DEVELOPMENT OF NIR SPECTROSCOPY MODELS FOR STARCH CONTENT PREDICTION AND ETHANOL PRODUCTION FROM MUTANT GRAIN SORGHUM

by

KAELIN E. SAUL

B.S., North Carolina State University, 2014

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Biological and Agricultural Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

Approved by:

Major Professor
Dr. Donghai Wang

Copyright

Abstract

The growing demands for renewable energy sources have led researchers to investigate other biomass sources, aside from maize. Grain sorghum is comparable to maize in its starch content and can be grown in regions with drier climates, where maize is a less suitable crop for these areas. In attempts to increase yield prior to harvest and for ethanol production, this study focuses on mutant grain sorghum. One hundred and nine mutant grain sorghum samples were analyzed for their chemical and physical properties and fermented into ethanol. The current method for starch analysis is time-consuming and tedious. Near infrared spectroscopy (NIR) models were developed as fast, cost-effective, and non-destructive methods for grain sorghum starch content analysis. Each mutated grain sorghum sample was scanned in a wavelength range from 4,000 to 10,000 cm^{-1} as a whole grain and in flour form. Partial Least Squares (PLS) regression method was used for NIR model development. The coefficients of determination (R^2) of 0.77 and 0.90 were achieved for starch content calibration and prediction models, respectively. This model demonstrates the possibility of a positive correlation between the actual and calculated values for starch content. Another PLS first derivative model with $R^2 = 0.95$ for calibration and a reduced wavelength range (4,000-5,176 cm^{-1}), using 39 of the original 109 samples (27 for calibration and 8 for validation), was created to predict the fermentation efficiency.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to start by thanking my major advisor, Dr. Donghai Wang, for his support and mentorship. His dedication to the field and quality of work are attributes that have been instilled into me. I would also like to thank Dr. Xiuzhi Sun and Dr. Lisa Wilken for serving on my committee and helping me through the thesis process.

I appreciate all of the advice and assistance from Dr. Ke Zhang for my research. I would also like to thank all of the students, specifically Bairen Pang and Meaghan Dunn, who helped me with my research. I would like to acknowledge all of my friends and the BAE graduate students for their continued support during my time at Kansas State University.

I would like to thank Michael Moore from the Grain Science department for allowing me accessibility to the lab to operate the RVA equipment. I would also like to thank Ravindra Thakkar from the Nanotechnology Innovation Center of Kansas State in the Department of Anatomy and Physiology for his SEM imaging services.

I am so thankful for the opportunity, presented to me by Dr. Mary Rezac and Keith Rutlin, to be a part of the National Science Foundation Integrative Graduate Education and Research Traineeship program. I would like to thank Dr. Naiqian Zhang and Dr. Joseph Harner for all of their support and for welcoming me into the Kansas State University BAE family. Lastly, I would like to thank Barbara Moore in the BAE department for helping me navigate through my graduate program and for always being there.

Dedication

I would like to dedicate this work to my sister, Amelia, who inspires me to be the best I can be every day and has always been there to support me. To my parents, Geoffrey and Brenda, thank you for all of your love and for always encouraging me to chase my dreams. For the inspiration from my grandmothers, Anita Kaelin Saul and Eileen Butler Chewens, who I am named after. I would also like to dedicate this work to my family, who have taught me what is truly important.

Chapter 1 - Introduction

1.1 Problem Statement

Despite policies being in place to encourage renewable energy, many obstacles prevent the production of biofuels on a larger scale. One of the main challenges is the use of biofuel feedstock with a minimal impact on human food and animal feed production. The main biomass source for first generation biofuels in the United States is maize. However, with an increase in bioethanol production, maize has become overused as a renewable source, which affects the amount of maize used for human food and animal feed consumption. An alternative crop that could be mixed with maize for ethanol production is grain sorghum. Grain sorghum has been used to produce gluten-free food products for people with celiac disease (Liu et al., 2013). Grain sorghum is also used for animal feed, but tannins can reduce the digestibility of starch and protein (Yan et al., 2009). While maize can produce a higher yield, it also requires more water and fertilizer than other crops, such as grain sorghum (Assefa et al., 2014). For this reason, grain sorghum, as opposed to maize, mostly thrives in the Central Plains, where there is a limited amount of rainfall and irrigation available.

Maize and grain sorghum are a part of the *Gramineae*, or grass family (Assefa et al., 2014). Both crops are comparable in starch content with maize ranging from ~63-73% (an average of 67%) and grain sorghum ranging from 64-74% with an average of 70% (Orman & Schumann Jr., 1991; Wu et al., 2007). According to Wang et al. (2008), there is a positive relationship between the amount of starch and the ethanol yield. Thus, in addition to maize, grain sorghum has the potential to become a supplemental feedstock for bioethanol production.

Chemical composition, especially starch content, is a major factor affecting ethanol yield. Even though, a higher amount of starch is correlated with a higher ethanol yield, there is not

necessarily a linear relationship between starch and fermentation efficiency (Wang et al., 2008). In addition, a statistical model could provide the opportunity to represent a relationship between fermentation efficiency and starch content.

One of the common methods for observing different components of cereal grains is infrared spectroscopy. While there has been extensive use of spectroscopy in the mid-infrared region, a more quantitative method for evaluation of chemical composition is near infrared (NIR) spectroscopy. NIR spectroscopy has been used for comparison of different cereal grains, but a majority of the studies emphasize grains other than grain sorghum. Thus, the use of NIR spectroscopy can potentially offer a high-throughput and cost-effective method for phenotyping grain sorghum, which would facilitate plant breeding and genetics studies, affected by the starch content and the fermentation efficiency.

1.2 Objectives

The ultimate goal of this research is to study the potential of mutant grain sorghum for bioethanol production and to create fast methods for sorghum starch analysis and prediction. The objectives that will help to achieve this goal are:

1. To evaluate the fermentation performance of mutant grain sorghum for ethanol production.
2. To develop NIR models for the prediction of starch content and fermentation efficiency of mutant grain sorghum.

1.3 Significance of Work

With the depletion of nonrenewable resources from the Earth, more sustainable practices must be implemented to address the energy needs of the world. According to Guo, Song, & Buhain (2015), it is believed that bioethanol and biodiesel will be the primary fuels for automobiles and larger vehicles by 2050. To meet these demands, many countries around the world have implemented policies to address this issue.

In United States, the Energy Independence and Security Act of 2007 mandates the reduction of gasoline use by 20% within 10 years and increase biofuel blending to 36 billion gallons by 2022 (Guo, Song, & Buhain, 2015). The European Union passed the Directive on the Promotion of Renewable Energy in 2009, which set up "20-20-20" objectives to be achieved by 2020: reduction of greenhouse gas emissions and energy use by 20% and increase the use of renewable energy sources by 20% of the total energy (Su, Zhang, & Su, 2015). Finally, in 1997, Brazil established the National Agency of Petroleum, Natural Gas, and Biofuels to focus on creating ethanol subsidies in the hopes of producing three times more ethanol by 2021 (Solomon et al., 2015). Brazil retracted caps on gasoline prices in 1997 since ethanol was able to join the competitive market with no subsidized credit (Solomon et al., 2015). Table 1 shows the world bioethanol as well as biodiesel production and its significant growth from 2009 to 2013 (F.O. Licht, 2014; Koizumi, 2015). Currently, the United States and Brazil make up 86% of the overall world production of bioethanol while the European Union produces the most biodiesel in the global market (Koizumi, 2015).

Table 1. Change in bioethanol production from 2009-2013 (F.O. Licht, 2014)

	(Million L)				
Year	**2009**	**2010**	**2011**	**2012**	**2013**
Bioethanol and biodiesel production total	90.87	104.57	108.51	109.65	117.72
World bioethanol production total	73.07	84.92	84.15	83.35	88.17
USA	40.73	50.09	52.81	50.35	50.40
Brazil	23.92	25.53	21.02	21.62	25.53
China	2.05	2.05	2.10	2.10	2.10
EU	3.55	4.14	4.39	4.51	4.55

While these policies have helped the movement towards renewable energy, many countries rely predominantly on one crop. For example, Brazil mainly uses sugarcane as their biomass source for ethanol production. The primary crop for the United States bioethanol production is maize, which ignited the Food versus Fuel debate. The reason for the debate is due to the multipurpose usefulness of maize as food and feed sources and for the conversion to bioethanol. With growing concerns for food security in the future, researchers have looked into other crops to meet growing demands for biofuels.

In comparison to maize, grain sorghum has a relatively high amount of starch and it is more tolerant to drought and heat stress when grown in drier climates (Yan et al., 2009). Grain sorghum production increased by 38% between 2014 and 2015, while maize production decreased by 4% (USDA-NASS, 2016). This shift in production demonstrates that there is a possibility for grain sorghum to be incorporated into maize ethanol plants to move towards less of a dependence on maize alone.

1.4 Literature Review

1.4.1 Biofuels

Aside from bioethanol, there are other types of biofuels that fall into different classifications based on the type of biomass. Biomass is any agricultural crop or renewable product used for biofuel production. The three categories, based on the type of biomass, are: first generation, second generation, and third generation of biofuels. First generation mainly includes plant seeds and grains, which compete with food and animal feed consumption. The second generation consists of cellulosic biomass and energy crops. Second generation feedstock does not compete with any food sources and the energy crops are only grown for renewable energy purposes. The third generation of biofuels is produced from algae. The type of biomass dictates the form of biofuel that can be produced. For example, used vegetable oil or oil seeds can be used to make biodiesel, but starchy grains or sugars can produce bioethanol and biobutanol. Figure 1 represents the different generations of biofuels and also demonstrates the difference between primary and secondary biofuels. Primary fuels do not undergo any modification and are completely natural whereas secondary biofuels undergo some form of processing, such as bioethanol or biogas (Nigam & Singh, 2011).

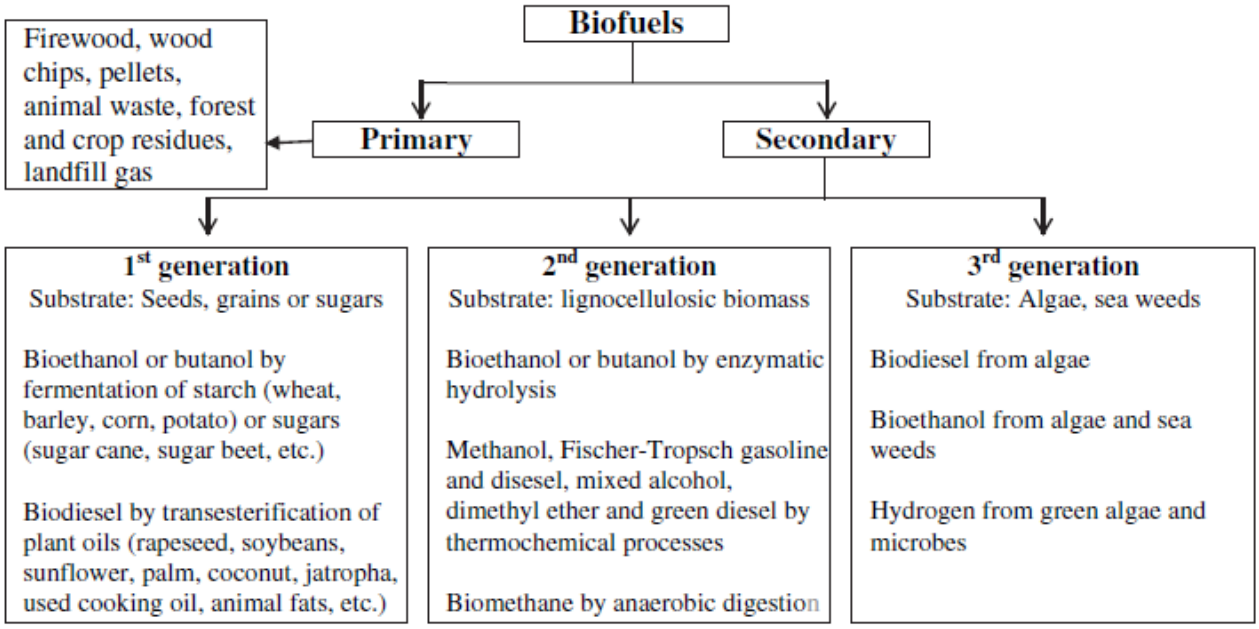


Figure 1. Categorization of biofuels (Nigam & Singh, 2011)

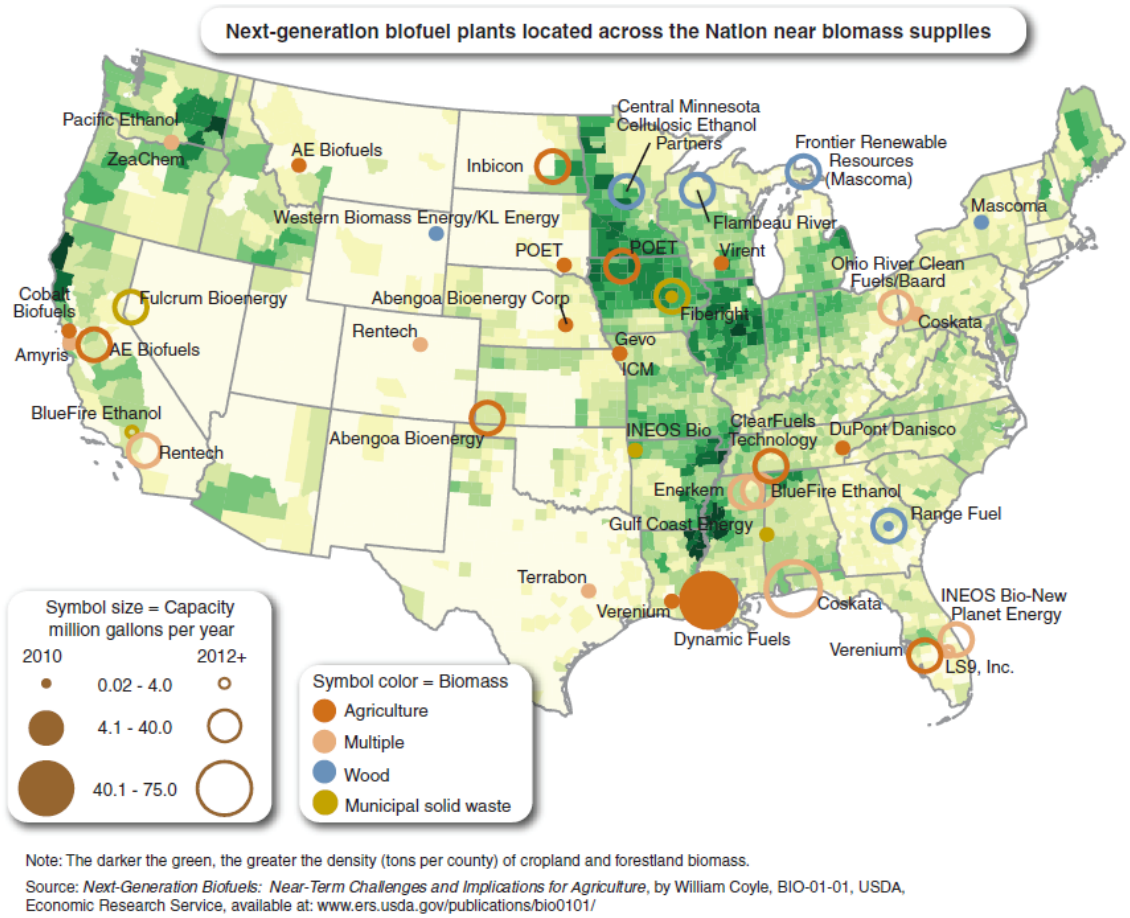
Due to the competition with food using first generation substrates, there has been a significant increase in research on second and third generation biofuels. However, due to challenges, such as pretreatment, it has been difficult for conversion of second and third generation biomass in large-scale biofuel production. Due to simplicity, the first ethanol plants used first generation biomass, mainly maize as their feedstock. Since maize has many uses, especially as a food source, some facilities mix grain sorghum, also known as milo, with the maize since the two crops are comparable in starch content. Thus, the focus of this research was to provide a method for a fast determination of the starch content and fermentation efficiency of mutant grain sorghum for ethanol plants.

1.4.1.1 Ethanol production

Due to technological advances and environmental concerns, there has been extensive research in the area of biofuels. However, the existence of bioethanol dates back to the nineteenth century. In 1859, Edwin Drake was responsible for the discovery of bioethanol, which occurred before petroleum (Songstad et al., 2009). Over time, there were actions that prevented bioethanol from being used in automobiles. Although Henry Ford intended for ethanol to become the main fuel source for the automobile, an imposed tax, set in place during the Civil War, limited the competition between gasoline and ethanol (Songstad et al., 2009; Dimitri & Effland, 2007). After the removal of this tax, gasoline already dominated the market and ethanol faded into the background until the Arabic oil embargo in 1974 (Songstad et al., 2009).

