
Figure 2.	Regions and associated counties included in study.....	14
Figure 3.	NDVI seasonal growth profiles for a Montana county.....	17
Figure 4.	An example of integrating the apparent growing season (AGS).....	19
Figure 5.	Regional wheat yield – Integrate 12 periods relationships.....	25
Figure 6.	Regional wheat yield –NDVI (sum through Aug.) relationships.....	26
Figure 7.	Regional wheat yield – NDVI critical period (7/6 – 8/30) relationships...	27
Figure 8.	Five farm sites included in study.....	39
Figure 9.	Farm spring wheat yield – NDVI (integrate AGS) relationships.....	49
Figure 10.	Farm spring wheat yield – NDVI (Sum Aug.) relationships.....	51
Figure 11.	Farm spring wheat yield – NDVI (End July) relationships.....	52
Figure 12.	Farm protein concentration – NDVI (Integrate AGS) relationships.....	56
Figure 13.	Farm protein concentration – NDVI (Ending Julian Day) relationships...	57
Figure 14.	Departure from 10-year average NDVI map for Site 2.....	66
Figure 15.	Site 2 departure from average data for the 1998 growing season.....	66

ABSTRACT

Crop condition monitoring and early season yield estimates could provide information for assessing crop stress, fertilizer demand, and productivity and assist in farm management decisions. I examined the application of the Advanced Very High Resolution Radiometer (AVHRR) Normalized Difference Vegetation Index (NDVI) for making crop yield estimates at regional, county, and farm levels in Montana. Seasonal growth profiles were examined to determine if yield and/or protein concentration could be accurately estimated prior to harvest.

Results are presented for six regions, thirty-nine counties, and five farms in Montana. Biweekly NDVI values were extracted to produce yearly growth profiles representing April through mid-September for the years 1989 through 1997. Wheat yield data were supplied from Montana Agricultural Statistics Service and from cooperating farmers. Protein concentration – NDVI parameter relationships were developed for four farm sites. NDVI values were integrated and summed in various ways across the growth profile to find the best model for wheat yield estimation. Sixteen NDVI growth profile parameters were computed at the region and county level whereas twenty-nine were computed at the farm level. A multiple linear regression model was used to determine overall relationships between NDVI parameters and yield or protein concentration. Regions, counties, and farms were included as indicator variables in the model, along with interaction terms, when significant (p -value = 0.05).

Results show correlation between yield and several NDVI growth profile parameters at the region, county, and farm level. Regions showed strong relationships between wheat yields and integrated NDVI over the entire growing season (adj. R^2 = 0.753, p = 0.0001) and with late season NDVI parameters (summation period 7/6 – 8/30 adj. R^2 = 0.686; summation through August adj. R^2 = 0.745). Counties exhibited similar, yet weaker relationships between wheat yield and NDVI parameters. Farms revealed strong relationships between integrated NDVI over the entire growing season and spring wheat yields (adj. R^2 = 0.628) that improved when adjusted to the apparent growing season (AGS) (adj. R^2 = 0.688). Protein concentration was also strongly correlated with integrated NDVI when adjusted for AGS (adj. R^2 = 0.789). Early season NDVI parameters showed weak relationships with spring wheat yields and percent protein content and thus, could not provide accurate yield and protein concentration estimates.

Results indicate a need for region, county, and farm-specific calibration of NDVI growth profiles and refinement of integration and summation periods to improve wheat yield and protein concentration estimates. The use of AVHRR-NDVI growth profiles at the regional level provided the best yield estimates. At the farm scale, the spatial resolution (1km^2) limited the certainty for accurate portrayal of field locations. However, our models provide a basis for further examination of time-series NDVI data with satellite/sensor systems with higher spatial and spectral resolution to be launched in the near future.

CHAPTER 1

GENERAL INTRODUCTION

Montana is largely an agricultural state within the northern Great Plains region. Agriculture accounts for over 30% of the state's basic industry employment, labor income, and gross sales (Montana Agricultural Statistics, 1997). Approximately 61% of Montana land is either farmed or ranched, with nearly 30% in cropland and an average farm size of 1040 ha. Wheat (spring, winter, and durum) is the primary crop, with barley ranking second (Montana Agricultural Statistics, 1997). Agricultural production in this region is characterized by risk due to weather, international markets, and consumer preference (Seielstad, 1995). While risk can never be eliminated, it can be minimized with access to timely information that would allow farm and ranch managers to monitor crop condition and make better, more informed management decisions.

Since the early 1970's, satellite remote sensing has been promoted as a potentially valuable tool for agricultural monitoring because of its frequent synoptic coverage and ability to "see" in many spectral wavelengths (Hinzman et al., 1986; Quarmby et al., 1993). Numerous studies have shown the possibilities of remotely monitoring phenological and physiological change in crop canopies at the global-, regional-, and farm-scales using various satellite, airborne and ground sensors (Pinter et al., 1981; Benedetti & Rossini, 1993; Fischer, 1994; Reed et al., 1994; Blackmer et al., 1996). In particular, spectral vegetation indices, which reduce multi-band observations to a single number, have been used to monitor crop development and estimate final yields (Weigand

& Richardson, 1990; Rasmussen, 1992; Groten, 1993; Quarmby et al., 1993; Doraiswamy & Cook, 1995). With several years of remotely sensed data now available, multi-temporal analyses of spectral data is developing as an exciting research area (Moran, 1996).

Presently, AVHRR-NDVI satellite imagery has relatively low spatial resolution (1km^2) but is available on a daily, weekly, or biweekly basis. This relatively high temporal resolution might allow for a near real-time monitoring and assessment of crop condition that is crucial in farm management. While AVHRR-NDVI might not be appropriate for assessing yield differences within fields due to low spatial resolution, a time-series of image data can provide a historical and near real-time record of crop performance of a region within a growing season. This information, used in conjunction with managers' knowledge of their own land, could prove to be beneficial not only to farm and ranch managers in assessing crop or rangeland performance, but also to providers of remote sensing services eager to find ways of ground-truthing their data. Together these two groups can advance the science and applications of remotely sensed data in agricultural production.

This study investigates the potential of the Normalized Difference Vegetation Index (NDVI) produced from the Advanced Very High Resolution Radiometer (AVHRR) for regional and farm-scale monitoring of wheat production in Montana. The goal was to determine whether AVHRR-NDVI time-series profiles could be used to estimate wheat productivity (yield) and quality (protein concentration) at the regional and/or farm scale and determine if early estimates of wheat yield and protein concentration would be useful to farmers and land managers.

Chapter 1 provides a literature review on the use of remotely sensed data for studying vegetation dynamics and, in particular, the use of AVHRR-NDVI satellite imagery for estimation of crop yields. Chapter 2 describes AVHRR-NDVI time series and wheat yield relationships at the regional and county scale. At these scales, regional patterns of productivity can be seen and addressed with AVHRR-NDVI. In Chapter 3, the capability of AVHRR-NDVI time series profiles for estimating wheat yields and protein concentration at the farm scale is examined.

Literature Review

Traditional methods of monitoring crops throughout the growing season include climatic or plant process models such as Growing Degree-Days (GDD) or the Crop Environment Resource Synthesis (CERES) Wheat model (Wiegand and Richardson, 1990). These process models use weather, soils, and other environmental data as response functions to describe development, photosynthesis, evapotranspiration, and yield for a specific crop. Though based on strong physiological and physical concepts, these models are poor predictors when spatial variability in soils, stresses, or management practices are present or when simultaneous, multiple variables affect yields (Wiegand, 1984; Wiegand & Richardson, 1990).

Weigand and Richardson (1990) recognized that plant development, stress response, and yield capabilities are expressed in plant canopies and could be observed using spectral analysis of various spectral vegetation indices. Spectral vegetation indices (VIs) are typically a sum, difference, or ratio of two or more spectral wavelengths

(Wiegand et al., 1991). VIs are often produced by ratios or combinations of red (R) and near-infrared (NIR) spectral bands (Wiegand & Richardson, 1990). Plant chlorophyll absorbs incident radiation in the visible red (0.6 - 0.7 μ m) and thus reflects very little in this wavelength (10%), while plant mesophyll causes strong reflectivity (40-60%) due to scattering in the near infrared (0.75 - 1.35 μ m) (Knipling, 1970). VIs are highly correlated with photosynthetic activity in non-wilted plant foliage and have been shown to be good predictors of plant canopy biomass, vigor, or stress (Tucker, 1979). VIs have been correlated to canopy development and used to quantify canopy responses and measure the amount of photosynthetically active plant tissue present (Hatfield, 1983; Wiegand et al., 1986b). Tucker et al. (1981) found that red and infrared spectral data measured with a hand-held radiometer were highly related ($r^2 = 0.86$) to canopy vigor (dry matter accumulation) of winter wheat and suggested that measurements taken from satellite systems could be used to monitor plant growth and development. Wiegand & Richardson (1990) found the Normalized Difference Vegetation Index (NDVI) to be a good measure of absorbed photosynthetically active radiation (APAR), and its relative magnitude and rate of change during maturation an indication of the relative number of fruit or seed to expect per unit area.

When sequential VI observations are taken frequently over a season, profiles can be developed that show the progression of crop canopy emergence, maturity, and senescence, which are factors that reflect crop performance and are related to crop yields (Pinter et al., 1981; Malingreau, 1989; Benedetti & Rossini, 1993). The integration of VIs over time should reveal the productive history of the canopy and tell us, by inference, about vegetation production (Malingreau, 1989). Seasonal growth profiles have been

related to specific physiological changes in crop canopies. Agronomic variables such as ear water concentration (EWC), crop senescence rates, early crop emergence, and final grain yields have been estimated using multi-temporal VI (Idso et al., 1980; Boissard et al., 1993; Benedetti & Rossini, 1993). Examination of crop cycles over many growing seasons and identification of critical times in crop growth cycles have been identified recently as potential research areas that could provide a basis for crop monitoring and prediction of final grain yield (Moran, 1996).

Studies that derived pre-harvest yield estimates typically correlated final grain yield with a single VI observation or time-integrated VI over a specific period (Idso et al., 1980; Weigand et al., 1991; Benedetti & Rossini, 1993; Quarmby et al., 1993; Doraiswamy & Cook, 1995). Grain yields have been found to be correlated with time of maximum NDVI (Barnett & Thomson, 1983), with time-integrated NDVI during the reproductive phase (Rasmussen, 1992), with crop senescence rates (Idso et al., 1980), and with phenological stages from anthesis-to-maturity (Boissard et al., 1993). Yield estimations made by examining early periods in the crop growth profile would be more useful for early season management. Yield predictions made by Rudorff and Batista (1991) correlated the Ratio Vegetation Index (RVI) at a period 50 – 60 days after wheat crop emergence (end of stem elongation and beginning of heading stage) and found that it could be used to explain 64% of the yield variability. Hershenhorn (1992) reported that end of May and end of June AVHRR NDVI were highly correlated with relative spring wheat yield, winter wheat yield, and range condition in the northeastern Montana region for 1985 through 1988. She hypothesized that if data became available in “real-time”, regional yield predictions based on NDVI could be made available in early season.

Another approach is to integrate the area under the seasonal growth profile and correlate this value with final grain yield. In this approach, assessment of the efficiency of the vegetative phase (accumulated biomass) using the integrated NDVI during the first phase of the development curve can be made and then yields adjusted as the NDVI evolves thereafter (Hatfield, 1983). In general, larger areas under the curve should correspond with more photosynthetic activity and biomass and result in larger yields. This has been shown to be particularly true for fodder crops, in which final yield is made of the entire above ground production (Malingreau, 1989; Benedetti & Rossini, 1993). However, for cereal crops whose yield depends on the efficiency of their leaves to assimilate CO² into storage organs, the VI link with yield is indirect and more difficult to measure (Benedetti & Rossini, 1993). One of the critical periods for wheat is grain-fill (flowering to physiologic maturity). During this period, the upper two leaves are the most actively photosynthesizing part of the plant and produce the majority of the substances stored in the grain (Benedetti & Rossini, 1993). Consequently, events (such as drought, rain, or fertilization) that occur during this time can greatly affect final grain production. For example, the addition of N fertilizer just prior to grain-fill has been shown to increase grain protein concentration in N deficient plants (Westcott et al., 1997). In Montana, N deficiency is frequently observed in years when growing season moisture is favorable for maximum plant growth resulting in depletion of available N. However, if crop growth is impeded due to moisture stress, N deficiency is less likely to occur and a mid-season application of N will not be cost effective. Close examination of VI profiles during this time would be important for understanding crop performance and estimating final grain yields and quality.

A few researchers have looked at integration of vegetation indices during critical periods of plant growth. Doraiswamy & Hodges (1991) found the area under the NDVI profile between silking and maturity was correlated with regional corn yields reported by the National Agricultural Statistics Service (NASS) (r^2 of 0.72). An integrated AVHRR-NDVI used by Benedetti & Rossini (1993) from the third week in April through mid June (flowering to physiologic maturity) was correlated with wheat yield in Italy ($r^2 = 0.515$). Use of comparative VI profile observations in early-season combined with examination of profiles during critical periods might be useful in the formation of crop yield estimates.

