



Michigan Technological University  
Create the Future Digital Commons @ Michigan Tech

---

Dissertations, Master's Theses and Master's Reports - Open

Dissertations, Master's Theses and Master's Reports

---

2012

## Analysis of spring break-up and its effects on a biomass feedstock supply chain in northern Michigan

Huihui Lin  
*Michigan Technological University*

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>

 Part of the [Civil and Environmental Engineering Commons](#)

Copyright 2012 Huihui Lin

---

### Recommended Citation

Lin, Huihui, "Analysis of spring break-up and its effects on a biomass feedstock supply chain in northern Michigan ", Master's Thesis, Michigan Technological University, 2012.  
<https://digitalcommons.mtu.edu/etds/252>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etds>

 Part of the [Civil and Environmental Engineering Commons](#)

ANALYSIS OF SPRING BREAK-UP AND ITS EFFECTS ON A BIOMASS  
FEEDSTOCK SUPPLY CHAIN IN NORTHERN MICHIGAN

By

Huihui Lin

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Environmental Engineering

MICHIGAN TECHNOLOGICAL UNIVERSITY

2012

© 2012 Huihui Lin

This thesis has been approved in partial fulfillment of the requirements for the Degree of  
MASTER OF SCIENCE in Environmental Engineering.

Department of Civil and Environmental Engineering

Thesis Advisor: *David Watkins*

Committee Member: *Veronica Griffis*

Committee Member: *James Pickens*

Department Chair: *David Hand*

## TABLE OF CONTENTS

LIST OF FIGURES .....	v
LIST OF TABLES .....	vii
ACKNOWLEDGMENTS .....	viii
ABSTRACT .....	ix
<b>CHAPTER 1. INTRODUCTION .....</b>	<b>1</b>
1.1. Background.....	1
1.2. Objectives and Significance .....	3
<b>CHAPTER 2. BACKGROUND: PROJECT INTRODUCTION.....</b>	<b>4</b>
2.1. CoEE Project Overview .....	4
2.2. CoEE Optimization Model .....	5
2.3. CoEE Simulation Model .....	5
<b>CHAPTER 3. SPRING BREAK-UP DATA.....</b>	<b>8</b>
3.1. Introduction .....	8
3.2. Data Sources .....	10
3.3. Factors Affecting Spring Break-up.....	11
<b>CHAPTER 4. SPRING BREAK-UP MODEL.....</b>	<b>17</b>
4.1. Introduction .....	17
4.2. Preliminary Data Analysis.....	17
4.3. Methodology .....	19
4.4. Model for Alcona County .....	22
4.5. Spatial Correlation .....	29
4.5.1. Simple Linear Regression .....	29
4.5.2. Reduced Major Axis Regression (RMA) .....	34

<b>CHAPTER 5. SIMULATION RESULTS .....</b>	<b>40</b>
5.1. Verification of Spring Break-up Model.....	40
5.2. Uncertainty Effects on Supply Chain Simulations .....	42
5.3. Impact of Rail Use and Rail Yard Storage.....	46
5.4. Variability in Simulation Results .....	51
<b>CHAPTER 6. CONCLUSION AND FUTURE WORK.....</b>	<b>53</b>
6.1. Conclusion.....	53
6.2. Recommendations for Future Work.....	53
<b>REFERENCES.....</b>	<b>55</b>
<b>APPENDIX.....</b>	<b>61</b>
Appendix A: Historical Spring Break-up Data .....	61
Appendix B: Mathematica Code for Simulation Results vs. Historical Data.....	62
Appendix C: Transformations and simple linear regression relationships.....	63
Appendix D: Mathematica Code used to obtain RMA regression equation .....	72
Appendix E: Transformations and RMA regression relationships .....	75

## List of Figures

FIGURE 1.1: COMPONENTS OF THE FEEDSTOCK SUPPLY CHAIN .....	2
FIGURE 2.1: HARVESTING REGION FOR FEEDSTOCK CoEE PROJECT .....	4
FIGURE 2.2: SCHEMATIC OF THE CoEE SIMULATION MODEL LOGIC .....	6
FIGURE 3.1 MAP OF ALCONA COUNTY SHOWING ROADS OF DIFFERENT TYPES .....	9
FIGURE 3.2: TYPICAL TREND OF SPRING BREAK-UP IN THE HARVESTING REGION .....	10
FIGURE 3.3: MAP OF CHARLEVOIX COUNTY AND ILLUSTRATION OF THE PROCEDURE FOR DEVELOPING AREA WEIGHTS OF STATIONS .....	14
FIGURE 4.1: HISTOGRAM OF SPRING BREAK-UP START DAY FOR (A) L.P. COUNTIES AND (B) U.P. COUNTIES.	18
FIGURE 4.2: HISTOGRAM OF SPRING BREAK-UP DURATION FOR (A) L.P. COUNTIES AND (B) U.P. COUNTIES ..	20
FIGURE 4.3: HISTOGRAM OF SPRING BREAK-UP END DAY FOR (A) L.P. COUNTIES AND (B) U.P. COUNTIES.....	20
FIGURE 4.4: OVERVIEW OF STUDY METHODOLOGY .....	21
FIGURE 4.5: MAP OF AVERAGE HISTORICAL START DAYS OF SPRING BREAK-UP AND YEARS OF RECORD .....	23
FIGURE 4.6: THE BOX-COX TRANSFORMATION PLOT OF ALCONA COUNTY’S START DAY .....	24
FIGURE 4.7: THE NORMAL DISTRIBUTION ESTIMATED FOR THE TRANSFORMED START DAY OF ALCONA COUNTY .....	25
FIGURE 4.8: SIMULATED END DAY (WITHOUT PRE-TRANSFORMATION) VS. HISTORICAL END DAY FOR ALCONA COUNTY .....	26
FIGURE 4.9: SIMULATED DATA VS. HISTORICAL START DAY (A) AND END DAY (WITH PRE-TRANSFORMATION) (B) FOR ALCONA COUNTY .....	28
FIGURE 4.10: LINEAR REGRESSION MODEL OF ANTRIM COUNTY TRANSFORMED START DAY VS. ALCONA COUNTY TRANSFORMED START DAY .....	31
FIGURE 4.11: REGIONAL GROUPINGS AND CORRELATION COEFFICIENTS OF TRANSFORMED SPRING BREAK-UP START DAY. ....	33
FIGURE 4.12: REGIONAL GROUPINGS AND CORRELATION COEFFICIENTS OF TRANSFORMED END DAYS OF SPRING BREAK-UP .....	34
FIGURE 4.13: RMA REGRESSION SCHEMATIC, USING THE EXAMPLE OF TRANSFORMED SPRING BREAK-UP START DAYS IN ANTRIM COUNTY AND ALCONA COUNTY. ....	35
FIGURE 4.14: COMPARISON OF SIMPLE LINEAR REGRESSION AND RMA REGRESSION .....	36
FIGURE 4.15: REGRESSION LINES RMA REGRESSION SCHEMATIC .....	37
FIGURE 4.16: EXAMPLE OF SIMULATING THE SPRING BREAK-UP START DAY OF CHIPPEWA COUNTY .....	39
FIGURE 5.1: SIMULATED START DAY OF CHARLEVOIX COUNTY FROM ARENA VS. HISTORICAL DATES .....	40
FIGURE 5.2: SIMULATED END DAY OF BENZIE COUNTY FROM ARENA VS. HISTORICAL DATES .....	41

FIGURE 5.3: SIMULATED DURATION OF CHIPPEWA COUNTY FROM ARENA VS. HISTORICAL DATES .....	41
FIGURE 5.4: COMPARISON OF TOTAL COST FOR DETERMINISTIC AND RANDOM SPRING BREAK-UP SCENARIOS .....	45
FIGURE 5.5: COMPARISON OF TOTAL ENERGY CONSUMPTION FOR DETERMINISTIC AND RANDOM SPRING BREAK-UP SCENARIOS.....	45
FIGURE 5.6: COMPARISON OF TOTAL EMISSION FOR DETERMINISTIC AND RANDOM SPRING BREAK-UP SCENARIOS .....	46
FIGURE 5.7: COMPONENTS OF TOTAL COST, TOTAL FUEL CONSUMPTION AND TOTAL GREENHOUSE GAS EMISSIONS FOR THE BASELINE SCENARIO.....	50
FIGURE C.1: RELATIONSHIP BETWEEN TRANSFORMED START DAYS OF ALPENA COUNTY AND ALCONA COUNTY .....	68
FIGURE C.2: RELATIONSHIP BETWEEN TRANSFORMED START DAYS OF PRESQUE ISLE COUNTY AND ALCONA COUNTY .....	69
FIGURE C.3: RELATIONSHIP BETWEEN TRANSFORMED START DAYS OF BENZIE COUNTY AND ALCONA COUNTY .....	69
FIGURE C.4: MAP OF LEELANAU COUNTY, ANTRIM COUNTY, MANISTEE COUNTY AND WEXFORD COUNTY	70

**List of Tables**

TABLE 3.1: CORRELATION COEFFICIENTS OF SPRING BREAK-UP TIMING AND TEMPERATURE ARE PRESENTED .....13

TABLE 3.2: CORRELATION COEFFICIENTS OF SPRING BREAK-UP TIMING AND PRECIPITATION ARE PRESENTED. ....15

TABLE 4.1: TO PRESENT ESTIMATION RESULTS OF ALCONA COUNTY’S SPRING BREAK-UP TIMING .....27

TABLE 4.2: TO PRESENT SAMPLE MEANS AND STANDARD DEVIATION FOR ALCONA COUNTY SPRING BREAK-UP TIMING .....29

TABLE 5.1: TO PRESENT MEANS AND STANDARD DEVIATIONS FOR THE START DAY OF CHARLEVOIX COUNTY .....40

TABLE 5.2: TO PRESENT MEANS AND STANDARD DEVIATIONS FOR THE END DAY OF BENZIE COUNTY .....41

TABLE 5.3: TO PRESENT MEANS AND STANDARD DEVIATIONS FOR THE DURATION OF CHIPPEWA COUNTY ...42

TABLE 5.4: TO PRESENT HISTORICAL SPRING BREAK-UP STATISTICS .....43

TABLE 5.5: TO PRESENT SUMMARY OF SIMULATION RESULTS .....44

TABLE 5.6: TO PRESENT SUMMARY OF SIMULATION RESULTS FOR RAIL YARD CAPACITY REDUCTION SCENARIOS .....48

TABLE 5.7: TO PRESENT SUMMARY OF SIMULATION RESULTS FOR RAIL USE REDUCTION .....50

TABLE 5.8: COMPARISON OF MODEL COMPUTATIONAL TIME IS PRESENTED. ....51

TABLE 5.9: RESULTS OF SIMULATION OF 300 REPLICATIONS ARE COMPARED .....52

TABLE A.1: HISTORICAL SPRING BREAK-UP DATA INCLUDES START DAY AND DURATION.....60

TABLE C.1: RESULTS OF THE BOX-COX TRANSFORMATION FOR SPRING BREAK-UP TIMING INCLUDE TRANSFORMATION PARAMETER AND DISTRIBUTION AFTER TRANSFORMATION. ....63

TABLE C.2: RELATIONSHIPS FOR TRANSFORMED START DAY OF SPRING BREAK-UP ARE DEVELOPED BY SIMPLE LINEAR REGRESSION. ....65

TABLE C.3: RELATIONSHIPS FOR TRANSFORMED END DAY OF SPRING BREAK-UP ARE DEVELOPED BY SIMPLE LINEAR REGRESSION. ....67

TABLE D.1: SUMS OF SQUARES AND OTHER QUANTITIES FOR TRANSFORMED START DAYS OF ALCONA COUNTY AND ANTRIM COUNTY ARE LISTED .....72

TABLE E.1: RELATIONSHIPS FOR TRANSFORMED START DAYS OF SPRING BREAK-UP ARE DEVELOPED BY RMA. ....75

TABLE E.2: RELATIONSHIPS FOR TRANSFORMED END DAYS OF SPRING BREAK-UP ARE DEVELOPED BY RMA. 76



# Acknowledgments

This Center of Energy Excellence (CoEE) research project was funded by the Michigan Economic Development Corporation and Frontier Renewable Resources, Inc. The project was conducted by Michigan Technological University and Michigan State University. I would like to express my gratitude to my advisor, Dr. David Watkins, for providing advice, guidance, inspiration and support through my two-year study towards this master's degree, and also for his detailed reviews and constructive comments on the manuscript. I really appreciate the assistance from my other committee members, Dr. Veronica Griffis and Dr. Jim Pickens, on this thesis. I am also grateful for the support from the optimization model team: Dr. Jim Friendewey, Dr. Jim Pickens, Chris Miller, and Eric Dunnack. Also, I am indebted to Dr. Dana Johnson, Dr. Greg Graman, Dr. Barry Solomon, Timothy Jenkins, Dr. Pasi Lautala and Dr. Robert Handler for their help on this project. Spring break-up data were compiled from county road commissions by Dr. William Knudson of Michigan State University and from the Michigan Department of Transportation by Dr. Mark Johnson of Michigan Technological University. Last but not least, I would like to thank the post-doctoral research assistants Haihong Huang and Yuejun Yin, who worked on the Arena simulation model for about one year before I took over.

# Abstract

Demand for bio-fuels is expected to increase, due to rising prices of fossil fuels and concerns over greenhouse gas emissions and energy security. The overall cost of biomass energy generation is primarily related to biomass harvesting activity, transportation, and storage. With a commercial-scale cellulosic ethanol processing facility in Kinross Township of Chippewa County, Michigan about to be built, models including a simulation model and an optimization model have been developed to provide decision support for the facility. Both models track cost, emissions and energy consumption. While the optimization model provides guidance for a long-term strategic plan, the simulation model aims to present detailed output for specified operational scenarios over an annual period. Most importantly, the simulation model considers the uncertainty of spring break-up timing, i.e., seasonal road restrictions. Spring break-up timing is important because it will impact the feasibility of harvesting activity and the time duration of transportation restrictions, which significantly changes the availability of feedstock for the processing facility.

This thesis focuses on the statistical model of spring break-up used in the simulation model. Spring break-up timing depends on various factors, including temperature, road conditions and soil type, as well as individual decision making processes at the county level. The spring break-up model, based on the historical spring break-up data from 27 counties over the period of 2002-2010, starts by specifying the probability distribution of a particular county's spring break-up start day and end day, and then relates the spring break-up timing of the other counties in the harvesting zone to the first county. In order to estimate the dependence relationship between counties, regression analyses, including standard linear regression and reduced major axis regression, are conducted.

Using realizations (scenarios) of spring break-up generated by the statistical spring break-up model, the simulation model is able to probabilistically evaluate different harvesting and transportation plans to help the bio-fuel facility select the most effective strategy. For

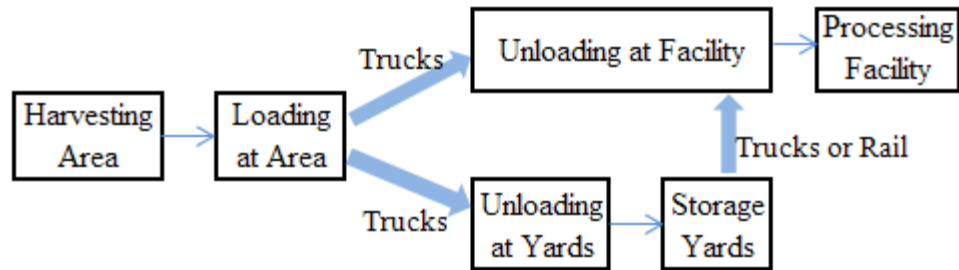
early spring break-up, which usually indicates a longer than average break-up period, more log storage is required, total cost increases, and the probability of plant closure increases. The risk of plant closure may be partially offset through increased use of rail transportation, which is not subject to spring break-up restrictions. However, rail availability and rail yard storage may then become limiting factors in the supply chain. Rail use will impact total cost, energy consumption, system-wide CO<sub>2</sub> emissions, and the reliability of providing feedstock to the bio-fuel processing facility.

# Chapter 1. Introduction

## 1.1. Background

Biomass is one of the renewable energy sources (including solar, wind, tides and geothermal heat). Biomass is popular and important, because it is easily stored compared to solar or wind energy, it can be generated in many locations compared to tides and geothermal heat, and it can be transported using currently available vehicles and infrastructure. Bio-fuel is a type of fuel, usually liquid, which is derived from biomass that is grown, harvested and replaced on a regular basis. The demand for bio-fuel is increasing mainly because of the shortage and high price of fossil fuels and the concerns brought on by greenhouse gas emissions in recent years. Meanwhile, the academic community has been increasingly interested in the bio-fuel industry as well. Some recent studies have focused on the influences of the raw material supply chain on the overall system economics (1) and energy performance (2), the management and value of waste biomass (3), the importance of bio-fuel for its positive effects on energy security issues, environmental concerns, foreign exchange savings, and socioeconomic issues related to the rural area (4), and the technology of conversion from biomass to designed bio-fuels (5, 6), among others.

A bio-fuel feedstock supply chain is comprised of several processes. These processes include harvesting, loading, storage, processing and transporting, as shown in Figure 1.1. The feedstock supply chain aims to achieve low-cost, time-effective, and reliable delivery of input to bio-fuel processing facilities. More specifically, the supply chain includes the collection activities of biomass in harvesting areas, loading and unloading of trucks and/or rail cars, possibly storage at multiple storage yards, and finally transportation to a facility for production. Typically, trucks are used for transporting biomass for short distances, and rail is used for long distances.



**Figure 1.1.** Components of the Feedstock Supply Chain (Adapted from "Feedstock Supply Chain Models for the Frontier Renewable Resources Bio-fuel Facility in Kinross, Michigan", 2012 (7))

The total cost of all the processes above can be identified as the total cost of the supply chain system. This cost component is considered in this thesis, neglecting the cost of operating the bio-fuel processing facility itself. The total supply chain system cost includes harvesting cost, loading and unloading costs, storage costs of both the facility yard and other log yards, and transport costs of both truck and rail. Similarly, carbon emissions are considered during the log harvesting, transportation, and storage activities. The unit for carbon emissions is CO<sub>2</sub>-equivalent greenhouse gas emissions per unit of biomass (e.g., kg/ton). Energy consumption associated with log harvesting, transportation, and storage yard activities is also considered. The unit for energy consumption is the energy consumed per unit of biomass (e.g., MJ/ton).

One of the nation's first commercial scale cellulosic ethanol processing facilities is planned to be built by Frontier Renewable Resources (Frontier) in the Township of Kinross in Chippewa County, Michigan. Frontier is owned by Mascoma Corporation, a renewable fuels company, and J.M. Longyear (8), a Michigan-based national leader in natural resources management. The planned ethanol plant will produce cellulosic fuels from non-food biomass feed stocks, e.g., hardwood pulpwood. As of September 2012, the Kinross facility is scheduled to open in 2014 at a Phase I capacity of 20 million gallons per year, with potential to expand in subsequent years (9).

## **1.2. Objectives and Significance**

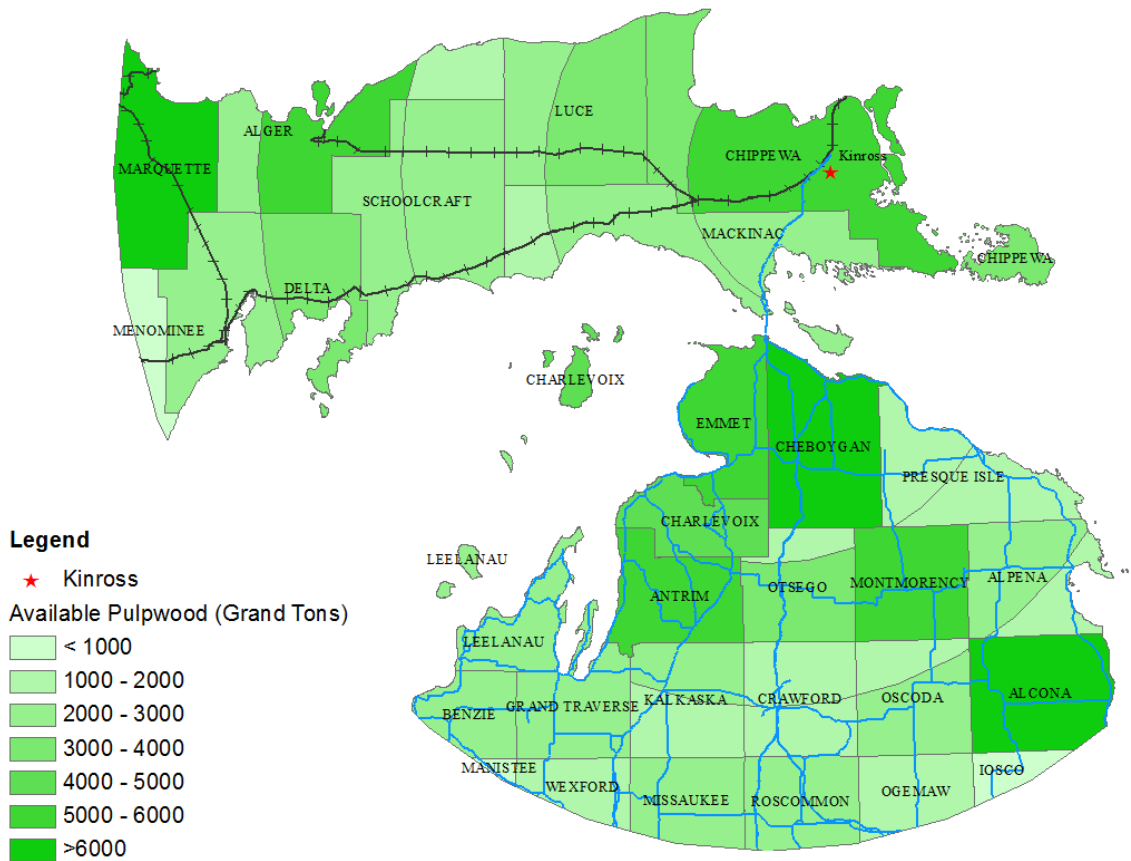
As is common with other bio-fuel supply chain analysis efforts (10, 11), this thesis will present a number of economic and environmental performance measurements relevant to strategic and operational planning. Unlike previous efforts, this study does the analysis considering uncertainty of the supply chain system resulting from environmental conditions. Specifically, a statistical model of spring break-up is developed for use in a feedstock supply chain simulation model. The spring break-up model is based on historical data at the county level, and an approach is developed to represent the spatial correlation of spring break-up timing across the region. Using scenarios generated by this model, supply chain simulation results are compared to results without randomness.

