

2011

Analyses of bioenergy systems: detecting hard-coding errors in spreadsheets, and comparing biofuel cropping systems

Vertika Rawat
Iowa State University

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



Part of the [Bioresource and Agricultural Engineering Commons](#)

Recommended Citation

Rawat, Vertika, "Analyses of bioenergy systems: detecting hard-coding errors in spreadsheets, and comparing biofuel cropping systems" (2011). *Graduate Theses and Dissertations*. 10226.

<http://lib.dr.iastate.edu/etd/10226>

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

**Analyses of bioenergy systems: detecting hard-coding errors in spreadsheets, and
comparing biofuel cropping systems**

by

Vertika Rawat

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

MASTER OF SCIENCE

Major: Agricultural Engineering

Program of Study Committee:

D. Raj Raman, Major Professor

Robert P. Anex

Matthew Z. Liebman

Tae H. Kim

Iowa State University

Ames, Iowa

2011

Copyright © Vertika Rawat, 2011. All rights reserved.

TABLE OF CONTENTS

CHAPTER 1. GENERAL INTRODUCTION	1
Objectives	2
Authors' Role	3
Literature Review for Chapter-2	4
Literature Review for Chapter-3	6
References	9
CHAPTER 2. DETECTING AND SUBCATEGORIZING HARD-CODING ERRORS IN BIOENERGY-RELEVANT SPREADSHEETS USING VISUAL BASIC FOR APPLICATIONS (VBA)	13
Introduction	13
Materials & Methods	15
Results & Discussion	19
Conclusions	24
Acknowledgements	24
References	25
CHAPTER 3. A COMPARATIVE ANALYSIS OF THE LIFE CYCLES OF MULTIPLE CROPPING SYSTEMS IN IOWA, REPRESENTED BY BOONE COUNTY	26
Introduction	26
Materials & Methods	28
Results & Discussion	36
Conclusions	45
Acknowledgements	46
References	47
CHAPTER 4. GENERAL CONCLUSION	51
APPENDIX A. HARD-CODING DETECTOR (HCD)	54

APPENDIX B. SUBCATEGORIZATION CODE (SubCat)	60
ACKNOWLEDGEMENTS	65

CHAPTER 1

GENERAL INTRODUCTION

Bioenergy production in the form of liquid fuels, heat and electricity constitutes about 14% of the global energy production (Yazan et al., 2010). Like any other technology, bioenergy has a spectrum of advantages and disadvantages associated with it. Biofuels, for instance, are expected to reduce the dependence on foreign oil considerably (United States Congress, 2007). However, the energy balance and carbon emission savings achieved by biofuels is still under question. Moreover, issues like “food vs. fuel” have been raised with authors like Ziegler (2007) claiming biofuels from crops as “*a crime against humanity*,” especially with respect to poor countries. However, amidst all this debate, global bioenergy development continues to take place and is being studied and investigated by scientists, engineers and government agencies worldwide. Simulation, modeling and analysis of bioenergy systems is an inherent and critical part of this process.

Much of the bioenergy simulation, modeling and analysis performed in academia or the industrial world takes aid of the spreadsheet software. Spreadsheets are widely used in many spheres of activity owing to their ease of use and multiple capabilities. However, related literature speaks that spreadsheets are vulnerable to a variety of errors. There is an elaborate taxonomy of spreadsheet errors given by Rajalingham et al. (2008). Many others like Powell et al. (2009) have come up with their own error taxonomies, specific to their breadth of analysis. To that end, there have been many attempts at detection, and removal or minimization of different kinds of errors in spreadsheets.

The second chapter of this thesis deals with the analysis of bioenergy-relevant spreadsheets with respect to their vulnerability to errors. In particular, we focused on the *hard-coding error*, which is a highly pervasive error in spreadsheets of many types. Using programs written in the VBA language embedded in Excel, we determined the rates of occurrence of hard-coding error in multiple bioenergy-relevant spreadsheets including key bioenergy spreadsheets like GREET and GBAMM. As spreadsheets today are used for critical

decision-making; our tools can help identify the weak zones of spreadsheets, so that users can remedy them and make spreadsheets much more reliable.

Results from bioenergy analyses have not always been convergent, and have sometimes caused controversies and polarized debates amongst the scientific community. A part of it can be attributed to the fact that bioenergy can be produced globally in numerous ways, using different feedstocks, different climates, and different cultivation methods (Whitaker et al., 2010). The third chapter of this thesis encompasses a life cycle analysis of six kinds of cropping systems for bioenergy production in the state of Iowa in the United States of America.

The concept of this project stemmed from an ongoing project called “COBS” or “Comparison of Biofuel Systems”, undertaken by researchers at Iowa State University. The cropping systems in COBS included both conventional cropping systems (C2: Corn in corn-soybean rotation, S2: Soybean in corn-soybean rotation, CC: Continuous Corn, CCW: Continuous Corn with a rye cover crop) and alternate cropping systems (P: unfertilized prairies, PF: fertilized prairies). The COBS project was established with the primary motive to examine the agronomic and ecological performance of the above cropping systems, and assess their feasibility as bioenergy systems.

The complexity and multiplicity of this experimental design motivated us to perform a life cycle analysis of these cropping systems, as a better estimator of the overall sustainability of the cropping systems. We have made an effort in this direction, and tried to model and analyze the aforementioned six cropping systems for the state of Iowa. This analysis is expected to help in a holistic assessment of these cropping systems as promising biofuel production systems for the future.

OBJECTIVES

The research objectives for this work were:

- To design a pair of VBA programs to automatically detect hard-coding errors in bioenergy-relevant spreadsheets, and then to subcategorize them

- To use the programs to characterize the cell error rates (CER) in bioenergy-relevant spreadsheets.
- To perform a life cycle analysis of six kinds of cropping systems in Iowa
 - To quantify the Fossil Energy Ratio (FER), Net Energy Yield (NEY), Global Warming Potential (GWP), and Eutrophication Potential (EP) of the six cropping systems for a period of 10 years
 - To quantify the Theoretical Biofuel Yield (TBY) of the six cropping systems for a period of 10 years

AUTHORS' ROLE

The research papers within this thesis are a result of the efforts of the first author, assisted and guided by the co-authors. All methods were performed by the first author, unless otherwise indicated.

Dr. D. Raj Raman (Associate Professor, Department of Agricultural & Biosystems Engineering, Iowa State University) was instrumental in providing a framework for the first paper, and guiding through the design and structure of the automated tools. Dr. Robert P. Anex (currently Professor, Biological Systems Engineering, University of Wisconsin-Madison) provided some direction along the technicalities and execution of the codes. Carol Faulhaber (MS student, Iowa State University) and Sami Khanal (PhD student, University of Wisconsin-Madison) were a valuable help in terms of troubleshooting.

The idea for the second paper was conceived by Dr. Robert P. Anex, who also was the key guide through the design and development of this paper. Dr. Matthew Z. Liebman (Professor, Department of Agronomy, Iowa State University) provided relevant information through the course of the project. David N. Sundberg (Agricultural Specialist, Department of Agronomy, Iowa State University) and Meghann E. Jarchow (PhD Candidate, Department of Agronomy, Iowa State University) also helped greatly to understand the related basics and supplied a lot of information. Dr. D. Raj Raman helped in the overall organization of all the thesis components.

LITERATURE REVIEW FOR CHAPTER 2

Errors in Spreadsheets – Spreadsheets are used in all possible spheres of intellectual and organizational activity, ranging from the mundane to the mission-critical (Powell et al., 2009). The spreadsheet software is invaluable in terms of its huge data storage, innovative features, user-friendly interface, and a broad range of tasks that it can perform. Since the inception of the computer era, users have been incessantly dealing with the problem of removal of errors (Powell et al., 2008). The spreadsheet software on account of its versatility and ease of use has garnered enough attention with respect to its vulnerability to errors, and the received wisdom is that errors are prevalent in spreadsheets of all kinds (Panko, 2005). Moreover, with the extensive use of spreadsheets in decision-making - errors in spreadsheets can translate into sub-optimal decisions being made, or even losses of millions of dollars in the business world (Galletta et al., 1997).

Taxonomy of Spreadsheet Errors – A substantial amount of research has gone into understanding how errors are created during the course of spreadsheet development, how deleterious their effect is on the spreadsheet, and how to mitigate them (Anderson and Bernard, 1988). These efforts culminated into the creation of a number of taxonomies of spreadsheet errors as a way to better understand the nature of errors, their commonalities and distinctions.

Galletta et al. (1993) were the first to conceptualize differences between classes of errors – *domain* errors and *device* errors. Domain points to a particular application area of spreadsheets, like accounting, whereas device refers to the spreadsheet technology. So a mistake in logic due to misunderstanding the concept of depreciation would be a domain error, but entering an incorrect reference in the depreciation equation would be a device error. However, it was Panko and Halverson (1996) who were the first ones to give a more complete and detailed classification. They distinguished between qualitative and quantitative errors. Quantitative errors lead to incorrect bottom-line values in the current model of spreadsheet, whereas qualitative errors degrade the quality of the model and can lead to erroneous results in the future runs. Panko and Halverson further divided quantitative errors

into *mechanical* errors (due to mistyping), *logic* errors (due to choosing the wrong function or creating the wrong formula), or *omission* errors (due to incomplete understanding of the model). Teo and Tan (1997) expanded the above taxonomy by the addition of two categories of errors: *jamming* errors (placing more than one parameter in a single cell) and *duplication* errors (when a single parameter is defined in more than one way). Rajalingham et al. (2000) have developed one of the most elaborate taxonomies till date, which is discussed at length in Chapter 2 of this thesis. However, a problem with elaborate taxonomies is that many categories tend to overlap with each other. Also, the usefulness of such a taxonomy in practice comes into question, if has not been tested for all kinds of errors listed on spreadsheets in use (Powell et al., 2008).

Detection of errors: The way spreadsheets have been examined for errors can be broadly categorized into two types. Either, the researcher asks the subjects to find errors deliberately made by him, or the experts look for errors in operational spreadsheets in the field audits (Powell et al., 2008). In the domain of research activity, most of these subjects know little about errors, and are given no information about the same. One exception is the two-part experiment by Teo and Tan (1997) who asked their subjects to first build a spreadsheet from a written problem description, and later change the embedded parameters in the spreadsheet and observe the impact. Teo and Lee-Partridge (2001) performed their testing on similar lines to assess the relation of error detection with the nature of error, prior practice and expertise.

Panko and Halverson (1997) performed an experiment on both individuals and groups to compare their error-finding abilities. The groups detected about two-thirds of all errors, whereas the individuals were able to find only one-third of errors. Purser and Chadwick (2006) conducted a web-based survey on students and professionals in two rounds. In the first round, no information was provided on the kinds of errors to find, and the second round involved the discussion of an error taxonomy prior to detection. Not surprisingly, the professionals were found to perform better than the inexperienced students; however, the knowledge of error types beforehand did not necessarily give better results. Clermont et al. (2002) used an auditing software in a field audit of three large spreadsheets. Butler (2000)

has described the auditing procedure used by HM Customs and Excise, which involves the use of SpACE, a software tool for government auditing of small-business tax returns. Powell et al. (2009) used two software tools (XL Analyst and Spreadsheet Professional) to audit their spreadsheets.

Hard-Coding Error: Rajalingham et al. (2008), in their detailed taxonomy of spreadsheet errors described hard-coding error as a kind of qualitative error which decreases the quality of the spreadsheet by making it less flexible. Powell et al. (2008) mentioned hard-coding input parameters into a formula a “risky” practice that would not be a cause of concern for the current version but could result in erroneous calculations in the subsequent versions of the model. Powell et al. (2009) audited 50 spreadsheets and placed the errors found across six well-defined error types of their self-designed taxonomy. They found that hard-coding errors (“placing numbers in a formula” in their words) were the most common of these errors in the audited spreadsheets.

It is apparent that software-based error detection in spreadsheets has not been very common. At the same time, manual efforts have not resulted in a very appreciable detection of spreadsheet errors despite consuming a substantial amount of time. Hence, we have taken a step in that direction by designing a pair of auditing tools to detect hard-coding error (a kind of qualitative error) in bioenergy-relevant spreadsheets, some designed in our lab, and some available in the public domain. The Chapter 2 of this thesis deals with the detection of hard-coding errors, its subsequent sub-categorization into four types, cell error rates, and related analysis.

LITERATURE REVIEW FOR CHAPTER 3

Biofuels: The issue of extensive use of fossil fuels in both developed and developing countries assumed paramount importance in the past decade. Global warming, dependence on foreign oil reserves, national energy security and sustainability are the top issues that have hence triggered the biofuel revolution (Cherubini and Stromman, 2010). Liquid biofuels, batteries, and hydrogen fuel cells are considered to be potentially viable technologies for road

transportation (Whitaker et al., 2010). The last two options have some logistical problems to overcome before they can be incorporated into the renewable fuel stream on a large scale. This leaves us with biofuels which have the maximum potential to replace petroleum fuels in the short-to-medium term as they are compatible with the current fuel infrastructure, as well as the current engine infrastructure. For the production of biofuels, agricultural biomass is used as the renewable carbon-based source. The prospect of large scale implementation of biofuel programme has thus brought different agricultural systems and their performances into focus (Gelfand et al., 2010).

Cropping Systems: The Energy Independence and Security Act of 2007 (EISA) mandates the US to produce 33 billion gallons of biofuels annually by 2021, of which 18 billion gallons per year is to be produced as “advanced biofuels.” Fuels obtained from a renewable carbon-source other than corn starch fall in the category of “advanced biofuels,” e.g., fuel from corn stover, or perennials like switchgrass or Miscanthus. This implies that cellulosic ethanol processes are going to play a critical role in the upcoming age of biofuels (Hill et al., 2006).

The corn-based cropping systems including the conventional corn-soybean rotation have played a pivotal role in the production of biofuel in the last decade. But these systems consume large amounts of N and P fertilizers, and are responsible for environmental burdens like soil erosion, extensive eutrophication and hypoxia (Landis et al., 2007). Moreover, the increasing demand for cellulosic ethanol from corn stover will cause farmers to harvest corn stover in addition to corn grain. Conventionally, most of the corn stover is retained on the field, and only 5% is harvested for animal feed and bedding (Glassner et al., 1999). However, removing stover from the fields in large quantities could give rise to increased rates of soil erosion and decreased soil organic carbon sequestration (Mann et al., 2002). It seems unlikely to achieve higher productivities from row-crop agriculture without a concomitant increase in ecological damage.

It is anticipated that perennials like Miscanthus and switchgrass will be major players in cellulosic ethanol production as they provide environmental benefits as well as economic

benefits (Heaton et al., 2004). Substituting annuals with perennials could be the solution to the rising ecological problems associated with row-crop agriculture, and will also aid in the preservation of natural biodiversity (Cook and Beyea, 2000). Also, herbaceous perennials have been shown to sequester carbon for multiple years, which could provide fiscal benefits if the carbon trading market attains some maturity (McLaughlin and Walsh, 1998).

Moreover, “diverse” species of prairies offer added advantages from an agronomical as well as ecological point of view (DeHaan et al., 2009). The diverse community is able to endure a wider range of negative environmental conditions in addition to having a stronger resistance to pest outbreaks (Tilman et al., 1997). Although there is a sizeable body of research that explores the ecological performance and productivity of diverse prairies, their efficacy as a promising biofuel feedstock hasn’t been researched to our knowledge. A direct assessment of these prairie systems and their comparison with the existing corn-based biofuel systems is essential for proper development of the biofuel industry.

