PROTEIN MAPPING SPRING WHEAT USING A MOBILE NEAR-INFRARED SENSOR AND TERRAIN MODELING

by

Corey Grant Meier

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

of

Master of Science

in

Soils

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2004

APPROVAL

of a thesis submitted by

Corey Grant Meier

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.

Daniel Long

Richard Engel

Approved for the Department of Land Resources and Environmental Sciences

John Wraith

Approved for the College of Graduate Studies

Bruce R. McLeod

STATEMENT OF PERMISSION TO USE

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

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

Corey Meier

April19, 2004

ACKNOWLEDGEMENTS

The financial support and dedication of instrumentation from Case New Holland is greatly appreciated, as is the expertise of personnel from Textron Systems, particularly Jessica Drake, for her instruction of modeling processes. The field data collected by Dr. Andrew Lenssen is much appreciated and finally, the contributions of data, insight and instruction by committee chairs, Dr. Dan Long and Dr. Rick Engel, in addition to committee members Dr. Bill Quimby and Dr. Gerry Nielsen is greatly appreciated. Finally, innumerable thanks to my wife, Noelle, for her encouragement and patience with this endeavor.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF TABLES - CONTINUED

Table Page

LIST OF FIGURES

Figure Page

LIST OF FIGURES - CONTINUED

Figure Page

ABSTRACT

The feasibility of using a combine harvester mounted, Case New Holland ProSpectra™ Grain Analyzer to determine protein content of hard red spring wheat (Triticum aestivum L.) was tested at a level of accuracy of 0.5% protein. Should the sensor meet the target accuracy, it would be useful for on-farm segregation of wheat by protein content and for field mapping protein. Geo-referencing protein values in the field, at harvest, enables creation of protein maps to be used as surrogates for maps of soil nitrogen for directing variable rate nitrogen applications. Models were developed in the lab using partial least squares regression to relate spectra in near infrared wavelengths to grain protein, then tested on a combine harvester. The sensor did not meet the target level of accuracy in the lab and performance degraded in the field. Poor field performance was due to the presence of contaminants present in grain samples at harvest and to sensor hardware problems that yielded inconsistent results. These results suggest that the sensor is not accurate for the purpose of segregating grain, but may be accurate enough to characterize broad patterns of variability in a field to direct soil sampling efforts, if hardware problems can be fixed to replicate the best lab results in the field. An approach to characterize topographic relief in fields using LandmapR™ software in a geographic information system was utilized to relate hillslope position to grain protein and yield variability. A direct relationship would allow partitioning fields into management zones for directing soil sampling efforts and variable rate applications of soil nitrogen based on hillslope. A five class categorical hillslope model and wetness index explained the most variability of the models tested, explaining less than 24% variance in protein and less than 18% variance in yield. Hot conditions during grain fill were likely responsible for the low variance. Yield and protein increased downslope, most likely in response to deeper soils at downslope positions that retained moisture later in the growing season for plant uptake. Hillslope position was deemed inadequate to predict grain protein and yield, but sufficient for directing soil sampling.

1. GENERAL INTRODUCTION

Efficient agriculture practices necessitate the ability to increase production, while effectively controlling input costs. The ability to accurately place fertilizer, herbicides and pesticides where they are most needed is a pursuit of maximum benefit for minimum cost. This style of farming, called precision agriculture, utilizes new technology to meet these goals so is a rapidly changing field of study.

Reliable technology allows producer to make informed management decisions based on data collected in the field. Dividing a field, characterized by variable production, into manageable zones of homogenous qualities, such as soil fertility is a method for producers to direct inputs to locations where they provide optimum benefits. To date, accurately characterizing soil fertility of fields has proven to be expensive at high spatial resolutions, such as soil sampling on a grid. More cost-effective methods, such as using yield monitors or remote sensing, do not reveal the whole picture of soil nutrition in a field, but are beneficial in directing soil sampling efforts.

An application of existing technology for field applications may provide an answer for determining soil nutrition in a cost-effective manner. Case New Holland, in union with Textron Systems, has developed a near infrared sensor that can be taken to the field to measure the protein, oil and moisture contents of grain. In Montana, where dryland wheat is grown, such a sensor could be used to help producers take advantage of protein premiums for their crops and assist them in developing variable rate nitrogen (N) fertilizer plans. In Montana, producers are paid a protein premium for wheat over 12%,

so having a protein sensor on the farm to help them segregate grain based on protein can help them realize new profits. The second application of this technology on the farm is in preparing a variable rate nitrogen application strategy. By mounting the sensor on a combine and geo-referencing protein readings with a global positioning satellite system, protein maps of a field can be prepared. Because nitrogen is a major component of grain protein, the protein map can be used as a surrogate for soil nitrogen mapping, thereby illustrating the variance of soil N and allowing the producer to prepare variable N application maps.

This thesis follows development of the Case/Textron ProSpectra™ Grain Analyzer (PGA) from development through application and concludes with an examination of protein and yield variability in fields under semiarid conditions, in response to topographic hillslope position. Models for converting raw spectra, collected by the PGA, to grain protein are developed and tested in a laboratory setting in Chapter 1. In Chapter 2, the PGA and models developed in the first chapter are tested in the field under normal harvest operating conditions on a combine harvester and a determination of the sensor for on farm applications is made. Chapter 3 studies the relationships of protein and yield to landscape position on several fields in northern Montana over a three-year period. The goal of studying these relationships being to assess if hillslope maps can explain sufficient protein and yield variability to justify their use for delineating soil N management zones.

2. CALIBRATION OF THE PROSPECTRA GRAIN ANALYZER

Introduction

The ProSpectra™ Grain Analyzer (PGA), is a joint venture of Case New Holland, Inc. and Textron Systems, Inc.. It is designed to be the first mobile sensor that functions on the principles of near-infrared reflectance (NIR) (Engel et al., 1997) for the purposes of determining protein, moisture, and oil content in grains as they are harvested. Introduction of the PGA into the agricultural marketplace has at least two potential benefits including segregating grain based on protein content (Case, 2000) and improving precision nutrient management practices for dryland wheat (Engel et al., 1999, Long et al., 2000, Long et al., 1998).

The ability to segregate wheat (*Triticum aestivum L.*) according to its protein content on the farm, or as it is harvested, would provide an opportunity to the farmer to receive a price premium on high protein grain. For example, over a 10-year period (1993-2002) the average price for hard red spring wheat in Montana was $4.40 and $4.96 per bushel at 13 and 14% protein, respectively. Hence, the premium over this period averaged $0.56 per bushel for wheat that was grown at 14% vs. 13% protein (data supplied by Montana Wheat and Barley Committee).

Development of a mobile grain protein sensor, such as the PGA, will enable producers to map grain protein levels in their fields in much the same way as yield mapping is currently conducted (Wibawa et al., 1993). Such maps can be used to index

soil nitrogen (N) fertility status across a field landscape indicating areas of a field where current N fertility programs are adequate or deficient for maximum economic return. In addition, protein maps in combination with yield maps can be used as a surrogate tool to intensive grid soil sampling and N testing (Engel et al., 1999; Long et al., 2000). This capability can speed adoption of precision nutrient management in regions of the Great Plains where this practice is not being utilized.

The PGA sensor operates by irradiating a grain sample with light from a tungsten/halogen bulb. Light covering a broad spectrum strikes the grain sample and a portion is reflected back to the sensor. The ratio of reflected to emitted light is a function of the chemical composition of the sample and forms the premise behind NIR spectroscopy. Murray and Williams (1987) have reviewed the chemical principles of NIR spectroscopy. A brief summary is included here. Atoms and molecules are in constant motion and vibrate at frequencies that correspond to wavelengths in the infrared portion of the electromagnetic spectrum (750-25,000nm). As a molecule is irradiated, it absorbs some of the energy from the radiation at wavelengths unique to that molecule and the remainder of the energy is reflected. The absorbence by a compound is proportional to the concentration of the absorbing molecules and is noted as $A=\log_{10}(1/R)$, where R is the energy reflected. An NIR sensor reads the wavelengths of reflected energy and infers that the absorbed energy corresponds to the presence of a particular molecule. The ratio of reflected to absorbed energy is expressed as A, the optical density, in the equation above. Several factors complicate the sensing process, including having a sample sufficiently thick to absorb or reflect all emitted light, refraction of light, absorption by

water, oil or other substances, particle size and shape and the crystalline structure of the sample. Temperature fluctuations can also affect accurate sensor readings, but can be compensated for through sensor design. Error due to the presence of moisture is compensated for using a moisture correction factor in the NIR calibration model.

Estimating grain quality through NIR spectroscopy requires developing calibration or spectroscopic models that relate the amount and intensity of reflected light to the chemical property being estimated, i.e. protein content. Spectroscopic models can be developed by a variety of statistical approaches including partial least squares analysis (PLS) (Hruschka, 1987; Martens and Naes, 1987), multiple linear regression (MLR) (Martens and Naes, 1987), or principal component regression (PCR) (Martens and Naes, 1987). When spectra are collected, optical density (log 1/R) is plotted on the y-axis and wavelengths, plotted on the x-axis. MLR can outperform PLS if there is extreme non-linearity and a surplus of x variables; however, PCR and PLS usually provide better results because MLR can not handle multi-collinearity between wavelengths (Bjorsvik and Martens, 1992; Hrushka, 1987; Naes and Martens, 1985). PLS treats multi-collinearity beneficially, as a method to reduce effects of noise. Another advantage of PCR and PLS over MLR is that all available wavelengths are used for model development. Doing so has a smoothing effect, thereby reducing noise and including more relevant data in the model (Bjorsvik and Martens, 1992; Martens and Jensen, 1983). PLS has advantages over PCR because the most pertinent data are concentrated in the first few factors. This reduces the effects of noise that are found in later factors. A more complete description of PLS can be found in Martens and Naes (1987).

The objectives of the research in this chapter are to i) develop calibration models for the PGA sensor using PLS analysis that can be used to predict the protein content of hard red spring wheat; ii) test the accuracy and reproducibility of these models; and iii) determine if the sensor meets Case and Textron accuracy requirements of 0.5% protein.

Materials and Methods

Instrument Description

Lab calibration and validation studies were conducted on a beta version (so named by Textron Systems) of the Case/New Holland PGA. The unit was designed to be mounted on a combine and to interface with an in-cab data collector. For this lab based study, the unit interfaced with a Pentium II, 233MHz laptop computer. The PGA measured reflectance from whole kernels in the visible (300-750nm) and near-infrared range (750-2600nm). It consisted of two components; the sensor probe and the base unit (Figure 1a). The base unit was a closed system (30 cm x 24 cm x 10 cm) that contained the power supply, electronics and an internal cooling unit. It was designed to withstand the rigors of field conditions at harvest including heat, dust, vibration and interference from other electronic instruments. The probe (7 cm x 7 cm x 5 cm)**,** contained a sapphire lens (2 cm dia.), tungsten energy emitting bulb, reference shutter and fiber optic pickup that carried reflected spectra to detection sensors in the base unit. During operation, the reference shutter closed over the lens every 15 minutes and several internal reference spectra were recorded. This information was used to compute optical density (log

internal reference spectra/sample spectra), or theoretically log(1/R) (Murray and Williams, 1987). The sensor also performed an internal recalibration after every reference spectra collection.

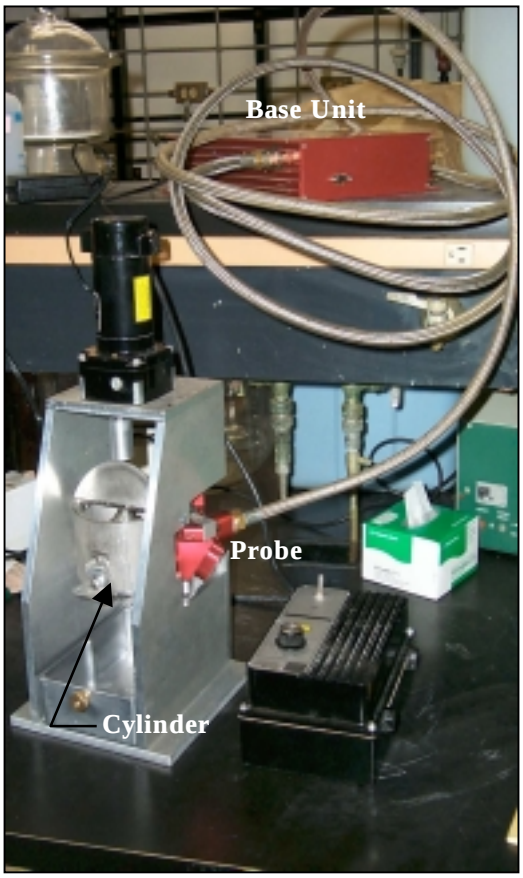
The PGA was designed to record spectra on a moving sample, i.e grain-flow through an auger. To maintain consistency in sample presentation between the field and the lab, the probe was connected to a 1000 ml cylinder and spinner assembly during the lab calibration phase of this study. This apparatus was designed to simulate the grain flow and packing conditions found on a combine (Figure 1b). Grain was placed in the cylinder and a center mounted impeller spun grain past the head at a rate similar to what would be found on the combine. The spinner was a closed system, to prevent ambient light from entering the cylinder. D2Tsc 3.0a (DSquared Development Inc., LaGrande, OR.), a proprietary software developed for the sensor was used to operate the PGA and record spectra reflected from the kernels.

Sample Selection

Grain samples for calibration and validation studies were derived from a small-plot field study conducted over a three-year period at the Northern Agricultural Research Center, in Havre, Montana. This study consisted of four spring wheat cultivars (Amidon, Hi-line, McNeal and Rambo) grown under a diverse set of water and N fertility regimes. Details of the experimental design were described by Engel et al. (1999). Grain samples from that study provided a robust data-set embodying a wide range of color, protein levels (Figure. 2a) and kernel sizes (Figure. 2b). Samples were cleaned prior to analysis

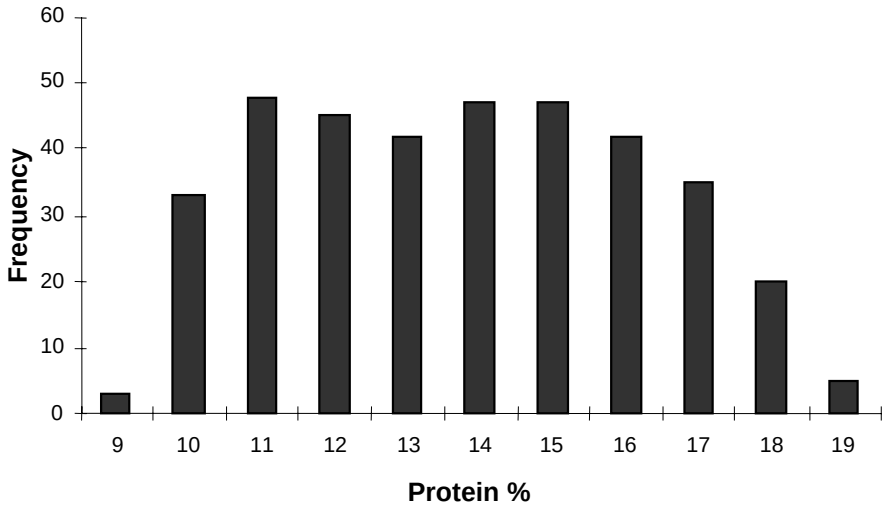
of grain protein. Reference protein values were determined through LECO dry combustion (Sweeney and Rexford, 1987) and corrected to a 12% moisture basis.

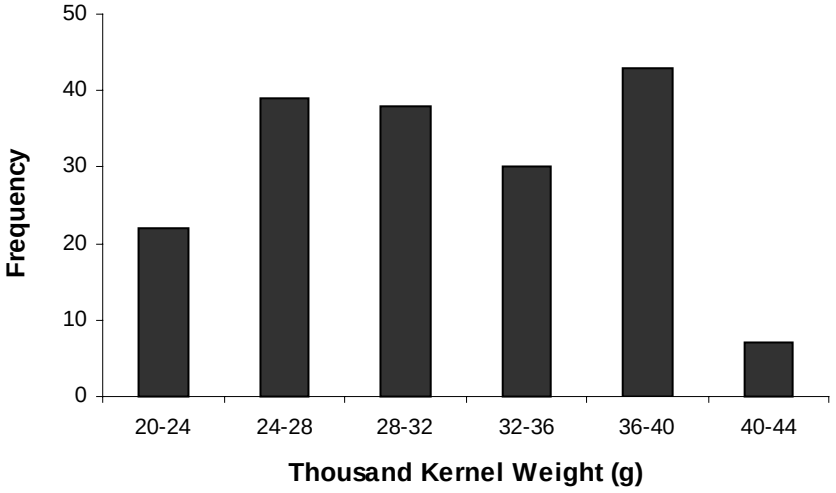
(a)



(b)

Figure 1. Close-up of ProSpectra Grain Analyzer (a) in lab configuration (b).





(b)

Figure 2. Distribution of protein values (a) and kernel sizes (b) in the calibration set of 366 samples.