It is shown that simulation results with uncertainty better represent expected supply chain performance and allow users to better understand the possible results of contract and resource allocation decisions.

# Chapter 2. Background: Project Introduction

## 2.1. CoEE Project Overview

The Feedstock Supply Chain Center of Energy Excellence (Feedstock CoEE) project aims to support the technical needs of Frontier. Efforts have focused on collecting data on biomass availability in the northern Michigan harvest region (as shown in Figure 2.1), estimating life-cycle impacts of the use phase of various harvesting and transportation technologies, and developing simulation and optimization models to support harvest and logistics decision making (7). A summary of the optimization and simulation models is presented herein.



**Figure 2.1:** Harvesting region for Feedstock CoEE project (data from MDOT)

## 2.2. CoEE Optimization Model

The optimization model is developed to identify the most efficient combination of harvesting and transportation options to supply the facility. It is created to evaluate a 20-year planning horizon, with the decision variables of how much to harvest from each harvest area and ownership type (e.g., industrial, private non-industrial, state, and federal) in each year of the planning horizon (7). The model is formulated as a large-scale linear (network) programming model, solved using the FICO® Xpress Optimization Suite. Results of the optimization model support long-term strategic planning decisions, such as harvesting contracts and storage yard locations. A method has also been developed to translate the annual decisions prescribed by the optimization model into seasonal, monthly, or weekly harvest and transportation decisions for input to the simulation model.

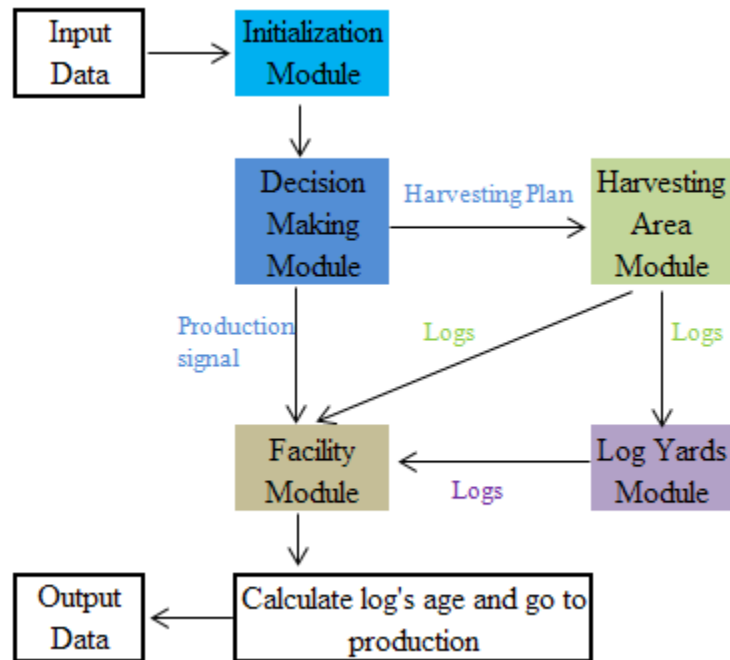
## 2.3. CoEE Simulation Model

While the optimization model provides guidance for an overall strategic plan, the simulation model gives annual detailed output for specified operational scenarios using a daily time step. Its main refinements include modeling spring break-up timing and the transportation and inventory decisions prior to and during spring break-up. Both the simulation and optimization modeling have made the same supply area assumptions -- the supply area is split into 43 harvest areas based on the intersection of 29 counties and 30-mile radius haul zones of the Kinross facility, and 3 harvest areas are assumed to be farther than 150 miles from facility (7).

The purpose of the CoEE simulation model is to evaluate the annual feedstock delivery cost, energy consumption, and greenhouse gas emissions of the Kinross supply chain for specified operational scenarios. The CoEE simulation model was developed using commercial software called Arena (12). The model consists of five sub-modules: initialization, decision making, harvesting areas, log yards and the processing plant. The simulation is driven by the daily demand of the processing facility, located in the Kinross Township of Chippewa County, and a specified harvest plan for each of the 46 harvest



areas. Thus, the supply chain system is simulated as a combination of “pull” from the facility and “push” from the harvest areas. Figure 2.2 illustrates the overall model logic, and briefly explains the logic of each sub-module, described separately below.



**Figure 2.2:** Schematic of the CoEE simulation model logic (Adapted from "Feedstock Supply Chain Models for the Frontier Renewable Resources Bio-fuel Facility in Kinross, Michigan", 2012 (7))

Initialization is the module that reads in data, including lifecycle cost, storage yard inventory, each harvesting site’s harvest plan, the transportation plan and spring break-up data.

The decision making module controls the simulation model in several ways, such as sending a signal to control production at the facility after checking the inventory of the facility yard, signaling harvesting sites to deliver harvested logs after activating the

corresponding harvest and transportation plans, and writing out the simulation results to the Excel output file.

"Decisions" are made after the initialization is completed and before activities in harvesting areas, log yards and the facility happen. The harvesting areas module creates entities, representing logs, in each of the 46 different harvesting areas. "Entity" is a concept defined by the software Arena. According to W. David Kelton, "entities are the dynamic objects in the simulation; they usually are created, move around for a while, and then are disposed of as they leave (12)". Most entities represent "real" things in a simulation (12). For our model, entities are the logs. After receiving the signal to send harvested logs out, the harvesting module will decide the destination, either log yards or facility yards. Log yards are placed into two categories: rail yards with rail access in the Upper Peninsula and truck yards without rail access in the Lower Peninsula, due to the Mackinac Bridge. For rail yards, logs are transported from harvesting sites by truck first, and then transported to the facility by rail. For truck yards, trucks are the only available transportation tool. For both rail and truck transportation activities includes loading, transporting and unloading. Finally, the facility module receives logs from the harvesting areas and storage yards. After the inventory information of the facility yard is checked by the decision making module, the facility either begins production, if it receives enough timber, or closes due to lack of timber. When the logs are utilized to produce bio-fuel, the logs' age is tracked as well.

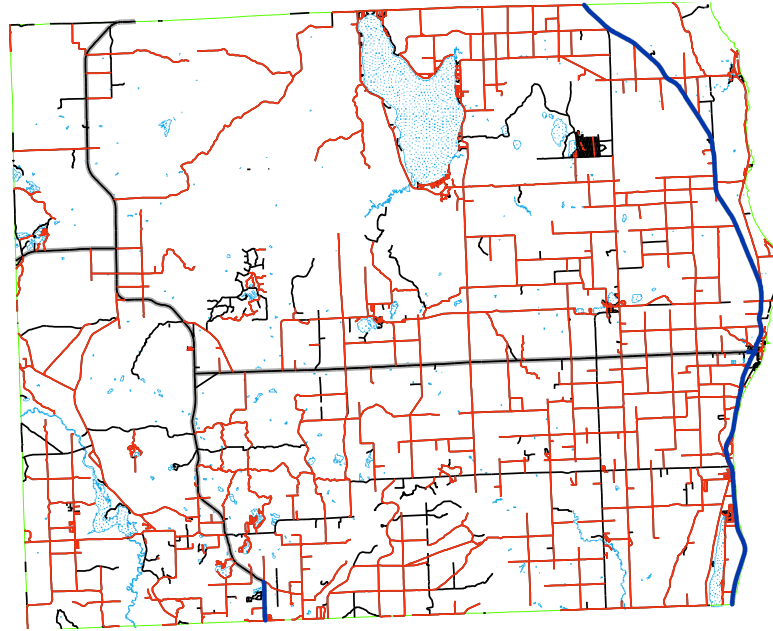
# Chapter 3. Spring Break-up Data

## 3.1. Introduction

Spring break-up load weight restrictions are required by Michigan state law in order to minimize traffic-induced road damage and extend the useful life of the road. As Truck Weight Laws issued by Michigan Department of Transportation (MDOT) indicate, pavement damage is directly related to axle loadings, and road agencies in each county can enact weight restrictions on roads that are not designed as all-season routes during the spring cycle of freezing and thawing. Weight restrictions involve lowering the axle-loading limits and reducing the maximum travel speed for certain vehicles, including the trucks that transport logs (13-15). Thus, spring break-up is expected to have a significant impact on the cost and reliability of biomass feedstock delivered to the Frontier facility.

Spring break-up usually begins in early March and may last over 3 months, during which period heavy loads cannot be shipped on most public roads and full loads of logs can only be hauled on major roads designated as Class A highways. For the purposes of the simulation model, it is assumed that storage yard inventory would build up before spring break-up and draw down during spring break-up, as the primary source of logs during road restriction would be those stored at rail yards in the Upper Peninsula of Michigan, truck yards in the Lower Peninsula of Michigan, and the yard at the Kinross facility. Truck yards are expected to be located on Class A highways, and rail yards built with railroad access.

A list of seasonal road restrictions for state roads was obtained from MDOT (13), and lists of local and county road restrictions were obtained from MDOT and the County Road Association of Michigan (CRAM) (16). A sample map of Alcona County with state roads and non-state roads is shown in Figure 3.1.

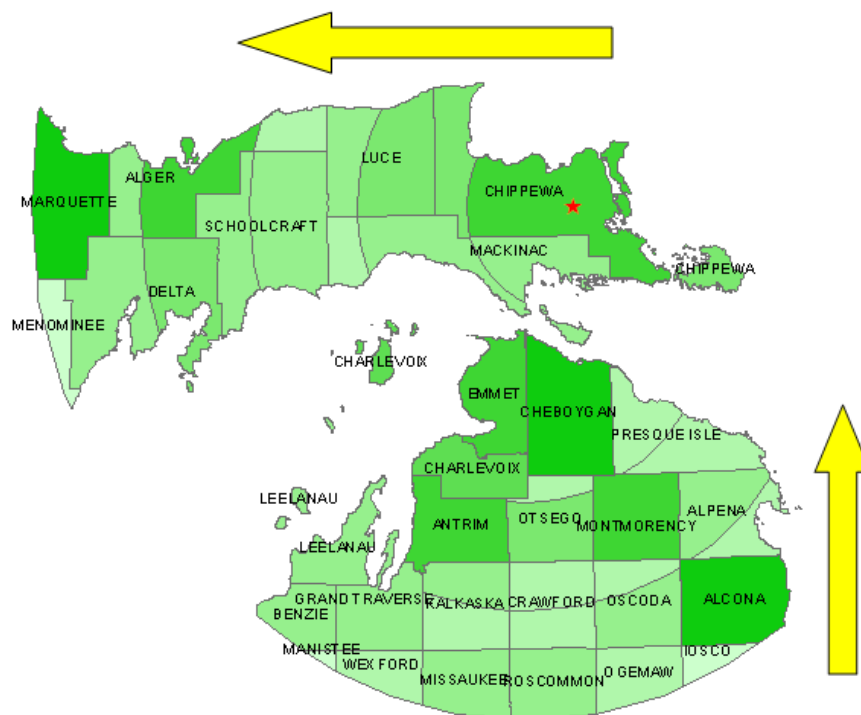


**Figure 3.1.** Map of Alcona County showing roads of different types (data from MDOT and CRAM)

As seen on the map above, the seasonal load restrictions are applied on most of the roads within Alcona County, indicating spring break-up will significantly affect log transportation. As each county road commission is able to implement load restrictions depending on its own conditions, the start day and duration of the seasonal load restrictions vary from county to county. In Alcona County, spring break-up usually starts in early March and ends in April. Further north, it starts in March or early April and usually ends in April or early May.

### 3.2. Data Sources

Historical spring break-up data for the years 2002 to 2010 (including partial data for years 2006 to 2010) was obtained from the MDOT and several county road commissions by Michigan State University (MSU) and Michigan Technological University (MTU), as shown in Appendix A. The data includes the start day and duration of weight restrictions in each county. According to MDOT, various factors such as the soil type, moisture conditions, and the weather are involved in determining the starting and ending dates of the spring break-up. Thus, it is extremely challenging to predict the spring break-up timing, even a short time in advance. However, some relationships are found based on the limited historical data; such as spring break-up start dates and durations in neighboring counties are highly correlated, and start dates and end dates vary geographically among counties in a consistent way. Counties located in the south usually start spring break-up earlier than ones in the north; while counties located in the east usually start earlier than ones in the west, as indicated in Figure 3.2.



**Figure 3.2.** Typical trend of spring break-up in the harvesting region

### 3.3. Factors Affecting Spring Break-up

According to MDOT, weather conditions are important in determining the spring break-up timing (7). In order to analyze the relationship between weather and spring break-up timing, air temperature data is obtained from the Weather Data Depot (17). According to science reports from the USDA Forest Service (18, 19), spring load restrictions should be put into effect when the ground begins to thaw. Pavement loses bearing capacity during this time period because as ice melts it makes the soil unconsolidated. Methods to identify the beginning of the thaw include field measurements (e.g., using a Light Weight Deflectometer (20, 21)) and model prediction. The modeling methods include Mn/DOT method, using air temperatures and reference temperatures to calculate daily freezing index and thawing index (22), and GEO-SLOPE numerical modeling method, simulating freezing and thawing front migration within a pavement structure (23), and so on.

Analysis of the correlation coefficients between the timing of spring break-up and temperature was part of the preliminary analysis, and indicates some potential for the use of seasonal climate forecasts. Analysis results are shown in Table 3.1. The historical monthly average temperature data was obtained from Weather Data Depot (13, 17). The average temperature in March is used to calculate the correlation with the start day, while the average temperature in April is used to calculate the correlation with the end day, and duration uses the average temperature in both March and April. The temperature in April is used for all counties; although some of them end spring break-up in early May, this late spring break-up could be caused by a cold or wet April.

As shown in Table 3.1, for the 27 counties which have historical spring break-up start day data, 26 have negative correlation coefficients for the start day, and one (Crawford County) has zero correlation. Since the one anomaly may be attributed to low statistical power because of the small sample sizes, it may be concluded the start day tends to be earlier when the weather is warmer.

Similarly, for the 27 counties which have historical spring break-up end day data, three (Antrim County, Delta County, and Otsego County) have positive correlation coefficients

for the end day, and 24 have negative coefficients. Since the three anomalies may be attributed to the small sample sizes, it may be concluded the end day tends to be earlier when the weather is warmer.

Finally, for the 27 counties which have historical spring break-up duration data, 26 have negative correlation coefficients for the spring break-up duration and temperature, and one (Delta County) has positive correlation. Since the one anomaly may be attributed to the small sample sizes, it may be concluded the spring break-up tends to be shorter when the weather is warmer.

As mentioned, analysis of correlations with climate data indicates some potential for the use of seasonal climate forecasts, but the decision was made to base the statistical spring break-up model on the historical start and end days because correlations with climate data were not very high.

**Table 3.1**

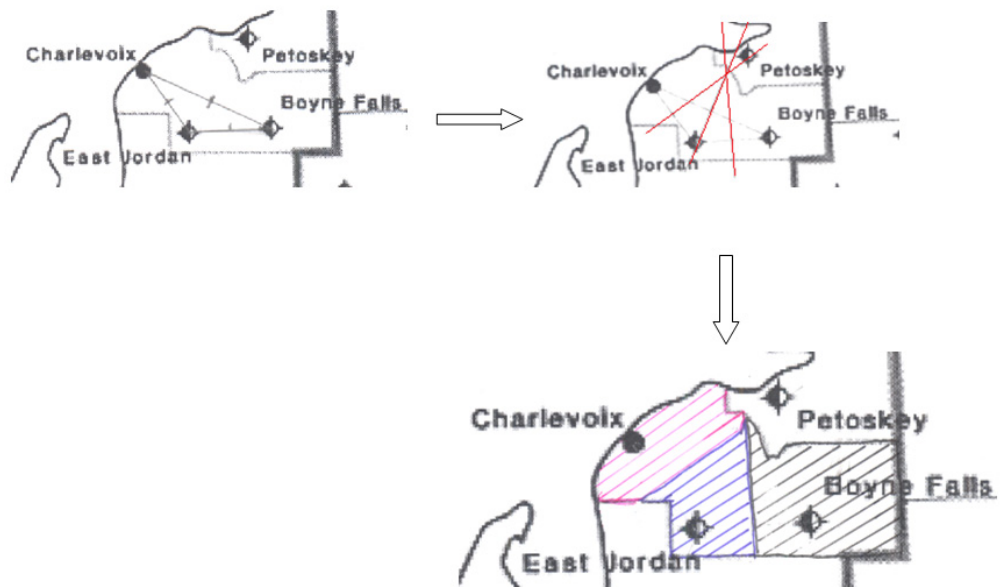
Correlation Coefficients of spring break-up timing and temperature are presented.

Correlation values with \* are significant at  $p=0.05$ , and others are at  $p=0.1$ .

H.A. #	County	corr. coef. for start day	corr. coef. for end day	corr. coef. for duration
1	Alcona	-0.24	-0.57	-0.43
2/3/4	Alger	-0.95*	-0.63	-1.00*
5/6	Alpena	-0.68	-0.61	-0.23
7	Antrim	-0.31	0.09	-0.24
8	Benzie	-0.76*	-0.20	-0.57
9	Charlevoix	-0.83*	-0.66	-0.73
10	Cheboygan	-0.81	-0.50	-0.87*
11/12	Chippewa	-0.85	-0.68	-0.78
13/14	Crawford	0.00	-0.71	-0.88*
15/16	Delta	-0.89*	0.18	0.22
17	Emmet	-0.51	-0.79	-0.59
18	Grand Traverse	-0.69	-0.34	-0.79
19	Iosco	-0.76	-0.71	-0.90*
20/21	Kalkaska	-0.49	-0.71	-0.86
22	Leelanau	no spring breakup data		
23/24	Luce	-0.80	-0.81	-0.35
25/26/27	Mackinac	-0.92*	-0.18	-0.45
28	Manistee	no spring breakup data		
29	Marquette	-0.87*	-0.62	-0.87*
30	Menominee	-0.88*	-0.74	-0.66
31	Missaukee	-0.58	-0.23	-0.79
32	Montmorency	-0.91*	-0.11	-0.30
33	Ogemaw	-0.05	-0.56	-0.58
34/35	Oscoda	-0.85	-0.56	-0.87*
36/37	Otsego	-0.22	0.03	-0.37
38/39	Presque Isle	-0.83*	-0.44	-0.23
40	Roscommon	-0.92*	-0.67	-0.79
41/42	Schoolcraft	-0.92*	-0.26	-0.30
43	Wexford	-0.62	-0.36	-0.72



As precipitation may affect spring break-up timing as well, another analysis is conducted to show the correlation between the spring break-up timing with rainfall and snowfall. The historical monthly precipitation is obtained from the Michigan State Climatologist's Office (24). Since there might be multiple stations that are collecting precipitation data and snowfall data in one county, a Thiessen polygon method (25) was used to weight the station data as shown in Figure 3.3. The example shows the map of Charlevoix County, which possesses three weather stations, in Charlevoix, East Jordan and Boyne Falls.



**Figure 3.3.** Map of Charlevoix County and illustration of the procedure for developing area weights of stations

Analysis results of the correlation coefficients between spring break-up timing and rainfall are shown in Table 3.2. Considering all 29 counties, six of them lack either spring break-up data or rainfall data. For rainfall analysis, 13 out of the remaining 23 counties have positive correlation coefficients for the start day, and 10 have negative correlations, while all 23 counties have positive correlations for the end day. For snowfall analysis, 15 out of the remaining 23 counties have positive correlation coefficients for the start day

and 8 have negative correlations, while 13 out of 23 counties have positive coefficients for end day. No definitive statement can be made regarding the relationships between rainfall and start day or snowfall and spring break-up timing; however, rainfall and spring break-up end day are positively correlated. In other words, spring break-up is expected to end later if more rainfall occurs in that month, which may be due partly to cloud cover.

**Table 3.2**

Correlation coefficients of spring break-up timing and precipitation are presented.

H.A. #	County	Rain		Snow	
		Corr. Coef. for start day	Corr. Coef. for end day	Corr. Coef. for start day	Corr. Coef. for end day
1	Alcona	-0.29	0.55	-0.26	0.38
2/3/4	Alger	0.34	0.77	-0.27	-0.9
5/6	Alpena	0.42	0.39	0.53	0.43
7	Antrim	0.51	0.65	-0.31	-0.24
8	Benzie	no rainfall/snowfall data			
9	Charlevoix	0.32	0.09	0.28	0.12
10	Cheboygan	0.2	0.68	0.85	0.25
11/12	Chippewa	0.46	0.9	0.78	-0.22
13/14	Crawford	0.36	0.22	0.71	0.28
15/16	Delta	no rainfall/snowfall data			
17	Emmet	0.33	0.32	-0.74	0.66
18	Grand Traverse	-0.73	0.8	0.29	-0.53
19	Iosco	-0.18	0.79	0.53	0.3
20/21	Kalkaska	-0.15	0.59	0.22	0.38
22	Leelanau	no spring break-up data			
23/24	Luce	no rainfall/snowfall data			
25/26/27	Mackinac	0.33	0.34	0.76	-0.86
28	Manistee	no spring break-up data			
29	Marquette	-0.12	0.02	0.16	-0.83
30	Menominee	no rainfall/snowfall data			

Table3.2 (Continued)

H.A. #	County	Rain		Snow	
		Corr. Coef. for start day	Corr. Coef. for end day	Corr. Coef. for start day	Corr. Coef. for end day
31	Missaukee	0.34	0.52	0.63	-0.1
32	Montmorency	-0.14	0.8	-0.38	-0.4
33	Ogemaw	-0.55	0.55	-0.32	0.11
34/35	Oscoda	-0.01	0.3	0.34	0.21
36/37	Otsego	-0.49	0.57	-0.37	0.21
38/39	Presque Isle	0.14	0.48	-0.27	0.39
40	Roscommon	-0.15	0.82	0.22	0.35
41/42	Schoolcraft	0.39	0.21	0.54	-0.54
43	Wexford	0.73	0.77	0.45	-0.18

In addition to weather conditions, one of the primary factors impacting the load carrying capacity of a road, and thus spring break-up timing, is the pavement and sub-grade materials. There are various categories of materials for road design (26). For example, pavements are mainly characterized as asphalt or concrete materials (27), and sub-grades are mainly characterized as soil, clay and other mixed materials (e.g. lime and cement) (28, 29). Variation in the properties of the different materials of pavement and sub-grade occurs due to environmental/climate changes, accordingly. Consequently, roads with materials of lower support capacity in spring require load restrictions earlier and/or for a longer duration. For instance, a typical low volume asphalt road's life will be increased by about 10 percent due to proper spring load restrictions (30), and thus road commissioners may choose to impose road restrictions in a conservative manner (i.e., start early and end late).

