Assessing Risk in Fixing Wind Energy Rates

by

Molly Meyers

A thesis submitted in partial fulfillment of
the requirements for the Degree of

MASTER OF SCIENCE
In
ENVIRONMENTAL SCIENCE AND POLICY

University of Wisconsin-Green Bay

June 2014

Approved:

Dr. John Katers, Major Professor

Dr. Gregory J. Davis, Director of Graduate Studies

Thesis Committee Members:
Dr. Steven J. Meyer
Robert D. Benninghoff

DEDICATION

I dedicate my thesis work to my major advisor, Dave Dolan. Dave was my professor, mentor, friend, and supervisor throughout my undergraduate and graduate career. Dave was always present, patient, and knowledgeable whether I had questions with Intro Stats, SAS programming, or researching obscure methods for my thesis data analysis. He was my inspiration to continue on to the Environmental Science & Policy graduate program after completing my undergraduate in statistics. I will be forever grateful to Dave for his impact on my academic career.

ACKNOWLEDGEMENTS

I wish to thank my thesis committee - Dave Dolan, John Katers, Rob Benninghoff, and Steven Meyer - for the generous gift of their time, patience and expertise. I would like to thank Wisconsin Public Service for providing support and funding to acquire the data set for this study. I wish to extend my gratitude to John Katers for his additional time spent in seeing my thesis through to completion. Thank you to my family and friends, especially Brian and Reece, for their years of support and patience throughout my academic career.

ABSTRACT

ASSESSING RISK IN FIXING WIND ENERGY RATES

Molly Meyers

Wind turbines are an increasing source of renewable electricity generation. This increase can be attributed in part to growth in voluntary renewable energy markets. Those who contribute to the voluntary market before a potential crossover point (where the cost of renewable energy is less than that of after than non-renewable generation) are providing a benefit to all who use the energy market in the future, because contributions to the voluntary market allow for the purchase of renewable capacity that will be operational for several decades. In order for contributors to the voluntary energy market to financially benefit from their contributions, the cost of wind energy could be fixed by establishing a green rate. In order to accurately fix a green rate based on wind, the uncertainty regarding monthly wind energy generation and capacity availability must be analyzed. This study explored methods that estimate monthly wind energy generation and capacity values with associated levels of certainty on those values.

The results suggest that even with robust data sets (20+ years), the sample size of the monthly data is too low to accurately estimate the quantiles needed for the 95% confidence level. The bootstrap method of simulating additional months succeeded in more completely populating the extreme quantiles and succeeded when compared with the holdout year. In regards to capacity, this study suggests that no capacity can be counted on from a single wind facility at a 95% level of certainty. In summary, it is concluded that a single wind project's contribution to a fixed green rate product must exclude any capacity claims and be that of an energy-producing asset only and that the P99.6 quantile should be utilized for confidently estimating monthly energy generation.

TABLE OF CONTENTS

DEDICATION II

ACKNOWLEDGEMENTS III

ABSTRACT IV

TABLE OF CONTENTS V

LIST OF TABLES VII

LIST OF FIGURES IX

LIST OF EQUATIONS XIV

CHAPTER 1-INTRODUCTION 1

CHAPTER 2-LITERATURE REVIEW 7

2.1 Compliance and Voluntary Renewable Energy Markets 7

2.2 Cost of Energy Generation 17

2.3 Carbon Policies 22

2.4 Developing a Wind Facility 27

CHAPTER 3-METHODS 37

3.1 Data and Site Information 37

3.2 Transforming Wind Speeds to Energy and Capacity 41

3.3 Methods – Empirical Energy Analysis 45

3.4 Methods – Bootstrapping Energy Analysis 50

3.5 Methods – Capacity Analysis 57

CHAPTER 4-RESULTS 60

4.1 Results – Empirical Energy Analysis 60

4.2 Results – Bootstrapping Energy Analysis 68

4.3 Results – Capacity Analysis 75

CHAPTER 5-DISCUSSION 80

5.1 Discussion – Energy Analysis 80

5.2 Discussion – Capacity Analysis 82

CHAPTER 6-CONCLUSION 84

REFERENCES 85

APPENDIX I. TABLES AND FIGURES – OBSERVED DATA 88

APPENDIX II. TABLES AND FIGURES – BOOTSTRAPPED DATA 95

APPENDIX III. TABLES AND FIGURES - CAPACITY ANALYSIS 102

APPENDIX IV. OBSERVED MONTHLY GENERATION OVER TIME 109

APPENDIX V. SAS PROGRAMMING 116

Empirical Data Analysis 116

Bootstrap Data Analysis 118

Capacity Data Analysis 122

LIST OF TABLES

Table 1. Forecasts for Annual U.S. Wind Capacity Additions (MW) (Wiser & Bolinger, 2013) 2

Table 2. Utility voluntary renewable energy programs in WI as of May 2012 (DOE, 2012) 12

Table 3. Number of Customer Participants (as of December 2012) (Heeter & Nicholas, 2013) 13

Table 4. Customer Participation Rate (as of December 2012) (Heeter & Nicholas, 2013) 13

Table 5. GreenChoice Commercial Offerings January 1, 2014 (Austin Energy, 2014)16

Table 6. Example of Capacity Factors at an Operational Wind Facility 20

Table 7. Summary of Data Verification 40

Table 8. Data set summary report (1980-2008) generated by Windographer 1.4 beta. 40

Table 9. Sample of generation factors based on wind speed and direction (modified from Benninghoff (2010)) 43

Table 10. Number of Months Correctly Estimated at Various Confidence Levels 55

Table 11. Summary Statistics of Observed Data (in MWh) 61

Table 12. Observed Quantiles - Observed Data (in MWh) 61

Table 13. Monthly Capacity Factors Calculated from Observed Monthly Quantiles 62

Table 14. Comparing 2007 Monthly Generation to Observed P95 Estimate (MWh) 66

Table 15. Comparing Observed Monthly Means (SNK test in PROC ANOVA) 67

Table 16. Summary Statistics of Bootstrapped Data (in MWh) 68

Table 17. Lognormal Quantiles - Bootstrapped Data (in MWh) 69

Table 18. Percent Difference between Bootstrapped Summary Statistics and Observed Summary Statistics. 69

Table 19. Percent Difference between Bootstrapped Quantiles and Observed Quantiles 70

Table 20. Capacity Factors Calculated from Bootstrapped Monthly Quantiles 70

Table 21. Comparing 2007 Monthly Generation to Bootstrap Estimate (MWh)..... 73

Table 22. Comparing Bootstrapped Log Transformed Monthly Means (SNK test in PROC ANOVA) 75

Table 23. Capacity Quantiles by Month (MW) 78

Table 24. Comparing Monthly Mean Wind Speeds (SNK test in PROC ANOVA)..... 79

Table 25. Years and their assigned values used in Figure 79 through Figure 90..... 109

LIST OF FIGURES

Figure 1. Relative Contribution of Generation Types in Annual Capacity Additions (Wiser & Bolinger, 2013) 1

Figure 2. Annual Installed Wind Capacity (MW) by Region 2005-2012 (Global Wind Energy Council, 2013) 2

Figure 3. Theoretical cost of electric generation over time 4

Figure 4. State RPS Policies and Non-Binding Renewable Energy Goals (as of June 2013) (Wiser & Bolinger, 2013) 8

Figure 5. The Impact of Sporadic Expiration and Renewal of the PTC on the Wind Installations (Schneider et al, 2013) 9

Figure 6. Estimated annual voluntary sales by market sector (Heeter & Nicholas, 2013) 10

Figure 7. Estimated green power sales by renewable energy source, 2012 (Heeter & Nicholas, 2013) 10

Figure 8. Residential and nonresidential voluntary sales, 2006-2012 (Heeter & Nicholas, 2013) 11

Figure 9. GreenChoice Prices Versus Fuel Charge (Ankrum, 2010) 14

Figure 10. Installed Wind Power Project Costs over Time (Wiser & Bolinger, 2013) 17

Figure 11. Average Levelized Long-Term Wind PPA Prices and Yearly Wholesale Electricity Prices over Time (Wiser & Bolinger, 2013) 18

Figure 12. Cumulative Wind Power Capacity Categorized by Owner Type (Wiser & Bolinger, 2013) 19

Figure 13. 2010 Wind Power Price as a Function of 2010 Capacity Factor (Wiser & Bolinger, 2011) 21

Figure 14. Average Cumulative Sample-Wide Capacity Factor by Calendar Year (Wiser & Bolinger, 2013) 22

Figure 15. Carbon Trading Programs (Grubb, 2012) 23

Figure 16. Volume of Carbon Dioxide Allowance Trades (Newell et al, 2013) 24

Figure 17. Carbon Dioxide Allowance Prices (Newell et al, 2013) 24

Figure 18. Annual Average Wind Speed at 80 Meters (AWS Truewind, 2010) 28

Figure 19. Location of Wind Power Development in the United States (Wiser & Bolinger, 2013) 28

Figure 20. Average wind project size (modified from AWEA (2009)) 29

Figure 21. Distribution of wind turbines installed in 2008 by capacity (modified from AWEA (2009)) 30

Figure 22. Weibull curves with average wind speed of 7 m/s and varying shape factors (k) (Windographer 1.4 beta) 33

Figure 23. Wind Capacity Data (1/1/07) 35

Figure 24. The approximate location of meteorological tower at the Sewage Treatment Plant in Superior, Wisconsin (Google Maps) 38

Figure 25. Satellite view of meteorological tower at the Sewage Treatment Plant in Superior, Wisconsin (Google Maps). 39

Figure 26. Technical Data and Power Curve for the GE 1.5sle (modified from GE Energy (2009)) 42

Figure 27. Summary of Days Missing per Month 46

Figure 28. Level of Certainty Need to Correctly Estimate Monthly Values 55

Figure 29. April Energy Generation - Observed Data 63

Figure 30. October Energy Generation - Observed Data 63

Figure 31. December Energy Generation - Observed Data 64

Figure 32. Monthly Energy Generation - All Months - Observed Data 64

Figure 33. Monthly Energy Generation - 2007 (Holdout Year) 65

Figure 34. April Energy Generation - Bootstrapped Data 71

Figure 35. December Energy Generation - Bootstrapped Data 72

Figure 36. Daily Energy Generation - All Months 74

Figure 37. Wind Speed Analysis - All Months 76

Figure 38. April Wind Speed Analysis 76

Figure 39. December Wind Speed Analysis 77

Figure 40. Monthly Energy Generation - All Months - Observed Data 88

Figure 41. January Energy Generation - Observed Data 88

Figure 42. February Energy Generation - Observed Data 89

Figure 43. March Energy Generation - Observed Data 89

Figure 44. April Energy Generation - Observed Data 90

Figure 45. May Energy Generation - Observed Data 90

Figure 46. June Energy Generation - Observed Data 91

Figure 47. July Energy Generation - Observed Data 91

Figure 48. August Energy Generation - Observed Data 92

Figure 49. September Energy Generation - Observed Data 92

Figure 50. October Energy Generation - Observed Data 93

Figure 51. November Energy Generation - Observed Data 93

Figure 52. December Energy Generation - Observed Data 94

Figure 53. Monthly Energy Generation - All Months - Bootstrapped Data 95

Figure 54. January Energy Generation - Bootstrapped Data 95

Figure 55. February Energy Generation - Bootstrapped Data 96

Figure 56. March Energy Generation - Bootstrapped Data 96

Figure 57. April Energy Generation - Bootstrapped Data 97

Figure 58. May Energy Generation - Bootstrapped Data 97

Figure 59. June Energy Generation - Bootstrapped Data 98

Figure 60. July Energy Generation - Bootstrapped Data 98
Figure 61. August Energy Generation - Bootstrapped Data 99
Figure 62. September Energy Generation - Bootstrapped Data 99
Figure 63. October Energy Generation - Bootstrapped Data 100
Figure 64. November Energy Generation - Bootstrapped Data 100
Figure 65. December Energy Generation - Bootstrapped Data 101
Figure 66. Wind Speed Analysis - All Months – Weibull 102
Figure 67. January Wind Speed Analysis - Weibull Distribution 102
Figure 68. February Wind Speed Analysis - Weibull Distribution 103
Figure 69. March Wind Speed Analysis - Weibull Distribution 103
Figure 70. April Wind Speed Analysis - Weibull Distribution 104
Figure 71. May Wind Speed Analysis - Weibull Distribution 104
Figure 72. June Wind Speed Analysis - Weibull Distribution 105
Figure 73. July Wind Speed Analysis - Weibull Distribution 105
Figure 74. August Wind Speed Analysis - Weibull Distribution 106
Figure 75. September Wind Speed Analysis - Weibull Distribution 106
Figure 76. October Wind Speed Analysis - Weibull Distribution 107
Figure 77. November Wind Speed Analysis - Weibull Distribution 107
Figure 78. December Wind Speed Analysis - Weibull Distribution 108
Figure 79. January Energy Generation Over Time 110
Figure 80. February Energy Generation Over Time 110
Figure 81. March Energy Generation Over Time 111
Figure 82. April Energy Generation Over Time 111

Figure 83. May Energy Generation Over Time 112

Figure 84. June Energy Generation Over Time 112

Figure 85. July Energy Generation Over Time 113

Figure 86. August Energy Generation Over Time 113

Figure 87. September Energy Generation Over Time 114

Figure 88. October Energy Generation Over Time 114

Figure 89. November Energy Generation Over Time 115

Figure 90. December Energy Generation Over Time 115

LIST OF EQUATIONS

Equation 1. Capacity Factor Calculation..... 20

Equation 2. Model converting wind speed and direction to energy generation (modified from GE Energy (2009))..... 43

Equation 3. Prorating monthly energy generation..... 46

Equation 4. Lognormal density function (modified from SAS Institute Inc. (2008)) 49

Equation 5. Binomial Distribution (modified from SAS Institute Inc. (2008)..... 54

Equation 6. Log transform (Aitchison & Brown, 1963) 57

Equation 7. Weibull density function (SAS Institute Inc., 2008) 59

CHAPTER 1-INTRODUCTION

Wind turbines are an increasing source of renewable electricity generation. As shown in Figure 1, wind installations accounted for 43% of new capacity additions in the U.S. in 2012 (Wiser & Bolinger, 2013). Wind energy generation is also on the rise throughout the rest of the world (Figure 2). Increases in wind energy generation are primarily dependent on state and federal policies, improved economic feasibility and growth in the voluntary renewable energy market (Swezey et al, 2007). Table 1 lists forecasts from fives sources for wind capacity installations in 2013, 2014 and 2015. Estimates for 2015 and beyond are uncertain and dependent in part on natural gas prices, fossil fuel plant retirements, federal policy decisions and future prospects for growth based on recent declines in the price of wind energy (Wiser & Bolinger, 2013).

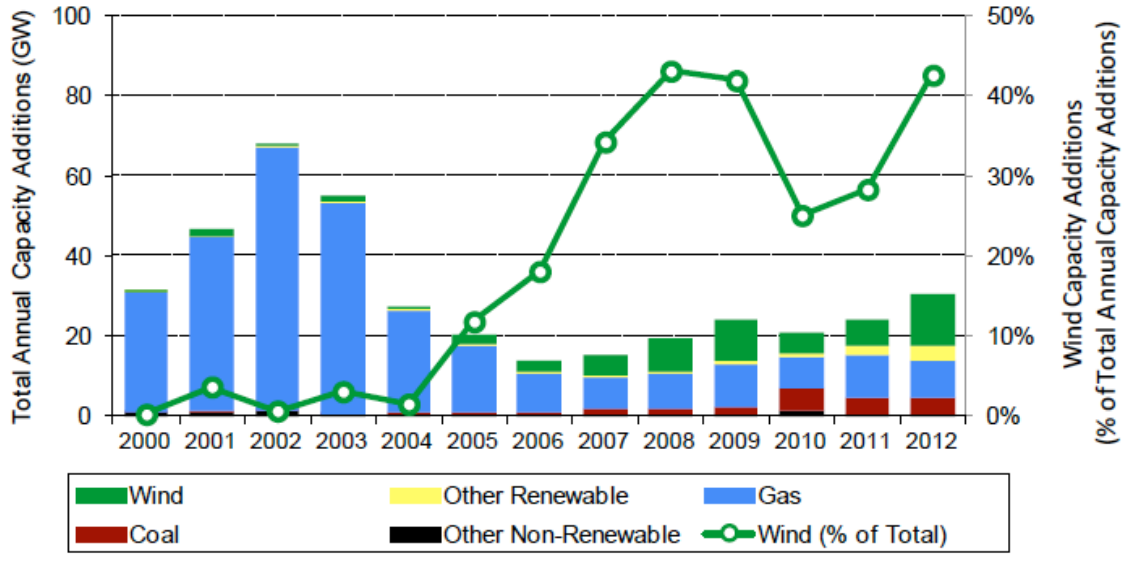


Figure 1. Relative Contribution of Generation Types in Annual Capacity Additions (Wiser & Bolinger, 2013)

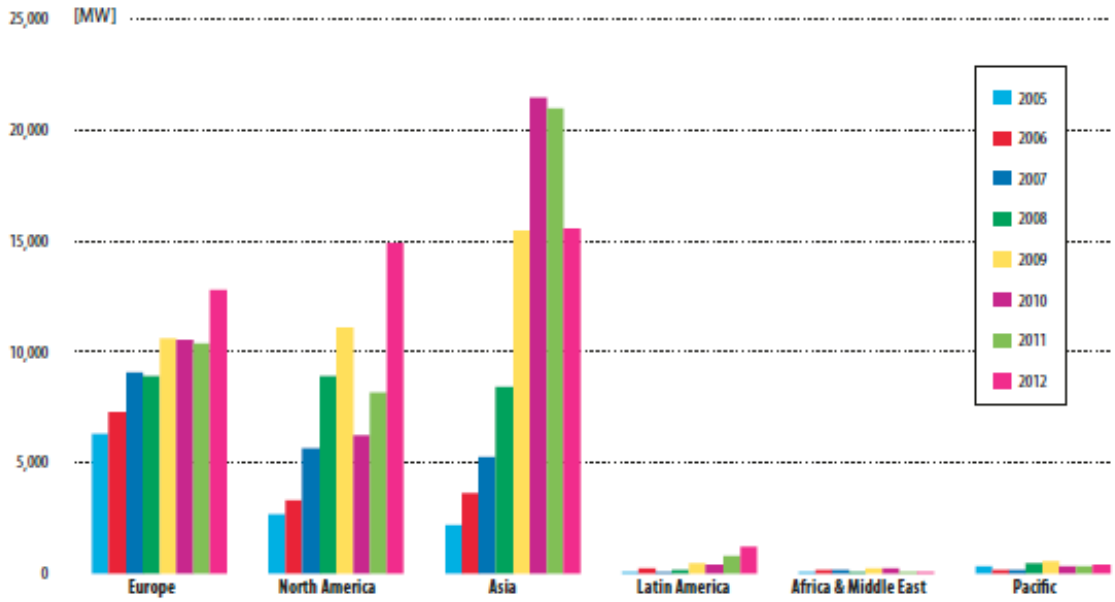


Figure 2. Annual Installed Wind Capacity (MW) by Region 2005-2012 (Global Wind Energy Council, 2013)

Table 1. Forecasts for Annual U.S. Wind Capacity Additions (MW) (Wiser & Bolinger, 2013)

Source	2013	2014	2015
Bloomberg NEF (2013a, 2013c)	2,800	8,000	3,200
IHS EER (2013)	2,000	6,000	7,300
Navigant (2013)	5,000	9,000	3,500
MAKE Consulting (2013)	3,500	7,700	4,500
EIA (2013b)	3,600	10,100	N/A

The voluntary renewable energy market, although not a significant driver in the current wind energy market, is an important component for utility companies to consider when growing their renewable generation or energy purchases for compliance. The voluntary renewable energy market allows contributors to purchase kilowatt or megawatt hours of renewable energy, also known as renewable energy credits (RECs) and their

purchase ensures renewable electricity is being generating on their behalf. However, the renewable energy is not directly serving their load, meaning no new "green" power lines are being run to their home or business. Instead, the contributors to the voluntary market are paying for renewable energy generation to serve the load for all that use the energy market. The price of a REC represents the incremental cost of renewable energy over conventional generation. In the voluntary market, contributors elect to pay this premium, while in the compliance market this incremental cost is recovered by utilities through their ratepayers. Currently, wind energy is the most cost-effective way to generate renewable energy for RPS compliance and to generate credits for the voluntary renewable energy market (Wiser & Bolinger, 2008).

The cost of non-renewable generation is expected to continue to increase due to the rising cost of fuel sources and expansions in carbon regulation (Bird et al, 2007). Therefore, renewable generation may eventually cost less than non-renewable generation. Past the crossover point (see Figure 3), renewable generation will be cheaper than non-renewable generation and all those who use the energy market will benefit from any renewable energy already in the energy mix. Therefore, those contributing to the voluntary market before the crossover point are providing a benefit to all who use the energy market in the future, because contributions to the voluntary market allow for the purchase of renewable capacity that will be operational for several decades.

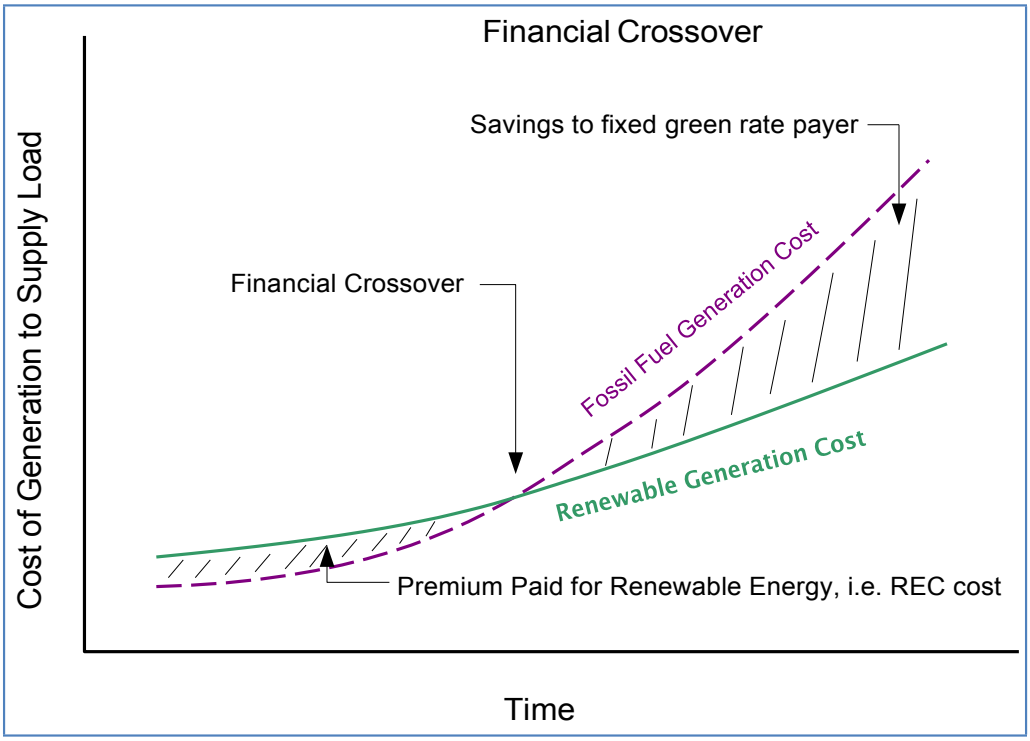


Figure 3. Theoretical cost of electric generation over time

If the price of a REC were fixed by establishing a "green rate," subscribers to the rate would pay a premium for their electricity now, but since their rate would be fixed over a period of time, they would pay a lower rate than non-subscribers after the crossover point. Since utility scale wind facilities are the most cost effective way to generate renewable energy, the "green rate" would essentially be a wind-based rate. Explaining how a wind rate could be quantified requires an understanding of the nature of wind energy and how it integrates with other electric generation.

The quantity of electricity generated by a wind facility, at any given moment, is primarily dependent on the magnitude of the wind speed. However, the direction of the wind, turbulence and air density also has an effect on electricity generation. Therefore, wind is classified as a "non-dispatchable" energy source because the magnitude of wind

speeds cannot be known far in advance. Dispatchability refers to the ability of an electric generator to be turned on or off, as well as increase or decrease generation at the request of an energy system operator. Given that wind is a non-dispatchable energy source, it has to be supplemented with dispatchable energy to ensure that enough electricity is generated at all times to reliably serve the load. Dispatchable energy can come from non-renewable (i.e. coal, natural gas, or nuclear) or renewable (i.e. landfill gas or biomass) sources. A green rate based on wind energy must include the cost of wind facility plus the supplemental dispatchable energy. The cost of the dispatchable energy is based on the capital investment, estimated life of the project, annual operation and maintenance fees and in most cases, fuel costs. The cost of wind energy is based on the capital investment, estimated life of the project, expected energy generation and annual operation and maintenance fees. The fuel, wind energy, is free, but there is uncertainty in its availability to generate energy.

Therefore, in order to accurately fix wind energy rates, this uncertainty must be analyzed statistically. To assess the risk involved in fixing wind energy rates, this study will explore methods that estimate monthly wind energy and capacity values with associated levels of certainty. More specifically, two main objectives are addressed:

1. The objective of the energy analysis is to determine a method for assessing statistical certainty around theoretical monthly energy generation at the 95% confidence level.
2. The objective of the capacity section is to determine a method for assessing statistical certainty around the amount of capacity available at the hourly level at the 95% confidence level.

When monthly energy and capacity and their corresponding probability are quantified, the amount of risk involved in offering a fixed green rate can be determined.

CHAPTER 2-LITERATURE REVIEW

This section will add more detail to topics discussed in the previous section and also provides background information regarding the methods used in the data analysis portion of this study (Sections 3-3.5).

2.1 Compliance and Voluntary Renewable Energy Markets

As previously mentioned, the growth in wind energy generation in the U.S. can be largely attributed to state and federal policies. Renewable portfolio standards (RPS) are the driving state policies. Although the individual RPS policies tend to differ in design and implementation, in general, an RPS is an energy policy that mandates or encourages electric utilities to increase the amount of renewable electricity they sell over a period of time (Wiser & Barbose, 2008). Figure 4 shows the states that have binding (29) and non-binding (7) RPS policies or goals as of June 2013.

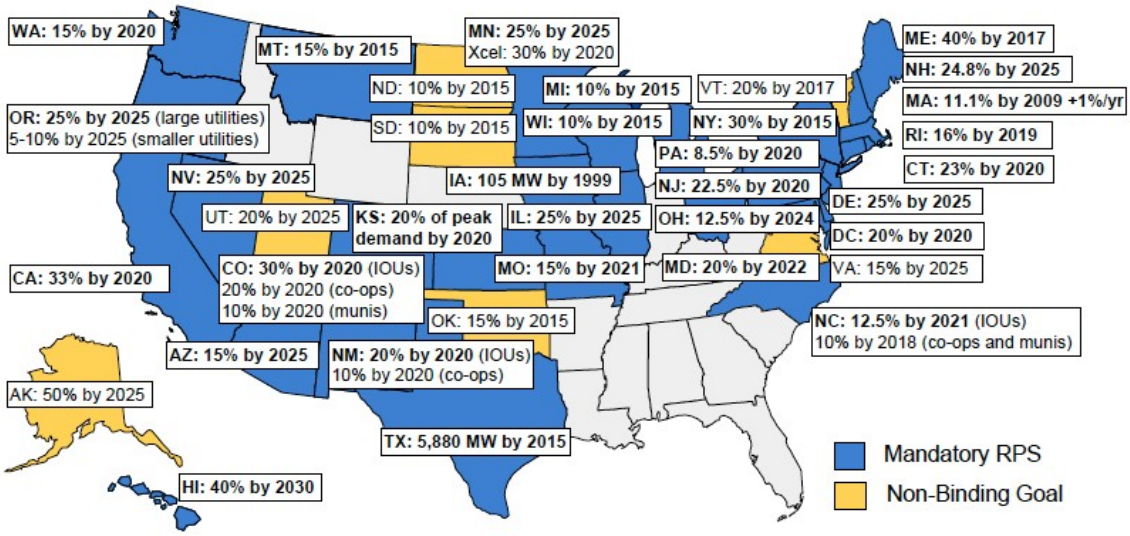


Figure 4. State RPS Policies and Non-Binding Renewable Energy Goals (as of June 2013) (Wiser & Bolinger, 2013)

Since renewable generation currently and historically costs more than non-renewable generation, RPS policies are in place to help stimulate the market for renewable energy. There is evidence that RPS policies are succeeding in stimulating the market for wind. In 2012, 83% of the wind power capacity installed in the U.S. was located in states with RPS policies (Wiser & Bolinger, 2013). However, the compliance market created by the RPS policies is not the only explanation for the increase in wind energy installations. Federal tax incentives for wind energy were set to expire at the end of 2012, driving a record setting year for annual wind power capacity installations in 2012 (Figure 5).

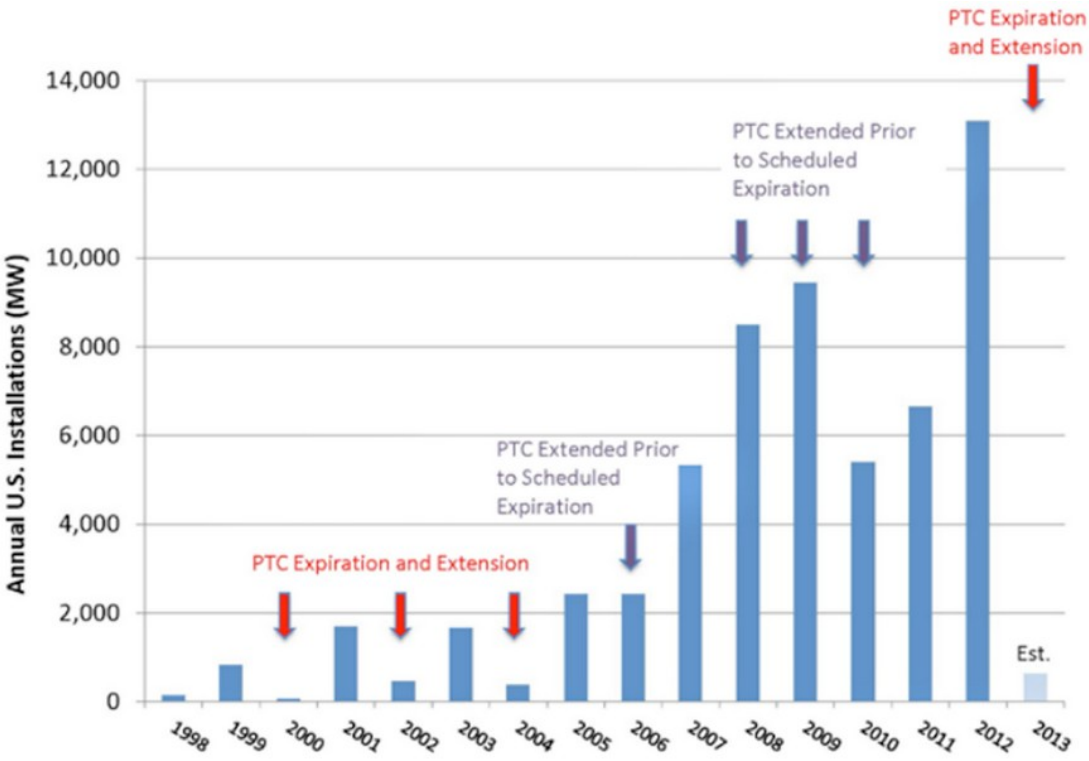
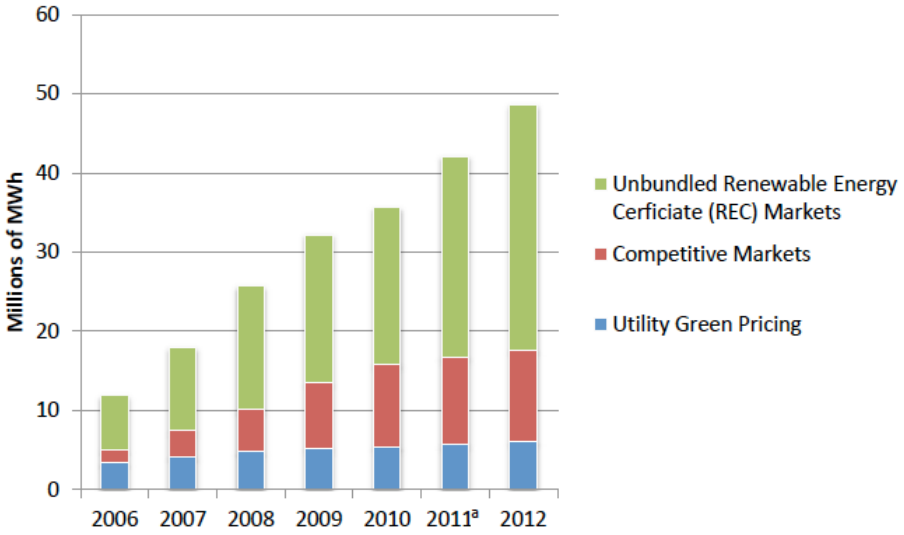


Figure 5. The Impact of Sporadic Expiration and Renewal of the PTC on the Wind Installations (Schneider et al, 2013)

In January 2013, congress passed a short-term extension of federal tax incentives; however, it will take time for the project pipeline to respond to the extension (Wiser & Bolinger, 2013). The voluntary market and the prospect of future carbon regulations also influence decisions regarding the type of new capacity installations.

The voluntary market, created by green pricing programs, allows customers to contribute to renewable energy installations beyond what is required in the compliance market. Figure 6 illustrates the increase in demand for renewable energy in the voluntary market over the past several years with wind being the primary renewable energy source

for meeting the demand (Figure 7). Sales to nonresidential customers continued to increase faster than sales to residential customers (Figure 8).



a Voluntary sales for 2011 are estimated as the mid-point of 2010 and 2012 sales.

