ABSTRACT

REDDING, ARTHUR. An Assessment of the Dilute Acid Pretreatment of Coastal Bermudagrass for Bioethanol Production. (Under the direction of Jay J. Cheng).

There is a clear interest domestically to examine alternative liquid fuels which are more sustainable and environmentally friendly than gasoline. Bioethanol is a leading candidate for this replacement, but limitations exist on current starch based production. As a result, lignocellulosics are being examined. Lignocellulosics require a pretreatment step to degrade the biomass enough to allow enzymes to access to the carbohydrates. Dilute acid pretreatment has been demonstrated across many lignocellulosic feedstocks as a leading method compared to other pretreatment options.

Coastal bermudagrass was identified as a promising lignocellulosic feedstock for bioethanol production. It is well suited for the Southeastern United States where it is currently grown for nutrient management in concentrated animal farming operations and as a source of hay. In a full factorial experimental design, dilute sulfuric acid pretreatment was used to pretreat coastal bermudagrass at 120, 140, 160, and 180ºC using 0.3, 0.6, 0.9, and 1.2% w/w sulfuric acid over residence times of 5, 15, 30, and 60 minutes. After enzymatic hydrolysis, the highest yield of total sugars (combined xylose and glucose) was 97%, which was shared by the pretreatment condition combinations of 0.9% acid at 160ºC for 15 minutes and 1.2% acid at 160ºC for 5 minutes. The prehydrolyzate liquor was analyzed for inhibitory compounds (furfural, hydroxymethylfurfural (HMF)) in order to assess potential risk for inhibition during fermentation. Accounting for the inhibitory compounds, a pretreatment with 1.2% acid at 140ºC for 30 minutes with a total sugar yield of 94% o r0.9% acid at 160 ºC for five minutes with a total sugar yield of 91% may be more favorable for fermentation

because furfural levels remain under the inhibitory threshold concentration of 1 g/L. Additionally, due to significant interactions between factors, there are likely optimal pretreatment condition combinations possible other than those found experimentally.

Both kinetic and multiple linear regression (MLR) models have been developed in other studies to describe dilute acid pretreatment, however no study has yet applied an artificial neural network (ANN). In this study, the utility of an ANN was assessed for modeling the dilute acid pretreatment of coastal bermudagrass using statistics that quantified the error between the predicted data and actual data and through a comparison with an MLR model. The statistics used were the coefficient of determination (R^2), the root mean squared error (RMSE), and the root percent deviation (RPD). A standard 2nd-order polynomial multiple linear regression (MLR) model was developed for comparison with the ANN model. Time (minutes), acid concentration (% w/w), and temperature (ºC) were input into the models to generate xylose in the prehydrolyzate (PreH), glucose in the PreH, furfural in the PreH, HMF in the PreH, xylose in the enzymatic hydrolyzate (EH), and glucose in the EH. It was found that the two types of models predicted most of the outputs closely with the exception of the xylose in the PreH, which the ANN predicted more accurately. An ANN model with six hidden layer neurons was found to be the best overall model and confirmed the utility of utilizing ANN modeling in the area of biomass pretreatment.

An Assessment of the Dilute Acid Pretreatment of Coastal Bermudagrass for Bioethanol Production

by
Arthur Redding

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Biological and Agricultural Engineering

Raleigh, North Carolina

2009

APPROVED BY:

_____	_____
Jay J. Cheng Committee Chair	Steven W. Peretti
_____	
Ratna R. Sharma-Shivappa	

DEDICATION

This thesis is dedicated to my parents and my grandparents for their support of my education. I certainly would not have made it this far without the foundation you provided to me as a child by reading to me and answering the myriad of questions I would ask each night before bed.

BIOGRAPHY

Arthur Redding was born in Washington DC, but grew up in the more rural Accokeek, MD. As a child, he spent the majority of his free time running through the woods, following creeks and streams, and climbing trees with his friends and sister. Arthur has also always valued imagination and the expression of creativity which has resulted in the pursuit of various creative endeavors. From an early age, Arthur questioned the world (and the people) around him trying to understand how things worked and fit together. This led to an interest in science which was nurtured throughout grade school by his parents and grandparents. He attended four years of Catholic high school which he feels strengthened his education by infusing it with studies on ethics and morality. After high school, Arthur applied to several universities with the intent of studying computer science however, upon accepting admittance to the University of Maryland in College Park, MD, he postponed declaring a major. In a decision he hoped would combine creativity and science, he decided to pursue a bachelor of science in Biological Resources Engineering. Over the course of his college career he worked in the Biological Resources Engineering machine shop where he was able to get hands-on experience with the manufacturing side of engineering development. After college, Arthur found employment in a mechanical engineering role as a facility engineer at a small contract manufacturing pharmaceutical company in Baltimore, MD. He lived and worked in Baltimore for two years, always with the idea of returning to school to obtain a Masters degree. In 2006 he decided he was interested in the idea of sustainable resources and sustainable engineering and eventually found his way into the Biological and

Agricultural Engineering department at North Carolina State University under the direction of Dr. Jay Cheng. He has happily pursued his Master of Science degree here in Raleigh, NC for the past two years and is presently looking for career opportunities related to bioprocessing.

ACKNOWLEDGMENTS

I first must thank my advisor Dr. Jay Cheng and my committee members, Dr. Ratna Sharma and Dr. Steve Peretti. Each of you has helped me on my way to creating this thesis through coursework, meetings, and answering questions. You have also put up with my sometimes optimistic view of what I could complete over a given amount of time, so thank you for your patience. I truly have learned a great deal over the past two years from all of you.

Dr. Cheng, I must thank you further. With your guidance, my written and oral communication skills have greatly improved over the past two years. Additionally, you are an interesting guy to sit down and have a beer with and I am grateful you took the time to do so. I have really enjoyed my last two years as a member of your research group.

This department would not be what it is without the community of graduate students, former and current, who have populated the halls of Weaver. Thank you all for making life outside of graduate study interesting. Billy, Hayes, Evan, and Rob, our weekends out and our trips were necessary distractions and I will always remember the fun times. Jess, I will miss throwing pottery (and gossip) immensely. Ziyu, Deepak, and Jiele, I could not have asked for better research partners. Thank you for helping me learn and for keeping our research work enjoyable. You are all my friends and you will always have a welcome place in my life.

I must also thank Holly, my girlfriend, who helped save me hours by writing sample labels for me. You made the long days in the lab more bearable with your company and I'm glad that you were there smiling even when I was not.

And finally, I would like to thank my family because without them, I wouldn't be anything.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

THE DILUTE ACID PRETREATMENT OF LIGNOCELLULOSIC BIOMASS FOR THE PRODUCTION OF BIOETHANOL: A REVIEW

1.1 ABSTRACT

There is a clear interest domestically to examine alternative liquid fuels which are more sustainable and environmentally friendly than gasoline. Bioethanol is a leading candidate for this replacement, but limitations exist on current starch based production. As a result, lignocellulosics are being examined. Lignocellulosics require a pretreatment step to degrade the biomass enough to allow enzymes to access to the carbohydrates. Dilute acid pretreatment alone and combined with steam explosion has been demonstrated across many lignocellulosic feedstocks as a leading method compared to other pretreatment options. However, because dilute acid has a tendency to generate compounds inhibitory to fermentation, care in analysis and process design must be taken to limit the creation and effects of those compounds. Although kinetic and regression models have been used in the past to assist in finding optimal process conditions, it is possible that neural networks could offer even further and easier generalization and prediction across unit operations and even feedstocks.

1.2 BACKGROUND

Liquid fuels made up 40 quadrillion Btu (40%) of the total energy consumed in the United States (U.S.) in 2007. Consumption of liquid fuel for the transportation sector in 2007 was 27.5 quadrillion Btu, 68% of the total liquid fuel consumption in the U.S., and is

projected to increase to 30.3 quadrillion Btu (73%) by 2030. During the same period of time, it is projected that domestic oil production will increase from 5.1 to about 7 million barrels a day while the consumption of liquid fuel for transportation will increase from 14.3 to 15.8 million barrels a day (EIA, 2009). In 2007, gasoline made up 65% of liquid transportation fuel use and for every 100 barrels of crude oil only about 47 barrels of gasoline are produced. If all 5.1 million barrels a day produced domestically were used to make gasoline it would only result in 2.4 million barrels a day of gasoline leaving a shortfall of 6.9 million barrels a day in demand. This gap between the domestic oil available for refinement into liquid fuels and the domestic demand for liquid fuels in the transportation sector is left to be filled by imported oil. Unfortunately, world demand for oil has grown over the last 20 years and is projected to continue to grow through the next 20 years. This is causing the price of oil to climb at increasing rate. It is projected that the price will increase by 25% from January 2009 to January 2011 (EIA, 2009). Given that oil production has already peaked domestically and is expected to peak worldwide within the next one to five decades, it is understood that the supply is limited. The price is only expected to increase further as a result of limited supply.

There are also environmental implications to using petroleum derived fuels. Oil is the result of natural processes converting organic matter (sequestered carbon) deposited millions of years ago and therefore burning oil releases carbon into the atmosphere which had been removed from the contemporary carbon cycle. Without a counterbalance, this carbon in the form of gasses builds up and has been linked to a variety of environmental impacts including increasing acidity in the oceans, melting of polar ice, and higher average temperatures

globally (IPCC, 2001). In 2007, 23% of the total carbon dioxide emitted by the US came from burning liquid fuels in the transportation sector (EIA, 2009). Using biological sources to produce fuels minimizes the global warming impact by creating a cycle where the carbon released while burning the fuel is equal to the carbon taken in by the next generation of biofuel crop (Difiglio, 1997).

1.3 POTENTIAL GASOLINE REPLACEMENTS

The liquid transportation fuel replacement candidates with the most potential include biomethanol, biobutanol, and bioethanol where the 'bio' prefix establishes these fuels are produced from a biomass source.

1.3.1 Biomethanol

Biomethanol is a one carbon alcohol historically generated as a byproduct of the conversion of wood into charcoal, earning it the nickname 'wood alcohol'. Currently, biomethanol is produced from biomass through a gasification process to generate synthesis gas followed by a reaction in the presence of a catalyst to produce methanol, however this process is not economically feasible. Additionally, methanol mixes well with water, is toxic to humans through ingestion and skin absorption, and it burns with an invisible flame making it potentially hazardous in comparison to other bio-alcohols (Demirbas, 2007).

1.3.2 Biobutanol

Biobutanol is a four carbon alcohol that is a metabolic product a number of microorganisms. Historically, industrial production has utilized *Clostridium acetobutylicum* because a starch feedstock could be used to produce acetone, butanol, and ethanol in a ratio of 3:6:1 with acetone being the primary product of interest initially (Antoni, 2007). Biobutanol has a number of advantages that make it an ideal gasoline replacement including energy content similar to gasoline, it can be used pure or as a mixture with gasoline in unmodified engines, and it separates from water (Durre, 2007). Some disadvantages researchers are currently trying to address are low limits of butanol tolerance in the production organisms, expensive downstream processing to separate the butanol from metabolic co-products, and the ability to use cellulosic feedstocks instead of starches (Ezeji et al., 2005; Antoni et al., 2007).

1.3.3 Bioethanol

Bioethanol is two carbon alcohol most notably generated from the six carbon sugar glucose by the yeast *Saccharomyces cerevisiae* as metabolic product. Five carbon sugars, like xylose, can also be converted to ethanol by genetically modified yeast strains and certain bacteria (Krishnan et al., 1999; van Maris, 2007; Zhang et al., 1995). Additionally, synthesis gas (syngas), a mixture of hydrogen, carbon monoxide, and carbon dioxide, generated from gasifying biomass can be converted to ethanol either through fermentation using bacteria or by chemical reaction using a catalyst (Henstra et al., 2007; Galvita et al., 2001). As a combustion fuel, ethanol burns cleaner than gasoline and when blended with gasoline it

provides oxygen, which aids in a more complete combustion. Because of these aspects, ethanol has been used as a gasoline additive (10% ethanol v/v) in the U.S. rather than the more toxic chemical fossil fuel derived methyl tertiary butyl ether (MTBE) (Browner, 2000). The main disadvantages to using ethanol are that it is hygroscopic and has about 66% of the energy as the same volume of gasoline. Despite this lack of comparable energy density, ethanol can still transport a vehicle 80% of the distance an equal volume of gasoline would be able to transport thanks to efficiency gains from more complete combustion (Lynd et al. 1991). In the end, the combination of a large scale production technology that is more mature than the alternative bio-alcohols, surmountable disadvantages, and the current use in domestic gasoline blends makes bioethanol the ideal candidate for a near term gasoline replacement.

1.4 FEEDSTOCK OPTIONS FOR BIOETHANOL PRODUCTION

As described above, bioethanol can be generated from syngas or from sugars. While syngas production provides the luxury of being nearly feedstock independent, the process currently has many issues keeping it from economic production with current technology (Lewis et al., 2008). This leaves bioethanol production from sugars. In general, there are three major sources of sugars which each require different process steps to first convert them into fermentable sugars and then into bioethanol.

1.4.1 Sugar Feedstocks

The simplest option is to just use sugar. Example feedstocks are sugar cane, sweet sorgum, and sugar beets. Sugars from these sources are already available in a fermentable state and as a result only mechanical processing is necessary to extract the sugars for fermentation by yeast. Brazil uses this platform very successfully and by coupling this technology with the appropriate infrastructure, Brazil is able to offer up to 100% ethanol fuels as a gasoline replacement inexpensively. There are some parts of the U.S. that can use sugar feedstocks for ethanol production, but the wider adoption of this sugar source is largely limited by climate constraints. Because of this, sugar feedstocks will not be able to meet the full requirements of gasoline replacement by bioethanol domestically.

1.4.2 Starch Feedstocks

Starch is a polymer of glucose molecules connected by α-1-4 and α -1-6 glycosidic bonds. This configuration makes the polymer appear like a set of stairs with each step a glucose molecule. This type of bonding between glucose molecules makes it easier for enzymes (amylases) to access and break these bonds. Therefore, starches are an ideal feedstock because after mechanical processing starches can be saccharified in a mixture of hot water and amylases. Examples of major starch feedstocks include food crops like barley, corn, potatoes, rice and wheat depending on locale. In the U.S. corn is currently the leading source for producing bioethanol (Solomon et al., 2007).

Although corn is easy to convert into bioethanol, there are limitations associated with using it that will keep it from being able to supply enough bioethanol to replace gasoline.

One of the most obvious problems with corn and many other starch sources is that they compete as food. This can negatively impact corresponding markets with inflated prices of the feedstock (Cohen, 2008). Another issue concerns the central locations where feedstocks are produced. In the case of corn in the U.S., the majority of production occurs in the central part of the country requiring transportation of the fuel to the coasts. Since ethanol is hygroscopic, it cannot be transported by pipeline like oil and must be transported by a more costly method in tankers (Wakeley et al., 2008). Corn also requires arable land and a significant amount of nitrogen to grow. Nitrogen fertilizer is produced using large amounts of natural gas and often much of the nitrogen in fertilizer is lost through run-off and volatilization nullifying much of the environmental benefit (Balkcom, 2003; Randall, 2003). Overall, corn is a good short-term source for bioethanol production, but it barely addresses environmental concerns and it will not be able to meet the supply requirements to replace gasoline alone.

1.4.3 Lignocellulosic Feedstocks

Lignocellulosic feedstocks offer a possible solution to the constraints faced with the sugar and starch feedstocks. Lignocellulose is the shorthand term for any biomass containing the three polymers hemicellulose, cellulose, and lignin. Hemicellulose is mostly made up by a backbone of the five carbon sugar xylose with side chains of other five and six carbon sugars like arabinose, glucose, and galactose. Cellulose is, like starch, a polymer of glucose molecules except that in cellulose the glucose bonds are β-1-4 glycosidic bonds. This means that instead of stairs, cellulose appears to be a straight line making it much more stable

(crystalline) and more difficult to degrade into glucose monomers. Lignin is a polymer made up of a variety of aromatic subunits and in a plant the general purpose of lignin is to prevent access to the structural carbohydrates (cellulose and hemicellulose) of the plant. Figure 1.1 below contains a simplified schematic of how lignin, hemicellulose, and cellulose are organized in lignocellulosic biomass. Example lignocellulosic feedstocks include agricultural, food processing and municipal wastes, perennial grasses, and woody biomass (Tan et al., 2008).

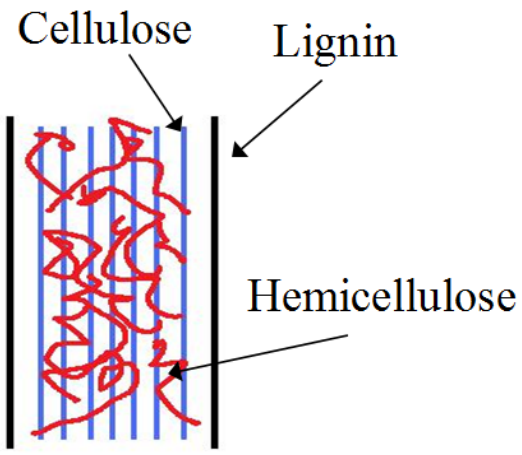

Figure 1.1 Organization of lignocellulosic biomass adapted from Mosier et al. (2005)

Most of the benefits attributed to lignocellulosic biomass are a result of the feedstock options. Lignocellulosic biomass is nearly ubiquitous on earth. This opens the possibility of local bioethanol production, which could keep costs to transport the fuel low. Additionally, many of the potential feedstocks are wastes or low-input crops that can be grown on marginal

lands. This keeps the feedstock prices much lower than corn and avoids fertilizer related environmental impacts. Since lignocellulosic feedstocks are not food resources nor do they require land used to grow food, there is no risk that either food or lignocellulosic feedstock prices will increase from competition.

The reason that lignocellulosic materials are not being widely used currently is because there are few barriers to using these feedstocks for bioethanol production. One barrier is that the enzymes (cellulases) needed to break down cellulose into glucose monomers are much more expensive than the amylases used to break down starches. The Department of Energy (DOE) funded four private companies in 2008 as part of a project to lower the price of cellulases per gallon of ethanol. While some decrease in the price was achieved previously, research into further lowering the price as well as alternative cellulase sources is still being pursued (DOE, 2008; Kumar et al., 2008). Another barrier is that the xylose component of the biomass is not fermentable using the same microorganism (*Saccharomyces cerevisiae*) used in current starch based ethanol production. Because hemicellulose can make up as much as 80% (w/w) of the biomass, fermenting the xylose can be important for making bioethanol production from lignocellulosic biomass economically feasible (Sun and Cheng, 2002). Research has yielded some gains in this area through the development of yeast and bacteria genetically altered to ferment xylose in addition to the application of some naturally occurring strains (Krishnan et al., 1999; van Maris, 2007; Zhang et al., 1995). The largest barrier keeping lignocellulosic biomass from being a feasible bioethanol feedstock is the recalcitrant nature of the biomass to grant access to the cellulose and hemicellulose components. A process step referred to as 'pretreatment' is required to disrupt the biomass

structure to make it accessible to hydrolysis by cellulases in order to release fermentable monomeric sugars the yeast or bacteria can use.

1.5 PRETREATMENT OPTIONS FOR LIGNOCELLULOSIC BIOMASS

Pretreatment, enzymatic hydrolysis, and fermentation are the three areas receiving the most attention through research as alluded to above. Of these, pretreatment is of significant importance because as a process step, it is upstream of both enzymatic hydrolysis and fermentation. A good pretreatment will disrupt the biomass enough to allow for the maximum hydrolysis of both the hemicellulose and cellulose components into monomeric sugars with minimal generation of enzymatic hydrolysis and fermentation inhibitors (Hu et al., 2008). The most often studied pretreatment technologies fall into four main categories which include physical pretreatments, chemical pretreatments, physio-chemical (combination) pretreatments, and biological pretreatments (Sun and Cheng, 2002).