Finally, some counties may just post spring break-up timing subjectively, such as making the decision depending on neighboring counties. For instance, from the data for 2006 to 2010, it appears that Presque Isle County always ends its load restrictions approximately 4 days after Cheboygan County removes its load restrictions.

# Chapter 4. Spring Break-up Model

## 4.1. Introduction

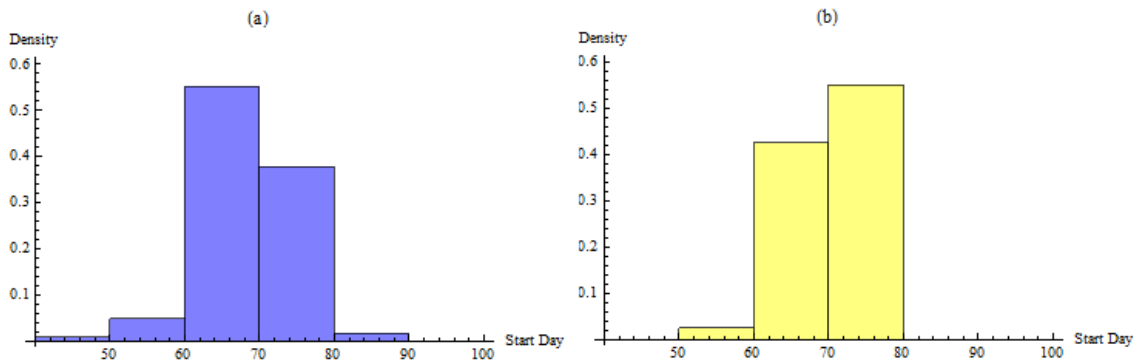
In the simulation model, there are three ways to input spring break-up data. The first way is to input the spring break up start day and end day for all 29 counties (spanning the 46 harvest areas) in the spreadsheet input file. The second way is to specify only the start day and end day of Alcona County's load restrictions, and the spring break-up model (which is coded in Arena) will randomly sample a single realization of the other counties' spring break up start dates and durations according to a statistical model of spatial correlations developed from the historical data. Alternatively, the probability distributions of Alcona County's start day and end day may be defined, and Arena will perform a Monte Carlo simulation, generating a specified number of spring break-up scenarios for the entire harvest region. All calculated spring break-up data are written back to the spreadsheet output file so that the user can check the results.

More details of the procedure for developing the statistical model of spring-break are given in section 4.3.

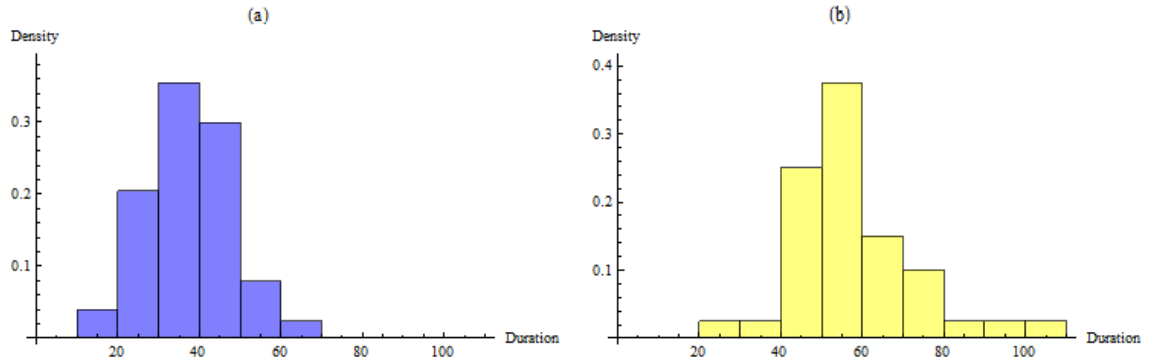
## 4.2. Preliminary Data Analysis

To better understand the historical spring break-up data, a preliminary data analysis is completed here. Figure 4.1(a) shows that all counties in the Lower Peninsula of Michigan start spring breakup between the middle of February and the end of March, and most of them start it in March, while Figure 4.1(b) shows all counties in the Upper Peninsula of Michigan start spring break-up in March as well, and most of them start it in the middle or late March. Figure 4.2(a) shows that the load restrictions in all counties in the Lower Peninsula of Michigan last from 20 days to 70 days, and most of them last for around one and a half months, while Figure 4.2(b) shows that the load restrictions in all counties in the Upper Peninsula of Michigan last from one month to three months, and most of them

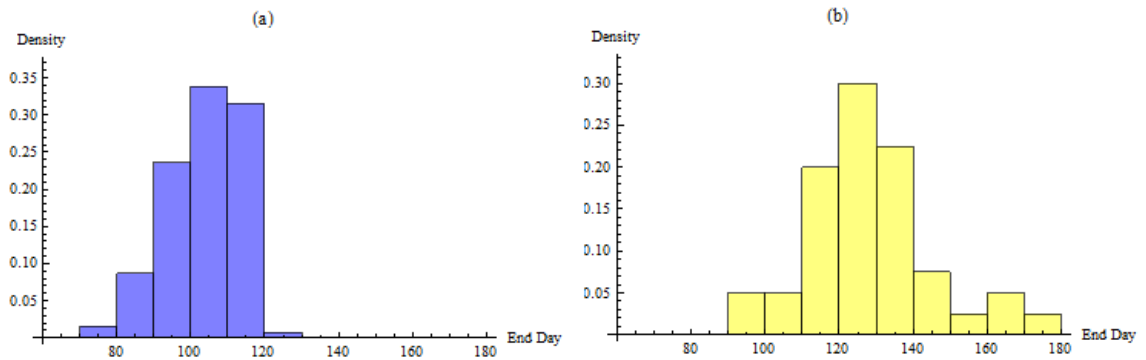
last for around two months. Another finding from the raw data is that load restriction duration is inversely related to the start day because load restrictions tend to end at about the same time every year. In other words, a county with early spring breakup is likely to have a long duration break-up, while a county starting spring break-up late is more likely to have a short duration break-up. The histogram of end day of load restrictions for counties in the Lower Peninsula of Michigan is shown in Figure 4.3(a), indicating that most of the counties end their spring break-up in April. The histogram of end day of load restrictions for counties in the Upper Peninsula of Michigan is shown in Figure 4.3(b), indicating that a large portion of the counties end their spring break-up in May. Furthermore, comparison of histograms indicates that counties in the Lower Peninsula are likely to start and end spring break-up earlier than counties in the Upper Peninsula. Also, counties from the Lower Peninsula have a much shorter spring break-up than counties in the Upper Peninsula. The average duration is 36.6 days in the Lower Peninsula compared to 57.6 days in the Upper Peninsula, and the longest duration is 63 days in the Lower Peninsula compared 104 days in the Upper Peninsula, based on the historical data.



**Figure 4.1.** Histogram of spring break-up start day for (a) L.P. counties and (b) U.P. counties.



**Figure 4.2.** Histogram of spring break-up duration for (a) L.P. counties and (b) U.P. counties



**Figure 4.3.** Histogram of spring break-up end day for (a) L.P. counties and (b) U.P. counties

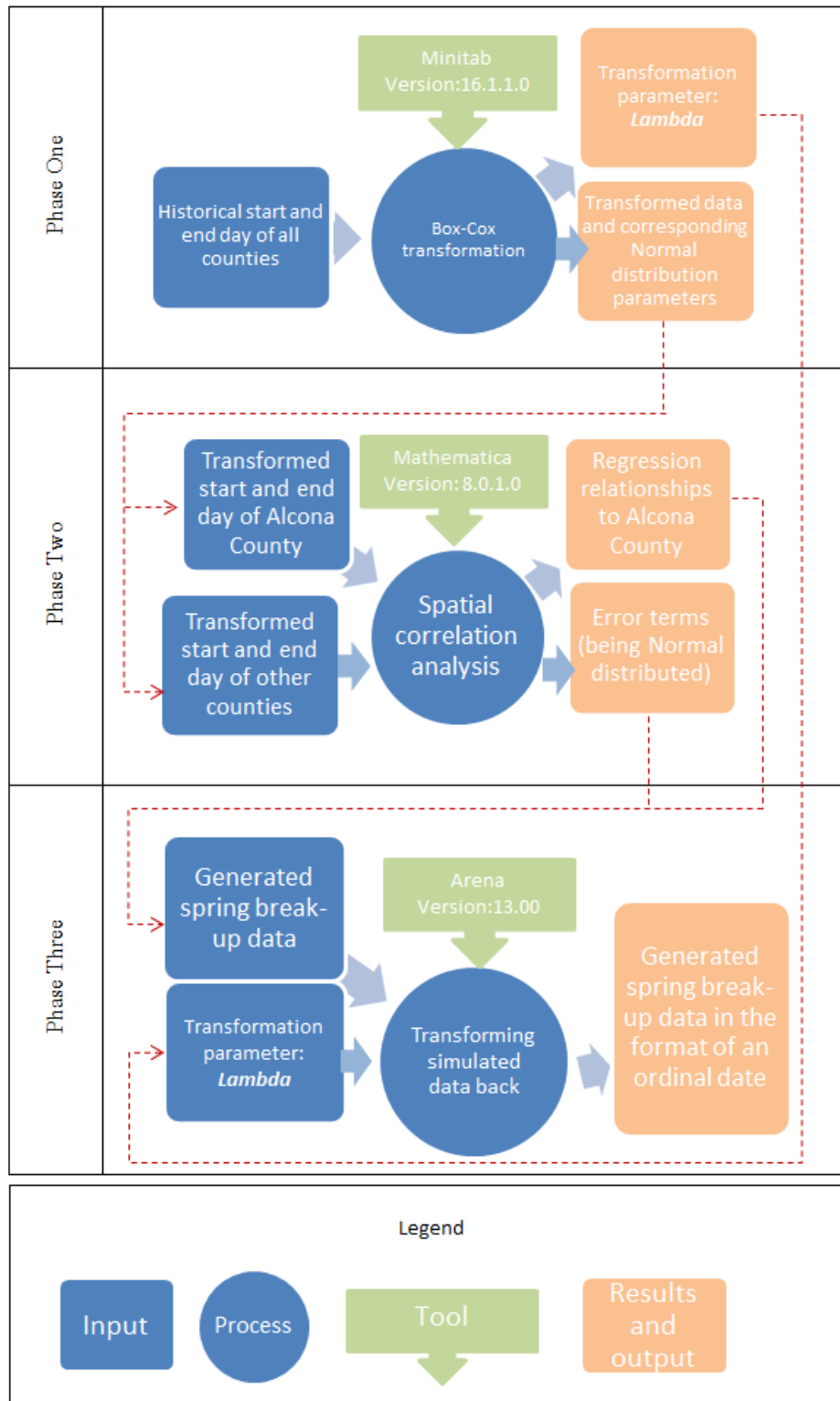
### 4.3. Methodology

Because developing the statistical model of spring break-up requires analysis of correlations, the methodology for the spring break-up model consists of three steps of related analysis:

- (1) Fitting probability distributions to historical data: transformation of the historical start and end day using the Box-Cox transformation, which improves skewed distributed data, in order to fit normal distributions to the transformed data, because regression analysis in step (2) assumes normality of the data sets;

- (2) Developing regression relationships to represent spatial correlations: selection of Alcona County as the first county of the spring break-up model, development of regression relationships of other counties to Alcona County using the transformed data from step (1), as well as estimation of the distribution of the error terms, which are also normally distributed according to step (1);
- (3) Transforming the generated spring break-up scenarios corresponding to the Box-Cox transformed data back to real space, in format of an ordinal date, for use in the simulation model.

Figure 4.4 outlines the steps and tools employed in each step.



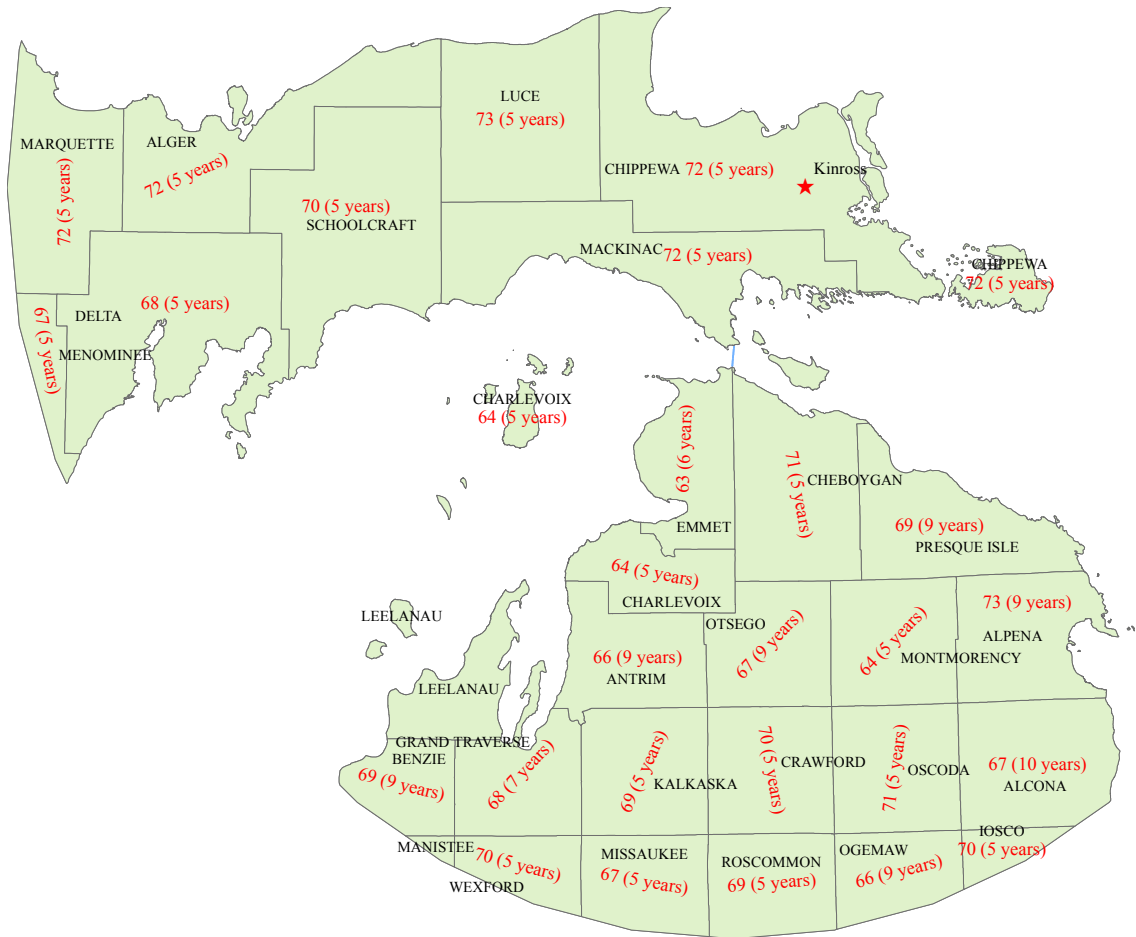
**Figure 4.4.** Overview of study methodology



#### 4.4. Model for Alcona County

The statistical model of spring break-up timing begins with a model of Alcona County's start day and end day. Start day and end day were found to be statistically independent, whereas start day and duration are significantly correlated, for the reason described previously. Alcona County is selected as the basis for other counties in the model because it tends to be one of the earliest counties that experiences spring break-up due to its geographic location, located in the southeast corner of the defined harvest area and adjacent to Lake Huron. Alcona County also has an inland lake, Hubbard Lake, which is among the 10 largest inland lakes in Michigan (31). Due to water's large heat capacity, winter temperature of counties near water bodies tend to be milder than counties without large water bodies (32).

A map, showing average start days of historical data and how many years the historical data has for each county, is presented in Figure 4.5. It shows that Alcona County starts road restrictions relatively early. Charlevoix County may seem to be another good choice for the baseline county, as its average start day is 64 (March 4 or 5), but there are only 5 years of record for Charlevoix County. The limited observations will impact the accuracy of the model. Thus, Alcona County is selected due to its early spring break-up start and its historical data length.



**Figure 4.5.** Map of average historical start days of spring break-up and years of record

As stated in section 4.3, the start and end days of Alcona County’s spring break-up have to be transformed so that they are approximately normally distributed. The Box-Cox transformation is a particularly useful approach to do that. Referring to Box and Cox (1964), the Box-Cox transformation is defined as:

$$Y^{(\lambda)} = \begin{cases} \frac{Y^\lambda - 1}{\lambda} & (\lambda \neq 0), \\ \ln Y & (\lambda = 0). \end{cases} \quad [4-1]$$

It transforms the response variable  $Y$  to an approximately normally distributed variable  $Y^{(\lambda)}$ , with parameter  $\lambda$  (33).

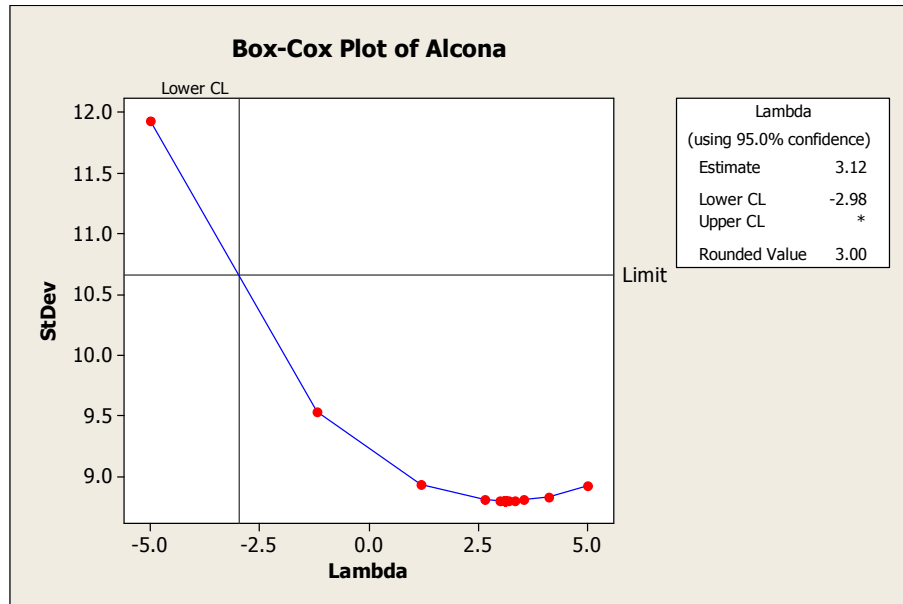
Statistical software Minitab is employed in order to find the transformation parameter  $\lambda$  and transform the data. Although  $Y^{(\lambda)}$  is defined a little differently from equation 4-1,

Minitab provides an effective way to find the best value for  $\lambda$  and an optimal power transformation, as presented in equation 4-2 (34):

$$Y^{(\lambda)} = \begin{cases} Y^\lambda & (\lambda \neq 0), \\ \text{Ln}Y & (\lambda = 0). \end{cases} \quad [ 4-2 ]$$

Since an analysis of variance is the same by a linear transformation, equation 4-1 is equivalent to equation 4-2.

To find the optimal  $\lambda$ , Minitab plots the standard deviation resulting from each transformation vs. possible values for  $\lambda$ , as shown in Figure 4.6, the Box-Cox transformation plot of Alcona’s spring break-up start day. The optimal  $\lambda$  is selected as the value with minimum residual standard deviation, subject to the constraint that the transformation parameter  $\lambda$  is limited to values between -5 and 5.