If VI growth profiles are used throughout a growing season for crop yield assessment, the only operational satellite/sensor-system currently capable of providing data with adequate frequency is the National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) (Rasmussen, 1992). Satellite data from the AVHRR has been used in numerous environmental modeling and global change studies because of its extensive global coverage, high temporal frequency (twice daily coverage), and low cost (Box et al., 1989; Soriano & Paruelo, 1992; Gutman et al., 1994; Wagner, 1998). However, the spatial resolution of AVHRR data is low at 1.1km^2 at nadir for Local Area Coverage (LAC). Low spatial resolution complicates the separation of crops from other land cover types when dimensions of fields are small (Fischer, 1994).

The AVHRR sensors are carried aboard the NOAA 11, 12, 14, and 15 polar orbiting satellites. The sensor collects data in five spectral bands ranging from visible red to thermal. One product produced from the data is the Normalized Difference Vegetation Index (NDVI), which is a normalized ratio of the red ($R = 0.58 - 0.68\mu\text{m}$) and near

infrared (NIR = 0.725 – 1.1 μ m) spectral wavelengths. NDVI is a unitless value defined as:

$$\text{NDVI} = (\text{NIR} - \text{R}) / (\text{NIR} + \text{R}).$$

NDVI is expressed in a range of -1 to 1, where increasing positive values are related to increasing green vegetation. Negative NDVI values indicate non-vegetated surface features such as water, clouds, or snow that have high red reflectance relative to infrared reflectance. Positive NDVI values near zero indicate rock and bare soil, which have similar reflectance in both the visible and infrared wavelengths (EDC, 1995; Lillesand & Keifer, 1994). A ratio of the vegetation index reduces the effects of changing illumination conditions, surface slope and aspect, and atmospheric perturbations while emphasizing varying vegetation density (Holben et al., 1986). NDVI data supplied from EROS Data Center are linearly re-scaled to byte data range values of 0 to 200, where the original NDVI value - 1.0 is equal to 0, NDVI value 0.0 is equal to 100, and NDVI value 1.0 is equal to 200 (EDC, 1995). The 1.1km pixels are resampled to 1km pixel resolution.

Atmospheric conditions can severely affect the amount of reflected radiation that reaches a satellite sensor. Haze from water vapor, clouds, and cloud shadows can diminish or scatter reflected radiation and introduce 'static' into NDVI data. To compensate, NDVI data are compiled biweekly and an algorithm is applied that retains the maximum NDVI value for each image pixel from that two-week period. Maximum Value Composite (MVC) NDVI imagery is used to create a relatively cloud-free data set by choosing NDVI pixels from days when radiance interference is lowest and sun angle is highest with the assumption that the selected pixel is most representative of actual ground reflectance (EDC, 1995; Holben, 1986).

Composite, multi-temporal imagery requires registration to a common map projection to ensure that, from image to image, each 1km pixel represents the same ground location (EDC, 1995). The Lambert Azimuthal Equal Area Projection is used for the AVHRR-NDVI data set distributed by the EROS Data Center. Images are registered to a hydrography base map developed from a U. S. Geological Survey 1:2,000,000-scale digital line graph (DLG). The base map is registered by using 250 ground control points with a root mean square error (RMSE) of less than one pixel, and registration is performed through image-to-image correlation (EDC, 1995).

Use of AVHRR-NDVI data in agricultural studies has been extensive. Methods for deriving phenological metrics have been devised that estimate the rate of green-up and senescence, time of onset of growing season, duration of growing season, and maximum NDVI in relation to agricultural monitoring (Reed et al., 1994; Idso et al., 1980; Badhwar, 1980). Techniques for discerning different agricultural crops within AVHRR's coarse pixels have been proposed (Fischer, 1994; Benedetti & Rossini, 1993). Multi-temporal measurements made from AVHRR-NDVI have been used to estimate and predict crop yields over large and small regions (Benedetti & Rossini, 1993; Doraiswamy and Cook, 1995; Quarmby et al., 1993; Rasmussen, 1992). While AVHRR-NDVI has inherent limitations, it is an inexpensive and readily available source of remotely sensed data that has provided much of the basis for monitoring land resources with spectral radiance measurements.

CHAPTER 2

REGIONAL AND COUNTY SCALE WHEAT YIELD ESTIMATION USING MULTI-TEMPORAL AVHRR-NDVI SATELLITE IMAGERY

Introduction

Techniques for quickly monitoring and estimating agricultural production during the growing season would be useful for land managers whose decisions depend on timely and accurate information (Rudorff & Batista, 1991). Satellite remote sensing is a promising tool that provides periodic coverage of land resources and has been highly correlated with many agronomic variables (Quarmby et al., 1993). These data not only provide information about vegetation condition over extensive regions during the growing season, but also provide volumes of historic data that could be used to compare to present conditions. Monitoring crop performance and making early season yield estimates could provide farm managers and consultants with a new method for assessing and comparing fertilizer demand, drought stress, or productivity among regions.

Agencies interested in world regional grain production have used satellite remote sensing to increase the accuracy of crop production predictions (Wagner, 1998). In particular, the AVHRR data have been used extensively due to their large ground coverage and timely acquisition. In some parts of the world, it is often the only data available. NDVI has gained most of the attention because it is correlated with agronomic variables such as live biomass, leaf area index, and stand density when normalized for the

effects of some atmospheric and ground conditions not associated with vegetation condition (Tucker, 1979; Weigand et al., 1991; Thoma, 1998). (For more detail on NDVI, see Ch. 1.) Growth profiles produced from NDVI data represent plant growth responses for the season and should provide an indirect estimate of final grain yield (Malingreau, 1989). Many studies have examined AVHRR-NDVI imagery as a tool for estimating final yields at the county or regional level (Rasmussen, 1992; Benedetti & Rossini, 1993; Groten, 1993; Quarmy et al., 1993; Doraiswamy & Cook, 1995). In most of these studies, a single NDVI value or an integrated NDVI over a period during the growing season has been correlated with final grain yield over a single season or a few years. Few researchers have examined long-term relationships of the NDVI seasonal growth profile and grain yield.

The objectives of this study were to (1) determine if region and county level AVHRR-NDVI growth profiles were related to the wheat grain yields reported by state agencies, (2) identify times in the growing season when useful yield estimations could be made using AVHRR-NDVI yearly growth profiles.

Methods and Materials

Montana physiography and climate

Montana is a region of extreme physiographic and climatic diversity. It is part of the Great Plains and the Middle and the Northern Rocky Mountain physiographic provinces, which have been affected by glacial and volcanic activity (Montagne et al., 1982). It is characterized by smooth to extremely dissected plains in the eastern part of

the state, while high mountains and inter-mountain valleys dominate western regions. Montana's climate is primarily influenced by cold, dry continental air masses originating in the Arctic, and cool, wet air masses coming from the Pacific Ocean in the west (USDA Soil Conservation Service, 1982). The plains east of the Continental Divide experience extreme seasonal temperature fluctuations while the mountains and valleys west of the continental divide have a more mild, maritime climate. Embedded in these areas are pockets of microclimates associated with local topography and aspect (Montagne et al., 1982).

Growing season duration across the state ranges from less than 32 days to more than 135 days depending on elevation, aspect, and humidity (USDA Soil Conservation Service, 1982; MAPS, 1990). Precipitation increases towards the west and with increasing elevation. In most of the mountains and western valleys precipitation averages from 300 to more than 1000mm annually while in the semiarid plains east of the Divide, annual precipitation ranges from 300 to 400mm. Soils west of the Divide are diverse due to variable topography. Mollisols are primarily associated with the valley regions where there is crop production, whereas Alfisols and Inceptisols dominate in the forested hills and in some valleys in the Northwest. Soils of the semiarid plains region are primarily Mollisols with ustic moisture regimes, high available water capacity, and high suitability for dryland grains (USDA Soil Conservation Service, 1982). However, the semiarid plains are prone to periods of drought, which cause severe economic impact (Hershenhorn, 1992). These great physiographical diversities and climatic fluctuations play a large part in vegetation density, composition, and ultimately the reflective response that is received by satellite sensors.

Montana non-irrigated agricultural lands

The majority of wheat in Montana is grown under dryland conditions (90+%) (Montana Agricultural Statistics Service, 1997). Hence, only dryland acres within Montana were selected for this study. A 1973 land use polygon coverage of Montana delineating dryland agricultural lands was extracted from the Montana Agricultural Potentials System (MAPS, 1990). The coverage was imported into Arc/Info™ GIS software and converted to an Arc/Info™ GRID format. The grid cells produced from the dryland agriculture coverage were approximately 20.7km² in size. The coverage was converted to a Lambert Azimuthal Equal Area (LAEZA) projection to correspond to the projection of the NDVI imagery. The dryland agriculture grid was used as a mask to extract corresponding AVHRR-NDVI grid cells from each NDVI biweekly image so that only NDVI data from predominantly dryland agricultural lands would be included in the study (Figure 1). This procedure created a subset of only dryland acres from each NDVI biweekly period. Furthermore, only regions and counties with at least 100 pixels (100 km²) of dryland agriculture were included in the study to avoid excess “pixel-mixing” in small areas and image geometric correction. Pixel-mixing is defined as having an NDVI response from other types of vegetation or objects that are contained in the area of interest. Regions were delineated as the Montana Agricultural Statistics Districts used by the Montana Agricultural Statistics Service (Montana Agricultural Statistics Service, 1997). Regions and their associated counties are presented in Figure 2. Counties and number of acres within each county included in this study are presented in Table 1.

Figure 1. MAPS Atlas coverage of areas where dryland agriculture dominates in Montana.

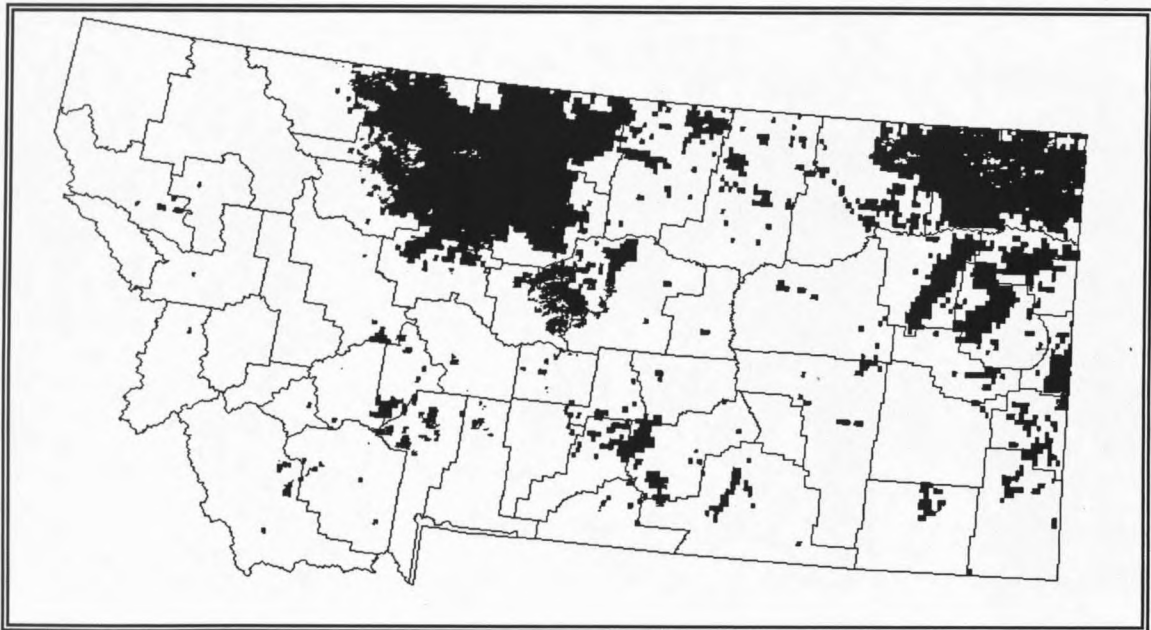


Figure 2. Regions and associated counties included in study, as defined by the Montana Agricultural Statistics Service.

