EVALUATING STANDARD WET CHEMISTRY TECHNIQUES AND NIR SPECTROSCOPIC MODELS FOR DETERMINING COMPOSITION AND POTENTIAL ETHANOL YIELDS OF MULTI-SPECIES HERBACEOUS BIOENERGY CROPS

A Thesis
Submitted to The Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Ewumbua Menyoli Monono

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Agricultural and Biosystems Engineering

July 2011

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Evaluating Standard Wet Chemistry Techniques and NIR Spectroscopic Models for Determining Composition and Potential Ethanol Yields of Multi-species Herbaceous Perennial Bioenergy Crops

By

Ewumbua Menyoli Monono

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

ABSTRACT

Monono, Ewumbua Menyoli, M.S., Department of Agricultural and Biosystems Engineering, College of Engineering and Architecture, North Dakota State University, July 2011. Evaluating Standard Wet Chemistry Techniques and NIR Spectroscopic Models for Determining Composition and Potential Ethanol Yields of Multi-Species Herbaceous Bioenergy Crops. Major Professor: Dr. Scott W. Pryor.

Herbaceous perennials represent a considerable portion of potential biomass feedstocks available for the growing bioenergy industry. Their chemical composition and biomass yields, which are important in determining ethanol potential on an area and mass basis, vary with plant variety and type, environment, and management practices. Therefore, a study was conducted to assess the variability of lignin and carbohydrate content, biomass yields, and theoretical ethanol yields on an area basis among different herbaceous perennial species combinations grown in Minot (2008) and Williston (2008, 2009, and 2010), North Dakota (ND).

After wet chemistry compositional analysis was done, the carbohydrate contents were used to determine theoretical ethanol potential on a mass basis. Using the dry-matter yield, the theoretical ethanol yield on an area basis was also calculated for these biomass species. Total carbohydrate content for the biomass samples in Williston and Minot varied from 45 to 61% dry basis. Analysis of Variance (ANOVA) at $\alpha = 0.05$ showed that carbohydrate content varied between years and environments. Also an interaction plot shows that no biomass species had consistently higher or lower carbohydrate content in the different environments. Switchgrass (*Panicum vigatum* L.) grown as single species or together with other perennial grasses had higher dry-matter yield and theoretical ethanol yield potential in Williston irrigated plots while mixtures containing intermediate or tall wheatgrass species (*Thinopyrum* spp.) produced better yields in Minot non-irrigated plots.

Variability in theoretical ethanol yield on a mass basis (3.7% coefficient of variation (CV) in Williston and 9.7% CV in Minot) was much less than the variability in dry-matter yields (27.5% CV in Williston and 14.8% CV Minot). Therefore, biomass production is much more important than composition in choosing species to grow for ethanol production.

Recently, many studies have focused on developing faster methods to determine biomass composition using near infrared (NIR) spectroscopy. Other NIR models have been developed on single biomass feedstocks but a broad-based model for mixed herbaceous perennials is yet to be developed. Therefore, NIR calibration models for lignin, glucan, and xylan were developed with 65 mixed herbaceous perennial species using a DA 7200 NIR spectrometer (950 – 1,650 nm) and GRAMS statistical software. The models for lignin and xylan had R^2 values of 0.844 and 0.872, respectively, upon validation and are classified as good for quality assurance purposes while glucan model had an R^2 of 0.81 which is considered sufficient for screening. The R^2 and the root mean square error of prediction (RMSEP) results showed that it is possible to develop calibration models to predict chemical composition for mixed perennial biomass when compared with results for models developed for single feedstock by Wolfrum and Sluiter (2009) and Liu et al. (2010). Studying the variability in predicting constituents using NIR spectroscopy over time (hours and days), it was observed that the average CV was between 1.4 to 1.6%. The average CV due to repacking (presentation) alone was 1.3%. The CVs for NIR predictions ranged between 1.4 to 5.7% while for wet chemistry ranged between 3.8 to 13.5%; hence, NIR predictions were more precise than wet chemistry analysis.

ACKNOWLEDGEMENTS

I would like to thank a number of people who have contributed toward this research. Firstly, thanks to my advisor Dr. Scott W. Pryor for the opportunity he offered me to work in renewable fuels and his guidance during the course of my research work.

Thanks to my committee members, Dr. Dennis Wiesenborn, Dr. Zhulu Lin, and Dr. Marisol Berti, for their help and corrections during my thesis proposal and defense. Additional thanks to Ms. Nurun Nahar, Dr. Darrin Haagenson, Dr. Bhavnita Dhillon, and Dr. Bishnu Karki for their assistance in understanding procedures and equipment in the pilot plant and laboratory.

I would also like to thank Dr. Paul Nyren, Mr. Gordon Bradbury, and Mr. Mark Halvorson from the Central Grassland Research and Extension Center in Streeter, and NDSU Research Extension Centers in Williston and Minot, respectively, for their support in obtaining biomass samples and information about their growing conditions. I also extend thanks to the department staff: Jana Daeuber, Lori Sholts, and Nancy Stroh for their help with computer and office needs.

Great thanks to my wife, Mrs. Esther Monono and uncles, Mr. Steve Ndima, Mr. Gustave Efotte, Mr. Molua Lambe, and Mr. Mbella Nganje for their family support and encouragement.

Finally, I would like to thank my friends, Mr. Ebot Tabe, Dr. Mafany Mongoh, Dr. Diomo Motuba, and Mr. and Mrs. Emmanuel and Amy Nojang for all their support.

DEDICATION

To my son, Etina Lifongo Monono.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

1.0. BACKGROUND

The increasing demand for energy, the instability and uncertainty in petroleum resources, and environmental concerns associated with the use of fossil fuels have made the development of alternative energy a priority. The United States (US) is the largest consumer of petroleum products using 19.1 million barrels per day (MBD) in 2010 where 11.8 MBD (49% net import) of the total consumption is imported (EIA, 2011). Petroleum products in the US contributed 41.7% of the total carbon dioxide (CO_2) emissions in 2010 (Lindstrom, 2011). Hence, the increase in demand for transportation fuel and other fossil fuels will also increase CO_2 emissions. Many renewable energy resources (solar, wind, biomass, geothermal, etc) have been identified to mitigate CO_2 emissions currently available, but ethanol produced by fermenting 6-carbon or 5-carbon sugars is the main alternative for transportation fuel. These simple sugars can be obtained from biomass, starchy grains, and sugar crop feedstocks. Presently, ethanol is produced from corn in the US, sugarcane in Brazil, and sugar beets in France. However, there is a growing concern for food competition (Mitchell, 2008) and issues of environmental degradation during the production of these food crops (Pimentel and Patzek, 2008). These competition and environmental problems have created an increased interest in the use of cellulosic feedstock materials for ethanol production.

Cellulosic materials including wood, grasses, and agricultural residues are estimated to be available at sustainable harvest rates of 1.36 billion Mg per year in the US (Perlack et al., 2005). This estimate takes into account biomass to be left in the field to resolve environmental problems such as erosion, soil fertility, and water quality. The 2007 Energy Independence and Security Act mandates that by 2022, 79.5 billion liters (21

billion gallons) per year (BLY) of biofuel production be derived from cellulosic biomass, which includes 60.6 BLY of cellulosic biofuels and 18.9 BLY of other advanced biofuels (EIA, 2008). Perennial grasses, a second generation energy crop, represent approximately 28% of the total 1.36 billion Mg/year of US biomass potential. Currently, there is no large scale economical production of perennial grasses in the US for biofuel. To achieve the potential biomass production, 20 million hectare (ha) of US cropland, idle cropland, and pasture would be used for the production of perennial grasses (Perlack et al., 2005).

Perennial grasses offer many advantages compared to other cellulosic biomass like wood and corn because they can grow in diverse geographic environments with high net energy yields (Downing et al., 1995). Relatively high yields can be obtained even when perennial grasses are planted on poor soils because of their low nutrient demand; they can also provide important soil and water conservation benefits to erosion-prone land (Agblevor and Besler, 1996; Downing et al., 1995). Even with biennial or annual harvesting, perennial grasses improve long-term sustainability on lands by reducing erosion and net greenhouse gas emissions. In addition, perennial grasses require less energy to maintain once they are established and annual or biennial harvesting can be done over a period of 10 to 20 years (Nyren et al., 2007).

The organic components (primarily cellulose, hemicellulose, and lignin) available in plant biomass are vital resources in biochemical and thermochemical conversion processes to produce biofuel or heat energy. The inorganic components present cannot be converted to energy or used for biofuel production. High amounts of inorganic components reduce the relative amount of organic components and increase the quantity of ash which is undesirable for thermal processes (Yaman, 2004). Cellulose and hemicellulose, which are

sugar polymers embedded in the plants cell walls, are broken into simple sugars through hydrolysis. These simple sugars are then fermented to ethanol by yeast. Lignin which is also a large portion of the organic constituent is mostly burned for process energy in a biochemical conversion process (McKendry, 2002a).

Research to improve pretreatment technology to increase yields of 6-carbon sugars, fermentation processes to increase utilization of 5-carbon sugars, and thermochemical processing to include use of lignin in biofuel production will make the use of cellulosic biomass feedstocks economically and environmentally better than the use of first generation feedstocks (Himmel et al., 2007). Knowledge of biomass chemical composition, both organic and inorganic components, is a vital step towards determining the expected yield of ethanol and/or heat energy from burning of biomass. The composition of biomass structure is built during the growth phase of the plant. Therefore, the conditions upon which the plant is subjected will have some influence on the proportion of its constituents. The rate of polysaccharide formation during photosynthesis and the amount of inorganic substance absorbed from the soils among many other factors, influence the heterogeneity and complex chemical composition of biomass species (Adler et al., 2009; Michalet et al., 2006).

Due to these differences in biomass composition, payment for feedstock should also be based on compositional value rather than weight only. Compositional analysis of biomass is usually determined with the standard wet chemistry protocols developed by the US Department of Energy's National Renewable Energy Laboratory (NREL). However, using these protocols is time-consuming, expensive, and laborious. Therefore, in recent years, studies have been focused on developing faster methods to determine biomass

composition using near infrared (NIR) spectroscopy. Presently, no robust model has been developed to predict composition of a variety of biomass species; rather, many different models have been built for single or specific biomass varieties (Liu et al., 2010). The objective will be to study the variability in the chemical composition of perennial biomass and also to determine the potential of developing robust NIR calibration models that will predict the chemical composition of multi-species herbaceous perennials.

2.0. LITERATURE REVIEW

2.1. Perennial grass as bioenergy crop

2.1.1. Growth, climate and ecological adaptability

Perennial grasses have always been used as feed for livestock ruminants all over the world due to their ability to grow and adapt in diverse climatic and ecological conditions. The advantages of perennials over grain crops as biofuels feedstock include that they require lower inputs, reduce net greenhouse gas emissions, and produce more energy per dry ton (Sanderson and Adler, 2008). Perennial grass species exhibit different growth potentials in different climatic and ecological conditions. Perennial grasses can be classified as C3 or C4 plants; C3 plants strive best in cold region and C4 types dominate the warm regions (Elberson, 2002). The names C3 and C4 are derived from the difference in photosynthetic pathways of how the plants assimilate carbon dioxide (CO_2) to form simple sugars. The C3 plants fix CO_2 directly from the atmosphere to form 3-phosphyglycerate with the aid of the RuBisCo enzyme while C4 plants have an additional pathway where phosphoenolpyruvate (PEP) carboxylase enzyme fixes CO_2 into oxaloacetate which is later decarboxylated to the normal C3 pathway. The C4 plants have higher maximum photosynthetic conversion efficiency than C3 plants because PEP carboxylase has higher affinity for CO_2 than RuBisCo, therefore, CO_2 will still be fixed at high temperature when the stomata opening reduces (Klooster and Palmer-Young, 2004). This will cause C4 plants in low-water environment to loss less water and get more CO_2 than the cool-season (C3) plants (Bloom et al., 1985). Some C3 plants include wild rye (*Elymus* spp.), alfalfa (*Medicago sativa*), sweet clover (*Melilotus officinalis*), and wheatgrass (*Thinopyrum* spp.) while some C4 plants include maize (*Zea mays*),

switchgrass (*Panicum vigatum* L.), big bluestem (*Andropogo gerardii*), and *Miscanthus* spp. (Elberson, 2002). As dry-matter yield is arguably the most important factor in determining ethanol yield per ha, planting grass species in areas where they adapt best and have high productivity is important.

It is worth noting that perennial grasses strive and adapt to various ecological conditions due to the presence of vegetative reproduction structures such as rhizomes, bulbs, and tubers that enable them to live from year to year (Chai et al., 2010). These structures are dormant in adverse climatic conditions (extreme cold or hot) but grow when conditions become favorable. The chances of continuous growth or production of perennial grass once the field is established is higher than if fields had trees and shrubs. The diverse reproductive options also allow them to tolerate wildfire as reproductive structures like rhizomes and roots are protected in the soil and will grow again as normal conditions resume (Chai et al., 2010).

2.1.2. Factors influencing the variability in perennial grass composition

Cellulose, hemicelluloses, and lignin are the major organic compounds present in the perennial grasses for the biochemical and thermal processes of ethanol production. The relative amount of these components varies among grass species. This variability is influenced by plant variety, environment, and management practices during growth. To be more specific, these factors include nutrient status of the soil (Gough et al., 2000), physiological ecotype as determined by difference in climate, latitude, and longitude (Casler et al., 2004; Partel et al., 2007), harvesting period (spring or fall) and minerals present in the soil (Adler et al., 2006) , period of morphological development (Sanderson and Wolf, 1995) and genetic information of grass species (Partel et al., 2007).

2.2. Cellulosic ethanol production from perennial grass and other biomass

2.2.1. Current production status, conversion technologies, and challenges

Ethanol is currently being produced in the US mainly using grain-based starch technologies while other countries such as Brazil and France use sucrose from sugarcane and sugar beet, respectively. However, ethanol can also be produced from biomass such as grasses, agricultural residues, and woody biomass. Biomass is from the photosynthetic reaction in plants between water, CO_2, and sunlight to produce glucose, the basic unit for cellulose. The sugar monomer units bind together by glycosidic bonds to form a complex carbohydrate structures that make up the biomass structure. Biomass is composed of cellulose, hemicellulose, lignin, extractives (chlorophyll, waxes, and minor components that dissolve in ethanol), and ash (inorganic substances) (Martone et al., 2009).

Cellulose is a polymer of β-D-glucose units joined together by glycosidic bonds to form a long, rigid structure. As cellulose constitutes 30 to 50% of most harvestable biomass material, it is considered the most abundant organic material on earth (McKendry, 2002a). Hemicellulose, which represents 20 to 35% of biomass structure, is not chemically homogenous like cellulose but is made of branched structures of pentoses (xylose and arabinose) and hexoses (mannose, glucose, and galactose) with xylose constituting the highest proportion of hemicellulose in perennial grasses (McKendry, 2002a). This branched, non-crystalline structure of hemicellulose makes it less compact and more readily hydrolyzed than cellulose.

Approximately 10 to 25% of biomass is lignin which has a dense, complex, aromatic heteropolymer structure (Martone et al., 2009). Lignin fills spaces between and around cellulose microfibrils and hemicellulose forming a complex chemical and structural

combination known as lignocellulose. The protective nature of lignin makes biomass recalcitrant to chemical or enzymatic degradation (Himmel et al., 2007).

Most of the remaining portion of biomass is composed of extractives and ash. Extractives are all components in biomass feedstock that are soluble in neutral organic solvents for example chlorophyll, nitrate/nitrites, protein, and waxes. Ash is the inorganic materials bound in the physical structure of biomass (Yaman, 2004). Some of the natural factors that have been identified in contributing to biomass recalcitrance are the degree of lignification, the arrangement and density of cellulosic bundles, heterogeneity and complexity of cell wall and its constituents, and the insolubility of biomass constituents (Himmel et al., 2007). Developing cost effective technologies to break these complex structures to simple sugars has been a major concern in cellulosic ethanol production.

Several technologies do exist to convert lignocellulosic biomass feedstocks into ethanol or other liquid fuels and these can be grouped in two main categories; the sugar platform (fermentation of biomass sugars) (Lin and Tanaka, 2006) and thermochemical platform (e.g. gasification followed by biological or chemical processing) (Perkins et al., 2008). However, these technologies are relatively new and therefore very expensive compared to the existing corn and sugar conversion technologies. Despite the high cost of these technologies, more than a dozen pilot plants or small scale ethanol plants of 7.6 to 11.4 million liters (L) per year are being developed to use these cellulosic ethanol technologies (Waltz, 2008). At the end of 2007, the US Department of Energy (DOE) provided research funds of $1 billion to develop cost effective conversion technologies with a goal of bringing production costs of cellulosic ethanol to $0.35 per L ($1.33 per gallon) by 2012 (Curtis, 2008). The success in this area of research will encourage

investors in the domain of cellulosic ethanol production which will set a pace in achieving the objective of 60.6 BLY of cellulosic biofuels by 2022.

Dwivedi et al. (2009) suggested that future conversion technologies for cellulosic ethanol will be likely based on the thermochemical platform rather than the sugar platform. Their rationale was that thermochemical technologies like gasification and catalytic conversion are more advanced and need little improvement compared to current biochemical conversion process. In spite of this, current research work on improving pretreatment and hydrolysis technologies may also promote the growth of biochemical conversion plant. Some cost effective innovations to enhance both biochemical and thermochemical technologies have been studied. These include; consolidated bioprocessing where cellulase production, cellulose hydrolysis, and fermentation are carried out in one reactor with a single organism (Cardona and Sánchez, 2007), and mobile fast pyrolysis by producing bio-oil at harvest sites to reduce transportation cost (Badger and Fransham, 2006).

2.2.2. Influence of feedstock variability on conversion technologies

Variation in yields of potential fermentable sugars for the production of ethanol was found to be influenced by the variation in carbohydrate composition and the efficiency of the acid/enzyme saccharification process (Dien et al., 2006). Regardless of the technology used, higher proportions of organic components (cellulose, hemicelluloses, and lignin) are needed to obtain higher biofuel yields. Sugar-based technologies will benefit first from higher cellulose and next from higher hemicellulose while thermochemical processing will value high lignin content because of its higher energy density. Lignin has been proven to have a negative effect on biochemical conversion process while it is a desirable component

for thermochemical conversion processes (Boateng et al., 2008). Both carbohydrate and lignin content increases with plant tissues maturity but extracting glucan become more difficult as lignin strongly protects cellulose and hemicelluloses. Increasing severity of pretreatment to obtain more accessible glucan will influence the yields of hemicellulose sugars (McKendry, 2002b). Since yield and conversion efficiency of biochemical process is greatly influenced by biomass composition, there is a need for genetic breeding to improve yield and compositional components that will help in improve biochemical conversion processes (Dien et al., 2006).

2.3. Techniques in determining biomass composition

As biomass composition is one of the most important parameters in cellulosic ethanol production, much interest is given in developing an effective tool to determine composition with accuracy and less energy input. Several techniques do exist and can be grouped into two major analytical procedures; i) standard wet chemistry, and ii) spectroscopic analysis.

2.3.1. Standard wet chemistry analysis

The International Energy Agency (IEA) validated a standardized procedure to determine chemical composition for biomass related materials which can be obtained from American Society for Testing and Materials (ASTM) (Milne et al., 1992). The Technical Association of the Pulp and Paper Industry (TAPPI) test method has also been developed for the paper and wood industries to determine many components including ash, lignin, and carbohydrates (IPS, 2010). Theander (1991) published comprehensive procedures of his methods for lignocellulosic compositional analysis before Milne et al. (1992) could publish the standard analytical procedures. The National Forage Testing Association (NFTA) also

developed standard procedure for preparing samples for compositional analysis, moisture and ash content determination. With the many procedures existing, the National Renewable Energy Laboratory (NREL) consolidated several procedures (ASTM, TAPPI and NFTA) to develop their own standard biomass analytical procedures (NREL, 2010). Recently, Sluiter et al. (2010) made a review of all the different methods that have been employed for biomass composition since 1922 and found the NREL methods were the most accurate though some suggestions for improvement were given.

