IOWA STATE UNIVERSITY
Digital Repository

2014

Impact of sampling interval on GIS data based productivity calculations and optimal configuration and lead time of corn stover biomass harvesting teams

Levi Powell
Iowa State University

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

Part of the Agriculture Commons, and the Bioresource and Agricultural Engineering Commons

**Impact of sampling interval on GIS data based productivity calculations
and
optimal configuration and lead time of corn stover biomass harvesting teams**

by

Levi John Powell

A thesis submitted to the graduate faculty

In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Industrial and Agricultural Technology

Program of Study Committee:
Matthew Darr, Major Professor
Brian Steward
Thomas Brumm

Iowa State University
Ames, Iowa
2014

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

CHAPTER 1:

GENERAL INTRODUCTION AND REVIEW OF LITERATURE

Introduction

Over the last decade, agricultural equipment has seen an exponential increase in the amount of technology and electronic systems being utilized in their functionality. This increase has been a natural progression in the quest to make agricultural equipment safer, more productive, and more efficient.

An added benefit to these new technologies is the potential to easily glean vast amounts of data about all the machine parameters and their utilization. This potential has gone seemingly untapped until recently. The onset of larger corporate farms and more industrialized agricultural production systems has created a need for managing large fleets of equipment. This need has driven the research and development of automated data logging and processing systems for agricultural equipment.

As equipment fleets grow larger and agricultural operations become more industrialized, there has been a push towards more standardization. This standardization processes involves evaluating the current performance of each machine based on its productivity and efficiency. Once the specific performance of each machine is determined, an average performance bench mark can be set for the entire fleet. This process highlights which machines are performing above the average, and which machines are performing below it. This drives further investigation as to why some machines are below the average, whether it is

machine problems, environmental factors or operator issues. These same questions can also be asked of the machines performing above the average. This performance evaluation process informs managers and personnel where they should focus their efforts and resources in order to learn from their machines that exhibit above average performance, and help to improve those machines with low performance.

Newly developed data logging and processing systems enable this exact process. They collect and process data on machine performance and report that data back to the operator, as well as the manager. This allows the operators and managers to make informed, data-driven decisions about machine operations and logistics. These systems can be tailored to any agricultural operation and are the next step to increasing the productivity and efficiency of agricultural equipment.

The potential of these systems has been realized by several companies in the developing cellulosic ethanol industry. As these companies grow their supply chains they have discovered the value and cost savings that a system like this can provide. Managing equipment on an industrial scale such as this has helped to drive further development of these systems as well as tailoring them to specific industries such as cellulosic ethanol. The flexibility of these data collection and processing systems will reshape the way large fleet management is conducted and maximize the productivity potential of the fleet as well as each machine within it.

Objectives

Objective 1: Determine effects of GIS data sampling interval on machine productivity calculations.

Currently, the CyCAN system records GIS data samples every second. The goal is to find out how decreasing the sampling rate will affect the machine productivity calculations. If data collection intervals can be increased with minimal effects to the productivity numbers, it will be advantageous to do so. This will involve performing the same productivity calculations on the same data set using alternative sampling intervals.

Objective 2: Determine the optimal configuration and lead time for a corn stover biomass harvesting team.

It was observed during the 2011 harvest that windrowers and balers would travel as teams from field to field. Since the baling operation is dependent on the windrowing operation, this created a waiting period for the balers while the windrowers were getting started. It also created a waiting period for the windrowers once they were finished; they would wait for the balers to finish before moving to the next field. The first task of this objective was to investigate the effect this wait time had on the productivity of both the windrowers and the balers. The second task was to investigate what performance gains could be made by splitting the windrowers and balers into separate teams, and what the optimal lead time should be between the two operations.

Literature Review: Current Knowledge

The Renewable Fuels Standard (RFS) was created in 2005 under the Energy Policy Act. This was the first renewable fuel volume mandate in the United States. The original RFS called for 7.5 billion gallons of renewable fuel to be blended in the U.S. fuel supply by 2012 (Renewable Fuel Standard 2013). This was the first step in the development of the cellulosic ethanol industry in the United States. In 2007 the volume mandate was revised and by 2022, 36 billion gallons of renewable fuel is to be blended into the nation's fuel supply (Renewable Fuel Standard 2013).

As cellulosic ethanol begins production in the United States, there will be an increasing demand for high quality biomass feedstock. This feedstock can come from dedicated energy crops, as well as agricultural residues. In the Midwest states, corn stover has been targeted has the primary source of cellulosic biomass. Corn stover is in abundance in the Midwest and has no impact on the nation's food supply. Perlack (2005) states "The single largest source of this current potential is corn residues or corn stover, totaling close to 75 million dry tons per year." (Perlack 2005)

Iowa is the nation's leader in corn production, harvesting corn on 13.1 million acres in 2013 (Crop Production 2013). This vast amount of corn production means Iowa holds a great potential for corn stover biomass production as well. Several companies have realized this and are working to tap into this potential. Poet-DSM Advanced Biofuels and DuPont Cellulosic Ethanol are both constructing cellulosic ethanol plants in central Iowa.

Both of these plants will use corn stover as their biomass feedstock. Poet-DSM is scheduled to open their 25 million gallon per year plant in Emmetsburg, Iowa in early 2014. The Poet-DSM

plant will require 280,000 dry tons of corn stover annually and will be collected within a 30 mile radius of the plant (Biomass Resources 2014).

The DuPont Cellulosic Ethanol Facility near Nevada Iowa is also currently under construction and scheduled to open in late 2014. The DuPont Facility will be a 30 million gallon per year plant requiring 375,000 tons of corn stover annually. This plant also has 30 mile radius collection area with 815,000 available corn acres. However with a sustainable harvest rate of 2 ton/acre only 190,000 acres will be required to meet the feedstock supply demands of the facility (Nevada Cite CE Facility 2013).

A third company has also realized the biomass supply potential of the Midwest, Abengoa Bioenergy Biomass of Kansas (ABBK) is also nearing completion of its 25 million gallon per year facility and is scheduled to begin production in late 2014. The Facility is located in Hugoton Kansas and will also use corn stover biomass as its primary feedstock. "The ABBK plant will utilize 350,000 tons of biomass to run this facility - approximately 15% of the available biomass within a 50-mile radius of the Hugoton plant." (Project Sustainability 2011)

The task of supplying a 25-30 million gallon per year cellulosic ethanol plant is challenging and multifaceted. The corn stover biomass supply for a plant of these capacities would require over 700,000 large square bales annually. With an average fall harvest window of 30 days in central Iowa, it takes a large fleet of equipment and a huge work force to accomplish this. An operation of this scale presents many logistical and organizational challenges. Understanding the logistical challenges of harvesting and transporting corn stover leads to solutions to improve efficiencies and drive down the overall cost, making corn stover a viable biomass feedstock (Nevada Cite CE Facility 2013).

Figure 1: 2012 Central Iowa Corn Stover Harvest

To understand the logistical challenges with harvesting corn stover, we have to examine the harvest equipment as a whole team, and how all the machines interact together. There are many ASABE standards already in place to calculate the productivities of individual machines in a variety of conditions (Hunt 2008). However, there is very little research that looks at how these productivities are affected when you place these machines in a team environment, specifically a biomass corn stover harvesting team. In this team environment, the productivity of one machine is usually impacted by, if not dependent on, the productivity of another.

This situation adds much more complexity when it comes to making purchasing decisions for equipment. Not only does the capacity and productivity of the single machine have to be considered, but how it interacts and affects the other machines in the team. This study looks at several of these machine relations and organizational issues associated with corn stover biomass harvesting equipment.

Iowa State University has been on the forefront of the rapidly expanding cellulosic industry and conducted research on all aspects of the supply chain including the harvest, storage, and transportation of corn stover biomass. Much of that work has been centered around collecting and analyzing GIS data from each stage of the supply chain. Work by Webster (2011) and Peyton (2012) showed the potential that GIS data had and the machine metrics information that could be gleaned from it.

Later work by Covington (2013) and Askey (2013) established and automated the sorting and processing of machine data into specific machine categories. These categories break down machine utilization and provide data back to operators and managers about that machines productivity. This type of feedback allows them to make informed data driven decisions about machine logistics. The work in this thesis explores more in depth the organizational issues that contribute to poor productivity of machines and entire harvesting teams.

Knowledge Gaps

There were several potential areas of study derived from this literature review. The first was the productivity of machines working within in a team and how that varies from the ASABE standard productivity of machines. While this is not a new problem, the impact of it has a much larger effect when conducting large commercial scale biomass harvests. The second was exploring which factors have the largest impact on a biomass harvesting team's productivity. Machine related issues specific to biomass harvesting have been explored but there is an organizational aspect that requires further research. Along with this is the question of "best practice" harvest team organization. What is the best configuration for a corn stover biomass

harvesting team? These three knowledge gaps are the focus of the research work presented in this thesis.

References

Askey, Jeffery C., Matthew Darr, Keith Webster, Benjamin Covington, and Jeremy Brue.

Automated Logistics Processing of GIS Data for Agricultural Harvest Equipment. Tech. no. 1596410. Print. 2013 ASABE Annual International Meeting

"Biomass Resources." *Project Liberty*. POET-DSM, 2014. Web. 09 Mar. 2014.

Covington, Benjamin R. "Assessment of Utilization and Downtime of a Commercial Level Multi-pass Corn Stover Harvesting System." Thesis. Iowa State Univeristy, 2013. Print.

Crop Production. Rep. no. 1936-3737. USDA, 8 Nov. 2013. Web. 1 Apr. 2014.

Darr, M.J. 2012 Sept. "CAN bus technology enables advanced machinery management." Resource Magazine. 19(5): 10-11-2012 St. Joseph,Michigan

"Nevada Site CE Facility." *DuPont BioFuel Solutions*. DuPont Industrial Biosciences, 2013. Web. 09 Mar. 2014.

Hunt, Donnell. *Farm Power and Machinery Management*. 10th ed. Long Grove, IL: Waveland, 2008. Print

Perlack, Robert D., Lynn L. Wright, Anthony F. Turhollow, and Robin L. Graham. Biomass as Feedstock for a Bioenergy and Bioproducts Indsutry: The Technical Feasibility of a Billion-Ton Annual Supply. 2005. Rep. Print. Joint Study by U.S. Department of Energy and U.S. Department of Agriculture

Peyton, Kevin S. "Geographic Information System Tools for the Analysis of Commercial Level Multi-pass Corn Stover Harvesting Systems." Thesis. Ames, Iowa: Iowa State University, 2012. Print.

"Project Sustainability." *Abengoa Bioenergy Biomass of Kansas*. Abengoa Bioenergy, 2011. Web.

    09 Mar. 2014.

"Renewable Fuel Standard (RFS)." *EPA*. Environmental Protection Agency, 2013. Web. 3 Mar.

    2014.

SMS Advanced. Vers. 12.0. Ames, IA: AgLeader, 2012. Computer software.

Sokhansanj, Shahab, Anthony Turhollow, Janet Cushman, and John Cundiff. "Engineering

    Aspects of Collecting Corn Stover for Bioenergy." <u>*Biomass and Bioenergy*</u> 23.5 (2002):

    347-55. Print.

Webster, Keith E. "Single-pass Corn Stover Harvest System Productivity and Cost Analysis."

    Thesis. Ames, Iowa: Iowa State University, 2011. Print.

Experimental Dataset & Collection

The challenge with a large harvesting operation is to determine the performance of the entire fleet of equipment, which is spread across a large geographical area. It is not feasible to go and observe the operation of each machine. This is where automated data logging and data processing systems come into play. These systems can be deployed on a piece of equipment and used to record data from the electronic controls on the machine. This data can then be analyzed and used to determine the productivity of the machine.

The network of electronic controllers on agricultural machines offers a gold mine of data that can be captured and recorded for use in evaluating the machine's performance. These electronic controllers communicate through a Controller Area Network or CAN bus. The CAN bus carries messages that control all the functions of the machine. These messages can be captured and recorded by the data logging system. This data, combined with the GPS position of the machine, provides a powerful GIS data set that can be used to understand where the machine was located and what operations it was performing.

For this research, data was collected using the CyCAN system developed at Iowa State University. The CyCAN system is an embedded hardware system that can be deployed on any agricultural machine with a CAN bus. It can be configured to fit any application; for this work, it was setup to log data from tractors used in a commercial scale corn stover biomass harvest. The CyCAN system records data from the tractor's CAN bus along with the GPS signal from the tractor's receiver. Currently, it collects this data at one sample per second (Webster 2011).
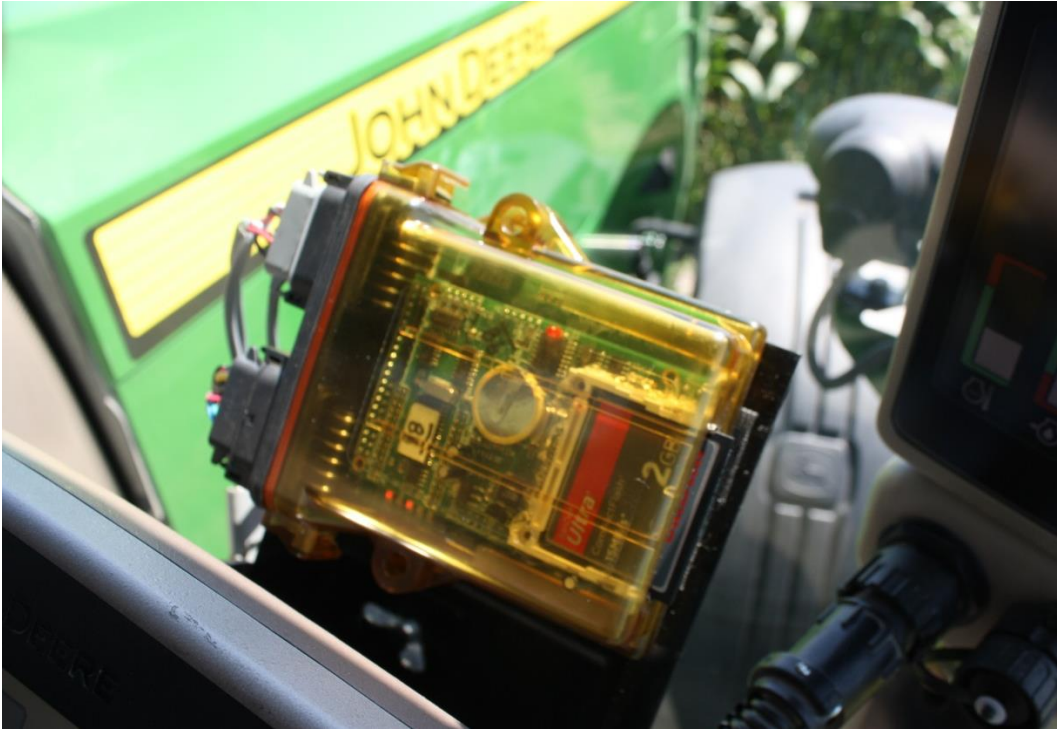
Figure 2: CyCAN Data logger used to record data from vehicles CAN bus

The CAN data and GPS coordinates are then extracted from the raw data recorded by the logger. This data is then separated and organized by field through a process called geofencing. Geofencing uses a geographical boundary for each field to determine which data points are from that field. Any data point with a GPS coordinate that falls inside the boundary for that field gets categorized under that field. This process spatially breaks up the data and allows for productivity analysis on a per-field basis.

The data for this research was collected from a commercial scale biomass corn stover harvest in central Iowa. Data was collected for a two year period and included the fall harvest season of 2011 and 2012. Both harvests occurred in the same geographical area of central Iowa. Both harvests included tractors from the three major North American manufacturers: John Deere, Case IH, and AGCO. All balers were either AGCO or Krone balers, and shredding

equipment was from the Hiniker Company of Mankato Minnesota. This research focuses on

crew organizational factors and their impact on machine productivity. For reasons of

confidentiality the specific equipment manufacturers are not disclosed within the work of this

thesis, however when applicable different manufactures are noted in generic terms.

CHAPTER 2:

IMPACT OF SAMPLING INTERVAL ON GIS DATA BASED

PRODUCTIVITY CALCULATIONS

Abstract

Analyzing GIS data collected from an agricultural machine is an invaluable tool in machine development, as well as fleet management. This data provides knowledge about machine performance characteristics that helps producers and managers make informed decisions to improve machine productivity. The frequency at which this data is collected, called the sampling interval, can have a significant impact on the machine productivity calculations. This study looks specifically at this impact and how increasing the sampling interval influences those productivity calculations. If the sampling interval can be increased without significantly changing the outcome of the productivity calculations, it would be advantageous to do so because it would decrease the amount of data generated and shorten data processing times. Currently data is collected at one hertz or once per second. Alternative sampling intervals of 2, 5, 10, 15, 30, and 60 seconds were tested in this study. 15 seconds was found to be the longest interval that did not significantly impact the results of the productivity calculations. The 15 second interval provides a 93% decrease in data which will significantly reduce data storage requirements and data processing time. As new telematics data collection systems enter the market place, reducing the amount of data means more reliable remote data transfer from these devices.

Introduction

With the increase in technology on agricultural equipment comes the opportunity to collect vast amounts of data with relative ease. The machines CAN bus communication system provides information about the entire machine and by looking at the right combination of functions it is possible to determine what operations the machine is performing. When this data is combined with a GPS location, it creates a GIS data set loaded with information about the machines, location, performance, and productivity. Providing this information to producers and managers allows them to make informed data driven decisions that improve machine productivity.

These systems also provide great data sets for researchers to glean very specific information about machine performance or other parameters that they may be testing. This sort of research based data set requires a high sampling rate in order to achieve the necessary resolution in the data so that accurate conclusions can be made from it. The current sampling interval used for this type of data collection is one hertz.

This was originally chosen because it was easy to calculate performance metrics on a time basis, but was also a high enough frequency that in-depth research analysis could be performed. As the productivity calculations have evolved, the data set used to make them has decreased; there are only a few key parameters used in these calculations now, mainly engine speed, PTO speed and vehicle speed. Also, the focus of these data collection systems has shifted away from research and more towards machine monitoring. The system now directly calculates productivity and other specific machine factors; there is no longer a need to log everything on the CAN bus, only the messages containing the specific machine parameters

required for the productivity calculations are logged. This has greatly decreased the data log file size but there is still need to further decrease it.