1.5.1 Physical Pretreatments

Physical pretreatments include various methods of mechanical alteration like hammer and ball mills (Millet et al., 1976). Mechanical alteration is the action of grinding or chipping materials into smaller pieces which, in the case of lignocellulosic biomass, disrupts the biomass structure and increases surface area. The smallest particles are the most susceptible to enzymatic hydrolysis and although small particle sizes are achievable, the amount of energy increases greatly as the size of particles decreases. As a result, physical pretreatment alone is not economically feasible at larger scales (Galbe and Zacchi, 2007).

1.5.2 Chemical Pretreatments

Chemical pretreatments involve using alkali reagents, dilute acids, peroxide, organic solvents, or ozone to disrupt the structure of lignocellulosic biomass (Sun and Cheng, 2002). The alkaline reagents most often used in pretreatment applications are sodium hydroxide and calcium hydroxide (lime). The major mechanism of alkaline pretreatments is the saponification of the ester bonds between lignin and hemicellulose leading to the delignification of the biomass (Tarkow and Feist, 1969). Lime pretreatments can potentially lower costs because disruption of biomass structure can occur at or near ambient temperatures, however these reactions require residence times on the order of hours to days (Mosier et al., 2005; Playne, 1984). Dilute acid acts to hydrolyze hemicellulose out of the solid biomass. Although dilute acid pretreatment greatly increases the success of enzymatic hydrolysis, disadvantages include the formation of inhibitory compounds, expensive equipment made from stainless steel to resist corrosion and downstream neutralization of the acid prior to fermentation because a low pH is inhibitory to the growth of yeast and other microorganisms (Mosier et al., 2005; Sun and Cheng, 2002). Dilute peroxide can be used to disrupt lignin and hemicellulose bonds resulting in the removal of hemicellulose from the solid biomass at near ambient conditions. This improves resulting glucose yields from enzymatic hydrolysis to above 90%, but only after at least 8 hours of pretreatment time at 30ºC (Azzam, 1989). Organic solvents, like methanol, ethanol, or acetone can be coupled with an acid to disrupt the lignin and hemicellulose bonding in biomass. These solvents and acids need to be removed to allow for further downstream processing and should be recovered for economical benefit (Galbe and Zacchi, 2007). Ozone, like lime, can also be

used to affect lignin and hemicellulose bonding at ambient conditions. However, the amount of ozone required to significantly treat the biomass is not cost effective (Sun and Cheng, 2002).

1.5.3 Physio-Chemical Pretreatments

Combinations of physical and chemical pretreatments have also been investigated with typical examples being steam explosion, ammonia fiber explosion (AFEX), and carbon dioxide explosion. The general idea shared by each of these variations is a pressurization and heating of the biomass, forcing an intermediate into the structure, followed by a rapid depressurization to ambient pressure, which causes the biomass structure to explode (Sun and Cheng, 2002; Galbe and Zacchi, 2007). Steam has the added effect of heat which promotes autohydrolysis, or the formation of organic acids from the biomass, which then help to break up the structure of the biomass through the removal of hemicellulose. Dilute acid has also been coupled with steam explosion to improve the hydrolysis of hemicellulose (Hu et al., 2008; Tucker et al., 2003). AFEX is done at lower temperature and pressure than steam explosion, but takes advantage of ammonia which acts like an alkali reagent and removes lignin from the biomass structure without degrading the carbohydrates (Hu et al., 2008). Carbon dioxide explosion is a similar notion steam explosion and AFEX, except with lower yields (Sun and Cheng, 2002; Dale and Moreira, 1982). Zheng et al. (1998) reported lower inhibitor levels for carbon dioxide explosion compared to steam explosion and AFEX.

1.5.4 Biological Pretreatments

Organisms which naturally break down lignocellulosic biomass have been researched to identify if they could be applied as a pretreatment in large scale bioethanol production. Varieties of white-rot fungus have been the most popular organism investigated for this application because of effective and preferential lignin degradation that has been observed. The main benefits of biological pretreatments include low energy requirements, lower equipment costs, and lower water requirements. Unfortunately, the fungi may also metabolize a portion of the carbohydrates and pretreatment times are on the order of days to weeks which makes this type of pretreatment unsuitable for industrial scale production (Taniguchi et al., 2005; Hatakka, 1983).

1.6 DILUTE ACID PRETREATMENT: BENEFITS AND APPLICATIONS

Dilute acid pretreatment has a number of benefits that currently make it a better choice compared to the other pretreatment options outlined above and as a result, dilute acid pretreatment is first in line for commercial application in the production of bioethanol from lignocellulosics (Taherzadeh and Karimi, 2008). One benefit from a process standpoint is that dilute acid generates separable streams. There is a liquid pre-enzymatic hydrolysis stream containing a majority of xylose, a liquid post-enzymatic hydrolysis stream containing a majority of glucose, and a solid stream containing a majority of lignin. This removes the need for complex or costly unit processes for separation and allows for product specific unit processes. Additionally, acids, like sulfuric acid and acetic acid, are less expensive compared to other chemicals, most specifically alkalis (Mosier et al., 2005; Hu et al., 2008).

Eggeman (2005) reported dilute sulfuric acid pretreatment as cheaper per gallon of ethanol produced than sodium hydroxide, lime, or AFEX. The last major benefit, as outlined in the examples below, is that researchers have shown that using dilute acid and dilute acid combined with steam explosion are effective pretreatments prior to enzymatic hydrolysis across many types of lignocellulosic feedstocks resulting in high yields of monomers from both the hemicellulose and cellulose components.

1.6.1 Agricultural Residues

Agricultural residues are ideal feedstocks because they are already associated with a cropping system and being wastes keeps the costs low. Corn stover is one of the most researched feedstocks in this category. In a summary of studies using different pretreatment technologies on corn stover, dilute acid stands out as the only pretreatment yielding great than 90% of both xylose and glucose (Galbe and Zacchi, 2007). The dilute acid pretreatment referenced is from a study done by Lloyd and Wyman (2005) where at solid loading of 10%, 91% of theoretical glucose and xylose were recovered after a pretreatment of the corn stover at 160ºC and 0.49% sulfuric acid over a residence time of 20 minutes and an enzymatic hydrolysis step. Tucker et al. (2003) showed that combining dilute sulfuric acid with steam explosion to pretreat corn stover at a solid loading of 46% yielded greater than 90% xylose and glucose for a pretreatment at 190ºC and 1.1% (w/w) acid for a residence time of 90 seconds. Another agricultural residue example, wheat straw, was pretreated with 0.75% sulfuric acid (v/v) at 121ºC for 60 minutes at about a 7% solid loading and was shown to yield 74% of total sugars after enzymatic hydrolysis (Saha et al. 2005). Saha et al. (2008)

improved the yield of total sugars from wheat straw to 84% using a microwave for heating to 160ºC with a sulfuric acid concentration of 0.5% (w/v) for 10 minutes.

1.6.2 Woody Biomass

The main sources of woody biomass can broadly be grouped as forest residuals and wood chips. In general, woody biomass tends to respond better to the combination of dilute acid pretreatment with steam explosion than just dilute acid alone (Hu et al., 2008). In an example of this dilute acid pretreatment coupled with steam explosion, Emmel et al. (2003) was able to recover 70% of hemicellulose sugars using 0.175% acid at 210ºC for a two minute pretreatment time. In the same study, a 90% cellulose conversion was achieved at 200ºC under similar pretreatment conditions. Examining only xylose recovery, Esteghlalian et al. (1997) were able to achieve close to 90% of the theoretical yield from poplar under pretreatment conditions between 170-180ºC using in excess of 0.9% sulfuric acid and residence times of minutes.

1.6.3 Herbaceous Biomass

Generally, switchgrass stands out as a popular choice of research because of a high yield of biomass per hectare (5.2 - 11.1 Mg/ha dry), the option to cultivate it on marginal land, and a potentially favorable net energy gain (Schmer et al., 2007). Young switchgrass pretreated with 1.5% acid at 120ºC for 60 minutes was shown to yield close to 80% glucose and above 90% non-glucose (xylose and arabinose) sugars. At 1.5% acid and 150ºC for 20 minutes (with 10 minute preheat) young switchgrass yielded about 85% glucose and 80%

non-glucose sugars (Dien et al., 2006). Switchgrass was also pretrated at 180ºC and 1.2% (w/w) sulfuric acid to yield 90% glucose from cellulose after enzymatic hydrolysis.

However, switchgrass is not the only possibility of an herbaceous energy crop and therefore other options more specific to certain locales have been investigated as well. Vegetative stage reed canary grass was pretreated with 1.5% acid at 120ºC for 60 minutes to yield above 80% glucose and 90% non-glucose sugars (Dien et al., 2006). Sun and Cheng (2005) were able to show the enzymatic glucose conversion of about 82% for coastal bermudagrass, a grass grown for hog waste nutrient management in the southeastern U.S., after pretreatments at 1.5% sulfuric acid (w/w) at 121ºC for 60 and 90 minutes. Corresponding xylose conversion, however, was only about 60%.

1.7 DEGRADATION PRODUCTS

Dilute acid, especially when coupled with high temperatures, unfortunately has a propensity to degrade biomass compounds into products which are inhibitory to downstream processes like enzymatic hydrolysis and fermentation. Inhibitory levels of these compounds, which will be discussed further below, are generally not very high and therefore restrict the severity of pretreatment conditions feasible for successful downstream process. A balance between severity and yield can be difficult to reach and some researchers have generated greater higher yields for monomeric sugars from hemicellulose and cellulose, only to find that the yeast are inhibited and those pretreatment conditions cannot be used to make bioethanol (Laser et al., 2002). The inhibitory levels of each of these compounds and the

corresponding degrees of inhibition are not as clear because of interaction effects between inhibitors and other factors like yeast concentration (Palmqvist and Hahn-Hagerdahl, 2000).

1.7.1 Sugar Degradation Products

During high severity pretreatment, as pentose polymers are hydrolyzed into monomers, those monomers can be further degraded into furfural. After continued exposure to severe conditions, furfural can be degraded into levulinic acid. The same process occurs with hexoses as well, generating hydroxymethylfurfural (HMF) and formic acid respective of exposure to pretreatment conditions. Equations 1 and 2 below outline the path of sugar degradation. Generally, furfural and HMF have been reported as causing a lag phase in

Hemicellulose → Pentose → Furfural → Levulinic Acid (1)

Cellulose and Hemicellulose → Hexose → HMF → Formic Acid (2)

yeast growth before sugar consumption begins because the yeast take up the two degradation compounds first (furfual faster than HMF) before moving on to converting sugars to ethanol. Levulinic and formic acids are generally shown as helpful at levels up to 100 mmol and inhibitory after levels from 100-200 mmol. This has been confirmed by experiments done by Larsson et al. (1998) which showed furfural and HMF as not inhibitory to ethanol yield, while weak acids (formic, levulinic, and acetic) at a combined concentration greater than 100 mmol did inhibit ethanol yield. Navarro (1994) showed the increased inhibition of yeast productivity and growth rate corresponding to higher levels of furfural. He also reported on

how increasing the initial yeast concentration could decrease the inhibitory effects of furfural. In a review of furfural and HMF inhibition studies, Almeida et al. (2009) shows that the pooled research on this topic confirms some level of inhibition to yeast and ethanol production at a variety of furfural, HMF, and initial yeast concentrations.

1.7.2 Lignin Degradation Products

Just as lignin molecules can include a variety of different chemical compounds, the resulting degradation compounds possible are just as varied. In general, lignin degrades into phenolic compounds with a variety of molecular weights (Klinke et al., 2004). Details regarding phenolic inhibition are limited due to a lack of accurate investigation, but it is suspected that low molecular weight phenolic compounds are more inhibitory to fermentation and that there are interaction effects with furfural and HMF which increase overall inhibitory effects of both compounds (Palmqvist and Hahn-Hagerdahl, 2000).

1.8 MODELING

Multiple linear regression models and kinetic models have been developed across many types of biomass and pretreatment condition ranges (temperatures, acid concentrations, and residence times). Kinetic models have been shown to accurately portray the relationship between inputs to pretreatment and enzymatic hydrolysis and corresponding product outputs (xylose, furfural, glucose). Artificial neural networks (ANN), which hold the potential of generalization beyond across these models as well as predictive application, have been neglected. Only one study, by O'Dwyer et al. (2008), has addressed the conversion of

lignocellulosic biomass into ethanol. This study uses an ANN to correlate biomass composition with enzymatic digestibility and with a successful generalization, makes a recommendation to use ANNs in further applications. In the conclusion of the review by Mosier et al. (2005) a recommendation is made for further development of predictive modeling.

1.9 CONCLUSION

Bioethanol appears to be a good gasoline replacement to start with until a better liquid alternative becomes more feasible. Research using dilute acid pretreatment, especially when coupled with steam explosion, across all types of lignocellulosic feedstocks should continue in order to refine optimal pretreatment conditions that maximize sugars and minimize inhibitory compounds. Commercial scale adoptions of dilute acid pretreatment should also be used to further refine process conditions for the benefit of process economics. It is possible that artificial neural networks (or other mathematical models underused in this field) can be used to assist with modeling and optimizing the complex relationships and should be further pursued in research.

REFERENCES

Almeida, J., Bertilsson, M., Gorwa-Grauslund, M., Gorsich, S., Gunnar, L., 2009. Metabolic effects of furaldehydes and impacts on biotechnological processes. *Applied Micribiol. Biotechnology.* 82: 625-638.

Antoni, D., Zverlov, V., Schwarz, W., 2007. Biofuels from microbes. *Applied Micriobiolgy Biotechnology*. 77:32-35.

Azzam, A.M., 1989. Pretreatment of cane bagasse with alkaline hydrogen peroxide for enzymatic hydrolysis of cellulose and ethanol fermentation. *Journal of Environmental Science and Health. Part B, Pesticides, Food Contaminants, and Agricultural Wastes.* 24 (4), 421-433.

Balkcom, K. S., Blackmer, A. M., Hansen, D. J., Morris, T. F., and Mallarino, A. P., 2003. Testing soils and cornstalks to evaluate nitrogen management on the watershed scale. *Journal of Environmental Quality.* 32 (3):1015-1024.

Browner, C., 2000. Remarks as prepared for delivery to press conference on March 20, 2000. Available from: http://www.epa.gov/otaq/consumer/fuels/mtbe/press34b.pdf.

Cohen, S., 2008. Power Politics: Energy News, Policy, and Analysis. Available from: http://www.thenewatlantis.com/blog/power-politics/ethanol-and-the-corn-we-eat

Dale, B.E., Moreira, M.J., 1982. A freeze-explosion technique for increasing cellulose hydrolysis. *Biotechnology and Bioengineering Symposium.* 12, 31-43.

Demirbas, A., 2007. Progress and recent trends in biofuels. *Progress in Energy and Combustion Science.* 33:1-18.

Department of Energy (DOE), 2008. News release for 2-26-2008 available from http://www.energy.gov/news/6015.htm

Difiglio, C., 1997. Using advanced technologies to reduce motor vehicle greenhouse gas emissions. *Energy Policy.* 25 (14-15): 1173-1178.

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

Durre, P., 2007. Biobutanol: An attractive biofuel. *Biotechnology Journal*. 2:1525-1534.

Dwyer J., Zhu, L., Granda, C., Chang, V., Holtzapple, M., 2008. Neural Network Prediction of Biomass Digestibility Based on Structural Features. *Biotechnolgy Progress.* 24: 283-292.

Eggeman, T. and Elander, R., 2005. Process and economic analysis of pretreatment technologies. *Bioresource Technology.* 96: 2019-2025.

Emmel, A., Mathias, A.L., Wypych, F., Ramos, L.P., 2003. Fractionation of Eucalyptus grandis chips by dilute acid-catalyzed steam explosion. *Bioresource Technology. 86*: 105-115.

Energy Information Administration (EIA), 2009. Annual energy outlook 2009. Available from http://www.eia.doe.gov/oiaf/aeo/index.html

Esteghlalian, A., Hashimoto, A.G., Fenske, J.J., Penner, M.H., 1997. Modeling and optimization of the dilute-sulfuric-acid pretreatment of corn stover, poplar and switchgrass. *Bioresource Technology.* 59, 129-136.

Ezeji, T., Karcher, M., Qureshi, N., Blachek, H., 2005. Improving performance of gas stripping-based recovery system to remove butanol from *Clostridium beijerinckii* fermentation. *Bioprocess Biosystems Engineering.* 27: 207-214.

Galbe, M., Zacchi, G., 2007. Pretreatment of lignocellulocis materials for efficient bioethanol production. *Adv Biochem Engin/Biotechnol.* 108:41-65

Galvita, V.V., Semin, G.L., Belyaev, V.D., Semikolenov, V.A., Tsiakaras, P., Sobyanin, V.A., 2001. Synthesis gas production by steam reforming of ethanol. *Applied Catalysis A: General.* 220, 123-127.

Hatakka, A.I., 1983. Pretreatment of wheat straw by white-rot fungi for enzymic saccharification of cellulose. Eur. J. *Appl. Microb. Biotechnol.* 18 (6): 350-357.

Henstra, A., Sipma, J., Rinzema, A., Stams, A., 2007. Microbiology of syngas fermentation of biofuel production. *Current Opinion in Biotechnology.* 18:200-206.

Hu, G., Heitmann, J., Roja, O., 2008. Feedstock pretreatment strategies for producing ethanol from wood, bark, and forest residues. *BioResources.* 3(1): 270-294.

IPCC, 2001: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T.,Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K., Maskell, and C.A. Johnson (eds.)]. Cambridge University Press, Cambridge, United Kingdom and NewYork, NY, USA, 881pp.

Klinke, H., Thomsen, A., Ahring, B., 2004. Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Appl. Microbiol. Biotechnol.* 66: 10-26.

Krishnan, M.S., Ho, N.W.Y., Tsao, G.T., 1999. Fermentation kinetics of ethanol production from glucose and xylose by recombinant *Saccharomyces* 1400 (pLNH33). *Applied Biochemistry and Biotechnology.* 77-79: 373-388.

Kumar, R., Singh, S., Singh, O.V., 2008. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *Journal of Industrial Microbiology and Biotechnology.* 35: 377-391.

Larsson, S., Reimann, A., Nilverbrant, N., Jonsson, L., 1999. Comparison of different methods for the detoxification of lignocellulose hydrolyzates of spruce. *Applied Biochemistry and Biotechnology*. 77-79: 91-103.

Laser, M.; Schulman, D.; Allen, S.G.; Lichwa, J.; Antal, M.J., Jr.; Lynd, L.R. A comparison of liquid hot water and steam pretreatments of sugar cane bagasse for bioconversion to ethanol. *Bioresource Technol.* 2002, *81*: 33-44.

Lewis, R.S., Frankman, A., Tanner, R.S., Ahmed, A., Huhnke, R.L., 2008. Ethanol via biomass-generated syngas. *International Sugar Journal.* 110: 150-+.

Lloyd, T., Wyman, C., 2005. Combined sugar yields for combined dilute acid pretreatment of corn stover followed by enzymatic hydrolysis of the remaining solids. *Bioresource Technology.* 96: 1967-1977.

Lynd, L.R., Cushman, J.H., Nichols, R.J., Wyman, C.E., 1991. Fuel ethanol from cellulosic biomass. *Science.* 251 (4999): 1318-1323.

Millet, M.A., Baker, A.J., Scatter, L.D., 1976. Physical and chemical pretreatment for enhancing cellulose saccharification. *Biotech. Bioeng. Symp*. 6: 125–153.

Moiser, N., Wyman, C., Dale, B., Elander, R., Lee, Y.Y., Holtzapple, M., Ladisch, M., 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology.* 96: 673-686.

Navarro, A., 1994. Effects of Furfural on Ethanol Fermentation by Saccharomyces cerevisiae: Mathematical Models. *Current Microbiology.* 19: 87-90.