There are some drawbacks with these standard methods; they cannot be used in a commercial setting requiring immediate results because the procedures are labor-intensive, time-consuming and expensive. For example, a complete analysis for a single sample will cost between $800 and $2,000 and the results will not be available for days or weeks (Hames et al., 2003). A fast analytical approach that is less expensive and less labor intensive can be developed by using multivariate analysis that relates chemical composition with spectroscopic data. The most laborious work with spectroscopic approach is during the development of statistical model but once a model is developed, chemical prediction can be obtained within seconds (Blanco and Villarroya, 2002). Though standard wet chemistry is laborious and expensive, it remains the reference standard for NIR calibration model development.

2.3.2. Spectroscopic analysis

Spectroscopy is a study of the interaction between radiation and matter as a function of wavelength or wave number. Light or mass spectra generated from spectroscopy is used to determine or analyze properties (atomic, molecular, or ionic) or concentration of a given chemical substance or sample (Siesler, 2008). Many different

types of spectroscopic tools have been used to determine biomass chemical composition. Pyrolysis-molecular beam mass spectroscopy (PMBMS) was used to develop a model to predict chemical composition of herbaceous materials, agricultural fiber, and woody materials (Agblevor et al., 1994; Kelley et al., 2004; Labbé et al., 2005). One of the main disadvantages that have made PMBMS less popular in developing models is the fact that the samples are destroyed during pyrolysis; this problem is not observed with vibrational (infrared and Raman) spectroscopy. Many studies have used infrared spectroscopic methods in developing models for different biomass species and these will be discussed in section 2.4.

2.4. Infrared spectroscopy

2.4.1. Basic concepts

Many different types of spectroscopy have been developed depending on the type of light source, nature of the energy intensity it measures (e.g. electromagnetic radiation, electrons, or sound) and the process of energy interactions; that is absorption, transmission, emission, and scattering (Hollas, 1996). Spectroscopic techniques have become popular in recent years due to the increasing demand for product quality improvement in the food, pharmaceutical, petrochemical, and process industries to replace the conventional and time consuming analytical techniques (Siesler, 2008). Organic and inorganic molecular bonds in a material or substance excite to higher energy state when illuminated with incident radiation. The energy transition from this excitement as light is absorbed by molecules causes vibration within molecular bonds which are detected by infrared (IR) spectroscopy. Molecules absorb radiation only at specific frequencies and these absorptions can be quantified to form a spectrum. The spectrum thus shows the interaction (absorbance or

transmittance) between light and a sample as a function of frequency or wavelength. The fundamental concept of infrared activity, leading to absorption of infrared radiation, is that bonds must be electrical dipole which changes at the same frequency. Thus, symmetrical bonds which have identical elements or groups on each sides like C-C, Cl-Cl, CH_3-CH_3 etc. are not absorbed by IR radiation (Workman, 1993). Another spectroscopy method that uses vibrational technique as IR is the Raman spectroscopy. A Raman spectrum is formed from scattered emissions due to temporal distortion of electrons distributed around molecular bonds. Hence, dipoles are momentarily formed and disappear upon relaxation making symmetric bonds detectable by Raman spectroscopy (Hollas, 1996).

Infrared radiation covers a section of an electromagnetic spectrum with a wavelengths ranging from 0.78 to 1,000 μm or wave number from 13,000 to 10 cm^{-1}. Wave number is the number of waves per unit length and is directly proportional to frequency and the associated energy while wavelength is the distance of a complete wave oscillation and is inversely proportional to the wave frequency and associated energy. Electromagnetic spectrum is measured with either wave number or wavelength and wave number can be converted to wavelength using the equation 1;

$$Wave\ number\ (cm^{-1}) = \frac{1}{Wave\ length\ (nm)} x\ 10^{-7} \qquad (1)$$

An IR spectrum has absorption intensity or percent transmittances on the y-axis with the wavelength or wave number on the x-axis. The transmittance (T) is the ratio of the light intensity to the sample (I) and the light intensity out of the sample (I_o) (Coates, 2000). The absorbance is derived from the transmittance through equation 2;

$$Absorbance = log_{10}\left(\frac{1}{T}\right) = -log_{10}T = -log_{10}\left(\frac{I}{I_0}\right) \qquad (2)$$

The conversion from transmittance to absorbance is reversible without losing any spectral information. Percent transmittance spectra whose peaks face downward are usually used for qualitative analysis while the absorbance spectra whose peak face upward are mostly used for quantitative analysis (Workman, 1993). The basis of quantitative analysis is that absorbance follows Beer's Law in equation 3:

$$A = mcd \tag{3}$$

where: A = sample absorbance at a given frequency,

m = molecular absoptivity at given frequency,

d = path length of light source in the sample, and

c = concentration of the sample.

The IR region is divided into three main sub regions namely; near-IR, mid-IR, and the far-IR with wavelength ranges of 0.78 to 2.5 μm, 2.5 to 50 μm and 50 to 1,000 μm, respectively (Hsu, 1997). Unlike the mid-IR and far-IR spectra, the near-IR (NIR) sampling requires minimal or no sample preparation. It can also be done with large samples (thickness of up to 1 cm) and offers fast quantitative analysis without any destruction of the sample. Also, the NIR region produces consistent fingerprints of compounds that can be more easily identified than with mid- and far-IR spectra (Siesler, 2008). Initially IR spectroscopy was usually used for qualitative analysis especially with the conventional dispersive instrument but the development of stronger computerized and statistical software alongside reliable IR instruments like the Fourier-Transformed Infrared (FTIR) spectrometer have improved the analytical procedure for quantitative data. Near-infrared (NIR) spectroscopy has been widely used to determine quantitative information for many

biological and clinical materials using a single NIR spectrum due to its unique absorption pattern (Chen et al., 2004).

2.4.2. Chemical bond and NIR absorption

As the electromagnetic energy of a particular wavelength is absorbed by a sample, molecular vibration occurs in the bond between two atomic centers in form of X-H, where X represents carbon (C), nitrogen (N), or oxygen (O) and H is hydrogen. The normal types of vibrational motion observed by infrared-active molecules are; stretching and bending vibrations (Coates, 2000), and to a lesser extent contracting vibrations (Smith, 1999). Bending vibrations appear to be more complex than stretching vibrations as they may occur out of plane or in plane. Out of plane vibrations consist of twisting and wagging while in plane vibrations consist of rocking and scissoring. Stretching vibrations are symmetrical or asymmetrical and occur at higher frequencies (shorter wavelength) than bending vibrations (Coates, 2000). Other molecular vibrations in the NIR region include the carbonyl C=O double-bond stretching vibrations and C-C stretching vibrations but these vibrations are sometime too weak or absent to be considered in analysis (Blanco and Villarroya, 2002). The bending and stretching of molecular bonds have an impact and contribute to the overall absorption spectrum (Elena, 2004).

These spectroscopic vibrations occur at particular frequencies producing overtones and combination bands. An overtone is any frequency produced that is higher than the fundamental frequency (lowest frequency in a harmonic series) while combination bands correspond to the interactions between stretching and bending vibrations of C-H, O-H, and N-H bonds associated with their frequencies and are used to identify specific compounds (Chen et al., 2004; Coates, 2000). In NIR, molecular absorption of overtones generally

occurs between 700 to 1,800 nm while combination bands occur between 1,800 to 2,700 nm (Shenk et al., 2008). Some authors defined overtone bands from 780 to 2,000 nm (Blanco and Villarroya, 2002), 1,300 to 1,900 nm (Kemeny, 2005), and combination bands from 1,900 to 2,500 nm (Blanco and Villarroya, 2002; Kemeny, 2005). This difference depends on the overtone order (first, second, or third) and the bond nature and strength considered in defining these regions.

When a molecule begins to vibrate as it absorbs light at a given frequency, it produces a peak in the infrared spectrum at the wavelength where the light was absorbed. Functional groups are assigned to different peaks; however, not every molecular vibration can be excited by infrared radiation (Smith, 1999). Different overtone bands are produced as the frequencies increase which leads to first, second, third or even higher order overtones. The first order overtone is 10 to 100 times weaker than the fundamental vibrational frequency while the second order overtone is weaker than the first (Blanco and Villarroya, 2002). Chen et al. (2004) demonstrated that models developed from combination spectra for glucose, urea, and lactate solute were better than that developed only from first overtone spectra. Hence, this denotes the importance of combination spectra in model development. From these overtones and combination bands, organic components like cellulose, lignin, and pectin have been calibrated. Table 1 shows NIR-assigned bands for general organic bonds and Table 2 shows tentative bands for cellulose and lignin (Shenk et al., 2008; Stuart, 2004).

Using only the NIR spectra region or part of the NIR region or extending to the visible region to develop calibration models to determine biomass composition varies from paper to paper. Robert and Cardet (1998) also distinguished two characteristic parts

important for carbohydrate NIR spectra; 1,100 to 1,800 nm for first and second overtone of O-H and C-H stretching and 1,800 to 2,500 nm for combination bands of O-H and C-H stretch. The first overtone and combination regions of 1,000 to 2,500 nm were used to develop models for biomass like corn stover, switchgrass and rice straw (*Oryza sativa* L.) [(Jin and Chen, 2007; Labbé et al., 2008; Liu et al., 2010; Sanderson et al., 1996; Ye et al., 2008)].

Table 1: Common NIR band of organic compounds (Stuart, 2004)

Wavelength (nm)	Assignments
2,200-2,450	Combination of C-H stretching
2,000-2,200	Combination of N-H and O-H stretching
1,650-1,800	First overtone C-H stretching
1,400-1,500	First overtone N-H and O-H stretching
1,300-1,420	Combination C-H stretching
1,100-1,225	Second overtone C-H stretching
950-1,100	Second overtone N-H and O-H stretching
850-950	Third overtone C-H stretching
775-850	Third overtone N-H and O-H stretching

Table 2: Tentative band assignments for cellulose and lignin (Shenk et al., 2008)

Type of material	Assignments	Wavelength (nm)
Cellulose	First overtone O-H stretch	1,490
	First overtone C-H stretch	1,780
	Second overtone C-O and O-H stretch	1,820
	C-H stretch and C-H deformation	2,335
	$C-H_2$ symmetry and $=CH_2$ deformation	2,347
	CH_2 second overtone	2,352
	C-H stretch and C-C stretch combination	2,488
Lignin	Second overtone C-H stretch	1,170
	First overtone O-H stretch	1,410
	C-H stretch combination	1,417
	First overtone O-H stretch	1,420
	C-H stretch combination	1,440
	First overtone C-H stretch	1,685

The use of the entire NIR spectra that will include both the first and second overtones and also third overtone in the visible region from 400 to 2,500 nm was recommended (Hames et al., 2003) and it was used in developing a model for corn stover (Wolfrum and Sluiter, 2009). Successful development of an NIR model requires understanding: 1) the basic fundamentals in spectroscopy, 2) the type of instrument (sensitivity and directivity) which will lead to higher scan resolution and spectra smoothing and 3) the skills and techniques for building calibration models. However, continuous improvement of a model is necessary. This is possible as new instruments and statistical softwares are being developed.

2.4.3. NIR instrumentation

There are many different types and brands of NIR instruments and each has claimed advantages. It is good for users to be knowledgeable about certain distinguishing characteristics of NIR before selecting the best instrument to use. Some of these characteristics include optical configurations, scan rates, source type, detector type, sample averaging technique, dust and water proofing, and vibration tolerance (Workman and Burns, 2008).

For research, an instrument with broad capabilities is always preferable as different parameters will be available to define the instrument to get the most reliable result. Most commercial IR instruments use dispersive or Fourier-transformed (FT) spectrometric techniques. Both dispersive and FT spectrometers have three basic components: radiation source; optical splitter that is monochromator (dispersive) or interferometer (FT); and detector (Griffiths and De Haseth, 2007). In the dispersive IR spectrometer, the monochromator disperses a broad spectrum of radiation and provides a continuous series of

electromagnetic energy bands of determinable wavelength range. The dispersive components in a monochromator are prisms or gratings which are used with variable slit mechanism, mirrors, and filters.

The interferometer in an FTIR spectrometer replaces the monochromator in a dispersive IR instrument with several advantages. The FTIR is very rapid and can average many scans given the same time to be used for a single scan in a conventional dispersive IR which causes little difference in repeatability (McCarthy and Kemeny, 2008). Higher energy throughput for the same resolution is observed with FTIR when compared to the earlier dispersive IR and this leads to higher wavelength accuracy and constant spectra resolution (Griffiths and De Haseth, 2007). These advantages allow FTIR to achieve the same signal-to-noise ratio (SNR) or sensitivity in a shorter time than a dispersive IR instrument. The SNR shows how much the vibrational signal produced has been corrupted by electronic or environmental noise. A higher ratio signifies that noise has a less important influence on the signal (Workman and Burns, 2008). The interferometer in the FTIR also ensures that a constant and accurate wave number, continuous spectra and negligible stray light from the source is achieved unlike conventional dispersive IR which requires constant changing of grating or filter (McCarthy and Kemeny, 2008). Mechanical simplicity of FTIR also generally allows lower breakdown rates than the conventional dispersive IR (Griffiths and De Haseth, 2007). Some research claimed that FTIR also differs from conventional dispersive IR with experimental results not only from the theoretical and mechanical difference illustrated above (Ye et al., 2008). It is therefore important to note that different results can be obtained due to difference in wavelength range (e.g. NIR or

mid-IR), instrument technology (FTIR vs. dispersive IR) and acquisition mode (transmittance vs. reflectance).

Recently, the development of a diode array (DA) spectrometer has been possible due to the availability of silicon-based sensors in linear arrays. Thus, the DA spectrometer is an advanced dispersive instrument which incorporates a diode array detector, as well as fiber optics that improve the energy throughput of the instrument. The array effectively contains hundreds of detectors that acquire a complete spectrum simultaneously as compared to the conventional dispersive spectrometer (McClure, 2001). The DA detectors enable many spectra to be collected from a single sample in a fraction of second. The diode also accumulates energy that enables spectrometer to produce spectrum during low energy measurements by exposing the diode arrays. The fiber optics collect most of the reflectance spectra from the sample directly to a fixed grating in a monochromator, hence improving the energy throughput (McClure, 2001). Due to these improves features; the DA is presently the most popular dispersive NIR instrument (Workman and Burns, 2008).

2.4.4. Sampling and model development

Developing a model to predict chemical composition or components of unknown samples depends on the samples used for calibration. The accuracy and precision of both NIR and reference methods largely depends on sample selection, sampling procedure, and sample preparation (Sun, 1997). Williams (2008) identified about 30 factors affecting the accuracy and precision of NIR analysis due to sampling, samples, and sample preparation and argued that most errors arise during sample preparation followed by sample presentation. This is mainly due to the difficulty in preparing a sample without altering its original form. Sample presentation which includes thickness of cell, particle size, and

surface presentation of sample to an NIR instrument impacts the spectra. Techniques to ensure uniformity in sample presentation are vital (Workman, 2008). Hein et al. (2009) conducted an experiment to examine how particle size, milling procedure, and quality of solid wood surface influence the prediction of chemical components in *Eucalyptus* wood using an NIR. They found out that sample preparation showed greatest influence in predicting chemical properties, followed by sample presentation (solid or milled wood) and lastly particle size (thin and thick powder).

Sample calibration set selection is the initial step in the model development process. Workman (2008) expressed that an ideal calibration set should be greater than 10 to15 samples per analytical term (glucan, xylan, lignin, etc.) and a minimum of 30 samples for preliminary study and up to 100 to 300 samples for a robust calibration set was proposed by other authors (Hames et al., 2003). The number used in other papers have been: 228 samples (Jin and Chen, 2007), 35 samples (Ye et al., 2008), 77 samples (Wolfrum and Sluiter, 2009) and 71 samples for a broad based model of 36 switchgrass samples and 35 corn stover sample (Liu et al., 2010). The most important aspect to consider is that the calibration set should include samples with a wide compositional variability and wide range of constituents (glucan, xylan, lignin etc). In order to get wide compositional variability in their calibration set, most of the papers used samples from varied locations and different botanic fractions (leaf, stems, nodes, pith, etc.). Wide composition variability is important because prediction can only be obtained for a new sample if variables (composition) lie within that covered by the calibration samples. Statistically, extrapolation outside developed range always leads to erroneous prediction (Sun, 1997). Using a larger number of samples strengthens the calibration set for statistical analysis.

Other processes in model development include: determining reference quantity, collecting NIR spectra data, developing the mathematical model, validating the model with samples of known concentration, routine analysis and modeling, transferring of calibration, and performing an actual feasibility study (Workman, 2008).

2.5. Statistical and mathematical analysis of NIR data

Spectral data need to be correlated to a property of interest. The process of using statistical and mathematical methods to correlate spectral information to chemical information is known as chemometrics. Since each spectrum consists of several hundred of data points, statistical analysis is used rather than univariate analysis of an NIR spectra (Workman, 2008). The multivariate regression analysis is used to build a calibration model that depicts the relationship between spectra and chemical properties. This calibration model is then used as a tool to predict components in unknown samples. Wavelengths with the highest molar absorptivity for a specific component are selected but with the complex matrix of NIR spectra, it is difficult to know if an ideal set of wavelengths have been selected for analysis. Of the many different types of multivariate analysis that exist, the principal component analysis (PCA) and the partial least squares (PLS) regression or principal component regression (PCR) analysis are mostly used for spectra analysis to solve this problem of complex spectra matrix. PCA identifies patterns in data and expresses the data in a way that highlights similarities and differences which remove or reduce large baseline variations (Sun, 1997). One of the disadvantages of PCA is that its results are difficult to interpret (Workman, 2008). PLS regression combines features from PCA and multiple regressions to predict a set of dependent variables from a large set of independent

variables (Helland, 1990). One of the disadvantages of both statistical models is the difficulty to transfer calibrations between instruments (Workman, 2008).

Many statistical software packages have been developed to perform these complex mathematical and statistical procedures. Understanding such software and which statistical tests to use is vital in obtaining the best analysis. Analysis of spectral data starts with preprocessing or pretreatment. Preprocessing is a step where statistical methods are used to transform raw spectral data to a processed data before model development (Sun, 1997). This process helps to reduce the degree of uninformative variance in the spectra due to noises, complex background and baselines. Noise and background problems usually occur when the spectra are altered due to alignment of grating, mirrors, lamps and attenuated reflectance cells. Without preprocessing, these spectra variation will lead to complicated and biased calibration models. Preprocessing methods commonly used include mean-centering and variance-scaling, multiplicative scatter (or signal) correction, orthogonal signal correction, instrument standardization, and smoothing functions (Xu et al., 2008) with smoothing or derivatives mostly used for univariate analyses (Boysworth and Booksh, 2008). Out of the regression methods, PLS has been mostly used to develop models to predict carbohydrate content in biomass (Hames et al., 2003; Jiang et al., 2010; Jin and Chen, 2007; Kelley et al., 2004; Liu et al., 2010; Sanderson et al., 1996; Ye et al., 2008). PLS is faster and can achieve higher values of predictive level while employing less spectral variance compared to the other methods (Bjørsvik and Martens, 2008).

Before model selection and validation, detection of outliers improves the model as observations that are distant from the rest of the data are removed (Workman, 2008). There are three statistical parameters often used to select or validate the performance of models.

These include; root mean squared error of calibration (RMSEC), root mean squared error of cross validation (RMSECV), and root mean squared error of prediction (RMSEP) (Boysworth and Booksh, 2008). These three are based on the calculation of root mean squared error (RMSE) in equation 4.

$$RMSE = \sqrt{\sum_{i=1}^{I} \frac{(c_i - \hat{c}_i)^2}{I}} \qquad (4)$$

where I represent the number of samples, c_i is the reference measured values, $\hat{c}_i$ is the NIR modeled values which differ in RMSEP, RMSECV and RMSEC.

The coefficient of determination (R^2) indicates the percent variation in the predictive value that is explained by the model. The coefficient of correlation (r) measures the correlation between the predictive value and the measured value (Workman, 2008). Higher R^2 and r, and lower RMSE is desired for the best model (Boysworth and Booksh, 2008).

