2009

Development of rapid methods to determine the quality of corn for ethanol production

Allison Palmer Burgers

Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/etd

Part of the Nutrition Commons

Recommended Citation

Burgers, Allison Palmer, "Development of rapid methods to determine the quality of corn for ethanol production" (2009). Graduate Theses and Dissertations. 10722.
http://lib.dr.iastate.edu/etd/10722

This Thesis is brought to you for free and open access by the Graduate College at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Development of rapid methods to determine the quality of corn for ethanol production

by

Allison Burgers

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Food Science and Technology

Program of Study Committee:
Charles R. Hurburgh, Jr., Major Professor
Jay-lin Jane
Thomas Brumm

Iowa State University
Ames, Iowa
2009

Copyright © Allison Burgers, 2009. All rights reserved.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>IV</td>
</tr>
<tr>
<td>GENERAL INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>LITERATURE REVIEW</td>
<td>2</td>
</tr>
<tr>
<td><em>Ethanol Production</em></td>
<td>2</td>
</tr>
<tr>
<td><em>Starch Overview</em></td>
<td>9</td>
</tr>
<tr>
<td><em>Near-Infrared Spectroscopy</em></td>
<td>16</td>
</tr>
<tr>
<td><em>Current Studies</em></td>
<td>25</td>
</tr>
<tr>
<td>THESIS ORGANIZATION</td>
<td>27</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>28</td>
</tr>
<tr>
<td><strong>PART I: PREDICTION OF CORN ETHANOL YIELD BY NEAR-INFRARED SPECTROSCOPY</strong></td>
<td></td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>35</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>36</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>37</td>
</tr>
<tr>
<td><em>Ethanol Yield Measurement</em></td>
<td>37</td>
</tr>
<tr>
<td><em>NIRS Calibration</em></td>
<td>38</td>
</tr>
<tr>
<td>RESULTS AND DISCUSSION</td>
<td>39</td>
</tr>
<tr>
<td>CONCLUSIONS</td>
<td>42</td>
</tr>
<tr>
<td>LITERATURE CITED</td>
<td>42</td>
</tr>
<tr>
<td><strong>PART II: APPLICATION OF ETHANOL YIELD PREDICTION TO CORN SAMPLES</strong></td>
<td></td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>55</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>56</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>57</td>
</tr>
<tr>
<td><em>Ethanol Yield Equation</em></td>
<td>57</td>
</tr>
<tr>
<td><em>Data from Previous Crop Years</em></td>
<td>58</td>
</tr>
</tbody>
</table>
ABSTRACT

As ethanol production greatly increased in recent years in the U.S., there has been interest to make the ethanol production process more efficient and economical, therefore maximizing profits. Measuring the amount of ethanol produced from a lot of corn takes days when done by conventional methods. There is a need to develop a rapid method of determining quality of corn for ethanol production. Near-infrared spectroscopy (NIRS) could be useful in this application. This thesis was intended to develop a rapid method using NIRS to predict ethanol production from corn. A partial least squares and a component calculation equation were developed to predict the ethanol yield of corn samples. It was determined that the component calculation was more accurate in validation and more practical for use by ethanol plants. This method uses a component calculation equation including protein, oil, and density values predicted by near-infrared spectroscopy (NIRS) to predict the ethanol yield in gallons per bushel at 15% moisture. Using this method instead of implementing new NIRS calibrations would save time and money involved with new calibrations. The component calculation equation was applied to Iowa corn data from previous crop years as well as a planting date study data. The equation produced expected results when applied to data from previous crop years with increasing ethanol yields from 2005-2008. When applied to the planting date study, the results showed significant losses in grain yield as well as a loss in grain quality for ethanol production at later planting dates. In summary, the component calculation was shown to be able to accurately and rapidly predict ethanol yield of corn samples.
Environmental and economic concerns have contributed to a recent increase in ethanol production in the U.S. Fuel ethanol production has dramatically increased, from 175 million gallons in 1980 to 9 billion gallons in 2008 (Figure 1) (Renewable Fuels Association, 2009). This increase is due to several proposed benefits of ethanol in the fuel industry including the decrease in gasoline use, the reduction of dependence of foreign oil, the lowering of net carbon dioxide (CO₂) emissions (which improves air quality), and assistance to the economy of the U.S., specifically the rural economy (Dale and Tyner, 2006; Renewable Fuels Association, 2009). The U.S. government as well as some state governments provides a subsidy for renewable fuels production.

**Figure 1.** Historic U.S. ethanol production.  
(Source: Renewable Fuels Association, 2009)
As ethanol production increased, needs for new or improved ethanol production technologies developed. Currently, the measurement of the ethanol yield in a lot of corn can take days. A rapid and accurate prediction of corn quality for ethanol production would be helpful to the ethanol industry. Near-infrared spectroscopy (NIRS) could be a useful technique in predicting the ethanol yield of corn, because it is rapid and non-destructive.

The following research was conducted on the use of NIRS in determining the quality of corn for ethanol production by developing a NIRS method to predict the ethanol yield from corn samples. This method will then be applied to corn samples and the impact on ethanol plant economics and procurement will be examined.

**LITERATURE REVIEW**

**Ethanol Production**

Ethanol from corn is produced by the breakdown of complex starch molecules into the monosaccharide glucose, followed by fermentation of the glucose. Starch is the energy storage unit for plants that is made-up of polymers of glucose units. Starch is present in the endosperm, which is the major structure of corn and other cereal grains. Starch consists of two components, amylose and amylopectin. Amylose is a linear molecule consisting of glucose units connected by alpha–1,4 linkages. Amylopectin is a highly branched molecule consisting of alpha–1,4 and alpha–1,6 linkages. The starch must be broken down into individual glucose monomers to be fermented into ethanol by yeast. The ethanol produced by fermentation is then concentrated by distillation.
In the U.S., corn is the main feedstock for ethanol accounting for 97% of ethanol produced (USDA, 2007). Sorghum, wheat and mixed grains, and processing waste are also used, but on a much smaller scale. The two types of industrial corn to ethanol processes used in the U.S. are wet-mill and dry-grind. Each of these processes produces different amounts of ethanol and coproducts from corn. This research focuses on the dry-grind process because it makes up most of the new plants being constructed today and accounts for greater than 70% of ethanol production in the U.S. (Mosier and Ilelej, 2006).

In the past, ethanol produced from dry-grind facilities was used for beverages and industrial uses. However, since the 1970’s, its use in the fuel industry has become more prominent as oil prices increased (Singh et al, 2001). In the fuel industry, ethanol is used as a gasoline extender, octane enhancer and an oxygenate, which adds oxygen to the gasoline allowing it to burn cleaner. Ethanol is sold as E-5 (5% ethanol), in which the ethanol acts as an oxygenate. It is also marketed as E-10 (10% ethanol), where the ethanol is a fuel extender. Ethanol can also be sold as E-85 (85% ethanol), which can be used in flexible fuel vehicles.

**Wet-milling**

A small amount, less than 30%, of ethanol produced in the U.S. is by the wet-mill process (O’Brien and Woolverton, 2008). Wet-milling is the more complex of the ethanol production processes and is commonly used to produce many more products than simply ethanol. Coproducts produced from wet-milling are more valuable and versatile. The wet-milling process focuses on separating the kernel into starch, germ, fiber, and protein and processing each part of the kernel independently (Rausch et al, 2007). This attempts to remove the maximum amount of starch by first soaking and steeping the kernel in water. Steeping along with mechanical shear removes the high protein germ from the high starch...
endosperm. Major coproducts include corn gluten meal and corn gluten feed. Products from starch include high fructose corn syrup, food additives, and biodegradable plastics (Mosier and Ilelej, 2006).

**Dry-grind**

The dry-grind process is used by smaller plants than wet-milling, for a smaller capital investment. For these reasons there has been a major growth in the dry-grind industry (Butzen et al, 2003). The dry-grind method processes and uses the entire kernel. The products are then separated at the end of the process. Non-fermentable parts of the kernel are carried through the process. The parts of the kernel that do not ferment are the germ, protein, vitamins, minerals, and fiber (Singh et al, 2001). The conventional dry-grind process has an average yield of 2.7-2.8 gallons of ethanol per bushel (56 lbs) (Mosier and Ilelej, 2006). There are five steps to the conventional dry-grind process: milling, liquefaction, saccharification, fermentation, and distillation and recovery (Figure 2).

![Figure 2. Conventional dry-grind ethanol process.](Source: Singh et al, 2001)
1. Milling

Before milling begins, the corn arrives at the plant and is received and inspected by the quality control department. The corn is then cleaned using screens, blowers, and magnets. The corn is now ready for the first step, milling, which involves grinding the corn into a corn flour using a hammermill (most commonly used in ethanol plants) or a rollermill. Kernels are reduced in size until they can pass through a set screen size. The ground corn usually has a mean diameter between 0.8 mm and 1.0 mm. This allows the starch to be exposed for efficient hydrolysis in the following steps (liquefaction and saccharification). Next, the flour is turned into a slurry by adding water. The slurry commonly consists of 20-40% solids (Dale and Tyner, 2006). Water recycled from later steps (backset) is often used to make the slurry.

2. Liquefaction

Next, alpha-amylase, a thermostable enzyme, is added to the slurry. Alpha-amylase is an endoenzyme that randomly cleaves the alpha-1,4 linkages of starch (Dale and Tyner, 2006). Once the slurry is made it is cooked using a jet-cooker which injects steam and heats the slurry to above 100º C. The heat and shear by the jet-cooker gelatinize the starch granules, while the enzyme begins to hydrolyze the starch into oligosaccharides or dextrins, which are short glucose chains.

The slurry is cooled to 80-90º C and more alpha-amylase is added to continue the liquefaction process for another thirty minutes. The corn slurry is now termed corn mash. Each bushel of corn produces approximately 22 gallons of mash (Bothast and Schlicher, 2005).

3. Saccharification
In this step, the oligosaccharides formed during liquefaction are further hydrolyzed to glucose monomers by the enzyme glucoamylase. The corn mash is first cooled to 30º C and then the second enzyme glucoamylase is added. Glucoamylase is an exo-enzyme that cleaves both alpha-1,4 and alpha-1,6 linkages (Dale and Tyner, 2006). This step can take place in a saccharification tank or in the same tank as the fermentation, which is called simultaneous saccharification and fermentation. The simultaneous method is most often used as it lowers both the chance of contamination and osmotic stress on the yeast, while being an overall more energy efficient method (Bothast and Schlicher, 2005).