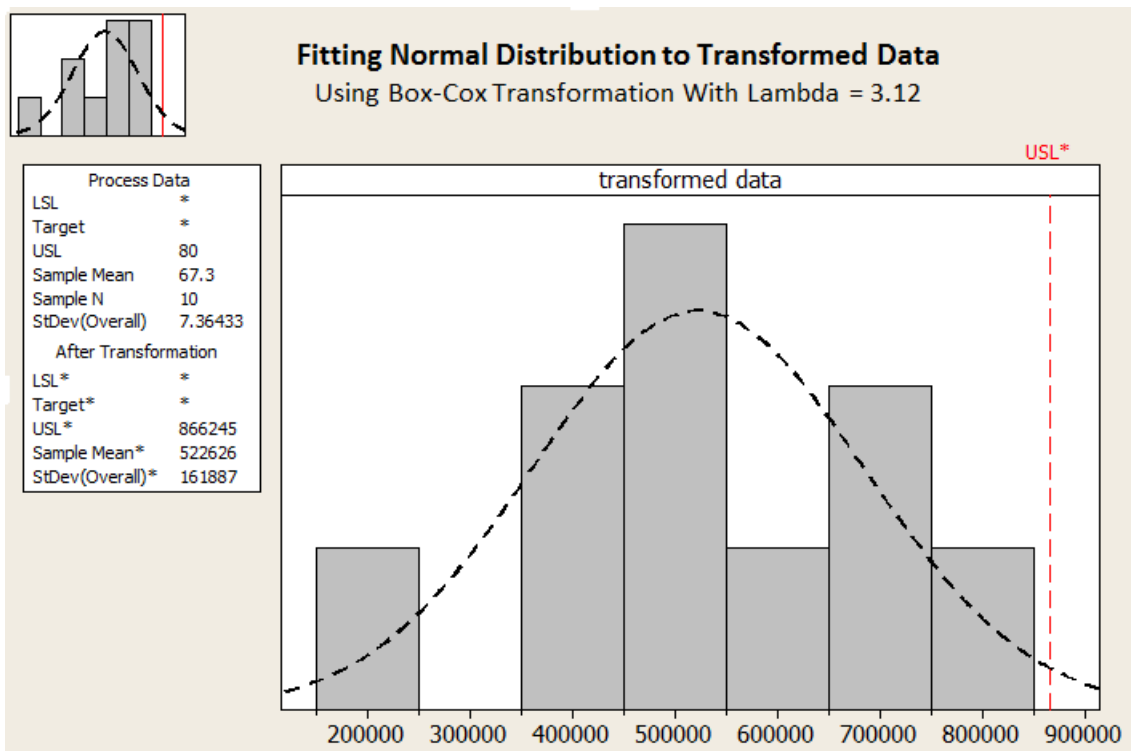


**Figure 4.6.** The Box-Cox transformation plot of Alcona County’s Start Day

Taking the start day of Alcona County, for instance, the best estimate of  $\lambda$  is 3.12, and it can be rounded to 3.0. In this example, both 3.12 and 3.0 are reasonable choices because they both fall within the 95% confidence interval (marked by vertical lines, Lower CL

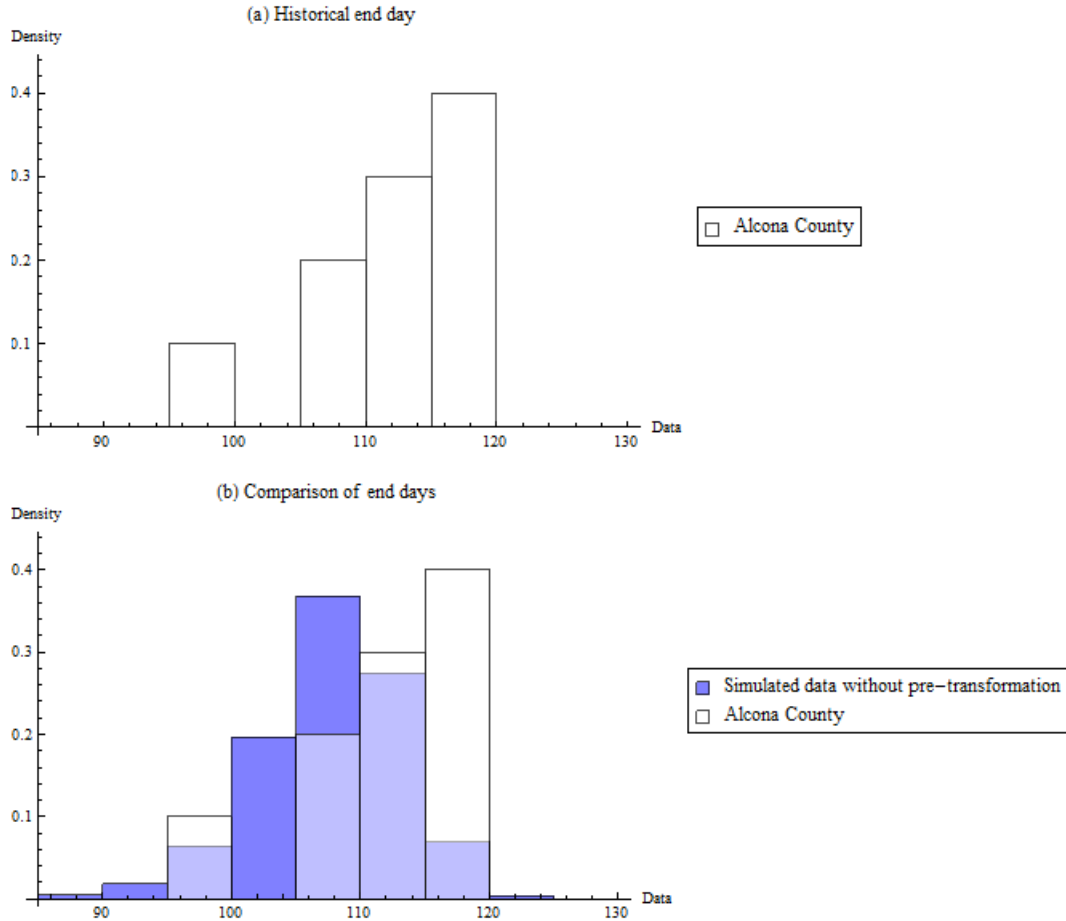
and Upper CL, although Upper CL does not appear because it is greater than 5). In our study, the original value of  $\lambda$ , 3.12, is used for the transformation.

After estimation of  $\lambda$ , a normal distribution is fit to the transformed data, again using Minitab, as shown in Figure 4.7. The plot shows less skew in the transformed data set than in the original observations, shown in the upper left corner of Figure 4.7. The sample mean and standard deviation of the Normal distribution fit to the transformed data are listed in the figure as well.



**Figure 4.7.** The Normal distribution estimated for the transformed start day of Alcona County

Similarly, estimating  $\lambda$  and fitting a Normal distribution to the transformed end day data are also conducted using Minitab. However, the historical data are so skewed (as shown in Figure 4.8(a)) that the simulation results from the Box-Cox transformation using Minitab are not ideal (as shown is Figure 4.8(b)).



**Figure 4.8.** Simulated end day (without pre-transformation) vs. Historical end day for Alcona County. Bars in blue represent simulated data, bars in white represent historical data, and bars in light blue represent overlay from simulated and historical data.

To adjust for this skew, pre-transformation of the historical end day,  $Y$ , is conducted before the Box-Cox transformation using the following equation:

$$T(Y) = (\ln Y)^{21.65} \quad [4-3]$$

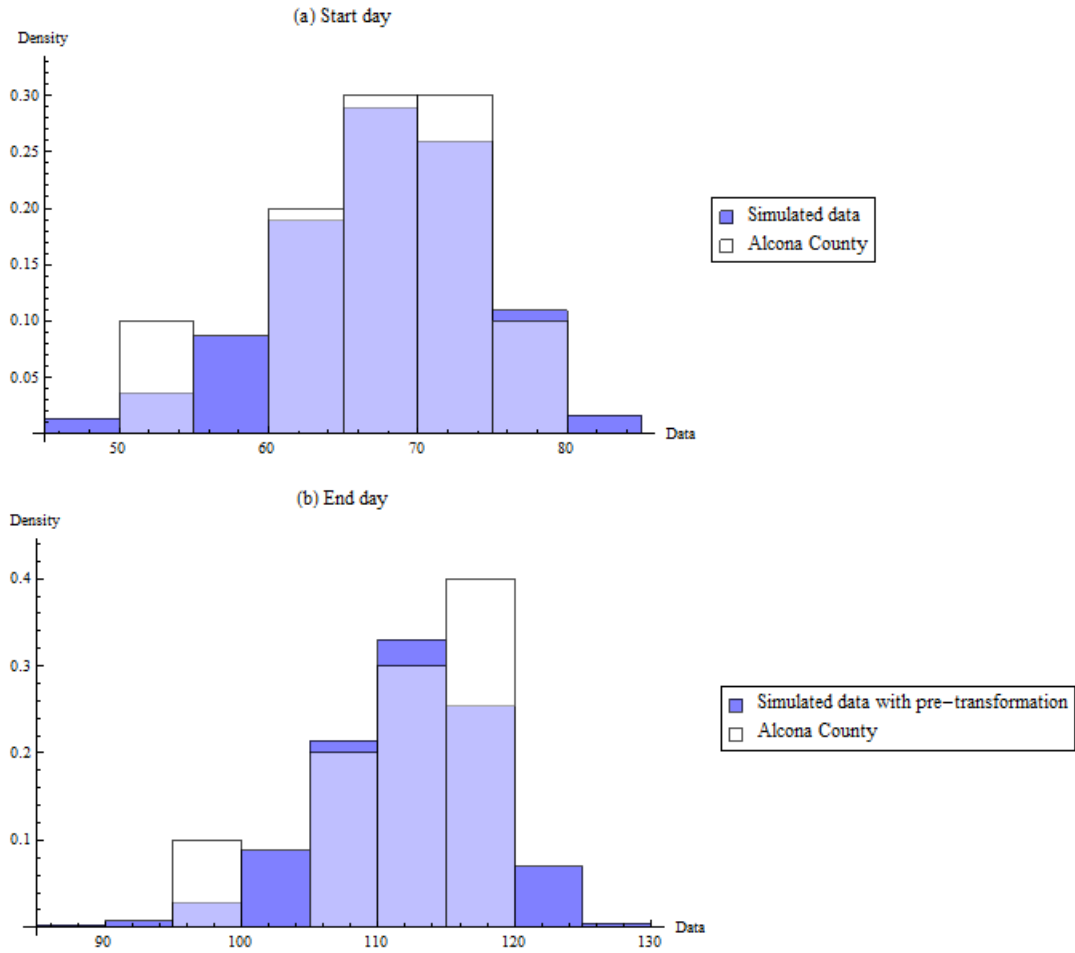
Natural log transformation is one of the most widely used transformations in statistical tests (35). The exponent 21.65 is obtained from iterations of the Box-Cox transformation with  $\ln Y$  via Minitab. The sample mean and standard deviation of the Normal distribution fit to the transformed data are obtained via Minitab as well. Results are listed in Table 4.1.

**Table 4.1**

To present estimation results of Alcona County's spring break-up timing.

<b>Alcona County</b>	<b>Pre-transformation</b>	<b>Transformation Parameter, <math>\lambda</math></b>	<b>Mean of Normal Distribution, <math>\mu</math></b>	<b>Standard Deviation of Normal Distribution, <math>\sigma</math></b>
Start Day	N/A	3.12	$5.226 \times 10^5$	$1.6189 \times 10^5$
End Day	$T(Y)=(\text{Ln}(Y))^{21.65}$	1	$3.944 \times 10^{14}$	$9.473 \times 10^{13}$

In order to check the consistency with historical data for Alcona County, 1000 samples are generated from the estimated Normal distribution and plotted as a histogram, as shown in Figure 4.9. (Code used to generate the figure and random data used in Mathematica 8 are provided in Appendix B.) The distributions of generated samples, shown in blue, are reasonably close to the Alcona County's historical data, shown in a lighter color. Generated sample mean and standard deviation are compared to historical data in Table 4.2. From Table 4.2, the means and standard deviations of simulated data are reasonably close to the historical data.



**Figure 4.9.** Simulated data vs. Historical start day (a) and end day (with pre-transformation) (b) for Alcona County. Bars in blue represent simulated data, bars in white represent historical data, and bars in light blue represent overlay from both.

**Table 4.2**

To present sample means and standard deviation for Alcona County spring break-up timing.

	Start Day		End Day	
	Mean	Standard Deviation	Mean	Standard Deviation
Historical data for Alcona County	67.30	7.36	111.90	6.56
Generated data from Normal distribution	67.22	7.46	111.91	6.21

## 4.5. Spatial Correlation

Spatial correlations are considered because start dates and end dates vary geographically among counties in a consistent way. In other words, the timing of load restrictions for each county is not independent of neighboring counties, based on correlation analyses of start day and end day. Furthermore, counties located in the south usually start spring break-up earlier than those farther north, while eastern counties usually start earlier than western counties. Although more sophisticated multivariate statistical models could be developed, due to the short record lengths, linear regression analysis is deemed adequate to develop relationships between Alcona County and the other 28 counties within the harvesting region.

### 4.5.1. Simple Linear Regression

Simple linear regression is a common approach to modeling the functional dependence of one variable on another, or in other words, the relationship between a scalar variable  $Y$  and one or more explanatory variables,  $X_i$ . In particular, a simple linear regression relationship is of the form:

$$Y_i = \beta + \alpha X_i + \varepsilon_i \quad i = 1, 2, \dots, N \quad [ 4-4 ]$$

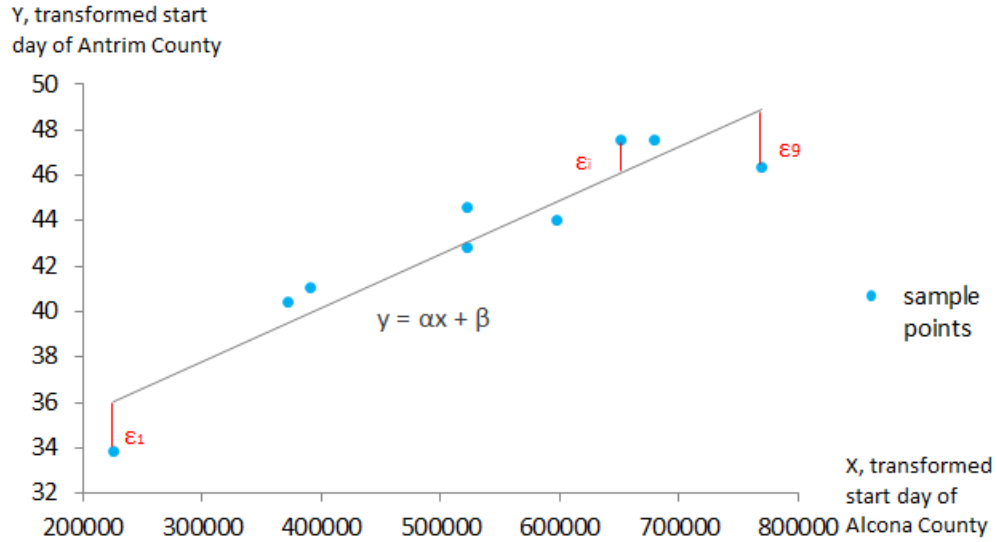


where  $Y_i$  is a random dependent variable and  $X_i$  is an observable independent variable. The parameters  $\beta$  and  $\alpha$  are the intercept and slope of the regression, and  $\varepsilon_i$  is a random variable called the error term, or residual, which captures all other factors influencing the linear relationship between the dependent variable  $Y_i$  and observation  $X_i$  (36).

A linear regression model may be fitted using an approach called Least Squares, which aims to minimize the sum of squares of the error terms. Consider the example of the transformed spring break-up start day of Alcona County,  $X_i$  (transformed using the Box-Cox transformation parameter  $\lambda=3.12$ ), and of Antrim County,  $Y_i$  (transformed using  $\lambda=0.9$ ), as shown in Figure 4.10. For any line  $y = \alpha x + \beta$ , the residual sum of squares (RSS) is defined to be the sum of squares of error term  $\varepsilon$ ; that is

$$RSS = \sum_{i=1}^n (\varepsilon_i)^2 = \sum_{i=1}^n (Y_i - (\beta + \alpha X_i))^2 \quad [ 4-5 ]$$

The RSS measures the vertical distance from each data point to the fitted line  $y = \beta + \alpha x$  and then sums the squares of these distances. The selected values  $\alpha$  and  $\beta$  are defined to be those values such that the line  $y = \beta + \alpha x$  minimizes the RSS. From this approach, it is found that the Y-intercept value equals 30.706 and the slope of the line is 2E-5, while the RSS reaches a minimum value of 55.21.



**Figure 4.10.** Linear regression model of Antrim County transformed start day vs. Alcona County transformed start day. Vertical distances are measured by RSS.

In this study, how well the line fits the data, or what percent of the variance can be defined by the line, is measured by Pearson’s correlation coefficient, defined as:

$$R_{XY} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{[(\sum X^2 - \frac{(\sum X)^2}{n})(\sum Y^2 - \frac{(\sum Y)^2}{n})]}} \quad [4-6]$$

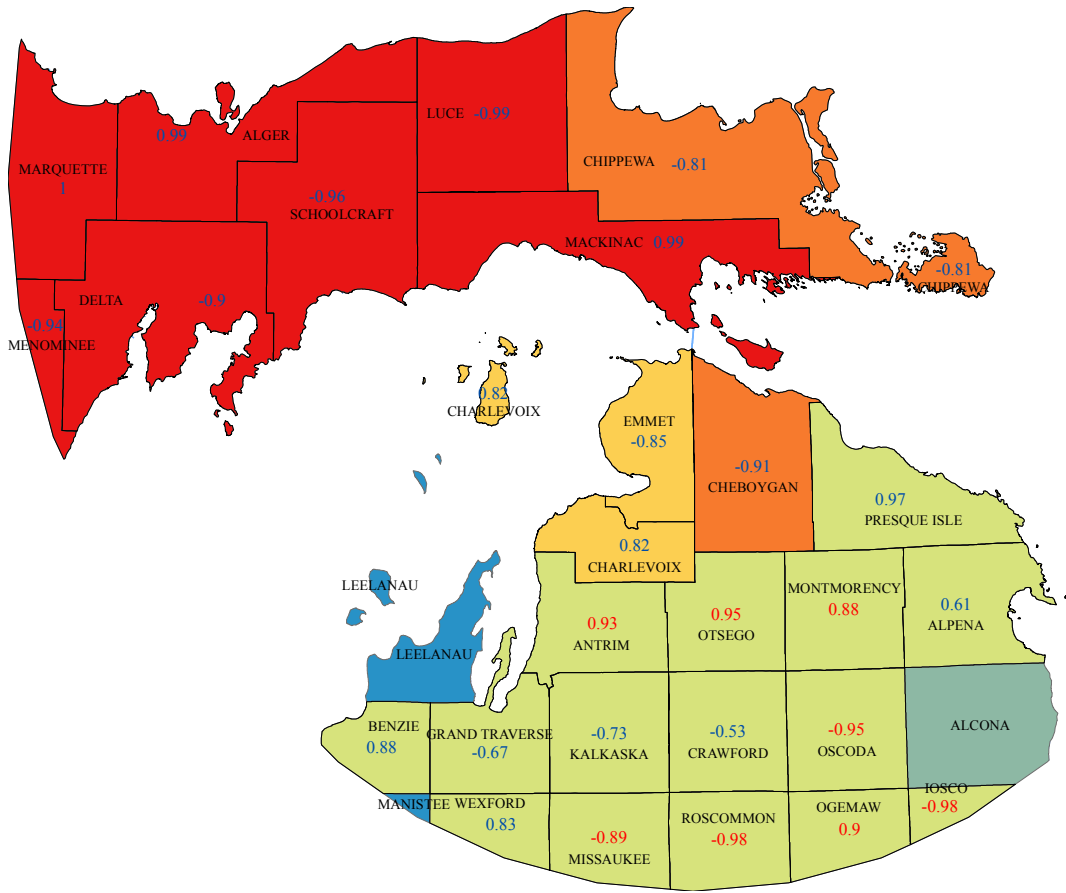
in which  $n$  refers to the number of pairs of data used in the regression (37). The correlation coefficient,  $R$ , ranges from -1 to 1. The portion of the variance explained by the line is given by  $R^2$ , where  $0 \leq R^2 \leq 1$ . Thus, if all variation in  $Y$  is explained by  $X$ , then  $R^2 = 1$ , but if the dependent variable  $Y$  and observation  $X$  have no linear relationship, then  $R^2 = 0$  (38).

From equation 4-6 above, the correlation coefficient  $R$  equals 0.934, and the coefficient of determination  $R^2$  is 0.872. Therefore, the transformed spring break-up start days of Alcona County and Antrim County are highly correlated, and it is concluded that the start day of spring break-up in Alcona County is a good predictor of the start day of spring

break-up in Antrim County, using the equation  $Y = 2.53E-5X + 30.706$  (for the transformed data).

Regression was done in Excel to relate the spring break-up start days of Alcona County to the other 27 counties. For most of the counties, this was done in three stages by first estimating the Box-Cox transformation parameters using Minitab, then developing regression models for counties with the highest correlation coefficients, and finally relating these “key counties” to the rest of the counties. The groupings and correlation coefficients are shown in the Figure 4.11. The reason, negative correlation coefficients are shown for some counties, is that regression relationships are inverted, because transformation parameter  $\lambda$  values have opposite signs. Therefore, these counties still show a strong correlation only signs have changed. All regression equations and transformation results are listed in Appendix C.

Although all counties could have been related to Alcona County in a single step, some counties’ start and end days were found to have statistically insignificant correlations with Alcona County’s spring break-up timing, perhaps due to their short record lengths. For instance, correlation of end day of Antrim and Alcona County is 0.55 at  $p=0.125$  level, but correlation of end day of Antrim and Cheboygan County is 0.884 at  $p=0.047$  level. The grouping of counties into “regions”, although done subjectively, was considered to be a more intuitive approach, with precedence in the hydrologic sciences (39). In Figure 4.11, correlation values in red indicate significance at the  $p < 0.05$  level, while values in blue indicate marginal significance at the  $p < 0.1$  level. The regional approach was also envisioned to provide a means for sampling from the regression residuals, as discussed in the next section.



**Legend**

- no data
- required
- according to Alcona
- according to Montmorency
- according to Presque Isle
- according to Chippewa

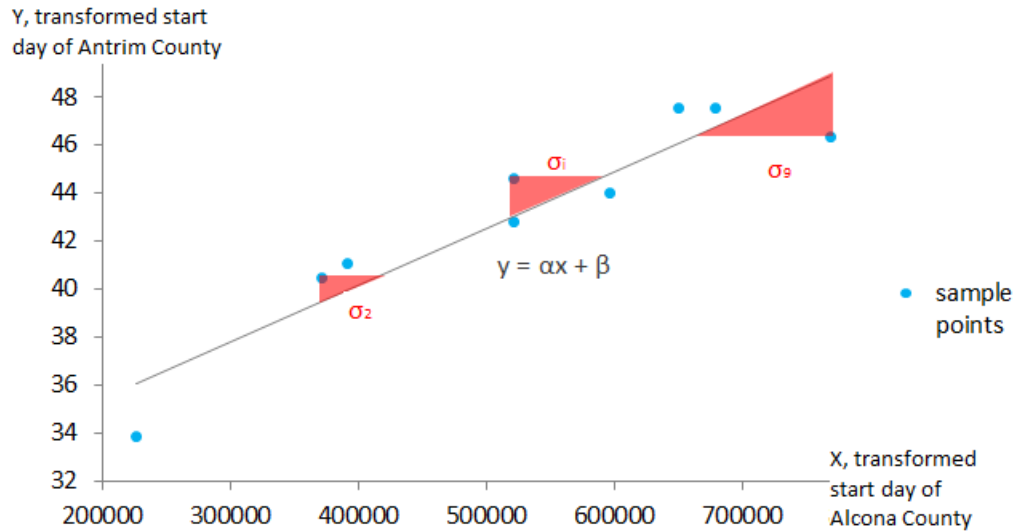
**Figure 4.11.** Regional groupings and correlation coefficients of transformed spring break-up start day.