3.0. PROBLEM STATEMENT AND OBJECTIVES

3.1. Problem statement

Meeting the 2022 mandate of 60.6 BLY of cellulosic biofuels production will likely require a wide variety of feedstocks that were identified in USDA/DOE study (Perlack et al., 2005). Perennial grasses, which represent as much as 28% of the total US biomass potential, grow in diverse environmental conditions around the US. Hence, many cellulosic ethanol plants in the US will likely use perennial grasses and other herbaceous perennials grown in their locality as part of their feedstock. Structural carbohydrate composition and biomass yields are vital in determining ethanol yield per acre. Chemical composition and biomass yields vary with plant variety and type, environment, and management practices during growth.

North Dakota has approximately 1.2 million ha (3 million acre) of Conservation Reserve Program (CRP) lands and more than 2.8 million ha (7 million acres) of erodible land. These CRP and erodible lands cannot be used for annual crop cultivation but they may be used as grassland for biomass energy crops. These grasslands available in North Dakota have given the state the potential to be the greatest producer of herbaceous perennial crops in the US (Nyren et al., 2009). Different grasses including switchgrass (SG), wheatgrass (WG), wild rye (WR), and big bluestem (BB) and also herbaceous perennials species including alfalfa (AL) and sweet clover (CL) are grown across North Dakota by NDSU Research Extension Centers (RECs) in Streeter, Minot, Carrington, Williston, and Hettinger. The goal of the research is to evaluate the best management practice for maximum biomass yields and quality from theses herbaceous perennials. These perennials are grown as individual or conglomerate bioenergy crops and have different

chemical composition and biomass yields, hence, will vary in ethanol production potential yearly and in different environment. Identifying individual and conglomerate herbaceous perennial species with the highest ethanol yield potential on a mass and area basis will be beneficial to agricultural producers and ethanol processors. There is no study that has evaluated the ethanol potential of mixed perennial biomass grown in ND.

Developing a fast and less labor-intensive method to determine biomass composition using NIR spectroscopy has become popular as the need of cellulosic biofuel production is increasing. This will enable feedstocks to be evaluated using compositional value rather than weight only. NIR spectroscopy models for single biomass feedstock like corn stover, rice straw, switchgrass have been developed (Jin and Chen, 2007; Labbé et al., 2008; Ye et al., 2008). Liu et al. (2010) also showed that it is possible to have a broad-based model for biomass material by conducting a study using switchgrass, corn stover, and wheat (*Triticum sativum*) straw. These materials are expected to have more dissimilarity in physical properties than different herbaceous perennial species. Hence, developing a model for individual or conglomerate herbaceous perennial is possible, but no such study has been reported. There are approximately 30 factors that affect the accuracy and precision of NIR calibration model (Williams, 2008; Williams and Norris, 2001). Most of these factors can be controlled especially during sample selection, preparation, and sampling. The experiments in this research will look at some of the factors that depend on methodology and instrument stability rather human error. Therefore, the variability of wet chemistry method, NIR spectroscopic sampling time, and effect of repacking on NIR calibration model will be studied. The degree to which these factors affect the model has not yet been reported.

3.2. Objectives

The objectives of this study are;

- ❖ Study the variability in the chemical composition of herbaceous perennial biomass harvested annually on irrigated plots in Williston, ND for three years (2008, 2009, and 2010); and on non-irrigated plots in Minot in 2008 and predict their potential in terms of ethanol production on a mass and area basis.
- ❖ Develop an NIR spectroscopy model to predict lignin, glucan, and xylan for individual and conglomerate herbaceous perennial biomass species grown in ND and use the model to test the accuracy and precision of the NIR calibration models with respect to sampling time and repacking.

4.0. MATERIALS AND METHODS

4.1. Biomass sample collection

Biomass samples were obtained through the Central Grassland Research Extension Center (CGREC) in Streeter, ND. The CGREC and 4 other NDSU Research Extension Centers have six grassland experimental plots in five locations (Williston, Carrington, Minot, Hettinger, and Streeter) across central and western ND. Williston has irrigated and non-irrigated plots. Fig. 1 is a map of ND showing the five locations of the experimental plots.

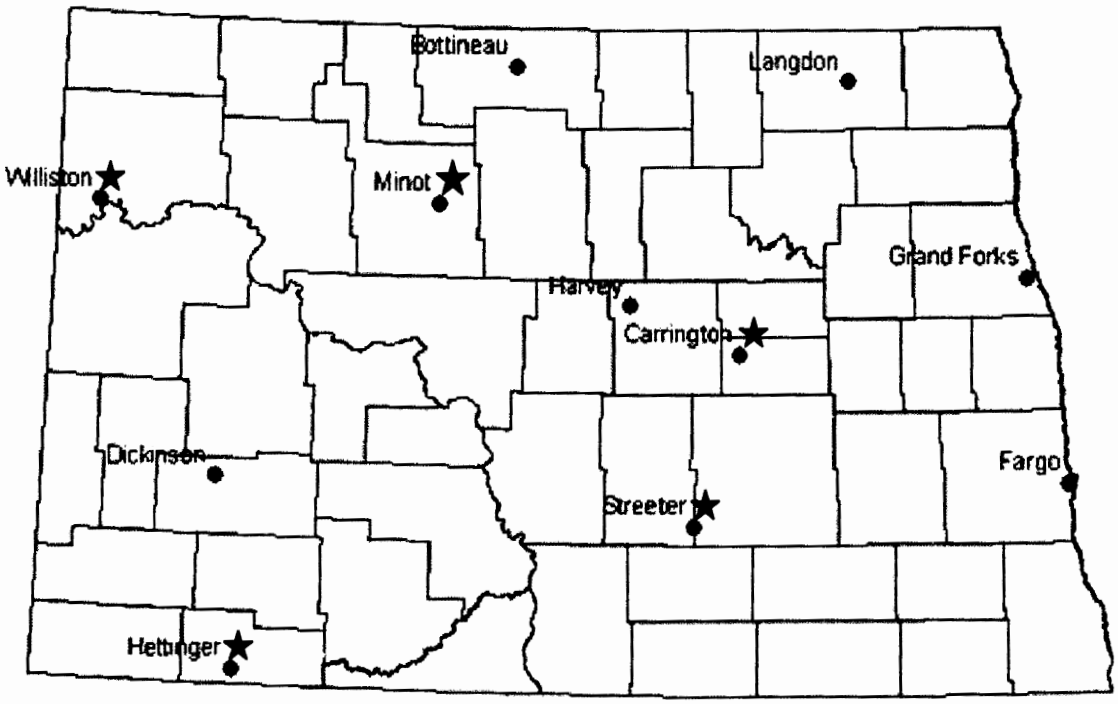


Figure 1: Locations of experimental grassland plots across central and western ND.

★ Location of experimental grassland plots

N.B: Map is from North Dakota Agricultural Weather Network (NDWAN), http://ndawn.ndsu.nodak.edu/

Single or conglomerate grass species planted on these 6 plots in the 5 locations in 2006 include: 2 varieties of switchgrass (SG) [Dakota and Sunburst]; 2 species of wheatgrass (WG) [Intermediate WG (*Thinopyrum intermedium*) and Tall WG (*Thinopyrum ponticum*)]; 2 species of wildyre (WR) [Magnar WR (*Elymus cinereus*) and Mustang WR (*Elymus angustus*)]; and big bluestem (BB). Alfalfa (AL) and sweet clover (CL), which are not grass species, were planted together with Intermediate WG to constitute another set of sample plots. A total of 10 single and conglomerate species treatments were planted in these 5 locations across ND. The detailed information about the 10 treatments and the number of samples obtained from the different locations each year of harvest are shown in Table 3. Biomass yields for the samples were also obtained from CGREC 2010 annual report (Nyren et al., 2010).

Table 3: Biomass samples obtained by treatment, location and year of harvest

No.	Species	Williston (irrigated)			Minot	Carrington	Hettinger	Streeter
		2008	2009	2010	2008	2008	2009	2008
1.	Tall Wheatgrass	√	√	√	√	√	√	
2.	Dakota Switchgrass	√	√	√	√			
3.	Intermediate Wheatgrass	√	√	√	√		√	
4.	Sunburst Switchgrass	√	√	√	√	√		√
5.	Tall & Intermediate Wheatgrass	√	√	√	√	√	√	
6.	Sunburst Switchgrass & Tall Wheatgrass	√	√	√	√	√		
7.	Magnar & Mustang Wild rye	√	√	√	√	√		
8.	Alfalfa, Sweet Clover & Intermediate Wheatgrass	√	√	√	√		√	
9.	Sunburst Switchgrass & Big bluestem	√	√	√	√	√		
10.	Sunburst Switchgrass & Mustang Wild rye	√	√	√	√			
	Total Samples	10	10	10	10	6	4	1

In Williston, biomass yields from non-irrigated plots were low (0.16 to 0.24 Mg ha^{-1}), therefore only samples from irrigated plots were obtained. In addition to rainfall, the irrigated plots were supplied additional water to give a total of 635 mm of water per year. Samples from other locations were not irrigated but had biomass yields of more than 0.36 Mg ha^{-1} (Nyren et al., 2010). Some treatments from Williston (2009) and Minot (2008) were manually separated into two botanic fractions (leaf and stem) to increase variability and number of samples. These species were Dakota SG, Sunburst SG, Magnar and Mustang WR, Sunburst SG and BB, and Sunburst SG and Mustang WR from Minot (2008) and also Sunburst SG and Mustang WR, and Intermediate WG from Williston (2009). These were samples that had nearly equal proportion of leaf and stem which could be separated. Therefore, 14 more samples (7 leaf and 7 stems) were obtained from this process which were combined with the 51 samples from Table 3 to make 65 samples. Of the 65 samples used for model development, 55 samples were used as calibration set and 10 samples as validation set. The calibration and validation samples were randomly selected using Microsoft Excel.

The complete set of samples from Williston (2008, 2009, and 2010) and Minot (2008) were used to determine the ethanol potential on an area (L ha^{-1}) and mass (L Mg^{-1}) basis for these single and conglomerate species treatment. Table 4 shows the total rainfall from April – September, the average air temperature, and growing degree days (GDD) based on 10°C (50°F) and 4.5°C (40°F) from June - September for Williston and Minot from 2008 to 2010. The GDD is based on 10°C and 4.4°C because no appreciable growth is observed in warm-season and cool-season grasses below these temperatures, respectively (Sanderson and Wolf, 1995). The growing season starts in April but effective growth of C4

warm-season grasses begins in June after weed control is done in May. Harvesting was done yearly in the 2nd and 3rd week of September.

Table 4: Average temperature (June-September), acumulated growing degree days based on 10°C (June – September) and 4.4 °C (April - September), and total rainfall (April – September) in Williston and Minot

	Williston			Minot		
Parameter	2008	2009	2010	2008	2009	2010
Average Maximum Temperature (°C)	25.8±4.3	25.3±1.2	24.3±3.9	24.5±3.9	24.0±1.1	23.4±4.0
Average Minimum Temperature (°C)	11.4±3.3	11.2±1.2	11.7±3.0	10.8±2.9	11.1±0.8	11.9±3.0
Growing Degree Days based on 4.4°C	2,050	1,994	1,980	1,883	1,871	1,945
Growing Degree Days based on 10 °C	1,053	1,013	990	929	934	948
Total Rainfall (mm)	194.3	234.7	339.3	401.3	297.7	473.2

Source: NDWAN at http://ndawn.ndsu.nodak.edu/.

4.2. Chemical composition determination

The biomass samples (whole) were ground using a Wiley Mill (Model 4, Thomas Scientific, Swedesboro, NJ, USA) passing through a 1-mm sieve. The ground samples were then sieved to 0.18 to 0.85 mm particle size using mesh #20 and #80 American Society for Testing and Material (ASTM) E-11 standard sieves. These samples (-20/80 particle size) were used for moisture, ash, compositional analysis, and NIR calibration model development. Compositional analysis methods below were done by using the NREL analytical procedures (Sluiter et al., 2005a; Sluiter et al., 2008; Sluiter et al., 2005b). The chemical composition and NIR calibration model development processes involve steps that are described in Fig. 2.

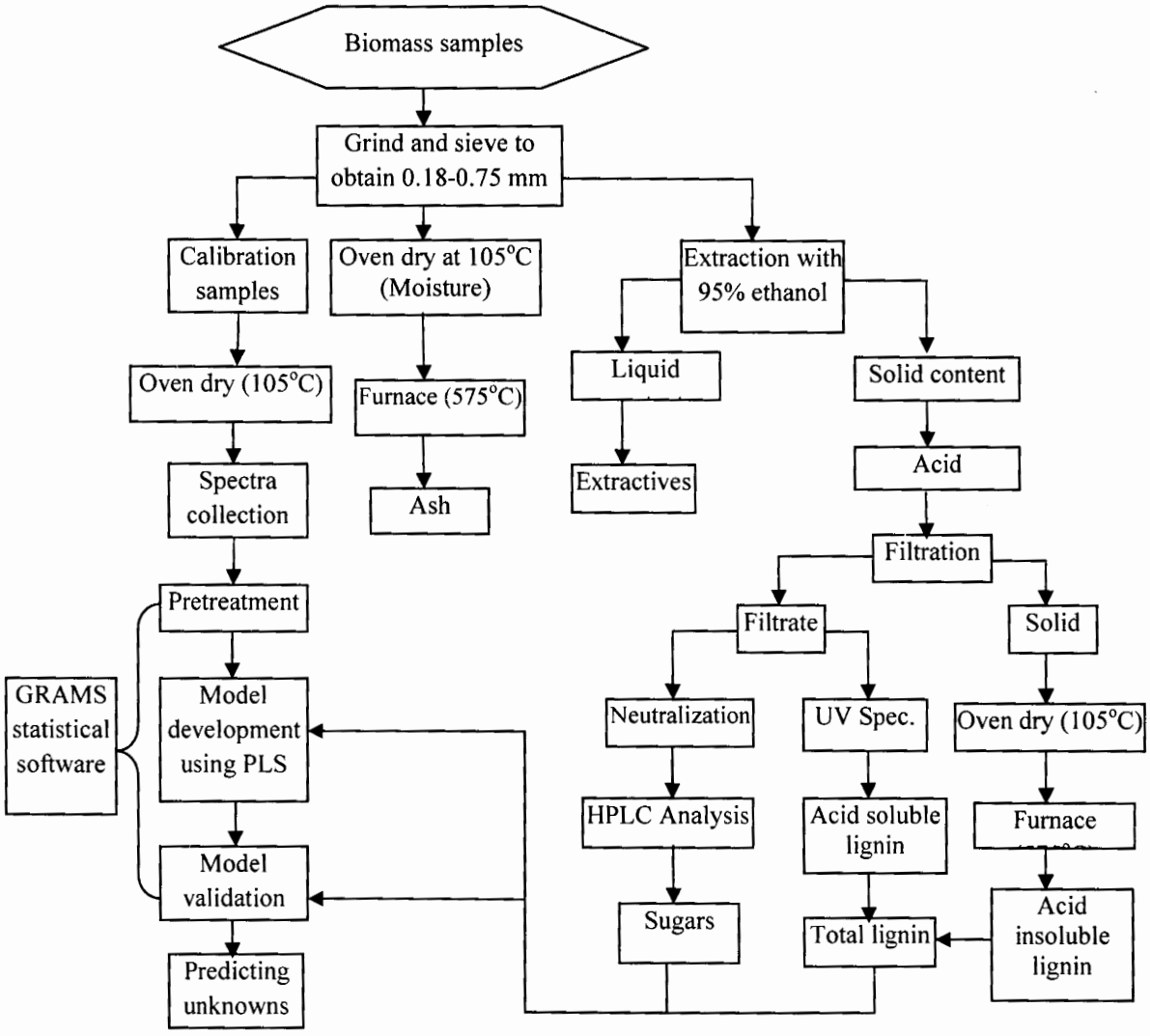


Figure 2: Flow chart of experimental procedure.

4.2.1. Moisture content and ash

To determine moisture content, 0.5 to 1.0 g samples were dried in a gravity convection oven (Precision Scientific, Inc; Winchester, IL, USA) at 105 °C overnight. After moisture determination, these samples were put in a furnace (Model 48000; Barnstead/Thermolyne; Dubuque, IA, USA) at 575 °C for 24 h for ash determination. Values for moisture and ash were calculated as percent on a wet basis.

4.2.2 Extractives

Non-structural components like chlorophyll, waxes, and other minor components in biomass were removed with 95% ethanol using an automated ASE 200 solvent extractor (Dionex; Sunnyvale, CA, USA) set at 100 °C, and 10.3 MPa for 5 min heating time and 7 min static time. Ethanol was removed from the extractives solution by putting the tubes in a 40 °C water bath (Thermo Electron Corporation; Marietta, Ohio, USA) for approximately 12 h. The weights of the extractives were measured after oven drying at 40 °C overnight. The insoluble portions (biomass samples free from extractives) were then air-dried and later acid-hydrolyzed for lignin and carbohydrate determination.

4.2.3. Carbohydrate and lignin quantification

Hydrolysis was done in two phases: 0.3 g of dried extractives-free biomass samples were first hydrolyzed with 3 ml of 72% sulfuric acid (EMD Chemicals Inc.; Darmstadt, Germany) at 30 °C for 1 h and the hydrolyzates were then diluted to 4% sulfuric acid by adding 84 mL of distilled water. The diluted hydrolyzate was further hydrolyzed in an autoclave (Consolidated Stills & Sterilizers; Boston, MA, USA) at 121 °C for 1 h. Sugar recovery standards (SRS) containing 0.3 g each of glucose, xylose, galactose (Mallinckrodt Baker; Phillipsburg, NJ, USA), and arabinose (Calibiochem; Darmstadt, Germany) were dissolved in 4% sulfuric acid and also hydrolyzed in an autoclave at 121 °C for 1 h. The SRS accounts for the loss of sugar during the second stage of hydrolysis in the autoclave. Samples were hydrolyzed in duplicate in order to obtain the standard deviation.

The hydrolyzed biomass samples were then filtered into liquid and solid fractions using ceramic filtering crucibles. The liquid portion or filtrate contains monosaccharides, acid soluble lignin, organic acids, proteins, and a small amount of ash. Approximately 10

mL of this liquid portion from each biomass sample and SRS were neutralized in 0.72 g calcium carbonate (EMD Chemicals Inc.; Darmstadt, Germany) to obtain a pH of 6.8 to 7.2. The supernatants, after centrifugation, were filtered into 1-ml high pressure liquid chromatography (HPLC) vials and were stored in the freezer before sugar analysis. Sugar concentrations in g L^{-1} were obtained using an HPLC (Waters Corporation; Milford, MA, USA) with an Aminex HPX-87P (300x7.8 mm) carbohydrate column (Bio-Rad Laboratories; Hercules, CA, USA) running at 85 °C. The HPLC system was equipped with an isocratic pump, autosampler, and a refractive index detector (RID, model 2414, Waters Corporation). Before running biomass and SRS samples, four sets of calibration standards having glucose, xylose, galactose, and arabinose were run with concentrations shown in Table 5. Injection volume into the carbohydrate column was 20 µL and sugars were eluted with 18 MΩ NANOpure water at a flow rate of 0.6 mL min^{-1}. Sugar peaks were detected by RID detector at 50 °C and quantified using the 4-point external standard curves derived from the calibration standards.

Table 5: Concentration of HPLC quantification standards

Sugar	Low (g L^{-1})	Medium (g L^{-1})	Medium High (g L^{-1})	High (g L^{-1})
Glucose	1.0	5.0	10.0	20.0
Xylose	0.5	2.5	5.0	10.0
Galactose	0.5	2.5	5.0	10.0
Arabinose	1.0	5.0	10.0	20.0

Filtrate (1-2 mL) was also used to determine acid soluble lignin (ASL) using a UV spectrophotometer (Varian; Santa Clara, CA, USA) at 205 nm. The solid components after hydrolysis were dried in an oven at 105 °C for 4 to 6 h, and then put into a furnace at 575

°C for 4 to 6 h to determine acid insoluble lignin (AIL). The values for AIL and ASL were added to obtain the lignin content for the biomass samples.