4. Fermentation

Fermentation is the process of converting the fermentable sugars glucose, maltose, maltotriose, and fructose to alcohol and carbon dioxide using yeast. Yeast and nutrients are added to the corn mash to begin fermentation. The yeast species *Saccharomyces cerevisiae* is often used for fermentation because of its efficient alcohol production and its ability to endure high alcoholic environments and osmotic stress (Butzen et al, 2003). *S. cerevisiae* ferments best at a temperature around 33º C and a pH of 4.0 (U.S. Grains Council, 2007). Higher temperatures will decrease the efficiency of the yeast. CO₂ and ethanol are produced on an almost equal basis. The reaction formula for this fermentation is shown below.

\[
\text{C}_6\text{H}_{12}\text{O}_6 + \text{H}_2\text{O} + \text{Yeast} \rightarrow 2\text{CO}_2 + 2\text{C}_2\text{H}_5\text{OH} + \text{H}_2\text{O} + \text{Heat}
\]

100 lb 48.8 lb 51.1 lb 17,000 BTU

(Singh et al, 2001)

Theoretically, 100 pounds of glucose will produce 51.1 pounds of ethanol and 48.8 pounds (lbs) of CO₂ (Singh et al, 2001). This uses the stoichiometric ratio of 1.047 g of
ethanol (molecular weight of 46.069) per 1.000 g of CO₂ (molecular weight of 44.010). A bushel (bu) of corn at 15% moisture equals 56.0 lbs or 47.6 lbs/bu dry matter (dm) (Murphy, 1993). Starch and simple sugars on average make up 74.6% dm of the corn kernel which equals 35.5 lbs per bushel of corn dry matter (Watson, 1987). At the rate fermentation rate described above (0.51), 18.1 lbs of ethanol is produced per bu of corn on a dry moisture basis. This will produce 17.3 lbs CO₂/bu dm corn. The remaining part of the kernel will produce approximately 17 lbs/bu of the coproduct distillers dried grains with solubles (DDGS) at a later step (RFA, 2009).

Fermentation can either be done by the continuous or batch method. Batch fermentation involves a whole tank being allowed to ferment before it is emptied. Whole tank fermentation usually takes around 52 hours (Dale and Tyner, 2006). Most dry-grind set-ups will have at least three batch fermentation tanks to keep the entire production continuous. Twenty percent of plants use continuous fermentation which cycles the mash through the fermentation vessels (Dale and Tyner, 2006).

The CO₂ produced during fermentation can be released or cleaned, compressed and sold for use in soft drinks industry or dry ice production. CO₂ can also be used in the flash freezing of meat, paper mills or other industrial uses (American Coalition for Ethanol, 2008).

Ethanol yield can be inhibited by bacterial contamination or wild yeast contamination, which competes with yeast for the nutrients and often produces lactic acid and acetic acid products. Lactic and acetic acid production will act to inhibit the yeast activity and growth. Antimicrobials and sanitation are used to prevent bacterial contamination.
Upon fermentation, the corn mash is called beer. The fermentation tank is then emptied into the beer well or holding tank. The beer well stores the beer and supplies the final steps with a continuous flow of beer.

5. Distillation and recovery

The beer at this point is 8-12% ethanol and the remaining consists of water and solid material. The distillation process allows the ethanol to be removed through a sequence of vaporizations and condensations. The distillation step usually contains two distillation columns and a stripping column. This process takes advantage of the fact that ethanol boils at 78\(^\circ\) C while water boils at 100\(^\circ\) C. The ethanol is boiled off to separate it from the non-fermentable parts of the corn. The ethanol leaves through the top of the distillation column as vapor and enters the condenser where it becomes a liquid. Water and the non-fermentables (protein, oil, fibers, and residual chemicals) exit through the bottom of the column.

Distillation produces 95% pure ethanol, which is additionally purified using molecular sieves. These sieves are made of microporous material, which separates the smaller water molecules from the larger ethanol molecules. The sieves absorb the remaining water and produce 99% pure ethanol (Kwiatkowski et al, 2006). The purified ethanol is then denatured by adding 5% unleaded gasoline which produces fuel grade ethanol and makes the ethanol undrinkable to humans. The addition of the gasoline avoids a large tax from the Bureau of Alcohol, Tobacco, and Firearms. The denatured ethanol is then held in a tank until it is transported to be used in the fuel industry.

After the removal of ethanol in the distillation step, water and non-fermentable solids remain. These remains are called whole stillage, which is then separated by centrifugation into the solid part called wet distillers’ grain and the liquid part called thin stillage. Around
30-50% of the thin stillage is used as backset (Rausch et al, 2007). The backset can be added to the ground corn at the beginning of the process or to the output of the liquefaction step as it provides nutrients essential to the yeast in the fermentation step.

The remaining thin stillage is evaporated into a syrup called distillers solubles. Water is recovered during this step and is recycled to the first step where it is mixed with the flour. The solubles are added to the wet distillers grain and dried to produce distillers dried grain with solubles (DDGS) which is used as animal feed. DDGS is a major coproduct of the dry-grind process. If this mixture is not dried it is known as wet distillers grains with solubles (WDGS). WDGS can be sold and used as animal feed, but has a much shorter shelf-life than DDGS (Butzen et al, 2003).

**Starch Overview**

Starch is formed in plastids and amyloplasts of plants during the process of photosynthesis in which sunlight is converted to chemical energy. The main function of starch in nature is as an energy storage unit in plants. Starch is made up of α-D-glucose molecules linked by alpha-1,4 and alpha-1,6 glycosidic bonds. The main components of starch are amylose and amylopectin, lipids and phosphates are also present in smaller portions.

Amylose molecules are primarily linear and consist of 500-20,000 glucose units connected by the 1-4 bonds (Figure 3) (Chaplin, 2008). Amylopectin has a highly branched structure which is linked by both the 1-4 and 1-6 bonds (seen in Figure 4). In amylopectin, one out of every twenty bonds is a 1-6 linked branch.
Normal starch is made up of 20-30% amylose and 70-80% amylopectin. This can vary depending on the type and hybrid of grain (Chaplin, 2008). Grains containing little or no amylose are given the name waxy (Bertoft, 2004).

Starch is densely packed into insoluble spherical granules which make the structure perfect for storage. Starch granules size and shape depends on the plant source. Corn starch granules are approximately 25 µm in diameter. Granules are semi-crystalline in nature. This semi-crystalline portion is associated with the amylopectin of the granule, while the amylose region is amorphous (Singh et al, 2003). Figure 5 shows how the amylose and amylopectin
are specially arranged within the granule. Figure 6 shows the appearance of corn starch granules by a scanning electron micrograph.

**Figure 5.** Conceptual model of starch granule showing amylose and amylopectin regions. (Source: Robertson et al, 2006)

**Figure 6.** Scanning electron micrograph (SEM) of corn starch granules. (Source: Singh et al, 2003)
Most people recognize starch as a source of food as it usually makes up a large portion of humans daily consumed calories. However, starch is also an important industrial product as it is flexible and can be used in a variety of ways. Starch for industrial uses mainly comes from the following plant sources: corn, wheat, tapioca, and potato (Van der Maarel et al, 2002). Industries which utilize starch include: pharmaceutical, food and beverage, cosmetics, paper, detergent, plastics, medical, textiles, mining, fuel, and building (Ellis et al, 1998; IENICA, 2003).

Starch is converted enzymatically, chemically, and physically to produce products for industrial use. Syrups, maltodextrins, cyclodextrins, organic acids, biopolymers, and alcohols (ethanol) are examples of converted starch products (Ellis et al, 1998). Ethanol has increasingly become one of the most important industrial starch products, as the need for a biorenewable fuel source has recently increased.

**Starch Hydrolysis**

The natural structure of starch must be destroyed for the industrial production of ethanol. The glycosidic bonds of starch must be taken apart, or hydrolyzed, in order for the yeast to be able to ferment the glucose monomers into ethanol. This disassembly of starch was first done by dilute acid hydrolysis. The glycosidic bonds hydrolyze at the low acidic pH which is present in the acid hydrolysis process. Acid hydrolysis results in unnecessary additional reactions beyond hydrolysis (Nigam and Singh, 1995).

The ethanol industry has since moved to using enzymatic hydrolysis, as the acid caused problems with the equipment, yield, and cost of the process (Robertson et al, 2006). Enzymatic hydrolysis is a less harsh method of starch degradation in which the enzymes alpha-amylase and glucoamylase are used for enzymatic hydrolysis. Enzymatic hydrolysis
conventionally involves a high temperature cooking step called liquefaction. This cooking step requires a large energy input, as the mash is heated to high temperatures between 140º C and 180º C (Matsumoto et al, 1981).

**Starch Measurement Methods**

Measuring the starch content of corn allows the ethanol plant to know how much ethanol can be produced from a specific lot of corn. Conventionally, starch measurement has been done by time consuming laboratory procedures. Two types of starch measurement methods are important to the ethanol industry: extractable starch measurement and fermentable starch measurement. Measuring total starch content has been proven not predictive of ethanol yield (Butzen et al, 2003). This could be due to amount of resistant starch in the corn that does not ferment to ethanol. Resistant starch is resistant to enzymatic degradation. Resistant starch is often used in food products as a functional food for its health benefits. However, it is not beneficial for industrial ethanol production as it inhibits the enzymatic hydrolysis and complete fermentation of starch to alcohol. In the fermentation of corn starch to ethanol, resistant starch includes starch that is entrapped in the protein matrix, maltodextrins, starch-lipid complex, and retrograded starch.

Resistant starch content can be determined indirectly by the difference between the total starch content and the digestible (non-resistant) starch content of the sample (Englyst, 1989). Other direct methods involve quantifying the amount of residue after enzymatic hydrolysis and removal of digestible starch (Berry, 1986; Goni et al, 1996). The AOAC Method 2002.02 and AACC Method 32-40 both use a direct method to measure resistant
starch. Results from resistant starch measurements will vary depending on the method (direct or indirect) used.

Extractable starch measures what is available from the wet-milling process. Extractable starch is often determined as the starch recovered from the 100g wet-milling laboratory process. This method is not an accurate measurement of the amount of starch available from the dry-grind process. Studies have shown that there is not a significant correlation between extractable starch and ethanol yield by the dry-grind method (Singh and Graeber, 2005). Extractable starch contents are lower than total starch contents and are often influenced by genotype, drying treatment, and environmental growing conditions (Rathore et al, 2006).

The 100g wet-milling procedure was developed by Eckhoff, et al. (1996) to reduce sample size and labor required to determine the wet-milling characteristics of corn. This method starts by steeping and grinding the corn sample. Then the course and fine fiber is removed using a sieve. The slurry is pumped onto a starch table which removes the protein and leaves wet starch. The table is dried and the starch is weighed, from this the total extractable starch dry weight is calculated. Protein, germ, and fiber contents can also be calculated by this method (Eckhoff et al, 1996).

Measuring fermentable starch best indicates the amount of ethanol produced by the dry-grind method (Bothast and Schlicher, 2005). The laboratory procedure for measuring fermentable starch is very similar to the traditional dry-grind procedure previously described. The conventional dry-grind corn laboratory fermentation procedure uses 150g sample of ground corn (Rathore et al, 2005). Water is added to create a 25% solids mash. The sample is then liquefied using alpha-amylase and heat. The temperature is lowered, pH adjusted, and
glucoamylase is added to begin the saccharification process. The mash is then cooled and yeast is added to begin fermentation, which is carried out for 72 hours at 30º C. After fermentation, a 5 mL sample is filtered and centrifuged for high performance liquid chromatography (HPLC) analysis which measures the amount of ethanol produced. Laboratory fermentation allows the corn substrate to be the limiting factor in fermentation by supplying other factors in excess.