Life Cycle Analysis: The production and use of biomass for bioenergy confers many advantages, but faces many repercussions as well. The energy-intensiveness of industrial agriculture due to its reliance on huge quantities of fossil inputs, which leads to adverse environmental impacts is well documented (Cleveland, 1995). The negative environmental impacts of energy crops should be completely assessed before final political decisions regarding them are made (Hanegraaf et al., 1998). Also, one of the requirements for the Environmental Protection Agency (EPA) is to assess the lifecycle GHG impacts associated with different types of renewable fuels (Energy Independence and Security Act, 2007). Another serious environmental issue is eutrophication, noticeably in the Midwestern region of the United States and the Gulf of Mexico, which could be exacerbated by an increase in the use of corn and soybeans for biofuels (Powers, 2007). Life cycle assessment (LCA) is a methodology widely used to examine all of the above - the energy balance and various ecological impacts of biofuel production (Whitaker et al., 2010). There is a global interest in developing a sustainability assessment protocol for biofuels, and LCA has been suggested as an appropriate method (Hill et al., 2006).

The third chapter of this thesis examines the life cycles of six cropping systems, developed for bioenergy production at an experimental site in Boone County, Iowa. Six kinds of cropping schemes were considered, which included both conventional energy crops like corn and soybeans, and alternatives like mixed prairies. A life cycle assessment of these systems was done to compare their energy use, greenhouse gas emissions and eutrophication potential in order to examine their performance as potential biofuel production systems in multiple dimensions.

REFERENCES

1. Anderson, K and Bernard, A. Spreading mistakes. *Journal of Accounting and EDP*, 4(1988): 42-50.
2. Galletta, F.; Abraham, D.; Louadi, M.E.; Leske, W.; Pollalis, Y.; and Sampler, J. An empirical study of spreadsheet error-finding performance. *Accounting, Management & Information Technology*, 3(1993): 79-95.
3. Cleveland, C.J. The direct and indirect use of fossil fuels and electricity in USA agriculture, 1910-1990. *Agriculture, Ecosystems and Environment*, 55(1995): 111-121.
4. Panko, R. and Halverson, P. Spreadsheets on Trial: a survey of research on spreadsheet risks. *Proceedings of the 29th HICSS Conference*, 1996: 326-335.
5. Galletta, D.F.; Hartzel, K.S.; Johnson, S.E.; Joseph, J.L.; and Rustagi, S. Spreadsheet Presentation and Error Detection: An Experimental Study. *Journal of Management Information Systems*. 13(1997): 45-63.
6. Panko, R. and Halverson, P. Are two heads better than one (at reducing errors in spreadsheet modeling?). *Office Systems Research Journal*, 15(1997): 21-32.
7. Teo, T. and Tan, M. Quantitative and Qualitative errors in spreadsheet development. *Proceedings of the 30th Hawaii International Conference on Systems Sciences*, 1997, 149-155.

8. Tilman, D.; Knops, J.; Wedin, D.; Reich, P.; Ritchie, M.; and Siemann, E. The Influence of Functional Diversity and Composition on Ecosystem Processes. *Science*, 277(1997), www.sciencemag.org accessed on 18th March, 2011.
9. Hanegraaf, M.C.; Biewinga, E.E.; and Bul, G.V.D. Assessing the ecological and economic sustainability of energy crops. *Biomass and Bioenergy*, 15(1998): 345-355.
10. McLaughlin, S.B. and Walsh, M.E. Evaluating environmental consequences of producing herbaceous crops for bioenergy. *Biomass and Bioenergy*, 14(1998): 317–324.
11. Panko, R. What we know about spreadsheet errors. *Journal of End-User Computing*, 10(1998): 15-21.
12. Glassner, D.; Hettenhaus, J.; and Schechinger, T. Corn stover potential: recasting the corn sweetener industry. In: *Janick J (ed) Perspectives on new crops and new uses*. ASHS, Alexandria, VA, 1999, pp 74–82.
13. Teo, T.S.H. and Lee-Partridge, J.E. Effects of error factors and prior incremental practice on spreadsheet error detection: an experimental study. *The International Journal of Management Science*, 29(2001): 445-456.
14. Mann, L.; Tolbert, V.; and Cushman, J. Potential environmental effects of corn (*Zea mays* L.) stover removal with emphasis on soil organic matter and erosion. *Agriculture, Ecosystems and Environment*, 89(2002): 149–166.
15. Heaton, E.A.; Clifton-Brown, J.; Voigt, T.; Jones, M.B.; and Long, S.P. Miscanthus for renewable energy generation: European Union experience and projections for Illinois. *Mitigation and Adaptation Strategies for Global Change*, 9(2004): 433–451.
16. Dalgaard, R.; Schmidt, J.; Halberg, N.; Christensen, P.; Thrane, M.; and Pengue, W.A. LCA of soybean meal. *The International Journal of Lifecycle Assessment*, 13(2006): 240-254.
17. Hill, J.; Nelson, E.; Tilman, D.; Polasky, S.; and Tiffany, D. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences of the United States of America*, 103(2006): 11206-11210.

18. Purser, M. and Chadwick, D. Does an awareness of differing types of spreadsheet errors aid end-users in identifying spreadsheet errors? *Proceedings of the European Spreadsheet Risk Interest Group Annual Conference, Cambridge, UK, 2006*, pp. 185-204.
19. Energy Independence and Security Act, 2007, http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_bills&docid=f:h6enr.txt.pdf accessed on 17th Dec, 2010.
20. Landis, A.E.; Miller, S.A.; and Theis, T.L. Life Cycle of the Corn-Soybean Agroecosystem for Biobased Production. *Journal of Environmental Science & Technology*, 41(2007): 1457-1464.
21. Powers, S. Nutrient loads to surface water from row crop production. *International Journal of Life Cycle Assessment*, 12(2007): 399-407.
22. United States Congress (2007) H.R. 2419-110th Congress: *Food, Conservation, and Energy Act of 2008*.
23. Rajalingham, K.; Chadwick, D.R.; and Knight, B. Classification of Spreadsheet Errors. *Computing Research Repository*, 2008, abs/0805.4224.
24. DeHaan, L.R.; Weisberg, S.; Tilman, D.; and Fornara, D. Agricultural and biofuel implications of a species diversity experiment with native perennial grassland plants. *Agriculture, Ecosystems and Environment*, 2009, doi:10.1016/j.agee.2009.10.017.
25. Powell, S.G.; Baker K.R.; and Lawson, B. Errors in operational spreadsheets. *Journal of Organizational and End User Computing*, 21(2009): 24-36.
26. Cherubini, F. and Stromman, A.H. Production of Biofuels and Biochemicals from Lignocellulosic Biomass: Estimation of Maximum Theoretical Yields and Efficiencies Using Matrix Algebra. *Energy Fuels*, 24(2010): 2657-2666.
27. Gelfand, I.; Snapp, S.S.; and Robertson, G.P. Energy Efficiency of Conventional, Organic, and Alternative Cropping Systems for Food and Fuel at a Site in the U.S. Midwest. *Journal of Environmental Science and Technology*, 44(2010): 4006-4011.
28. Whitaker, J.; Ludley, K.E.; Rowe, R.; Taylor, G.; and Howard, D.C. Sources of variability in greenhouse gas and energy balances for biofuel production: a systematic review. *GCB Bioenergy*, 2(2010): 99-112.

29. Yazan, D.M.; Garavelli, A.C.; Petruzzelli, A.M.; and Albino, V. The effect of spatial variables on the economic and environmental performance of bioenergy production chains. *International Journal of Production Economics* (2010), doi:10.1016/j.ijpe.2010.07.017.
30. Ziegler, J. The impact of biofuels on the right for food. Report No. A/62/289 (United Nations General Assembly, New York). /<http://www.righttofood.org/new/PDF/A62289.pdf> accessed on 21 Jan, 2011.

CHAPTER 2. DETECTING AND SUBCATEGORIZING HARD-CODING ERRORS IN BIOENERGY-RELEVANT SPREADSHEETS USING VISUAL BASIC FOR APPLICATIONS (VBA)

Modified from a paper published in *The Journal of Applied Engineering in Agriculture*

Vertika Rawat, D. Raj Raman, Robert P. Anex

INTRODUCTION

The versatility of spreadsheets has led to their extensive application at all levels of organizations. Because of their wide use, concerns have been raised about the integrity and validity of spreadsheets, as stated by Galletta et al. (1997), and many other authors including Powell et al. (2009) have shown that spreadsheets are highly vulnerable to errors. Users cannot readily detect the majority of such errors, which could result in potentially devastating miscalculations in many settings. The typical approach to debugging spreadsheets involves doing hand calculations to verify the results – unfortunately, this approach is time consuming and is frequently skipped or done cursorily. Furthermore, even if the spreadsheet is providing correct results with one set of input data, hidden errors can mean that when inputs change, incorrect values result.

With the pervasiveness of spreadsheet use, they are increasingly being used for mission-critical applications. Consequently, errors in spreadsheets can lead to making sub-optimal decisions as discussed by Teo and Lee-Partridge (2001). These errors cost the organizations that rely on them millions of dollars (EUSPRIG, n.d.). Panko (1999) showed that human-based code inspection – either in groups or individually – was only 60 – 80% effective at capturing errors in spreadsheets. Panko did not estimate the cost of such inspections, which would likely show human error detection is extremely expensive. A systematic and automated method of error detection could serve to reduce error rates and make spreadsheets more reliable.

A first step in developing any type of automated error detection system is to characterize the types of errors that can occur. To this end, Rajalingham et al. (2008) proposed an elaborate taxonomy for spreadsheet errors, wherein errors are broadly categorized as *system-generated* or *user-generated*. User-generated errors are further decomposed into *qualitative* or *quantitative errors*. Quantitative errors are numerical errors that lead to incorrect bottom-line values, as opposed to qualitative errors, which do not immediately produce incorrect numeric values but degrade the quality of the model.

Quantitative errors are further subdivided into *accidental errors* (due to typing errors), *omission errors* (failure to consider one or more important parameters), *alteration errors* (making changes to the model) and *duplication errors* (re-creating elements of the model). They could also fall into the categories of *domain knowledge errors* (stemming from a lack of knowledge), *mathematical representation errors* (due to inaccurate construction of a formula) or *logic/syntax errors* (due to erroneous logic or syntax).

Qualitative errors are trifurcated into *structural errors* (resulting from flaws in the design or lay-out of the model), *temporal errors* (from the use of data which has not been updated), and *maintainability errors* (from spreadsheet features which make it difficult to be modified). An extremely common maintainability error is the *hard-coding error*. Hard-coding errors (HCE) are defined in the literature as the use of raw numerical value(s) in cell formulae. For example, “=A3*2.204” or “=C7/365” are both HCE, whereas “2.204” or “365” or “1” coded into a cell are not, because the numerical value is not embedded in a formula. It is noteworthy formula cells have a disproportionately high share of errors, e.g., Powell et al. (2008) stated that approximately 80% of errors documented by EUSPRIG (European Spreadsheet Risks Interest Group) occurred in formula cells. The term *hard-coded* applies because it renders the formula, and hence the whole spreadsheet, inflexible to changing values in future scenarios. Updating a model containing HCE is time consuming because of the dispersion of numerical data throughout the spreadsheet.

Powell et al. (2009) applied a spreadsheet auditing protocol to 50 diverse operational spreadsheets, and reported that hard-coding errors were the most common (43.5% of erroneous cells), followed by logic errors (28.6% of erroneous cells) and reference errors

(22.1% of erroneous cells). The remaining categories in their own interim error taxonomy including copy/paste, omission, and data input errors together accounted for less than 5% of erroneous cells. In addition to their high frequencies of occurrence, hard-coding errors are cumbersome to detect manually.

However, hard-coding is vulnerable to automated detection, and in this paper we report on the results of a spreadsheet auditing effort in which hard-coding errors were automatically identified and subcategorized, thus addressing a need for such information identified by prior workers (e.g., Powell et al., 2008). An attempt to find errors using multiple manual strategies, was made by Galletta et al. (1997), but did not prove to be very effective. We have tried to take a step forward in that direction and have developed programs for hard-coding error detection in-house. As we scrutinized our audit results, we realized the importance of subcategorizing hard-coding errors when dealing with engineering spreadsheets, and added a second program with subcategorization capabilities. These capabilities extend beyond what is available commercially (e.g., auditing and error-checking tools available in Excel, and from third-party firms such as XL Analyst, <http://www.codematic.net>, and Spreadsheet Professional, <http://www.spreadsheetinnovations.com/>). We then applied the pair of programs to multiple bioenergy-relevant spreadsheets.

MATERIALS & METHODS

Overview

Both programs were written in Microsoft Excel Visual Basic for Applications (VBA). The first program identified hard-coding errors and presented a summary of error statistics and a detailed error report on a new worksheet tab. This tab was labeled HCER (Hard-Coding Error Report). The first program also flagged error cells in the respective worksheets using shading and font bolding to make it easy for users to locate them. The second program scanned the HCER summary, and subcategorized the errors into four unique types.

Algorithms

The first program (namely “HCD” – Hard-Coding Detector) stores all worksheet names in the workbook in a string array. Worksheets that are strictly charts/graphs are automatically skipped. The program displays the worksheet count and queries the user to see if there are any protected sheets in the workbook. If there are any protected worksheets, the cell shading and bolding functions are disabled. Because the detection algorithm can be misled by worksheet names containing numbers (e.g., “TAB_44”), the user is prompted to enter new names for any such worksheets, and the program assigns the new names. The program uses built-in functions to find row and column bounds of data for each worksheet, thus greatly reducing runtime. On each worksheet, the program loops through all cells in within the data bounds. Once a formula cell is found, the formula is stored in a string and parsed. If a number is encountered as the string is parsed, a check is made on the preceding element. If the predecessor turns out to be a letter, the program assumes that a cell address is specified, not an unreferenced numerical value. If this is the case, the program checks the successor string element too, skipping the successive string elements, as long as they are numbers. However, if the predecessor to a number was not a letter, a hard-coding error is flagged. A counter variable keeps track of the number of such instances. If a hard-coding error has been detected, the numerical value is checked to see if it is equal to one. If this numeral is “1” there is another counter variable that keeps track of the number of unity occurrences. The loop continues until the last element of the formula string of the concerned cell. At the end of this, the two counter variables are compared. If they are equal, the cell is solely suffering from the “unity” error.

The second program (namely “SubCat”) subcategorizes the hard-coding errors detected by the first program, by reading the formula of the faulty cell into a string, and then parsing it. When it runs into a number, it employs the same strategy described earlier to distinguish valid cell addresses from unreferenced numerical values. When the program locates an unreferenced numerical value, it subcategorizes according to the taxonomy shown in table 2.1. If a cell contains multiple subcategories, each type is captured and reported. While HCE already exist in the taxonomies on spreadsheet errors (e.g., Rajalingham et al., 2008; Powell et al., 2009), this is the first time they are being subcategorized to our knowledge.

Table 2.1. Taxonomy of hard coding errors implemented by program, indicating type, description, example, and comments.

Error Type	Description	Example	Comments
Unity errors	the presence of the value 1 as an unreferenced numerical value in the cell formula	$B6 = (1-A5)/B14$	The first “1” is the error, the second occurrence is not flagged because it is part of the cell address
Power of 10	the presence of numbers like 10, 100, 1000 and so on as an unreferenced value	$C21 = (B10-A2)*100$	The “100” is the error, the “10” in B10 is not flagged because it is part of the cell address
Commonly used unit conversions	the presence of common unit conversion factors relevant to bioenergy	$D4 = (C4*3.785)/ (B2-D4)$	The “3.785” is most likely a gallon-to-liter conversion
Other unidentified numerals	the presence of numerical values other than unity, power of 10, and unit conversion factors, as unreferenced values	$G4 = (D4 - 13.9)*(8 + G2)$	Both the “8” and the “13.9” are errors of this type

Interface

A series of dialog boxes are used for the primary user interface for the first program (the second program does not require any such dialog boxes). Message boxes, input boxes and radio buttons are used as follow: (a) to display the total number of worksheets in the workbook (b) to respond to whether there are any protected worksheets in the workbook (c) to display the tab names of worksheets which contain number(s) (d) to enter the new tab names for worksheets with numbers (e) to choose the background color and font of the cells to be flagged.