Palmqvist, E. and Hahn-Hagerdah, B., 2000. Fermentation of lignocellulosic hydrlyzates. II: inhibitors and mechanisms of inhibition. *Bioresource Technology.* 74: 25-33.

Playne, M.J., 1984. Increased digestibility of bagasse by pretreatment with alkalis and steam explosion. *Biotechnology and Bioengineering.* 26 (5): 426-433.

Randall, G.W., Vetsch, J.A., and Huffman, J.R., 2003. Nitrate losses in subsurface drainage from a corn-soybean rotation as affected by time of nitrogen application and use of Nitrapyrin. *J. Environ. Qual.* 32: 1764–1772.

Saha, B.C., Iten, L.B., Cotta, M.A., Wu, Y.V., 2005. Dilute acid pretreatment, enzymatic saccharification and fermentation of wheat straw to ethanol. *Process Biochemistry.* 40 (12): 3693-3700.

Saha, B.C., Biswas, A., Cotta, M.A., 2008. Microwave pretreatment, enzymatic saccharification and fermentation of wheat straw to ethanol. *Journal of biobased materials and bioenergy.* 2 (3): 210-217.

Schmer, M.R., Vogel, K.P., Mitchell, R.B., Perrin, R.K., 2008. Net energy of cellulosic ethanol from switchgrass. *Proc. Natl. Acad. Sci. U. S. A.* 105 (2): 464-469.

Solomon, B., Barnes, J., Halvorsen, K., 2007. Grain and cellulosic ethanol: History, economics, and energy policy. *Biomass and Bioenergy.* 31: 416-425.

Sun, Y., Cheng, J., 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology.* 83: 1-11.

Sun, Y., Cheng, J., 2005. Dilute acid pretreatment of rye straw and bermudagrass for ethanol production. *Bioresource Technology.* 96: 1599-1606.

Tan, K.T., Lee, K.T., Mohamed, A.R., 2008. Role of energy policy in renewable energy accomplishment: The case of second-generation bioethanol. *Energy Policy*. 36 (9): 3360-3365.

Taniguchi, M., Suzuki, H., Watanabe, D., Sakai, K., Hoshino, K., Tanaka, T., 2005. Evaluation of Pretreatment with Pleuotus ostreatus for Enzymatic Hydrolysis of Rice Straw. *Journal of Bioscience and Bioengineering.* 100 (6): 637-643.

Tarkow, H., Feist, W.C., 1969. In: A Mechanism for Improving the Digestibility of Lignocellulosic Materials with Dilute Alkali and Liquid NH3 Advance Chemistry Series 95. American Chemical Society, Washington, DC, 197–218.

Taherzadeh, M., Karimi, K., 2008. Pretreatment of lignocellulosic wastes to improve ethanol and biogass production: a review. *Int. J. Mol. Sci.* 9: 1621-1651.

Tucker, M. Kim, K., Newman, M., Nguyen, Q., 2003 Effects of temperature and moisture on dilute-acid steam explosion pretreatment of Corn Stover and Cellulase Enzyme Digestibility. *Applied Biochemistry and Biotechnology.* 105-108: 165-177.

van Maris, A., Winkler, A., Kuper, M., de Laat, W., Dijken, J., Pronk, J., 2007. Development of Efficient Xylose Fermentation in Saccharomyces cerevisiae: Xylose Isomerase as a Key Component. *Advanced Biochemial Engineering/Biotechnology.* 108: 179-204.

Wakeley, H., Griffin, M., Hendrickson, C., Matthews, H.S., 2008. Alternative Transportation Fuels: Distribution Infrastructure for Hydrogen and Ethanol in Iowa. *J. Infastruc. Syst.* 14: 262-271.

Zhang, M., Eddy, C., Deanda, K., Finkelstein, M., Picataggio, S., 1995. Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*. *Science*. 267, 240–243.

Zheng, Y.Z., Lin, H.M., Tsao, G.T., 1998. Pretreatment for cellulose hydrolysis by carbon dioxide explosion. *Biotechnology Progress.* 14, 890-896.

CHAPTER 2

HIGH TEMPERATURE DILUTE ACID PRETREATMENT OF COASTAL BERMUDAGRASS

2.1 ABSTRACT

Dilute sulfuric acid pretreatment was used to pretreat coastal bermudagrass prior to enzymatic hydrolysis. After both steps, the highest yield of total sugars (combined xylose and glucose) was 97%, which was shared by the pretreatment condition combinations of 0.9% acid at 160ºC for 15 minutes and 1.2% acid at 160ºC for 5 minutes. The prehydrolyzate liquor was analyzed for inhibitory compounds (furfural, hydroxymethylfurfural (HMF)) in order to assess potential risk for inhibition during fermentation. Accounting for the inhibitory compounds, a pretreatment with 1.2% acid at 140ºC for 30 minutes with a total sugar yield of 94% or 0.9% acid at 160 ºC for five minutes with a total sugar yield of 91% may be more favorable for fermentation because furfural levels remain under the inhibitory threshold concentration of 1 g/L. From this study, it can be concluded that dilute sulfuric acid pretreatment can be successfully applied to coastal bermudagrass to achieve high yields of monomeric glucose and xylose with acceptable levels of inhibitory compound formation.

2.2 INTRODUCTION

Due to a lack of resource sustainability of fossil fuels as well as negative environmental effects from emissions, research is being done to find a liquid fuel for use as a gasoline replacement. In the United States (U.S.) ethanol appears to be a good near term

option. Ethanol is already mixed with gasoline at a level of 10% and support from vehicle manufacturers has resulted in vehicles that can use up to an 85% ethanol-gasoline mixture. Currently, ethanol is mainly produced from corn starch in the U.S., but competition with corn for food and limitations on arable land for expanding the cultivation of corn place a ceiling on the ethanol producible lower than what is required to replace gasoline (Sun and Cheng, 2002). This has spurred research into cellulosic feedstocks for use in ethanol production. Lignocellulosic material is abundant nearly everywhere opening the potential for different geographic regions to take advantage of locally abundant cellulosic feedstocks for use in ethanol production. While corn ethanol simply requires a hot water pretreatment followed by enzymatic hydrolysis using inexpensive amylases, lignocellulosic materials require a harsher pretreatment step followed by hydrolysis using more costly cellulases (Solomon et al., 2007). Currently, pretreatment, enzymatic hydrolysis, and fermentation are the key areas being researched to make lignocellulosic feedstocks competitive with corn for the production of ethanol and this study focuses on the pretreatment step. Of the many pretreatment technologies covered by a number of review papers, dilute acid pretreatment stands out having been already examined in a large number of studies as well as appearing more economically feasible at larger scale than other current pretreatment technologies (Sun and Cheng, 2002; Mosier et al. 2005; Eggeman and Elander 2005).

In the southeastern U.S., coastal bermudagrass (*Cynodon dactylon* L.) is grown by hog producers as part of required nutrient management practices to mitigate nitrogen and phosphorus losses from hog waste. With a yield of 6-10 dry tons per acre each year and an active cropping system in place, bermudagrass is potentially a viable lignocellulosic

feedstock option for ethanol production in the southeastern U.S. (Holtzapple et al., 1994). Sun and Cheng (2005) investigated the dilute sulfuric acid pretreatment of coastal bermudagrass at 121ºC over a range of acid concentrations (0.6% - 1.5% w/w) and residence times (30 - 90 minutes). Optimal pretreatment conditions were found to be 1.2% acid for 60 minutes yielding about 70% of the total sugars from the biomass. Other studies using dilute acid pretreatment with different feedstocks have shown that increasing temperature increases sugar yields and can decrease the residence time and acid concentration required to generate the same yields at lower temperatures and longer residence times (Lloyd and Wyman, 2005; Saha et al., 2005; Saha et al., 2008). Based on results from Sun and Cheng (2005) and the review of research done on other feedstocks using dilute acid at temperatures higher than 121ºC, this study was performed to examine if higher sugar yields from coastal bermudagrass could be generated at higher temperatures, to identify the effects of temperature, concentration, residence time, and the interactions between these factors on sugar yields, and to assess what pretreatment conditions would be best in the pretreatment of coastal bermudagrass.

2.3 MATERIALS AND METHODS

2.3.1 Biomass Handling and Storage

Coastal bermudagrass was obtained in 2007 from Central Crops Research Station located in Clayton, NC, courtesy of Dr. Joe Burns of the Crop Science Department at North Carolina State University. The collected bermudagrass was stored in loose bale at ambient room temperature in a large plastic bag. When required for pretreatment, the bermudagrass

was ground to particle sizes no greater than 2 mm using a Wiley Laboratory Mill (Thomas, Model No. 4) fitted with a 2 mm screen. The resulting ground biomass was stored in sealed bags at ambient room temperature in the lab until used.

2.3.2 Composition Analysis of Raw Biomass

Prior to any pretreatment, the bermudagrass was analyzed for its composition. The moisture content was determined using a 105ºC oven to dry the biomass according to the procedure outlined in NREL laboratory analytical procedure (LAP) "Determination of Total Solids in Biomass" (Sluiter, 2005a). Extractive components like waxes, chlorophyll, non-structural sugars, and other minor components were quantified using a Soxhlet extractor, a 2:1 toluene-ethanol mixture, and 24 hour reflux time (Silverstein, 2004; Han and Rowell, 1997). The extractive-free biomass was then examined for the structural carbohydrates glucan, xylan, arabinan, and galactan as well as acid insoluble lignin (AIL) and acid soluble lignin (ASL) using the NREL LAP "Determination of Structural Carbohydrates and Lignin in Biomass" (Sluiter, 2006). The ash content of the biomass was found using a benchtop muffle furnace in accordance with NREL LAP "Determination of Ash in Biomass" (Sluiter, 2005b).

2.3.3 Pretreatment

A factorial experimental design was developed to examine the effect of the pretreatment conditions, reaction temperature, acid concentration and residence time, on the sugar yields in the both the prehydrolyzate and the hydrolyzate liquor as well as the generation of sugar degradation products in the prehydolyzate liquor. Factor levels of the

experimental design are based on literature review and prior work on the dilute acid pretreatment of bermudagrass done by Sun and Cheng (2005). Sulfuric acid concentrations of 0.3, 0.6, 0.9, and 1.2% (w/w) were examined at temperatures of 120, 140, 160, and 180ºC and residence times of 5, 15, 30, and 60 minutes.

Dilute sulfuric acid pretreatment was performed in vessels constructed from 306 stainless steel pipe pieces with a 0.75 inch inner diameter by 4 inch long purchased from McMaster Carr. Sulfuric acid was prepared on a w/w basis from a 96% stock solution beforehand and stored in a reagent bottle for continued use. Six vessels (two treatments in triplicate) were loaded at a ratio of 10 parts acid to one part raw biomass (w/w) (30 ml acid, 3 g biomass). The 7th vessel was loaded with de-ionized (DI) water instead of dilute acid in the same 10% solids loading as the other vessels and sealed with a cap containing a plug k-type thermocouple (Omega, model TC-K-NPT-G-72-SMP). This vessel served as a temperature monitor for the approximate internal temperature of the other reaction vessels.

A Fisher Scientific High-Temp Bath (Fisher Scientific, model 160A) was used to indirectly heat the vessels with silicone oil as a heat transfer fluid. A k-type thermocouple temperature probe was submerged in the oil in order to monitor the oil temperature and both this probe and the probe in the vessel were connected to an Extech Easyview Thermometer (Extech, model 11A) which displayed the temperature readings. The oil was heated to 20ºC higher than the desired set-point for the reaction to shorten preheat time. Once the oil reached the desired temperature, the seven vessels were added to the bath and the time of this addition was recorded. The blank vessel's internal temperature was then monitored along with the temperature of the oil until both reached the desired reaction temperature at which

time it was inferred that all of the vessels were at the desired temperature. Heating was manually controlled using the dial on the oil bath to turn the heating element on and off in order to keep the time it took to heat the vessels to the desired temperature to 12 minutes plus or minus 2 minutes. Once the proper internal temperature was reached, the residence time began and upon completion of the residence time, the vessels were quickly removed from the oil bath and placed in a bin of cool water where they were cooled to less than 40ºC within five minutes.

After pretreatment, the vessels were opened up for the filtration step. A Welch Duo-Seal vacuum pump (Sargent-Welch Scientific Co., model 1400) was used to pull a vacuum across a Whatman P8 filter paper in a standard vacuum flask and buchner filter funnel setup. The contents of each vessel were emptied on to a pre-weighed filter paper associated with that particular sample. 60 ml of DI water at room temperature was used to rinse the biomass and ensure the collection of all the sugars hydrolyzed in the pretreatment step. The total amount of filtrate and wash water was measured in a graduated cylinder and a portion was stored in labeled 15 ml tubes and frozen at -20ºC. The solids on the filter paper were rinsed with an additional 140 ml of room temperature DI water and then placed into an appropriately labeled bag for storage at 4ºC. These bags were weighed before and after the addition of biomass for use in calculating solid loss later. The filter paper was dried for 24 hours at 105ºC and then weighed to account for biomass remaining on the filter paper.

2.3.4 Enzymatic Hydrolysis

Enzymatic hydrolysis was preformed in 50 ml centrifuge tubes maintained at 55ºC and 155 rpm agitation for 72 hours by an automated shaking water bath (New Brunswick, model C76). Each tube was first loaded with the retained rinsed pretreated biomass weighed to account for moisture so that each tube received the equivalent of 0.5 grams of dry biomass. Novozymes North America (Franklinton, NC) provided the cellulase, NS-50013, and the cellobiase, NS-50010, which were determined using an assay procedure derived from Ghose (1987) to have activities of 76.44 FPU/ml (filter paper units, where one unit equates to 1 μmol of cellulose in the form of filter paper converted to glucose per minute) and 283.14 CBU/ml (cellobiase units, where one unit equates to 1 μmol of cellobiose converted to glucose per minute) respectively. The enzymes were loaded in excess at 40 FPU of cellulase per gram of dry biomass and 70 CBU of cellobiase per gram of dry biomass to avoid any limitation in monomeric sugar production caused by enzyme deficiency. Sodium azide, 0.3% (w/v), was added to inhibit microbial growth. Accounting for the amount of moisture in the biomass, the sodium azide, and the enzymes, the total amount of liquid in each tube was brought up to 15 ml using a 0.05 M sodium citrate buffer to maintain a pH of 4.8 during the enzymatic hydrolysis step.

2.3.5 Sugar Analysis

The hydrolyzate and prehydrolyzate were analyzed using a high performance liquid chromatography system (HPLC) to quantify the sugar monomers. The HPLC was also used to quantify prehydolyzate sugar degradation products. An Aminex HPX-87P column

coupled with a refractive index detector was used to distinguish amounts of glucose, xylose, galactose, arabinose in the hydrolyzate samples. Samples were prepared from the pH 4.8 hydrolyzate, diluted by four times, and filtered through a 2 μm syringe filter. The column was run at 80ºC with a flow rate of 0.6 ml/min using HPLC grade water as a mobile phase. Samples injections were 10 μL and the runtime was 35 minutes. A 25 minute post-runtime was included to clear late-eluting compounds from the column. For the prehydrolyzate samples, an Aminex HPX-87H column was used to quantify the levels of glucose, xylose, furfural, 5-hydroxyfuranmethal (HMF), forminc acid, and levulinic acid. Samples were prepared from the prehydrolyzate which had a pH of less than two and filtered through a 2 μm syringe filter. This column was run at 65ºC with a flow rate of 0.7 mL/min. Sample injections were run for 51 minutes. Total sugars, monomeric sugars, and degradation products were calculated on a per gram of raw biomass basis.

2.3.6 Lignin Analysis of Pretreated Biomass

The retained solid biomass was analyzed for acid insoluble lignin and acid soluble lignin composition in the same manner as the raw biomass was analyzed using the NREL LAP "Determination of Structural Carbohydrates and Lignin in Biomass" (Sluiter, 2006).

2.3.7 Statistical Analysis

Although triplicates were collected, the data had a number of outliers which appeared to be due to unequal heating during pretreatment, enzymatic hydrolysis error, or analysis error. As a result many triplicates had two sets of measured values very close and one set

with a clear difference. To remedy this, one sample was removed systematically from each triplicate. The remaining data was run through the GLM procedure in SAS 9.1.3 (SAS Institute Inc., Cary, NC) to perform an analysis of variance (ANOVA) to identify statistically significant effects from each of the factors as well as interactions of the factors. The interaction effect combinations included two-way interactions between temperature and concentration, temperature and time, concentration and time, and the three-way interaction between temperature, concentration, and time. The data was adjusted using Tukey's adjustment and significant differences were evaluated where $p < 0.01$ ($\alpha = 0.01$). The SAS code is included for reference in Appendix A. MATLAB 7.6.0.324 R2008a (The MathWorks, Natick, MA) was used to generate contour plots for visual analysis of the data.

2.4 RESULTS AND DISCUSSION

2.4.1 Composition of Coastal Bermudagrass

The bermudagrass was examined for carbohydrate, lignin, extractive, and ash composition by Wang (2009) and Sun and Cheng (2005) with the resulting dry weight percentages shown in table 2.1. The two sets of composition data are different with the set from Sun and Cheng reporting a greater amount of sugar with 6.77% more glucan and 3.49% more xylan than the set reported by Wang. This difference can likely be attributed to the difference in harvest time, crop inputs, or the fact that Wang did the analysis on extractive-free biomass rather than the raw biomass. In removing the extractives, non-structural carbohydrates could have also been removed. This study was run in parallel with the study done by Wang (2009) and used the same bermudagrass stock and composition data.

Table 2.1 Composition of Coastal Bermudagrass

Component	% weight of dry biomass (Wang)	% weight of dry biomass (Sun and Cheng)
Glucan	25.59	32.36
Xylan	15.88	19.37
Arabinan	1.95	4.33
Galactan	1.46	1.09
Acid insoluble lignin	15.37	20.33
Acid soluble lignin	3.96	ND
Extractives	4.17	ND
Ash	6.6	4.17

Glucan, xylan, arabinan, and galactan can be converted respectively to glucose, xylose, arabinose, and galactose using anhydro corrections of 0.90 for six-carbon sugars and 0.88 for five-carbon sugars which accounts for the removal of a water molecule. In this study arabinose and galactose concentrations in both prehydrolyzate and hyrolyzate are negligible. References to 'total sugars' will only include the sum of glucose and xylose for simplification. The theoretical maximum for glucose is 284.3 mg per g of raw biomass. The theoretical maximum for xylose is 180.5 mg per g of raw biomass. The theoretical maximum for total sugars is therefore 464.8 mg per g of raw biomass. Further yield calculations will be based upon these theoretical values.

2.4.2 Dilute Acid Pretreatment

The pretreatment step had two major outputs, the pretreated bermudagrass solids and the prehydrolyzate liquor. The solids were analyzed for both acid insoluble lignin and acid

soluble lignin. The prehydrolyzate liquor was analyzed for carbohydrates and carbohydrate degradation products.

Lignin Reduction

Dilute acid pretreatment works primarily by hydrolyzing hemicellulose into sugars (mostly xylose) from the biomass structure (Mosier et al., 2005). However, lignin also undergoes simultaneous degradation and accumulation reactions during an acid pretreatment (Pearl, 1967). Yang and Wyman (2004) showed that the net prevailing reaction in a batch reactor appeared to correspond to the severity of the acid pretreatment. While intermediate pretreatment severities tended to correspond with higher measured lignin removal (20-30%), there was also data showing little or no lignin removal. Additionally, both lower and higher severity acid pretreatments tended to show either no lignin removal or apparent lignin accumulation above that measured in the initial biomass composition analysis. Yang and Wyman (2004) suggested that these unclear lignin measurements were a complication of using a batch reactor. They showed that using a flowthrough reactor where solubilized lignin was removed from the system as the pretreatment progressed resulted in measured lignin removal of up to 85% that was clearly positively linearly correlated with pretreatment severity. Since batch reactors were used in this study, the lignin values measured after pretreatments follow the general trend presented by Yang and Wyman (2004) for batch reactors, which made it difficult to elucidate the specific effects of the pretreatment on the lignin portion of the biomass. As a result, acid soluble lignin and acid insoluble lignin data have not been reported in the body of this report, but have instead been included in Appendix

B for reference, however, it can be inferred that acid pretreatment in a batch reactor disrupts the structure of lignin.