Other laboratory methods of determining fermentable starch contents involve calculating the loss in the weight of the sample from the beginning of the fermentation to the end instead of using HPLC analysis. The loss of weight is then converted to grams of ethanol, as the CO₂ produced is proportional to the amount of ethanol produced in nearly equal amounts (stoichiometric ratio of 1.047 ethanol: 1.000 CO₂) (ICIA, 2008). This is presented as ethanol yield in gallons of 200 proof undenatured ethanol per bushel of corn, at 15% moisture. This method is used by the Illinois Crop Improvement Association, Inc. (ICIA) laboratory (ICIA, 2008).

**Highly Fermentable Hybrids**

High extractable starch hybrids for wet-milling and high fermentable starch hybrids for dry-grind have been developed especially for the ethanol industry. These two types of hybrids are similar, but used in different industries. Both classes are bred for their increased ability to ferment into ethanol, one by direct fermentation and one by wet separation before fermentation.

Several seed companies (Pioneer Hi-Bred, Monsanto, and Syngenta) are interested in identifying and marketing hybrids with enhanced ethanol production. Hybrids titled High Total Fermentables (HTF), by the company Pioneer Hi-Bred, indicate above average ethanol
yields for the dry-grind process (Butzen and Haefele, 2008). HTF yields 4% more ethanol than other hybrids (Bothast and Schlicher, 2005). Using HTF hybrids could increase the profitability of a 40 million gallon per year (MGY) plant profitability by one to two million dollars per year (Butzen et al, 2003). Monsanto has a list of hybrids for dry grind ethanol production named “Processor Preferred Fermentable Corn” also known as HFC. According to Monsanto, using HFC hybrids increased ethanol yield by 2.7% (Bothast, 2005). Syngenta has labeled their highly fermentable hybrids “NK Brand Extra Edge”, which have been on the market since 2004 and are specially marketed for use in the dry-grind industry (Bothast, 2005).

Near-Infrared Spectroscopy

There is a need for a rapid method for determining the quality of corn for dry-grind ethanol production. NIRS is able to measure organic substances in seconds instead of the hours it takes other chemical laboratory methods. NIRS is an accurate, rapid, nondestructive, and relatively low cost technology that could be used to determine the amount of fermentable starch and therefore the quality of a lot of corn for ethanol production.

Using NIRS to scan corn at the receiving end of the ethanol plant will allow the plant to identify high and low ethanol yielding lots of corn. The plant can than discourage accepting lower yielding lots, offer premiums for high yielding lots, or rank and segregate corn for ethanol production, therefore increasing the plants overall efficiency, ethanol yield, and profit.

NIRS Theory
NIRS is used for qualitative and quantitative analysis in a variety of industries including: pharmaceuticals, agriculture, food, medical science, chemical science, fuel, environmental, polymers, and the textile industry. The following research will focus on the use of NIRS in the grain industry specifically for the measurement of corn. NIRS is commonly used to analyze corn samples to predict their moisture, density, protein, oil, and starch contents (Illinois Crop Improvement Association, 2008; Orman and Schumann, 1991).

NIRS is an analytical absorption technique based on Beer’s law. Beer’s law states that there is a linear relationship between concentration and absorbance (Siesler, 2002). Beer’s law: \( A = \log \left( \frac{1}{T_\lambda} \right) \), where \( T \) equals transmittance, \( \lambda \) is wavelength, and \( A \) equals absorbance.

**Figure 7.** Electromagnetic spectrum.
(Source: Kovalenko, 2006)

NIRS utilizes the electromagnetic spectrum in the near-infrared (NIR) range 700-2500 nm (14000-4000 cm\(^{-1}\)) (Figure 7) (Kovalenko, 2006). In NIRS, samples are irradiated with NIR light. The amount of NIR energy absorbed by a sample depends on the sample composition (Kovalenko, 2006). The molecules of the sample vibrate according to their
vibrational energy levels once the NIR light is absorbed. Harmonic systems have equal distances between energy levels while anharmonic systems, which pertain to actual molecules, have unequal distances between levels, these systems (Figure 8) (Pasquini, 2003).

![Figure 8. Harmonic (A) and anharmonic (B) models of vibrational energy levels. (Source: Pasquini, 2003)](image)

The NIR range of the spectrum includes overtones of OH, NH, and CH vibrations. Combinations modes and stretching/bending vibrations are present in the NIR region. Overtone transitions, unequal distances between energy levels, along with the combination modes and stretching/bending vibrations all contribute to the broad, overlapping, and combination absorption bands in this region of the spectrum. These factors make NIR spectral data difficult to interpret by simple methods.

There are many advantages to using NIRS. Little to no sample preparation needed. Also, Analysis is completed in seconds, when conventional laboratory methods of analysis can take hours and days to complete. In addition to saving time, NIRS can also save money that is involved in the extensive chemical laboratory procedure. Samples can be repeatedly scanned, as the NIR radiation does not alter the sample. Limitations of NIRS include large
initial expense of NIRS machine, updating calibrations, and the need for a skilled chemometrician to develop and maintain calibrations.

**NIR Spectrometers**

The two types of spectrometers used to analyze absorbance values are transmittance and reflectance. Transmittance spectrometers measure the amount of light that passes through a sample. Transmittance spectrometry uses the function $A = \log \left( \frac{1}{T} \right)$. While reflectance spectrometers measure how much light is reflected or bounces off of the sample and use the function $A = \log \left( \frac{1}{R} \right)$ (Kovalenko, 2006). Where ‘A’ represents absorbance, ‘T’ represents transmittance, and ‘R’ represents reflectance.

Spectrometers commonly consist of five parts: light source, wavelength selector, sample compartment, detector, and a signal processor. Figure 9 shows a simple configuration of typical spectrometers (both reflectance and transmittance shown). Tungsten halogen lamps are commonly used as light sources for NIR machines. These lamps are used due to their long lifetime, inexpensive price, and they meet NIR requirements (Siesler, 2002).

![Figure 9. Configuration of a spectrometer. (Source: Reich, 2005)](image-url)
The wavelength selector can either be a filter or a monochromator, simple or complex. The purpose of a monochromator is to spread out the NIR radiation by wavelength and to vary the wavelength over the range by scanning (McClure, 2001; Burns, 2005). Figure 10 shows two different types of monochromators. Monochromators commonly consist of an entrance slit, lens or mirror, dispersing element (prism or reflective grating), focusing element (mirror), and an exit slit (Burns, 2005).

![Figure 10. Two types of monochromators. (Source: Burns, 2005)](image)

The sample compartment is where the sample is held during scanning. It is located between the monochromator and the detector. The detector is a photosensitive material which produces a signal when light meets it. Detectors are made of silicon, PbS, or indium gallium arsenide (InGaAs) depending on the NIR region (Pasquini, 2003). The signal processor is a computer which transforms the electric signal into a digital signal. Chemometrics and statistical computer programs are then used to interpret the digital data.

**Chemometrics**

Chemometrics is the application of mathematical and statistical procedures used to analyze and interpret data. NIR spectra are difficult to interpret due to complicated...
overlapping bands. Chemometrics is needed to interpret the complicated spectral data. Reference samples with known chemistry from laboratory measurements are needed for calibration model creation along with spectra from the samples.

Calibration is the building of a model to relate NIR spectral data to the reference data. This process involves developing a calibration equation using chemometrics. The equation developed is then used to predict the amount of a certain constituent in the samples. There are four main steps in developing a calibration model. The first step involves choosing samples to include in the calibration set. The calibration set must include samples that are representative of the entire population including future samples and include a large range of the constituent of interest. The calibration set samples are scanned on the NIR spectrometer to obtain the spectra. The spectral data can then be preprocessed to remove noise. The next step is to obtain reference data on samples in the calibration set. Chemical or other laboratory methods are used to determine how much of the constituent is in the samples. It is important to have an accurate reference method, as the calibration model is based on this. Calibration model creation is the next step. Model creation uses multivariate methods to relate spectral data to the reference data. The final step is model validation with independent samples not include in the calibration set. Validation is used to determine how well the model predicts values of new samples.

Spectra Pretreatment

Spectral data prepossessing is often needed to remove noise and other unnecessary data caused by light scattering and other interactions. The following will briefly describe some common data pretreatments as this is not a focus of the research conducted.
Multiplicative scatter correction (MSC) and standard normal variant (SNV) are two similar and commonly used spectral pretreatment methods. MSC is a preprocessing method commonly used to correct light scattering effects. SNV like MSC corrects light scattering while centering the spectra (Kovalenko, 2006).

Smoothing is a technique used to reduce noise and is often combined with other preprocessing methods. First and second derivatives are preprocessing methods that are used following smoothing to correct overlapping peaks and enhance relevant peaks. A spectrum that has been treated using the first derivative method is the slope of the raw untreated spectrum. The second derivative spectrum is then the slope of the first derivative. Derivatives are combined with smoothing because their use increases the noise in the spectrum.

**Principal Component Analysis**

Principal component analysis (PCA) is a data analysis method which is used to identify data patterns, clustering and outliers and reduce the number of variables without losing valuable information (Smith, 2002). This method places spectral data into orthogonal components named principal components (PC’s). The first PC (PC₁) represents the most variation of all the linear combinations, then PC₂ represents the second greatest variation, and this continues for all of the PC’s (Reich, 2005). The outliers need to be removed to create a good calibration model. It is important not to remove too many samples that represent a significant amount of the variability. Mean centering and autoscaling preprocessing are required before PCA can be conducted.

The number of principal components (PC’s) used to create a calibration model is important to the precision of the model. When the number of PC’s is too small, some important data which is needed to predict may be missing. When this happens it is called
underfitting the model. The model can be overfit by using too many PC’s and including too much noisy, unnecessary data (Xixiong et al, 2007). Software is often used to cross validate the model and choose the best number of components.

**Calibration Development**

A calibration model is developed using the calibration set, reference data, and multivariate methods to relate the spectral data to the reference data. Once created the model is used to predict values of a certain constituent (protein, oil, starch) in the samples. Common linear methods for calibration model development include multiple linear regression (MLR), principle component regression (PCR), and partial least squares (PLS). Artificial neural networks (ANNs) is a nonlinear method, used for calibration model development. The following research applies MLR and PLS methods for calibration model development.

**MLR**

MLR is a purely older linear calibration method. MLR uses the following equation: \( y = b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n \). Where \( b_0 \) is the \( y \)-intercept and \( b_1 \) to \( b_n \) are the regression coefficients for \( n \) wavelengths, \( x \) is equal to log \( 1/R \). For development of a component calibration, as in the following research, \( x \) is equal to the corresponding component (protein, oil, starch or density). The regression coefficients are found when \( y \) is the dependent variable and \( x \) is the independent variable. In MLR calibration, the original wavelengths are used as variables. MLR provides good models but is sometimes not as good at predicting new samples.

**PLS**

PLS is a linear calibration method which is useful for a high number of collinear variables. The goal is to extract latent variables (components) that represent as much variation as possible. The components are then used as variables in linear regression (Tobias, 1997). PLS components are found by maximizing the covariance between \( y \) and functions of
Three to seven components usually account for 99% of the variation (Pirouz, 2006). The PLS provides good models while using a lower number of PC’s than PCR, for this reason PLS is often used over PCR for NIRS calibration (Pasquini, 2003).