Recently, telematics has become the focus of large fleet management research. These systems allow data to be collected on a machine and then wirelessly transferred back to a central server location, where the data is stored, organized, and processed. This allows for real-time analysis of machine productivity throughout the day, but more importantly it gives managers the information they need to make informed data driven decisions about machine logistics.

The large data log file sizes make it difficult to transmit over the wireless networks that these systems utilize. By increasing the sampling interval, it would greatly reduce the amount of data that is generated and therefore decrease the amount of data that needs to be transferred. This increases the speed and reliability of the data transfer process and decreases the risk of lost or corrupted data. This study focuses on increasing the sampling interval time as a solution to decreasing the data log file size.

Research Objective

The objective of this research is to determine the effects of GIS data sampling interval on machine productivity calculations. Currently the CyCAN system records GIS data samples every second. If data collection intervals can be increased with minimal effects to the productivity metrics, then it will be advantageous to do so. This will involve performing the same productivity calculations on the same data set using alternative sampling intervals.

Materials

Machine Productivity Calculations

To analyze a machine's productivity accurately, a fundamental understanding of the machine and its intended functionality has to first be established. From there, its functionality can be broken down into specific parameters to determine which ones are critical to the task that it is performing. For instance, for a large square baler to operate, it has to have PTO power input from the tractor, as well as be towed by the tractor through the field. From this, it is determined that in order for the baler to be operational, the tractor has to have the PTO engaged and be moving. These are both parameters which can be recorded off of the tractor's CAN bus. Recent work by Benjamin Covington at Iowa State University has taken this one step further and assigned threshold values for these parameters, in order to determine which state of operation the machine is in (Covington 2013).

Covington defines three machine states; Production, Idle, and Transport. These states were used to evaluate the productivity of the machine, which in this case was a large square baler that is harvesting corn stover. To determine which state the baler is in, three key machine parameters are used, PTO speed, engine speed, and vehicle ground speed. Based on his research, Covington developed thresholds for each of these parameters, and the combination of all of these parameters and their thresholds defines which state the machine is in. These thresholds and cutoff points are represented in the flow chart in Figure 3 (Covington 2013).

UP 2011 model
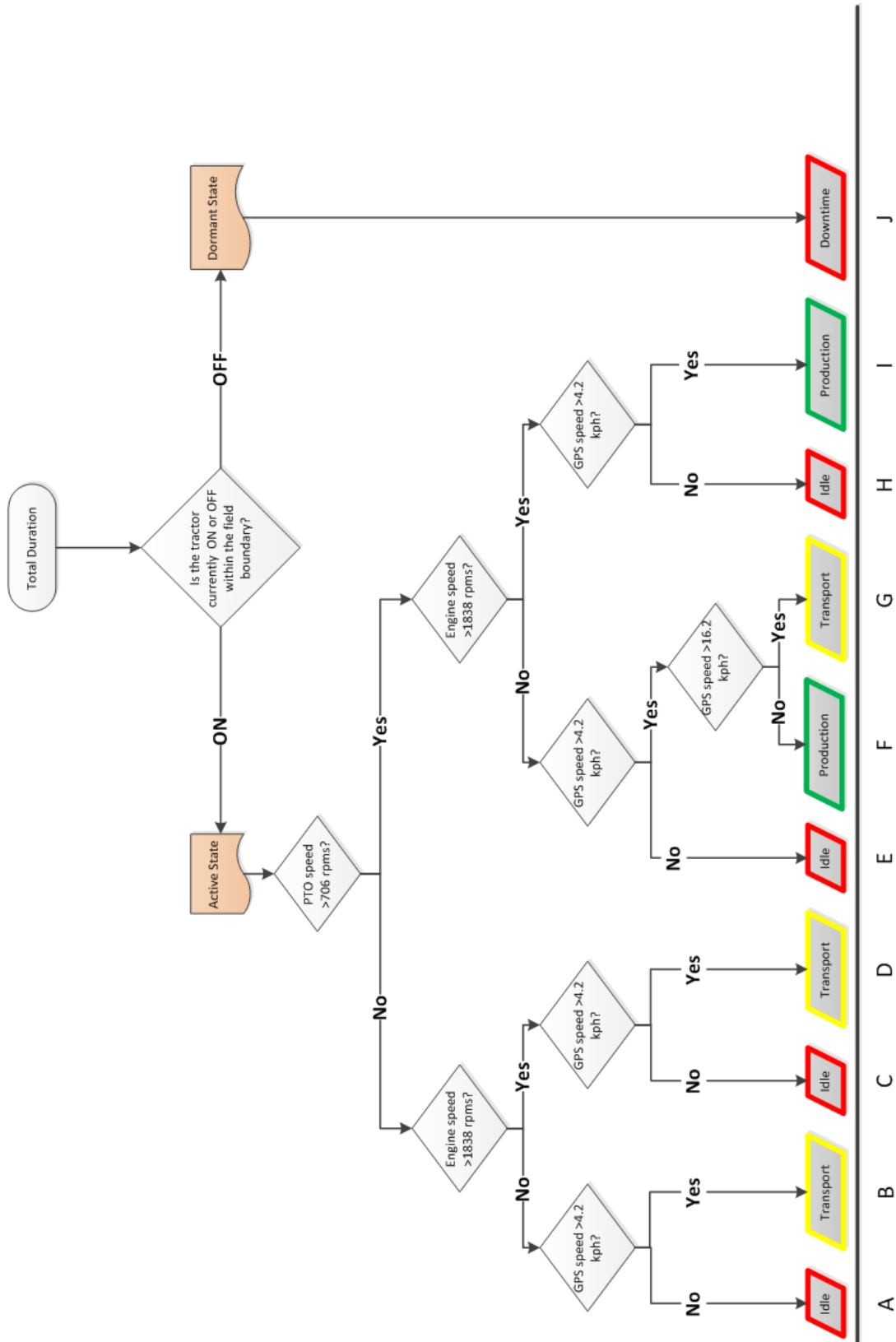
Figure 3: Utilization Paradigm, Covington 2013

In this flow chart, each of the three key machine parameters are described and their thresholds are listed. It also illustrates the decisions that are made at those threshold points to determine the current machine state. Data is recorded from the tractor and baler and then processed through this matrix. Each data point is then filtered down to the correct "bucket" or machine state that it belongs in. This decision matrix was developed for a data sampling rate of one second, which allows the data points to very easily be converted back to a time basis. By totaling up the data points from each "bucket", a total number of data points for each machine state can be determined. Since each data point represents one second, it is easy to calculate the amount of time the baler spent in Active, Idle, or Transport. The productivity of the baler is then calculated by finding the ratio of each machine state to the total number of seconds from all the machine states, this is shown here in Equation 1.

$$\% \text{ Productivity} = \frac{Production}{(Production + Idle + Transport)} \times 100$$

Equation 1: Productivity Ratio Equation

An automated data processing program was developed to rapidly quantify productivity for multiple machines. The program sorts the extracted CAN data from the CyCAN system into the appropriate "bucket" based off the threshold limits. It then sums up the buckets for each machine state to get the total for each state. The program then calculates the productivity for each machine, as well as extracts other parameters from the raw data, such as number of bales or number of acres. The output file can be customized depending on what the intended use is, but the most basic output from the program is a summary table of all the desired parameters and productivities sorted by machine. This can then be used to compare machines and harvest teams to start benchmarking the performance of each. This is a valuable tool to help increase

the productivity of the entire harvesting operation. (Askey 2013)

Data Analysis

The focus of this research study was to evaluate the possibility of increasing the

sampling interval on CAN bus data collection systems in order to decrease the log file size. To

do this, a set of alternative sampling intervals was developed and evaluated to find the longest

possible sampling interval that does not have a significant effect on the machine productivity

calculations.

For this analysis only data from the large square balers was used. The automated data

processing script was setup specifically for balers at this point, so no shredder data was used.

Data from both the 2011 and 2012 harvest seasons was used in this analysis. For the first

portion of this chapter, only 2011 data is discussed. The 2012 data is presented later in the

chapter. There were 7,500 acres in the 2011 harvest; out of this data set, 16 fields were chosen

to be used for this analysis. These 16 fields had complete data log files with no missing or

corrupted data. The fields ranged in size from 70 to 120 acres. There were 82 fields totaling

9,300 acres from the 2012 data set that were able to be used for this analysis; the others were

excluded due to missing or corrupted data.

Alternative Sampling Intervals

The current sampling interval for the CyCAN system is set at one hertz. After evaluating

the GIS data it was decided that alternative sampling intervals needed to be less than 30

seconds, as this was believed to be the maximum within the acceptable range. A set of

sampling intervals was chosen that covered the range of acceptable values, as well as outside of

the range, to test that the cutoff point of 30 seconds was justified. The alternative intervals

selected for testing where as follows; 2, 5, 10, 15, 30, and 60 seconds. The potential impact on

data log file size of these sampling intervals is shown below in Table 1. Column two is the

percent decrease in data log file size.

Table 1: Percent Decrease in Data Log File Size

Sampling Interval	Percent Decrease
1	0%
2	50%
5	80%
10	90%
15	93%
30	97%
60	98%

To apply these alternative sampling intervals, an additional script was written and added

to the data processing software. The raw data log was collected with a one second sampling

interval; it was then down-sampled based on the desired sampling interval to be tested. For

example, if the desired sampling interval was two seconds, the script would select every other

data point and discard the rest. Using that same raw data log, this process was repeated for all

the alternative sampling intervals. Figure 4 illustrates the down-sampled data sets from an

example field used in this analysis. For each alternative sampling interval, the number of data

points is shown for that specific machine state. The smaller down-sampled files for each

interval were then run through the same automated data processing software to calculate the

machine productivities.
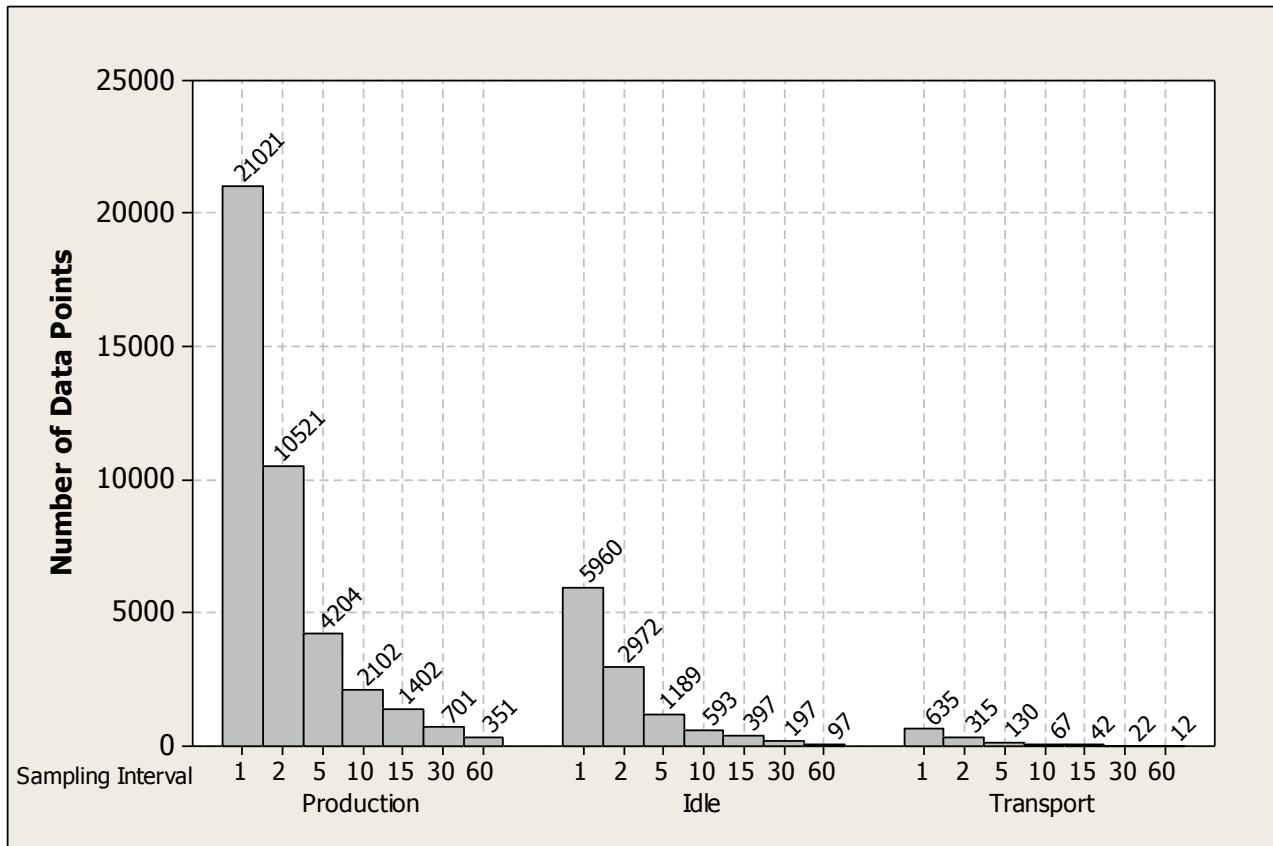
Figure 4: Example; Down Sampled Field Data Set for Alternative Sampling Intervals

Productivity Comparisons, Ratio Basis

To evaluate the effect that the alternative sampling intervals had on the productivity

calculations two primary methods were used. The first was a productivity ratio comparison,

where machine productivity was calculated for each sampling interval using Equation 1. This is

simply a ratio of each machine state time to the total amount of time, to give a percent

productive, idle, or transport. Since this method of analysis is based on the ratio of a machine

state to the total number of data points, it is still able to be used the same way on the down

sampled data sets. The percent productivity can be calculated in the same way for each

sampling interval just as it is for the raw data set.

To evaluate the results of each sampling interval, the percent productivity calculated for each sampling interval was compared to that of the raw one second interval. The difference in percentage points was calculated between each sampling interval and the raw data file. This percentage point difference is the amount of error for that sampling interval. This is illustrated in Table 2 which is the results from the example field F092. The table shows that as sampling interval is increased, the calculated productivities start to diverge away from the actual. The productivity calculation for F092 had a max percentage point difference of 0.2% while the idle calculation had a max percentage point difference of 0.5%. Both of these larger differences occurred at the 60 second sampling interval.

Table 2: Productivity Ratio Results for Example Field F092

Sampling Interval	Productivity	Idle	Transport	Productivity Difference	Idle Difference	Transport Difference
1	76.1%	21.6%	2.3%	0.0%	0.0%	0.0%
2	76.2%	21.5%	2.3%	0.1%	-0.1%	0.0%
5	76.1%	21.5%	2.4%	0.0%	-0.1%	0.1%
10	76.1%	21.5%	2.4%	0.0%	-0.1%	0.1%
15	76.2%	21.6%	2.3%	0.1%	0.0%	0.0%
30	76.2%	21.4%	2.4%	0.1%	-0.2%	0.1%
60	76.3%	21.1%	2.6%	0.2%	-0.5%	0.3%

Overall in this example field, the divergence was fairly small but in some field data sets there was a much larger impact from increasing the sampling interval. This process was repeated for all 16 fields in the 2011 data set and the results summarized by sampling interval. This is shown below in Figure 5.
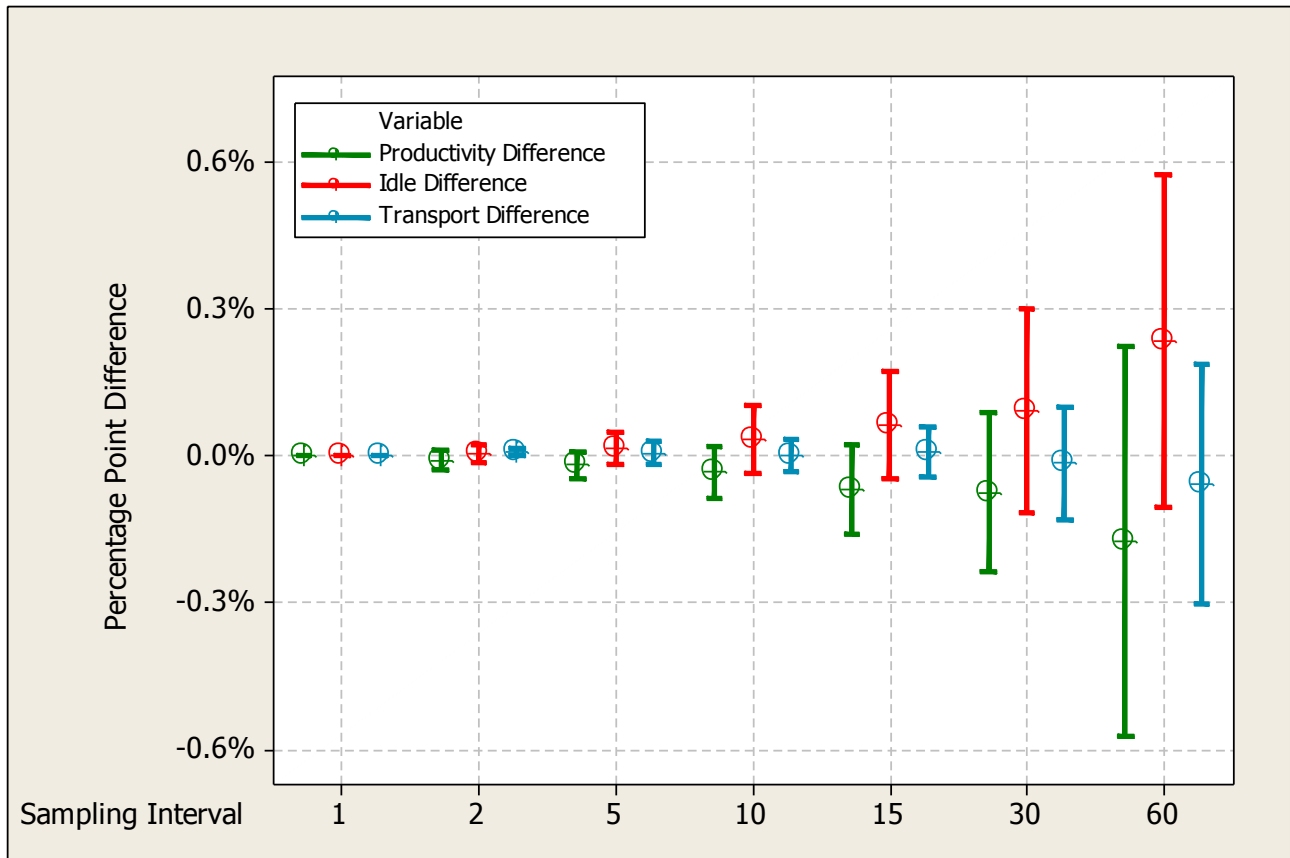
Figure 5: Productivity Ratio Percentage Point Difference by Sampling Interval

The same trend holds true here for the summarized data, as sampling interval increases so does the error. The means diverge slightly, but what is more interesting is the increase in variance. As the sampling interval increases, the variance becomes much greater in each machine state. Overall, this was still a fairly accurate way to calculate machine productivity, even at the longer sampling intervals, there was still less than 1.2 percentage point total variance.