Acid Hydrolysis of Biomass

Xylose is the major carbohydrate in the prehydrolyzate liquor and the largest component of hemicellulose. The highest yield of xylose measured in the prehydolyzate, 83%, was generated after a pretreatment at 1.2% acid, 140ºC, and 30 minutes residence time. However, xylose yields at 140ºC for 30 and 60 minutes and 160ºC for 5 and 15 minutes at concentration of 0.9% acid as well as 140ºC for 60 minutes and 160ºC for 5 minutes at concentration of 1.2% acid were found to be statistically similar to the highest yield. Figure 2.1 contains four subplots corresponding to each of the four acid concentration levels. Each subplot shows the monomeric xylose measured in the prehydrolyzate for each combination of temperature and residence time. An examination of the plots shows indications of interaction effects that could explain why several sets of pretreatment combinations generate similarly high xylose yields. Table 2.2 shows the ANOVA output for several of the major response variables considered in this study, including xylose. The ANOVA output for xylose confirms the significance of each of the factors as well as each of the interaction terms and shows that temperature and acid concentration offer the greatest influence, about 25% each, in the variance of the xylose yield. In general, the importance of temperature, acid concentration, time and interactions between these factors is not too surprising because the relationship has already been investigated using modified Arhenius equations in other studies that have proposed these equations as rate constants in kinetic models to describe the acid

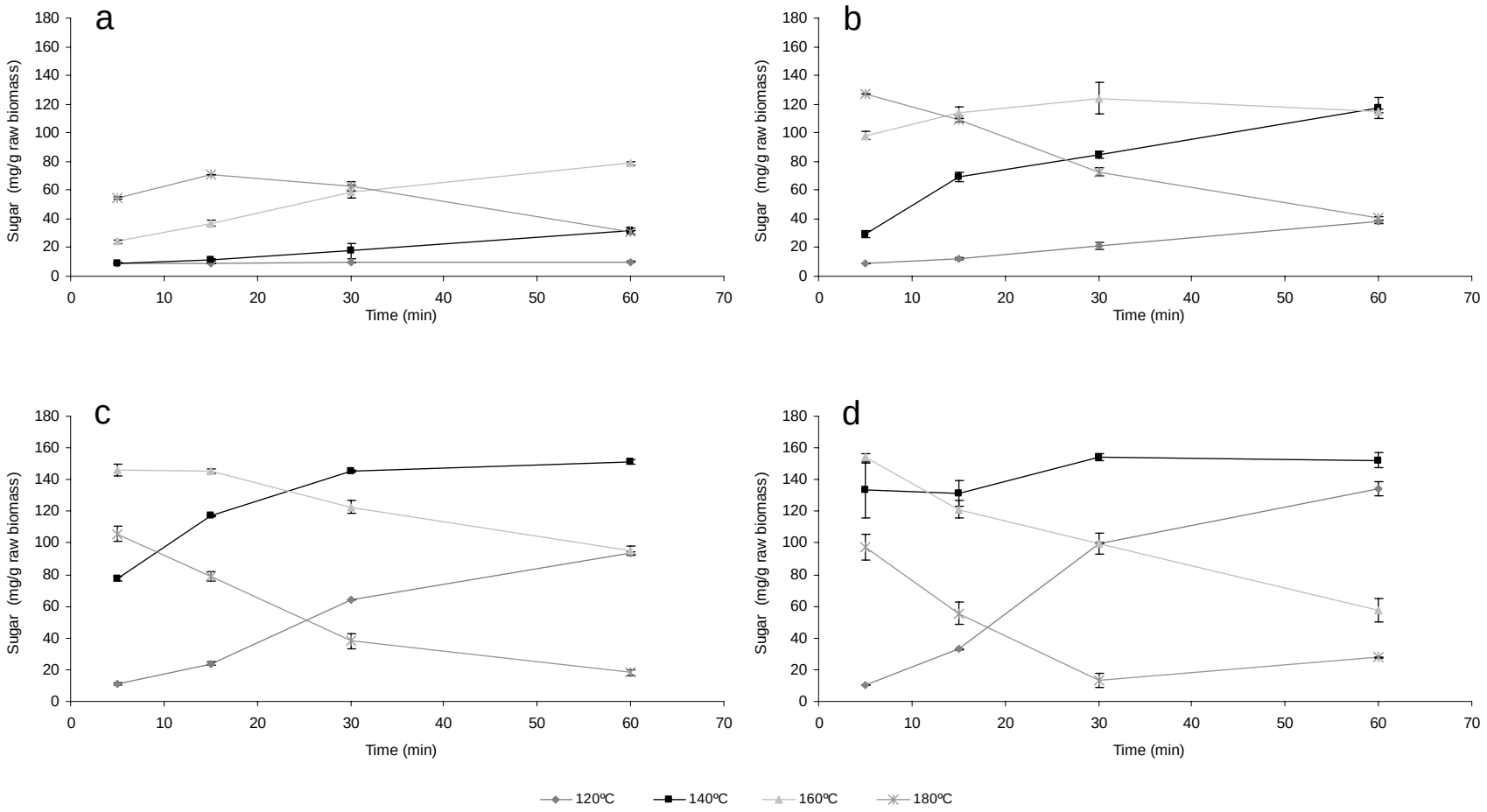

Figure 2.1 Monomeric xylose measured in the prehydrolyzate after pretreatment as a function of time and temperature for a) 0.3 %, b) 0.6 %, c) 0.9 %, and d) 1.2 % sulfuric acid concentration (w/w).

Table 2.2 ANOVA Responses for xylose, furfural, and hydroxymethylfurfural (HMF) in the prehydrolyzate (PreH), glucose in the enzymatic hydrolyzate (EH), and total reducing sugars (total glucose and xylose).

		Type III Sum of Squares					P-Value				
Source	**DF**	**PreH Xylose**	**PreH Furfural**	**PreH HMF**	**EH Glucose**	**Total Sugars**	**PreH Xylose**	**PreH Furfural**	**PreH HMF**	**EH Glucose**	**Total Sugars**
Temp	3	76499.71	27467.55	1026.81	184369.60	448978.77	<.0001	<.0001	<.0001	<.0001	<.0001
Conc	3	72045.48	2779.12	151.33	52055.52	277733.29	<.0001	<.0001	<.0001	<.0001	<.0001
Temp*Conc	9	792.68	1579.07	178.09	14867.76	21775.18	<.0001	<.0001	<.0001	<.0001	<.0001
Time	3	55443.03	4388.32	93.62	41856.66	172873.57	<.0001	<.0001	<.0001	<.0001	<.0001
Temp*Time	9	52388.85	2750.29	100.43	20023.35	166082.95	<.0001	<.0001	<.0001	<.0001	<.0001
Conc*Time	9	1840.37	278.40	25.75	4730.23	21412.97	<.0001	<.0001	0.0294	<.0001	<.0001
Temp*Conc*Time	27	36623.21	1110.38	61.57	10667.38	83095.35	<.0001	<.0001	0.0288	<.0001	<.0001
Error	64	1104.53	238.82	81.29	2075.85	21183.18					
Total	127	296737.85	40591.96	1718.89	330646.35	1213135.26					

hydrolysis of hemicellulose into xylose (Carrasco and Roy, 1992; Baugh and McCarty, 1988).

Figure 2.1 also highlights the loss of xylose after certain pretreatment conditions. This loss can most likely be attributed to the degradation of xylose into furfural as a result of harsh pretreatment conditions. The major degradation products will be addressed later in more detail. The loss of xylose is first evident at 0.3% and 180ºC after 15 minutes. At 0.6% and 160ºC after 30 minutes there is a slow decrease in xylose levels as well as a much more rapid decrease at 180ºC after 5 minutes. At both 0.9% and 1.2% the rapid degradation of xylose is clear at both 160ºC and 180ºC after 5 minutes.

The prehydrolyzate can also contain glucose. Table 2.3 shows the measured glucose monomers liberated from the bermudagrass during the pretreatment averaged across the residence times for each acid concentration and temperature combination. The time variable is omitted because it does not influence glucose yield as strongly as the other factors (ANOVA not shown), but the standard deviation is included so that the general effect of time on glucose yield in the prehydrolyzate can still be inferred. A clear trend is visible showing that the increase of temperature or acid concentration results in the increase of glucose in the prehydrolyzate. This follows the logic that as the pretreatment conditions become harsher, the hemicellulose is hydrolyzed faster and more completely exposing the cellulose to acid hydrolysis as well. Similar trends are reported in a study on the dilute acid pretreatment of corn stover done by Lloyd and Wyman (2005).

Table 2.3 Glucose in the prehydrolyzate averaged over the time variable.

Concentration, % w/w	Temperature, ºC	Average Glucose Pre-Hydrolyzate, mg/g raw biomass
0.3	120	5.93 ± 0.25
0.3	140	6.47 ± 0.74
0.3	160	9.09 ± 2.72
0.3	180	13.30 ± 4.10
0.6	120	6.44 ± 0.90
0.6	140	10.71 ± 3.79
0.6	160	21.67 ± 6.48
0.6	180	24.05 ± 10.91
0.9	120	8.08 ± 2.14
0.9	140	20.22 ± 9.13
0.9	160	32.15 ± 5.16
0.9	180	34.18 ± 3.14
1.2	120	10.55 ± 4.90
1.2	140	28.30 ± 6.48
1.2	160	37.29 ± 2.17
1.2	180	43.43 ± 15.93

Degradation of Carbohydrates

Under harsh conditions, carbohydrates are degraded into furfural and HMF which in turn are degraded into levulinic acid and formic acid respectively. In this study, even the harshest conditions in the factorial did not generate levulinic acid and formic acid levels in the prehydrolyzate close to reported inhibitory levels (>200 mM) while furfural and HMF were measured well beyond reported inhibitory levels (>1.0 g/L) (Palmqvist and Hahn-Hagerdal, 2000; Navarro, 1994; Almeida et al., 2009; Delgenes et al., 1996; Taherzadeh et al., 1997). For this reason, figure 2.2 shows only the furfural and HMF levels. There is one subplot for each acid concentration and one bar for each combination of time and temperature. For ease of comparison, the individual amounts of furfural and HMF have been

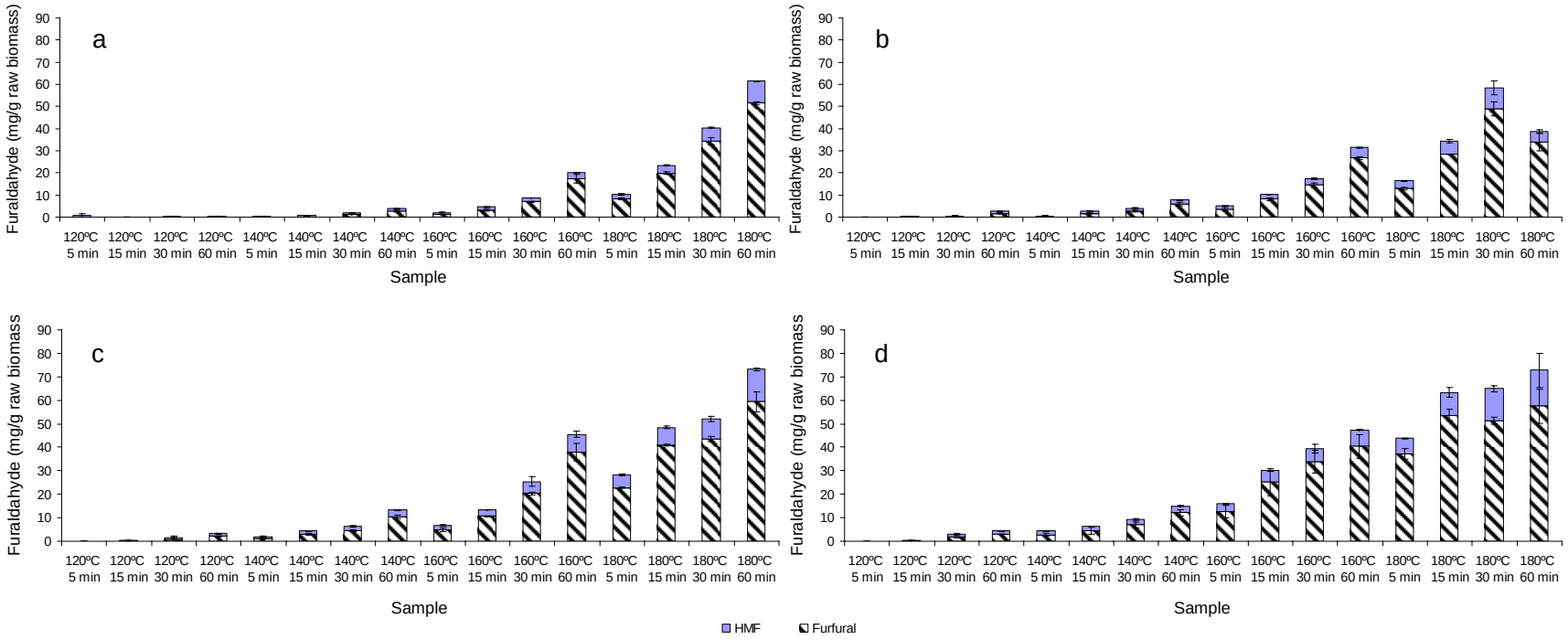

Figure 2.2 Combined furfural and hydroxymethylfurfural measured in the prehydrolyzate after pretreatment as a function of time and temperature for a) 0.3 %, b) 0.6 %, c) 0.9 %, and d) 1.2 % sulfuric acid concentration (w/w).

summed as a single bar in each subplot of figure 2.2 where 10 mg of furfural or HMF each corresponds to 1 g/L of that compound for the 10% solid loading used in this study. As inferred previously, the degradation product generation inversely reflects the xylose losses seen in figure 2.1 which also helps explain why much more furfural than HMF is generated. The general trend seen in each of the plots is supported by the ANOVA output for furfural and HMF in table 2.2, which shows that temperature is the dominating significant effect for both response variables. All of the factors and the interaction terms are significant for furfural and most are significant for HMF except for the interaction between concentration and time and the three-way interaction term.

2.4.3 Enzymatic Hydrolysis

The enzymatic hydrolyzate was analyzed for glucose and xylose monomers. Figure 2.3 shows the levels of monomeric glucose measured in the hydrolyzate for each combination of the factors temperature, concentration, and time. The highest yield of glucose measured in the hydolyzate was 95% and was generated after a pretreatment with 0.6% sulfuric acid at 160ºC for 30 minutes. Examining both figure 2.3 and the output of a statistical multiple comparison of the entire data set (not shown), there are 18 other pretreatment condition sets that produce glucose yields that are similar to the 95% yield. As with the xylose generation in the prehydrolyzate, there are interaction effects contributing to the multiple pretreatment combinations with similar results, however, this is only part of the explanation. The number of pretreatment combinations that give high glucose yields is also increased by the fact that hemicellulose removal in the form of either xylose or furfural

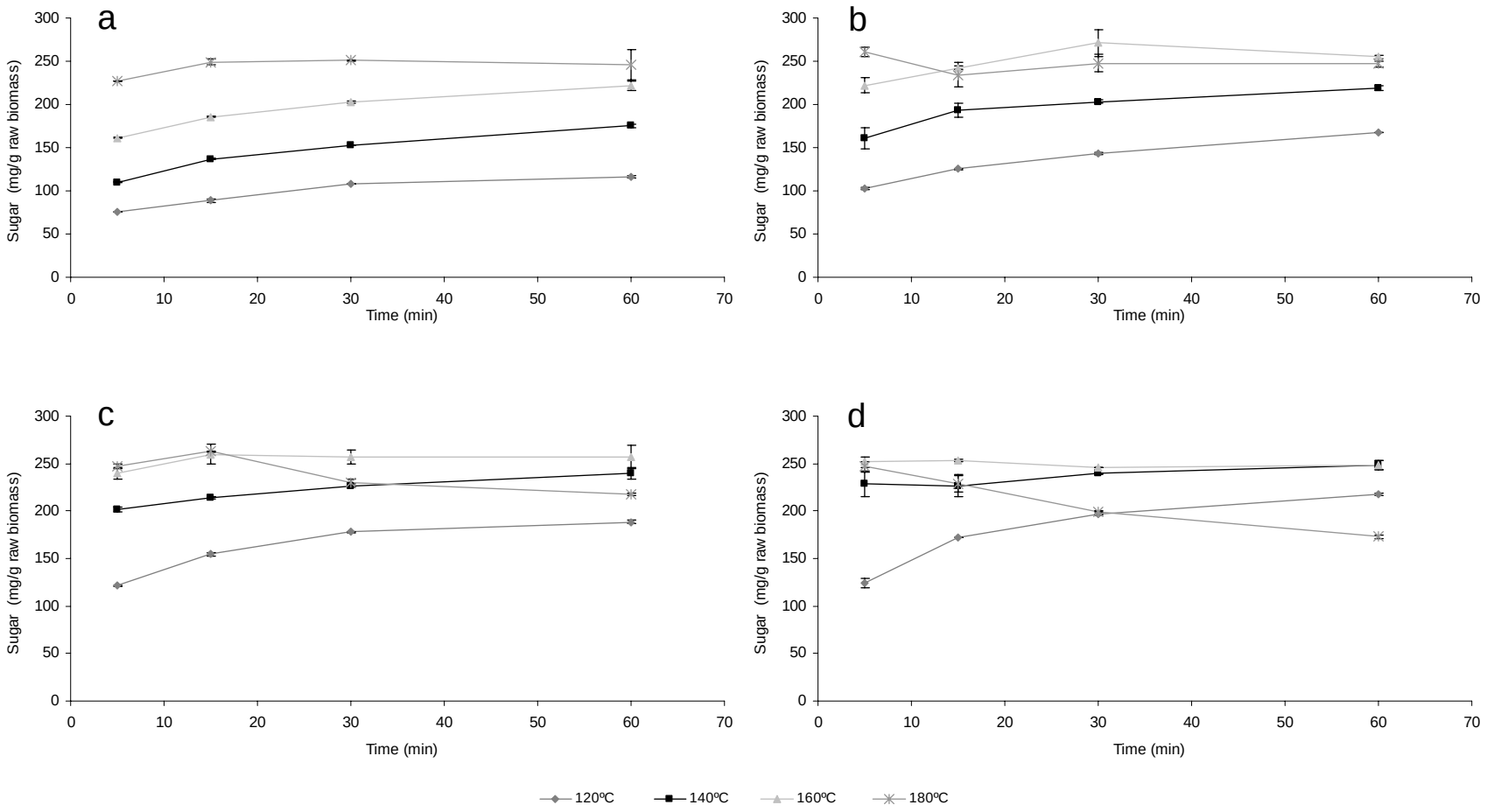

Figure 2.3 Monomeric glucose measured in the hydrolyzate after enzymatic hydrolysis as a function of time and temperature for a) 0.3 %, b) 0.6 %, c) 0.9 %, and d) 1.2 % sulfuric acid concentration (w/w).

generation can lead to high glucose yield in the hydrolyzate. This can be observed upon a review of figures 2.1 and 2.2 in comparison with figure 2.3 and this makes sense because the success of the acid pretreatment is based around the concept of hemicellulose removal in order to maximize glucose yield after enzymatic hydrolysis. Since furfural levels are substantially influenced by temperature and xylose levels are influenced by temperature and acid concentration, it would be expected that glucose in the hydrolyzate would be influenced most substantially by temperature followed by concentration. The ANOVA output in table 2.2 supports this claim. All of the factors and interaction terms significantly contribute to the glucose yield in the hydrolyzate, however, the temperature and concentration terms provide the most influence over the variance making up 55% and 15% respectively. Figure 3.2 shows a loss in glucose yield under a few acid concentrations and residence times, but only for 180ºC. The fact that this apparent degradation happens under harsher conditions than xylose degradation and only at the highest temperature is reasonable because glucose is a more stable carbohydrate than xylose (Baugh and McCarty, 1988).