From the sugar values obtained after compositional analysis for samples in Williston and Minot, theoretical ethanol yield in gallons per dry ton (GDT) were calculated using the Department of Energy (DOE) theoretical ethanol calculator (DOE, 2009) and converted to L Mg^{-1}. The dry mass percentage sugar yield were input in their respective columns and the calculation of theoretical ethanol yield in GDT on the DOE calculator was derived from the formula; (1.11 lbs of C6 sugar or 1.136 lbs of C5 sugar /lb of polymeric sugar) x (lbs of polymeric sugar/100 lbs of biomass) x (0.51 lbs of ethanol/lb of C6 or C5 sugar) x (1 gal of ethanol/6.55 lbs of ethanol) x (2000 lbs of biomass/ton biomass). The theoretical ethanol yield on an area basis (L ha^{-1}) was obtained by multiplying the theoretical ethanol yield on a mass basis (L Mg^{-1}) and the biomass (dry-matter) yields in Mg ha^{-1}.

During carbohydrate and lignin quantification, six samples were analyzed concurrently and chemical analysis for one control sample was repeated several times with some of the batches. This sample (Alkar Tall WG, Williston 2009 harvest) was repeated 8 times and the results were used to analyze reproducibility of compositional analysis methods.

4.3. NIR spectra acquisition

Samples from the -20/80 ASTM E-11 sieve were oven dried overnight before spectra acquisition. The samples were dried in order to prevent moisture content interfering with spectra. A Perten Diode Array (DA) 7200 NIR spectrometer was used to collect spectra for all samples. The spectra were collected within 950 to 1,650 nm wavelength

range with a scan resolution of 5 nm. A non-NIR absorbing Teflon sampling cup (diameter = 75 mm, depth = 7.6 mm) was used. Approximately 6 g of each ground sample was scanned in duplicate and in duplicate again after repacking. The four spectra obtained for each sample were averaged during exportation of spectra data. Averaging the spectra helps reduce the total number of spectra and processing time during calibration development. The data were exported to Thermo Scientific GRAMS Suite statistical software for model development.

4.4. Multivariate calibration

Partial least square (PLS) regression was used to build a multivariate calibration for lignin, glucan, and xylan in GRAMS IQ v9.0 (Thermo Fisher Scientific; Waltham, MA, USA). There were two versions of the PLS algorithm; PLS-1 and PLS-2. Though PLS-1 is slower than PLS-2, PLS-1 was used because it is more accurate when concentrations range of constituents are more than 1% (GRAMS, 2010). PLS-1 differs from PLS-2 because it calculates separate loading and score vector for each constituent.

A spectral region of 1,100 to 1,650 nm was selected because Shenk et al., 2008 showed that lignin, glucan and xylan are absorbed above 1,100 nm. When using PLS method, it is usually difficult to choose the correct number of loading vectors (factors) to use in the model. Choosing the appropriate diagnostic type is important in determining the correct factors. The diagnostic types test the calibration model and there are four types in the GRAMS IQ: self-prediction, leverage validation, cross validation, and calibration. Cross validation was used because it optimizes the model with variability in constituent values. The process in cross validation is done by removing and predicting one sample until all samples have been left out and/or predicted at least once. Before a preprocessing

method was selected, the spectra were variance-scaled. Variance scaling gives all values equal weight by emphasizing small variations in spectra data. Standard normal variate (SNV) and detrending preprocessing method was then used. The SNV method is similar to multiplicative signal correction (MSC) but the detrending removes additional variation in baseline shift and curve linearity. Also, Savitzky-Golay first-order derivative method, which is a part of GRAMS, was used with the SNV detrending method. Outliers were identified and excluded using spectral residual and concentration residual plots (GRAMS, 2010). Removing outliers increased the efficiency of the calibration model.

4.5. Testing the variability of NIR spectroscopy and sample size

The NIR calibration models developed to predict lignin, glucan, and xylan were installed into the NIR spectrometer. This model was used to test variability in compositional prediction based on sampling time and repacking. Three grass samples that had the best prediction for at least two of the three constituents were used for this variability study. These samples were oven dried and kept in the desiccator. To test the variability due to sampling time and repacking, prediction was done 10 times after every 30 minutes with no repack, 10 times on the same day with repack, and 10 times daily with repack.

4.6. Statistical analysis

Two-way Analysis of Variance (ANOVA) from SAS 9.2 was used to compare mean carbohydrate content for biomass species from different environments and years. The regression coefficient (R^2), and R/SEP [Ratio of the range of dataset (R) to standard error of prediction (SEP)] were used for model evaluation based on standards developed by the American Association of Cereal Chemists (AACC, 2000; Williams, 2001). Coefficient of

variation (CV) was used to evaluation variability of NIR spectroscopy based on standards from Williams (2001).

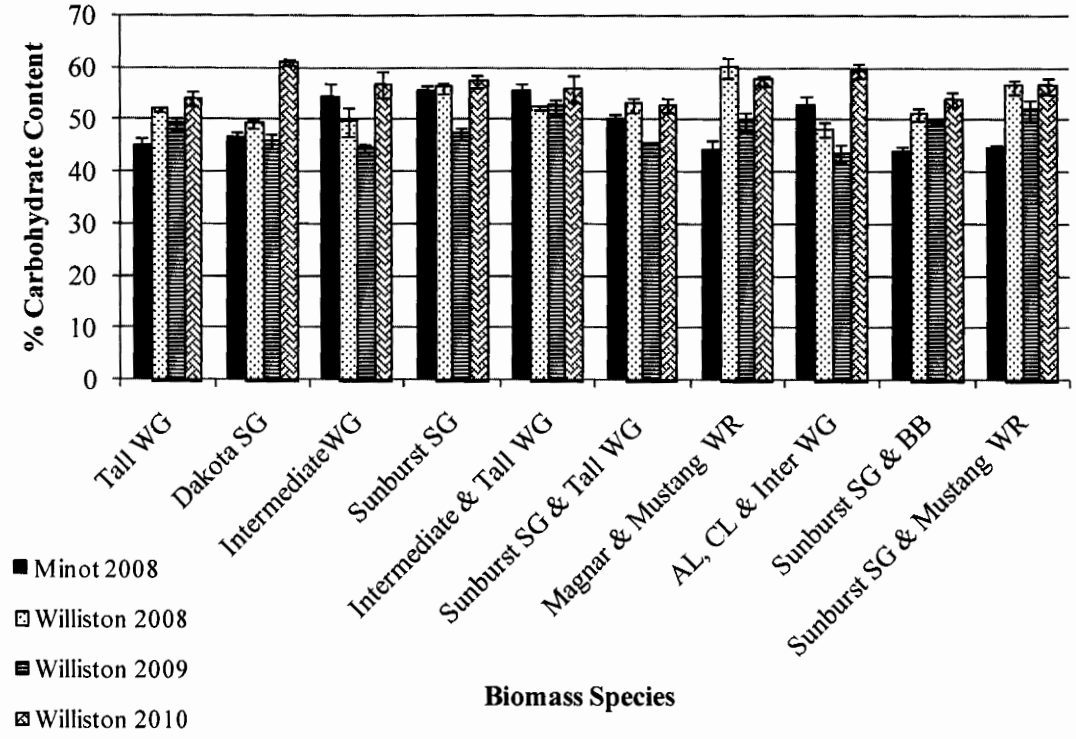
5.0. RESULTS AND DISCUSSION

5.1. Chemical composition and theoretical ethanol yield

5.1.1. Carbohydrate content variability with environment, year and species

The theoretical amount of ethanol on a mass basis (L Mg^{-1}) would be most important to ethanol producers because it indicates the yield of ethanol produced per Mg (dry mass) of biomass received. This value is proportional to the amount of polymeric sugars (glucan, xylan, arabinan, and galactan) present in the feedstock. Biomass with higher concentrations of structural carbohydrates will have a higher theoretical ethanol yield. Fig. 3 shows the total carbohydrate content (sum of glucan, xylan, arabinan, and galactan) obtained from compositional analysis of biomass samples harvested from the irrigated Williston plots in 2008, 2009, and 2010 and also from the non-irrigated Minot plot in 2008. The standard deviations for the total carbohydrate contents in Fig. 3 were calculated by using the propagation of error from standard deviations of glucan, xylan, arabinan, and galactan (Bevington and Robinson, 1992) .

Mass closure for compositional analysis of these biomass samples was 92.3±4.3%. The constituents included in this closure were structural carbohydrate (glucan, xylan, arabinan, and galactan values), lignin, extractives, and ash. The total carbohydrate content for all biomass samples varied between 43.5 and 61% with glucan values varying between 27.0 and 37.0%, xylan between 9.9 and 17.6%, arabinan between 2.5 and 9.5%, and galactan between 0.7 and 4.7%. In the DOE biomass database and Liu et al.(2010), a total carbohydrate content of 45 to 60% for switchgrass and big bluestem was reported (DOE, 2006).



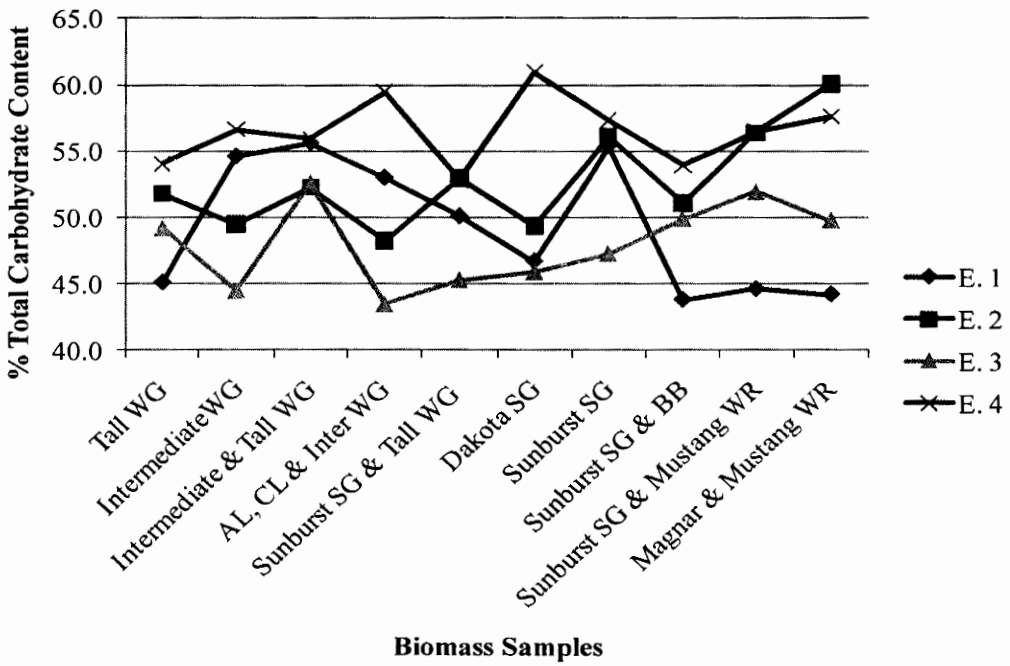
WG – Wheatgrass; SG – Switchgrass; WR – Wild rye; AL– Alfalfa, CL – Sweet clover; BB – Big bluestem.

Figure 3: Carbohydrate content for biomass samples from Minot non-irrigated plot and Williston irrigated plots. Error bars represent standard deviation.

Looking at Fig. 3, one can observe that total carbohydrate content varies yearly (bars 2-4) for the same biomass variety harvested in 2008, 2009, and 2010 in Williston and also carbohydrate yields vary with location (bars 1 and 2) for biomass samples from Minot and Williston harvested in 2008. Putting together the samples from Williston for the three years and Minot as four different environments, a two-way ANOVA was conducted in SAS 9.2 to test if these variations were statistically significant. This was done by using the biomass samples as blocks and the four different growing environments as treatments and the results are shown in Table 6. Also an interaction plot for the four environments was plotted as shown in Fig. 4.

Table 6: Result of a two-way ANOVA conducted at α=0.05 with mean total carbohydrate content of biomass samples from Williston harvested in 2008, 2009, and 2010 and Minot in 2008 as four different growing environments

Source	DF	Type I SS	Mean Squares	F-value	p-value
Environment	3	854.1	284.7	33.1	< 0.001
Species	9	192.6	21.4	2.5	0.023
Environment* Species	27	800.5	29.6	3.4	0.001



WG – Wheatgrass; SG – Switchgrass; WR – Wild rye; AL– Alfalfa, CL – Sweet clover; BB – Big bluestem.

Figure 4: Interaction plot for biomass samples from four different growing environments where Environment (E) 1-4 are Minot in 2008 and Williston in 2008, 2009, and 2010, respectively.

With a p-value of 0.001 for the interaction between the environment and species, the results in Table 6 show that mean total carbohydrate content for each biomass samples in the different environment were significantly different. Since there was interaction among biomass species and the environment, it means no particular biomass species consistently

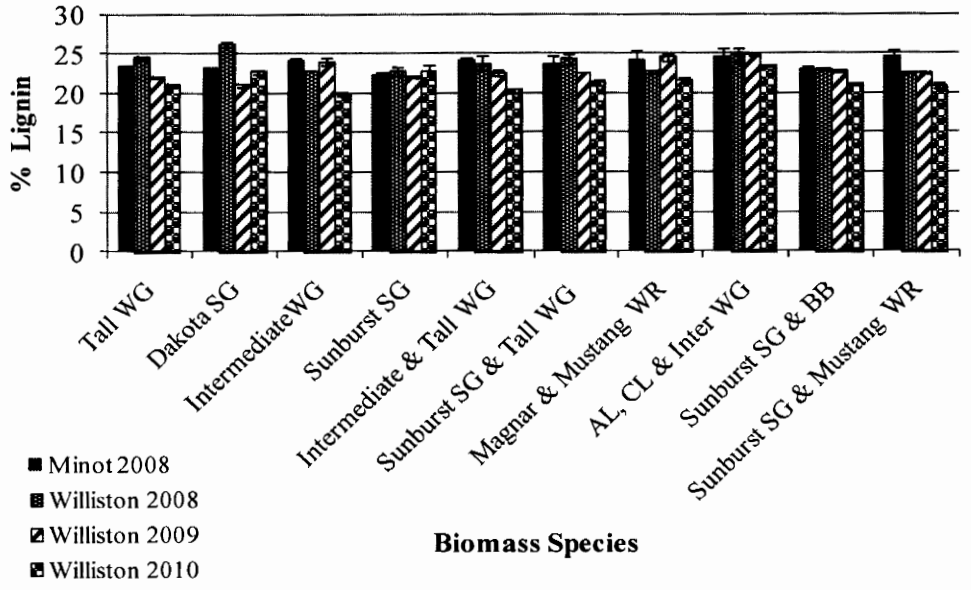
had highest or lowest total carbohydrate content in the different environments. Looking at Fig. 4, one can observe that Environment 3 (Williston in 2009) had lower carbohydrate content than Environment 1 (Minot in 2008) for six samples (both wheatgrass and switchgrass). Temperature and moisture normally affects the chemical composition of biomass. Environment 1 (Minot in 2008) had lower GDD values and water supply than Environment 2, 3, and 4 but had higher carbohydrate content than in Environment 3 even for some warm-season grasses like Dakota and Sunburst SG. Lower carbohydrate content was observed in some warm-season grasses like wild rye and big bluestem when combined with Sunburst SG and also Tall WG; a cool-season plant in Environment 1. This inconsistency with plant species and type makes it difficult to identify a particular species that had higher or lower carbohydrate content because of environmental condition.

Research work has been done on switchgrass and the results show variability in carbohydrate yield between location (Sanderson and Wolf, 1995), variety (Adler et al., 2009), and year (Adler et al., 2006). The plots in Environment 2, 3, and 4 in Williston were irrigated; therefore water availability will have limited effect on the difference in cellulose and hemicellulose content of the biomass. It can be observed from Fig. 4 that total carbohydrate content was lower for all biomass samples (except combination Intermediate and Tall WG) from Environment 2 (Williston in 2009) than those in Environment 3 and 4 (Williston in 2008 and 2010, respectively). It is difficult to conclude if temperature influenced the cellulose and hemicellulose amount because the GDD were higher in Environment 2 than in Environment 3 and Environment 4, but, Environment 4 in 2009 had lowest carbohydrate content irrespective of the species (warm-season or cool-season grasses). Sanderson and Wolf (1995) showed that celluloses concentrations increases

linearly from 800 to 1,000 GDD for switchgrass varieties but little or no increase can be observed beyond 1,000 GDD. The GDD for warm-season grasses in Table 4 showed that GDD for the three years were above 1,000 except in 2010 that had a value of 990. Cool-season grasses like wheatgrass had higher GDD above 1,600 but it is difficult to correlate the variation in carbohydrate content in the different environment or years to temperature. Other factors may have also influence this variation. In most cases, the GDD significantly affect the morphological development and maturity time of plant (Sanderson, 1992).

One factor that has been reported to negatively influence the total sugar yield upon hydrolysis is lignin content (Boateng et al., 2008; Dien et al., 2006). The lignin content for biomass species in Williston and Minot varied from 19.9 – 26.2% as shown in Fig. 5. Apart from Dakota and Sunburst SG, biomass samples from Environment 4 (Williston in 2010) had lower lignin content than from Environment 2 and 3(Williston in 2008 and 2009, respectively). This correlates with the higher carbohydrate yield observed from samples from Environment 4 in Fig. 4. A two-way ANOVA conducted at α=0.05 for lignin content show that lignin content varied (p-value = 0.001) for the different environments. This shows that the lignin content influenced the total carbohydrate content because same statistical results were observed for the mean total carbohydrate content in Table 6. If average total carbohydrate content is calculated for the four environments, it is difficult to conclude which biomass species as individual or conglomerate species had the highest carbohydrate content. For example, Intermediate and Tall WG had the highest mean total carbohydrate value of 54.0% but the carbohydrate content varied from 55.6%, 52.3%, 52.6%, and 55.9% for Environment 1, 2, 3, and 4, respectively, while sunburst SG and big bluestem with the lowest mean total carbohydrate content of 49.6% had an average total

carbohydrate content of 43.6%, 51.1, 49.9, and 53.9% for Environment 1, 2, 3, and 4, respectively. However, there was no clear trend in the variability among environments or years. In terms of carbohydrate content, it may not matter the variety of herbaceous perennial grown.