Error Statistics/Output

After HCD is finished running on all the selected worksheets, it displays the total number of cells checked, the number of cells with hard-coding errors, and the corresponding cell error rate (CER) of the audited workbook, in a popup box. The Cell Error Rate (CER), a generic term coined by Panko and Halverson (1996), refers to the frequency of error cells as a percentage of total cells in consideration. The complete error statistics also including the number of cells with hard-coding errors that are uniquely unity errors can be viewed on the

Hard-Coding Error Report tab (denoted by “HCER”). The HCER worksheet is created by the program after the last used worksheet of the workbook.

To facilitate rapid review of all errors, the HCER also presents a list of each error detected, sorted by worksheet, indicating both cell reference and the equation in the cell. If the only hard-coding error in a cell is a unity error, it is displayed with a grey fill to be easily distinguished from others. This helps the user to rapidly scan through the HCER by overlooking the grey-fill cells suffering from only unity errors (unique), as they tend to be far less dangerous than the others. The *unity-error (unique)* distinction is made because in certain formulae in engineering spreadsheets – such as when converting dry-basis to wet-basis moisture content – the use of a numerical one is justified and not indicative of a typical hard-coding error, as mentioned by Powell et al. (2008).

SubCat creates a subcategorization table (next to the report generated by the first program) on the same HCER worksheet. The subcategorization statistics include the frequency of (1) unity errors, (2) power of 10 conversions, (3) commonly used unit conversions, and (4) other unidentified numerals. The table shows the number of instances of each of the above type in each faulty cell detected by HCD.

We report the frequencies of errors identified by our programs distributed both across error types and across spreadsheets. To our knowledge, this is the first data on analysis of hard-coding errors and their subcategorization to appear in research literature on spreadsheet errors.

Audit of Bioenergy-Relevant Spreadsheets

The programs were used on the following six diverse workbooks related to simulation, modeling and analysis of bioenergy systems: The *Cob-Cost* workbook designed by Carol Faulhaber (MS student, Iowa State University) computes amortized grassroots capital cost of corn-cobs storage systems. The *Simple Framework for Analyzing Anaerobic Digestion (S-FAAD)* workbook, by Faulhaber and Raj Raman evaluates the economic viability of anaerobic digestion using a set of operating parameters and scale factors. The *Framework for the Evaluation of Bioenergy Feedstocks (FEBEF)* was developed by Raj Raman and Katrina

Christiansen (Ph.D. student, Iowa State University) to provide insight into the relative costs and lifecycle impacts of algae, switchgrass, Miscanthus, and corn. The *GREET-BESS Analysis Meta-Model (GBAMM)* (Energy and Resources Group, University of California, Berkeley, CA) compares life cycle global warming intensity estimates for corn ethanol as computed in BESS (Biofuel Energy Systems Simulator) and GREET (Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation) to understand why the results from the two models are so disparate. The *Ethanol-Profitability DI-10* workbook (Ag Decision Maker, Iowa State University Extension, Ames, IA), presents an economic model of a typical northern Iowa corn ethanol plant to help track its profitability of corn ethanol production. The *GREET Model* (Argonne National Laboratory, Argonne, IL) is a comprehensive model evaluating the energy use and emissions for diverse scenarios; GREET is available for the research community to use, and has been used in hundreds of refereed journal articles.

RESULTS & DISCUSSION

Sample Results

Figures 2.1 and 2.2 illustrate output from running HCD on a sample spreadsheet. Figure 2.3 illustrates the output from running SubCat on a sample spreadsheet.

Subcategorization of Hard-Coding Errors

In light of the large number of hard-coding errors detected in the sample spreadsheets, it appeared useful to further subcategorize them into the following:

- Unity errors
- Power of 10 conversions
- Commonly used unit conversion factors
- Other unidentified numerals

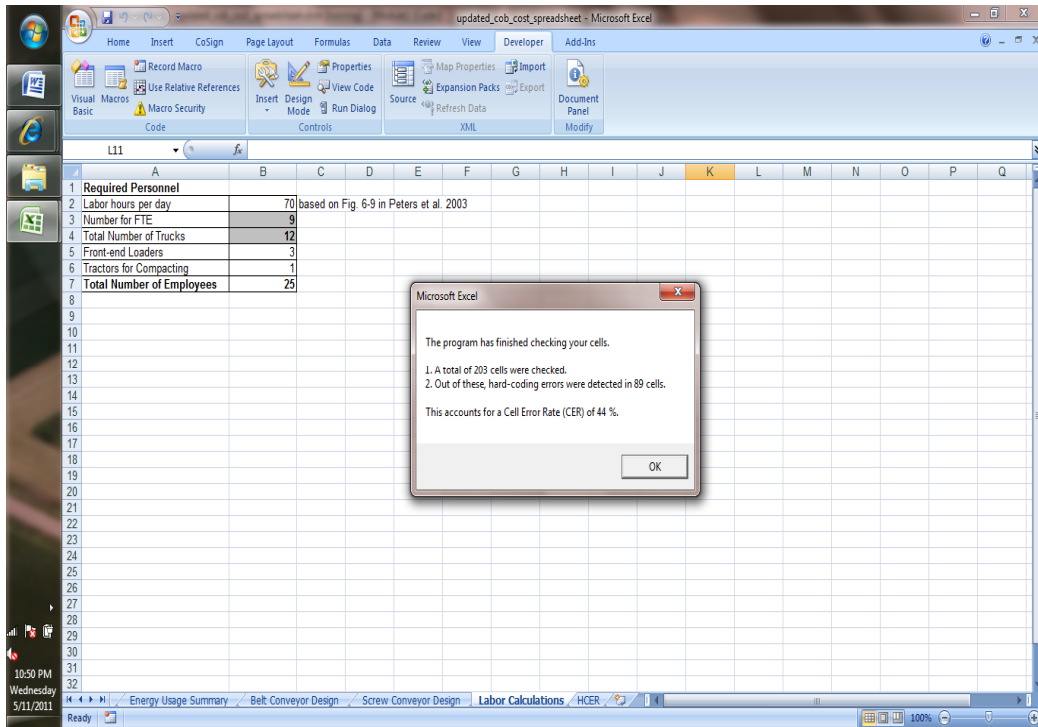


Figure 2.1. Screenshot of the final pop-up message box produced by HCD

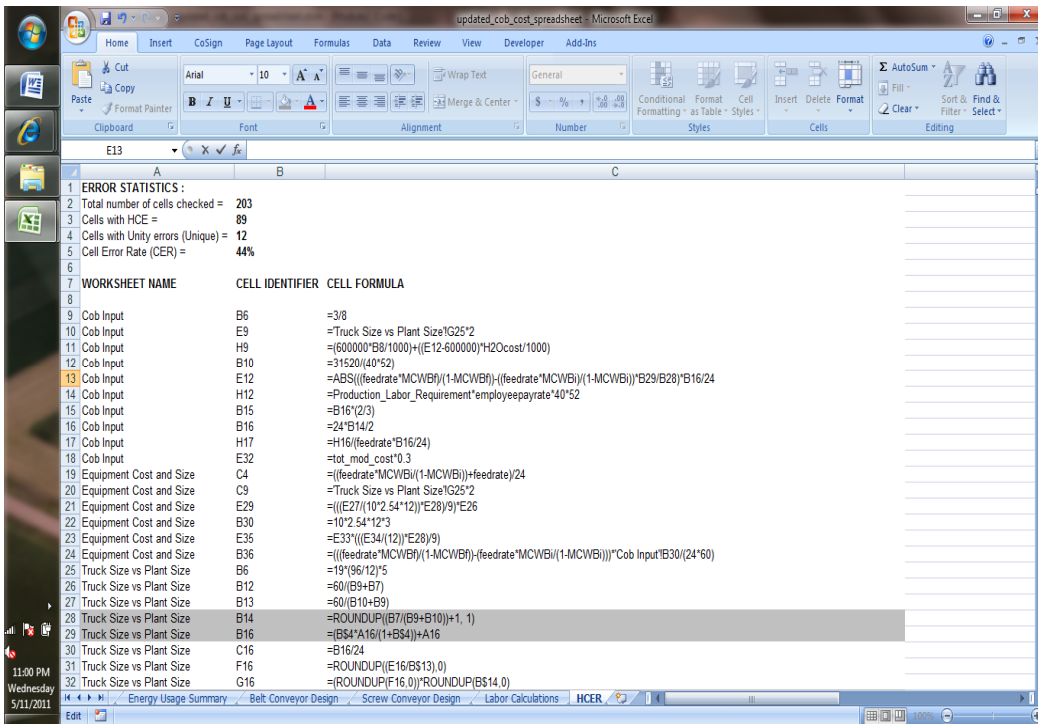


Figure 2.2. Screenshot of the Hard-Coding Error Report [Summary error statistics are shown at top of page, while specific error instances are listed below. Unity cells (unique) are marked with a grey fill.]

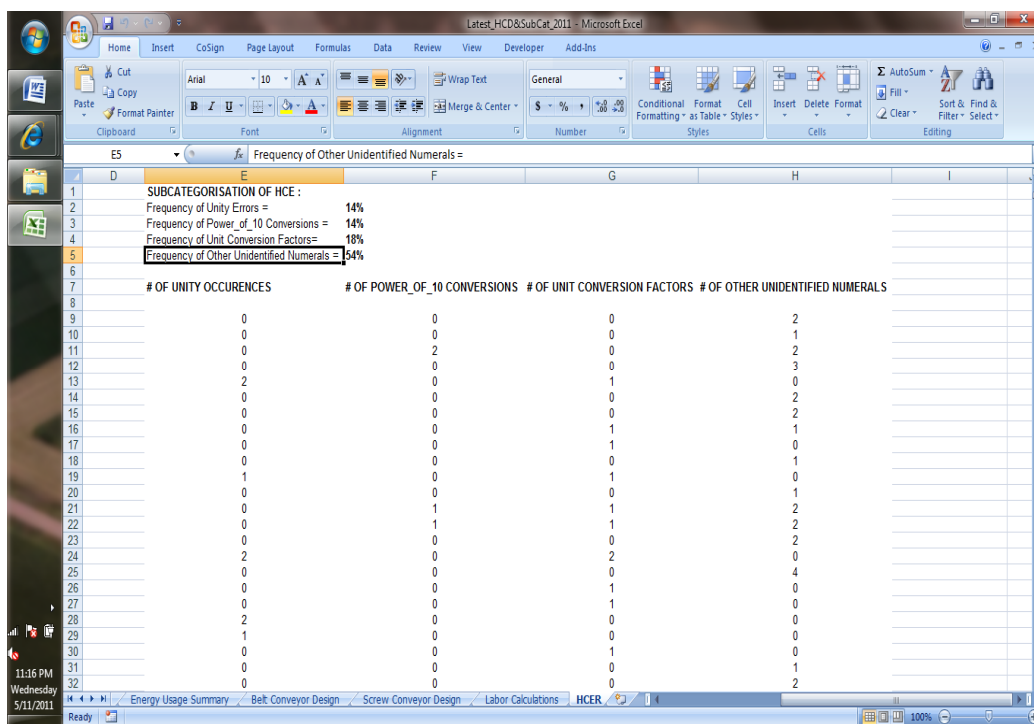


Figure 2.3. Screenshot of the results of sub-categorization of hard-coding errors [Summary statistics are shown at top of page, while a matrix of instances of each subcategory is shown below.]

Although power of 10 conversions can be unit conversion factors as well, they form a class of their own and have an overwhelming occurrence rate compared to the other commonly used unit conversion errors. For this reason, we chose to separate them from the other commonly used unit conversion errors.

Audit Results from Bioenergy-Relevant Spreadsheets

The results of the audits are shown in tables 2.2 and 2.3. Table 2.2 shows the Cell Error Rate (CER) of hard-coding errors in the tested spreadsheets ranged from 11% to 44%. The workbook with the lowest CER (FEBEF, 11%) originally had a 45% CER; the 11% reported reflected a major effort to remove hundreds of instances of hard-coding errors. If we had not actively improved FEBEF based on the audit, the minimum observed CER would have been 22% (GBAMM and Ethanol Profitability). While eliminating HCE instances from FEBEF, one of the co-authors of this article found a hard-coded cell which also contained a serious

mathematical representation error, that caused significant mistakes in the bottom-line values in that spreadsheet. This reveals yet another facet of hard-coding errors – namely their ability to mask other kinds of errors and consequently, be damaging to the spreadsheet. Although we only explored six spreadsheets, a total of nearly seventy-two thousand formulae cells were checked. Interestingly, but perhaps not surprisingly, we observed HCE frequencies similar to the 43.5% reported by Powell et al. (2009).

Table 2.2. Cell error rates (CER) of hard-coding errors (HCE) in the six tested spreadsheets

Workbook tested	Total number of cells checked	Number of cells with HCE	CER (%)
Cob Cost	203	89	44
GBAMM	462	100	22
S-FAAD	702	181	26
FEBEF	844	90	11
Ethanol Profitability	2757	608	22
GREET	66945	26867	40

Table 2.3. Sub-categorization of hard-coding errors from six tested spreadsheets [Each of the percentages is specific to the spreadsheet, i.e., 14% unity errors mean 14% of the total number of HCE instances in Cob Cost were unity errors.]

Workbook tested	Unity errors (%)	Power of 10 (%)	Commonly used Unit Conversions (%)	Other Unidentified Numerals (%)
Cob Cost	14	14	18	54
GBAMM	7	69	4	20
S-FAAD	21	7	46	26
FEBEF	94	0	0	6
Ethanol Profitability	8	77	15	0
GREET	47	25	9	19

Table 2.3 provides distribution statistics of the hard-coding errors in the six spreadsheets. The values in table 2.3 are the frequencies of each type as a percentage of the total hard-coding instances in the respective workbooks. Figure 2.4 provides a pie-chart representation of frequencies of each subcategory of hard-coding error for all six spreadsheets put together. For all the six spreadsheets, subtotals of instances of unity errors, power of 10 conversions, unit conversions, and other unidentified numerals were computed. The total count of hard-coding errors was obtained by summing up the four subtotals. Next, the frequencies of each

of the subcategories were calculated as a percentage of the total instances of hard-coding errors.

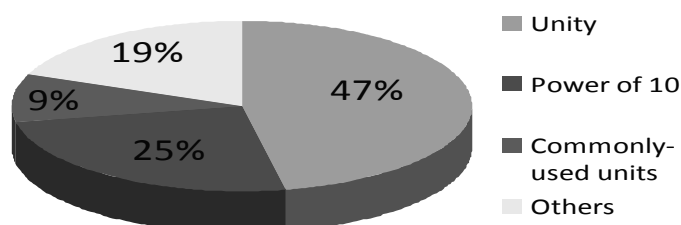


Figure 2.4. Distribution of sub-categories of HCE, showing preponderance of unity and power of 10 errors in the spreadsheets tested

Both GBAMM and Ethanol-Profitability workbooks suffered from high rates of Power of 10 conversions (69% and 77% respectively of the total instances of HCE), which justifies their separate categorization from unit conversions. Reflecting the effort to rid FEBEF of power of ten and unit conversion errors, unity errors predominate in FEBEF (at a rate of 94% of the total instances of HCE in FEBEF). The Cob Cost, S-FAAD and GREET had unity errors exceeding 14%, while the Ethanol-Profitability Workbook and GBAMM had fewer than 10%. Other unidentified numeral hard-coding errors formed a significant mass of errors in most of the spreadsheets with a frequency reaching as high as 54% of total HCE instances in the Cob-Cost Workbook. In some instances, their frequency even exceeds those of the commonly used unit conversion factors. Future versions of this program could allow users to specify additional numerical values used heavily in their spreadsheets.