Figure 6. Estimated annual voluntary sales by market sector (Heeter & Nicholas, 2013)

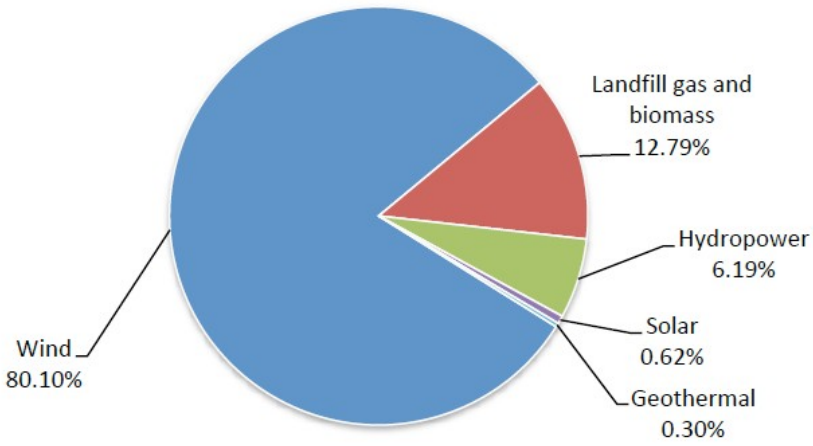


Figure 7. Estimated green power sales by renewable energy source, 2012 (Heeter & Nicholas, 2013)

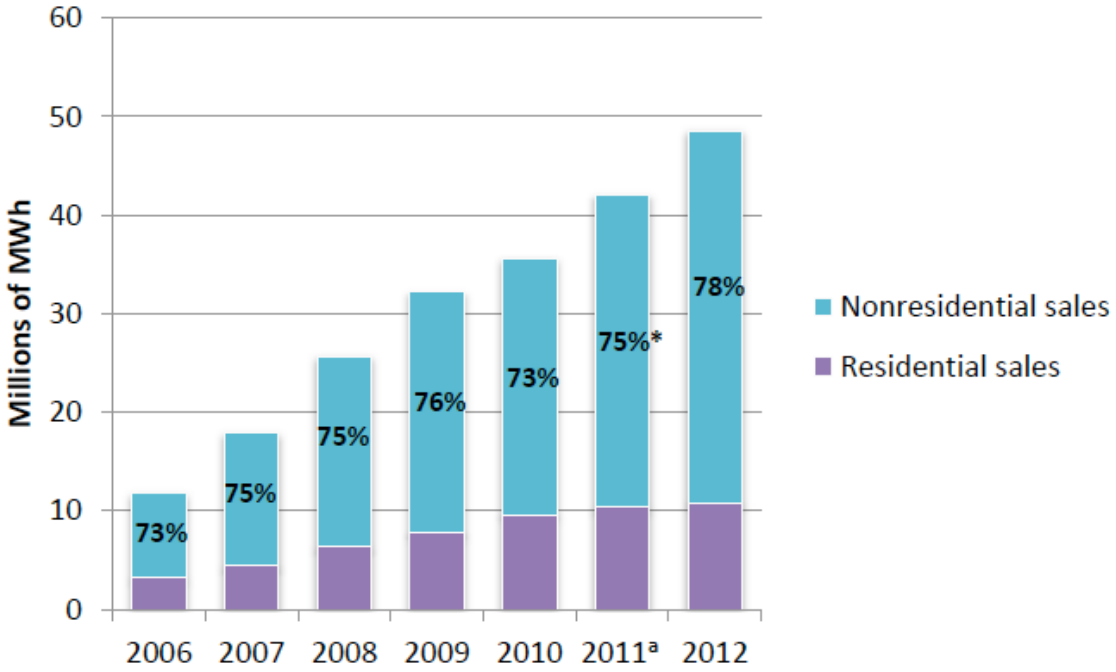


Figure 8. Residential and nonresidential voluntary sales, 2006-2012 (Heeter & Nicholas, 2013)

Table 2 lists information about the voluntary green pricing programs in Wisconsin. Looking at the "Resource Mix" column, it is apparent that wind energy is primarly used to meet the needs of the utilities' individual green pricing programs. There are approximately 50,000 customers contributing to voluntary renewable energy programs in Wisconsin.

Table 2. Utility voluntary renewable energy programs in WI as of May 2012 (DOE, 2012)

Utility	Program Name	Resource Mix	Price	Started
Alliant Energy	Second Nature	wind, landfill gas	2.0¢/kWh	2000
Dairyland Power Cooperative	Evergreen Renewable Energy Program	hydro, wind, landfill gas, biogas	1.5¢/kWh	1998
Great River Energy: Head of the Lakes	Wellspring Renewable Wind Energy Program	wind	0.4¢/kWh	1997
Madison Gas and Electric	Green Power Tomorrow	wind solar	4.0¢/kWh	1999
WPPI Energy	Renewable Energy Program	small hydro, wind, biogas	1.0¢/kWh	2001
We Energies	Energy for Tomorrow	landfill gas, PV, hydro, wind	1.39¢/kWh	1996
Wisconsin Public Service	Solar Wise for Schools	PV in schools	Contribution	1996

Table 3 and Table 4 highlight leading green pricing programs in the U.S. as of December 2012 (data more current than in Table 2). The National Renewable Energy Laboratory (NREL) has been analyzing data on green pricing programs since 2000. Table 3 lists the 10 green pricing programs with the highest number of participants. Note that We Energies (WI) is on the list with 20,066 participants. When NREL assessed the customer participation rate for green pricing programs, another Wisconsin utility made the list with 9.6% of Madison Gas and Electric's (MGE) customers participating in their green pricing program.

Table 3. Number of Customer Participants (as of December 2012) (Heeter & Nicholas, 2013)

Rank	Utility	Participants
1	Portland General Electric	87,987
2	PacifiCorp - Blue Sky Usage, Block and Habitat	87,919
3	Xcel Energy	61,315
4	Sacramento Municipal Utility District	55,207
5	Puget Sound Energy	34,962
6	Connecticut Light and Power Co. / United Illuminating	27,664
7	Iberdrola: NYSEG and RG&E	21,201
8	We Energies	20,066
9	National Grid	18,302
10	Dominion Virginia Power	15,179

Table 4. Customer Participation Rate (as of December 2012) (Heeter & Nicholas, 2013)

Rank	Utility	Program(s)
1	City of Palo Alto (California)	18.2%
2	Portland General Electric	12.4%
3	Madison Gas & Electric Co.	9.4%
4	Sacramento Municipal Utility District	9.2%
5	City of Naperville (Illinois)	7.6%
6	Pacific Power (Oregon)	7.6%
7	Silicon Valley Power	6.9%
8	River Falls Municipal Utilities	6.1%
9	Stoughton Utilities	5.1%
10	Cuba City Light & Water	5.0%

Among these various green pricing programs, there is one example of a program that has offered a fixed rate, Austin Energy's GreenChoice program. GreenChoice was established in 2000 and offers a 10-year fixed Green Power Charge to subscribers, replacing the variable Fuel Adjustment Charge paid by nonsubscribers. This rate has been offered in batches and, in early years, it took about 6 months for a rate to sell out. Those who signed on for the first two batches have saved a lot of money over the 10-year

program – "up to $1,485 for Batch 1 participants who signed on by September 30, 2000, according to Austin Energy spokesman Carios Cordova" (Ankrum, 2010). The rate has become increasingly expensive, with the 2009 offering priced at more than 3 times the first batch. Not surprisingly, this batch of the GreenChoice rate has attracted minimal subscribers and 95% of the rate was still available after more than a year on the market while all previous batches have sold out (Ankrum, 2010). Figure 9 shows the increases in batch rates along with the fluctuations in fuel charges from 2000 through 2010.

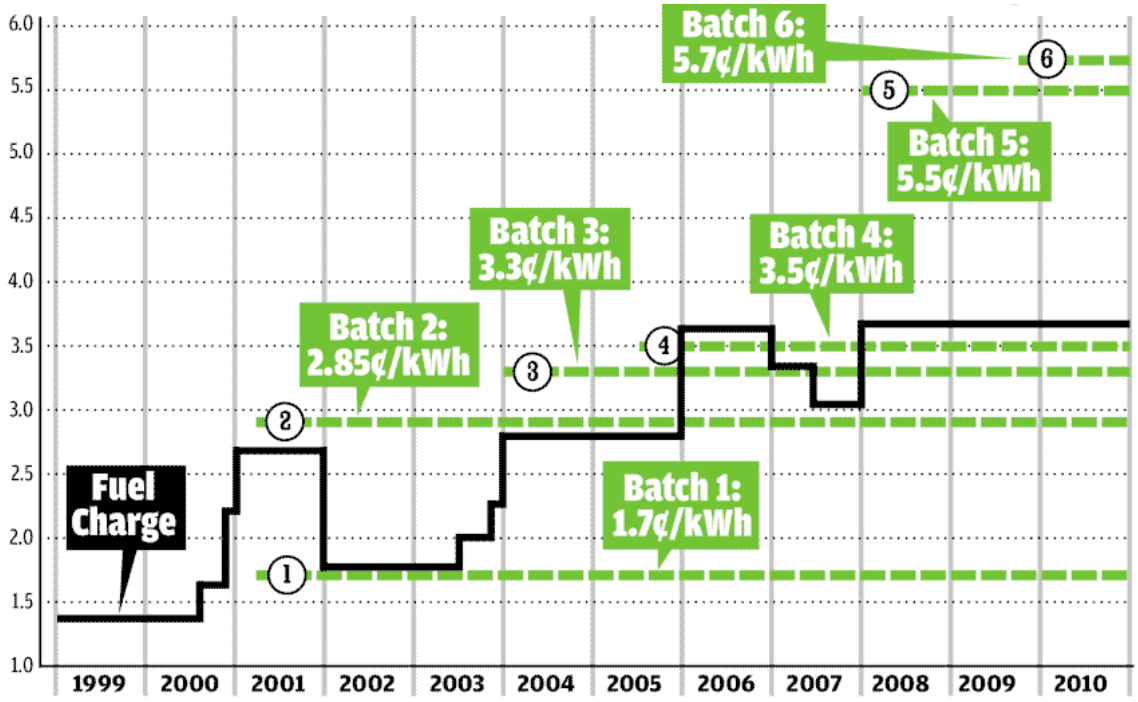


Figure 9. GreenChoice Prices Versus Fuel Charge (Ankrum, 2010)

Since 2010, the GreenChoice program has been restructured. For residential customers the cost is 1 cent more per kWh with no contract. The extra fee assures participates that renewable energy is being generated on their behalf from Texas wind

power. There are also three options for commercial customers (Table 5). One option includes a fixed rate; however, the contract length is only three years versus the previous programs' 10-year contracts.

Table 5. GreenChoice Commercial Offerings January 1, 2014 (Austin Energy, 2014)

GreenChoice Products	BusinessCents	Energizer	Patron
GreenChoice charge Replacing the Power Supply Adjustment (PSA) charge	PSA plus 1¢ 4.709¢ / kWh adjusted upon change in PSA	PSA plus 1¢ 4.709¢ / kWh adjusted upon change in PSA	4.9¢ / kWh fixed rate until December 2016
Eligibility	Commercial account in good standing not subscribed to GC B4, B5, B6.14 or B6.21	Commercial account in good standing not subscribed to GC B4, B5, B6.14 or B6.21	Commercial 1.2 M kWh annually or greater on one or a combination of accounts not subscribed to GC B4, B5, B6.14 or B6.21
Minimum Subscription	100% of account kWh	100% of account kWh	Minimum of 100,000 kWh blocks monthly on each account
Contracts	No contract required (Letter of Acceptance to confirm billing in January 2014)	12-month contract Allows qualification for • EPA Green Power Partner Program • City of Austin's Green Business Leaders Program	3-year contract Allows qualification for • EPA Green Power Partner Program • City of Austin's Green Business Leaders • LEED credits • Austin Energy Green Building ratings
Commitment	No penalty for unsubscribing and may resubscribe January 1 of following year	12-month commitment May continue on a month-to-month basis	Committed until December 2016 May resubscribe to new product after term expiration

2.2 Cost of Energy Generation

This section will look at the cost of wind over time and offer some price comparisons with conventional generation costs. The cost of wind turbines have decreased 20-30% since late 2008, in part because of continued engineering advances yielding increases in hub heights and rotor diameters. This downward trend can also be attributed to increased competition among manufacturers and a move to a buyer's market. These price reductions influence total project costs and wind power prices (Wiser & Bolinger, 2013). Figure 10 shows the installed wind power prices over time in 2012 dollars per kilowatt-hour (kWh).

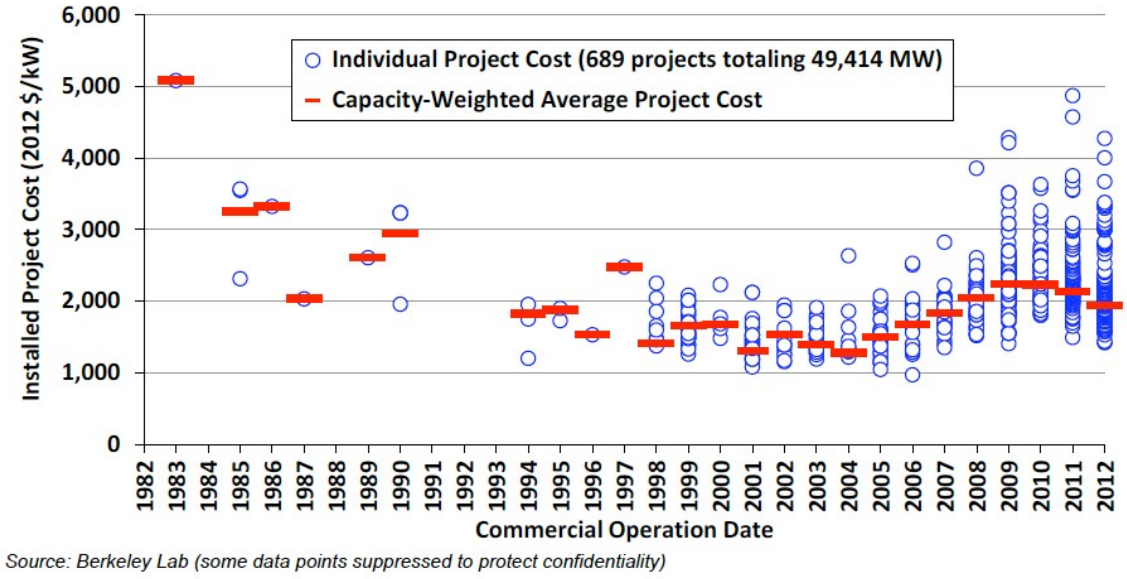


Figure 10. Installed Wind Power Project Costs over Time (Wiser & Bolinger, 2013)

From 2003-2008 wind power prices were very competitive with wholesale power prices. The sharp decrease in 2009 (driven by lower natural gas prices) moved the

average wind power purchase agreement (PPA) prices out of the range of wholesale power prices across the entire U.S (Figure 11). Utilities often enter into PPA's with independent power producers (IPP's) to benefit from the purchase of renewable energy without having to own the project. As of 2012, IPP's own more than 80% of the cumulative installed wind power capacity in the U.S. (Figure 12).

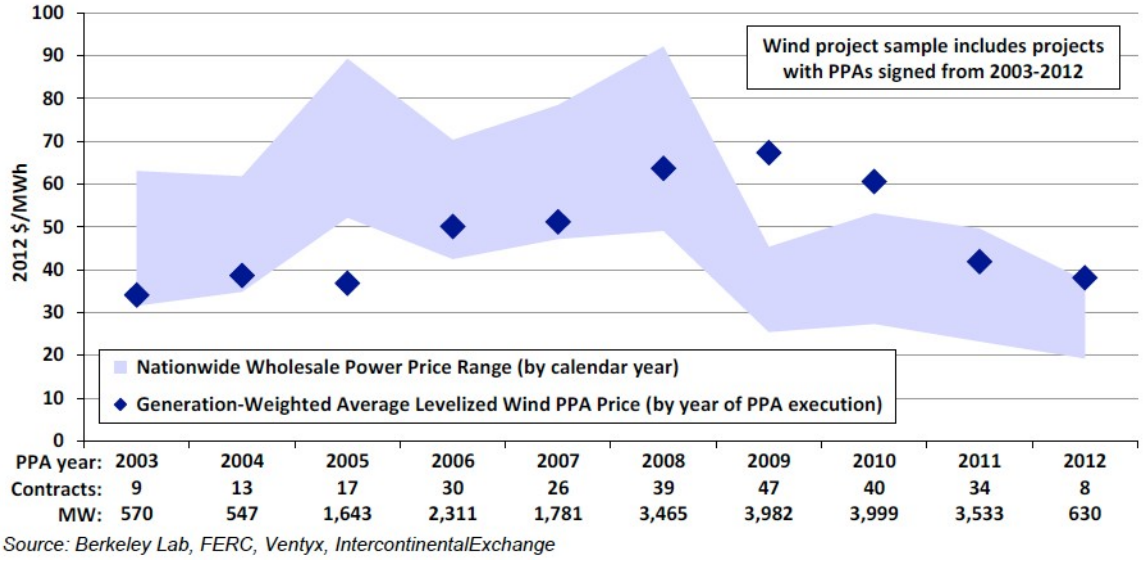


Figure 11. Average Levelized Long-Term Wind PPA Prices and Yearly Wholesale Electricity Prices over Time (Wiser & Bolinger, 2013)

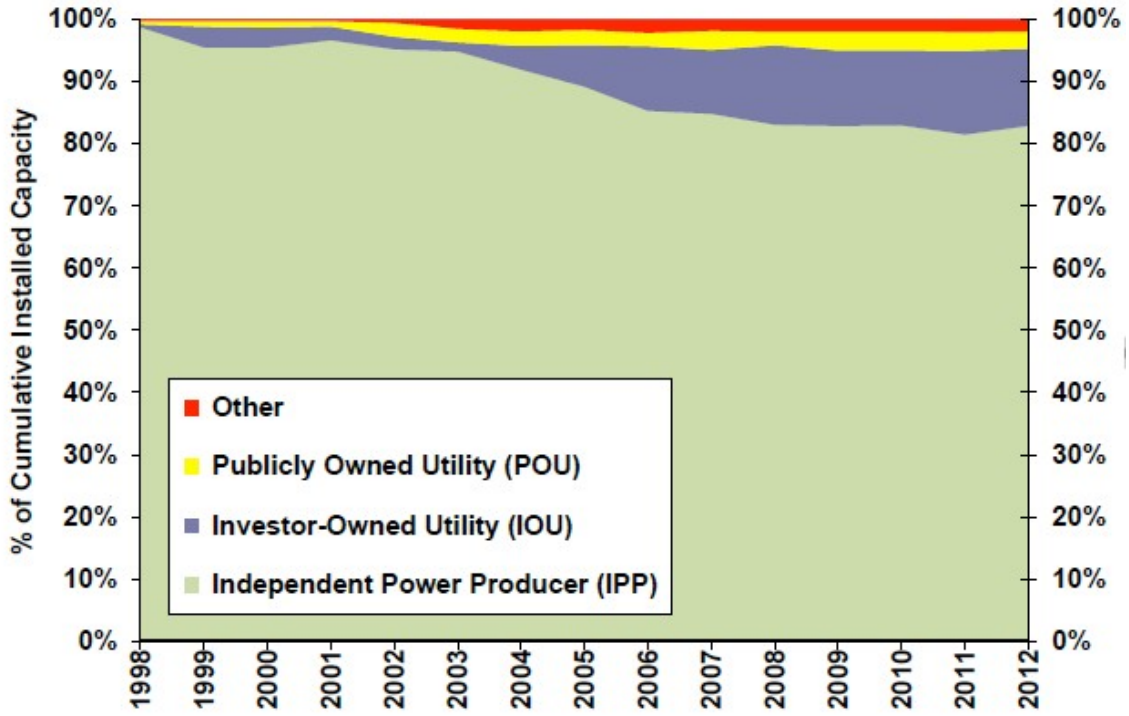


Figure 12. Cumulative Wind Power Capacity Categorized by Owner Type (Wiser & Bolinger, 2013)

Price is not the only factor controlling wind power installations. Utilities need a reliable source of dispatchable power to serve the load and there is the need for new or upgraded transmission to bring the wind power from where it is generated to the load (Wiser & Bolinger, 2009). It's also important to note that neither the wind nor wholesale electricity prices in Figure 11 capture the entire social costs of power generation. For example, sources receive various types of financial incentives, transmission costs are not fully reflected and fossil fuel generation does not fully account for the associated environmental and social costs.

The capacity factor of a wind facility also influences the cost of generation. A capacity factor is measure of how much electricity is generated compared to its potential

maximum over a given period of time. Equation 1 shows the formula for calculating an annual capacity factor.

$$cf = x / \{(8760 \text{ hr/year}) * (y)\}$$

Where:

cf = capacity factor (%)

x = the amount of megawatt hours (MWh) generated over a year

y = the rated capacity in megawatts (MW) of the electricity generating facility

Equation 1. Capacity Factor Calculation

Capacity factors are a useful tool to summarize average annual wind resource and have a strong influence on the cost of electricity produced and the profitability of the wind facility (Logan & Kaplan, 2008). In Table 6, the P50 value of 38.0% can be interpreted as follows; there is a 50% probability that the annual capacity factor will exceed 38.0% percent.

Table 6. Example of Capacity Factors at an Operational Wind Facility

Probability of Exceedance	Annual Capacity Factor
P50	38.0%
P75	34.6%
P90	31.5%
P95	29.7%
P99	26.2%

Capacity factors at operational wind facilities tend to range from 15% to 50%, with very few facilities with capacity factors below 25% (Wiser & Bolinger, 2009). Wind facilities with higher capacity factors are favored because they offer more energy output per MW installed. However, the capacity factor is a product of the wind resource at the site and the sites with the more favorable wind resource are not always available for development due to siting or transmission restrictions (Wiser & Bolinger, 2009). Figure 13 illustrates wind power prices as a function of the capacity factor of the project.

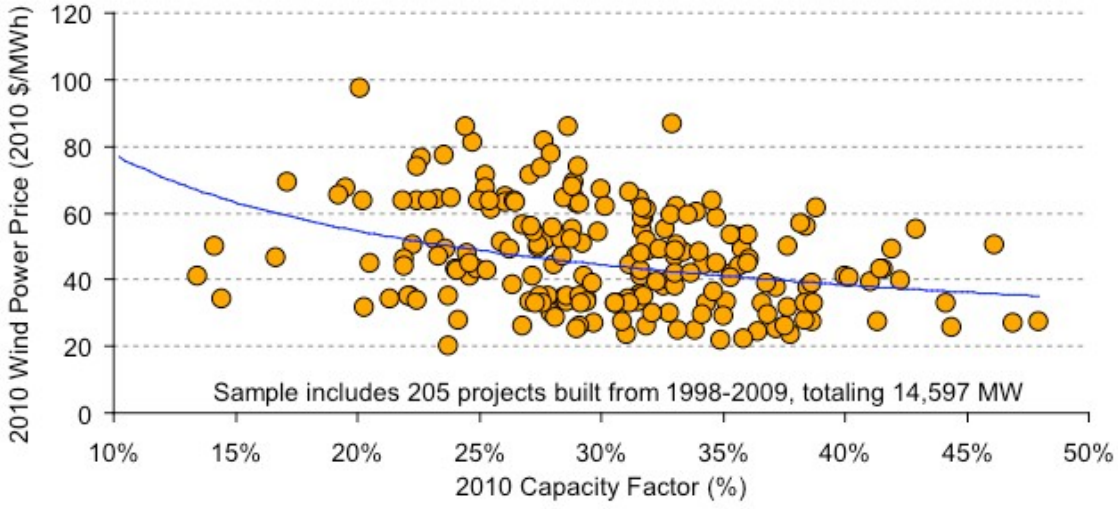


Figure 13. 2010 Wind Power Price as a Function of 2010 Capacity Factor (Wiser & Bolinger, 2011)

It may be expected that capacity factors would gradually increase as larger turbines are added to the system. However, the trends in Figure 14 are influenced by wind power curtailment, inter-year variability of the wind resource and the average quality of the wind resource at project locations. The strength of the wind resource varies as a result to

weather patterns, such as El Nino/La Nina and its influence on annual average capacity factors is evident by the green line in Figure 14.

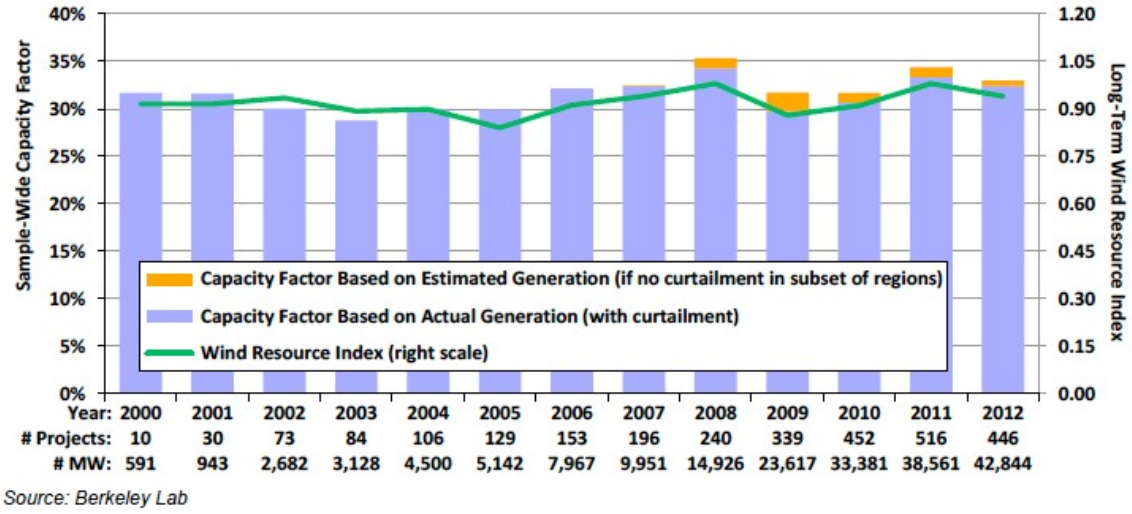


Figure 14. Average Cumulative Sample-Wide Capacity Factor by Calendar Year (Wiser & Bolinger, 2013)

2.3 Carbon Policies

In addition to the effects of the compliance and voluntary markets, some utilities and energy markets are planning electric capacity installations based on the future prospects of carbon legislation. There are various types of carbon policies, but in general, they aim to reduce the amount of carbon dioxide emitted into the atmosphere in attempts to mitigate global climate change. Two general types of policies are carbon taxes and carbon cap and trade programs. Carbon taxes have existed internationally for nearly 20 years and are most commonly placed on gasoline, coal and natural gas (Sumner et al, 2009). There are also carbon cap and trade programs that establish emission limits. Figure 15 shows cities, regions and developing economies that are modeled after Europe

in creating carbon emission trading programs. The filled circles represent the amount of carbon capped and the open blue circles represent the national emission volume – both circle types are shown on the same scale.

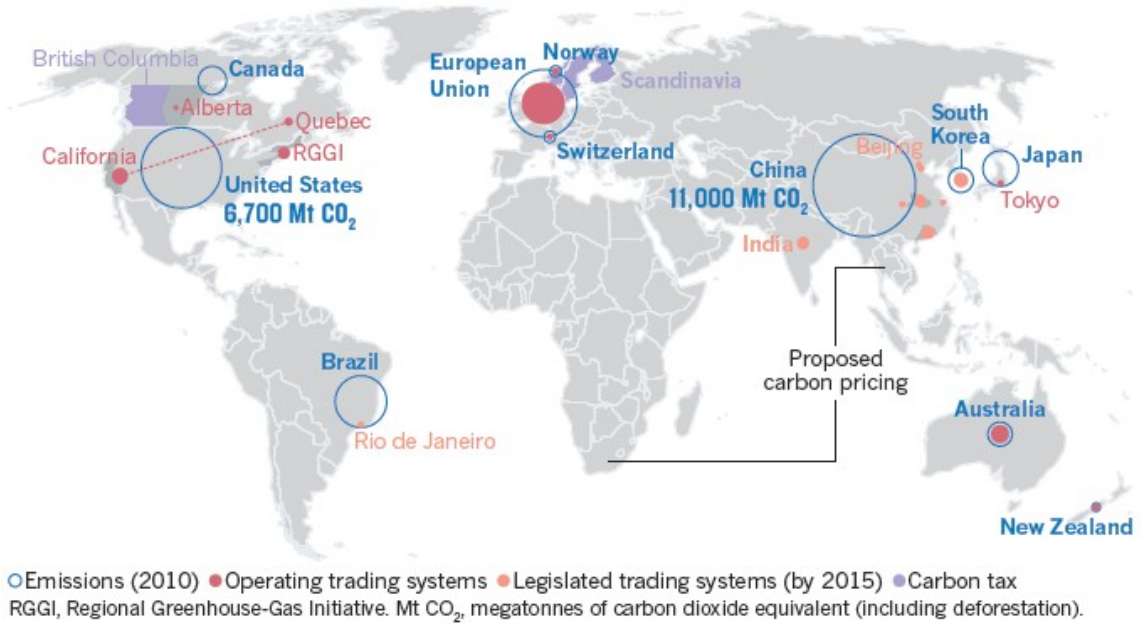


Figure 15. Carbon Trading Programs (Grubb, 2012)

At the end of 2012, most of the carbon trading around the world took place in five markets: the European Union's Emission Trading System (ETS), the Clean Development Mechanism (developed under the Kyoto Protocol), RGGI (northeastern U.S.), New Zealand's Emissions Trading Scheme and voluntary markets (Newell et al, 2013). Figure 16 and Figure 17 show how these five markets compare over time in volume of trades and the cost of allowances.

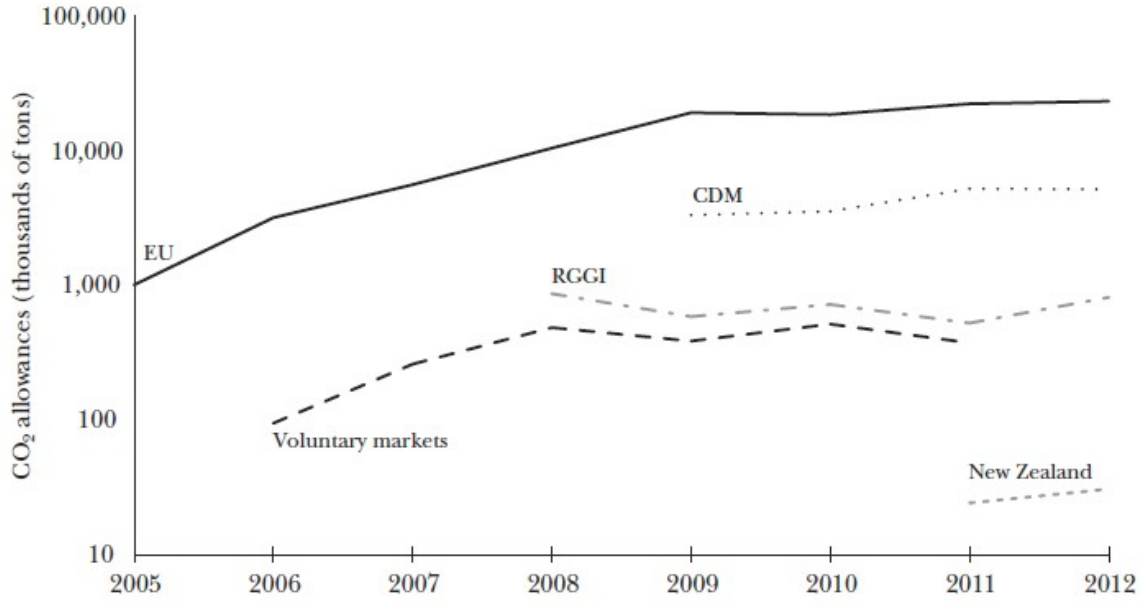


Figure 16. Volume of Carbon Dioxide Allowance Trades (Newell et al, 2013)

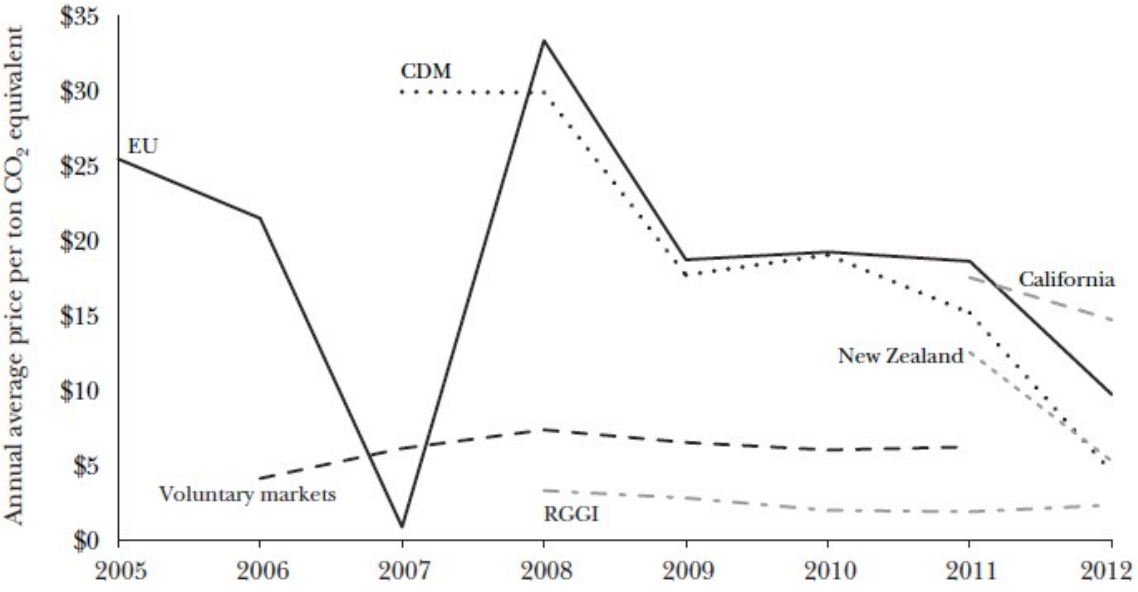


Figure 17. Carbon Dioxide Allowance Prices (Newell et al, 2013)

Nine U.S. states are currently participating in market-based carbon emissions reduction program called the Regional Greenhouse Gas Initiative (RGGI). The program sought to reduce emissions from large electricity generators by 10 percent below 2009 levels by 2018. State governments receive the revenue from auctioned allowances and used the majority of the funds in local renewable energy and energy efficiency projects. After several years RGGI exceeded its initial reduction targets, mostly due to fuel switching from coal-fired power to lower priced natural gas. Total emissions for the then ten-state region were 33 percent below the program cap of 188 million short tons in 2011. New Jersey left RGGI in 2011 and from 2012-2013 the cap was 165 million short tons for the nine-state region. Even though RGGI exceeded reduction targets, allowance prices did not collapse due to the established floor price. However, trading had generally occurred near the floor price and allowances have gone unsold at auctions (Newell et al, 2013). RGGI completed a comprehensive program review in 2012 and implemented a new 2014 cap of 91 million short tons. The cap then declines 2.5 percent every year from 2015 to 2020. The idea is that as the amount of allowances decrease, the price of the pollution permits will increase and power generators will choose to invest in renewable energy over purchasing permits or facing penalties (Pool, 2010). Currently, policy only affects nine states; however, if pending federal legislation were adopted, a similar model could be in place for all states.

The American Power Act of 2010 (sponsored by John Kerry and Joseph Lieberman) and the American Clean Energy and Security Act of 2009 (sponsored by Henry Waxman and Edward Markey) both aimed to create a carbon market along with other energy

initiatives. This is the third time such a policy has been proposed in federal legislation. Both bills passed in the House, but were defeated in 2010 by the Senate.

Due to these many defeats by Congress, President Obama stated his intent to use his authority under the Clean Air Act and a 2007 Supreme Court decision to issue new regulations on carbon emissions during his 2014 State of the Union Address. President Obama ordered the Environmental Protection Agency (EPA) to submit a draft of a regulation setting national standards for carbon emissions by June 1, 2014. The regulation would primarily affect the 600 coal-fired power plants in the U.S., which will spur strong political debate in the 2014-midterm elections. Initial suggestions state that the regulation will order states to create and implement their own policies for meeting the standard. Contingent on how regulations are written, states could comply by establishing cap and trade programs. The EPA's task has an ambitious timeline because President Obama intends for the majority of the regulatory process to be completed prior to the end of his term in January of 2017. After the draft is submitted in June 2014, the President wants the final version of the regulation by June 2015. Next, the states must submit plans for implementation by June 2016 (Davenport, 2014).

These existing and pending carbon policies will influence investments in renewable energy. If the pending legislation were implemented, the cost of conventional generation will likely increase, narrowing the gap between the cost of wind power and conventional generation.