Collection of Spectra

A total of 669 grain samples were used during the calibration and validation phases (Figure. 3). The approach used to calibrate the instrument was similar to that

described by Hruschka (1987). Three hundred and sixty six samples were dedicated to calibrating the instrument and three hundred samples formed a validation set, used to test the accuracy of the instrument models developed during the calibration phase. The three remaining samples served as a check set, that was used to test instrument precision or reproducibility. The three check samples had protein concentrations of 11.5, 14.4, and 18.4%.

The PGA was turned on 30 minutes prior to commencing the collection of spectra to allow the unit to warm-up. Each sample was mixed and a 250g sub-sample was placed in the spinner assembly. The spinner was set at 70 RPM to replicate flow characteristics of grain in the bypass unit, or clean grain auger, on the combine. Five spectra were collected per sample as the grain flowed past the probe. The process took approximately 10 seconds to complete per spectra. Broken kernels, which can cause errors in prediction (Martens and Jensen, 1983), were not a problem due to the short time the samples remained in the spinner.

Raw spectra were stored on a laptop PC using D2Tsc software (DSquared Development Inc., LaGrande, OR). D2Tsc also performed a transformation, proprietary in nature, of the raw spectra to sample spectra. Sample spectra were then used to perform the calibration and validation tests. The D2Tsc software collected 100 spectra per file. Of those, the first 15 spectra (3 samples) were derived from scans of the check samples representing low (11.5%), medium (14.4%) and high (18.4%) protein levels.

The lens of the sensor probe was periodically cleaned to prevent a buildup of dust that could impair the performance of the sensor (Williams, 1975). The lens and entire

spinner unit was thoroughly cleaned after every 20 samples, or 100 scans, to reduce the potential for cross-contaminating the samples.

Of the five spectra collected per sample, three were used to form a set of calibration spectra and the remaining two formed a prediction spectra set.

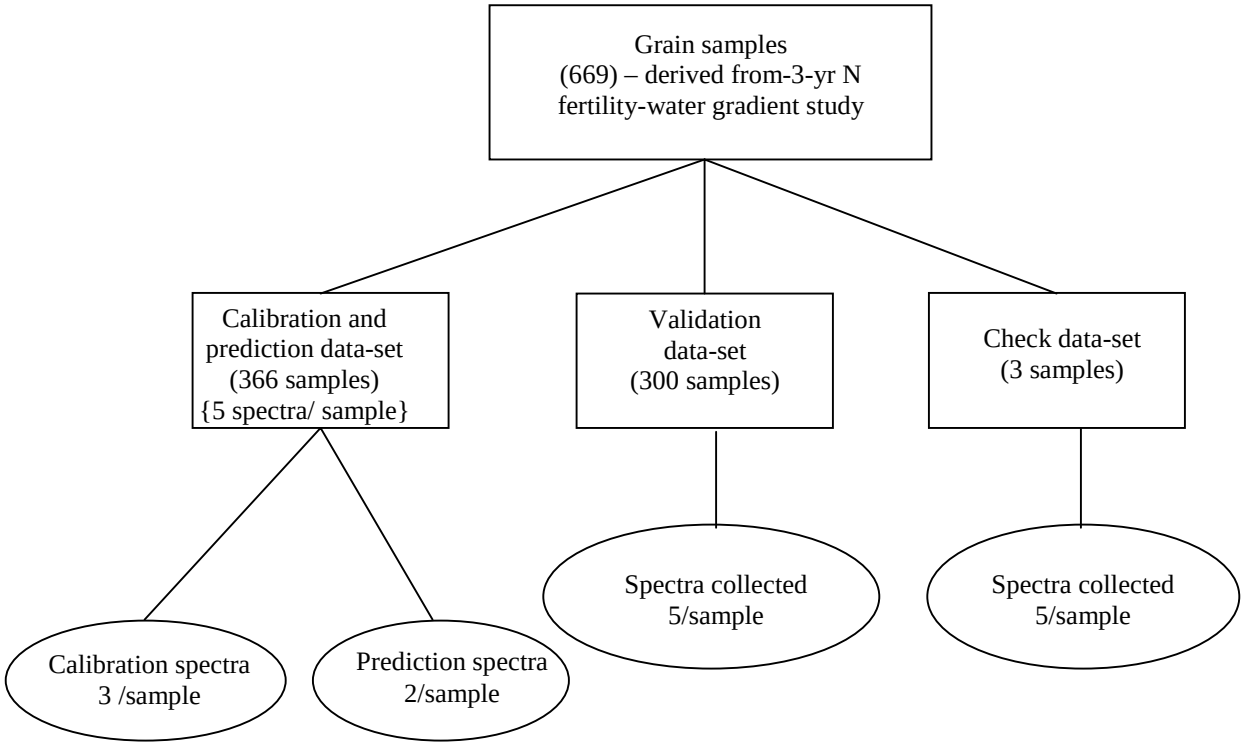


Figure 3. Distribution of samples and spectra for calibration and validation phases of this study.

Model Development

Partial least squares regression was used to develop models to predict protein from grain spectra collected with the PGA. Sample spectra were smoothed to reduce noise prior to regression analysis. That was done using a combination of binning and

Gaussian smoothing (Hrushka, 1987). Model development was conducted in two stages. During the first stage MatLab 5.3 (The MathWorks, Inc.) and PLS Toolbox 2.0 (Wise and Gallagher, 1998) were used to identify promising models and wavelengths. For the sake of expediency, a subset of 150 samples (50 from each year of the small plot studies) was used from the calibration set that encompassed the full range of protein values found in the whole set. MatLab, in conjunction with PLS Toolbox, generated model combinations of bins, Gaussian smoothing factors and up to 20 PLS factors. Models were generated beginning with the entire reflected spectrum (collected from 600.5nm-1080nm in 0.5nm increments), then the spectral range was progressively narrowed in 5 nm increments of wavelengths to find the best range for the model. That range was further refined, to 1 nm increments, until the most relevant range of wavelengths was found. Six promising models resulted and were recorded. Grain moisture content was automatically detected by the PGA and built into the modeling process during spectra collection.

During the second stage, using the six promising models derived in MatLab as guidelines, Delight 2.3b (DSquared Development, Inc, LaGrande, OR.) and Dementia 1.1a (DSquared Development, Inc., LaGrande, OR.) were used to develop final models specific to wavelengths of 0.5 nm, on the entire calibration set of grain samples. The desired number of PLS factors, or latent variables, in the final models was determined by examining the prediction residual sum of squares (PRESS) (Geladi and Kowalski, 1986). Partial least squared factors were added to the models until the PRESS values ceased to decrease by 0.4% protein.

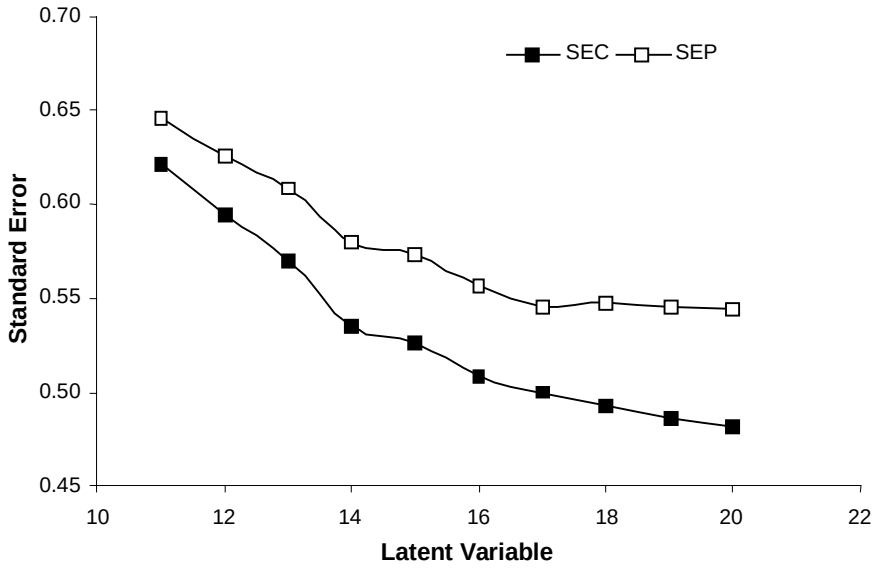


Figure 4. Standard error of calibration (SEC) and prediction (SEP) vs. number of latent variables for Model 1. The optimum latent variables are determined at the point where SEP no longer decreases. In this case, 17 latent variables are desirable.

Models were further refined using a number of latent variables equal to the best PRESS estimate plus four latent variables higher than the estimate and four less than estimate. That left nine models per each set of bin, smooth combination and wavelength range. Standard errors of calibration (SEC) and prediction (SEP) were then calculated for each model using Systat 7.0 (SPSS, Inc, 1997). SEC was calculated, based on the calibration set of spectra and SEP was calculated based on the prediction set. SEC and SEP were plotted against the number of latent variables to determine the most efficient model (Figure. 4). The most efficient model was defined at the point where SEP no longer improved with an increasing number of latent variables.

.

Results and Discussion

Model Selection

Six models to predict protein content resulted from the calibration process (Table 1). Models 1 and 2 were the best of the six, evident from the low SEC and high R^2. Model 1 was a combination of 2 bins, a Gaussian smoothing factor of 2 with 17 latent variables that performed best on wavelengths beginning at 604nm and continuing to 1075.5nm in 0.5nm increments. Model 2 did not involve the use of any binning and smoothing so included more noise. It too operated with 17 latent variables and on similar wavelengths.

Table 1. Summary statistics for calibration models developed for the PGA.

	Calibration		**Gaussian Smoothing Factors**		**Light Spectrum Range**
Model	**SEC†**	**R^2**	**Bin**	**Smooth**	**nm**
1	0.500	0.958	2	2	604 – 1075.5
2	0.399	0.973	0	0	603-1073.5
3	0.517	0.955	2	2	601-1024.5
4	0.530	0.953	2	4	600.75-1079.75
5	0.531	0.953	4	2	611.5-1075
6	0.564	0.947	6	2	626-1070

†SEC-Standard Error of Calibration

Upon examination of the calibration statistics for the two models, Model 2 fit the data best, with a higher R^2 and lower SEC, suggesting that smoothing the data not only removed noise, but useful data as well. Model performance was evaluated further on the prediction and validation sets. Summary statistics for evaluating the prediction set found in Table 2 include: 1) coefficient of determination (R^2); 2) standard error of prediction (SEP); 3) standard deviation of reference protein values (SD_x), determined by LECO

analysis; 4) bias, defined as the mean difference between reference and model (PGA) protein values; 5) absolute bias, defined as the mean of absolute difference between reference and model generated values; 6) ratio of performance to deviation (RPD), defined as the ratio of SD_x to SEP, based on PGA models (Williams, 1987).

Table 2. Summary statistics of calibration and prediction for the PGA.

	Calibration		Prediction					
Model	**R^2**	**SEC†**	**R^2**	**SEP‡**	**SD_x§**	**RPD¶**	**Bias**	**Abs. Bias#**
1	0.958	0.500	0.953	0.531	2.442	4.586	0.041	0.424
2	0.973	0.399	0.952	0.536	2.442	4.552	0.038	0.427

†SEC-Standard Error of Calibration ‡SEP-Standard Error of Prediction
§SD_x-Standard Deviation of Reference Values ¶RPD=SD_x/SEP
#Abs. Bias.-Absolute Bias

SEP and R^2 values were derived from a linear regression and SD_x was calculated on protein values found for the reference samples. RPD was calculated as a dimensionless ratio used to express the ability of a model to produce a given value beyond what can be attributed to chance. A value of 1 would indicate a lack of modeling power. A value in excess of 2.5 would be indicative of the model being suitable for screening and breeding programs and values greater than 5 suggest the model would be acceptable for quality control (Williams, 1987). Bias expresses accuracy, but can theoretically be 0 if residuals fall equally above and below 0, so the absolute bias is more indicative of the magnitude of residuals produced by the model.

Models 1 and 2 performed equally well on the prediction set of spectra as indicated by similar SEP and R^2 values (Figure 5 and Table 2). Both models explained a

high level of protein variability. RPD values for both models were similar, falling just short of 5.0. Biases from both models were low at 0.04 suggesting that the models over-predicted for as many samples as they under-predicted. Using the values for absolute bias, both models can be said to be accurate to approximately 0.4% protein, or to 0.5% protein by SEP.

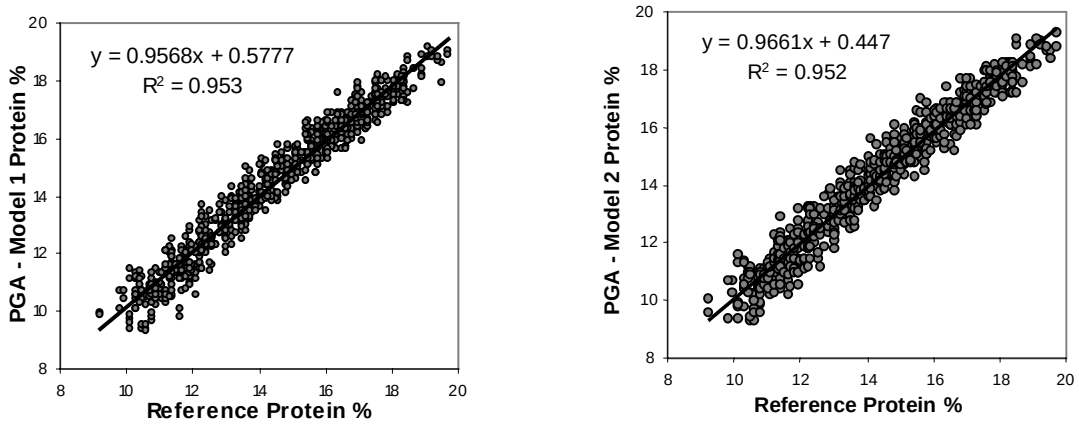


Figure 5. PGA protein (y-axis) vs. reference protein (x-axis) relationships for prediction spectra data-set.

Further analysis was conducted on a separate set of validation samples and spectra. Summary statistics of validation are found in Table 3. SEP and R^2 values were derived from a linear regression of PGA protein vs. reference protein for the same samples. Unlike the prediction statistics, validation statistics demonstrate the predictive abilities of the two models. The models performed similarly with standard errors of validation (SEV) near 1.06, R^2s of 0.81 and RPD of 2.3 (Figure 6 and Table 3).

Validation statistics indicate the performance of both models degraded substantially when compared to the prediction statistics. Bias values for both models

increased by close to 1% protein, indicating that the PGA was under-predicting. Finally, examination of the values for the absolute bias, of 0.906 for Model 1 and 0.893 for Model 2, indicates that the accuracy of both models decreased and that Model 2 performs slightly better, in a predictive capacity.

Table 3. Summary statistics of validation for the PGA.

Model	R^2	SEV†	SD_x‡	Validation RPD§	Bias	Abs. Bias¶	W/I Err.#
1	0.813	1.056	2.441	2.312	0.142	0.906	0.729
2	0.812	1.059	2.441	2.305	0.121	0.893	0.850

†SEV-Standard Error of Validation Values
‡SD_x-Standard Deviation of Reference
§RPD=SD_x/SEP
¶Abs. Bias.-Absolute Bias
#W/I Err.-Within Sample Error

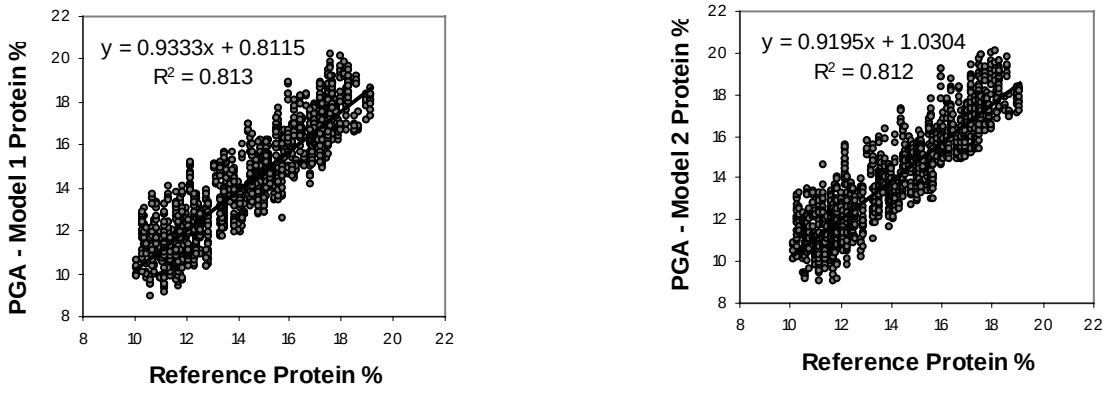


Figure 6. PGA protein (y-axis) vs. reference protein (x-axis) relationships for validation spectra data-set.

Precision

Two methods were used to assess the precision, or reproducibility of the models. One was by examining the range of predicted values among the five spectra collected per sample and expressed as within sample error. Within sample error is the mean range of predicted protein values obtained for the five spectra per sample. The second was by examining the residuals (actual – predicted protein values) of the check samples for trends of consistent over-predicting or under-predicting over time. The standard error of a single test (SET) (Williams, 1987), expresses the contribution of PGA performance to overall predictive error (SEV). The SET is the standard error found for the check samples, whereas, the SEV is the standard error for the samples composing the validation sample set.