To reduce the frequency of hard-coding errors and their impacts, spreadsheet authors can create an “Assumptions” tab in the beginning of the spreadsheet. By listing necessary conversions and other important constants used during the course of development of the spreadsheet, and then assigning them a brief but descriptive moniker (e.g., “Acresperha”, “rhoH2O” etc.) using the “define name,” functionality in Excel, one can get rid of hard-coding errors substantially. When dealing with unit conversions in particular, users can also

use the built-in “CONVERT” function of Microsoft Excel. Using 3.875 L/gal in one cell and 3.785 L/gal (the correct value) in another cell of the same spreadsheet leads to a 2.4% quantitative error, and constitutes a duplication error because the same value is being coded as multiple values in the spreadsheet. The “CONVERT” function will help to maintain consistency throughout the workbook and duplication errors can be avoided.

CONCLUSIONS

The frequency of hard-coded cells or the CER of hard-coding errors in the tested bioenergy-relevant spreadsheets ranged from 11 – 44%. This turns out to be a high error rate, especially since each occurrence is an opportunity for more serious numerical errors. We recommend the replacement of hard-coded values by unique descriptive monikers, as discussed before. By systematically using these named factors in equations, most hard-coding errors can be eliminated. Factors that occur rarely, perhaps in only one or two cells, can similarly be replaced by a named factor, but the cost-benefit ratio is questionable. Having a small fraction (e.g., less than 1%) of cells with such errors is probably not a major problem for most spreadsheets, especially if an auditing program such as the ones describe here are used to rapidly review any HCE instances. Along with structuring spreadsheets to make computations easy to follow, and clearly listing units on all quantities, elimination (or at least minimization) of hard-coding errors must be considered another fundamental part of good spreadsheet practices.

ACKNOWLEDGEMENTS

Partial financial support for this work was provided by the USDA Higher Education Challenge Grant Award # 2006-38411-17034. The authors would like to thank Carol Faulhaber and Sami Khanal for the technical guidance provided through the course of this project.

REFERENCES

1. EUSPRIG. (n.d.). European Spreadsheet Risks Interest Group stories. Available at: www.eusprig.org/stories.htm. accessed on 9th Jan, 2010.
2. Galletta, D.F.; Hartzel, K.S.; Johnson, S.E.; Joseph, J.L.; and Rustagi, S. Spreadsheet Presentation and Error Detection: An Experimental Study. *Journal of Management Information Systems*, 13(1997): 45-63.
3. Panko, R.R. and Halverson, R.P. Spreadsheets on Trial: a survey of research on spreadsheet risks. *Proceedings of the 29th HICSS Conference, Wailea HI, USA*, 2(1996): 326-335.
4. Panko, R.R. Applying code inspection to spreadsheet testing. *Journal of Management Information Systems*, 16(1999): 159-176.
5. Powell, S.G.; Baker, K.R.; and Lawson, B. A critical review of the literature on spreadsheet errors. *Decision Support Systems*, 46(2008): 128-138.
6. Powell, S.G.; Baker, K.R.; and Lawson, B. Errors in operational spreadsheets. *Journal of Organizational and End User Computing*, 21(2009): 24-36.
7. Rajalingham, K.; Chadwick, D.R.; and Knight, B. Classification of Spreadsheet Errors. *Computing Research Repository*, 2008, abs/0805.4224.
8. Teo, T.S.H. and Lee-Partridge, J.E. Effects of error factors and prior incremental practice on spreadsheet error detection: an experimental study. *The International Journal of Management Science*, 29(2001): 445-456.

CHAPTER 3: A COMPARATIVE ANALYSIS OF THE LIFE CYCLES OF MULTIPLE CROPPING SYSTEMS IN IOWA, REPRESENTED BY BOONE COUNTY

INTRODUCTION

Energy security and climate change due to continued use of fossil fuels are creating enormous concern at the global level as they impact multiple facets of our existence (Whitaker et al., 2010). Hence it is imperative to look into renewable energy options - one of which is biofuels. There are multiple options for the production of biofuel, with some agricultural systems being cost-effective but not so environment-friendly and some which may prove to be environment-friendly but are not so cost-effective yet. For example, annuals like corn and soybeans are proven biofuel feedstocks, but they can have a detrimental effect on the environment, which is more pronounced in continuous corn systems than corn-soybean systems (Miller et al., 2006, Landis et al., 2007). One of the common problems with row-crop agriculture is excessive nitrate leaching into water bodies as compared to perennials (McIsaac et al., 2010). Contrarily, there is yet no definite infrastructure for converting perennial crops into biofuels on a commercial scale. However, perennials can produce appreciable amounts of biomass while providing beneficial ecological services like carbon sequestration (Post and Kwon, 2000). This makes research into alternative cropping systems and their comparison with current biomass cropping systems desirable.

Per the Energy Independence and Security Act of 2007 (EISA) - the larger fraction of the total biofuel production by 2021 in the US is to be produced as “advanced” biofuels; i.e., biofuels produced from feedstocks other than corn starch. Hence, active research on cellulosic biofuel feedstocks and related conversion processes is underway as there is no single “best” way to produce biofuel yet. The biofuel industry in the US is expanding rapidly (Wang et al., 2007), but there is uncertainty regarding the relative efficacy of different cropping systems. A stronger understanding of the opportunities and tradeoffs for alternative cropping systems to provide biofuel feedstocks would be useful to policymakers and other analysts. An integrated approach to study and compare life cycles of multiple cropping

systems with multiple feedstocks provides a unique and holistic evaluation of the cropping systems as biofuel production systems.

The idea for this project was conceived from an ongoing project called COBS (Comparison of Biofuel Systems). The COBS project is a large scale experiment undertaken by researchers at Iowa State University in 2008 that investigates six kinds of cropping systems on the basis of their agronomic and ecological performance characteristics. Our aim was to make a direct comparison of the systems in COBS on the basis of their energy use, productivity and environmental impacts during their life cycle.

The dependence of many forms of conventional agriculture on fossil fuels is well known to the scientific community (Cleveland, 1995). Nitrogen fertilizer application to the soil, for instance, accounts for more than 30% of the total fossil energy use in biomass production (Kim et al., 2009). The adverse environmental impacts of energy crops due to their reliance on fossil energy should be completely assayed before their deployment on a large scale is considered (Hanegraaf et al., 1998). For example, eutrophication as a result of heavy use of fertilizers in row-crop agriculture is one serious environmental issue, which has caused hypoxia in the Gulf of Mexico (Powers, 2007). Also, one of the requirements for the Environmental Protection Agency (EPA) is to assess the lifecycle greenhouse gas emissions of renewable fuels (Energy Independence and Security Act, 2007). Life cycle analysis (LCA) is a computational tool that can be used to examine the energy balance and various other ecological impacts of biofuel production, and has been touted as an appropriate method to do the same (Hill et al., 2006).

A life cycle analysis of the six cropping systems in COBS was performed to assess their sustainability in the long term use. The objectives of this project were to quantify life cycle Fossil Energy Ratio (FER), Net Energy Yield (NEY), Global Warming Potential (GWP), and Eutrophication Potential (EP) of the COBS cropping systems for a period of 10 years. We also sought to determine the Theoretical Biofuel Yield (TBY) for the systems for a 10 year period, as a theoretical estimate of the maximum productivity of the system.

MATERIALS & METHODS

Experimental layout - The COBS Project was started in 2008, and conducted at South Reynoldson Farm, Boone County, IA (41°55'13" N, 93°44'54" W). It used a randomized block design experiment with four replications. Each block was divided into six plots [each plot: 61 m (200 ft) x 27 m (90 ft)] corresponding to the six treatments.

The six cropping systems in COBS under investigation were: 1) *C2*: Corn in Corn-Soybean rotation, 2) *S2*: Soybean in Corn-Soybean rotation, 3) *CC*: Continuous corn, 4) *CCW*: Continuous corn with a rye cover crop, 5) *P*: Diverse tallgrass prairie without nitrogen fertilization, and 6) *PF*: Diverse tallgrass prairie with nitrogen fertilization. In *C2* and *S2*, corn was grown for grain, and the stover was retained on the field. In *CC* and *CCW*, corn was grown for grain, and 60-70% of the stover was also harvested. All systems were managed without tillage.

Overview of LCA – We performed a Life Cycle Analysis to understand the performance of these cropping systems from multiple perspectives.

- *Goal of LCA* - To quantify life cycle Fossil Energy Ratio (FER), Net Energy Yield (NEY), Global Warming Potential (GWP), and Eutrophication Potential (EP) of the six cropping systems for a period of 10 years.
- *Functional Unit* – The following functional units were chosen:
 - (i) 1 kg of harvested aboveground biomass
 - (ii) 1 MJ of energy in harvested aboveground biomass
 - (iii) 1 ha of arable, non-irrigated land.

Having multiple functional units in an LCA enables us to look at the scenarios from multiple perspectives of energy, biomass and land. Hence, we have reported our results as (i) per unit of biomass (kg^{-1}), (ii) per unit of output energy (MJ^{-1}), (iii) per unit of land area (ha^{-1})

Although we studied the feasibility of the cropping systems as potential biofuel feedstock production systems, we did not choose refined biofuel as our functional unit. The benefits of including storage, transport, and conversion are offset by a huge increase in uncertainty (e.g., Cook and Beyea, 2000). This is because we are dealing with diverse cropping patterns including first and second-generation feedstocks. The second-generation cellulosic biofuel industry is still in its infancy, and it is not clearly known how conversion techniques will perform for these feedstocks in terms of efficiency, yield and value of co-products. Annual crops like corn and soybeans have a well defined post-harvest scheme (Dien et al., 2002) because of their use on a commercial scale but lignocellulosic feedstocks (perennials) have a poorly defined infrastructure for conversion into biofuels (Zheng et al., 2009).

- *Scope of LCA - Cradle-to-farm gate*

As mentioned above, the inclusion of post-harvest mechanisms by making an array of assumptions for such a diverse scenario would not result in a credible analysis. We have drawn the system boundary at the farm-gate immediately following biomass harvest. Although biomass storage and its transportation costs to the biorefinery are significant, we chose not to model beyond harvesting, because we are looking at multiple feedstocks like corn, soybeans and mixed prairies. The standard storage and transportation data for corn and soybean is readily available, but is hard to find for prairie species because they are not used as feedstocks on a commercial scale. There is no reliable source of data for standard storage conditions of a mixture of prairies.

- *Impact Categories* – We considered the following two potential impacts of the systems on the environment:

- a. *Global Warming Potential (GWP)* – One of the primary motivations to explore renewable energy sources is the emissions of carbon dioxide and other greenhouse gases by fossil fuels. However, modern agricultural systems also contribute to greenhouse gas emissions owing to the substantial use of fossil energy for production of fertilizers and pesticides, as fuel for field operations, and emissions of carbon dioxide and nitrous oxide from the soil (Gelfand et

al., 2010). Hence, it seemed pertinent to do a comparative analysis of greenhouse gas emissions by the six cropping systems in COBS. Cropping systems with large biomass productivity but with lesser GWP will be more competitive than the others. It is to be noted that the GWP here reflects the gross GHG emissions. Due to incomplete information on emissions from soil and soil carbon assumptions, calculation of a “net” GWP is not possible.

- b. *Eutrophication Potential (EP)* - One of the significant negative environmental impacts of Midwestern row crop agriculture is off-site transport of nitrogen and phosphorus (Smith et al., 2008). Increasing biofuel production could conceivably add to this problem, and one of the goals of the COBS Project is to see whether there are high-yielding cropping methods that can operate with low offsite transport of environmentally-relevant nutrients. The problem of eutrophication becomes all the more pronounced with systems like continuous corn and corn-soybean rotations which have high nitrogen requirements (Powers, 2007); eutrophication is expected to be significantly lower in prairie systems. Comparing systems on a eutrophication potential basis enables a richer understanding of the tradeoffs between systems.

Modeling Approach – The latest version 7.3 of SimaPro (*System for Integrated Environmental Assessment of Products*, developed by PRe Consultants, Amersfoort, The Netherlands) was employed to model the six cropping systems in COBS. SimaPro comes integrated with the ecoinvent database (developed by Competence Centre of the Swiss Federal Institute of Technology Zürich (ETH Zurich) and Lausanne (EPF Lausanne), the Paul Scherrer Institute (PSI), the Swiss Federal Laboratories for Materials Testing and Research (Empa), and the Swiss Federal Research Station Agroscope Reckenholz-Tänikon (ART)), and is considered to be the most successful LCA software worldwide. We created a new project called ‘COBS’ in SimaPro, and within ‘COBS’ new processes were created for each of the cropping systems. We built these processes taking all the material flows in and

out of the cropping systems from the already built-in processes of SimaPro database. The following methods in SimaPro were used to assess the following:

- *Cumulative Energy Demand Method*: To assess the cumulative fossil energy use (FEU) and the total energy use (TEU)
- *IPCC 2007 GWP 100a*: To assess the Global Warming Potential (GWP)

In a separate part of the COBS experiment, investigators measured tile drain leachate from all plots in 2009 and 2010, and determined mass loss rates for nitrate and phosphate in the tile drain flows. These values were used in lieu of the SimaPro-generated values for EP, because the SimaPro assumptions were not appropriate to these systems. For example, the EDIP/UMIP 97 method of SimaPro calculates EP based on the “N and P content in organisms.” The values hence calculated were not reflective of the actual leaching data collected from the COBS plots, because the SimaPro assumptions did not take essential parameters like slope of land, precipitation and general climate, crop type, growing season, type of cropping system, presence/absence of cover crop, soil type and erosion, presence/absence of soil conservation practices into consideration. The leaching data is courtesy of Dr. Matthew Helmers, Associate Professor, Agricultural and Biosystems Engineering, Iowa State University, Personal Communication, 2011. The calculation of EP for the systems from the leaching data was done following the guidelines from the works of Adisa Azapagic et al. (2004).

Scale-up – Most of our material flows into the cropping systems were taken from the field logs associated with the COBS experiments. However, we added energy inputs associated with anticipated practices for each system, if it were deployed at full-scale in Iowa. This ensured that we did not confine ourselves to the experimental site at the Boone County, and effectively modeled the agricultural systems in COBS on a large scale.

- *Fertilizers*: For all the cropping systems, we adopted most of the application rates of N, P, and K from the COBS field log. The fertilizer application rates are shown in Table 3.1 below. Some adjustments were made to the fertilizer application rates so that they are representative of the standard farming practices in Iowa. As per the field log, only one application of P and K was

made to the prairie plots in 2008 and none later. However, the prairie plots continued to give better yields every next year. Besides, a soil test should be taken every three years to maintain P & K concentrations in the soil for successful large-scale switchgrass production (USDA, Technical Note No. 3, 2009). Hence, we assumed the annual rates of P and K application to be used in our model as:

$$P \text{ (as P2O5)} = 70/3 \text{ lb/a} = 23.3 \text{ lb/a} = 26 \text{ kg/ha}$$

$$K \text{ (as K2O)} = 130/3 \text{ lb/a} = 43.3 \text{ lb/a} = 48.7 \text{ kg/ha}$$

Standard data for fertilizing is available for corn and soybean farming. The N fertilizer application rates in the COBS experiment were comparable to the general application rates given by NASS-USDA. However, in 2009, 168 kg/ha of P fertilizer was applied to C2, S2, CC, and CCW, which is more than double the application rate given in NASS-USDA statistics for both corn and soybeans. In 2008 however, 78 kg/ha of P fertilizer was applied to both corn and soybeans which is comparable to the 72 kg/ha given by NASS-USDA for the state of Iowa. So we have assumed an annual application rate of 78 kg/ha of P fertilizer to both corn and soybeans in our model. Also, the K fertilizer application in COBS was about 0.5 times more than the recommended rates. So we assumed an annual K application rate of 96 kg/ha for both corn and soybeans (NASS-USDA, Agricultural Chemical Applications, Iowa, 2005).