The 2011 data set only contained 16 fields so this analysis was repeated on the 2012 data set which was much larger; these results are presented later in the chapter.

Productivity Comparisons, Time Basis

The second method for evaluating the impact of the increased sampling interval was a time-based analysis. On the machine performance reports, it is necessary to have total amounts of time that the machine spent in each state of production, idle, and transport. There are also other performance numbers, such as bales per hour of production, or acres per hour of production, that require the use of a total amount of time.

To calculate these total amounts of time the number of data points for each machine state is multiplied by the sampling interval with which it was collected. Figure 6, below, shows this graphically, for the production time of example field F092. The down-sampled data sets were multiplied by their respective sampling interval, shown as Production Calculated. The product of this operation is representative of the number of seconds that the machine spent in that state. This number is very similar to that of the raw data collected at one second intervals, but there is some error associated to this process. The amount of error varies in magnitude and direction, depending on the sampling interval, the field, and the machine status.
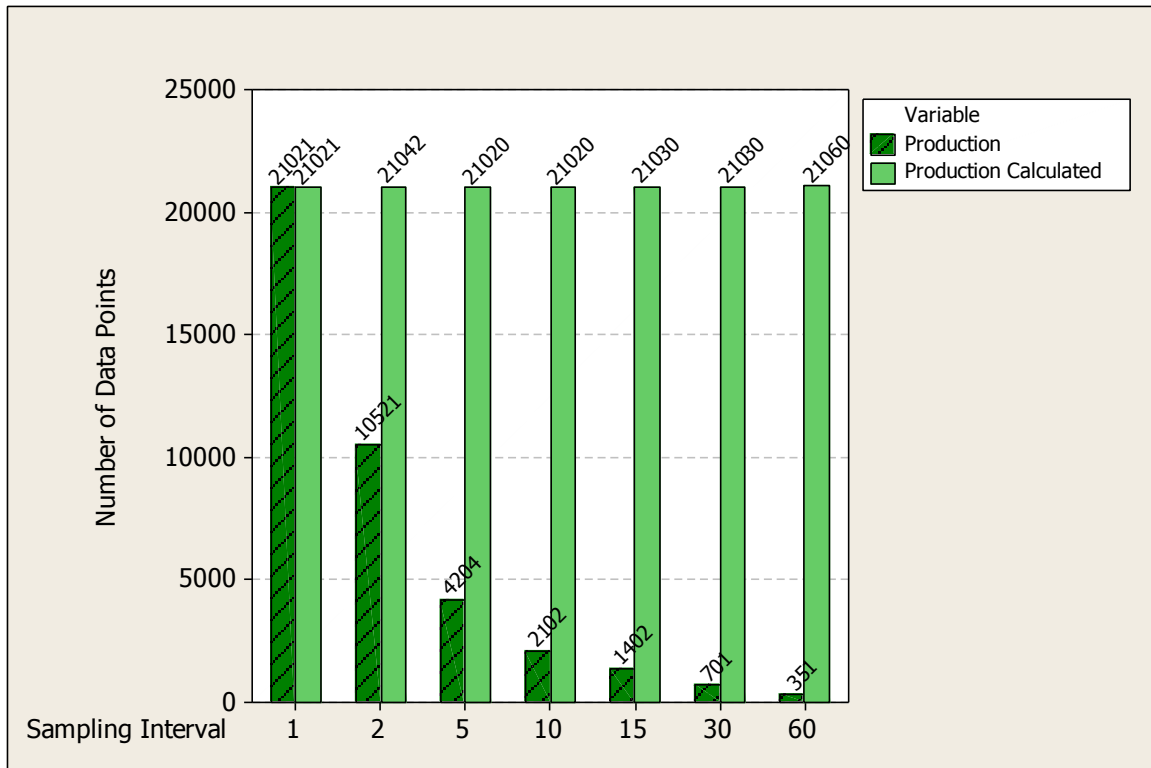
Figure 6: Down Sampled Data Multiplied by Sampling Interval.

This error was further quantified on a percent basis; this was done by finding the difference between the calculated time and the actual time of the raw data and then dividing that difference by the actual time. This is shown in Equation 2.

$$\% \ Error = \frac{(Down \ Sampled \ Data \ \times Sampling \ Interval) - Original \ Data}{Original \ Data}$$

Equation 2: Percent Error calculation for time based productivity

The percent error was calculated for each field, machine state, and sampling interval. The data was categorized by sampling interval but also by machine state just as for the ratio based productivity analysis. Figure 7 summarizes these results and shows a very similar trend to

the ratio analysis: as sampling interval increases so does the variance of the error. The time-based percent error is much larger than that of the ratio based with some error stretching out beyond 5%.
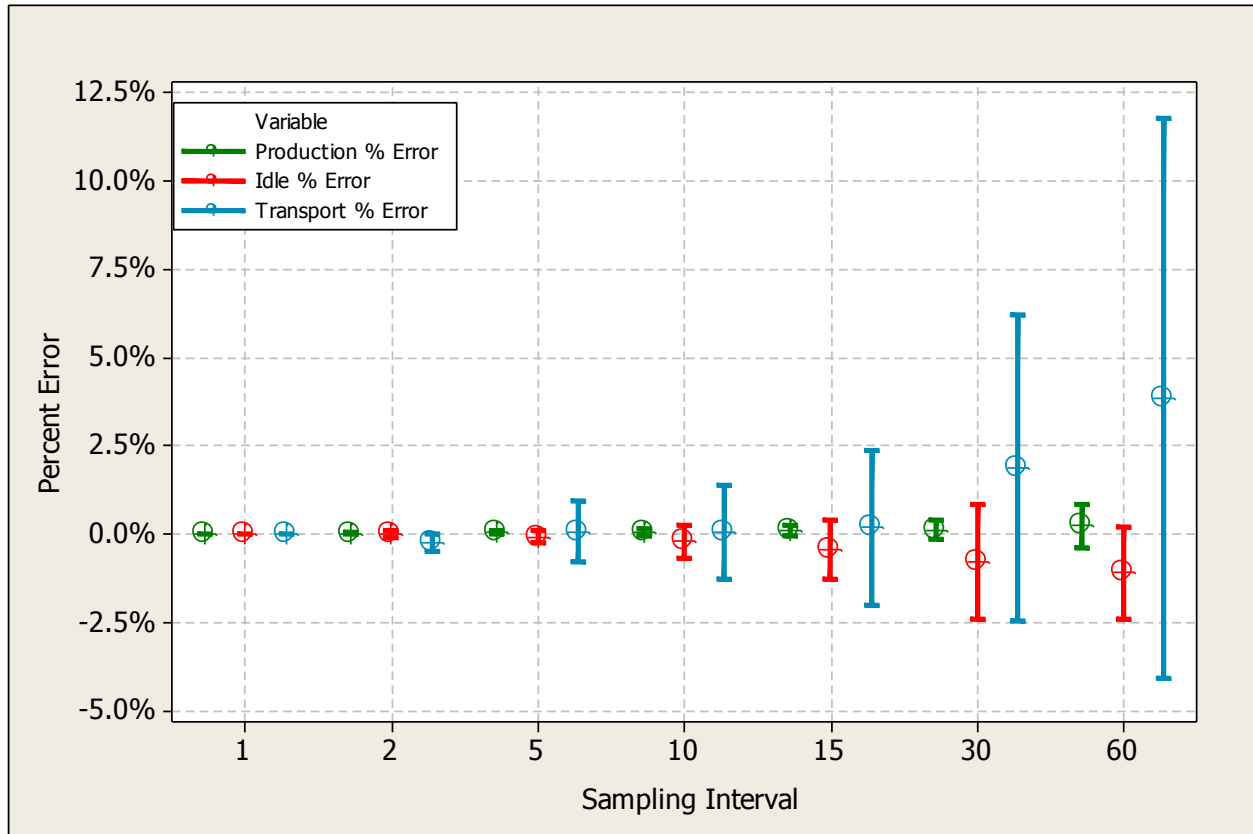
Figure 7: Time Based Productivity Error

This larger magnitude in error is due to the multiplication process in estimating the total amounts of time for each sampling interval. If there is a small error in the point count of each machine state for a given sampling interval, that error is compounded and magnified when multiplying by the sampling interval to get the estimated time.

There is much higher variation and error in the idle and transport state than there is in the production state. This is strictly due to the amount of data points there are in each state. The highest percentage of data points fall into the production state, while the lowest falls into transport. If it is a small field or a field with very little transport time, there will not be many data points in the transport state. This number becomes even lower when being down-sampled for each sampling interval. The smaller this data point count gets, the more likely it is for error to be induced when multiplying back by the sampling interval.

2012 Data Analysis

The data collection for the 2012 harvest was much more intense, with loggers on every machine in the fleet. There were 82 fields totaling 9,300 acres that were able to be used in this analysis, the other fields from the 2012 harvest were excluded due to lost or corrupted data.

The larger data set of the 2012 harvest allowed for more in-depth analysis of both the productivity calculation methods. It also created more confidence in the outcome with a much larger sample size.

Ratio Basis

The productivity ratio analysis was conducted the same as it was for the 2011 data. The data log files were down-sampled using the same alternative sampling intervals of; 2, 5, 10, 15, 30, and 60 seconds. The results were summarized in the same format and organized by sampling interval.

The same trends that were observed in the 2011 data were also present in the 2012 data. As the sampling interval increases, so does the error and variability. The magnitude,

however, was slightly larger than that of the 2011 data. Figure 8 below shows this graphically with some points stretching out beyond 0.5% difference.
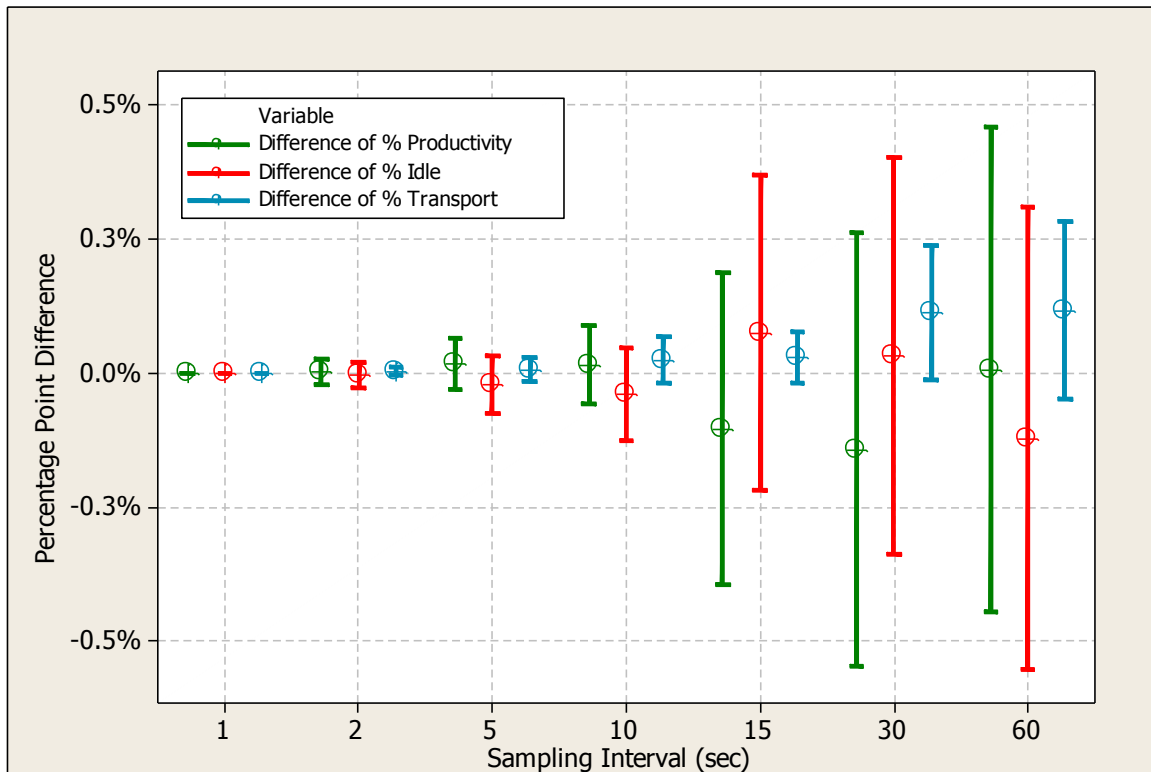
Figure 8: 2012 Productivity Ratio Percentage Point Difference by Sampling Interval

As sampling interval increases, the variance for each mean difference also increases. Because of this increase in variance, a statistical difference cannot be shown between the means for the different sampling intervals. This provides no way to distinguish which sampling interval had a significant impact on the productivity calculation.

This shifted the focus to assessing the standard deviation of each sampling interval, rather than the means. This method looks at how accurately each sampling interval can predict the actual productivity, and further investigates the trend seen in the summary chart of Figure 8. The trend visually observed in the chart shows an increase in variance as sampling interval

increases. A test for equal variances between each sampling interval was performed. The test was performed on one machine state at a time. Figure 9 shows the results from the equal variance test of percent productivity difference.
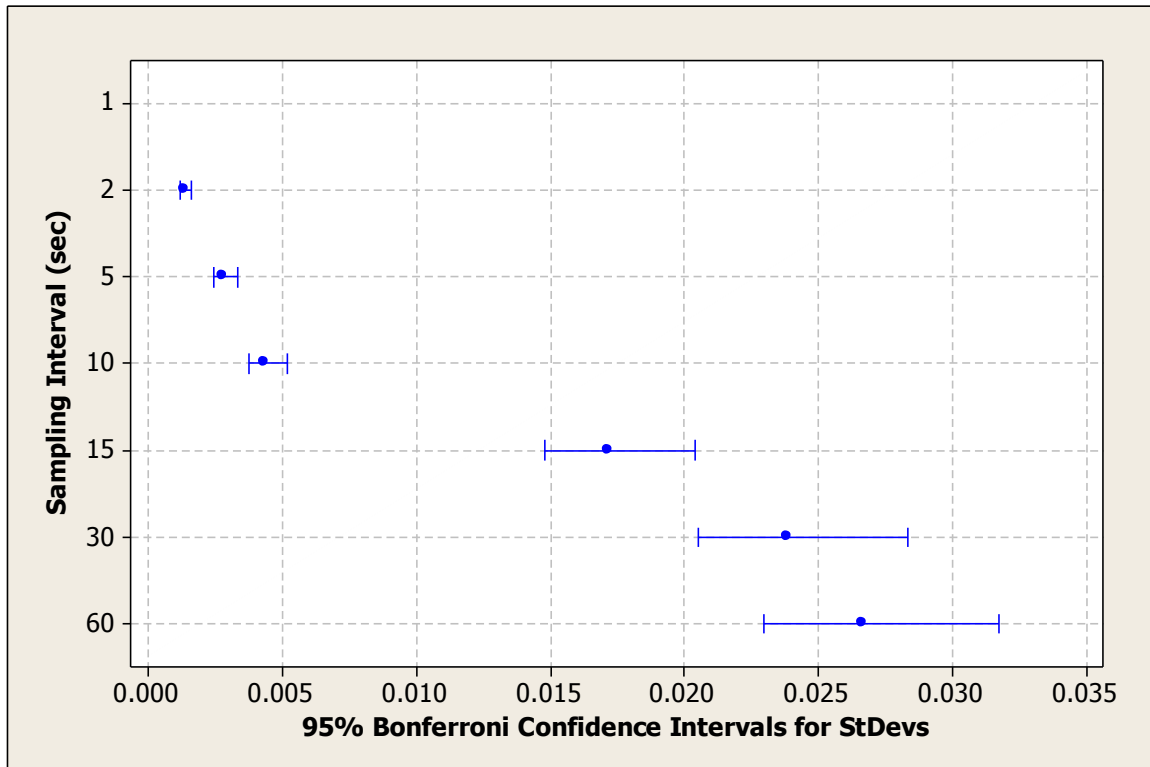
Figure 9: Equal Variance Test for Productivity Ratio

There is a clear separation between the standard deviations of the 10 and 15 second sampling intervals. Based on these results, there is enough evidence to show a statistical difference between the standard deviation of the two sampling intervals. This break point falls right in the middle and breaks the sampling intervals into two clusters. The lower cluster, which goes up to the 10 second interval, has very small standard deviation, less than 0.005 percent.

The upper cluster has significantly higher standard deviations than that of the lower but is still not greatly significant. Refer to Appendix A for boxplots of machine state productivity differences to visualize the standard deviations.

Figure 9 only shows the standard deviations for the difference in percent productivity. All the machine state calculations need to be considered to determine which one if any of the sampling intervals has an effect.

To compare only the sampling intervals, the machine states were combined. To do this, the mean of the standard deviation was found for each machine state at each sampling interval. An ANOVA was then performed on these mean values to determine the relation between the different sampling intervals. The ANOVA process pooled the mean standard deviation of each sampling interval into a category, this category contained the mean standard deviation of the production, idle, and transport machine states. Table 3, below, shows the Tukey grouping of the sampling intervals.