2.4 Developing a Wind Facility

The basics of the current wind market have been discussed. The development of a wind facility will now be discussed. The wind resource, a description of the wind magnitude and frequency, is a determining factor regarding where a wind facility will be located. Wind resources in the U.S. and around the globe are well known and documented. Figure 18 shows the average annual wind speed at a height of 80 meters. It should be noted that wind facilities are not necessarily concentrated where the best wind resources exist. In some cases, there is limited transmission and load near these high wind areas, limiting wind development in the area. Figure 19 shows the location of wind power development as of 2008. Comparing Figure 19 and Figure 20 it is clear that there are underutilized wind resources throughout the Dakotas, Nebraska and Kansas.

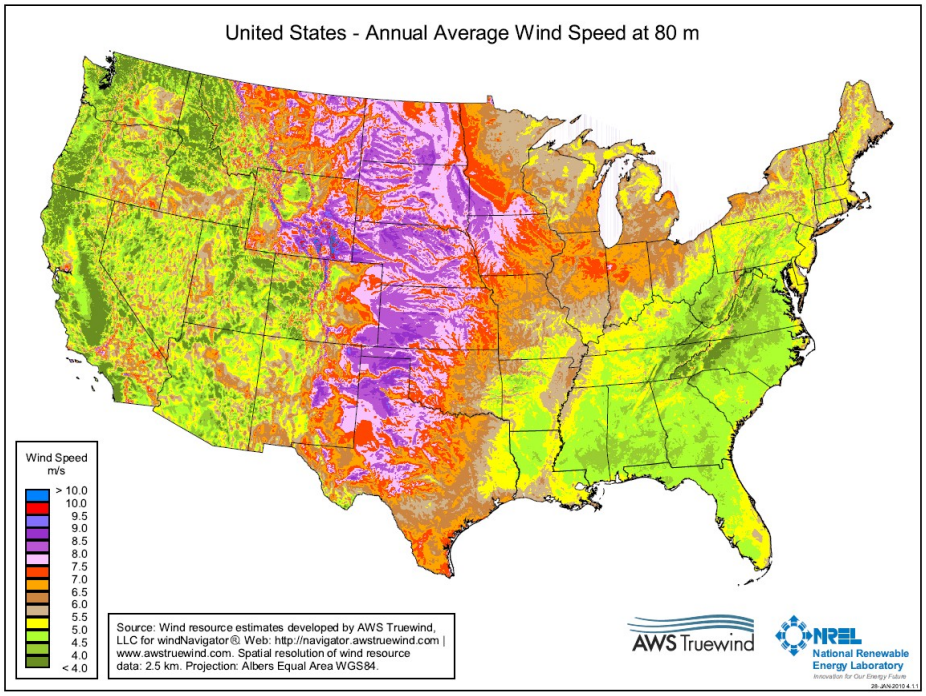
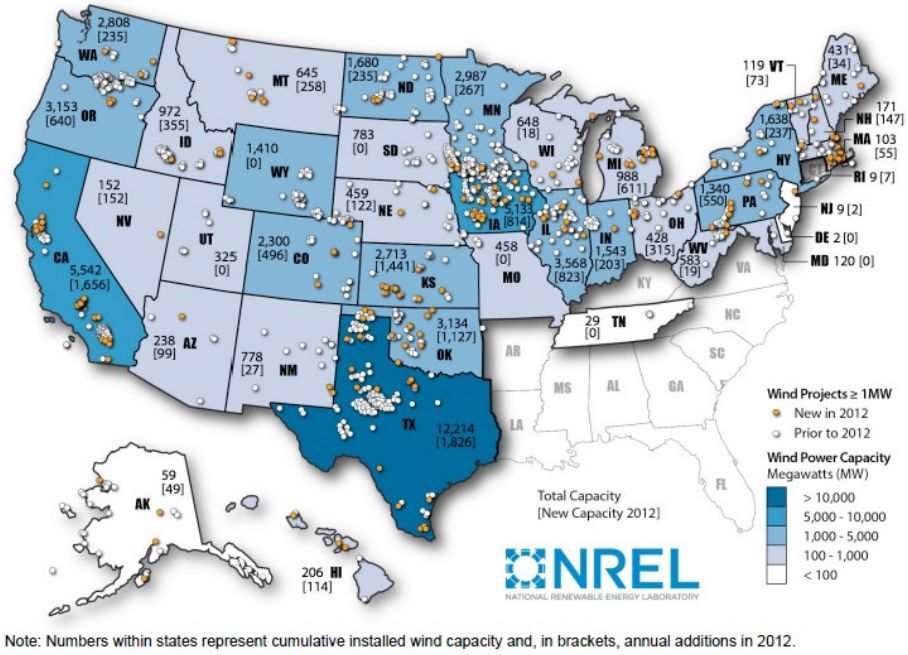


Figure 18. Annual Average Wind Speed at 80 Meters (AWS Truewind, 2010)



Note: Numbers within states represent cumulative installed wind capacity and, in brackets, annual additions in 2012.

Figure 19. Location of Wind Power Development in the United States (Wiser & Bolinger, 2013)

Figure 20 shows a general increasing trend in the average total capacity of wind facilities, peaking at about 100 MW.

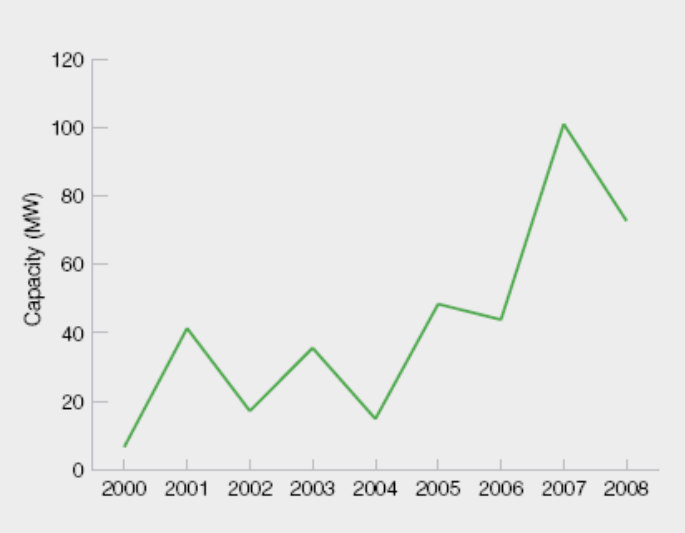


Figure 20. Average wind project size (modified from AWEA (2009))

Figure 21 shows the distribution of the size of turbines installed in 2008, with the 1.5 MW turbine being the mostly commonly used size of wind turbine.

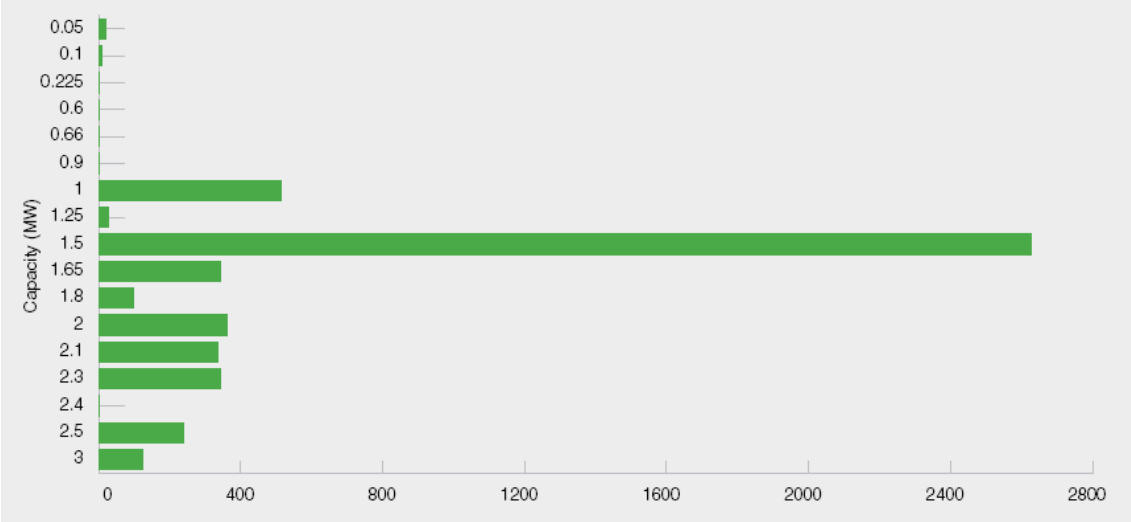


Figure 21. Distribution of wind turbines installed in 2008 by capacity (modified from AWEA (2009)

There are many factors that go into developing a wind facility, but much of it starts with the wind resource assessment (WRA). In general, the areas best suited for wind facilities are known (Figure 18); however, a WRA evaluates the site-specific wind resource to assist in turbine layout and estimating annual energy output from the facility. To perform a WRA for a utility scale wind facility, typically 2-3 meteorological towers are placed on the potential site for approximately one or two years. Meteorological towers are erected to collect wind speed and other meteorological data for sites under consideration for wind power development. Wind speed, wind direction and temperature are recorded in 1 to 10 minute intervals. These data are useful, but a year or two of data is not enough to determine the feasibility of a utility scale wind project (Bailey & McDonald, 1997). To gain better understanding of the wind resource, the onsite data are then correlated with 10-20 years of airport data. After the data are correlated and any adjustments are made, estimates of annual energy generation and potential annual variation are then calculated using both data sets. The methods used for estimating annual

energy generation vary slightly between wind developers and are dependent on the developer's wind software of choice (Thorgersen & Sorensen, 2007).

There are factors beyond wind magnitude that can increase or decrease electricity generation at a wind facility; for example the effects of wind direction, air density, topography and turbulence. Directional specific losses, also referred to as wake affects, occur when wind turbines extract energy from the wind and create a wake downstream where wind speed is subsequently reduced. To minimize the occurrence of these losses, detailed planning takes place when determining the layout of a wind facility. The predominant wind direction(s), topography and high turbulence areas are all factors that affect the layout.

For several reasons, this study does not replicate the typical data analysis methods used in WRA. First, typical WRA calculations are done on an annual basis. This study analyzes data on a monthly and hourly basis. The same methods cannot be used because the annual distribution of the data varies significantly from the monthly and hourly distributions (these distributions are shown in the Results, Sections 3.1 and 3.2). Second, certain steps of the data analysis methods are proprietary and difficult to replicate since they are performed within software programs specifically made for wind analysts. One wind software program, Windographer, was used on a free trial basis, but some of the data analysis methods required for this study were not available, so it could not be used exclusively. Other software programs were explored (Wind Farm and Wind Farmer) but neither allowed the importing of the 20+ years of data utilized for this study. Also, it was cost prohibitive to utilize features beyond the free trial version.

This study does explore the use of some typical data analysis tools. For example, the Weibull probability density function (pdf) is commonly used in the wind energy industry for characterizing the wind resource at the site of a potential wind facility (Ackerman & Soder, 2002; Kantar & Usta, 2007; and Carta, Ramirez, & Bueno, 2008). A pdf as defined by Aliaga & Gunderson (2003) is "The probability distribution of a continuous random variable X is a curve such that the area under the curve over an interval is equal to the probability that the random variable X is in the interval. The values of a continuous probability distribution must be at least 0 and the total area under the curve must be 1."

The use of the Weibull pdf is desired to represent the data because in the wind speed distribution, the mean is not a very descriptive measure of the wind resource. The mean and corresponding variance measures such as the standard deviation and confidence intervals are the most meaningful when the assumptions of normality are met. Wind speeds are not normally distributed and confidence intervals assume normal distribution or a normal variance from the mean. For example, the curves in Figure 22 all represent wind speed distributions that have means of 7 m/s, but the shape of the distributions (which dictate energy generation) are very different.

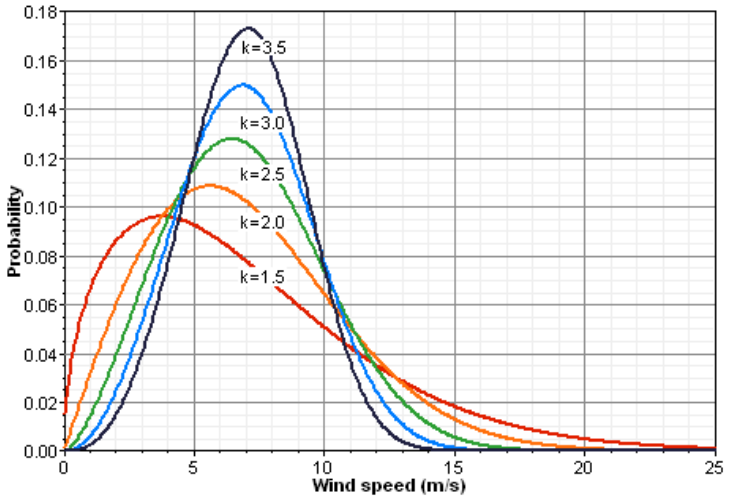


Figure 22. Weibull curves with average wind speed of 7 m/s and varying shape factors (k) (Windographer 1.4 beta).

A pdf is useful because it fits a representative curve to observed data and a function, with specific parameters, is typically generated. A p-value is typically available for how well the pdf fits the empirical data. The curves in Figure 22 are also representative of the Weibull pdf.

Weibull is commonly used and accepted throughout the wind energy field; however, some of the recent wind literature suggests that Weibull does not accurately represent data with a high frequency of low wind speeds (Carta et al, 2008 and Kantar & Usta, 2007). Those two studies present other pdfs, which claim to characterize wind speed distributions with a better fit than Weibull. The Carta et al (2008) paper titled "A joint probablity density function of wind speed and direction for wind energy analysis" was of specific interest to this research, as it offered the inclusion of wind direction. This study intially planned on including the risk of potential changes in wind direction and its effect on wind energy generation. This risk is important because wind facilities are arranged in

a manner where significant changes in the expected wind direction will negatively impact electricity generation. Therefore, it is necessary to account for not only the wind speed probabilities but also to account for the wind turbine array and wake effects (Benninghoff, 2008). The corresponding author, Carta, was contacted to seek assistance with implementing the model suggested in the paper. No assistance was received and without knowledge of the software utilized to perform the analysis (Mathcad), it was soon realized that this method was too complex to pursue. Given that there were no assurances that it would prove to be useful in this study, since the paper was only published recently and the model has not yet been thoroughly tested in the field, this effort was stopped.

Energy and capacity are the main sections throughout the rest of this thesis, as estimates for both values are necessary for structuring an electricity rate. The energy section will evaluate methods for assessing probabilities around monthly wind electricity generation. Capacity differs from energy, as it refers to the amount of energy available to serve the load at any given moment. Similar to the energy section, the capacity section will evaluate methods for assessing probabilities around monthly wind capacity availability. For the purposes of this study, it is important to have a good understanding of the difference between energy and capacity. Energy is typically measured in kilowatt hours (kWh) or megawatt hours (MWH), where 1000 kWh equals 1 MWH. Similarly, typical units for capacity are kilowatts (kW) or megawatts (MW), where 1000 kW equals 1 MW. Because this study involves utility scale wind, the larger units are utilized (MW and MWH). Figure 23 shows the amount of capacity available throughout the day on January 1, 2007 (data set information in section 2.1).

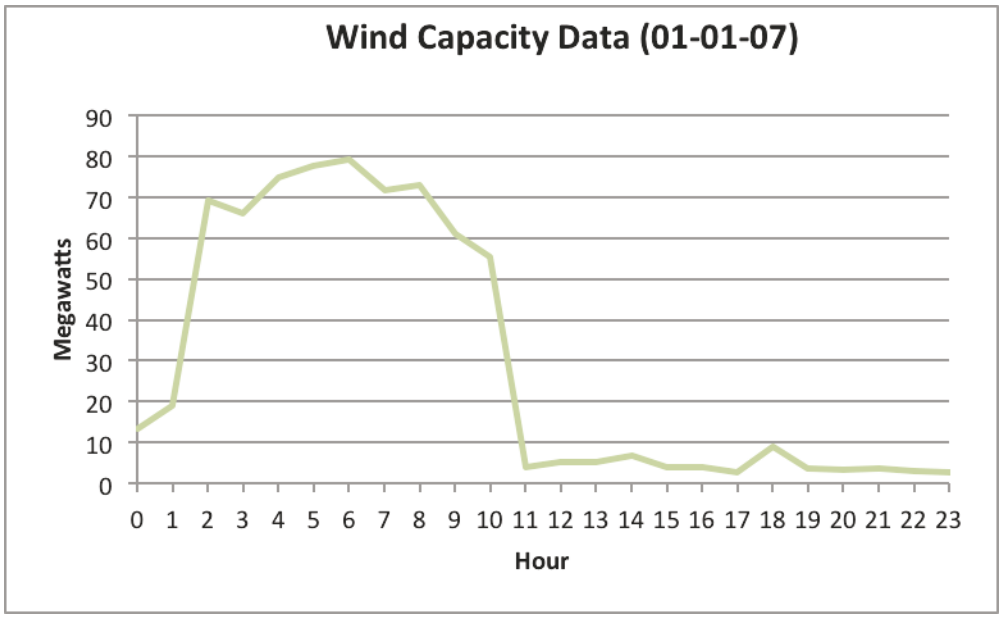


Figure 23. Wind Capacity Data (1/1/07)

At a point during the sixth hour, about 80 MW of capacity was available. Energy differs from capacity, as it is essentially a sum of the available capacity over a given period of time. For example, if the wind speed and wind direction remained relatively constant during hour six, 80 MWH would be generated in that hour.

There has been some research done regarding the amount of capacity available at wind facilities. A study by Archer & Jacobson (2007) aimed to show that if wind projects were aggregated, they could be treated as a source of baseload power. Their study included data from 19 commercially developed wind projects across New Mexico, Colorado, Oklahoma, Texas and Kansas. When the wind projects were combined in blocks of 1, 3, 7, 11, 15 and 19, firm capacities at the 87.5% confidence level were respectively as follows: 0%, 4%, 6%, 10%, 11% and 15%. More specifically as it relates to this research, Archer & Jacobson found that a firm capacity of 0% (no capacity) could

be relied on at the 87.5% confidence level from a single wind project. However, they are arguing for aggregating the wind facilities, where the best scenario (19 projects) offers an estimated firm capacity of 15% at the 87.5% confidence level.

Next, the Methods section discusses details regarding the data used for this study and the methods used for converting hourly wind speeds to theoretical energy generation and capacity values.

CHAPTER 3-METHODS

In order for contributors to the voluntary renewable energy market to financially benefit from a crossover point, a long-term, fixed green rate must be established. Before a fixed green rate can be established, the method for assessing the risk of fixing such a rate needs to be explored. The methods chapter is divided into the following sections:

3.1 Data and Site Information
3.2 Transforming Wind Speeds to Energy and Capacity
3.3 Empirical Energy Analysis
3.4 Bootstrapping Energy Analysis
3.5 Capacity Analysis

The objective of the energy analysis section is to determine a method for assessing statistical certainty around theoretical monthly energy generation (MWh) at the 95% confidence level. The objective of the capacity section is to determine a method for assessing statistical certainty around the amount of capacity (in MW) available at the hourly level at the 95% confidence level. Statistical Analysis Software (SAS) version 9.1.3 was used for the data analysis.

3.1 Data and Site Information

Since the objectives of this study center around assessing statistical certainty, the volume and completeness of the data set utilized was a priority. Unlike the actual development of a utility scale wind project, the magnitude of the annual average wind speed at the site was not of great interest. Given that this study is addressing statistical certainty, the goal was to obtain a large, continuous data set from a single location. The

data set selected contains approximately 22 years of hourly wind speed data measured at the Superior Sewage Treatment Plant in Superior, Wisconsin (see the "A" marker in Figure 24 and Figure 25 for approximate meteorological tower location). Chinook Wind obtained the raw wind data from the site in Superior through the Wisconsin Department of Natural Resources (WDNR) and the Environmental Protection Agency (EPA).

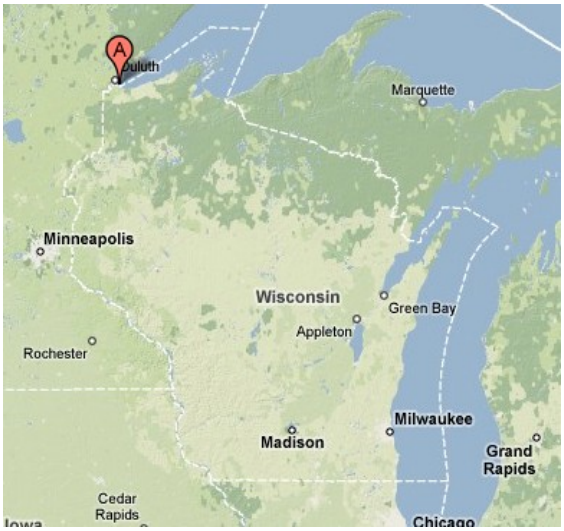


Figure 24. The approximate location of meteorological tower at the Sewage Treatment Plant in Superior, Wisconsin (Google Maps)

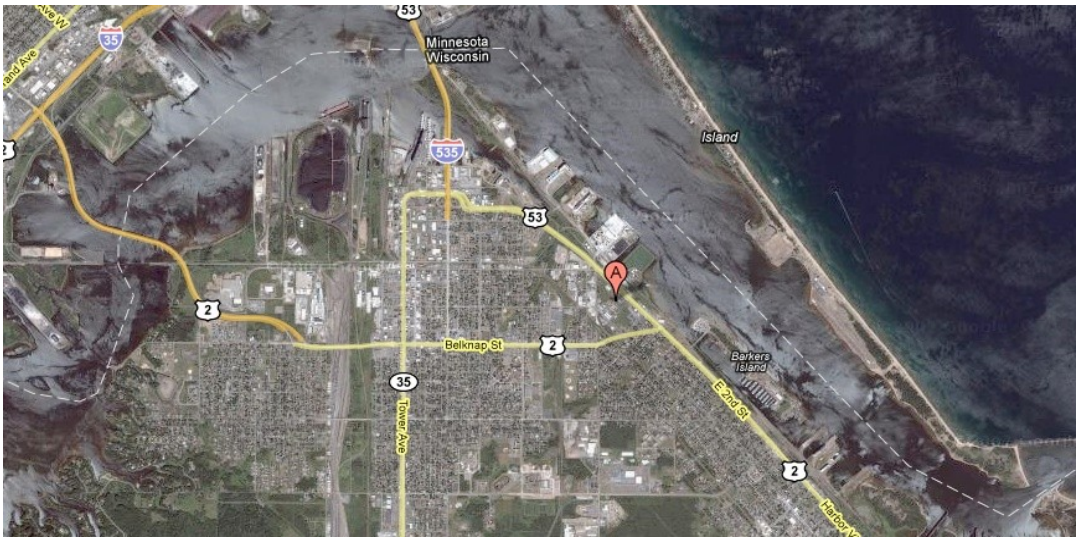


Figure 25. Satellite view of meteorological tower at the Sewage Treatment Plant in Superior, Wisconsin (Google Maps).

Chinook Wind is a consulting firm based in Everson, Washington, that specializes in wind energy applications. Their team, consisting of engineers, meteorologists and wind analysts, acquired and validated the data. There is a standard process for validating the data and it involves data screening and data verfication (Bailey & McDonald, 1997). Data screening uses algorithms to screen the data for suspect values (questionable or erroneous), while data verfication requires a case-by-case decision regarding the suspect values (Bailey & McDonald, 1997). Table 7 summarizes the data verification done by Chinook Wind.

Table 7. Summary of Data Verification

Quality Code	# of hours	Description
0	187,801	no change was made to the data
1	214	one-hour gap of missing data was filled using persistence (using previous hour value)
2	7,912	treated as calm (wind speed set to 0)
3	7,393	treated as ice (wind speed set to 0)

About 8% of the missing data was deemed appropriate for replacement. Wisconsin Public Service acquired the cleaned and validated data from Chinook Wind for use in this research.

Monitoring began at this site on October 2, 1980 and the data set concludes on August 31, 2008. This is a range of about 27 years, but approximately 20 percent of the hourly wind speed and direction data are missing, i.e., the recovery rate is about 80 percent, which is equivalent to approximately 22 years of hourly data (Table 8).

Table 8. Data set summary report (1980-2008) generated by Windographer 1.4 beta.

Data Column Properties

Label	Units	Height	Possible Records	Valid Records	Recovery Rate (%)	Mean	Min	Max	Std. Dev
Wind Speed	m/s	80 m	244,688	196,452	80.29	5.18	0.06	31.05	3.10
Direction	°	17.2 m	244,688	195,202	79.78	172.6	1.0	360.0	106.8
Temperature	°F		244,688	109,478	44.74	42.5	-36.0	101.0	20.3

3.2 Transforming Wind Speeds to Energy and Capacity

The model below (Equation 2) is used to convert hourly wind speed (meters/second) and direction data to theoretical megawatt hours (MWh) generated and capacity available at a 99 MW wind facility.

$$MWh = \begin{cases} ((1500 / (1+ 2.72 \text{ ^ } (6.26 - 0.75 * (WS))))*GF*66/1000) & \text{If } 3.5 \leq WS \leq 25 \\ 0 & \text{otherwise} \end{cases}$$

Where: WS = Wind speed
GF = Generation factor
MWh = Megawatt hours

Equation 2. Model converting wind speed and direction to energy generation (modified from GE Energy (2009))

This research uses 99 MW (equal to 66 turbines rated at 1.5 MW) for the size of the wind facility, as it is representative of the average size of a utility scale wind project in 2007 (Wiser & Bolinger, 2008). The model incorporates the turbine size (1.5 MW) and quantity of turbines (66). This model uses the assumption that wind speed and direction remain constant for the hour. This is not unusual, as hourly averages are often used during data processing and reporting in wind resource assessments (Bailey & McDonald, 1997).

This equation was derived by fitting the power curve of a General Electric (GE) 1.5sle wind turbine to a continuous function. Figure 26 shows technical data and the power curve for the GE 1.5sle.

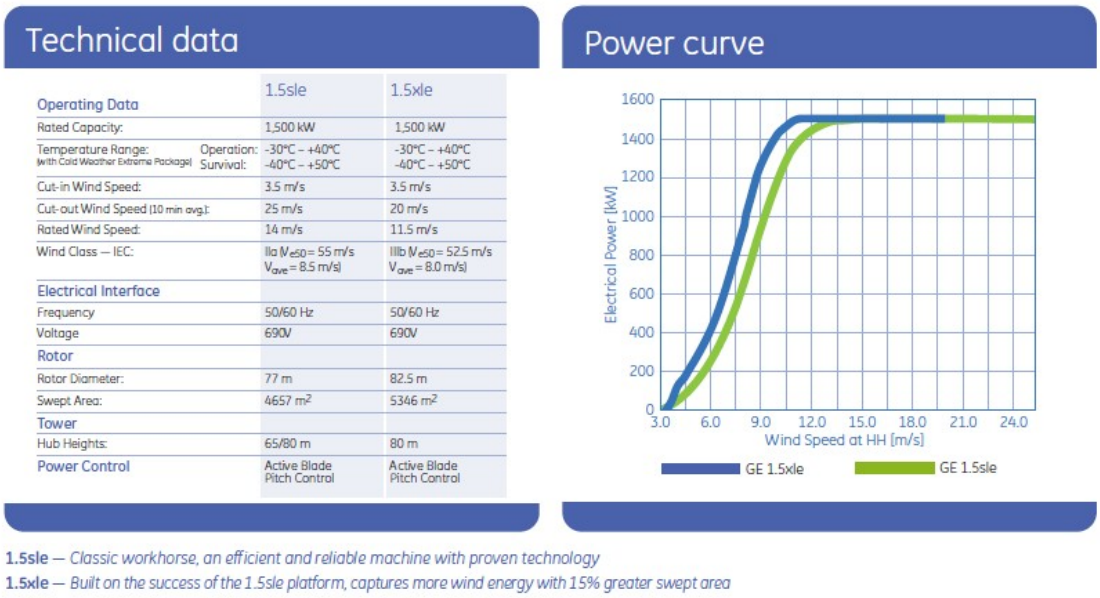


Figure 26. Technical Data and Power Curve for the GE 1.5sle (modified from GE Energy (2009))

The power curve represents the wind turbines conversion of wind speed (which is a measure of the energy in the wind) to electricity generation by the wind turbine. The GE 1.5sle turbine was chosen as it is commonly used, proven technology within the wind industry (GE Energy, 2009).

The generation factor is used to account for directional specific losses; it is representative of the portion of energy left after removing directional specific losses. Table 9 shows a sample of generation factors used.

Table 9. Sample of generation factors based on wind speed and direction (modified from Benninghoff (2010)).

Wind Speed (m/s)	Wind Direction (Degrees)							
	0.00	**22.50**	**45.00**	**67.50**	**90.00**	**112.50**	**135.00**	**157.50**
3.5	59%	72%	69%	69%	75%	66%	67%	65%
4.0	71%	84%	79%	81%	84%	78%	79%	77%
4.5	73%	85%	81%	82%	85%	80%	81%	78%
5.0	77%	88%	84%	85%	88%	83%	84%	82%
5.5	80%	90%	87%	87%	90%	85%	86%	84%
6.0	82%	90%	87%	88%	91%	86%	87%	85%
6.5	83%	91%	88%	89%	91%	87%	87%	86%
7.0	83%	91%	88%	89%	91%	87%	88%	86%
7.5	83%	91%	88%	89%	91%	87%	88%	87%
8.0	84%	91%	89%	89%	92%	88%	88%	87%
8.5	84%	91%	89%	89%	91%	87%	88%	87%
9.0	85%	93%	90%	91%	93%	89%	89%	88%
9.5	87%	94%	91%	92%	94%	90%	91%	90%
10.0	89%	95%	93%	94%	95%	92%	93%	92%
10.5	91%	96%	95%	95%	96%	94%	94%	93%
11.0	94%	98%	96%	97%	97%	96%	96%	95%
11.5	96%	99%	98%	98%	99%	97%	98%	97%
12.0	98%	99%	99%	99%	99%	98%	99%	98%
12.5	99%	100%	99%	99%	100%	99%	99%	99%
13.0	99%	100%	99%	99%	100%	99%	99%	99%
13.5	100%	100%	100%	100%	100%	100%	100%	100%
14.0	100%	100%	100%	100%	100%	100%	100%	100%
14.5	100%	100%	100%	100%	100%	100%	100%	100%
15.0	100%	100%	100%	100%	100%	100%	100%	100%
15.5	100%	100%	100%	100%	100%	100%	100%	100%
16.0	100%	100%	100%	100%	100%	100%	100%	100%
16.5	100%	100%	100%	100%	100%	100%	100%	100%
17.0	100%	100%	100%	100%	100%	100%	100%	100%
17.5	100%	100%	100%	100%	100%	100%	100%	100%
18.0	100%	100%	100%	100%	100%	100%	100%	100%
18.5	100%	100%	100%	100%	100%	100%	100%	100%
19.0	100%	100%	100%	100%	100%	100%	100%	100%
19.5	100%	100%	100%	100%	100%	100%	100%	100%
20.0	100%	100%	100%	100%	100%	100%	100%	100%
20.5	100%	100%	100%	100%	100%	100%	100%	100%
21.0	100%	100%	100%	100%	100%	100%	100%	100%
21.5	100%	100%	100%	100%	100%	100%	100%	100%
22.0	100%	100%	100%	100%	100%	100%	100%	100%
22.5	100%	100%	100%	100%	100%	100%	100%	100%
23.0	100%	100%	100%	100%	100%	100%	100%	100%
23.5	100%	100%	100%	100%	100%	100%	100%	100%
24.0	92%	94%	99%	99%	94%	100%	100%	97%
24.5	81%	65%	72%	81%	64%	98%	94%	85%
25.0	35%	24%	32%	37%	26%	54%	60%	49%

The full table continues to the direction of 337.5 degrees. These generation factors were acquired from an operational wind facility in Wisconsin (Benninghoff, 2008). The direction specific losses are site specific and primarily dependent on the prominent wind directions of the site. These generation factors could be applied to the data in this study because the prominent wind directions were very similar in both sites. To interpret Table 9 further, when the speed is 3.5 m/s and coming from 90 degrees the maximum amount of energy is generated for that wind speed. The generation factor applied to the entire 99 MW facility is 74%; meaning 26% of the energy is lost to directional specific losses. This occurs because the wind speed is very low and not enough energy can be regained by the wind before it reaches the next turbine. Conversely, when the generation factor equals 100%, there are no directional specific losses. This situation occurs when the wind speed is high enough that after wind turbines extract some energy from the wind, the wind can regain enough energy back before it reaches the turbines downstream. In other words, the wind that hits the turbines downstream has enough energy to generate the same amount of energy as the first turbines that extracted energy from that wind stream.

When MWh is equal to zero (see Equation 2), the wind speed is either below the cut-in (3.5 m/s) or higher than the cut-out (25 m/s). The cut-in value refers to the minimum wind speed necessary for the turbine to generate energy. The cut-out value refers to the maximum operational wind speed. For the GE 1.5sle, if the wind speed averages 25 m/s over 10 minutes, the turbine will shut down, minimizing risk of damage due to extreme wind speeds. There are no generation factors listed below 3.5 m/s and above 25 m/s because the turbines are not generating power, regardless of the direction the wind originates.

As of now, the hourly wind speed and direction data has been converted to hourly energy generation. Next, these data will be processed and analyzed by month.

3.3 Methods – Empirical Energy Analysis

After the hourly wind speed values were converted to hourly energy generation, the hourly values were summed to create daily and monthly energy generation values. Gaps in the data produced some incomplete months, which had to be left out for analyzing monthly energy generation. The days and hours that make up the months that had to be excluded for the monthly energy analysis (when complete months were necessary) were able to be utilized in later data analysis (sections 3.4 and 3.5). Figure 27 shows the frequency of days missing on the monthly level. For example, 53 months were missing either one or two complete days of data. To preserve the data, those months that were only missing one or two days of data were prorated using Equation 3.

$$p = o * (t \div c)$$

Where:

p = prorated monthly generation

o = observed monthly generation

t = total days in the month

c = number of days with complete data

Equation 3. Prorating monthly energy generation

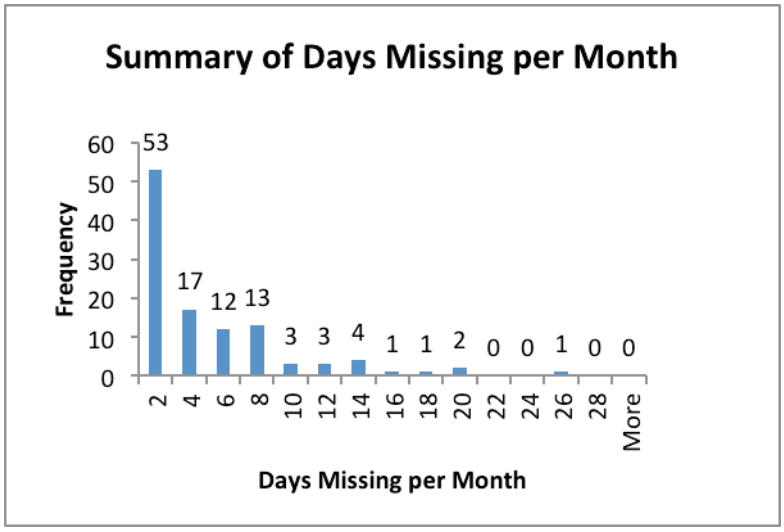


Figure 27. Summary of Days Missing per Month

After the wind speed data were converted to energy, the data were screened for monthly outliers and trending. A total of 12 months were removed; 7 months were deemed outliers and 5 months were causing a trend. The outliers were defined as months that were significantly different among that month. For example, three September months differed significantly from the rest of the population of September months. To identify outliers, the UNIVARIATE procedure (PROC UNIVARIATE) within SAS was used. PROC UNIVARIATE provides a selection of descriptive measures and statistical methods, which can be used to summarize, visualize, analyze and model the statistical distributions of numeric variables (SAS Institute Inc., 2008). PROC UNIVARIATE also reports outliers and extreme observations. Another SAS procedure, PROC REG, was used to assess trending. PROC REG is a general-purpose regression procedure that offers tests of linear hypothesis, predicted values, residuals, confidence limits, correlations and a variety of other analysis (SAS Institute Inc., 2008). The month of March showed a significant decreasing trend over time. The reason for the decrease over time is unknown and it was beyond the scope of this study to create a model to account for trending. Instead, it was decided to remove five of the earlier months for March (1983, 1984, 1985, 1987 and 1988). By removing data from those years, the significant trend for March disappears. Also, this removes days and months that are not representative of the recent conditions, keeping a portion of the high generating days and months out of the data set. This left 221 complete months available for further analysis, with an average of 18 months of data for each month of the year.