Table 3.1. Rates of N, P, K fertilizers (kg/ha) applied to each cropping system

[C2: Corn in Corn-Soybean rotation, S2: Soybean in Corn-Soybean rotation, CC: Continuous corn, CCW: Continuous corn with a rye cover crop, P: Diverse tallgrass prairie without nitrogen fertilization, PF: Diverse tallgrass prairie with nitrogen fertilization]

	C2		S2		CC	CCW	P	PF
	Corn	Soybean	Soybean	Corn				
N	138	0	0	140	155	188	0	84
P	78	78	78	78	78	78	26	26
K	96	96	96	96	96	96	48.7	48.7

- *Herbicides:* As per the field logs, glyphosate was the only pesticide applied to C2, S2, CC and CCW systems. However, atrazine, acetochlor and S-metolachlor have been the most common pesticides for corn production in the Mid-western United States (Landis et al., 2007). Hence, in addition to glyphosate, the model assumed atrazine (0.72 kg/ha/yr), acetochlor (1.94 kg/ha/yr), and S-metolachlor (2.16 kg/ha/yr) for corn in C2, corn in S2, CC and CCW (Minnesota Agricultural Statistics, NASS-USDA, 2004).

For soybeans, pendimethalin is commonly applied (Soybean Production Practices, Wisconsin, NASS-USDA, 2006). So, in addition to glyphosate, the model assumed pendimethalin (1.13 kg/ha/yr) for soybeans in C2 and soybeans in S2.

No herbicides were added to the P and PF systems in the COBS experiment. In our model too, we assumed none, as perennials require minimal inputs and management, in part because the diversity in plant species itself acts as a mechanism to combat pest infestation (Lewis et al., 1997).

- *Insecticides:* No insecticides were added to any of the COBS plots. P and PF plots would not require any insecticides because of the reasons mentioned above. Stacked gene varieties i.e. varieties containing biotech traits for both herbicide and insect resistance were applied to 61% of the corn acreage in Iowa (Biotechnology Varieties, Iowa and U.S., NASS-USDA). In the COBS experiment too, insect-resistant hybrid of corn was used. Hence, no insecticide was added to corn in the model. However, the model assumed lambda-cyhalothrin (0.024 kg/ha/yr) for soybean which is used to control soybean aphids (Agricultural Chemical Use Estimates, NASS-USDA, 2006).
- *Liming:* For the state of Iowa, the use of aglime (80% limestone+20% dolomite) is anticipated, as it is commonly available in the Mid-Western United States (West and Bride, 2005). The model assumed a lime application rate of 469 kg/ha for corn, based upon the GREET model (by Argonne

National Laboratory, Illinois). However, the frequency of application would be different for each system. Corn requires substantial amounts of N-fertilizer, which results in soil acidity relatively quicker than the rest of the systems. In light of this, we estimated the frequency of lime applications for the systems depending upon the amount of N-fertilizer application to each system. For continuous corn systems (CC and CCW), we assumed a period of 4 years. For C2 and S2 rotations, we assumed every seven years, since the N-fertilizer input to these systems is ~37-46% of that applied to the continuous corn systems. For unfertilized prairies (P), we didn't assume liming because it doesn't receive any N-fertilizer, so the soil PH is unlikely to drop on its own. However, for fertilized prairies (PF), we assumed a time period of 10 years as it receives ~ 44% the N-fertilizer applied to the continuous corn systems.

Calculations – We quantified the energy outputs of our agricultural systems by the energy content of harvested aboveground biomass. Published energy densities were used for corn grain, corn stover and soybean grain (Cruse et al., 2010), and bomb-calorimeter-based experimental values were used for unfertilized prairies and fertilized prairies. We took this experimental approach for prairies because although energy densities of individual prairie grasses are tabulated, the energy density of this precise mixture of grasses is not known. We restricted the “output energy” of our systems to the energy content of harvested biomass. Once the energy outputs were known, the fossil energy ratio (FER) was calculated as the ratio of the total energy output to the total fossil energy use of the system. We analyzed Global Warming Potential (GWP) and Eutrophication Potential (EP) as potential environmental impacts of the systems.

We also used the Theoretical Biofuel Yield (TBY) as another parameter to assess the productivity of the system. We conducted a series of experiments using the Ankom²⁰⁰ Fiber Analyzer and related methodology to find out the NDF, ADF and ADL values for feedstock samples from all treatments. Only corn stover was used from the corn plots. The cellulose and hemicellulose percentages were ascertained using the above values. Next, the TBY

(shown in Figure 3.1) for the cropping systems was calculated in terms of liters per ha (Theoretical Ethanol Yield Calculator, Information Resources, USDA). Since there is a lot of uncertainty revolving around which production processes will be used for which feedstocks, the TBY will be a measure of the maximum productivity of the system irrespective of the conversion processes used to convert the feedstock into biofuel. The TBY gives us the theoretical maxima of biofuel yields in terms of quantity of biofuel for all the feedstocks from the six cropping systems, except corn grains and soybean grains, for which we have realized ethanol yield of 2.8 gallons of ethanol per bushel of corn (University Extension, Iowa State University) and realized biodiesel yields of 1.5 gallons per bushel of soybeans (NBB).

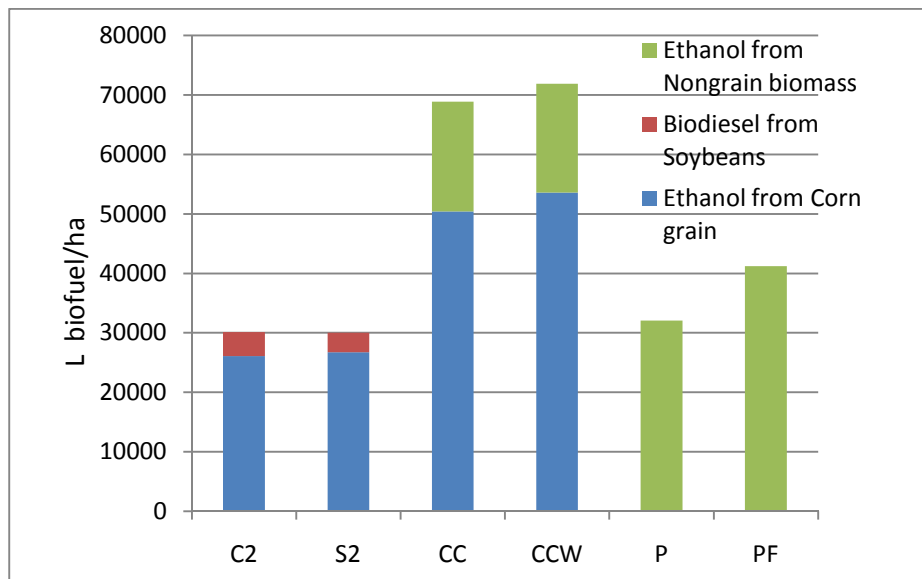


Figure 3.1. Theoretical Biofuel Yield (L/ha) vs. Cropping Systems for a period of 10 years

[C2: Corn in Corn-Soybean rotation, S2: Soybean in Corn-Soybean rotation, CC: Continuous corn, CCW: Continuous corn with a rye cover crop, P: Diverse tallgrass prairie without nitrogen fertilization, PF: Diverse tallgrass prairie with nitrogen fertilization]

Allocation issue: Since we harvested two products (corn grain and corn stover) in CC and CCW systems, it was necessary to allocate inventory data among the co-products. Generally, the three approaches followed to perform allocation are: economic allocation, energetic allocation, and system expansion. In energetic and economic allocation, the energy requirements and environmental impacts are distributed between the co-products on the basis

of energy content and economic factors respectively. In system expansion, however, one or more co-products are substituted by external product(s) in the analysis. For COBS, the method of energetic allocation was adopted. Because the cropping systems in this study were evaluated on the basis of their efficacy as bioenergy systems, and not on the basis of their cost structure; it is appropriate to perform energetic allocation between the co-products (corn grain and corn stover).

RESULTS & DISCUSSIONS:

The fossil energy use (FEU), global warming potential (GWP), and eutrophication potential (EP) per kg harvested product of the six cropping systems for a period of 1 year is shown in Table 3.2.

Table 3.2. Fossil Energy Use (MJ/kg), Global Warming Potential (g CO₂ eq. /kg), Eutrophication Potential (g NO₃ eq. /kg) for the six cropping systems

[C2: Corn in Corn-Soybean rotation, S2: Soybean in Corn-Soybean rotation, CC: Continuous corn, CCW: Continuous corn with a rye cover crop, P: Diverse tallgrass prairie without nitrogen fertilization, PF: Diverse tallgrass prairie with nitrogen fertilization]

	C2		S2		CC		CCW		P	PF
	Corn	Soybean	Soybean	Corn	Grain	Stover	Grain	Stover		
FEU	2.78	1.6	2.39	2.27	1.46	2.44	1.66	2.81	0.612	1.68
GWP	279	102	156	234	148	247	171	289	42.9	239
EP	3.50	2.38	8.45	0.60	1.33	2.23	0.93	1.58	0.069	0.13

The fossil energy use for soybean production in C2 (1.6 MJ/kg) was very close to the 1.54 MJ/kg value reported by USDA (2009). To our knowledge, there is no literature reporting life cycle results of mixed prairies. However, the fossil energy requirement for unfertilized prairies was found to be close to the 0.72 MJ/kg value given by Perlack et al. (1992) for switchgrass monocultures. The GWP of corn grain in our corn systems ranged between 148 – 279 g CO₂/kg of grain which fairly overlaps the 254 – 824 g CO₂/kg of grain given by Kim et al. (2009). The FEU values for corn grain in corn systems ranged from 1.46 – 2.78 MJ/kg of grain which fairly coincides with the 2.1 - 3.3 MJ/kg range found by Kim et al. (2009) for

different locations in the US Corn Belt. The FEU values for corn stover were between 2.44 – 2.81 MJ/kg of stover which turns out to be much higher than the 0.85 - 0.98 MJ/kg range given by Kim et al. (2009).

This is because results are heavily influenced by the allocation scheme used. In the case of energetic allocation between corn grain and corn stover, the larger share of the life cycle inventory is allocated to the corn stover because it has a larger higher heating value (HHV) than corn grain (~18 MJ/kg compared to ~17.6 MJ/kg). Moreover, the results per kg of the harvested material also depend on the relative amounts of co-products harvested. In COBS, approximately 65% of the stover was harvested in CC and CCW, and the rest retained. Hence the share of corn stover per unit mass in general turned out to be higher than that of the grain.

The fossil energy use on a per unit biomass basis was largest for the corn crop, followed by soybeans, followed by prairies. Most of the fossil energy used in agriculture comes indirectly in the form of fertilizers and pesticides, and directly as diesel fuel for field operations, and natural gas for seed drying operations. Corn production requires large amount of inputs and management, and hence has a large fossil energy use. Since soybeans are in rotation with corn in C2 and S2, no N-fertilizer was added to them, which resulted in a lower fossil energy use than corn. Fertilized prairies have a decent fossil energy use despite a significantly lower application of P and K fertilizer inputs than the corn-based systems, and absence of pesticides. This is because of the use of N-fertilizer in considerable amount (84 kg/ha/yr) on fertilized prairies. Unfertilized prairies on the other hand have a strikingly low fossil energy use due to minimal P and K inputs, absence of N-fertilizer and pesticides, and hence minimal management in the form of farm operations. The variation in Global Warming Potential follows the same trend, as more fossil energy inputs lead to more greenhouse gas emissions of nitrous oxide post N-fertilizer application, and more carbon dioxide emissions through field operations, added to the indirect emissions during the manufacture of agricultural chemicals.

One of the functional units for our analysis is 1 MJ of harvested biomass, and the SimaPro results per MJ of harvested products for the cropping systems are reported in Table 3.3.

Table 3.3. Fossil Energy Use (MJ/MJ), Global Warming Potential (g CO₂ eq./MJ), Eutrophication Potential (g NO₃ eq./MJ) for the six cropping systems

[C2: Corn in Corn-Soybean rotation, S2: Soybean in Corn-Soybean rotation, CC: Continuous corn, CCW: Continuous corn with a rye cover crop, P: Diverse tallgrass prairie without nitrogen fertilization, PF: Diverse tallgrass prairie with nitrogen fertilization]

	C2	S2	CC	CCW	P	PF
FEU	0.131	0.122	0.103	0.118	0.0356	0.0968
GWP	12.6	11.8	10.4	12.2	2.5	13.7
EP	0.078	0.063	0.094	0.067	0.004	0.007

It is important to note here that the cropping systems in consideration have different rotational periods. C2 and S2 comprise a corn-soybean rotation, and have a 2 year cycle. CC and CCW comprise a single annual corn crop, and so have a rotational period of 1 year. However, there is no agreement regarding the longevity of prairie grass stands. Moreover, no such information is available for a mixture of prairies. For the six cropping systems to be compared from multiple perspectives, it is necessary to have a common time frame for all of them. Hence, we *estimated* a time period of 10 years which lies well within the range projected by Parrish and Fike (2005) for productive switchgrass stands. The total environmental impacts (fossil energy use, energy output, fossil energy ratio, total energy use, net energy yield, global warming potential, eutrophication potential for the six cropping systems for a period of 10 years is shown in Table 3.4. As mentioned before, the energy output (EO) of the systems refers to the product of the energy content/caloric value per unit biomass and the total amount of harvested biomass in the system. For C2 and S2, we summed up five years of corn data and five years of soybean data.

Table 3.4. Fossil Energy Use (GJ/ha), Total Energy Use (GJ/ha), Energy Output (GJ/ha), Fossil Energy Ratio, Net Energy Yield (GJ/ha), Global Warming Potential (Mg CO₂ eq. /ha), and Eutrophication Potential (Mg NO₃ eq. /ha) for a period of 10 years

[C2: Corn in Corn-Soybean rotation, S2: Soybean in Corn-Soybean rotation, CC: Continuous corn, CCW: Continuous corn with a rye cover crop, P: Diverse tallgrass prairie without nitrogen fertilization, PF: Diverse tallgrass prairie with nitrogen fertilization]

	C2	S2	CC	CCW	P	PF
FEU	186	185	346	404	44	171
TEU	197	196	364	424	46	176
EO	2,836	3,042	3,365	3,413	1,230	1,764
FER (EO/FEU)	15.2	16.4	9.7	8.4	27.9	10.3
NEY (EO-TEU)	2639	2846	3001	2989	1184	1588

Table 3.4. (Continued)

GWP	17.9	17.9	35.1	41.5	3.1	24.2
EP	0.220	0.190	0.317	0.227	0.005	0.013

The graph in Figure 3.2(a) below depicts the variation in Fossil Energy Ratio (FER) with the cropping systems in consideration. The fossil energy ratio of a system is defined as the ratio of the renewable output energy to the non-renewable or fossil energy going into the system (USDA, 2009).