Table 3: Tukey Method Grouping of Ratio Difference Standard Deviations

Sampling Interval	N	Mean Std Deviation	Grouping		
1	3	0.0000	A		
2	3	0.0011	A		
5	3	0.0025	A	B	
10	3	0.0040	A	B	
15	3	0.0124	A	B	C
30	3	0.0177		B	C
60	3	0.0206			C

Means that do not share a letter are significantly different

For a Tukey grouping such as this, the means that do not share a letter are significantly different from each other. Based off of this chart and looking at the letter C, the 2, 5, 10, and 15 second intervals are significantly different from the 30 and 60 second intervals. Results also indicate that the 2, 5, and 10 second intervals are statistically different from the 15, 30, and 60 second intervals when looking at the letter A. This second division point is similar to what was seen in the Bonferroni chart with a distinct split between the 10 and 15 second intervals. This ANOVA comparison, however, contained a combination of all the machine states, whereas the Boneferroni only contained the productivity state.

The objective of this research was to find the longest sampling interval that did not significantly impact the productivity calculations. From this study of the productivity ratio differences, it appears that either the 10 or the 15 second interval does not significantly affect the productivity calculations. In order to make a final selection, though, the time-based productivity calculations still have to be considered.

Time Basis

The time based productivity numbers are an important tool in evaluating the factors that play into the productivity of a machine. It is necessary to know exactly how much time a machine spent in production, idle, or transport rather than just a percentage ratio of the total time. The time-based analysis was also repeated on the 2012 data, which consisted of 82 fields.

The data was down-sampled using the script that was added to the automated data processer. Each of the alternative sampling intervals was applied and the results summed for each machine state. The total for each machine state was then multiplied by its respective sampling interval, the same as for the 2011 data set. The error was then calculated by taking

the difference of the calculated time and the actual time and dividing that difference by the actual time to get a percent error. This was done for all the fields in the data set and the results are displayed in Figure 10.
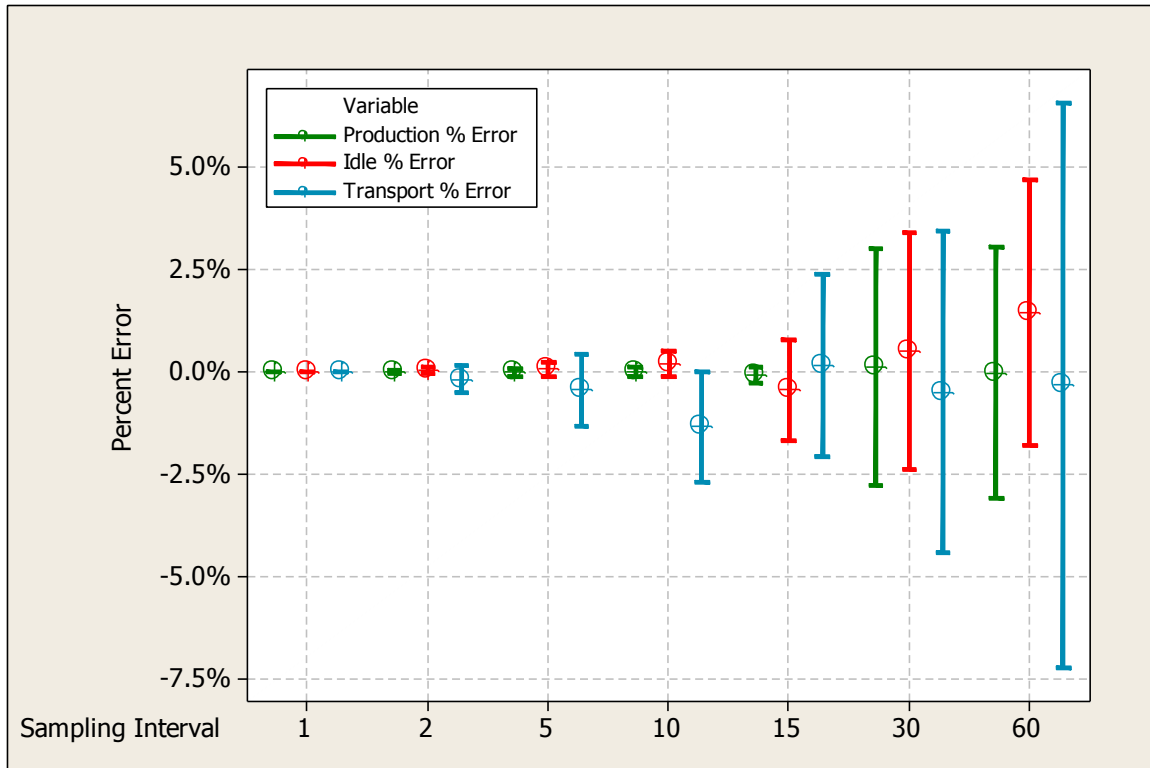
Figure 10: 2012 Time Based Productivity Percent Error

The interval plot shows the same trend as before in the other data sets, as sampling interval increases so does the variance. This data set, however, shows a more prevalent increase in variance after the 15 second sampling interval. The mean percent error is again not able to be proven to be statistically different, so an analysis of the standard deviations was performed.

Upon visual observation, there appears to be an increasing trend in the variance as the sampling interval increases. To confirm this, an equal variance test was performed individually

on each machine state. The results for the productivity machine state are seen below in Figure 11. To see the equal variance tests for idle and transport refer to Appendix B.
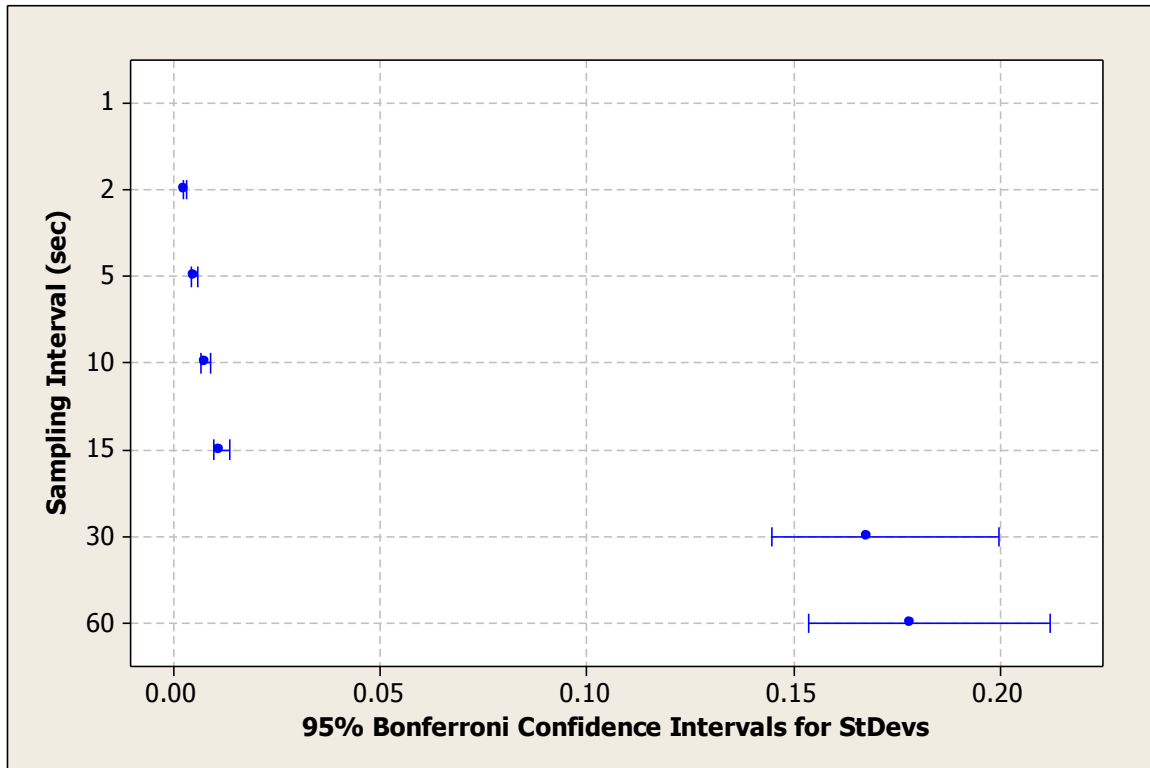
Figure 11: Equal Variance Test for Production Percent Error

The equal variance test for production percent error shows the definite break between the 15 and 30 second intervals that was observed in Figure 10. This break point is similar to that seen in the productivity ratio analysis and further supports the significant difference between the 15 and 30 second sampling intervals. Again, this test was only for the productivity state and did not include the idle or transport machine states. To draw a conclusion, data from all the

machine states needs to be taken into account. The same process as in the ratio analysis was repeated for this analysis.

The mean standard deviation was found for each sampling interval and each machine state. An ANOVA was then performed on these mean standard deviations to find the relation between the sampling intervals. The ANOVA process pooled the mean standard deviation of all machine states for each sampling interval, and then compared all the sampling intervals. The Tukey method grouping results are shown below, in Table 4.

Table 4: Tukey Method Grouping of Mean % Error Standard Deviation

Sampling Interval	N	Mean	Grouping		
1	3	0.0000	A		
2	3	0.0091	A		
5	3	0.0224	A		
10	3	0.0354	A	B	
15	3	0.0721	A	B	
30	3	0.1882		B	C
60	3	0.2561			C

Means that do not share a letter are significantly different

In the Tukey grouping, the means that do share a letter are statistically different. From this table, looking at letters A and C, the 30 and 60 second intervals are statistically different from the 2, 5 ,10, and 15 second intervals. This further confirms the breakpoint that was observed between the 15 and 30 second intervals in the Bonfferoni chart, as well the interval plot. The one second and 15 second intervals are grouped together in group C, which indicates that they are not significantly different.

Results and Conclusion

From the results of the ratio-based and time-based productivity analysis, it was concluded that there was an increase in variability at the 30 second sampling interval. This increase was found in both sets of analysis and shown in the Bonferroni interval plots. It was also further proven in the Tukey method grouping of the ANOVA results that there is a significant difference that occurs between the 15 and 30 second sampling intervals. With the 15 second sampling interval the productivity of the machine can still be calculated accurately. This is true of both the productivity ratio and time based calculations.

It is also important to note that in the Tukey chart for each analysis there was not a statistical difference between the 1, 2, 5, 10, and 15 second intervals. It cannot be said that the 1 and 15 second intervals are the same, only that they are not statistically different. This scenario occurred in both the ratio and time-based analysis and further confirms the break point between the 15 and 30 second intervals.

Shifting Transition Point Error

The error associated with increasing the sampling interval comes from shifting transition points. A transition point is the point in time which the machine switches from one state to another. For example, when the machine is sitting in the field idling, it is in "idle". However as soon as the baler is engaged and it starts to produce bales, it is in the "production" state. The point at which it changes from idle to production is a transition. Throughout the course of a day, the machine will change states multiple times and have many transition points. The exact number of transitions varies greatly and is dependent on a large number of factors such as, breakdowns, field size, weather, operator, machine type, and many more. There is no way to

predict the number of transitions a machine will have in any one day due to the large variety of influential factors that cause a transition.

When sampling at one hertz, it is easy to precisely tell when a machine changes state. But as sampling interval increases it is much harder to distinguish when exactly a transition occurs. This decrease in precision is what causes the error in both the ratio and time-based productivity calculations. The difference is that in the time-based calculations that error is being multiplied by a factor equal to its sampling interval.

For instance if a machine transitions from idle to production right after a sixty second interval data point is collected, the machine state will not change to production until the next data point is collected. This means that there is 59 seconds of production time that is actually being recorded as idle time. This is an extreme case scenario, but this is exactly what is occurring at each transition point as the sampling interval is increased. Refer to Appendix C for a graphical representation of this.

Ideal Sampling Interval

Based on the results from this study, 15 seconds was shown to be the ideal sampling interval for collecting GIS data in a commercial scale corn stover biomass harvest. This will decrease the data log file sizes by 93% while having no significant impact on the machine productivity calculations. This significant decrease in file size will drastically decrease data storage space needs, as well as data processing times. As the industry demand for telematics systems increases, the 15 second sampling interval will be the best balance of accurate productivity calculations and reduced data file sizes. This will greatly increase the reliability of the wireless data transfer from these units, and allow for on board data processing. Telematics

is the next step in large fleet management solutions. It offers real time data acquisition and

processing with instant access to machine productivities, which allows operators and managers

make informed data driven decisions.

References

Askey, Jeffery C., Matthew Darr, Keith Webster, Benjamin Covington, and Jeremy Brue.

   Automated Logistics Processing of GIS Data for Agricultural Harvest Equipment. Paper.

   no. #1596410, 2013 ASABE Annual International Meeting

Covington, Benjamin R. "Assessment of Utilization and Downtime of a Commercial Level Multi-

   pass Corn Stover Harvesting System." Thesis. Iowa State University, 2013. Print.

Webster, Keith E. "Single-pass Corn Stover Harvest System Productivity and Cost Analysis."

   Thesis. Ames, Iowa: Iowa State University, 2011. Print.

CHAPTER 3:

OPTIMAL CONFIGURATION AND LEAD TIME OF CORN STOVER

BIOMASS HARVESTING TEAMS

Abstract

Harvest team organization for biomass feedstock production has a major impact on the productivity of the team, as well as the productivity of the equipment. Organization is one of the biggest areas for improvement and can have an almost instant effect on productivity levels of windrowing and baling systems. When these operations are joined together as one unit, the baling productivity is limited while waiting on the windrowers. This decreases the overall productivity of the harvest team and adds additional cost to the biomass feedstock supply chain.

In order to be effective, the two operations need to be split and the windrowers should maintain a lead time over the balers. This lead time needs to be long enough to cushion the effect of any mechanical problems and the slower transport speeds of the windrowers, but yet short enough that, if a rain event were to occur, the windrowers could be shut down and give the balers time to catch up and harvest the ground that has already been completed by the windrowers.

By splitting the two operations, significant gains can be seen in productivity. In order for harvesting operations to run smoothly, windrowing and baling operations have to be managed separately, with sufficient lead time maintained between the two operations.

Introduction

The task of supplying a 25-plus million gallon per year ethanol plant is challenging and multifaceted. A corn stover biomass supply chain, totaling 740,000 large square bales, will be needed each year for the plant to operate. These bales all have to be harvested within the average 30 day fall harvest window of central Iowa. To accomplish this, it will take a large fleet of equipment and an immense work force. The key to making all this function smoothly is careful planning and good organization.

Organization has a major impact on the productivity of a harvest team, as well as the productivity of the equipment. Organization is one of the biggest areas for improvement and can have an almost instant effect on a harvest team's productivity levels. This analysis looks at the possible gains in productivity by separating the windrowing and baling operations. This analysis used data from a 7,500 acre corn stover biomass harvest in central Iowa conducted in the fall of 2011. This analysis was one of the first to use machine parameters, such as PTO speed, engine speed, and vehicle ground speed, to sort GIS data into the specific categories.

There were 6 fields with complete sets of windrower and baler data from the 2011 harvest; these are the fields used in the 2011 analysis for the first portion of this chapter. Because of this small data set, it was decided to repeat this analysis on the 2012 harvest data set. There were 126 fields totaling 14,100 acres used in this analysis; the others were not used due to missing or corrupted data. The 2012 data analysis uses the same concepts from the 2011 analysis, but uses updated processes. The results of this analysis are presented in the second portion of this chapter.

Research Objective

The objective of this research was to determine the optimal configuration and lead time for a corn stover biomass harvesting team. It was observed during the 2011 harvest that windrowers and balers would travel as teams from field to field. This created a wait time for the balers at the begging of the field, as well as a wait time for the windrowers at the finish of the field. The first task of this objective was to investigate the effect this wait time had on the productivity of both the windrowers and the balers. The second task was to investigate what performance gains could be made by splitting the windrowers and balers into separate teams, and what the optimal lead time should be between the two operations.

2011 Configuration and Lead Time Analysis

Methods

To analyze the productivity of a harvest team or a machine, it is necessary to classify how its time is spent. General operational segments of machinery systems can be subdivided into three discrete categories: Production, Idle, and Transport. These specified productivity categories organize the data from each machine and make it easy to classify which data points fall into the production category, which is the only category in which productive work is being done.

To qualify as production, a machine has to be in working mode, this means that the machine is in the field doing work and being productive. If the machine is idle, it is not in working mode, but is in the field with the engine running. Transport includes time when the

machine is moving through the field but not doing work, as well as time on the road. For a harvest team to be the most productive, they need to maximize their production time and minimize their idle and transport time. Transport is one area that is more or less fixed because machines can only travel down the road so fast. This time can be optimized, however, with good organization and by ensuring the operator knows where to go and the shortest distance to get there. Idle time is more controllable and should be minimized in order to maximize overall productivity.

The factors that influence the amount of production time are primarily related to equipment. If larger, higher capacity equipment is used, it will increase the productivity. In addition, good maintenance and servicing will help keep the equipment operating at peak performance and maximum capacity. To make any significant improvements in this area requires engineering and development work to increase the size and capacity of the equipment. Reducing idle time can have a direct and measurable positive impact on overall productivity.

Idle time is influenced by many factors and is the source of all wasted time and cost of a harvest team. Every machine hour that falls into the idle category is a reduction in the productivity, and adds excess cost to the whole harvesting system. Idle time can be influenced by poor quality equipment or poorly maintained equipment. Time spent on breakdowns is lost productive time, as well as the input cost of the operator who is just sitting. One of the biggest influencing factors on idle time is organization. There are two main operations within a harvest team, windrowing and baling. The organization and management of these two operations has a major impact on the overall productivity of the team.

Experimental Field

Data for this study was collected from windrowers and balers in a commercial scale corn stover harvest in Central Iowa. Data was recorded from equipment using the CyCAN data logging system developed at Iowa State University. This system records desired CAN parameters every second and ties them to a GPS coordinate. This style of spatial data collection shows the researcher where the machine was, and what it was doing, according to the CAN parameters selected to be logged. (Webster 2011)

Field Analysis Methods and Machine States

AgLeader SMS Advanced software was used to analyze the GIS data from each machine and determine the time it first entered and last exited the field. The data was analyzed based on three perimeters; PTO engagement status, engine speed, and vehicle ground speed. By filtering a field dataset for specific combinations of these parameters, it was determined whether the machine was in production, Idle, or transport (Peyton 2012, Covington 2013).