The remaining complete months and the prorated months were combined and analyzed by month using PROC UNIVARIATE. This procedure produces summary

statistics such as the mean, standard deviation, maximum, minimum and quantiles. Quantiles or percentiles quantitatively describe the data by dividing it into sections. For example, Table 12 contains monthly quantile data. The 50% quantile (P50) value for January of 7,765 MWh should be interpreted as approximately 50% of the January months generated more than 7,765 MWh. Similarly, it could be stated that approximately 50% of the January months generated less than 7,765 MWh. After the quantiles were calculated for each month, capacity factors were calculated for the quantiles. The capacity factors were calculated by dividing the quantile value by the maximum potential energy generation for that month.

PROC UNIVARIATE also has a histogram feature that plots the frequency distribution of the data and has the option to fit various pdfs to the frequency distribution. Pdfs are useful in that they offer descriptive parameters to represent the data. For example, the normal distribution is a pdf with the parameters (μ, σ), where μ is the mean and σ is the standard deviation. The histogram feature in PROC UNIVARIATE has six well known pdfs to choose from for finding a fit to a frequency distribution. For the distribution of monthly energy generation, the lognormal pdf was ultimately the best fit. When SAS tests if the chosen pdf (lognormal for example) fits the observed frequency distribution, the null hypothesis is: "The distribution is lognormal." In this case, high p-values (high meaning close to 1) would provide greater evidence that the data fit the lognormal distribution. The p-value represents the chance of getting the observed results, computed under the assumption that the null hypothesis is true. For this study, a significance level of 0.05 was chosen. The lognormal distribution is a pdf of a continuous random variable whose logarithm is normally distributed (SAS Institute Inc., 2008).

The fitted lognormal density function is shown in Equation 4 (modified from SAS Institute Inc. (2008)):

$$p(x) = \begin{cases} \frac{100h\%}{\sigma\sqrt{2\pi}(x-\theta)} \exp\left(-\frac{(\log(x-\theta)-\zeta)^2}{2\sigma^2}\right) & \text{for } x > \theta \\ 0 & \text{for } x \leq \theta \end{cases}$$

Where:

θ = threshold parameter

ζ = scale parameter $(-\infty < \zeta < \infty)$

σ = shape parameter $(\sigma > 0)$

h = width of histogram interval

Equation 4. Lognormal density function (modified from SAS Institute Inc. (2008))

The threshold parameter θ must be less than the minimum data value. The threshold parameter is often set to equal 0, but can be estimated as well. When the threshold parameter is estimated, it is used to displace the distribution's position along the x-axis. The scale parameter is closely representative of a mean value and the shape parameter describes the variance of the data (SAS Institute Inc., 2008).

As part of assessing the risk of fixing a green rate, it was decided to estimate a value for monthly energy generation for each month and compare that value to a holdout year. It was decided to compare the P95 quantiles with a year of observed data (2007) that was not included in the analysis. This is similar to assuming that the energy company based the rate on expecting energy generation from the wind facility to be greater than or equal

to the P95 values. All data from 2007 was left out of the energy and capacity analysis and the monthly energy generation from 2007 was calculated for comparison with the observed monthly P95 quantile. The year 2007 was chosen as the holdout year because it is the most recent complete year of data in the set. Using a recent year allows for comparisons between energy output and locational marginal prices (LMP) of energy in the Midwest Independent Transmission System Operator (MISO), as LMP records are only available for recent years.

Stratifying the data by month was done for the ease of data analysis and to simplify interpretations for the green rate analysis. This is because within electricity markets, certain energy and capacity estimates are made on the monthly level and these estimates would affect the green rate. This study decided to explore the difference among months, to see if there is a distinct grouping or pattern of energy generation. The means of the monthly energy data were compared using the analysis of variance SAS procedure (PROC ANOVA), to test if, for example, January energy generation varies significantly from August energy generation and all other months for that matter. The Student-Newman-Keuls (SNK) multiple range test, an option within PROC ANOVA, was used to find group differences after the analysis of variance was found significant (Cody & Smith, 2006).

3.4 Methods – Bootstrapping Energy Analysis

Bootstrapping is comprised of various resampling processes that are useful in simulating a larger sample size enabling further statistical analysis (Efron & Tibshirani,

1993). There was a need to bootstrap the data because the low number of observed sample months resulted in the inability to get good estimates of the P95 and P99 quantiles. The P95, P99 and minimum value are essentially equal for all months and more granular estimates are needed. Also, the increased sample size aides in understanding the frequency distribution of each month.

The observed data are randomly sampled and processed based on the statistical parameters of interest. For this study, the parameter of interest is monthly energy generation. To simulate additional months, a series of steps are performed:

1. Summing the observed hourly energy data for each day, resulting in a new variable, daily energy output (days with less than 24 hours of data were excluded).

2. Stratifying the days by months, creating data sets containing daily energy generation for each month. Each month had approximately 700 days of energy generation.

3. Days are randomly sampled from each monthly data set, without replacement using the PROC SURVEY SELECT in SAS, based on the program presented in Cassell (2007). The number of days randomly sampled was equal to the number of days in the individual month (i.e. 31 for January, 28 for February and so on). This process is repeated 2000 times for each month. The 2000 simulations were chosen based on recommendations from Efron and Tibshirani (1993). They stated that a high number (1000 -2000) of simulations is necessary for high and low quantile estimations, such as the 95% quantiles this study is utilizing.

a. To use January as an example, 31 days are randomly sampled from the total set of 758 observed January days of energy generation and combined to create a new, simulated or bootstrapped month. Next, these 31 days are put back in the January data set and 31 days were randomly sampled again and combined to create a new, bootstrapped month. This process was repeated until there were 2000 bootstrapped months. The process is exactly the same for the other months, except the number of days that are randomly sampled and pooled is equal to the number of days in that specific month.

The analysis to follow is very similar to the analysis done for the observed data. The bootstrapped data will be referred to as the simulated months. The simulated months were analyzed by month using PROC UNIVARIATE. This procedure produces summary statistics such as the mean, standard deviation, maximum, minimum and quantiles. After the quantiles were calculated for each month, capacity factors were calculated for the quantiles. The capacity factors were calculated by dividing the quantile value by the maximum potential energy generation for that month. The summary statistics and quantiles from the observed and simulated data were compared to see how the bootstrapped months compared to the observed data.

The histogram option within PROC UNIVARIATE was utilized to create graphs of the monthly frequency distribution. For the distribution of the simulated monthly energy generation, the lognormal pdf was ultimately the best fit (as it was with the observed monthly data).

As part of assessing the risk of fixing a green rate, it was decided to calculate the amount of energy that could be counted on each month, with a probability of 0.95. This value would then be compared to the monthly energy generation of the holdout year to see if the prediction succeeded. All data from 2007 were left out of the energy and capacity analysis and the monthly energy generation from 2007 was calculated for comparison with the simulated monthly energy quantiles. Comparisons with the holdout year is similar to assuming that the energy company based the green rate on the fact that they were 95% confident that X amount of energy would be generated in that month. If the amount is less than X, the energy company loses money. If the amount is greater than X, the energy company can use that energy in a different market. It would be in the best financial interest of the energy company to underestimate (versus overestimate) the amount of energy generated each month, to mitigate the risk of having to go into the market to get power when prices could be high.

To ensure that there is a 95% probability that all 12 months will be correctly estimated, it was necessary to calculate an additional quantile, P99.6. The use of the P99.6 quantile was determined by using the binomial distribution (Equation 5). The binomial distribution can be used for calculating the probability of success when there are only two defined outcomes. For this study the two defined outcomes are correctly estimating monthly energy generation, or not correctly estimating monthly energy generation.

$$f(k;n,p) = \Pr(K = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k = 0, 1, 2, ..., n$, where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

p = the probability of success in each independent trial

n = the number of trials

k = the number of successful trials

f(k; n,p) = The probability of getting exactly k successes in n trials

Equation 5. Binomial Distribution (modified from SAS Institute Inc. (2008)

Equation 5 was used in calculations for Table 10 and Figure 28; where $n = 12$ (number of months), p (probability of success for an individual month) was varied between the values on the left column in Table 10 (0.87, 0.88 and so on through 0.999) and k (number of months successfully estimated) was varied between 9, 10, 11 and 12. The outputs of these analyses are the values shown in Table 10. The objective was to find when *f(k; n,p)* was approximately equal to 0.95 (the confidence level of interest for this study) and those values are shown in red.

Table 10. Number of Months Correctly Estimated at Various Confidence Levels

	Probability of Correctly Estimating the listed number of Months			
	12/12 Months	11/12 Months	10/12 Months	9/12 Months
P87	0.188	0.525	0.802	0.940
P88	0.216	0.569	0.833	0.954
P89	0.247	0.613	0.862	0.965
P90	0.282	0.659	0.889	0.974
P91	0.322	0.705	0.913	0.982
P92	0.368	0.751	0.935	0.988
P93	0.419	0.797	0.953	0.992
P94	0.476	0.840	0.968	0.996
P95	0.540	0.882	0.980	0.998
P97	0.694	0.951	0.995	1.000
P98	0.785	0.977	0.998	1.000
P99	0.886	0.994	1.000	1.000
P99.5	0.942	0.998	1.000	1.000
P99.6	0.953	0.999	1.000	1.000
P99.7	0.965	0.999	1.000	1.000
P99.8	0.976	1.000	1.000	1.000
P99.9	0.988	1.000	1.000	1.000

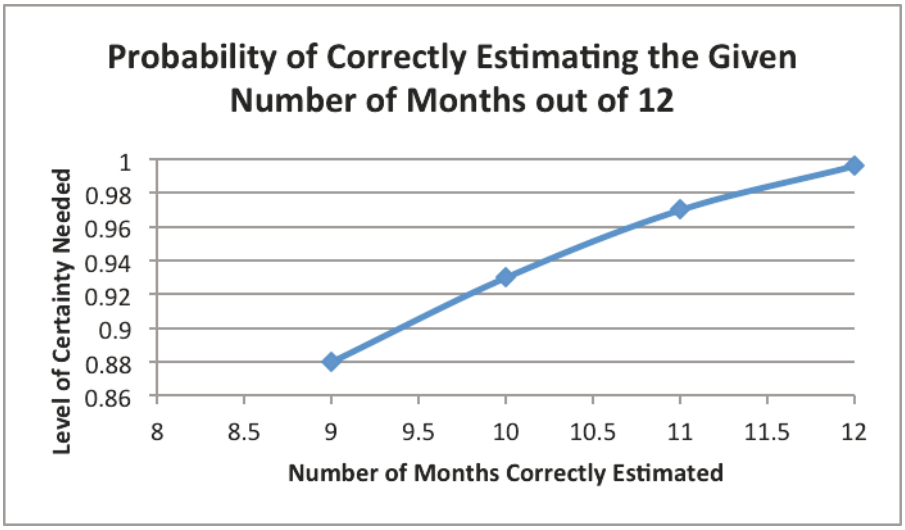


Figure 28. Level of Certainty Need to Correctly Estimate Monthly Value

In regards to later comparisons with the monthly energy output of the holdout year, if the amount of energy at the P99.6 quantile is less than the energy output of the month it is being compared with, the outcome is defined as a success. If the amount of energy at the P99.6 quantile is greater than the energy output of the month it is being compared with, the outcome is defined as a failure. Table 10 and Figure 28 show the probability of getting a certain number of months correct versus the quantile used. For example, if the requirement is to be 95% certain of getting all months correct, the P99.6 should be used. This study is using the P99.6 quantile because it is desired to successfully estimate monthly energy generation for each month (12/12). If the requirement was less conservative, where it was only necessary to be 95% certain of getting at most one month wrong (11/12), the P97 quantile could be used.

Using the P99.6 quantile was not necessary for the observed data as no new results would occur. The P99.6 quantile occurs between the P99 and minimum observed value and the P99 and minimum observed value were equal for every month, meaning the P99.6 would be equivalent to the same value (see Table 11 and Table 12). It is very likely that the P99 and minimum are equal, given the small sample size of the observed months.

Similar to the analysis of the observed monthly energy data, the means of the simulated monthly energy data were compared using the analysis of variance SAS procedure (PROC ANOVA) to test if, for example, January energy generation varies significantly from August energy generation and all other months for that matter. Since all the simulated months fit the lognormal distribution and PROC ANOVA uses the assumption that the means are normally distributed, it was decided to log transform the data prior to running this analysis. Log transforming the lognormal data moves the data

from a log distribution to a normal distribution. Equation 6 is used to perform this transformation, which includes a threshold parameter specific to each month.

$$y = \log (x - t)$$

Where:

y = the normalized variate of the monthly energy generation

x = monthly energy generation

t = the threshold parameter specific to that month

Equation 6. Log transform (Aitchison & Brown, 1963)

After the data were transformed into normal space, the SNK test within PROC ANOVA was used to find group differences.

3.5 Methods – Capacity Analysis

For the capacity analysis, the objective is to determine the amount of capacity available in any hour, stratified by month, with various levels of certainty. For example, this study is striving to fill in the blanks of statement for each month: "There is a 95% probability that the wind facility will generate MW in any given hour and a 50% probability that the wind facility will generate MW in any given hour."

Capacity and wind speeds are analyzed in this section. Wind speeds were initially analyzed using the Weibull pdf because there are methods available for estimating energy generation using the Weibull parameters (Thorgersen & Sorensen, 2007). Also, wind

speeds are easily converted to real-time capacity availability. However, this study was unable to utilize the models correlating Weibull parameters with energy output because the Weibull pdf did not accurately fit the data. Nonetheless, the methods and results are described here to add to the findings in recent literature stating that caution should be exercised when using Weibull to characterize wind speeds shown.

Several pdfs were explored in addition to Weibull prior to using results from an empirical distribution. The wind speed data and capacity data were analyzed by month using PROC UNIVARIATE in SAS. The observed summary statistics and quantiles are reported in the Results section 4.3 and in Appendix III. Within PROC UNIVARIATE there is the option of selecting a specific pdf to characterize the data and this study attempted to represent the hourly wind speed data with the following probability density functions: Weibull, Gamma, Beta and Lognormal. For simplicity and discussion purposes, all the histograms in the capacity section show only the fit of the Weibull curve to the observed data. When SAS tests if the chosen pdf (Weibull for example) fits the observed frequency distribution, the null hypothesis is: "The distribution is Weibull." In this case, high p-values (high meaning close to 1) would provide greater evidence that the data fit the Weibull distribution. The p-value represents the chance of getting the observed results, computed under the assumption that the null hypothesis is true. For this study, a significance level of 0.05 was chosen. The parameters of the pdf from various samples can be compared and tested to determine any significant differences among distributions, if the pdf fits the data with a high level of confidence. This would be useful in determining if the wind speed distribution varies on an annual or monthly basis. The three parameter Weibull pdf is shown in Equation 7.

$$\frac{c}{\sigma}\left(\frac{x-\theta_0}{\sigma}\right)^{c-1}\exp\left(-\left(\frac{x-\theta_0}{\sigma}\right)^{c}\right)$$

Range : $x > \theta_0$

	Location	Scale	Shape
Parameter	θ_0	σ	*C*

Equation 7. Weibull density function (SAS Institute Inc., 2008)

The parameters are often identified differently. For example, the shape factor is usually referred to as "k". This study will utilize the parameter notation shown in the table above. The location parameter is representative of the minimum value for the data, the scale parameter is a weighted mean and the shape parameter describes the sharpness of the peak of the curve. This study uses a two parameter Weibull, which is what is normally used. A two parameter Weibull fixes the location at zero, as opposed to estimating the location (SAS Institute Inc., 2008).

Similar to the analysis done in the energy sections, the means of the monthly wind speed data were compared using the analysis of variance SAS procedure (PROC ANOVA), to test if, for example, average January wind speeds vary significantly from average August wind speeds and all other months for that matter. The SNK multiple range test was used to find group differences after the analysis of variance was found significant.

CHAPTER 4-RESULTS

The Results section is divided into three sections, similar to the Methods sections 3.3 – 3.5. For simplicity and consistency, only the graphs from April, December and the annual level will be shown within the text. April and December were chosen because they are the windiest and calmest months, respectively. Appendix I includes all figures from the empirical energy analysis, Appendix II includes the figures from bootstrapped data and Appendix III includes all figures from the capacity analysis. Each month was analyzed for trending and the results can be found in Appendix IV, Figure 79 through Figure 90.

4.1 Results – Empirical Energy Analysis

Table 11 summarizes basic information on the observed data on the monthly level and the quantiles are listed by month in Table 12. The quantiles can be interpreted as the percent of observed values that fall above the given quantile. For example, in January 95% of the observed months exceed 4,017 MWh. It should be noted that the P95 values, P99 values and minimum observed month are essentially identical for all months. This is result of the number of observed months in the data set. The P95 value is less than 95% of the data and at most there are 22 months per month (see Table 11). Considering 5% of 22 equals 1.1 or approximately 1 month, which results in 95% of the months being greater than that one month. That month then serves as the P95, P99 and minimum value for 9 out of the 12 months. In January, February and June the P95 and P99 vary slightly; however the P99 and minimum values are equal.

Table 11. Summary Statistics of Observed Data (in MWh)

Month	# of complete months	Mean	Standard Deviation	Maximum	Minimum
January	21	7,899	2,350	12,285	3,906
February	22	8,558	2,436	13,600	5,315
March	17	11,980	3,337	19,290	6,538
April	19	15,335	3,882	25,105	8,894
May	19	12,700	2,996	19,157	8,322
June	22	9,988	3,283	16,331	4,385
July	16	7,964	2,059	13,167	4,769
August	16	8,137	2,849	14,464	4,569
September	16	8,557	2,155	12,714	5,481
October	19	11,356	4,105	17,470	4,181
November	17	11,096	3,718	19,156	5,846
December	17	7,548	2,580	10,619	2,957

Table 12. Observed Quantiles - Observed Data (in MWh)

	P1	P5	P10	P25	P50	P75	P90	P95	P99
January	12,285	12,271	10,204	9,463	7,765	6,069	5,640	4,017	3,906
February	13,600	13,139	12,645	9,266	8,051	6,895	5,917	5,320	5,315
March	19,290	19,290	17,243	13,572	11,196	10,102	7,507	6,538	6,538
April	25,105	25,105	22,235	17,598	14,400	12,101	11,436	8,894	8,894
May	19,157	19,157	17,289	14,097	13,241	9,998	8,454	8,322	8,322
June	16,331	14,014	13,656	12,525	10,757	6,756	5,816	4,677	4,385
July	13,167	13,167	10,365	9,218	7,438	6,682	5,646	4,769	4,769
August	14,464	14,464	13,263	8,925	7,473	6,144	5,616	4,569	4,569
September	12,714	12,714	12,051	9,970	8,496	6,925	5,658	5,481	5,481
October	17,470	17,470	17,051	14,665	11,942	7,591	4,724	4,181	4,181
November	19,156	19,156	16,351	13,658	11,015	8,472	6,002	5,846	5,846
December	10,619	10,619	10,504	9,488	8,654	5,804	3,197	2,957	2,957

The capacity factors in Table 13 were calculated using the quantiles in Table 12. April offers the highest capacity factors and December has the lowest.

Table 13. Monthly Capacity Factors Calculated from Observed Monthly Quantiles

	P50	P75	P90	P95
January	10.54%	8.24%	7.66%	5.45%
February	12.10%	10.36%	8.89%	8.00%
March	15.20%	13.72%	10.19%	8.88%
April	20.20%	16.98%	16.04%	12.48%
May	17.98%	13.57%	11.48%	11.30%
June	15.09%	9.48%	8.16%	6.56%
July	10.10%	9.07%	7.67%	6.47%
August	10.15%	8.34%	7.62%	6.20%
September	11.92%	9.72%	7.94%	7.69%
October	16.21%	10.31%	6.41%	5.68%
November	15.45%	11.89%	8.42%	8.20%
December	11.75%	7.88%	4.34%	4.01%
Average Annual	**13.89%**	**10.80%**	**8.74%**	**7.58%**

Figure 29, Figure 30, Figure 31 and Figure 32 show the histogram output from PROC UNIVARIATE. The P95 line was added to Figure 29 to serve as visual connection between the quantiles and the frequency distribution.

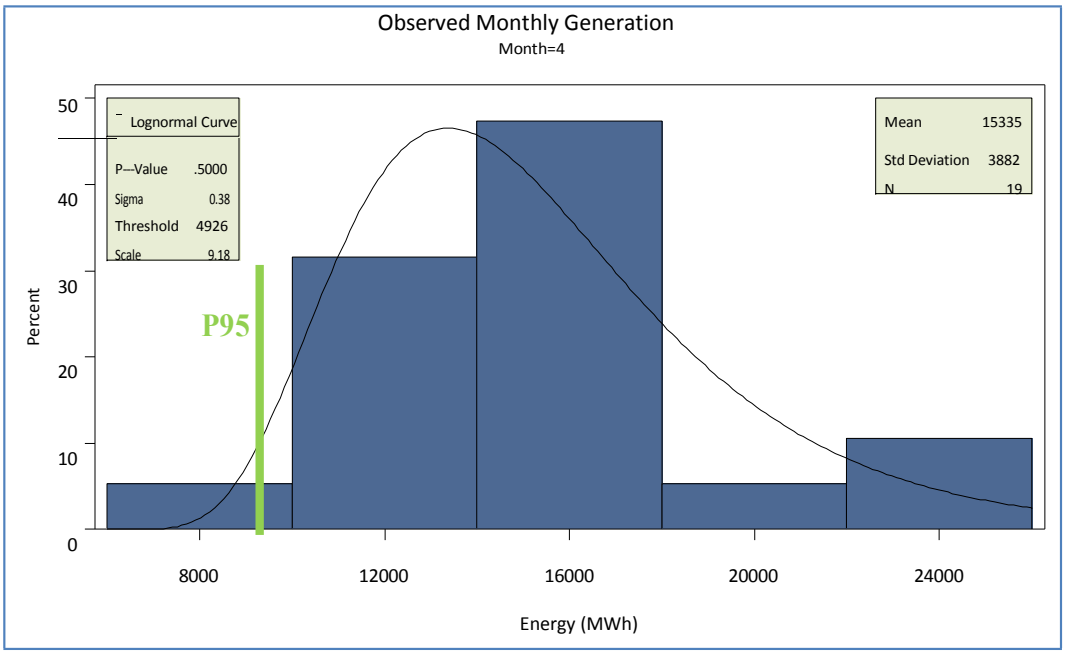


Figure 29. April Energy Generation - Observed Data

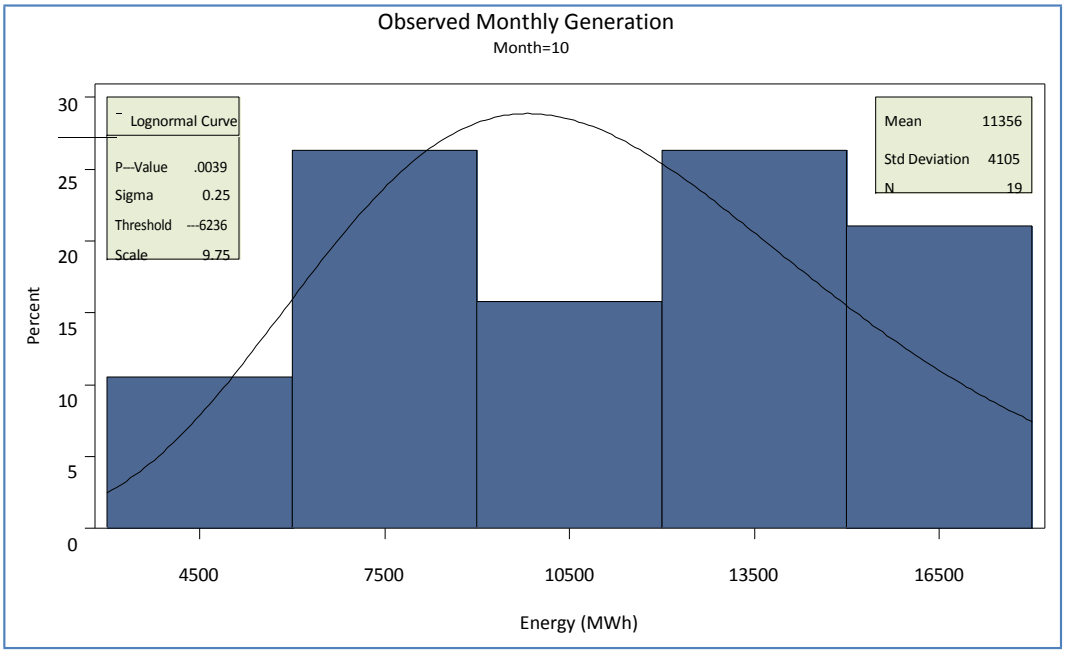


Figure 30. October Energy Generation - Observed Data

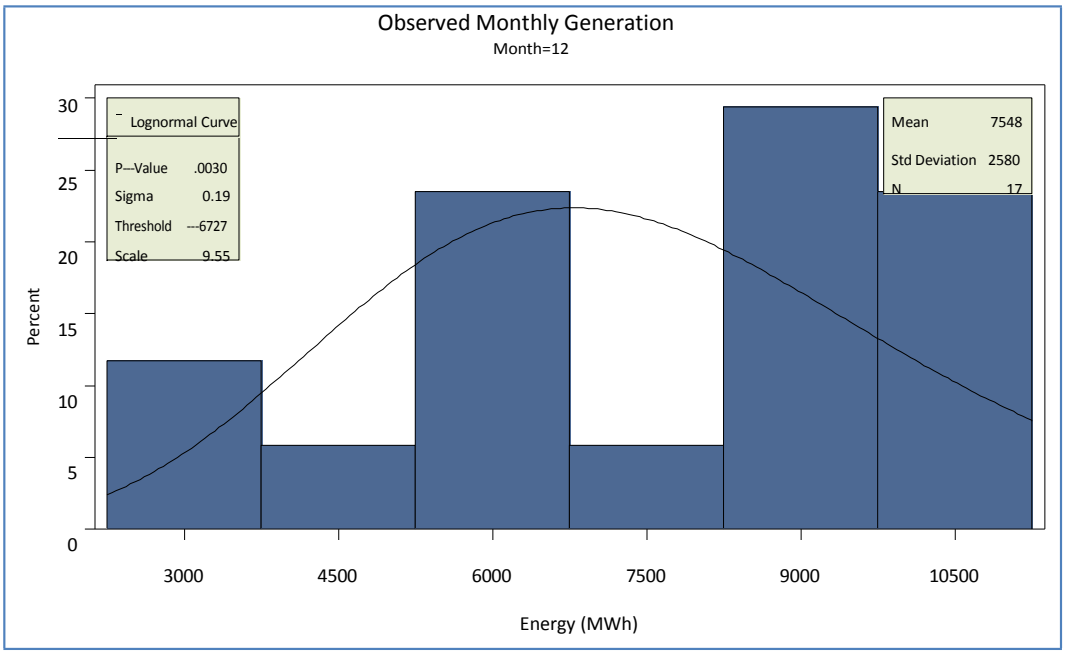


Figure 31. December Energy Generation - Observed Data

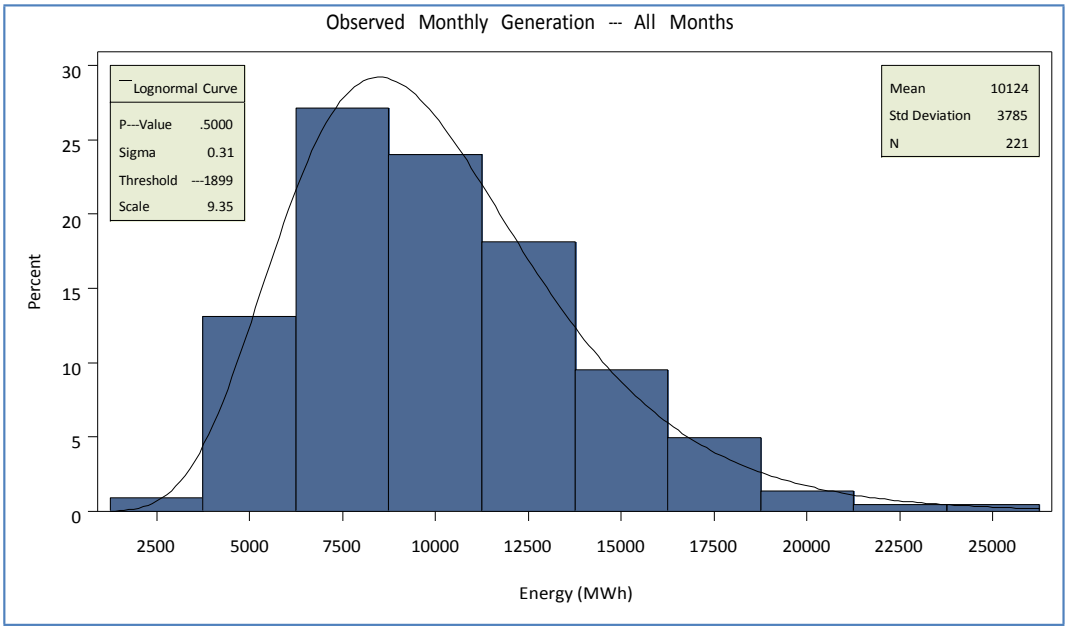
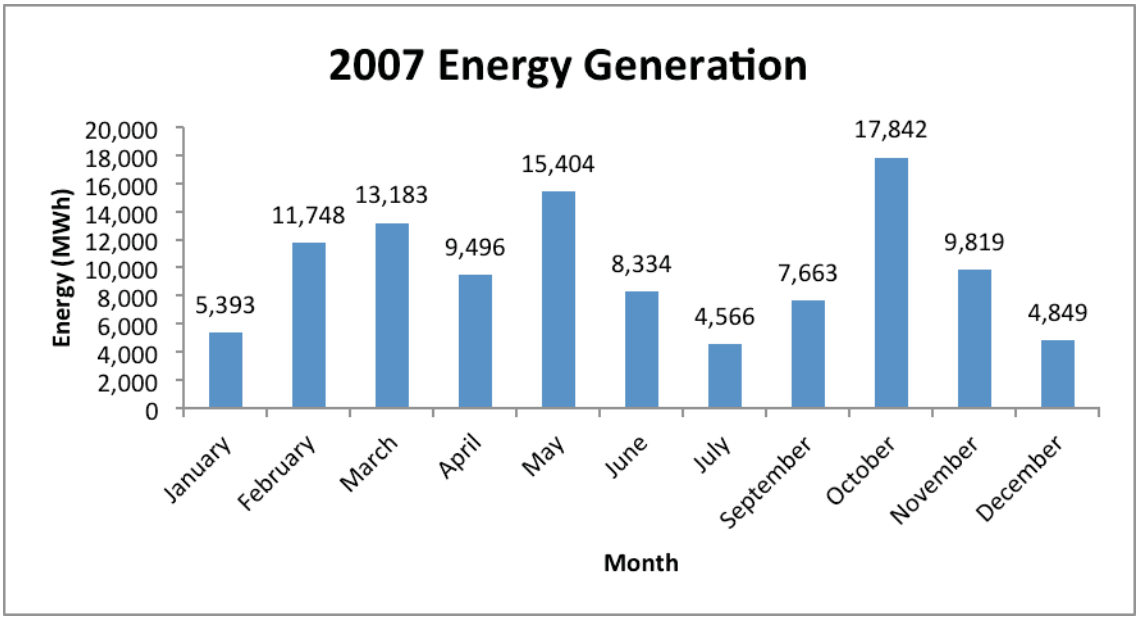


Figure 32. Monthly Energy Generation - All Months - Observed Data

Each of the figures show the lognormal curve fit to the distribution; however, the fit for October and December is very poor. The distribution of October is shown in addition because the distribution of October and December is unique from the 10 other months. Both months show a higher frequency of months with high energy generation, whereas the 10 other months all have a higher frequency of months with low energy generation. This has implications on the results of bootstrapping, discussed in Section 5.1. The histograms for the rest of the observed months are available in Appendix I. Eight of the observed months fit the lognormal distribution (p-value > 0.05) and four of the observed months did not fit the lognormal distribution (p-value < 0.05).

Figure 33 shows the observed energy generation by month for the holdout year, 2007. It should be noted that there are no data shown for August 2007, as more than two days of data were missing from that month in the observed data set.



*Data not available for the month of August

Figure 33. Monthly Energy Generation - 2007 (Holdout Year)

Table 14 shows the comparison between the observed P95 values and the holdout year. Recall that the P95 values are essentially equivalent to the minimum observed values. The July 2007 generation was less than the P95 estimate for July because July 2007 had the lowest energy generation of all the July's on record within the data set.

Table 14. Comparing 2007 Monthly Generation to Observed P95 Estimate (MWh)

	2007	Observed - P95	Difference (2007 - P95)
January	5,393	4,017	1,376
February	11,748	5,320	6,428
March	13,183	6,538	6,645
April	9,496	8,894	602
May	15,404	8,322	7,082
June	8,334	4,677	3,657
July	4,566	4,769	(203)
August	full month not available	-	-
September	7,663	5,481	2,182
October	17,842	4,181	13,661
November	9,819	5,846	3,973
December	4,849	2,957	1,892
Totals	108,297	61,002	47,295

*negative values represent a month where the 2007 generation was less than the observed P95 value

Table 15 shows the output from the SNK option in PROC ANOVA. The SNK grouping gives the same letter to means that are not significantly different. For example,

April is the only month that is significantly different from all others months. All months besides April share some similarities to another month.

Table 15. Comparing Observed Monthly Means (SNK test in PROC ANOVA)

SNK Grouping			Mean (MWh)	N	Month
	A		15042.7	20	April
	B		12835.5	20	May
	B				
C	B		12047.2	18	March
C	B				
C	B		11679.9	20	October
C	B				
C	B	D	11025	18	November
C		D			
C	E	D	9915.7	23	June
	E	D			
	E	D	8696.3	23	February
	E	D			
	E	D	8504.2	17	September
	E				
	E		8136.8	16	August
	E				
	E		7784.7	22	January
	E				
	E		7763.9	17	July
	E				
	E		7397.9	18	December

*Means with the same letter are not significantly different.