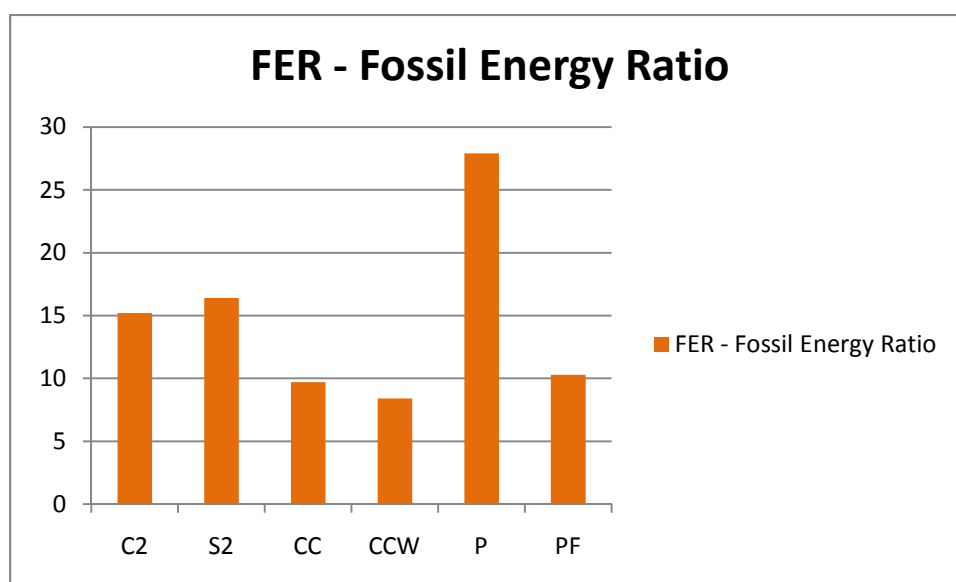


Figure 3.2(a). Fossil Energy Ratio of the cropping systems for a period of 10 years

[C2: Corn in Corn-Soybean rotation, S2: Soybean in Corn-Soybean rotation, CC: Continuous corn, CCW: Continuous corn with a rye cover crop, P: Diverse tallgrass prairie without nitrogen fertilization, PF: Diverse tallgrass prairie with nitrogen fertilization]

As can be seen, the fossil energy ratios of C2 and S2 both of which comprise a corn-soybean rotation are almost the same. CC and CCW, which represent the continuous corn systems, have smaller fossil energy ratios than the corn-soybean rotations. Although the biomass productivities of CC and CCW are higher than those of C2 and S2 because of harvesting grain and stover both, CC and CCW consume a much larger proportion of fossil energy inputs which results in lower fossil energy ratios. Most notably, it is due to more N-fertilizer application in case of continuous corn systems than corn soybean rotations (Feng et al., 2010). Unfertilized prairies have the highest fossil energy ratio, which can be ascribed to the fact that they have high biomass production as compared to their significantly low fossil

inputs. Fertilized prairies, on the other hand, give a greater energy output than unfertilized prairies in terms of the harvested biomass, but have a smaller fossil energy ratio because of the inclusion of N-fertilizer in the system.

The graph in Figure 3.2(b) below depicts the variation in Net Energy Yield (NEY) with the cropping systems in consideration. Interestingly, the Net Energy Yield analysis doesn't reproduce the same results as those of Fossil Energy Ratio.

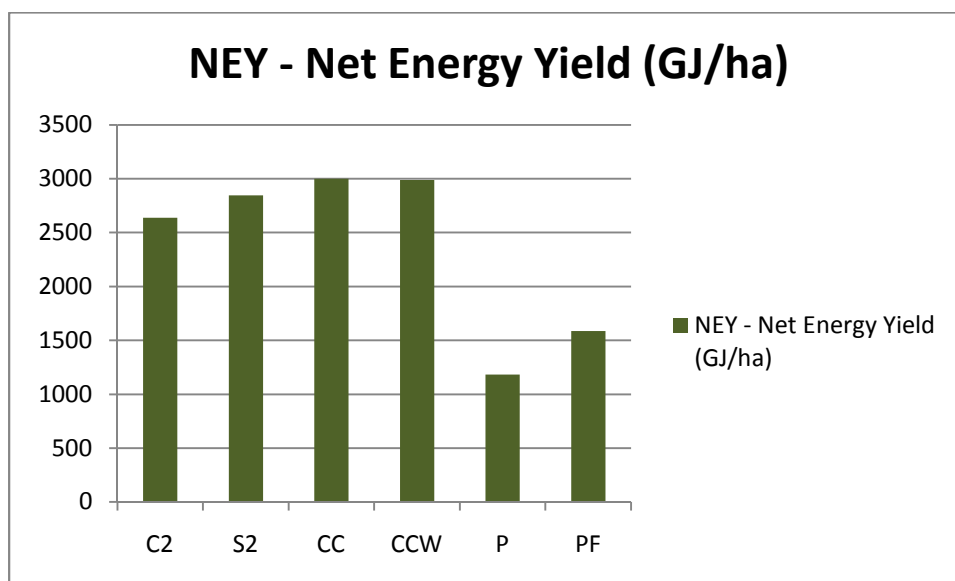


Figure 3.2 (b). Net Energy Yield (GJ/ha) of the cropping systems for a period of 10 years

[C2: Corn in Corn-Soybean rotation, S2: Soybean in Corn-Soybean rotation, CC: Continuous corn, CCW: Continuous corn with a rye cover crop, P: Diverse tallgrass prairie without nitrogen fertilization, PF: Diverse tallgrass prairie with nitrogen fertilization]

The continuous corn systems (CC and CCW) have the highest NEY, followed by the corn-soybean rotations (C2 and S2), followed by the mixed prairies (P and PF). The NEY has been calculated as the difference of the energy output (EO) and the total energy use (TEU) of the system. The total energy use includes nuclear energy and multiple forms of renewable energy in addition to the fossil energy, which is automatically calculated by SimaPro. CC and CCW consume the maximum amounts of fossil energy, but they very high biomass yields compared to the rest of the systems, resulting in the maximum NEY. Unfertilized prairies (P) have the highest fossil energy ratio but the lowest net energy yields. They have a significantly high FER as compared to the rest of the systems owing to the minimal use of

fossil energy inputs in the form of chemical additives. Since FER is a ratio, if the denominator (FEU) is close to zero, it will give a dramatically high value even if the numerator (EO) is small in absolute terms. Fertilized prairies have a greater NEY than unfertilized prairies despite more consumption of fossil energy in the form of N-fertilizer. This is because the biomass yields from fertilized prairies were about 43% higher than the yields from unfertilized prairies.

Figure 3.3 below shows the varying trends in the global warming potential (GWP) across the six cropping systems for a period of 10 years. The prairie bars can be seen to have a different fill than the rest of the cropping systems because we recognize that these data do not include carbon stored in the soil and so tend to overestimate the GWP of prairie cropping systems that have the ability to deposit significant amounts of carbon belowground. Prairies have been shown to sequester carbon for multiple years (McLaughlin and Walsh, 1998).

Therefore, the GWP values reported in our analysis reflect gross GHG emissions for all systems, and not a net GHG flux because they do not take into account natural sinks like soil carbon sequestration that have been identified in previous studies e.g. Adler et al. (2007). If carbon sequestration were taken into account, it is expected to offset some portion of the gross emissions in prairie systems and would result in a substantial reduction in their GWP as compared to the corn-based systems.

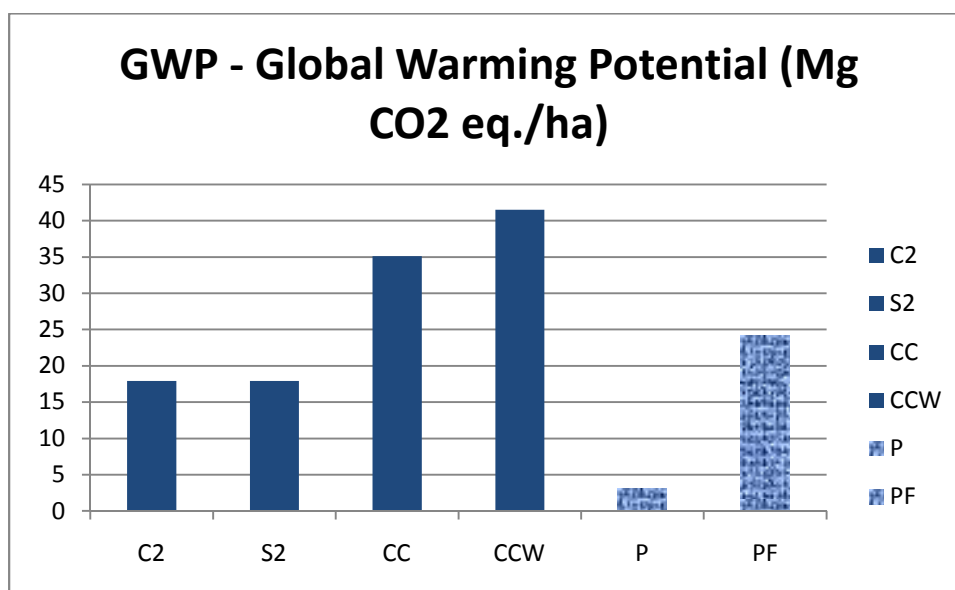


Figure 3.3. Global Warming Potential (Mg CO₂ eq./ha) of the cropping systems for a period of 10 years

[C2: Corn in Corn-Soybean rotation, S2: Soybean in Corn-Soybean rotation, CC: Continuous corn, CCW: Continuous corn with a rye cover crop, P: Diverse tallgrass prairie without nitrogen fertilization, PF: Diverse tallgrass prairie with nitrogen fertilization]

As shown in Figure 3.3, the continuous corn systems have the highest gross GHG emissions followed by the corn-soybean rotations followed by the unfertilized prairie systems. This is closely related to the rationale for fossil energy ratios i.e. larger fossil energy inputs result in larger greenhouse gas (GHG) emissions. Agricultural operations result in the production of greenhouse gases predominantly through fossil fuel use, use of nitrogen fertilizer, and soil disturbance in the form of tillage (Johnson et al., 2007). Fossil fuel used in farm machinery for the field operations results in direct emissions of carbon dioxide to the atmosphere. Indirectly, carbon dioxide is also produced in the manufacture of nitrogen fertilizer, when natural gas is combined with atmospheric nitrogen to yield ammonia and carbon dioxide. Following application, nitrogen fertilizer causes emissions of nitrous oxide (N₂O), a greenhouse gas 298 times more potent than carbon dioxide (McSwiney et al., 2010). The continuous corn systems have a much higher N-fertilizer use followed by the corn-soybean rotation followed by the prairies. CCW has a higher GWP than CC as it requires more N-fertilizer to replenish the amount consumed by the rye cover crop during its growth. Fertilized prairies have a significantly higher value of greenhouse gas emissions than unfertilized prairies, again due to the use of N-fertilizer. Although fertilized prairies have the same application rates of P and K as unfertilized prairies, but P and K fertilizers have an insignificant GHG impact (Feng et al., 2010), as can also be seen from Figure 3.3.

As noted earlier, since carbon sequestration is ignored, the GWP estimates are incomplete, and must be considered preliminary. Data are forthcoming as the COBS experiment matures but will not be presented in this thesis.

Figure 3.4 below depicts the variation in eutrophication potential with respect to the six cropping systems. As mentioned before, the eutrophication potential values listed in our analysis are derived strictly from the observed tile drain leachate values from the systems in the COBS experiment.

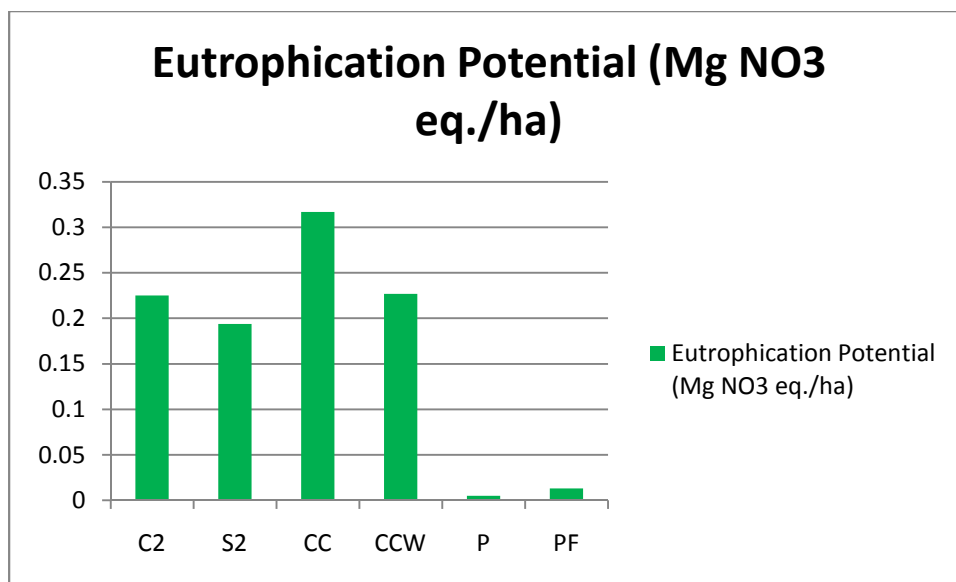


Figure 3.4. Eutrophication Potential (Mg NO₃ eq./ha) of the cropping systems for a period of 10 years

[C2: Corn in Corn-Soybean rotation, S2: Soybean in Corn-Soybean rotation, CC: Continuous corn, CCW: Continuous corn with a rye cover crop, P: Diverse tallgrass prairie without nitrogen fertilization, PF: Diverse tallgrass prairie with nitrogen fertilization]

The continuous corn systems exhibit the highest eutrophication potential followed by the corn-soybean rotation followed by the prairie systems. The EP for CC (31.7 kg NO₃ eq./ha) and CCW (22.7 kg NO₃ eq./ha) lies well within the 38.5+- 15.9 kg NO₃ eq./ha range for corn found by Miller et al. (2006) using Monte Carlo simulations in her model. The eutrophication potential for soybeans in S2 (30.9 kg NO₃ eq./ha) and soybeans in C2 (8.77 kg NO₃ eq./ha) were also found to be within the 20.8+-16.5 kg NO₃ eq./ha range found by Miller et al. (2006). It is important to note here that these values most likely underestimate the actual EP since the nutrient data used does not take nutrient loss through soil erosion into account. Although, due to flat fields used in the COBS experiment, the nutrient loss by erosion is expected to be very small in all treatments (smallest in prairie systems), yet these data should be considered incomplete.

CC and CCW have a higher EP due to the use of N-fertilizer every year compared to the use of N-fertilizer every other year by C2 and S2 systems, as no nitrogen is applied to soybeans. CCW has a lower EP than CC most likely because of the presence of rye cover crop which causes less nutrients leaching out of the system in spring. The use of a cover crop during the off growing season reduces nutrient losses from soil (Kim et al., 2009). The prairie systems

have the least EP not only because of minimum use of N and P fertilizers amongst all the systems, but also because of their extensive root structure which keeps the nutrients intact in soil, and continuous soil cover that perennials provide (McLaughlin and Walsh 1998, Parrish and Fike 2005). However, we can see that fertilized prairies have a greater EP than that of unfertilized prairies, which is likely due to more leachate coming out of PF due to use of N-fertilizer (84 kg/ha/yr), as compared to none in unfertilized prairies.

Likelihood of Error in Results: The GWP and energy use (FEU and TEU) of the six cropping systems were ascertained using the SimaPro software. When systems are modeled in SimaPro, all the inputs (fertilizers, pesticides, diesel etc.) going into the systems become “processes” of the system. Inventory data for each of these processes is extracted out of the ecoinvent database which is an integral part of the SimaPro software. Most of these LCI data are for Switzerland (CH) or other Western European countries (RER). We chose RER over CH since the farming practices in the US are more similar to the rest of Europe than Switzerland. As far as the lifecycle inventory data for the US is concerned, it is in its development phase. However, since there would still be considerable differences pertaining to the agro-climatic zones and industrial practices between Europe and the US, these results (GWP, FER derived from FEU, and NEY derived from TEU) are the most sensitive to error.

As discussed before, the calculation of EP was based on the actual leaching data from the COBS plots, but since this data does not take nutrient loss through erosion into account, the ascertained values are likely to underestimate the actual EP values. On the other hand, the calculation of energy output of the prairie systems was based on the caloric values of the prairie biomass, so there is a high confidence associated with these results. Also, the use of published energy densities of corn grain and corn stover to calculate the energy output of the corn-based systems would involve minimal error since there is not much variation in their published values. The calculation of TBY was done with the help of a series of Ankom Fiber Analyzer experiments and related methodology, and USDA guidelines, so these are also expected to have very low error rates.