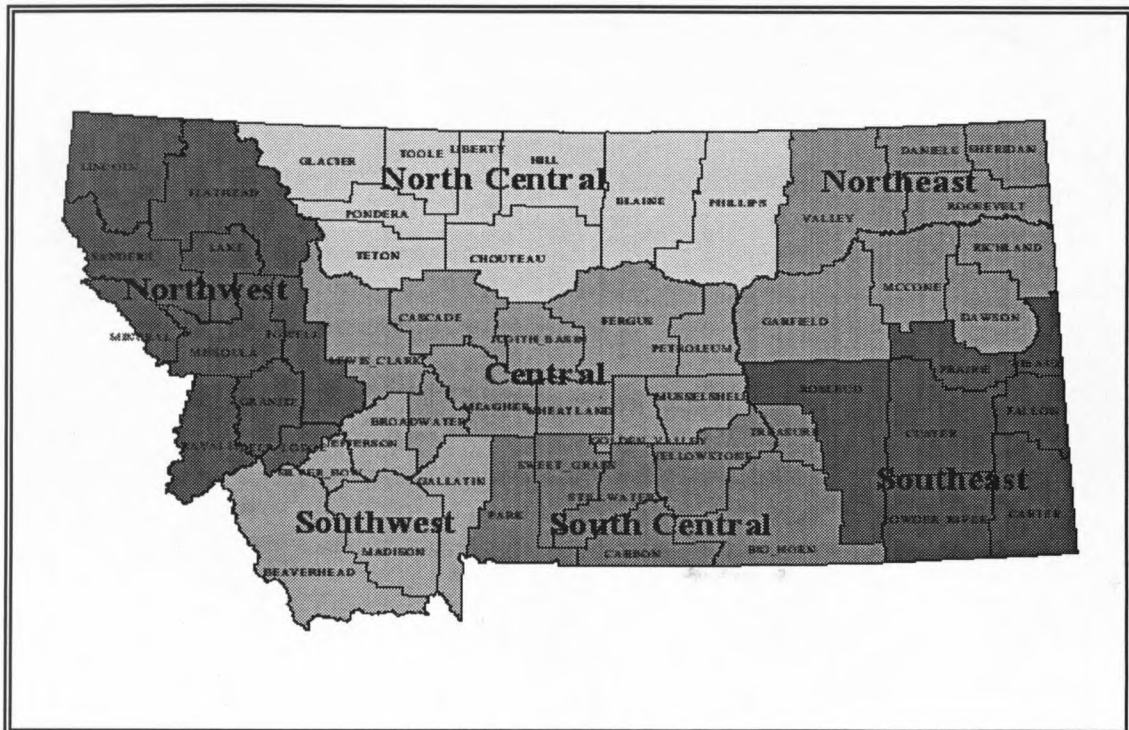


Table 1. Counties, pixels, and area in dryland agriculture included in the study.

Counties	Pixels	Area included in study (hectares)	% of county in dryland agriculture
Beaverhead	244	24,400	1.70%
Big Horn	626	62,600	4.78%
Blaine	2014	201,400	18.22%
Broadwater	324	32,400	10.05%
Carter	548	54,800	6.39%
Cascade	2033	203,300	29.37%
Choteau	7424	742,400	73.12%
Daniels	3558	355,800	95.20%
Dawson	2264	226,400	36.73%
Fallon	1161	116,100	27.45%
Fergus	2556	255,600	23.20%
Gallatin	812	81,200	11.74%
Garfield	324	32,400	2.60%
Glacier	2109	210,900	27.09%
Golden Valley	223	22,300	7.31%
Hill	5887	588,700	77.68%
Judith Basin	1759	175,900	36.13%
Lewis and Clark	222	21,200	2.47%
Liberty	3316	331,600	88.97%
Madison	167	16,700	1.82%
McCone	2429	242,900	35.36%
Musselshell	151	15,100	3.09%
Park	255	25,500	3.56%
Phillips	1272	127,200	9.29%
Pondera	3083	308,300	71.97%
Powder River	462	46,200	5.43%
Prairie	455	45,500	10.17%
Richland	223	22,300	4.17%
Roosevelt	4985	498,500	80.23%
Rosebud	365	36,500	2.80%
Sanders	124	12,400	1.70%
Sheridan	4009	400,900	89.99%
Stillwater	720	72,000	15.44%
Teton	2260	226,000	38.04%
Toole	4317	431,700	85.48%
Valley	3354	335,400	26.10%
Wheatland	245	24,500	6.59%
Wibaux	1057	105,700	45.80%
Yellowstone	1037	103,700	15.81%

Montana county yield data

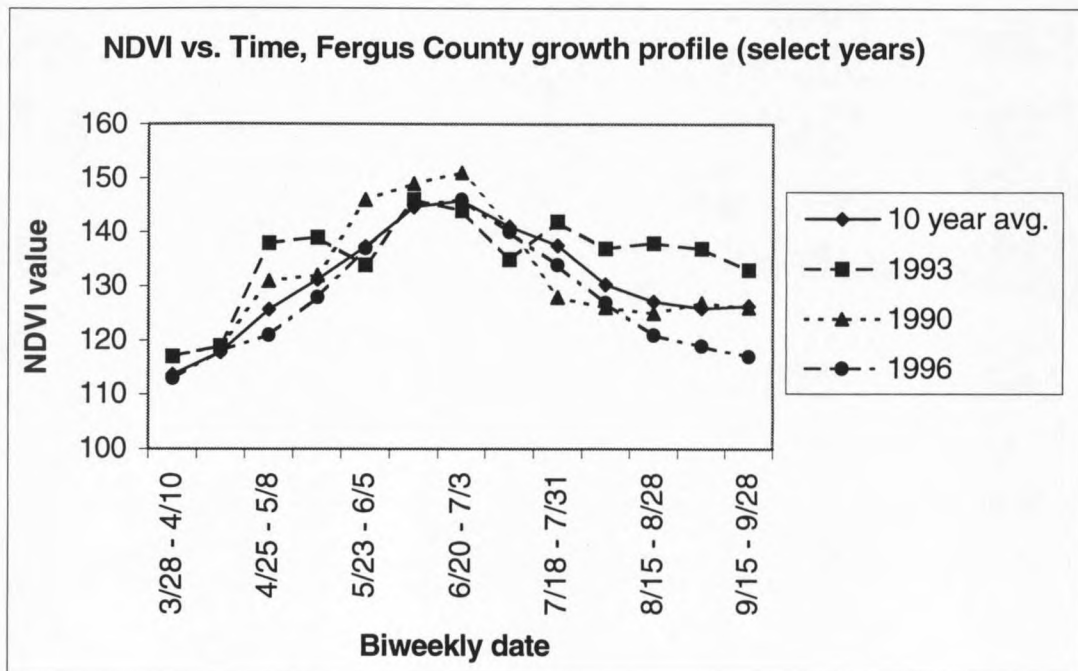
Dryland wheat yield data used in this project were collected from the Montana Agricultural Statistics Service in Helena, Montana for the years 1989 through 1997 (Montana Agricultural Statistics Service, 1997). Dryland wheat yields include spring, winter, durum, and other wheat grown in Montana. Inclusion of all types of wheat was acceptable in this study because the delineation of dryland agricultural lands from MAPS Atlas was also inclusive of these types of wheat. The Montana Agricultural Statistics Service estimates regional and county yields using probability sample surveys of all known producers several times each year. County estimates are supplemented with additional questionnaires to improve coverage in all counties (Lund, pers. comm., 1998). Wheat yield estimates are made in March for intentions to plant, in June for what was actually planted, and in September for what was actually harvested. Counties are not published when fewer than 200 hectares are planted or if one report has 60% or more of the acreage (Lund, pers. comm., 1998). This study included 6 regions (Northwest region excluded) and 39 counties in Montana. Regional analyses were based on the production districts defined by the Montana Agricultural Statistics Service.

AVHRR-NDVI processing

AVHRR-NDVI satellite data were supplied from Earth Resources and Observation System (EROS) data center in Sioux Falls, SD. The NDVI data used are biweekly maximum value composites for April through September of the years 1989 – 1997 (see Chap. 1 for more detail). AVHRR-NDVI data were imported and converted into Arc/Info™ GRID format in a LAZEA projection.

An Arc/Info™ polygon coverage of Montana counties was obtained from the Natural Resources Information System (NRIS) GIS clearinghouse and converted to a LAZEA projection. Each region and county polygon was extracted from the original coverage. Individual region and county polygons were used to extract NDVI data from each of the dryland biweekly periods by selecting only the NDVI cells contained within the region or county polygon. An average NDVI value of the selected cells was used as the representative NDVI value across the region or county for that biweekly period. Plotting these NDVI values against biweekly periods for each year produced region and county annual growth profiles (Figure 3).

Figure 3. NDVI seasonal growth profiles for a Montana county.



Dips within growth profiles can represent contamination from clouds, haze, or other types of noise that were not removed during compositing process. They can also represent real changes in vegetation occurring on the ground, such as drought stress, insect infestation, or other factors that will reduce chlorophyll or mesophyll content in plants.

NDVI growth profile analysis

Several parameters of the NDVI growth profiles were examined to find the best model for yield estimation. Since NDVI growth profiles reflect plant productivity in a season, we hypothesized that the area under the growth profile was related to final wheat grain yield. The growth profiles of NDVI values were integrated from approximately April 1 through September 15 (12 biweekly periods) using a trapezoidal approximation.

Integrate 12 periods =

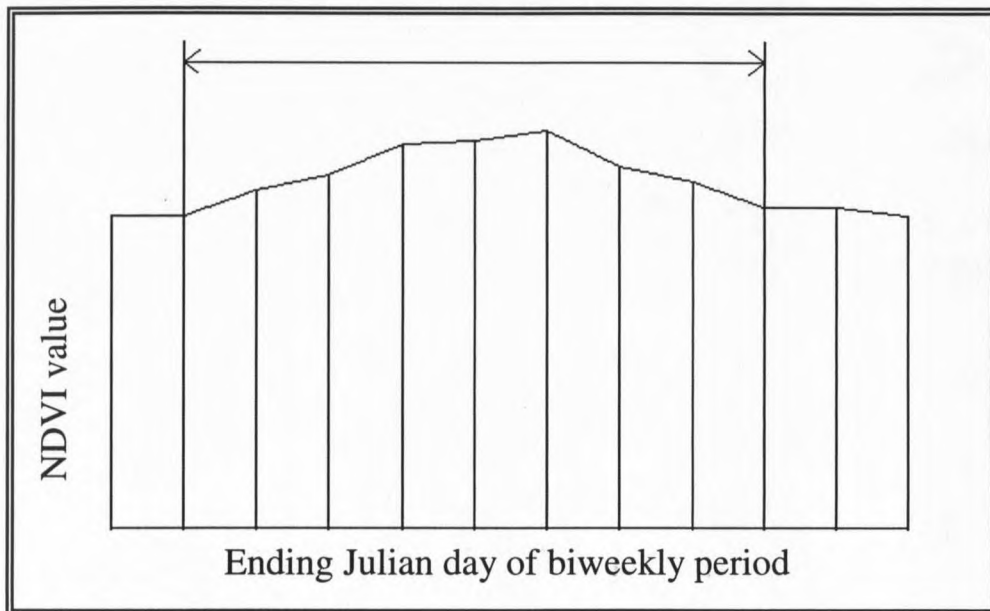
$$I \cong I_h = \frac{h}{2}(y_0 + 2y_1 + 2y_2 \dots + 2y_{n-2} + 2y_{n-1} + y_n)$$

Where h is the distance between intervals, y is the height of the rectangle, and n is the number of intervals. Integration of the area is then approximated from the area of a trapezoid and the summation process (Dorn & McCracken, 1972).

For Montana's climate, April through mid September usually includes the emergence, maturation, and senescence phases for wheat. However, the growing season fluctuates across the state and over years and late emergence or early senescence might be reflected in NDVI growth profiles. By integrating NDVI only over the dates that reflect actual growing season, wheat production might be better represented for that region or

county (Groten, 1993; Quarmby et al., 1993). Therefore, the apparent growing season (Integrate AGS) was determined using seasonal fluctuations in the incline and decline of NDVI values across the seasonal growth profile. The growing season was defined using the approach of Groten (1993), beginning with two consecutive biweekly periods with positive NDVI increments. The defined minimum increase was an increment of ≥ 1 digital value followed with an increment of ≥ 3 digital values. This could be seen as the “steepening” of the growth curve when photosynthetic activity is first increasing. Conversely, end of season was estimated as a decrease of ≤ 2 followed by a decrease of ≤ 1 digital value. This could often be seen as the “flattening of the curve” when photosynthetic activity has decreased substantially (Figure 4).

Figure 4. An example of integrating the apparent growing season (Integrate AGS). Integration of the area under a curve is calculated using a trapezoidal approximation



Summations of NDVI growth profiles were examined as an alternative method for assessing seasonal growth profiles. A summation of the entire NDVI growth profile from April 1 – September 15 (Sum 12 periods) was calculated to approximate integration of NDVI across the growing season.