**Validation**

Validation of the model is conducted to show how well the model predicts values of new samples. Once the calibration equation is determined it is used in the validation of the model. Validation includes predicting the amount of the constituent in a sample from the validation set and then comparing the value to the reference value. The differences between the predicted and the reference values are calculated and validation statistics are used to evaluate the predictability of the model.

Validation can be done by using an external validation set or cross validation. An external validation set is made up of new samples that were not included in the calibration set. These new samples are used to predict the amount of the constituent from the calibration equation created earlier. Cross validation uses samples that are in the calibration set. A subset of samples is left out during model creation and then the left out samples are predicted using the calibration equation. The model uses one subset to predict the one that is left out. Full cross validation or leave-one-out cross validation leaves only one sample out of the model and then predicts it. This continues until all samples have been left out once (used as the validation set) and predicted, the results are then averaged and validation statistics can be calculated. Igne et al (2007) found with triticale that cross validation is best used at the time of calibration but new samples should be used to validate the model when available.

Validation statistics are used to evaluate how well the model predicts, some commonly used statistics will be defined. The coefficient of determination ($R^2$) is the square
of the regression coefficient \((R)\). \(R^2\) shows the proportion of the variability that is accounted for in the model. \(R^2\) ranges from 0 to 1, where a \(R^2\) of close to 1 indicates a robust calibration model. The standard error of prediction (SEP or SECV when using cross validation) shows the precision of the predictions. The SEP is the standard deviation of the differences between the predicted and the reference values. The bias is the mean difference between the predicted and the reference data. The bias shows the accuracy of the model. The relative performance determinant (RPD) is also used as a model evaluation statistic. The RPD is equal to the standard deviation of the reference data divided by the SEP.

**Current Studies**

NIRS has been used to predict the outcomes of both wet-mill and dry-grind processes. This research and the technology developed from it are important as it enables the ethanol industry to predict the yield from processes and rank and select the most profitable lots of corn to use for ethanol production.

Pioneer Hi-Bred International Inc. along with FOSS (FOSS Group, www.foss.dk) introduced a direct (from spectra) NIRS calibration method which predicts the amount of ethanol yield from a corn sample (Bryan, 2003). They plan to use this to identify their HTF hybrids or high-yielding corn at ethanol plants. They offer this calibration to the ethanol industry and are developing licensing terms. This method applies only to the FOSS Infratec 12xx series of near infrared transmission analyzers (Pioneer Hi-Bred, 2008).

Monsanto began the initiative “Fuel your profits”, to increase profits of ethanol plants and growers. Monsanto provides participating ethanol plants with a proprietary NIRS calibration tool, developed under ISO 17025 compliance, to allow managers to identify
highly fermentable hybrids (Bernick, 2004; Anderson, 2003). Both Pioneer’s and Monsanto’s work on their NIRS calibrations have not been published.

A wet-mill study focused on the prediction of extractable starch yield using NIRSystem 6500 reflectance spectrophotometer. This study used the 100g laboratory procedure (Eckhoff et al, 1996) as a reference method for starch yield. The calibration had a R^2 of 0.77 (Paulsen et al, 2003). They determined that the calibration would be useful in separating high and low starch yielding lots of corn. A similar study was conducted in 2004 using the Infratec 1229 NIT spectrometer to predict extractable starch yield (Paulsen and Singh, 2004). A calibration for extractable starch was developed having a R^2 of 0.79 and a standard error of prediction (SEP) of 1.24.

A study was conducted by Rathore et al (2006), to develop a calibration for ethanol yield using a fourier-transform near-infrared (FT-NIR) spectrometer. They developed models using PCA, PLS, and discriminate PLS. Their reference method used was a conventional dry-grind fermentation procedure which involved a HPLC test of the fermented corn samples. The discriminate PLS model had the best R^2 of 0.82. They concluded that more research was needed in this area, to determine what properties make corn highly fermentable.

Rathore, et al. (2005), determined the products of fermentation (ethanol, total soluble sugars, glycerol, and organic acids) from the dry-grind process. The study used a 150g corn dry-grind fermentation laboratory procedure including simultaneous saccharification and fermentation as the reference method. The research used NIRS at the end of the process as a process quality control, as opposed to screening the feedstocks at the beginning of the process. The authors successfully identified wavelength regions to predict ethanol and total
soluble sugars but not the other components. This research would be used as an on-line quality control measure for the fermentation part of the ethanol production process.

It is clear after studying the previous, somewhat conflicting, work on this topic, that more research was needed. The genetics company calibrations are proprietary, developed on the respective company databases. There is still a need for a universally accurate method for the rapid determination of the quality of corn for ethanol production.

**THESIS ORGANIZATION**

This thesis is organized into three sections. The first being the general introduction and literature review covering ethanol production, a starch overview, starch measurement methods, near-infrared spectroscopy, and current studies about NIRS and ethanol production. The next section is part one of the research titled, “Prediction of corn ethanol yield by near-infrared spectroscopy”. The first part of the research involves the initial development of a method to rapidly predict the ethanol yield of corn samples to be used for ethanol production. The second part of the research is titled, “Application of ethanol yield prediction to corn samples.” Part two of the research involves the application of the method developed in part one to corn samples from studies conducted at Iowa State University. The results from this research are prepared for publication by the American Association of Cereal Chemist (AACC) in *Cereal Chemistry*. 
REFERENCES


calibration methods for the prediction of protein, oil, and starch in maize grain. *Journal 


Paulsen, M. R., Pordesimo, L. O., Singh, M., Mbuvi, S. W., & Ye, B. 2003. Maize starch 


Pioneer Hi-Bred. 2008. Near infrared technology application for ethanol yield potential in 
corn grain. Pioneer Hi-Bred, a DuPont Company. Available at 

Pirouz, D. M. 2006. An overview of partial least squares. The Paul Merage School of 
Accessed 06/15/2008.

interpretation of fermentation in a dry grind corn process. St. Joseph, MI: American 
Society for Agricultural Engineers (ASAE). ASAE Paper Number: 053041

in dry-grind corn processing using fourier-transform near-infrared spectroscopy. 
American Society of Agricultural and Biological Engineers. Paper No. 063047. UILU 
No. 2006-7001.


PART I: PREDICTION OF CORN ETHANOL YIELD BY NEAR-INFRARED SPECTROSCOPY

A paper to be submitted to Cereal Chemistry

Allison Burgers¹ and Charles R. Hurburgh, Jr.²

ABSTRACT

A rapid method to predict ethanol yield from a lot of input corn is important to ethanol plants. Near-infrared spectroscopy (NIRS) was applied for this purpose. Near-infrared spectra from the Infratec 1229 Grain Analyzer and ethanol yield measurements were obtained from 249 calibration and 80 validation corn samples. A NIRS calibration relating ethanol yield to NIRS was developed. A component calculation using multiple linear regression from current NIRS measurements (protein, oil, starch and density) was also used to predict ethanol yield. The final calibration included 283 samples and had a coefficient of determination \(R^2\) of 0.88 and a standard error of cross validation (SECV) of 0.031 gallons/bushel (gal/bu). The final component equation included 293 samples and had an \(R^2\) of 0.74 and SECV of 0.042 gal/bu. The component calculation was more stable in external validation, and therefore will be simpler to maintain.

¹ Graduate Student, Department of Food Science and Human Nutrition, Iowa State University, Ames, IA 50011
² Major Professor, Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, IA 50011
INTRODUCTION

Measuring the starch content of corn allows the ethanol plant to predict the amount of ethanol that can be produced from a specific lot of corn. Development of rapid starch measurement methods is important in maximizing the ethanol yield from corn. The traditional measurement of starch content has been done through laboratory procedures by measuring extractable starch (Eckhoff et al, 1996), fermentable starch (Rathore et al, 2005), and total starch contents (McCleary et al, 1997). Extractable starch measurement is suitable for the wet-mill industry but does not reflect the ethanol yield of a dry-grind process (Singh and Graeber, 2005). Measuring total starch content has also been found to not be predictive of ethanol yield because not all starch is converted to ethanol during the production process (Zhao et al, 2008).

Dry-grind ethanol production makes up 70% of ethanol produced in the U.S. (Mosier and Ilelej, 2006). In 2008, Iowa currently had 35 dry-grind plants producing 3.8 billion gallons of ethanol (Iowa Corn Promotion Board, 2008). Fermentable starch indicates the amount of ethanol produced by the dry-grind method (Bothast and Schlicher, 2005). The laboratory procedure for measuring fermentable starch is very similar to the traditional dry-grind ethanol procedure. It uses 150g sample of ground corn fermented for 72 hours (Rathore et al, 2005). After fermentation, HPLC is used to measure the concentration of ethanol produced. An alternative method measures the amount of ethanol produced by determining the loss of weight in the sample from beginning to end of fermentation because the CO₂ produced relates to the amount of ethanol produced in nearly equal amounts (Illinois
Crop Improvement Association, 2008). This method also takes 60-80 hours of fermentation time.

Near-infrared spectroscopy (NIRS) is a rapid nondestructive technique that is able to measure organic substances in seconds instead of hours or days. This is an analytical technique that utilizes the near-infrared region of the spectrum (700-2500 nm). NIRS is commonly used in the grain industry to simultaneously predict amounts of moisture, protein, oil, and starch and density (Illinois Crop Improvement Association, 2008; Orman and Schumann, 1991).

Paulsen et al, (2003) developed an NIRS method for determining the extractable starch content in corn. Developing a method to rapidly predict fermentable starch content in corn using NIRS would be helpful to the dry-grind ethanol industry.

The objective of this study was to develop a widely applicable NIRS ethanol yield calibration based on the Illinois Crop Improvement Association (ICIA) method, and then compare the ethanol yield calibration to a regression against NIRS predicted values of protein, oil, starch, and density.

MATERIALS AND METHODS

Ethanol Yield Measurement

The ethanol fermentation test was done by the ICIA laboratory. The laboratory fermentation test uses a typical dry-grind process with simultaneous saccharification and fermentation, which takes 64 hours. Ethanol yield is calculated by the loss in weight of the sample from the beginning of the fermentation to the end. The loss of weight is then
converted to grams of ethanol produced. This uses the assumption that the CO₂ formed is proportional (stoichiometric ratio of 1.047 g ethanol: 1.000 g CO₂) to the amount of ethanol produced (Illinois Crop Improvement Association, 2008) as the chemical equation for fermentation is:

\[ C_6H_{12}O_6 \text{ (Glucose)} + H_2O \text{ (Water)} + \text{Yeast} \rightarrow 2 \text{ CO}_2 + 2 \text{ C}_2\text{H}_5\text{OH} \text{ (Ethanol)} + H_2O \]

(Singh et al, 2001). Ethanol yield is presented as gallons of 200 proof undenatured ethanol per bushel of corn (56 lb), at 15% moisture.

The initial calibration set included 249 corn samples which were analyzed for ethanol yield by the ICIA method. These samples represented the germplasm population of a major corn seed supplier and are typical of Midwest corn. Table 1 shows the NIRS predicted values of protein, oil, starch, and density at 15% moisture of these samples. They were quite typical of average dent corn.