With an increase in oil prices, there became a greater interest in alternative fuel sources, such as ethanol. In 2007, the Energy Independence and Security Act (EISA) stated that the United States needed to increase its biofuel usage to 36 billion gallons by the year 2022 (Coyle, 2010). In order to meet this high demand, new energy crops and advanced technology, with the help of ample funding, were on the verge of discovery.

In Figure 2, there are a large number of next-generation biofuel plants opening across the United States (Coyle, 2010). The map also illustrates the utilization of different types of biomass across the nation. The two most common biofuel sources are from agriculture or biomass from multiple categories. It is also interesting to note the changes between 2010 (filled circles) and 2012 (open circles).



Note: The darker the green, the greater the density (tons per county) of cropland and forestland biomass.

Source: *Next-Generation Biofuels: Near-Term Challenges and Implications for Agriculture*, by William Coyle, BIO-01-01, USDA, Economic Research Service, available at: www.ers.usda.gov/publications/bio0101/

Figure 2. United States map of biofuel plants (Coyle, 2010)

1.4.2 Grain sorghum

Since maize is the driving force behind the food versus fuel debate, other alternative energy crops are necessary to meet the bioethanol demands in the United States. In order to address these bioenergy needs, other types of feedstock, such as grain sorghum, are under investigation. Grain sorghum is a more cost effective crop for semiarid regions in the United States, including the Southwest and the Midwest (Yan et al., 2011). The primary use of grain sorghum is as animal feed, which does not interfere with human food consumption. Similar to maize, sorghum contains approximately 70% starch (Sun et al., 2014). Also, sorghum starch generally contains about 70-80% amylopectin and 20-30% amylose (Yan et al., 2011).

1.4.3 Starch

Due to the importance of starch in grain sorghum, it is vital to have a greater understanding of its physical and chemical properties. Starch primarily consists of two types of D-glucose polymers, called amylose and amylopectin, which make up 98-99% of the dry weight of starch (Copeland et al., 2009; Tester, Karkalas, & Qi, 2004). Amylose is a linear chain with 1-4 alpha-glucan linkages while amylopectin is mostly branched with 1-6 alpha-glucan branch points off of the 1-4 alpha-glucan chain. Figures 3 and 4, pictured below, demonstrate the structural difference between these two components.

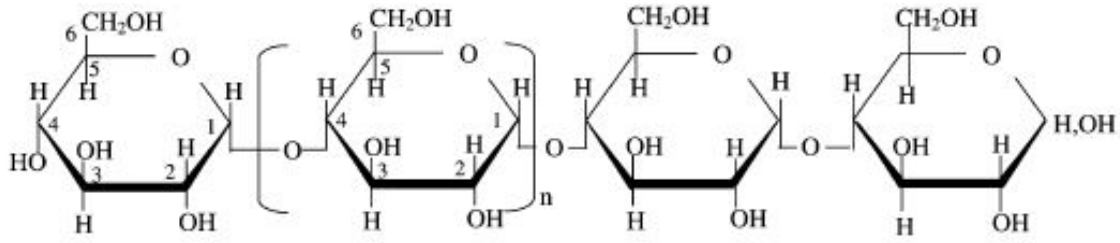


Figure 3. Amylose structure

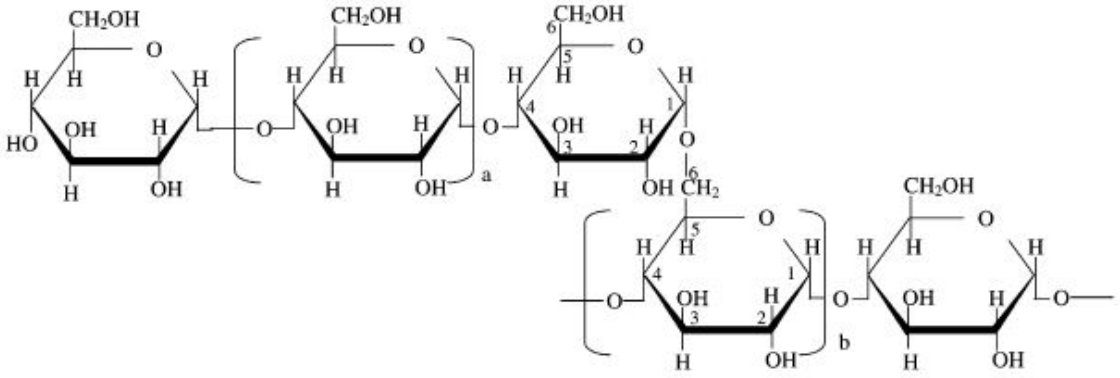


Figure 4. Amylopectin structure

In addition to the difference in the structure, amylose and amylopectin also vary in their size and multitude within the starch granule. The molecular weight of amylose can be as large as 1,000,000, but this size is smaller than amylopectin's molecular weight that is around 10^8 (Delcour & Hoseney, 2010). Tester, Karkalas, and Qi (2004) explains how there is a variation in the amount of amylose and amylopectin based on the type of cereal grain. Generally, the amount of amylose in normal starch does not vary for wheat and sorghum, but can vary for rice starch (Delcour & Hoseney, 2010). For this reason, there is the potential for wheat and sorghum to have "waxy" starches, which contain a higher amylopectin content (Singh et al., 2003).

Amylopectin is important since it dictates the degree of crystallinity within the starch granules. Native starch is only considered semi-crystalline since the amount of amylose creates the amorphous, not crystalline, regions in the granule (Singh, Dartois, & Kaur, 2010). The degree of polymerization (DP) describes the relative number of individual glucose molecules within one polymer. For amylose, the DP is no more than 10,000 and the DP of amylopectin is above 1,000,000 (Copeland et al., 2009).

Due to the branched nature of amylopectin, there are three types of chains that help to distinguish the level of substitution. The C-chain contains the only reducing group and both alpha-1,4 linkages as well as alpha-1,6 branch points. The B-chain contains the two types of linkages, alpha-1,4 and alpha-1,6, but it does not consist of the reducing group. The A-chain is the only chain that is not branched and only involves alpha-1,4 linear linkages (Copeland et al., 2009; Delcour & Hoseney, 2010). The amylopectin and amylose structures of starch are broken down during the dry-grind ethanol process.

1.4.4 Dry-grind ethanol process

In order to understand the details of this process, it is important to first learn about the conversion of starch to ethanol during fermentation. The sorghum starch is converted to simple sugars by using starch-degrading enzymes before yeast converts these monosaccharides into ethanol. A highly profitable by-product, called distillers' dried grain solubles (DDGS), of ethanol production that is left over after the distillation can be sold as animal feed by the ethanol companies. Overall, there are five steps in the dry-grind process: milling, liquefaction, saccharification, fermentation, and distillation (Mosier & Ileleji, 2014).

1.4.4.1 Milling

Once the grain complies with quality standards, hammer mills or roller mills grind the grain kernel down to the appropriate size. In general, hammer mills are more common when the grains do not contain an outer covering, called a husk or hull (Kelsall & Lyons, 2003). Another difference is that the hammer mills operate with a compressive force while the roller mill is more focused on shearing the hull off first before applying a compressive force. In this study, a lab-scale cyclone mill is used to grind the grain into flour, which is similar to a hammer mill. The purpose of milling is to break down the grain into small particles in order to allow for water penetration during cooking (Kelsall & Lyons, 2003).

1.4.4.2 Liquefaction

Prior to cooking, flour is mixed with water to form a slurry. It is important that the water is in contact with all of the flour before gelatinization. Gelatinization occurs when the starch granules swell due to water absorption at certain temperatures, which leads to the loss of its

crystalline structure (Kelsall & Lyons, 2003). With the change in crystallinity, the slurry can undergo further hydrolysis. The addition of starch-degrading enzyme, α-amylase, at this point is necessary to begin the process of breaking down starch. Alpha-amylase only hydrolyzes the α-1,4 glucosidic linkages within the starch polymers, which resulted in dextrins (Delcour & Hoseney, 2010). Dextrins are short chains with a varying number of glucose molecules. According to Zhao et al. (2008), α-amylase behaves as a shear thinning fluid, which helps to reduce the viscosity of the gelatinized starch during liquefaction. The purpose of liquefaction, also known as the cooking step, is to split the hydrogen bonds of starch molecules in order to disrupt the granules (Kelsall & Lyons, 2003).

1.4.4.3 Saccharification

Saccharification is the process of hydrolyzing dextrins into individual glucose molecules. An enzyme that aids in this break down is glucoamylase. Glucoamylase consecutively hydrolyzes the remaining α-1,4 glucosidic linkages. Also, this enzyme completes the hydrolysis at the β-1,6 branch points, but at a slower pace than the α-1,4 bonds (Delcour & Hoseney, 2010). Before the addition of glucoamylase, there must be a decrease in the temperature and the mash must undergo a pH adjustment with use of sulfuric acid or backset stillage (Kelsall & Lyons, 2003). The reason for these changes is because the temperature and pH are based on the essential enzymes in each stage of the process.

Another way of completing this step is to skip the saccharification tank and send the mash directly to the fermenter before adding the enzymes. This alternative method is known as simultaneous saccharification and fermentation (SSF). One of the main reasons that this

procedure is becoming more common is because there is less likelihood of microbial contamination (Bothast & Schlicher, 2005).

1.4.4.4 Fermentation

Finally, fermentation is the step when the glucose is converted into ethanol. One mole of glucose is converted into two moles of ethanol and two moles of carbon dioxide. The main component for this conversion is the microorganism, which in this case, is yeast. Yeast, specifically *Saccharomyces cerevisiae*, consumes the sugars in order to produce ethanol. Also, it is important to note that yeast produces carbon dioxide as a by-product. Thus, weight loss, due to carbon dioxide, can be measured over the 72-hour period of fermentation. The weight loss curve can also be used to calculate the ethanol yield.

1.4.4.5 Distillation

After fermentation, the mixture contains ethanol and the leftover solids and water. In order to separate the ethanol from the fermentation broth, there is a process called distillation. Distillation requires the boiling of the mixture before ethanol and water evaporate from the mixture and proceed through cooling coils into a flask. Water and ethanol both evaporate out because their vaporization temperatures are 100 and 78 °C, respectively (Bothast & Schlicher, 2005). Once distillation is over, more than 4 % water is left in the ethanol, which forms an azeotrope (Swain, 2003; Bothast & Schlicher, 2005). Industry regulations require further processing, with a molecular sieve system, in order to produce pure ethanol (Bothast & Schlicher, 2005). The starch content from wet chemistry analysis and fermentation efficiency,

based on starch content, were used to create near infrared spectroscopy prediction models in this study.

1.4.5 Near infrared (NIR) spectroscopy

NIR is only a small portion of the electromagnetic spectrum. The use of NIR spectroscopy focuses on the wavelength range between 800 to 2500 nm (12500 to 4000 cm-1), which is also known as the near infrared region (Livermore, Wang, & Jackson, 2003). A smaller wavelength in this region will be closer to the high end towards the visible spectrum and a larger wavelength is found closer to the start of the mid-infrared section (Liebmann, Friedl, & Varmuza, 2010; Livermore, Wang, & Jackson, 2003).

One of the main purposes of using the near infrared wavelength range is that spectroscopy in this region allows for the detection as well as determination of different components. The most important factors to examine in cereal grains are moisture, protein content, and starch (Liebmann, Friedl, & Varmuza, 2010). In this study, the primary component of importance was starch, which was found in the initial mutant grain sorghum samples.

This type of spectroscopy is beneficial in multiple ways. The main reason to use NIR is to determine the quality of the materials for process optimization. According to Livermore, Wang, and Jackson (2003), the time for testing takes only a few seconds and there is no harm to the sample, which is more cost efficient. Another advantage is that the quantity of organic material needed for NIR is relatively small. For these reasons, this study was able to use the NIR equipment for a limited amount of material per sample while still allowing the time needed to test a large number of different samples.

There are many different types of spectrometers used for observation in the near infrared region, including: filter photometers, scanning dispersive spectrophotometers, and detector array dispersive spectrophotometers (Livermore, Wang, & Jackson, 2003). For this research, a Fourier transform (FT) NIR spectrometer was used. Figure 5 below demonstrates how this spectrometer works once the light source hits the Michelson inferometer. The three parts that make up the inferometer are the fixed mirror, the moving mirror, and the beam splitter. The beam splitter allows some of the light to be reflected on the fixed mirror while there is a transmission of the rest of the light to the moving mirror. As the light source is focused toward to the beam splitter from the fixed and moving mirrors, there is interference. The moving mirror shifts back and forth, which leads to a variance in the interference. This difference in the interference causes the intensity of light hitting the detector to change over time. FT-NIR spectrometers are beneficial due to high optical resolution and high ratios of signal to background noise (Livermore, Wang, & Jackson, 2003).

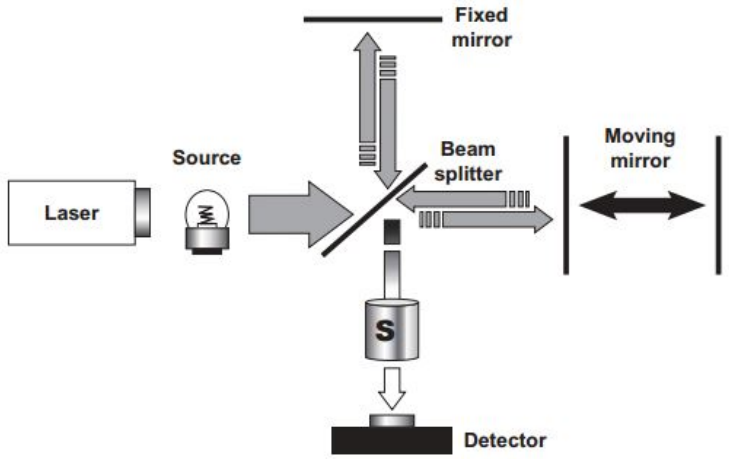


Figure 5. Diagram of FT-NIR spectrometer (Livermore, Wang, & Jackson, 2003)

1.4.5.1 Previous work using NIR

Pohl and Senn (2011) examined 10 wheat varieties, 24 rye varieties, and 6 triticale varieties, for a total of 480 samples, by using a diode array spectrometer for whole grain samples and a dispersive near infrared monochromater for flour samples. This study used partial least squares regression (PLS) for their NIR model and found that it was only an adequate method for starch analysis, not fermentation, due to the inability to find final ethanol yields. Thus, NIR would be beneficial for initial testing of incoming grain and for development of better breeding programs. Similarly, Kim and Williams (1990) investigated wheat, barley, and maize ground samples with a Pacific Scientific research composition analyzer, a Pacific Scientific feed-quality analyzer, and a DICKEY-john grain analysis computer. As a result of this study, there was only a strong predictability for protein and energy, not starch.

Cozzolino, Roumeliotis, and Eglinton (2013) observed the performance of 130 barley whole grain samples using a rapid visco analyzer (RVA) and near infrared reflectance (NIR) spectroscopy. With use of partial least squares (PLS) regression, this study found that RVA and NIR could be useful tool for discovering the starch pasting properties for genotype selection. Another study conducted by Lin et al. (2014) also looked at 277 samples of whole grain barley, but their focus was on protein content. As opposed to only using one type of calibration like the previous studies, this paper includes the use of partial least squares (PLS), least squares support vector machine regression (LSSVR), and radial basis function (RBF) neural network. However, LSSVR was the most accurate multivariate calibration model by establishing a strong correlation between the first derivative spectra and the protein content. These experiments also demonstrated how NIR was an accurate method for prediction of whole grain barley protein content.

Additionally, Bao, Cai, and Corke (2001) utilized RVA and NIR in their investigation of 162 ground rice samples. A modified PLS regression model correlated the first derivative spectra and the scatter correction for the standard normal variance and de-trend (SNVD). The RVA pasting results revealed that there were strong relationships in the scatter plots and high coefficients of determination for setback (SB) and breakdown (BD), which are significant parameters in the determination of high quality in rice. Thus, NIR has the potential to predict these values, which will be beneficial for rice breeding programs. The study also showed how NIR could be used with rice flour samples to accurately measure starch values for SB, BD, the apparent amylose content, and the gelatinization peak temperature.

Orman and Schumann Jr. (1991) intended to predict composition, such as starch, oil, and protein, of 156 out of 500 whole grain maize samples through NIR calibration models using a monochrometer infrared spectrophotometer. The best and lowest error for prediction was found for the diffuse reflectance of the grain flour in comparison to the whole grain spectra. Thus, the authors concluded that the flour diffuse reflectance method was the best for the prediction of starch, protein, and oil. Hao, Thelen, and Gao (2012) examined 222 samples of maize flour to determine the potential for using NIR and the bootstrapping method to estimate ethanol yields. The average value of root mean square error of prediction (RMSEP) for the reduced wavelength range and original spectra validation was 0.56%. The bootstrapping method with optimum wavelength intervals could be helpful in determining compositional factors contributing to the ethanol yield for maize in the future.

Due to the large amount of literature published in the past ten years, NIR is a technological tool that is gaining popularity for prediction of compositional analysis. However, there are very limited studies on the prediction for starch content and fermentation efficiency for

sorghum, especially mutant grain sorghum. Thus, there is research needed to develop rapid methods for the prediction of starch and ethanol fermentation efficiency of mutant grain sorghum using NIR technology.