4.2 Results – Bootstrapping Energy Analysis

Table 16 lists summary statistics for the bootstrapped data on the monthly level and the quantiles are listed by month in Table 17. The summary statistics of the observed data (Table 11) were subtracted from the summary statistics of the bootstrapped data (Table 16) and Table 18 shows the percent difference between the observed and bootstrapped summary statistics. The quantiles of the observed data (Table 12) were subtracted from the quantiles of the bootstrapped data (Table 17) and Table 19 shows the percent difference between the observed and bootstrapped quantiles. In both, the negative values represent a case were the bootstrap value was less than the observed value. The capacity factors in Table 20 were calculated using the quantiles in Table 17. Similar to the observed data, April offers the highest capacity factors and December has the lowest.

Table 16. Summary Statistics of Bootstrapped Data (in MWh)

Month	Mean	Standard Deviation	Maximum	Minimum
January	8,292	1,979	16,280	3,117
February	8,756	2,149	16,295	3,416
March	12,326	2,697	23,965	4,467
April	16,131	3,336	27,875	6,680
May	13,587	2,846	27,012	4,386
June	9,996	2,486	20,211	3,676
July	7,837	1,860	15,220	3,130
August	8,232	2,128	15,794	2,371
September	8,679	1,965	19,380	3,896
October	12,158	2,486	22,679	5,097
November	11,766	2,840	23,782	4,459
December	7,695	1,774	15,132	3,056

Table 17. Lognormal Quantiles - Bootstrapped Data (in MWh)

	P1	P5	P10	P25	P50	P75	P90	P95	P99
January	13,669	11,806	10,900	9,507	8,122	6,891	5,899	5,354	4,420
February	14,539	12,557	11,586	10,087	8,584	7,238	6,145	5,541	4,499
March	19,348	17,024	15,863	14,036	12,160	10,435	9,000	8,191	6,769
April	24,477	21,831	20,478	18,301	16,001	13,819	11,952	10,877	8,941
May	20,888	18,510	17,311	15,409	13,435	11,599	10,055	9,178	7,622
June	16,517	14,342	13,260	11,564	9,833	8,250	6,941	6,207	4,921
July	12,723	11,091	10,279	9,009	7,713	6,530	5,552	5,004	4,044
August	13,948	11,992	11,034	9,551	8,064	6,729	5,645	5,045	4,009
September	13,873	12,126	11,261	9,911	8,542	7,297	6,273	5,702	4,705
October	18,619	16,485	15,417	13,736	12,008	10,416	9,091	8,344	7,029
November	19,302	16,757	15,500	13,542	11,561	9,765	8,293	7,473	6,046
December	12,515	10,846	10,033	8,785	7,543	6,439	5,551	5,062	4,224

Table 18. Percent Difference between Bootstrapped Summary Statistics and Observed Summary Statistics

Month	Mean	St Dev	Max	Min
1	6.51%	-15.97%	32.52%	-20.20%
2	0.59%	-12.64%	19.82%	-35.73%
3	2.32%	-17.02%	24.24%	-31.68%
4	7.23%	-16.56%	11.03%	-24.89%
5	0.01%	-24.89%	16.44%	-47.30%
6	-1.92%	-27.56%	22.22%	-16.17%
7	0.94%	-13.77%	15.59%	-31.45%
8	-0.52%	-24.46%	9.20%	-48.11%
9	0.42%	-7.27%	52.43%	-28.92%
10	4.09%	-41.51%	27.11%	21.91%
11	2.17%	-24.13%	24.15%	-23.73%
12	-0.99%	-31.32%	42.50%	3.35%

*negative values represent a statistic where the bootstrap estimate was the given percent less than the observed statistic

Table 19. Percent Difference between Bootstrapped Quantiles and Observed Quantiles

	P1	P5	P10	P25	P50	P75	P90	P95	P99
January	11.3%	-3.8%	6.8%	0.5%	4.6%	13.5%	4.6%	33.3%	13.2%
February	6.9%	-4.4%	-8.4%	8.9%	6.6%	5.0%	3.9%	4.2%	-15.3%
March	0.3%	-11.7%	-8.0%	3.4%	8.6%	3.3%	19.9%	25.3%	3.5%
April	-2.5%	-13.0%	-7.9%	4.0%	11.1%	14.2%	4.5%	22.3%	0.5%
May	9.0%	-3.4%	0.1%	9.3%	1.5%	16.0%	18.9%	10.3%	-8.4%
June	1.1%	2.3%	-2.9%	-7.7%	-8.6%	22.1%	19.4%	32.7%	12.2%
July	-3.4%	-15.8%	-0.8%	-2.3%	3.7%	-2.3%	-1.7%	4.9%	-15.2%
August	-3.6%	-17.1%	-16.8%	7.0%	7.9%	9.5%	0.5%	10.4%	-12.3%
September	9.1%	-4.6%	-6.6%	-0.6%	0.5%	5.4%	10.9%	4.0%	-14.2%
October	6.6%	-5.6%	-9.6%	-6.3%	0.6%	37.2%	92.4%	99.6%	68.1%
November	0.8%	-12.5%	-5.2%	-0.8%	5.0%	15.3%	38.2%	27.8%	3.4%
December	17.9%	2.1%	-4.5%	-7.4%	-12.8%	10.9%	73.6%	71.2%	42.9%

*negative values represent a quantile where the bootstrap estimate was the given percent less than the observed quantile

Table 20. Capacity Factors Calculated from Bootstrapped Monthly Quantiles

	P50	P75	P90	P95
January	11.03%	9.36%	8.01%	7.27%
February	12.90%	10.88%	9.24%	8.33%
March	16.51%	14.17%	12.22%	11.12%
April	22.45%	19.39%	16.77%	15.26%
May	18.24%	15.75%	13.65%	12.46%
June	13.79%	11.57%	9.74%	8.71%
July	10.47%	8.86%	7.54%	6.79%
August	10.95%	9.14%	7.66%	6.85%
September	11.98%	10.24%	8.80%	8.00%
October	16.30%	14.14%	12.34%	11.33%
November	16.22%	13.70%	11.63%	10.48%
December	10.24%	8.74%	7.54%	6.87%
Average Annual	**14.26%**	**12.16%**	**10.43%**	**9.46%**

Figure 34 and Figure 35 show the histogram output from PROC UNIVARIATE. Both figures show a good fit to the lognormal distribution, as do the rest of the months (see Figure 54 through Figure 65 in Appendix II).

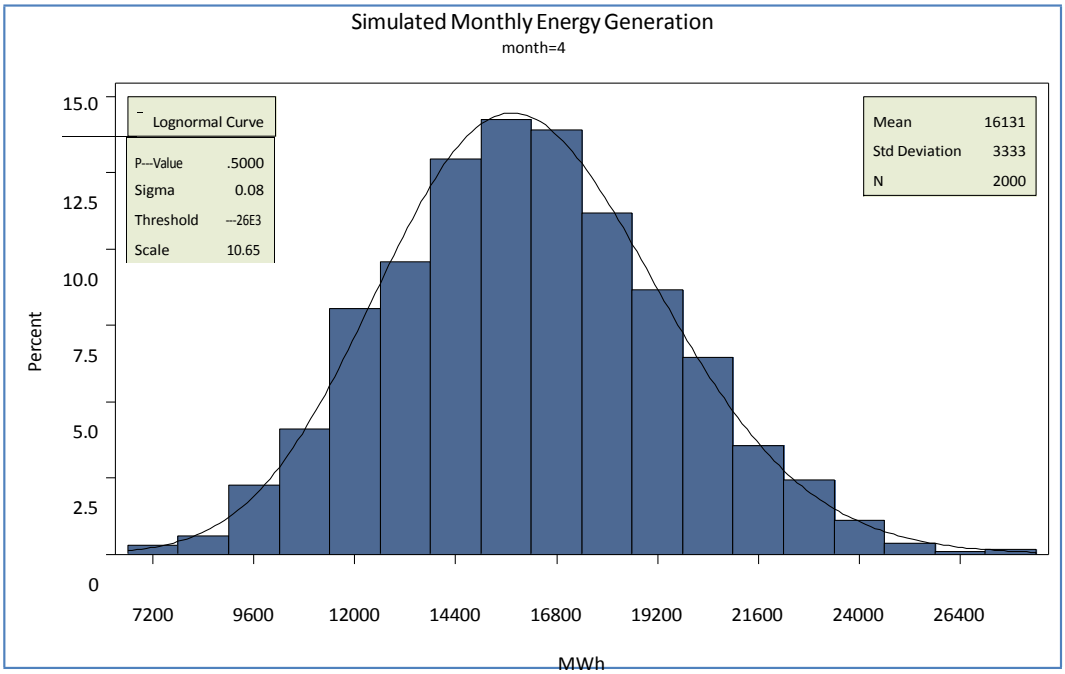


Figure 34. April Energy Generation - Bootstrapped Data

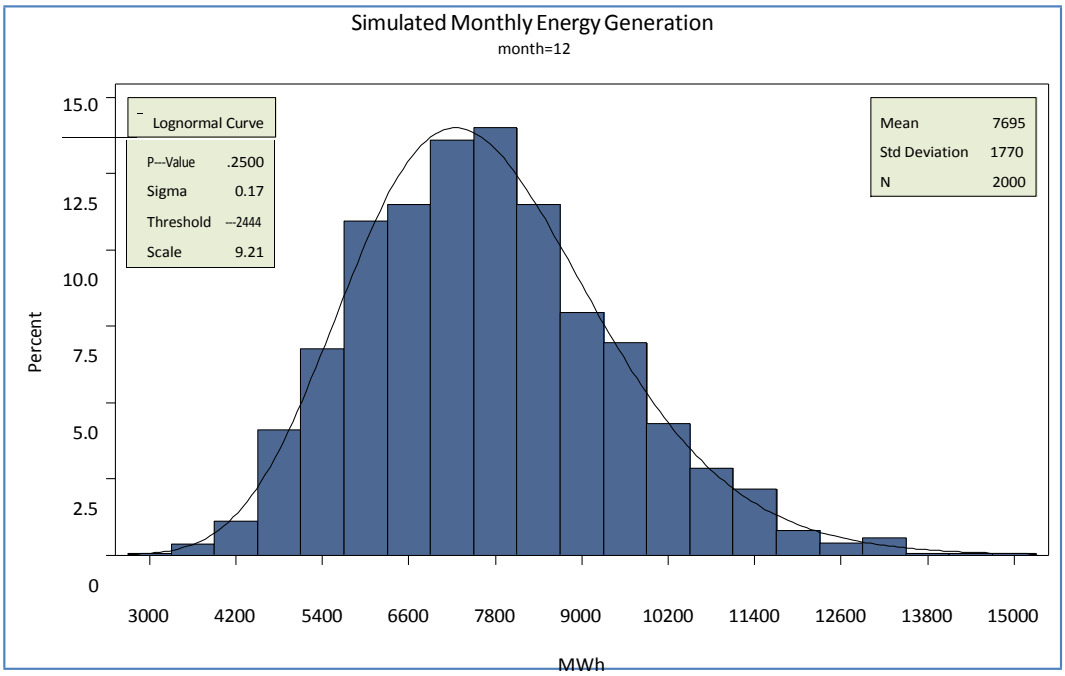


Figure 35. December Energy Generation - Bootstrapped Data

Table 21 shows the comparison between the bootstrapped P99.6 values and the holdout year. All months in 2007 have a higher energy generation than the P99.6 values. It should be noted that there are no data shown for August 2007, as more than two days of data were missing from the month in the observed data set.

Table 21. Comparing 2007 Monthly Generation to Bootstrap Estimate (MWh)

	2007	Bootstrap - P99.6 95% Confidence Level	Difference (2007 - P99.6)
January	5,393	4,059	1,334
February	11,748	4,079	7,669
March	13,183	6,238	6,945
April	9,496	8,339	1,157
May	15,404	6,874	8,530
June	8,334	4,429	3,905
July	4,566	3,486	1,080
August	data not available	-	-
September	7,663	4,444	3,219
October	17,842	6,401	11,441
November	9,819	5,509	4,310
December	4,849	3,899	950
Total	**108,297**	**57,756**	**50,541**

Figure 36 shows the distribution of daily energy generation. These are the data that were sampled from for the bootstrap procedure. This figure includes days from all months, but this distribution was very similar when broken down to the monthly level. Note the range of daily energy generation spans the theoretical potential, starting at zero and approaching the maximum potential of 2376 MWh/day.

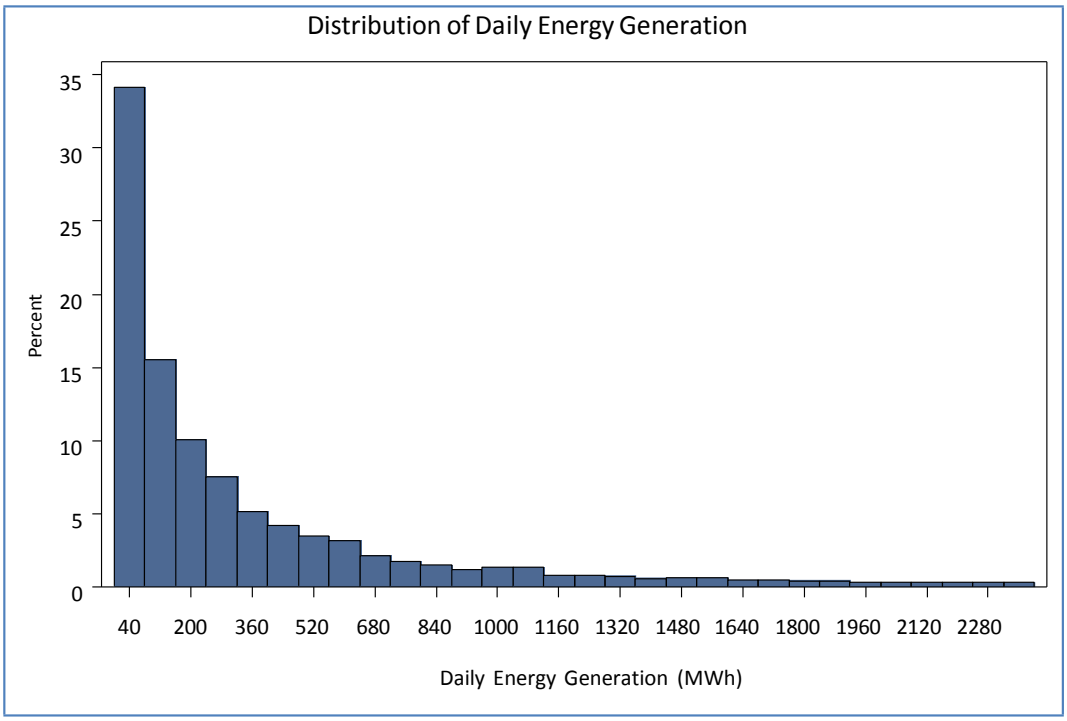


Figure 36. Daily Energy Generation - All Months

Table 22 shows the results of the SNK test done with PROC ANOVA on the bootstrapped log-transformed monthly means. The results suggest that July and September are not significantly different from each other, but do differ from all other months. In addition, the results suggest that February and August are not significantly different from each other, but do differ from all other months. The remainder of the months not previously mentioned are grouped alone, suggesting that they are significantly different from all other months.

Table 22. Comparing Bootstrapped Log Transformed Monthly Means (SNK test in PROC ANOVA)

SNK Grouping	Mean	N	Month
A	10.651299	2000	April
B	10.181724	2000	May
C	9.979557	2000	March
D	9.91145	2000	October
E	9.861831	2000	November
F	9.826668	2000	June
G	9.525942	2000	July
G	9.523994	2000	September
H	9.476131	2000	August
H	9.47349	2000	February
I	9.317846	2000	January
J	9.209035	2000	December

*Means with the same letter are not significantly different.

4.3 Results – Capacity Analysis

The annual distribution of wind speeds is shown in Figure 37 and the monthly distributions for April and December are shown in Figure 38 and Figure 39, respectively. All three distributions do not fit the Weibull curve (p-value < 0.05), nor do the rest of the months (see Figure 67 through Figure 78 in Appendix III).

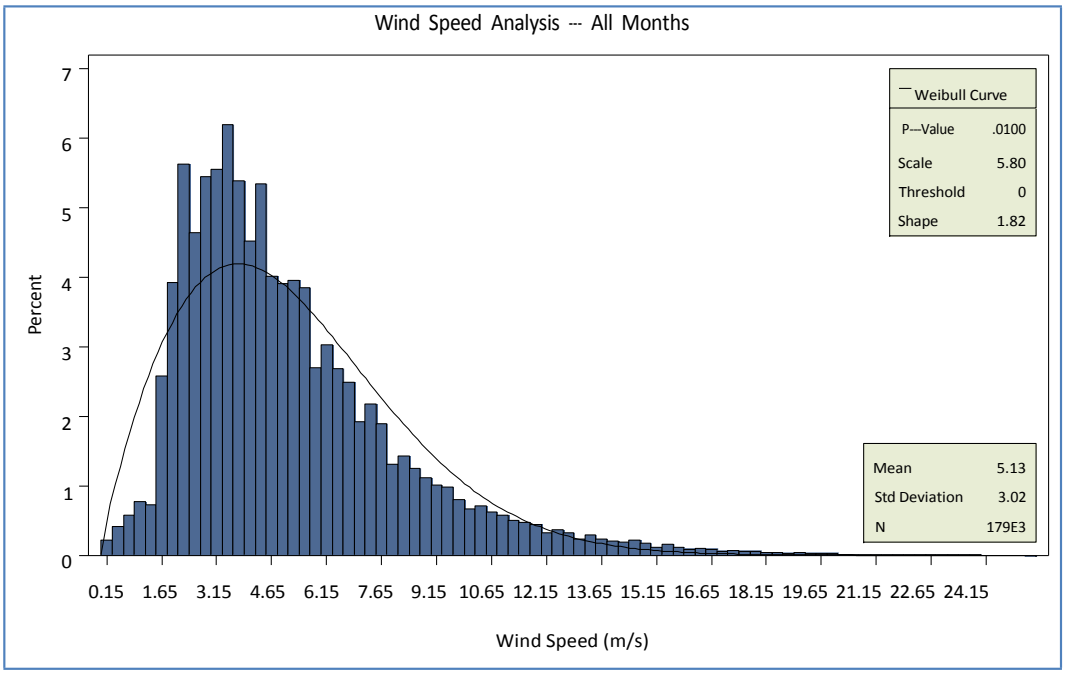


Figure 37. Wind Speed Analysis - All Months

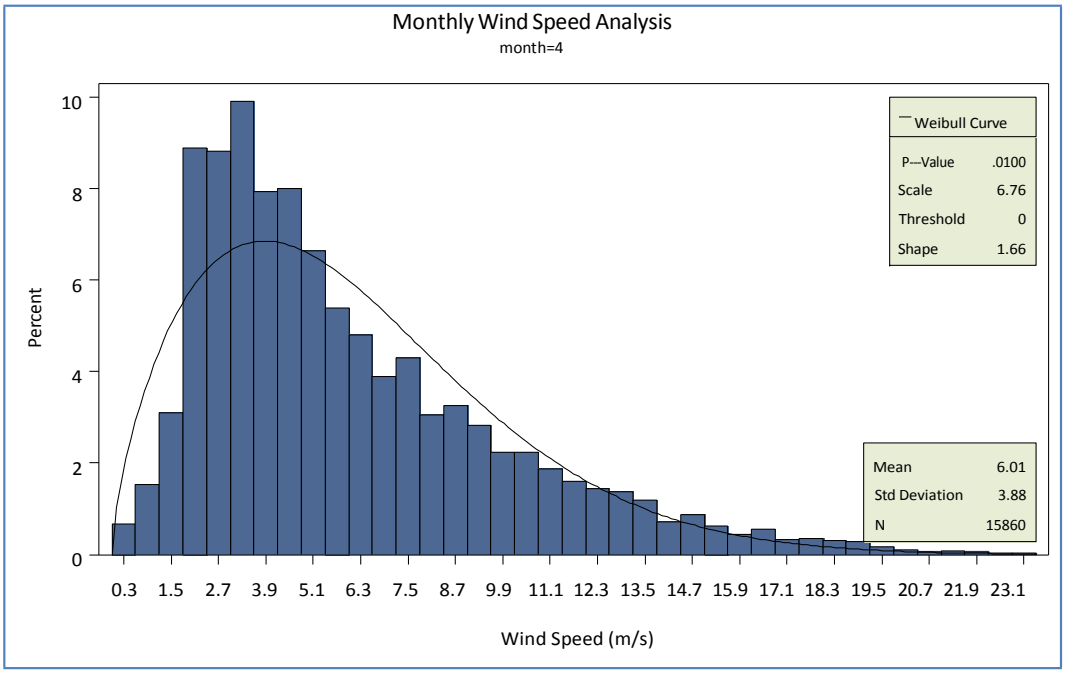


Figure 38. April Wind Speed Analysis

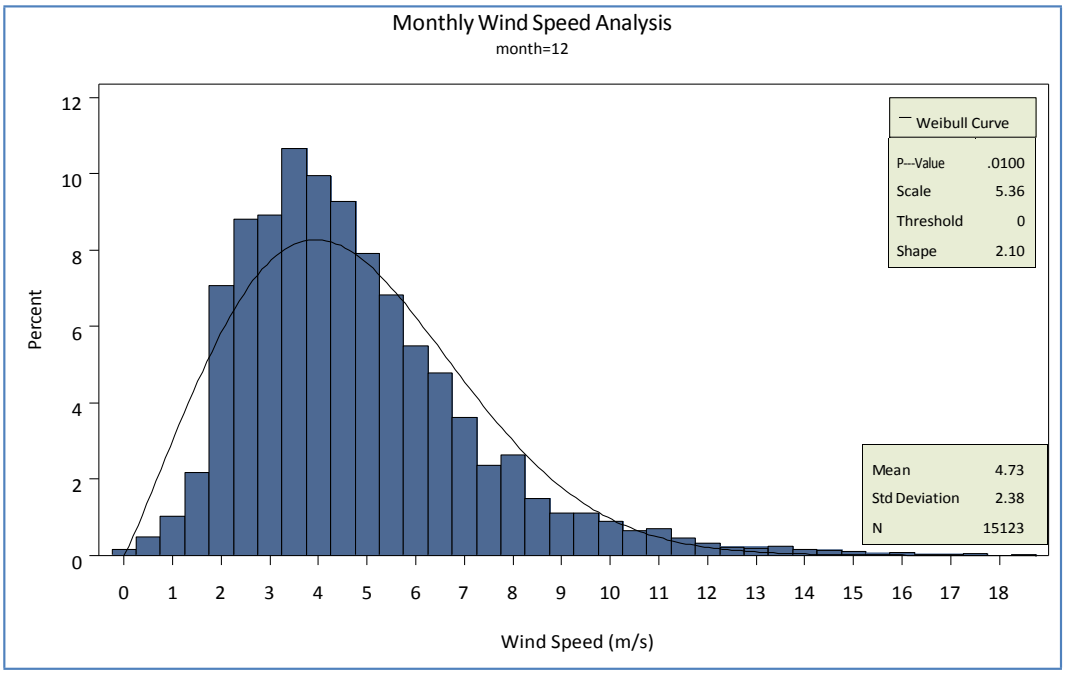


Figure 39. December Wind Speed Analysis

Table 23 shows the capacity quantiles by month for the observed data. When the values are not equal to 0, the given portion of capacity can be relied on. For example, in April, 1.6 MW of capacity or greater is available 69% of the time. No capacity is available in any month between the P70 and P99 quantiles.

Table 23. Capacity Quantiles by Month (MW)

Month	P1	P5	P25	P50	P55	P61	P62	P63	P64	P65	P66	P67	P68	P69	P70	P99
January	97.1	65.8	12.7	3.4	2.4	1.8	1.8	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
February	98.2	75.5	15.2	3.9	3.1	2.1	2.0	1.9	1.8	1.7	0.0	0.0	0.0	0.0	0.0	0.0
March	98.9	92.9	22.8	4.8	3.7	2.4	2.3	2.1	2.1	2.0	1.9	1.8	1.6	0.0	0.0	0.0
April	98.9	97.2	39.1	5.6	4.1	3.0	2.5	2.3	2.2	2.1	2.0	1.9	1.8	1.6	0.0	0.0
May	98.7	93.6	25.4	4.8	3.7	2.4	2.3	2.2	2.1	2.0	1.9	1.9	1.7	0.0	0.0	0.0
June	98.2	85.2	15.6	3.4	2.5	1.9	1.8	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
July	94.1	61.9	11.4	3.0	2.2	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
August	95.8	67.2	12.1	3.1	2.1	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
September	97.0	68.9	13.8	3.4	2.5	1.9	1.8	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
October	98.5	86.2	22.2	5.0	3.8	2.8	2.4	2.2	2.1	2.1	1.9	1.9	1.8	1.6	0.0	0.0
November	98.9	92.7	20.3	4.6	3.5	2.3	2.2	2.1	2.0	1.9	1.8	1.7	0.0	0.0	0.0	0.0
December	94.7	60.4	11.7	3.5	2.9	2.0	1.9	1.9	1.8	1.7	0.0	0.0	0.0	0.0	0.0	0.0

Table 24 shows the output from the SNK option in PROC ANOVA. In this table the means of the monthly wind speeds are being compared. The SNK grouping gives the same letter to means that are not significantly different. April, June, February, September and July all differ significantly from each other and from all other months. All months besides these share some similarities to another month(s).

Table 24. Comparing Monthly Mean Wind Speeds (SNK test in PROC ANOVA)

SNK Grouping		Mean	N	Month
	A	6.01132	15860	April
	B	5.54625	16019	May
	B			
C	B	5.48187	12793	March
C				
C	D	5.41397	15551	November
	D			
	D	5.39701	15974	October
	E	5.02327	15810	June
	F	4.95366	14729	February
	G	4.83617	12441	September
	H	4.73377	15123	December
	H			
	H	4.71488	16363	January
	H			
	H	4.69583	13820	August
	I	4.62692	14671	July

*Means with the same letter are not significantly different.

CHAPTER 5-DISCUSSION

The following sections will discuss the ability of the methods used in this study to predict monthly energy and capacity values with the desired level of certainty. The implications of the results on the fixing a wind energy rated are also addressed.

5.1 Discussion – Energy Analysis

The distribution and number of observed months proved to limit the type of statistical analysis that could be performed. There were not enough months to get an accurate summary of the monthly distribution. This conclusion is also suggested by Moon & Haynes (2006) who state that the growing body of research is showing the need for wind data covering a span of 30 years or more to accurately represent the wind resource. The highest number of complete months available for this study was 22 (see Table 11). When represented in a histogram, eight of the months in the observed data follow the lognormal distribution (see observed data in Figure 41 through Figure 52 in Appendix I) and the remaining four could not be characterized by a known probability density function. Since there was not a pdf that could be utilized for every month, the observed quantile estimates were compared to the monthly energy generation of the holdout year. The holdout year exceeded the observed P95 values in 10 out of 11 months (August 2007 unavailable for analysis, Table 14). The P95 value for July overestimated the July 2007 generation by 202 MWh. This is not unexpected, as July 2007 has the lowest energy generation of all the July's in the observed data set. The analysis of the observed data suggests a need for

more data to ensure the extreme quantiles are well populated and to achieve general understanding of the monthly distributions. To address the lack of data and a probability density function, it was decided to simulate additional months through bootstrapping.

The bootstrap method of simulating additional months succeeded in more completely populating the extreme quantiles, as would be expected since 2000 months were simulated for each month of the year. The summary statistics and quantiles of the simulated months were compared to the observed months. Several patterns emerged during this comparison. First, the mean was slightly overestimated by the bootstrap for 11 out of 12 months. The highest overestimate occurred for October and was a modest 6.6%. The standard deviation was underestimated for all 12 months, primarily because the high number of samples in the simulated data drove the standard deviation down. Efron & Tibshirani (1993) offer methods for correcting the standard deviations, but implementing those methods within SAS proved to be very challenging and the results would not substantiate the effort. Future research should attempt to correct this issue. As expected, the monthly maximums of the simulated months were greater than that of the observed. Similarly, the monthly minimums of the simulated months were less than that of the observed, for 10 out of 12 months. Together this illustrates an increase in the number of months in the extreme ends of the distribution, which was desired. Also, the comparison of the P99.6 bootstrap quantiles to the holdout year yielded success for all 11 months compared (Table 21). Recall that in order to be 95% certain that the energy generation would exceed the estimated value for 12 out 12 months, it was necessary to use the P99.6 quantile and not the P95 (Table 10). The results gave more conservative energy estimates

than the observed quantiles for 8 out of 11 months and performed better than the observed based on how success was defined.

The monthly means were analyzed to see if the grouping by a single month was arbitrary, when in fact there were larger seasonal trends. This was not the case. The only discernible conclusion from this analysis when comparing the observed and the bootstrapped results is that April is significantly different from all other months. All other groups in the observed results show too much overlap for meaningful interpretation. The bootstrap results present a clearer grouping (Table 22), but more evidence would be needed before confidently concluding that July and September or August and February, could be grouped for predictions regarding energy generation.

5.2 Discussion – Capacity Analysis

All 12 months, along with the combined annual distribution did not fit the Weibull distribution (see Figure 66 through Figure 78 in Appendix III). This study confirms findings in recent literature which state that caution should be exercised when using Weibull to characterize wind speeds (Garcia-Bustamante et al, 2008). In the methods within wind literature and in some wind software, the Weibull parameters are used to estimate annual and monthly energy generation (Thorgersen & Sorensen, 2007). If the Weibull distribution does not provide a good fit to the data, the parameters will not properly describe the data, leading to compromised energy generation estimates.

The analysis continued without using a pdf to represent the data. The monthly capacity quantiles were analyzed to estimate the available capacity. In April, 1.6 MW of

capacity or greater is available 69% of the time. No capacity is available in any month between the P70 and P99 quantiles. Arguably, this small amount of available capacity could result from the poor wind resource at the site from which this data was recovered. However, the study done by Archer & Jacobson (2007) estimated the firm capacity of a single wind project in the "Midwest" at an 87.5% probability was zero. This further suggests that very limited amounts of capacity, if any, can be relied upon with a high level of certainty. Therefore, since no capacity can be counted on with the required level of certainty, it is concluded that a single wind project's contribution to a fixed green rate product must be that of an energy-producing asset only.

CHAPTER 6-CONCLUSION

To assess the risk involved in fixing wind energy rates, this study explored methods that estimate monthly wind energy and capacity values with associated levels of certainty. Even with the robust data set used, the sample size of the monthly data proved too low to accurately estimate the quantiles needed for this study. The bootstrap method of simulating additional months succeeded in more completely populating the extreme quantiles and succeeded in comparisons with the holdout year. This study suggests that no capacity can be counted on from a single wind facility at a 95% level of certainty and this assertion was confirmed within the wind literature. This study also adds to the growing body of research that questions and cautions the use of Weibull to estimate wind energy generation. In summary, it is concluded that a single wind project's contribution to a fixed green rate product must exclude any capacity claims and be that of an energy-producing asset only and that the P99.6 quantile should be utilized for confidently estimating monthly energy generation.

REFERENCES

Ackerman, T., & Soder, L. (2002). An Overview of wind enery-status 2002. *Renewable and Sustainable Energy Reviews , 6*, 67-128.

Aliaga, M., & Gunderson, B. (2003). *Interactive Statistics. Second Edition.* Upper Saddle River: Prentice Hall.

Ankrum, N. (2010). *Austin Energy's Wrinkle in Time: Early GreenChoice customers' low rates will soon be gone with the wind.* Austin Chronicle. Retrieved March 2014 from: http://www.austinchronicle.com/news/2010-11-12/austin-energys-wrinkle-in-time/

Archer, C. L., & Jacobson, M. Z. (2007). Supplying Baseload Power and Reducing Transmission Requirements by Interconnecting Wind Farms. *Journal of Applied Meteorology and Climatology* , 46: 1701-1717.

Austin Energy. (2014). GreenChoice 2014 Renewable Energy. Retrieved March 2014 from: http://austinenergy.com/wps/portal/ae/Programs/GreenChoice/

AWEA. (2009). *Annual Statistics on U.S. Wind Energy.* Retrieved from http://www.awea.org/publications/reports/AWEA-Annual-Wind-Report-2009.pdf

AWS Truewind. (2010, April 7). *Wind Maps and Wind Resource Potential Estimates*. Retrieved May 2010, from Wind Powering America: http://www.windpoweringamerica.gov/wind_maps.asp

Bailey, B. H., & McDonald, S. L. (1997). *Wind Resource Assessment Handbook.* National Renewable Energy Laboratory.

Benninghoff, R. D. (2008, May). Personal Communication.

Bird, L., & Lokey, E. (2007). Interaction of Compliance and Voluntary Renewable Energy Markets. *The Electricity Journal , 21*, 18-30.

Bird, L., Holt, E., & Carroll, G. (2007). *Implications of Carbon Regulation for Green Power Markets.* National Renewable Energy Laboratory.

Bird, L., Kreycik, C., & Friedman, B. (2009). *Green Power Marketing in the United States: A Status Report (2008 Data).* National Renewable Energy Laboratory.

Carta, J. A., Ramirez, P., & Bueno, C. (2008). A joint probablity density function of wind speed and direction for wind energy analysis. *Energy Conversion & Management* , 49: 1309-1320.

Cassell, D. (2007). Don't Be Loopy: Re-Sampling and Simulation the SAS Way. *SAS Global Forum. Statistics and Data Analysis*.

Cody, R. P., & Smith, J. K. (2006). *Applied Statistics and the SAS Programming Language. 5th Edition.* Upper Saddle River: Pearson Prentice Hall.

Davenport, C. (2014, February 4). E.P.A. Staff Struggling to Create Pollution Rule. *The New York Times.* Retrieved March 2014 from: http://www.nytimes.com/2014/02/05/us/epa-staff-struggling-to-create-rule-limiting-carbon-emissions.html?_r=0

DOE. (2012, May). Retrieved March 2014, from U.S. Department of Energy, Green Power Markets, Green Pricing Utility Programs by State: http://apps3.eere.energy.gov/greenpower/markets/pricing.shtml?page=1

DSIRE. (2010, May). Retrieved May 27, 2010, from Rules, Regulations and Policies for Renewable Energy: http://www.dsireusa.org/summarytables/rrpre.cfm

DSIRE. (2014, Feb). Retrieved Feb 17, 2014, from Summary Maps: http://www.dsireusa.org/documents/summarymaps/RPS_map.pdf

Efron, B., & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap.* Chapman & Hall/CRC.

Garcia-Bustamante, E., Gonzalez-Rouco, J. F., Jimenez, P. A., Navarro, J., & Montavez, J. P. (2008). The Influence of the Weibull Assumption in Monthly Wind Energy Estimation. *Wind Energy* , 11: 483-502.

Global Wind Energy Council. (2013). Retrieved February 17, 2014 from Global Wind Report 2012: http://www.gwec.net/wp-content/uploads/2012/06/Annual_report_2012_LowRes.pdf

GE Energy. (2009). Retrieved April 29, 2010, from 1.5 MW Wind Turbine : http://www.gepower.com/prod_serv/products/wind_turbines/en/downloads/GEA14954C15-MW-Broch.pdf

Grubb, M. (2012). Cap and Trade Finds New Energy. *Nature*, 491: 666-667.