Table 2.4 shows the xylose found in the hydrolyzate. Like the glucose levels in Table 2.3, the xylose levels in table 2.4 have been averaged across the residence times for each acid concentration and temperature combination. The standard deviation is included so that the general effect of time on xylose levels in the hydrolyzate can still be inferred. A clear trend is visible showing that the increase of temperature or acid concentration results in the decrease of xylose in the hydrolyzate. Only 0.3% acid at 120ºC does not appear to follow this trend and this is most likely explained by a weak acid hydrolysis during the pretreatment that did not adequately disrupt the biomass. Generally, incomplete acid hydrolysis is what

leaves hemicellulose in the solids. Under harsher pretreatments more hemicellulose is removed leaving less in the solids. This correlates to the removal of hemicellulose from the biomass as represented in figures 2.1 and 2.2. Similar trends are reported in a study on the dilute acid pretreatment of corn stover done by Lloyd and Wyman (2005).

Table 2.4 Xylose in the hydrolyzate averaged over the time variable.

Concentration, % w/w	Temperature, ºC	Average Xylose Hydrolyzate, mg/g raw biomass
0.3	120	18.67 ± 7.41
0.3	140	33.61 ± 8.17
0.3	160	28.57 ± 8.07
0.3	180	14.45 ± 5.99
0.6	120	31.14 ± 7.19
0.6	140	27.25 ± 5.85
0.6	160	17.76 ± 4.54
0.6	180	9.14 ± 3.66
0.9	120	30.79 ± 5.32
0.9	140	18.77 ± 4.87
0.9	160	10.50 ± 4.15
0.9	180	3.22 ± 1.64
1.2	120	29.11 ± 6.51
1.2	140	15.37 ± 3.68
1.2	160	6.36 ± 2.42
1.2	180	0.84 ± 0.97

2.4.4 Total Sugars

All of the available sugars from the biomass must be considered for fermentation into ethanol in order to make cellulosic biomass a feasible choice for ethanol production (Sun and Cheng, 2002). In this study, the total sugars available for fermentation from coastal bermudagrass are the summation of the xylose and glucose monomers found in both the

prehydrolyzate and the hydrolyzate. The overall highest yield of total sugars is 97% which is shared by 0.9% acid at 160ºC for 15 minutes and 1.2% acid at 160ºC for 5 minutes. Examining all of the pretreatment conditions with total sugar yields 80% and greater highlights that while sugar production between these pretreatment conditions remains relatively close, the generation of furfurals varies from 3mg/g raw biomass to 37 mg/g raw biomass (equivalent to concentrations of 0.3 g/L to 3.7 g/L respectively). If the threshold concentration of furfurals where yeast inhibition may begin is assumed to be 1 g/L, based on the other studies cited previously, then this list in can be narrowed from twenty entries to seven entries. In this shorter list, the pretreatment condition combination of 1.2% acid at 140ºC for 30 minutes appears to be the best choice based on total sugar yield.

2.4.5 Fermentation Simulation

To help verify the best pretreatment parameters, a fermentation simulation step was used to approximate ethanol production relative to changes in ethanol yield by the organism and inhibition caused by furfurals. This step examined the co-fermentation of the PreH and EH streams together.

All ethanol conversion values were taken from the study done by Krishan et al. (1999) which examined the co-fermentation characteristics of *Saccharomyces* 1400(pLNH33) for ethanol production from varying mixtures of xylose and glucose in the media. The ethanol conversion of this particular organism was found to vary from 0.40 to 0.48 grams of ethanol per gram of sugar depending on the ratio of xylose to glucose in the media. A greater ratio of

xylose resulted in a lower conversion by the organism and a summary of the xylose ratios and corresponding ethanol conversions from the study are given in table 3.1 below.

Table 2.5 Ethanol conversion by *Saccharomyces* 1400(pLNH33) based on the ratio of xylose to glucose in the growth media.

% Glucose in Media	% Xylose in Media	Ethanol Conversion g EtOH / g sugar
100	0	0.48
50	50	0.46
33	66	0.44
0	100	0.40

A specific relationship between the combined level of furfural and HMF and the rate of ethanol production in *Saccharomyces cerevisiae* was identified in a study done by Tazherzadeh et al. (1997). This study supported same trend outlined in the other studies mentioned previously by showing the inhibition of the yeast beyond the furfural and HMF concentration of 1 g/L. An equation for this relationship was approximated from a diagram in the publication using Microsoft Excel to generate a 3rd-order polynomial regression trend line, shown in figure 2.4, and was applied to data in this study. For the simulation, it was assumed that the recombinant yeast from Krishan et al. (1999) would be inhibited by the combined furfural and HMF to the same degree as the natural yeast.

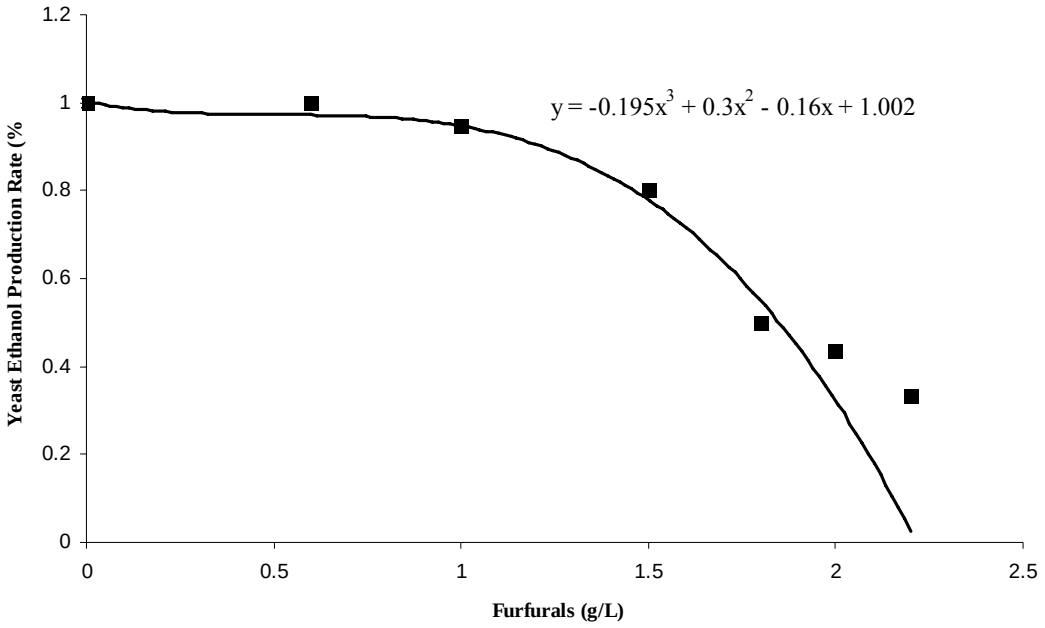

Figure 2.4 The effect of combined furfural and HMF concentration on yeast ethanol production rate.

Table 2.6 shows the output of the fermentation simulation with the approximated ethanol yields exhibiting the expected relationship with combined furfural levels. The combined furfural levels under 10 mg/g (1 g/L) combined with the sugar yields over 90% give ethanol yields greater than 80%. This confirms that a pretreatment with 1.2% acid at 140ºC for 30 minutes is an optimal option as well as 0.9% acid at 160ºC for 5 minutes. The simulation also supports the use of 1 g/L as threshold for furfural production when assessing the data for optimal pretreatment conditions.

Table 2.6 Simulated fermentation output using experimental pretreatment conditions.

Temp. °C	Conc. % w/w	Time, min	Average Furfural, mg/g raw biomass	Average HMF, mg/g raw biomass	EtOH Yield, %	Total Sugar Yield, %
140	0.6	60		0.8	71	80
160	0.6	15	8.15	2.15	74	85
160	0.6	30	14.49	2.91	52	94
180	0.6	5	12.90	3.55	57	91
140	0.9	30	4.35	1.99	79	89
140	0.9	60	10.28	2.95	74	94
160	**0.9**	**5**	**4.76**	**1.85**	**81**	**91**
160	0.9	15	10.61	2.86	76	97
120	1.2	60	3.01	1.38	75	84
140	1.2	5	2.78	1.56	77	87
140	1.2	15	4.46	1.81	76	85
140	**1.2**	**30**	**6.98**	**2.30**	**83**	**94**
140	1.2	60	12.13	2.89	69	96
160	1.2	5	12.67	3.14	65	97

2.4.6 Operating Region for Optimal Pretreatments

The data presented in this study has identified two optimal pretreatment parameter options, however if a contour plot for the total sugar data is overlaid on a contour plot of the furfurals data, regions containing pretreatment condition combinations with the potential for high total sugar production and furfural production around 1 g/L can be seen that were not covered experimentally in this study. Figure 2.5 shows four contour plots with total sugar yields represented by the solid lines and total furfurals generated represented by the dashed lines. Each plot is at one level of the time factor. Favorable operating regions in between 10 mg/g raw biomass of furfurals and 400 mg/g raw biomass of total sugars can be found on each plot. These regions do not extend below 0.7% acid, above 165ºC, or below 125ºC on

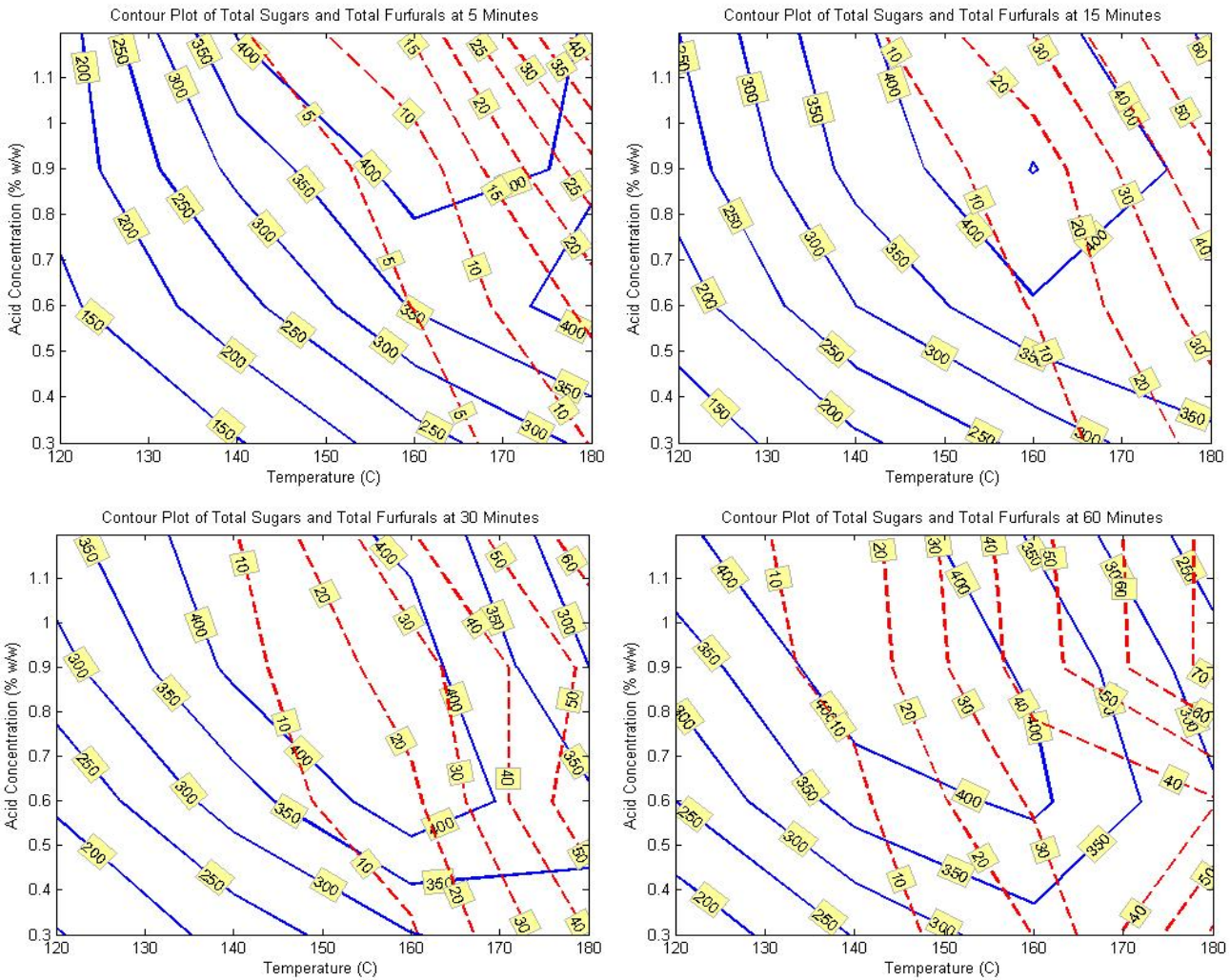

Figure 2.5 The overlay of contour plots for total sugar yield and total furfural generation as a function of acid concentration and temperature at each time level. (Dashed line - furfural, Solid line - total sugars)

any of the plots within the range of the data investigated. The trend of the plots suggests that at higher acid concentrations than examined in this study it may be possible for the favorable region to extend to temperatures lower than 125ºC and at lower residences times than examined in this study, the favorable region might extend to higher temperatures than 165ºC. It would also appear that 0.7% acid is minimum concentration to still be able to operate in the favorable region and operating between 0.9% and 1.2% acid gives the most options for different temperature and time combinations.

2.4.7 Conclusion

Dilute sulfuric acid pretreatment of coastal bermudagrass at high temperatures is a technically feasible way to disrupt the biomass for enzymatic hydrolysis and can result in overall yields above 90% total sugars. The highest total sugar yield was found to be 97% at 0.9% acid at 160ºC for 15 minutes and 1.2% acid at 160ºC for 5 minutes, but if the formation of the sugar degradation products furfural and HMF are accounted for, then a slightly lower total sugar yield of 94% or 91% after a pretreatment with 1.2% acid at 140ºC for 30 minutes or 160 ºC for five minutes respectively may prove more optimal for fermentation. The experimental data also suggests both statistically through the significance of interaction effects and visually in contour plots, that there are other non-experimentally found pretreatment combination options that would yield high total sugars. These options could be useful for the future economic analysis of coastal bermudagrass because there may be several different operating points which share similar yields and costs and this could give some flexibility to the design of a process.

The data in this study follows the trends of the data presented in other similar studies looking at the dilute acid pretreatment of herbaceous biomass or crop residues (Lloyd and Wyman, 2005; Sun and Cheng, 2005; Guo et al., 2008). In general, as pretreatment severity increases, hemicellulose is more completely hydrolyzed from the solid biomass. The more complete hydrolysis of hemicellulose appears to correlate with the higher glucose yields after enzymatic hydrolysis even with higher generation of furfurals. Temperature significantly drives the degradation of xylose into furfural and therefore at higher temperatures there is an increased rate of xylose degradation. Additionally, the glucose found in the prehydrolyzate increases and the xylose found in the solid biomas decreases as pretreatment severity increases. One study in particular this data trends well with is a study on the dilute acid pretreatment of corn stover done by Lloyd and Wyman (2005). Although further research would be required, this might suggest that coastal bermudagrass and corn stover could be pretreated together giving more flexibility in the design of a cellulosic ethanol plant.

REFERENCES

Almeida, J., Bertilsson, M., Gorwa-Grauslund, M., Gorsich, S., Gunnar, L., 2009. Metabolic effects of furaldehydes and impacts on biotechnological processes. *Applied Micribiol Biotechnology* 82: 625-638.

Baugh, K., McCarty, P., 1988. Thermochemical pretreatment of licnocellulose to enhance methane fermentation: I. monosaccharide and furfurals hydrothermal decomposition and product formation rates. *Biotechnology and Bioengineering* 31:50-61.

Carasco, F., Roy, C., 1992. Kinetic study of dilute-acid prehydrolysis of xylan-containing biomass. *Wood Science and Technology* 26: 189-208.

Delgenes, J., Moletta, R., Navarro, J., 1996. Effects of lignocellulose degradation products on ethanol fermentation of glucose and xylose by *Saccharomyces cerevisiae, Zymomonas mobilis, Pichia stipitis,* and *Candida shehatae*. *Enzyme and Microbial Technology* 19: 220-225.

Eggeman, T., and Elander, R., 2005. Process and economic analysis of pretreatment technologies. *Bioresource Technology* 96: 2019-2025.

Ghose, T.K., 1987. Measurement of cellulases activities. *International Union of Pure and Applied Chemistry.* 59, 257-268.

Guo, G., Chen, W., Chen, W., Hwang, W., 2008. Characterization of dilute acid pretreatment of silvergrass for ethanol production. *Bioresource Technology* 99: 6046-6053.

Han, J., Rowell, J. 1997. Chemical composition of agro-based fibers, in: Rowell, R.M., Young, A. R., Rowell, J., (Eds), Paper and composites from agro-based resources. CRC Lewis Publishers, New York, pp 81-134.

Holtzapple, M.T., Ripley, E.P., Nikolaou, M., 1994. Saccharification, fermentation, and protein recovery from low-temperature afex-treated coastal bermudagrass. *Biotechnol. Bioeng.* 44 (9), 1122-1131.

Krishnan, M.S., Ho, N.W.Y., Tsao, G.T., 1999. Fermentation kinetics of ethanol production from glucose and xylose by recombinant *Saccharomyces* 1400 (pLNH33). *Applied Biochemistry and Biotechnology.* 77-79: 373-388.

Lloyd, T., Wyman, C., 2005. Combined sugar yields for combined dilute acid pretreatment of corn stover followed by enzymatic hydrolysis of the remaining solids. *Bioresource Technology* 96: 1967-1977.

Moiser, N., Wyman, C., Dale, B., Elander, R., Lee, Y.Y., Holtzapple, M., Ladisch, M., 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology.* 96: 673-686.

Navarro, A., 1994. Effects of Furfural on Ethanol Fermentation by Saccharomyces cerevisiae: Mathematical Models. *Currecnt Microbiology* 19: 87-90.

Palmqvist, E. and Hahn-Hagerdah, B., 2000. Fermentation of lignocellulosic hydrolyzates. II: inhibitors and mechanisms of inhibition. *Bioresource Technology* 74: 25-33.

Pearl, I. A., 1967. The Chemistry of Lignin. Marcel Dekker Inc., New York.

Saha, B.C., Iten, L.B., Cotta, M.A., Wu, Y.V., 2005. Dilute acid pretreatment, enzymatic saccharification and fermentation of wheat straw to ethanol. *Process Biochemistry* 40 (12): 3693-3700.

Saha, B.C., Biswas, A., Cotta, M.A., 2008. Microwave pretreatment, enzymatic saccharification and fermentation of wheat straw to ethanol. *Journal of biobased materials and bioenergy* 2 (3): 210-217.

Silverstein, R. A., 2004. A comparison of chemical pretreatment methods for converting cotton stalks to ethanol. MS thesis. Raleigh, NC: North Carolina State University, Department of Biological and Agricultural Engineering.

Sluiter, A., 2005a. Determination of total solids in biomass. NREL Biomass Analysis Technology Team Laboratory Analytical Procedure #001. NREL, Golden, CO. Available at: www.nrel.gov/biomass/analytical_procedures.html#lap-001.