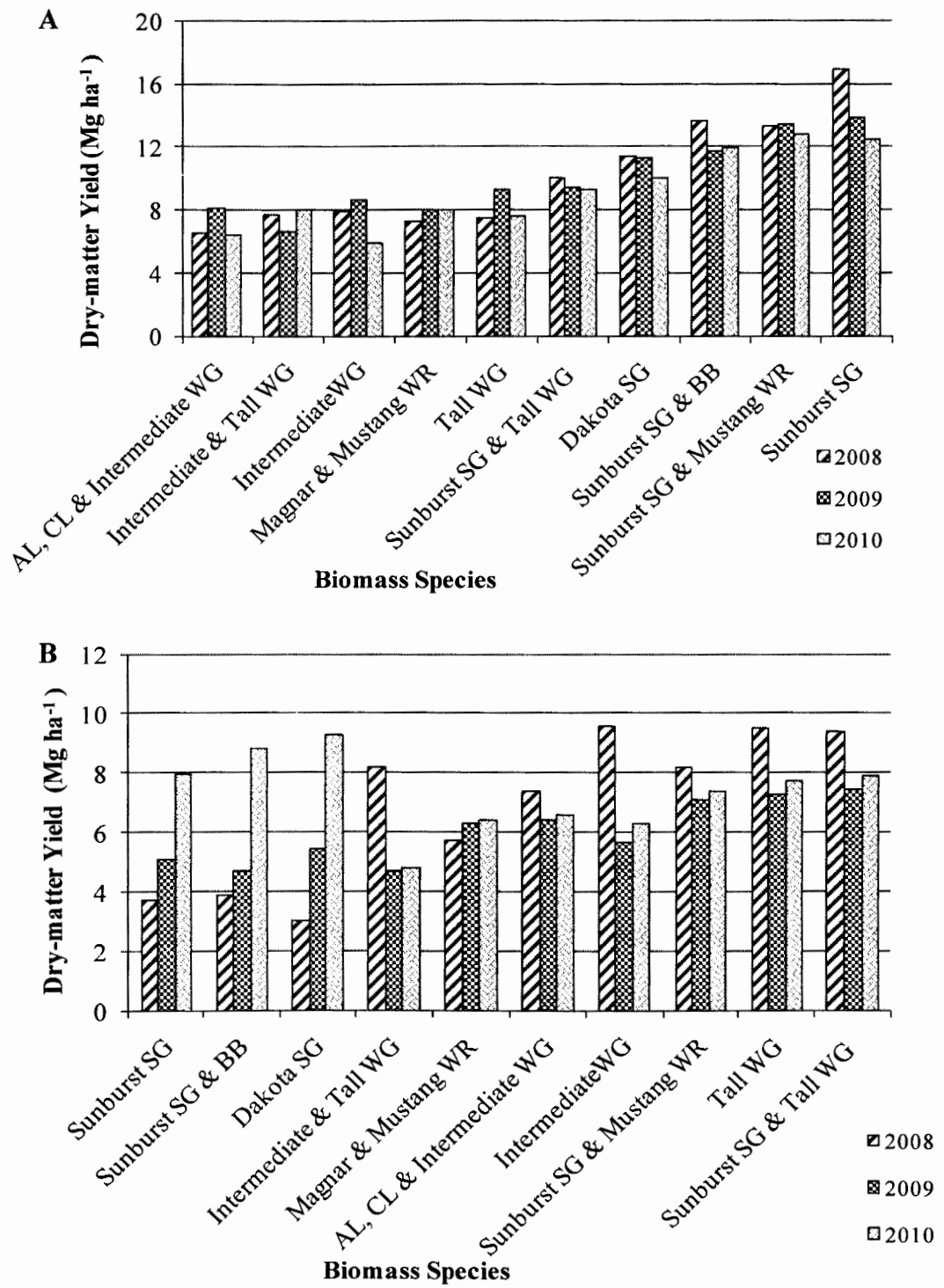


WG – Wheatgrass; SG – Switchgrass; WR – Wild rye; AL– Alfalfa, CL – Sweet clover; BB – Big bluestem.

Figure 5: Lignin content for biomass samples from Minot non-irrigated plot and Williston irrigated plots. Error bars represent standard deviation.

5.1.2. Biomass (dry-matter) yields

The NDSU-CGREC 2010 annual report gave the biomass yields after harvesting was done yearly during the 2nd and 3rd week of September (Nyren et al., 2010). These biomass yields were multiplied by the total solid content to obtain yields in dry Mg ha^{-1}. The dry-matter yield for biomass species from the Williston irrigated plots and Minot non irrigated plot for the years 2008, 2009, and 2010 are shown in Fig. 6.



WG – Wheatgrass; SG – Switchgrass; WR – Wild rye; AL– Alfalfa, CL – Sweet clover; BB – Big bluestem

Figure 6: Dry-matter yield for biomass species grown in (A) Williston irrigated plots from 2008 to 2010 and (B) Minot non-irrigated plots from 2008 to 2010 (Nyren et al., 2010).

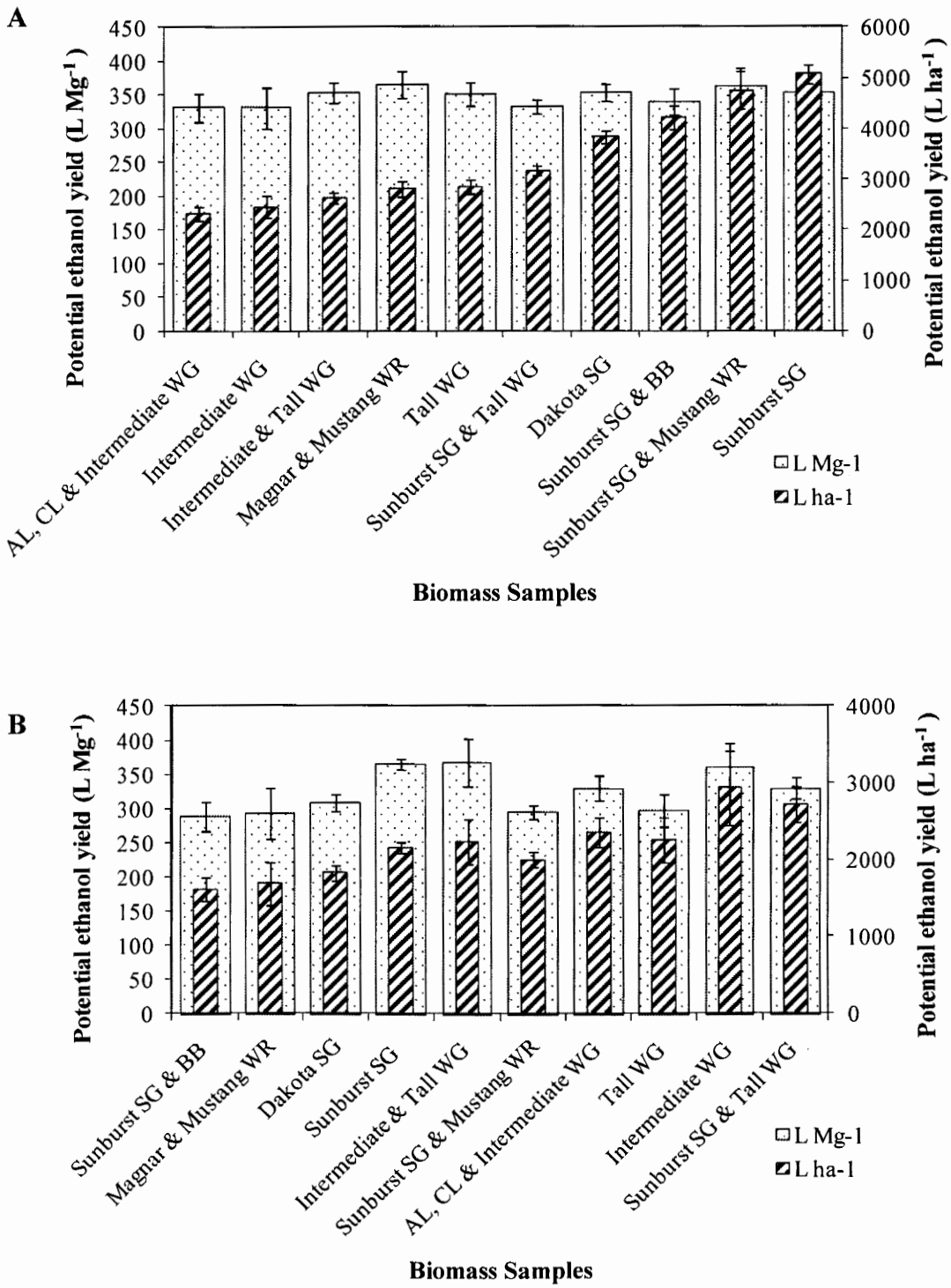
In Fig. 6A, switchgrass varieties as single or conglomerate species showed higher dry-matter yields in Williston irrigated plot than treatments without switchgrass. Sunburst SG had the highest average yields of 14.4 Mg ha^{-1} but the annual dry-matter yields fell from 16.9 Mg ha^{-1} in 2008 to 12.4 Mg ha^{-1} in 2010. Even though Sunburst SG yield decreased over the years, the 2010 dry-matter yield of 12.0 Mg ha^{-1} was only slightly lower than the highest yield of 12.7 Mg ha^{-1} from sunburst SG & Mustang WR for that same year. This decrease in dry-matter yields over the years was also seen in Dakota SG, Sunburst SG and Tall WG, and Sunburst SG and BB. With the dry-matter yields of wheatgrass species as single or conglomerate species stands, it was difficult to explain the variability observed in Fig. 6A. Magnar and Mustang WR showed a slight increase in dry-matter yields from 7.2 Mg ha^{-1} in 2008 to 8.0 Mg ha^{-1} in 2009 and 2010. In addition to rainfall, the irrigated plots in Williston are supplied with water to give a total of 635 mm of water per year. Water supply in Williston was almost constant yearly; therefore, temperature likely influenced the variability of dry-matter yields over the years. The difference in yields with switchgrass could be the warmer temperatures and higher GDD in 2008 compared to 2009 and 2010 shown in Table 4. Although some studies have reported that GDD do not significant influence on switchgrass yield in monoculture or mixture (Heaton et al., 2004; Wang et al., 2010), generally, C4 (warm-season) plants like switchgrass and big bluestem thrive best in warmer temperatures. The switchgrass samples as single or mixed feedstock has better dry-matter yields in 2008 with the warmest temperatures.

Fig. 6B showed that average dry-matter yields for wheatgrass species were higher in Minot (non-irrigated) but switchgrass and big bluestem yields were increasing from

2008 – 2010. Switchgrass species may have been slow to establish in Minot. Generally, treatments with switchgrass have a higher dry-matter yield on wetter soils than wheatgrass species. Elberson (2002) also reported that switchgrass yields were higher when irrigated or on wetter soils (2002). The high rainfall in Minot of 473.2 mm in 2010 as compared to 401.3 and 297.7 mm in 2008 and 2009, respectively, showed in Table 4 likely contributed to the 40 to 50% increase in dry-matter yield of switchgrass species from 2008 and 2009 to 2010. Even though the GDD was sufficient for biomass growth in all 3 years in Minot, it is worth noting from Table 4 that GDD was highest in 2010 than 2008 and 2009. The higher rainfall and heat accumulation may have caused SG yields in 2010 to be higher than in 2008 and 2009. Variability in yearly dry-matter yield was higher in Minot than in Williston from Fig. 6. This is because the Minot plots were not irrigated and depend solely on rainfall for water which varied yearly. The higher variability in water likely had a greater influence in dry-matter yield than temperature. Although there was no significant difference in mean temperature between location at $\alpha = 0.05$ (p-value = 0.722) and yearly (p-value = 0.996), small differences in temperature and their monthly distribution was observed and Myneni et al. (1997) reported that small differences in temperature can largely influence photosynthesis. This must have also contributed to the yearly variations in these two locations especially in Williston irrigated plots.

5.1.3. Theoretical ethanol yield

Analyzing the ethanol yield on an area basis (L ha^{-1}) is economically important because it relates to how much land is needed to feed a biorefinery of a given size. The average theoretical ethanol yield on a mass and area basis for three-year period in Williston and one-year period for Minot are shown in Figures 7A and 8B.



WG – Wheatgrass; SG – Switchgrass; WR – Wild rye; AL– Alfalfa, CL – Sweet clover; BB – Big bluestem

Figure 7: Potential ethanol yield on a mass and area basis for (A) Williston irrigated plots, three-year average and (B) Minot non-irrigated plots. Error bars represent standard deviation.

The value for theoretical ethanol yield on an area basis is obtained from both the total carbohydrate content and dry-matter yield. The Department of Energy (DOE) theoretical ethanol yield calculator (DOE, 2009) was used to determine ethanol yield on a mass basis (gal/t) by using carbohydrate content obtained from chemical composition and these values were converted to L Mg^{-1}. Biomass samples in Fig. 7 were arranged from the smallest to the largest amount of theoretical ethanol yield on an area basis. The switchgrass treatments as single or conglomerate species showed greatest potential for ethanol per ha in the Williston irrigated plot; wheatgrass treatments as single or conglomerate species showed the greatest potential for ethanol per ha in Minot. The lowest theoretical ethanol yield from Minot was 288 L Mg^{-1} while from Williston was 331.3 L Mg^{-1}. Theoretical ethanol yields in Williston varied from 331.3 to 365.1 L $Mg^{-1}$and 2,324.4 to 5,081.7 L ha^{-1} while in Minot they varied from 288 to 366.8 L Mg^{-1} and 1,614.8 to 2,719.8 L ha^{-1}. The theoretical ethanol yield per acre in Williston was almost twice that from Minot. A similar trend was seen with dry-matter yields in Figures 6A and 6B. The theoretical ethanol yield on an area basis is a product of biomass yield (Mg ha^{-1}) and theoretical ethanol yield on a mass basis (L Mg^{-1}). Table 7 shows the CV of biomass and theoretical ethanol yield of biomass species from Williston and Minot.

The CV for theoretical ethanol yield in L $ha^{-1}$was almost the same as that of biomass yield for both Williston and Minot. The variability in carbohydrate content among biomass species was almost nine times lower than the variability in biomass yields in Williston while in Minot it was two times lower than the biomass yields. The lower variability in biomass yield from Minot (15.0% vs 27.6% in Williston) can be attributed to the higher biomass yields of switchgrass treatments seen on the wetter soils of Williston.

Table 7: Biomass and theoretical ethanol yield for biomass species from Williston and Minot

Parameters	Williston (2008, 2009, and 2010)			Minot (2008)		
	Biomass yields (Mg ha^{-1})	Theoretical ethanol yield (L Mg^{-1})	Theoretical ethanol eield (L ha^{-1})	Biomass yields (Mg ha^{-1})	Theoretical ethanol yield (L Mg^{-1})	Theoretical ethanol yield (L ha^{-1})
Mean	9.8	347.5	3,411.4	6.7	323.0	2,179.1
SD	2.7	12.7	992.0	1.0	31	424
CV (%)	27.6	3.7	29.1	15.0	9.8	19.5

With plenty of water in Williston, the difference in biomass yields among species was larger than the less controlled environment in Minot. Table 7 also shows us that variability among biomass species in terms of potential ethanol yield in L ha^{-1} was mainly due to the variability in biomass yields because difference in composition among species were relatively small. Since variation on an area basis is far greater than variation on a mass basis, biomass production is much more important than composition in choosing species for ethanol production.

5.2. Development and variability test of NIR calibration model

5.2.1. Model evaluation

NIR calibration models were developed to predict lignin, glucan, and xylan contents for herbaceous perennial biomass species. Fifty-five biomass samples were used as a calibration set and ten biomass samples were used as validation set. A wide variability in constituent values lead to a more robust calibration model. Table 8 shows the wet chemistry data range for lignin, glucan, and xylan for all 65 biomass samples used for NIR calibration and validation and the data is found in the appendix.

Table 8: Summary of wet chemistry data for all 65 biomass samples

Constituent	Range (% w/w)	Mean (% w/w)	St. Dev. (% w/w)
Lignin	19.9 – 26.2	23.1	1.4
Glucan	21.2 – 40.7	31	3.2
Xylan	7.4 - 17.6	12.9	2.2

The values for all three constituents had been reported for several varieties of switchgrass and big bluestem in the DOE biomass database although a lower lignin content (13%) and a higher xylan content (24%) were reported for some species (DOE, 2006). This variability was due to differences in locations, species mixtures, year of harvest, and botanic fractions.

During model development, some grass samples were detected as outliers and were removed by using spectral and concentration residual plots (GRAMS, 2010). Due to the outliers removed, the number of samples used for the calibration model was 45, 46, and 46 for lignin, glucan, and xylan respectively. Outliers may have arisen from either erroneous wet chemistry results or spectra sample consistency is expected to be high because of homogenization through grinding. Variability in wet chemistry method was determined using a control sample that was analyzed eight times during chemical composition processing. The analysis is discussed in section 5.2.2. Also, some biomass samples were affected by static electricity during spectra collection and this may have influenced their spectra data as reported (Williams, 2008; Williams and Norris, 2001).

Figures 8A and 8B show the correlation between the wet chemistry and the predicted values for both the calibration and validation sets. The R^2 for the validation set was lower than the R^2 for the calibration set in all three constituents. The prediction for the

calibration set was done with the same biomass samples used to develop the model, therefore, a better prediction is expected than with the validation set for which a separate biomass sample was used. The American Association of Cereal Chemists (AACC) developed evaluation criteria with R/SEP that is used to validate NIR calibrations (AACC, 2000). They stated that a model with R/SEP ≥ 4 is fair and acceptable for screening, R/SEP ≥ 10 indicates a good and acceptable model for quality control, and R/SEP ≥ 15 indicates a very good model acceptable for research quantification.

Williams (2001) also gave some guidelines using R^2 that can be used to validate NIR models. He advised that a model with R^2 of 0.66 – 0.81 should be used for screening and approximate calibration; a model with R^2 of 0.83 – 0.90 is good for quality assurance purpose but should be used with caution; and a model with $R^2 \geq 0.92$, is very good for quality assurance. Table 9 shows the parameters obtained in this study to evaluate the NIR model developed.

Table 9: Statistics for model calibration and validation

Parameter	Lignin	Glucan	Xylan
Model evaluation			
Number of samples	45	46	46
SEP (%)	0.48	1.17	0.70
R/SEP	10.64	7.66	10.47
Model validation			
R^2	0.84	0.81	0.87
RMSEP	0.54	1.12	0.97

SEP = Standard error of prediction; R/SEP = Ratio of R (range of the validation dataset) to SEP; R^2 = Coefficient of determination; RMSEP = Root mean square error of prediction.

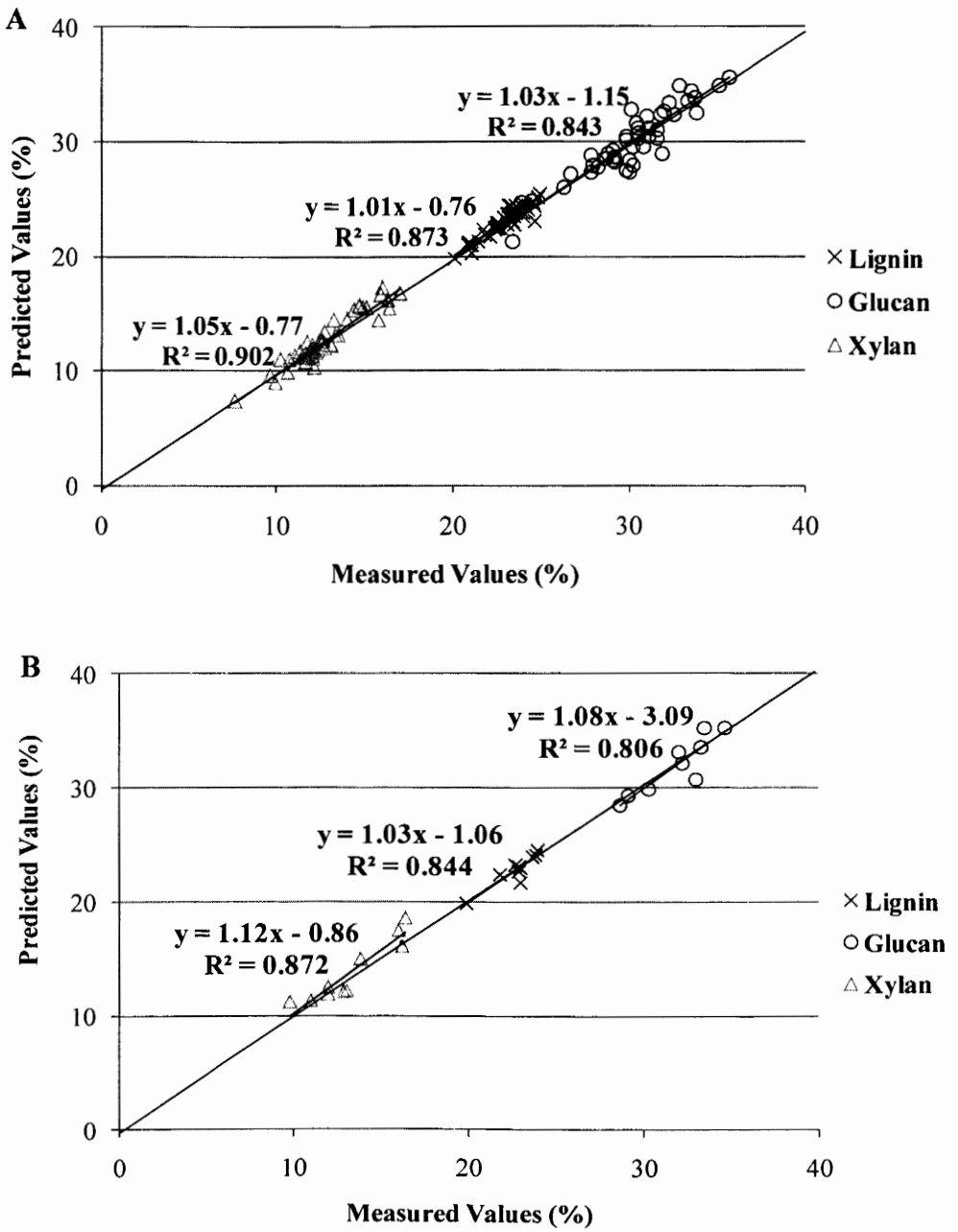


Figure 8: Chemical composition of perennial biomass species predicted by NIR calibration model versus measured values from wet chemistry method: (A) calibration set and (B) validation set.