Examination of the mean within sample error for the validation sample set indicated that Model 1 was more precise than Model 2 (Table 3). Model 2 was no more precise than it was accurate, when comparing absolute bias with average within sample error. By the same standard, Model 1 was slightly more precise, leading to the conclusion that smoothing to reduce noise was beneficial. The lower SET for Model 1 implied that it was slightly better.

The SET is used as a means for interpreting the values for SEV, presented in Table 4. The true test error (TTE) is defined by equation 1 (Williams, 1987) as an expression of the error not attributable to the reproducibility of the instrument being tested (PGA), or the instrument providing the reference values (LECO).

$$TTE = \sqrt{[(SEV)^2 - (SET_{PGA})^2 - (SET_{LECO})^2]} \quad \text{Eq. [1]}$$

The LECO is precise to a relative standard deviation of 0.3%, or an SET of 0.003 for 1% protein. This is insignificant so an SET for the LECO is not included. The TTE is a measure of procedural error. If the SEV and an SET are high, a low TTE results, indicating that there is room for improvement in sampling procedure, or other factors influencing the SEV. The high TTE relative to the SET values found in Table 4, indicate that most of the error can be attributed to model, or PGA hardware performance.

Table 4. Summary statistics of reproducibility for the PGA.

			Precision		
Model	**SEV†**	**R^2**	**SET_{pga}‡**	**SET_{Leco}§**	**TTE¶**
1	1.056	0.940	0.699	n/a	0.792
2	1.059	0.938	0.711	n/a	0.785

†SEV-Standard Error of Validation
‡SET_{pga}-Standard Error of a Single Test for the PGA
§SET_{pga}-Standard Error of a Single Test for the LECO
¶TTE-True Test Error

Because the standard errors for the validation set were much higher than those for the prediction set, the distribution of residuals from the check samples was examined for temporal trends, to attempt to identify any decline or sudden shift in sensor performance during spectra collection of the validation set (Figure 7). Residuals ranged from –2 to +2% protein and were not consistent over time. Residuals less than zero (PGA over-predicted protein values) tended to occur, almost exclusively on May 10 and June 19, suggesting swings in PGA performance. Such swings contribute to the SET, so a portion of the poor accuracy of the PGA (high SEV) is attributed to a lack of precision of the

instrument, suggesting that there is a hardware problem with the instrument.

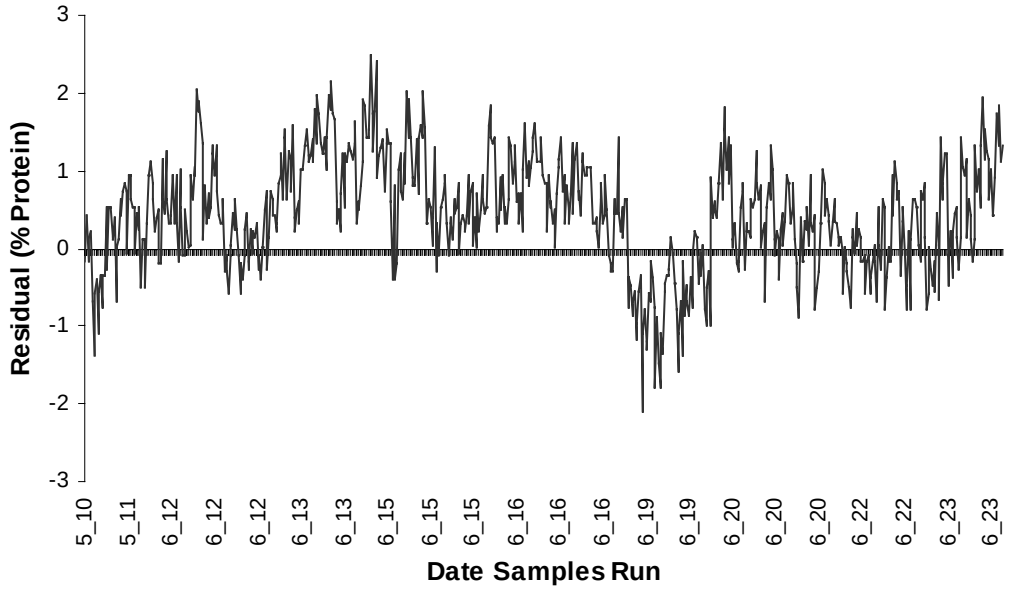


Figure 7. Temporal distribution of check sample residuals, using model 1, run during validation spectra collection. Residuals are plotted by month and date run, 2000.

Conclusions

The PGA is an experimental sensor that measures protein in whole grain kernels of wheat based on reflectance of near-infrared radiation. Two models (Models 1 and 2) for predicting the protein content of hard red spring wheat were developed for this instrument during the calibration phase of this study using partial least squares modeling procedures. Both models utilized spectra within a similar range of wavelengths (approximately 603-1074nm) to explain variability of protein in the sample. Model 1 incorporated smoothing to reduce noise, while Model 2 did not. Models 1 and 2 performed well during the calibration and prediction phases of this study, explaining

greater than 95% of protein variability. Standard error of prediction were approximately 0.5% protein, meeting manufacturer's guidelines; however, the models did not perform as well on a validation data set, explaining only 81% of the protein variability and were only accurate to 1.1% protein by standard error of validation.

Using within sample error as a measure of precision revealed that the smoothing techniques used by Model 1 increased precision, but an error of 0.7% protein was higher than the targeted accuracy of 0.5%. A lack of reproducibility could explain the low accuracy results. The true test error indicated that the high standard error of validation was due to model or sensor performance and not procedural in nature and examination of the temporal distribution of residuals suggested that sensor performed inconsistently. One possible explanation is that the internal shutter, that the sensor used for recalibrating itself, malfunctioned.

Because the PGA obtained statistics of error of double the company's target of 0.5% protein, it is not well suited to the functions of segregating grain or field mapping for precise applications of nutrients.

3. FIELD EVALUATION OF THE PROSPECTRA™ GRAIN ANALYZER

Introduction

The ProSpectra™ Grain Analyzer (PGA) is an experimental analyzer that uses near-infrared reflectance spectroscopy to measure protein, oil, and moisture content in grain crops as they are harvested (Engel et al., 1997). Textron Systems, Inc. and Case New Holland, Inc. are leading the efforts to develop and market this analyzer that is deigned to be stable in harsh environments that are encountered on combines (Von Rosenberg et al., 2000). In Montana, interest focused on the ability of the PGA to measure protein content of hard red spring wheat (*Triticum aestivum L*). This potentially has two applications. First, the analyzer could be used to segregate grain according to protein content as it is harvested (Case, 2000). Second, protein data collected by the PGA would enable production of protein maps of farm fields when integrated with GPS technology. Development of protein maps has been proposed as a surrogate for soil nitrogen (N) testing, for use in making N fertilizer recommendations (Engel et al., 1999; Long et al., 2000).

In 1999 Montana State University entered into a contract with Case Corporation to evaluate the PGA. The evaluation was conducted in two phases. The first was to develop a model for the PGA to predict the protein content of hard red spring wheat and test it in the lab. This was summarized in Chapter 1 of this thesis. The second phase, that comprises this chapter, involved testing the PGA in the field at harvest. The objectives

were to determine the accuracy and precision of protein estimates made by the PGA on a combine, contrast accuracy of estimates made by the sensor at two locations on the combine and determine the effect of contaminants on grain protein estimates made by the PGA.

Material and Methods

The field testing protocol for the PGA occurred in four steps that are outlined below:

I. Harvest

i.- Spectra from gamma and beta units are collected in the field and accuracy of both sensors is evaluated. Grab samples of wheat are collected in the field as sensors collect spectra

ii.- Spectra are collected in the lab on grab samples and accuracy of the beta unit is evaluated.

II. Contaminants

- Effects of contaminants on beta unit accuracy and precision is assessed on an independent set of wheat samples.

III. Validation

- Accuracy and precision of the beta unit is evaluated using an independent set of wheat samples. Comparisons are made between results from this validation set and the validation set used in Chapter 1.

IV. Precision

- Beta unit precision is assessed using residuals of check samples run, beginning with the validation set used in Chapter 1, through the validation set used in Chapter 2, to determine if mechanical/electronic failure of the sensor has occurred.

Instrumentation Description and Installation

Beta and gamma versions of the PGA (so named by Textron Systems) were field tested on a 1660 Case IH combine. The beta unit was identical to the one used during the calibration phase (Chapter 1). The gamma unit was added to serve as a standard to compare the beta unit performance to. The gamma was similar to the beta but incorporated additional protection against electromagnetic interference. Both units functioned similarly and were designed to record reflected spectra from a grain flow via a sensor probe. In the lab, the probe was mounted to a spinner assembly that created a grain flow, while in the field the probe for the beta and gamma units were mounted to the bubble auger and bypass unit of the clean grain elevator, respectively. During harvest, the outside ambient air temperature ranged from as high as 49^{o}C to as low as 16^{o}C. Although, the PGA was designed to operate over a temperature range of –34^{o}C to +65^{o}C (Von Rosenberg et al., 2000) efforts were made to minimize temperature extremes, and to provide an environment similar to the lab. Therefore, the base units of both sensors were mounted on the back wall, inside the combine cab. By maintaining the operating environment in the cab near room temperature (~22^{o}C) and away from exposure to direct

sunlight, the intent was to eliminate the possibility of declining performance due to overheating. Therefore, if a problem with sensor performance was identified, temperature related problems with the base units could be discounted. Each PGA was connected to a PII, 233MHZ laptop computer for collecting spectra. D2Tsc 3.0a software (DSquared Development Inc., LaGrande, OR.) was used to collect and record raw spectra.

Field Site Description

Field tests of beta and gamma PGAs were conducted at two dryland fields seeded to McNeal spring wheat (Field 1 = 40.5 and Field 2 = 25.3 ha). The sites were located in Philips County, Montana and harvested in August 2000. The soils series at both field sites were fine, montmorillonitic, frigid Aridic Argiborolls. The fields were part of a larger study that included uniform vs. variable N application rate strips, ranging from 0 to 90 kg N /ha. This resulted in a broad range of protein values (Figure 8).

Grain and Spectra Sample Collection in the Field and Analyses

Collecting grain spectra and grab samples in the field involved communication between two individuals, one to operate the computer software inside the combine cab and a second to collect grab samples from the bubble grain auger. This was accomplished using two-way hand held radios. A third individual operated the combine. Software supplied by Textron was programmed to record raw spectra from the sensors at

30-second intervals. All spectra were saved with a time stamp. The clocks for both laptops were synchronized to each other and to a GPS receiver (Trimble Ag 132), using GPS time for the purposes of cross referencing sensor results and geo-referencing each spectra. At the moment the two sensors collected spectra, a one liter grab sample was collected from the end of the bubble auger. Grab samples were placed in plastic Ziplock bags, labeled and stored for protein analysis. A total of 914 and 293 samples were collected in Field 1 and Field 2, respectively. When the sensors acquired spectra, the position of the combine was recorded using an AgNavigator (Ashtech Inc., Sunnyvale, CA) with positional data supplied by the GPS receiver and an OmniStar 3000 differential signal receiver. The GPS time-stamped spectra were geo-referenced using the positions recorded in the AgNavigator.

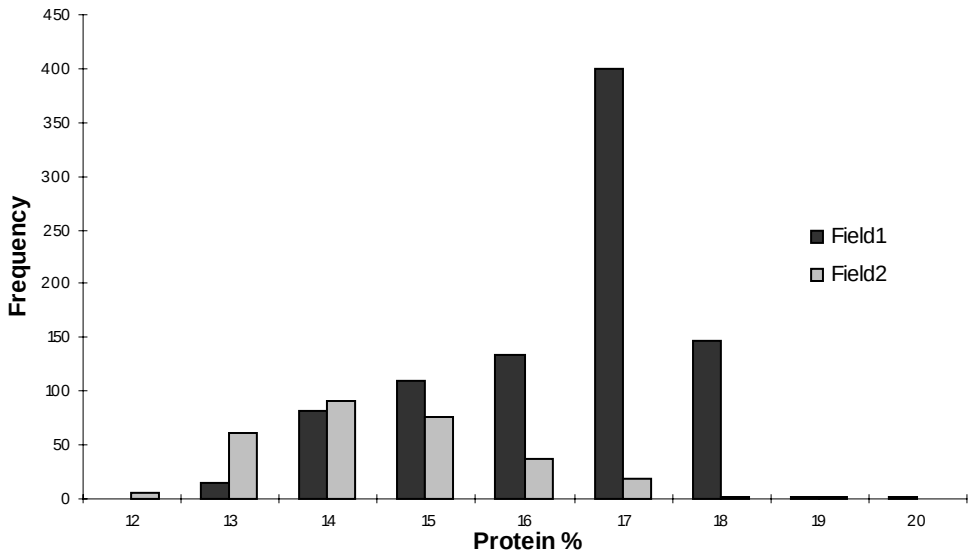


Figure 8. Distribution of protein values in Field 1 (914 samples) and Field 2 (293 samples).

Grain grab samples were analyzed for protein content by the Cereal Quality Lab located at Montana State University (MSU) using an InfraTec grain analyzer. Grain protein values reported here were corrected to 12% moisture content.

Where the DT2sc software automatically transformed raw spectra to sample spectra during the calibration and validation processes, proprietary code, in MatLab 5.3 (The MathWorks, Inc., 1999) transformed raw spectra collected in the field into sample spectra after harvest. Models 1 and 2, developed in Chapter 1, were applied to the spectra collected in this chapter for analysis.

Post-harvest Lab Tests on Beta PGA

Following harvest, the beta unit was returned to the lab and reconnected to the spinner described in Chapter 1. Spectra were recorded for the grab samples collected during harvest. Instrumentation and data collection protocols were similar to those outlined in Chapter 1; however, only one spectra per sample was recorded and samples were not cleaned to be consistent with the field protocol.

To assess if contaminants in grain affected performance of the PGA, 33 grab samples were spiked with chaff and foreign material (including dust, insects and weed seeds) at rates of 2.5, 5.0, 7.5, 10.0, and 15.0% of the grain weight. A preliminary investigation established that contaminant levels in combined grain frequently range from 0.9 to 9.5%, hence the rates chosen bracketed this contaminant range. The spectra collection protocol was the same as that outlined for the validation set in Chapter 1. Five

spectra per sample were collected and a portion of the clean grain samples was used to determine check sample protein values.

To assess if the sensor was adversely affected by field conditions, the beta sensor was tested in the lab following harvest. To do this, an independent set of 194 cleaned samples of hard red spring wheat were tested by the beta unit in the lab. Protein distribution of the samples is presented in Figure 9. One measure of accuracy was bias, calculated by taking the mean of the differences between Infratec protein and model protein values. Absolute bias was also used because bias can theoretically be 0 if residuals fall equally above and below 0. The absolute bias was more indicative of the magnitude of residuals produced by the sensor and was calculated by taking the mean of the absolute differences between Infratec protein and model protein values.

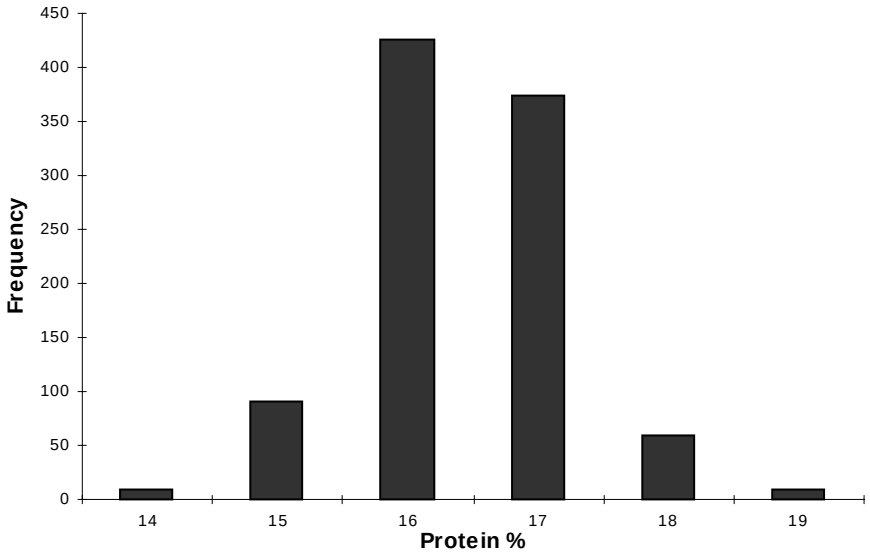


Figure 9. Distribution of protein values in validation sample set (194 samples).

Precision of the PGA was evaluated after harvest by examining the reproducibility of protein values for three check samples (11.5, 14.4 and 18.4% protein). These were the same check samples used in Chapter 1 and their protein values were determined through LECO dry combustion (Sweeney and Rexford, 1987). Spectra were recorded on the check samples for every seventeen grain samples sensed by the beta unit, as during the calibration phase (Chapter 1). Five spectra per sample were collected. Check sample residuals (reference – predicted protein values) were inspected over time for irregularities in sensor performance. Within sample error, measured as the mean range of protein values of five spectra per grain sample, was also calculated as a measure of precision. Statistical analyses were performed using Systat 7.0 (SPSS, Inc., 1997).