Similarly, estimation of the Box-Cox transformation together with regression and correlation analysis was conducted on the historical end day data. The regional groupings and correlation coefficients are shown in Figure 4.12. Correlation values in red indicate significance at the  $p < 0.05$  level, while values in blue indicate marginal significance at the  $p < 0.1$  level. All regression equations are listed in Appendix C.



with error (41). The purpose of this section is to evaluate whether or not the simple linear regression used in the former section is an optimal (or at least acceptable) choice.

The underlying assumption for simple linear regression is that deviations between sample points and the fitted line happen only in the direction of vertical axis, as pictured in Figure 4.9. In contrast, RMA regression assumes that deviations between observations and fitted data values occur in both vertical and horizontal directions. While simple linear regression tries to minimize the sum of Y-direction error terms, the RMA regression minimizes the product of the X and Y deviations from the fitted line, or the triangular areas shown in red in Figure 4.13.



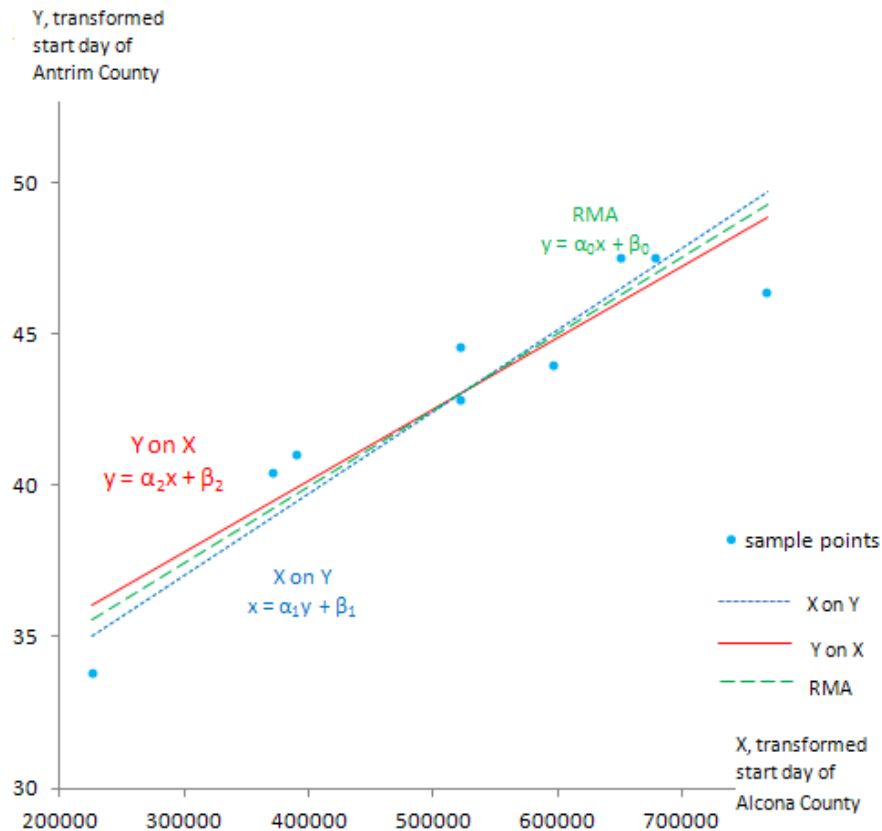
**Figure 4.13.** RMA regression schematic, using the example of transformed spring break-up start days in Antrim County and Alcona County.

Letting  $\sigma_i$  be the measure of deviation for sample point  $(X_i, Y_i)$ , then for any line  $y = \alpha x + \beta$ , the aim of RMA regression is to minimize the sum of the  $\sigma_i$ , or the sum of the product of  $(X_i - X)$  and  $(Y_i - Y)$ , as follows:

$$\text{Objective function: } \sum_{i=1}^n (\sigma_i) = \sum_{i=1}^n \left[ \left( X_i - \frac{Y_i - \beta}{\alpha} \right) * (Y_i - (\beta + \alpha X_i)) \right] \quad [4-7]$$

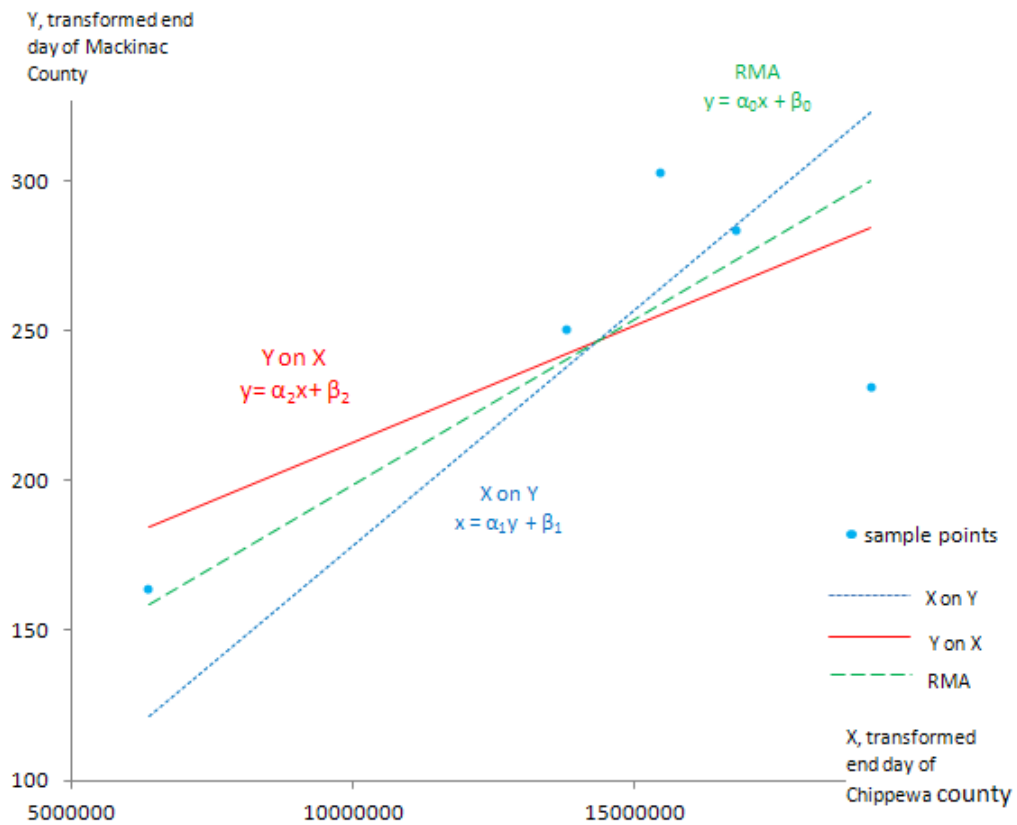
Thus, the linear fit is achieved by minimizing the product of deviations between observations and fitted values in both X- and Y-directions.

A comparison of simple linear regression and RMA regression is shown in Figure 4.14, presenting the transformed historical start day of spring break-up for Alcona County and Antrim County. Simple linear regression, which is termed "Y on X" as it minimizes the deviations in the Y-direction, is shown in red. The inverse regression result, "X on Y", with the deviations in the X-direction minimized, is shown in blue. The RMA regression is shown as the dashed green line. All three lines intersect at the bi-variant mean  $(\bar{X}, \bar{Y})$ . These three regression lines are very close to each other because the two variables are highly correlated. More extreme examples are found in the data set, however.



**Figure 4.14.** Comparison of simple linear regression and RMA regression

For example, in Figure 4.15, Mackinac County's transformed end day is plotted against that of Chippewa County. In this case, the regression lines of X on Y, Y on X, and the RMA line are much more distinct than the previous example because the correlation coefficient is only 0.703. This illustrates that RMA regression is more suitable for the statistical model of spring break-up, because the X variable (e.g., Alcona County's spring break-up start day in the first example) is chosen arbitrarily.



**Figure 4.15.** Regression lines RMA regression schematic

The calculation procedure for RMA regression and code in Mathematica 8 used to determine the RMA regression equation are shown in Appendix D. All results, as slopes and intercepts, of the RMA regression analysis are listed in Appendix E.



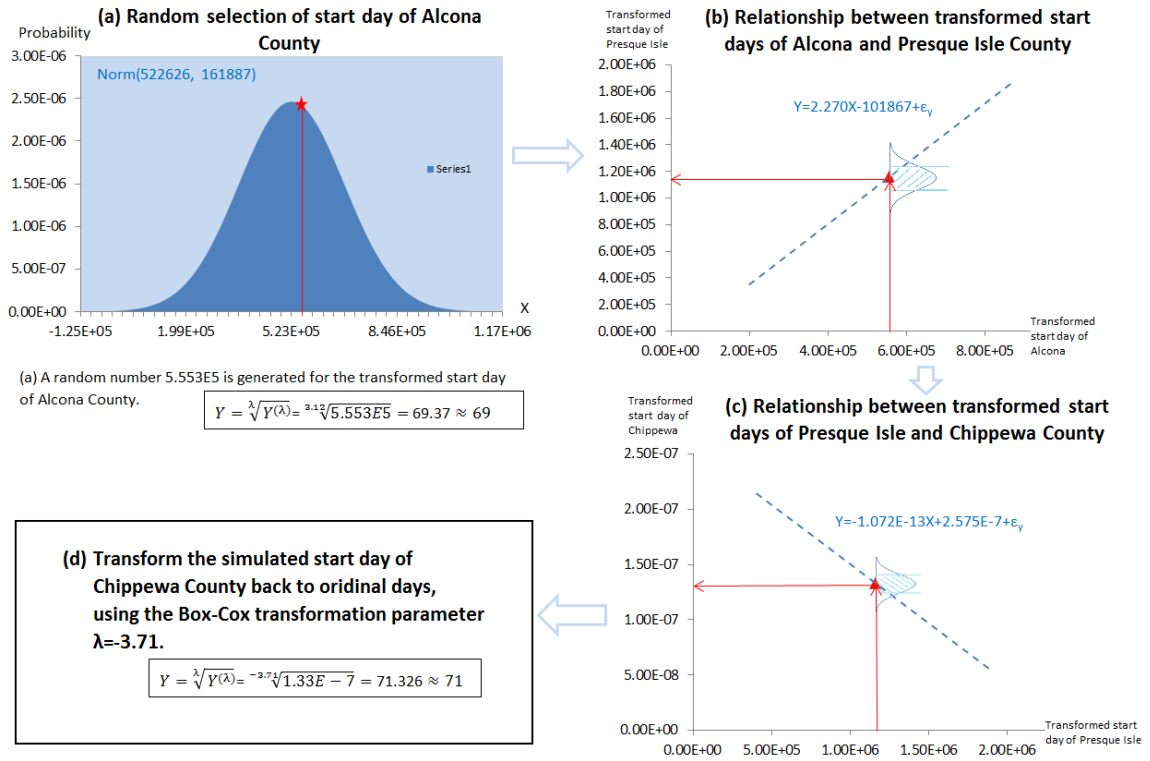
Attempting to re-create historical variability in the simulated scenarios, an error term  $\varepsilon_i$  is added to the RMA regression relationship as follows:

$$Y_i = \beta_0 + \alpha_0 X_i + \varepsilon_i \quad i = 1, 2, \dots, N, \quad [4-8]$$

where  $\alpha_0$  and  $\beta_0$  are the RMA regression slope and intercept, respectively, and the error term  $\varepsilon_i$  is independent with distributions  $Y_i$ . In the utilization of RMA regression, two assumptions are made to consider the error term:

- For a given  $X$ , the  $Y$ 's value is predictable with a random error  $\varepsilon_i$ , where  $X_i \sim \text{Normal}(\mu_x, \sigma_x)$ ,  $Y_i \sim \text{Normal}(\mu_y, \sigma_y)$ ,  $\varepsilon_i$  is  $\text{Normal}(0, \sigma_y)$ . The error term  $\varepsilon_i$  is independent with distributions of both  $X_i$  and  $Y_i$ .
- The value of  $\varepsilon_i$  will be truncated within the 68% probability range. In other words,  $\varepsilon_i$  is always greater than  $-\sigma_y$  and less than  $\sigma_y$ . This constraint is based on experiments which showed allowing a larger range tended to overinflate the variance and could lead to implausible results, such as the end day coming before the start day.

Based on these two assumptions, the spring break-up algorithm can generate potential spring break-up scenarios. An example is given in Figure 4.16, illustrating the procedure for estimating a selected county's spring break-up timing from Alcona County and another baseline county.

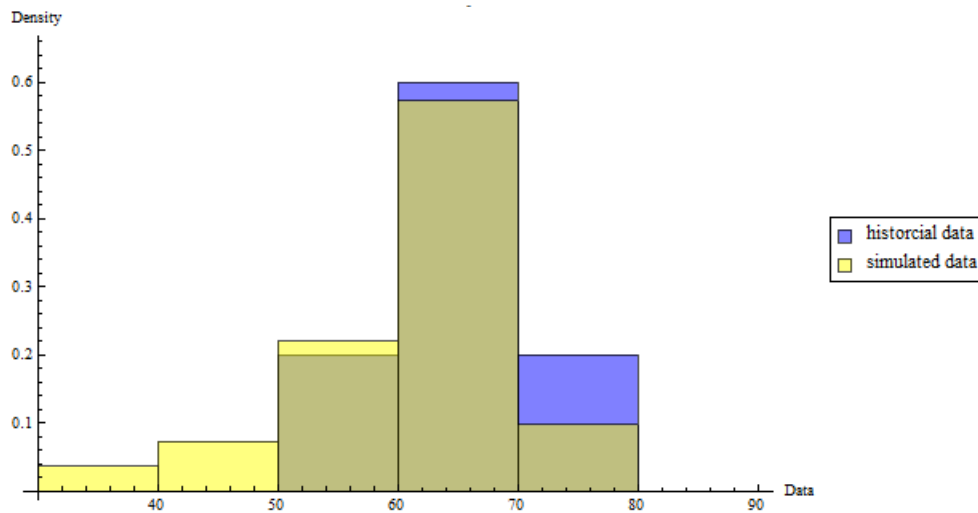


**Figure 4.16.** Example of simulating the spring break-up start day of Chippewa County

# Chapter 5. Simulation Results

## 5.1. Verification of Spring Break-up Model

For verification purposes, a comparison of historical data and simulation results for 100 replications is presented in the form of histograms and summary statistics. Several examples are presented below, such as the start day of Charlevoix County, the end day of Benzie County and the duration of Chippewa County.

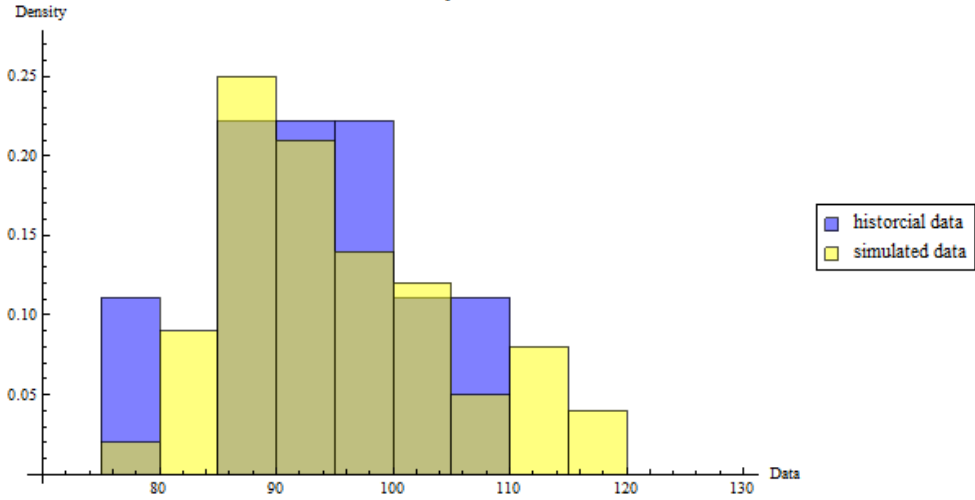


**Figure 5.1.** Simulated start day of Charlevoix County from Arena vs. Historical dates

**Table 5.1**

To present means and standard deviations for the start day of Charlevoix County.

	Mean	Standard Deviation
Simulated start day	60.90	8.21
Historical start day	64.40	5.13

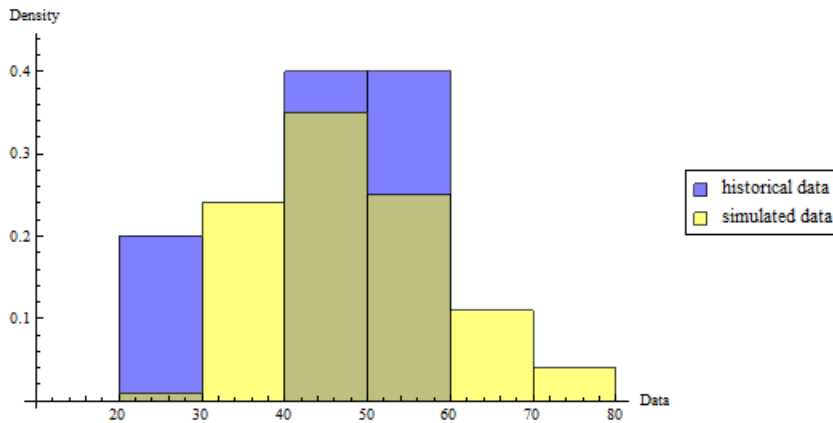


**Figure 5.2.** Simulated end day of Benzie County from Arena vs. Historical dates

**Table 5.2**

To present means and standard deviations for the end day of Benzie County.

	Mean	Standard Deviation
Simulated end day	94.79	9.87
Historical end day	92.56	8.53



**Figure 5.3.** Simulated duration of Chippewa County from Arena vs. Historical dates

**Table 5.3**

To present means and standard deviations for the duration of Chippewa County.

	<b>Mean</b>	<b>Standard Deviation</b>
Simulated duration	48.03	10.89
Historical duration	47.6	11.06

For these three randomly selected examples, simulated mean start day, end day and duration are similar to historical means, but the standard deviations show some discrepancies. This may be because the historical observations are too limited for accurate fitting of distributions, especially for the tails of each distribution, or because the historical data tend to come from truncated distributions.

## **5.2. Uncertainty Effects on Supply Chain Simulations**

After developing and verifying the spring break-up statistical model, the Arena supply chain simulation model is able to consider the uncertainty arising from the timing of load restrictions. For comparison, results of the random (stochastic) simulation are compared to a deterministic scenario, with no uncertainty in spring break-up timing. The deterministic scenario assumes that all counties in the Upper Peninsula of Michigan start spring break-up on the same day and end it on the same day, as do counties in the Lower Peninsula. All the specified start days and durations are taken from average historical statistics, as listed in Table 5.4.

**Table 5.4**

To present historical spring break-up statistics.

	<b>Average start day</b>	<b>Average duration</b>
Counties in the L.P.	68	37
Counties in the U.P.	71	58

Results from simulations of 100 random spring break-up scenarios and the deterministic scenario are summarized in the following table and figures.

Table 5.5 gives the reliability for the supply chain supporting the processing facility in both days and years. The supply chain is considered reliable for a given year if there is no plant closure in that year. As the reliability results show, for the fixed spring break-up, the supply chain supports the processing facility perfectly. If is known exactly when the spring break-up begins and ends, proper harvest plans and transportation plans can be made in advance to make sure log inventories are sufficient to support production throughout the time of load restrictions. However, in reality, spring break-up occurs randomly. When spring break-up happens earlier than the expected start day, the supply chain may not have prepared enough inventory, and thus the facility production may need to be reduced during spring break-up. As shown in Table 5.5, the reliability is somewhat lower when spring break-up happens randomly, and in fact at fixed production levels, plant closures (for 1 or more days) may be expected in 41% of the years. In practice, production may be scaled back rather than the plant closing, but this decision is not supported by the simulation model.

**Table 5.5**

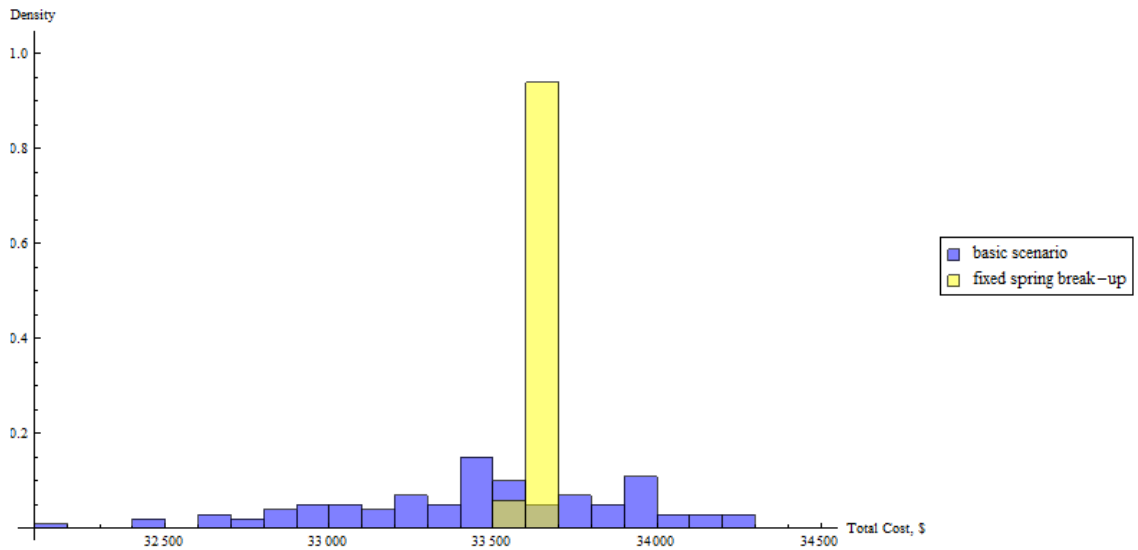
To present summary of simulation results.