Based on AACC standard, R^2 in Table 9 indicate that the models for lignin and xylan are acceptable for quality assurance but should be used with caution. Even though the R^2 values were slightly lower than those reported on single feedstock like corn stover

[lignin – 0.92, glucan – 0.85, and Xylan – 0.90] (Wolfrum and Sluiter, 2009); and rice straw [lignin - 0.89] (Jin and Chen, 2007); and switchgrass [lignin – 0.98, glucose – 0.93, xylose – 0.94] (Sanderson et al., 1996), it is important to note that the models developed were for multi-species biomass. Also, RMSEP values were small even though the models were developed with multiple feedstocks. The RMSEP values for lignin, glucan, and xylan of 1.12, 0.92, and 1.03 respectively was reported for a model developed for corn stover only (Ye et al., 2008). These values reported in Ye et al. (2008) were higher than the values reported in this study except the value for glucan. Both RMSEP and R^2 values show that the models were good when compared to the single biomass models that have been reported.

Comparing the models for lignin, glucan and xylan that was developed, RMSEP value indicate that lignin model is better that xylan model while R^2 and R/SEP values suggest that xylan model is better. Both lignin and xylan models had R/SEP values slightly above 10 which indicate that they are appropriate for quality assurance use. Based on R^2 and R/SEP criteria, the glucan model appears to be suitable for screening and approximation purposes only. Liu et al. (2010) developed a broad-based model for corn stover and switchgrass which had slightly better R/SEP values than the models in this study. The difference with the models in this study compared to that of Liu et al. (2010) were; (1) the wavelength range was between 950 to 1,650 nm as opposed to 1,000 to 2,500 nm for Liu et al. (2010), hence, glucan and xylan were developed with only the first and second overtone bands between 950 to 1,650 nm. The absence of the combination band (1,800 to 2,500 nm) of O-H and C-H stretching, which is important for carbohydrate prediction, may have reduced the performance of glucan and xylan models (Robert and

Cadet, 1998). (2) A diode array NIR spectrometer (advanced dispersive spectrometer) was used as opposed to FT-NIR. Both instruments have higher energy throughput than conventional dispersive but an FT-NIR can collect spectra faster and at lower resolution than a DA spectrometer. Even though these differences were observed, the R^2, RMSEP, and R/SEP from this study showed that it is possible to develop satisfactory calibration models to predict chemical composition for mixed perennial biomass using NIR.

5.2.2. Variability of NIR spectroscopy

After validating the calibration models, the models were combined and installed into the DA 7200 Perten spectrometer. Three biomass samples were used for the instrument and model variability (accuracy and precision) experiments. Samples were selected based on how best the models predicted at least two of the 3 constituents. The samples selected were: sample 1 - Intermediate & Tall WG (Williston 2010); sample 2 - Dakota SG stem only (Minot 2008); and sample 3 - Sunburst SG (Streeter 2008). Fig. 9 shows a bar chart comparing the predicted and measured values for these three samples. The predicted values were mean values of 10 daily spectroscopic readings. The daily data was used because it contains variability in both repacking and time.

Apart from xylan in sample 1, the differences between the NIR predicted and measured values were less than 2%. Glucan had smaller difference of 0.4% and 0.2% in samples 1 and 3 respectively, which indicated that the model accurately predicted glucan for these two samples. Lignin had a difference greater than 1% for all three samples. With these differences between the predicted and measured values, it is difficult to say which of the two measurements gives the true value for chemical composition. Generally, the wet

chemistry is considered to be the standard measurement but there is uncertainty in getting an accurate value with the method as will be explained later in this section.

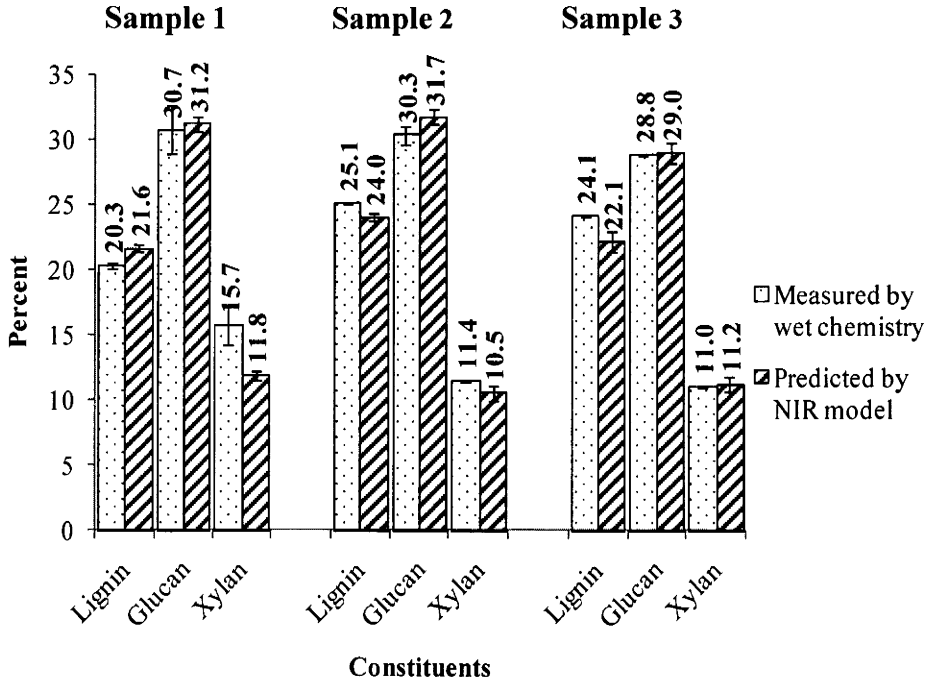


Figure 9: Comparison of NIR predicted and measured values for three selected biomass samples. Error bars represent standard deviation.

The developed spectroscopic models were used to study prediction variability in terms of time and presentation (repacking). Tables 10, 11, and 12 summarize results for prediction done 10 times after every 30 minutes with no repack, 10 times with repack on the same day, and 10 times daily with repack, respectively.

Table 10: NIR prediction for three biomass samples 10 times after every 30 minutes with no repack

Parameters	Sample 1			Sample 2			Sample 3		
	Lignin	Glucan	Xylan	Lignin	Glucan	Xylan	Lignin	Glucan	Xylan
Mean (%w/w)	21.36	30.45	12.35	24.41	31.30	10.92	22.31	29.20	11.42
SD (%w/w)	0.50	0. 35	0.32	0.26	0.18	0.16	0.36	0.23	0.15
CV (%)	**2.33**	**1.14**	**2.60**	**1.08**	**0.58**	**1.45**	**1.62**	**0.78**	**1.27**

Table 11: NIR prediction for three biomass samples done 10 times with repack on the same day

Parameters	Sample 1			Sample 2			Sample 3		
	Lignin	Glucan	Xylan	Lignin	Glucan	Xylan	Lignin	Glucan	Xylan
Mean (%w/w)	21.85	30.65	12.06	24.00	31.13	11.46	22.24	27.48	12.06
SD (%w/w)	0.6	0.51	0.37	0.36	0.43	0.44	0.88	0.56	0.44
CV (%)	**2.75**	**1.66**	**3.06**	**1.49**	**1.40**	**3.85**	**3.95**	**2.05**	**3.67**

Table 12: NIR prediction for three biomass samples 10 times daily with repack

Parameters	Sample 1			Sample 2			Sample 3		
	Lignin	Glucan	Xylan	Lignin	Glucan	Xylan	Lignin	Glucan	Xylan
Mean (%w/w)	21.61	31.17	11.83	24.01	31.68	10.47	22.11	29.01	11.23
SD (%w/w)	0.30	0.61	0.39	0.31	0.51	0.64	0.79	0.77	0.44
CV (%)	**1.38**	**1.95**	**3.29**	**1.27**	**1.59**	**6.15**	**3.57**	**2.66**	**3.95**

Williams (2001) gave some guidelines for interpreting CV in a repeatability or reproducibility test as shown in Table 13. The interpretation of CV depends on the situation and source of data because in some case, CV > 5% is satisfactory.

Table 13: Guidelines for interpreting coefficient of variation

CV Values (%)	Methods for compositional analysis	
	Wet Chemistry	NIR prediction
0 – 0.5	Exceptional	/
0.6 – 1.0	Excellent	Exceptional
1.1 – 2.0	Very good	Excellent
2.1 – 3.0	Good	Very good
3.1 - 4.0	Fair	Good
4.1 – 5.0	Poor	Fair
5.1+	Needs investigation	Poor

Source: (Williams, 2001).

Except the CV for xylan content of sample 2 (Table 10), all the CVs were less than 4%. This indicates that the repeatability of NIR prediction was good to exceptional according to Table 13. The repeatability of NIR prediction without repacking (Table 10)

was better than prediction done with repacking (Tables 11 and 12). The CVs in Table 10 were lower than CVs (Tables 11 and 12) except for the CV of lignin in sample 1. This shows that the NIR instrument is generally stable throughout the day as there was little sample to sample variability.

Looking at Tables 10 and 11, it is clear that repacking adds some variability in prediction. The average of CVs in Table 10 was 1.4% while in Table 11, it was 2.7%. This 1.3% increase in CV can be attributed to repacking only. Although it is recommended to leave the DA spectrometer overnight, the stability of the instrument over days rather than hours as the instrument is turned on and off was also tested and the results are shown on Table 12. Averaging the CVs in Tables 12 was 2.9% which was slightly higher the average CV of 2.7% in Table 11. Looking at this average change in CV in Table 11 and 12 indicate that daily variability was approximately 0.2%. Williams and Norris (2001) listed sources of instrument variability to be; change in wavelength and photometric scales, instrument temperature and humidity control, and mathematical treatment of spectral signal. This study shows that variability from repacking alone (1.3% CV change) had approximately the same effect as instrument variability of prediction over time (1.4 to 1.6% CV).

The assumption in developing and validating an NIR model is that the wet chemistry is the true value for composition. The uncertainty in determining an accurate measured value from NREL wet chemistry method has been reported (Templeton et al., 2010). The research reported standard deviations of 1 to 3% in composition for glucan, xylan, lignin and extractives when same biomass sample was analyzed several times from different analysts and laboratories. A 98% confidence interval of ±1% for lignin and ±1.5% for glucan and xylan for biomass composition have been published (Hames, 2010).

A control sample [Tall WG (Williston 2009)] was also used in this study to investigate the uncertainty during compositional analysis. After 8 trials, a 95% confidence interval for lignin, glucan, and xylan after wet chemistry were 23.3±0.7%, 32.7±1.1%, and 13.2±1.7% respectively. The range was 22.3 – 24.1%, 30.8 – 35.0%, 10.7 – 15.8% for lignin, glucan, and xylan, respectively. Therefore, one value in the lower or upper range can be off the true value significantly. Getting a true value to use as the reference value for NIR model development requires several wet chemistry analysis but the method is very expensive and time consuming. Fig. 10 shows the interval plot constructed from MINITAB 16 using the wet chemistry and NIR prediction data for the control sample.

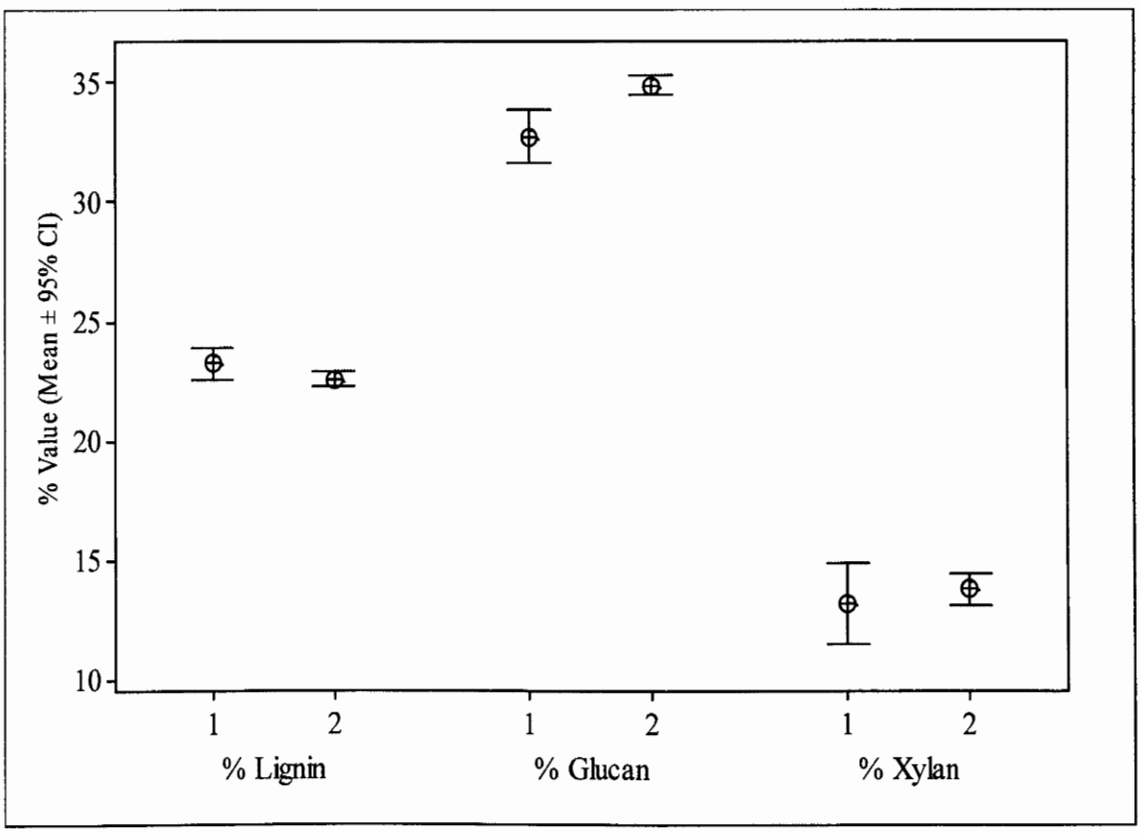


Figure 10: A 95% confidence interval plots to estimate the true values for lignin, glucan, and xylan using wet chemistry and NIR prediction for tall wheatgrass (control sample) after eight trials. 1 – Wet chemistry range; 2 – NIR prediction range.

NIR prediction was done daily (eight times) with repack on the control sample. Fig. 10 shows that confidence interval for lignin and xylan were in the same range for both the wet chemistry analysis and NIR model prediction. The prediction interval for glucan using both methods showed slightly different ranges of values. For all constituents, it might be preferable to choose the interval for NIR prediction for the true composition because the models were developed with 45 to 46 biomass samples. Once an NIR model is developed, the precision of the method is stronger than wet chemistry method as seen in the confidence interval in Fig. 10. Williams and Norris (2001) also confirmed that NIR prediction is often superior to the wet chemistry method in terms of precision. In terms of accuracy, it is difficult to know the true value. Wet chemistry method is used as the reference method to NIR model development. If wet chemistry analysis is done once, the probability of getting an inaccurate value is higher than in NIR. The intensive labor in the wet chemistry process is more liable to errors than NIR spectroscopy. The error can occur from any of the steps during the process. In some cases, histogram peaks and boundaries are difficult to be identified for a particular sugar or sample at the HPLC.

The CV can also be used to show the NIR prediction is more reliable than wet chemistry method. Using wet chemistry method, CV for lignin, glucan, and xylan were 3.4%, 4.0%, and 13.5% respectively for the control sample, while, with the NIR prediction, the CV was 1.8%, 1.4%, and 5.7% for lignin, glucan, and xylan respectively. From the guidelines in Table 13, it is observed that the repeatability of the chemical composition was fair for lignin and glucan while xylan was out of range. The CV from wet chemistry analysis for constituents with low concentration (< 5%) like galactan and arabinan had very high CV > 40%. Calibration models were not developed for galactan and arabinan because of the large

standard errors in their wet chemistry results. Therefore, it is important to pay close attention to all steps during wet chemistry analysis in order to have a reliable reference data.

6.0. CONCLUSION AND RECOMMENDATIONS

6.1. Conclusion

Herbaceous perennials have great potential in the cellulosic ethanol industries because they grow in diverse environments. Different perennial biomass species show different yield potentials in cellulose and hemicelluose formation with variation in climatic and environmental conditions. Maximizing production in cellulosic ethanol industries will need grass species with high biomass yields and high structural carbohydrate contents. This study demonstrated that the variability in carbohydrate yields among different biomass species is much less than the variability in dry-matter yields in different environments and over several years. That is, if these different perennial biomass species are grown under different environmental conditions, the proportion of cellulose and hemicelluose on a mass basis will not differ enough to compensate for lower dry-matter yields. As variation on an area basis is far greater than variation on a mass basis; thus, biomass production is much more important than composition in choosing species for ethanol production. Switchgrass varieties as single or mixed biomass species had better dry-matter yield and ethanol potential (area basis) on irrigated plots in Williston while wheatgrass species were better on non-irrigated plots in Minot. Therefore, switchgrass should be planted in area with plenty of rainfall or water but both switchgrass and wheatgrass species can be planted in area where rainfall amount greatly fluctuates yearly.

This study also demonstrated that it is possible to develop an NIR calibration model for mixed perennial biomass species as the values of R^2, RMSEP, R/SEP were good when compared to models developed for single feedstocks. The models developed for lignin and xylan were good for quality assurance based on their R^2 and R/SEP while glucan was

appropriate for screening. The NIR instrument in predicting chemical constituent was stable over time (hours and days) as the CV was lowest without repacking. The variability due to repacking alone was approximately the same as the variability observed by the instrument over time. Variability, as measured by CV, was generally lower for repeated NIR predictions than for wet chemistry analysis while their means were generally in the same range. Factors like sieve size, ambient temperature, sample cell type, and moisture affecting the variability of NIR calibration are easily controllable; therefore, if reliable reference data is obtained, an accurate calibration model for mixed biomass species can be developed.

6.2. Recommendations

Further study should be done with Minot samples harvested in 2009 and 2010, and also beyond 2010 in both Williston and Minot in order to increase the dataset for the variability study of the total carbohydrate content and dry-matter yield among perennial biomass species. Switchgrass species as single or mixed feedstock had an increasing trend in dry-matter yield from 2008 to 2010 in the non-irrigated plots in Minot. It is therefore important to study this trend alongside their carbohydrate content.

Further research to improve these models for mixed perennial biomass can be done by using an NIR instrument that will have faster scan rate and lower resolution like the FT-NIR. The NIR instrument should be able to collect spectra up to 2,500 nm in which the combination bands from 1,800 to 2,500 nm will be included and also to use more samples than was used in this research. Increasing the number of samples (80 to 100) might make the models to be more robust.

REFERENCES

AACC. 2000. Near-Infrared methods—Guidelines for model development and maintenance (AACC Method 39-00). American Association of Cereal Chemists, St. Paul, Minnesota, USA.