Results and Discussion

Sensor Accuracy in the Field

Beta and gamma PGA performance was assessed in the field and following harvest. Lab calibration models were applied to the sample spectra collected in the field during the grain harvest. Predicted protein values were then compared against protein levels of the grab samples provided by the MSU-Cereal Quality Lab. Following harvest, grab samples were run the through the PGA spinner assembly in the lab and new spectra collected. Calibration models were applied to those sample spectra and predicted protein levels were compared to MSU-Cereal Quality Lab values. Results are reported by day and field to illustrate temporal and location trends.

The beta unit performed poorly in the field compared to the lab, as evidenced by the validation statistics (Table 5). Field statistics showed R^2 values lower than 0.1, standard errors consistently higher than 1.0 and RPD values of approximately 1. A large disparity in protein predictions between the two models exists. Predicted protein values for Model 1 were 12 to 14% above the reference. Predicted values for Model 2 displayed a negative bias, although closer to the 0.1% bias and 0.9% absolute bias found for the validation sample set in Chapter 1 (Table 3). It can be assumed that the noise Model 1 attempted to correct for was, in fact, useful spectral data. Day to day differences in R^2 and SEp values were noticeable (Figures 10 and 11). For example, both models estimated protein lower for a portion of Field 1, harvested on August 17, than for preceding days. In fact, where some values were 2 to 4% above the reference in Field 1, on August 17, they were 10 to 16 percent above the reference on other days. Day to day variability suggested differences in the physical performance of the sensor. The range of predicted protein for any given value of reference protein was evidence that the models, or sensors were not performing as well in the field as the beta unit had in the lab.

Statistics for the gamma unit indicate it performed better than the beta at harvest (Table 6); however, it did not perform as well as the beta unit in the lab (Table 3, Chapter 1). Statistics show R^2 values lower than 0.3, standard errors close to 1 and RPD values ranging between 1 and 1.3. With exception of the standard errors, they indicate the gamma did not perform well at harvest. Unlike the beta unit, Models 1 and 2 performed similarly. Both models had biases of approximately 2-3%, similar to the absolute bias results, indicating that the models under predicted actual protein. Bias for both models

was approximately double those for the validation statistics for the beta (Table 6). Noise correction included in Model 1 did not appear to have any noticeable effect on field sample results. Day to day differences in R^2 and SEp values were noticeable, reinforcing the beta unit results.

Table 5. Summary statistics for beta unit performance in the field.

Date	Field	Performance (Field) R^2	SEp†	SD_x‡	RPD§	Bias	Absolute Bias
----------Model 1----------							
8-15	1	0.002	1.204	1.202	0.998	-12.0	12.0
8-16	1	0.004	1.284	1.285	1.000	-12.5	12.5
8-17	1	0.057	1.028	1.055	1.026	-12.4	12.4
8-17	2	0.028	1.244	1.259	1.012	-14.3	14.3
----------Model 2----------							
8-15	1	0.007	1.200	1.202	1.002	-2.5	2.9
8-16	1	0.054	1.251	1.285	1.027	-1.5	1.8
8-17	1	0.037	1.038	1.055	1.016	-1.3	2.0
8-17	2	0.006	1.257	1.259	1.002	-3.3	3.3

Model 1-B2S2 lv17 range=604-1075.5 nm
Model 2-B0S0 lv17 range=603-1073.5 nm
†SEp-Standard Error of Prediction §RPD=SD_x/SEp
‡SD_x-Standard Deviation of Reference Values

The predictive ability of the gamma unit using Model 1 and 2 is illustrated in Figures 12 and 13. The distribution of predicted protein values is similar and both models predicted low values for many samples below 12% and over-predicted for many values greater than 12% protein. Day to day variations in performance were not as apparent as noted for the beta unit, but a difference in performance from Field 1 to Field 2 on August 17 is apparent (Figures 12 and 13) with a broader range of protein values predicted in Field 2. The more consistent predictive ability of the gamma suggests that the hardware problem found to exist in the beta is not present in the gamma. The

difference in predicted values found on August 17 suggests that some variability in harvest samples has a negative effect on model performance.

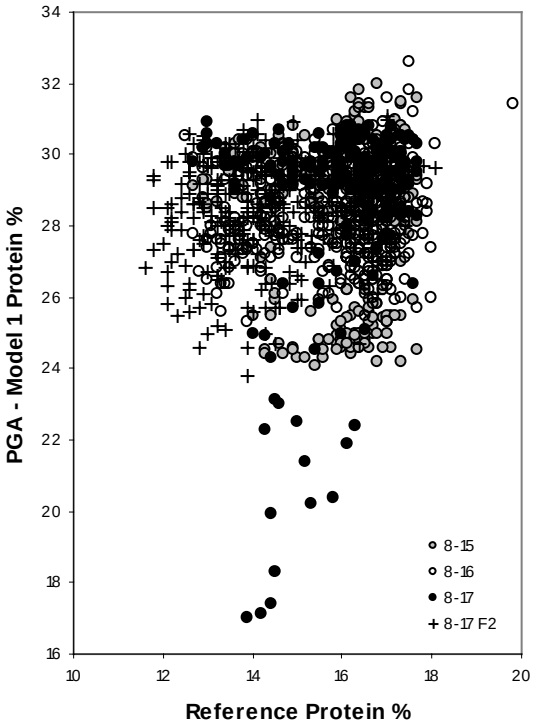


Figure 10. PGA (beta unit) predicted protein using calibration Model 1 in the field vs. reference protein.

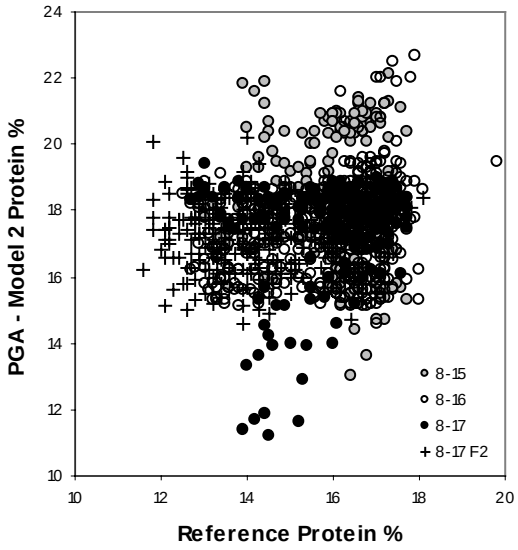


Figure 11. PGA (beta unit) predicted protein using calibration Model 2 in the field vs. reference protein.

Table 6. Summary statistics for gamma unit performance in the field.

Date of harvest	Field	Performance (Field) R^2	SEp†	SD_x‡	RPD§	Bias	Absolute Bias
----------Model 1----------							
8-15	1	0.197	1.079	1.202	1.114	2.3	2.5
8-16	1	0.305	1.072	1.285	1.199	2.8	2.9
8-17	1	0.234	0.923	1.055	1.143	2.0	2.2
8-17	2	0.234	1.106	1.259	1.138	2.1	2.4
----------Model 2----------							
8-15	1	0.134	1.121	1.202	1.072	2.5	2.6
8-16	1	0.339	1.045	1.285	1.230	2.0	2.2
8-17	1	0.237	0.921	1.055	1.145	1.9	2.0
8-17	2	0.212	1.122	1.259	1.122	1.8	2.1

Model 1-B2S2 lv17 range=604-1075.5 nm
Model 2-B0S0 lv17 range=603-1073.5 nm
†SEp-Standard Error of Prediction §RPD=SD_x/SEp
‡SD_x-Standard Deviation of Reference Values

Post Harvest Evaluation

Protein estimates by the beta unit were more accurate in the lab, than in field. R^2 values ranging from 0.4 to 0.8 for the beta and gamma units in the lab were higher (Table 7) and SEp values of 0.6 to 0.9 were lower than for both units in the field. RPD values increased slightly and bias values ranged from 1 to –2.2, indicating the models over-predicted protein. When compared to validation statistics in Table 3, Chapter 1, the bias became negative, the absolute bias doubled, R^2 values decreased and RPD values decreased, all indicating a decrease in performance, possibly attributable to the presence of chaff and other foreign material in the field samples.

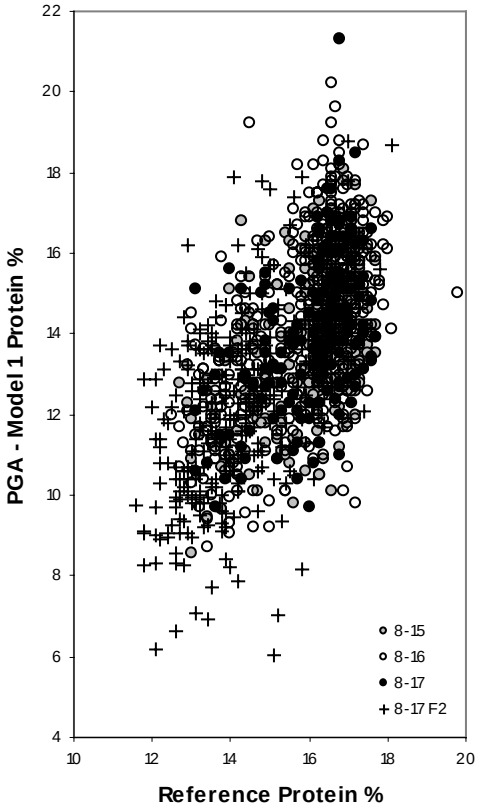


Figure 12. PGA (gamma unit) predicted protein using calibration Model 1 in the field vs. reference protein.

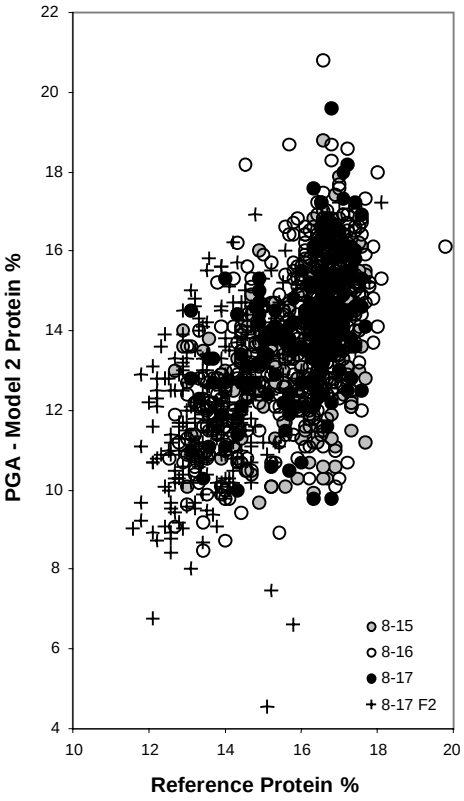


Figure 13. PGA (gamma unit) predicted protein using calibration Model 2 in the field vs. reference protein.

SEp values were lower than SEV values, possibly due to the narrower range of protein of the field samples, indicated by a lower SD_x and lower RPD values. Examination of Figure 14 revealed fluctuations in performance, in the form of two data clouds of predicted results. Since these patterns were consistent between models and each field and harvest date are present in each cloud, this fluctuating performance was likely due to inconsistent sensor performance. A line representing a perfect relationship of predicted to reference values has one cloud of data surrounding it. A second cloud of

data is present immediately above the first, indicating that the sensor consistently over-predicted for some spectra.

Table 7. Summary statistics for beta unit performance on field samples in the lab

Date of harvest	Field	Performance (Lab) R^2	SEp†	SD_x‡	RPD§	Bias	Absolute Bias
----------Model 1----------							
8-15	1	0.771	0.576	1.202	2.087	-1.7	1.8
8-16	1	0.539	0.879	1.285	1.462	-1.9	2.0
8-17	1	0.421	0.858	1.055	1.223	-0.9	1.1
8-17	2	0.577	0.822	1.259	1.532	-1.7	1.8
----------Model 2----------							
8-15	1	0.737	0.617	1.202	1.948	-1.7	1.7
8-16	1	0.587	0.833	1.285	1.543	-2.2	2.2
8-17	1	0.495	0.802	1.055	1.315	1.0	1.1
8-17	2	0.537	0.860	1.259	1.464	-1.7	1.8

Model 1-B2S2 lv17 range=604-1075.5 nm
Model 2-B0S0 lv17 range=603-1073.5 nm
†SEp-Standard Error of Prediction §RPD=SD_x/SEp
‡SD_x-Standard Deviation of Reference Values

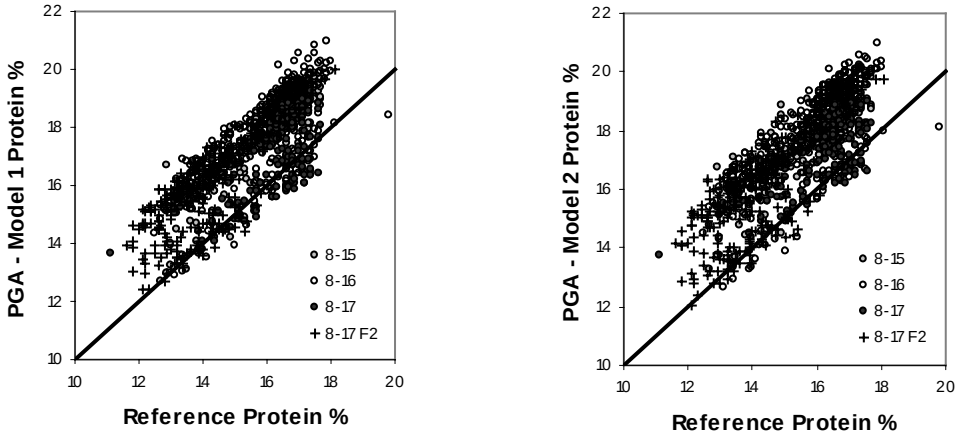


Figure 14. Beta unit lab collected spectra on field samples. PGA protein vs reference protein.

Contaminant Effects on Model Performance

As expected, the presence of contaminants in the grain samples decreased the predictive ability of Models 1 and 2 (Table 8). R^2 and RPD values decreased and SEp and bias values increased as contamination increased, leading to the conclusion that chaff/foreign matter negatively affect accuracy. Model 1 outperformed Model 2 on clean grain, however, Model 1 was more sensitive to the effects of increasing contaminant levels than Model 2 as evidenced by higher values for SEp and lower R^2.

In the model building process, the noise that Model 1 eliminated through binning and smoothing was actually useful data when contaminants were present. Within sample error increased for Model 1 at the 2.5% level of contamination and at 5% for Model 2, then held constant for both models with increasing contaminant levels. This higher threshold for contaminants makes Model 2 more desirable for use in the field.

Post Harvest Validation for Accuracy

The beta PGA was tested on a set of validation samples to determine if the poor performance of the sensor in the field continued after harvest. If not, it could be assumed that the calibration conditions did not accurately imitate field conditions due to the flow characteristics of the grain past the sensor probe, such as error induced by a non-constant packing density (Tkachuk, 1987). However, if poor sensor performance did continue after harvest, that would be indicative of engineering flaws manifesting themselves under the harsher operating environment of field conditions. Statistical results of model

performance on the validation set are summarized in Table 9. The R^2 values for both models were much lower than observed in the lab. RPD values were much lower also, and absolute bias increased, indicating a decline in sensor performance, illustrated in Figure 15. The lower SEV for the post harvest validation set, versus the pre-harvest validation set, is due to the narrower range of reference protein values, indicated by a lower SD_x.

Table 8. Summary statistics of beta unit performance with chaff/foreign materials added to clean samples.

	Performance (Lab)						
Chaff %	**R^2**	**SEp†**	**SD_x‡**	**RPD§**	**Bias**	**Absolute Bias**	**W/I Err.¶**
----------Model 1----------							
0	0.808	0.573	1.304	2.276	-1.808	1.808	0.709
2.5	0.706	0.709	1.304	1.839	-3.359	3.359	0.872
5	0.572	0.856	1.304	1.523	-3.297	3.305	0.900
7.5	0.538	0.890	1.304	1.465	-3.561	3.561	0.888
10	0.508	0.918	1.304	1.420	-3.589	3.589	0.909
15	0.418	0.998	1.304	1.307	-3.819	3.819	0.884
----------Model 2----------							
0	0.786	0.605	1.304	2.155	-2.486	2.486	0.763
2.5	0.771	0.627	1.304	2.080	-3.491	3.491	0.741
5	0.730	0.680	1.304	1.918	-3.536	3.536	0.906
7.5	0.703	0.712	1.304	1.831	-3.855	3.855	0.881
10	0.721	0.692	1.304	1.884	-3.949	3.949	0.841
15	0.682	0.737	1.304	1.769	-4.303	4.303	0.878

Model 1-B2S2 lv17 range=604-1075.5 nm
Model 2-B0S0 lv17 range=603-1073.5 nm
†SEp-Standard Error of Prediction ‡SD_x-Standard Deviation of Reference Values
§RPD=SD_x/Sep ¶W/I Err.-Within Sample Error

Table 9. Summary statistics of validation for the beta unit after harvest.