Since we were particularly interested in early season estimation of wheat yields, we examined NDVI growth profiles in early season and as they progressed throughout the season. We hypothesized that high NDVI values early in the season would be a good predictor of a high yield year, in the absence of a disastrous climatic event, and that this would carry through as the season progressed. Monthly NDVI parameters were examined successively throughout the season. Summations through each consecutive month (Sum April through Sum Aug.) were included in yield – NDVI parameter assessments to detect which NDVI growth profiles reflect final wheat yields. Another approach investigated the relationship of wheat yield and end of the month NDVI values. End of the month NDVI values were found to be highly correlated to wheat yields in Northeast Montana (Hershenhorn, 1992) and suggested that early season end of the month NDVI could be used to make yield estimations two to three months prior to harvest. End of the month NDVI values (End April through End Aug.) were included in the analysis with wheat yields for each region and county.

Integrating NDVI over select dates within the growing season has been shown to be related to wheat yields in some studies (Doraiswamy & Cook, 1995; Quarmby et al., 1993; Benedetti & Rossini, 1995). This study examined three summation periods during mid- to late-season. The periods chosen were based on those proposed by Doraiswamy & Cook (1995) that suggested integration around the time of maximum NDVI encompassed

a “critical period” in grain production. The summation periods include four NDVI biweekly periods (total of 8 weeks) that approximately correspond to flowering through maturity (sum 6/4 – 8/2; sum 6/22 – 8/16; sum 7/6 – 8/30).

Because there are differences among regions and counties due to climate, topography, farming practices, and other factors, regions and counties were included as indicator variables in multiple linear regression using SAS statistical software (SAS Institute Inc., Cary, NC). Indicator variables reduce the error effects caused by location differences and show only the overall relationship of NDVI variables and yield, regardless of regional or county effects. In addition, interactions between NDVI parameters and regions or counties were included as a test for heterogeneity of slope. The model equation for a NDVI parameter with indicator variables and interaction was:

$$\text{Wheat yield} = \text{NDVI parameter} + \text{region} + \text{NDVI parameter} * \text{region}$$

where NDVI parameters have ratio-interval values and regions and counties have classificatory values.

The inclusion of the region, county, or interaction was determined by significance (p-value value) in the full model, with a minimum p-value of 0.05 for inclusion of the parameter in the final model. For models that found no significant difference including interactions and regions or counties, a simple regression analysis was performed between wheat yield and the NDVI parameter.

Results

Wheat yield – Regional NDVI parameters relationships

Results of wheat yield – NDVI parameter relationships are presented in Table 2. The table includes the adjusted R^2 , p-value, and mean square error (MSE) of each wheat yield – NDVI parameter relationship. Regions and interaction p-values are presented to show when they were included in a model. Degrees of freedom associated with each model are also presented.

Table 2. Wheat yield – NDVI parameter relationships for regional study sites. Corrected total SS is 3941.49.

NDVI Parameter	Adjusted R^2	p-value, full model	MSE, overall	p-value, interaction	p-value, region	DF
Integrate 12 periods	0.753	0.0001	18.36	0.034*	0.0001*	42
Integrate AGS	0.738	0.0001	19.51	0.038*	0.0001*	42
Sum 12 periods	0.745	0.0001	18.93	0.042*	0.0001*	42
Sum through April	0.440	0.620	41.67	0.973	0.0001*	47
Sum through May	0.454	0.234	40.63	0.241	0.0001*	47
Sum through June	0.454	0.234	40.63	0.743	0.0001*	47
Sum through July	0.463	0.135	39.92	0.644	0.0001*	47
Sum through Aug.	0.748	0.0001	18.73	0.0001*	0.0001*	42
End of April	0.444	0.442	41.36	0.899	0.0001*	47
End of May	0.563	0.0006	32.46	0.631	0.0001*	47
End of June	0.471	0.087	39.34	0.602	0.0001*	47
End of July	0.555	0.001	33.06	0.452	0.0001*	47
End of Aug.	0.697	0.0001	22.53	0.046*	0.0001*	42
Sum (6/4 – 8/2)	0.587	0.0001	30.71	0.175	0.0001*	47
Sum (6/22 – 8/16)	0.644	0.0001	26.49	0.160	0.0001*	47
Sum (7/6 – 8/30)	0.686	0.0001	23.35	0.055	0.0001*	47

* Region and/or interaction term included in final model

Our results show that integrated NDVI (Integrate 12 periods) and apparent growing season (Integrate AGS) relationships with yield were both significant predictors of yield (p-value = 0.0001), though the R^2 was slightly lower for the apparent growing period (Integrate 12 periods adj. $R^2 = 0.753$, Integrate AGS adj. $R^2 = 0.738$). Summation of the NDVI over the growing season (Sum 12 periods) showed a similar relationship with yield (adj. $R^2 = 0.745$) that was also significant (p-value = 0.0001). The regions and interactions were significant (p-value = 0.0001) and included in the final model for these parameters. Inclusion of the region and interaction terms reduced the degrees of freedom but resulted in overall higher adj. R^2 values, which account for fewer degrees of freedom.

Both monthly parameters (sum through each month and end of the month) show trends of increasing association with final yield as the season progresses (Table 2). However, end of the month May NDVI showed higher correlation (adj. $R^2 = 0.563$) and was more significant (p-value = 0.0006) than end of the month June (adj. $R^2 = 0.471$, p-value = 0.087) or July (adj. $R^2 = 0.555$, p-value = 0.001) NDVI. End of the month NDVI values show higher correlation and are more significant than summation through those months with the exception of August (end Aug. adj. $R^2 = 0.697$, p-value = 0.0001; sum Aug. adj. $R^2 = 0.748$, p-value = 0.0001). Regions for these parameters were all significantly different (p-value = 0.0001) and included in the final model. The significance of the interactions for monthly parameters increased as the season progressed. Interactions were not significant in early season (April) but increased throughout the season. Only August parameters showed significant interactions of p-value < 0.05 (Table 2). For these parameters, interaction terms were included in the final model.

Summations of critical periods, (sum 6/4 – 8/2; sum 6/22 – 8/16; sum 7/6 – 8/30), showed increasing correlation as summation periods approached July – August with strongest correlation for the summation period of 7/6 – 8/30 (adj. $R^2 = 0.686$, p-value = 0.0001). Regions were all significantly different (p-value = 0.0001) and included in the model while interactions were not significant.

Analysis of best model scatterplots

Scatterplots of the three best fitting models are presented in figures 5 through 7. Figure 5 represents the fit of the predicted slopes to the reported yield values based on integrated NDVI over the entire growing season (Integrate 12 periods). This relationship (adj. R^2 value = 0.753, p-value = 0.0001) is the strongest of all NDVI parameters examined. Regional differences and interactions were both significant and included in the final model (p-values = 0.0001 and 0.034, respectively). Interaction is evident by the differences in slopes among regions. For most regions, predicted yield slopes are positive and show strong relationship with reported yield values.

Figure 6 represents the fit of the predicted slope to the reported yield values based on summation through August (Sum through Aug.). The relationship is the second strongest relationship between wheat yield and an NDVI parameter (adj. $R^2 = 0.748$, p-value = 0.0001). Regional and interaction terms were both significant and included in this model (p-value = 0.0001 and 0.0001, respectively). For the summation through August NDVI parameter, most regions show a strong positive slope and good relationship between predicted and reported yield values. The Southeast region shows a much flatter slope and little variation in reported yields.

Figure 5. Regional wheat yield – Integrate 12 periods relationships. Model includes interaction terms. $\text{Adj. } R^2 = 0.753$, $p\text{-value} = 0.0001$. Regression equations are: Central yield = $(-216.32) + (0.021 + -0.02)$ (Integrate 12 periods), North Central yield = $(-220.01) + (0.013)$ (Integrate 12 periods), Northeast yield = $(-180.6) + (0.035)$ (Integrate 12 periods), South Central yield = $(-118.67) + (0.007)$ (Integrate 12 periods), Southeast yield = $(-49.66) + (0.004)$ (Integrate 12 periods), Southwest yield = $(-388.84) + (0.021)$ (Integrate 12 periods).

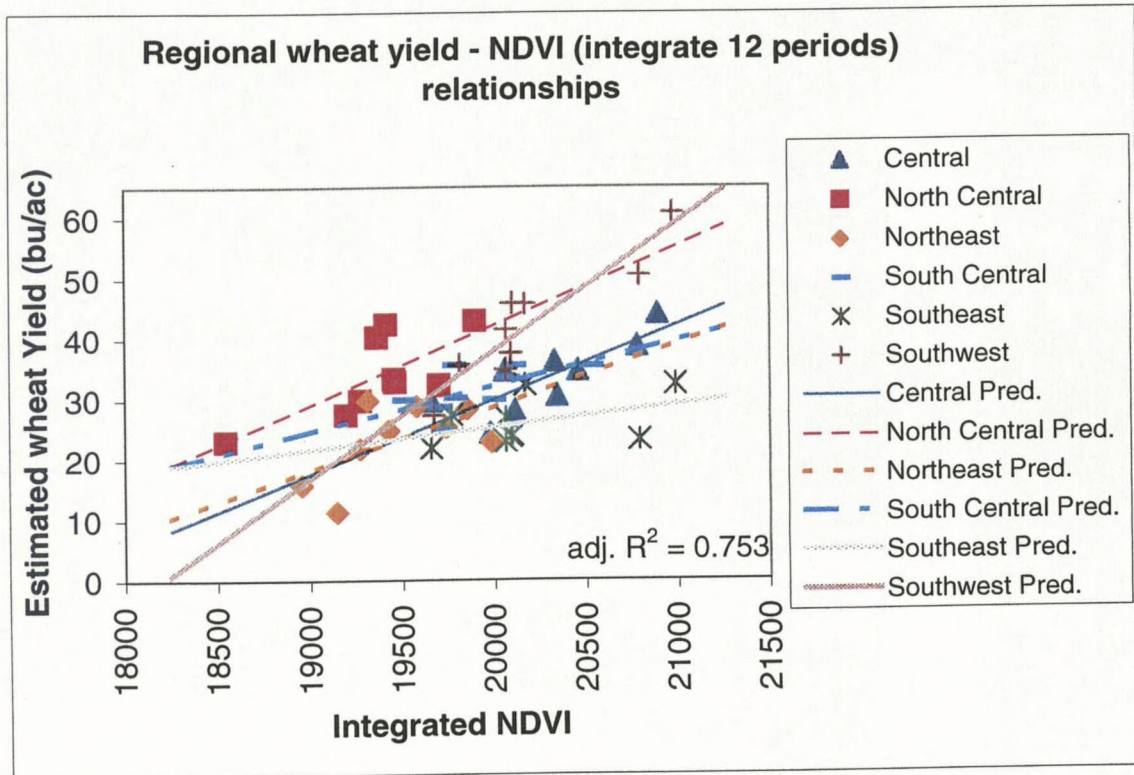


Figure 7 represents the fit of the predicted slopes to actual wheat yield values based on a NDVI critical period (Sum 7/6 – 8/30). This was the strongest relationship of all the critical periods ($\text{adj. } R^2 = 0.686$, $p\text{-value} = 0.0001$). Regions proved to be significantly different but interaction was not significant in this model ($p\text{-value} = 0.0001$ and 0.055 , respectively), as can be seen by the parallel slopes across regions. The summation values for this critical period predicted less variable yields across regions than found in reported yields. However, reported yields tend to fall within their predicted range, with Southwest

reported and predicted yields falling within the highest yield range and the Northeast and Southeast reported and predicted yields falling within the lowest yield range.

Figure 6. Regional wheat yield – NDVI (sum through Aug.) relationships. Model includes interaction terms. Adj. $R^2 = 0.748$, p-value = 0.0001. Regression equations are: Central yield = $(-224.17) + (0.179)$ (sum through Aug.), North Central yield = $(-209.98) + (0.178)$ (sum through Aug.), Northeast yield = $(-178.20) + (0.146)$ (sum through Aug.), South Central yield = $(-118.95) + (0.107)$ (sum through Aug.), Southeast yield = $(17.05) + (0.007)$ (sum through Aug.), Southwest yield = $(-408.10) + (0.314)$ (sum through Aug.).