References

Assefa, Y., Roozeboom, K. L., Thompson, C., Schlegel, A., Stone, L., & Lingenfelser, J. (2013). Chapter 1 – Introduction. *Corn and Grain Sorghum Comparison: All Things Considered* (pp. 1-2). Waltham, MA: Academic Press.

Bao, J. S., Cai, Y. Z., & Corke, H. (2001). Prediction of rice starch quality parameters by near-infrared reflectance spectroscopy. *Journal of Food Science*, *66*(7), 936-939.

Bothast, R. J., & Schlicher, M. A. (2005). Biotechnological processes for conversion of corn into ethanol. *Applied microbiology and biotechnology*, *67*(1), 19-25.

Copeland, L., Blazek, J., Salman, H., & Tang, M. C. (2009). Form and functionality of starch. *Food Hydrocolloids*, *23*(6), 1527-1534.

Coyle, W. (2010). Next-generation biofuels: Near-term challenges and implications for agriculture. *Amber Waves, 8*(2), 20-27.

Cozzolino, D., Roumeliotis, S., & Eglinton, J. (2013). Exploring the use of near infrared (NIR) reflectance spectroscopy to predict starch pasting properties in whole grain barley. *Food Biophysics*, *8*(4), 256-261.

Crop production 2015 summary (2016). No. 1057-7823). Washington, D.C.: USDA-NASS.

Delcour, J.A. & Hoseney, R.C. (2010). Chapter 2 - Starch. *Principles of cereal science and technology* (3rd ed., pp. 23-51). St. Paul, MN, U.S.A.: AACC International, Inc.

Dimitri, C., & Effland, A. (2007). Fueling the automobile: an economic exploration of early adoption of gasoline over ethanol. *Journal of Agricultural & Food Industrial Organization*, *5*(2).

F.O. Licht. (2014). World Ethanol & Biofuels Report. *12(15)*. Ratzeburg, Germany: Informa UK Ltd.

Guo, M., Song, W., & Buhain, J. (2015). Bioenergy and biofuels: History, status, and perspective. *Renewable and Sustainable Energy Reviews*, *42*, 712-725.

Hao, X., Thelen, K., & Gao, J. (2012). Prediction of the ethanol yield of dry-grind maize grain using near infrared spectroscopy. *Biosystems engineering*, *112*(3), 161-170.

Kelsall, D.R. & Lyons, T.P. (2003). Chapter 2 - Grain dry milling and cooking procedures: extracting sugars in preparation for fermentation. In K. A. Jacques, T. P. Lyons & D. R. Kelsall (Eds.), *The alcohol textbook* (4th ed., pp. 9-21). Thrumpton, Nottingham, U.K.: Nottingham University Press.

Kim, H. O., & Williams, P. C. (1990). Determination of starch and energy in feed grains by near-infrared reflectance spectroscopy. *Journal of Agricultural and Food Chemistry*, *38*(3), 682-688.

Koizumi, T. (2015). Biofuels and food security. *Renewable and Sustainable Energy Reviews, 52*, 829-841.

Liebmann, B., Friedl, A., & Varmuza, K. (2010). Applicability of near-infrared spectroscopy for process monitoring in bioethanol production. *Biochemical Engineering Journal*, *52*(2), 187-193.

Lin, C., Chen, X., Jian, L., Shi, C., Jin, X., & Zhang, G. (2014). Determination of grain protein content by near-infrared spectrometry and multivariate calibration in barley. *Food chemistry*, *162*, 10-15.

Liu, L., Maier, A., Klocke, N. L., Yan, S., Rogers, D. H., Tesso, T., & Wang, D. (2013). Impact of deficit irrigation on sorghum physical and chemical properties and ethanol yield. *Transactions of the ASABE*, *56*(4), 1541-1549.

Livermore, D., Wang, Q., & Jackson, R. S. (2003). Chapter 12 - Understanding near infrared spectroscopy and its applications in the distillery. In K. A. Jacques, T. P. Lyons & D. R. Kelsall (Eds.), *The alcohol textbook* (4th ed., pp. 145-170). Thrumpton, Nottingham, U.K.: Nottingham University Press.

Mosier, N. S., & Ileleji, K. E. (2014). How fuel ethanol is made from corn. *Bioenergy: Biomass to Biofuels*, 379.

Nigam, P. S., & Singh, A. (2011). Production of liquid biofuels from renewable resources. *Progress in energy and combustion science*, *37*(1), 52-68.

Orman, B. A., & Schumann Jr, R. A. (1991). Comparison of near-infrared spectroscopy calibration methods for the prediction of protein, oil, and starch in maize grain. *Journal of agricultural and food chemistry*, *39*(5), 883-886.

Pohl, F., & Senn, T. (2011). A rapid and sensitive method for the evaluation of cereal grains in bioethanol production using near infrared reflectance spectroscopy. *Bioresource technology*, *102*(3), 2834-2841.

Shewayrga, H., Sopade, P. A., Jordan, D. R., & Godwin, I. D. (2012). Characterisation of grain quality in diverse sorghum germplasm using a Rapid Visco-Analyzer and near infrared reflectance spectroscopy. *Journal of the Science of Food and Agriculture*, *92*(7), 1402-1410.

Singh, J., Dartois, A., & Kaur, L. (2010). Starch digestibility in food matrix: a review. *Trends in Food Science & Technology*, *21*(4), 168-180.

Singh, N., Singh, J., Kaur, L., Sodhi, N. S., & Gill, B. S. (2003). Morphological, thermal and rheological properties of starches from different botanical sources. *Food Chemistry*, *81*(2), 219-231.

Solomon, B. D., Banerjee, A., Acevedo, A., Halvorsen, K. E., & Eastmond, A. (2015). Policies for the sustainable development of biofuels in the pan American region: a review and synthesis of five countries. *Environmental management*, *56*(6), 1276-1294.

Songstad, D., Lakshmanan, P., Chen, J., Gibbons, W., Hughes, S., & Nelson, R. (2011). Historical perspective of biofuels: learning from the past to rediscover the future. In *Biofuels* (pp. 1-7). Springer New York.

Su, Y., Zhang, P., & Su, Y. (2015). An overview of biofuels policies and industrialization in the major biofuel producing countries. *Renewable and Sustainable Energy Reviews*, *50*, 991-1003.

Sun, Q., Han, Z., Wang, L., & Xiong, L. (2014). Physicochemical differences between sorghum starch and sorghum flour modified by heat-moisture treatment. *Food chemistry*, *145*, 756-764.

Swain, R.L. (2003). Chapter 23 - Development and operation of the molecular sieve: an industry standard. In K. A. Jacques, T. P. Lyons & D. R. Kelsall (Eds.), *The alcohol textbook* (4th ed., pp. 337-341). Thrumpton, Nottingham, U.K.: Nottingham University Press.

Tester, R. F., Karkalas, J., & Qi, X. (2004). Starch—composition, fine structure and architecture. *Journal of Cereal Science*, *39*(2), 151-165.

Wang, D., Bean, S., McLaren, J., Seib, P., Madl, R., Tuinstra, M., ... & Zhao, R. (2008). Grain sorghum is a viable feedstock for ethanol production. *Journal of industrial microbiology & biotechnology*, *35*(5), 313-320.

Yan, S., Wu, X., MacRitchie, F., & Wang, D. (2009). Germination-Improved Ethanol Fermentation Performance of High-Tannin Sorghum in a Laboratory Dry-Grind Process 1. *Cereal chemistry*, *86*(6), 597-600.

Yan, S., Wu, X., Bean, S. R., Pedersen, J. F., Tesso, T., Chen, Y., & Wang, D. (2011). Evaluation of waxy grain sorghum for ethanol production.

Zhao, R. Z., Bean, S. R., Wang, D., & Wu, X. (2008). Assessing fermentation quality of grain sorghum for fuel ethanol production using rapid visco analyzer. In *2008 Providence, Rhode Island, June 29–July 2, 2008* (p. 1). American Society of Agricultural and Biological Engineers.

Zhou, X., Yang, Z., Haughey, S. A., Galvin-King, P., Han, L., & Elliott, C. T. (2015). Classification the geographical origin of corn distillers dried grains with solubles by near infrared reflectance spectroscopy combined with chemometrics: A feasibility study. *Food chemistry*, *189*, 13-18.

Chapter 2 - Materials and Methods

2.1 Mutant grain sorghum

A total of 109 mutant grain sorghum samples were obtained from Plant Stress and Germplasm Development Unit of USDA-ARS, Lubbock, TX. Mutant sorghums were used for this study since the genotype was found to have a significant correlation to ethanol yield and fermentation efficiency (Wu et al., 2007). The samples were cleaned by removing foreign materials and small broken kernels. The cleaned samples were milled into flour through a 0.5 mm screen by an Udy cyclone sample mill (Udy, Ft. Collins, CO).

2.2 Chemical composition of mutant grain sorghum

2.2.1 Moisture content

The moisture content was adapted from the AACC 44-15.02 standard moisture air-oven method (AACC International, 1999). All of the metal dishes, along with their lids, were placed in the oven at 130 °C for 2 h for sterilization. Two grams of each mutant sorghum flour sample, with one replicate, were placed in metal dishes. Each sample was covered with the lid until the samples were ready to go into the oven. The lids were removed and put underneath the corresponding dish before being placed in the oven. The samples were left in the oven at 130 °C for 3 h. Once the samples were done in the oven, they were quickly put in an airtight desiccator for cooling. The dishes were immediately weighed, after reaching room temperature, for the most accurate reading.

2.2.2 Total starch content

Megazyme starch assay kits with two starch enzymes, α-amylase and amyloglucosidase, (Megazyme International Ltd., Ireland) were used for the starch content measurement following the standard AACC 76-13.01 method (AACC International, 1999). The Mega-Calc software (Megazyme International Ltd., Ireland) determined the total starch in the grain sorghum flour. This software uses the absorbance data from experimentation and the moisture content to calculate the total amount of starch on a dry basis (db).

2.2.3 Protein content

A 2400 CHNS/O Series II Analyzer (PerkinElmer Inc., Shelton, CT) was used to determine the elemental composition of mutant grain sorghum, including: carbon, hydrogen, nitrogen, sulfur, and oxygen. The operating mode was CHNS so the combustion and reduction temperatures were set at 975 °C and 500 °C, respectively. With the oxygen valve turned off, the instrument runs four blanks followed by ~3.5 mg of sulfamic acid sulfur conditioning reagent (also known as a conditioner) measured into an aluminum cup before running a blank and one more conditioner in the autosampler. Next, three 2.0-2.5 mg cystine standards (also known as k-factors) were weighed in aluminum cups and placed in the autosampler with the oxygen valve on. Two consecutive k-factors must pass with all of the elements within an acceptable range and with only a small variation between the two standards before running four blanks with the oxygen off. Then, the oxygen needs to be turned back on before continuing with the samples. The blanks need to finish before placing a sample (2.0 – 2.5 mg) into the autosampler. All of the samples were weighed in aluminum cups using an AD 6 Autobalance (PerkinElmer Inc., Shelton, CT). The amount of nitrogen was multiplied by 6.25 to calculate the protein content.

2.3 Morphological properties of whole grain sorghum

The morphology of mutant sorghum was studied using the scanning electron microscope (SEM) with an accelerating potential of 5 kV (Hitachi S-3500N, Hitachi Science Systems, Ltd., Japan). Only six of the ten selected with the highest and lowest fermentation efficiencies were analyzed for SEM imaging. In preparation, a few grains were selected from each sample and were cut with a razor blade. This method was used to provide control when cutting the grain into cross-sections. Grain fragments with flat surfaces were selected with tweezers and placed on a sticky carbon film. Before inserting the samples into the microscope, a Desk II sputter coater covered each sample with a layer of gold and palladium under vacuum conditions (Denton Vacuum, LLC, Moorestown, NJ). The purpose of this step is to increase the signal sensitivity and create better resolution images. After the sputter coating, four samples at a time were loaded into the scanning electron microscope for imaging. The samples images were taken at magnifications of 500x and 5000x for the opaque regions of the endosperm.

2.4 Thermal properties

Differential scanning calorimetry (DSC) machine (DSC-Q200, TA Instruments Inc., New Castle, DE) was used to determine the thermal properties, such as the transition temperatures and changes in enthalpy. The method followed the same process as described by Wu et al. (2007) and Wang et al. (2008). Five samples with the lowest and highest fermentation efficiencies were selected for this analysis. Approximately 10 mg of flour was weighed in a stainless steel pan before adding ~35 μL of distilled water to the pan. The pans were sealed and mixed well before being kept in a 4°C refrigerator overnight. A sealed empty stainless steel pan was used as a

reference. All samples were held at 20°C for 1 min and then heated from 20 to 120 °C at 10°C/min.

2.5 Pasting properties

A Rapid Visco Analyzer (RVA) (RVA-3c, Newport Scientific Ltd., Warriewood, Australia) was used to look at the pasting properties of grain sorghum flour. The same five samples with the lowest and highest fermentation efficiencies, also used with DSC, were selected for this analysis. The method was based on the approved AACC standard method (76-21.01) for wheat and rye flour or starch (AACC International, 1999). Grain sorghum flour (3.5 g, 14% moisture content) was added to approximately 25 mL of distilled water in a canister. The slurry was mixed at 50 °C for 1 min before the temperature increased to 95 °C. The paste was held at 95 °C for 2.5 min before decreasing back to 50 °C, where the slurry was held for 2 min. For a uniform mixture, the speed started at 960 revolutions per minute (rpm) before decreasing to 160 rpm after 10 sec, which was the speed for the remaining duration of the test.

2.6 Ethanol fermentation

Only 39 of the original 109 samples were used for fermentation. The process of liquefaction started by adding 100 mL broth to 30 g (db) of grain sorghum flour in a 250 mL flask. Twenty microliters of α-amylase, called Liquozyme (Novozymes, Franklinton, NC) and 0.1 g of potassium phosphate, were stirred into 100 mL distilled water, preheated to 70 °C. This fermentation broth was gradually added to each flask to ensure that all of the flour was fully saturated using a spatula. This step is important because α-amylase will begin to break down the starch into glucose. The flasks were then carefully placed in a water bath rotary shaker at 70 °C

and 180 rpm, which began the process of gelatinization. Immediately after putting the flasks in the rotary shaker, the temperature was increased to 90 °C, until it reached 90 °C, and then, the temperature was reduced to 86 °C for 1 h.

Once the flasks cooled down to room temperature, the pH was adjusted from ~ 5.0-6.0 to 4.2 using 2.0 N HCl. The activated yeast was prepared with approximately 40 min left of pH adjustment. One gram of activated dry yeast was weighed into an autoclaved 125 mL flask before adding 19 mL of precultured broth under the purifier class II biosafety cabinet (Delta Series, Labconco Corporation, Kansas City, MO). Then, the activated yeast flask were placed in the incubator shaker for 30 min at 200 rpm and 38 °C. Under the biosafety hood, 100 µL of glucoamylase, also known as Spirizyme (Novozymes, Franklinton, NC), 0.3 g yeast extract, and 1.0 mL activated yeast, once it is done in the shaker, were inserted to each flask. For fermentation, all the flasks were placed into the incubator shaker at 30 °C and 150 rpm for a total time of 72 h. Weights were recorded for each flask at certain time increments, which indicated the conversion to ethanol through loss due to carbon dioxide, a by-product of fermentation. Thus, weight loss and ethanol yield curves can be expressed by this data.

After fermentation, contents were washed from each flask using distilled water and were poured into 500 mL distillation flasks with two to three drops of antifoam. These flasks were placed on a heating element and connected to a cooling coil apparatus. The heating element was turned on high to boil the contents and cause the evaporation of ethanol and water. The ethanol-water mixture was collected in a 100 mL volumetric flask. The remaining contents, called dried distillers' grain solubles (DDGS), from the distillation flask were transferred to glass jars and placed in a freezer at -20 °C. High Liquid Pressure Chromatography (HPLC) was used to determine the actual ethanol yield using an established standard curve. The fermentation

efficiency was calculated by dividing the actual ethanol yield by the theoretical yield. Based on these results, ten samples with the lowest and highest fermentation efficiencies were selected for the investigation of the differences in the morphological, thermal, and pasting properties.

2.7 Chemical composition of DDGS

The glass jars with the DDGS were removed from the freezer and thawed until the contents turned into a liquid form. Then, the jars were placed in a constant temperature cabinet oven (Blue M Electric Company, Blue Island, IL) at 49 °C for 48 h. The dried DDGS was ground with a pestle in a 750 mL porcelain mortar (Cat. No. 60325, CoorsTek, Golden, CO) into smaller pieces before undergoing further grinding in a Micro-Mill (Bel-Art Products Scienceware, Pequannock, NJ). The ground samples were then analyzed to determine the moisture content based on the same adaptation of the oven-dry AACC 44-15.02 standard procedure for the ground grain sorghum (AACC International, 1999). The DDGS samples were sent to the Kansas State University Department of Animal Science and Industry Analytical Laboratory (Manhattan, KS) for crude protein content analysis. Similar to the sorghum flour, the DDGS starch content was determined with the Megazyme total starch assay kit (Megazyme International Ltd., Ireland).