To determine how long equipment was waiting at the start and finish of each field, a time log was constructed. Three time categories were established; Begin, Working, and Finish. "Begin" runs from the time the machine first pulled into the field until the time it first entered the production state. "Work" runs from the time production first started in the field until the time it last ended. "Finish" runs from the time production last ended until the time the machine last exited the field.

To find when the machine entered and exited the field, a query was done of the data points directly surrounding the field entrance. In the query results, the point with the lowest date and time is when the machine entered the field, and the point with the greatest date and

time is when the machine exited the field. An example of this process is shown below in Figure 12.

Figure 12: Field Enter and Exit Time Query (Covington 2013)

This example of the field enter and exit time query shows how the process is performed. By examining the minimum and maximum points for this field, it shows that, in this case, the baler entered the field at 5:58 PM and exited the field at 6:48 PM the following day (Covington 2013).

To determine when the baler first began work and finished work, the field dataset was queried for all production data points. A production data point is defined as a point that has: a speed greater than 2 kilometers per hour, an engine speed greater than 1500 rpm, and a PTO speed greater than 500 rpm. These parameters are the primary indicators that the machine was moving and doing work. From the production query results, the point with the lowest date and time is when the machine began work and the point with highest date and time is when the machine ended work for that field.

The visual playback function within SMS Advanced was also utilized to verify that the times from the queries coincide with the machine's movements in the field. The playback

function allows one to double check that the machine's movements follow a typical harvesting pattern, and that it was not doing anything out of the ordinary.

This analysis was focused on looking at the idle time at the beginning and end of a field. Any idle time that occurred in between the first start up and the last shut down for that field was included as working time. The only exception to this was overnight time. If an overnight event occurred somewhere during the duration of harvesting the field, the time that the tractor was turned off overnight was subtracted from the production dataset. By doing this, it gives an accurate amount of time that the machine spent in the field and could have been in production.

After this analysis, a time log for the field, including when each machine arrived, when it began work, when it finished work, and when it exited the field, was established. This time log was then converted to a decimal military time basis. By converting the times to this format, it is possible to graph them in a zero hour format. Every day begins at zero and the time accumulates up throughout the course of the harvest day. This was then used to establish a zero hour time log for each field. Whichever machine entered the field first started the zero hour for that field, and then all events were cumulative on top of that.

If a field took more than one day to complete, the time simply kept accumulating. For the next day, 24 hours was added to the times and if a third day occurred, 48 hours was added to the times. As was stated earlier, the overnight time was accounted for and removed from the working calculations.

This process was replicated on all the fields included within this dataset. Graphs were made to visualize the timeline for each field and make it apparent when equipment was waiting

at the beginning and finish of the field. To evaluate the loss in productivity due to this time spent waiting, a simple equation was used. The productivity loss was calculated by taking the sum of the begin time and the finish time, divided by the total time spent in the field.

$$Productivity\ loss\ (\%) = \frac{(Begin\ Time + Finish\ Time)}{Total\ Time}$$

Equation 3: Productivity Loss Percentage

Equation 3 yields a percentage of time that was spent waiting at the beginning and finish of the field; in other words, a percentage of time that was unproductive. This is a quick way to see the amount of possible improvement that could be made by understanding and addressing these harvest team organizational factors.

Data Analysis

Two treatment factors were developed to categorize and analyze the data set. Treatment 1 covers the effects of having the windrowing and baling operations together and the impacts it has on harvest team productivity. Treatment 2 covers the effects of separating the windrowing and baling operations and the potential gains in harvest team productivity.

Treatment 1, Operations Together

During the 2011 harvest, it was observed that the harvest teams would move as one unit from field to field. This created a backup at the start and finish of a field. The windrowers would begin while the balers waited for them to successfully windrow a portion of the field.

When the windrowers were finished, they would wait for the balers to finish, and all move to the next field together.

This action was also seen and quantified when examining the GIS data. The baler can cover more acres per hour than the windrower, therefore it would wait until the windrower had a sufficient area of the field completed before it would start baling. Alternatively, instances were measured where the baler would start at the same time as the windrower, but run at a reduced capacity equal to that of the windrower.

Both of these actions have a negative impact on the productivity of the baler because if it is waiting and with the engine running, it is burning fuel and accruing engine hours. If it is operating at a reduced capacity, it is not being as productive as it possibly could. Both of these actions result in increased production costs for the biomass feedstock supply chain.

The GIS data also confirmed a similar story for the windrower. In some instances, the windrower would get far enough ahead of the baler that it would finish first, but then it would sit at the field edge and wait for the baler to finish so they could move as a unit to the next field. This organizational issue is a reduction in the productivity of the windrower because it could have proceeded to the next field and maintained its lead time on the baler.

Example Case 1: Poor Organizational Logistics

Field 2 is an example of why moving the balers and windrowers together has a negative influence on their productivity. In this field, the baler arrived ahead of the windrower and had to wait until the windrower arrived and began working. The windrower then finished before the baler and waited for 4 hours until the baler finished. Both machines then waited in the field for

an additional hour before moving to the next harvest field. Figure 13 shows this machine interaction graphically, and highlights the time wasted due to poor organization.

Figure 13: Work Flow Timeline for Field 2
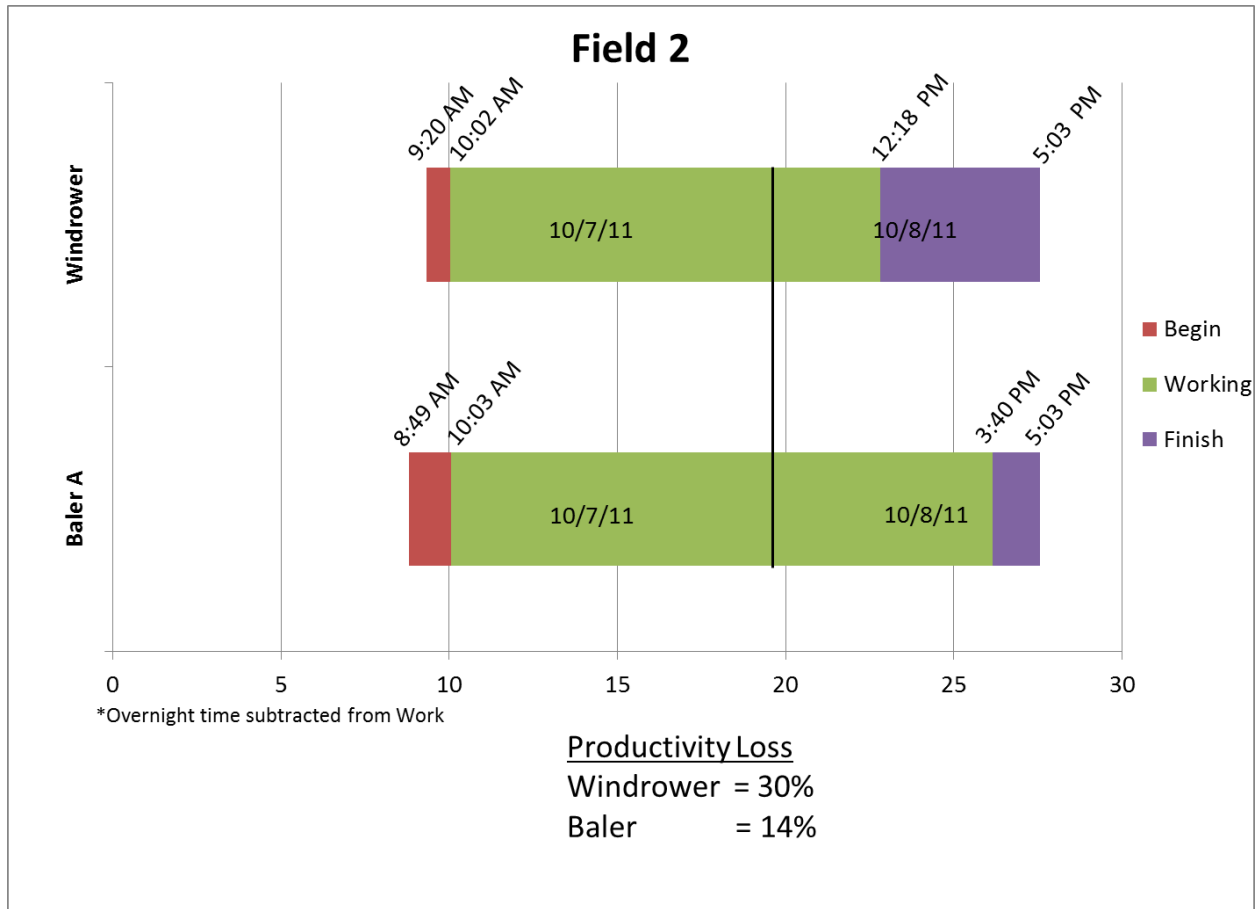
There is another productivity-related issue occurring here as well; the baler has a higher capacity than the windrower. Since they started at the same time, they should have finished at almost the same time. However, the baler finishes much later, which indicates it was either running at a reduced capacity, stopping and waiting on the windrower, or a combination of the two.

The productivity loss in this field for the windrower was 30%, and for the baler there was a 14% loss. These are large numbers when you start considering all the costs associated with idle equipment time. These losses could have been greatly reduced or eliminated if the crew had been organized differently.

Productivity Analysis

Figure 14 shows the total unproductive time caused by waiting that each machine had in each field that was analyzed. These are unproductive hours that could have been spent windrowing or baling. Some of this time is necessary and cannot be eliminated because it is needed to perform regular maintenance and to transition the machinery from field mode to road mode. Beyond that, it is all excess time that can be eliminated with better organization.
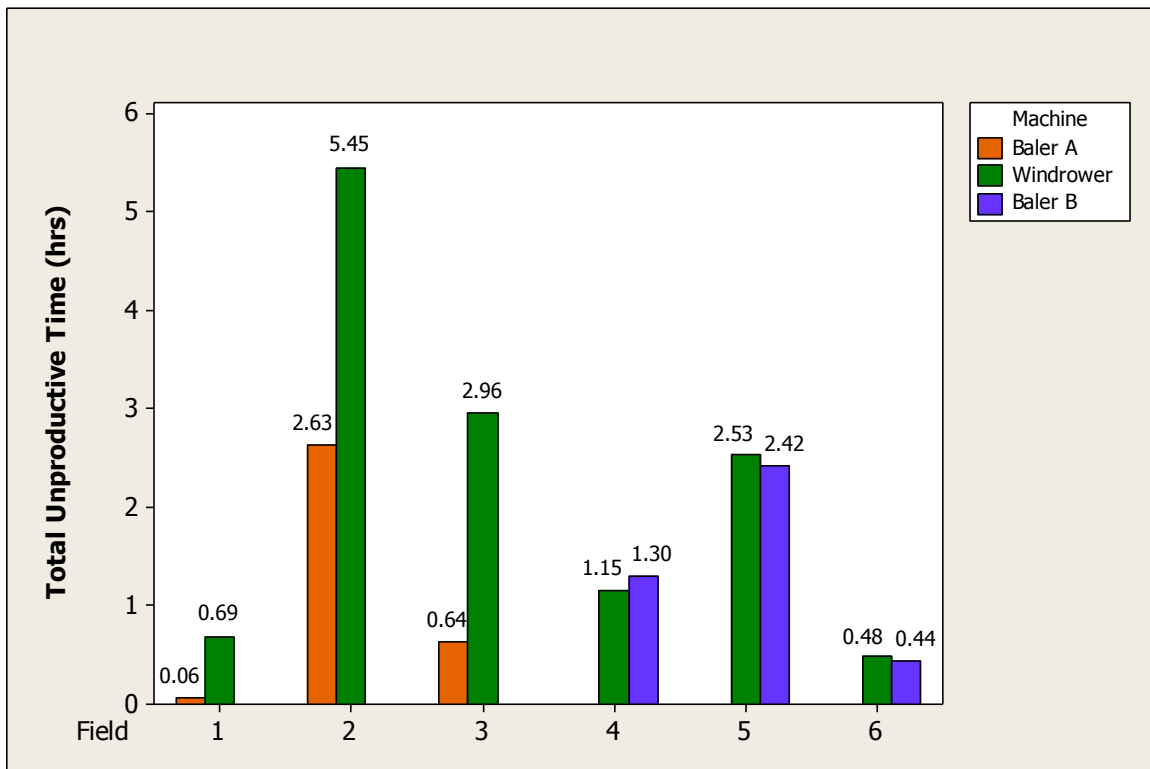

Figure 14: Total Unproductive Machine Time

In Figure 14, Field 1 and 6 can be seen to have low amounts of idle time; the idle times from these fields is an ideal amount of idle time for that equipment. The majority of that time should be spent at the beginning doing any servicing and preparing the machine for field operations. However, once a field is finished, the machine should transition into road mode and there should be minimal wait time once the field is completed. This is especially true of the baler because it does not require any machinery configuration changes to transition to road mode. The stalk chopping windrowers do require some time to transition to road mode, but after observing them during the 2011 harvest, this time should only be around 15 – 20 minutes.

Four out of the six fields that were analyzed had what would be classified as excessive idle times. This works out to be 66% of the fields. If 66% of all the fields harvested during the 2011 harvest season had excessive idle times such as these, there is a huge opportunity to increase harvest team productivity by addressing this organizational factor.

Treatment 2, Operations Separated

After observing the harvest teams during the 2011 harvest and data from the fields in this data set, it became clear that the windrowing and baling operations needed to be separated. They can still all be part of the same harvest team, but the two operations need to be managed separately. By separating the two operations, it is possible to cut out much of the idle time at the beginning and end of a field.

This will have an almost instant effect on the team's productivity. Figure 15 shows the plots of productivity loss for each machine, with the average listed in the middle. These productivity losses could also be thought of as potential productivity gains for each machine. By minimizing the organizational issues and staggering the windrowing and baling operations,

Baler A, for example, could have a 6% increase in productivity during the next harvest. More significantly the windrower could see a 14% increase in productivity. As was stated earlier, there is a small amount of unproductive idle time that is necessary to transition the equipment in and out of road mode, but the rest can be eliminated with better organization.
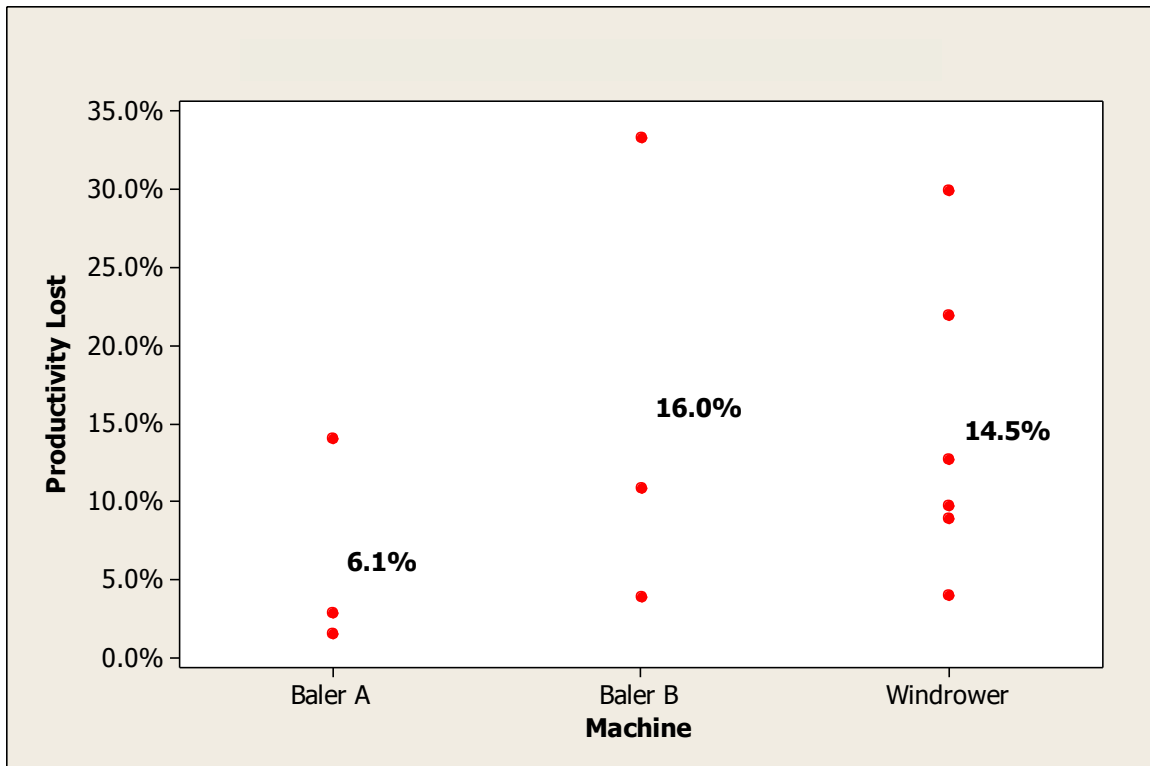
Figure 15: Productivity Loss per Machine, Loss Equals Potential Gains