	<b>Reliability, % days</b>	<b>Reliability, years</b>
Random spring break-up	99.57%	59/100
Deterministic spring break-up	100%	1/1

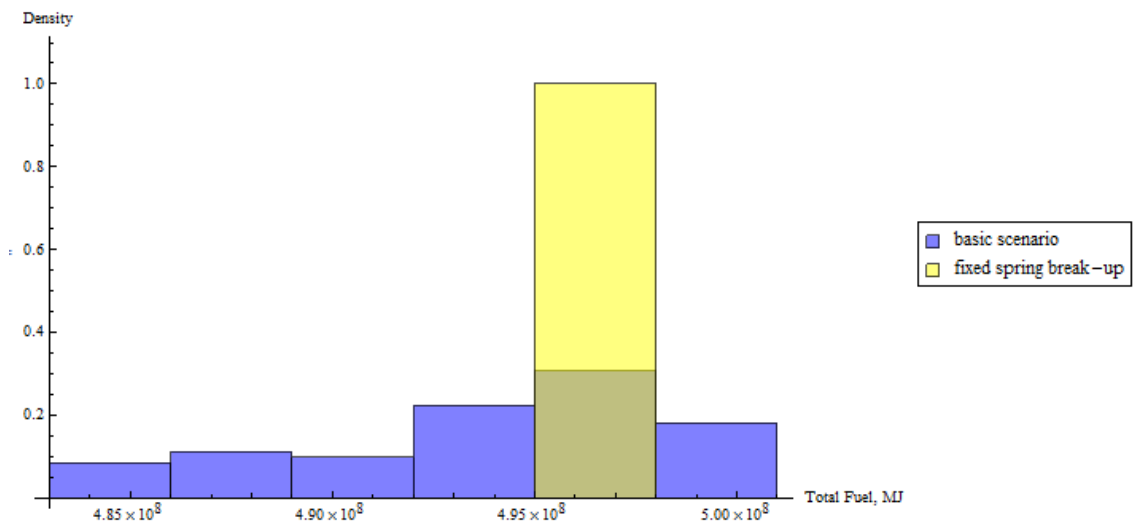
In addition to reliability information, the total cost, energy and emissions results for the random spring break-up scenarios and deterministic scenarios are shown in Figure 5.4 to 5.6. (As shown, there is actually a little randomness in the deterministic spring break-up scenario, due to randomness in the transportation loading and unloading activities.) Based on the histograms of total cost shown in Figure 5.4, some scenarios have higher total cost than the deterministic scenario, which might be due to the increased storage cost caused by late spring break-up; however, some total cost results are smaller than the deterministic scenario, which might be due to the decreased storage cost caused by early spring break-up. However, no cost for facility closure is specified, which could significantly increase the real cost when a long spring break-up occurs.

Similar to total cost, the histograms of total energy consumption and greenhouse gas emissions, as shown in Figure 5.5 and 5.6, indicate significant variability for the random spring break-up scenarios. Total energy consumption and greenhouse gas emissions are lower when spring break-up is longer, during which time more rail transportation is used. Rail consumes only 0.397 MJ fuel per ton-mile, while trucks consume 2.1 MJ fuel per ton-mile; and rail discharges only 0.035 kg greenhouse gas per ton-mile, while trucks discharge 0.183 kg per ton-mile. Also, when spring break-up is longer than expected, fewer logs are harvested due to load restrictions, and thus the fuel consumption and greenhouse gas emissions due to harvesting activities are reduced as well.

Based on these comparisons, it is clear that the random spring break-up model provides more comprehensive insight into the risks associated with various supply chain decisions.

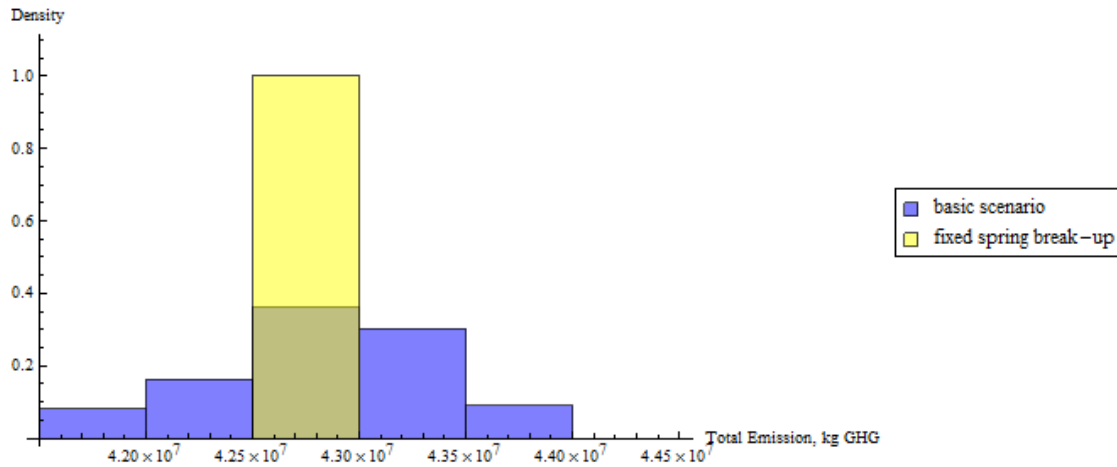


**Figure 5.4.** Comparison of Total Cost for deterministic and random spring break-up scenarios



**Figure 5.5.** Comparison of Total Energy Consumption for deterministic and random spring break-up scenarios





**Figure 5.6.** Comparison of Total Emission for deterministic and random spring break-up scenarios

The real benefit of the simulations with random spring break-up comes from the ability to adjust decisions that lead to an increase in reliability, and to evaluate associated tradeoffs. For example, the reliability can be improved by adjusting harvesting plans, transportation plans, or both. Compared with the baseline harvesting plan in the previous example, if just 13% more logs are harvested during February, the daily and annual reliability will increase to 99.66% and 65.57%, respectively. However, this increase in reliability comes at the expense of a 0.45% increase in total cost, a 0.38% increase in total emissions, and a 0.36% increase in fuel consumption. The effects of changing the transportation plan are illustrated in the next section.

### 5.3. Impact of Rail Use and Rail Yard Storage

During the spring break-up, truck loads are limited in order to minimize the damage to seasonal roads, and the only access to logs for the facility will be either from its own storage yard, logs stored in a truck yard along a Class A highway, logs stored with railway access, or logging jobs that are taking place on a Class A highway. Increased rail use is needed to carry logs to the facility during this time period. Available rail yards, and their storage capacities, will impact the overall performance of the supply chain as well. The aim of

this analysis is to show the impacts on the overall supply chain due to rail use and rail yard storage.

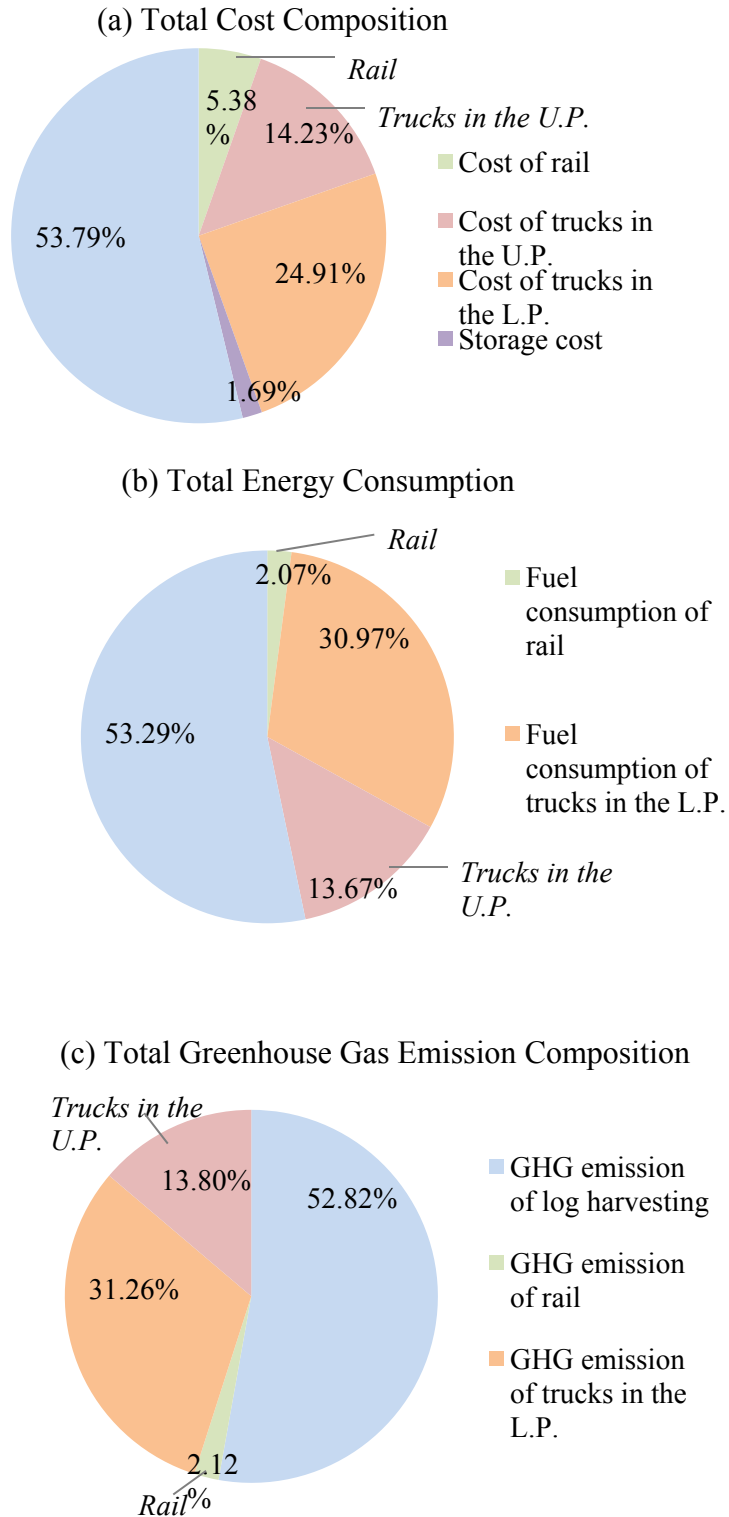
As a sensitivity analysis, rail yard capacity is reduced by 15%, 30% and 60% for comparison with the baseline scenario, in which it is assumed that the rail yards are able to hold logs from 150 rail cars, and each rail car carries 80 tons of logs. In other words, the rail yards' capacity is 12,000 tons for a baseline scenario, 10,200 tons for the 15% reduction scenario, 8,400 tons for the 30% reduction scenario, and 4,800 tons for 60% reduction scenario.

Results in Table 5.6 show that reducing the rail yard capacity reduces the reliability of providing feedstock to the processing facility. The reliability does not decrease rapidly, however, because even with reduced rail yard capacity, logs may also be kept in the facility yard in advance of spring break-up. The averages of total cost, total energy consumption and total greenhouse gas emission increase as the rail yard capacity decreases. This is mainly because rail use is also decreasing, which means more trucks have to be employed, and rail is a more efficient transporter than truck. Total cost does not decrease rapidly, however, because transportation cost of rail and trucks from the Upper Peninsula make up a relatively small proportion of the overall cost, according to the total cost composition of the baseline scenario, as shown in Figure 5.7 (a). Total energy consumption and total greenhouse gas emissions do not decrease rapidly either, due to the similar reasons, as shown in composition plotting in Figure 5.7 (b) and (c).

**Table 5.6**

To present summary of simulation results for rail yard capacity reduction scenarios.

		<b>Baseline scenario</b>	<b>15% reduced rail yard capacity</b>	<b>30% reduced rail yard capacity</b>	<b>60% reduced rail yard capacity</b>
	Reliability, days	96.63%	96.37%	96.12%	95.61%
	Reliability, years	0.00%	0%	0%	0%
Total Cost, K\$	Average	32547	32559	32564	32572
	Standard Deviation	498.26	497.41	497.18	497.48
Total Energy Consumption, MJ	Average	$4.695 \times 10^8$	$4.707 \times 10^8$	$4.716 \times 10^8$	$4.736 \times 10^8$
	Standard Deviation	$6.571 \times 10^6$	$6.566 \times 10^6$	$6.669 \times 10^6$	$6.591 \times 10^6$
Total Emission, kg GHG	Average	$4.053 \times 10^7$	$4.064 \times 10^7$	$4.072 \times 10^7$	$4.089 \times 10^7$
	Standard Deviation	$5.664 \times 10^5$	$5.659 \times 10^5$	$5.654 \times 10^5$	$5.680 \times 10^5$



**Figure 5.7.** Components of total cost, total fuel consumption and total greenhouse gas emissions for the baseline scenario

Similarly, another sensitivity analysis is conducted in which rail use is reduced by 10%, 25%, 50% and 100% reduction (no rail scenario) for comparison with the baseline scenario, in which it is assumed that there are 882 available rail trips annually. In other words, the number of available rail trips is 882 for a baseline scenario, 794 for the 10% reduction scenario, 662 for the 25% reduction scenario, 441 for the 50% reduction scenario and none for the “no rail” scenario.

Results in Table 5.7 show that reducing the rail use reduces the reliability of providing feedstock to the processing facility. The reliability does not decrease rapidly, however, because even with reduced rail availability, the trucks which in the baseline scenario transport logs from harvesting areas to rail yards, are able to transport logs to the facility yard directly, and thus logs may be kept in the facility yard instead of rail yards prior to spring break-up. As before, the averages of total cost, total energy consumption and total greenhouse gas emissions increase as the rail use decreases

**Table 5.7**

To present summary of simulation results for rail use reduction.

		<b>Baseline scenario</b>	<b>10% reduced rail use</b>	<b>25% reduced rail use</b>	<b>50% reduced rail use</b>	<b>no rail use</b>
	Reliability, days	95.41%	95.30%	95.25%	94.17%	92.38%
	Reliability, years	0%	0%	0%	0%	0%
Total Cost, K\$	Average	31908.23	31964.38	32057.06	32146.77	32416.2
	Standard Deviation	455.92	469.89	451.81	459.23	462.62
Total Energy Consumption, MJ	Average	4.646x10 <sup>8</sup>	4.661x10 <sup>8</sup>	4.689x10 <sup>8</sup>	4.774x10 <sup>8</sup>	5.024x10 <sup>8</sup>
	Standard Deviation	6.272x10 <sup>6</sup>	6.518x10 <sup>6</sup>	6.256x10 <sup>6</sup>	6.638x10 <sup>6</sup>	6.727x10 <sup>6</sup>

Table 5.7 (Continued)

		<b>Baseline scenario</b>	<b>10% reduced rail use</b>	<b>25% reduced rail use</b>	<b>50% reduced rail use</b>	<b>no rail use</b>
Total Emission, kg GHG	Average	4.012x10 <sup>7</sup>	4.024x10 <sup>7</sup>	4.049x10 <sup>7</sup>	4.123x10 <sup>7</sup>	4.341x10 <sup>7</sup>
	Standard Deviation	5.413x10 <sup>5</sup>	5.625x10 <sup>5</sup>	5.400x10 <sup>5</sup>	5.733x10 <sup>5</sup>	5.810x10 <sup>5</sup>

## 5.4. Variability in Simulation Results

To use Arena, the computer system must meet these hardware requirements: hard drives with at least 250MB free disk space, at least 1GB RAM, and processor 2GHz or faster (12). However, regardless of the performance of system conditions, only one simulation can be run at a time in Arena. Computer speed and memory impact the model computational time significantly; as an example, the performance of two systems is compared in Table 5.8.

**Table 5.8**

Comparison of model computational time is presented.

	<b>RAM memory</b>	<b>Processor type</b>	<b>Processor speed</b>	<b>Simulation time for 100 replication</b>
Computer 1	4.00 GB	Intel® Core™ i5-2500 CPU	3.30GHz	2 hours
Computer 2	2.00 GB	Intel® Core™ 2 CPU 6600	2.40GHz	6 hours

To check model accuracy, a simulation of 300 replications is run, and distributional results are compared, as shown in part in Table 5.9. The table shows very little variability in the mean values, and some variability in the standard deviations, between sets of 100

replications. Thus, considering both the computational time and simulation accuracy, a replication number of 100 is recommended.

**Table 5.9**

Results of simulation of 300 replications are compared.

County	replications	Start Day		End Day	
		Mean	Stdv	Mean	Stdv
Alcona	1-100	66.33	7.15	108.54	5.06
	101-200	66.84	6.53	109.02	5.21
	201-300	66.98	6.42	108.28	4.70
Chippewa	1-100	69.68	4.79	117.71	10.16
	101-200	70.86	6.25	117.98	9.70
	201-300	70.35	5.55	117.88	7.88
Benzie	1-100	66.25	5.54	88.67	6.96
	101-200	66.67	6.09	88.86	6.12
	201-300	65.82	6.73	88.13	6.17

# Chapter 6. Conclusion and Future Work

## 6.1. Conclusion

A supply chain simulation model coupled with an optimization model has been developed to better understand the supply chain system for the CoEE project. A spring break-up statistical sub-model is employed to simulate the spring break-up timing in a practical way. The model considers the load restriction timing distributions of Alcona County, and then relates other counties to Alcona County. Variable timing of spring break-up and its effects on the biomass supply chain are presented comprehensively by the simulation model.

Sensitivity analyses demonstrate the potential utility of the spring break-up model in supporting decisions under uncertainty. If the simulation is run with fixed spring break-up timing, the results are impractical as the results are based on perfect knowledge. In practice, uncertainty in spring break-up reduces the reliability of providing feedstock to the processing facility.

Also, so as to better understand the expected supply chain system performance, results of the total cost, energy consumption, greenhouse gas emissions and reliability are compared for a number of scenarios of rail use and rail yard storage capacity. The results show that less rail use or smaller rail yard capacity leads to higher total cost, more energy consumption and emissions, and lower reliability.

## 6.2. Recommendations for Future Work

One factor that is currently considered in the optimization model, but not the simulation model, is different land ownership categories that are expected to affect the timing and spatial distribution of feedstock availability. The optimization model includes four distinct ownership categories: federal, state, private industrial and private non-industrial. Different types of land owners are expected to have different levels of willingness to



harvest their timber at offered prices, so it is very likely that timber will be acquired at different costs from different land owners. Therefore, one recommendation for the simulation model is to include the types of land ownership in order to adjust harvest plans with more accurate consideration of harvest costs.

The statistical spring break-up model developed in this study has the potential to be extended as a seasonal forecast model for Frontier and county road commissions, as spring break-up timing has been shown to be associated with weather conditions, and seasonal climate forecasts are improving in many parts of the world (22, 42). It may take many years, however, for Frontier and road commissions to determine how and to what extent they will rely on the forecasts for their decisions. In the meantime, the compilation of spring break-up data should continue, and the model should be routinely updated, as the model is currently based on a very small sample of historical data (5-9 years, depending on the county). In addition, having recent spring break-up data will be even more important if a climate change trend becomes apparent in the region.

Another direction of future work is to use the simulation model as an educational and communication tool, e.g., helping loggers better plan timber harvests, helping storage yard managers maintain proper inventory levels, and helping processing facility managers make decisions regarding production. In turn, future work may focus on representing aspects of communication across the supply chain, as effective communication among supply chain agents is expected to improve responsiveness and reduce total supply chain costs.

There are two additional limitations on the simulation model. First, the model does not consider wood from Canada. Chippewa County, where the bio-fuel processing facility is to be located, is adjacent to the Canadian province of Ontario. Therefore, forests in Canada might be a potential biomass source. Another limitation is that competition with other mills in the area has not been considered. There will be a reduced quantity of available biomass for conversion to ethanol if pulp and paper mills or energy producers are demanding biomass within the region as well. Future work may focus on how to include these considerations in the simulation model.

# References

1. Athanasios AR, Athanasios JT, Ilias PT. Logistics issue of biomass: The storage problem and the multi-biomass supply chain. *Renewable and sustainable energy reviews*. 2009;13:887-894.
2. Yu Y, Bartle J, Li CZ, Wu HW. Mallee Biomass as a key bioenergy source in Western Australia: Importance of biomass supply chain. *Energy & Fuels*. 2009;23:3290-3299.
3. Iakovou E, Karagiannidis A, Vlachos D, Toka A, Malamakis A. Waste biomass-to-energy supply chain management: a critical synthesis. *Waste Management*. 2010;30:1860-1870.
4. Demirbas A. The importance of bioethanol and biodiesel from biomass. *Energy Sources, part B*. 2008;3:177-185.
5. Demirbas. Conversion of biomass using glycerin to liquid fuel for blending gasoline as alternative engine fuel. *Energy Conversion and Management*. 2000;41(16):1741-1748.
6. Aresta A, Dibenedetto A, Carone M, Colonna T, Fragale C. Production of biodiesel from macroalgae by supercritical CO<sub>2</sub> extraction and thermochemical liquefaction. *Environmental Chemistry Letters*. 2005;3(3):136-139.
7. Watkins D, et al. Feedstock Supply Chain Models for the Frontier Renewable Resources Biofuel Facility in Kinross, Michigan. Michigan Technological University. 2012.
8. J.M. Longyear [Internet]. Marquette(MI): JML Heirs, LLC. 2012 [cited 2012 June 30]. Available from: <http://jmlongyear.com/index.asp?cookies=True>
9. Frontier Renewable Resources[Internet]. Marquette(MI): Frontier Renewable Resources, LLC. 2012 [cited 2012 June 30]. Available from: <http://www.frontier-renewable.com>
10. Banerjee A, Nogueer JL. Biofuels Supply Chain Characterization. Massachusetts Institute of Technology. 2007.
11. Christopher M, Gattorna J. Supplychaincost management and value-based pricing. *Industrial Marketing Management*. 2005;34(2):115-121.