2.8 Near infrared spectroscopy

Prior to milling, whole grain samples were scanned using Antaris II Fourier Transform Near Infrared (FT-NIR) Analyzer spectrophotometer (Thermo Fisher Scientific Inc., Madison, WI). This process was also repeated after milling for the flour samples. Whole grain samples (~20 g) and flour samples (~10 g) were loaded into a circular sample cup. For each sample, the

cup was very lightly shaken for a more even distribution before loading. The spectrophotometer was operated at a resolution of 4 cm^{-1} and a wavelength range of 4,000-10,000 cm^{-1}. The spectrum for each sample was obtained as the average of 32 scans and the data was recorded as log (1/R), when R is reflectance. In modeling, the samples were divided into calibration and validation in a ratio of 4:1. The spectroscopic method and data used TQ Analyst software (Thermo Fisher Scientific Inc., Waltham, MA).

2.8.1 Model Parameters

The NIR data was analyzed in the TQ Analyst software (Thermo Fisher Scientific Inc., Waltham, MA). The method for the quantitative analysis was partial least squares (PLS) and multiplicative signal correction (MSC) was selected for the pathlength type. The modeling was completed for the starch content and the fermentation efficiency. The number of factors for all of the PLS models did not exceed a total of 4 so the model would not become too complex.

2.8.2 Model Development: Starch Content and Fermentation Efficiency

First, all of the starch values were included in the model to gain an understanding of the initial standard error of calibration (SEC), standard error of prediction (SEP), and their correlation coefficients. Then, outliers were determined based on the differences between the calculated values from the software and actual values from experimentation. Outliers were removed individually and the model was recalibrated each time. Once the initial model, prior to removal of outliers, was fully calibrated, the optimum wavelength range was selected to remove the noise. Finally, the last two models represented the first derivative of the entire spectra and

the reduced spectra for the selected wavelength range. The same method was repeated for the other parameter, fermentation efficiency.

References

AACC International. (November 3, 1999). AACC International Approved Methods of Analysis, 11th Ed. Method 76-21.01. General Pasting Method for Wheat or Rye Flour or Starch Using the Rapid Visco Analyser. AACC International, St. Paul, MN, U.S.A. http://dx.doi.org/10.1094/AACCIntMethod-76-21.01

AACC International. (November 3, 1999). AACC International Approved Methods of Analysis, 11th Ed. Method 44-15.02. Moisture—Air-Oven Methods. AACC International, St. Paul, MN, U.S.A. http://dx.doi.org/10.1094/AACCIntMethod-44-15.02

AACC International. (November 3, 1999). AACC International Approved Methods of Analysis, 11th Ed. Method 76-13.01. Total Starch Assay Procedure (Megazyme Amyloglucosidase/α-Amylase Method). AACC International, St. Paul, MN, U.S.A. http://dx.doi.org/ 10.1094/AACCIntMethod-76-13.01

Wu, X., Zhao, R., Bean, S. R., Seib, P. A., McLaren, J. S., Madl, R. L., ... & Wang, D. (2007). Factors impacting ethanol production from grain sorghum in the dry-grind process 1. *Cereal Chemistry*, *84*(2), 130-136.

Chapter 3 - Results and Discussion

3.1 Chemical composition and morphological properties

Moisture content of the mutant sorghum was between 7% and 10% (wet basis, wb) with average of 8.5% (wb) (Appendix A, Table A.2). Starch content for the mutant sorghum ranged from around 63% to 71% (dry basis, db) with average of 67.3% (db), which is a lower result than in the literature (Table A.3). The average and standard deviation compositional data for each sample can be found in Appendix A. According to Sun et al. (2014), grain sorghum normally contains around 70% starch. The amount of total starch can vary due to irrigation and environmental conditions (Liu et al., 2013; Yan et al., 2009).

The standard conversion factor for crude protein content was calculated by multiplying the percentage of nitrogen by 6.25. Protein values for the mutant sorghum ranged from 11 to 18% (Table A.4). However, it is important to note that any values above 16% crude protein are not comparable to other results (Wong et al., 2010; Wu et al., 2010). The reason for this discrepancy could mean that mutant grain sorghum has a potential to reach a higher protein content than other varieties. While a high protein content is generally related to a lower starch content, the protein digestibility, which is lower for grain sorghum with tannins, can be positively correlated to the fermentation efficiency (Wang et al., 2008; Wu et al., 2007). A summary of the composition analysis of the mutant grain sorghum can be found in Table A.1.

Figure 6 shows SEM images of the mutant sorghum endosperm for the low and high fermentation efficiency samples at 500x and 5000x magnification. The use of these images emphasized the morphological differences between the samples in the opaque region of the grain sorghum kernel.

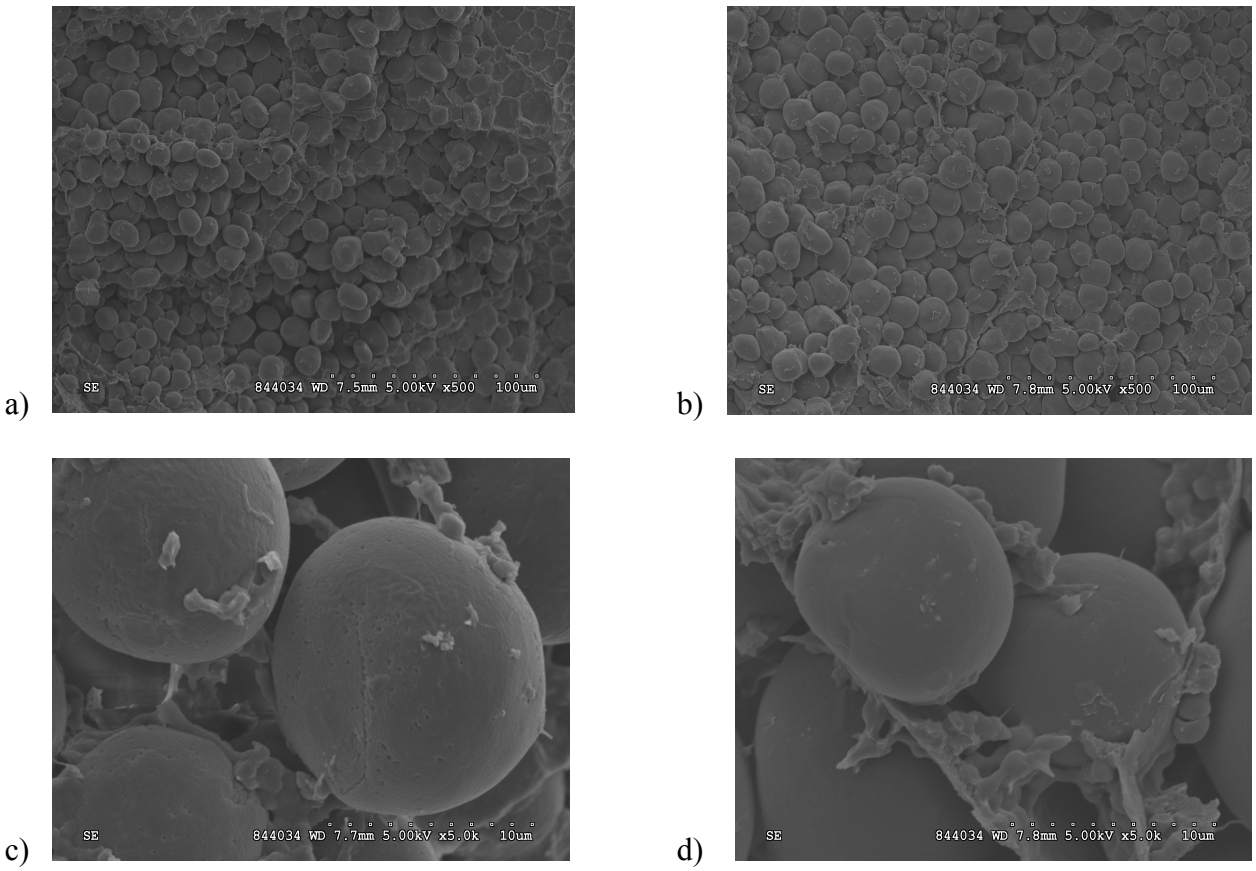


Figure 6. a) Low efficiency sample in opaque region at 500x magnification; b) High efficiency sample in opaque region at 500x magnification; c) Low efficiency sample in opaque region at 5000x magnification; d) High efficiency sample in opaque region at 5000x magnification

SEM images at 500x magnification showed the higher efficiency samples with a weaker protein matrix than samples with low fermentation efficiency (Figures 6a & 6b). This result is in agreement with previous studies that a strong protein matrix can reduce the efficiency of starch hydrolysis (Wu et al., 2007). At a higher magnification of 5000x, the opaque regions demonstrated larger starch granules for low efficiency samples and the high efficiency samples also depicted the protein matrix between the starch granules (Figures 6c & 6d). There were many small holes seen on the surface of the lower efficiency samples, which is similar to the germinated grain sorghum samples (Yan et al., 2009). Ai et al. (2011) describes how these tiny

holes could be from amylase hydrolysis, and only the grain sorghum, not maize, exhibited these small pinholes. However, Yan et al. (2009) also states that the germinated grain sorghum had higher fermentation efficiencies, which differs from these results.

3.2 Thermal and pasting properties

DSC peak temperature values for the ten mutant sorghum samples (5 lowest and 5 highest fermentation efficiency) ranged from 76.1-78.5 °C with an average of 77.5 °C. The average peak temperatures for the lowest and highest fermentation efficiency samples had no observed difference. This result could be due to the similarities in the mutant grain sorghum composition. The enthalpies of gelatinization varied from 4.8 to 6.6 J/g with an average of 5.4 J/g for low and high efficiency samples, respectively. Liu et al. (2013) explains how the higher peak temperatures are correlated with a lower fermentation efficiency and more energy is required to start the process of gelatinization. Figure 7 depicts all of the DSC curves for the high and low fermentation efficiency samples from 60 to 85 °C. These curves showed the temperature required for gelatinization of the starch in each sample.

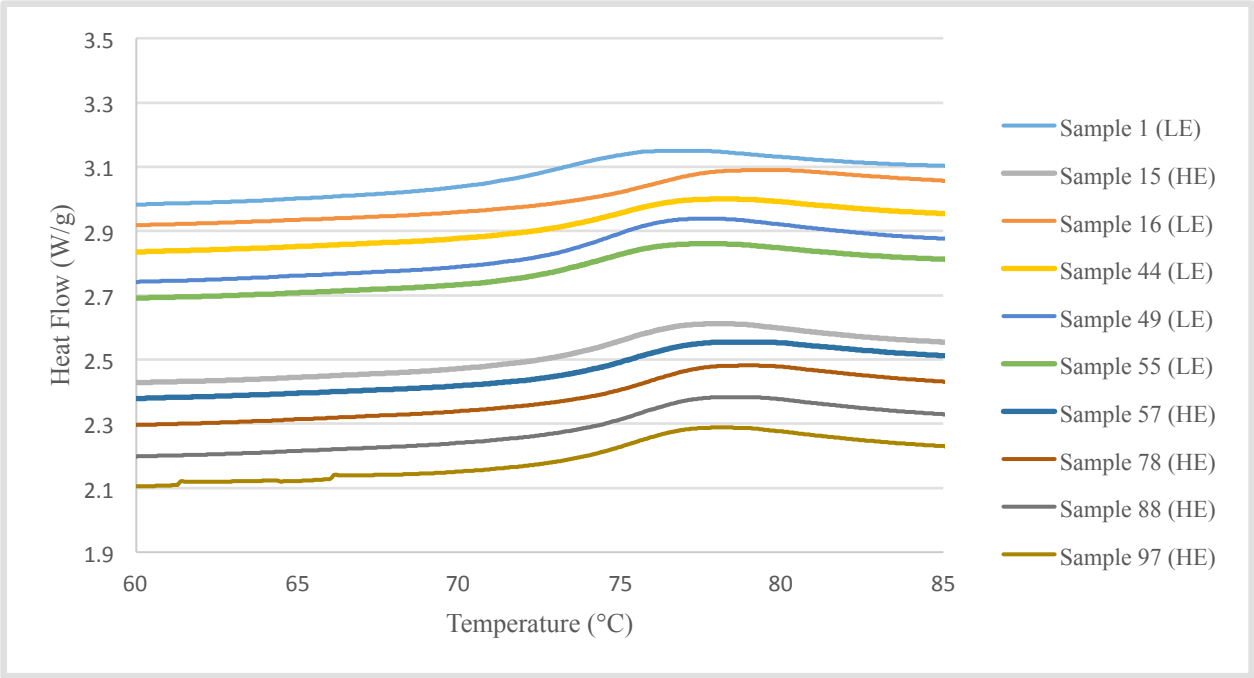


Figure 7. DSC curves for low (LE) and high (HE) fermentation efficiency samples

All of the samples had slight amylose-lipid complex peaks from 100.0 to 101.2 °C. However, both of the averages for the low and high efficiency samples were about 100 °C, with a standard deviation of 0.25 J/g and 0.42 J/g, respectively. The average enthalpy value (0.25 J/g) of the amylose peaks for the lower fermentation efficiency samples demonstrated how the presence of amylose requires more energy for gelatinization initiation. On the other hand, with a high fermentation efficiency, the average amylose enthalpy was 0.20 J/g, indicating that it was easier for the gelatinization process to begin. Other research on waxy starches, primarily composed of amylopectin, relates a lower amylose-lipid complex with a higher fermentation efficiency (Wu et al., 2007).

All of the RVA results can be found in Table A.5 of Appendix A. The peak viscosities ranged from 1141-1759 cP with an average of 1522 cP and peak times ranged from 6.3 to 6.8 min (an average of 6.4 min). The final viscosities were between 2328 and 3963 cP (an average

of 3218 cP) and the setback values were from 1244 to 2192 cP with an average of 1842 cP. The average peak pasting temperature was 74.1 °C, with averages of 79.4 and 68.8 °C for the lowest and highest fermentation efficiency samples, respectively. The results were expected because the samples with high efficiency have lower peak pasting temperatures, indicating the high efficiency samples are easier to cook and hydrolyze. These values are in agreement with Shewayrga et al. (2011) except that the range for the peak pasting time was less for this study and the pasting temperatures were higher than their reported study. The peak viscosity values were also high, but there was a positive correlation between the starch content and viscosity, which indicated that a sample with a higher starch content are generally more viscous (Shewayrga et al., 2011). Based on the viscosity values, sample 1 has a low starch content, which correlates to the lowest peak viscosity (1141 cP), and sample 78 has the highest starch content and a high peak viscosity of 1718 cP. Figures 8 and 9 illustrate the RVA curves based on the ten samples with the lowest and highest fermentation efficiencies.

