

# **Integrating Traditional Handheld Remotely Sensed Measurement Techniques with Chemometric Analyses in the Classroom**

## Authors

**Kevin P. Price**<sup>1</sup>, Thomas J. Prebyl<sup>2</sup>, Nan An<sup>3</sup>, Nathan R. Myers<sup>4</sup>, Yared M. Assefa<sup>3</sup>, Bethany S.P. Grabow<sup>3</sup>, Mark J. Davis<sup>3</sup>, Andrew N. Tucker<sup>3</sup>, Lisa Auen<sup>3</sup>

## Affiliations

(1) Agronomy and Geography, Kansas State University, Manhattan, KS USA

(2) Biology, Kansas State University, Manhattan, KS USA

(3) Agronomy, Kansas State University, Manhattan, KS USA

(4) Geography, Kansas State University, Manhattan, KS USA

## Abstract

The industrial application of Near Infrared (NIR) analyses for characterizing the chemical constituents of a substance started in the 1950's. The use of remotely sensed spectral measurements and chemometric analysis however is a relatively new field of research. In the spring of 2009, Dr. Kevin Price at Kansas State University (KSU) offered the first KSU course in Spectroscopy (hyperspectral analysis) and Chemometric Analysis. In preparing such a course, one will soon find that there are no textbooks on this topic; therefore, the first 1/3 of the semester was spent reading assigned articles to acquaint students with the basic concepts in the field. Later students were introduced to hyperspectral analysis methods including: Principal Components Analysis, Multiple Regression Analysis, 3-D visualization methods, and Derivative Analysis. Students used a hand-held 2000+ band spectroradiometer to collect spectra from compounds such as: salt, sugar (granular and powdered), flour, baking powder, baking soda, garlic powder, black pepper, and paprika. Students entered their spectral measurements into the ViewSpec program to display, statistically analyze and export their spectra to a spreadsheet file format. Students were required to complete a report describing observed spectral differences among compounds and display the spectra in various graphic formats.

Students, in groups of 2 or 3 used the GRAMS chemometrics software to analyze spectra of dry and live grasses that were placed on light and dark soil backgrounds, and to estimate crop leaf and stem N, P, and K nutrient contents. Student hands-on projects proved to be intellectually stimulating and all the students took an active role in their respective groups. Student findings showed considerable promise in using spectroradiometer measurements and chemometric analysis methods for quantifying rangeland grass cover over dark and light soils, and leaf nutrient content of Great Plains crops.

## Background

In this paper authors discuss the challenges, successes and findings of a Chemometrics course offered for the first time in the Spring of 2009 at Kansas State University (KSU) through the Departments of Geography and Agronomy. Dr. Price was the instructor and the co-authors on the paper are the students who helped make the class possible and produced project results that will be summarized at the end of the paper.

About 10 years ago, Dr. Price took a week long Hyperspectral Analysis training course offered at the University of Colorado. At the time, Price was a professor in Geography at the University of Kansas (KU). The first semester after taking the course, he offered a seminar on hyperspectral remote sensing, which was well attended and the student response was very enthusiastic. An important lesson learned

from this experience was that the benefits realized from teaching this seminar would have an expanding effect in that students who took the course passed on the information to other students who used it in their research and in developing lab exercises in hyperspectral analysis for the advanced remote sensing course at KU.

In 2008, Dr. Price took a chemometrics training course from Analytical Spectral Devices, Inc., and in the spring of 2009, after accepting a joint appointment in Agronomy (80%) and Geography (20%) at KSU, offered a seminar on spectrometry and chemometrics. In preparing for the seminar, he was surprised to find no textbooks covering the course topic. During a review of the literature he also found multiple published papers on spectrometry and chemometrics, but as would be expected, most authors presented their findings assuming their readers had at least a basic understanding of the topics. A search of the World Wide Web also produced no courses covering the topics. It soon became clear that this is a field in which few teaching materials exist. It is therefore our hope that this paper will provide some useful information to those desiring to teach a spectrometer and chemometrics course in the future.

### Course Structure

The basic outline of the course included:

- Introduction - What is spectroscopy and how does it relate to remote sensing chemometrics?
- Introduction to the literature on basic principles related to the topic
- Selection and discussion of relevant literature in topics of interest to students in the class (This section was designed to help students better understand how this science and technology can be used within their areas of interest)
- Lecture on hyperspectral datasets including imagery and spectroradiometer measurements
- Lectures by the instructor and guest lectures on mathematical and statistical methods using in spectroscopy and chemometrics including: basic uni- and multivariate statistical analysis (computation of centrality, variance, co-variance, correlation, regression, multi-regression, principal components analysis (PCA), and computations of first and second derivatives and their interpretation)
- In class demonstration of a spectroradiometer (2,150 bands from ~380-2,500 nm) and student hands-on operation of the instrument allowing students to perform calibration steps and collection and storage of spectra for various plant types.
- Development of group projects
- Work on laboratory exercise that involves the analysis of what we called our Kitchen Chemical Samples
- Introduction to the ASD software “ViewSpec”
- Introduction of the chemometrics software program “Grams”
- Laboratory exercises using the “ViewSpec” and “Grams” to produce Chemometric models
- Hands-on laboratory work on student projects that involved use of the spectroradiometer, the “ViewSpec” spectroradiometer analysis software, and “Grams” chemometric modeling software.
- Student Reports

A particularly useful article for introducing students to the topic was the paper titled, *Introduction to NIR Technology* by Analytical Spectral Devices. This paper introduces the reading to the relevant terminology, principles of electromagnetic energy interaction with matter, how energy interacts at the atomic level and how molecular structure influence wavelength. This paper also has a list of the

absorption regions for many molecules. This paper helped students understand how the spectral properties of matter could be used to characterize the molecular make of matter.

Students were assigned to do a literature search on the application of spectrometer and chemometrics in their particular areas of interest. In this way, students could quickly see that these technologies had exciting applications within their field of interest which was focused mainly on crop nutrient assessment and grassland and crop soil residual quantification.

A few formal lectures were given in the course on basic and advanced statistical analysis methods. Students needed to understand how PCA is used to reduce spectral data dimensionality and orthogonalize these datasets to remove multicollinearity so the variables can be used in multiple regression analysis to build chemometric models. These topics were covered in a more conceptual manner leaving the theoretical math and proofs to be taught in an advanced statistic course.

A guest speaker was invited to class to discuss the use of first and second derivatives in spectrometry and how such methods can be used to identify useful regions of the electromagnetic spectrum such as the red edge and unique absorption features for characterizing the chemical properties of the substance being analyzed.

Within a few weeks of the beginning of class students and instructor began discussing possible subgroup projects. The class was divided into groups of 2 to 3 students and after a topic of interest for each group was identified, testable hypotheses were developed and methodologies were outlined.

#### Hands on Spectroradiometer Measurements and Analysis: Kitchen Chemical Samples

Hands-on experience with a full spectrum spectroradiometer is something few students are able to get. It was determined that such experience would be provided in a somewhat controlled environment (the instrument remained in the lab and was used under supervision by the instructor or graduate students with training in the use and proper operation of such a highly technical and expensive instrument). Students were given a sample of ten commonly used kitchen cooking spices and ingredients and asked to collect 10 spectral reading from each of the ingredients (Figure 1). They were then instructed to import their spectral readings into ViewSpec Pro spectral analysis software, to convert their measurements to values of reflectance and then computed means and standard deviation for each of the 10 substances. ViewSpec Pro was then used to plot their means in a line plot so they could compare spectral curves (Figure 2).