CONCLUSIONS:

In this study, the LCA methodology was used to model and evaluate six kinds of cropping systems, experimentally established in Boone County, Iowa. The systems were compared on the basis of fossil energy ratio (FER), net energy yield (NEY), global warming potential (GWP), eutrophication potential (EP), and theoretical biofuel yield (TBY). C2 and S2 were found to have almost similar fossil energy ratios and environmental impacts, which is understandable as C2 and S2 are the two phases of a single system. The continuous corn systems, on the other hand, represent the lowest fossil energy ratios, coupled with higher environmental impacts. CCW leads the systems with a GWP of 41.5 Mg CO₂ eq./ha, whereas CC leads the systems with an EP of 0.317 Mg NO₃ eq./ha. It is however important to note that although continuous corn systems exhibit the highest negative environmental impacts; they have the maximum theoretical biofuel yield (TBY) and the highest net energy yield (NEY), both of which reflect their highest productivities as biofuel feedstock production systems. Contrary to this, the prairie systems (P and PF) have low net energy yields, but have minimum detrimental environmental impacts with respect to eutrophication. Unfertilized prairies have the least negative environmental impact in terms of GWP, EP and FER; but have the lowest net energy yield of all systems.

The results lead to contemplation that for the rapid expansion of the biofuel industry, the corn-based systems may be a good choice for the present and the near future. The nutrient exports and soil erosion concerns about corn are well documented, but given its high productivities and operational biofuel conversion infrastructure, it continues to be a successful biofuel crop. However, considering environmental impacts, the corn-soybean system seems to be a better alternative than continuous corn systems. Continuous corn systems with stover removal have a higher Theoretical Biofuel Yield (TBY) than corn-soybean rotations, because of the contribution of ethanol from stover as well. But due to higher fertilizer use, they do not perform well from an ecological point of view. Moreover, TBY is just a theoretical estimate of the system productivities, and wouldn't reflect a true picture as long as the cellulosic ethanol technologies become competitive with the existing ones. Prairies are a good choice of feedstocks for bioenergy from an ecological point of view.

Not only do they consume less fossil energy in the form of fertilizers and pesticides, they cause minimum damage to the environment in terms of GWP and EP. As discussed before, the calculations of GWP and EP are based on incomplete data, and therefore subject to revision.

From an agro-ecological perspective, fertilized prairie systems (PF) perform the best with moderate biomass yields and significantly lower environmental impacts as compared to the corn systems. The additional use of N-fertilizer give fertilized prairies a lower fossil energy ratio and higher negative environmental impacts than the unfertilized prairies, but fertilized prairies have a higher theoretical biofuel yield and a higher net energy yield on account of higher biomass production per unit land. However, the net energy yield (NEY) of fertilized prairies is still well below those of corn systems. Therefore, improvement of prairie strains to give yields comparable to the corn systems, and development of feasible cellulosic ethanol production technologies may be the key to a major development of the prairie systems as potential biofuel production systems in the future, and consequently reduce the ecological pressure on row-crop agriculture.

It is important to note, however, that this was a preliminary study, and the models in our analysis were predominantly based on early data from the COBS experiment. As the experiment goes on, the prairie yields are expected to increase further which would reduce their already small negative environmental impacts of global warming and eutrophication per unit mass. Moreover, the Net Energy Yields (NEY) and Fossil Energy Ratios (FER) would increase dramatically when aggregated over a period of 10 years. This would very likely change the relative comparison of the cropping systems with each other, and would bring down the scales more in favor of perennials.

ACKNOWLEDGEMENTS

Partial financial support for this work was provided by the USDA Higher Education Challenge Grant Award # 2006-38411-17034.

REFERENCES

1. Perlack, R.D.; Ranney, J.W.; and Wright, L.L. Environmental Emissions and Socioeconomic Considerations in the Production, Storage, and Transportation of Biomass Energy Feedstocks. ORNL/TM-12030. Oak Ridge, TN: Oak Ridge National Laboratory, 1992.
2. Cleveland, C.J. The direct and indirect use of fossil fuels and electricity in USA agriculture, 1910-1990. *Agriculture, Ecosystems and Environment*, 55(1995): 111-121.
3. Lewis, W.J.; Lenteren, J.C.V.; Phatak, S.C.; and Tumlinson, J.H. A total system approach to sustainable pest management. *Proc. Natl. Acad. Sci. USA*, 94(1997): 12243–12248.
4. Hanegraaf, M.C.; Biewinga, E.E.; and Bul, G.V.D. Assessing the ecological and economic sustainability of energy crops. *Biomass and Bioenergy*, 15(1998): 345-355.
5. Cook, J. and Beyea, J. Bioenergy in the United States: progress and possibilities. *Biomass and Bioenergy*, 18(2000): 441-455.
6. Post, W.M. and Kwon, K.C. Soil Carbon Sequestration and Land-Use Change: Processes and Potential. *Global Change Biology*, 6(2000): 317-328.
7. Dien, B.S.; Bothast, R.J.; Nichols, N.N.; and Cotta, M.A. The U.S. corn ethanol industry: An overview of current technology and future prospects. *International Sugar Journal*, 104(2002): 204-211.
8. Azapagic, A.; Perdan, S.; and Clift, R. Sustainable Development in Practice - Case Studies for Engineers and Scientists. *John Wiley & Sons*, 2004, ISBN 0-470-85609-2.
9. Crop Profile for Field Corn in Pennsylvania, 2004.
(<http://www.ipmcenters.org/cropprofiles/docs/pacornfield.pdf>) accessed on 28 Jan, 2011.
10. Minnesota Agricultural Statistics, NASS-USDA, 2004.
(http://www.nass.usda.gov/Statistics_by_State/Minnesota/Publications/Annual_Statistical_Bulletin/agstatbk2004/p031.pdf) accessed on 3rd Feb, 2011.

11. Parrish, D.J. and Fike, J.H. The Biology and Agronomy of Switchgrass for Biofuels. *Critical Reviews in Plant Sciences*, 24(2005): 423-459.
12. West, T.O. and McBride, A.C. The contribution of agricultural lime to carbon dioxide emissions in the United States: dissolution, transport, and net emissions. *Agriculture, Ecosystems and Environment*, 108(2005): 145–154.
13. Agricultural Chemical Use Estimates, NASS-USDA, 2006
(<http://www.ncagr.gov/stats/release/ChemUse2006.pdf>) accessed on 3rd Feb, 2011.
14. Hill, J.; Nelson, E.; Tilman, D.; Polasky, S.; and Tiffany, D. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences of the United States of America*, 103(2006): 11206-11210.
15. Miller, S.A.; Landis, A.E.; and Theis, T.L. Use of Monte Carlo Analysis to Characterize Nitrogen Fluxes in Agroecosystems. *Journal of Environmental Science and Technology*, 40(2006): 2324-2332.
16. Soybean Production Practices, Wisconsin, NASS-USDA, 2006.
(http://www.nass.usda.gov/Statistics_by_State/Wisconsin/Publications/Miscellaneous/chemuse_soybeans_06.pdf) accessed on 4th Feb, 2011.
17. Energy Independence and Security Act, 2007, http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_bills&docid=f:h6enr.txt.pdf accessed on 17th Dec, 2010.
18. Adler, P.R.; Grosso, S.J.D.; and Parton, W.J. Life Cycle Assessment of Net Greenhouse Gas Flux for Bioenergy Cropping Systems. *Ecological Applications*, 17(2007): 675-691.
19. Johnson, J.M.F.; and Franzluebbers, A.J.; Weyers, S.L.; and Reicosky, D.C. Agricultural opportunities to mitigate greenhouse gas emissions. *Journal of Environmental Pollution*, 150(2007): 107–124.
20. Landis, A.E.; Miller, S.A.; and Theis, T.L. Life Cycle of the Corn-Soybean Agroecosystem for Biobased Production. *Journal of Environmental Science & Technology*, 41(2007): 1457-1464.

21. Powers, S. Nutrient loads to surface water from row crop production. *International Journal of Life Cycle Assessment*, 12(2007): 399-407.
22. Wang, M.; Wu, M.; and Huo, H. Life-cycle energy and greenhouse gas emission impacts of different corn ethanol plant types. *Environmental Research Letters*, 2(2007): 024001 (13pp)
23. Smith, D.R.; Livingston, S.J.; Zuercher, B.W.; Larose, M.; Heathman, G.C.; and Huang, C. Nutrient losses from row crop agriculture in Indiana. *Journal of Soil and Water Conservation*, 63(2008): 396-409.
24. Energy Life-Cycle Assessment of Soybean Biodiesel, USDA, (845): 2009.
25. Kim, S.; Dale, B.E.; and Jenkins, R. Life cycle assessment of corn grain and corn stover in the United States. *International Journal of Life Cycle Assessments*, 14(2009): 160-174.
26. USDA, Technical Note No. 3, 2009, Planting and Managing Switchgrass as a Biomass Energy Crop.
27. Zheng, Y; Pan, Z; and Zhang, R. Overview of Biomass Pretreatment for Cellulosic Ethanol Production. *International Journal of Agricultural & Biological Engineering*, 2(2009): 51-68.
28. Cruse, M.J.; Liebman, M.; Raman, D.R.; and Wiedenhoef, M.H. Fossil Energy Use in Conventional and Low-External-Input Cropping Systems. *Agronomy Journal*, 102(2010): 1-8.
29. Feng, H.; Rubin, O.D.; and Babcock, B.A. Greenhouse gas impacts of ethanol from Iowa corn: Life cycle assessment versus system wide approach. *Biomass and Bioenergy*, 34 (2010): 912-921.
30. Gelfand, I.; Snapp, S.S.; and Robertson, G.P. Energy Efficiency of Conventional, Organic, and Alternative Cropping Systems for Food and Fuel at a Site in the U.S. Midwest. *Journal of Environmental Science and Technology*, 44(2010): 4006-4011.
31. McIsaac, G.F.; David, M.B.; and Mitchell, C.A. Miscanthus and Switchgrass Production in Central Illinois: Impacts on Hydrology and Inorganic Nitrogen Leaching. *Journal of Environmental Quality*, 39(2010): 1790-1799.

32. McSwiney, C.P.; Bohm, S.; Grace, P.R.; and Robertson, G.P. Greenhouse Gas Emissions Calculator for Grain and Biofuel Farming Systems. *Journal of Natural Resources and Life Sciences Education*, 39(2010): 125–131.
33. Mitchell, R.; Vogel, K.; Berdahl, J.; and Masters, R. Herbicides for Establishing Switchgrass in the Central and Northern Great Plains. *BioEnergy Research*, 3(2010): 321-327(7).
34. Whitaker, J.; Ludley, K.E.; Rowe, R.; Taylor, G.; and Howard, D.C. Sources of variability in greenhouse gas and energy balances for biofuel production: a systematic review. *GCB Bioenergy*, 2(2010): 99-112.
35. Helmers, M.J. Associate Professor, Agricultural and Biosystems Engineering, Iowa State University, Personal Communication, 2011.
36. Biotechnology Varieties, Iowa and U.S., NASS-USDA (http://www.nass.usda.gov/Statistics_by_State/Iowa/Publications/Annual_Statistical_Bulletin/2010/65_10.pdf) accessed on 12th Feb, 2011.
37. National Biodiesel Board (NBB), <http://biodiesel.org> accessed on 15th March, 2011.
38. Ecoinvent database, <http://www.ecoinvent.org/> accessed on 16th Oct, 2010.
39. SimaPro (*System for Integrated Environmental Assessment of Products*), PRe Consultants, Amersfoort, The Netherlands, <http://www.pre-sustainability.com/content/sima-pro-installation-registered-users> accessed on 16th Oct, 2010.
40. University Extension, Iowa State University, <http://www.extension.org/pages/14044/corn-ethanol-production> accessed on 6th March, 2011.
41. United States Department of Agriculture (USDA), Energy Efficiency and Renewable Energy, Information Resources, Theoretical Ethanol Yield Calculator, http://www1.eere.energy.gov/biomass/ethanol_yield_calculator.html accessed on 14th Feb, 2011.

CHAPTER - 4

GENERAL CONCLUSION

Errors in important spreadsheets can affect decision-making, and can result in monetary losses too. The received wisdom is that errors are prevalent in spreadsheets of all nature. Since the computer era began, there have been numerous attempts at detection of errors in spreadsheets. While manual efforts have not been very efficient and successful, software-based inspection hasn't been very common. We tapped this need and opportunity to design an automated tool, using the built-in VBA programming capability, to detect hard-coding errors in bioenergy-relevant spreadsheets. Hard-coding error is a qualitative error, can mask other kinds of errors, and has high occurrence rates in spreadsheets, as hard-coding numbers into an Excel formula is a common human practice, or rather a malpractice!

It should be noted that though we used our tools on a set of bioenergy-relevant spreadsheets, our tools are not spreadsheet-specific, and can be used on all Microsoft Excel spreadsheets. The cell error rates for hard-coding error in the tested spreadsheets were found to range from 11-44%, consistent with the values in the related literature. While scanning through the error reports generated by our programs, we realized that hard-coding errors could be further classified into four prominent sub-types depending upon the hard-coded value in the Excel cell. To our knowledge, this is the first time hard-coding errors have been sub-categorized, and hence we have contributed four additional types of errors (within the hard-coding error) to the existing literature on spreadsheet errors. Also, the beauty of our tool lies in the report generated by it, which allows the user to see the formula of the hard-coded cells with their cell addresses, sorted by worksheet, and distributed by the four sub-types. It also lets the user know the frequencies of each sub-type as a percentage of the total hard-coding errors in the spreadsheet, and facilitates the user to quickly decide which hard-coded cells can be potential sites for errors, and accordingly fix them.

Bioenergy development certainly provides a remedial solution to the US to reduce its dependence on foreign oil supplies, and reduce greenhouse gas emissions. However, like any

other technology, it comes with its own set of doubts and disadvantages. The issue of carbon savings achieved by biofuels is still surrounded by skepticism. With the motive of increasing its annual biofuel production, the US government is providing massive funding to educational institutions and government organizations to look into renewable energy options and their relative performance. Nowadays, agricultural land is increasingly being used for bioenergy production, and there are multiple choices of feedstocks and cropping systems for the same. Corn grain and soybeans have been the forerunners in the ethanol and biodiesel market respectively. However, with the prospect of cellulosic ethanol production, corn systems with stover removal, and herbaceous perennials like *Miscanthus* and switchgrass are being studied heavily. This is where a life cycle analysis is extremely helpful to assess the biofuel production systems and the related processes of conversion of feedstocks into biofuel from multiple perspectives.

The third chapter of this thesis dealt with the life cycle analysis of six kinds of cropping systems, experimentally established in Boone County, Iowa. The systems were analyzed from cradle-to-farm gate, and their fossil energy ratio, net energy yield, global warming potential, eutrophication potential, and theoretical biofuel yield were quantified for a period of 10 years. Life cycle analyses help in a mature assessment of the sustainability of cropping systems as bioenergy systems. While the continuous corn systems had the maximum yields, they impacted the environment most severely. The corn-soybean rotation seemed to be a better alternative than continuous corn in terms of ecological performance as its fossil energy use, greenhouse gases and eutrophication potential were lower than those of the continuous corn systems. Mixed prairies, on the other hand, were not so competitive with corn in terms of biomass yields, but they exhibited a significantly low eutrophication potential compared to the corn-based systems, while consuming minimal fossil energy inputs for their production. The literature reports of several such analyses involving corn, soybeans and prairie monocultures like *Miscanthus* and switchgrass, however comparative analyses for cropping systems with multiple feedstocks for the state of Iowa have not been very common. Moreover, the potential of mixed prairies as biofuel production systems has not been assessed to our knowledge. This is where our analysis adds to the current literature on life

cycles of bioenergy cropping systems. As the cellulosic ethanol technologies become economically viable with definite infrastructure, the scope of the analysis can be widened to perform a cradle-to-pump or cradle-to-wheel analysis of biofuels from different cropping systems with multiple feedstocks, to assist a more holistic understanding of the lifecycle effects of each system.