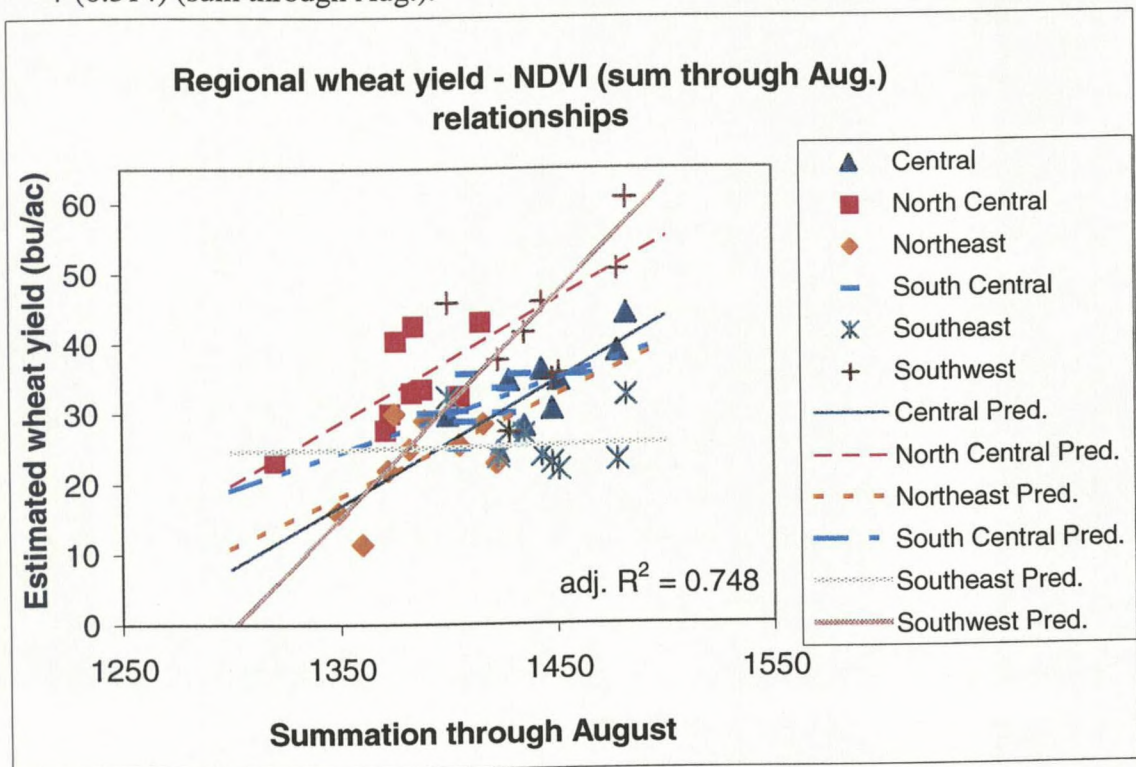
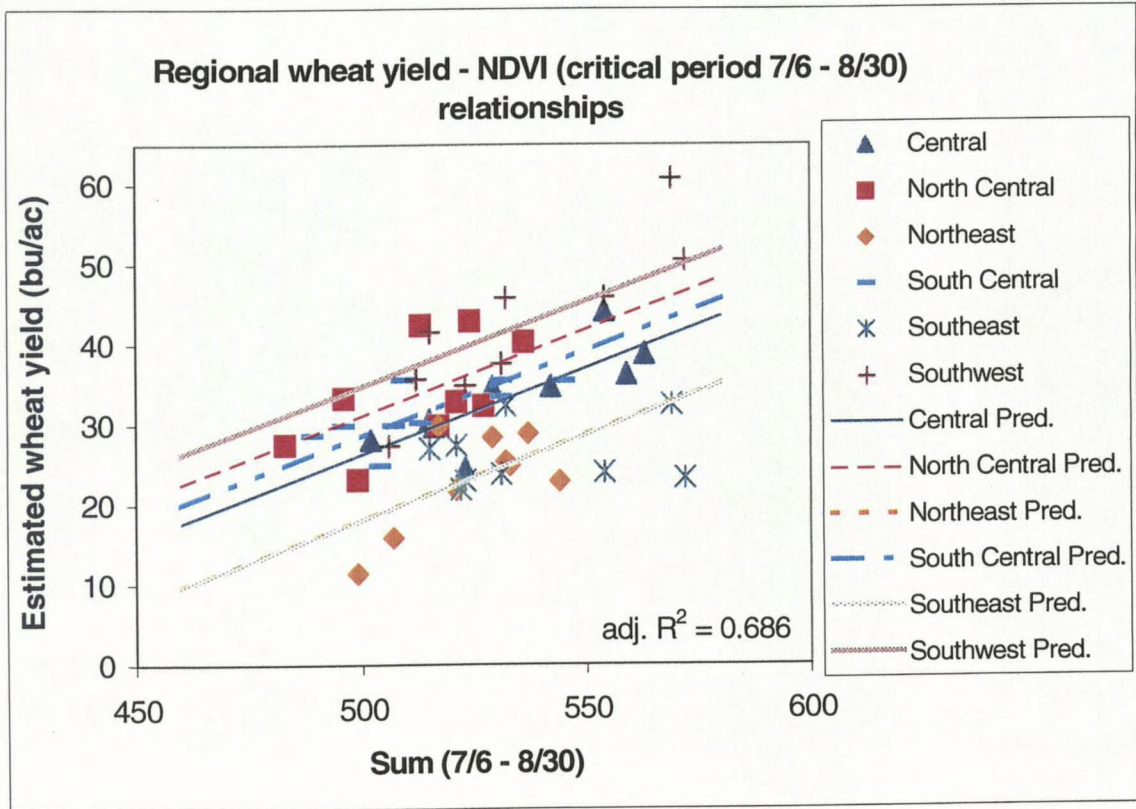


Figure 7. Regional wheat yield – NDVI critical period (7/6 – 8/30) relationships. Adj. $R^2 = 0.686$, p-value = 0.0001. Regression equations are: Central yield = $(-80.61) + (0.214) (\text{sum } 7/6 - 8/30)$, North Central yield = $(-75.7) + (0.214) (\text{sum } 7/6 - 8/30)$, Northeast yield = $(-88.75) + (0.214) (\text{sum } 7/6 - 8/30)$, South Central yield = $(-78.3) + (0.214) (\text{sum } 7/6 - 8/30)$, Southeast yield = $(-88.85) + (0.214) (\text{sum } 7/6 - 8/30)$, Southwest yield = $(-72.18) + (0.214) (\text{sum } 7/6 - 8/30)$.



County Yield – NDVI relationships

Results of wheat yield – NDVI parameter relationships at the county scale are presented in Table 2. The table includes the adjusted R^2 , p-value, and mean square error (MSE) of each wheat yield – NDVI parameter relationship. Counties and interaction p-values are presented to show when they were included in a model. Degrees of freedom associated with each model are also presented.

Table 3. Wheat yield – NDVI parameter relationships for county study sites. Corrected total SS is 20868.35.

NDVI Parameter	Adjusted R ²	p-value, full model	MSE, overall	p-value, interaction	p-value, county	DF
Integrate 12 periods	0.566	0.0001	30.11	0.123	0.0001*	262
Integrate AGS	0.536	0.0001	32.04	0.004*	0.0001*	224
Sum 12 periods	0.548	0.0001	31.32	0.229	0.0001*	262
Sum through April	0.254	0.282	51.69	0.996	0.0001*	262
Sum through May	0.316	0.0001	47.41	0.468	0.0001*	262
Sum through June	0.319	0.0001	47.19	0.867	0.0001*	262
Sum through July	0.300	0.0001	48.54	0.995	0.0001*	262
Sum through Aug.	0.546	0.0001	31.45	0.250	0.0001*	262
End of April	0.256	0.411	51.56	0.992	0.0001*	262
End of May	0.386	0.0001	42.57	0.693	0.0001*	262
End of June	0.288	0.0005	49.34	0.473	0.0001*	262
End of July	0.400	0.0001	41.62	0.182	0.0001*	262
End of Aug.	0.436	0.0001	39.11	0.605	0.0001*	262
Sum (6/4 – 8/2)	0.414	0.0001	40.60	0.533	0.0001*	262
Sum (6/22 – 8/16)	0.485	0.0001	35.67	0.680	0.0001*	262
Sum (7/6 – 8/30)	0.511	0.0001	33.90	0.660	0.0001*	262

* Regional and/or interaction term included in final model

Integration of NDVI over the entire growing season (Integrate 12 periods) showed a slightly stronger relationship with wheat yields than using an integration of the apparent growing season (Integrate AGS) with adjusted R² values of 0.566 and 0.536, respectively. Summation of NDVI over the growing season (Sum 12 periods) resulted in a similar, but slightly lower relationship (adj. R² = 0.548) than using integration over the growing season (Integrate 12 periods). All of these relationships were highly significant (p-value = 0.0001). Counties proved to be significantly different and were included in the final model, while only the apparent growing season (Integrate AGS) showed a significant interaction (p-value = 0.004).

As with the regional monthly data, county monthly data showed increasing correlation as NDVI values were cumulated toward the end of the season (Table 3). Summation through August showed the strongest relationship with wheat yields of all the monthly parameters (Sum through Aug. adj. $R^2 = 0.546$, $p\text{-value} = 0.0001$). While most of the monthly adjusted R^2 values were low, significance was high for all months ($p\text{-value} = 0.0001$) with the exception of both April parameters. Interactions of the monthly parameters were not significant and did not show the trend of increasing significance through the growing season that was seen in the regional monthly data. Counties, however, were all significantly different ($p\text{-value} = 0.0001$) and included in the final model.

Critical periods showed a moderate and increasing relationship with yield as dates approached July and August (Table 3). The strongest correlation occurred with the summation date of 7/6 – 8/30 (adj. $R^2 = 0.511$, $p\text{-value} = 0.0001$). For the county NDVI parameters, integrated NDVI for the entire growing season (Integrate – 12 periods), sum of the entire growing season (Sum – 12 periods), summation through late season (Sum through Aug.), and the critical period (Sum 7/6 – 8/30) showed the strongest relationships and most potential for wheat yield estimation (Table 3). However, these relationships were much lower than those at the regional level were.

Discussion

Satellite remote sensing of regional crop dynamics has the potential of bringing timely data about crop performance to those who need information to manage the land. The interaction of incident energy (sunlight) on a crop canopy as detected by a satellite sensor is shown to be highly related to crop health and vigor in many studies throughout the world and in Montana (Tucker et al, 1980; Malingreau, 1989; Weigand & Richardson, 1990; Hershenthorn, 1992; Reed et al., 1994; Thoma, 1998). The basis for its use is the generalization of "the better the growing conditions, the healthier and more vigorous the plants will be" (Wiegand et al., 1984). This study was designed to evaluate the potential of AVHRR-NDVI seasonal growth profiles for real-time crop monitoring and yield estimation at the regional and county level. Our approach was to investigate times during the growing season when NDVI growth profile parameters were related to wheat yield for regions and counties in Montana.

Regional yield – NDVI relationships

As predicted in our hypothesis, AVHRR-NDVI seasonal growth profiles from April through September showed strong relationships with final grain yields (adj. $R^2 = 0.753$ for Integrate 12 periods, adj. $R^2 = 0.38$ for Integrate AGS, and adj. $R^2 = 0.745$ for Sum 12 periods). This is an indication that AVHRR NDVI growth profiles are representing seasonal biomass production and those seasons with higher biomass result in higher wheat yields that are represented by larger areas under the seasonal curve (integrated NDVI) or larger summations over the season. Seasonal biomass – NDVI

relationships have been observed in many studies (Hatfield, 1983; Benedetti & Rossini, 1993; Groten, 1993; Quarmby et al., 1993; Thoma, 1997). In 1997, Thoma found that as biomass increased throughout the season in Montana rangelands, NDVI increased accordingly resulting in a strong positive relationship between live green vegetation and NDVI response ($r^2 = 0.715$) (Thoma, 1998). Similarly, Benedetti & Rossini (1993) found NDVI growth profiles to be representative of seasonal wheat phenology and photosynthetic efficiency and, thus, were able to estimate wheat yields with a summation of NDVI (R^2 of 0.515).

For this study, "high biomass years" typically are years of higher grain production, most likely because of favorable growing conditions. In 1993 and 1995, precipitation was abundant throughout the growing season for most parts of the state (Montana Agricultural Statistics Service, 1997). These were also years of higher than average yield across much of the state. This effect can be seen in the predicted slopes of the wheat yield – integrated NDVI scatterplot for some regions (Figure 5). Most regions show a moderate to strong positive relationship between predicted wheat yield and integrated NDVI (Integrate – 12 periods).

There are times when integrated NDVI values will be high despite a low reported yield or low despite larger reported yields. Both NDVI and yield are influenced by seasonal events such as drought stress, insect infestation, or nutrient deficiency that affect the wheat crop. Examination of the yearly NDVI growth profiles at different times in the season could be important for understanding wheat production in a particular year. Monthly NDVI parameters revealed increasing adj. R^2 values as dates approached August (Table 2), suggesting that as the growing season progressed toward harvest, our ability to

estimate wheat yield increased. These results disagree with Hershenhorn (1992), who found that end of the month May and June NDVI were highly correlated with winter wheat yield in Northeastern Montana ($R = 0.996$ and 0.866 , respectively), and that the relationship decreased toward August and September. However, the study only included four years of NDVI and yield data, which might not be enough to characterize wheat yields over longer periods of time. Our end of May relationship with yield was higher than end of June or July, though not as high as that of Hershenhorn (Table 2). This might indicate that vegetation condition in May, during early vegetative growth, is a critical period that could be monitored for signs of stress to the crop.