12. Kelton WD, Sadowski RP, Swets NB. Simulation with Arena. 5th ed. Columbus (OH): McGraw Hill; 2009. 636 p.
13. Michigan Department of Transportation (MDOT) [Internet]. Lansing (MI): State of Michigan. 2011 [cited 2012 May 31]. Available from:  
<http://www.michigan.gov/mdot/>
14. Michigan State Policy [Internet]. Lansing (MI): State of Michigan. 2011 [cited 2012 May 31]. Available from: [http://www.michigan.gov/msp/0,1607,7-123-1586\\_1710-87560--,00.html](http://www.michigan.gov/msp/0,1607,7-123-1586_1710-87560--,00.html)
15. Michigan Legislature Website [Internet]. Michigan Legislative Council, the Michigan House of Representatives, the Michigan Senate. Michigan Vehicle Code (Excerpt) Act 300 of 1949; 2011 [cited 2012 April 30]. Available from:  
<http://www.legislature.mi.gov/%28S%28bx5zhk55hb0b5545ubt2f555%29%29/mileg.aspx?page=getobject&objectname=mcl-257-722&queryid=1945485&highlight=seasonal+weight>
16. County Road Association of Michigan [Internet]. Lansing (MI): County Road Association of Michigan. 2009 [updated 2009 Nov, cited 2012 June 30]. Available from: <http://www.micountyroads.org/>
17. Weather Data Depot [Internet]. State College (PA): EnergyCAP Enterprise. 2012 [cited 2012 April 30]. Available from: <http://www.weatherdatadepot.com/>
18. Kestler, MA, Knight T, Krat AS. Thaw Weakening and Load Restriction Practices on Low Volume Roads. ERDC/CRREL Technical Report TR-00-6. Engineer Research and Development Center (ERDC), Cold Regions Research and Engineering Laboratory (CRREL). 2000.
19. Kestler MA, Richard LB, etc. Keeping Springtime Low-Volume Road Damage to a Minimum. Transportation Research Record, 2011;2205:155–164.
20. Steinert BC, Humphrey DN, Kestler MA. Field and Laboratory Evaluation of the Portable Falling Weight Deflectometer: A Study for the New England Transportation Consortium. Department of Civil and Environmental Engineering, University of Maine, Orono. 2005.

21. Berg RL, Kestler MA, Eaton RA, Benda CC. Estimating When to Apply and Remove Spring Load Restrictions. ASCE Cold Regions Engineering Conference, Bangor, Maine. 2006.
22. Minnesota Department of Transportation (Mn/DOT) [Internet]. Minneapolis (MN): Mn/DOT. Guidelines for Seasonal Load Limit Starting and Ending Dates; 2004 [cited 2012 June 30]. Available from: [http://www.mrr.dot.state.mn.us/pavement/pvmtdesign/techmemos/tm03\\_02\\_mrr\\_01.pdf](http://www.mrr.dot.state.mn.us/pavement/pvmtdesign/techmemos/tm03_02_mrr_01.pdf)
23. GEO-SLOPE. Thermal Modeling with Temp/W: An Engineering Methodology. 3rd ed. Calgary (Alta): GEO-SLOPE International Ltd. 2007.
24. Michigan State Climatologist's Office [Internet]. Lansing (MI): Department of Geography, MSU. 1990 [updated 2009 June 15, cited 2012 June 30]. Available from: <http://climate.geo.msu.edu/mi-map.html>
25. Wurbs RA, James WP. Water Resources Engineering. 1st ed. Upper Saddle River (NJ): Prentice Hall; 2002. 828p.
26. ARA Inc, ERES Consultants Division. Guide for Mechanistic–Empirical Design of New and Rehabilitated Pavement Structures. Final report, NCHRP Project 1-37A. Transportation Research Board of the National Academies, Washington D.C.; 2004.
27. Gibbons J. Pavements and Surface Materials. Uconn Extension Land Use Educator. 1999 [cited 2012 June]. Available from: [http://nemo.uconn.edu/publications/tech\\_papers/tech\\_paper\\_8.pdf](http://nemo.uconn.edu/publications/tech_papers/tech_paper_8.pdf)
28. Indiana Department of Transportation (INDOT) [Internet]. Highway Certified Technician Program Training Manual; 2011 November [cited 2012 June]. Available from: <http://www.mtu.edu/gradschool/administration/academics/thesis-dissertation/pdfs/procedures.pdf>
29. American Association of State Highway and Transportation Officials (AASHTO). AASHTO Guide for Design of Pavement Structures 1993. 2nd ed. American Association of State Highway and Transportation Officials. Washington (DC): Amer Assn of State Hwy; 1993. 640p.

30. Ovik JM, Deusen AD, Siekmeier JA. Improved Spring Load Restriction Guidelines Using Mechanistic Analysis. Minnesota Department of Transportation. 2000. Available from: <http://www.mtu.edu/gradschool/administration/academics/thesis-dissertation/pdfs/procedures.pdf>
31. Michigan Department of Natural Resources (DNR) [Internet]. Lansing (MI): State of Michigan. Lake maps by county; 2001 [cited 2012 June]. Available from: [http://michigan.gov/dnr/0,1607,7-153-30301\\_31431\\_32340---,00.html](http://michigan.gov/dnr/0,1607,7-153-30301_31431_32340---,00.html)
32. Dodson S. Introduction to Limnology. 1st ed. Columbus (OH): McGraw Hill; 2004. 416p.
33. Box GEP, Cox DR. An Analysis of Transformation. *Journal of the Royal Statistical Society Series B (Methodological)*. 1964;26(2):211-252.
34. Minitab® 16.1.1 User Manual. Minitab Inc.
35. McDonald JH. Handbook of Biological Statistics. 2nd ed. Baltimore (MD): Sparky House Publishing; 2009.
36. Casella G, Berger RL. Statistical Inference. 2nd ed. Pacific Grove (CA): Duxbury Press; 2001. 660p.
37. Francis DP, Coats AJ, Gibson D. How high can a correlation coefficient be? *International Journal of Cardiology*. 1999;69(2):185-199.
38. Motulsky H. Analyzing Data with GraphPad Prism. Graphpad Software Inc. 1999.
39. Stedinger JR, Lettenmaier DP, Vogel RM. Multisite ARMA(1,1) and Disaggregation Models for Annual Streamflow Generation. *Water Resource Research*. 1985;21(4):497-509.
40. Douglas CM, George CR, Norma FH. Engineering Statistics. 4th ed. Montgomery (MD): John Wiley & Sons Ltd; 2007. 487p.
41. Ayers GP. Comment on Regression Analysis of Air Quality Data. *Atmospheric Environment*. 2002;93(2):153-154.
42. Baiz S, Tighe SL, Mills B, Haas CT, Huen K. Using Road Weather Information Systems (RWIS) to Control Load Restrictions on Gravel and Surface-Highways. Highway Innovation Infrastructure Funding Program, Ministry of Transportation of Ontario, Canada. 2007.

43. John CD. *Statistics and Data Analysis in Geology*. 3rd ed. Montgomery (MD): John Wiley & Sons Ltd; 2002. 656p. Kansas Geological Survey. The University of Kansas.

# Appendix

## Appendix A: Historical Spring Break-up Data

**Table A.1**

Historical spring break-up data includes start day and duration.

H.A. # and County		Year: 20-									
		10	09	08	07	06	05	04	03	02	01
1 Alcona	Start Day	62	68	77	71	68	73	61	74	52	67
	Duration	36	50	43	45	40	37	50	45	63	47
2/3/4 Alger	Start Day	67	76	78	71	69					
	Duration	43	55	56	53	49					
5/6 Alpena	Start Day	65	68	77	71	67	74	75	80	77	
	Duration	31	46	38	41	41	35	31	35	36	
7 Antrim	Start Day	62	68	71	67	65	73	61	73	50	
	Duration	24	40	42	27	38	29	43	21	56	
8 Benzie	Start Day	67	68	72	68	67	73	61	73	72	
	Duration	11	29	34	22	20	25	31	18	31	
9 Charlevoix	Start Day	62	68	70	57	65					
	Duration	30	51	40	52	43					
10 Cheboygan	Start Day	67	72	77	71	67					
	Duration	25	43	37	36	37					
11/12 Chippewa	Start Day	67	76	78	72	68					
	Duration	19	41	38	30	35					
13/14 Crawford	Start Day	67	65	72	81	65					
	Duration	29	56	49	48	56					
15/16 Delta	Start Day	60	72	72	68	67					
	Duration	51	50	56	54	70					
17 Emmet	Start Day	61	68	70	60	60	60				
	Duration	28	40	43	61	36	49				
18 Grand Traverse	Start Day	69	68	74	64	67	73	61			
	Duration	15	40	32	29	27	25	50			

Table A.1 (Continued)

H.A. # and County		Year: 20-									
		10	09	08	07	06	05	04	03	02	01
19	Start Day	67	68	77	71	69					
Iosco	Duration	25	43	38	40	27					
20/21	Start Day	68	65	77	71	65					
Kalkaska	Duration	24	46	37	43	38					
22	Start Day	No Spring Break-up Data									
Leelanau	Duration	No Spring Break-up Data									
23/24	Start Day	68	77	78	74	68					
Luce	Duration	46	59	57	90	76					
25/26/27	Start Day	67	76	77	72	69					
Mackinac	Duration	33	60	86	74	104					
28	Start Day	No Spring Break-up Data									
Manistee	Duration	No Spring Break-up Data									
29	Start Day	67	76	78	71	68					
Marquette	Duration	47	63	66	57	61					
30	Start Day	60	70	71	68	66					
Menominee	Duration	43	48	50	53	49					
31	Start Day	64	65	72	71	65					
Missaukee	Duration	15	35	38	22	30					
32	Start Day	57	65	77	60	60					
Montmorency	Duration	32	54	33	30	36					
33	Start Day	67	64	77	66	67	68	61	76	49	
Ogemaw	Duration	29	51	40	41	34	41	50	39	59	
34/35	Start Day	67	70	77	71	72					
Oscoda	Duration	32	48	43	45	32					
36/37	Start Day	62	68	72	68	68	73	61	76	53	
Otsego	Duration	27	44	48	26	35	31	45	22	60	
38/39	Start Day	64	68	77	69	66	73	59	73	75	
Presque Isle	Duration	32	47	40	41	42	36	45	40	39	
40	Start Day	64	68	77	68	67					
Roscommon	Duration	25	38	32	39	29					
41/42	Start Day	57	77	77	72	67					
Schoolcraft	Duration	62	62	78	49	57					
43	Start Day	64	68	73	72	72					
Wexford	Duration	25	38	36	25	22					



## Appendix B: Mathematica Code for Simulation Results vs. Historical Data

(1) For creating the comparison graph of start day:

```
data=RandomReal[NormalDistribution[522626,161887],1000];  
data1=data^(1/3.12);  
(*historical data*)  
data2={62,68,77,71,68,73,61,74,52,67};  
Histogram[{data1,data2},{45,85,5},"Probability",AxesLabel->{Data,Density},  
ChartStyle->{Blue,White}, ChartLegends->{"simulated data","County Alcona"},  
PlotLabel->"Start Day"]  
Mean[data1] //N  
StandardDeviation[data1] //N  
Mean[data2] //N  
StandardDeviation[data2] //N
```

(2) For creating the comparison graph of end day:

```
data=RandomReal[NormalDistribution[3943850000000000, 947260000000000],1000];  
data1=Exp[data^(1/21.65)];  
data2={97,117,119,115,107,109,110,118,114,113};  
Histogram[{data1,data2},{85,130,5},"Probability",AxesLabel->{Data,Density},  
ChartStyle->{Blue,White}, ChartLegends->{"simulated data","County Alcona"},  
PlotLabel->"End Day"]  
Mean[data1] //N  
StandardDeviation[data1] //N  
Mean[data2] //N  
StandardDeviation[data2] //N (*historical data*)
```

## Appendix C: Transformations and simple linear regression relationships

**Table C.1**

Results of the Box-Cox transformation for spring break-up timing include transformation parameter and distribution after transformation.

Area #	County Name, (Y)	Start Day		End Day	
		$\lambda$	Distribution after transformation	$\lambda$	Distribution after transformation
1	Alcona	3.12	Norm(522626, 161887)	1	$\text{Ln}(Y)^{21.65} \sim \text{Norm}(3.94385\text{E}14, 9.4726\text{E}13)$
2/3/4	Alger	-5.00	Norm(5.35011E-10, 1.52333E-10)	0.00	Norm(4.80472, 0.0725125)
5/6	Alpena	0.47	Norm(7.49173, 0.204328)	5.00	Norm(1.56251E10, 3.04358E9)
7	Antrim	0.90	Norm(43.1245, 4.33934)	2.41	Norm(66935.2, 12984.1)
8	Benzie	5.00	Norm(1.60763E9, 4.1976E8)	-0.71	Norm(0.0403566, 0.00262747)
9	Charlevoix	5.00	Norm(1.16255E9, 4.26811E8)	0.47	Norm(8.96747, 0.333643)
10	Cheboygan	-5.00	Norm(5.84371E-10, 1.58528E-10)	3.09	Norm(1.8157E6, 387581)
11/12	Chippewa	-3.71	Norm(1.31282E-7, 3.01057E-8)	3.44	Norm(1.43249E7, 4.41341E6)
13/14	Crawford	-5.00	Norm(6.53599E-10, 2.48784E-10)	4.70	Norm(2.88462E9, 9.52551E8)
15/16	Delta	1.00	Norm(67.8, 3.76773)	-0.37	Norm(0.168758, 0.00427354)
17	Emmet	-5.00	Norm(1.05413E-9, 2.26955E-10)	0.96	Norm(87.1476, 9.27338)
18	Grand Traverse	-3.10	Norm(2.13975E-6, 4.62184E-7)	0.68	Norm(22.5978, 1.32135)
19	Iosco	-5.00	Norm(5.98306E-10, 1.42101E-10)	2.00	Norm(9724.54, 1697.01)
20/21	Kalkaska	-5.00	Norm(6.67038E-10, 2.10725E-10)	5.00	Norm(1.4047E10, 4.33679E9)
22	Leelanau		assume same as <b>Antrim County's</b> <sup>(2)</sup>		
23/24	Luce	0.39	Norm(5.32741, 0.126309)	-2.02	Norm(4.99521E-5, 1.08714E-5)
25/26/27	Mackinac	-1.95	Norm(0.000239581, 2.55535E-5)	1.11	Norm(246.572, 45.3304)

Table C.1 (Continued)

Area #	County Name, (Y)	Start Day		End Day	
		$\lambda$	Distribution after transformation	$\lambda$	Distribution after transformation
28	Manistee		assume same as <b>Wexford County's</b> <sup>(2)</sup>		
29	Marquette	-5.00	Norm(5.44694E-10, 1.63064E-10)	-0.38	Norm(0.157652, 0.00480212)
30	Menominee	4.70	Norm(3.93222E8, 9.31917E7)	5.00	Norm(2.03972E10, 4.31548E9)
31	Missaukee	-5.00	Norm(7.4522E-10, 1.68338E-10)	-1.12	Norm(0.00622604, 0.000802048)
32	Montmorency	1.00	Norm(63.8, 7.91833)	-2.79	Norm(2.83177E-6, 9.37396E-7)
33	Ogemaw	0.00	Norm(4.18385, 0.131)	3.24	Norm(3.8991E6, 736502)
34/35	Oscoda	-5.00	Norm(5.55235E-10, 1.31533E-10)	3.23	Norm(4050350, 940177)
36/37	Otsego	0.31	Norm(3.67408, 0.124475)	0.23	Norm(2.9038, 0.0635796)
38/39	Presque Isle	3.30	Norm(1.21618E6, 318317)	5.00	Norm(1.55268E10, 3.41793E9)
40	Roscommon	-5.00	Norm(6.83403E-10, 2.02085E-10)	4.03	Norm(1.207E8, 3.35034E7)
41/42	Schoolcraft	3.59	Norm(4.42099E6, 1.38272E6)	-5.00	Norm(3.01889E-11, 1.38117E-11)
43	Wexford	5.00	Norm(1.69412E9, 2.52101E8)	-3.08	Norm(7.6328E-7, 1.77104E-7)
44	area 1>150 miles		assume same as <b>Marquette County</b> <sup>(2)</sup>		
45	area 2>150 miles		assume same as <b>Marquette County</b> <sup>(2)</sup>		
46	area 3>150 miles		assume same as <b>Marquette County</b> <sup>(2)</sup>		

**Table C.2**

Relationships for transformed start day of spring break-up are developed by simple linear regression.

Harvesting Area #	County Name (Y)	Based on (X)	Omitted Year/Years	Corr Coef	Equation
1	Alcona	from <b>Alcona County's Spring Break-up Model</b>			
2/3/4	Alger	Chippewa		0.99	$y = 0.0051x - 1E-10$
5/6	Alpena	Alcona	2002 <sup>(1)</sup>	0.61	$y = 1E-06x + 6.8322$
7	Antrim	Alcona		0.93	$y = 2E-05x + 34.138$
8	Benzie	Alcona	2002 <sup>(1)</sup>	0.88	$y = 2712.2x + 4E+07$
9	Charlevoix	Montmorency		0.82	$y = 4E+07x - 2E+09$
10	Cheboygan	Presque Isle		-0.91	$y = -5E-16x + 1E-09$
11/12	Chippewa	Presque Isle		-0.81	$y = -9E-14x + 2E-07$
13/14	Crawford	Alcona		-0.53	$y = -1E-15x + 1E-09$
15/16	Delta	Chippewa		-0.90	$y = -1E+08x + 86.107$
17	Emmet	Montmorency		-0.85	$y = -4E-11x + 3E-09$
18	Grand Traverse	Alcona		-0.67	$y = -2E-12x + 3E-06$
19	Iosco	Alcona		-0.98	$y = -1E-15x + 1E-09$
20/21	Kalkaska	Alcona		-0.73	$y = -1E-15x + 1E-09$
22	Leelanau	assume the same start day as <b>Antrim County's</b> <sup>(2)</sup>			
23/24	Luce	Chippewa		-0.99	$y = -4E+06x + 5.8861$
25/26/27	Mackinac	Chippewa		0.99	$y = 874.31x + 0.0001$
28	Manistee	assume the same start day as <b>Wexford County's</b> <sup>(2)</sup>			
29	Marquette	Chippewa		1.00	$y = 0.0054x - 2E-10$
30	Menominee	Chippewa		-0.94	$y = -3E+15x + 8E+08$
31	Missaukee	Alcona		-0.89	$y = -1E-15x + 1E-09$
32	Montmorency	Alcona		0.88	$y = 5E-05x + 35.489$
33	Ogemaw	Alcona		0.90	$y = 7E-07x + 3.8213$

Table C.2 (Continued)

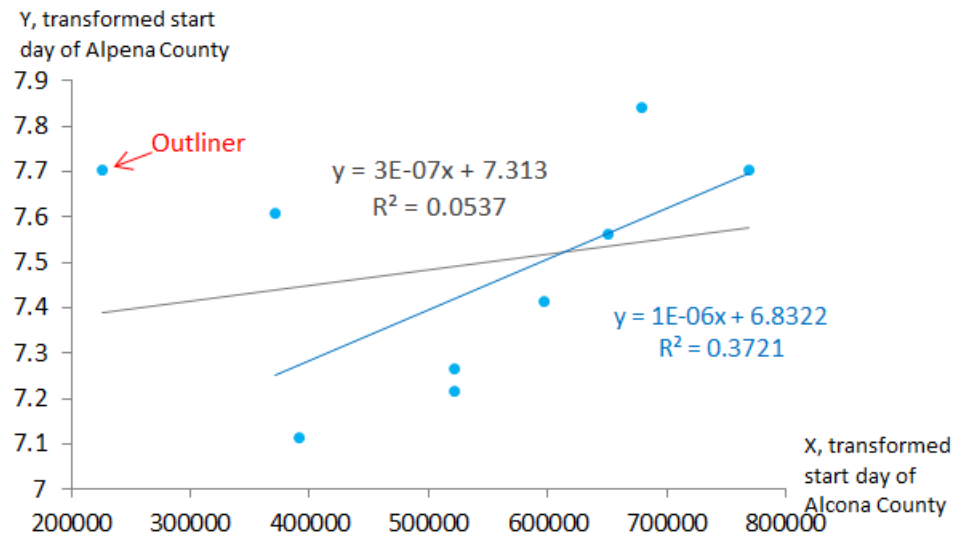
Harvesting Area #	County Name (Y)	Based on (X)	Omitted Year/Years	Corr. Coef.	Equation
34/35	Oscoda	Alcona		-0.95	$y = -9E-16x + 1E-09$
36/37	Otsego	Alcona		0.95	$y = 7E-07x + 3.3132$
38/39	Presque Isle	Alcona	2002 <sup>(1)</sup>	0.97	$y = 2.2132x - 69902$
40	Roscommon	Alcona		-0.98	$y = -1E-15x + 1E-09$
41/42	Schoolcraft	Chippewa		-0.96	$y = -5E+13x + 1E+07$
43	Wexford	Alcona		0.83	$y = 2521.9x + 3E+08$
44	area 1>150 miles	assume the same start day as			<b>Marquette County</b> <sup>(2)</sup>
45	area 2>150 miles	assume the same start day as			<b>Marquette County</b> <sup>(2)</sup>
46	area 3>150 miles	assume the same start day as			<b>Marquette County</b> <sup>(2)</sup>

**Table C.3**

Relationships for transformed end day of spring break-up are developed by simple linear regression.