Heeter, J., & Nicholas, T. (2013). *Status and Trends in the U.S. Voluntary Green Power Market (2012 Data).* National Renewable Energy Laboratory.

Kantar, Y. M., & Usta, I. (2007). Analysis of Wind Speed Distributions: Wind Distribution Function Derived from Minimum Cross Entropy Principles as a Better Alternative to Weibull Function. *Energy Conversion & Management*, 49: 962-973.

Logan, J., & Kaplan, S. M. (2008). *Wind Power in the United States: Technology, Economic and Policy Issues.* CRS Report for Congress. Order Code RL34546.

Moon, D. A., & Haynes, S. E. (2006). The Long-Term Wind Resource. *AWEA WINDPOWER 2006.* WindLogics Inc.

Newell, R.G., Pizer, W.A., & Raimi, D. (2013). Carbon Markets 15 Years after Kyoto: Lessons Learned, New Challenges. *Journal of Economic Perspectives*, 27 no. 1: 123-146.

Pool, S. (2010). *The Proof Is in the Pudding: Regional Greenhouse Gas Initiative Shows Pollution Pricing Works.* Retrieved March 2014 from: http://www.americanprogress.org/issues/2010/03/rggi_roadmap.html

RENEW Wisconsin. (2009, July). Retrieved May 2010, from At A Glance --Utility Green Pricing Programs in Wisconsin: http://www.renewwisconsin.org/grpowr/greenpow.htm

SAS Institute Inc. (2008). Retrieved April 30, 2010, from SAS OnlineDoc® 9.1.3. Cary, NC: SAS Institute Inc.: http://support.sas.com/onlinedoc/913/docMainpage.jsp

Schneider, J., Dutzik, T. & Sargent, R. (2013). *Wind Energy for A Cleaner America II: Wind Energy's Growing Benefits for Our Environment and Our Health.* Enviroment Ohio. Retrieved March 2014 from: http://www.environmentohiocenter.org/sites/environment/files/reports/OH_WindEnergy_scrn_0.pdf

Swezey, B., Aabakken, J., & Bird, L. (2007). *A Preliminary Examination of the Supply and Demand Balance for Renewable Electricity.* National Renewable Energy Laboratory.

The American Power Act. (2010, May). *The American Power Act: Short Summary.* Retrieved May 2010, from John Kerry - U.S. Senate: http://kerry.senate.gov/americanpoweract/pdf/APAShortSummary.pdf

Thorgersen, M. L., & Sorensen, T. (2007). Measure-Correlate-Predict Methods: Case Studies and Software Implementation. *EMD International A/S*.

Wiser, R., & Barbose, G. (2008). *Renewable Portfolio Standards in the United States: A Status Report with Data Through 2007.* Lawerence Berkely National Laboratory.

Wiser, R., & Bolinger, M. (2008). *Annual Report on U.S. Wind Power Installation, Cost and Performance Trends: 2007.* U.S. Department of Energy: Energy Efficiency and Renewable Energy.

Wiser, R., & Bolinger, M. (2009). *2008 Wind Technologies Market Report.* U.S. Department of Energy: Energy Efficiency and Renewable Energy.

Wiser, R., & Bolinger, M. (2011). *2010 Wind Technologies Market Report.* U.S. Department of Energy: Energy Efficiency and Renewable Energy.

Wiser, R., & Bolinger, M. (2013). *2012 Wind Technologies Market Report*. U.S. Department of Energy: Energy Efficiency and Renewable Energy.