The monthly NDVI parameter results disagree with our hypothesis that early season NDVI parameters are a good indicator of final wheat yields. Early season parameters might reflect current crop condition and provide an estimate of wheat yield potential, but subsequent events such as drought, heat stress, insect infestation, or disease that occur later in the season when grain is forming, are unpredictable. Thus early season NDVI parameters make poorer estimates of final wheat yield. In addition, it is not until August that interactions between regions and NDVI become significant. This means that in early season, the wheat yield – NDVI parameter relationships were similar across regions but as the season progressed, regional differences became more pronounced, and the yield – NDVI parameter relationships became different as well. This is consistent with expected responses because in early season, all regions are very green due to spring precipitation and vegetation emergence. As the season progresses, differences among regional crop maturation and senescence rates become more pronounced in accordance with climatic, topographic, and other site-specific characteristics among regions. This

affect can be seen in the shapes, lengths, and amplitudes of regional NDVI growth profiles. Some regions or seasons exhibit short, high growth curves while others have lower and longer growth curves. While these curves might exhibit similar integrated or summed NDVI, the shape, length and amplitude of the NDVI growth curve might reflect much more about crop performance, particularly during the senescence, or grain-filling period. Thus, an increase in the NDVI does not afford the same increase in wheat yield across regions. This suggests a need for region-specific (perhaps site-specific) examination of NDVI growth profiles. Still, early yield estimation is difficult using any technique due to unpredictable events later in the season, and these estimates, though not strong, might be as good as many traditional yield estimate techniques.

Critical periods examined on or around the time of maximum NDVI show significant relationship with wheat yields (Table 2). These summations occur toward the later part of the growing season in July and August, where we have already found good relationship between wheat yield and late season NDVI. Our results are similar to those obtained by Benedetti & Rossini (1993) and Doraiswamy & Cook (1995), who found NDVI summation periods around the time of maximum NDVI (end of vegetation phase) could be used to estimate wheat grain yields ($R^2 = 0.515$ and $R^2 = 0.57$, respectively). These critical periods correspond to the emergence of the flag leaf and the beginning of grain filling in wheat when mid-season fertilizer correction could be applied.

County yield – NDVI relationships

County yield – NDVI parameter relationships showed more variability and had lower adj. R^2 than at the regional level (Tables 2 and 3). The overall trends in yield –

NDVI parameter relationships were, however, very similar, with integrated NDVI (Integrate 12 periods), apparent growing season (Integrate AGS), summation of the growing season (Sum 12 periods), and summation through August NDVI (Sum through Aug.) models having the highest relationships with final wheat grain yield. Counties in all models proved to be significantly different, which would be expected across this diverse region. Interactions between counties and NDVI parameters, however, were not present with the exception of apparent growing season (Integrate AGS).

Summation of critical periods, through each month, and end of the month wheat yield – NDVI relationships again showed increasing correlation as August approached (Table 4). Sum through April and end of April NDVI showed low correlation and significance with final grain yield and appears to be too early to predict yields. As maturity progresses in the crop, correlation increased substantially (end April adj. $R^2 = 0.256$, end Aug adj. $R^2 = 0.436$). Studies have shown that for grain crops, climatic condition during grain filling is a major factor contributing to final grain yields (Idso et al., 1980; Weigand, 1984; Benedetti & Rossini, 1993). Pre-harvest yield estimates of grain crops using vegetation indices are shown to be more accurate when tied to agrometeorological models that consider climatic condition during critical periods during crop growth (Maas, 1988; Rudorff & Batista, 1991; Quarmby et al., 1993). Historic and real-time climatic data could provide a means of deciphering fluctuations in NDVI that are related to crop condition. Crop growth profiles tied with climatic data could be used to monitor for signs of a more productive year in early season (using NDVI seasonal crop growth profile comparisons) while yield estimates could be updated as the season progressed, providing better yield estimates (Maas, 1988; Malingreau, 1989).

Sources of error

Regional and county wheat yields are related to climate, topography, and soils of the region plus the agricultural practices, wheat types and cultivars used by the land manager. While crops integrate and reflect these factors in their canopies (Malingreau, 1989; Weigand & Richardson, 1990), sensing their effects remotely over time, and interpreting those effects are difficult. Mixtures of crop canopies with other terrestrial elements within a 1km AVHRR pixel will confound interpretation further. The coarse resolution of NDVI or the biweekly compositing might not adequately reflect subtle shifts in the crop canopy that affect final wheat yields.

Another source of error could be related to the delineation of dryland agricultural lands. The MAPS Atlas data used to extract dryland agricultural areas in Montana was based on Landsat imagery from the early 1970's. While it might be better than using the entire NDVI average over a region or county, it might not be accurate enough to use with the agricultural statistics provided by state agencies. We are also reminded that the crop yield values reported by the Montana Agricultural Statistics Service are, themselves, estimates of actual regional and county yields and their accuracy is not well documented.

Another source of error could be from differences in AVHRR sensor sensitivity. It is well documented that the AVHRR sensors degrade over time and satellite orbits are altered due to drift (EDC, 1995). Since data in 1992 were collected aboard the NOAA-11 satellite and 1997 data collected aboard NOAA-14, the sensors might need calibration to precisely compare temporal patterns in reflectance data.

Conclusions

This study illustrates several methods for modeling wheat yield at the regional and county level with a time-series of AVHRR-NDVI remotely sensed data. Our results indicate that NDVI crop growth profiles can provide good estimates of regional yield during the later part of the growing season, prior to harvest. Wheat yield estimates made at the end of the growing season might allow state agencies to improve the accuracy of regional and county yield statistics, but are too late for aiding early to mid-season management decisions. Early season estimates, though not strong, could provide crop yield estimation when little other data is available to land managers for yield estimation. Although reasonable wheat yield estimates were obtained for many of the NDVI models, over and under estimations of yield with NDVI parameters need to be investigated. Models developed from the growth profile, such as slope or amplitude, could also be examined to find the best relationship with yield. Since much of the final grain yield depends on water availability and N status, inclusion of climatic data would be useful for modifying yield predictions.

AVHRR biweekly composite imagery provides the best source of frequent and historical data at low cost, making it ideal for satellite imagery users. However, higher resolution imagery would provide more accuracy with respect to specific wheat producing areas. Commercial and government satellites launched starting in 1999-2000 will have higher spatial and spectral resolution, better positional accuracy, improved processing, and be available at the temporal frequency needed to monitor fields site-specifically. The cost of these products might be the determining factor for many agricultural applications.

CHAPTER 3

FARM-SCALE WHEAT YIELD AND PROTEIN CONCENTRATION ESTIMATION USING MULTI-TEMPORAL AVHRR-NDVI SATELLITE IMAGERY

Introduction

Farm managers have always monitored their crops closely throughout the growing season for signs of nutrient deficiency, heat and water stress, insect infestation, weeds, and disease. With large acreages, however, it can be difficult to find the time to monitor all fields and even more difficult to follow change throughout the growing season and from year to year. Satellite remote sensing has the potential to monitor crop condition over extensive areas during the growing season and provide spectral reflection information about crop performance over seasons and years (Reed et al., 1994, Wade et al., 1994; Schepers et al., 1996). An examination of historic crop production could provide clues into future crop trends and give an indication of potential crop yields and qualities during the growing season (Burgan et al., 1996; Wagner, 1998).

So far, applications for agriculture have been limited primarily to regional crop production forecasts and resource surveys with little widespread, long-term use of remote sensing by individual farmers (Jackson et al., 1986; Spry et al., 1996; Senft, 1996; Clarke, 1997). Reasons include unfamiliarity with remote sensing products and lack of useful, inexpensive farm-scale products (UMAC, 1998). Inexpensive remote sensing methods

developed to monitor crop productivity throughout a season and estimate pre-harvest yield at the farm scale could provide information about crop condition early enough in the season so that farm managers could take in-season, corrective measures (Aase, et al., 1984; Westcott et al., 1997).

Many studies have used AVHRR-NDVI satellite imagery to examine crop production and dynamics (Malingreau, 1989; Fischer, 1994; Reed et al., 1997) and estimate crop yields (Rasmussen, 1992; Groten, 1993; Potdar, 1993; Quarmby et al., 1993; Idso et al., 1980). Benedetti & Rossini (1993) examined AVHRR-NDVI crop growth profile – crop physiology relationships of specific fields in Italy to use at the regional scale for national production forecast organizations. Quarmby et al. (1993) used SPOT HRV imagery as an intermediate step to link field ground measurements to AVHRR pixels in Greece. In that study, integrated NDVI profiles were used to estimate wheat yields as a proportion of a reference year. The estimates were to be used by national agencies for an early warning system for crop failure. Few studies have applied NDVI seasonal crop growth profiles for estimation of wheat yields at the farm scale, specifically for use by farmers.

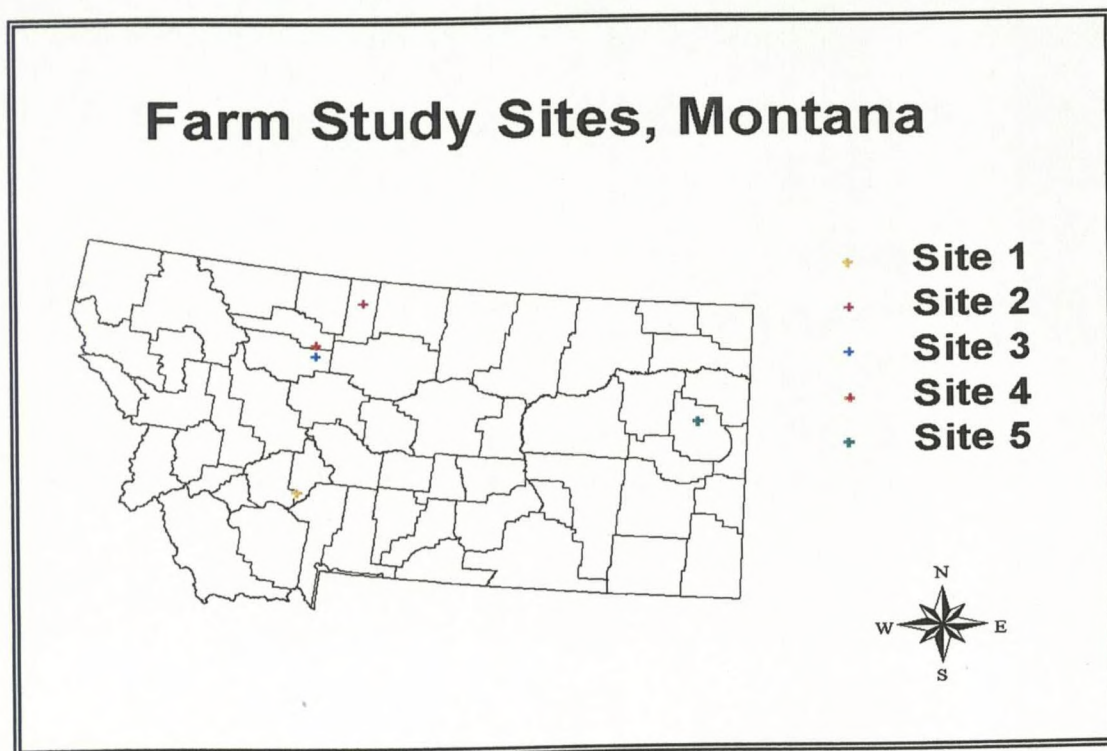
This research examines the potential of using readily available AVHRR-NDVI satellite data for estimation of wheat yields and protein concentration prior to harvest at the farm level. The goals are to (1) determine if the NDVI growth could be used to estimate yields and protein concentration for farms in Montana, (2) determine times in the growing season that produce the most useful estimations of yield and protein concentration, and (3) determine the utility of real-time delivery of satellite yield predictions for farm management.

Methods and Materials

Study sites

Study sites, yield, and protein concentration data used in this project were from the farms of cooperating producers who work with Montana State University to enhance farm management. Five study sites were evaluated at locations in Central, North Central, Southwestern, and Northeastern Montana (Figure 8). Study site descriptions were obtained from MAPS Atlas (1990) and from producers.

Figure 8. Five farm sites included in study.



Site 1 is in Southwestern Montana at approximately 46.0 north and 111.6 west. The landscape is nearly level to gently sloping, and moderately dissected, the majority of which is in range or dryland crop production. Elevation is about 1280 – 1430m with a growing season that averages 100 – 120 days. Soils in this region tend to be deep, consisting of calcium-rich Aridisols. Mean annual precipitation is 300 – 350mm, the majority of which falls in May and June (MAPS Atlas, 1990). Production for this site consists of spring wheat that was planted in alternate years from 1989 to 1997.