Students were also asked to compute the first derivative using the mean spectra for baking soda and brown sugar (Figure 3). From the first derivative graph it can be seen that the two compounds have different regions of the spectrum where the greatest changes in the spectral patterns can be observed. For example, for brown sugar the most dynamic region is in the visible (400 – 730 nm) portion of the spectrum, while baking soda exhibits the greatest derivative variation throughout the NIR (731- 1300 nm) and MIR (1301 – 2500 nm) regions of the spectrum.



Figure 1. Students took 10 spectroradiometer readings from each of the 10 kitchen spices and ingredients pictured above. Notice that multiple substances that appear white to the human eye were intentionally used to show students that each has unique spectral properties regardless of the fact that they appear similar in color or lack thereof.

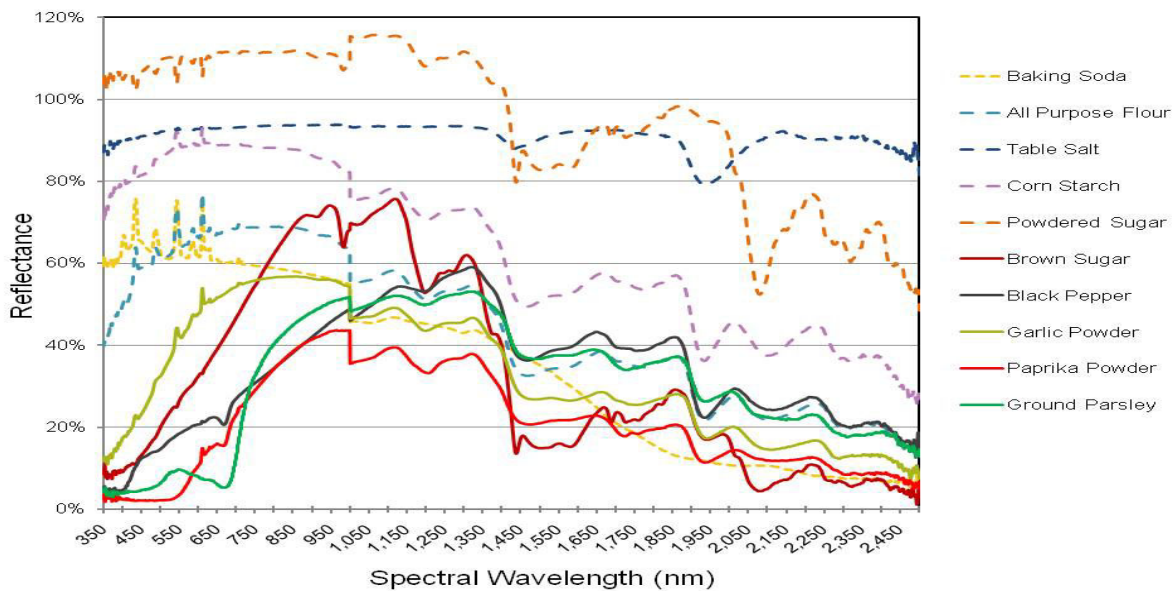


Figure 2. Spectral response curves for the 10 kitchen ingredients shown in Figure 1. The lines displayed in dashed lines are the five white ingredients. From this graph, students were able to observe the unique spectral characteristics of each of the ingredients, and how the spectra of the white ingredients even displayed considerable variation and the powdered sugar reflected more electromagnetic energy in the visible and near infrared region than the Spectralon<sup>®</sup> calibration panel. Notice each compound has unique absorption features.

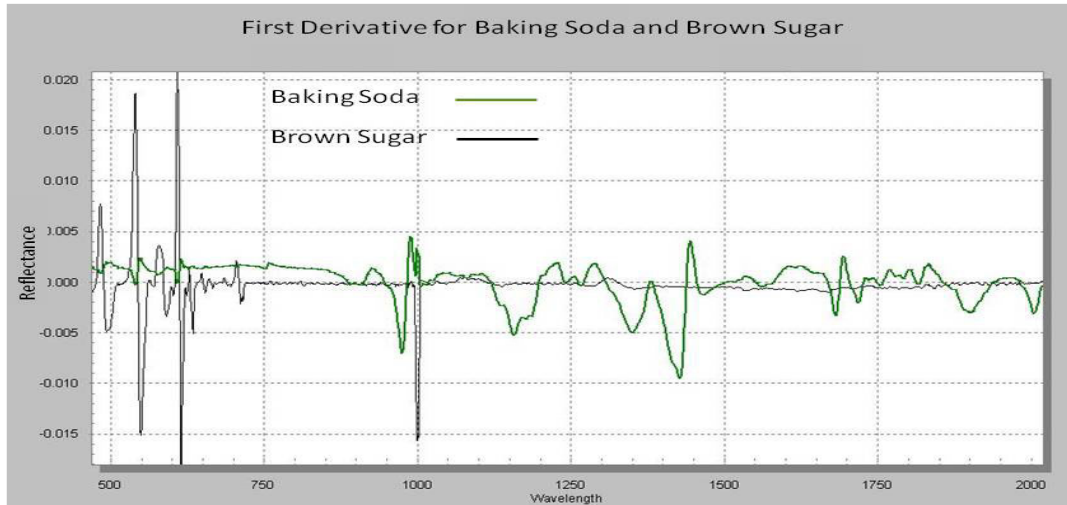


Figure 3. A comparison of first derivative curves for baking soda and brown sugar (Wavelengths are reported in nm). Note that the derivative curves for the two substances vary greatly among bands suggesting that spectral regions where these components most strongly reflect and absorb energy varies greatly. This suggests that differences in these components such as compound purity might be ascertained using different regions of the spectrum.

## STUDENT PROJECTS SUMMARIES

### **PROJECT 1 - Modeling Nitrogen Content of Grain Sorghum Flag Leaves**

#### Introduction

Grain Sorghum (*Sorghum bicolor*) is one of the most drought and stress tolerant crops grown in Kansas. For this reason, much of the sorghum is grown in high risk environments where other crops are more likely to fail or be unprofitable. Efficient sorghum cropping systems should produce high yields and use nutrients efficiently while limiting the risks and giving farmers more flexibility in making fertility decisions.

The price of nitrogen (N) fertilizer has increased substantially in recent years. Practices that allow farmers to assess crop production potential as late in the growing season as possible after planting but before applying costly inputs like fertilizer, can increase the potential for a profitable return on those inputs in risky environments. Currently, most sorghum growers routinely apply all their N fertilizer prior to planting, sometimes as much as 6 months prior.

The current KSU nitrogen recommendation is yield goal based and performs well when the grower is able to predict yields six months or more in advance of harvest. Because long range weather and yield predictions are not reliable, risk can be reduced by deferring applications until later in the season when yield can be more accurately predicted. Therefore a question this project addresses is whether spectral remote sensing methods can be used to assess N status for grain sorghum crop management planning?