Sluiter, A., 2005b. Determination of ash in biomass. NREL Biomass Analysis Technology Team Laboratory Analytical Procedure #005. NREL, Golden, CO. Available at: www.nrel.gov/biomass/analytical_procedures.html#lap-005.

Sluiter, A., 2006. Determination of structural carbohydrates and lignin in biomass. NREL Biomass Analysis Technology Team Laboratory Analytical Procedure #002. NREL, Golden, CO. Available at: www.nrel.gov/biomass/analytical_procedures.html#lap-002.

Solomon, B., Barnes, J., Halvorsen, K., 2007. Grain and cellulosic ethanol: History, economics, and energy policy. *Biomass and Bioenergy* 31: 416-425.

Sun, Y., Cheng, J., 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology.* 83: 1-11.

Sun, Y., Cheng, J., 2005. Dilute acid pretreatment of rye straw and bermudagrass for ethanol production. *Bioresource Technology.* 96: 1599-1606.

Taherzadeh, M., Eklund, R., Gustafsson, L., Niklasson, C., Liden, G., 1997. Characterization and fermentation of dilute-acid hydrolyzates from wood. *Ind. Eng. Chem. Res.* 36: 4659-4665.

Wang, Z., 2009. Alkaline pretreatment of coastal bermudagrass for bioethanol production. MS thesis. Raleigh, NC: North Carolina State University, Department of Biological and Agricultural Engineering.

Yang, B., Wyman, C. E., 2004. Effect of xylan and lignin removal by batch and flowthrough pretreatment on the enzymatic digestibility of corn stover cellulose. *Biotechnology and Bioengineering* 86 (1): 88-95.

CHAPTER 3

AN ASSESSMENT OF AN ARTIFICIAL NERUAL NETWORK FOR MODELING THE DILUTE ACID PRETREATMENT OF COASTAL BERMUDAGRASS

3.1 ABSTRACT

The utility of an artificial neural network (ANN) was assessed for modeling the dilute acid pretreatment of coastal bermudagrass using statistics that quantified the error between the predicted data and actual data as well as a comparison with a multiple linear regression (MLR) model. The statistics used were the coefficient of determination (R^2), the root mean squared error (RMSE), and the root percent deviation (RPD). A standard 2nd-order polynomial multiple linear regression (MLR) model was developed to predict the same data for use as a relative comparison of ANN success. Time (minutes), acid concentration (% w/w), and temperature (ºC) were input into the models to generate xylose in the prehydrolyzate (PreH), glucose in the PreH, furfural in the PreH, HMF in the PreH, xylose in the enzymatic hydrolyzate (EH), and glucose in the EH. It was found that the two types of models predicted most of the outputs closely with the exception of the xylose in the PreH, which the ANN predicted more accurately. An ANN model with six hidden layer neurons was found to be the best overall model and confirmed the utility of utilizing ANN modeling in the area of biomass pretreatment.

3.2 INTRODUCTION

3.2.1 Background

Domestic concerns over increasing oil prices, the limited supply of oil, and global warming are providing an impetus to find a domestically producible and environmentally friendly liquid fuel to replace gasoline. Ethanol produced from cellulosic biomass has the potential to meet this need. Cellulosic biomass is ubiquitous and a variety of feedstocks are possible depending on locale. Additionally, cellulosic feedstocks provide non-food crop options which have the potential to be grown on marginal land. This keeps crop input costs low and avoids the current competitions between food and energy that ethanol from corn encounters.

Much of the research being done on the conversion of cellulosic biomass to ethanol concerns three parts of the overall process: pretreatment, enzymatic hydrolysis and fermentation. Pretreatment is of particular interest because it is upstream of both enzymatic hydrolysis and fermentation and can affect both processes. An effective pretreatment renders the biomass susceptible to enzymatic hydrolysis for maximum fermentable sugars without providing inhibition of either enzymatic hydrolysis or fermentation. Many of the most popular pretreatment methods being investigated involve a combination of heating and either an acid or alkali chemical (Mosier et al., 2005). Of the many pretreatment options, dilute sulfuric acid pretreatment is a near term option that has been investigated for a large number of feedstocks and appears to be less expensive at scale when compared to other pretreatment technologies (Eggeman and Elander, 2005; Mosier et al., 2005).

3.2.2 Modeling Dilute Acid Pretreatment

Modeling is an attractive way for researchers to find optimal pretreatment conditions without performing vast numbers of experiments. Initial work on a dilute acid pretreatment model was done by Saeman (1945) who examined the acid hydrolysis of cellulose into glucose as well as the degradation of glucose. His model was based on the idea that the overall relationship for both of these reactions was first order (3.1) with rate constants that could be described by the Arrhenius relationship between reaction rate and temperature (3.2) where k is the rate constant (min^{-1}), t is time (min), A_0 is a pre-exponential

$$x = x_0 \cdot e^{(-k \cdot t)} \tag{3.1}$$

$$k = A_0 \cdot e^{(\frac{-E}{R \cdot T})} \tag{3.2}$$

constant (min^{-1}), T is temperature (K), R is the gas constant, and E is the activation energy. The constants in this equation can be approximated empirically with the help of linearization and regression. This model has since been applied to the hydrolysis of hemicellulose into xylose and the degradation of xylose into furfural and other degradation products. Kobayashi and Sakai (1955) introduced the idea that a larger portion of hemicellulose tended to hydrolyze fast, while a smaller portion hydrolyzed slower and this version of the model appears to be a popular option (Carrasco and Roy, 1992; Jacobsen and Wyman, 2002; Yat et al., 2007). For further development of a kinetic model for dilute acid pretreatment it has been suggested that the formation of oligomers, transport limitations, nonhomogenous

reactions, and relative binding of the hemicellulose portion to other components of the biomass (lignin) might need to be considered (Jacobsen and Wyman, 2000; Maloney et al., 1985). Although this illuminates a number of options for further pursuit of more complex kinetic models, there are other simpler options for easy modeling of dilute acid pretreatment for prediction and optimization.

A standard modeling technique often used as part of a statistical analysis or optimization instead of a kinetic model is a multiple linear regression (MLR). An MLR is used to solve for coefficients for each term in a linear (often second order polynomial) equation that is then used to describe the response of a variable to input values. This type of model can be used to create response surfaces or contour plots which are used to visualize optimum response regions. Additionally, this type of model can be used to statistically identify both the inputs and interactions between inputs that are significant to the response. However each response variable needs to be handled with a separate set of coefficients and drastically increasing the number of inputs makes the model much more complicated and can make working with the MLR more difficult. In the approximation of more complex and nonlinear systems the MLR may fall short because of the linear structure around which it is developed.

Another type of model which has not been tried in dilute acid pretreatment is the artificial neural network (ANN). An ANN offers the benefit of modeling a complex system without the underlying mathematical descriptions as well as the option to model nonlinear systems easily and could save time and resources when searching for optimum pretreatment

conditions (O'Dwyer et al., 2008). This study aims to assess the utility of using a simple ANN to model a dilute acid pretreatment step as applied to coastal bermudagrass.

3.2.3 Artificial Neural Network

An ANN is mathematical way of simulating the way biological neurons in a brain learn the relationship between input and output data after training with example input and output data sets. Studies in pharmaceuticals (Takayama et al., 1999), agriculture (Edriss et al., 2008), and biofuels (Rajendra et al., 2008; O'Dwyer et al., 2008) have examined the potential for prediction and optimization possible with an ANN. Figure 3.1 below shows an example of the architecture for a typical feedforward backpropagation ANN, which includes an input

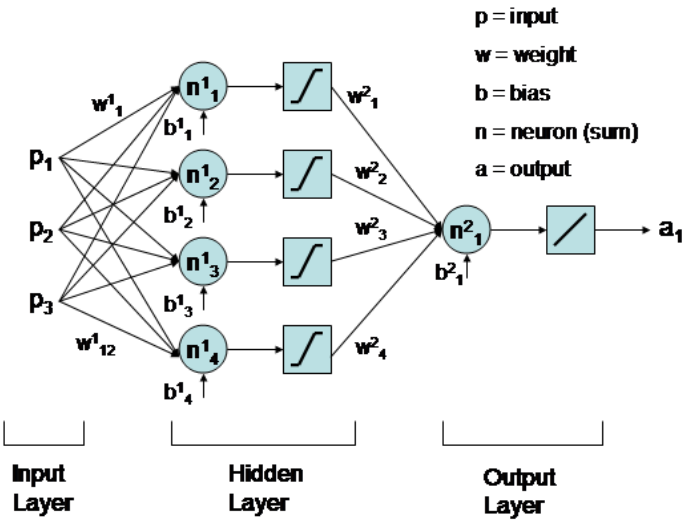

Figure 3.1 Basic schematic for an artificial neural network (ANN).

layer with three inputs, a hidden layer that contains four neurons, layer weights, biases, and sigmoid transfer functions, and an output layer that includes layer weights, one neuron, and a

linear transfer function. A feedforward backpropagation ANN feeds inputs forward through the architecture of the model to generate outputs and adjusts the weight values to minimize the difference between the predicted and the actual output. This process is called 'learning' and it highlights the importance of training the network well. Equation 3.3 shows the basic formula for how each neuron and transfer function set operates to produce an output

$$a = f(w \cdot p + b) \tag{3.3}$$

$$a = \frac{1}{1 + e^{(w \cdot p + b)}} \tag{3.4}$$

and equation 3.4 shows this formula as applied though a sigmoid function, where *a* is the output, *w* is the weight, *p* is the input, and *b* is the bias. Optimal training depends on several factors including the number neurons specified for a hidden layer, the number of hidden layers, the transfer functions, the training algorithm, and the method employed to avoid over-fitting the data. Generally for fewer than five inputs the optimal network performance occurs when the number of neurons in the hidden layer are double the number of inputs (Priddy et al., 2005). Too many neurons can result in over-fitting and too few can result in under-fitting. Additionally, the simpler the ANN will learn best, so minimizing the number of hidden layers is ideal. One of the most commonly used transfer functions is the sigmoid and when paired with a linear transfer function, the ANN can be used to approximate almost any problem (Priddy et al., 2005; Demuth et al., 2008). Once trained optimally, the ANN can be used as a generalized model to accurately predict data within the range of training.

3.3 MATERIALS AND METHODS

3.3.1 Dilute Acid Pretreatment

The data used in this study is from a full factorial experimental design that was developed to examine the effect of dilute acid pretreatment conditions (reaction temperature, acid concentration, and residence time) on the coastal bermudagrass sugar yields in the both the prehydrolyzate (PreH) and the hydrolyzate (EH) as well as the generation of sugar degradation products in the PreH. The values of the design are based on literature reviewed and prior work on the dilute acid pretreatment of bermudagrass done by Sun and Cheng (2005). Sulfuric acid concentrations of 0.3, 0.6, 0.9, and 1.2% (w/w) were examined at temperatures 120, 140, 160, and 180^{o}C and residence times of 5, 15, 30, and 60 minutes. Each pretreatment combination was preformed in triplicate.

Table 3.1 Composition of Coastal Bermudagrass.

Component	% weight of dry biomass
Glucan	25.59
Xylan	15.88
Arabinan	1.95
Galactan	1.46
Acid insoluble lignin	15.37
Acid soluble lignin	3.96
Extractives	4.17
Ash	6.6

Coastal bermudagrass was obtained in 2007 from Central Crops Research Station located in Clayton, NC courtesy of Dr. Joseph Burns of the Crop Science Department at

North Carolina State University and table 3.1 outlines the composition of the biomass. The bermudagrass was ground to particle sizes no greater than 2mm and stored in sealed bags at ambient room temperature in the lab until used. Stainless steel vessels were loaded with 3 g of ground bermudagrass and 30 ml of dilute sulfuric acid before being mixed and sealed. The vessels were heated indirectly in an oil bath with approximately a 12 minute heating period to get to temperature before a planned residence time at that temperature. After pretreatment, the vessels were removed and placed in cold water for immediate cooling prior to being vacuum filtered. The solids were rinsed with 60 ml of water to capture all the hydrolyzed sugars from the pretreatment step. This filtrate, the PreH, was stored at -20^{o}C for analysis later. The solids were then rinsed with another 140 ml of water and stored in sealed plastic bags at 4^{o}C for hydrolysis.

3.3.2 Enzymatic Hydrolysis

Enzymatic hydrolysis was preformed in 50 ml centrifuge tubes maintained at 55ºC and 155 rpm agitation for 72 hours by an automated shaking water bath (New Brunswick, model C76). Each tube was first loaded with the retained rinsed pretreated biomass weighed to account for moisture so that each tube received the equivalent of 0.5 grams of dry biomass. Novozymes North America (Franklinton, NC) provided the cellulase, NS-50013, and the cellobiase, NS-50010, which were determined using an assay procedure derived from Ghose (1987) to have activities of 76.44 FPU/ml (filter paper units, where one unit equates to 1 μmol of cellulose in the form of filter paper converted to glucose per minute) and 283.14 CBU/ml (cellobiase units, where one unit equates to 1 μmol of cellobiose converted to

glucose per minute) respectively. The enzymes were loaded in excess at 40 FPU of cellulase per gram of dry biomass and 70 CBU of cellobiase per gram of dry biomass to avoid any limitation in monomeric sugar production caused by enzyme deficiency. Sodium azide, 0.3% (w/v), was added to inhibit microbial growth. Accounting for the amount of moisture in the biomass, the sodium azide, and the enzymes, the total amount of liquid in each tube was brought up to 15 ml using a 0.05 M sodium citrate buffer to maintain a pH of 4.8 during the enzymatic hydrolysis step.

3.3.3 Sugar Analysis

The EH and PreH were analyzed using a high performance liquid chromatography system (HPLC) to quantify the sugar monomers. The HPLC was also used to quantify prehydolyzate sugar degradation products. An Aminex HPX-87P column was used to distinguish amounts of glucose, xylose, galactose, arabinose in the hydrolyzate samples. The column was run at 80ºC with a flow rate of 0.6 ml/min using HPLC grade water as a mobile phase. Samples injections were 10 µL and the runtime was 35 minutes. A 25 minute post-runtime was included to clear late-eluting compounds from the column. For the prehydrolyzate samples, an Aminex HPX-87H column was used to quantify the levels of glucose, xylose, furfural, 5-hydroxyfuranmethal (HMF), formic acid, and levulinic acid. This column was run at 65ºC with a flow rate of 0.7 mL/min. Sample injections were run for 51 minutes. Total sugars, monomeric sugars, and degradation products were calculated on a per gram of raw biomass basis.

3.4 MODEL

Several ANNs with differing numbers of hidden layer neurons and one standard second-order MLR were developed for relative comparison as non-kinetic empirical model options for describing dilute acid pretreatment. The experimental data, 64 data points, was divided so that 75% was used to train the model and 25% was used to test the generalization ability of the model. The training data was chosen to best fill the input data space (see figure 3.2) which is an effective approach for generating computer models (Simpson et al., 2001). Two sets of data were used for testing with one set being the remaining initial experimental data at each temperature and concentration combination for the 15 minute residence time representing an intermediate slice through the input data space and the second set being data collected from the experimentally determined optimal region (see figure 3.2). Although triplicates of the experimental data were collected, the data had a number of outliers which appeared to be due to consistent heating inequality during pretreatment, enzymatic hydrolysis error, or analysis error. As a result many triplicates had two sets of measured values very close and one set with a clear difference. To remedy this, one sample was removed systematically from each triplicate. The remaining sets of duplicates were averaged and used to train and test the models in this study. The response variables of interest were the xylose in the PreH, the Glucose in the PreH, the Furfural in the PreH, the HMF in the PreH, the xylose in the EH, and the glucose in the EH.

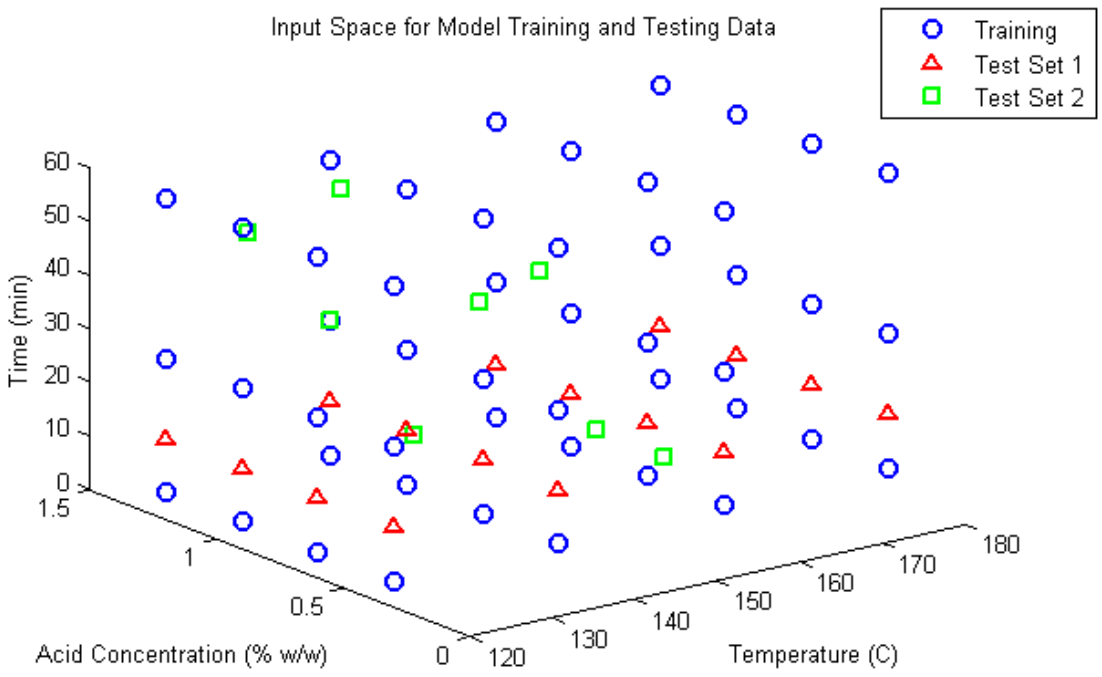

Figure 3.2 Three dimensional plot of the input space showing the spread of the training and testing data used in model development.

3.4.1 Artificial Neural Network Model

MATLAB version 7.6.0.324 R2008a (The MathWorks, Natick, MA) was used to generate a feed-forward back-propagation neural network which modeled both the pretreatment and the enzymatic hydrolysis steps. Given the enzymatic hydrolysis in this study was unconstrained by enzyme loading, this step was used as an indication of the success of the pretreatment, rather than a investigation of enzymatic hydrolysis. As mentioned in the introduction, there are a number of different parameters that can be altered in the development of an ANN, however for simplicity, the ANN was limited to fairly

'default' options. Only a single hidden layer was used. The training algorithm *trainbr* was used to train the ANN in MATLAB. *Trainbr* employs the Levenberg-Marquardt algorithm, a mixture of the Gauss-Newton algorithm and gradient descent for minimization, as a training algorithm and a Bayesian regularization algorithm that limits the magnitude of the weights in order to limit over-fitting. A script was developed in MATLAB that trained the ANN over a range from three to twelve neurons using the default feed forward network learning rate in MATLAB of 0.01, the sum of squares error as a performance function, and allowing 250 epochs (iterations) to reach convergence. The network was retrained 50 times for each number of neurons to account for variation in training outcome. The hidden layer transfer function used was a sigmoid and the output layer transfer function was linear. In each network, the three inputs, the pretreatment conditions, i.e. temperature (^{o}C), acid concentration (% w/w), and time (minutes), were mapped to six outputs, xylose from the PreH, glucose from the PreH, xylose from the EH, glucose from the EH, furfural in the PreH, and HMF in the PreH. All of the data input into the neural network was preprocessed to normalize means and standard deviations and to scale each input between -1 and 1 as the *trainbr* algorithm works best with data in this format. The MATLAB code for the ANN can be referenced in Appendix E.