APPENDIX A. HARD-CODING DETECTOR (HCD)

Option Explicit

Sub hcd()

Dim final_tabnames() As String

Dim round1, round2, freq, freq_final As Long, totnum, x, y, i, j, k, u, v, a, lastrow, lastcol, see, cells_checked, first, second, diff, ratio, remainder, unity_count, unity_check, error_check As Single, ident, str, str1, prechr, nxtchr, check, ratio_str, remainder_str, user_response As String

Dim vr As Boolean ' flag true if an error exists

k = 8

a = 0

cells_checked = 0

unity_count = 0

Worksheets("HCER").Range("A1:C2008").ClearContents 'Clears the content of the error report before every run

Worksheets("HCER").Range("A1:C2008").Interior.ColorIndex = 2

Worksheets("HCER").Cells(1, 1).Value = "ERROR STATISTICS :"

Worksheets("HCER").Cells(1, 1).Font.Bold = True

Worksheets("HCER").Cells(2, 1).Value = "Total number of cells checked ="

Worksheets("HCER").Cells(3, 1).Value = "Cells with HCE ="

Worksheets("HCER").Cells(4, 1).Value = "Cells with Unity errors (Unique) ="

Worksheets("HCER").Cells(5, 1).Value = "Cell Error Rate (CER) ="

Worksheets("HCER").Cells(7, 1).Value = "WORKSHEET NAME"

Worksheets("HCER").Cells(7, 1).Font.Bold = True

Worksheets("HCER").Cells(7, 2).Value = "CELL IDENTIFIER"

Worksheets("HCER").Cells(7, 2).Font.Bold = True

Worksheets("HCER").Cells(7, 3).Value = "CELL FORMULA"

Worksheets("HCER").Cells(7, 3).Font.Bold = True

ReDim final_tabnames(Worksheets.count)

totnum = Worksheets.count 'Counting the total number of spreadsheets

MsgBox ("Your workbook has a total of " & (totnum - 1) & " spreadsheets containing data.")
 'Displays the total number of spreadsheets

user_response = InputBox("Press 'Y' if you have any protected worksheets in your workbook")

For round1 = 1 To (totnum - 1)

check = InputBox("Please enter 'Y' if you want your program to run on the following
 spreadsheets: " & Worksheets(round1).Name)

If check = "Y" Or check = "y" Then

a = a + 1

final_tabnames(a) = Worksheets(round1).Name

u = 1

v = 0

While (v = 0) And (u < Len(final_tabnames(a))) ' Finding the number of
 worksheet tab names containing number

If Asc(Mid(final_tabnames(a), u, 1)) > 47 And Asc(Mid(final_tabnames(a),
 u, 1)) < 58 Then

v = 1

MsgBox ("This worksheet tab name - " & final_tabnames(a) &
 "can create a problem. Please suggest another name for the same.
 Please note that the new name should not have any numbers.")

final_tabnames(a) = InputBox("Please enter the new worksheet tab
 name here.")

Worksheets(round1).Name = final_tabnames(a) 'Assigning
 the new name entered by the user to the worksheet tab

End If

u = u + 1

Wend

End If

Next round1

For round2 = 1 To a 'Loop through all the worksheets selected by the user

Worksheets(final_tabnames(round2)).Activate

If WorksheetFunction.CountA(Cells) > 0 Then

```
lastrow = Cells.Find(What:="*", After:=[A1], SearchOrder:=xlByRows,
SearchDirection:=xlPrevious).row      'Search for any entry, by searching
backwards by Rows
```

```
lastcol = Cells.Find(What:="*", After:=[A1], SearchOrder:=xlByColumns,
SearchDirection:=xlPrevious).Column    'Search for any entry, by searching
backwards by Columns
```

```
End If
```

```
For x = 1 To lastrow
```

```
    For y = 1 To lastcol
```

```
        unity_check = 0
```

```
        error_check = 0
```

```
        see = 0
```

```
        vr = False      ' set flag to false (presume innocent)
```

```
        str = Worksheets(final_tabnames(round2)).Cells(x, y).Formula
```

```
        If Not IsEmpty(Cells(x, y)) And Mid(str, 1, 1) = "=" Then    'check if
equation
```

```
            cells_checked = cells_checked + 1
```

```
            i = 2
```

```
            While i <= Len(str)      ' loop through entire string
```

```
                str1 = Mid(str, i, 1)    ' parse string elements one by one
```

```
                If (Asc(str1) > 47) And (Asc(str1) < 58) Then    ' if it's a
number
```

```
                    prechr = Mid(str, i - 1, 1)    ' get preceding value
```

```
                    If (Asc(prechr) < 65) And (Asc(prechr) <> 36)
Then      'if the preceding value is an
alphabet
```

```
                        vr = True
```

```
                    ElseIf (Asc(prechr) > 90) And (Asc(prechr) <
97) Then
```

```
                        vr = True
```

```
                    ElseIf (Asc(prechr) > 122) Then
```

```
                        vr = True
```

Else

vr = False

End If

If vr = True Then

error_check = error_check + 1

see = 1

End If

If vr = True And i = Len(str) And Mid(str, i, 1) =
"1" And Mid(str, i - 1, 1) <> "1" Then
unity_check = unity_check + 1

If vr = True And i <> Len(str) And Mid(str, i, 1)
= "1" And Mid(str, i - 1, 1) <> "1" And Mid(str, i
+ 1, 1) <> "." Then unity_check = unity_check +
1

If (i < Len(str)) And vr = False Then

nxtchr = Mid(str, i + 1, 1) 'Get the
succeeding value

While (Asc(nxtchr) > 47) And
(Asc(nxtchr) < 58) And (i < Len(str))

i = i + 1 'Keeps on
skipping as long as numerical
values are encountered

If (i < Len(str)) Then nxtchr =
Mid(str, i + 1, 1)

Wend

End If

End If

i = i + 1

Wend

If see = 1 Then

If user_response <> "Y" Then

Worksheets(final_tabnames(round2)).Cells(x,
y).Interior.ColorIndex = 15 'Highlights the
faulty cells with a grey fill

Worksheets(final_tabnames(round2)).Cells(x,
y).Font.Bold = True 'Turns the font
in the faulty cells to bold

End If

k = k + 1

Worksheets("HCER").Cells(k, 1).Value =
final_tabnames(round2) 'Displays a report of the
hard-coding errors in all the spreadsheets

If y <= 26 Then

j = y + 64 'Conversion of column
number back to column alphabets

ident = Chr(j)

End If

If y > 26 Then 'for double-alphabet
column address

ratio = Int(y / 26)

ratio = ratio + 64

ratio_str = Chr(ratio)

remainder = y Mod 26

remainder = remainder + 64

remainder_str = Chr(remainder)

ident = ratio_str & remainder_str

End If

Worksheets("HCER").Cells(k, 2).Value = ident & x
'Displays the address of the faulty cell

Worksheets("HCER").Cells(k, 3).Value = "" & str
'Displays the formula of the faulty cells

If error_check = unity_check Then

```

Worksheets("HCER").Cells(k,
1).Interior.ColorIndex = 15

Worksheets("HCER").Cells(k,
2).Interior.ColorIndex = 15

Worksheets("HCER").Cells(k,
3).Interior.ColorIndex = 15

unity_count = unity_count + 1

End If

End If

End If

Next y

Next x

Next round2

freq = ((k - 8) / cells_checked) * 100      'Calculates the frequency of hard coding errors
freq_final = Round(freq, [2])

Worksheets("HCER").Cells(2, 2).Value = cells_checked
Worksheets("HCER").Cells(3, 2).Value = k - 8
Worksheets("HCER").Cells(4, 2).Value = unity_count
Worksheets("HCER").Cells(5, 2).Value = freq_final & " %"

Worksheets("HCER").Cells(2, 2).Font.Bold = True
Worksheets("HCER").Cells(3, 2).Font.Bold = True
Worksheets("HCER").Cells(4, 2).Font.Bold = True
Worksheets("HCER").Cells(5, 2).Font.Bold = True

MsgBox ("The program has finished checking your cells." & vbNewLine & vbNewLine & "1. A total of " &
cells_checked & " cells were checked." & vbNewLine & "2. Out of these, hard-coding errors were detected in "
& k - 8 & " cells." & vbNewLine & vbNewLine & "This accounts for a Cell Error Rate (CER) of " & freq_final
& " %.")

End Sub

```

APPENDIX B. SUBCATEGORIZATION CODE (SubCat)

Option Explicit

Sub subcat()

Dim i, count, lastrow, row, pot, track, unitconv, unity, others, str1, temp_num, tot, sum_unity, unity_freq, sum_unitconv, unitconv_freq, sum_pot, pot_freq, sum_others, others_freq, temp1 As Single, str, temp, prechr, nxtchr As String

sum_pot = 0

sum_unitconv = 0

sum_others = 0

sum_unity = 0

Worksheets("HCER").Activate

Worksheets("HCER").Range("E1:H65536").Interior.ColorIndex = 2

Worksheets("HCER").Range("E1:H65536").ClearContents

Worksheets("HCER").Cells(1, 5).Value = "SUBCATEGORISATION OF HCE :"

Worksheets("HCER").Cells(1, 5).Font.Bold = True

Worksheets("HCER").Cells(2, 5).Value = "Frequency of Unity Errors ="

Worksheets("HCER").Cells(3, 5).Value = "Frequency of Power_of_10 Conversions ="

Worksheets("HCER").Cells(4, 5).Value = "Frequency of Unit Conversion Factors="

Worksheets("HCER").Cells(5, 5).Value = "Frequency of Other Unidentified Numerals ="

Worksheets("HCER").Cells(7, 5).Value = "# OF UNITY OCCURENCES"

Worksheets("HCER").Cells(7, 5).Font.Bold = True

Worksheets("HCER").Cells(7, 6).Value = "# OF POWER_OF_10 CONVERSIONS"

Worksheets("HCER").Cells(7, 6).Font.Bold = True

Worksheets("HCER").Cells(7, 7).Value = "# OF UNIT CONVERSION FACTORS"

Worksheets("HCER").Cells(7, 7).Font.Bold = True

Worksheets("HCER").Cells(7, 8).Value = "# OF OTHER UNIDENTIFIED NUMERALS"

Worksheets("HCER").Cells(7, 8).Font.Bold = True

lastrow = Worksheets("HCER").Range("C" & Rows.count).End(xlUp).row

For row = 9 To lastrow

str = Range("C" & row).Formula

pot = 0

unity = 0

unitconv = 0

others = 0

i = 2

While i <= Len(str) 'loop through entire string

count = 0

str1 = Mid(str, i, 1) 'parse string elements one by one

If (Asc(str1) > 47) And (Asc(str1) < 58) Then

prechr = Mid(str, i - 1, 1) 'get preceding value

If (Asc(prechr) < 65) And (Asc(prechr) <> 36) Then 'if the preceding value is an alphabet

count = 1

ElseIf (Asc(prechr) > 90) And (Asc(prechr) < 97) Then

count = 1

ElseIf (Asc(prechr) > 122) Then

count = 1

Else

count = 2

End If

If (i <= Len(str)) And count = 2 Then

If i < Len(str) Then nxtchr = Mid(str, i + 1, 1) 'Get the succeeding value

While (Asc(nxtchr) > 47) And (Asc(nxtchr) < 58) And (i < Len(str))

i = i + 1 'Keeps on skipping as long as numerical values are encountered.

If (i < Len(str)) Then nxtchr = Mid(str, i + 1, 1)

```

Wend

End If

If (i <= Len(str)) And count = 1 Then
    temp = Mid(str, i, 1)
    track = 0
    If i < Len(str) Then nxtchr = Mid(str, i + 1, 1)    'Get the succeeding value
        While (i < Len(str) And Asc(nxtchr) > 47 And Asc(nxtchr) < 58)
            Or (i < Len(str) And Asc(nxtchr) = 46)
                temp = temp + Mid(str, i + 1, 1)
                i = i + 1    'Keeps on skipping as long as numerical
                            values are encountered
                If (i < Len(str)) Then nxtchr = Mid(str, i + 1, 1)
            Wend
        temp_num = Val(temp)
        If temp <> 0 Then temp1 =
            Application.WorksheetFunction.Log(temp)
        If temp = "1" Then
            unity = unity + 1
            track = 1
        ElseIf temp <> 0 And temp1 = Int(temp1) Then
            pot = pot + 1
            track = 1
        ElseIf temp = "2.54" Or temp = "0.3937" Or temp = "2.303" Then
            unitconv = unitconv + 1
            track = 1
        ElseIf temp = "3.785" Or temp = "0.2642" Then
            unitconv = unitconv + 1
            track = 1
        ElseIf temp = "2.202" Or temp = "0.454" Then

```

```
unitconv = unitconv + 1
```

```
track = 1
```

```
ElseIf temp = "0.239" Or temp = "4.184" Then
```

```
unitconv = unitconv + 1
```

```
track = 1
```

```
ElseIf temp = "0.4047" Or temp = "2.47" Then
```

```
unitconv = unitconv + 1
```

```
track = 1
```

```
ElseIf temp = "0.6214" Or temp = "1.6" Then
```

```
unitconv = unitconv + 1
```

```
track = 1
```

```
ElseIf temp = "3600" Or temp = "24" Or temp = "60" Or temp =  
"365.25" Or temp = "365" Then
```

```
unitconv = unitconv + 1
```

```
track = 1
```

```
End If
```

```
If track = 0 Then others = others + 1
```

```
End If
```

```
End If
```

```
i = i + 1
```

```
Wend
```

```
Worksheets("HCER").Cells(row, 5) = unity
```

```
Worksheets("HCER").Cells(row, 6) = pot
```

```
Worksheets("HCER").Cells(row, 7) = unitconv
```

```
Worksheets("HCER").Cells(row, 8) = others
```

```
Next row
```

```
For i = 9 To lastrow
```

```
sum_unity = sum_unity + Cells(i, 5).Value
```

```
sum_pot = sum_pot + Cells(i, 6).Value  
sum_unitconv = sum_unitconv + Cells(i, 7).Value  
sum_others = sum_others + Cells(i, 8).Value
```

```
Next i
```

```
tot = sum_unity + sum_pot + sum_unitconv + sum_others  
unity_freq = (sum_unity / tot) * 100  
unity_freq = Round(unity_freq, [2])  
pot_freq = (sum_pot / tot) * 100  
pot_freq = Round(pot_freq, [2])  
unitconv_freq = (sum_unitconv / tot) * 100  
unitconv_freq = Round(unitconv_freq, [2])  
others_freq = (sum_others / tot) * 100  
others_freq = Round(others_freq, [2])
```

```
Worksheets("HCER").Cells(2, 6).Value = unity_freq & "%"  
Worksheets("HCER").Cells(2, 6).Font.Bold = True  
Worksheets("HCER").Cells(3, 6).Value = pot_freq & "%"  
Worksheets("HCER").Cells(3, 6).Font.Bold = True  
Worksheets("HCER").Cells(4, 6).Value = unitconv_freq & "%"  
Worksheets("HCER").Cells(4, 6).Font.Bold = True  
Worksheets("HCER").Cells(5, 6).Value = others_freq & "%"  
Worksheets("HCER").Cells(5, 6).Font.Bold = True
```

```
End Sub
```

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to everyone who helped me through my graduate studies at Iowa State University.

First and the foremost, I want to thank my advisor, Dr. D. Raj Raman, for his incredible support and guidance right from the beginning through to the end. I sincerely appreciate his mentoring skills, well-rounded personality, patience, self-esteem, and innate ability to deal with tough situations. I would like to thank Dr. Robert P. Anex for the knowledge, ideas, critiques, and extensive assistance, especially through the second project. I would like to thank Dr. Matthew Z. Liebman for expert guidance at some critical junctures through the second project. I would also like to thank Dr. Matthew J. Helmers for providing data relevant to the second project.

I wish to express appreciation to the faculty and the staff in the Department of Agricultural & Biosystems Engineering for the assistance provided throughout my studies.

I would also like to thank Sami Khanal and Carol Faulhaber for the help and guidance provided through the first project. I would like to thank Dave Sundberg and Meghann Jarchow for their extensive assistance through my second project. I would like to express my deepest gratitude to my wonderful support systems of friends.

Finally, I wish to thank my parents, Urmila and S.S. Rawat, for their unconditional love and support, and for being my everyday inspiration.