## Methods

An Analytical Spectral Devices, Inc. (ASD) FieldSpec Pro FR (Boulder Colorado) was used to collect spectral readings from 21 sorghum flag leaves that has been oven dried and finely ground. The instrument was set to take the average of 10 reading for each of the 21 samples, so each of the 21 values represents an average of 10 spectral readings. The spectral measurements were entered into the ViewSpec Pro software converted to values of reflectance (Figure 4), examined for erroneous readings and exported into a file format compatible for entry into the Thermo Science GRAMS chemometric software. The laboratory derived N and spectral values were used as inputs into the GRAMS software to create a multiple regression model for predicting leaf N.

## Results and Conclusion

Figure 5 shows the multiple regression model developed by the GRAMS software. In building the model, seven spectral principal components were generated and used to explain 80% of the variance in leaf N ( $r = 0.89$ ). Based on an analysis of the factor loadings, the red and NIR wavelengths were most highly correlated with the components that most strongly predicted leaf N. The method shows great promise, but will need to be tested under actual field conditions to determine the robustness of the model in the field.

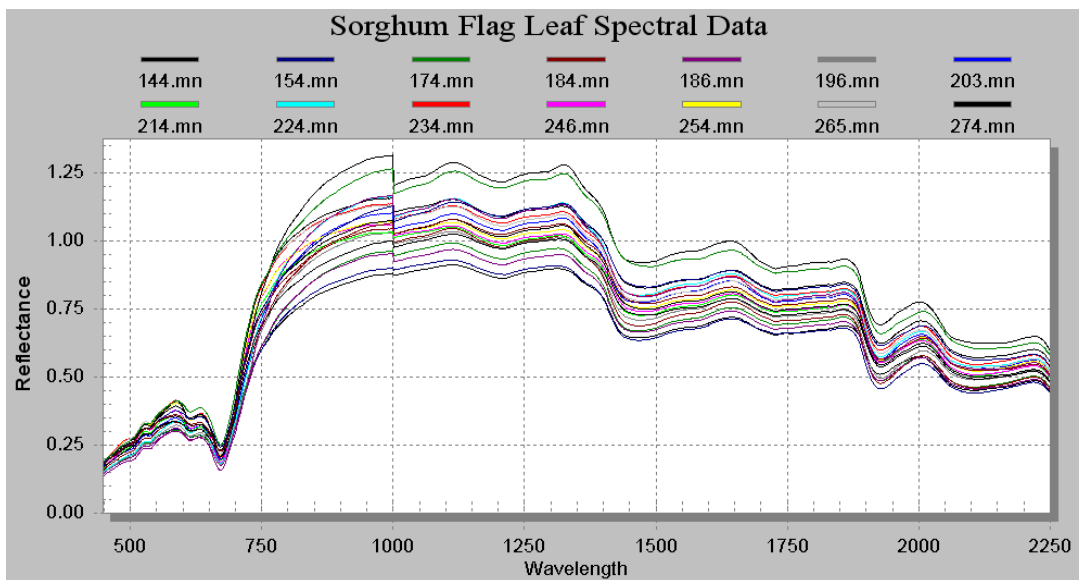


Figure 4. The mean spectral response curves for the 21 spectral measurements of the sorghum flag leaves (the legend only lists 12 of the 21 spectra and the numbers used in the legend are simple a numerical file names given to the mean of each sample). The curves show a spectral gradient that is later shown to be correlated to variation in N concentrations. (The spectral anomaly seen at 1000 nm is due to a sensor detector change in this region of the spectrum and heterogeneity in the dry ground leaf material).

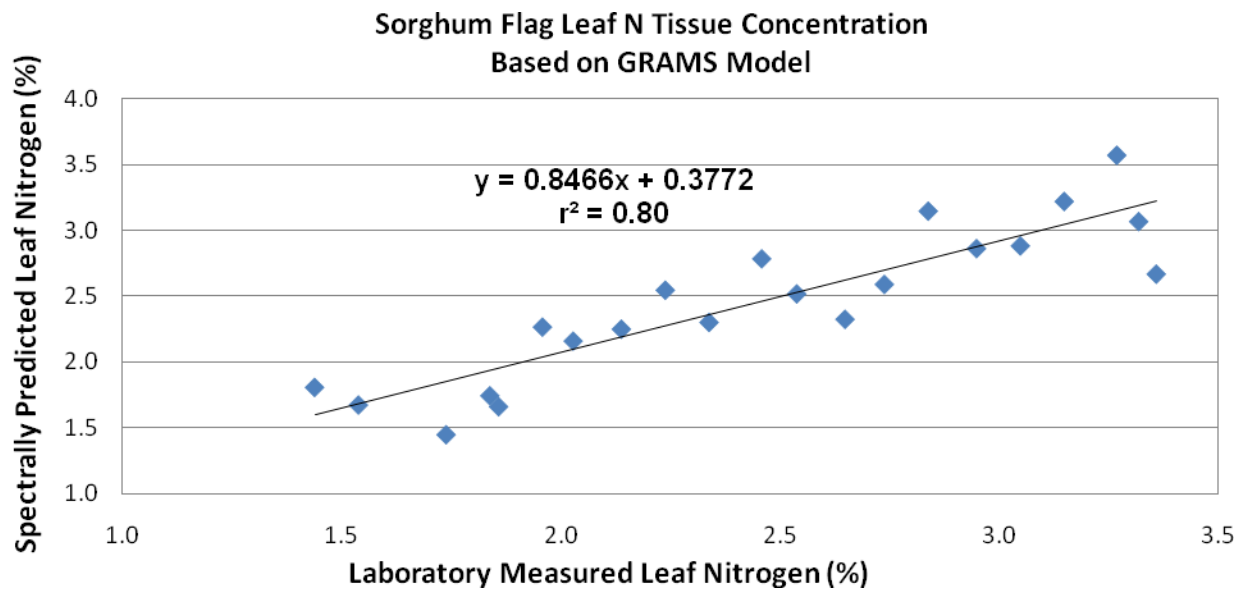


Figure 5. Shows the GRAMS regression analysis results in which spectral components were used to model sorghum leaf N. The model explained about 80% of the variation in leaf N.

## PROJECT 2 - Predicting Cover of Senescent Vegetation Using Hyperspectral Data

### Introduction

Many grassland animal species, especially birds and small mammals utilize senescent vegetation for nesting cover. Senescent vegetation is also an important component in building organic matter into soils and this dry matter is necessary for carrying grassland fires, which are necessary for maintaining the Great Plains prairie ecosystem. Unfortunately, quantifying standing dead grass residue is difficult and not feasible over large geographic areas.

In this project, we look at the use of hyperspectral measurements for estimating percent cover of grass on dark and light soil backgrounds. We hypothesized that the presence of green vegetation and variable soil backgrounds will diminish our ability to spectrally quantify dry grass cover. As part of the study we attempt to determine the optimal regions of the spectrum for quantifying dry grass cover.