Case 2: Improved Organizational Logistics

Field 1 is a great example of how the harvesting operations should go. Figure 16 shows the windrower arrived early in the morning, and had some begin time in order to transition from road to field operational mode. It then began work and the baler arrived to the field almost three hours later. The baler immediately began work and worked straight through until

the field was complete. The windrower finished and moved to the next field while the baler was

still working, then the baler finished and moved to the next field.
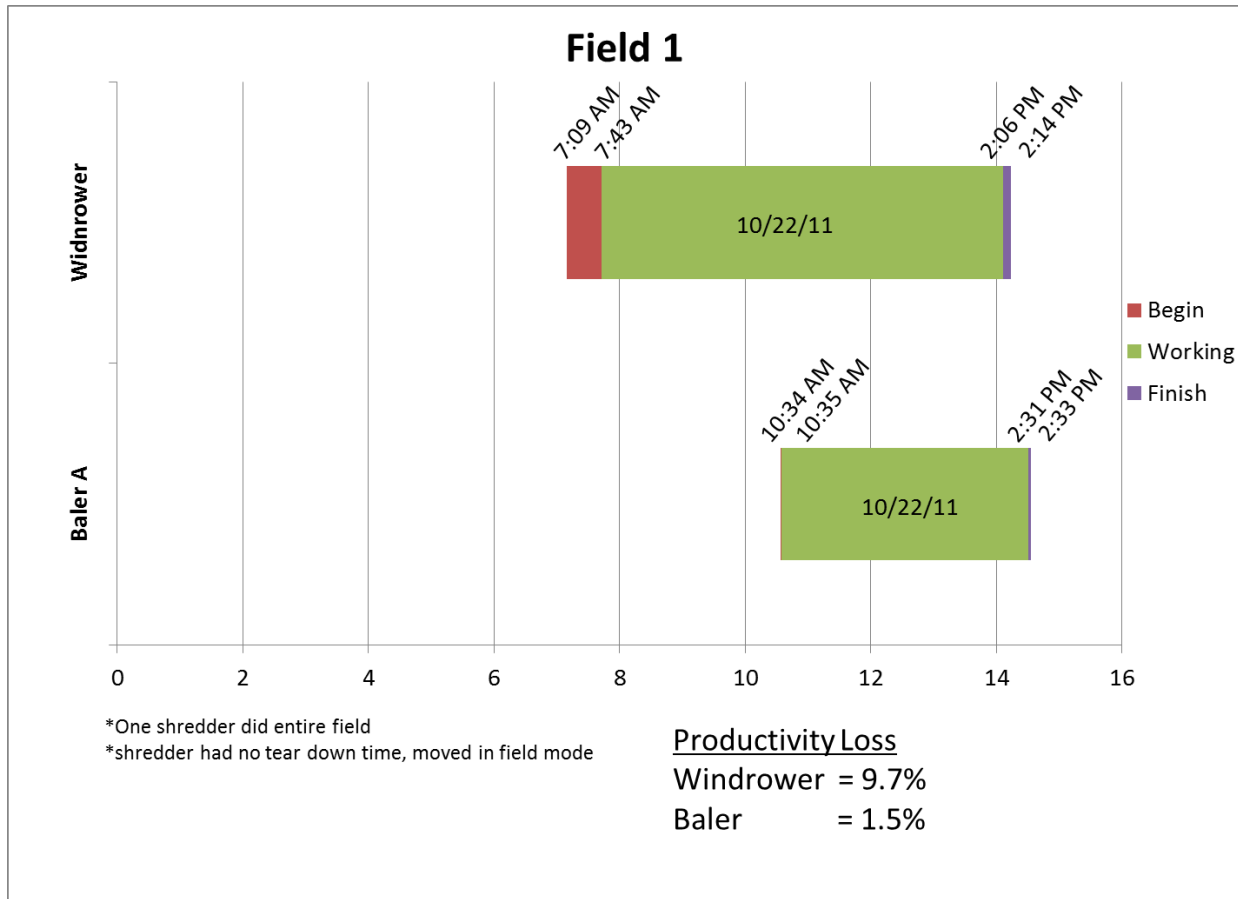
Figure 16: Workflow Timeline for Field 1

This is a good example of how the windrowing and baling operations should be

staggered, so that the windrower has a head start on the baler. The windrower had a 9.7%

productivity loss, which is understandable since it had setup time, while the baler did extremely

well and only had a 1.5 % productivity loss. This staggering arrangement allows for each

machine to run at full capacity and never have to wait on the other.

As Figure 16 shows, though, the baler almost caught up to the windrower by the time it completed the field. It's likely that, at the next field, the baler did catch the windrower and either had to go back to running at a reduced capacity, or had to stop and wait. This difference in capacity can be solved by adding additional windrowing equipment to the harvest team in order to equalize the operational capacities between windrowing and baling.

Lead Time

If the operational capacities are equalized there is still a certain amount of lead time that needs to be maintained between the windrower and the baler. This lead time is critical to making the harvesting operation run smoothly and prevents any piece of equipment from having to sit idle waiting on another. The lead time acts as a cushion or operational slack between the windrowing and baling operations. Even if the capacities are equalized, there are other factors that can cause equipment to get backed up.

Mechanical breakdowns, slower transport speeds, and longer road to field transition times are all factors that could impact the windrowers lead time. This is why it's important for the windrower to have a large enough lead time that these factors don't slow it down to the point that the baler catches up with it throughout the course of a harvest day.

However this issue of lead time can also work the other way. If the lead time becomes too large, there is a risk of rain and weather events that could occur before the baler can harvest the material on the area the windrower has already covered. This makes the issue of lead time much more complicated. There is a balance between staying far enough ahead so that baler does not catch up, while at the same time not getting too far ahead to the point of risking windrows being deteriorated by weather related events.

Equal Operational Capacities

Field 1 was the most ideal scenario of proper lead time and exemplary of how a field should be harvested. In Figure 16, the windrower had a three hour lead time on the baler. In this case, however, they did not have equal capacities, so the baler was caught up to the windrower after four hours of operation. The Windrower had a working time of 6.39 hours while the baler had a working time of 3.93 hours.

$$^{3.93}/_{6.39} = .615 = 61.5\%$$

In this case, the baler had a 61.5% capacity advantage over the windrower, but there was only one windrower and one baler. To equalize the capacities, you would need 1.5 windrowers per baler. An Ideal harvest team setup would be three windrowers and two balers. This would result in equal capacities for windrowing and baling operations.

Lead time Determination

As we begin to look at lead time throughout the course of the day, and not just for one field, things like transportation speed, setup, teardown, and mechanical problems now become factors to consider. The windrowers use much smaller tractors than the balers, so their road speed is not as fast. The windrowers also have more setup and teardown time because they have to reconfigure from road mode to field mode, whereas the balers can just simply pull into a field and start working. These factors all impact the lead time for the windrowers.

Mechanical problems are usually seen as down time for a single machine, and a loss in productivity. In a commercial harvest system, however, mechanical problems cause the operational capacities to become unbalanced. If a windrower goes down because of mechanical

problems, the balers then have greater capacity than the windrowers, and will eventually catch up.

Field 1 was a very ideal field, and it had a one-to-one ratio of windrowers to balers. In this field, the windrower was about 60% slower. To balance the capacities, it was determined that 3 windrowers and 2 balers would be needed for a harvest team. If one windrower were to break down, the operational capacities would be unbalanced and the balers would be at a 60% advantage again. If the balers are back to a 60% advantage over the windrower, they will catch up after only four hours of operation. This is where the lead time comes into play.

If there is enough lead time or operational slack between the windrowing and baling operations, it would allow for time to get the windrower operational again before the balers catch up. At a 60% advantage, it would take the balers four hours to catch up, if they were running at full capacity. As a lead time, four hours seems pretty reasonable and would allow for most minor repairs to be made on the windrower. Four hours also equates to a little under half of a 10 hour working day, which is typical of fall harvest operations.

With this data set, it is also possible to examine the Begin and Finish times of the windrower. Since windrowing is the first step in the system, it does not have to wait on any other operation in order to perform its task. Thus, examining the distribution of the Begin times should yield an idea of how long it takes to reconfigure the windrower from road to field mode. Figure 17 shows the distribution for this dataset and the average Begin time of .26 hours, or about 15 minutes. This average is pulled down slightly due to two fields in the data set that did not have any setup time because the windrower was transported in field mode. This analysis, combined with in field observations from the 2011 harvest, yields an average reconfiguration

time of 15-20 minutes. This reconfiguration time also applies to the Finish time at the end of

the field. This means there is a total of 30 – 40 minutes per field required just for the

windrower to reconfigure between road and field mode. This is 30 - 40 minutes of delay per

field that the baler does not experience. Estimating that the crew will only move fields once per

day, this is almost an additional hour that should be added to the windrowers' lead time.
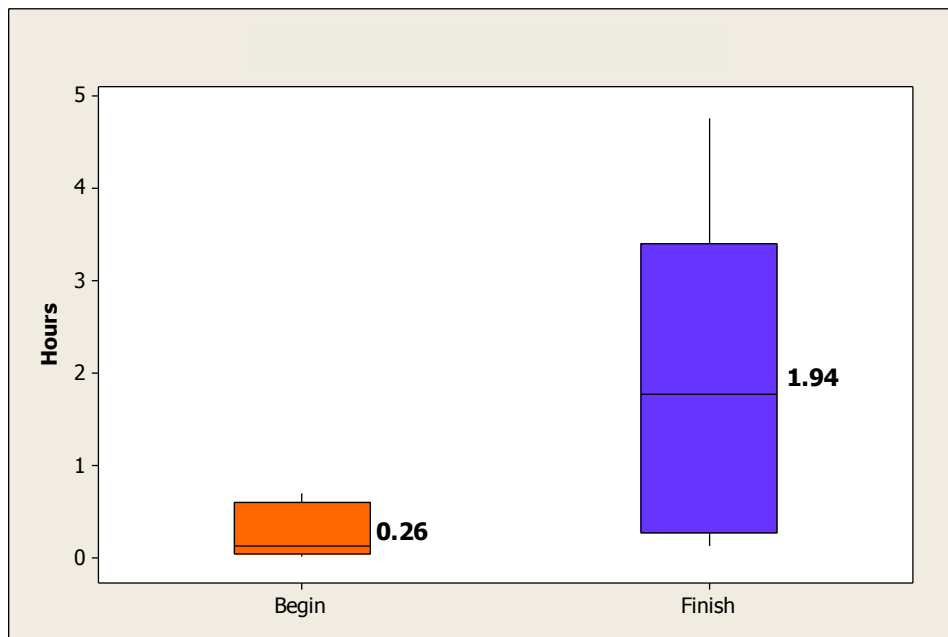
Figure 17: Windrower Begin and Finish times

There is no data in this data set from which an analysis on the difference in transport

speeds between the windrower and the baler can be performed. All that is known is that the

windrowers used smaller tractors that did not travel as fast. To account for this difference in

road speed, the windrowers' lead time should be increased; however, a specific amount cannot

be determined without collecting additional data.

Taking all these factors into consideration, a necessary lead time for the windrowers over the balers was determined. Mechanical breakdowns are the largest factor, and it was determined that 4 hours of lead time would provide a good operational cushion to repair most breakdowns. Reconfiguration and transport time also have an impact, and they were determined to require 1-2 hours of additional lead time. This brings the total lead time for the windrower to 5-6 hours. This 5-6 hour lead time is long enough to cushion the effect of any mechanical problems, reconfiguration, and transport time. Yet it is short enough that, if there is the possibility of a rain or weather event, the windrowers could be shut down in order to allow the balers to catch up. This duration of lead time would give the balers about half a day to catch up and finish the ground that was already completed by the windrowers.

2012 Configuration and Lead Time Analysis

The configuration and lead time analysis was repeated on the 2012 fall harvest data set. The purpose was to confirm the results from the 2011 analysis on a larger data set, and make comparisons between the two. Updated methods developed from the 2011 analysis were used to sort and process the new data.

Methods

Data for the 2012 harvest was collected using the same CyCAN system developed at Iowa State University. Loggers were installed on all of the equipment used in the 2012 harvest and monitored more frequently to ensure quality data collection. Data processing was accomplished using data processing software developed at Iowa State University. This automated processing system takes the place of the hand analysis that was previously done

using AgLeader's SMS Advanced software. This greatly increased the efficiency and accuracy of the data processing portion of this analysis (Askey 2013).

Data from the loggers was organized by field using the geofencing process. This process organizes the data points into their respective field using the GPS coordinates of the data points. A geographical boundary is established for each field, that boundary is then used to sort the data points. Any points with GPS coordinates that fall within that boundary are categorized in that field.

To find the field enter and exit times the system scans the data set of each field and finds the data points with the minimum and maximum date and time. This works much the same way as doing a manual data query in SMS Advanced but takes a fraction of the time. The first production time and the last production time are also found in a similar manner. The system again scans the data set for each field and locates the data points containing the first and last time that the machine was in the production state within that field. The production state is defined as ground speed greater than 2 kilometers per hour, Engine speed greater than 1500 rpm's, and PTO speed greater than 500 rpm's.

The field enter and exit times as well as the production start and stop times were already built into the data processing software. To complete this analysis however additional information was needed. To account for the overnight time the same way it was in the original 2011 analysis the data had to be summarized by day. For fields that took multiple days to complete the working time for each day had to be found. In the previous analysis this was done by hand in SMS Advanced, the updated processing software looks for the first and last key on time of each day. The key on event is the first time the machines key is powered on for the day,

and similarly for the key off event it is the last time the machines key was turned off for the day. This idea of machine "On" time was originally developed by Benjamin Covington of Iowa State University and implemented into the data processing software for this analysis. (Askey 2013, Covington 2013)

The final piece added to the processing software was to calculate the begin and finish times for each field. The begin time was calculated by taking the difference of the field enter time and the first production time. Similarly, the finish time was found by taking the difference of the last production time and the field exit time. With this in place all the necessary data could be gathered to complete the configuration and lead time analysis.

The 2012 harvest contained similar equipment to that of the 2011 harvest except there were greater numbers of each machine. There were also four different crews operating the equipment during this harvest season. This adds another influencing factor to the analysis and is examined later on in this section.

Data Analysis

The 2012 analysis was completed using the same concepts and analyzing the same factors as the 2011 analysis. The biggest importance of the 2012 analysis is that it contains a much larger data set with which to draw conclusions from.

The first item covered was the begin and finish time of machines. The machines are grouped as Baler A, Baler B, and Windrower. Baler A and Baler B are different manufactures of large square balers but each group contains multiple machines which were operated by different crews. Baler A and Baler B are the same respective manufacturers from the 2011

study. The windrower group contains all the windrowers from all the crews. All the windrowers were the same make and model but they again were operated by different crews.
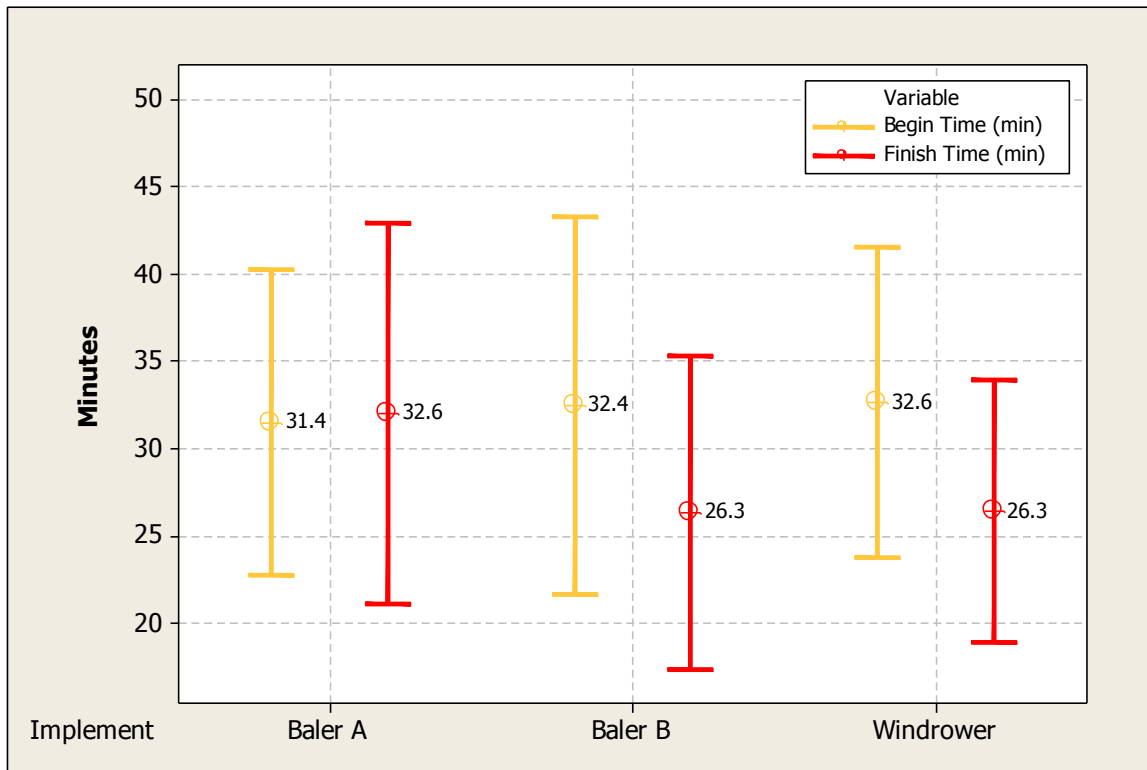
Figure 18: Begin and Finish Time by Machine for All Crews

In Figure 18, the begin and finish times for the 2012 harvest are significantly lower than that of the 2011 harvest which shows a great improvement in the overall organization of the crews. The average baler begin and finish times are around 30 minutes for both brands. The windrowers also showed a dramatic improvement with begin and finish times also averaging around 30 minutes. This is very good for the windrowers as this time includes their transition from road to field mode.