Site 2 is in North Central Montana. It is located at 48.6 north and 111.0 west. The landscape is nearly level to sloping with elevations of about 1040 – 1070m. Land use consists primarily of range and dryland agriculture. Climate is continental with cold winters and hot summers with a growing season of 100 – 110 days. Mean annual precipitation is 300 – 350mm, with the majority occurring in April through July. Soils are typically deep Argiborolls. The management of Site 2 has varied from year to year. From 1989 through 1995, alternate crop fallow farming was practiced while in 1996 and 1997 it was in annual crop. In 1989, spring wheat and barley were planted. In 1990 and 1991, only spring wheat was planted. From 1992 through 1994 winter wheat was planted and since 1995, spring wheat has been grown. This complicates the process of making year-to-year comparisons. To compensate, yield estimates equivalent to spring wheat yields were derived. The ratio used to convert winter wheat yields to spring wheat yields was 1.12:1 based on conversions derived from Brown & Carlson (1990). The ratio for barley to spring wheat was 1.6:1. Also, annual crop years were excluded since a continuous canopy cover would naturally have a higher NDVI value than a partial canopy

cover due to alternate crop-fallow. Annual cropped years (1996-1997) for Site 2 were not evaluated in this study (MAPS, 1990; Mattson, pers. comm., 1998).

Sites 3 and 4 are in Central Montana. Site 3 is located approximately at 47.8 north and 111.7 west, while site 4 is at 48.48 north and 111.7 west. The landscape is level to sloping, moderately dissected, with some bench and terrace lands. Dryland agriculture is the predominant land use in this area. Elevations range from 1040 to 1220m and result in a growing season of 115 – 125 days. Soils consist of Argiborolls and Camborthids with some Entisols near flood plains and terraces. Mean annual precipitation ranges from 250 – 350mm with the majority falling during April through July (MAPS, 1990). Both sites have been in an alternate crop-fallow system for all years. Crops have varied over the years between winter wheat, spring wheat, and barley. Yields were converted to spring wheat yield equivalents for consistency with other study sites.

Site 5 is in Northeastern Montana and is located at 47.4 north and 104.9 west. The landscape consists of rolling hills and soils consist primarily of Argiborolls. Climate is continental with cold winters and warm summers, with a growing season range of 110 – 130 days. Eighty percent of the precipitation occurs during the growing season and averages between 300 – 350mm per year (MAPS, 1990). Production consists of alternate crop-fallow of spring wheat over all years.

Wheat yield and protein concentration data were averages from fields considered representative of each farm. The largest, most homogenous fields were chosen for examination and ranged from a full section of land (256 hectares) to smaller fields (38 hectares). Geographic coordinates of the center of each field were used as the center point for NDVI pixel value extraction. Yield data were obtained from the producers for the

five farm study sites while protein concentration data were available from sites 1 through 4. Yield and protein concentration data were an average for each study site location.

Satellite data processing and analysis

NDVI data produced from the AVHRR satellite sensor were obtained from EROS Data Center in Sioux Falls, SD. NDVI data were in the form of biweekly maximum value composites (MVC). Compositing NDVI helps reduce atmospheric contaminants in the imagery (Holben, 1986). (See Chapter 1 for more detail on NDVI).

NDVI data from the years 1989 through 1997 and over the months April through mid-September were examined for this study. Data were imported and converted into Arc/Info7 GRID™ format in a Lambert Azimuthal Equal Area (LAZEA) projection. A point coverage for each farm site location was produced from the geographic coordinates provided from the farmers. The coverages were created in Arc/Info7™ and projected to a LAZEA projection to be consistent with the projection of the NDVI grids. Farm point coverages were converted into 3km x 3km grid cells and used as a “mask” to extract NDVI data in a 3km x 3km-averaged cell region. AVHRR data have resolution of 1km² and registration is at best plus or minus one pixel. Using a 9km² average is needed to overcome the limits of AVHRR-NDVI imagery and account for image-to-image geometric misregistration of the data. These averaged NDVI values are considered representative of the real NDVI values at the field point locations and surrounding areas.

Growth profile analysis

NDVI growth profiles were integrated from approximately April 1 through September 15 and consisted of 12 biweekly periods (Integrate 12 periods), integrated apparent growing season (Integrate AGS) determined using the approach of Groten (1993), and summation of NDVI over 12 biweekly periods (Sum 12 periods). End of the month NDVI (End April through End August), summation through each month of NDVI (sum through end of April to August), and summation of NDVI through critical periods as proposed by Doraiswamy & Cook (1993) (sum 6/ 4-8/2; 6/22-8/16, 7/6-8/30) were also calculated. This was to determine if, as the growing season progresses, there are times when final wheat yield and/or protein concentration could be estimated. Integration of NDVI and statistical analyses were identical to those discussed in Chapter 2.

Additional NDVI parameters were included that examined the growth profiles for each site and further explored early season integration when management decisions could be made and applied. NDVI growth profiles were integrated from the beginning of the apparent growing season (AGS) to the maximum NDVI value for each season (Integrate to maxNDVI). This should reveal the crop productivity during the "green-up" of the crop when photosynthetic activity is increasing to maximum canopy cover (Malingreau, 1989). Conversely, maximum NDVI to the end of the apparent growing season (Integrate after maxNDVI) should reveal information about the senescence of the crop canopy and length of the grain-filling period. The integration of these periods was examined to determine which part of the growing season curve is more correlated to final wheat yield or protein concentration in spring wheat.

Since we were particularly interested in examining the early growing season for yield and protein concentration estimation, the first six biweekly periods were integrated successively and used in the analysis with yield and protein concentration data.

Calculation began according to the beginning of the apparent growing season (AGS) and integrated through the first two biweekly periods (Integrate first 2 periods), then first three biweekly periods (Integrate first 3 periods), and so on through the first 6 biweekly periods (Integrate first 6 periods).

A variation of the critical periods defined by Doraiswamy & Cook (1993) was also examined. Time of maximum NDVI in a growth profile typically corresponds to time of maximum photosynthesis and marks the end of the vegetative phase and beginning of the reproductive/maturation phase (Benedetti & Rossini, 1993; Groten, 1993). This phase in the growth cycle might be useful to examine because stresses that occur during this time are shown to strongly influence final wheat grain yield and protein concentration. Additionally, it is at this time that an application of supplemental N would be considered. The approach examined individual farm maximum NDVI (maxNDVI) and summation of maximum NDVI plus the NDVI value from one, two, and three biweekly periods (MaxNDVI+1, 2, and 3 periods).

The final NDVI parameters examined were time of apparent growing season onset (Onset Julian Day) and end of apparent growing season (Ending Julian Day). Time of NDVI onset and end of growing season is considered to be the period when "apparent" growth begins and ends based on work by Groten (1993). The time of onset or end of season might be considered reflective of final grain yield if crops show patterns that relate to climatic conditions such as, early/dry season or late/wet season relationships. For

example, sufficient moisture early in the growing season is important for soil moisture storage, good seed germination, and early vegetative growth. In a warm, dry year crop emergence might come early but be an indication of early season drought and lower crop yields. In a wet early season, time of emergence might be delayed due to cool, wet conditions but could be related to a "bumper crop" because of the increase in stored soil moisture for later crop use. The end of the apparent growing season could also indicate whether a growing season has been prolonged due to favorable climatic conditions, and correspond to a longer grain filling period in wheat.

Reported yield data and protein concentration were included with NDVI parameters in a multiple linear regression model using SAS statistical software (SAS Institute Inc., Cary, NC). Farms were included as indicator variables to reduce error variance and increase the precision of the NDVI parameter. Interactions were examined to test for heterogeneity of slopes. The model equation for a NDVI parameter with indicator variables and interaction was:

$$\text{Wheat yield} = \text{NDVI parameter} + \text{farm} + \text{NDVI parameter} * \text{farm}$$

where NDVI parameters have ratio-interval values and farms have classificatory values.

The inclusion of farms or interaction terms in the final model was determined by significance (p-value) in the full model, with a minimum p-value of 0.05 for inclusion of the parameter in the final model. For models that found no significant difference including interactions and farms, a simple regression analysis was performed between wheat yield and the NDVI parameter.

Results

Spring wheat yield – NDVI parameters relationships

Results of farm wheat yield – NDVI parameter relationships are presented in Table 4. The table includes the adjusted R^2 , p-value, and mean square error (MSE) of each wheat yield – NDVI parameter relationship. Farms and interaction p-values are presented to show when they were included in a model. Degrees of freedom associated with each model are also presented.

Adjusted coefficients of determination for models ranged from no relationship to moderate relationship with the strongest correlation occurring between spring wheat yields and apparent growing season (Integrate AGS) (adj. $R^2 = 0.688$, p-value = 0.0001). Integrated NDVI over the entire growing season (Integrate 12 periods) and summation through the growing season (Sum 12 periods) revealed lower correlation (adj. $R^2 = 0.628$ and 0.613, respectively) and were both highly significant (p-value = 0.0001). Farms were significantly different and included in the final model while interactions were not significantly different for these NDVI parameters (Table 4).

Early season integrated NDVI parameters indicated no relationship with spring wheat yields. Successive integration through the first six biweekly periods (Integrate first 2 periods through first 6 periods) and integration to maximum NDVI (Integrate – maxNDVI) relationships were very low, as well as the maximum NDVI (Table 4). Integration from maximum NDVI through the end of the growing season (Integrate from maxNDVI) revealed a moderately stronger and significant relationship (adj. $R^2 = 0.367$, p-value = 0.0001).

Table 4. Spring wheat yield – NDVI parameter relationships for farm study sites.
Corrected total SS is 12133.50.

NDVI Parameter	Adjusted R ²	p-value, full model	MSE	p-value, interaction	p-value, farm	DF
Integrate 12 periods	0.628	0.0001	115.56	0.512	0.0003*	34
Integrate AGS	0.688	0.0001	97.01	0.068	0.0002*	34
Sum 12 periods	0.613	0.0001	120.15	0.401	0.0003*	34
Integrate first 2 periods	0.0	0.683	317.89	0.542	0.089	38
Integrate first 3 periods	0.0	0.558	316.40	0.339	0.108	38
Integrate first 4 periods	0.003	0.298	310.21	0.349	0.113	38
Integrate first 5 periods	0.003	0.293	310.02	0.574	0.119	38
Integrate first 6 periods	0.024	0.171	303.74	0.640	0.123	38
Integrate to maxNDVI	0.065	0.0610	290.77	0.493	0.119	38
Integrate from maxNDVI	0.367	0.0004	196.76	0.055	0.024*	34
Sum through April	0.0	0.4263	313.96	0.174	0.090	38
Sum through May	0.360	0.0041	199.20	0.033*	0.026*	30
Sum through June	0.242	0.017	235.68	0.461	0.042*	34
Sum through July	0.089	0.0343	283.33	0.432	0.080	38
Sum through Aug.	0.619	0.0001	118.50	0.503	0.0003*	34
End of April	0.226	0.045	240.72	0.106	0.031*	34
End of May	0.093	0.031	282.02	0.919	0.066	38
End of June	0.324	0.001	210.42	0.549	0.031*	34
End of July	0.510	0.0001	152.24	0.148	0.019*	34
End of Aug.	0.352	0.0006	201.53	0.066	0.010*	34
Maximum NDVI	0.243	0.0007	235.56	0.698	0.085	38
MaxNDVI+1 period	0.305	0.0001	216.27	0.801	0.070	38
MaxNDVI+2 periods	0.383	0.0002	191.94	0.943	0.037*	34
MaxNDVI+3 periods	0.421	0.0001	180.04	0.949	0.016*	34
Sum (6/4 – 8/2)	0.092	0.032	282.31	0.293	0.098	38
Sum (6/22 – 8/16)	0.052	0.084	294.93	0.129	0.083	38
Sum (7/6 – 8/30)	0.028	0.053	302.20	0.054	0.069	38
Onset Julian Day	0.042	0.108	298.01	0.828	0.099	38
Ending Julian Day	0.447	0.0001	171.88	0.437	0.006*	34

* Farm and/or interaction term included in final model

This trend could also be seen in early season monthly NDVI parameters.

Correlations were low until later in the season in July or August. Summation through August (Sum through Aug.) revealed a stronger relationship than summation through the entire growing season (Sum 12 periods) (adj. $R^2 = 0.619$ and 0.613 , respectively), while NDVI at the end of July (End July) had a stronger relationship with spring wheat yields than NDVI at the end of August (adj. $R^2 = 0.510$ and 0.352 , respectively). For most NDVI monthly parameters, farms were significantly different and included in the final model. Interactions were not significant in any model with the exception of summation through May (Sum through May) (p-value = 0.033).

Relationships between critical periods and yield were usually not significant. For critical periods adopted from Doraiswamy & Cook (1993), there was no correlation with spring wheat yields. Relationships were stronger for summation from maximum NDVI plus one, two, and three biweekly periods (MaxNDVI + 1, 2, and 3 periods). Correlation increased as biweekly periods were added to the maximum NDVI with the strongest relationship occurring with maximum NDVI plus three biweekly periods (adj. $R^2 = 0.421$, p-value = 0.0001). For this parameter, farms were significantly different and included in the model (p-value = 0.016).