### Methodology

Spectroradiometer measurements were taken across a cover gradient of senescent (dry), green, and mixed (combinations of senescent and green grass) smooth brome (*Bromus inermis*) leaf clippings. Spectra were collected using an ASD FieldSpec 3 spectrometer with a spectral range of 350 nm to 2500 nm and a sampling interval of about 1 nm. The end of the spectroradiometer fiber optics cable was mounted in the laboratory in a nadir viewing position high enough above the target area to yield a 50 cm diameter field of view (Figure 6). Digital pictures of each grass leaf clipping configuration that was sampled with the spectroradiometer were collected and used to estimate cover. The grass clippings were distributed across a dark and light soil background. The soils and leaf clippings were illuminated using a 14.5 volt 50 watt ASD Pro Lamp mounted on a tripod positioned approximately

30 cm above the target. Laboratory lights were turned off while spectral readings were being taken to eliminate incidental light contamination (Figure 7 a and b).

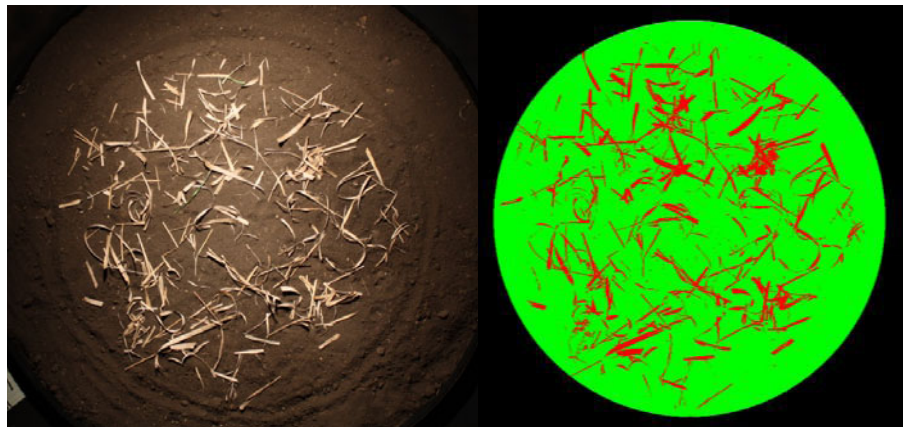


Figure 6. Shows the mounting stand with the pistol grip in which the fiber optic cable of the spectroradiometer is inserted. (Spectroradiometer is not shown)

### Data Collection

Spectroradiometer readings were collected for three grass blade configurations on two soil background types (dark and light) and across a 0% to 150% cover gradient using approximately 10% cover intervals (e.g., 10%, 20% . . . 150%). The initial cover estimates were made ocularly then verified using the classification of the digital images. The spectroradiometer was configured to take the average of 10 readings each time a sample was taken. The average of two samples for each grass cover configuration was used to derive the spectral characteristics of each grass cover interval and soil

background.



a)

b)

Figure 7 a and b. a) Dark soil background with senescent smooth brome grass blades laying on top of the soil. The target is illuminated using a 14.5 volt 50 watt lamp emitting light in the 300 – 2500 nm region of the spectrum. b) Example of the grass blade classification results for the senescent grass in the picture to the left (red is grass and green is the soil background).

A more rigorous estimation of the grass blade cover was derived using the digital pictures that were classified using the ENVI image processing software. The spectral statistics were derived using a supervised method for generating spectral statistics and submitted to a Maximum Likelihood Classifier to assign pixels to grass or background classes.

## Spectral Analysis

The spectral measurements were entered into the ViewSpec Pro software converted to values of reflectance that were displayed to inspect for differences among the samples to see how different the curves were for samples on different soil backgrounds and varying cover classes. The displayed curves were used to look for broad regions (e.g., ~ 50 - 100 nm) of the spectrum that might be used on spaceborne remote sensing systems for future modeling of senescent grass cover. The spectral measurements were exported from the ViewSpec Pro software and imported into the Thermo Science GRAMS chemometric software where a partial least square (PLS) regression equation was generated and used to estimate grass cover. The model was built using wavelengths ranging from 400 nm to 2250 nm, with the exception of the atmospheric water absorption bands around 1400 and 1900 nm.

The factor weight output from the PLS regression was used along with the visual inspection of the spectral response curves to select regions of the spectrum that provided the best discrimination among varying grass cover amounts. The average reflectance values in these selected regions were computed and exported to Microsoft Excel where they were plotted against grass cover estimates derived from the classification of the digital images.

## Results

Figure 8a shows the spectral response curves for dry, green and mixed dry and green grass on top of a dark soil. These curves show a definite spectral gradient associated with varying grass leaf cover configurations (dry, green, dry/green mix).

Spectral discrimination is greater among green and mixed grass cover than on dry grass cover. All grass cover types (green, mixed, and dry) were discriminable on the dark soil background. Figure 8b shows the spectral response curves for the same grass cover gradients laid against a light soil. Spectral discrimination is greatly diminished for dry grass when laid against the light soil where spectral contrast between the grass and soil is reduced. The region of the spectrum where spectral discrimination was most prominent also contracted to an area of the spectrum of about 1000-1350 nm.

Thermo Science GRAMS PLS regression analysis results show that the spectral region from 1250-1300 nm was particularly sensitive to changes in grass cover in this experiment. Spectroradiometer band values in this middle infrared (MIR) region were averaged and regressed against the grass cover that was estimated from the digital images. From this analysis, two distinct trends (models) emerged -- one for grasses on dark soils and one on the light soils (Figure 9). The MIR values explained 91% ( $r^2 = 0.91$ ) of the grass cover variance on the dark soils and 95% ( $r^2 = 0.95$ ) on the light soils. The difference in the slopes of the regression lines for the dark and light soils indicate that different regression models must be used to accurately estimate grass cover on dark versus light soils.

Figure 10 a and b provides a comparison between grass cover estimates using the MIR wavelengths and the Normalized Difference Vegetation Index (NDVI) that is frequently used to estimate biomass and net primary productivity of vegetation.

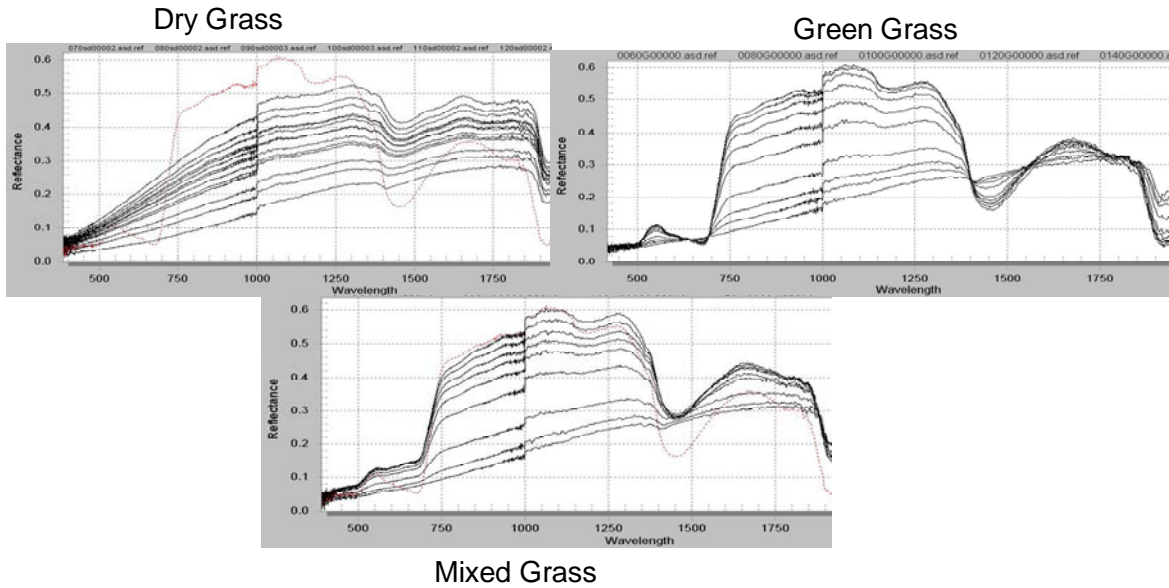


Figure 8a. Spectral response curves for the three grass treatments over a dark soil background. The red line plotted on the dry and mixed grass graphs serves as a reference for the maximum cover of green grass. As grass cover increases the reflectance values in the NIR/MIR regions of the spectrum increases.