APPENDIX I. TABLES AND FIGURES – OBSERVED DATA

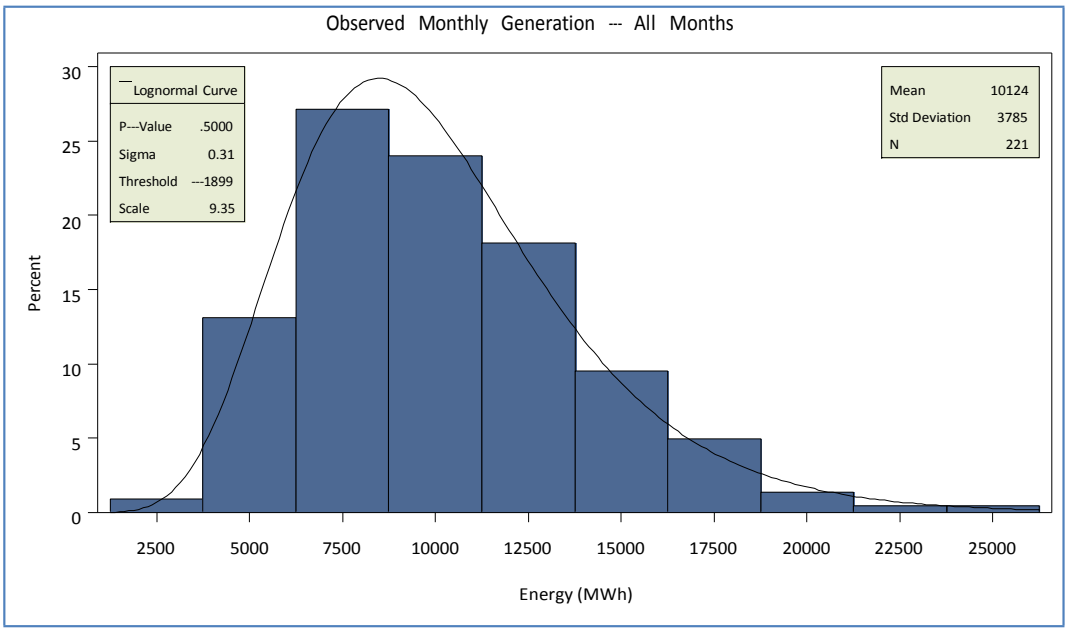


Figure 40. Monthly Energy Generation - All Months - Observed Data

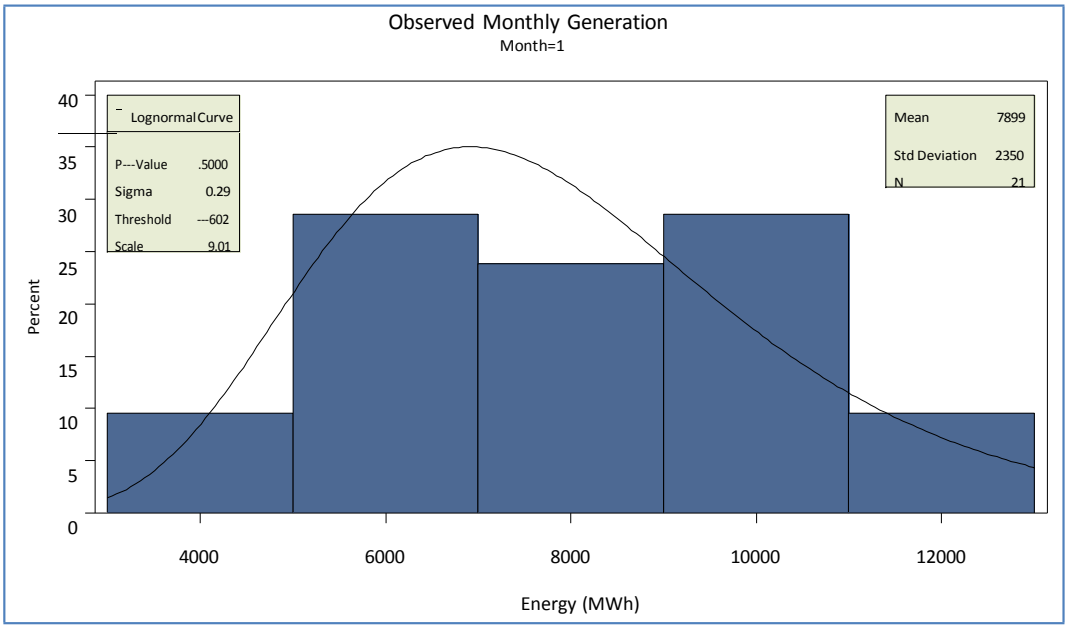


Figure 41. January Energy Generation - Observed Data

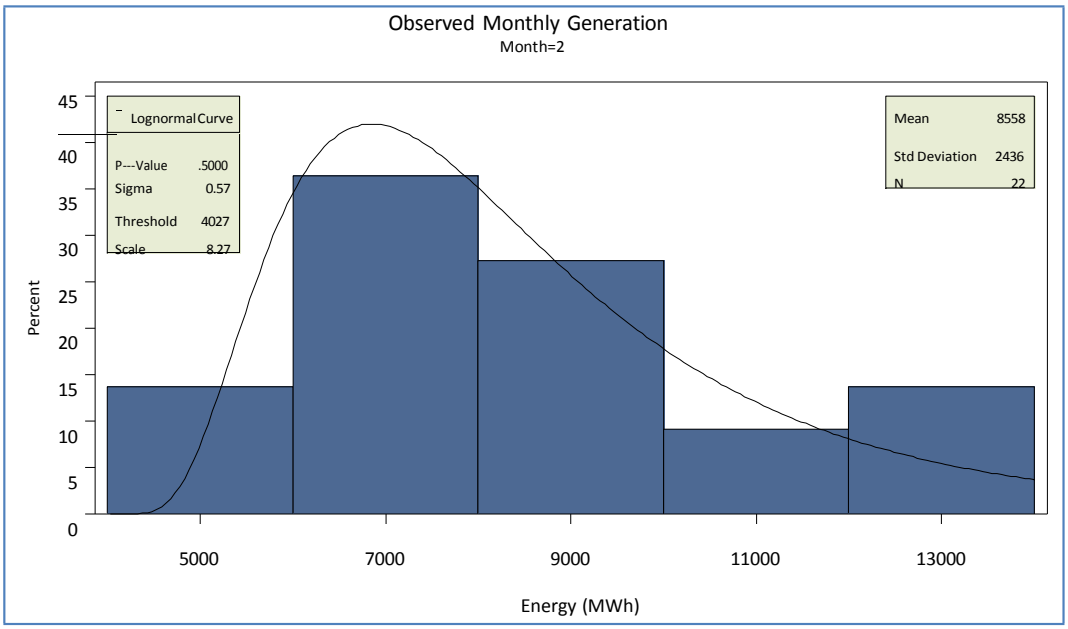


Figure 42. February Energy Generation - Observed Data

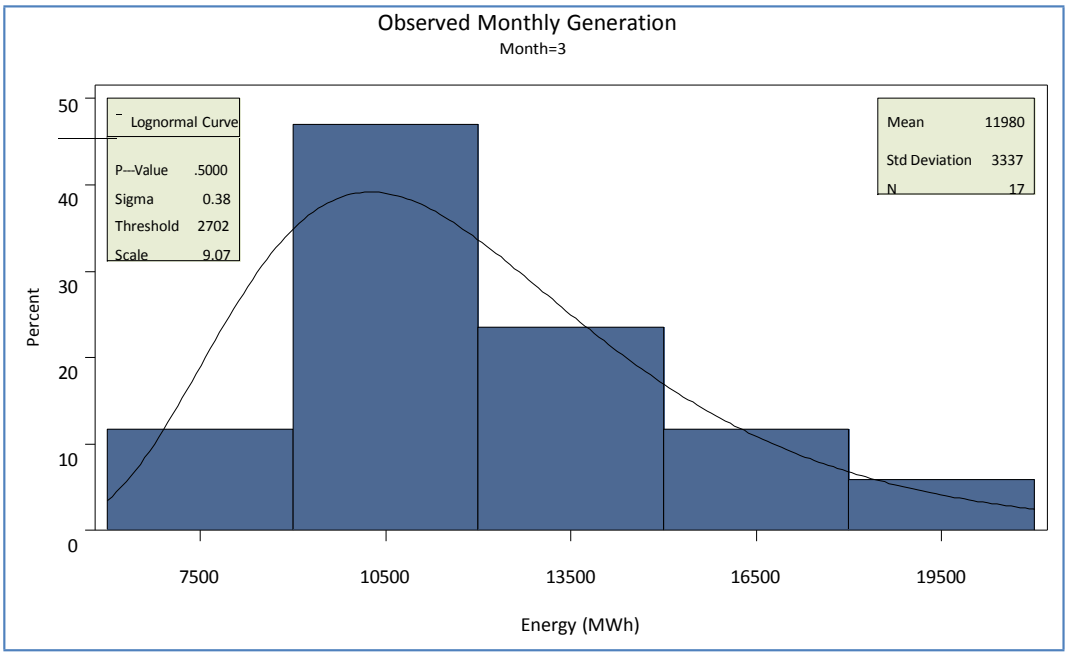


Figure 43. March Energy Generation - Observed Data

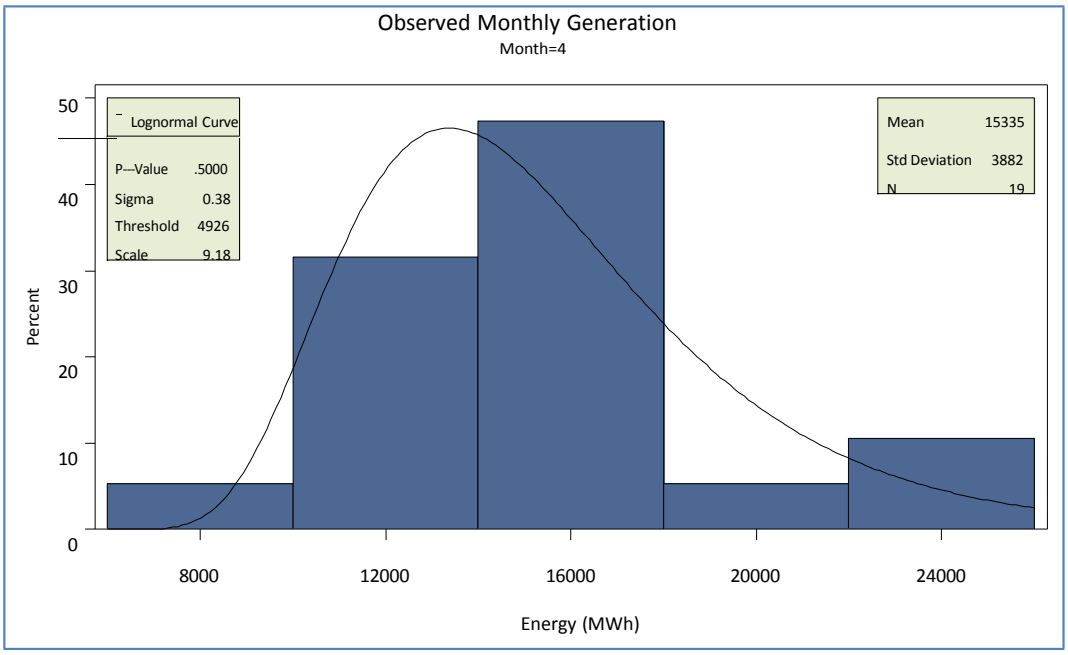


Figure 44. April Energy Generation - Observed Data

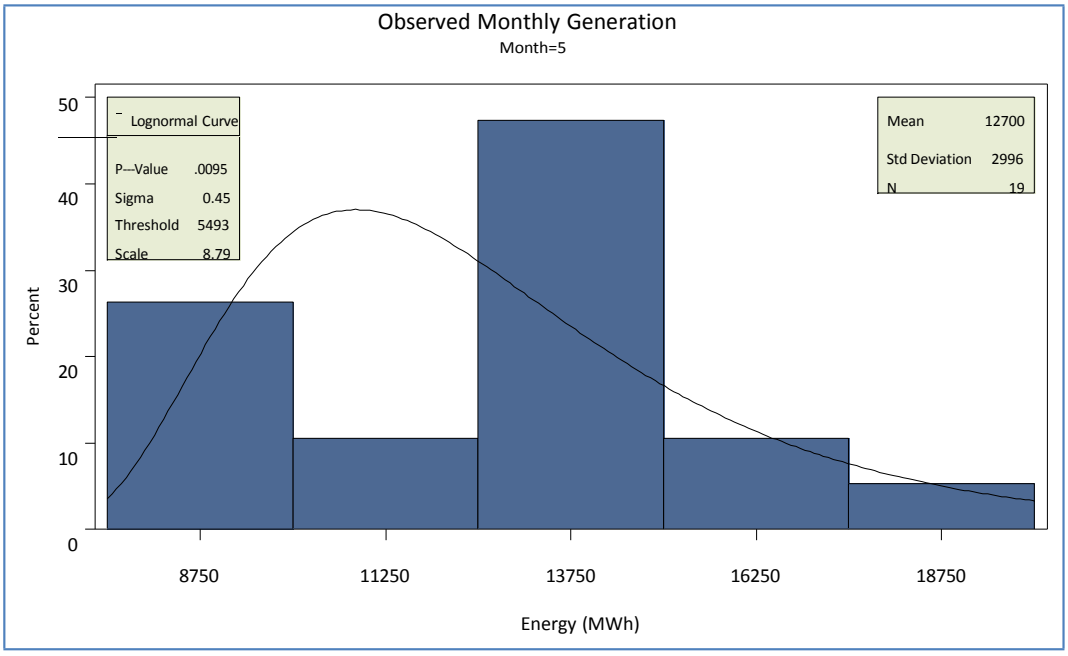


Figure 45. May Energy Generation - Observed Data

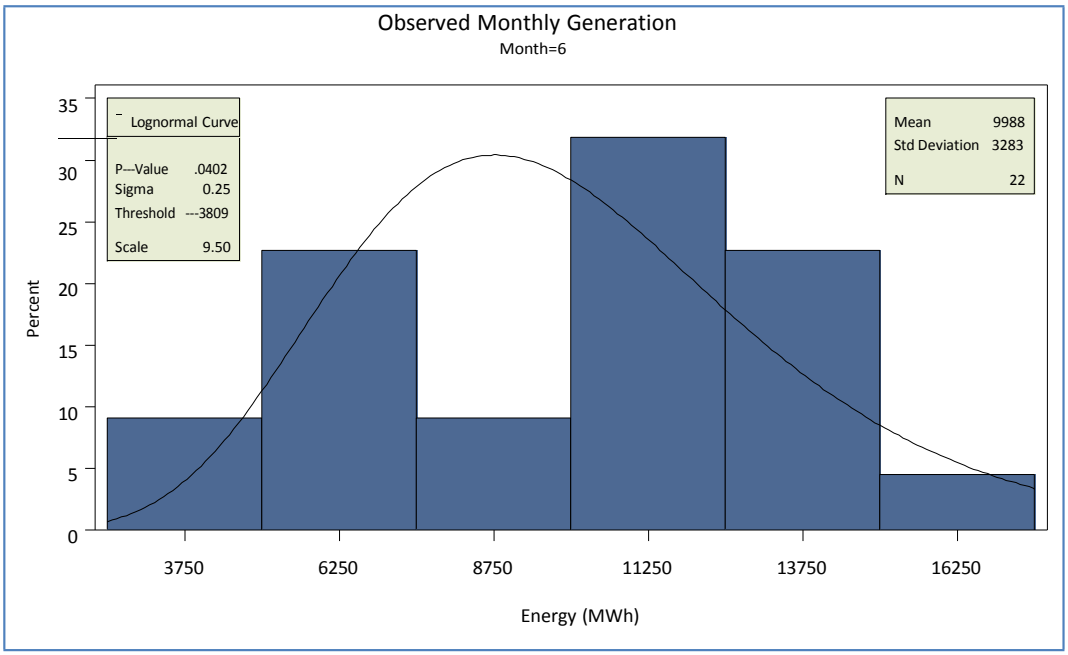


Figure 46. June Energy Generation - Observed Data

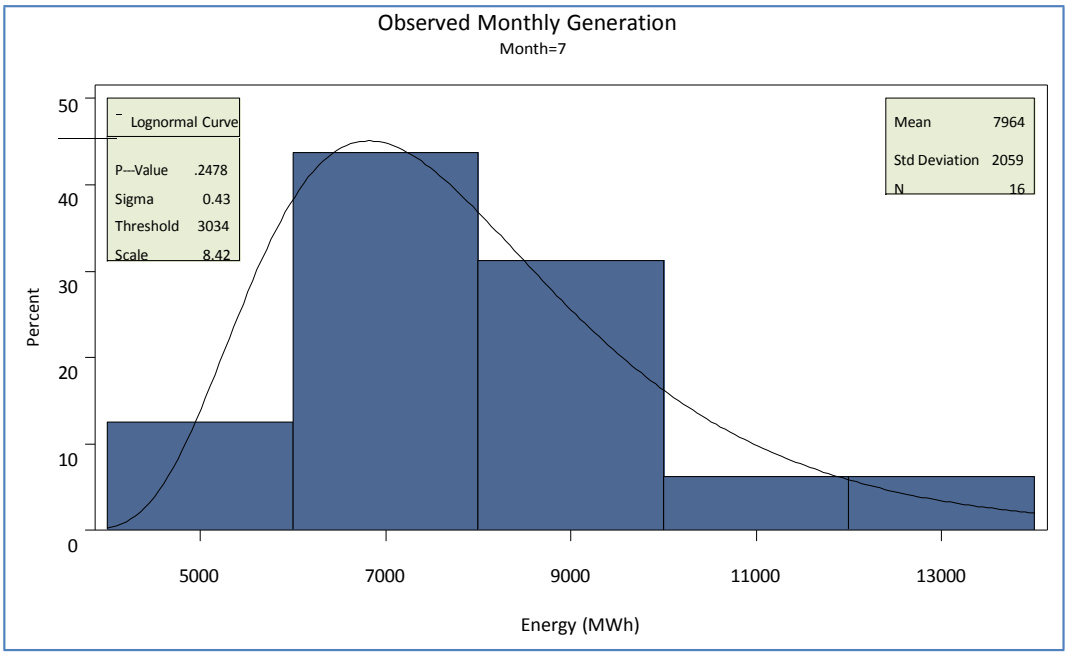


Figure 47. July Energy Generation - Observed Data

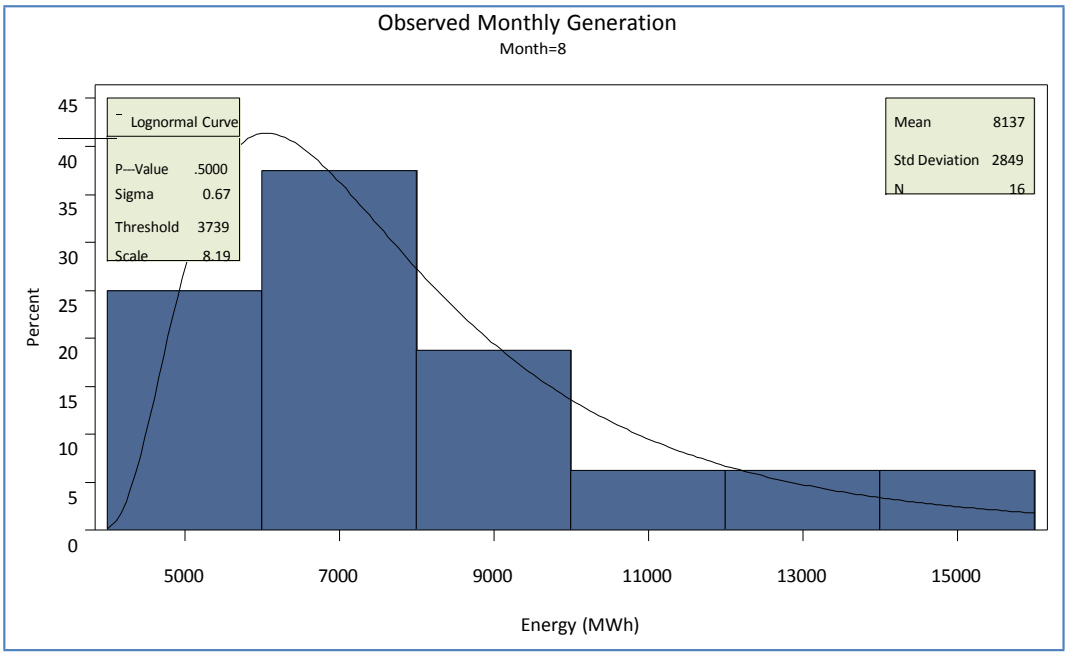


Figure 48. August Energy Generation - Observed Data

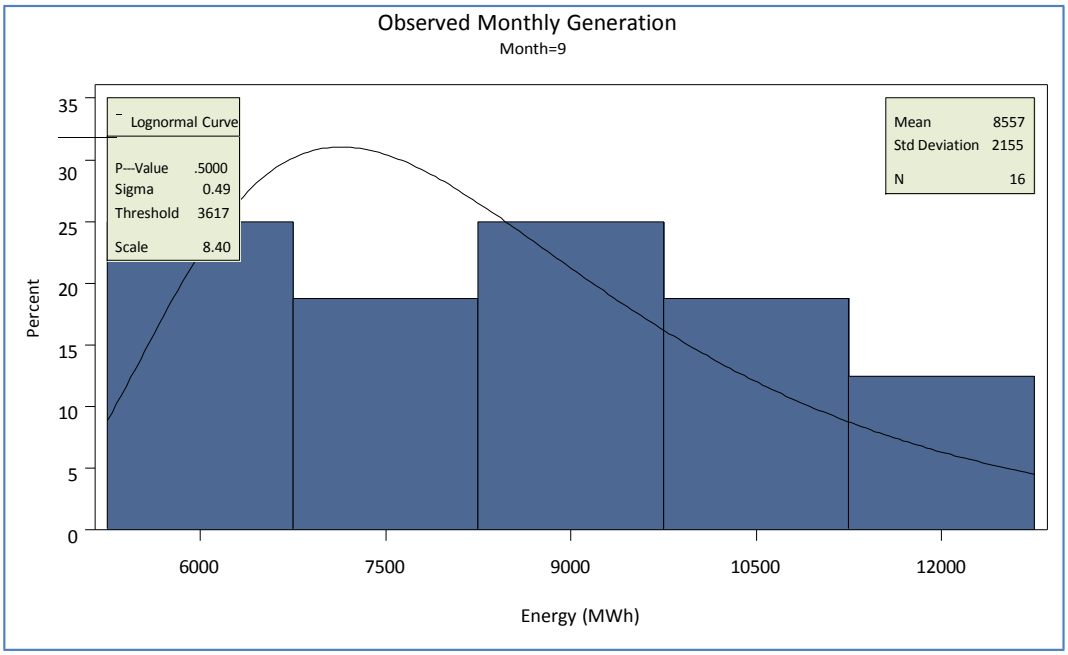


Figure 49. September Energy Generation - Observed Data

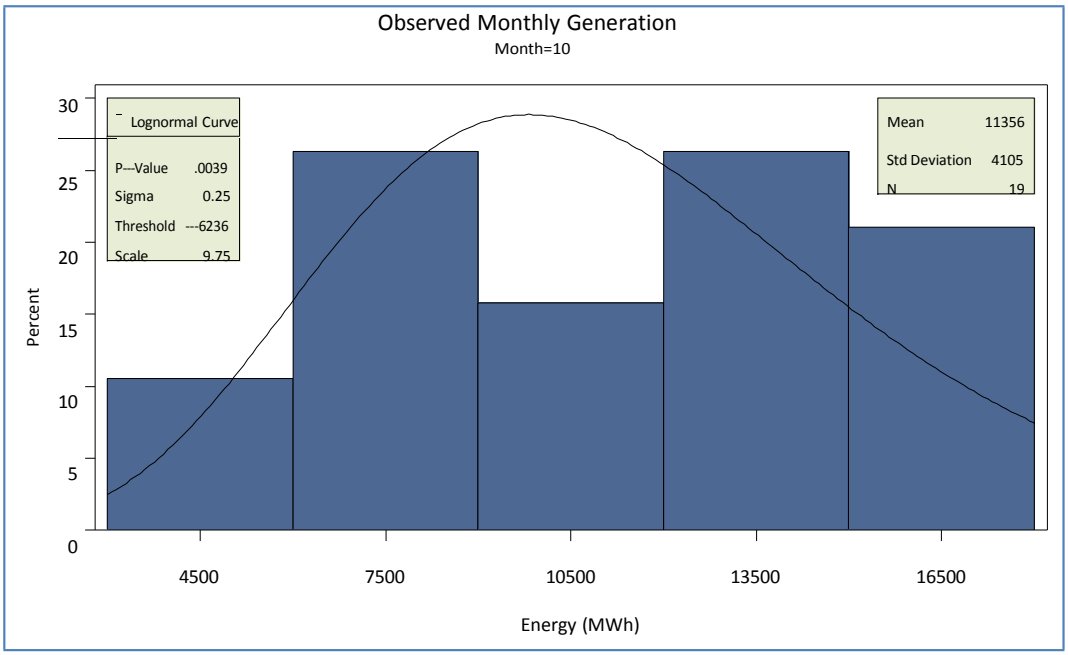


Figure 50. October Energy Generation - Observed Data

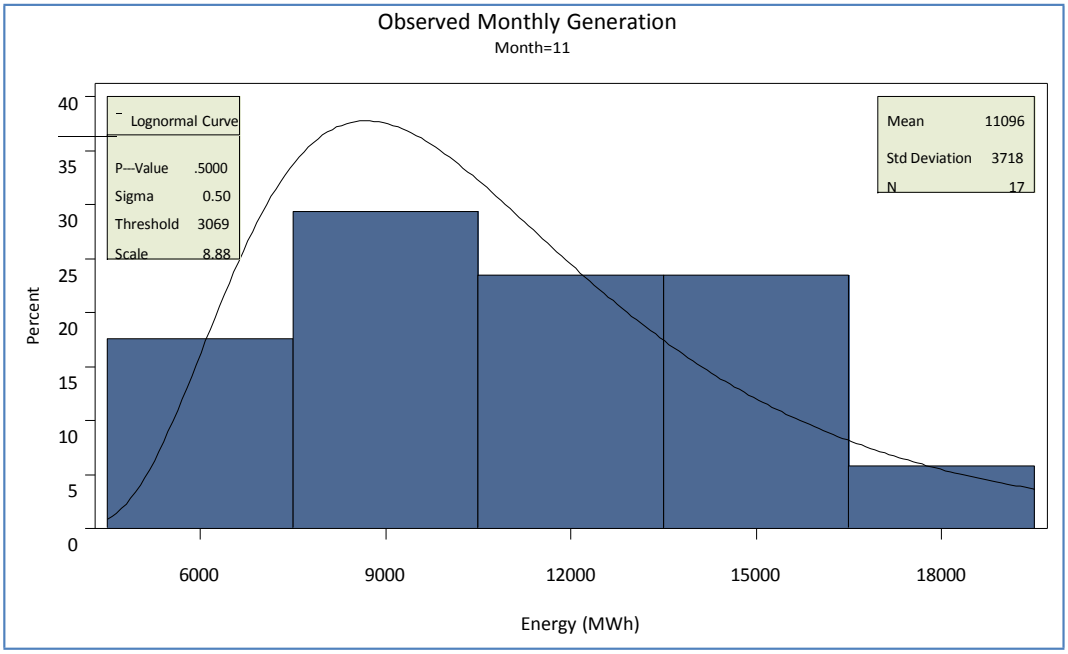


Figure 51. November Energy Generation - Observed Data

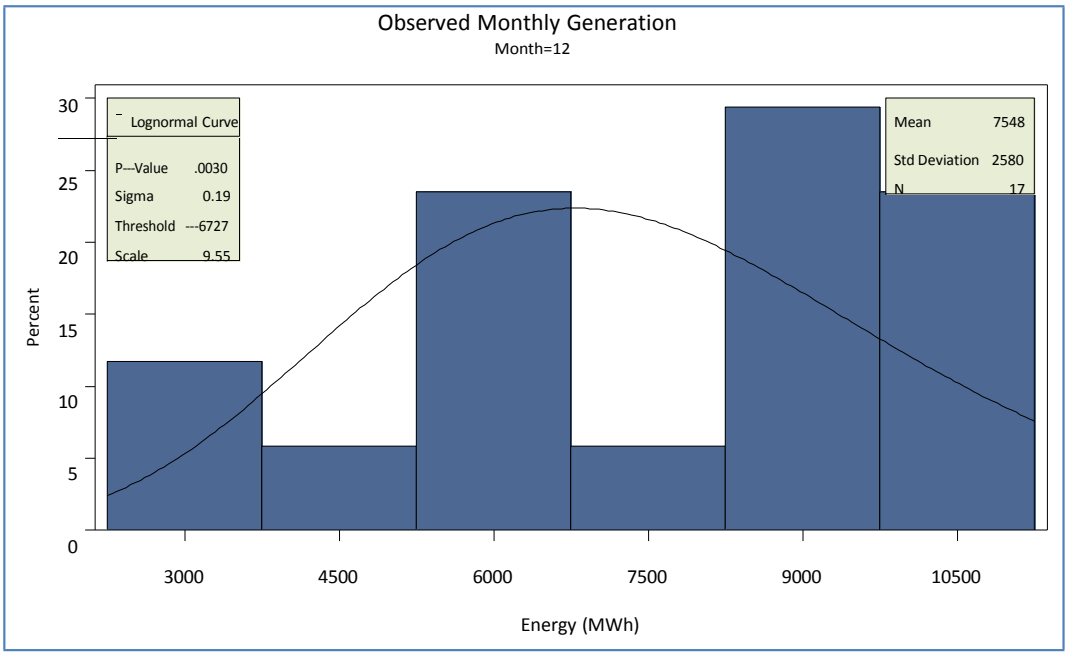


Figure 52. December Energy Generation - Observed Data

APPENDIX II. TABLES AND FIGURES – BOOTSTRAPPED DATA

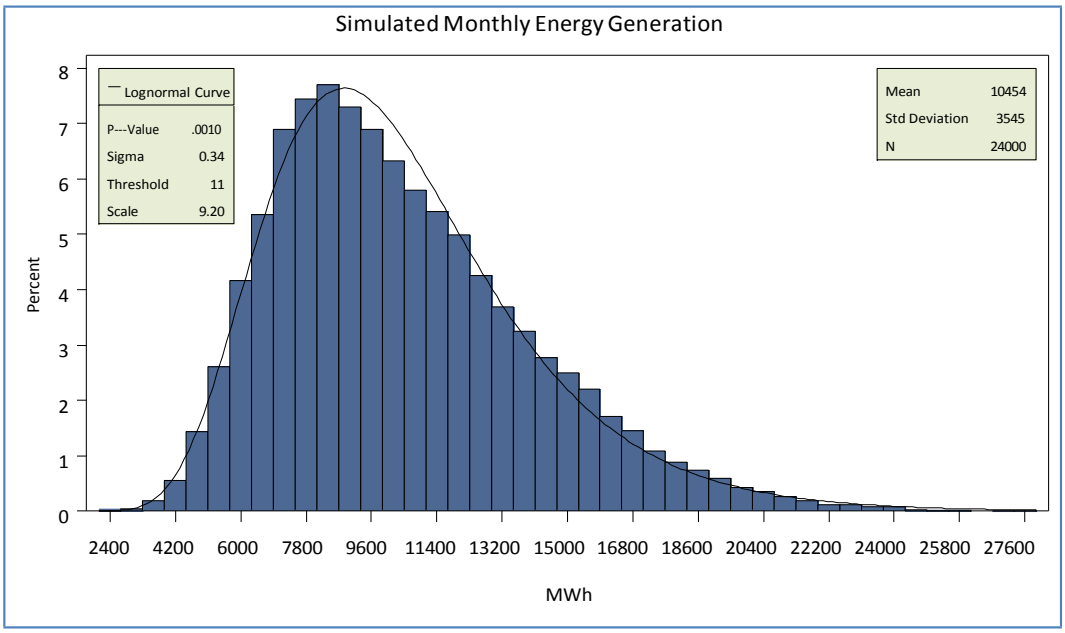


Figure 53. Monthly Energy Generation - All Months - Bootstrapped Data

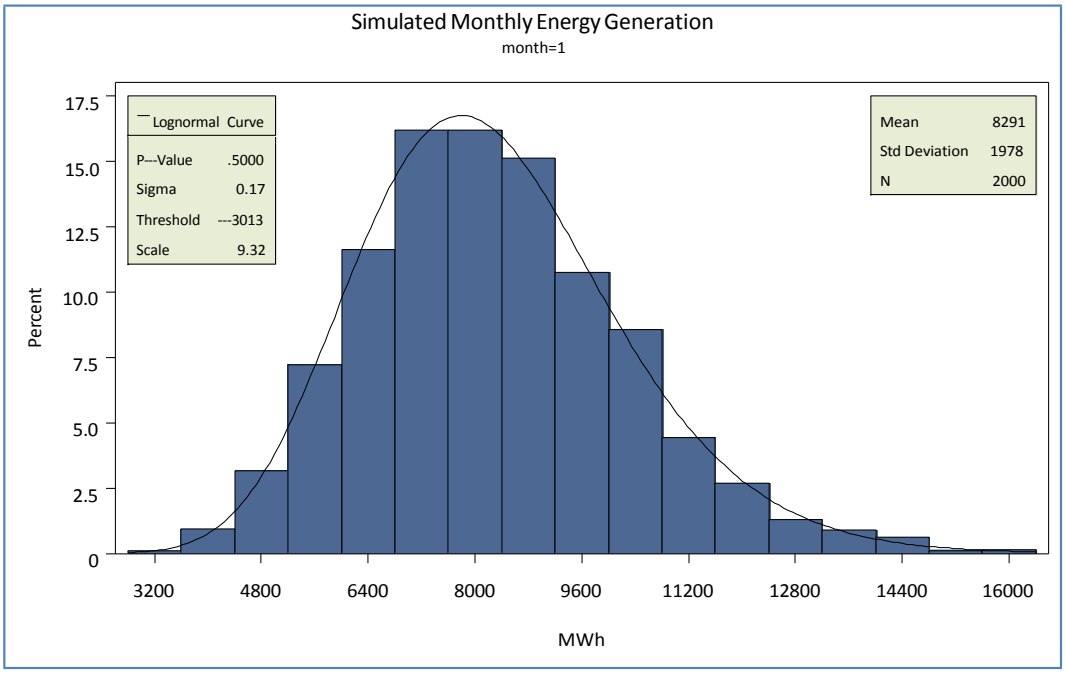


Figure 54. January Energy Generation - Bootstrapped Data

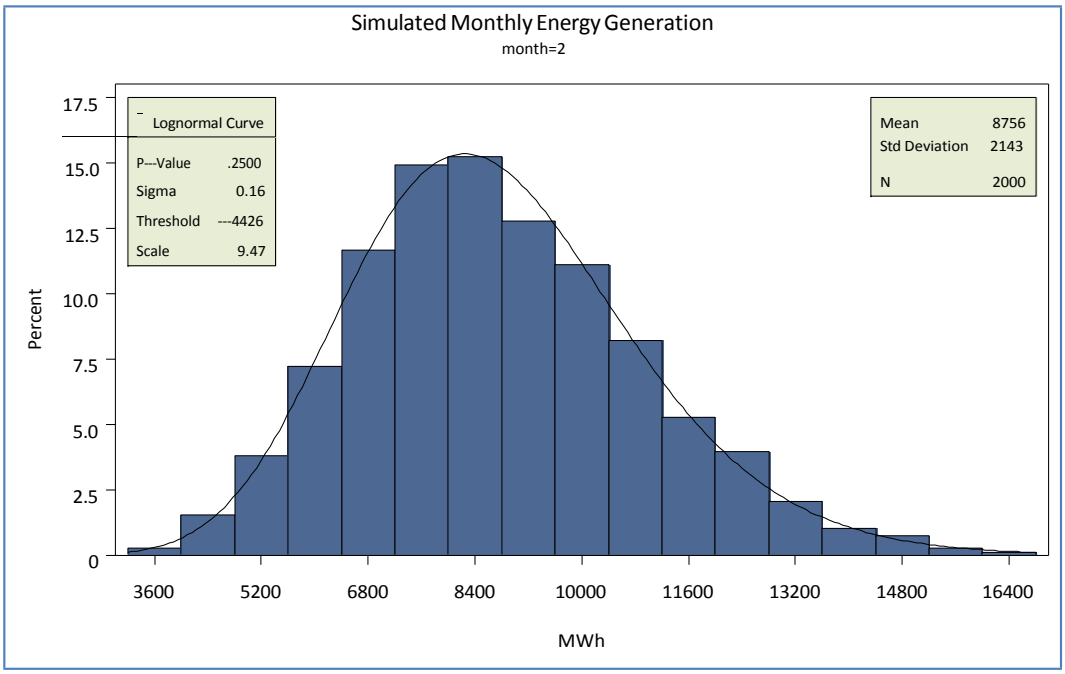


Figure 55. February Energy Generation - Bootstrapped Data

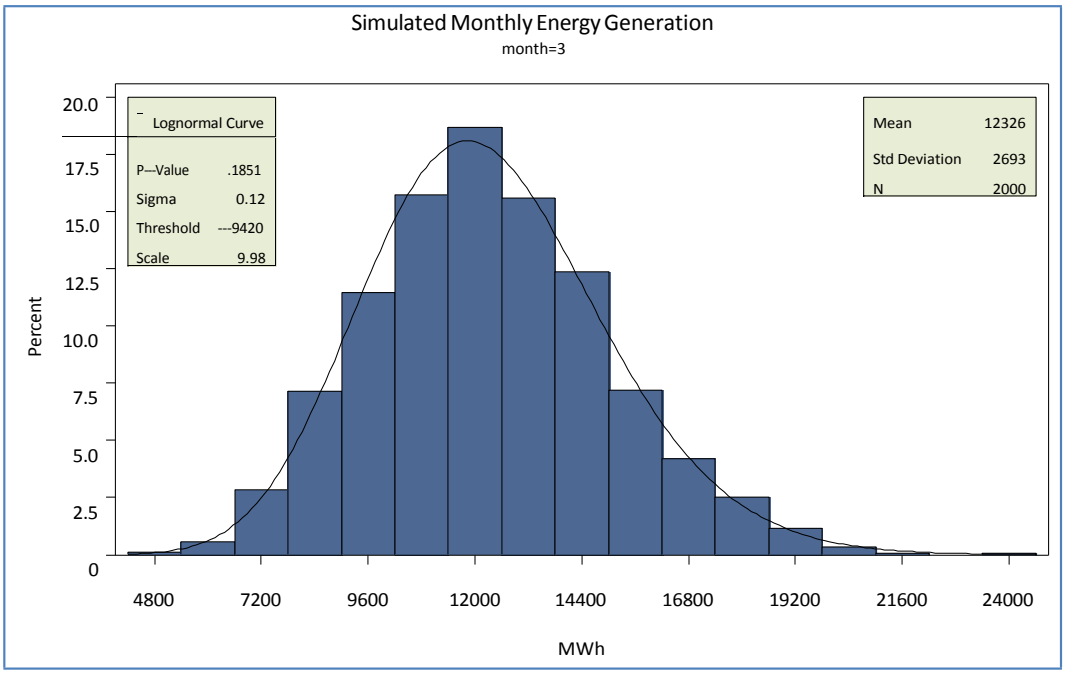


Figure 56. March Energy Generation - Bootstrapped Data

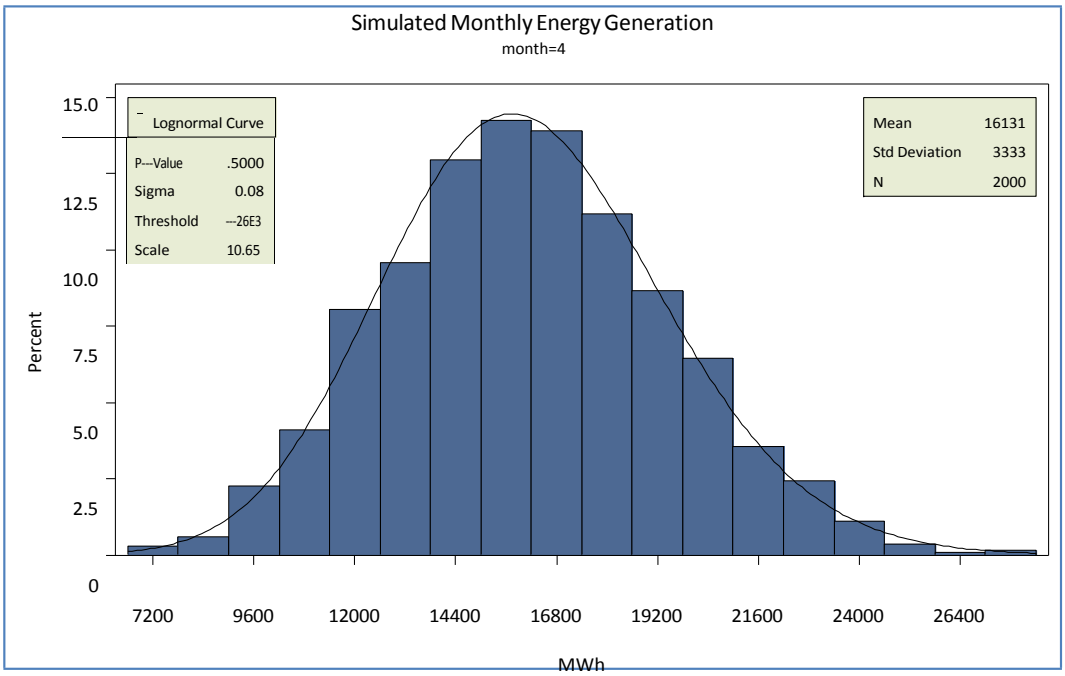


Figure 57. April Energy Generation - Bootstrapped Data

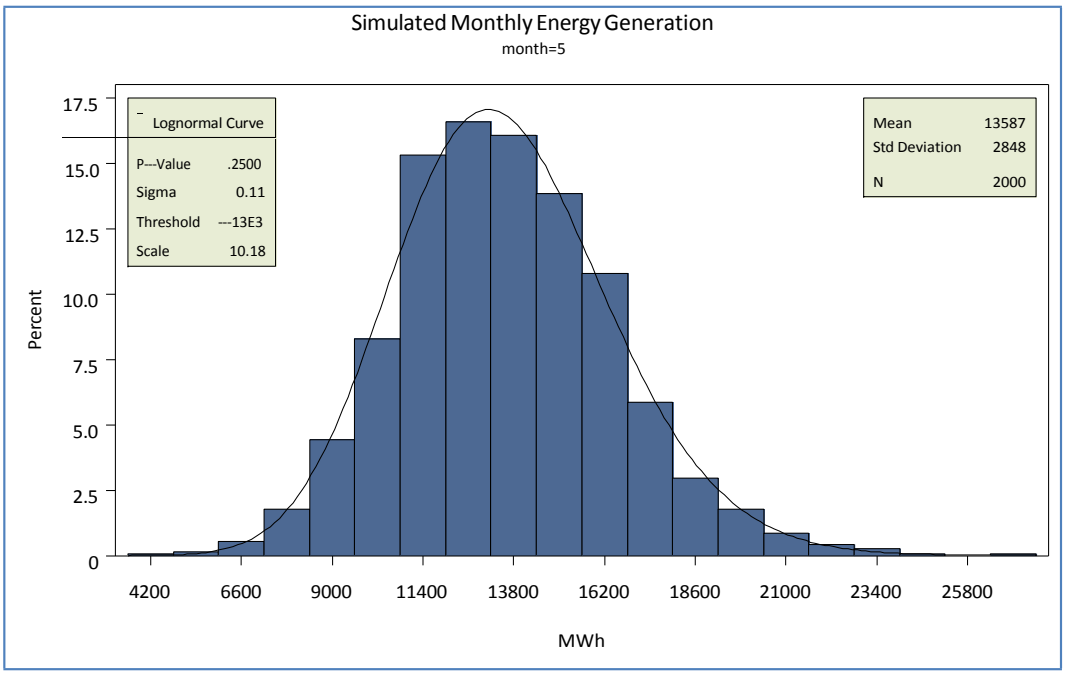


Figure 58. May Energy Generation - Bootstrapped Data

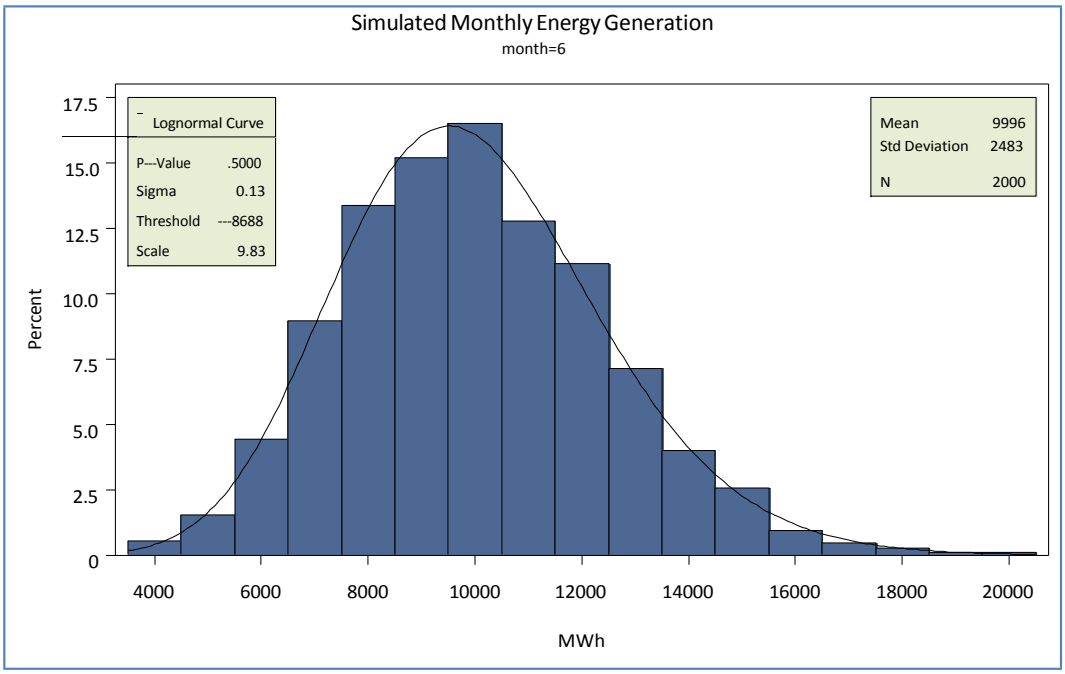


Figure 59. June Energy Generation - Bootstrapped Data

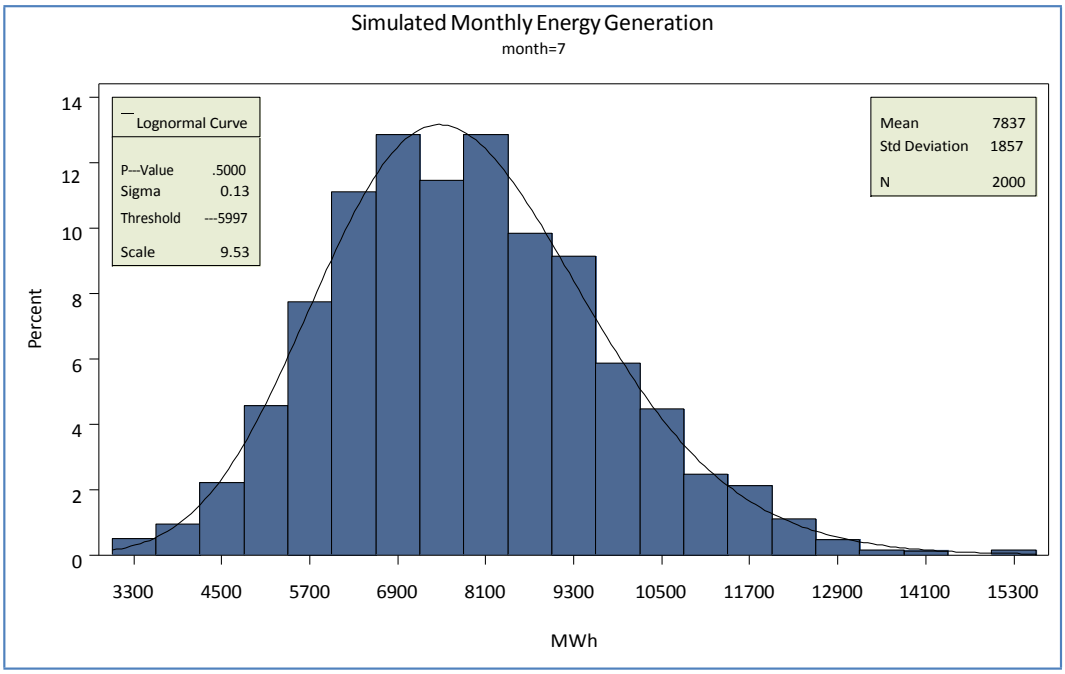


Figure 60. July Energy Generation - Bootstrapped Data

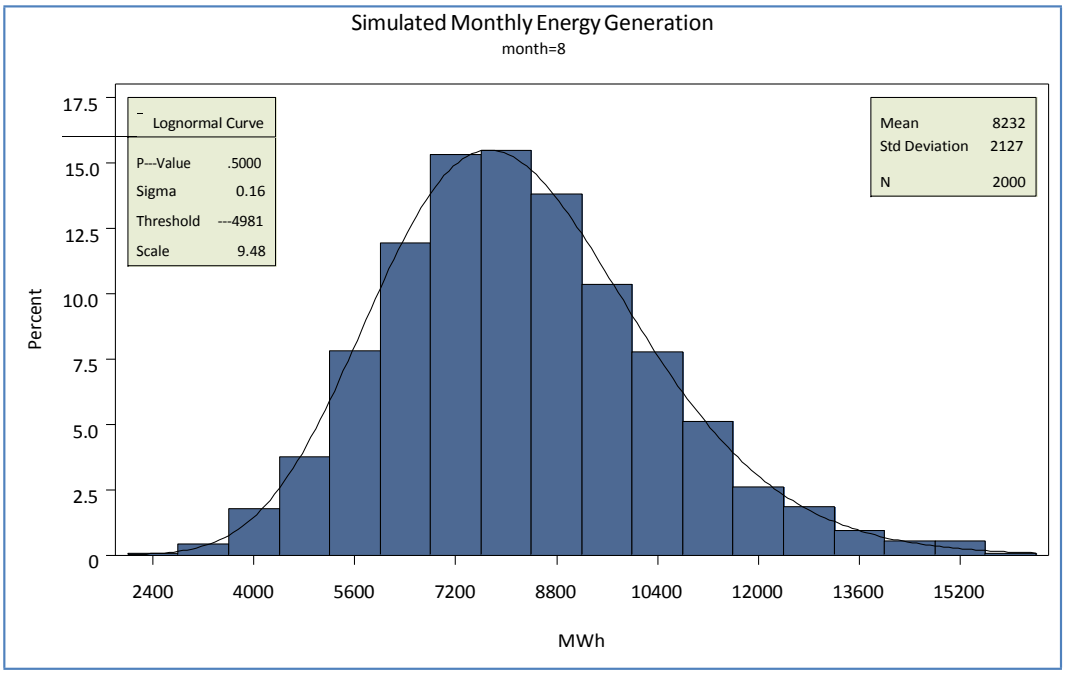


Figure 61. August Energy Generation - Bootstrapped Data

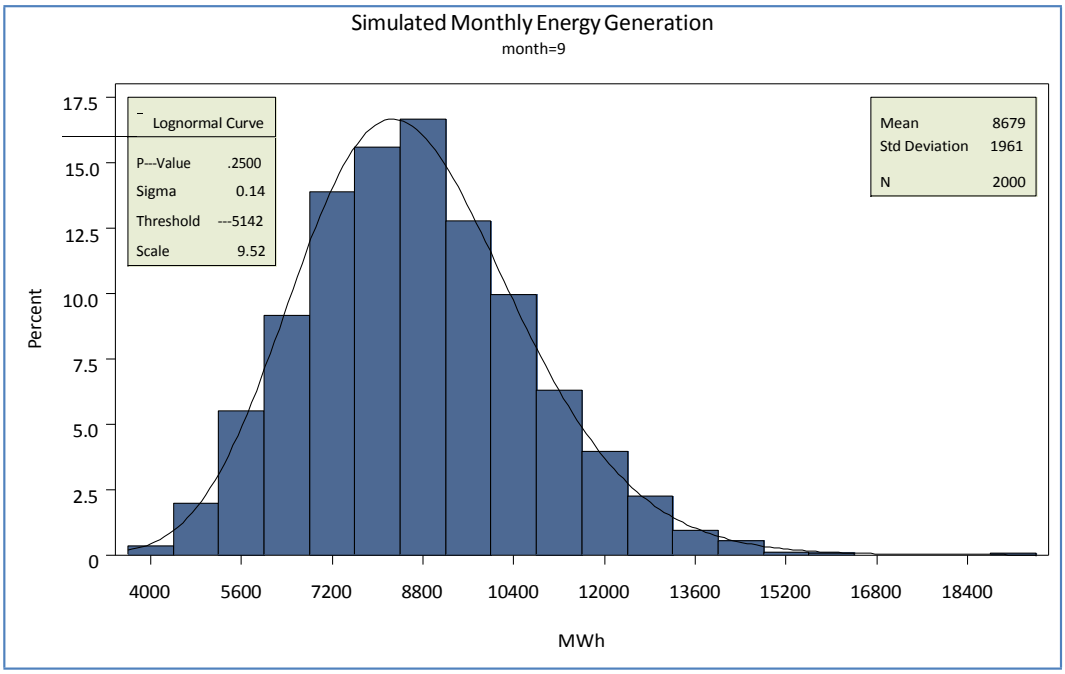


Figure 62. September Energy Generation - Bootstrapped Data

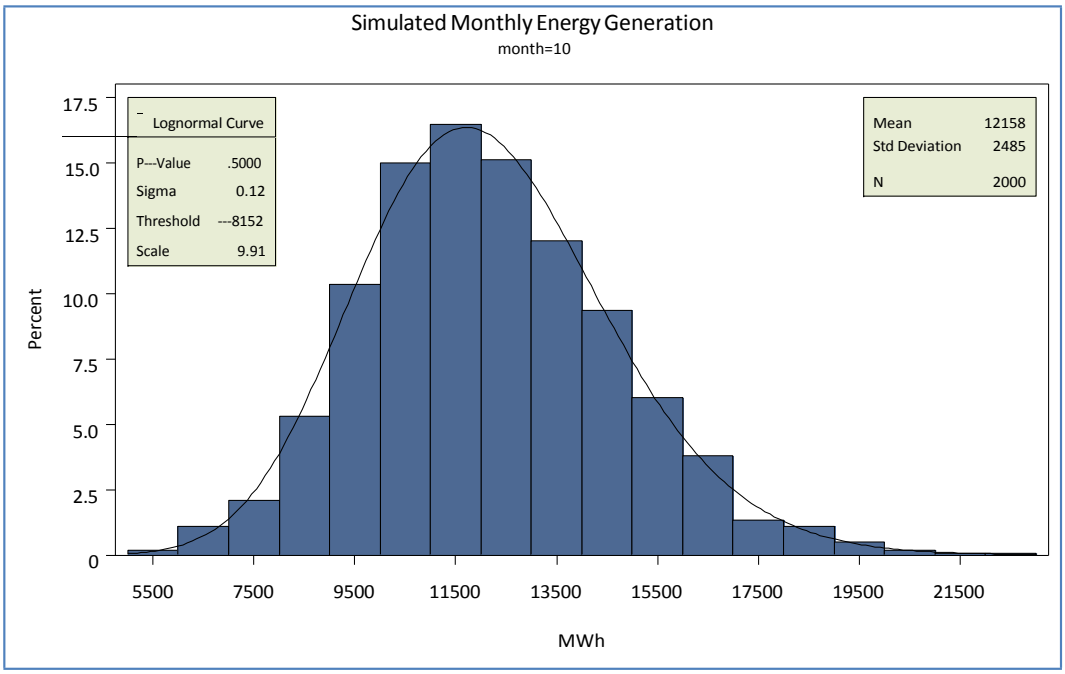