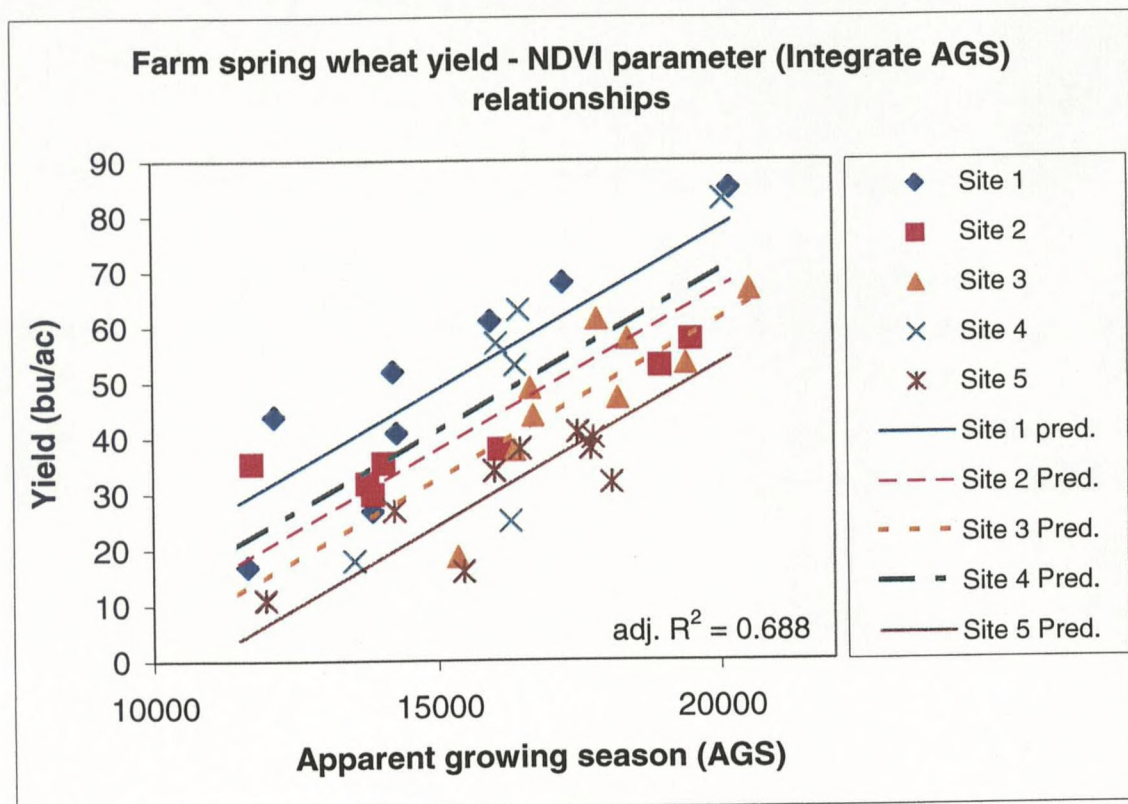
Onset Julian date and ending Julian date show similar characteristics to other NDVI parameters. Relationships with yield are stronger in the later part of the growing season with onset Julian day adj. $R^2 = 0.042$, and for ending Julian day adj. $R^2 = 0.447$.

Analysis of best model scatterplots

Scatterplots of the three best fitting models are presented in figures 9 through 11. Figure 9 represents the fit of the predicted slopes to the reported yield values based on integrated NDVI for the apparent growing season (Integrate AGS). This relationship (adj. R^2 value = 0.688, p-value = 0.0001) is the strongest of all NDVI parameters examined. Differences among farms were significant and included in the final model (p-value = 0.0002). Interactions were not included (p-value = 0.068) which can be seen by the parallel slopes of the prediction lines.

Figure 9. Farm spring wheat yield – NDVI (integrate AGS) relationships.

Adj. R^2 = 0.688, p-value = 0.0001. Regression equations are: Site 1 yield = $(-38.48) + (0.006)$ (Integrate AGS), Site 2 yield = $(-49.34) + (0.006)$ (Integrate AGS), Site 3 yield = $(-54.68) + (0.006)$ (Integrate AGS), Site 4 yield = $(-46) + (0.006)$ (Integrate AGS), Site 5 yield = $(-63.14) + (0.006)$ (Integrate AGS).



Predicted yield slopes are positive and show moderate to strong relationship with reported yield values. The scatterplot reveals the range in reported yields across farms is broad. Site 1 shows the largest range in actual yield values from 17 to 85 bu/ac and Site 5 has the smallest range from 11 to 41(bu/ac). Our model predicts Site 1 would produce the highest yields for a given AGS value, while Site 5 would produce the lowest.

Figure 10 represents the fit of the predicted slopes to the reported spring wheat yield based on the NDVI parameter, summation through August (Sum Aug.). The model produced the third strongest relationship among NDVI parameters (adj. $R^2 = 0.619$, p-value = 0.0001). Farms proved to be significantly different in this model (p-value = 0.0003). Predicted yields again show positive slope with Site 1 predicted to have the highest yields and Site 5 to have the lowest yields for a given summation through August value. Predicted yield slopes are positive and show moderate to strong relationship with reported yield values. The scatterplot reveals the range in reported yields across farms is broad. Site 1 shows the largest range in actual yield values from 17 to 85 bu/ac and Site 5 has the smallest range from 11 to 41(bu/ac). Our model predicts Site 1 would produce the highest yields for a given AGS value, while Site 5 would produce the lowest.

Figure 10. Farm spring wheat yield – NDVI (Sum Aug.) relationships.

Adj. $R^2 = 0.619$, p -value = 0.0001. Regression equations are: Site 1 yield = $(-404.17) + (0.338)$ (Sum Aug.), Site 2 yield = $(-412.72) + (0.338)$ (Sum Aug.), Site 3 yield = $(-426.27) + (0.338)$ (Sum Aug.), Site 4 yield = $(-417.31) + (0.338)$ (Sum Aug.), Site 5 yield = $(-429.91) + (0.338)$ (Sum Aug.).

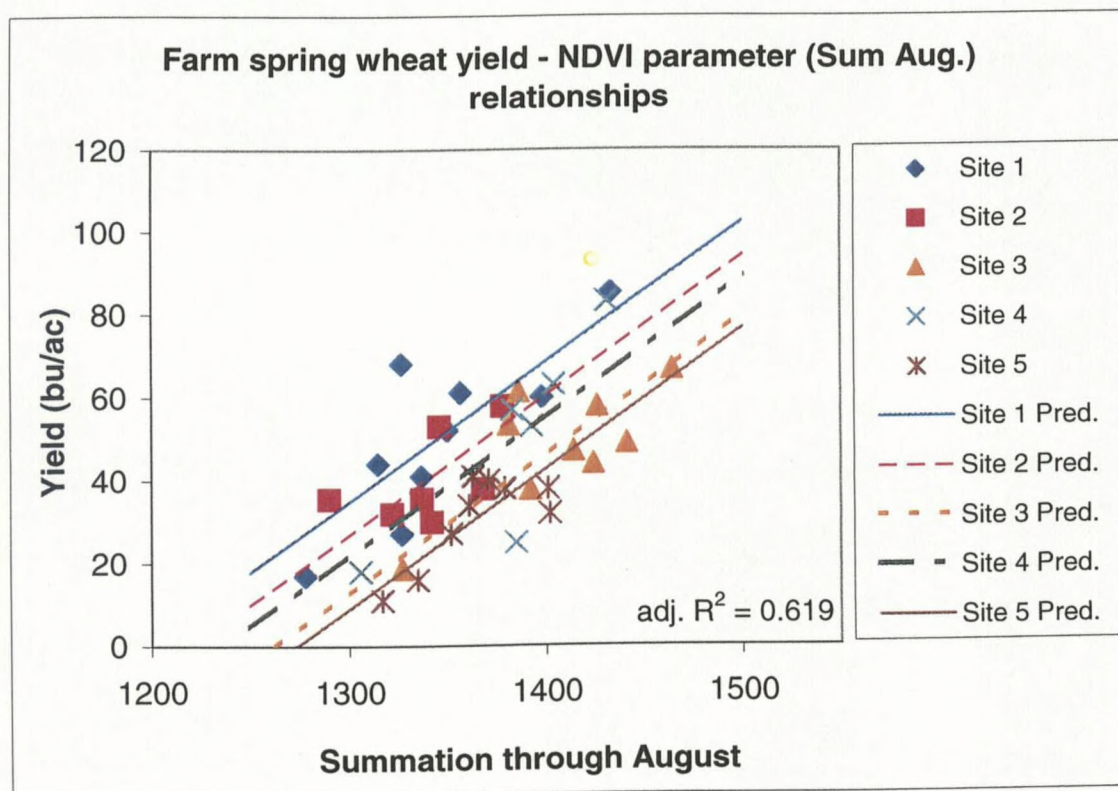
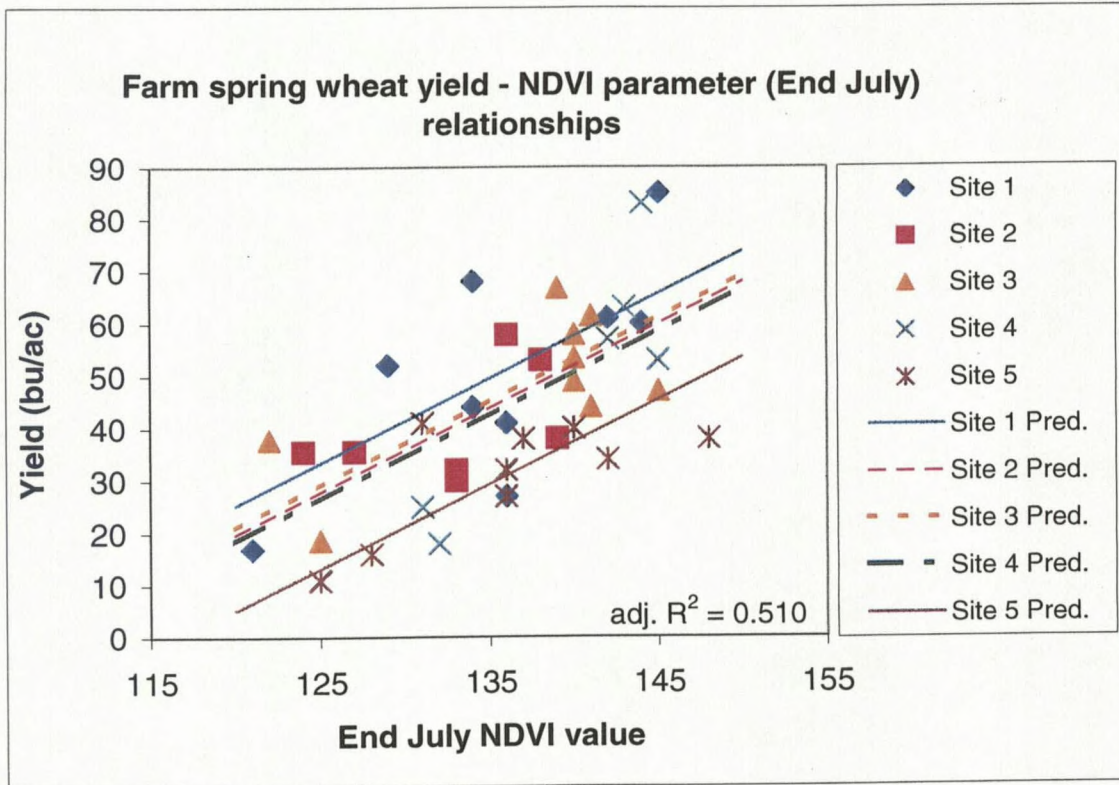


Figure 11 represents the fit of the predicted slopes to the reported yield values based on the NDVI parameter, End of July. This model produced the fifth strongest relationship with an adj. R^2 of 0.510, p -value = 0.0001. Differences among farms were significant enough to be included in the final model (p -value = 0.019) but interactions were not (p -value = 0.148). In this model, much more scatter is present between predicted spring wheat yield slopes and reported values reported by farms. The range in the NDVI parameter, End July, is much narrower than other parameters.

Figure 11. Farm spring wheat yield – NDVI (End July) relationships.

Adj. $R^2 = 0.510$, p-value = 0.0001. Regression equations are: Site 1 yield = $(-177.72) + (1.609) (\text{End July})$, Site 2 yield = $(-183.32) + (1.609) (\text{End July})$, Site 3 yield = $(-181.97) + (1.609) (\text{End July})$, Site 4 yield = $(-184.61) + (1.609) (\text{End July})$, Site 5 yield = $(-197.85) + (1.609) (\text{End July})$.



Protein concentration – NDVI parameters relationships

Results of protein concentration – NDVI parameter relationships are presented in Table 5. The table includes the adjusted R^2 , p-value, and mean square error (MSE) of each protein concentration – NDVI parameter relationship. Farms and interaction p-values are presented to show when they were included in a model. Degrees of freedom associated with each model are also presented.