		Validation					
Model	**R^2**	**SEV†**	**SD_x‡**	**RPD§**	**Bias**	**Absolute Bias**	**W/I Err.¶**
1	0.149	0.693	0.751	1.084	-0.8	1.5	0.743
2	0.197	0.674	0.751	1.114	-1.1	1.4	0.882

Model 1-B2S2 lv17 range=604-1075.5 nm
Model 2-B0S0 lv17 range=603-1073.5 nm
†SEV-Standard Error of Validation ‡SD_x-Standard Deviation of Reference Values
§RPD=SD_x/Sep ¶W/I Err.-Within Sample Error

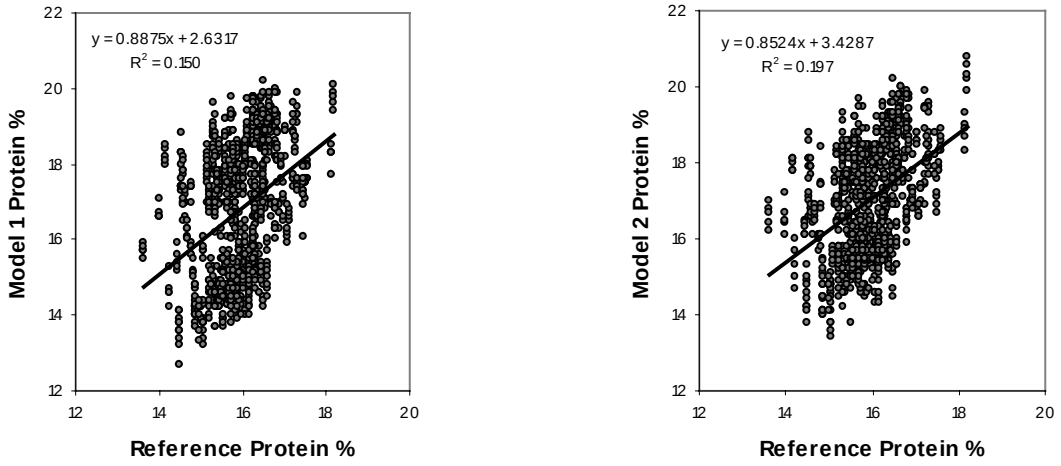


Figure 15. Post-harvest validation spectra set. PGA protein vs. reference protein.

Bias values shifted from positive to negative values, indicating that the sensor began over-predicting protein after harvest. Examination of residuals for the same set of spectra exhibited that most predicted protein values exceeded the actual (Figure 16), indicative of change in hardware performance.

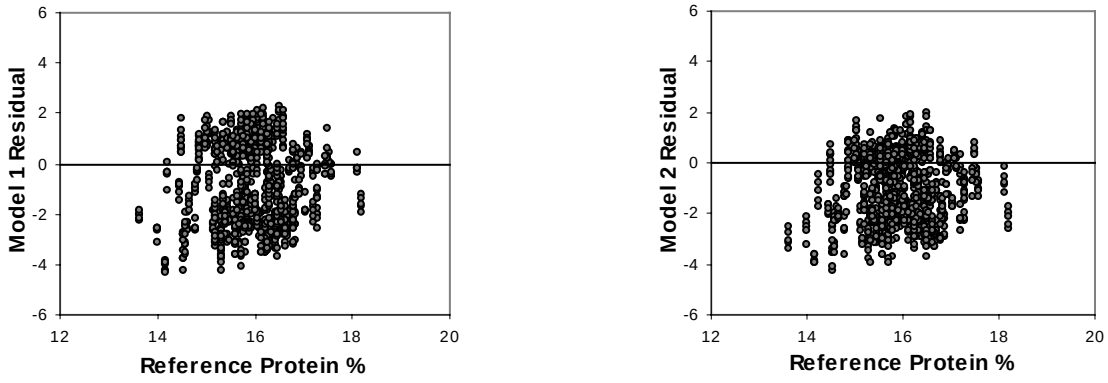


Figure 16. Residuals of post-harvest validation spectra set. Residuals vs. reference protein.

Precision of the instrument was assessed by examining the within sample error (Table 9). Values of 0.743 and 0.882 for Models 1 and 2 respectively were similar to those of 0.729 and 0.850 found in the pre-harvest validation set, presented in Table 3, Chapter 1. This indicates that reproducibility of the instrument is unchanged.

Post Harvest Evaluation of Precision

The temporal distribution plot of check sample residuals for Model 1 indicated episodic shifts in values occurred on October 17, 29 and 30 (Figure 17). The sudden shift in residuals suggests that a problem exists with the sensor and explains the degradation in sensor performance from the initial validation set to the last. The day-to-day inconsistent performance of the sensor could also explain some of the inaccuracy noted in the field. Using criteria to establish a True Test Error (TTE) for the check samples, from Chapter 1 (equation 1) was not possible because the SETs for the check samples, of 1.438 for Model 1 and 1.410 for Model 2, were larger than the SEVs for the models (Table 9).

This indicates that error was associated with the instrument and not lab procedure. In the initial validation phase, the SET from check samples was approximately 0.7 (Table 4, Chapter 1); whereas, it was approximately 1.4 for check samples run during the last validation phase.

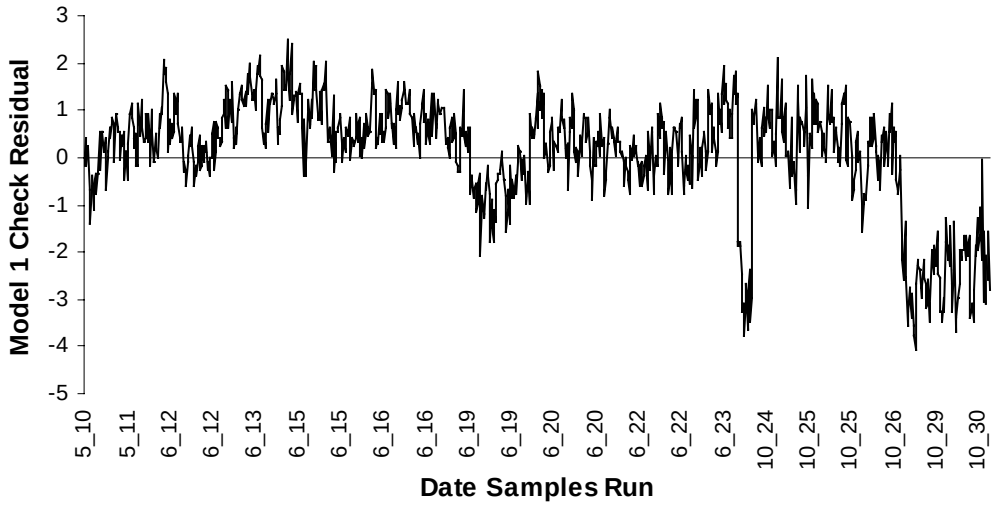


Figure 17. Temporal distribution of check sample residuals (run in lab), using Model 1 from calibration through final validation phases. Residuals are plotted by month and date run, 2000. The sensor was on the combine between June and October.

Conclusions

Two protein prediction models for the PGA were developed and reported in Chapter 1. During the field and field related evaluation phase, reported in this chapter, these same models were tested on beta and gamma versions of the PGA. The results of these investigations revealed that the PGA was not able to accurately predict grain protein within the specifications (<0.5%) established by Case New Holland, Inc. and Textron Systems, Inc. Beta and gamma PGA performance in the field was poor

compared to its performance during the lab calibration phase with R^2 values below 0.1 and SEp values ranging from 1 to 1.3. Biases ranged from –12 to –14 for Model 1 and –1 to –3% protein for Model 2 and RPD values were lower with values of 1 compared to 2.3 for the validation set in Chapter 1.

The apparent degradation in performance in the field may be due to differences in grain presentation from the lab to the combine. Modeling results could be improved by redesigning the lab based spinner assembly to more closely approximate presentation conditions found on the combine. Likewise, the manner of grain presentation on the combine could be engineered to hold a grain sample stationary in a cell long enough to acquire spectra before releasing it back into the combine. Such conditions would be easier to replicate in the lab than the existing continuous grain flow conditions.

Examination of the check sample residuals indicated that the hardware problem identified in Chapter 1 worsened after harvest. More field testing of the sensor is required to isolate those engineering aspects that perform poorly under normal harvest conditions

Post harvest evaluation of the beta sensor revealed that chaff and other foreign material encountered during harvest negatively affect the predictive capabilities of the models and that a hardware problem provided inconsistent results. Overall, Model 2, predicting protein without inclusion of binning and smoothing corrections for noise, provided the best results. Future work should focus on isolating and negating, or compensating for, sources of noise ranging from electronic interface to sample contaminants. Additionally, multiple sensors should be tested to assess how

representative tests are of the product as a whole and to how well models transfer from one sensor to another.

Because the PGA obtained statistics of error of double the company's target of 0.5% protein, it was not well suited to the functions of segregating grain or field mapping for precise applications of nutrients, but could be practical for directing soil sampling efforts if the hardware issues can be resolved.

4. SPATIAL VARIABILITY OF WHEAT YIELD AND PROTEIN IN RELATION TO TOPOGRAPHY UNDER SEMIARID CONDITIONS

Introduction

Hard Red Wheat (HRW) (*Triticum aestivum* L.) is an important cereal crop in the northern Great Plains (USDA, 2001). Because of its high protein content, millers and bakers blend HRW with lower quality wheat to increase the gluten content of the resulting flour. To ensure a consistent supply of high protein wheat, growers of HRW are offered protein premiums that serve as a financial incentive for maximizing the protein concentration of their grain. For example, the average value of Hard Red Spring Wheat in the Pacific Northwest Market for the 10-yr period (1988-97) was \$3.30 bu^{-1} at 13% protein vs. \$4.00 bu^{-1} at 15% protein (data supplied by Montana Wheat and Barley Committee).

Nitrogen (N) fertilization is the most important cropping input that affects grain protein levels in dryland wheat production systems. Growers may apply N to increase protein levels to capture premiums, or avoid discounts, based on this price-determining factor. Because farm fields are known to possess spatial variability in soil fertility and other factors that influence grain protein concentration, spatial variability in these factors should also mean that large differences in grain protein exist spatially across field landscapes. Therefore, the potential exists to improve wheat protein levels by applying N fertilizers in accordance with spatial patterns in soil fertility.

The site-specific management practice of variable-rate N application offers the opportunity to control grain protein levels that vary spatially within landscapes resulting from site-to-site differences in plant available moisture and soil N nutrition levels. To be effective, however, detailed geographic information about these soil-controlling factors is needed at the scale of fertilizer application. Once the relationships between crop response and soil or other terrain factors have been established, they can be used in determining field management zones for precision N application and other site-specific management purposes.

Obtaining this information by means of grid soil sampling (Mulla et al., 1992) is cost prohibitive in the Northern Plains where dollar returns are usually <$150 ac^{-1} (Engel et al., 1999). Soil surveys and topographic maps published at the 1:24,000-scale are readily available for this purpose (Carr et al., 1991; Karlen et al., 1990; and Ciha, 1984), but important soil and landscape differences may not be retrieved in sufficient detail (Long et al., 2000). Furthermore, yield maps derived from on-combine sensors provide a means to identify field areas where crop production was limited, but do not reveal the biophysical factors that limited production.

Using topography to capture grain yield and protein variability can provide a basis for delineating management zones needed for directing soil sampling efforts and variable-rate N applications. Topographic variation at spatial scales of <10-m may be characterized by taking measurements along transects over farm fields using the Global Positioning System (GPS) and survey-grade GPS receivers to collect an irregular network of easting, northing, and elevation points. These point coordinates then can be

interpolated in a geographic information system (GIS) to produce a digital elevation model (DEM) at fine grid point spacing. In addition, terrain-modeling software is commercially available for deriving slope steepness, slope orientation, isolation, plan and profile curvature, hillslope distribution, and other terrain attributes from a DEM.

One rationale for using terrain attributes as a basis for management zones is that topography influences soil development and soil fertility-related attributes. Under semiarid conditions, shallow, weakly developed soils tend develop on convex, upper slope positions whereas deep, well-developed soils develop on concave, lower slope positions (Gessler et al, 2000; Aandahl, 1948). In addition, soil and tillage erosion results in transport of soil materials from upper slopes and deposition in lower slopes. Consequently, soils in upper slopes are shallower, and more coarsely textured than in lower slopes, which are deeper, and more finely textured (Stewart et al., 2002; Moore et al., 1993; Hamblin et al., 1988; Aandahl, 1948; Machado et al., 2002; Moulin et al., 1994; Afyuni et al., 1993; Papendick et al., 1971). Plant available water also tends to be less in the upper-slope positions because of coarser textured soils that have less storage capacity (Stewart et al., 2002; Fiez et al., 1994; Nolan et al., 2000) and soil erosion of upper slopes is often severe enough to expose an underlying, root restricting calcic horizon (Pennock et al., 1987; Moulin et al., 1994).

Deep soils retain more moisture that plants utilize later in the growing season, which boosts grain yields (Ciha, 1984; Smika and Greb, 1973). Deeper, downslope soils also collect more moisture over winter and through the growing season by means of overland- and subsurface-flow from upslope positions (Ciha, 1984; Hanna et al., 1982;

Pennock et al., 1987). By influencing the plant's ability to uptake N, this redistribution of soil moisture has the effect of increasing yields downslope and protein upslope (Terman et al., 1969; Rawls et al., 1982; Walker et al., 1968). Higher protein values are found on water shedding, upper slopes because N in the plant is not diluted by a larger vegetative biomass as much as it is on lower slopes where yields and biomass are greater (Terman et al., 1969; Fiez et al., 1994). Locations of these deeper, wetter soils in lower-slopes, protected from wind and evaporation, also result in higher yields and lower protein values compared to upslope, summit positions (Aandahl, 1948; Moore et al., 1993). Greater evaporation and soil temperatures in summit positions, due to less organic cover, may also contribute to lower soil water and hence, higher grain protein concentration, especially under conventional vs. conservation tillage management (Smika and Greb, 1973).

Much research has been directed to investigating relationships between topography and yield or protein along linear transects in which topography was characterized by manually delineated hillslope classes (Ciha, 1984; Afyuni et al., 1993; Aandahl, 1948), or computer generated classes of slope, plan and profile curvature (Burt and Butcher, 1985; Pennock et al., 1987; Nolan et al., 1999; Nolan et al., 2000; Kravchenko and Bullock, 2002). These methods attempted to explain the spatial variability of yield and protein as a response to soil moisture distribution. Newer computer modeling techniques allow for the creation of categorical hillslope position delineation (MacMillan, 2000; MacMillan et al., 2000) with allowance for analysis at broad, whole field scales of observation. In addition, they provide for computation of the

compound topographic index, which is a quantitative measure of slope and catchment area that has been used to explain moisture distribution, but on a broader scale (Pennock, et al., 1987; Gessler et al., 2000; Burt and Butcher, 1985).

An improved understanding of how topographic variability influences grain yields and grain quality in semiarid environments can lead to development of management zone determination methods for directing soil sampling efforts, variable-rate fertilizer applications, and segregation of low vs. high quality grain during machine harvest. The objectives of this study were to identify relationships between topography and grain protein or grain yield, as well as illustrate quantitative hillslope classification and terrain modeling with results from dryland wheat fields in northern Montana.

Materials and Methods

Data Collection and Site Description

Eighteen farm fields in north central Montana were selected for this study, each comprising a 12- to 100-ha portion of dryland wheat fields in Liberty, Hill, and Phillips Counties of northern Montana (Table 10). These fields, planted with Hard Red Spring Wheat, were chosen for their rolling topography in glacial till hummocky terrain and availability of baseline data of elevation, grain yield, grain protein and soil nitrogen, collected from 1997 to 2000. A crop-fallow production regime was followed for all of the fields, except for the Mattson field, which was continuously cropped.

Table 10. Study area farm and field names and locations in northern Montana.