3.4.2 Multiple Linear Regression Model

The experimental data was run through the RSREG procedure using SAS 9.1.3 (SAS Institute Inc., Cary, NC) to estimate the coefficients for the second-order polynomial (quadratic) equations used to approximate the relationship of temperature, acid concentration,

and time with each response variable. The SAS code for this procedure and a table with the estimated coefficients can be found in Appendix C for reference. Although some of the terms were found to be insignificant ($p>0.01$), these terms were left in the models and the full models were used in the assessment of the MLR method. The coefficient estimates were transferred to a script written in MATLAB which plugged the values into the second-order polynomial equations and simulated the response variable outputs for each set of pretreatment condition inputs. The MATLAB script is included for reference in Appendix D.

3.4.3 Model Assessment Statistics

MATLAB was used to compute coefficient of determination (R^2), the root mean squared error (RMSE), and the root percent deviation (RPD) for the training and testing sets of each model. These statistics were used to assess the performance of model predictions compared to the actual data. The R^2 represents how well the approximated function predicts the response versus just using the response mean. Values closest to 1 are best and the R^2 was computed using the formula in equation 3.5. The RMSE is a representation of the difference between the predicted and actual values and gives a sense of how close the predicted values are to the observed values in the units of those values. Lower values of RMSE are best and it was calculated using the formula in equation 3.6. The RPD represents the percent that the error was of the value being estimated. Lower values are best and it was calculated using the formula in equation 3.7. In equations 3.5, 3.6, and 3.7 the variable n is the number of samples, y_{pi} is the predicted response, y_{pm} is the predicted response mean, y_{ai} is the actual response, and y_{am} is the actual response mean.

$$R^2 = \frac{\sum_{i=1}^{n}(y_{pi} - y_{pm})^2}{\sum_{i=1}^{n}(y_{ai} - y_{am})^2} \tag{3.5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{pi} - y_{ai})^2} \tag{3.6}$$

$$RPD = \frac{100}{n}\sum_{i=1}^{n}\frac{|y_{pi} - y_{ai}|}{|y_{ai}|} \tag{3.7}$$

3. 5 RESULTS AND DISCUSSION

3.5.1 Training Data Set

The R^2, RMSE, and RPD statistics from the training data are shown in table 3.2 for each model. Multiple ANN models are presented to show the effect of increasing the number of neurons in the hidden layer on the ability of the model to fit and predict the response variables. Generally, as the number of hidden layer neurons in the ANN increases, the R^2 values increase slightly and the RMSE and RPD values decrease slightly. This suggests that a more complex ANN is better at fitting the training data, but a testing set is needed to address the possibility of over-fitting. Overall, the statistics between models remain fairly close with exceptions being that the MLR model does not fit the training data for xylose level in the PreH as well as the ANN models and the ANN models do not fit the HMF data as well as the MLR model. Additionally, the RPD statistic for furfural formation is greater than 200% in all of the models which demonstrates that none of the models predict the furfural formation well. From the training data, the ANN model with 12 hidden layer neurons appears to be the best predictor.

Table 3.2 Artificial neural network (ANN) and multiple linear regression (MLR) R^2, RMSE, and RPD values from training data.

Statistic	Model	Xylose PreH	Glucose PreH	Xylose EH	Glucose EH	Furfural PreH	HMF PreH
R^2	**MLR**	0.72	0.81	0.74	0.91	0.94	0.92
	ANN 3 Neuron	0.94	0.75	0.75	0.93	0.94	0.82
	ANN 4 Neuron	0.95	0.76	0.79	0.95	0.94	0.83
	ANN 5 Neuron	0.96	0.82	0.85	0.96	0.94	0.84
	ANN 6 Neuron	0.97	0.84	0.90	0.97	0.94	0.85
	ANN 9 Neuron	0.99	0.89	0.92	0.99	0.95	0.87
	ANN 12 Neuron	0.99	0.92	0.93	0.99	0.96	0.89
RMSE (mg/g)	**MLR**	25.95	5.47	5.64	14.97	4.49	1.07
	ANN 3 Neuron	12.40	6.36	5.57	13.18	4.63	1.69
	ANN 4 Neuron	10.71	6.22	5.02	10.89	4.58	1.64
	ANN 5 Neuron	9.85	5.30	4.20	9.67	4.50	1.57
	ANN 6 Neuron	8.41	4.98	3.52	8.02	4.31	1.52
	ANN 9 Neuron	4.80	4.20	3.22	5.63	3.93	1.36
	ANN 12 Neuron	3.49	3.60	2.87	4.10	3.50	1.25
RPD (%)	**MLR**	82.91	32.64	331.82	7.11	372.66	45.71
	ANN 3 Neuron	32.98	33.89	96.45	6.03	326.49	80.97
	ANN 4 Neuron	24.93	33.37	94.19	4.76	341.76	92.69
	ANN 5 Neuron	22.52	27.34	104.03	4.31	273.09	83.34
	ANN 6 Neuron	18.75	25.43	114.58	3.42	233.79	72.96
	ANN 9 Neuron	11.48	20.32	71.22	2.41	256.62	65.04
	ANN 12 Neuron	8.31	17.12	57.26	1.67	221.74	65.63

3.5.2 First Testing Data Set

Table 3.3 contains the R^2, RMSE, and RPD statistics for the first testing set of data. This test set was a slice through the input space for each temperature and concentration combination at the 15 minute time point. In general the same trends that were seen in the training data statistics are seen in the testing data statistics. Again, most notable is the fact that the xylose level in the pretreatment is more accurately predicted by the ANN versus the MLR. For the MLR and the ANN models with six or fewer neurons, the R^2 and RMSE are close to those found for the training data, which suggests that these models are predicting

consistently and generalize well. The ANN models with more than six neurons showed substantial increases in the RMSE for PreH xylose (~3 times) and EH glucose (~3 times) which could represent over-fitting. This observation supports the earlier expectation that approximately double the number of inputs is the optimum number of neurons to use in the hidden layer and as represented by the first testing set, the ANN model with six neurons appears to be the best predictor.

Table 3.3 Artificial neural network (ANN) and multiple linear regression (MLR) R^2, RMSE, and RPD values from first testing data set.

Statistic	Model	Xylose PreH	Glucose PreH	Xylose EH	Glucose EH	Furfural PreH	HMF PreH
R^2	**MLR**	0.73	0.93	0.73	0.93	0.98	0.98
	ANN 3 Neuron	0.92	0.80	0.69	0.90	0.95	0.92
	ANN 4 Neuron	0.94	0.80	0.75	0.92	0.94	0.91
	ANN 5 Neuron	0.94	0.86	0.81	0.92	0.96	0.94
	ANN 6 Neuron	0.94	0.89	0.86	0.93	0.98	0.96
	ANN 9 Neuron	0.95	0.94	0.87	0.93	0.98	0.97
	ANN 12 Neuron	0.95	0.96	0.88	0.93	0.98	0.97
RMSE (mg/g)	**MLR**	23.60	4.75	6.70	14.39	4.30	0.41
	ANN 3 Neuron	13.10	7.65	5.57	17.42	4.50	0.82
	ANN 4 Neuron	11.76	7.58	5.02	15.95	4.79	0.84
	ANN 5 Neuron	11.56	6.28	4.20	15.87	4.22	0.71
	ANN 6 Neuron	11.85	5.85	3.52	14.91	3.92	0.58
	ANN 9 Neuron	12.30	4.45	3.22	15.44	4.29	0.55
	ANN 12 Neuron	12.35	3.81	2.87	14.99	4.02	0.51
RPD (%)	**MLR**	13.65	14.52	15.20	5.00	327.55	51.82
	ANN 3 Neuron	8.66	11.45	12.68	2.58	183.60	27.89
	ANN 4 Neuron	5.84	11.38	12.60	2.04	199.96	33.11
	ANN 5 Neuron	6.08	8.76	8.38	2.00	162.63	29.77
	ANN 6 Neuron	6.05	7.73	6.41	1.70	102.38	25.80
	ANN 9 Neuron	7.12	5.43	6.35	1.81	123.51	20.19
	ANN 12 Neuron	6.27	4.82	6.26	1.78	115.42	16.94

3.5.3 Second Testing Data Set

One further set of testing data was obtained using the same methods as described above for eight data points located in an optimal (high total sugar, low furans) operating region for the dilute acid pretreatment of coastal bermudagrass. This region was found using a contour plot created from the total set of experimental data (data and contour plot not shown) and this testing set is meant to help better assess how close the models fulfill the intended role of predicting data of interest in the optimal region. Table 3.4 contains the R^2, RMSE, and RPD statistics for this second testing set. The R^2 statistic is much lower for each model in the second testing set, however this reduction from the first testing set is amplified by the small sample size and lack of symmetry in the response. The differences between the models are most evident through the statistics from this testing set. Examining the R^2 and RMSE values shows that any of the ANN models predict the xylose and furfural levels in the PreH best, the xylose and glucose in the EH are predicted slightly better by the MLR, and the glucose and HMF in the PreH are predicted closely by all models. Looking at the RPD values shows that the errors made in the MLR model are on average a larger percentage of the value being predicted than the errors made in any of the ANN models. Although there are signs of slight over-fitting when looking at the decrease and then increase in RMSE statistics for the xylose, furfural, and HMF in the PreH in the ANN models, in general the ANN models do no vary substantially between different numbers of neurons. This may speak to the effectiveness of the Bayesian regularization algorithm in limiting the effects of over-fitting. The six neuron ANN appears to be the most attractive predictor based on the simplicity of this ANN versus 12 neurons as well as the low RMSE and the RPD statistics.

Table 3.4 Artificial neural network (ANN) and multiple linear regression (MLR) R^2, RMSE, and RPD values from second testing data set.

Statistic	Model	Xylose PreH	Glucose PreH	Xylose EH	Glucose EH	Furfural PreH	HMF PreH
R^2	**MLR**	0.60	0.25	0.48	0.60	0.77	0.85
	ANN 3 Neuron	0.67	0.27	0.06	0.22	0.59	0.24
	ANN 4 Neuron	0.60	0.24	0.02	0.32	0.52	0.14
	ANN 5 Neuron	0.52	0.31	0.12	0.33	0.57	0.32
	ANN 6 Neuron	0.45	0.30	0.13	0.34	0.59	0.44
	ANN 9 Neuron	0.24	0.29	0.28	0.47	0.66	0.50
	ANN 12 Neuron	0.20	0.43	0.49	0.38	0.73	0.47
RMSE (mg/g)	**MLR**	32.31	7.82	1.94	6.56	8.21	5.03
	ANN 3 Neuron	9.26	7.16	3.55	12.57	6.14	4.81
	ANN 4 Neuron	7.85	6.89	4.23	10.83	6.01	4.87
	ANN 5 Neuron	8.54	7.24	3.81	10.19	5.55	5.02
	ANN 6 Neuron	9.04	7.20	3.89	9.51	5.34	5.12
	ANN 9 Neuron	11.31	7.04	3.83	8.13	6.42	5.15
	ANN 12 Neuron	12.08	6.63	3.80	8.42	6.54	5.22
RPD (%)	**MLR**	5.44	21.21	17.09	2.73	184.35	50.73
	ANN 3 Neuron	0.83	3.09	2.58	0.75	35.94	8.57
	ANN 4 Neuron	0.76	3.03	2.79	0.65	34.99	8.66
	ANN 5 Neuron	0.81	3.26	2.49	0.59	30.68	8.86
	ANN 6 Neuron	0.86	3.29	2.50	0.57	27.81	8.76
	ANN 9 Neuron	1.01	3.19	2.66	0.45	31.39	8.78
	ANN 12 Neuron	1.09	3.06	2.89	0.45	31.16	9.18

A comparison of the predictions of a standard 2nd-order polynomial and the predictions of a feedforward backpropagation ANN with six neurons in the hidden layer versus the actual experimental values is shown in the plots in figure 3.3. In the plots it is easy to visualize how similar these two predictive models are excluding the xylose levels in the PreH where the errors in the MLR predictions are noticeably larger than the errors in the ANN predictions. This is especially true at higher levels of xylose in the PreH and indicates an advantage to using an ANN over an MLR in predicting optimal pretreatment parameters.

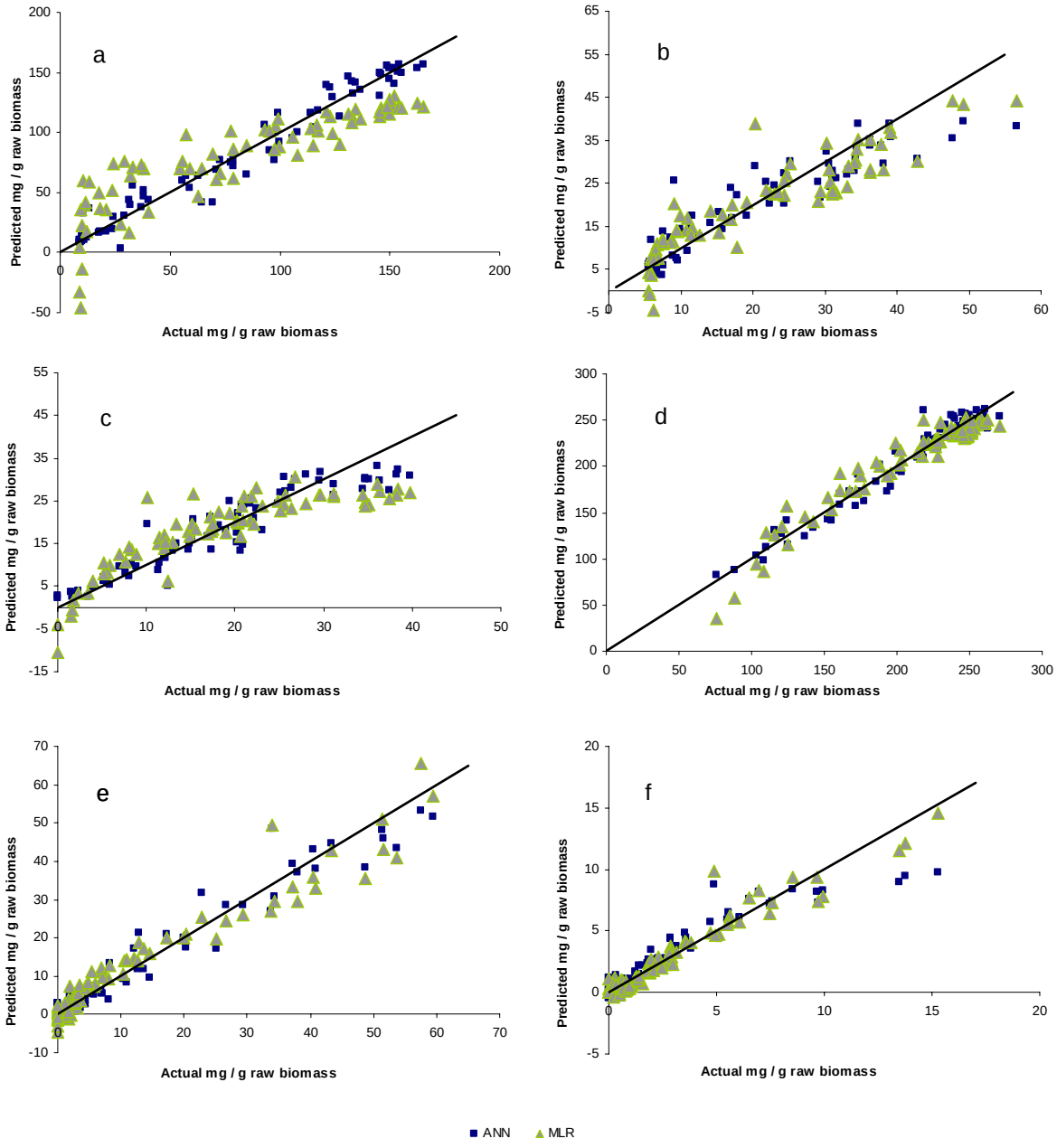

Figure 3.3 Combined training and testing predicted values versus actual experimental values for an ANN with 6 neurons in the hidden layer and an MLR where (a) xylose from pre-hydrolyzate, (b) glucose from pre-hydrolyzate , (c) xylose from enzymatic hydrolysis, (d) glucose from enzymatic hydrolysis, (e) furfural, and (f) HMF.

3.5.4 Conclusion

This study was able to demonstrate the utility of a six neuron feedforward backpropagation ANN for modeling a dilute acid pretreatment step at least as well as a standard 2nd-order polynomial MLR method. There is also support that the ANN model can predict the xylose levels in the PreH with superior accuracy over the MLR model, which may be due to nonlinear approximation inherent with the ANN architecture used. Though this particular example was limited in scope to the dilute acid pretreatment of coastal bermudagrass, an ANN could be expanded to include other input variables like the feedstock (composition), reactor type, and solid loading to improve the ability of the model to generalize a wider array of pretreatment situations. Changing the acid concentration input to a pH input would potentially allow for the use of a single ANN (after proper training) with multiple pretreatment methods. Additionally, outputs could be added to account for changes in biomass structural features after pretreatments rather than just monomeric sugar yields. A widely generalized complex model trained to output post-pretreatment biomass structural features could be used in conjunction with another ANN, like that presented by O'Dwyer et al. (2008), which models the digestibility of the biomass during enzymatic hydrolysis. Together these models could aid in the development of a generalized (across pretreatments) economic model.

REFERENCES

Carasco, F., Roy, C., 1992. Kinetic study of dilute-acid prehydrolysis of xylan-containing biomass. *Wood Science and Technology* 26: 189-208.

Demuth, H., Beale, M., Hagan, M., 2008. *Neural Network Toolbox 6: User's Guide*. Version 6. Natick, MA: The MathWorks.

Edriss, M., Hosseinnia, P., Edrisi, M., Rahmani, H., Nilforooshan, M., 2008. Prediction of second parity milk performance of dairy cows from first parity information using artificial neural network and multiple linear regression methods. *Asian Journal of Animal and Veterinary Advances* 3(4): 222-229.

Eggeman, T. and Elander, R., 2005. Process and economic analysis of pretreatment technologies. *Bioresource Technology* 96: 2019-2025.

Ghose, T.K., 1987. Measurement of cellulases activities. *International Union of Pure and Applied Chemistry.* 59, 257-268.

Jacobsen, S. and Wyman, C., 2000. Cellulose and hemicellulose hydrolysis models for application to current and novel pretreatment processes. *Applied Biochemistry and Biotechnology* 84-86: 81-96.

Jacobsen, S. and Wyman, C., 2002. Xylose monomer and oligomer yields for uncatalyzed hydrolysis of sugarcane bagasse hemicellulose at varying solids concentration. *Ind. Eng. Chem.Res.*41 (6): 1454-1461.

Kobayashi, T., and Sakai, Y., 1956. Hydrolysis rate of pentosan of hardwood in dilute sulfuric acid. *Bull .Agr. Chem. Soc. Japan* 20: 1-7.

Maloney, M., Chapman, T., Baker, A., 1985. Dilute acid-hydrolysis of paper birch-kinetics studies of xylan and acetyl-group hydrolysis. *Biotechnology and Bioengineering* 27: 355-361.

Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y., Holtzapple, M., Ladisch, M., 2005. Feature of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology* 96: 673-686.

O'Dwyer J., Zhu, L., Granda, C., Chang, V., Holtzapple, M., 2008. Neural Network Prediction of Biomass Digestibility Based on Structural Features. *Biotechnolgy Progress* 24: 283-292.

Priddy, K. and Keller, P. 2005. *Artificial Neural Networks: An Introduction*. Bellingham, WA: SPIE Press.

Rajendra, M., Jena, P., Raheman, H., 2009. Prediction of optimized pretreatment process parameters for biodieseal production using ANN and GA. *Fuel* 88: 868-875.