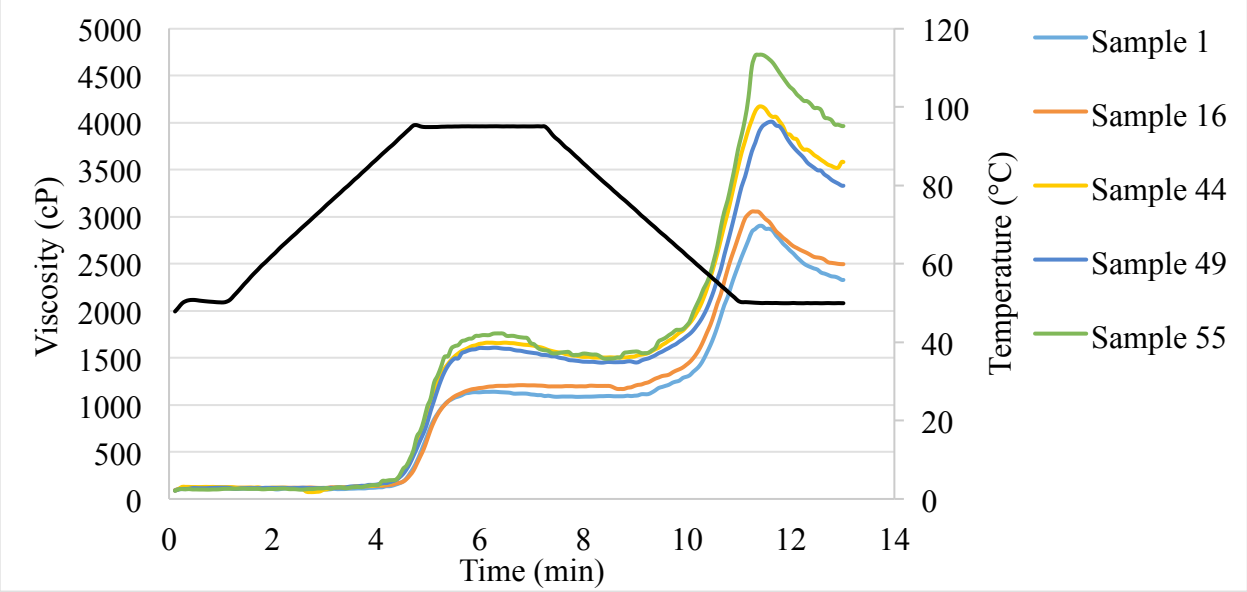


Figure 8. RVA curves for low fermentation efficiency samples

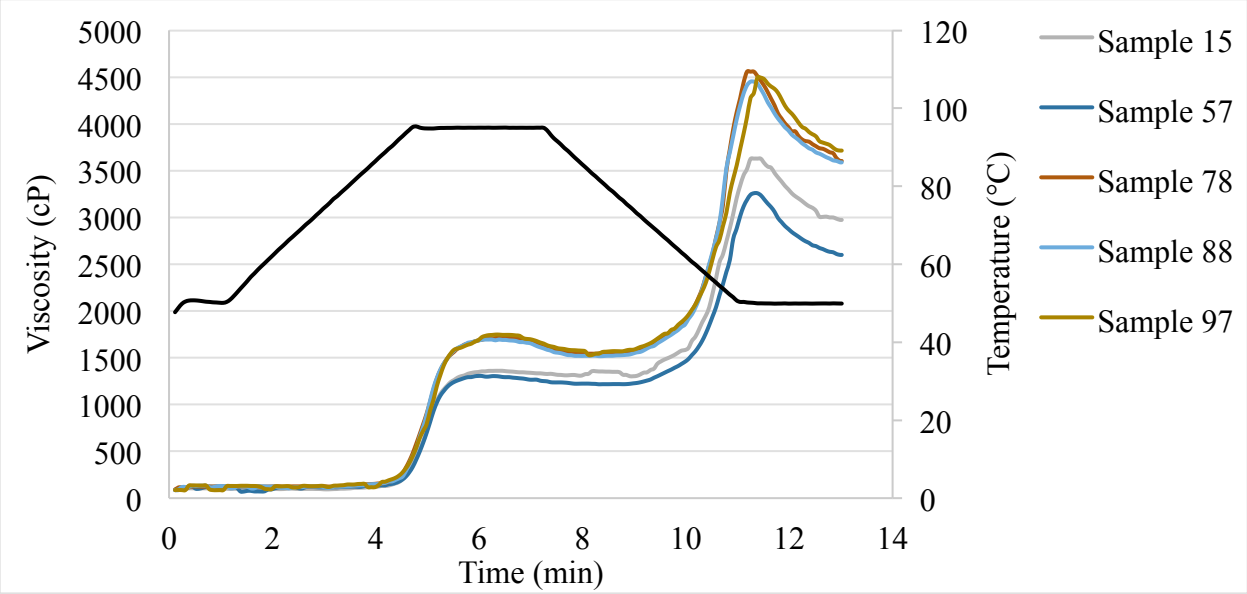


Figure 9. RVA curves for high fermentation efficiency samples

The first peak appeared around 5 min is representative of the peak viscosity related to starch gelatinization. The pasting temperature is the temperature before the sudden increase in viscosity, which happened around 4 min. The viscosity reaches a maximum around 11 min, which is known as the final viscosity. Setback is measured as the difference between the final and peak viscosity values. According to Zhao et al. (2008), starch content as well as ethanol yield had a strong relationship with final viscosity and setback. The results of this study showed that sample 1, with the lowest ethanol yield, also had the lowest final viscosity and setback of all ten samples. For the highest efficiency samples, sample 97 had a higher ethanol yield, the second highest final viscosity, and the highest setback. However, not all of these samples supported this correlation. These results may indicate other factors may also affect the pasting viscosity and more research needs to be completed before drawing conclusions.

3.3 Ethanol yield and fermentation efficiency

The actual ethanol yields varied from ~12-15 % (v/v) and an average of 13.4% (v/v) with a standard deviation of ~4.0% (Appendix A, Table A.6). These results are comparable to the ethanol yields (ranging from ~11-17%, v/v) found in a previous study that added sweet sorghum juice, instead of water, to different amounts of grain sorghum flour (Appiah-Nkansah et al., 2015). Furthermore, the fermentation efficiency was between 88-95%, with an average of 91% and a standard deviation of 2.2% (Table A.6). Yan et al. (2011) reported that the fermentation efficiency of waxy grain sorghum were in the range between 86-92%. The high ethanol fermentation efficiency indicates that the mutant sorghum has great potential for biofuel production.

Figure 10 expresses the correlation between ethanol yield and starch content. Based on the linear regression trend line, the R^2 value was around 0.70, which means that the line of best fit represents 69.5% of the data points. Thus, there was a positive correlation between the ethanol yield and the starch content. This linear relationship between the total starch content and the ethanol yield was supported by Cremer et al. (2014), but the R^2 value (0.74) was slightly higher.

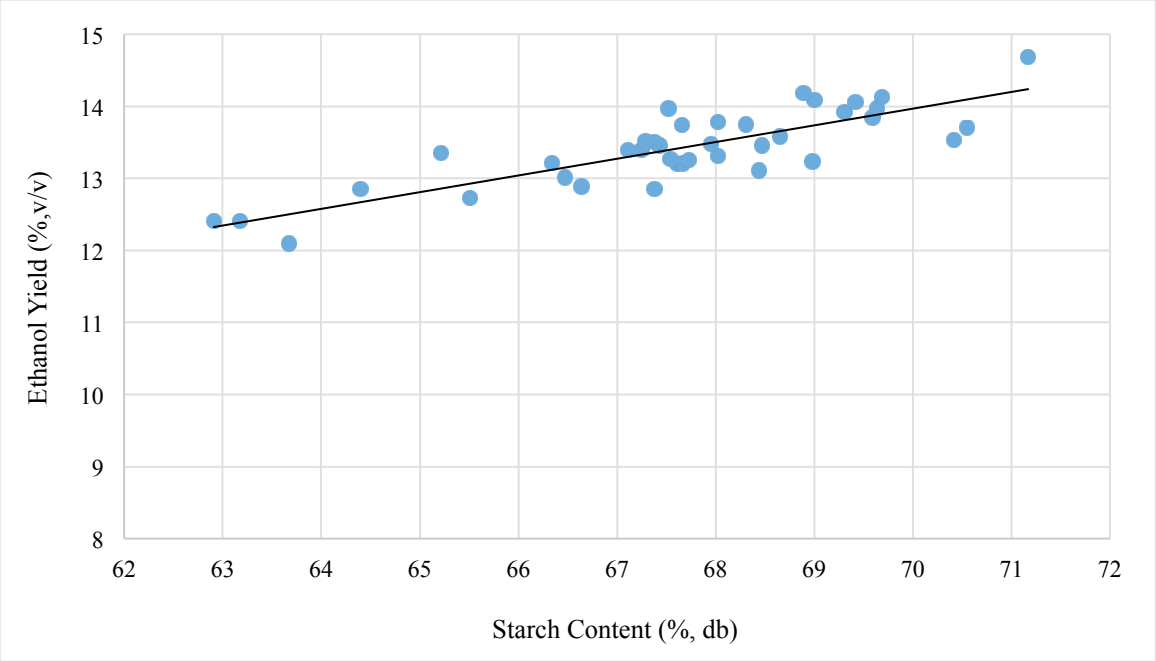


Figure 10. Correlation between ethanol yield and starch content

On the other hand, the starch content and fermentation efficiency were not found to be correlated based on Figure 11 and the studies completed by Wu et al. (2007). The R^2 value found for the trend line in Figure 11 was 0.047, which means that there was no relationship between the two parameters. However, the fermentation efficiency is an important factor in determining how much ethanol was produced compared to how much starch was available before fermentation.

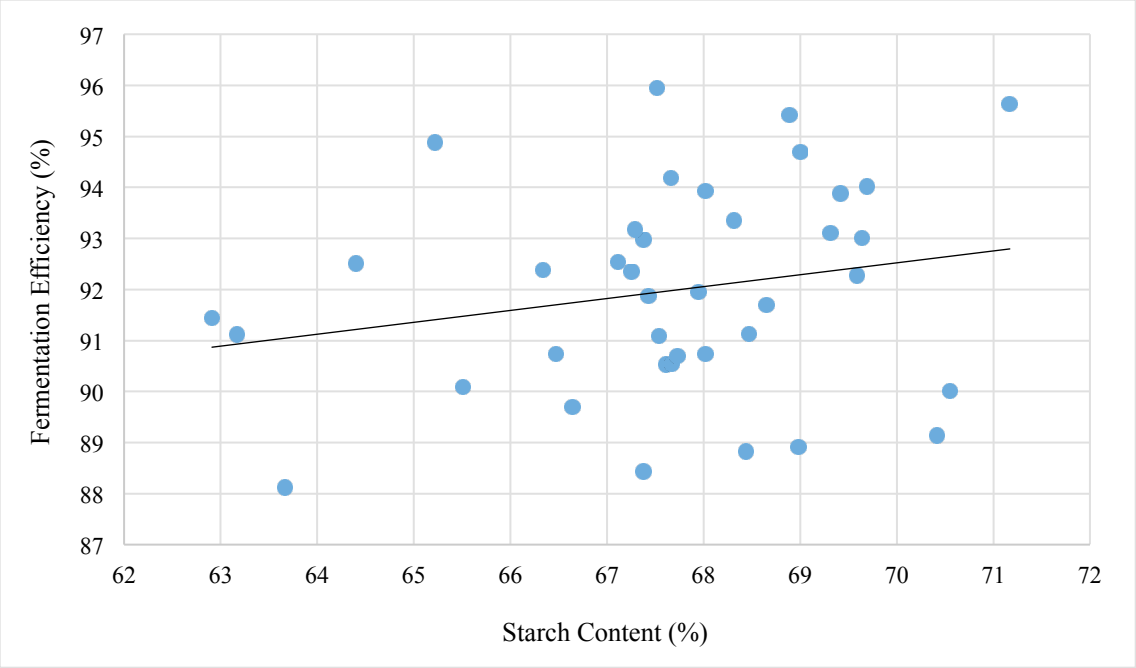


Figure 11. Correlation between fermentation efficiency and starch content

After recording the weight loss over a 72 h period, the weight loss could be converted into the ethanol yield. Figure 12 depicts the changes in ethanol yield with respect to time for all thirty-nine samples. The purpose of the ethanol yield curve is to determine when glucose is no longer undergoing conversion into ethanol.

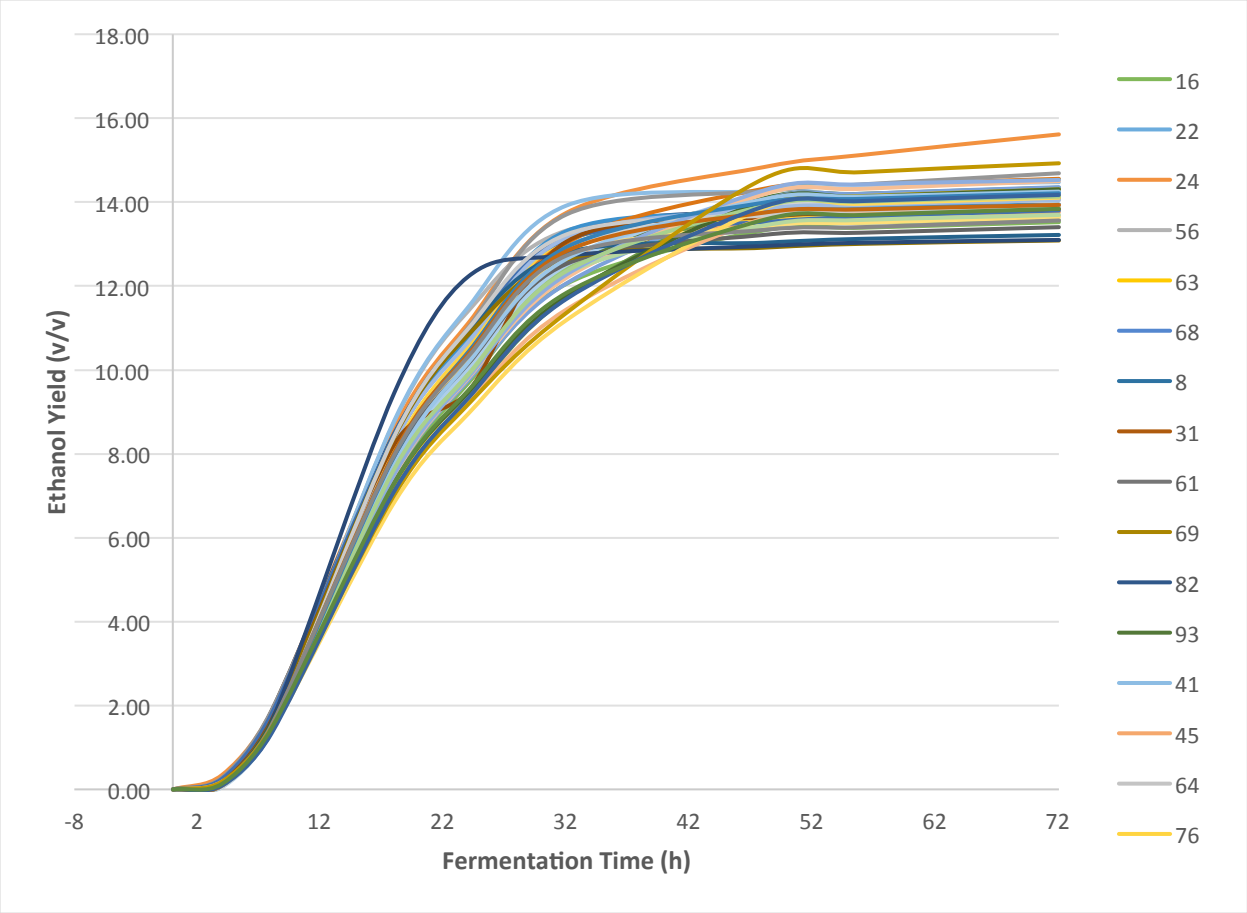


Figure 12. Ethanol yield with respect to fermentation time

Based on the ethanol yield curves, fermentation was done around 48 h, except for sample 82, which was complete at ~32 h. The time required for fermentation is significant for industrial application. If the bioconversion takes place in less time, the ethanol plant can stop fermentation and move the fermented product to distillation, which would cut costs and save energy for the next batch. Wu et al. (2007) describes how the grain sorghum samples took over 60 h before the completion of fermentation, but some of the samples were more similar to this research by taking only 40 h. However, germinated sorghum was reported to only take a total of 36 h to reach completion (Yan et al., 2009).

3.4 Chemical composition of DDGS

DDGS is the solid by-product left over after fermentation that is generally sold as feed by the ethanol plants. DDGS can have up to triple its original protein content after fermentation, which is significant since feed with higher protein content is more desirable for ruminant animals as it helps build muscle (Wang et al., 2008). In order to ensure successful ethanol conversion, all of the samples used for fermentation were analyzed for moisture content, starch content, and crude protein content. The representation of this compositional data is shown in Table A.7. For moisture content, the range was from ~11% (wb) to ~15.5% (wb) and an average of 14.2% (wb) with a standard deviation of 6.1%. Starch content values were between 0.9% (db) and ~2.0% (db), with a standard deviation of 16.9% (db) and an average starch content of 1.5% (db). The same calculation from the original mutant flour samples was repeated for the crude protein in the DDGS. The values for protein were ~36-42% (wb), with an average protein value of 38% (wb) and a standard deviation of 3.6%. Based on Wang et al. (2008), the starch content and the protein content are inversely related so if a higher starch content will produce a higher ethanol yield and a lower protein content in the DDGS.

3.5 Near infrared spectroscopy modeling

3.5.1 Starch content

Based on the starch content, the Partial Least Squares (PLS) regression model, with 77 samples for calibration, 22 samples for validation, and 10 samples removed as outliers, produced a root mean square of calibration (RMSEC) value of 0.83% with a coefficient of determination (R^2) of 0.77 and a root mean square error of prediction (RMSEP) value of 0.77% with an R^2 value of 0.90, for calibration and prediction models, respectively. These results demonstrate the

possibility of a positive correlation between the calculated and actual values for starch content (Figure 13). The "actual" values refer to the experimental data from wet chemistry and can be found in Table A.3 whereas the "calculated" are the predicted values determined by the regression model. The PLS model was chosen so the spectra could be represented by quantitative analysis through the use of factors. Each factor is a component chosen based on its importance to the data and the creation of the model. Figure 13 shows that 4 factors were used for simplicity in order to form the model.