Farm	Field	Year	Latitude	Longitude	Crop Rotation
Mattson		1997	48° 39' 50"	-110° 59' 45"	Annual Wheat
Grass	Russell	1998	48° 20' 44"	-110° 05' 39"	Fallow-Wheat
Grass	Sand	1998	48° 20' 19"	-110° 05' 25"	Fallow-Pea-Wheat
Peterson	Home	1998	48° 49' 59"	-110° 03' 33"	Fallow-Wheat
Anderson	East	1999	48° 25' 34"	-107° 35' 17"	Fallow-Wheat
Grass	Home	1999	48° 19' 54"	-110° 04' 39"	Fallow-Wheat
Grass	State-W	1999	48° 18' 35"	-110° 07' 35"	Fallow-Lentil-Wheat
Peterson	WF	1999	48° 49' 54"	-110° 05' 22"	Fallow-Wheat
Peterson	WW	1999	48° 49' 54"	-110° 04' 58"	Fallow-Wheat
Peterson	Larson	1999	48° 47' 47"	-110° 04' 20"	Fallow-Lentil-Wheat
Anderson	West	2000	48° 25' 53"	-107° 35' 29"	Wheat-Fallow
Grass	State-E	2000	48° 18' 47"	-110° 06' 27"	Fallow-Wheat
Grass	Sprague	2000	48° 21' 12"	-110° 07' 06"	Fallow-Pea-Wheat
Kaercher		2000	48° 31' 43"	-109° 55' 57"	Wheat-Fallow
Peterson	SE	2000	48° 49' 55"	-110° 04' 31"	Canola-Fallow-Wheat
Peterson	SW	2000	48° 49' 51"	-110° 04' 42"	Canola-Fallow-Wheat
Peterson	NE	2000	48° 50' 24"	-110° 04' 35"	Fallow-Wheat
Peterson	NW	2000	48° 50' 17"	-110° 04' 45"	Fallow-Wheat

Yield data were collected using farmer-owned combine harvesters equipped with either a mass flow yield sensor (Model 2000, AgLeader, Inc. Ames, IA) or a volumetric yield sensor (Model Ceres II, RDS Technology Ltd., Gloucestershire, UK) calibrated to within ±2 %. The AgLeader sensor was mounted at the top of the clean grain elevator whereas the RDS sensor was mounted in the middle of the clean grain elevator. Yield data were collected at 1-sec intervals and geo-referenced using a differential global

positioning system (DGPS) consisting of a differential signal receiver (Model 3000, Omnistar, Inc. Houston, TX) and 12 channel, L1, C/A code Ashtech AgNavigator (Ashtech Inc. Sunnyvale, CA) GPS receiver with ±1-m of positional accuracy.

Grain protein information was derived from 1-liter grab samples, collected manually from the bubble auger of the combine clean grain elevator at 30- to 60-sec intervals. As each sample was collected, the location of the combine in the field was recorded using the data-logging module of the AgNavigator. Following harvest, a Clipper Office Tester (A.T. Ferrell Company, Inc., Bluffton, IN) was used in the laboratory to clean the grain of dirt, stones, chaff and other foreign material. Near-infrared spectroscopy (Infratec Model 1226, Foss North America, Eden Prairie, MN) was then used to determine the protein concentration of each cleaned grain sample. Protein measurements were corrected to a 12% moisture basis to compensate for error introduced by the presence of moisture in the sample.

During harvest, the protein and yield observations were geo-referenced with respect to the GPS time when the grain was sampled at the top of the clean grain elevator, not when the crop was cut at the combine's header; therefore, a post-processing operation was applied to the recorded yield rates using Microsoft Excel (Microsoft Corp., Bellevue, WA). Each record was lagged along the path the combine traveled by 11 seconds, corresponding to a typical delay between when the header engages a crop and grain reaches the top of the clean grain elevator. This effectively forced the sampled yield data to conform to the geographic position where the header engaged the crop.

A pair of survey-grade GPS receivers (Model 3151R, Novatel, Inc. Calgary, AB, Canada) was used to obtain elevation measurements along a series of transects at each study site. To obtain elevation data, one of the receivers was mounted to a vehicle that traversed a finely spaced grid of points referenced by easting and northing UTM coordinates. Positional accuracies of these readings after post-processing were within ±4-cm in both horizontal and vertical planes relative to the location of the other receiver at a stationary base station. Elevation values were collected at least 30-m beyond the boundary of a field to avoid the negative influence of edge effects on data interpolation.

Terrain Modeling

Elevation point data for each field were used to construct a 10-m Digital Elevation Model (DEM) using ArcView Spatial Analyst ver. 3.2 (ESRI, Redlands, CA) in accordance with methods proposed by MacMillan (2000). A DEM was created by adding elevation point data to ArcView, creating a shapefile from it and using the Inverse Distance Weighted (IDW) function to create an interpolated DEM in an ArcView grid format. The resulting DEMs contained artifacts owing to the data collection interval, which was more intense in the direction of travel, represented as a linear pattern on the DEM (Fig. 18).

Figure 18. Plan view of Mattson field (with hill-shading) prior to smoothing of the associated Digital Elevation Model.

The image processing procedure of low pass filtering, or smoothing, was applied to remove linear artifacts. In smoothing, the elevation value at the center of a 7×7 template is computed as the average elevation level in a neighborhood defined by the template dimensions (MacMillan, 2000; Burrough and McDonnell, 1998). The result was visually examined for artifacts using a hillshade image generated with ArcView Spatial Analyst. Since elevation data were collected on north-south transects, the lighting for the hillshade was generated from a 315° azimuth. If artifacts were still visible, a 5×5-neighborhood smooth was run and another hillshade image was generated. If artifacts were again present, a second 5×5-neighborhood smooth was run and another hillshade generated to look for artifacts. In all cases, no more than two 5×5 neighborhood smooths were required in addition to the initial 7×7 smooth.

The continuous, ratio-interval values of a final, smoothed 10-m DEM (Figure 19) were grouped into 15 landscape element classes (Table 11) using the landscape modeling routines available in the LandMapR™ software package (Landmapper Environmental Solutions, Edmonton, AB.). The classification procedures used in LandMapR are based on those originally proposed by Pennock et al. (1987) and later improved by MacMillan et al. (2000). The DEMs were converted to a format that the LandMapR software could recognize, namely an ASCII file with associated metadata header. This involved the use of an intermediate conversion step, utilizing the commercial GIS software package Idrisi32 (Clark Labs, Worcester, MA).

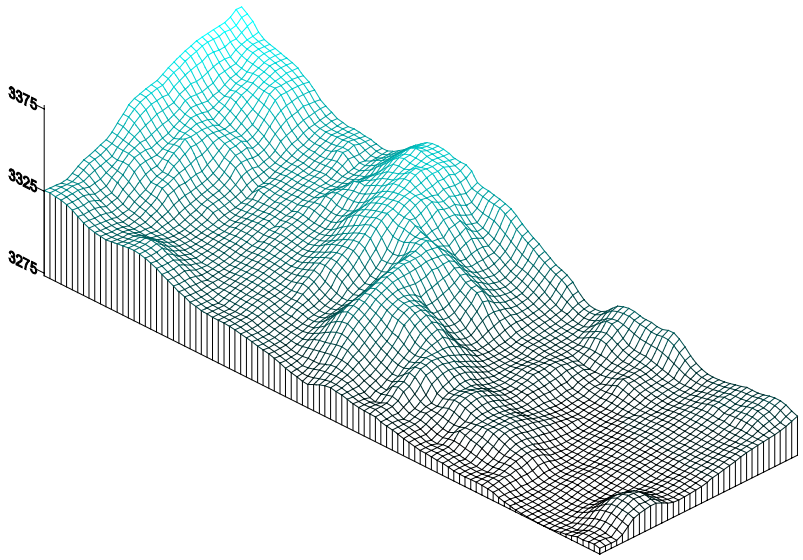


Figure 19. Block diagram of Mattson field with surface depiction of a smoothed Digital Elevation Model.

Table 11. Fifteen landscape elements created by the LandMapR software and associated slope and flow region class models for three, four and five class management zones.

Landscape Element	Landscape Class Model			
	3-slope	**4-slope**	**4-flow**	**5-slope**
1. Level crest	Upper	Upper	Flat	Interfluve
2. Divergent shoulder	Upper	Upper	Divergent	convex
3. Upper depression	Lower	lower level and depression	Depression	toe slope and depression
4. Backslope	mid-slope	mid-slope	Flat	mid-slope
5. Divergent backslope	mid-slope	mid-slope	Divergent	mid-slope
6. Convergent backslope	mid-slope	mid-slope	Convergent	mid-slope
7. Terrace	mid-slope	mid-slope	Flat	mid-slope
8. Saddle	mid-slope	mid-slope	Flat	mid-slope
9. Midslope depression	Lower	lower level and depression	Depression	toe slope and depression
10. Footslope	Lower	Lower	Convergent	concave
11. Toeslope	Lower	Lower	Flat	toe slope and depression
12. Fan	Lower	Lower	Divergent	toe slope and depression
13. Lower slope mound	Lower	Lower	Divergent	toe slope and depression
14. Lower level slope	Lower	lower level and depression	Depression	toe slope and depression
15. Lower depression	Lower	lower level and depression	Depression	toe slope and depression

In ArcView, each DEM was exported as a binary file that was then imported into Idrisi32 where a metadata header file was created for each binary grid. Then the grid was exported to an ASCII text format file that LandMapR could read. The LandMapR command routines were run on the ASCII text files in Visual Fox Pro (Microsoft Corp., Bellevue, WA) that generated database tables that included several ratio-valued terrain attributes including the compound topographic index. Database tables were then imported into ArcView from which 10-m grids were created of the 15-category hillslope

model (Figure 20) as well as wetness area and the compound topographic index (Figure 21).

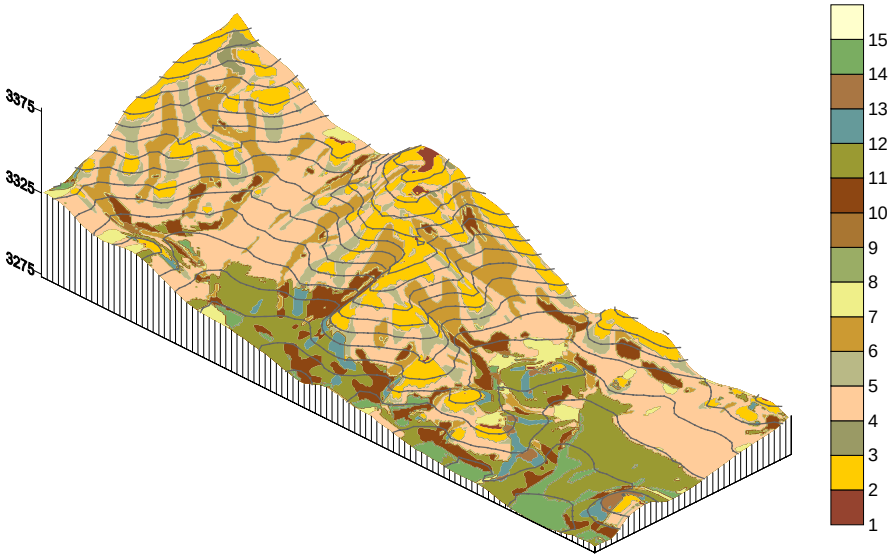


Figure 20. Fifteen element hillslope model computed from a DEM for the Mattson field.

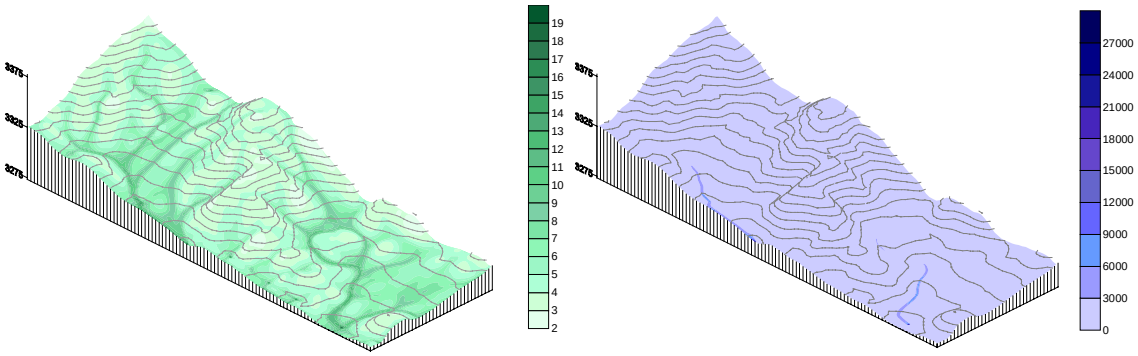


Figure 21. Secondary terrain attributes computed from DEM for the Mattson field: compound topographic index (left) and wetness area (right).

The Compound Topographic Index (CTI) is a quantification of position of a site in the local landscape, which can be computed from a DEM:

$$CTI = \ln (A_s / \tan \beta) \quad \text{Eq. [2]}$$

where A_s is the specific catchment area expressed as m^2 per unit width orthogonal to the flow direction and β is the slope angle in degrees. The CTI, also known as the steady-state wetness index, is a significant predictor of soil properties (Moore et al., 1993; Gessler et al., 2000), so should relate to crop productivity, especially in semiarid environments where water is a strong determinant of crop growth.

To simplify to a form that would be easier for producers to apply to site-specific field management, the 15 classed, landscape elements were grouped into three and four class management zones (Table 11 and Figures 23 and 24) in accordance with MacMillan et al. (2000). For example, a management map with three zones was created by grouping the landscape elements into upper-, middle-, and lower-slopes (Figure 24). Management maps with four zones were derived by grouping into upper-, middle-, and lower-slopes, and depressions; and also by grouping into flow patterns as for depressions and flat, convergent and divergent slopes (Figure 25). From Ciha (1984), a map with five zones consisting of depressions and interfluves and convex, middle, concave and toe slopes (Table 11 and Figure 22).

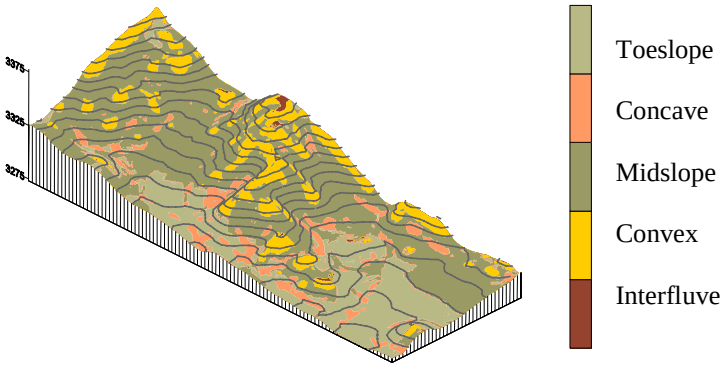


Figure 22. Mattson field– 5 element model

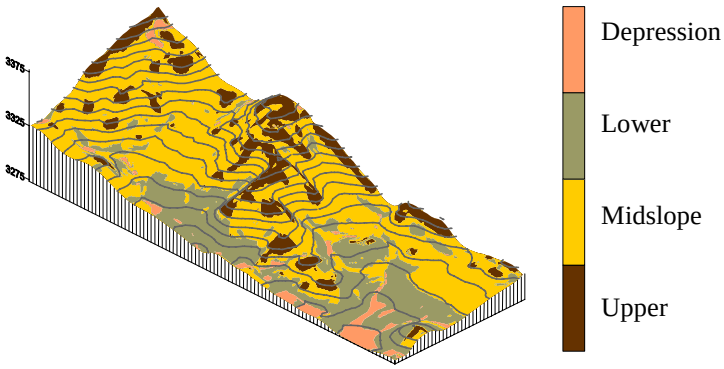


Figure 23. Mattson field – 4 element model

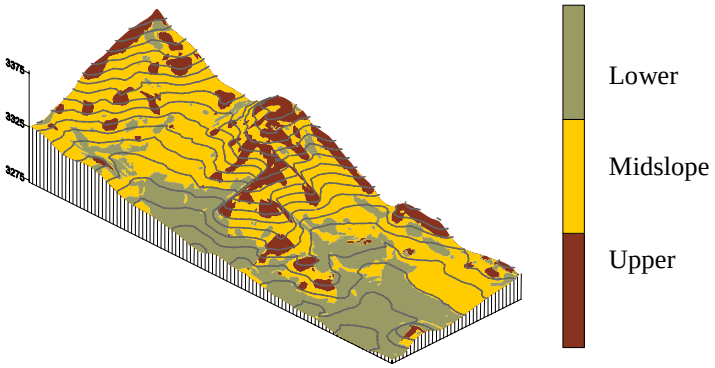


Figure 24. Mattson field – 3 element model

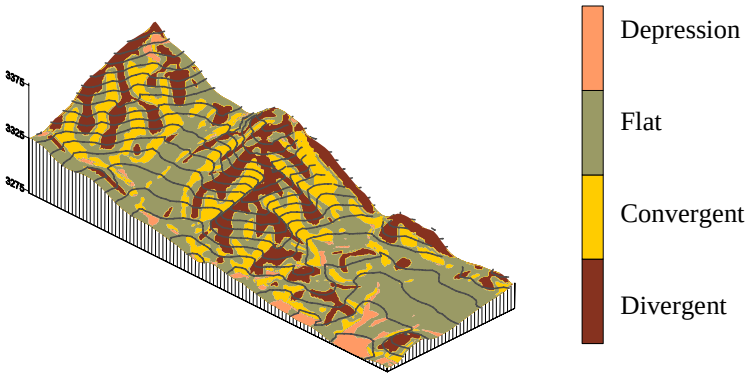


Figure 25. Mattson field – flow element model

Statistical Analyses

Lagged yield point data were cleaned of erroneously low values that occurred whenever the combine entered the crop and the threshing mechanism was not at full capacity, giving the impression of lower than actual yields being present. In addition, all grids were clipped to field boundaries to delete extraneous data that was collected when the combine traveled outside of the field boundary. Edited point data sets were saved as ArcView Shapefiles for final analysis. The grain sampled for protein concentration would not have been negatively affected by the dynamics of grain flow, so the resulting point data could be converted directly to Shapefiles.