The next item evaluated was the productivity loss for each machine. Recall that the productivity loss is the sum of the begin and finish times divided by the total time spent in the field, shown here by Equation 3.

$$Productivity\ loss\ (\%) = \frac{(Begin\ Time + Finish\ Time)}{Total\ Time}$$

Equation 3: Percent Productivity Loss

Again overnight time was accounted for and not included in the total time. The total time was found by summing the entire working time from each day work performed in that field. The working time is defined as the period of time between the first "key ON" and the last "key OFF" time for that machine on any given day. Figure 19 below shows the resulting productivity loss for each machine category.
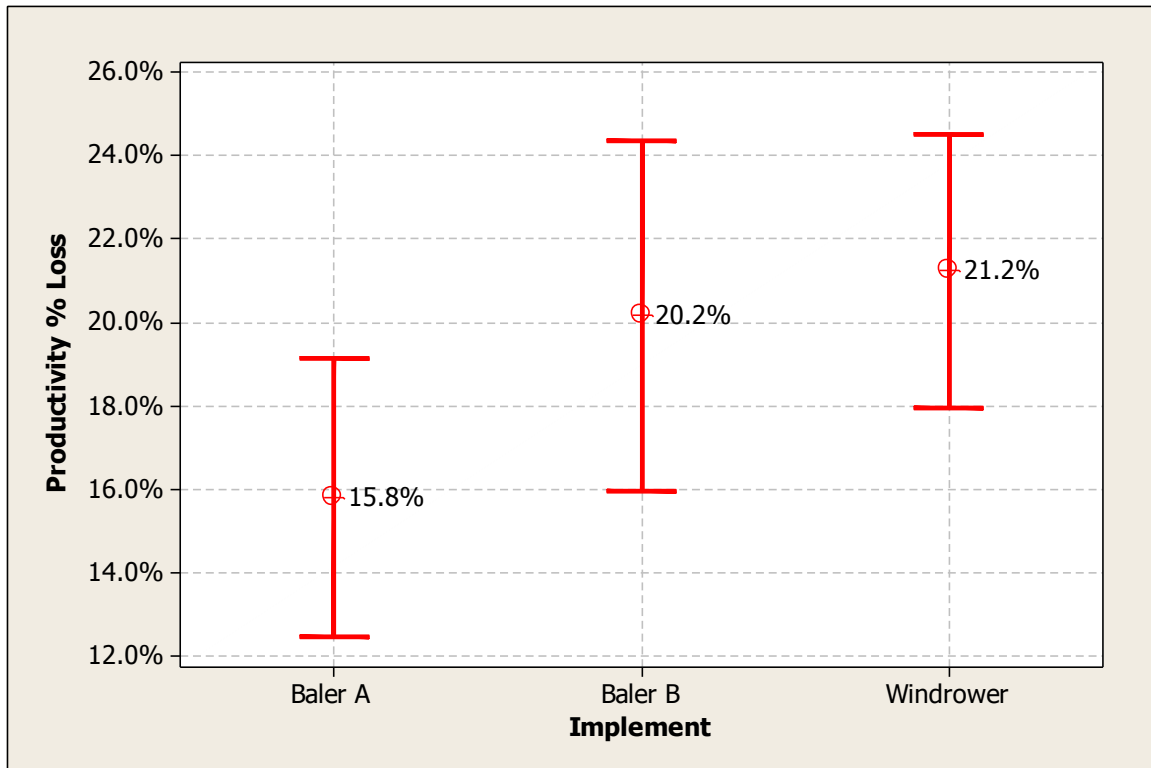
Figure 19: Percent Productivity Loss by Machine for all Crews

The productivity losses summarized above are slightly increased over that of the 2011 study. This is not unexpected though as this analysis contains a far greater amount of fields and machines. These productivity losses translate into possible productivity gains if they can be overcome with better organization.

The windrowers will always have some percentage of productivity loss due to their time required to transition from road to field mode and vice versa. So while that time is somewhat fixed in these calculations they are penalized for it as productivity loss. This was done because as the equipment evolves new systems could be engineered to decrease this transition time. As this happens the equipment can still be evaluated using the same process and acts as baseline to evaluate these design improvements as well as different manufacturers.

Figure 19 shows a 4.4% increase in productivity loss in baler B over baler A. This could be a difference in manufacture but is more likely associated with crew organization. To investigate this further the baler productivity loss and start times were summarized by crew. There were 4 different crews operating in the 2012 harvest which adds an additional layer to this organizational analysis.

Figure 20 below shows the productivity loss broken down by baler and by crew. Only two crew operated baler brand B, while all the crews operated baler brand A. By examining Figure 20, it can be seen that crew 2 and 3 had the lowest productivity loss for baler A. When looking at baler B crew 3 had the lowest productivity loss, while crew 1 had a higher productivity they were much more consistent in it. The fact that these productivity losses vary so much by crew highlights the fact that this is indeed an organizational issue.
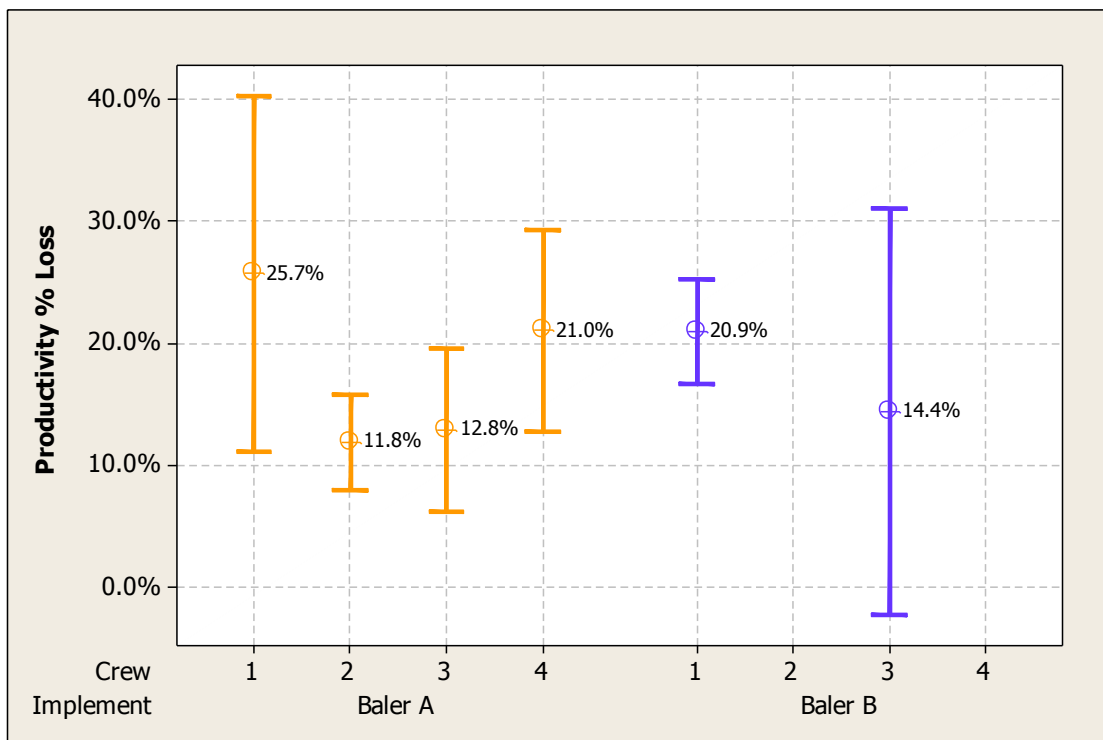
Figure 20: Baler Productivity Loss by Crew

Digging deeper into the differences between crews the begin and finish times for each crew were evaluated. Figure 21 shows the breakdown of this separated by crew and by baler. Crew 3 had the shortest average begin time at 17 minutes, while crew 2 had the shortest average finish time at 16.3 minutes.

Figure 21: Begin and Finish Time For Balers by Crew

Results and Conclusion

There are two main operations within a harvest team, windrowing and baling. It has been shown that the organization and management of these operations has a major impact on the productivity of the harvest team, as well as the productivity of the equipment.

It was shown that the productivity losses vary greatly by crew, which confirms that these issues are related to crew organization. Crews with poor organization had lo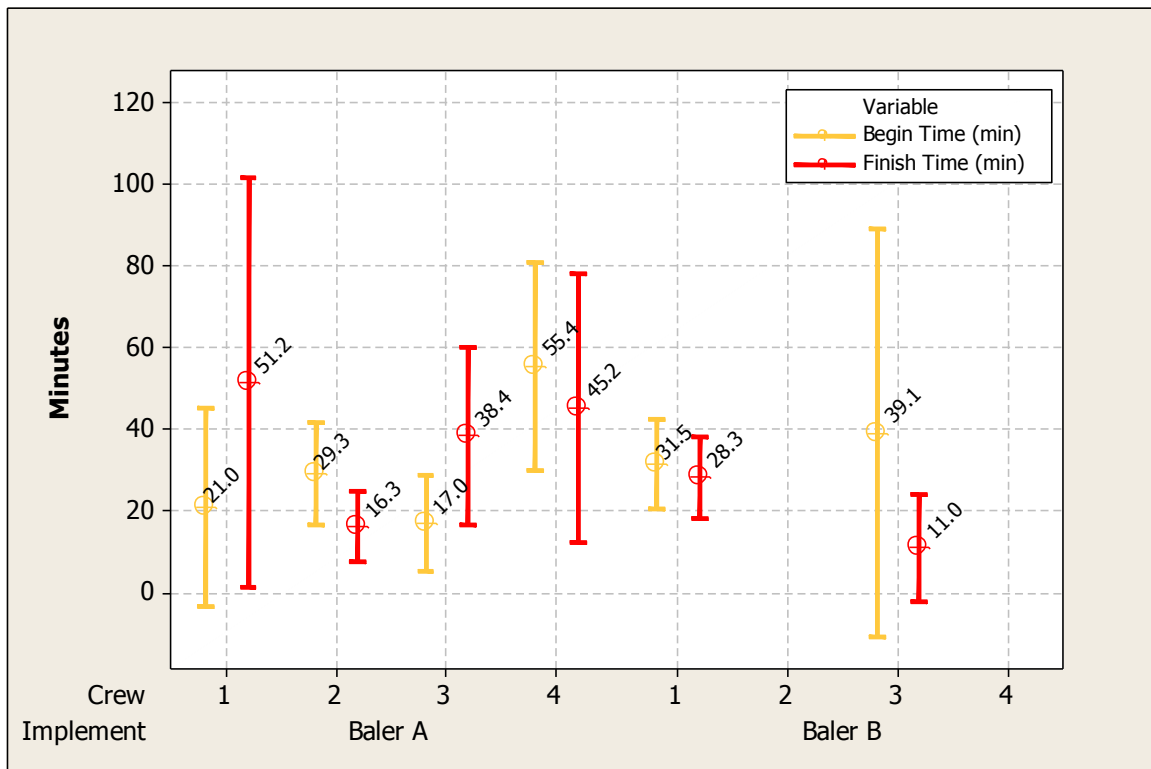nger begin and finish times and therefore higher productivity loss. The amount of productivity loss translates into possible productivity gains, which can be achieved with better organization of the crews.

By separating the windrowing and baling operations and staggering them, there can be significant increases in the overall productivity of the harvest team. A 5-6 hour lead time should be maintained between the two operations. This is long enough to cushion the effect of any mechanical problems or transportation and transition times of the windrower, but yet short enough that if there is the possibility of a rain event that the windrowers could be shut down and give the balers time to catch up and finish the areas that have already been windrowed.

This analysis showed the impact that harvest team organization has on productivity, and how much it can be improved by separating the two operations. Further analysis could be done on the amount of idle time that occurs during the active period which this analysis did not cover. There could be more potential productivity gains to be made from better understanding these events as well.

References

Askey, Jeffery C., Matthew Darr, Keith Webster, Benjamin Covington, and Jeremy Brue.
Automated Logistics Processing of GIS Data for Agricultural Harvest Equipment. Paper.
no. #1596410, 2013 ASABE Annual International Meeting

Covington, Benjamin R. "Assessment of Utilization and Downtime of a Commercial Level Multi-
pass Corn Stover Harvesting System." Thesis. Iowa State University, 2013. Print.

Peyton, Kevin S. "Geographic Information System Tools for the Analysis of Commercial Level
Multi-pass Corn Stover Harvesting Systems." Thesis. Ames, Iowa: Iowa State University,
2012. Print.

SMS Advanced. Vers. 12.0. Ames, IA: AgLeader, 2012. Computer software.

CHAPTER 4:

GENERAL CONCLUSIONS

As the world population and demand for energy continues to grow renewable fuels and biobased sources of energy will play a large role in meeting these increasing needs. In the U.S., the governments Renewable Fuels Standard and RFS2 have mandated the development and production of biobased fuels. The first and most prevalent of these was corn grain ethanol. It has been widely adopted with plants in operation throughout the Midwest.

More recently the focus has turned the cellulose based feed stocks that can also be converted into ethanol. Here in central Iowa corn stover has been selected as the primary feedstock. It is readily available and very abundant in the high yielding corn fields of central Iowa.

As cellulosic ethanol begins production in central Iowa and throughout the Midwest the supply chain to support it has grown significantly. A 25 million gallon per year plant will require over 740,000 large square bales per year. With short harvest windows in the fall it takes a large fleet of equipment and personal to accomplish a harvest of this scale.

Mobile data logging devices have played a key role in helping companies and research institutes better understand the productivity and efficiencies of these corn stover harvesting crews. By collecting and analyzing data about machine logistics, researchers are better able to provide solutions to managers and producers about how to improve the organization and productivity of the entire harvesting team.

Chapter 2 "Impact of Sampling Interval on GIS Data based Productivity Analysis" looked at the data logging devices used to capture this data. The data is collected from the equipment's vehicle CAN bus and then evaluated to determine its productivity. To evaluate the equipment in a corn stover harvesting operation the data is sorted and categorized based on three key machine parameters, engine speed, vehicle ground speed, and pto speed. Different combinations of these parameters determine whether the machine is in the Production, Idle, or Transport state. From this the machine's productivity can be calculated.

This chapter looked at increasing the sampling interval of the data loggers as a solution to decreasing the data log file sizes. Smaller file sizes are needed to decrease the risk of lost or corrupted data during data transfer operations. After evaluating alternative sampling intervals of 2, 5, 10, 15, 30, and 60 seconds it was found that 15 seconds was the optimal sampling interval. The 15 second interval was the longest interval that did not significantly affect the productivity calculations of the machines. The 15 second interval will yield a 93% decrease in the amount of data generated and is successful solution to decreasing the data log file size.

Chapter 3 "Optimal Configuration and Lead Time of Corn Stover Biomass Harvesting Teams" looks at another aspect of increasing the productivity of corn stover biomass harvesting teams. This chapter was focused around the organization of harvest teams and the impact it has the productivity of the entire team as well as individual machines. The analysis was completed in 2011 and repeated in 2012 on a larger data set.

The windrowing and baling operations of corn stover harvesting are interconnected. The baling operation is dependent on the windrowing and any decreases in the productivity of the windrowers also affect the balers. This dependency also creates other inefficiencies in the

harvesting operation. If the units travel between fields together the balers have to wait for the shredders to complete an area of the field before they can begin work. Also since they have a higher capacity they will eventually catch the windrowers. This time spent waiting at the beginning of a field is productivity loss for the 2012 harvest they had an average productivity loss of 16%. This can be overcome with better organization.

It was shown that the windrowing and baling operations need to be separated and that the windrowers should maintain a lead time over the balers. This lead time should be 5-6 hours. This is a long enough time to overcome any mechanical breakdowns and the transition time of the shredders. It is also short enough that in the case of a potential rain event the windrowers can be stopped and the balers would have time complete the ground they have already covered.

This lead time only works if operational capacities are equalized. The baler field capacity was found to be 62% greater than that of the windrower. In order to equalize these capacities the optimal harvest team configuration would consist of 3 windrowers and 2 balers. In this configuration, the capacities are nearly equalized and the 5-6 hour lead time can be maintained.

Limitations of Results

There are several limitations to the results of this study that should be taken into consideration. These results are only applicable in a commercial scale corn stover harvest that utilizes similar equipment and operational organization. If advancements in machine capacities are made this study will need be revisited and possibly modified. This study is also geographically specific. These results may shift and not hold true in other regions of the country

or world. This is due to differences in field sizes and other environmental factors such as terrain or cropping style. These items must be considered when applying these results to other harvesting operations.

There are many more potentials for improving the productivity of biomass harvesting teams and the results of this study are another step in the right direction. Continued research could further increase overall productivities, driving down the cost of the feedstock supply chain. This makes cellulosic ethanol an economic and environmentally viable solution for our nation's energy needs and helps decrease our dependence on foreign oil.