Saeman, J., 1945. Kinetics of wood saccharification: hydrolysis of cellulose and decomposition of sugars in dilute acid at high temperature. *Industrial Engineering Chemistry* 37: 42-52.

Sanderson, M.A., Reed, R.L., McLaughlin, S.B., Wullschleger, S.D., Conger, B.V., Parrish, D.J., Wolf, D.D., Taliaferro, C., Hopkins, A.A., Ocumpaugh, W.R., Hussey, M.A., Read, J.C., Tischler, C.R., 1996. Switchgrass as a sustainable bioenergy crop. *Bioresour. Technol.* 56 (1): 83-93.

Simpson, T., Peplinski, J., Koch, P., Allen, J., 2001. Metamodels for computer-based engineering design: survey and recommendations. *Engineering with Computers* 17: 129-150.

Sun, Y. and Cheng, J., 2005. Dilute acid pretreatment of rye straw and bermudagrass for ethanol production. *Bioresource Technology* 96:1599-1606.

Takayama, K., Fujikawa, M., Nagai, T., 1999. Artificial neural network as a novel method to optimize pharmaceutical formulations. *Pharmaceutical Research* 16: 1-6.

Yat, S., Berger, A., Shonnard, D., 2007. Kinetic characterization for dilute sulfuric acid hydrolysis of timber varieties and switchgrass. *Bioresource Technology* 99 (9): 3855-3863.

CHAPTER 4

CONCLUSIONS

4.1 SUMMARY AND CONCLUSIONS

Coastal bermudagrass is an appropriate choice as a lignocellulosic feedstock to use in bioethanol production when coupled with dilute sulfuric acid pretreatment. Experimentally, it was established that the dilute acid pretreatment of coastal bermudagrass at temperatures higher than 120ºC can enhance the effectiveness of the pretreatment and can lower the acid concentration and pretreatment time required for an effective pretreatment. Coupled with this finding, it was demonstrated that harsher pretreatments will quickly begin to degrade the sugars in coastal bermudagrass into furfurals. As a result, it was a set of pretreatment conditions of intermediate severity that were found which gave high sugar yields (>80%) and generated low amounts of degradation products. The highest total sugar yield was found to be 97% at both 0.9% acid at 160ºC for 15 minutes and 1.2% acid at 160ºC for 5 minutes. A simple fermentation simulation illuminated that a slightly lower total sugar yield of 94% after a pretreatment with 1.2% acid at 140 ºC for 30 minutes may be the more optimal choice because of higher ethanol yield due to lower concentration of furfurals (< 1 g/L yeast inhibition threshold).

From the experiments, multiple sets of pretreatment conditions were found to yield similar amounts of sugars. This illuminated the interactions between temperature, acid concentration, and time, which led to the development of a model for use in predicting the sugar and furfural yields after dilute acid pretreatments conditions not examined

experimentally. In an assessment of an artificial neural network (ANN), it was found that the ANN was able to predict sugar yields well based on statistics that describe how well the predicted values match the actual values. When compared to a more standard method, the multiple linear regression (MLR) model, it was found that the ANN predict most variables closely, except for the xylose level in the prehydrolyzate, which it can predict more accurately than an MLR.

4.2 SUGGESTIONS FOR FUTURE WORK

Further research can be conducted related to each of the major subjects addressed in this study, i.e. pretreatment, enzymatic hydrolysis, fermentation, and modeling. Pretreatment work could include testing alternative acids, like acetic acid, testing larger biomass particle sizes, and testing larger solid loadings. It would also be interesting to try and quantify lignin degradation products in order to correlate them with either pretreatment severity or the formation of sugar degradation products. In this study, it was assumed that enzymatic hydrolysis was ideal, but testing could be done to check the minimum enzyme loading, to examine the relationship between degradation products and enzyme activity, and to look at enzymatic hydrolysis under more realistic conditions. Fermentation using organisms that convert both five and six carbon sugars would be also be useful, especially to check against expectations of inhibition from furfurals. Finally, an economic model which examines energy and chemical costs would be useful for establishing the true optimal pretreatment conditions. This type of economic model could be built from data generated from a more

generalized artificial neural network which would be developed to take into account other feedstocks and other pretreatment methods.

APPENDICES

APPENDIX A: ANOVA SAS CODE

```
data pretreatment;
input temp conc time TRS@;
datalines;
Temp. Data     Conc. Data     Time Data     Sugar Data
;

proc glm data=pretreatment;

   model trs= temp|conc|time;
   lsmeans temp conc time temp*conc temp*time conc*time temp*conc*time
   /pdiff=all;

run;
```

APPENDIX B: LIGNIN DATA

Table A.1 Average percentage change in lignin measured in biomass between raw and pretreated biomass. (negative sign denotes more lignin present in pretreated biomass than orginally found in raw biomass).

		Acid Concentration, %(w/w)			
Temp. ºC	**Time, min**	**0.3%**	**0.6%**	**0.9%**	**1.2%**
120	5	-13.2	-4.9	-0.9	-9.4
120	15	-12.2	-4.9	0.9	-2.6
120	30	-7.2	-2.6	1.3	1.2
120	60	-5.3	4.1	2.7	-4.8
140	5	-4.4	-6.5	7.4	13.6
140	15	-4.8	-4.2	12.9	10.3
140	30	-4.7	-1.3	3.0	9.9
140	60	-2.0	0.8	-2.0	10.2
160	5	-4.4	12.3	10.9	7.6
160	15	11.2	6.0	12.5	0.5
160	30	7.1	16.3	7.5	-3.7
160	60	0.7	2.9	-8.7	-14.6
180	5	15.5	13.1	11.2	15.9
180	15	4.3	6.2	5.6	8.7
180	30	2.9	-4.4	-6.4	-12.9
180	60	-2.5	-12.1	-9.5	-12.4

APPENDIX C: MLR SAS CODE (RSREG)

```
data pretreatment;
input temp conc time TRS@;
datalines;
Temp. Data     Conc. Data     Time Data     Sugar Data
;

PROC RSREG data=pretreatment out=response;
    MODEL TRS=temp conc time/ predict;
run;
```

Table A.2 Estimates for multiple linear regression coefficients from the RSREG procedure in SAS.

Coeff.	Variable	PreH Xylose	PreH Glucose	EH Xylose	EH Glucose	PreH Furfural	PreH HMF
B1	**intercept**	-1592.7	-130.53	-45.454	-1320.3	217.06	51.124
B2	**temp**	18.73	1.5303	0.96307	14.941	-3.0633	-0.68388
B3	**conc**	443.79	-25.411	43.509	552.63	-51.112	-15.946
B4	**time**	5.8851	0.95603	0.50176	5.8603	-0.86785	-0.17364
B5	**temp* temp**	-0.053741	-0.004509	-0.003175	-0.036614	0.010443	0.002247
B6	**conc* temp**	-1.6487	0.35974	-0.33241	-2.1777	0.36334	0.10742
B7	**conc* conc**	-85.242	-2.4301	-1.7084	-101.38	5.6041	1.2927
B8	**time* temp**	-0.037095	-0.003983	-0.003098	-0.026061	0.009272	0.001755
B9	**time* conc**	-0.18281	-0.063478	-0.14415	-0.74124	0.035897	0.036147
B10	**time* time**	-0.001935	-0.003713	6.44E-05	-0.013492	-0.003999	-0.00092

APPENDIX D: MATLAB CODE FOR MLR

```
%Input data from files into matrices in MATLAB for manipulation
p = load('NNModelInputPartial2.txt'); %training dataset:inputs
t = load('NNModelOutputPartial2.txt'); %training dataset:outputs
b = load('RSREGEstimate.txt');  %coefficient estimates
t0 = load('NNModelOutputPartial2NOZERO.txt'); %data with zeroes replaced by
                                                %lowest measureable threshold
                                                %of 0.1 for division purposes
p = transpose(p);
t = transpose(t);
t0 = transpose(t0);


%initialize counters and matrices
temps = transpose(p(:,1));
concs = transpose(p(:,2));
times = transpose(p(:,3));
yMLR=[];
n=1;
l=1;
counter = size(temps,2);

%MLR Equation populated with coefficient estimates from SAS and run in a
%loop for each set of pretreatment conditions
h=1;
for h=1:counter
   temp=temps(h);
   conc=concs(h);
   time=times(h);
   yMLR(:,n) = b(:,1) + (b(:,2).*temp) + (b(:,3).*conc) + (b(:,4).*time) + (b(:,5).*temp.*temp) +
(b(:,6).*conc.*temp) + (b(:,7).*conc.*conc) + (b(:,8).*time.*temp) + (b(:,9).*time.*conc) +
(b(:,10).*time.*time);

   h=h+1;
   n=n+1;
end


%RMSE calculation
net_error = yMLR-t;
MSE_PreH_X = mse(net_error(1,:));
MSE_PreH_G = mse(net_error(2,:));
MSE_EH_X = mse(net_error(3,:));
MSE_EH_G = mse(net_error(4,:));
MSE_Furfural = mse(net_error(5,:));
MSE_HMF = mse(net_error(6,:));

RMSE_PreH_X=sqrt(MSE_PreH_X)
RMSE_PreH_G=sqrt(MSE_PreH_G)
RMSE_EH_X=sqrt(MSE_EH_X)
RMSE_EH_G=sqrt(MSE_EH_G)
```

```
RMSE_Furfural=sqrt(MSE_Furfural)
RMSE_HMF=sqrt(MSE_HMF)


%RSquare calculation
for column = 1:1:6
   r = corrcoef(yMLR(column,:),t(column,:));
   r = r(1,2);
   R_Square(1,column) = r^2;
end
R_Square


%RPD calculation
RPD = 100*sum(abs(net_error)./abs(t0),2)/counter
```

APPENDIX E: MATLAB CODE FOR ANN

```
%Input, transpose, and scaling of model inputs (p) and model outputs (t)
%Training Data
p = load('NNModelInputPartial2.txt');
t = load('NNModelOutputPartial2.txt');
t0 = load('NNModelOutputPartial2NOZERO.txt');
p = transpose(p);
t = transpose(t);
t0 = transpose(t0);

%First Test Set
p1 = load('NNModelInputCheck2.txt');
t1 = load('NNModelOutputCheck2.txt');
t10 = load('NNModelOutputCheck2NOZERO.txt');
p1 = transpose(p1);
t1 = transpose(t1);
t10 = transpose(t10);

%Second Test Set
p2 = load('NNModelInputCheck4.txt');
t2 = load('NNModelOutputCheck4.txt');
p2 = transpose(p2);
t2 = transpose(t2);
t20= t2;

%initalize values, counters, and storage matrices
R_Square = [];
RMSE = [];
Summary = [];
n=0;
counter = size(p,2);
counter1 = size(p1,2);
counter2 = size(p2,2);

%Starts of a loop that runs through and records average ANN performance
%statistics for different numbers of neurons (in this case from 3 to 12 neurons)
for neuron=3:1:12

%Below are function to create new ANN and train it
    net = newff(p,t,neuron,[],'trainbr');
    net.inputs{1}.processFcns = {'mapminmax','processpca','mapstd'};
    net.inputs{1}.processParams{2}.maxfrac = 0.001;
    net.outputs{2}.processFcns = {'mapminmax','mapstd'};
    net.trainParam.epochs = 7000;
    net.divideFcn ='';
    net.trainParam.show = 10;
    net.performFcn = 'mse';
    [net,tr]=train(net,p,t);
```

```
% The 'rerun' loop is how many iterations the ANN is run for a particular
% number of neurons
for rerun = 1:1:1

  net = init(net);
  [net,tr]=train(net,p,t);

  %simulate network with training data
  y = sim(net,p);

  % Compute the R_Square and RMSE statistics for the training data
  net_error = t-y;

  MSE_PreH_X = mse(net_error(1,:));
  MSE_PreH_G = mse(net_error(2,:));
  MSE_EH_X = mse(net_error(3,:));
  MSE_EH_G = mse(net_error(4,:));
  MSE_Furfural = mse(net_error(5,:));
  MSE_HMF = mse(net_error(6,:));

  RMSE_PreH_X=sqrt(MSE_PreH_X);
  RMSE_PreH_G=sqrt(MSE_PreH_G);
  RMSE_EH_X=sqrt(MSE_EH_X);
  RMSE_EH_G=sqrt(MSE_EH_G);
  RMSE_Furfural=sqrt(MSE_Furfural);
  RMSE_HMF=sqrt(MSE_HMF);

  % Matrix for recording the RMSE after each run of the training set
  % in order to average them later
  RMSE(rerun,:) =
[RMSE_PreH_X,RMSE_PreH_G,RMSE_EH_X,RMSE_EH_G,RMSE_Furfural,RMSE_HMF];

  for row = 1:1:6

    %RSquare calculation
    r = corrcoef(y(row,:),t(row,:));
    r = r(1,2);
    R_Square(rerun,row) = r^2;

  end

  %RPD calculation and storage matrix for training set
  RPD(rerun,:) = transpose(100*sum(abs(net_error./t0),2)/counter);

  % Simulate network with first testing set
      y1 = sim(net,p1);

  % Compute the R_Square and RMSE statistics for the first testing set
  net_error = t1-y1;

  MSE_PreH_X = mse(net_error(1,:));
```

```
MSE_PreH_G = mse(net_error(2,:));
MSE_EH_X = mse(net_error(3,:));
MSE_EH_G = mse(net_error(4,:));
MSE_Furfural = mse(net_error(5,:));
MSE_HMF = mse(net_error(6,:));

RMSE_PreH_X=sqrt(MSE_PreH_X);
RMSE_PreH_G=sqrt(MSE_PreH_G);
RMSE_EH_X=sqrt(MSE_EH_X);
RMSE_EH_G=sqrt(MSE_EH_G);
RMSE_Furfural=sqrt(MSE_Furfural);
RMSE_HMF=sqrt(MSE_HMF);

% Matrix for recording the RMSE after each run of the first test set
% in order to average them later
Test1RMSE(rerun,:) =
[RMSE_PreH_X,RMSE_PreH_G,RMSE_EH_X,RMSE_EH_G,RMSE_Furfural,RMSE_HMF];

for row = 1:1:6

   %R Square of first testing set
   r = corrcoef(y1(row,:),t1(row,:));
   r = r(1,2);
   Test1R_Square(rerun,row) = r^2;

end

%RPD calculation and storage matrix for first test set
RPD1(rerun,:) = transpose(100*sum(abs(net_error./t10),2)/counter);

%Simulation with second test set
      y2 = sim(net,p2);

% Compute the R_Square and RMSE statistics for the second testing set
net_error = t2-y2;

MSE_PreH_X = mse(net_error(1,:));
MSE_PreH_G = mse(net_error(2,:));
MSE_EH_X = mse(net_error(3,:));
MSE_EH_G = mse(net_error(4,:));
MSE_Furfural = mse(net_error(5,:));
MSE_HMF = mse(net_error(6,:));

RMSE_PreH_X=sqrt(MSE_PreH_X);
RMSE_PreH_G=sqrt(MSE_PreH_G);
RMSE_EH_X=sqrt(MSE_EH_X);
RMSE_EH_G=sqrt(MSE_EH_G);
RMSE_Furfural=sqrt(MSE_Furfural);
RMSE_HMF=sqrt(MSE_HMF);

% Matrix for recording the RMSE after each run of the second test set
% in order to average them later
```

```
    Test2RMSE(rerun,:) =
[RMSE_PreH_X,RMSE_PreH_G,RMSE_EH_X,RMSE_EH_G,RMSE_Furfural,RMSE_HMF];

    for row = 1:1:6

        %R Square of second testing set
        r = corrcoef(y2(row,:),t2(row,:));
        r = r(1,2);
        Test2R_Square(rerun,row) = r^2
    end

    %RPD calculation and storage matrix for second test set
    RPD2(rerun,:) = transpose(100*sum(abs(net_error./t20),2)/counter);

end

%compile averages and standard deviations of Neural Net Runs
R2_Average =
[mean(R_Square(:,1)),mean(R_Square(:,2)),mean(R_Square(:,3)),mean(R_Square(:,4)),mean(R_Square(:,5)),m
ean(R_Square(:,6))];
R2_STD =
[std(R_Square(:,1)),std(R_Square(:,2)),std(R_Square(:,3)),std(R_Square(:,4)),std(R_Square(:,5)),std(R_Square(:
,6))];
RMSE_Average =
[mean(RMSE(:,1)),mean(RMSE(:,2)),mean(RMSE(:,3)),mean(RMSE(:,4)),mean(RMSE(:,5)),mean(RMSE(:,6)
)];
RMSE_STD =
[std(RMSE(:,1)),std(RMSE(:,2)),std(RMSE(:,3)),std(RMSE(:,4)),std(RMSE(:,5)),std(RMSE(:,6))];
RPD_Average =
[mean(RPD(:,1)),mean(RPD(:,2)),mean(RPD(:,3)),mean(RPD(:,4)),mean(RPD(:,5)),mean(RPD(:,6))];

R2_Average1 =
[mean(Test1R_Square(:,1)),mean(Test1R_Square(:,2)),mean(Test1R_Square(:,3)),mean(Test1R_Square(:,4)),m
ean(Test1R_Square(:,5)),mean(Test1R_Square(:,6))];
R2_STD1 =
[std(Test1R_Square(:,1)),std(Test1R_Square(:,2)),std(Test1R_Square(:,3)),std(Test1R_Square(:,4)),std(Test1R_
Square(:,5)),std(Test1R_Square(:,6))];
RMSE_Average1 =
[mean(Test1RMSE(:,1)),mean(Test1RMSE(:,2)),mean(RMSE(:,3)),mean(Test1RMSE(:,4)),mean(Test1RMSE(
:,5)),mean(Test1RMSE(:,6))];
RMSE_STD1 =
[std(Test1RMSE(:,1)),std(Test1RMSE(:,2)),std(Test1RMSE(:,3)),std(Test1RMSE(:,4)),std(Test1RMSE(:,5)),st
d(Test1RMSE(:,6))];
RPD1_Average =
[mean(RPD1(:,1)),mean(RPD1(:,2)),mean(RPD1(:,3)),mean(RPD1(:,4)),mean(RPD1(:,5)),mean(RPD1(:,6))];

R2_Average2 =
[mean(Test2R_Square(:,1)),mean(Test2R_Square(:,2)),mean(Test2R_Square(:,3)),mean(Test2R_Square(:,4)),m
ean(Test2R_Square(:,5)),mean(Test2R_Square(:,6))];
R2_STD2 =
[std(Test2R_Square(:,1)),std(Test2R_Square(:,2)),std(Test2R_Square(:,3)),std(Test2R_Square(:,4)),std(Test2R_
Square(:,5)),std(Test2R_Square(:,6))];
```

```
    RMSE_Average2 =
[mean(Test2RMSE(:,1)),mean(Test2RMSE(:,2)),mean(Test2RMSE(:,3)),mean(Test2RMSE(:,4)),mean(Test2R
MSE(:,5)),mean(Test2RMSE(:,6))];
    RMSE_STD2 =
[std(Test2RMSE(:,1)),std(Test2RMSE(:,2)),std(Test2RMSE(:,3)),std(Test2RMSE(:,4)),std(Test2RMSE(:,5)),st
d(Test2RMSE(:,6))];
    RPD2_Average =
[mean(RPD2(:,1)),mean(RPD2(:,2)),mean(RPD2(:,3)),mean(RPD2(:,4)),mean(RPD2(:,5)),mean(RPD2(:,6))];

    n=n+1;


% Summary of averages for each statistic and each number of neurons

Summary(n,:)=[R2_Average R2_STD RMSE_Average RMSE_STD R2_Average1 R2_STD1
RMSE_Average1 RMSE_STD1 R2_Average2 R2_STD2 RMSE_Average2 RMSE_STD2 RPD_Average
RPD1_Average RPD2_Average]


end
```