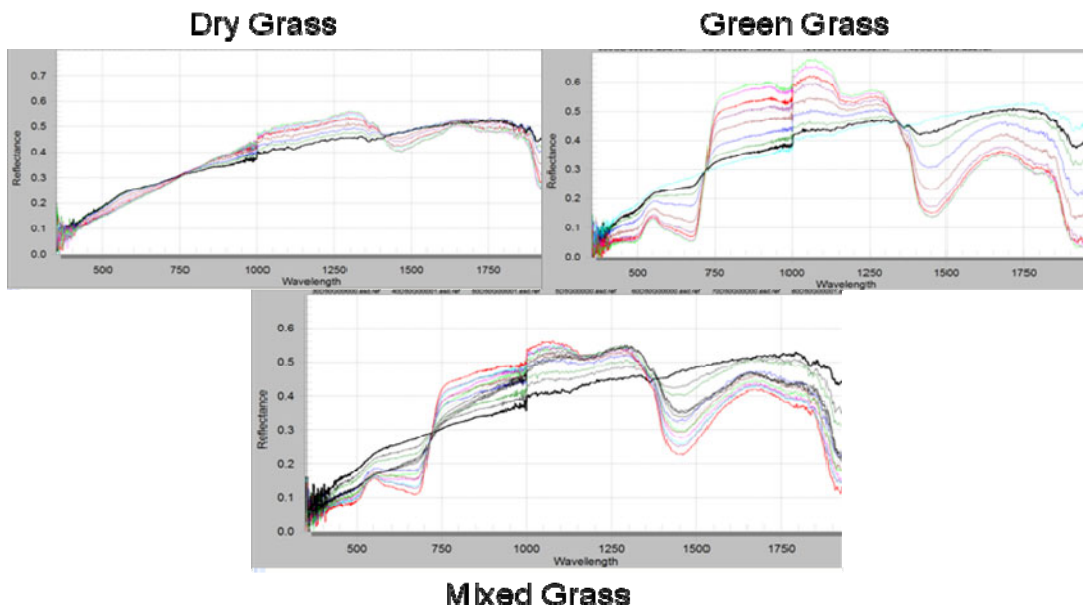


Figure 8 b. Spectral response curves for the three grass treatments over a light soil background. The red line plotted on the dry and mixed grass graphs serves as a reference for the maximum cover of green grass. As grass cover increases the reflectance values in the NIR/MIR regions of the spectrum increases. The bold dark line indicates the soil with no grass cover.

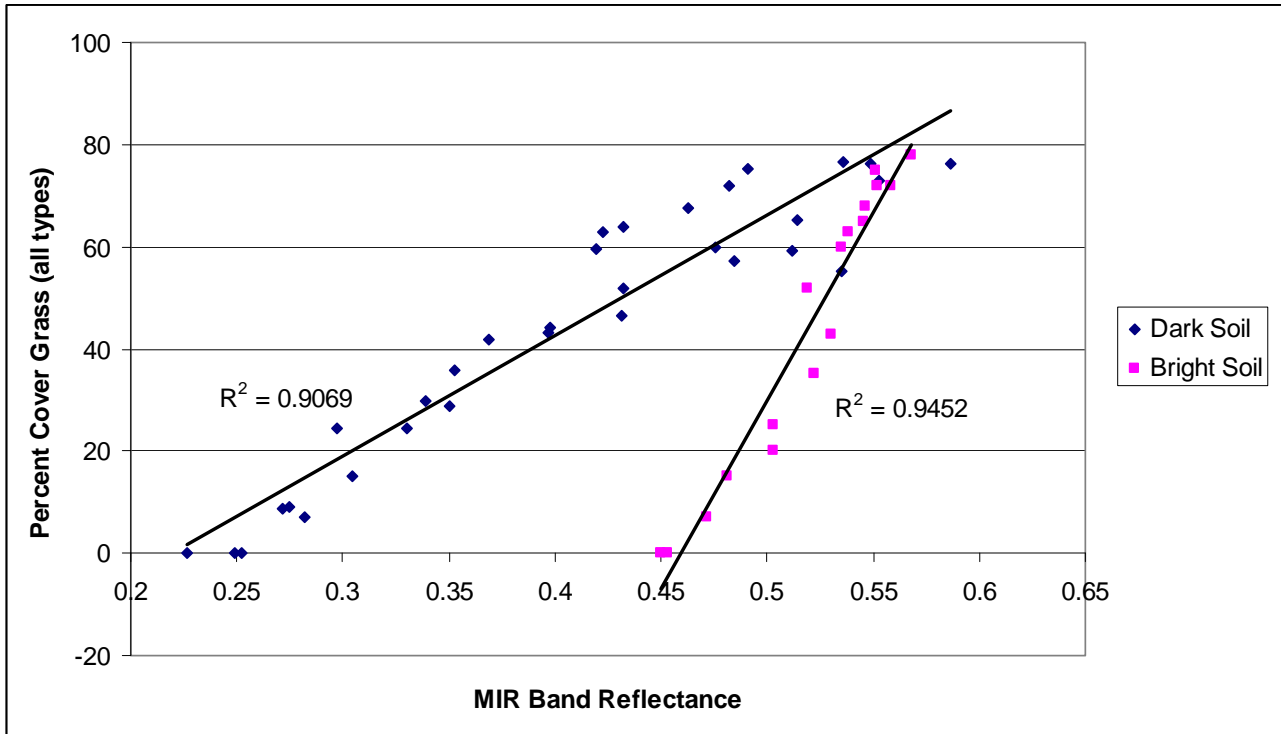
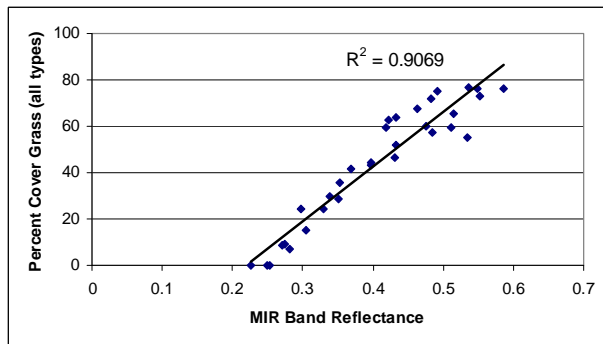
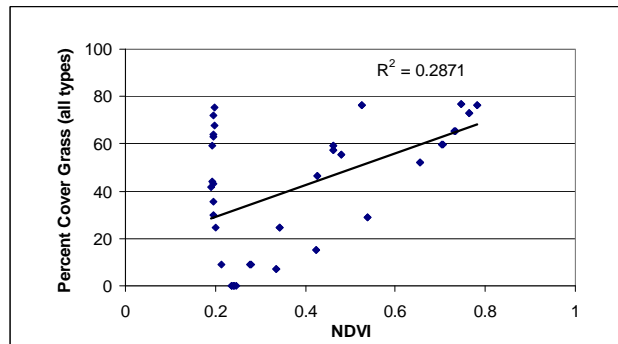


Figure 9. Regression models for the combined dry, green and mixed grass types regressed against the MIR wavelengths aggregated over the 1250-1300 nm region of the spectrum. The two regression lines highlight the effects of the soil background of the regression models.



a)



b)

Figure 10 a and b. a) This graph is used again to illustrate the relationship between MIR values and all three grass cover types on the dark soil as estimated using the digital images. b) This graph is used to illustrate the relationship between NDVI values and all three grass cover types on the dark soil as estimated using the digital images.