Typically, the number of data points for yield far exceeded the number of data points for grain protein concentration (i.e. n=30,000 vs. n=1,000). To reduce the amount of yield data to a manageable level for statistical analysis on a personal computer, the "Get Grid Value" script available for ArcView was used to extract values from a yield map to within 1-m of the same points where grain had been sampled by hand from the combine during harvest (Long et al., 2002). This sampling procedure was assumed to be adequate if the variances computed for the full and the reduced data series were approximately equal in value. Values of hillslope class, wetness area and wetness index were also extracted to the protein points using the Get Grid Value function.

A table of point values of the crop, topographic, and soil variables were exported from ArcView to enable their statistical analysis within the statistical software package, SYSTAT ver. 7.0 (SPSS, Chicago, IL). Descriptive statistics were computed for grain yield and protein within each field and its landscape elements. The hypothesis that

topography captures spatial variability in grain yield and grain protein was examined by investigating the correlation between these crop variables and slope, wetness index and wetness area. Linear regression and Analysis of Variance (ANOVA) were performed using the General Linear Model (GLM) procedures available in SYSTAT.

Results and Discussion

Growing Season Precipitation

Weather records for nearby National Oceanographic and Atmospheric Administration stations indicated that total growing season (April to July) precipitation timing and amounts varied widely from year to year (Table 12). The NOAA weather stations included: (1) Chester- located 21-km south of the Mattson farm; (2) Wildhorse- 10-km north of the Peterson farm; (3) Fort Assinniboine- 32-km northeast of the Grass farm and 11-km northwest of the Kaercher farm; and (4) Malta 7E- 11-km west of the Anderson farm.

Total seasonal precipitation was near normal (95% to 105%) in 1997 at Chester (Mattson farm), in 1999 at Wildhorse (Peterson farm), and in 2000 at Fort Assinniboine (Grass and Kaercher farms). As indicated by records for Wildhorse and Malta 7E, seasonal rainfall was likely below normal at the Peterson farm in 1998 (91% of long-term) and in 2000 (76% of long-term); and Anderson farm in 2000 (79% of long-term). In contrast, it was likely above-normal at the Grass farm in 1998 and 1999 (>135% of

long-term), and at the Anderson farm in 1999 (109% of long-term) as revealed by records for Fort Assinniboine and Malta 7E.

Moisture conditions in several of the farm-years were indicated below normal during July, which is the critical grain filling month prior to physiological maturity of spring wheat. For example, July precipitation was abnormally low at Wildhorse in 1998, 1999, and 2000 (25% to 82% of long-term), and at Fort Assinniboine and Malta 7E in 1999 (38% and 56% of long-term). In contrast, rainfall in July was greater than normal at Chester in 1997 (141% of long-term), and at Fort Assinniboine in 1999 and 2000 (112% and 135% of long-term).

Table 12. Monthly growing season precipitation and percentages of long-term precipitation taken from the nearest NOAA weather stations.

Station	Farm	Year	Precipitation (mm)									
			Apr		May		June		July		Total	
			GS	%	GS	%	GS†	%	GS	%	GS†	%
Chester	Mattson	1997	30	15	625	139	411	75	523	141	159	101
Wildhorse	Peterson	1998	53	34	365	95	721	141	155	43	129	91
	Peterson	1999	333	215	381	99	447	87	30	82	146	103
	Peterson	2000	224	145	353	92	409	80	91	25	108	76
Ft. Assinn.	Grass	1998	168	66	366	72	1295	271	401	112	223	139
	Grass	1999	465	183	579	114	996	208	137	38	218	136
	Grass/ Kaercher	2000	183	72	396	78	460	96	485	135	152	95
Malta 7E	Anderson	1999	284	124	848	149	693	106	257	56	208	109
	Anderson	2000	165	72	277	49	630	96	429	93	150	79

†GS-Growing season precipitation in millimeters
‡%-Percent of long term precipitation

Yield and Protein Measurements

Values in mean grain yield and grain protein varied substantially from field to field likely reflecting differences in cropping intensity and amount and timing of growing season precipitation (Tables 13 and 14). Mean wheat grain yields ranged from a low of 573 kg ha^{-1} at the Grass farm (State W field) in 1999 to a high of 3197 kg ha^{-1} at the Anderson farm (East field) in the same year. Most fields yielded less than 2351 kg ha^{-1} (35 bu ac^{-1}), which is generally considered typical for north central Montana in years of normal rainfall. Mean protein levels at the Peterson farm ranged from a low of 131 mg kg^{-1} in 1998 (Home field) to a high of 195 mg kg^{-1} in 2000 (SE field). Nearly all site-years exceeded a protein concentration of 140 mg kg^{-1}, which is generally considered the normal level for HRSW in this region.

Of interest is the coefficient of variation (CV), which provides a relative measure of the variability in a data series and is given by:

$$CV = sd/mean \qquad \text{Eq. [3]}$$

where *sd* is the standard deviation of the data series. Values near zero indicate low spatial variability in a data series (uniformity) and values near 1.0 indicate high spatial variability (heterogeneity). Grain yield CV values were consistently low to moderately low (CV values 0.11 to 0.33) except for three of the 18 fields (CV values 0.46 to 1.33). Grain protein CV values were consistently low (CV values 0.02 to 0.16) in all fields, which indicated a substantial lack of within-field variability.

In general, values of mean grain yield, ranked in ascending order across the farm-years, correspond more with decreasing cropping intensity than with increasing total

seasonal rainfall (Table 13). Due to greater cropping frequency, which reduces stored soil moisture, the yield potential of annual cropping is 60 to 75% of conventional, alternate fallow-wheat cropping (Brown et al., 1985). Therefore, below average grain production (<2351 kg ha^{-1} or <35 bu ac^{-1}) tended to be associated with fields in which the cropping frequency was higher (i.e. fallow-lentil-wheat and annual wheat) than the traditional fallow-wheat system.

Table 13. Grain yield statistics by field.

Farm	Field	Year	ppt†	Crop Rotation	*n*	mean	SD‡	CV§	Diff¶
						Grain Yield (kg ha^{-1})			
Grass	State-W	1999	136	Fallow-Lentil-Wheat	107	573	539	0.13	-1778
Grass	Sprague	2000	95	Fallow-Pea-Wheat	109	600	800	1.33	-1751
Mattson		1997	101	Annual Wheat	320	936	283	0.30	-1415
Peterson	SW	2000	95	Canola-Fallow-Wheat	182	1053	493	0.47	-1298
Peterson	SE	2000	95	Canola-Fallow-Wheat	121	1209	499	0.41	-1142
Peterson	NE	2000	95	Fallow-Wheat	143	1272	326	0.26	-1079
Peterson	Home	1998	139	Fallow-Wheat	160	1440	429	0.30	-911
Peterson	NW	2000	95	Fallow-Wheat	177	1472	262	0.18	-879
Grass	Sand	1998	139	Fallow-Pea-Wheat	200	1676	397	0.24	-675
Grass	State-E	2000	95	Fallow-Wheat	139	1946	888	0.46	-405
Peterson	Larson	1999	136	Fallow-Canola-Wheat	163	1974	586	0.30	377
Kaercher		2000	95	Fallow-Wheat	365	2028	464	0.23	-323
Peterson	WW	1999	82	Fallow-Wheat	211	2029	498	0.25	-322
Grass	Home	1999	136	Fallow-Wheat	160	2132	702	0.33	-219
Anderson	West	2000	79	Fallow-Wheat	591	2523	524	0.21	172
Peterson	WF	1999	136	Fallow-Wheat	237	2574	292	0.11	223
Grass	Russell	1998	139	Fallow-Wheat	207	2642	384	0.15	291
Anderson	East	1999	109	Fallow-Wheat	275	3197	580	0.18	846

†ppt-Gowing season percentage of long term precipitation
‡SD-Standard deviation §CV-Coefficient of Variation
¶Diff-Difference from average grain yield of 2352 kg ha^{-1} (35bu ac^{-1})

The lower grain production in these fields, that likely resulted from water stress, did not correlate with above normal (>140 mg g^{-1} or >14.0%) grain protein levels, as was

expected. There was no trend for cropping system or seasonal precipitation to result in higher or lower grain protein levels (Table 14). In dryland environments where water is limiting to yield, grain protein concentration is inversely related with grain yield. Normally, this inverse protein-yield relationship has been attributed to dilution of a limited amount of grain N by a much larger biomass accumulation, particulary when water does not limit yield (Terman et al, 1969).

Table 14. Grain protein concentration statistics by field.

Farm	Field	Year	ppt	Crop Rotation	*n*	mean	SD	CV	Diff
						Concentration (mg^{-1} g)			
Peterson	Home	1998	139	Fallow-Wheat	160	131	15	0.11	0.11
Grass	Russell	1998	139	Fallow-Wheat	207	138	14	0.10	0.10
Anderson	East	1999	109	Fallow-Wheat	275	147	10	0.07	0.07
Peterson	NW	2000	95	Fallow-Wheat	177	151	24	0.16	0.16
Anderson	East	2000	109	Fallow-Wheat	591	156	13	0.08	0.08
Mattson		1997	101	Annual Wheat	320	158	17	0.11	0.11
Grass	Sand	1998	139	Fallow-Pea-Wheat	200	159	13	0.08	0.08
Peterson	WF	1999	136	Fallow-Wheat	237	162	9	0.06	0.06
Peterson	WW	1999	136	Fallow-Wheat	211	164	13	0.08	0.08
Grass	Sprague	2000	95	Fallow-Pea-Wheat	109	166	10	0.06	0.06
Grass	Home	1999	136	Fallow-Wheat	160	167	6	0.04	0.04
Kaercher		2000	95	Fallow-Wheat	365	169	4	0.02	0.02
Peterson	Larson	1999	136	Fallow-Canola-Wheat	163	171	10	0.06	0.06
Grass	State-E	2000	95	Fallow-Wheat	139	172	6	0.03	0.03
Peterson	NE	2000	95	Fallow-Wheat	143	172	25	0.15	0.15
Grass	State-W	1999	136	Fallow-Lentil-Wheat	107	173	12	0.07	0.07
Peterson	SW	2000	95	Canola-Fallow-Wheat	182	188	7	0.04	0.04
Peterson	SE	2000	95	Canola-Fallow-Wheat	121	195	7	0.04	0.04

†ppt-Gowing season percentage of long term precipitation
‡SD-Standard deviation §CV-Coefficient of Variation
¶Diff-Difference from average grain protein of 140mg g^{-1} (14%)

Relationships Between Crop Yield, Protein and Landscape Elements

The hypothesis that grain yield was related to landscape elements was examined using the coefficient of determination (r^2), which measured the amount of variance explained by classificatory predictor variables. Regressions that incorporated landscape elements as a predictor of grain yield had r^2 values that were either insignificant or significantly low in all field sites (Table 15). In particular, the slope-region models had r^2 values that were significant in only six to seven of the 18 field sites vs. the flow-region model that had r^2 values that were significant in nine of the field sites. At best, however, only 18% of the variance in grain yield was explained by landscape elements, which was for the Peterson Larson field in 1999. In all fields, the relationship was so weak or inconsistent as to indicate that landscape elements apparently may not be useful in predicting grain yields.

Similar to results for grain yield, landscape elements as a regression estimator of grain protein concentration had r^2 values that were insignificant to significantly low in all field sites (Table 16). Though the total number of significant landscape-protein relationships exceeded that for grain yield (55 vs. 29), the slope and flow region models explained no better than 26% of the variance in grain protein, which was only in the Peterson SE field in 2000. Again, the amount of explained variance in grain protein was low in all fields, which does not confirm the hypothesis that topography captures spatial patterns in grain protein concentration.

Table 15. Coefficients of determination (r^2) for landscape class models predicting grain yield.

Farm	Field	Year	*n*	3-slope	4-slope	4-flow	5-slope
Mattson		1997	320				
Peterson	Home	1998	160				
Grass	Russell	1998	207	0.078	0.083	0.073	0.10
Grass	Sand	1998	200	0.036	0.11	0.10	0.058
Peterson	WF	1999	237	0.041	0.046	0.064	0.043
Peterson	WW	1999	211		0.057	0.067	
Peterson	Larson	1999	163	0.11	0.17	0.14	0.18
Grass	Home	1999	160	0.044			0.068
Grass	State-W	1999	107	0.096	0.10	0.11	0.098
Anderson	East	1999	139				
Peterson	SE	2000	275				
Peterson	SW	2000	121				
Peterson	NE	2000	182				
Peterson	NW	2000	143		0.13		
Grass	State-E	2000	177			0.062	0.17
Grass	Sprague	2000	109			0.080	
Anderson	West	2000	591			0.17	
Kaercher		2000	367				

*Listed r^2 values are significant at $P<0.05$.

Relationships Between Crop Yield, Protein and Computed Terrain Attributes

Regression analyses of grain yield were not successful in demonstrating strong relationships with the computed terrain attributes of slope, wetness area, or compound topographic index (Table 17). Though many regressions were statistically significant, coefficients of determination (r^2) were generally low, indicating that little of the variance in grain yield was explained by these computed terrain attributes. Grain yield correlated best with slope ($r^2 = 0.21$) and CTI ($r^2 = 0.18$) in the Grass Sand field in 1998. For nearly all fields, however, computed terrain attributes explained less than 10% of the variance in grain yield.

Table 16. Coefficients of determination (r^2) for landscape class models predicting grain protein concentration.

Farm	Field	Year	*n*	3-slope	4-slope	4-flow	5-slope
Mattson		1997	320	0.031	0.031	0.065	0.045
Peterson	Home	1998	160		0.071		
Grass	Russell	1998	207	0.030			
Grass	Sand	1998	200	0.096	0.096	0.088	0.14
Peterson	WF	1999	237	0.073	0.12	0.11	0.073
Peterson	WW	1999	211	0.049	0.059	0.035	
Peterson	Larson	1999	163	0.095	0.233	0.15	0.21
Grass	Home	1999	160	0.14	0.15	0.11	0.17
Grass	State-W	1999	107	0.11	0.11	0.091	0.12
Anderson	East	1999	139	0.056			
Peterson	SE	2000	275	0.26	0.26	0.13	0.26
Peterson	SW	2000	121				
Peterson	NE	2000	182	0.16	0.16	0.045	0.16
Peterson	NW	2000	143	0.055	0.062	0.057	0.11
Grass	State-E	2000	177			0.051	0.098
Grass	Sprague	2000	109	0.14	0.14	0.12	0.15
Anderson	West	2000	591	0.020	0.017	0.017	
Kaercher		2000	367	0.17	0.14	0.10	0.14

*Listed r^2 values are significant at $P<0.05$.

Use of slope, compound topographic index, or wetness area as a regression estimator of grain protein concentration also produced poor results. The listed r^2 values are either insignificant or are significantly low (Table 18). At best, slope explained only 15% of the variance in grain protein concentration at only one field site (Grass Sprague 2000) and the compound topographic index explained 26% at another (Peterson SE 1999). Unfortunately, less than 10% of the variance in grain protein concentration was explained in the remaining field sites.

Table 17. Coefficients of determination (r^2) for the relationship between grain yield and computed terrain attributes.

Farm	Field	Year	*n*	Slope	CTI†	WA‡
Mattson		1997	320		0.023	
Peterson	Home	1998	160		0.028	
Grass	Russell	1998	207	0.025	0.057	
Grass	Sand	1998	200	0.21	0.18	0.034
Peterson	WF	1999	237	0.029	0.010	0.022
Peterson	WW	1999	211			
Peterson	Larson	1999	163	0.05	0.11	
Grass	Home	1999	160			
Grass	State-W	1999	107		0.074	0.053
Anderson	East	1999	139			
Peterson	SE	2000	275			
Peterson	SW	2000	121			
Peterson	NE	2000	182	0.041		
Peterson	NW	2000	143	0.065		
Grass	State-E	2000	177	0.091	0.042	
Grass	Sprague	2000	109	0.18		0.043
Anderson	West	2000	591	0.013	0.01	
Kaercher		2000	367			

*Listed values are significant at P<0.05.
†CTI-Compound topographic index ‡WA-Wetness area

Average Yield and Protein Levels By Hillslope Position

To evaluate changes by hillslope position, the mean values in grain yield and grain protein concentration are listed in Table 19 with respect to the simple three-class, slope-region model. Both grain yield and grain protein concentration tended to increase in order of upper, middle and lower slopes in nearly all fields. The change in yield is likely a response to increasing water in the downslope direction, caused by redistribution of soil moisture by overland and subsurface flow (Ciha, 1984; Hanna et al., 1982); however, many studies have shown grain yield and grain protein concentration to be negatively correlated, especially in dryland environments where water stress is a common

occurrence (Smika and Greb, 1973; McNeal et al., 1972; Terman 1979). This was indicated for the Mattson field (1997) and Grass State-W field where the grain yields were observed to be extremely low (Table 19). Increases in protein concentration are attributable to water stress, which produces a disequalibrium within the cereal plant, where a large pool of N is available for remobilization to grain relative to a severely reduced amount of grain (Seles and Zentner, 1999). So an increase in wheat yields, corresponding to increases in protein, in most fields in the downslope direction towards water-catching topography is unexpected.

Table 18. Coefficients of determination (r^2) for the relationship between grain protein and computed terrain attributes.