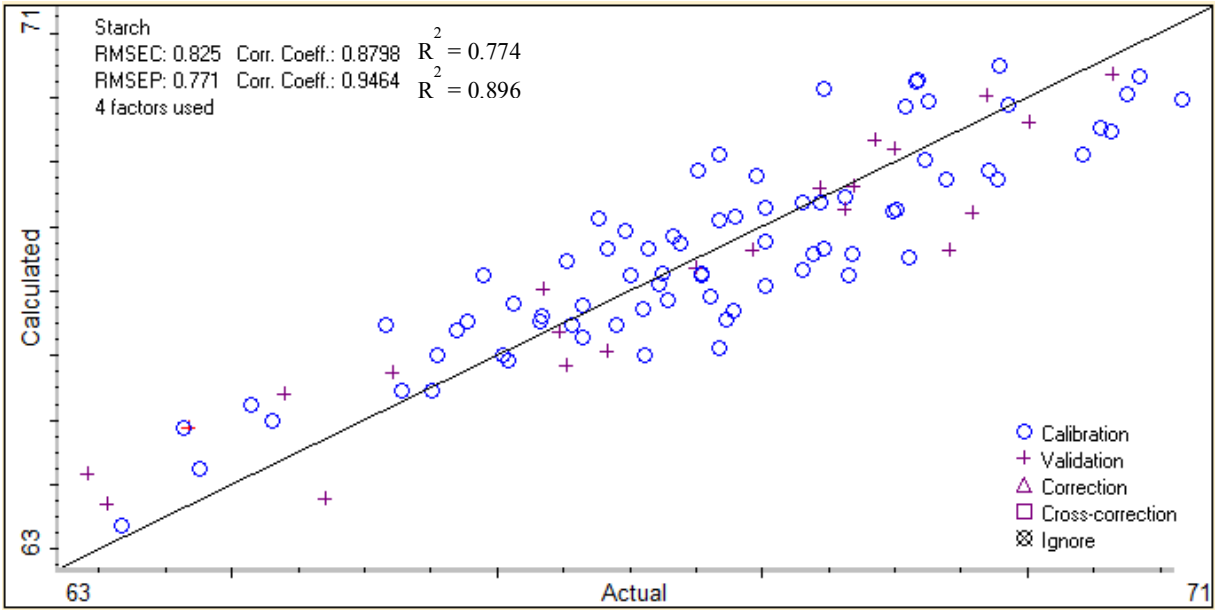


Figure 13. Scatter plot of NIR predicted values vs. actual starch values from 4,000-10,000 cm^{-1}

Figure 14 shows the NIR loading spectra from 4,000-10,000 cm^{-1}. The jagged parts in the spectra were sections of noise, which can interfere with the accuracy of the model. For this reason, it would be more beneficial to concentrate on a specific wavelength range. Selection of this range was based on the suggestion of the TQ Analyst software and other studies that suggested starch characteristics fall in the 1100-2300 nm (9,091-4348 cm^{-1}) range (Liebmann, Friedl, & Varmuza, 2010; Pohl & Senn, 2011; Hao, Thelen, & Gao, 2012).

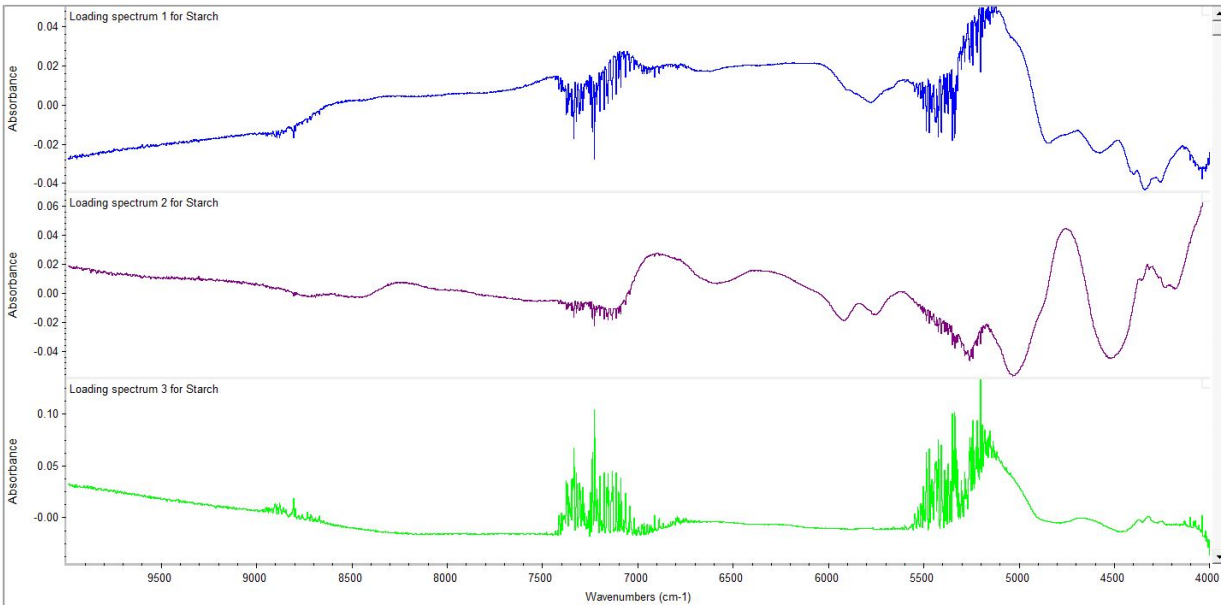


Figure 14. Loading NIR spectrum

By reducing the wavelength range from 4,000-5,176 cm^{-1}, the RMSEC decreased to 0.84% with an R^2 increases to 0.78 and the RMSEP increased to 0.89% with a slightly lower R^2 value of 0.84. Ten samples were ignored as outliers (77 for calibration and 22 for validation), but some of these samples were different from the first model. A maximum of 4 factors was used to create this model, seen in Figure 15 below. This scatter plot was similar to the graph for the whole range by illustrating a positive relationship between the actual and calculated values.

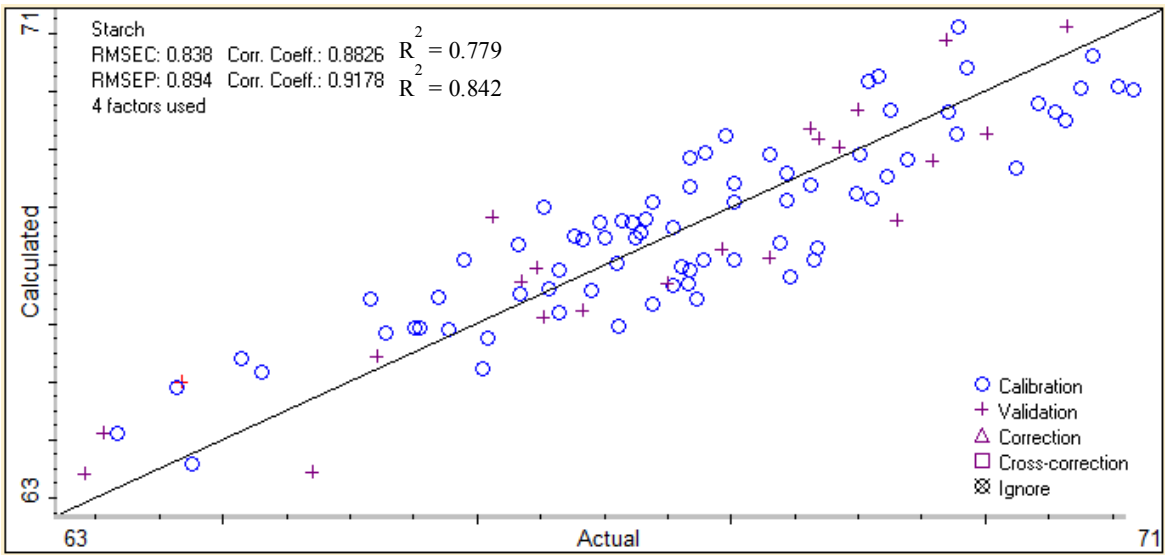


Figure 15. Scatter plot of NIR predicted values vs. actual starch values from 4,000-5,176 cm^{-1}

In order to limit the effect of the noise on the spectra and decrease background shifts, the first derivative was used to create two models for the full wavelength range and the reduced wavelength range (Lin et al., 2014). Figure 16 below show the scatter plot for the first derivative of the full range (4,000-10,000 cm^{-1}). This model showed how the RMSEC value decreased to 0.73% with R^2 equal to 0.83 and the RMSEP increased 0.92% (R^2 = 0.84).

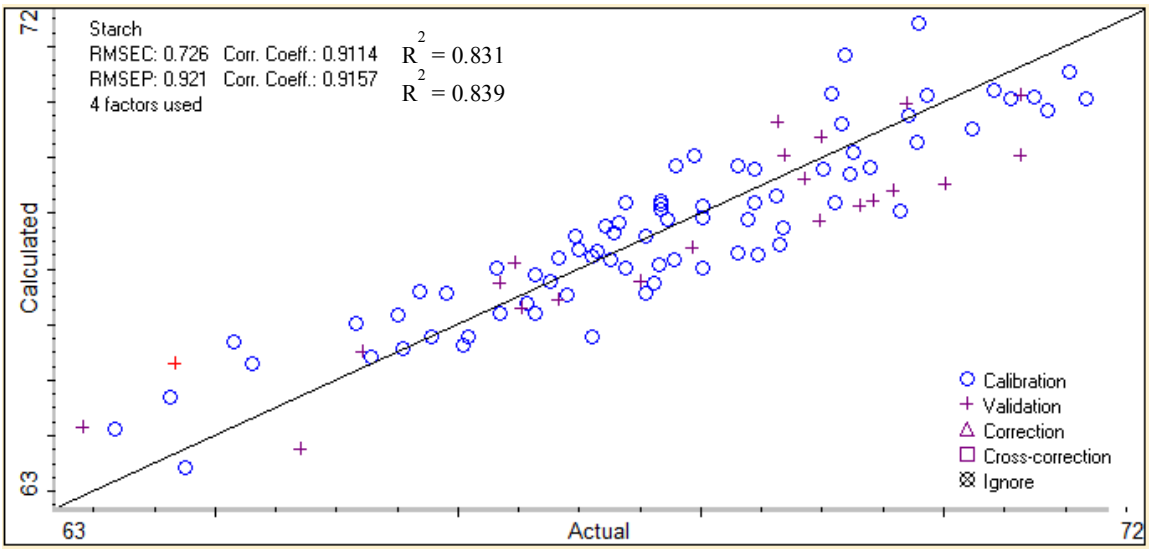


Figure 16. Scatter plot of NIR predicted values vs. actual starch values for the first derivative from 4,000-10,000 cm^{-1}

Figure 17 below show the scatter plot for the reduced range (4,000-5,176 cm^{-1}) of the first derivative. The RMSEC and RMSEP values decreased from 0.82% (R^2 = 0.79) and 0.85% (R^2 = 0.86). Based on these values, the first derivative in the limited range is a better model in comparison to the original curve.

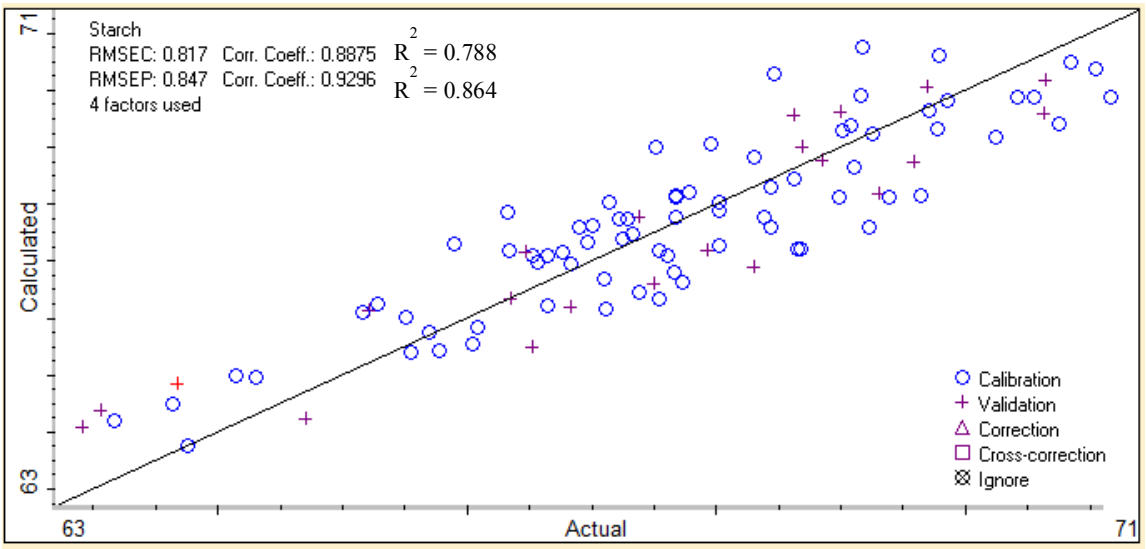


Figure 17. Scatter plot of NIR predicted values vs. actual starch values for the first derivative from 4,000-5,176 cm^{-1}

3.5.2 Fermentation efficiency

Another PLS model, using 39 of the original 109 samples (27 for calibration, 8 for validation and 4 ignored), was created for the fermentation efficiency, which had a RMSEC of 1.70% (R^2 = 0.26) and a RMSEP of 1.22% (R^2 = 0.75), seen in Figure 18. Based on this model, there does not appear to be any relationship for the fermentation efficiency and the spectra. However, limiting the range of wavelengths to avoid the background noise and taking the first derivative could reduce the errors and improve the coefficients of determination.

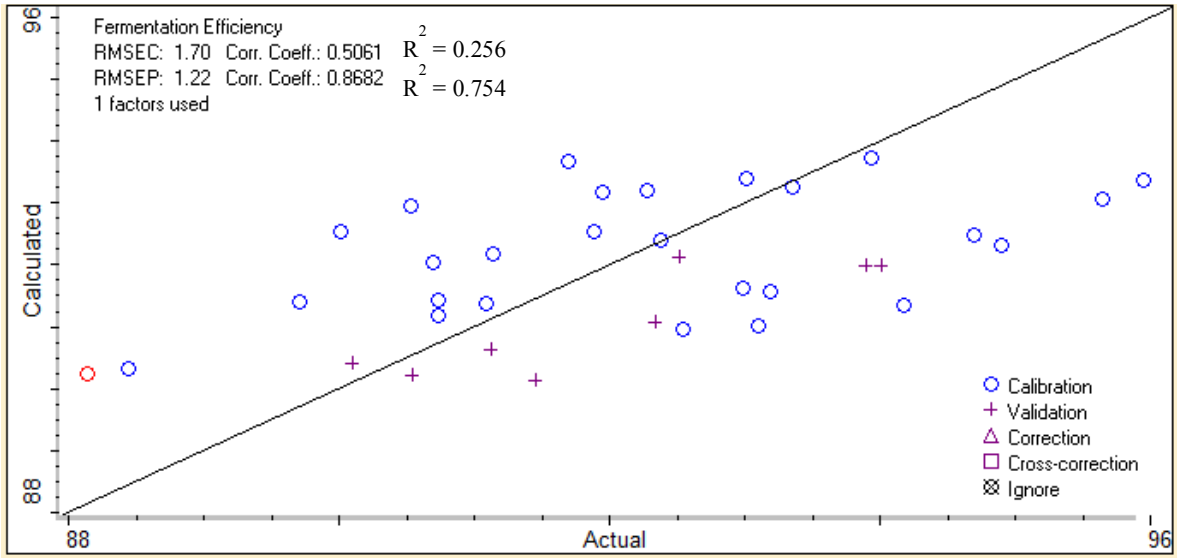


Figure 18. Scatter plot of NIR predicted values vs. actual fermentation efficiency values from 4,000-10,000 cm^{-1}

In Figure 19, the RMSEC was 1.64%, which is less than the full wavelength range, and the R^2 value increased to 0.42. On the other hand, the RMSEP increased to 1.84% and the R^2 value decreased to 0.57. The R^2 value is still really low, which is probably due to a large ratio of noise.

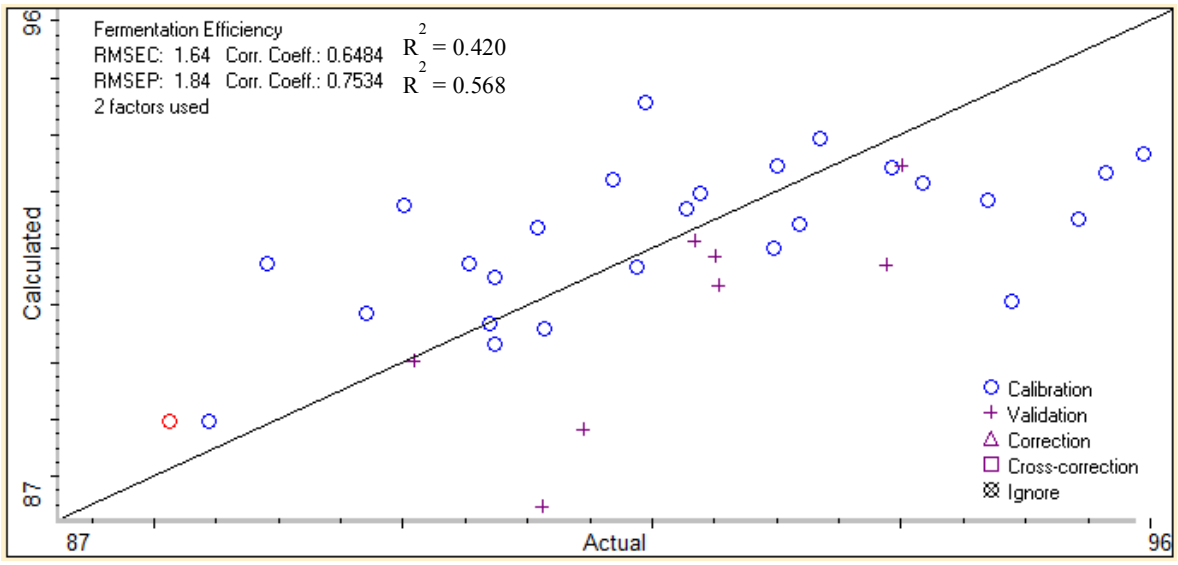


Figure 19. Scatter plot of NIR predicted values vs. actual fermentation efficiency values from 4,000-5,176 cm^{-1}

After taking the first derivative, figures 20 and 21 appear to have a linear relationship between the calculated and the actual fermentation efficiency values within the model. Figure 20 demonstrates how the RMSEC decreased to 0.98% ($R^2 = 0.76$) and the RMSEP increased to 1.74% ($R^2 = 0.38$). Even though the error for calibration was reduced, the error of the prediction is still very high in comparison to the full wavelength range. The coefficient of determination was also significantly lower than the original model.