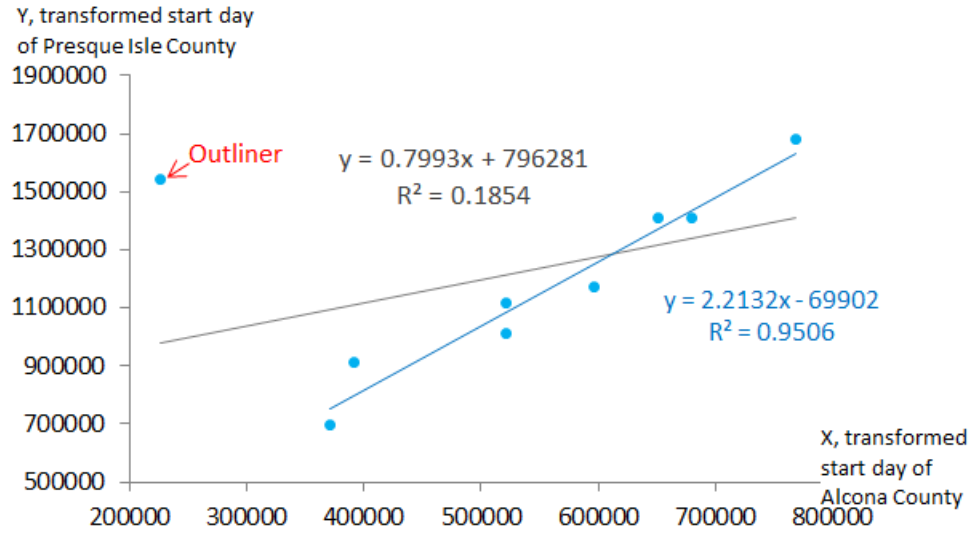
<b>Harvesting Area #</b>	<b>County Name (Y)</b>	<b>Based on (X)</b>	<b>Corr. Coef.</b>	<b>Equation</b>
1	Alcona	from <b>Alcona County's Spring Break-up Model</b>		
2/3/4	Alger	Chippewa	0.87	$y = 1E-08x + 4.5982$
5/6	Alpena	Alcona	0.97	$y = 4E-05x + 1E+09$
7	Antrim	Cheboygan	0.88	$y = 0.0324x + 7983.9$
8	Benzie	Alcona	-0.75	$y = -2E-17x + 0.0484$
9	Charlevoix	Alcona	0.85	$y = 3E-15x + 7.9606$
10	Cheboygan	Alcona	0.96	$y = 3E-09x + 483639$
11/12	Chippewa	Cheboygan	0.94	$y = 9.9172x - 4E+06$
13/14	Crawford	Alcona	0.94	$y = 8E-06x - 3E+08$
15/16	Delta	Chippewa	-0.68	$y = -7E-10x + 0.1784$
17	Emmet	Alcona	0.83	$y = 7E-14x + 62.193$
18	Grand Traverse	Alcona	0.69	$y = 1E-14x + 18.906$
19	Iosco	Alcona	0.98	$y = 2E-11x + 4743.6$
20/21	Kalkaska	Alcona	0.96	$y = 4E-05x - 1E+09$
22	Leelanau	assume the same end day as <b>Antrim County's</b> <sup>(2)</sup>		
23/24	Luce	Chippewa	-0.63	$y = -2E-12x + 8E-05$
25/26/27	Mackinac	Chippewa	0.70	$y = 8E-06x + 135.4$
28	Manistee	assume the same end day as <b>Wexford County's</b> <sup>(2)</sup>		
29	Marquette	Chippewa	-0.93	$y = -1E-09x + 0.1726$
30	Menominee	Chippewa	0.80	$y = 941.74x + 7E+09$
31	Missaukee	Alcona	-0.91	$y = -6E-18x + 0.0086$
32	Montmorency	Alcona	-0.71	$y = -5E-21x + 5E-06$
33	Ogemaw	Alcona	0.90	$y = 7E-09x + 1E+06$
34/35	Oscoda	Alcona	0.99	$y = 8E-09x + 949364$
36/37	Otsego	Alcona	0.59	$y = 4E-16x + 2.7559$
38/39	Presque Isle	Alcona	0.92	$y = 4E-05x + 7E+08$
40	Roscommon	Alcona	0.98	$y = 3E-07x + 9E+06$
41/42	Schoolcraft	Chippewa	-0.75	$y = -2E-18x + 6E-11$
43	Wexford	Alcona	-0.96	$y = -1E-21x + 1E-06$
44	area 1>150 miles	assume the same end day as <b>Marquette County's</b> <sup>(2)</sup>		
45	area 2>150 miles	assume the same end day as <b>Marquette County's</b> <sup>(2)</sup>		
46	area 3>150 miles	assume the same end day as <b>Marquette County's</b> <sup>(2)</sup>		

(1). Year 2002 is omitted for Alpena County, because the outlying observation from year 2002 ( $52^{3.12}$ ,  $37^{0.47}$ ) appears to deviate from other points markedly,  $R^2$  increases from 0.053 to 0.372 without the outlier.



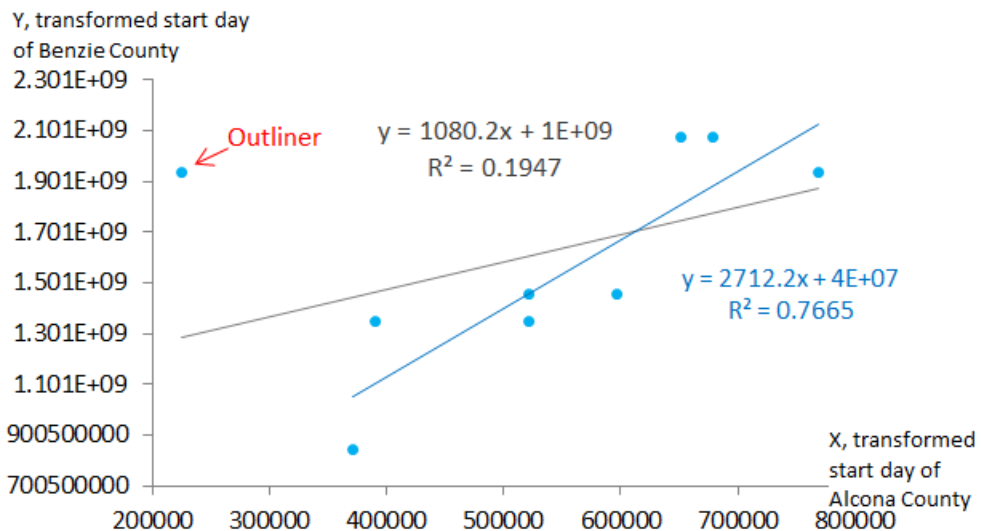
**Figure C.1.** Relationship between transformed start days of Alpena County and Alcona County

Similarly, year 2002 is omitted for Presque Isle County as well, because the outlying observation from year 2002 ( $52^{3.12}$ ,  $75^{3.3}$ ) appears to deviate from other points markedly. Without the outlier,  $R^2$  increases from 0.1854 to 0.9506, which indicates highly correlated relationship.



**Figure C.2.** Relationship between transformed start days of Presque Isle County and Alcona County

Again, year 2002 is omitted for Benzie County because the outlying observation from year 2002 ( $52^{3.12}$ ,  $72^5$ ) appears to deviate from other points markedly. Without the outlier,  $R^2$  increases from 0.1947 to 0.7665.

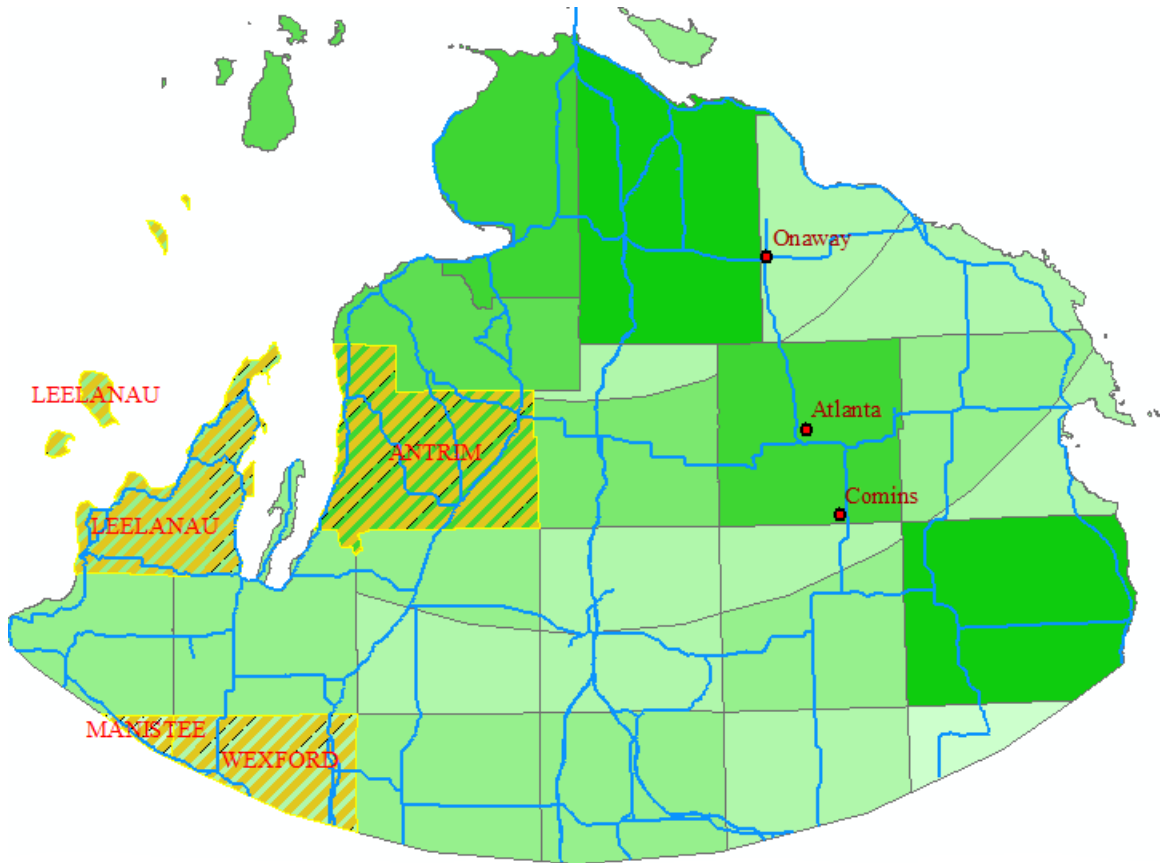


**Figure C.3.** Relationship between transformed start days of Benzie County and Alcona County



(2). The start day and end day of spring break-up of Leelanau County is assumed to be the same as the Antrim County's, due to a lack of historical data for Leelanau County. As stated in Chapter 3, the timing of spring break-up is partly determined by the weather conditions, and Leelanau County has very similar weather to Antrim County because of their close geographic location, as shown in Figure C.3. Leelanau County is adjacent to Antrim County to the west.

Similarly, the spring break-up of Manistee County is assumed to be started on the same day and ended on the same day with Wexford County due to a lack of historical data for Manistee County. Manistee County is adjacent to Wexford County to the west as shown in Figure C.3.



**Figure C.4.** Map of Leelanau County, Antrim County, Manistee County and Wexford County

Harvesting areas 44 to 46 are assumed to have the same spring breakup timing as Marquette County because we do not know the exact locations of these three harvesting areas. These three areas may be specified by the user for harvests beyond the 150-mile radius haul zone in the Michigan's Upper Peninsula in the Arena simulation model.

## Appendix D: Mathematica Code used to obtain RMA regression equation

(1) The RMA regression calculation procedure is as follows, using the transformed start days of Alcona County and Antrim County as an example:

**Table D.1**

Sums of squares and other quantities for transformed start days of Alcona County and Antrim County are listed.

	Alcona County, X	Antrim County, Y
Sum	$\sum X = 4728120$	$\sum Y = 388$
Mean	$\bar{X} = 525346.71$	$\bar{Y} = 43.12$
Variance of the sample mean	$s_X^2 = 29400019704$	$s_Y^2 = 18.83$
Standard error of the mean	$S_X = 171464.34$	$S_Y = 4.34$
Sum of squares	$SS_X = 235200157635$	$SS_Y = 150.64$
Covariance		$COV_{XY} = 694979.59$
Sum of products		$SP_{XY} = 5559836.70$
Correlation coefficient		$\gamma = 0.9340598$

\* Sum:  $\sum X = X_1 + \dots + X_i + \dots + X_n$ , for  $i = 1, 2, \dots, n$ .

Mean:  $\bar{X} = \sum X / n$ , where  $n$  is the sample size.

Variance of the sample mean:  $s_X^2 = \frac{\sum_{i=1}^n X_i^2 - n\bar{X}^2}{n-1}$ , which can be regarded as the average squared deviation of all possible observations from the sample mean.

Standard error of the mean:  $s_X = \sqrt{s_X^2}$ , which provides a statistic that describes dispersion or spread of data around the sample mean.

Sum of squares:  $SS_X = \sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}$ , which is a generated intermediate quantity, for example the variance can be found by division by  $(n-1)$ :  $s_X^2 = \frac{SS_X}{n-1}$ .

Sum of products:  $SP_{XY} = \sum_{i=1}^n (X_i Y_i) - \frac{\sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n}$ , which is generated in order to calculate covariance.

Covariance:  $Cov_{XY} = \frac{SP_{XY}}{n-1}$ , which is a joint variation of two variables about their common mean.

Correlation coefficient:  $\gamma_{XY} = \frac{Cov_{XY}}{s_X s_Y}$  or  $= \frac{SP_{XY}}{\sqrt{SS_X SS_Y}}$ , which is a dimensionless measure method of two variables to estimate the degree of interrelation between these two variables. (43)

For Y on X regression (simple linear regression) line,  $Y = \alpha_2 X + \beta_2$  is obtained by  $\alpha_2 = SP_{XY} / SS_X = 5559836.70 / 235200157635 = 2.364E-5$ , and  $\beta_2 = \bar{Y} - \alpha_2 \bar{X} = 30.706$ , which are the same parameters from section 4.4.1.

For X on Y regression (inverse linear regression) line,  $X = \alpha_1 Y + \beta_1$  is obtained by  $\alpha_1 = SP_{XY} / SS_Y = 5559836.70 / 150.64 = 36908.378$ , and  $\beta_1 = \bar{X} - \alpha_1 \bar{Y} = -1066309.186$ .

For RMA regression line,  $\alpha_0 = S_Y / S_X = 4.34 / 171464.34 = 2.53075E-5$ , and  $\beta_0 = \bar{Y} - \alpha_0 \bar{X} = 29.8293$ , then  $Y = 0.0000253075 * X + 29.8293$ .

(2) To get the RMA regression parameters using Mathematica 8 ( $\alpha$  and  $\beta$  are the slope and intercept, respectively.):

*x* := Data required here; for example,  
*x* := {912838.427458795, 1115011.66764939, 1680426.00244157, 1170043.47613226, 1010403.83818223}; for the transformed spring break-up start day of the Alcona County's;

*y* := Data required here;  
*n* := 5; (\*size of the sample\*)  
*xsum* := Total[*x* /. {*x*\_, \_} -> *x*]  
*ysum* := Total[*y* /. {*y*\_, \_} -> *y*]  
*xmean* := Mean[*x*]  
*ymean* := Mean[*y*]  
*xvar* := Variance[*x*]  
*yvar* := Variance[*y*]  
*xs* := StandardDeviation[*x*]  
*ys* := StandardDeviation[*y*]  
*spxy* := Total[*x* \* *y* /. {*x*\_, \_} -> *x*] - *xsum* \* *ysum* / *n*  
*gamma* := Correlation[*x*, *y*]  
*alpha0* := (*ys* / *xs*) \* If[*gamma* > 0, 1, -1]  
*beta0* := *ymean* - *alpha0* \* *xmean*

$x_{mean}$  //N  
 $y_{mean}$ //N  
 $y_s$  //N  
 $x_s$  //N  
 $\beta_0$ //N (\*intercept\*)  
 $\alpha_0$ //N (\*slope \*)

## Appendix E: Transformations and RMA regression relationships

**Table E.1**

Relationships for transformed start days of spring breakup are developed by RMA.

Harvesting Area #	County Name (Y)	Lambda	Based on (X)	Omitted Year/Years	Corr. Coef.	Slope	Intercept
1	Alcona	3.12	from <b>Alcona County's Spring Break-up Model</b>				
2/3/4	Alger	-5.00	Chippewa		0.99	5.17E-03	-1.44E-10
5/6	Alpena	0.47	Alcona	2002	0.61	1.84E-06	6.4274
7	Antrim	0.90	Alcona		0.93	2.53E-05	2.98E+01
8	Benzie	5.00	Alcona	2002	0.88	3097.8255	-1.77E+08
9	Charlevoix	5.00	Montmorency		0.82	5.39E+07	-2.28E+09
10	Cheboygan	-5.00	Presque Isle		-0.91	-5.32E-16	1.21E-09
11/12	Chippewa	-3.71	Presque Isle		-0.81	-1.07E-13	2.57E-07
13/14	Crawford	-5.00	Alcona		-0.53	-1.80E-15	1.66E-09
15/16	Delta	1.00	Chippewa		-0.90	-1.54E+08	88.0324
17	Emmet	-5.00	Montmorency		-0.85	-4.28E-11	3.74E-09
18	Grand Traverse	-3.10	Alcona		-0.67	-3.28E-12	3.93E-06
19	Iosco	-5.00	Alcona		-0.98	-1.05E-15	1.19E-09
20/21	Kalkaska	-5.00	Alcona		-0.73	-1.52E-15	1.52E-09
22	Leelanau		assume the same start day as <b>Antrim County's</b>				
23/24	Luce	0.39	Chippewa		-0.99	-4.29E+06	5.89E+00
25/26/27	Mackinac	-1.95	Chippewa		0.99	879.3394	1.24E-04
28	Manistee		assume the same start day as <b>Wexford County's</b>				
29	Marquette	-5.00	Chippewa		1.00	5.45E-03	-1.70E-10
30	Menominee	4.70	Chippewa		-0.94	-3.39E+15	8.38E+08
31	Missaukee	-5.00	Alcona		-0.89	-1.40E-15	1.53E-09
32	Montmorency	1.00	Alcona		0.88	5.73E-05	31.7234
33	Ogemaw	0.00	Alcona		0.90	7.73E-07	3.7777
34/35	Oscoda	-5.00	Alcona		-0.95	-9.70E-16	1.10E-09
36/37	Otsego	0.31	Alcona		0.95	7.26E-07	3.2927
38/39	Presque Isle	3.30	Alcona	2002	0.97	2.2700	-1.02E+05
40	Roscommon	-5.00	Alcona		-0.98	-1.46E-15	1.50E-09
41/42	Schoolcraft	3.59	Chippewa		-0.96	-5.21E+13	1.13E+07
43	Wexford	5.00	Alcona		0.83	3.03E+03	-3.09E+06
44	area 1 >150 miles		assume the same start day as <b>Marquette County's</b>				
45	area 2 >150 miles		assume the same start day as <b>Marquette County's</b>				
46	area 3 >150 miles		assume the same start day as <b>Marquette County's</b>				

**Table E.2**

Relationships for transformed end day of spring breakup are developed by RMA.

<b>Harvesting Area #</b>	<b>County Name (Y)</b>	<b>Lambda</b>	<b>Based on (X)</b>	<b>Corr. Coef.</b>	<b>Slope</b>	<b>Intercept</b>
1	Alcona	4.45	<b>from Alcona County's Spring Break-up Model</b>			
2/3/4	Alger	0.00	Chippewa	0.86	1.65E-08	4.5677
5/6	Alpena	5.00	Alcona	0.97	3.75E-05	8.69E+08
7	Antrim	2.41	Cheboygan	0.88	3.67E-02	2.55E+02
8	Benzie	-0.71	Alcona	-0.75	-2.72E-17	5.11E-02
9	Charlevoix	0.47	Alcona	0.85	3.06E-15	7.79E+00
10	Cheboygan	3.09	Alcona	0.96	3.59E-09	4.28E+05
11/12	Chippewa	3.44	Cheboygan	0.94	1.06E+01	-4.83E+06
13/14	Crawford	4.70	Alcona	0.94	8.81E-06	-5.17E+08
15/16	Delta	-0.37	Chippewa	-0.68	-9.96E-10	1.83E-01
17	Emmet	0.96	Alcona	0.83	7.97E-14	5.70E+01
18	Grand Traverse	0.68	Alcona	0.69	1.43E-14	1.72E+01
19	Iosco	2.00	Alcona	0.98	1.63E-11	4.59E+03
20/21	Kalkaska	5.00	Alcona	0.96	4.13E-05	-1.89E+09
22	Leelanau	assume the same end day as <b>Antrim County's</b>				
23/24	Luce	-2.02	Chippewa	-0.63	-2.79E-12	9.00E-05
25/26/27	Mackinac	1.11	Chippewa	0.70	1.10E-05	88.3276
28	Manistee	assume the same end day as <b>Wexford County's</b>				
29	Marquette	-0.38	Chippewa	-0.93	-1.12E-09	0.1738
30	Menominee	5.00	Chippewa	0.80	1176.4351	3.54E+09
31	Missaukee	-1.12	Alcona	-0.91	-6.81E-18	8.86E-03
32	Montmorency	-2.79	Alcona	-0.71	-7.29E-21	5.65E-06
33	Ogemaw	3.24	Alcona	0.90	7.64E-09	8.95E+05
34/35	Oscoda	3.23	Alcona	0.99	8.14E-09	9.06E+05
36/37	Otsego	0.23	Alcona	0.59	6.33E-16	2.65E+00
38/39	Presque Isle	5.00	Alcona	0.92	4.10E-05	-6.14E+08
40	Roscommon	4.03	Alcona	0.98	2.94E-07	7.14E+06
41/42	Schoolcraft	-5.00	Chippewa	-0.75	-2.83E-18	7.08E-11
43	Wexford	-3.08	Alcona	-0.96	-1.53E-21	1.36E-06
44	area 1>150 miles	assume the same end day as <b>Marquette County's</b>				
45	area 2>150 miles	assume the same end day as <b>Marquette County's</b>				
46	area 3>150 miles	assume the same end day as <b>Marquette County's</b>				