Figure 10 b shows the relationship between NDVI and all three grass cover types. Note three distinct linear patterns with the dry grass represented by the set of points on the left, mixed grass in the middle and green grass on the right side of the graph.

## Conclusions

Our results show how only two complicating factors—presence of green vegetation and varying soil background caused considerable variation in the ability to spectrally discriminate differences in dry grass cover. Different regression models are required for dark versus light soil background types. The spectral region of 1250- 1300 nm was identified as optimal for estimating grass cover. NDVI was shown to be a very poor estimator of dry grass cover, and a different model is required for green and mixed grass cover estimates.

## **PROJECT 3 - Hyperspectral Analysis of Sorghum Morphological Parts and their Nutrient Content**

### Introduction

Laboratory chemical analysis of plant tissue nutrients is a laborious and time consuming process. Development of a chemometric method for rapidly and non-destructively quantifying plant tissue nutrients would avoid the complex sample preparation process, reduce the need for chemicals, and could be used to quantify multiple nutrient properties simultaneously thus saving time and money.

The objective of our project is to evaluate the use of chemometric analysis methods for quantifying plant tissue nitrogen, phosphorus and potassium in sorghum leaf, stem and roots.

### Material and Method

In 2008, sorghum hybrid plants were grown in a green house under different fertilizer rate treatments. After harvest, the leaves, stems and roots were dried and ground to fine texture for subsequent chemical analysis. A tissue sample of the roots, leaves and stems was sent to the soil chemical analysis laboratory in the Department of Agronomy at KSU for chemical analysis. Fifteen ground tissue samples for the leaves, stems and roots parts, representing the high to low nutrient gradient, were placed in petri dishes and taken to the Ecology and Agriculture Spatial Analysis Laboratory (EASAL) in Agronomy at KSU for hyperspectral and chemometric analyses. Twenty spectroradiometer readings (each reading was generated from the average of 10 readings automatically taken by the spectroradiometer) were collected at random reading locations within each petri dish for a total of 300 (15 x 20) samples per plant part or total of 900 (3 tissue types x 300) readings.

The spectral data were imported into ViewSpec Pro where they were transformed to reflectance values and viewed for errors. The reflectance values were then exported to the Thermo Science GRAMS software program for model development.

### Result and Discussion

#### 1. Average Spectral Character of Leaves, Stems and Roots

The spectral character of each sorghum morphological part was averaged and plotted as shown in Figure 11. As can be seen in the figure, the ground and dried plant tissue for the three morphological parts are spectrally unique mostly in the 500 to 1250 nm region of the spectrum. In this region, sorghum stem has a higher reflectance than the root and the leaf parts. Leaf and root could be distinguished from one another in the region between 700 and 1250. In this region the leaf has high reflectance than the root.

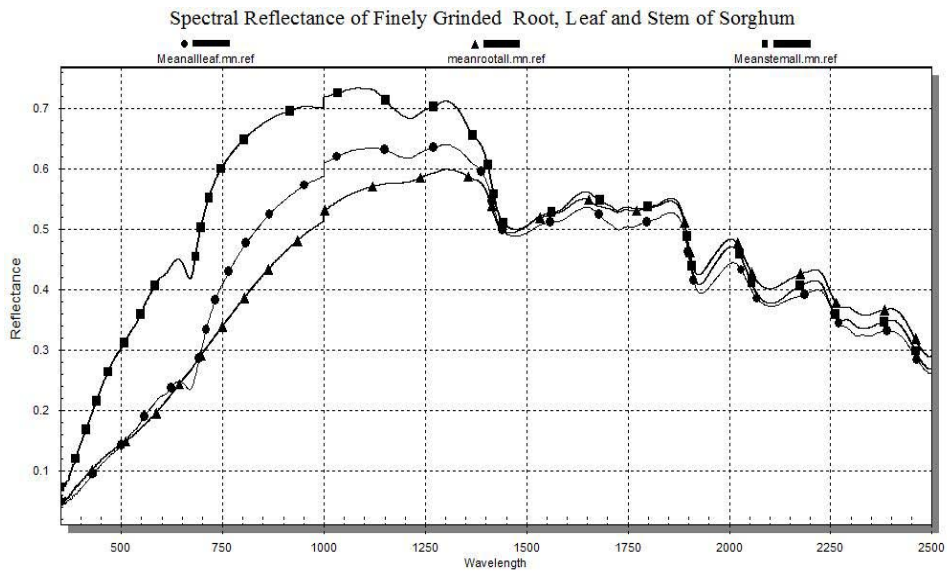


Figure 11. Shows the spectral response curves for dried and ground sorghum leaf (●), stem (■) and root (▲) tissues.

## 2. Leaf at different nutrient contents

The spectral response curves for dry sorghum leaf tissue across a nitrogen gradient are shown in Figure 12. Figure 13 shows regression analysis results for the relationship between spectroradiometer measurements and leaf N, P, and K.