**NIRS Calibration**

NIR spectra ranging from 850 nm to 1048 nm in 2 nm increments was obtained using the Infratec 1229 Grain Analyzer (FOSS Group, www.foss.dk). A partial least squares (PLS) calibration model was developed using the results from the fermentation test as the reference data. The initial model included 237 corn samples after removing samples with noisy spectra. A component calculation using multiple linear regression (MLR) analysis was done including progressive combinations (Table 2) of protein (%), oil (%), starch (%), and density (grams/cubic centimeter(g/cc)) at 15% moisture (predicted by Iowa State University Grain Quality Laboratory calibrations, Table 3) against ethanol yield in gallons/bushel (gal/bu) (1 bu = 56 lbs at 15% moisture). All calculations were done with The Unscrambler 9.6 (CAMO Inc., www.camo.com).
The first validation was done using leave-one-out cross validation. Statistics included the coefficient of determination ($R^2$), standard error of calibration (SEC) for the validation of the constituent model, standard error of prediction (SEP) for the independent validation sets and the standard error of cross validation (SECV) was used for cross validation of the NIRS calibration. The SEP (or SECV when using cross validation), is the standard deviation of the differences between the reference values and the predicted values and describes the precision of the predictions.

Independent validation was conducted using 80 wide ranging samples external to the calibration set. Validation was done in two groups; 55 samples which are normally grown in Iowa from the 2002-2007 crop years, then 15 inbred lines and 10 samples from the 2008 crop year. The inbred lines were specialty samples that included high oil samples which would usually not be used by the ethanol industry.

The 80 validation samples were evaluated for ethanol yield at the ICIA laboratory. Ten repeat samples were included to establish the repeatability of the ICIA laboratory procedure. The validation samples were then pooled with the original calibration samples to develop a final calibration and final component equation.

RESULTS AND DISCUSSION

The mean ethanol yield of the initial 249 corn samples was 2.74 gal/bu with a standard deviation of 0.069 gal/bu, and a range of 2.55 to 2.89 gal/bu (Figure 1). The PLS calibration model for fermentable starch gave a SECV of 0.025 gal/bu and $R^2$ of 0.86 (Figure 2). The PLS results were compared with the component calculation results of combinations
of protein, oil, starch, and density. The results for the best eight combinations (Table 4) were not statistically different. The combination protein, oil, and density was chosen for the prediction equation because these components have the most reliable calibrations, as opposed to total starch. The reference laboratory method for the starch calibration is not as reproducible as the other constituents. Cross validation for the protein, oil, density combination was conducted (Figure 3) and the results compared to the PLS model. The calculation model had a SECV of 0.030 gal/bu and $R^2$ of 0.79. The component calculation and the PLS model had similar results.

In the initial validation (55 of the 80 samples) external samples representing a wide range of constituents the PLS calibration had poor results for predicting corn ethanol yield with a SEP=0.40 gal/bu and $R^2$=0.28. These results were inconsistent with the initial cross validation of the calibration. The protein-oil-density component model gave consistent results to the original cross validation with a SEP=0.044 gal/bu and $R^2$=0.88. In the validation, the component model provided consistent results with initial cross validation. It was clear from the first external validation that a calibration approach would take repeated updates to remain useful.

The protein-oil-density calculation equation was then used to predict the ethanol yield of all 80 validation samples. These predicted values were then compared with the reference value measured by ICIA. The validation had a SEP of 0.086 gal/bu and $R^2$ of 0.73 (Figure 4). When the specialty samples were removed the 80-sample validation had a SEP of 0.056 gal/bu and $R^2$ of 0.79. The 2008 crop samples added some diversity not present in the previous data.
Ten of the validation samples were sent to the ICIA laboratory to be analyzed at different times to test the reproducibility of the laboratory method. The ICIA laboratory method was very reproducible with a coefficient of variation (CV) of 1.1% (Table 5). The standard deviation among replicates was nearly the same as the SECV of the either the calibration or the calculation, which means the NIRS accuracy likely will not get any better than presently reported.

All samples with reference values were pooled (316 samples) and new coefficients were calculated for the PLS calibration and component equation. The high oil samples were again removed. The final ethanol yield calibration included 283 samples and had a $R^2$ of 0.88 and SECV of 0.031 gal/bu. The final protein-oil-density component equation included 293 samples and had an $R^2$ of 0.74 and SECV of 0.042 gal/bu. The ethanol yield equation follows:

$$\text{Ethanol Yield} = 3.227 - 0.0624 \times \text{Protein (}) \% \text{) + 0.0296 } \times \text{Oil (}) \% \text{) + 0.104 } \times \text{Density (g/cc) (Table 6).}$$

This equation can be used to predict the ethanol yield of typical corn samples with the following constituent ranges: protein 5.0-12.9%, oil 2.6-6.4%, density 1.185-1.328 g/cc all at 15% moisture. The calibration could also be used but likely would need constant updates to include a large number and wider variety of samples. The protein, oil, and density calibrations include a large number of samples (Table 3) over many years. This equation applies to only one NIRS model with one set of constituent calibrations. The impact of other protein, oil, and density calibrations in other NIRS models was not studied.
CONCLUSIONS

A PLS ethanol yield calibration was developed using ethanol yield reference data from ICIA. The ICIA fermentable starch laboratory method was determined to be reproducible with a CV of 1.1%. Using the component calculation with predicted values of protein, oil, and density to predict ethanol yield, resulted in similar validation statistics as the new NIRS calibration. The PLS model was slightly more precise. The protein-oil-density component calculation performed better than the NIRS calibration in external validation.

Using the protein-oil-density component calculation is more practical because calibrations for these parameters are currently in use on NIRS machines. Implementing the calculation model would be easier than implementing a new NIRS calibration for fermentable starch. Any accurately calibrated NIRS unit can theoretically use the constituent regression, although a sensitivity analysis was not done in this study.

Using the MLR component combination protein, oil, and density in the calculation equation produces consistent and reliable results in validation. In the future, this equation could used to screen large number of corn samples for ethanol yield. Ranking corn for ethanol yield can be rapid, inexpensive and accurate.

LITERATURE CITED


Table 1. Near-Infrared Spectroscopy (NIRS) Predicted Protein, Oil, Starch, and Density Data of Initial 249 Corn Samples

<table>
<thead>
<tr>
<th></th>
<th>Protein&lt;sup&gt;a&lt;/sup&gt; (%)</th>
<th>Oil&lt;sup&gt;a&lt;/sup&gt; (%)</th>
<th>Starch&lt;sup&gt;a&lt;/sup&gt; (%)</th>
<th>Density&lt;sup&gt;b&lt;/sup&gt; (g/cc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>8.4</td>
<td>3.8</td>
<td>60.1</td>
<td>1.276</td>
</tr>
<tr>
<td>St. Dev.&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.8</td>
<td>0.6</td>
<td>1.1</td>
<td>0.018</td>
</tr>
<tr>
<td>Min.&lt;sup&gt;d&lt;/sup&gt;</td>
<td>11.4</td>
<td>6.1</td>
<td>62.6</td>
<td>1.322</td>
</tr>
<tr>
<td>Max.&lt;sup&gt;e&lt;/sup&gt;</td>
<td>6.3</td>
<td>3.1</td>
<td>55.9</td>
<td>1.195</td>
</tr>
</tbody>
</table>

<sup>a</sup> NIRS predicted (by Iowa State University Grain Quality Laboratory calibrations) values at 15% moisture

<sup>b</sup> Grams/cubic centimeter

<sup>c</sup> Standard Deviation

<sup>d</sup> Minimum value

<sup>e</sup> Maximum value
Table 2. Progressive Combinations of Components Used for Multiple Linear Regression

<table>
<thead>
<tr>
<th>Combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein</td>
</tr>
<tr>
<td>Protein Oil</td>
</tr>
<tr>
<td>Protein Oil Starch</td>
</tr>
<tr>
<td>Protein Oil Density</td>
</tr>
<tr>
<td>Protein Oil Starch Density</td>
</tr>
<tr>
<td>Protein Starch</td>
</tr>
<tr>
<td>Protein Starch Density</td>
</tr>
<tr>
<td>Protein Density</td>
</tr>
<tr>
<td>Oil</td>
</tr>
<tr>
<td>Oil Starch</td>
</tr>
<tr>
<td>Oil Starch Density</td>
</tr>
<tr>
<td>Oil Density</td>
</tr>
<tr>
<td>Starch</td>
</tr>
<tr>
<td>Starch Density</td>
</tr>
<tr>
<td>Density</td>
</tr>
</tbody>
</table>
Table 3. Iowa State University Grain Quality Laboratory Near-Infrared Spectroscopy Calibration Models

<table>
<thead>
<tr>
<th>Calibration Name&lt;sup&gt;a&lt;/sup&gt;</th>
<th>cornprot2005 &lt;sub&gt;_model&lt;/sub&gt;</th>
<th>cornoil2005 &lt;sub&gt;_model&lt;/sub&gt;</th>
<th>cornstarch2005 &lt;sub&gt;_model&lt;/sub&gt;</th>
<th>corndens2003 &lt;sub&gt;_model&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein&lt;sup&gt;b&lt;/sup&gt; (%)</td>
<td>0.17</td>
<td>0.20</td>
<td>0.63</td>
<td>0.19</td>
</tr>
<tr>
<td>Oil&lt;sup&gt;b&lt;/sup&gt; (%)</td>
<td>3086</td>
<td>3659</td>
<td>2599</td>
<td>2582</td>
</tr>
<tr>
<td>Starch&lt;sup&gt;b&lt;/sup&gt; (%)</td>
<td>8.81</td>
<td>4.30</td>
<td>60.35</td>
<td>1.28</td>
</tr>
<tr>
<td>Density&lt;sup&gt;c&lt;/sup&gt; (g/cc)</td>
<td>2.06</td>
<td>1.45</td>
<td>3.98</td>
<td>0.49</td>
</tr>
<tr>
<td>SEP&lt;sup&gt;d&lt;/sup&gt;</td>
<td>11.69</td>
<td>13.87</td>
<td>24.79</td>
<td>0.23</td>
</tr>
</tbody>
</table>

<sup>a</sup> Calibration ID used by Iowa State University Grain Quality Laboratory

<sup>b</sup> 15% moisture basis

<sup>c</sup> Grams/cubic centimeter

<sup>d</sup> Standard error of prediction

<sup>e</sup> Number of samples

<sup>f</sup> Mean of reference samples used for calibration

<sup>g</sup> Standard deviation of reference samples used for calibration

<sup>h</sup> Range of reference samples used for calibration
Table 4. Prediction Results for the Component Calculation Using Combinations of Protein, Oil, Starch and Density