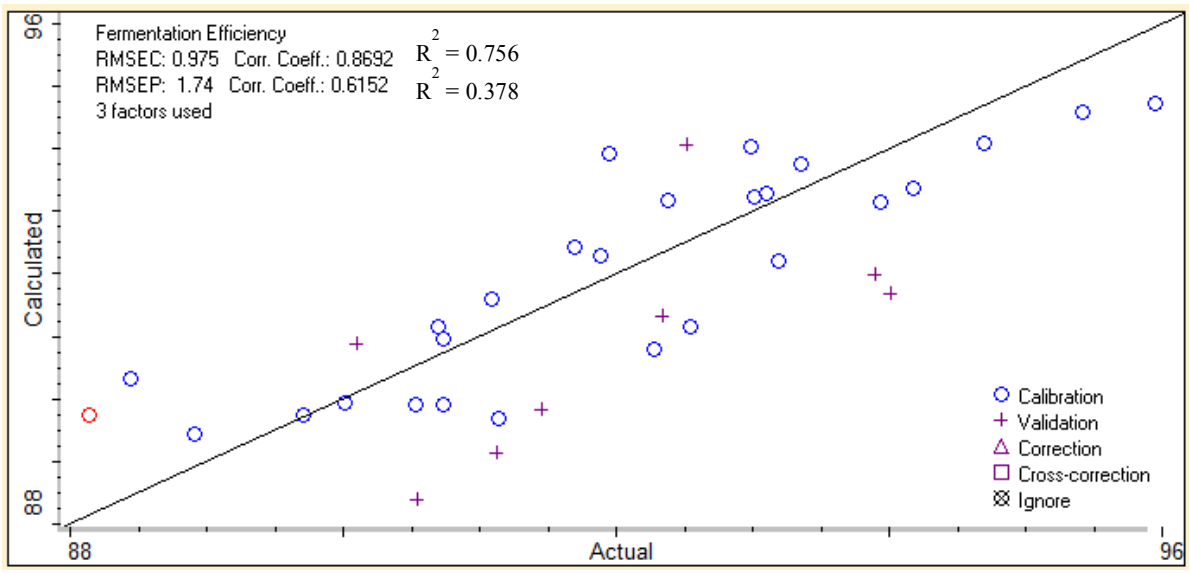


Figure 20. Scatter plot of NIR predicted values vs. actual fermentation efficiency values for the first derivative from 4,000-10,000 cm^{-1}

Finally, Figure 21 shows how the RMSEC and RMSEP both decreased to 0.49% and 1.12%, respectively. The R^2 value for the RMSEC increased significantly to 0.95 and the RMSEP R^2 value increased to 0.78.

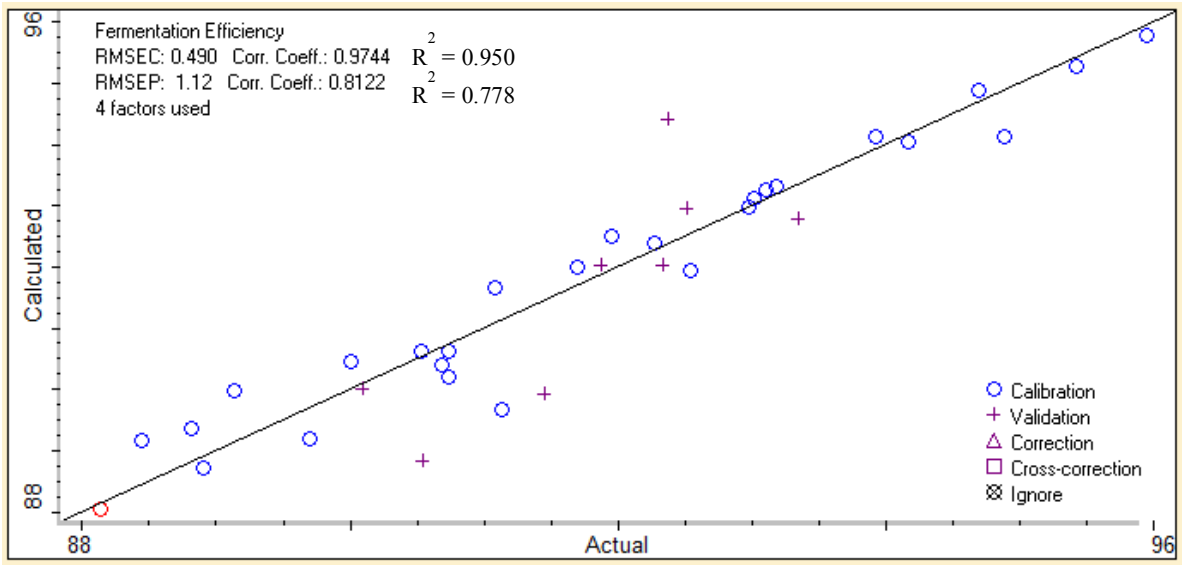


Figure 21. Scatter plot of NIR predicted values vs. actual fermentation efficiency values for the first derivative from 4,000-5,176 cm^{-1}

3.5.3 Modeling Discussion

After comparison, the best representation for the starch content was the full wavelength model due to the higher R^2 value (~0.90) for the error of prediction, which demonstrates robustness of the model. However, Kim and Williams (1990) stated that the best calibration model for starch in barley was the first derivative model and the second derivative was the best for maize starch and wheat. Also, this study states how their models produced high correlation coefficients and low standard errors, which is comparable to the calibration models in this research. According to Lin et al. (2014), which studied the protein content of barley, the SEP and R^2 values was significantly higher than the current model. However, their model was validated with the testing of additional samples, which has not been completed in this research.

Based on the fermentation efficiency results, the best calibration model is the first derivative within the reduced wavelength range. This model has a coefficient of determination equal to 0.78 and a value of 1.12 as its standard error of performance (SEP). There has been little literature on the use of NIR to determine fermentation efficiency, but other studies describe their method for ethanol production, specifically ethanol yield. For research completed by Pohl and Senn (2011), triticale was most similar to the mutant grain sorghum with an average starch content of ~67% (db). The results for the Foss 5000 instrument show a RMSEC of 0.69%, a calibration R^2 of 0.609, and a RMSEP of 0.65% for the ethanol yield of triticale. While the standard error of calibration R^2 was higher in this study (0.95), their RMSEC was higher and their RMSEP was lower with the use of 5 factors. The use of 5 factors creates a higher level of complexity for their model since the model in Figure 21 only used 4 factors. Overall, the models demonstrated the possibility of using NIR model prediction for mutant sorghum starch content and fermentation efficiency.

References

Ai, Y., Medic, J., Jiang, H., Wang, D., & Jane, J. L. (2011). Starch characterization and ethanol production of sorghum. *Journal of agricultural and food chemistry*, *59*(13), 7385-7392.

Appiah-Nkansah, N. B., Saul, K., Rooney, W. L., & Wang, D. (2015). Adding sweet sorghum juice into current dry-grind ethanol process for improving ethanol yields and water efficiency. *International Journal of Agricultural and Biological Engineering*, *8*(2), 97.

Cremer, J. E., Liu, L., Bean, S. R., Ohm, J. B., Tilley, M., Wilson, J. D., ... & Wang, D. (2014). Impacts of kafirin allelic diversity, starch content, and protein digestibility on ethanol conversion efficiency in grain sorghum. *Cereal Chemistry*, *91*(3), 218-227.

Hao, X., Thelen, K., & Gao, J. (2012). Prediction of the ethanol yield of dry-grind maize grain using near infrared spectroscopy. *Biosystems engineering*, *112*(3), 161-170.

Kim, H. O., & Williams, P. C. (1990). Determination of starch and energy in feed grains by near-infrared reflectance spectroscopy. *Journal of Agricultural and Food Chemistry*, *38*(3), 682-688.

Liebmann, B., Friedl, A., & Varmuza, K. (2010). Applicability of near-infrared spectroscopy for process monitoring in bioethanol production. *Biochemical Engineering Journal*, *52*(2), 187-193.

Lin, C., Chen, X., Jian, L., Shi, C., Jin, X., & Zhang, G. (2014). Determination of grain protein content by near-infrared spectrometry and multivariate calibration in barley. *Food chemistry*, *162*, 10-15.

Liu, L., Maier, A., Klocke, N. L., Yan, S., Rogers, D. H., Tesso, T., & Wang, D. (2013). Impact of deficit irrigation on sorghum physical and chemical properties and ethanol yield. *Transactions of the ASABE*, *56*(4), 1541-1549.

Pohl, F., & Senn, T. (2011). A rapid and sensitive method for the evaluation of cereal grains in bioethanol production using near infrared reflectance spectroscopy. *Bioresource technology*, *102*(3), 2834-2841.

Shewayrga, H., Sopade, P. A., Jordan, D. R., & Godwin, I. D. (2012). Characterisation of grain quality in diverse sorghum germplasm using a Rapid Visco-Analyzer and near infrared reflectance spectroscopy. *Journal of the Science of Food and Agriculture*, *92*(7), 1402-1410.

Sun, Q., Han, Z., Wang, L., & Xiong, L. (2014). Physicochemical differences between sorghum starch and sorghum flour modified by heat-moisture treatment. *Food chemistry*, *145*, 756-764.

Wang, D., Bean, S., McLaren, J., Seib, P., Madl, R., Tuinstra, M., ... & Zhao, R. (2008). Grain sorghum is a viable feedstock for ethanol production. *Journal of industrial microbiology & biotechnology*, *35*(5), 313-320.

Wong, J. H., Marx, D. B., Wilson, J. D., Buchanan, B. B., Lemaux, P. G., & Pedersen, J. F. (2010). Principal component analysis and biochemical characterization of protein and starch reveal primary targets for improving sorghum grain. *Plant Science*, *179*(6), 598-611.

Wu, X., Zhao, R., Bean, S. R., Seib, P. A., McLaren, J. S., Madl, R. L., ... & Wang, D. (2007). Factors impacting ethanol production from grain sorghum in the dry-grind process. *Cereal Chemistry*, *84*(2), 130-136.

Wu, X., Jampala, B., Robbins, A., Hays, D., Yan, S., Xu, F., ... & Wang, D. (2010). Ethanol fermentation performance of grain sorghums (Sorghum bicolor) with modified endosperm matrices. *Journal of agricultural and food chemistry*, *58*(17), 9556-9562.

Yan, S., Wu, X., MacRitchie, F., & Wang, D. (2009). Germination-Improved Ethanol Fermentation Performance of High-Tannin Sorghum in a Laboratory Dry-Grind Process. *Cereal chemistry*, *86*(6), 597-600.

Yan, S., Wu, X., Bean, S. R., Pedersen, J. F., Tesso, T., Chen, Y., & Wang, D. (2011). Evaluation of waxy grain sorghum for ethanol production.

Zhao, R. Z., Bean, S. R., Wang, D., & Wu, X. (2008). Assessing fermentation quality of grain sorghum for fuel ethanol production using rapid visco analyzer. In *2008 Providence, Rhode Island, June 29–July 2, 2008* (p. 1). American Society of Agricultural and Biological Engineers.

Chapter 4 - Conclusions and Recommendations

4.1 Conclusions

Ethanol fermentation was completed in about 48 h for most of the sorghum samples, which means that the fermentation time can be shortened from the original 72 h period. Thus, the mutant grain sorghum was advantageous for ethanol fermentation. The average starch content was 67.6% (db) with a standard deviation of 5.8% and the average ethanol yield was 13.4% (v/v) with a standard deviation of ~4.0%. There was a positive correlation ($R^2 = 0.70$) between the ethanol yield and the starch content. The average fermentation efficiency was 91% with a standard deviation of 2.2% and there was no relationship between the fermentation efficiency and starch content.

Near infrared spectroscopy has the potential to create a more cost-effective, quick method for prediction of the starch and fermentation efficiency. The PLS model for starch exhibited robustness with a high R^2 of 0.77 and 0.90 and low error values of 0.83% and 0.77%, for calibration and prediction models, respectively. For the fermentation efficiency, the RMSEC was 0.49% with a R^2 value of 0.95 and the RMSEP was 1.12% with a R^2 value was 0.78 for the first derivative PLS model (4,000-5,176 cm^{-1}). In future studies, it would be beneficial for the NIR spectra to be applied to another software to determine statistical significance and more analysis for fermentation in order to expand the basis for the model.

4.2 Recommendations

For fermentation studies, more research is needed to analyze the relationship among the genetic and functional properties, processing methods, and final ethanol yield and efficiency for grain sorghum. For studies on NIR models, more analysis is necessary to increase model

accuracy of fermentation efficiency models. Further statistical analysis is required in order to determine the significance of the model. Also, more mutant grain sorghum samples are needed to increase the model accuracy for both the starch content and fermentation efficiency models. Another recommendation is to combine starch content and ethanol yield into one model that could provide a fast way for ethanol plants to predict the starch content, ethanol yield, and fermentation efficiency.

Appendix A - Data Tables

Table A.1. Compositional analysis for mutant grain sorghum

Composition	Range	Average
Moisture (%, wb)	7-10	8.45
Starch (%, db)	63-71	67.27
Protein (%, wb)	11-28	14.83

Table A.2. Average and standard deviation data for moisture content (MC) with one replicate

Sample	Line	Average MC (%)	Standard Deviation (%)
1	25M2-0003	8.47	0.16
2	25M2-0054	8.10	0.80
3	25M2-0063	9.01	0.65
4	25M2-0083	8.05	1.34
5	25M2-0113	7.61	2.92
6	25M2-0176	7.95	0.55
7	25m2-0183	8.77	2.39
8	25m2-0193	9.14	1.47
9	25m2-0236	9.15	0.77
10	25m2-0275	7.89	1.02
11	25m2-0301	6.97	2.96
12	25m2-0315	8.97	0.21
13	25m2-0322	7.35	0.93
14	25m2-0323	7.40	0.59
15	25m2-0363	8.32	1.06
16	25m2-0365	8.08	2.06
17	25m2-0378	8.83	0.11
18	25m2-0439	7.66	0.34
19	25M2-0450	7.76	1.77
20	25M2-0466	8.78	0.70
21	25M2-0485	9.14	1.39

22	25M2-0488	8.73	2.40
23	25M2-0510	8.95	0.72
24	25M2-0516	8.68	1.82
25	25M2-0527	8.70	0.36
26	25M2-0535	7.62	0.32
27	25M2-0574	8.38	1.37
28	25M2-0579	7.95	0.14
29	25M2-0580	8.95	2.66
30	25M2-0581	7.55	1.42
31	25M2-0585	9.24	2.13
32	25M2-0601	8.38	1.21
33	25M2-0632	9.03	1.49
34	25M2-0637	8.06	0.33
35	25M2-0655	9.20	2.34
36	25M2-0664	8.26	0.60
37	25M2-0668	8.10	2.04
38	25M2-0696	8.64	0.94
39	25M2-0713	8.01	1.97
40	25M2-0714	8.66	0.40
41	25M2-0720	9.57	1.01
42	25M2-0731	7.82	1.10
43	25M2-0736	8.96	2.38
44	25M2-0740	8.79	0.53
45	25M2-0741	7.66	1.25
46	25M2-0768	8.49	0.07
47	25M2-0771	8.54	0.34
48	25M2-0775	8.01	1.36
49	25M2-0788	9.11	2.18
50	25M2-0897	9.16	0.15
51	25M2-0912	8.30	0.45
52	25M2-0940	8.78	0.08