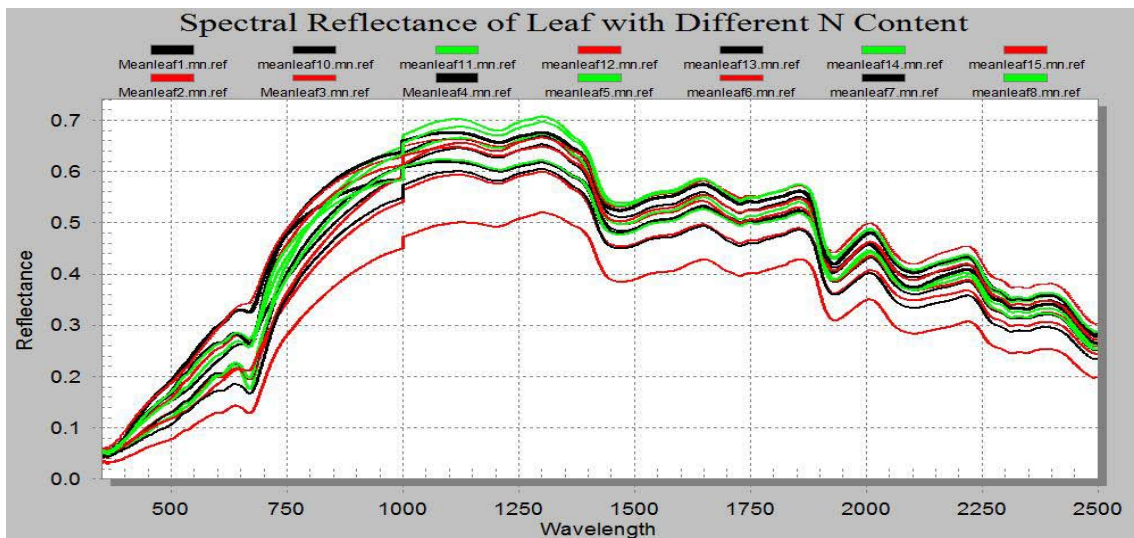
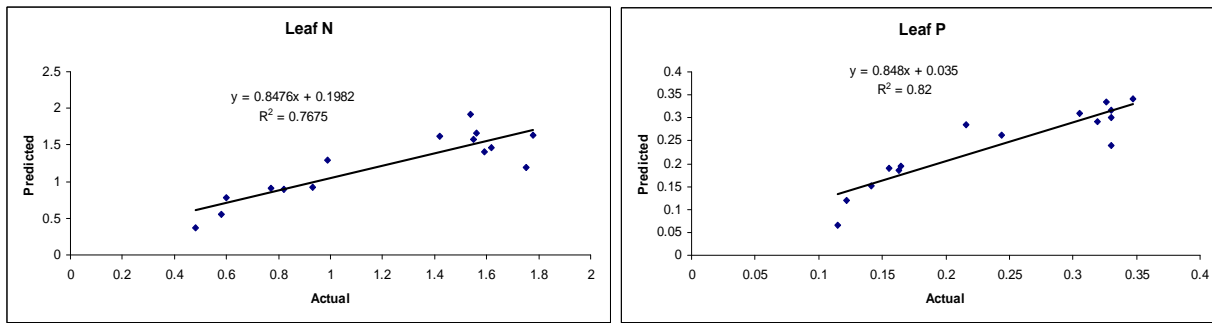
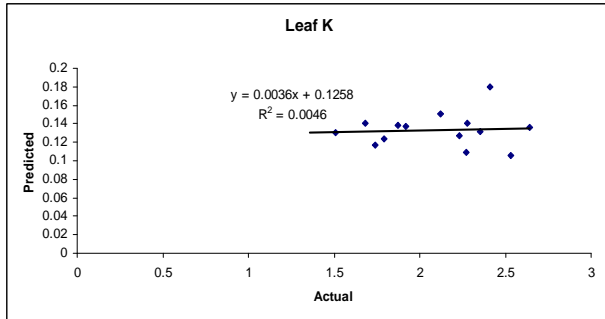


Figure 12. The spectral characteristic of sorghum leaf tissue across a nitrogen gradient.



a)

b)



c)

Figure 13. Results from GRAMS regression analysis in which the relationship between leaf N (graph a), P (graph b) and K (graph c) are examined. The results show a good fit for N ( $R^2 = 0.77$ ), and P ( $R^2 = 0.82$ ) and no relationship for K ( $R^2 = 0.0046$ ).

### 3. Root at different nutrient contents

The spectral response curves for dry sorghum root tissue across a nitrogen gradient are shown in Figure 14. Figure 15 shows regression analysis results for the relationship between spectroradiometer measurements and root N, P, and K. In general, the spectral/nutrient relationship is weaker for the root tissue than leaf tissue.

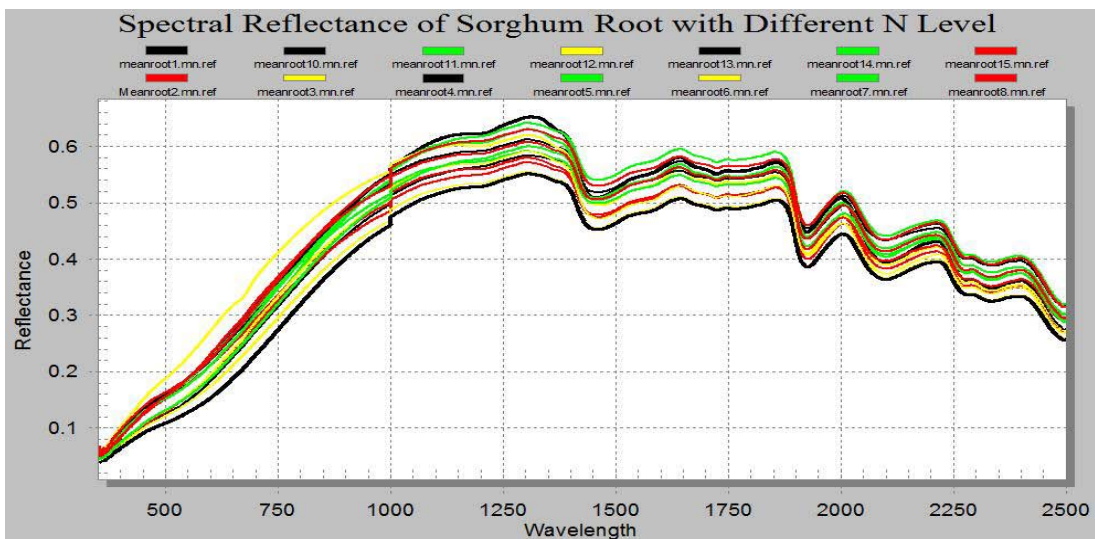
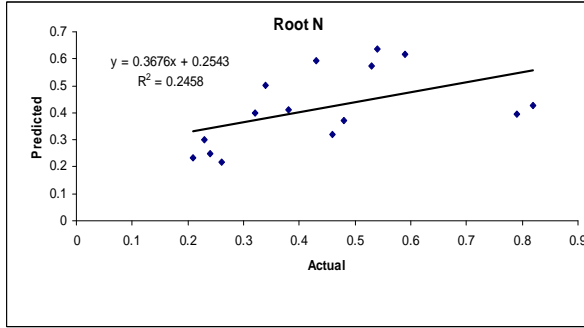
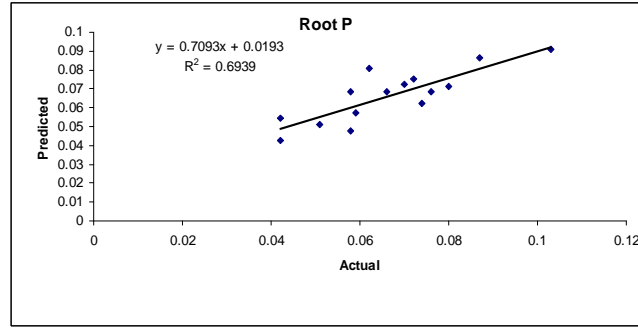


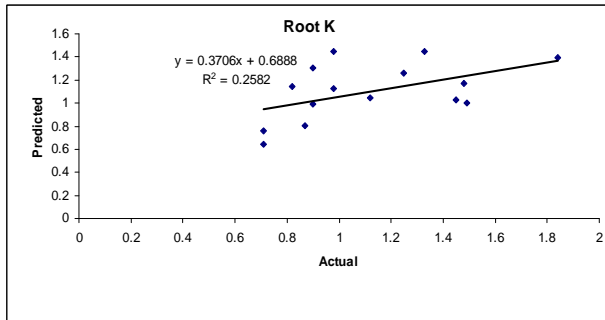
Figure 14. The spectral characteristic of sorghum root tissue across a nitrogen gradient.



a)



b)



c)

Figure 15. Results from GRAMS regression analysis in which the relationship between root N (graph a), P (graph b) and K (graph c) are examined. The results show a weak fit for N ( $R^2 = 0.25$ ), a good fit for P ( $R^2 = 0.69$ ) and weak relationship for K ( $R^2 = 0.26$ ).