<table>
<thead>
<tr>
<th>Combination</th>
<th>SECV&lt;sup&gt;a&lt;/sup&gt; (gal/bu)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>R&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oil Starch Density</td>
<td>0.030</td>
<td>0.80</td>
</tr>
<tr>
<td>Protein Oil Starch Density</td>
<td>0.030</td>
<td>0.80</td>
</tr>
<tr>
<td>Protein Oil Starch</td>
<td>0.030</td>
<td>0.80</td>
</tr>
<tr>
<td>Protein Oil Density</td>
<td>0.030</td>
<td>0.79</td>
</tr>
<tr>
<td>Protein Starch Density</td>
<td>0.031</td>
<td>0.78</td>
</tr>
<tr>
<td>Protein Starch</td>
<td>0.031</td>
<td>0.79</td>
</tr>
<tr>
<td>Oil Starch</td>
<td>0.031</td>
<td>0.79</td>
</tr>
<tr>
<td>Protein Oil</td>
<td>0.032</td>
<td>0.78</td>
</tr>
<tr>
<td>Protein</td>
<td>0.033</td>
<td>0.76</td>
</tr>
</tbody>
</table>

<sup>a</sup> Standard error of cross validation

<sup>b</sup> Gallons/bushel

<sup>c</sup> Coefficient of determination
**Table 5. Reproducibility of Illinois Crop Improvement Association Laboratory Fermentable Starch Method**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Rep 1 Ethanol Yield (gal/bu)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Rep 2 Ethanol Yield (gal/bu)</th>
<th>Rep 3 Ethanol Yield (gal/bu)</th>
<th>St. Dev.&lt;sup&gt;b&lt;/sup&gt; (gal/bu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.52</td>
<td>2.43</td>
<td>-</td>
<td>0.063</td>
</tr>
<tr>
<td>2</td>
<td>2.67</td>
<td>2.69</td>
<td>-</td>
<td>0.013</td>
</tr>
<tr>
<td>3</td>
<td>2.53</td>
<td>2.50</td>
<td>-</td>
<td>0.025</td>
</tr>
<tr>
<td>4</td>
<td>2.42</td>
<td>2.45</td>
<td>-</td>
<td>0.021</td>
</tr>
<tr>
<td>5</td>
<td>2.45</td>
<td>2.42</td>
<td>-</td>
<td>0.019</td>
</tr>
<tr>
<td>6</td>
<td>2.84</td>
<td>2.81</td>
<td>-</td>
<td>0.025</td>
</tr>
<tr>
<td>7</td>
<td>2.75</td>
<td>2.8</td>
<td>2.70</td>
<td>0.032</td>
</tr>
<tr>
<td>8</td>
<td>2.76</td>
<td>2.73</td>
<td>2.71</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td><strong>St. Dev.</strong></td>
<td></td>
<td></td>
<td><strong>0.027</strong></td>
</tr>
<tr>
<td>Average</td>
<td>Ethanol Yield (gal/bu)</td>
<td></td>
<td></td>
<td><strong>2.62</strong></td>
</tr>
<tr>
<td></td>
<td><strong>CV&lt;sup&gt;c&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td><strong>1.05%</strong></td>
</tr>
</tbody>
</table>

<sup>a</sup> Gallons/bushel

<sup>b</sup> Standard deviation

<sup>c</sup> Coefficient of variation
<table>
<thead>
<tr>
<th>Calibration Name</th>
<th>Component</th>
<th>Original Equation Coefficients</th>
<th>Final Equation Coefficientsa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n\textsuperscript{b}</td>
<td>237</td>
<td>283</td>
</tr>
<tr>
<td></td>
<td>B0</td>
<td>3.20</td>
<td>3.23</td>
</tr>
<tr>
<td>cornprot2005_model\textsuperscript{c}</td>
<td>Protein\textsuperscript{d}</td>
<td>-0.066</td>
<td>-0.062</td>
</tr>
<tr>
<td>cornoil2005_model\textsuperscript{c}</td>
<td>Oil\textsuperscript{d}</td>
<td>-0.020</td>
<td>-0.030</td>
</tr>
<tr>
<td>corndens2003_model\textsuperscript{c}</td>
<td>Density\textsuperscript{d}</td>
<td>0.129</td>
<td>0.104</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Final equation includes initial calibration set and the external validation set

\textsuperscript{b} Number of samples

\textsuperscript{c} Calibration ID used by Iowa State University Grain Quality Laboratory

\textsuperscript{d} 15% moisture
Figure 1. Reference ethanol yield (gallons/bushel) distribution of initial 249 corn samples.
**Figure 2.** Partial least squares cross validation for near-infrared spectra against reference corn ethanol yield (gallons/bushel).
Figure 3. Component calculation cross validation for protein, oil, and density against reference corn ethanol yield (gallons/bushel).
Figure 4. Predicted ethanol yield vs. reference corn ethanol yield for 80 external validation samples.
PART II: APPLICATION OF ETHANOL YIELD PREDICTION TO CORN SAMPLES

A paper to be submitted to Cereal Chemistry

Allison Burgers¹ and Charles R. Hurburgh, Jr.²

ABSTRACT

A near-infrared spectroscopic (NIRS) method to determine the dry-grind ethanol yield of corn samples was developed in 2009 at the Iowa State University Grain Quality Laboratory. This method uses a component calculation equation with NIRS predicted protein, oil, and density values to predict the ethanol yield of corn samples. The ethanol yield equation was applied to data from previous crop years (2005-2008) corn samples from Iowa as well as to samples obtained from a planting date study conducted by Iowa State University Extension in 2008. When applied to data from previous crop years the equation produced expected results with an average of 2.80 gallons/bushel (gal/bu) and a range of 0.25 gal/bu. The ethanol yield increased steadily from crop years 2005-2008. In the planting date study, there were significant losses in grain yield per acre as well as a loss in ethanol yield per bushel for grains produced with later planting dates. These losses would create economic loss from grain yield decrease and decrease in ethanol production.

¹ Graduate student, Department of Food Science and Human Nutrition, Iowa State University, Ames, IA 50011
² Major Professor, Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, IA 50011
INTRODUCTION

The production of fuel ethanol in the U.S. has increased rapidly from 175 million gallons in 1980 to 9 billion in 2008 (Renewable Fuels Association, 2009). There is now a need to make the ethanol production process more efficient through input corn quality control. The majority of ethanol produced in the U.S. is done by the dry-grind process (Mosier and Ilelej, 2006). The dry-grind method processes the entire kernel. The products are then separated at the end of the process. Dry-grind ethanol plants in Iowa produce 2.80 gallons per bushel (gal/bu) on average (Iowa Renewable Fuels Association, 2009).

A rapid method to determine the ethanol yield of corn samples was developed at the Iowa State University (ISU) Grain Quality Laboratory (GQL) in 2009 (Burgers and Hurburgh, 2009). This method uses a multiple linear component calculation equation (coefficient of determination of ($R^2$) =0.74 and standard error of cross validation (SECV)=0.042 gal/bu) including protein, oil, and density values predicted by near-infrared spectroscopy (NIRS) to predict the ethanol yield in gallons per bushel (gal/bu) at 15% moisture. This method could be used by ethanol plants to estimate ethanol production at corn receiving.

Using NIRS to predict the quality of corn, in terms of the concentration of components, is an accepted method in the grain industry and is commonly used to accurately predict protein, oil, and starch contents in corn (Orman and Schumann, 1991). NIRS has begun to be used by the ethanol industry. Calibrations have been developed to predict extractable starch for the wet-milling industry (Paulsen et al, 2003; Paulsen and Singh, 2004). Pioneer Hi-Bred International Inc. (www.pioneer.com) along with Foss North America (FOSS Group, www.foss.dk) has introduced a NIRS calibration which predicts directly from
the spectra the amount of ethanol yield in a corn sample (Bryan, 2003). This is used to identify Pioneer highly fermentable (HTF) hybrids at ethanol plants.

The effect of delayed corn planting is a concern to corn users because later planted corn does not receive the same light and heat units over the season as earlier planted corn. The yield of the corn can be decreased with delayed planting dates (Lauer, 2007). These are important to the ethanol industry economic and procurement needs.

The objective of this study was to apply the equation developed by the ISU GQL to typical samples that would be received by ethanol plants, as well as data from a planting date study conducted by Iowa State University Extension in 2008; and to examine the impact of this data on ethanol plant economics and procurement.

MATERIALS AND METHODS

Ethanol Yield Equation

Ethanol Yield Equation was developed using current ISU GQL calibrations for protein, oil, and density (Burgers and Hurburgh, 2009). The equation is only applicable at this point to ISU GQL calibrations for Infratec NIRS machines. The impact of using other calibrations has not been tested, although in theory this should work. Statistics for these calibrations are shown in Table 1. The ethanol yield equation gave accurate consistent results in validation. The initial model included 293 and had a SECV of 0.042 gal/bu and $R^2$ of 0.74. The reference samples were fermented by Illinois Crop Improvement Association laboratory method (Illinois Crop Improvement Association, 2008). Ethanol Yield was predicted by the
ethanol yield equation using the coefficients shown in Table 2, and can be used on corn samples within the constituent ranges shown in Table 2.

**Data from Previous Crop Years**

The component calculation was used to predict the ethanol yield from Iowa corn samples (2121 samples) from four crop years (2005-2008). These samples originated from 0.25-0.50 acre hybrid comparison trial data planted by grower groups, suppliers, and ISU Extension. Hybrids included in this analysis were those typically offered for sale in Iowa and therefore likely were representative of average Iowa corn production. These samples were analyzed on their respective harvest dates on the Infratec 1229 Grain Analyzer (FOSS Group, www.foss.dk) which was calibrated by ISU GQL. The Infratec calibrations were not changed over the years represented in this study. The Infratec units were calibrated annually and subjected to rigorous quality control and validation procedures outlined by AACC Method 39-00. For more information on near-infrared technology see Williams and Norris (2001). The ethanol yield equation was applied to the component results. JMP 6.0.2 (SAS Institute Inc., www.sas.com) was used for significance tests.

**Planting Date Study**

In 2008, Iowa State University Extension conducted a planting date study at six Iowa locations (Table 3), measuring the impact of planting date on grain yield. Samples were retained and scanned on the Infratec 1229 Grain Analyzer. Protein, oil, and density values were predicted using ISU GQL corn calibrations (Table 1). The three northern locations (northwest, north, and northeast) results were analyzed as this section represents the majority of corn grown in Iowa (USDA, 2008). This is only one year of a planned multiple year study,
but the analysis demonstrated the economic influences of yield and quality of corn for ethanol production.

RESULTS AND DISCUSSION

Data from Previous Crop Years

The component calculation (Table 2) was used to predict ethanol yield on the data from crop years 2005 to 2008. There were differences between years, and across samples. The ethanol yield data was further analyzed to show the range and to estimate the significance of year to year variations (Table 4).

The averages and ranges of the ethanol yield data were calculated (Figure 1). There was a steady and significant increase (student’s t-test p<0.05) in the ethanol yield per bushel over the four years 2005-2008, with 2008 being the highest, as expected. The range was consistently close to 0.25 gal/bu (8.9%) every year.