Adler, P. R., Sanderson, M. A., Boateng, A. A., Weimer, P. J., and Jung, H.-J. G. 2006. Biomass yield and biofuel quality of switchgrass harvested in fall or spring. *Agron J.* 98(6): 1518-1525.

Adler, P. R., Sanderson, M. A., Weimer, P. J., and Vogel, K. P. 2009. Plant species composition and biofuel yields of conservation grasslands. *Ecological Applications.* 19(8): 2202-2209.

Agblevor, F. A., and Besler, S. 1996. Inorganic compounds in biomass feedstocks: effect on the quality of fast pyrolysis oils. *Energy & Fuels.* 10(2): 293-298.

Agblevor, F. A., Evans, R. J., and Johnson, K. D. 1994. Molecular-beam mass-spectrometric analysis of lignocellulosic materials : I. Herbaceous biomass. *Journal of Analytical and Applied Pyrolysis.* 30(2): 125-144.

Badger, P. C., and Fransham, P. 2006. Use of mobile fast pyrolysis plants to densify biomass and reduce biomass handling costs--A preliminary assessment. *Biomass and Bioenergy.* 30(4): 321-325.

Bevington, P. R., and Robinson, D. K. 1992. *Data reduction and error analysis for the physical sciences.* second ed. Boston, Massachusetts: WCB McGraw-Hill.

Bjørsvik, H.-R., and Martens, H. 2008. Data analysis: calibration of NIR instrument by PLS regression. In *Handbook of Near-Infrared Analysis*, 189-206. Boca Raton, Fl: CRC Press.

Blanco, M., and Villarroya, I. 2002. NIR spectroscopy: a rapid-response analytical tool. TrAC Trends in Analytical Chemistry, 240-250.

Bloom, A. J., Chapin, F. S., and Mooney, H. A. 1985. Resource limitation in plants-an economic analogy. *Annual Review of Ecology and Systematics*. 6: 363-392

Boateng, A. A., Weimer, P. J., Jung, H. G., and Lamb, J. F. S. 2008. Response of thermochemical and biochemical conversion processes to lignin concentration in alfalfa stems. *Energy & Fuels*. 22(4): 2810-2815.

Boysworth, M. K., and Booksh, K. S. 2008. Aspesct of multivariate calibration applied to near-infrared spectroscopy. In *Handbook of Near-Infrared Analysis*, 207-230. Boca Raton, Fl: CRC Press.

Cardona, C. A., and Sánchez, Ó. J. 2007. Fuel ethanol production: Process design trends and integration opportunities. *Bioresource Technology*. 98(12): 2415-2457.

Casler, M. D., Vogel, K. P., Taliaferro, C. M., and Wynia, R. L. 2004. Latitudinal adaptation of switchgrass populations. *Crop Science*. 44: 293-303.

Chai, Q., Jin, F., Merewitz, E., and Huang, B. 2010. Growth and Physiological Traits Associated with Drought Survival and Post-drought Recovery in Perennial Turfgrass Species. *Journal of the American Society of Horticultural Science*. 135(2): 125-133.

Chen, J., Arnold, M. A., and Small, G. W. 2004. Comparison of combination and first overtone spectral regions for near-infrared calibration models for glucose and other biomolecules in aqueous solutions. *Analytical Chemistry*. 76(18): 5405-5413.

Coates, J. 2000. Interpretation of infrared spectra, a pratical approach. *Encyclopedia of Analytical Chemistry*: 10,815-10,837.

Curtis, B. 2008. U.S. ethanol industry: the reflection point. *United States Department of Energy (USDOE), Washington D.C.*

Dien, B. S., Jung, H. J. G., Vogel, K. P., Casler, M. D., Lamb, J. F. S., Iten, L., Mitchell, R. B., and Sarath, G. 2006. Chemical composition and response to dilute-acid pretreatment and enzymatic saccharification of alfalfa, reed canarygrass, and switchgrass. *Biomass and bioenergy*. 30(10): 880-891.

DOE. 2006. *Biomass feedstock composition and property database.*

DOE. 2009. *Theoretical ethanol yield calculator.*

Downing, M., Walsh, M., and McLaughlin, S. 1995. Perennial grasses for energy and conservation: evaluating some ecological, agricultural, and economic issues. *Center for Agriculture, Food and Environment, Tufts University, Medford, MA.*

Dwivedi, P., Alavalapati, J. R. R., and Lal, P. 2009. Cellulosic ethanol production in the United States: Conversion technologies, current production status, economics, and emerging developments. *Energy for Sustainable Development*. 13(3): 174-182.

EIA. 2008. Energy independence and security act of 2007: summary of provisions.

EIA. 2011. How dependent are we on foreign oil?

Elberson, H. W. 2002. Switchgrass as an alternative energy crop in Europe: Initiative of a productivity network. *Final Report FAIR 5-CT 97-3701.* www.switchgrass.nl.

Elena, P. 2004. The evolution of applied harmonic analysis: model of the real world. *Springer Science and Business Media.*: 335.

Gough, L., Osenberg, C. W., Gross, K. L., and Collins, S. L. 2000. Fertilization effects on species density and primary productivity in herbaceous plant communities. *OIKOS.* 89: 428-439.

Curtis, B. 2008. U.S. ethanol industry: the reflection point. *United States Department of Energy (USDOE), Washington D.C.*

Dien, B. S., Jung, H. J. G., Vogel, K. P., Casler, M. D., Lamb, J. F. S., Iten, L., Mitchell, R. B., and Sarath, G. 2006. Chemical composition and response to dilute-acid pretreatment and enzymatic saccharification of alfalfa, reed canarygrass, and switchgrass. *Biomass and bioenergy*. 30(10): 880-891.

DOE. 2006. *Biomass feedstock composition and property database.*

DOE. 2009. *Theoretical ethanol yield calculator.*

Downing, M., Walsh, M., and McLaughlin, S. 1995. Perennial grasses for energy and conservation: evaluating some ecological, agricultural, and economic issues. *Center for Agriculture, Food and Environment, Tufts University, Medford, MA.*

Dwivedi, P., Alavalapati, J. R. R., and Lal, P. 2009. Cellulosic ethanol production in the United States: Conversion technologies, current production status, economics, and emerging developments. *Energy for Sustainable Development*. 13(3): 174-182.

EIA. 2008. Energy independence and security act of 2007: summary of provisions.

EIA. 2011. How dependent are we on foreign oil?

Elberson, H. W. 2002. Switchgrass as an alternative energy crop in Europe: Initiative of a productivity network. *Final Report FAIR 5-CT 97-3701.* www.switchgrass.nl.

Elena, P. 2004. The evolution of applied harmonic analysis: model of the real world. *Springer Science and Business Media.*: 335.

Gough, L., Osenberg, C. W., Gross, K. L., and Collins, S. L. 2000. Fertilization effects on species density and primary productivity in herbaceous plant communities. *OIKOS*. 89: 428-439.

GRAMS. 2010. GRAMS chemometrics with GRAMS IQ: training course. T. F. Scientific, Ed., Cheshire, WA14 5TP, UK.

Griffiths, P. R., and De Haseth, J. A. 2007. Fourier transform infrared spectrometry. *Second Edition. John Wiley & Sons, New York.*

Hames, B. R. 2010. Biomass compositional analysis for energy applications. In *Biofuels*, 145-167. J. R. Mielenz, Ed.: Humana Press.

Hames, B. R., Thomas, S. R., Sluiter, A. D., Roth, C. J., and Templeton, D. W. 2003. Rapid biomass analysis. *Applied Biochemistry and Biotechnology.* 105(1): 5-16.

Heaton, E., Voigt, T., and Long, S. P. 2004. A quantitative review comparing the yields of two candidate C4 perennial biomass crops in relation to nitrogen, temperature and water. *Biomass and Bioenergy.* 27(1): 21-30.

Hein, P. R. G., Lima, J. T., and Chaix, G. 2009. Effects of sample preparation on NIR spectroscopic estimation of chemical properties of Eucalyptus urophylla S.T. Blake wood. *Holzforschung.* 64(1): 45-54.

Helland, I. S. 1990. Partial least squares regression and statistical models. *Scandinavian Journal of Statistics.* 17(2): 97-114

Himmel, M. E., Ding, S.-Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., and Foust, T. D. 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science.* 315(5813): 804-807.

Hollas, J. M. 1996. Modern spectroscopy. *John wileys & Sons Ltd, England.*

Hsu, C.-P. S. 1997. Infrared spectroscopy. In *Handbook of Instrumental Techniques for Analytical Chemistry*, 247-283: Prentice-Hall, Inc.

IPS. 2010. TAPPI test method standards. Intergrated Paper Service. Accessed July 8, 2010 http://www.ipstesting.com/IPSFinder/Method/TAPPIPulpandPaperTestingMethods/tabid/94/Default.aspx.

Jiang, W., Han, G., Zhang, Y., and Wang, M. 2010. Fast compositional analysis of ramie using near-infrared spectroscopy. *Carbohydrate Polymers*. 81(4): 937-941.

Jin, S., and Chen, H. 2007. Near-infrared analysis of the chemical composition of rice straw. *Industrial Crops and Products*. 26(2): 207-211.

Kelley, S. S., Rowell, R. M., Davis, M., Jurich, C. K., and Ibach, R. 2004. Rapid analysis of the chemical composition of agricultural fibers using near infrared spectroscopy and pyrolysis molecular beam mass spectrometry. *Biomass and Bioenergy*. 27(1): 77-88.

Kemeny, G. 2005. Applications for the intergrating spehere in the near infrared spectral region. *Application Notes, PIKE tecnologies: Spectroscopic Creativity*.

Klooster, B., and Palmer-Young, E. 2004. Water stress marginally increases stomatal density in *E. canadensis*, but not in *A. gerardii*. *Tillers*. 5: 35-40.

Labbé, N., Rials, T. G., Kelley, S. S., Cheng, Z.-M., Kim, J.-Y., and Li, Y. 2005. FT-IR imaging and pyrolysis-molecular beam mass spectrometry: new tools to investigate wood tissues. *Wood Science and Technology*. 39(1): 61-76.

Labbé, N., Ye, X. P., Franklin, J. A., Womac, A. R., Tyler, D. D., and Rials, T. G. 2008. Analysis of switchgrass characteristics using near infrared spectroscopy. *BioResources*. 3(4): 1329-1348.

Lin, Y., and Tanaka, S. 2006. Ethanol fermentation from biomass resources: current state and prospects. *Applied Microbiological Biotechnology*. 69: 627-642.

Lindstrom, P. 2011. Environment. Washignton DC: US Energy Information Administration.

Liu, L., Ye, X. P., Womac, A. R., and Sokhansanj, S. 2010. Variability of biomass chemical composition and rapid analysis using FT-NIR techniques. *Carbohydrate Polymers*. 81(4): 820-829.

Martone, P. T., Estevez, J. M., Lu, F., Ruel, K., Denny, M. W., Somerville, C., and Ralph, J. 2009. Discovery of Lignin in Seaweed Reveals Convergent Evolution of Cell-Wall Architecture. *Current Biology*. 19(2): 169-175.

McCarthy, W. J., and Kemeny, G. J. 2008. Fourier Transform spesctrophotometers in the near-infrared. In *Handbook of Near-Infrared Analysis*, 79-92. Boca Raton, Fl: CRC Press.

McClure, W. F. 2001. Near-Infrared Instrumentation. In *Near-Infrared Technology in the Agricultural and Food Industries*, 109-128. St. Paul, Minnesota: American Association of Cereal Chemists.

McKendry, P. 2002a. Energy production from biomass (part 1): overview of biomass. *Bioresource Technology*. 83(1): 37-46.

McKendry, P. 2002b. Energy production from biomass (part 2): conversion technologies. *Bioresource Technology*. 83: 47-54.

Michalet, R., Brooker, R. W., Cavieres, L. A., Kikvidze, Z., Lortie, C. J., Pugnaire, F. I., Valiente-Banuet, A., and Callaway, R. M. 2006. Do biotic interactions shape both sides of the humped-back model of species richness in plant communities? *Ecology Letters*. 9(7): 767-773.

Milne, T. A., Chum, H. L., Agblevor, F., and Johnson, D. K. 1992. Standardized analytical methods. *Biomass and Bioenergy*. 2(1-6): 341-366.

Mitchell, D. 2008. A note of rising food prices. Policy research working paper 4682. *Development Prospect Group. Washington DC, USA: The World Bank.*

Myneni, R. B., Keelling, C. D., Tucker, C. J., Asrar, G., and R., N. R. 1997. Increase plant growth in the northern high latitudes from 1981 to 1991. *Nature*. 386(17): 698-702.

NREL. 2010. *Standard biomass analytical procedures*.

Nyren, P., Wang, G., Patton, B., Bradbury, G., Eriksmoen, E., Halvorson, M., and Aberle, E. 2010. Perennial forages for biofuel production in central and western North Dakota. North Dakota State University.

Nyren, P., Xue, Q., Aberle, E., Bradbury, G., Eriksmoen, E., Halvorson, M., Nichols, K., Liebig, M., Bohn, R., Nyren, A., and Patton, B. 2009. Composition and Production of Perennial Grasses for Biofuel Production in Central and Western North Dakota. North Dakota State University.

Nyren, P. E., Eriksmoen, E., Bradbury, G., Halverson, M., Aberle, E., Nichols, K., Liebig, M., and Patton, B. 2007. The Evaluation of Selected Perennial Grasses for Biofuel Production in Central and Western North Dakota. North Dakota State University.

Partel, M., Laanisto, L., and Zobel, M. 2007. Contrasting plant productivity–diversity relationships across latitute: The role of evoluntionary history. *Ecology*. 88(5): 1091-1097.

Perkins, C. M., Woodruff, B., Andrews, L., Lichty, P., lancaster, B., Bingham, C., and Weimer, A. W. 2008. Synthensis gas production by rapid solar thermal gasification

of corn stover. *14th Biennial CSP SolarPACES (Solar Power and Chemical Energy Systems) Symposium, 4-7 March, Las Vegas, Nevada.*

Perlack, R. D., Wright, L. L., Turhollow, A. F., Graham, R. L., Stokes, B. J., and Erbach, D. C. 2005. Biomass as feedstock for a bioenergy and bioproducts industry: The technical feasibility of a billion-ton annual supply.: US Department of Agriculture and US Department of Energy.

Pimentel, D., and Patzek, T. W. 2008. Ethanol production using corn, switchgrass and wood; biodiesel production using soybean In *Biofuels, Solar and Wind as Renewable Energy Systems*, 373-394. D. Pimentel, Ed. New york: Springer.

Robert, C., and Cadet, F. 1998. Analysis of near-infrared spectra of some carbohydrates. *Applied Spectroscopy Reviews*. 33(3): 253 - 266.

Sanderson, M. A. 1992. Morphological development of switchgrass and kleingrass. *Agronomy Journal*. 84(3): 415-419.

Sanderson, M. A., and Adler, P. R. 2008. Perennial forages as second generation bioenergy crops. *International Journal of Molecular Sciences*. 9(5): 768–788.

Sanderson, M. A., Agblevor, F., Collins, M., and Johnson, D. K. 1996. Compositional analysis of biomass feedstocks by near infrared reflectance spectroscopy. *Biomass and Bioenergy*. 11(5): 365-370.

Sanderson, M. A., and Wolf, D. D. 1995. Switchgrass biomass composition during morphological development in diverse environments. *Crop Science*. 35(5): 1432-1438.

Shenk, J. S., Workman, J., Jr., and Westerhaus, M. O. 2008. Application of NIR spectroscopy to agricultural products. In *Handbook of Near-Infrared Analysis*, 347-386. Boca Raton, Fl: CRC Press.

Siesler, H. W. 2008. Basic principles of near-infrared spectroscopy. In *Handbook of Near-Infrared Analysis*, 7-19. Boca Raton, Fl: CRC Press.

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., and Templeton, D. 2005a. Determination of ash in biomass., National Renewable Energy Laboratory (Technical Report NREL/TP-510-42622) Revised January 2008.

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D., and Crocker, D. 2008. Determination of structural carbohydrates and lignin in biomass., National Renewable Energy Laboratory (Techninal Report NREL/TP-510-42618) Revised January 2010.

Sluiter, A., Ruiz, R., Scarlata, C., Sluiter, J., and Templeton, D. 2005b. Determination of extractives in biomass., National Renewable Energy Laboratory (Technical Report NREL/TP-510-42619) Revised January 2008.

Sluiter, J. B., Ruiz, R. O., Scarlata, C. J., Sluiter, A. D., and Templeton, D. W. 2010. Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. *Journal of Agricultural and Food Chemistry*. 58(16): 9043-9053.

Smith, B. 1999. Infrared spectral interpretation, a systematic approach. *CRC Press, Boca Raton, Fl.*

Stuart, B. 2004. Infrared spectroscopy: fundametals and applications. *Analytical techniques in science. John wileys & Sons Ltd, England.*

Sun, J. 1997. Statistical analysis of NIR data: Data pretreatment. *Journal of Chemometrics*. 11(6): 525-532.

Templeton, D. W., Scarlata, C. J., Sluiter, J. B., and Wolfrum, E. J. 2010. Compositional analysis of lignocellulosic feedstocks. 2. method uncertainties. *Journal of Agricultural and Food Chemistry*. 58: 9054-9062.

Theander, O. 1991. Chemical analysis of lignocellulose materials. *Animal Feed Science and Technology*. 32(1-3): 35-44.

Waltz, E. 2008. Cellulosic ethanol booms despite unproven business models. *Nat Biotech*. 26(1): 8-9.

Wang, D., Lebauer, D., and Dietze, M. 2010. A quantitative review comparing the yield of switchgrass in monocultures and mixtures in relation to climate and management factors. *GCB Bioenergy*. 2: 16-25.

Williams, P. 2008. Sampling, sample preparation, and sample selection. In *Handbook of Near-Infrared Analysis*, 267-306. Boca Raton, Fl: CRC Press.

Williams, P. C. 2001. Implementation of near-infrared technology. In *Near-Infrared Technology*, 145-169. St. Paul, Minnesota: American Association of Cereal Chemists.

Williams, P. C., and Norris, K. 2001. Variables affecting near-infrared spectroscopic analysis. In *Near-Infrared Technology*, 171-185. St. Paul, Minnesota: American Association of Cereal Chemists.

Wolfrum, E., and Sluiter, A. 2009. Improved multivariate calibration models for corn stover feedstock and dilute-acid pretreated corn stover. *Cellulose*. 16(4): 567-576.

Workman, J., Jerome J. 2008. NIR spectroscopy calibration basics. In *Handbook of Near-Infrared Analysis*, 123-150. Boca Raton, Fl: CRC Press.

Workman, J., Jerome J., and Burns, D. A. 2008. Commercial NIR Instrumentation. In *Handbook of Near-Infrared Analysis*, 67-78. Boca Raton, Fl: CRC Press.

Workman, J. J. 1993. A review of process near infrared spectrometry: 1980-1994. *Near Infrared Spectrometry*. 1: 221-245.

Xu, L., Zhou, Y.-P., Tang, L.-J., Wu, H.-L., Jiang, J.-H., Shen, G.-L., and Yu, R.-Q. 2008. Ensemble preprocessing of near-infrared (NIR) spectra for multivariate calibration. *Analytica Chimica Acta*. 616(2): 138-143.

Yaman, S. 2004. Pyrolysis of biomass to produce fuels and chemical feedstocks. *Energy Conversion and Management*. 45(5): 651-671.

Ye, X. P., Liu, L., Hayes, D., Womac, A., Hong, K., and Sokhansanj, S. 2008. Fast classification and compositional analysis of cornstover fractions using Fourier transform near-infrared techniques. *Bioresource Technology*. 99(15): 7323-7332.