#### 4. Stem at different nutrient contents

The spectral response curves for dry sorghum stem tissue across a nitrogen gradient are shown in Figure 16. The regression analysis results for the relationship between spectroradiometer measurements and stem nutrients show a good fit for N and poor fits for P and K (Figure 17).

#### Summary

Chemometric derived models can be used to predict sorghum nutrients in leaves, stems, and roots to varying degrees. Nitrogen and phosphorus were most strongly correlated with spectral measurements, while potassium was mostly weakly correlated. The strongest relationships were generally with leaf tissue.

#### Overall Paper Summary

Educational recourses for use in a remote sensing chemometrics course are difficult to find, with no textbooks on the subject identified. Those who wish to teach such a course will have to rely on the use of articles written on the topics of interest.

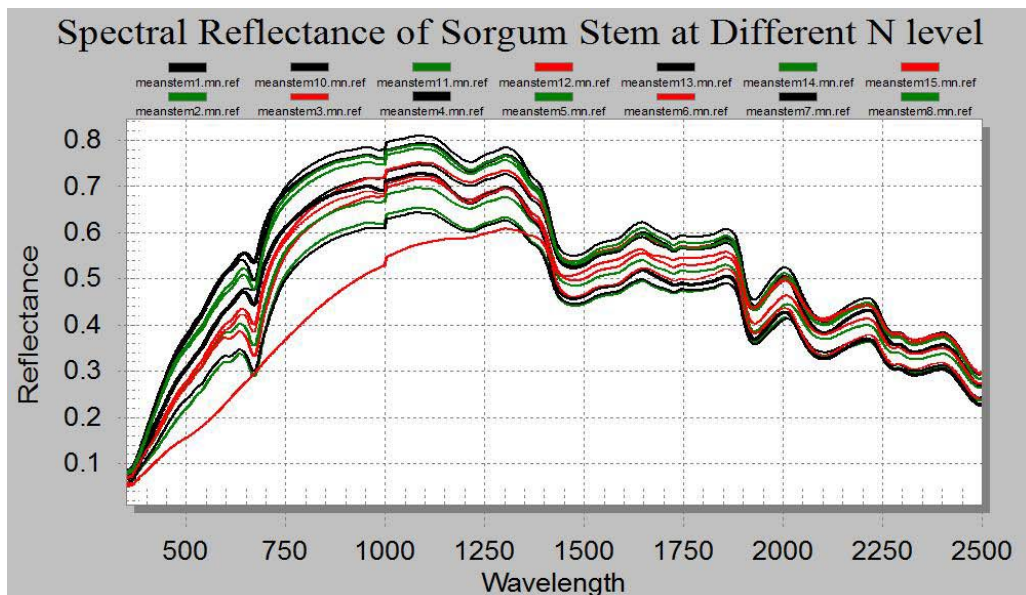


Figure 16. The spectral characteristics of sorghum stem tissue across a nitrogen gradient.

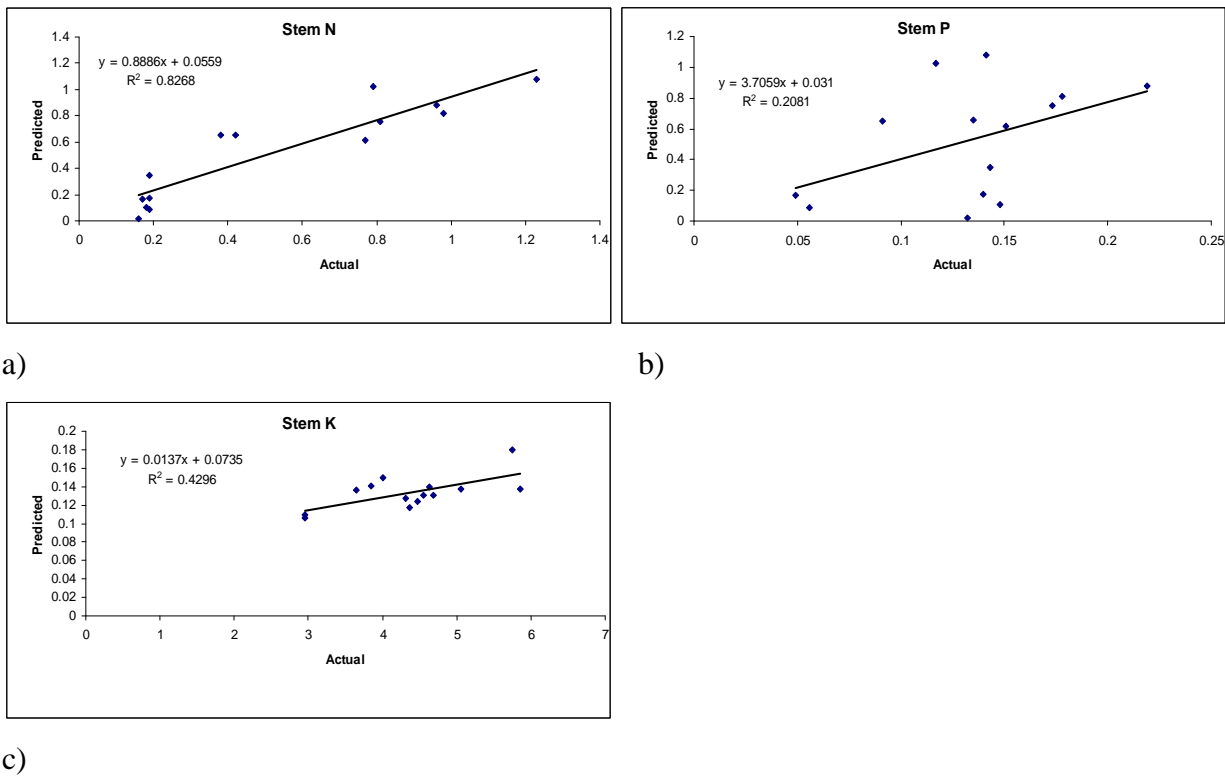


Figure 17. Results from GRAMS regression analysis in which the relationship between root N (graph a), P (graph b) and K (graph c) are examined. The results show a strong fit for N ( $R^2 = 0.82$ ), a weak fit for P ( $R^2 = 0.20$ ) and moderate linking for K ( $R^2 = 0.43$ ).

Based on student feedback, class members liked the course format and appreciated the opportunity to get to actually hands-on experience using a spectroradiometer, and the analysis capabilities of ViewSpec Pro and GRAMS chemometrics software.

Student projects generated interesting results and gave students an opportunity to address research questions of interest to them and others.

#### Literature Cited

Introduction to NIR Technology. 2005. Analytical Spectral Devices, Inc. 10 pp.