The average ethanol yield increased by 0.04 gal/bu from 2005 to 2008. Using the assumptions from the ISU Extension model ‘Ethanol Profitability’ (Hofstrand, 2008) a 100 million gallon per year (MGY) producing ethanol plant would see an increase of $2,028,571 per year of ethanol produced. The average ethanol yield over all the data from past crop years was 2.80 gal/bu, at current ethanol prices a 100 MGY ethanol plant would yield $142,000,000 of ethanol per year with a standard deviation of $2,130,000 per year. The model uses spot bid daily corn prices (No. 2 yellow) and ethanol daily price at northern Iowa ethanol plants as reported by USDA Ag Market in ‘Iowa Ethanol Plant Report’ and then converted into monthly average prices (Hofstrand, 2008).
**Planting Date**

Most locations had increased protein content and decreased grain yield with later planting dates (Table 5). At most locations, starch content, ethanol yield, ethanol amount and overall grain yields decreased with later planting dates. The Humeston (S) and Crawfordsville (SE) locations were the exceptions; there were fewer planting dates and fewer useable samples because of weed problems and excessive rainfall in the plots. The southern locations results are not used in this analysis.

The three northern locations showed similar trends for all constituents specifically ethanol yield, and grain yield. Grain yield and ethanol yield decreased with delayed planting. This data is summarized in Table 6.

The grain yield and ethanol yield data were converted to a percent of maximum yield to normalize the data for location differences and actual planting dates (Figure 2 and Figure 3). The percent grain yield vs. planting date model showed a very good fit with a R\(^2\) of 0.92 (Figure 2). The ethanol yield model had a R\(^2\) of 0.63 (Figure 3). Both grain and ethanol yields peaked around the early May planting date and decreased thereafter. This analysis updated for more years and hybrids could be used by ethanol plants to predict the amount of grain or ethanol yield per bushel from average planting dates in the area. If yields are reduced they will have to pay more to obtain grains from areas further away.

The ethanol amount in gallons per acre (gal/acre) was calculated by multiplying the grain yield (bu/acre) by the ethanol yield (gal/bu). The loss in grain yield, ethanol yield, and ethanol amount due to delayed planting dates would have an economic impact on the corn ethanol industry. Losses were calculated by subtracting the yield or amount at the planting date in question from the highest yielding planting date. Average was calculated by
subtracting the latest planting date from the highest yielding planting date. Loss increases with later planting dates were seen in grain yield and ethanol yield and therefore ethanol amount at the northern locations (Table 7).

Approximately 6.33 gal/acre (average grain yield (186.3 bu/acre)*average ethanol yield loss (0.034 gal/bu)) of the loss was due to reduced ethanol yield per bushel, and 99.7 gal/acre was due to reduced grain yield per acre. The corn quality loss for ethanol production was not as important as the overall grain yield loss, in ethanol yield per acre. However, the ethanol plant would experience the ethanol yield loss more directly than the grain yield loss.

Corn and ethanol prices were taken from the previously described economic model ‘Ethanol Profitability’. (Hofstrand, 2008). For a case study example:

Corn price March 2009 = $3.62/bu
Ethanol Price March 2009 = $1.42/gal

Using these prices and the loss in quality of corn at the latest planting date causes a $0.05/bu loss for ethanol production (Table 8). A 100 MGY producing ethanol plant uses 35,714,286 bushels of corn per year when operating at 100% capacity and an efficiency of 2.80 gal/bu (Hofstrand, 2008). Delayed planting (mid-June) would cause a $1,785,714 loss per year for a 100 MGY ethanol plant; a 50 MGY plant would lose $892,857 due to loss in corn quality from delayed planting.

CONCLUSIONS

The equation for predicting ethanol yield in corn samples was applied to data from previous crop years from 2005 to 2008. The results were consistent with yield reports used in
the ethanol industry with an overall average of 2.80 gal/bu and a range close to 0.25 gal/bu. The ethanol yield increased from 2005 to 2008 as expected, as hybrids have been developed especially for ethanol fermentation.

The equation was also applied to samples from an ISU Extension planting date study. For the samples in this study, delayed planting dates caused an increase in protein content with losses in grain yield. Later planting dates also saw decreases in starch content, ethanol yield, and ethanol amount. These losses contribute to a significant economic loss in grain yield and corn quality for ethanol production.

LITERATURE CITED


### Table 1. Protein, Oil, and Density Near-Infrared Spectroscopy Calibration Statistics

<table>
<thead>
<tr>
<th>Calibration Name</th>
<th>cornprot2005_model</th>
<th>cornoil2005_model</th>
<th>corndens2003_model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protein (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEP^b</td>
<td>0.17</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td>n^c</td>
<td>3086</td>
<td>3659</td>
<td>2582</td>
</tr>
<tr>
<td>Reference Mean^f</td>
<td>8.81</td>
<td>4.30</td>
<td>1.28</td>
</tr>
<tr>
<td>Reference St. Dev.^g</td>
<td>2.06</td>
<td>1.45</td>
<td>0.49</td>
</tr>
<tr>
<td>Reference Range^h</td>
<td>11.69</td>
<td>13.87</td>
<td>0.23</td>
</tr>
</tbody>
</table>

^a Calibration ID used by the Iowa State University Grain Quality Laboratory

^b Standard error of prediction

^c Number of samples

^d Protein, oil and density at 15% moisture

^e Grams/cubic centimeter

^f Mean of reference samples used in calibration

^g Standard deviation of reference samples used in calibration

^h Range of reference samples used in calibration
Table 2. Ethanol Yield Equation Coefficients and Constituent Ranges

<table>
<thead>
<tr>
<th>Calibration Name</th>
<th>Component</th>
<th>Units</th>
<th>Equation Coefficients</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>B0</td>
<td></td>
<td></td>
<td>3.227</td>
<td></td>
</tr>
<tr>
<td>cornprot2005_model</td>
<td>Protein</td>
<td>% at 15% M</td>
<td>-0.062</td>
<td>5.0-12.9%</td>
</tr>
<tr>
<td>cornoil2005_model</td>
<td>Oil</td>
<td>% at 15% M</td>
<td>-0.030</td>
<td>2.6-6.4%</td>
</tr>
<tr>
<td>corndens2003_model</td>
<td>Density</td>
<td>g/cc at 15% M</td>
<td>0.104</td>
<td>1.185-1.328</td>
</tr>
</tbody>
</table>

\(^a\) Calibration ID used by Iowa State University Grain Quality Laboratory

\(^b\) Coefficients used to predict ethanol yield

\(^c\) Percent at 15 percent moisture

\(^d\) Grams/cubic centimeter at 15 percent moisture
Table 3. Iowa Locations of Iowa State University Extension Planting Date Study

<table>
<thead>
<tr>
<th>Area</th>
<th>Location</th>
<th>Number of Hybrids</th>
</tr>
</thead>
<tbody>
<tr>
<td>North West</td>
<td>Doon</td>
<td>2</td>
</tr>
<tr>
<td>North</td>
<td>Kanawha</td>
<td>1</td>
</tr>
<tr>
<td>North East</td>
<td>Nashua</td>
<td>2</td>
</tr>
<tr>
<td>South West</td>
<td>Lewis</td>
<td>1</td>
</tr>
<tr>
<td>South</td>
<td>Humeston</td>
<td>1</td>
</tr>
<tr>
<td>South East</td>
<td>Crawfordsville</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 4. Variety Trial Iowa Corn Samples 2005-2008

<table>
<thead>
<tr>
<th>Calibration Name&lt;sup&gt;c&lt;/sup&gt;</th>
<th>cornprot20&lt;sup&gt;05_model&lt;/sup&gt;</th>
<th>cornoil20&lt;sup&gt;05_model&lt;/sup&gt;</th>
<th>cornstarch200&lt;sup&gt;5_model&lt;/sup&gt;</th>
<th>corndens2003&lt;sup&gt;_model&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficients&lt;sup&gt;f&lt;/sup&gt;</td>
<td>3.23</td>
<td>-0.062</td>
<td>-0.030</td>
<td>0.104</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>B0</th>
<th>Protein&lt;sup&gt;g&lt;/sup&gt;</th>
<th>Oil&lt;sup&gt;h&lt;/sup&gt;</th>
<th>Starch&lt;sup&gt;h&lt;/sup&gt;</th>
<th>Density&lt;sup&gt;i&lt;/sup&gt;</th>
<th>Ethanol Yield&lt;sup&gt;j&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n&lt;sup&gt;g&lt;/sup&gt;</td>
<td>(%)&lt;sup&gt;h&lt;/sup&gt;</td>
<td>(%)&lt;sup&gt;h&lt;/sup&gt;</td>
<td>(%)&lt;sup&gt;h&lt;/sup&gt;</td>
<td>(g/cc)&lt;sup&gt;ii&lt;/sup&gt;</td>
<td>(gal/bu)&lt;sup&gt;j&lt;/sup&gt;</td>
</tr>
<tr>
<td>2005</td>
<td>437</td>
<td>7.5&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.5&lt;sup&gt;a&lt;/sup&gt;</td>
<td>61.0&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.271&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.79&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>2005 St. Dev.&lt;sup&gt;k&lt;/sup&gt;</td>
<td>0.73</td>
<td>0.21</td>
<td>0.73</td>
<td>0.019</td>
<td>0.047</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>796</td>
<td>7.4&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.6&lt;sup&gt;b&lt;/sup&gt;</td>
<td>60.6&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.279&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.79&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>2006 St. Dev.</td>
<td>0.60</td>
<td>0.19</td>
<td>0.98</td>
<td>0.021</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>536</td>
<td>7.3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.6&lt;sup&gt;c&lt;/sup&gt;</td>
<td>61.1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.271&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.80&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>2007 St. Dev.</td>
<td>0.58</td>
<td>0.18</td>
<td>0.57</td>
<td>0.019</td>
<td>0.037</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>352</td>
<td>6.8&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3.4&lt;sup&gt;d&lt;/sup&gt;</td>
<td>61.6&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.25&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.83&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>2008 St. Dev.</td>
<td>0.53</td>
<td>0.23</td>
<td>0.52</td>
<td>0.023</td>
<td>0.035</td>
<td></td>
</tr>
<tr>
<td>All Years</td>
<td>2121</td>
<td>7.3</td>
<td>3.5</td>
<td>61.0</td>
<td>1.27</td>
<td>2.80</td>
</tr>
<tr>
<td>All Years St. Dev.</td>
<td>0.65</td>
<td>0.20</td>
<td>0.85</td>
<td>0.023</td>
<td>0.042</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a-d</sup> Numbers with different letter superscripts within the same column are significantly different by students t-test (p<0.05)

<sup>c</sup> Calibration ID used by Iowa State University Grain Quality Laboratory

<sup>f</sup> Equation coefficients used to predict ethanol yield

<sup>g</sup> Number of samples

<sup>h</sup> Protein, oil, starch, density and ethanol yield are all at 15% moisture

<sup>i</sup> Grams per cubic centimeter

<sup>j</sup> Gallons/bushel

<sup>k</sup> Standard deviation
Table 5. Planting Date Trends

<table>
<thead>
<tr>
<th>Location</th>
<th>Protein</th>
<th>Grain Yield</th>
<th>Ethanol Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>North West</td>
<td>Up</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>North</td>
<td>Up</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>North East</td>
<td>Up</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>South West</td>
<td>Up</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>South</td>
<td>Down</td>
<td>Up</td>
<td>Up</td>
</tr>
<tr>
<td>South East</td>
<td>None</td>
<td>Down</td>
<td>Down</td>
</tr>
</tbody>
</table>
Table 6. Grain and Ethanol Yield Data for Northern Iowa Locations