APPENDIX

Table A1: Compositional analysis data for Williston samples harvested in 2008 showing mean and standard deviation

Sample Name	% TS	% Ext	% Lig	% Glu	% Xyl	% Ara	% Gal	% Ash	Mass closure
Tall WG	95.0 ±0.1	8.9 ±0.1	24.5 ±0.2	34.4 ±0.4	12.0 ±0.1	4.1 ±0.2	1.3 ±0.0	6.2 ±0.1	91.4
Dakota SG	93.9 ±0.1	6.5 ±0.1	26.2 ±0.4	30.9 ±0.5	12.3 ±0.1	5.2± 0.0	0.9± 0.6	5.1 ±0.0	87.1
Intermediate WG	95.1 ±0.0	9.0 ±0.1	22.7 ±0.0	27.3 ±3.2	14. ±2.3	5.1 ±2.1	2.7 ±0.0	6.2 ±0.0	87.3
Sunburst SG	94.1 ±0.1	13.5 ±0.3	22.8 ±0.6	34.9 ±0.8	16.7 ±0.2	3.8 ±0.5	0.7 ±0.4	5.7 ±0.2	98.0
Intermediate & Tall WG	95.6 ±0.1	10.5 ±0.2	23.7 ±1.0	31.9 ±0.2	15.4 ±0.1	3.3 ±0.4	1.7 ±0.2	6.2 ±0.0	92.7
Sunburst SG & Tall WG	94.8 ±0.2	10.5 ±0.2	24.2 ±0.6	33.5 ±1.2	12.4 ±0.6 1	5.2 ±0.4	1.9 ±0.0	5.8 ±0.0	93.5
Magnar & Mustang WR	94.8 ±0.4	10.0 ±0.2	22.6 ±0.2	32.3 ±1.5	16.3 ±0.9	7.7 ±0.1	3.8 ±0.0	6.6 ±0.0	99.3
Alfalfa, sweet clover & Inter WG	94.1 ±0.0	10.8 ±0.1	24.7 ±0.9	30.4 ±1.1	10.8 ±0.4	4.8 ±0.1	2.2 ±0.4	6.2 ±0.1	89.8
Sunburst SG & Big bluestem	94.0 ±0.0	11.1 ±1.2	22.9 ±0.1	35.6 ±2.9	10.3 ±0.2	3.5 ±0.7	1.8 ±0.0	5.6 ±0.0	90.8
Sunburst SG & Mustang WR	94.2 ±0.1	12.2 ±0.3	22.5 ±0.1	36.9 ±0.3	12.2 ±0.6	5.0 ±1.2	2.3 ±1.0	5.7 ±0.0	96.8

TS- Total Solid; Ext- Extactives; Lig – Lignin; Xyl – Xylan; Ara – Arabinan; Gal – Galactan; WG – Wheatgrass; SG – Switchgrass; WR – Wild rye.

Table A2: Compositional analysis data for Williston samples harvested in 2009 showing mean and standard deviation

Sample Name	% TS	% Ext	% Lig	% Glu	% Xyl	% Ara	% Gal	% Ash	Mass closure
Tall WG	97.8 ±0.1	8.2 ±0.1	23.6 ±0.3	32.1 ±1.8	12.1 ±1.1	2.8 ±0.4	0.6 ±0.0	7.6 ±0.2	86.80
Dakota SG	95.5 ±0.3	21.2 ±0.2	21.0 ±0.2	29.5 ±2.0	10.2 ±0.8	4.1 ±0.6	1.1 ±0.0	6.9 ±0.3	93.99
Intermediate WG	94.2 ±0.3	9.6± 0.3	24.0 ±0.5	27.8 ±0.6	12.7 ±0.6	3.0 ±0.1	1.0 ±0.0	10.5 ±0.3	88.50
Sunburst SG	97.0 ±0.1	8.6± 0.1	22.0 ±0.1	29.9 ±0.9	13.0 ±0.5	3.3 ±0.3	1.1 ±0.1	6.9 ±0.1	84.68
Intermediate & Tall WG	95.9 ±0.1	18.2 ±0.5	22.5 ±0.2	33.9 ±2.6	14.6 ±1.1	3.2 ±1.0	0.9 ±0.1	6.9 ±0.6	100.50
Sunburst SG & Tall WG	97.0 ±0.1	10.5 ±0.2	22.4 ±0.0	30.2 ±0.0	11.8 ±0.1	2.5 ±0.0	0.8 ±0.1	10.0 ±0.2	88.14
Magnar & Mustang WR	96.8 ±0.5	7.2± 0.3	24.5 ±0.4	31.2 ±1.8	13.2 ±1.1	3.5 ±0.1	1.9 ±1.2	6.6 ±0.2	87.96
Alfalfa, sweet clover & Inter WG	94.2 ±0.2	12.2 ±0.0	23.9 ±0.1	28.6 ±1.7	9.9± 0.58	2.9 ±0.7	2.1 ±1.1	7.2 ±0.5	86.71
Sunburst SG & BB	94.3 ±0.1	11.1 ±0.2	22.7 ±0.1	32.5 ±0.1	13.4 ±0.4	3.1 ±0.0	0.9 ±0.0	6.5 ±0.0	90.17
Sunburst SG & Mustang WR	94.9 ±0.2	8.7± 0.3	22.5 ±1.1	31.8 ±0.4	11.8 ±0.2	5.9 ±0.4	2.7 ±0.0	12.1 ±0.0	96.19
Intermediate WG leaf	94.2 ±0.2	9.3± 0.6	22.2 ±0.4	26.0 ±0.4	12.2 ±0.2	3.9 ±0.2	1.4 ±0.1	12.0 ±0.0	87.01
Intermediate WG stem	94.7 ±0.0	7.1± 0.5	23.5 ±0.3	27.5 ±1.5	11.9 ±0.5	3.1 ±0.2	1.5 ±0.8	6.3 ±0.4	80.91
Sunburst SG & Mustang WR leaf	94.2 ±0.1	10.7 ±0.1	23.1 ±1.1	35.2 ±0.8	11.3 ±0.8	8.8 ±0.4	1.7 ±0.2	8.5 ±0.2	99.32
Sunburst SG & Mustang WR stem	94.8 ±0.2	12.0 ±0.4	23.9 ±0.3	35.8 ±0.0	12.8 ±1.5	4.2 ±0.1	0.4 ±0.1	4.8 ±0.1	93.96

TS- Total Solid; Ext- Extactives; Lig – Lignin; Xyl – Xylan; Ara – Arabinan; Gal – Galactan; WG – Wheatgrass; SG – Switchgrass; WR – Wild rye.

Table A3: Compositional analysis data for Williston samples harvested in 2010 showing mean and standard deviation

Sample Name	% TS	% Ext	% Lig	% Glu	% Xyl	% Ara	% Gal	% Ash	Mass closure
Tall WG	94.3 ±0.1	9.4 ±0.1	21.0 ±0.2	29.6 ±1.8	15.4 ±1.2	7.4 ±0.2	1.6 ±1.5	8.1 ±0.1	92.4
Dakota SG	94.6 ±0.4	5.5 ±0.2	22.6 ±0.2	32.6 ±0.1	16.7 ±0.5	9.5± 0.3	2.2 ±2.0	6.4 ±0.5	95.5
Intermediate WG	94.2 ±0.0	10.7 ±0.1	19.8 ±0.2	29.3 ±1.8	16.1 ±0.9	9.0± 0.4	2.2 ±1.9	9.5 ±0.1	96.5
Sunburst SG	94.6 ±0.0	8.3 ±0.3	22.4 ±1.0	32.1 ±0.9	17.5 ±0.4	5.4± 0.6	1.0 ±0.4	7.5 ±0.3	95.0
Intermediate & Tall WG	94.6 ±0.1	10.0 ±0.7	20.3 ±0.2	30.7 ±1.9	15.7 ±1.2	4.7± 0.2	1.8 ±1.5	9.5 ±0.7	95.6
Sunburst SW & Tall WG	95.2 ±0.2	9.9 ±0.3	21.1 ±0.3	30.5 ±1.1	15.6 ±0.5	4.3± 1.0	2.1 ±1.5	7.8 ±0.1	90.7
Magnar & Mustang WR	94.3 ±0.2	8.5 ±0.1	21.3 ±0.4	32.8 ±1.3	16.6 ±1.2	5.7± 1.2	2.1 ±1.9	6.3 ±0.4	93.1
Alfalfa, sweet clover & Inter WG	94.1 ±0.1	8.1 ±0.4	23.5 ±0.1	33.8 ±2.1	11.7 ±0.8	9.3± 0.3	2.6 ±2.0	5.6 ±0.1	96.6
Sunburst SG & BB	94.3 ±0.1	9.8 ±0.7	21.1 ±0.2	30.8 ±0.8	17.2 ±0.7	5.6± 0.7	0.3 ±0.3	7.4 ±0.1	92.0
Sunburst SG & Mustang WR	94.9 ±0.1	9.7 ±0.5	20.9 ±0.4	32.3 ±0.4	15.5 ±1.1	6.7± 2.0	1.7 ±1.0	7.8 ±0.0	94.5

TS- Total Solid; Ext- Extactives; Lig – Lignin; Xyl – Xylan; Ara – Arabinan; Gal – Galactan; WG – Wheatgrass; SG – Switchgrass; WR – Wild rye.

Table A4: Compositional analysis data for Hettinger samples harvested in 2009 and for Carrington and Streeter samples harvested in 2008 showing mean and standard deviation

Sample Name	% TS	% Ext	% Lig	% Glu	% Xyl	% Ara	% Gal	% Ash	Mass closure
Hettinger									
Tall WG	95.2 ±0.0	7.4 ±0.0	25.4 ±1.2	28.5 ±0.7	11.8 ±0.4	5.9 ±0.2	1.2 ±0.1	4.7 ±0.1	84.8
Intermediate WG	95.0 ±0.0	8.9 ±0.4	23.1 ±0.0	32.5 ±0.5	12.5 ±0.4	4.3 ±1.2	1.1 ±1.0	5.3 ±0.0	87.7
Intermediate & Tall WG	95.7 ±0.1	7.7 ±0.2	24.7 ±0.1	28.0 ±1.1	11.1 ±0.6	2.6 ±0.5	1.0 ±0.6	5.1 ±0.1	80.2
Alfalfa, sweet clover & Inter WG	95.6 ±0.1	8.5 ±0.2	24.6 ±0.5	32.1 ±1.4	14.8 ±0.4	3.4 ±0.2	1.6 ±0.0	5.5 ±0.1	90.4
Carrington									
Tall WG	91.9 ±0.1	6.4 ±0.4	22.5 ±0.4	32.3 ±2.1	12.6 ±1.0	3.6 ±0.8	0.8 ±0.1	6.5 ±0.1	84.6
Intermediate WG	94.3 ±0.2	7.0 ±0.1	20.4 ±0.1	40.7 ±0.7	15.7 ±2.1	5.6 ±1.9	0.8 ±0.8	4.9 ±0.0	95.1
Sunburst SG	94.0 ±0.1	10.9 ±0.4	22.1 ±0.2	30.3 ±0.4	11.7 ±0.5	3.6 ±0.7	0.7 ±0.3	4.9 ±0.2	84.1
Sunburst SG & Tall WG	91.3 ±0.1	8.5 ±0.6	24.5 ±0.4	30.7 ±0.1	12.1 ±0.1	5.5 ±0.3	1.0 ±0.1	5.5 ±0.1	87.7
Magnar & Mustang WR	90.8 ±0.3	7.2 ±0.2	25.4 ±0.4	33.4 ±0.8	11.0 ±0.3	4.2 ±2.0	0.8 ±0.3	5.3 ±0.3	87.2
Sunburst SG & BB	91.4 ±0.0	10.2 ±0.4	21.8 ±0.5	29.9 ±1.6	11.2 ±0.5	6.2 ±2.3	0.9 ±0.1	5.5 ±0.0	85.6
Streeter									
Sunburst SG	93.7 ±0.1	9.2 ±0.2	24.1 ±0.1	28.8 ±1.4	11.0 ±0.6	3.9 ±1.8	0.9 ±0.4	7.7 ±0.0	85.7

TS- Total Solid; Ext- Extactives; Lig – Lignin; Xyl – Xylan; Ara – Arabinan; Gal – Galactan; WG – Wheatgrass; SG – Switchgrass; WR – Wild rye.

Table A5: Compositional analysis data for Minot samples harvested in 2008 showing mean and standard deviation

Sample Name	% TS	% Ext	% Lig	% Glu	% Xyl	% Ara	% Gal	% Ash	Mass closure
Tall WG	94.0 ±0.1	11.0 ±0.0	23.4 ±0.1	29.0 ±0.3	11.4 ±1.9	4.0 ±0.7	0.7 ±0.6	7.2 ±0.0	86.7
Intermediate WG	93.5 ±0.1	7.8 ±0.2	24.0 ±0.5	33.9 ±1.5	13.1 ±1.4	4.8 ±1.2	2.7 ±1.2	7.2 ±0.1	93.6
Intermediate & Tall WG	92.8 ±0.3	8.1 ±0.8	24.1 ±0.4	33.1 ±0.3	15.0 ±2.4	4.8 ±0.1	2.7 ±2.6	7.7 ±0.2	95.5
Sunburst SG & Tall WG	92.9 ±0.1	8.8 ±0.1	23.7 ±1.0	31.2 ±0.9	14.6 ±0.5	3.3 ±0.3	1.0 ±0.1	8.0 ±0.1	90.4
Alfalfa Sweet Clover & Intermediate WG	92.8 ±0.1	7.2 ±0.0	24.6 ±1.0	31.5 ±1.3	12.2 ±0.8	4.2 ±0.3	2.1 ±0.4	6.9 ±0.0	88.5
Dakota SG	93.5 ±0.1	13.4 ±0.6	23.2 ±0.0	28.5 ±0.8	12.2 ±0.5	4.6 ±0.4	1.4 ±0.4	8.3 ±0.2	91.6
Dakota SG -Leaf only	92.0 ±0.1	16.0 ±0.8	21.7 ±0.3	21.2 ±0.5	7.4± 0.0	2.9 ±0.1	1.4 ±0.1	11.7 ±0.2	82.3
Dakota SG-Stem only	92.6 ±0.1	7.1 ±0.1	25.1 ±0.1	20.3 ±0.7	11.4 ±0.7	3.0 ±0.1	1.4 ±0.6	6.3 ±0.1	84.5
Sunburst SG	92.8 ±0.1	13.3 ±0.2	22.3 ±0.2	35.8 ±1.0	13.1 ±0.1	4.9 ±0.0	1.6 ±0.1	8.7 ±0.0	99.6
Sunburst SG-Leaf only	93.5 ±0.2	15.4 ±0.8	22.5 ±0.3	24.8 ±0.3	9.6± 0.2	3.4 ±0.1	1.7 ±0.1	9.5 ±0.3	86.8
Sunburst SG-Stem only	93.6 ±0.1	14.3 ±0.4	22.9 ±0.0	31.6 ±1.6	12.6 ±0.8	3.6 ±0.7	1.2 ±0.1	4.8 ±0.2	91.1
Magnar & Mustang WR	93.3 ±0.1	9.9 ±0.2	24.2 ±1.1	27.0 ±2.3	11.0 ±1.3	4.5 ±1.2	1.8 ±1.0	8.3 ±0.2	86.8
Magnar & Mustang WR-Leaf only	93.4 ±0.1	11.7 ±0.4	22.2 ±0.1	27.3 ±0.5	14.3 ±0.4	6.9 ±1.3	1.5 ±0.6	8.9 ±0.1	92.7
Magnar & Mustang WR-tem only	93.6 ±0.2	8.9 ±0.3	23.8 ±0.4	30.4 ±0.1	12.6 ±0.0	5.1 ±0.1	2.6 ±0.1	7.1 ±0.1	90.5
Sunburst SG & BB	93.2 ±0.1	12.1 ±0.7	22.9 ±0.4	28.1 ±1.5	10.8 ±1.5	3.4 ±0.2	1.5 ±0.1	7.5 ±0.1	86.3
Sunburst SG & BB-leaf only	92.2 ±0.2	13.6 ±0.4	21.8 ±0.2	24.8 ±1.1	9.0± 0.3	3.2 ±0.1	1.7 ±0.1	10.4 ±0.0	84.6
Sunburst SG & BB-Stem only	92.9 ±0.2	11.9 ±0.9	24.4 ±1.5	28.2 ±1.2	11.3 ±0.0	2.8 ±0.0	1.0 ±0.1	5.1 ±0.0	84.6
Sunburst SG & Mustang WR	93.1 ±0.0	10.3 ±0.0	24.5 ±0.9	28.3 ±0.5	11.5 ±0.3	3.4 ±0.1	1.3 ±0.7	9.0 ±0.2	88.4
Sunburst SG & Mustang WR-leaf only	92.6 ±0.6	9.7 ±0.3	24.0 ±0.4	28.8 ±0.8	12.3 ±0.3	3.7 ±1.0	1.3 ±0.0	8.8 ±0.0	88.5
Sunburst SG & Mustang WR-stem only	93.4 ±0.1	7.8 ±0.3	23.4 ±0.5	29.5 ±0.9	12.9 ±0.3	2.8 ±0.4	1.0 ±0.1	6.8 ±0.0	84.1

TS- Total Solid; Ext- Extactives; Lig – Lignin; Xyl – Xylan; Ara – Arabinan; Gal – Galactan; WG – Wheatgrass; SG – Switchgrass; WR – Wild rye.

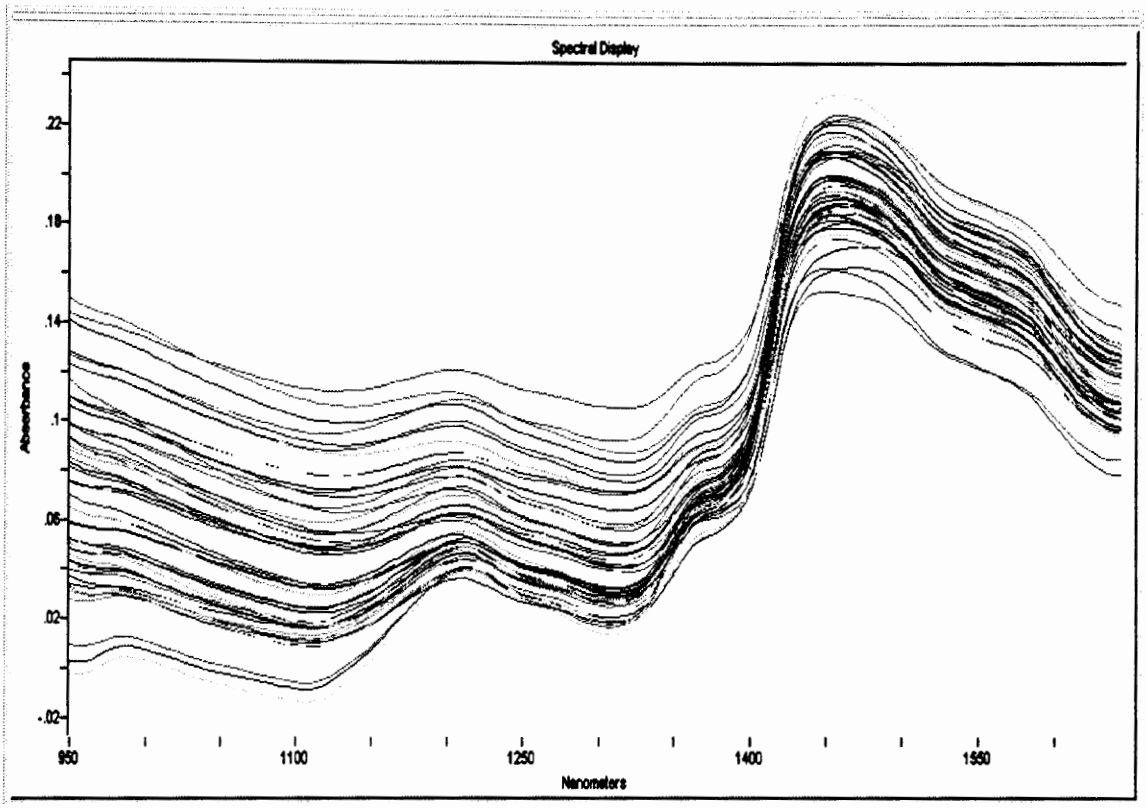


Figure A1: Spectra display of the fifty-five biomass samples used as calibration set.