APPENDIX A: PERCENT PRODUCTIVITY DIFFERENCE DATA

Boxplot of Productivity % Difference

Test for Equal Variances for Idle % Difference

Test for Equal Variances for Transport % Difference

Bartlett's Test	
Test Statistic	1134.40
P-Value	0.000
Levene's Test	
Test Statistic	44.88
P-Value	0.000

95% Bonferroni Confidence Intervals for StDevs
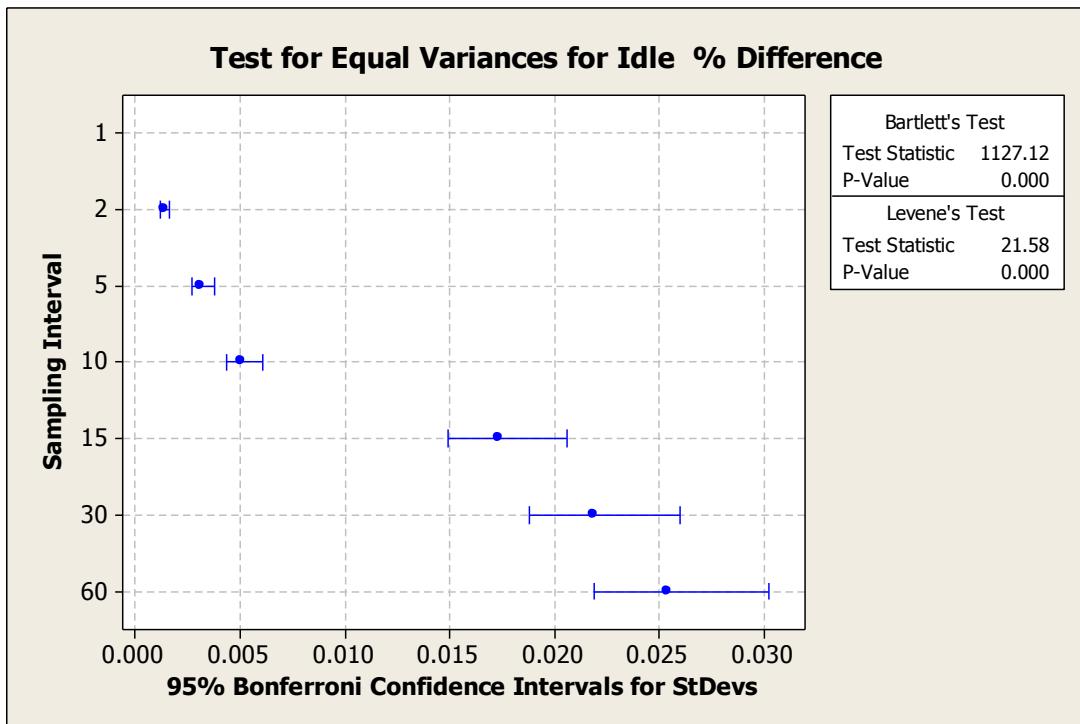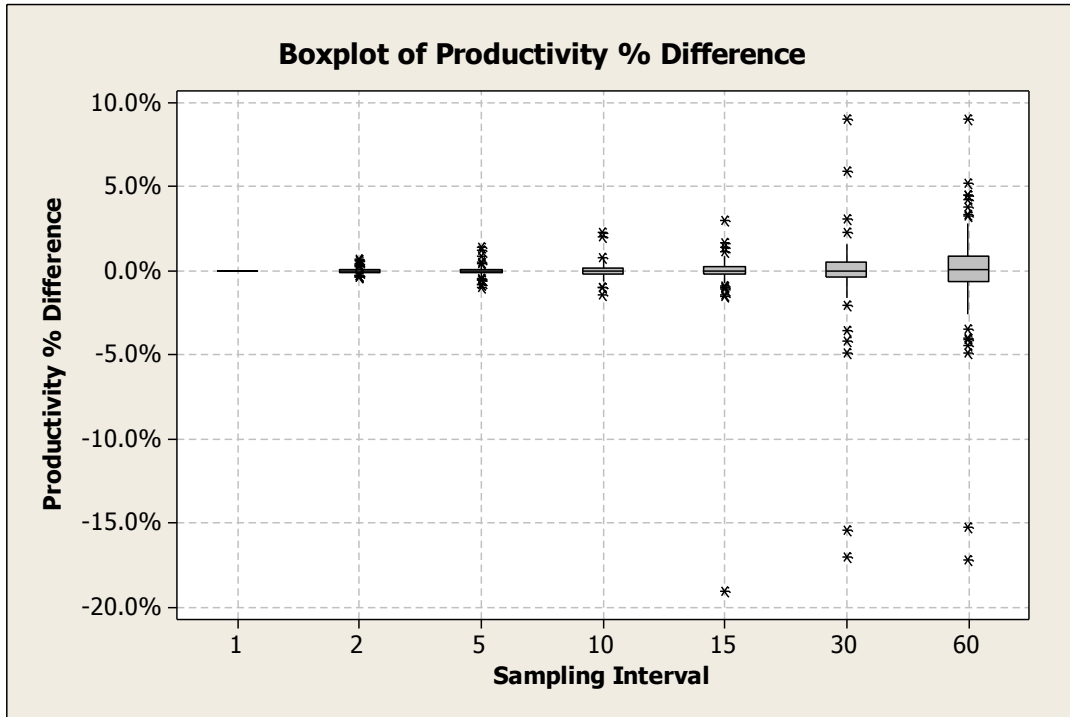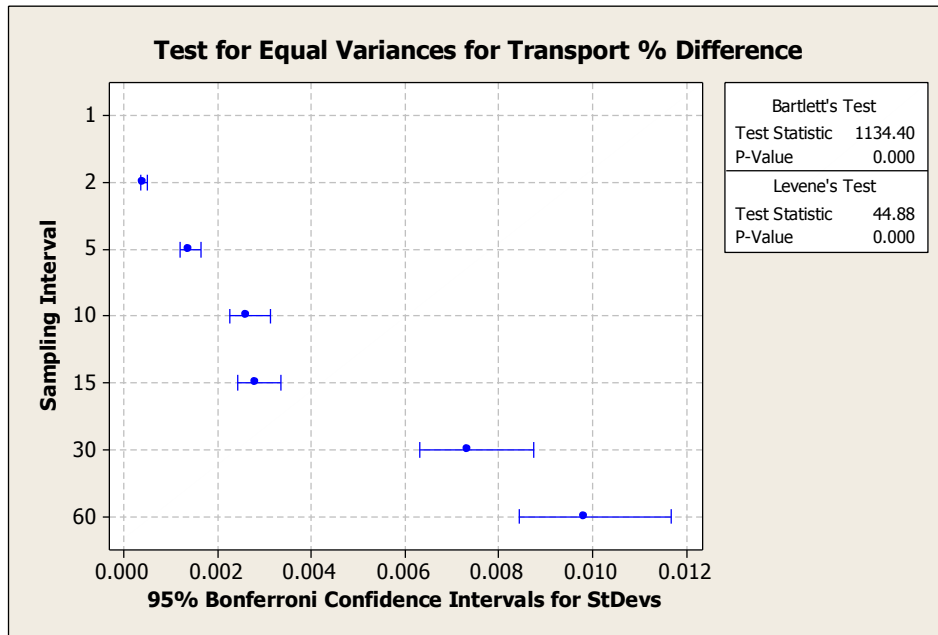
One-way ANOVA: StDev versus Sampling Interval

```
Source            DF   SS        MS        F     P
Sampling Interval  6  0.0012931  0.0002155  6.24  0.002
Error             14  0.0004837  0.0000345
Total             20  0.0017768
```

S = 0.005878   R-Sq = 72.78%   R-Sq(adj) = 61.11%

```
                   Individual 95% CIs For Mean Based on
                   Pooled StDev
Level N   Mean     StDev    -------+---------+---------+---------+--
 1    3  0.000000  0.000000  (------*------)
 2    3  0.001078  0.000558  (------*------)
 5    3  0.002462  0.000947   (------*-------)
10    3  0.004048  0.001274    (------*------)
15    3  0.012443  0.008338          (------*-------)
30    3  0.017695  0.009011            (-------*------)
60    3  0.020645  0.009396              (-------*------)
                            -------+---------+---------+---------+--
                             0.000    0.010     0.020     0.030
```
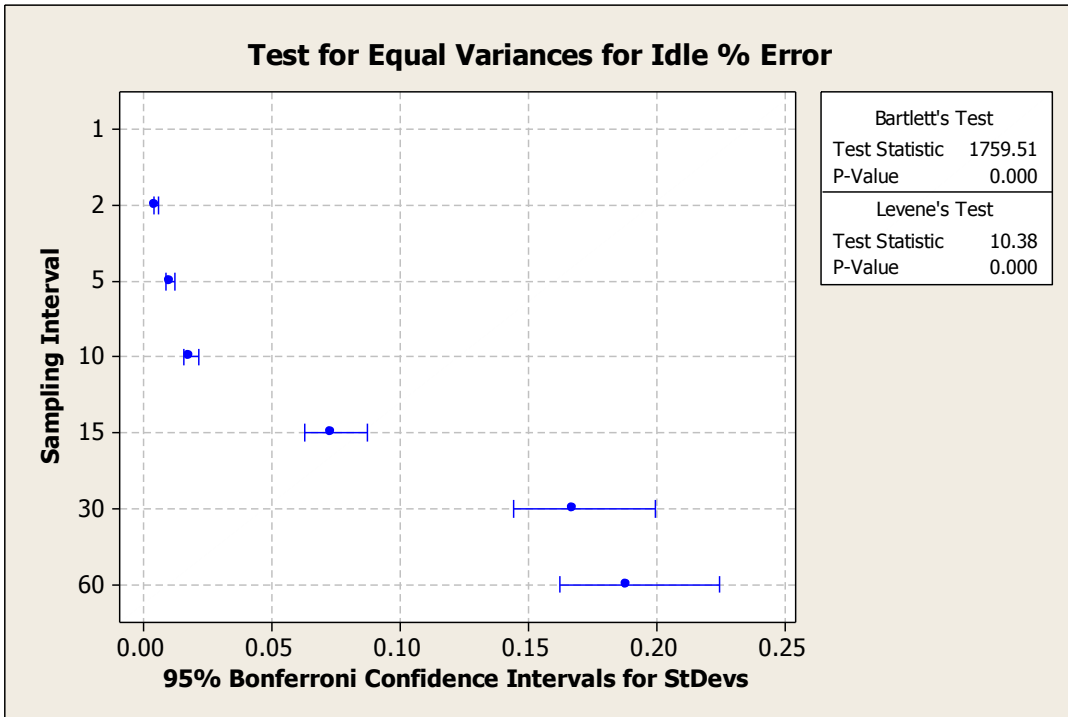
Pooled StDev = 0.005878

Grouping Information Using Tukey Method

```
Sampling
Interval N    Mean  Grouping
60     3  0.020645  A
30     3  0.017695  A B
15     3  0.012443  A B C
10     3  0.004048    B C
 5     3  0.002462    B C
 2     3  0.001078      C
 1     3  0.000000      C
```

Means that do not share a letter are significantly different.

APPENDIX B: TIME BASED PRODUCTIVITY DATA

Boxplot of Production % Error

Test for Equal Variances for Idle % Error

Test for Equal Variances for Transport % Error

One-way ANOVA: Std Deviation versus Sampling Interval

```
Source             DF    SS       MS      F     P
Sampling Interval   6  0.17835  0.02973  9.16  0.000
Error              14  0.04545  0.00325
Total              20  0.22380
```

S = 0.05698   R-Sq = 79.69%   R-Sq(adj) = 70.99%

```
                    Individual 95% CIs For Mean Based on
                    Pooled StDev
Level N   Mean     StDev   -------+---------+---------+---------+--
1     3  0.00000  0.00000  (------*------)
2     3  0.00905  0.00910  (------*------)
5     3  0.02244  0.02552   (------*------)
10    3  0.03535  0.03874     (-------*------)
15    3  0.07211  0.05981        (------*------)
30    3  0.18817  0.03478                  (------*------)
60    3  0.25614  0.12532                       (------*------)
                           -------+---------+---------+---------+--
                            0.00     0.10      0.20      0.30
```
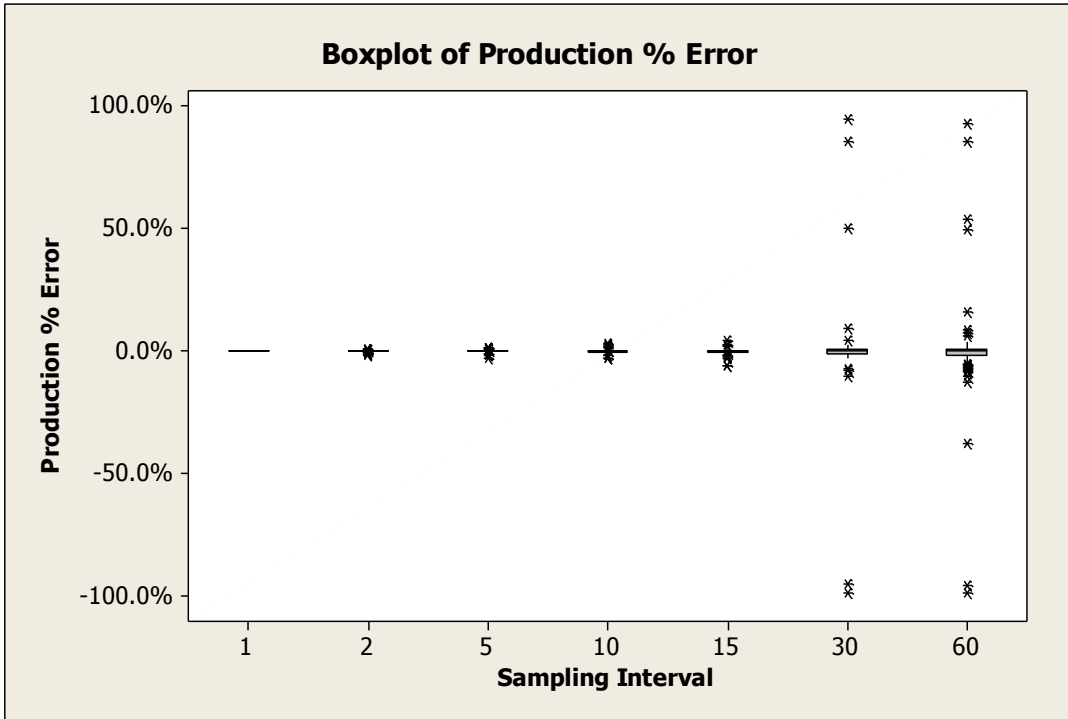
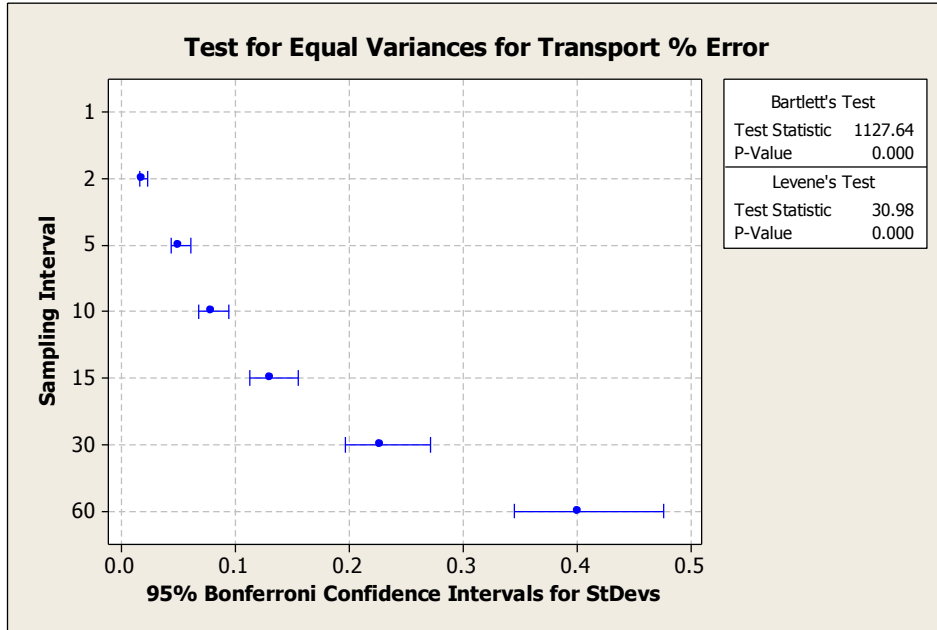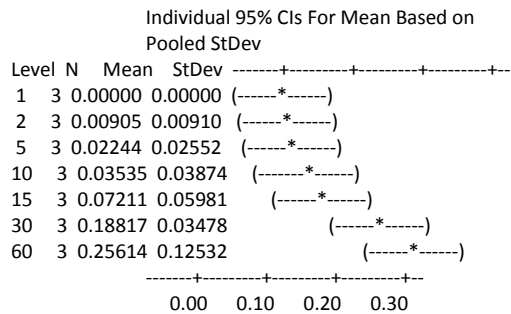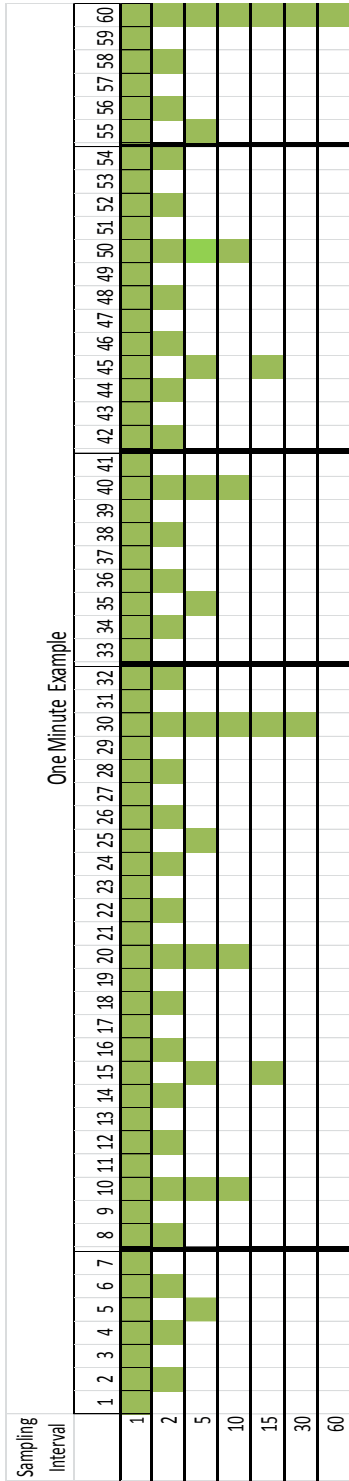Pooled StDev = 0.05698

Grouping Information Using Tukey Method

```
Sampling
Interval N    Mean  Grouping
60       3  0.25614  A
30       3  0.18817  A B
15       3  0.07211  B C
10       3  0.03535  B C
5        3  0.02244    C
2        3  0.00905    C
1        3  0.00000    C
```

Means that do not share a letter are significantly different.

APPENDIX C: SHIFTING TRANSITION POINT DOT PLOT

ACKNOWLEDGEMENTS

I would first like to thank our Heavenly Father, it is by your wonderful grace and mercy that we are able to serve you and honor you in all that we do.

I would like to thank DuPont Industrial Biosciences for funding this research project and giving me the opportunity to be on the forefront of a revolutionary industry. I would also like to thank my Professor Dr. Matt Darr for giving me the opportunity to work alongside you on this project and pushing the limits of agricultural technology. The knowledge and skills that you have taught me during my time at Iowa State will last me a lifetime.

I would like to thank my committee member Dr. Brian Steward for your continued support throughout my academic career. The skills and knowledge learned in you undergraduate classes have paid dividends during my graduate career. I would also like to thank Dr. Tom Brumm of my committee. Your support and guidance throughout my academic career at Iowa State has been outstanding and I will forever be grateful.

I owe a big thanks to the research team that worked on this project, especially Benjamin Covington, Brittany Schon, Jeff Askey, Kevin Peyton, Keith Webster, Ryan Bergman, and Drew Sprangers. Without you guys none of this would have happened. We have worked against all the odds sometimes and thanks to the perseverance and commitment from every one of you, we always managed to get the job done. It has been my pleasure to work alongside each of you and I am truly grateful to call you my friends.

Finally I would like to thank my family. You have been there since the start and it is because of you that I am where I am today. Your love and support has inspired me to be the

person I am today. A special thanks goes to my beautiful fiancée, I am so thankful to have you in my life. I can't thank you enough for your patience and understanding during my research; I love you and can't wait to start our adventure as husband and wife.