<table>
<thead>
<tr>
<th>Location</th>
<th>Planting Date&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Mean Grain Yield (bu/acre)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Grain Yield Range (bu/acre)</th>
<th>St. Dev.&lt;sup&gt;c&lt;/sup&gt; (bu/acre)</th>
<th>Mean Ethanol Yield (gal/bu)&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Ethanol Yield Range (gal/bu)</th>
<th>St. Dev. (gal/bu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NW</td>
<td>4/23/2008</td>
<td>206.3</td>
<td>36.8</td>
<td>9.2</td>
<td>2.79</td>
<td>0.025</td>
<td>0.008</td>
</tr>
<tr>
<td>N</td>
<td>4/23/2008</td>
<td>179.2</td>
<td>20.1</td>
<td>9.8</td>
<td>2.81</td>
<td>0.025</td>
<td>0.011</td>
</tr>
<tr>
<td>NE</td>
<td>4/16/2008</td>
<td>183.9</td>
<td>83.1</td>
<td>21.6</td>
<td>2.81</td>
<td>0.079</td>
<td>0.024</td>
</tr>
<tr>
<td>NW</td>
<td>4/30/2008</td>
<td>204.3</td>
<td>33.1</td>
<td>10.7</td>
<td>2.79</td>
<td>0.050</td>
<td>0.016</td>
</tr>
<tr>
<td>N</td>
<td>4/30/2008</td>
<td>180.1</td>
<td>18.6</td>
<td>9.6</td>
<td>2.79</td>
<td>0.028</td>
<td>0.012</td>
</tr>
<tr>
<td>NE</td>
<td>4/30/2008</td>
<td>193.4</td>
<td>74.5</td>
<td>21.8</td>
<td>2.81</td>
<td>0.078</td>
<td>0.026</td>
</tr>
<tr>
<td>NW</td>
<td>5/8/2009</td>
<td>211.2</td>
<td>26.8</td>
<td>7.4</td>
<td>2.78</td>
<td>0.032</td>
<td>0.011</td>
</tr>
<tr>
<td>N</td>
<td>5/9/2008</td>
<td>178.2</td>
<td>16.1</td>
<td>8.3</td>
<td>2.81</td>
<td>0.013</td>
<td>0.006</td>
</tr>
<tr>
<td>NE</td>
<td>5/13/2008</td>
<td>189.7</td>
<td>67.9</td>
<td>17.9</td>
<td>2.81</td>
<td>0.097</td>
<td>0.028</td>
</tr>
<tr>
<td>NW</td>
<td>5/19/2008</td>
<td>207.4</td>
<td>26.7</td>
<td>7.5</td>
<td>2.77</td>
<td>0.033</td>
<td>0.010</td>
</tr>
<tr>
<td>N</td>
<td>5/22/2009</td>
<td>172.5</td>
<td>15.1</td>
<td>6.8</td>
<td>2.80</td>
<td>0.045</td>
<td>0.020</td>
</tr>
<tr>
<td>NE</td>
<td>5/25/2008</td>
<td>181.8</td>
<td>69.9</td>
<td>20.9</td>
<td>2.80</td>
<td>0.102</td>
<td>0.030</td>
</tr>
<tr>
<td>NW</td>
<td>5/25/2008</td>
<td>186.1</td>
<td>73.9</td>
<td>23.8</td>
<td>2.75</td>
<td>0.069</td>
<td>0.023</td>
</tr>
<tr>
<td>N</td>
<td>6/1/2008</td>
<td>157.4</td>
<td>34.7</td>
<td>14.8</td>
<td>2.79</td>
<td>0.050</td>
<td>0.023</td>
</tr>
<tr>
<td>NE</td>
<td>6/11/2008</td>
<td>131.8</td>
<td>82.9</td>
<td>20.8</td>
<td>2.77</td>
<td>0.083</td>
<td>0.023</td>
</tr>
</tbody>
</table>

<sup>a</sup> Planting date month/day/year  
<sup>b</sup> Bushels/acre at 15% moisture  
<sup>c</sup> Standard deviation  
<sup>d</sup> Gallons per bushel at 15% moisture
Table 7. Losses in Grain and Ethanol Yield by Planting Date for Northern Iowa Locations

<table>
<thead>
<tr>
<th>Location</th>
<th>Planting Date</th>
<th>Grain Yield Loss (bu/acre)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Ethanol Yield Loss (gal/bu)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Ethanol Amount Loss (gal/acre)&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>4/23/2008</td>
<td>1.0</td>
<td>0.00&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>4/30/2008</td>
<td>0.0</td>
<td>0.015</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>5/9/2008</td>
<td>2.0</td>
<td>0.003</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>5/22/2008</td>
<td>7.6</td>
<td>0.010</td>
<td>20.4</td>
</tr>
<tr>
<td></td>
<td>6/1/2008</td>
<td>22.8</td>
<td>0.020</td>
<td>64.1</td>
</tr>
<tr>
<td>North West</td>
<td>4/23/2008</td>
<td>4.9</td>
<td>0.0</td>
<td>7.8</td>
</tr>
<tr>
<td></td>
<td>4/30/2008</td>
<td>6.9</td>
<td>0.0</td>
<td>16.2</td>
</tr>
<tr>
<td></td>
<td>5/8/2008</td>
<td>0.0</td>
<td>0.012</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>5/19/2008</td>
<td>3.8</td>
<td>0.019</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>5/28/2008</td>
<td>25.1</td>
<td>0.034</td>
<td>80.5</td>
</tr>
<tr>
<td>North East</td>
<td>4/16/2008</td>
<td>9.6</td>
<td>0.009</td>
<td>22.2</td>
</tr>
<tr>
<td></td>
<td>4/30/2008</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>5/13/2008</td>
<td>3.1</td>
<td>0.007</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>5/25/2008</td>
<td>12.3</td>
<td>0.013</td>
<td>30.9</td>
</tr>
<tr>
<td></td>
<td>6/11/2008</td>
<td>61.6</td>
<td>0.049</td>
<td>173.8</td>
</tr>
<tr>
<td>Average Loss&lt;sup&gt;e&lt;/sup&gt;</td>
<td></td>
<td>36.5</td>
<td>0.034</td>
<td>106.1</td>
</tr>
</tbody>
</table>

<sup>a</sup> Bushels/acre at 15% moisture  
<sup>b</sup> Gallons/bushel at 15% moisture  
<sup>c</sup> Gallons/acre at 15% moisture  
<sup>d</sup> A loss of 0 highest yielding planting date  
<sup>e</sup> Overall loss is the average (of the locations) at the latest planting date subtracted from the average at the highest yielding planting date
Table 8. Economic Loss by Planting Date Compiled Northern Iowa Locations

<table>
<thead>
<tr>
<th>Planting Date</th>
<th>$/acre loss in corn</th>
<th>$/bu(^a) loss in ethanol</th>
<th>$/acre loss in ethanol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18.57</td>
<td>0.00</td>
<td>14.16</td>
</tr>
<tr>
<td>2</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>6.20</td>
<td>0.01</td>
<td>3.36</td>
</tr>
<tr>
<td>4</td>
<td>28.66</td>
<td>0.02</td>
<td>31.17</td>
</tr>
<tr>
<td>5</td>
<td>132.08</td>
<td>0.05</td>
<td>150.72</td>
</tr>
</tbody>
</table>

\(^a\) U.S. Dollars

\(^b\) Bushels of corn

\(^c\) Prices based on current corn and ethanol prices as reported by USDA Ag Market (USDA, 2009)
*Error bar indicates plus and minus the standard error of the mean

a-c Years with different letters are significantly different by student’s t-test (p<0.05)

**Figure 11.** Ethanol yield mean over four years, 2005-2008.
Figure 2. Compiled percent maximum grain yield vs. planting date for 3 northern Iowa locations.
Figure 3. Compiled percent maximum ethanol yield vs. planting date for 3 northern Iowa locations.
GENERAL CONCLUSIONS

In part I, an initial PLS calibration was developed from samples and reference data from Illinois Crop Improvement Association. A component calculation was also developed using MLR and combinations of protein, oil, starch, and density predictions from Iowa State University Grain Quality Laboratory calibrations. The combination protein, oil, and density was chosen as the best combination as these have more reliable calibrations. The component calculation performed better than the PLS model in validation using a new sample set representing a wide variety of constituents. A final component equation was developed using all of the reference and NIRS data available.

In part II of the research, the component equation was applied to Iowa corn data from previous crop years and data from a 2008 Iowa State University Extension planting date study. When applied to data from previous crop years the equation was useful in estimating the range and average ethanol yield over the years 2005-2008. When applied to the planting date data the equation was useful in estimating ethanol yield and loss in ethanol yield due to delayed planting. This information was then useful when estimating economic losses to the ethanol plant.

The equation developed in part I performed well in cross validation and external validation. In part II, the equation was applied to actual field trial data. It was shown that this equation could be useful when applied to data from previous crop years and study data. Using the component equation instead of implementing a new NIRS calibration would be easier and more inexpensive for ethanol plants as calibrations for these constituents
are already in use. This technology would make ranking corn for ethanol yield, rapid, inexpensive and accurate for all genetics.

A limitation of this research is the use of NIRS instruments and the initial cost of the NIRS machine. NIRS calibration development and maintenance requires a skilled chemometrician. The reference method used in this research is also a limitation as it measures ethanol yield by the amount of weight lost from beginning to the end of the fermentation. Other volatile compounds besides CO₂ could be lost during the fermentation and contribute to error. However, this method showed to be reproducible in our research and provides results consistent with industry. In the future, other reference methods for ethanol yield fermentation could be studied.

The planting date study in part II of the research had a low number of samples, hybrids, and useable locations. Conclusions made in this section apply only to the samples in the study. More samples and hybrids could be added to this study in the future to make more certain conclusions.

Currently the ethanol yield equation can only be applied to NIRS predicted values of protein, oil and density using Iowa State University Grain Quality Laboratory NIRS calibrations. It will be important to examine the effects of using calibrations other than the ones used in this research. This research could also be expanded upon to also estimate the amount and the DDGS produced by the process and estimate the protein content and nutritional quality of the protein in the final DDGS. The ethanol yield equation could be field tested onsite at an ethanol plant to compare with plant output.
ACKNOWLEDGEMENTS

I would like to take this opportunity to express my appreciation to those who helped me during the process of conducting my research and the writing of this thesis. First, thank you to my major professor Dr. Charles R. Hurburgh Jr. for his support and guidance of my research. I would like to thank the members of my program of study committee Dr. Jay-lin Jane and Dr. Thomas Brumm. I would also like to thank Dr. Benoit Igne and Glen Rippke for help with the NIRS work. A special thanks to all the members of Iowa State University Grain Quality Laboratory for their daily support. Thank you to POET, LLC for research support and funding. Finally, I would like to thank my friends and family for their continued support.