53	25M2-0964	8.57	2.48
54	25M2-0991	8.64	1.70
55	25M2-1025	8.70	0.13
56	25m2-1036	9.45	0.75
57	25M2-1077	8.92	0.10
58	25m2-1079	7.68	1.34
59	25M2-1080	8.70	0.65
60	25M2-1082	7.99	1.47
61	25M2-1083	8.32	1.23
62	25M2-1091	9.06	0.55
63	25M2-1093	7.78	1.85
64	25M2-1098	7.39	0.41
65	25M2-1111	7.69	0.29
66	25M2-1129	8.51	0.25
67	25M2-1131	7.25	0.63
68	25M2-1135	8.23	0.08
69	25M2-1141	8.61	1.60
70	25M2-1159	8.70	0.38
71	25M2-1183	8.84	2.27
72	25M2-1194	8.50	0.51
73	25M2-1201	8.77	0.02
74	25M2-1343	8.20	0.23
75	25M2-1435	9.29	0.26
76	25M2-1449	8.66	2.02
77	25M2-1506	8.75	0.83
78	25m2-1522	8.76	2.71
79	25M2-1542	9.93	0.47
80	25m2-1553	8.76	2.91
81	25M2-1645	8.57	0.88
82	25M2-1682	8.09	1.11
83	25M2-1683	9.16	0.81

84	25M2-1705	8.57	1.74
85	25M2-1707	8.66	0.90
86	25M2-1716	8.49	0.35
87	25M2-1719	8.87	0.29
88	25M2-1728	8.08	0.86
89	25M2-1737	8.22	1.82
90	25M2-1784	8.74	0.72
91	25M2-1818	8.93	0.51
92	25m2-2001	8.35	0.70
93	25m2-2034	8.38	0.78
94	25m2-2082	8.59	0.75
95	25m2-2095	8.50	0.23
96	25m2-2134	7.89	0.41
97	25m2-2147	7.92	1.28
98	25m2-2265	7.91	1.23
99	25m2-2285	8.65	1.57
100	M2P0049	9.01	0.20
101	M2P0114	8.15	0.42
102	M2P0799	8.55	0.22
103	M2P0810	7.79	1.01
104	M2P0841	8.59	0.67
105	M2P0872	7.87	1.00
106	M2P0965	8.53	0.08
107	mut1104	9.50	2.80
108	BTx623	7.87	0.38
109	BTx623	8.70	0.18

Table A.3. Average and standard deviation data for starch content with one replicate

Sample	Line	Average Starch (db, %)	Standard Deviation (%)
1	25M2-0003	63.67	0.53

2	25M2-0054	66.32	0.14
3	25M2-0063	66.56	0.93
4	25M2-0083	64.40	0.46
5	25M2-0113	63.06	1.18
6	25M2-0176	65.28	1.08
7	25m2-0183	65.90	0.47
8	25m2-0193	68.02	0.31
9	25m2-0236	69.16	0.00
10*	25m2-0275	31.88	141.42
11	25m2-0301	65.55	0.71
12	25m2-0315	69.59	0.37
13	25m2-0322	67.10	0.00
14	25m2-0323	65.51	0.02
15	25m2-0363	68.89	0.85
16	25m2-0365	67.38	0.30
17	25m2-0378	69.79	0.04
18	25m2-0439	62.91	1.35
19	25M2-0450	64.71	0.67
20	25M2-0466	69.25	1.07
21	25M2-0485	67.54	0.34
22	25M2-0488	67.73	0.35
23	25M2-0510	69.64	0.58
24	25M2-0516	67.33	1.14
25	25M2-0527	65.78	0.74
26	25M2-0535	66.83	0.19
27	25M2-0574	69.78	0.02
28	25M2-0579	71.04	0.93
29	25M2-0580	68.85	0.73
30	25M2-0581	68.44	0.15
31	25M2-0585	68.65	0.66
32	25M2-0601	67.50	0.69

33	25M2-0632	67.61	0.80
34	25M2-0637	69.71	0.14
35	25M2-0655	67.14	0.68
36	25M2-0664	67.78	1.22
37	25M2-0668	63.63	0.26
38	25M2-0696	70.63	0.16
39	25M2-0713	65.88	0.97
40	25M2-0714	66.12	0.94
41	25M2-0720	67.80	1.59
42	25M2-0731	66.53	0.42
43	25M2-0736	66.04	0.25
44	25M2-0740	70.42	0.53
45	25M2-0741	67.11	1.12
46	25M2-0768	67.22	0.41
47	25M2-0771	66.96	0.30
48	25M2-0775	66.64	0.72
49	25M2-0788	68.98	0.18
50	25M2-0897	70.85	0.06
51	25M2-0912	68.02	0.65
52*	25M2-0940	-	-
53	25M2-0964	64.15	1.03
54	25M2-0991	69.31	0.84
55	25M2-1025	68.44	0.19
56	25m2-1036	67.67	0.47
57	25M2-1077	65.22	0.79
58	25m2-1079	63.80	0.73
59	25M2-1080	69.01	0.28
60	25M2-1082	66.83	0.20
61	25M2-1083	68.47	0.35
62	25M2-1091	70.01	0.31
63	25M2-1093	67.38	0.69

64	25M2-1098	67.25	0.87
65	25M2-1111	64.31	0.98
66	25M2-1129	69.39	0.28
67	25M2-1131	67.00	1.62
68	25M2-1135	67.54	0.74
69	25M2-1141	67.93	0.38
70	25M2-1159	66.76	0.60
71	25M2-1183	68.31	0.05
72	25M2-1194	69.42	0.32
73	25M2-1201	67.67	0.57
74	25M2-1343	68.69	0.46
75	25M2-1435	69.10	0.21
76	25M2-1449	67.29	0.51
77	25M2-1506	68.38	0.30
78	25m2-1522	71.17	0.71
79	25M2-1542	68.62	0.14
80	25m2-1553	70.64	0.42
81	25M2-1645	70.55	0.25
82	25M2-1682	68.02	0.32
83	25M2-1683	67.67	0.16
84	25M2-1705	68.63	0.47
85	25M2-1707	69.86	0.90
86	25M2-1716	65.16	0.49
87	25M2-1719	66.09	0.40
88	25M2-1728	67.52	0.66
89	25M2-1737	69.23	0.52
90	25M2-1784	68.47	0.73
91	25M2-1818	69.69	0.41
92	25m2-2001	66.47	0.40
93	25m2-2034	68.31	0.67
94	25m2-2082	66.34	0.95

95	25m2-2095	65.69	0.62
96	25m2-2134	66.30	1.88
97	25m2-2147	69.00	0.53
98	25m2-2265	67.66	0.55
99	25m2-2285	63.17	0.13
100	M2P0049	66.90	0.47
101	M2P0114	70.24	0.66
102	M2P0799	66.52	0.89
103	M2P0810	67.95	0.02
104	M2P0841	66.35	0.05
105	M2P0872	66.64	0.41
106	M2P0965	69.08	0.60
107	mut1104	68.68	1.18
108	BTx623	69.18	0.96
109	BTx623	70.75	0.65

* Experimental errors

Table A.4. Nitrogen and crude protein data

Sample	Line	Nitrogen (%)	Crude Protein (%)
1	25M2-0003	2.52	15.75
2	25M2-0054	2.28	14.25
3	25M2-0063	2.67	16.69
4	25M2-0083	2.48	15.50
5	25M2-0113	2.77	17.31
6	25M2-0176	2.64	16.50
7	25m2-0183	2.81	17.56
8	25m2-0193	2.02	12.63
9	25m2-0236	2.79	17.44
10	25m2-0275	2.44	15.25
11	25m2-0301	2.45	15.31
12	25m2-0315	2.27	14.19

13	25m2-0322	2.22	13.88
14	25m2-0323	2.37	14.81
15	25m2-0363	2.27	14.19
16	25m2-0365	2.34	14.63
17	25m2-0378	1.98	12.38
18	25m2-0439	2.94	18.38
19	25M2-0450	2.90	18.13
20	25M2-0466	2.34	14.63
21	25M2-0485	3.16	19.75
22	25M2-0488	2.34	14.63
23	25M2-0510	2.13	13.31
24	25M2-0516	2.74	17.13
25	25M2-0527	2.42	15.13
26	25M2-0535	2.45	15.31
27	25M2-0574	1.94	12.13
28	25M2-0579	1.99	12.44
29	25M2-0580	2.11	13.19
30	25M2-0581	2.31	14.41
31	25M2-0585	2.27	14.19
32	25M2-0601	2.33	14.56
33	25M2-0632	2.80	17.50
34	25M2-0637	2.06	12.88
35	25M2-0655	2.50	15.63
36	25M2-0664	2.39	14.91
37	25M2-0668	1.90	11.88
38	25M2-0696	2.12	13.25
39	25M2-0713	2.25	14.06
40	25M2-0714	2.42	15.13
41	25M2-0720	2.07	12.94
42	25M2-0731	2.80	17.47
43	25M2-0736	2.45	15.31

44	25M2-0740	2.26	14.13
45	25M2-0741	2.42	15.13
46	25M2-0768	2.45	15.28
47	25M2-0771	2.44	15.25
48	25M2-0775	2.12	13.25
49	25M2-0788	2.30	14.38
50	25M2-0897	2.84	17.75
51	25M2-0912	2.40	15.00
53	25M2-0964	2.43	15.19
54	25M2-0991	2.30	14.38
55	25M2-1025	2.12	13.25
56*	25m2-1036	4.48	28.00
57*	25M2-1077	3.50	21.88
58	25m2-1079	2.26	14.13
59	25M2-1080	2.23	13.94
60	25M2-1082	2.56	16.00
61	25M2-1083	2.47	15.44
62	25M2-1091	2.61	16.31
63	25M2-1093	2.29	14.31
64	25M2-1098	2.33	14.56
65	25M2-1111	2.87	17.94
66	25M2-1129	1.82	11.38
67	25M2-1131	2.20	13.75
68	25M2-1135	2.17	13.56
69	25M2-1141	2.20	13.75
70	25M2-1159	2.36	14.75
71	25M2-1183	2.34	14.63
72	25M2-1194	2.32	14.50
73	25M2-1201	2.75	17.19
74	25M2-1343	2.01	12.56
75	25M2-1435	2.46	15.38

76	25M2-1449	2.42	15.13
77	25M2-1506	2.40	15.00
78	25m2-1522	1.90	11.88
79	25M2-1542	2.23	13.94
80	25m2-1553	2.29	14.31
81	25M2-1645	2.22	13.88
82	25M2-1682	2.11	13.19
83	25M2-1683	2.40	15.00
84	25M2-1705	2.16	13.50
85	25M2-1707	3.07	19.19
86	25M2-1716	2.23	13.94
87	25M2-1719	2.72	17.00
88	25M2-1728	2.05	12.81
89	25M2-1737	2.34	14.63
90	25M2-1784	1.97	12.31
91	25M2-1818	1.98	12.38
92	25m2-2001	2.53	15.81
93	25m2-2034	2.10	13.13
94	25m2-2082	2.23	13.94
95	25m2-2095	2.30	14.38
96	25m2-2134	2.13	13.31
97	25m2-2147	2.14	13.38
98	25m2-2265	2.68	16.75
99	25m2-2285	2.73	17.06
100	M2P0049	2.40	15.00
101	M2P0114	1.98	12.38
102	M2P0799	2.22	13.88
103	M2P0810	1.95	12.19
104	M2P0841	2.45	15.31
105	M2P0872	2.28	14.25
106	M2P0965	1.96	12.25

107	mut1104	2.15	13.44
108	BTx623	1.90	11.88
109	BTx623	1.84	11.50

* Experimental errors

Table A.5. RVA data for the lowest and highest fermentation efficiency samples

Sample	Peak Visc. (cP)	Minimum Visc. (cP)	Breakdown Visc. (cP)	Final Visc. (cP)	Setback Visc. (cP)	Peak Time (min)	Pasting Temp. (°C)
Sample 1	1141	1084	57	2328	1244	6.317	92
Sample 15	1362	1301	61	2974	1673	6.450	92.1
Sample 16	1211	1163	48	2495	1332	6.783	92.75
Sample 44	1663	1500	163	3581	2081	6.250	73.4
Sample 49	1607	1435	172	3329	1894	6.250	86.25
Sample 55	1759	1487	272	3963	2476	6.317	52.8
Sample 57	1311	1213	98	2600	1387	6.250	60.3
Sample 78	1718	1540	178	3604	2064	6.583	90.35
Sample 88	1702	1515	187	3589	2074	6.317	50.95
Sample 97	1748	1523	225	3715	2192	6.317	50.1

Table A.6. Ethanol yields and fermentation efficiencies (sorted based on fermentation efficiency)

Sample	Line	Actual Ethanol Yield (%,v/v)	Theoretical Yield (%, v/v)	Flour Total Starch (%, db)	Fermentation Efficiency (%)
1*	25M2-0003	12.10	13.73	63.67	88.13
16*	25m2-0365	12.85	14.53	67.38	88.44
55*	25M2-1025	13.11	14.76	68.44	88.82
49*	25M2-0788	13.23	14.88	68.98	88.91
44*	25M2-0740	13.54	15.19	70.42	89.14
105	M2P0872	12.89	14.37	66.64	89.70
38	25M2-0696	13.70	15.22	70.55	90.01
14	25m2-0323	12.73	14.13	65.51	90.09
33	25M2-0632	13.20	14.58	67.61	90.53
56	25m2-1036	13.21	14.59	67.67	90.54
22	25M2-0488	13.25	14.61	67.73	90.69
92	25m2-2001	13.01	14.34	66.47	90.73
82	25M2-1682	13.31	14.67	68.02	90.73
68	25M2-1135	13.27	14.57	67.54	91.08
99	25m2-2285	12.41	13.62	63.17	91.12
61	25M2-1083	13.46	14.77	68.47	91.13
18	25m2-0439	12.41	13.57	62.91	91.45
31	25M2-0585	13.58	14.81	68.65	91.69
69	25M2-1141	13.46	14.65	67.43	91.88
103	M2P0810	13.47	14.65	67.95	91.95
12	25m2-0315	13.85	15.01	69.59	92.27
64	25M2-1098	13.39	14.50	67.25	92.34
94	25m2-2082	13.22	14.31	66.34	92.38
4	25M2-0083	12.85	13.89	64.40	92.51
45	25M2-0741	13.39	14.47	67.11	92.54
63	25M2-1093	13.51	14.53	67.38	92.98
23	25M2-0510	13.97	15.02	69.64	93.01

54	25M2-0991	13.92	14.95	69.31	93.10
76	25M2-1449	13.52	14.51	67.29	93.18
93	25m2-2034	13.75	14.73	68.31	93.35
72	25M2-1194	14.06	14.97	69.42	93.89
8	25m2-0193	13.78	14.67	68.02	93.93
91	25M2-1818	14.13	15.03	69.69	94.01
98	25m2-2265	13.74	14.59	67.66	94.17
97*	25m2-2147	14.09	14.88	69.00	94.69
57*	25M2-1077	13.35	14.07	65.22	94.89
15*	25m2-0363	14.18	14.86	68.89	95.42
78*	25m2-1522	14.68	15.35	71.17	95.64
88*	25M2-1728	13.97	14.56	67.52	95.95

*Samples with the lowest and highest fermentation efficiency selected for morphological, thermal, and pasting experiments

Table A.7. Composition of Distillers' Dried Grain Solubles (DDGS)

Sample	Line	Average MC (wb, %)	Average Starch (db, %)	Crude Protein (wb, %)
1	25M2-0003	13.83	1.67	40.48
4	25M2-0083	14.15	1.11	39.26
8	25m2-0193	14.41	1.28	38.11
12	25m2-0315	13.89	1.03	40.06
14	25m2-0323	14.70	1.53	38.89
15	25m2-0363	13.59	1.57	38.68
16	25m2-0365	13.47	0.92	40.09
18	25m2-0439	13.16	1.44	39.79
22	25M2-0488	14.17	1.49	39.19
23	25M2-0510	14.13	1.45	38.13
31	25M2-0585	15.10	1.58	39.31
33	25M2-0632	12.48	1.54	39.25
38	25M2-0696	15.35	1.00	37.64

44	25M2-0740	15.25	1.71	37.49
45	25M2-0741	13.43	1.53	39.56
49	25M2-0788	15.32	1.92	36.63
54	25M2-0991	13.70	1.89	38.06
55	25M2-1025	14.70	1.34	38.86
56	25m2-1036	14.28	1.33	37.67
57	25M2-1077	13.99	1.27	39.23
61	25M2-1083	13.81	1.61	39.66
63	25M2-1093	14.43	1.31	37.81
64	25M2-1098	13.98	1.95	37.28
68	25M2-1135	14.55	0.99	36.52
69	25M2-1141	13.62	1.65	39.82
72	25M2-1194	14.21	1.60	38.49
76	25M2-1449	13.09	1.59	39.93
78	25m2-1522	14.88	1.49	35.75
82	25M2-1682	14.17	1.64	36.91
88	25M2-1728	15.37	1.75	37.13
91	25M2-1818	14.32	1.50	36.63
92	25m2-2001	14.81	1.58	39.03
93	25m2-2034	14.85	1.45	37.08
94	25m2-2082	14.23	1.60	37.57
97	25m2-2147	14.38	1.58	36.77
98	25m2-2265	15.34	1.52	40.28
99	25m2-2285	11.24	1.25	41.78
103	M2P0810	15.43	1.80	36.55
105	M2P0872	15.56	1.34	39.81