Farm	Field	Year	*n*	Slope	CTI	WA
Mattson		1997	320	0.11	0.085	
Peterson	Home	1998	160			0.032
Grass	Russell	1998	207	0.052	0.029	0.044
Grass	Sand	1998	200			
Peterson	WF	1999	237	0.028	0.10	0.045
Peterson	WW	1999	211	0.035	0.037	0.022
Peterson	Larson	1999	163	0.16	0.16	
Grass	Home	1999	160	0.097	0.088	
Grass	State-W	1999	107		0.058	0.048
Anderson	East	1999	139		0.037	
Peterson	SE	2000	275	0.029	0.26	0.10
Peterson	SW	2000	121	0.059	0.058	
Peterson	NE	2000	182	0.057	0.062	
Peterson	NW	2000	143	0.053	0.045	0.065
Grass	State-E	2000	177	0.13		
Grass	Sprague	2000	109	0.15	0.10	
Anderson	West	2000	591			
Kaercher		2000	367	0.11	0.097	0.010

*Listed values are significant at P<0.05.
†CTI-Compound topographic index ‡WA-Wetness area

Table 19. Average grain yield and grain protein concentration by landscape position.

Farm	Field	Grain Yield *			Grain Protein *		
		Upper	Middle	Lower	Upper	Middle	Lower
		----------kg ha^{-1}----------			----------mg g^{-1}----------		
Mattson		909a	944a	913a	167 a	157 b	156 c
Peterson	Home	1256a	1452a	1458a	132a	130a	135a
Grass	Russell	2344 a	2647 b	2772 c	137 a	136 b	141 c
Grass	Sand	1648 a	1561 b	1779 c	156 a	156 a	164 b
Peterson	WF	2290 a	2533 b	2685 c	164 a	164 a	159 b
Peterson	WW	1976a	2068a	1884a	164 a	163 a	170 b
Peterson	Larson	1471 a	1854 b	2164 c	165 a	170 b	173 c
Grass	Home	2266 a	2231 b	1993 c	162 a	166 b	169 c
Grass	State-W	547 a	377 b	883 c	172 a	177 b	167 c
Anderson	East	3056a	3272a	3112a	170 a	172 b	174 c
Peterson	SE	1232a	1200a	1216a	144 a	144 a	154 b
Peterson	SW	1106a	1225a	1181a	187a	186a	192a
Peterson	NE	1408a	1262a	1213a	187 a	186 a	192 b
Peterson	NW	1491a	1496a	1376a	161 a	172 b	179 c
Grass	State-E	2061a	1906a	1926a	170a	172a	174a
Grass	Sprague	567a	638a	562a	163 a	163 a	171 b
Anderson	West	2423a	2563a	2539a	153 a	157 b	158 b
Kaercher		1978a	2036a	2042a	167 a	168 a	171 b

* Means followed by the same letter within a row are not significantly different at the 5% level of probability ($P<0.05$).

Positive yield-protein relationships have been attributed to changes in the N supply when N was more limiting than plant available water (Terman et al. 1969; Terman 1979; and Gauer et al. 1992). Furthermore, plants located on deeper soils in toeslopes and depressions draw upon stored reserves of water primarily below 60-cm (Smika and Greb, 1973), which increases the yield potential and thus the plant's demand for N and its capacity to uptake soil N. Therefore, the positive yield-protein relationship may have

resulted from soil N fertility and water availability that both increased as soils deepened downslope.

Conclusions

LandmapR and ArcView software were used to generate terrain attribute maps, including classed landscape elements and ratio-valued terrain attributes for eighteen farm fields. Yield and protein values for hard red spring wheat were collected over a three year period on these fields and correlations between these variables and predictor terrain variables examined. Yield was expected to increase and protein decrease in response to increasing plant available water in the downslope, direction. Instead, both yield and protein tended to increase downslope, likely due to increases in plant available water and nitrogen fertility in that direction. Climatic data from nearby weather stations did not indicate that growing season precipitation was unusually low; however, below average yields apparently may been caused by intensive cropping in some fields that depleted stored soil moisture.

Landscape elements, as defined by flow and slope regions, did not capture a majority of the variance in grain yield or grain protein. Apparently, a significant amount of variance remained within these elements at smaller spatial scales, indicating that other biophysical variables may underly the spatial variability. Thus, the question remains whether computed terrain attributes are effective for determining field management zones in glacial till hummocky terrain found in northern Montana and other semiarid environments. The newest terrain modeling approaches and PC-based commercial

software packages are now becoming available that may open new avenues for furthering this landscape-based research. Revisiting the concepts and methods outlined in this chapter and expanding them to include sample measurements of soil profile N, moisture, texture and structure as predictor variables may clarify the effects of topography on soil formation and protein and yield interactions at the field scale.

REFERENCES CITED

Aandahl, A.R. 1948. The characterization of slope positions and their influence on the total nitrogen content of a few virgin soil of western Iowa. Soil Science Society Proceedings, 1948: 449-454.

Afyuni, M.M., Cassel D.K. and Robarge W.P. 1993. Effect of landscape position on soil water and corn silage yield. Soil Science Society of America Journal, 57:1573-1580.

Bjorsvik, H.R. and Martens, H. 1992. Data Analysis: Calibration of NIR Instruments by PLS Regression, in Handbook of Near-Infrared Analysis, edited by Burns, D.A. and Ciurczak, E.W., M Dekker, New York, 159-180.

Brown, P.L., A.L. Black, C.M. Smith, J.W. Enz, and J.M. Caprio. 1985. Soil water guidelines and precipitation probabilities in Montana and North Dakota. Montana Coop. Ext. Serv. Bull. 356.

Burrough, P.A., and McDonnell, R.A. 1998. Chapter 8: Spatial analysis using continuous fields, in Principles of Geographical Information Systems. Oxford University Press, New York, NY., 183-198.

Burt, T.P. and Butcher, D.P. 1985. Topographic controls of soil moisture distributions. Journal of Soil Science, 36:469-486.

Carr, P.M., Carlson, G.R., Jacobsen, J.S., Nielsen, G.A. and Scogley, E.O. 1991. Farming soil, not fields: A strategy for increasing fertilizer profitability. Journal of Production Agriculture, 4:57-61.

Case Corporation. 2000. Press Release.

Ciha, A.J. 1984. Slope position and grain yield of soft white winter wheat. Agronomy Journal, 76:193-196.

Engel, R. E., Long, D. and Carlson, G. 1997. On-the-Go Grain Protein Sensing is Near. Better Crops with Plant Food, 81(4):20-23

Engel, R.E., Long, D.S., Carlson, G.R. and Meier, C. 1999. Method for Precision Nitrogen Management in Spring Wheat: I Fundamental Relationships. Precision Agriculture, 1:327-338.

Fiez, T.E., Miller, B.C. and Pan, W.L. 1994. Winter wheat yield and grain protein across varied landscape positions. Agronomy Journal, 86:1026-1032.

Furman, W. B. 1976. Continuous Flow Analysis: Theory and Practice. Marcel Dekker, Inc., New York.

Gauer, L.E., Grant, C.A., Gehl, D.T. and Bailey, L.D. 1992. Effects of nitrogen fertilization on grain protein content, nitrogen uptake, and nitrogen use efficiency of six spring wheat (Triticum aestivuum L.) cultivars in relation to estimated soil moisture supply. Canadian Journal of Plant Sciences, 72:235-241.

Geladi, P. and Kowalski, B.R. 1986. Partial Least-Squares Regression: A Tutorial. Analytica Chimica Acta, 185:1-17.

Gessler, P.E., Chadwick O.A., Chamran F., Althouse L. and Holmes K. 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. Soil Science Society of America Journal, 64:2046-2056.

Goos, R.J., Westfall, D.G., Ludwick, A.E. and Goris, J.E. 1982. Grain protein content as an indicator of N sufficiency for winter wheat. Agronomy Journal, 74:103-133.

Hamblin, A., Richards, Q. and Blake, J. 1988. Crop growth across a toposequence controlled by depth of sand over clay. Australian Journal of Soil Research, 26:623-635.

Hanna, A.Y., Harlan, P.W. and Lewis, D.T. 1982. Soil available water as influenced by landscape position and aspect. Agronomy Journal, 74:999-1004.

Hrushka, W.R. 1987. Data Analysis: Wavelength Selection Methods, in Near-Infrared Technology in the Agriculture and Food Industries, edited by Williams, P. and Norris, K., American Association of Cereal Chemists, Inc. St. Paul, Minnesota, 35-55.

Karlen, D.L., Sadler, E.J. and Busscher, W.J. 1990. Crop yield variation associated with coastal plain soil mapping units. Soil Science Society of America Journal, 54:859-865.

Kibite, S., and Evans, L.E. 1984. Causes of negative correlations between grain yield and grain protein concentration in common wheat. Euphytica, 33:801-810.

Kravchenko, A.N. and Bullock, D.G. 2002. Spatial variability of soybean quality data as a function of field topography: I. Spatial data analysis. Crop Science, 42: 804-815.

Long, D.S. and Engel, R.E. 1998. Grain Protein Management Using Precision Farming Methods. Proceedings of the Wheat Protein Symposium, Saskatoon, SK, March 9-10, 169-179.

Long, D. S., Engel, R.E. and Carlson, G.R. 2000. Method for Precision Nitrogen Management in Spring Wheat: II. Implementation. Precision Agriculture, 2:25-38.

Long, D.S., Meier, C., Engel, R.E. and Lenssen, A. 2002. Landscape elements as a basis for nitrogen management zones in Montana. Proceedings of the Great Plains Soil Fertility Conference, March 5-6, 2002, Denver, CO.

Machado, S., Bynum, E.D. Jr., Archer, T.L., Lascano, R.J., Wilson, L.T., Bordovsky, J., Segarra, E., Bronson, K., Nesmith, D.M. and Xu, W. 2002. Spatial and temporal variability of grain yield: Implications for site-specific farming. Crop Science, 42:1564-1576.

MacMillan, R.A. 2000. A protocol for preparing digital elevation (DEM) data for input and analysis using the landform segmentation model (LSM) programs, prepared by Landmapper Environmental Solutions. Prepared for the Soil Variability Analysis to Enhance Crop Production (SVAECP) Project.

MacMillan, R.A., Pettapiece, W.W., Nolan, S.C. and Goddard, T.W. 2000. A generic procedure for automatically segmenting landforms into landscape elements using DEMs, heuristic rules and fuzzy logic. Fuzzy Sets and Systems, 113(1):81-109.

Martens, H. and Jensen, S.A. 1983. Partial Least Squares Regression: A New Two-Stage NIR Calibration Method. Proceedings of the Seventh World Cereal and Bread Congress, Prague, CZ, June 28-July 2, 607-647.

Martens, H. and Naes, T. 1987. Multivariate Calibration by Data Compression, in Near-Infrared Technology in the Agriculture and Food Industries, edited by Williams, P. and Norris, K., American Association of Cereal Chemists, Inc. St. Paul, Minnesota, 57-87.

MathWorks. 1999. Learning MatLab, Version 5.3 (MathWorks, Inc., Natick, MA).

McNeal, F.H., Berg, M.A., McGuire, C.F., Stewart, V.R. and Baldridge, D.E. 1972. Grain and plant nitrogen relationships in eight spring wheat crosses, Triticum aestivuum L. Crop Science, 12:599-602.

Moore, I.D., Gessler, P.E., Nielsen, G.A. and Peterson, G.A. 1993. Soil attribute prediction using terrain analysis. Soil Science of America Journal, 57:443-452.

Moulin, A.P., Anderson, D.W. and Mellinger, M. 1994. Spatial variability of wheat yield, soil properties and erosion in hummocky terrain. Canadian Journal of Soil Science, 74:219-228.

Mulla, D.J. 1992. Mapping and managing spatial patterns in soil fertility and crop yield. Pg. 15. *In* P.C. Robert et al. (ed.) Proceedings of first workshop on soil specific crop management. 14-16 April 1992, Minneapolis, MN. ASA-CSSA-SSSA. Madison, WI.

Murray, I. and Williams, P.C. 1987. Chemical Principles of Near-Infrared Technology, inNear-Infrared Technology in the Agriculture and Food Industries, edited by Williams, P. and Norris, K., American Association of Cereal Chemists, Inc. St. Paul, Minnesota, 17-34.

Naes, T. and Martens, H. 1985. Comparison of Prediction Methods for Multicollinear Data. Communications in Statistics: Simulation and Computation, 14(3):545-576.

National Oceanic and Atmospheric Administration (NOAA). 2003. http://www.ncdc.noaa.gov/oa/ncdc.html. National Climatic Data Center.

Nolan, S.C., Goddard, T.W., Penney, D.C. and Green, F.M. 1999. Yield response to nitrogen within landscape classes. Proceedings of the 4th International Conference on Precision Agriculture, St. Paul, MN, July 19-22, 1998, 479-485.

Nolan, S.C., Goddard, T.W., Lohstraeter, G. and Coen, G.M. 2000. Assessing management units on rolling topography. Proceedings of the 5th International Conference on Precision Agriculture, Minneapolis, MN, July 16-19, 1999.

Olson, R.A., Frank, K.D., Deibert, E.J., Dreier, A.F., Sander, D.H. and Johnson, V.A. 1976. Impact of residual mineral N in soil on grain protein yields of winter wheat and corn. Agronomy Journal, 68:769-772.

Papendick, R.I., Cochran, V.L. and Woody, W.M. 1971. Soil water potential and water content profiles with wheat under low spring and summer rainfall. Agronomy Journal, 63:731-734.

Pennock, D.J., Zebarth B.J. and DeJong, E. 1987. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. Geoderma, 40:297-315.

Rawls, W.J., Brackensiek, D.L. and Saxton, K.E. 1982. Estimation of soil water properties. Transactions of the American Society of Agricultural Engineers, 25(5):1316-1320.

Selles F. and Zentner, R.P. 1999. Environmental factors affecting wheat protein, in Wheat Protein Production and Marketing: Proceedings of the Wheat Protein Symposium, Saskatoon, SK, March 9-10, 1998, 139-150.

Smika, D.E. and Greb, B.W. 1973. Protein content of winter wheat grains as related to soil and climatic factors in the semiarid Central Great Plains. Agronomy Journal, 65:433-436.

SPSS, Inc. 1997. Systat 7.0: Statistics (SPSS, Inc., Chicago, IL).

Stewart. C.M., McBratney, A.B. and Skerritt, J.H. 2002. Site-specific Durum wheat quality and its relationship to soil properties in a single field in Northern New South Wales. Precision Agriculture, 3:155-168.

Stone, J.R., Gilliam, J.W., Cassel, D.K., Daniels, R.B., Nelson, L.A. and Kleiss, H.J. 1985. Effect of erosion and landscape position on the productivity of piedmont soils. Soil Science of America Journal, 49:987-991.

Sweeney, R.A. and P. R. Rexford. 1987. Comparison of LECO FP-228 "Nitrogen determinator" with AOAC copper catalyst Kjeldahl method for crude protein. J. Assoc. of Anal. Chem. 70, 1028-1030.

Terman, G.L. 1979. Yields and protein content of wheat grain, as effected by cultivar, N, and environmental growth factors. Agronomy Journal, 71:437-440.

Terman, G.L., Ramig, R.E., Dreier, A.F. and Olson, R.A. 1969. Yield-protein relationships in wheat grain as affected by nitrogen and water. Agronomy Journal, 61:755-759.

Tkachuk, R. 1987. Analysis of Whole Grains by Near-Infrared Reflectance, in Near-Infrared Technology in the Agriculture and Food Industries, edited by Williams, P. and Norris, K., American Association of Cereal Chemists, Inc. St. Paul, Minnesota, 233-240.

USDA, Montana Agricultural Statistics Service. 2001. Montana Crop Production: Annual Summary, 2000

Von Rosenberg Jr., C.W., Abbate, A. & Drake, J. 2000. A Rugged Near-Infrared Spectrometer for the Real-time Measurement of Grains During Harvest. Spectroscopy, 15(6):34-38.

Walker, P.H., Hall, G.F. and Protz, R. 1968. Relation between landform parameters and soil properties. Soil Science Society of America Proceedings, 32:101-104.

Western, A.W., Grayson, R.B., Bloschl, G., Willgoose, G.R. and McMahon, T.A. 1999. Observed spatial organization of soil moisture and its relation to terrain indices. Water Resources Research, 35:797-810.

Wibawa, W.D., Duduzile, L.D., Swenson, L.J., Hopkins, D.G. and Dahnke, W.C. 1993. Variable Fertilizer Application Based on Yield Goal Soil Fertility and Soil Map Unit. Journal of Production Agriculture, 6:25-261

Williams, P.C. 1975. Application of Near Infrared Reflectance Spectroscopy to Analysis of Cereal Grains and Oilseeds. Cereal Chemistry, 52:561-576.

Williams, P.C. 1987. Variables Affecting Near-Infrared Reflectance Spectroscopic Analysis, in Near-Infrared Technology in the Agriculture and Food Industries, edited by Williams, P. and Norris, K., American Association of Cereal Chemists, Inc. St. Paul, Minnesota, 143-167.

Wise, B.M. and Gallagher, N.B. 1998. PLS_Toolbox Version 2.0 (Eigenvector Research, Inc., Manson, WA).