Figure 63. October Energy Generation - Bootstrapped Data

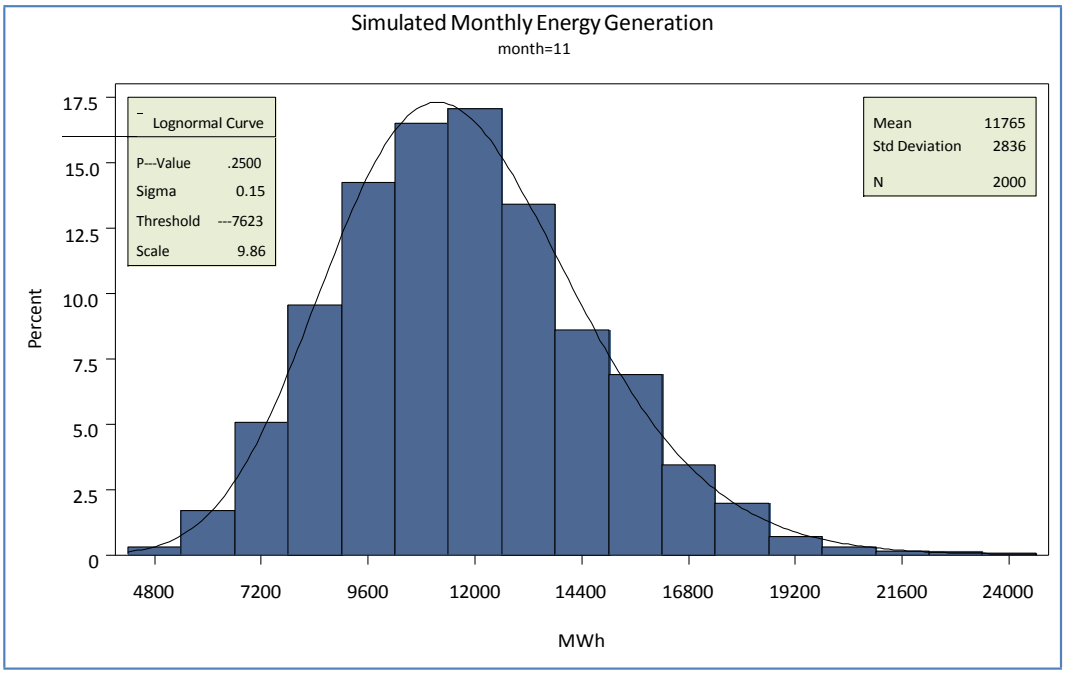


Figure 64. November Energy Generation - Bootstrapped Data

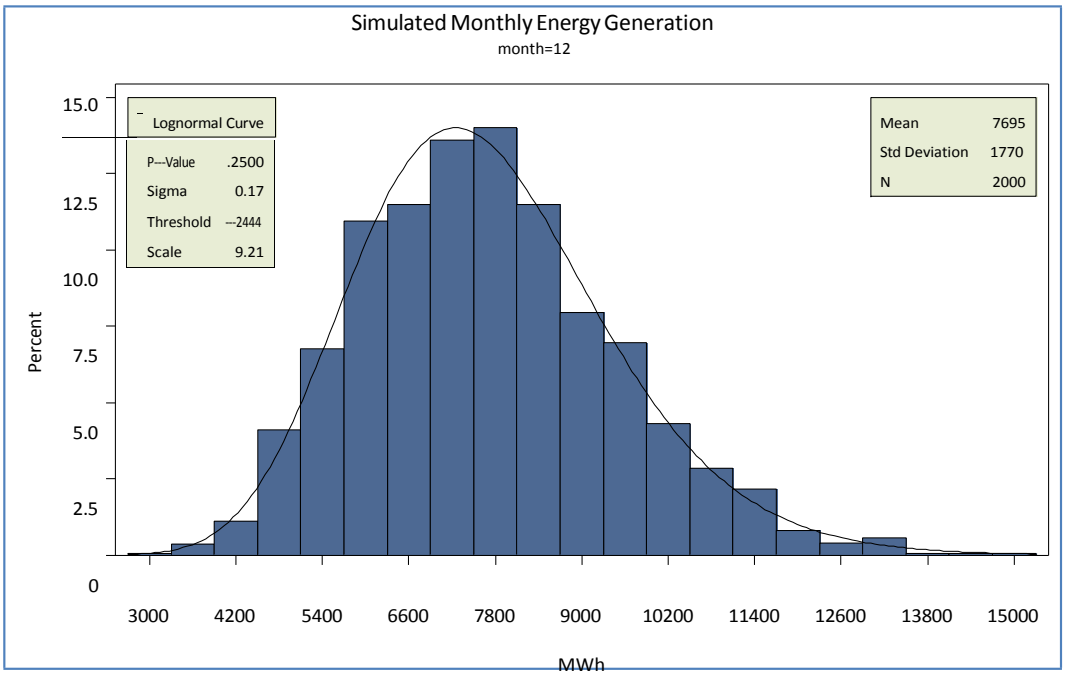


Figure 65. December Energy Generation - Bootstrapped Data

APPENDIX III. TABLES AND FIGURES - CAPACITY ANALYSIS

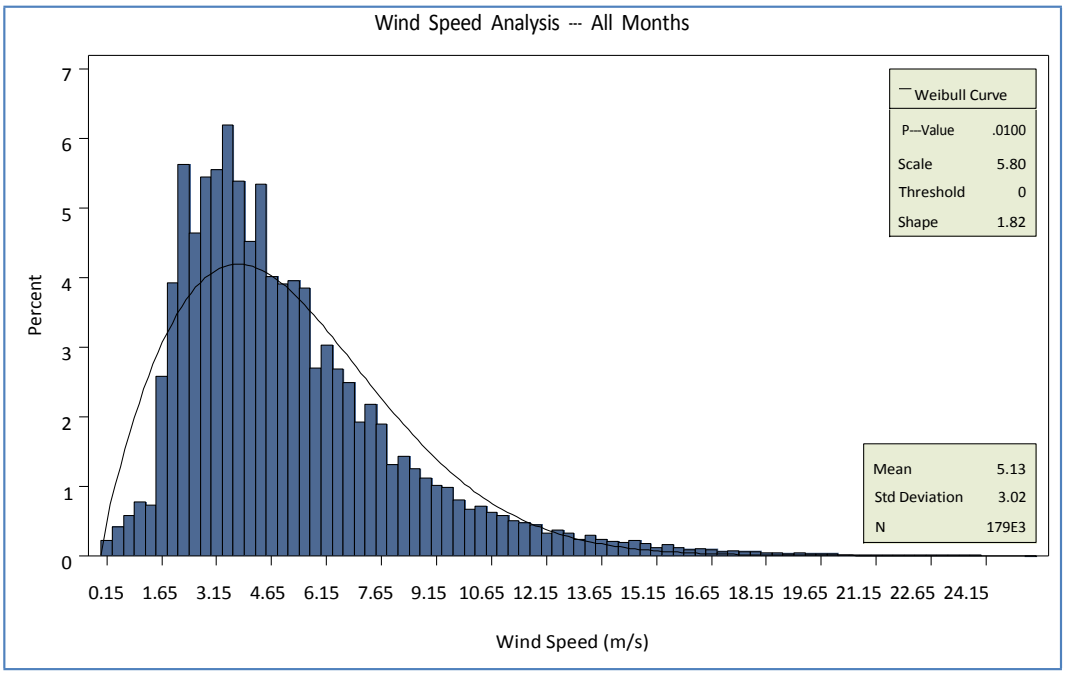


Figure 66. Wind Speed Analysis - All Months – Weibull

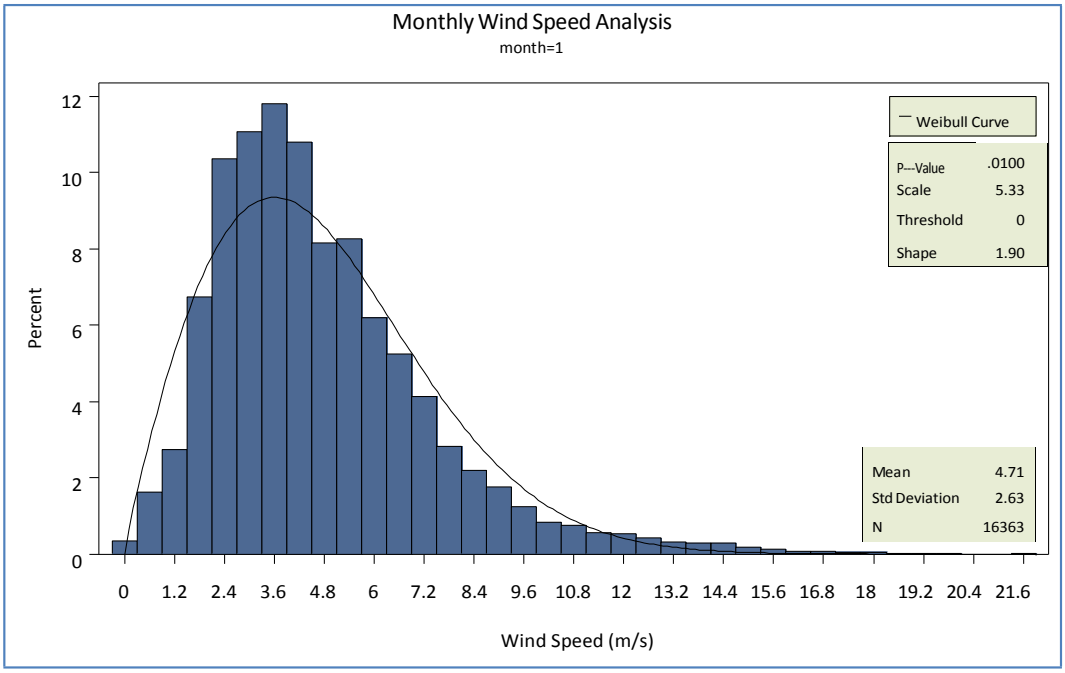


Figure 67. January Wind Speed Analysis - Weibull Distribution

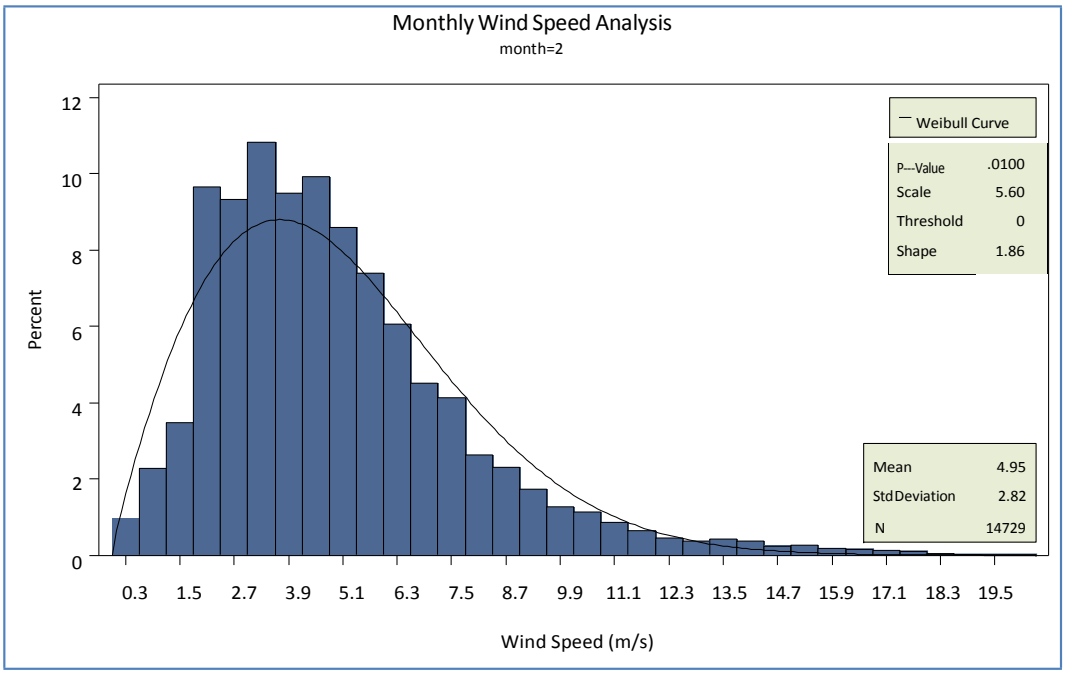


Figure 68. February Wind Speed Analysis - Weibull Distribution

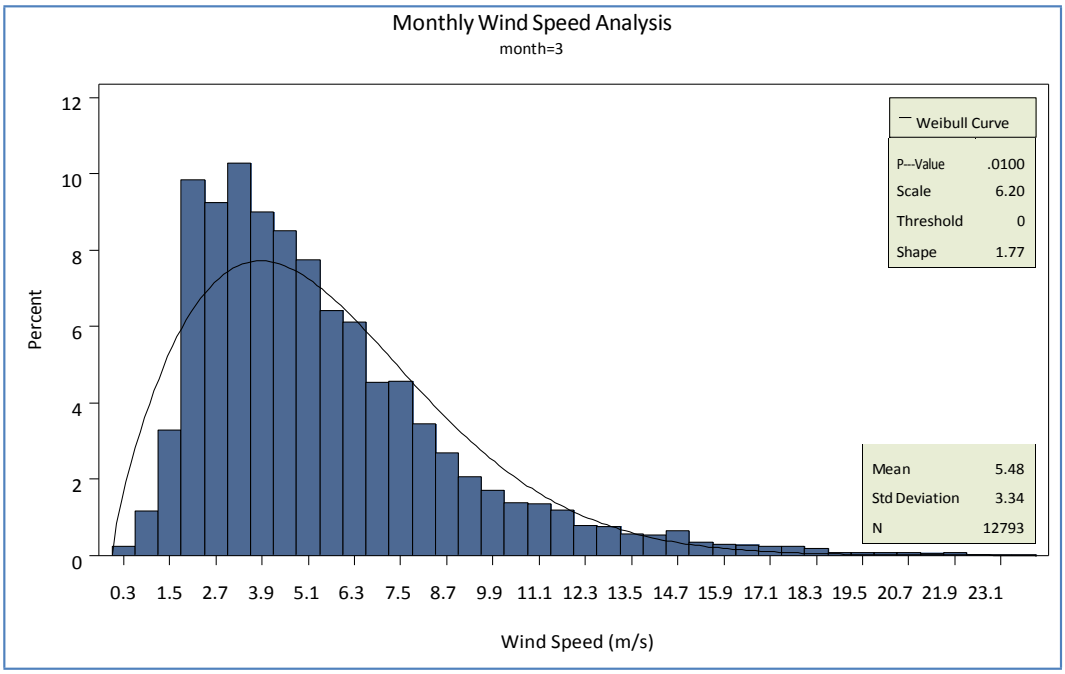


Figure 69. March Wind Speed Analysis - Weibull Distribution

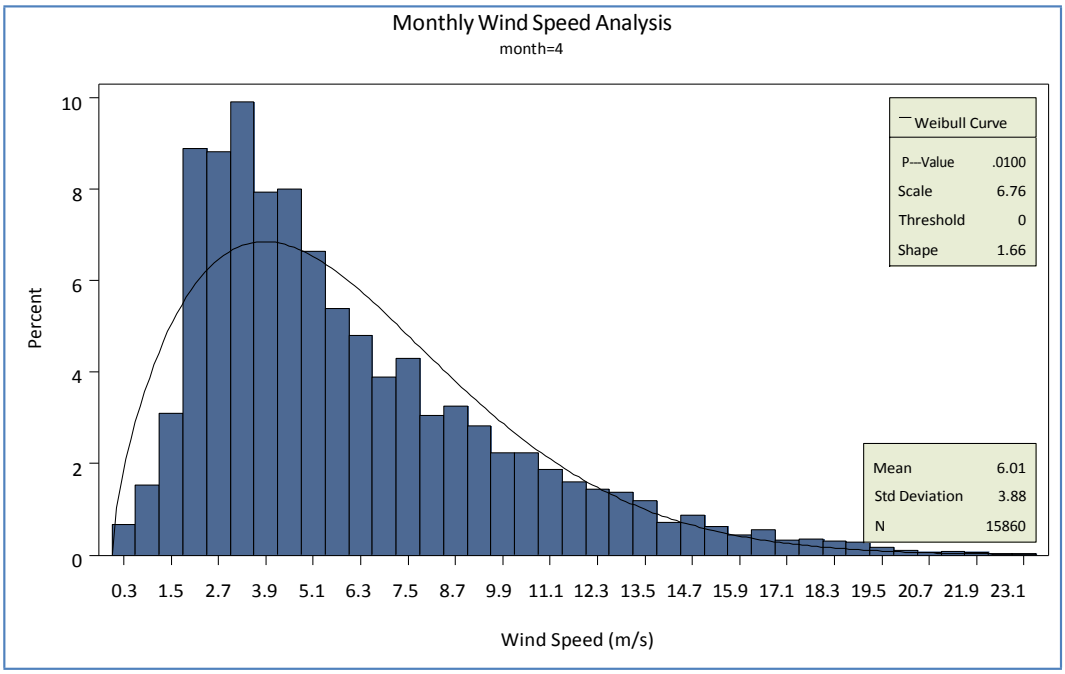


Figure 70. April Wind Speed Analysis - Weibull Distribution

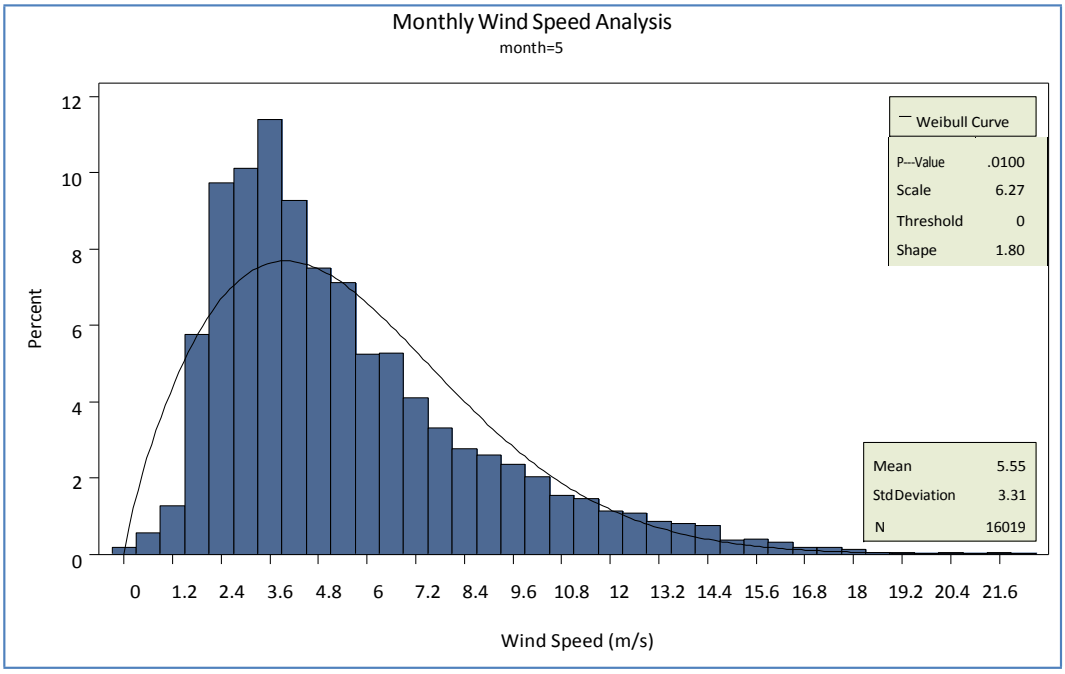


Figure 71. May Wind Speed Analysis - Weibull Distribution

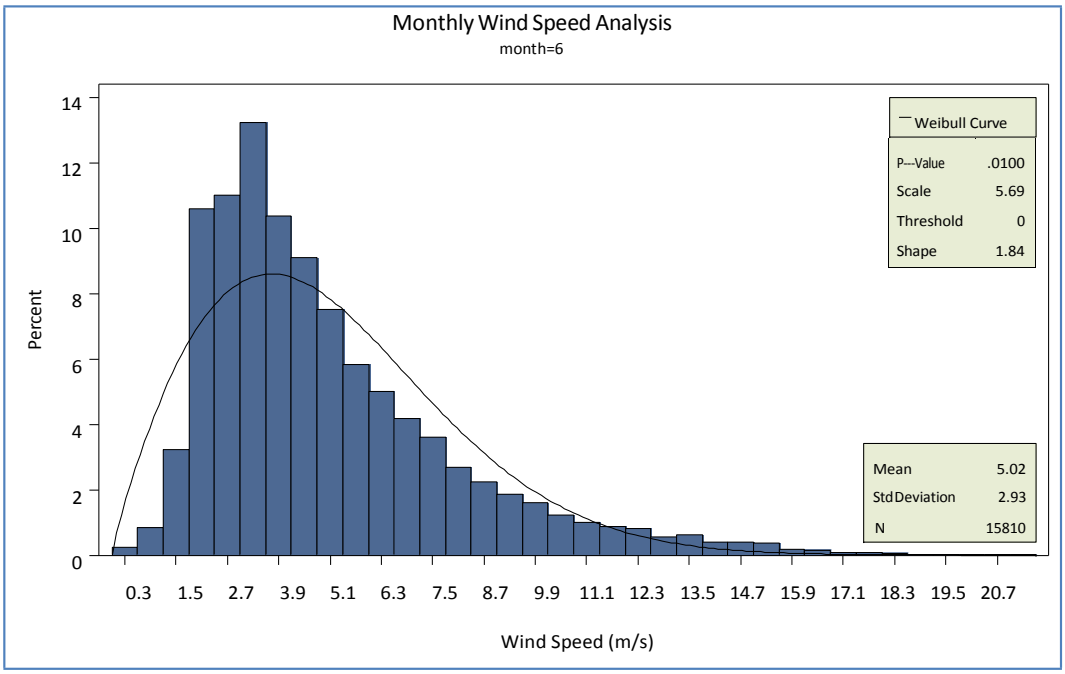


Figure 72. June Wind Speed Analysis - Weibull Distribution

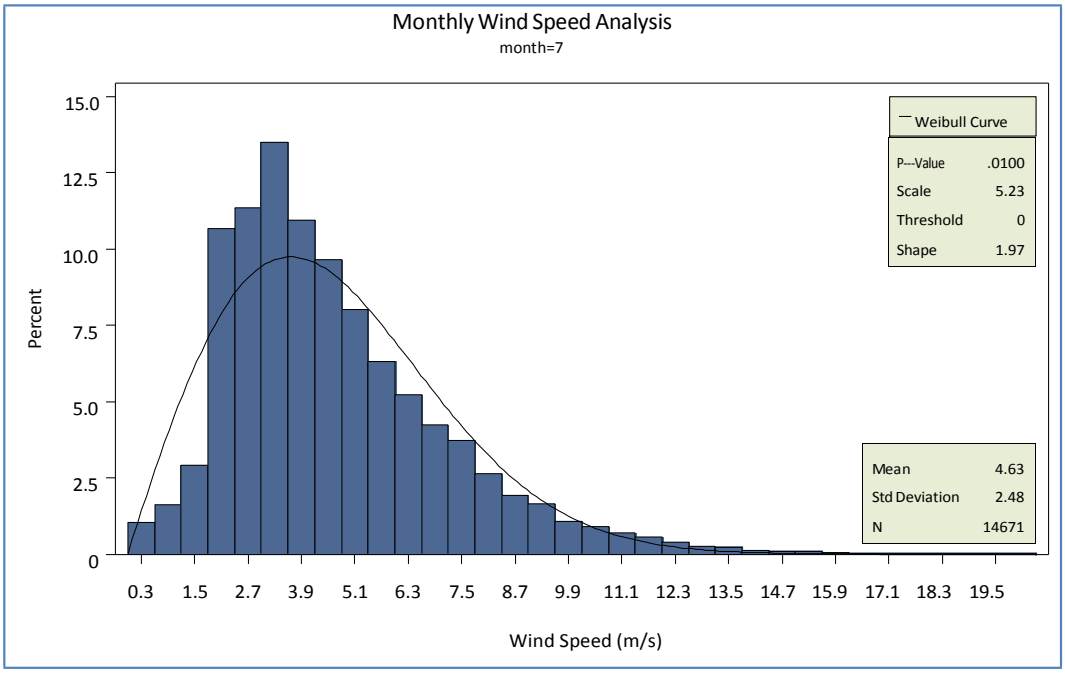


Figure 73. July Wind Speed Analysis - Weibull Distribution

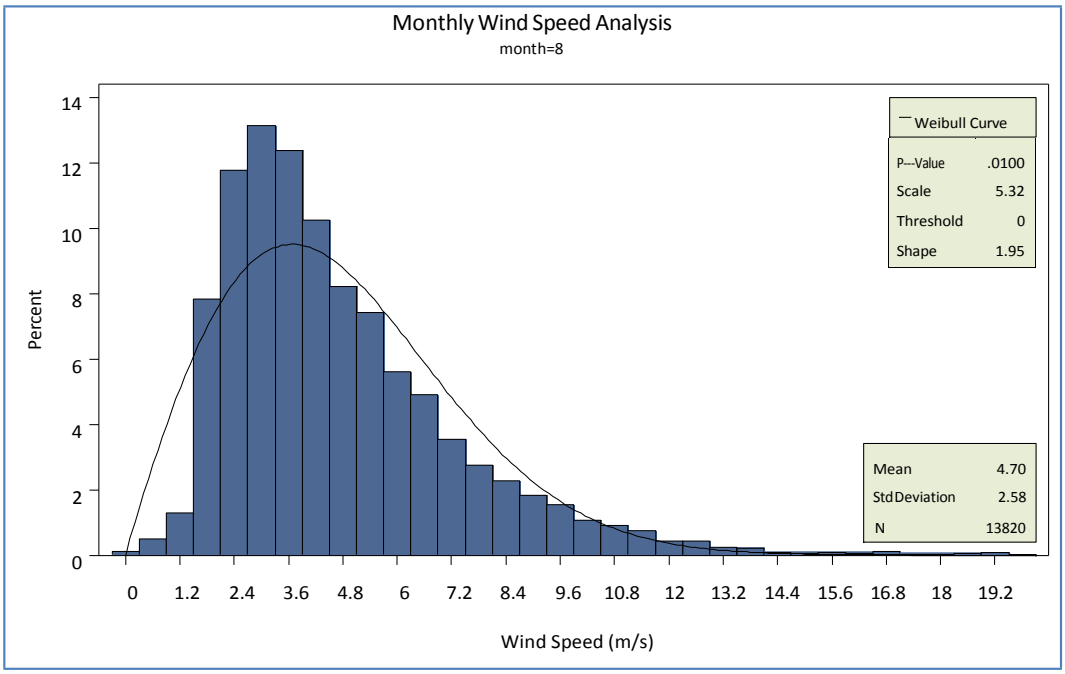


Figure 74. August Wind Speed Analysis - Weibull Distribution

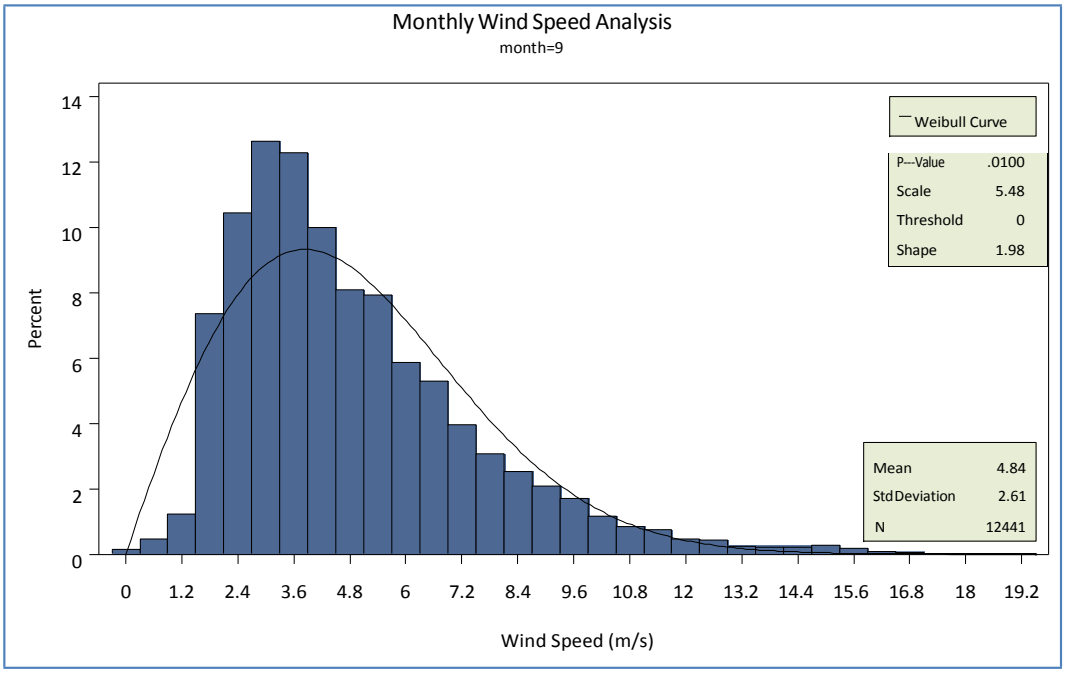


Figure 75. September Wind Speed Analysis - Weibull Distribution

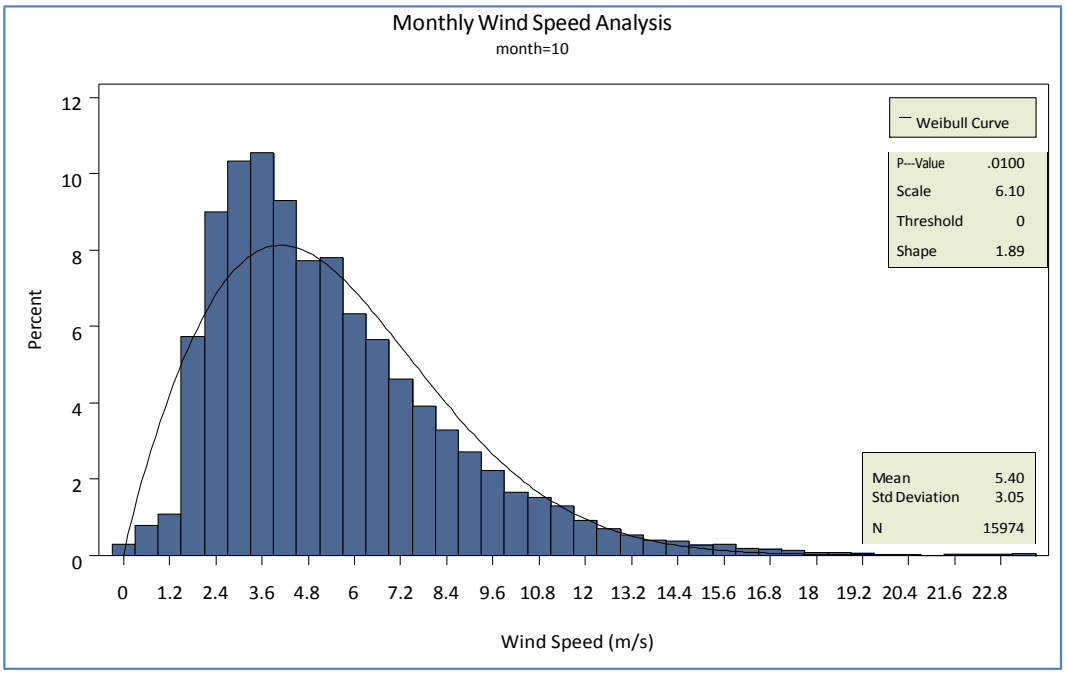


Figure 76. October Wind Speed Analysis - Weibull Distribution

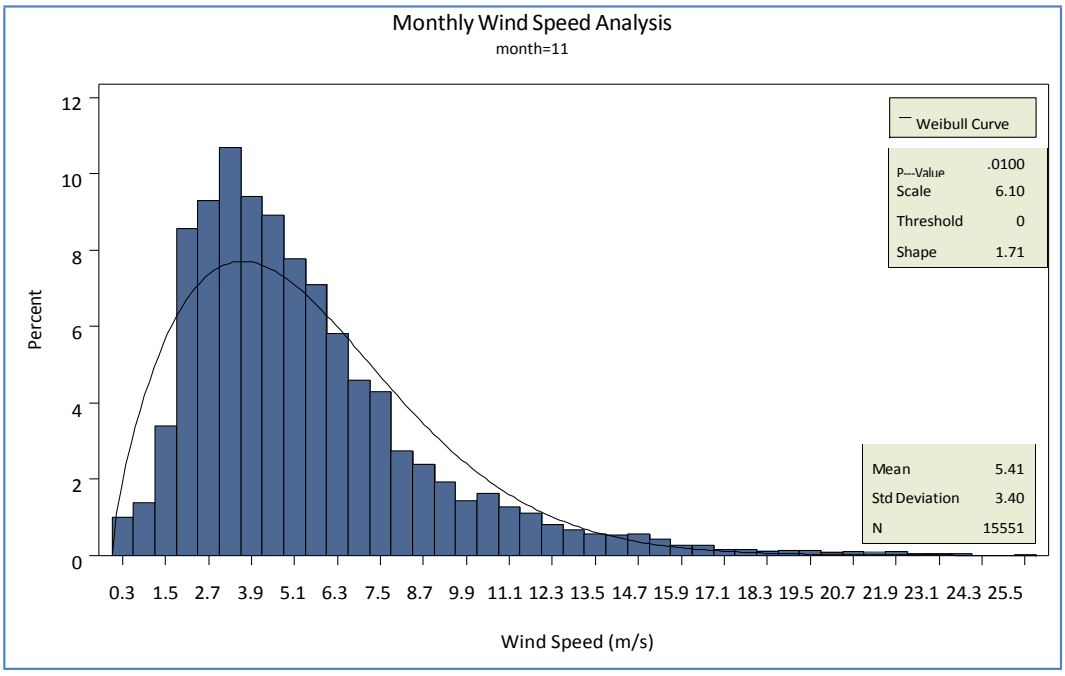


Figure 77. November Wind Speed Analysis - Weibull Distribution

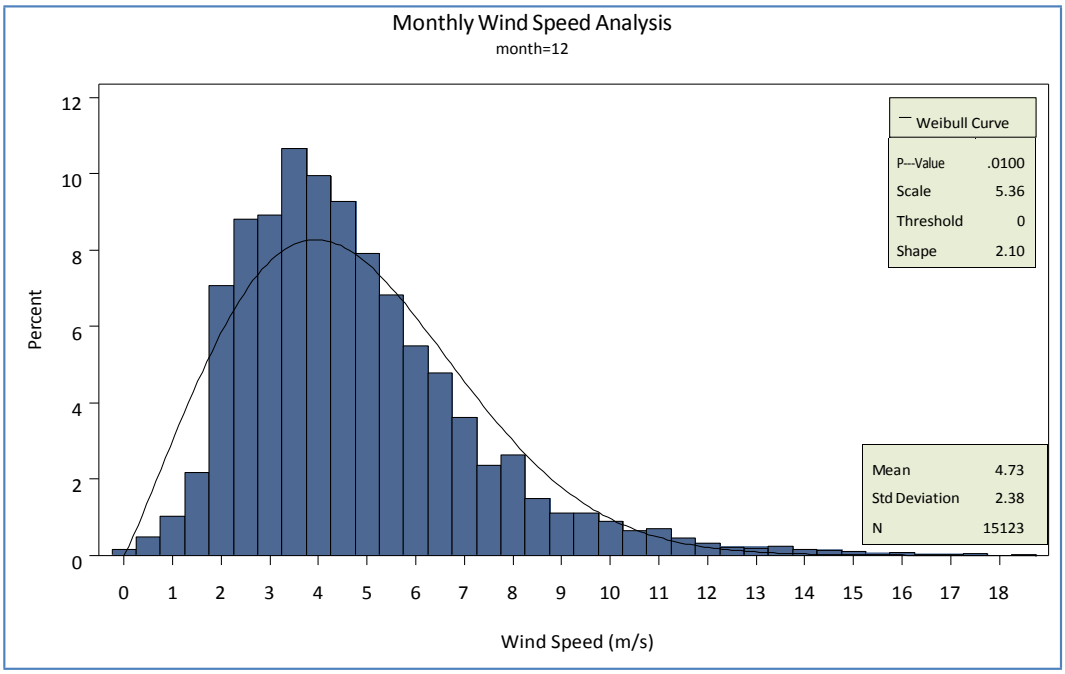


Figure 78. December Wind Speed Analysis - Weibull Distribution

APPENDIX IV. OBSERVED MONTHLY GENERATION OVER TIME

Table 25. Years and their assigned values used in Figure 79 through Figure 90

Assigned Value for Regression Analysis	Year
1	1983
2	1984
3	1985
4	1986
5	1987
6	1988
7	1989
8	1990
9	1991
10	1992
11	1993
12	1994
13	1995
14	1996
15	1997
16	1998
17	1999
18	2000
19	2001
20	2002
21	2003
22	2004
23	2005
24	2006
25	2007
26	2008

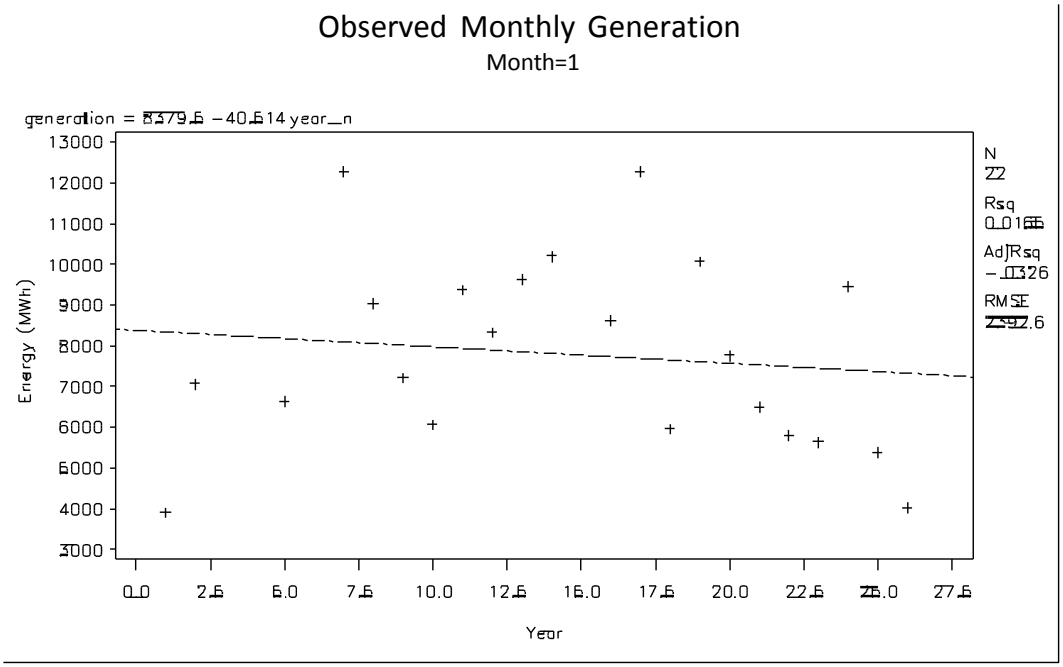


Figure 79. January Energy Generation Over Time

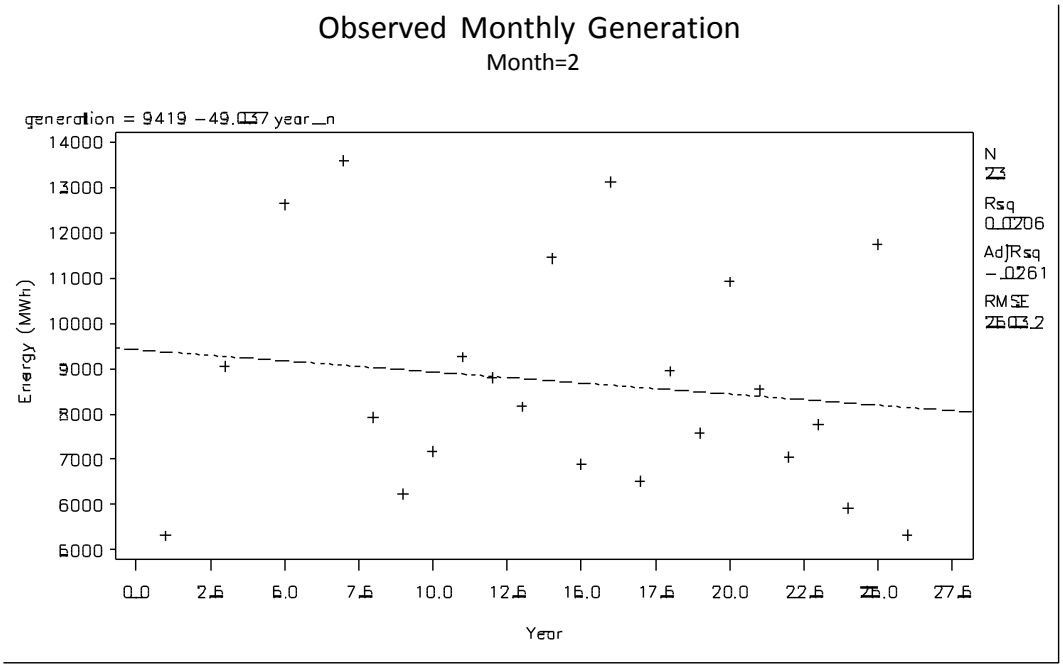


Figure 80. February Energy Generation Over Time

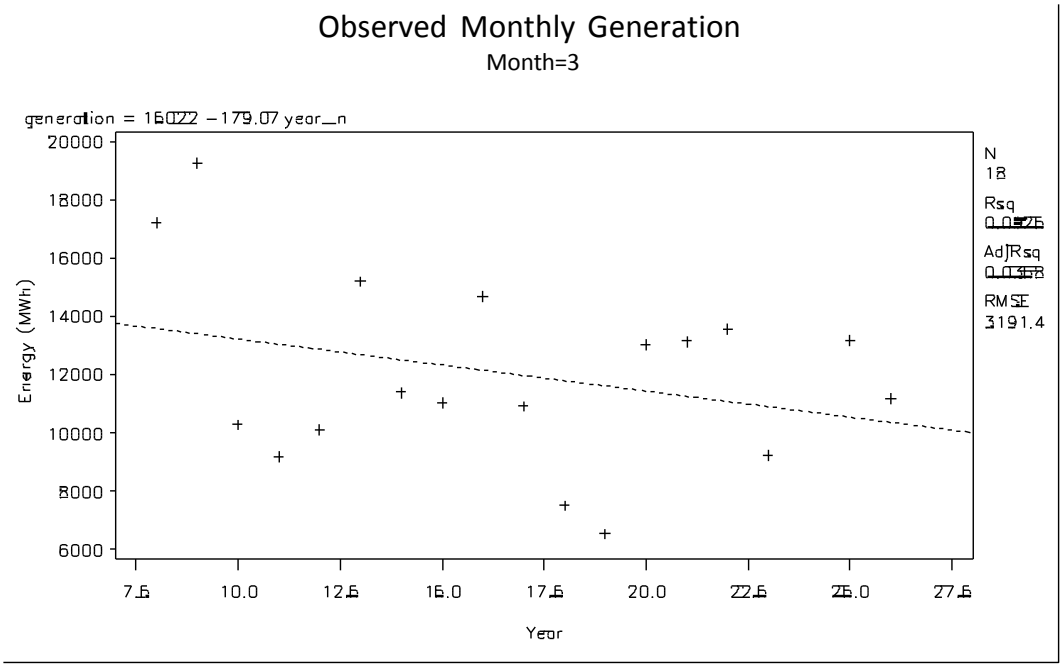


Figure 81. March Energy Generation Over Time

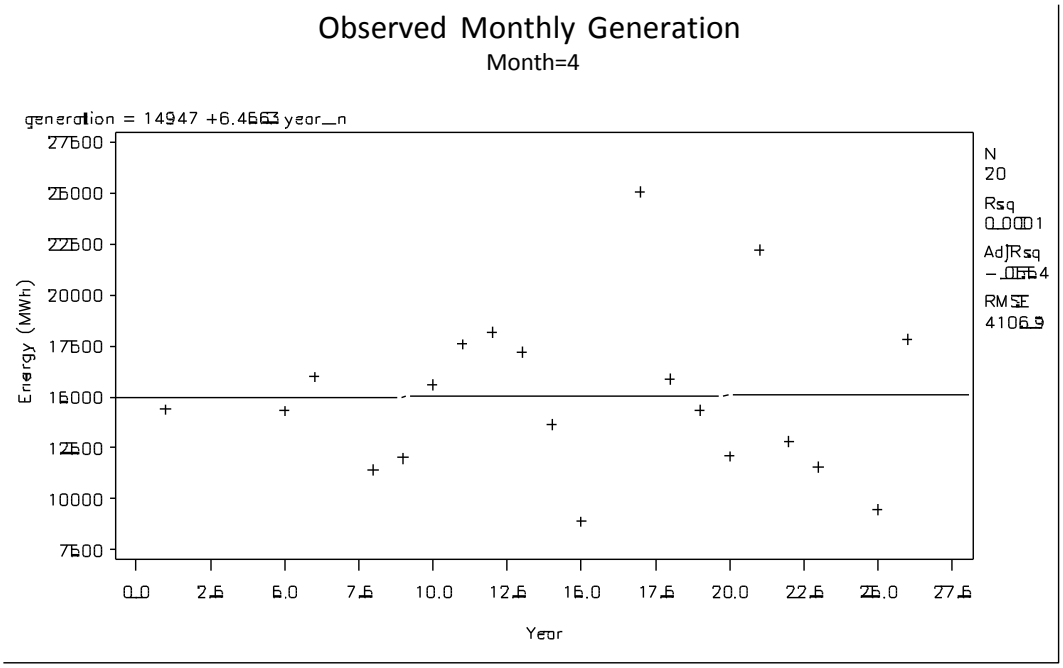


Figure 82. April Energy Generation Over Time

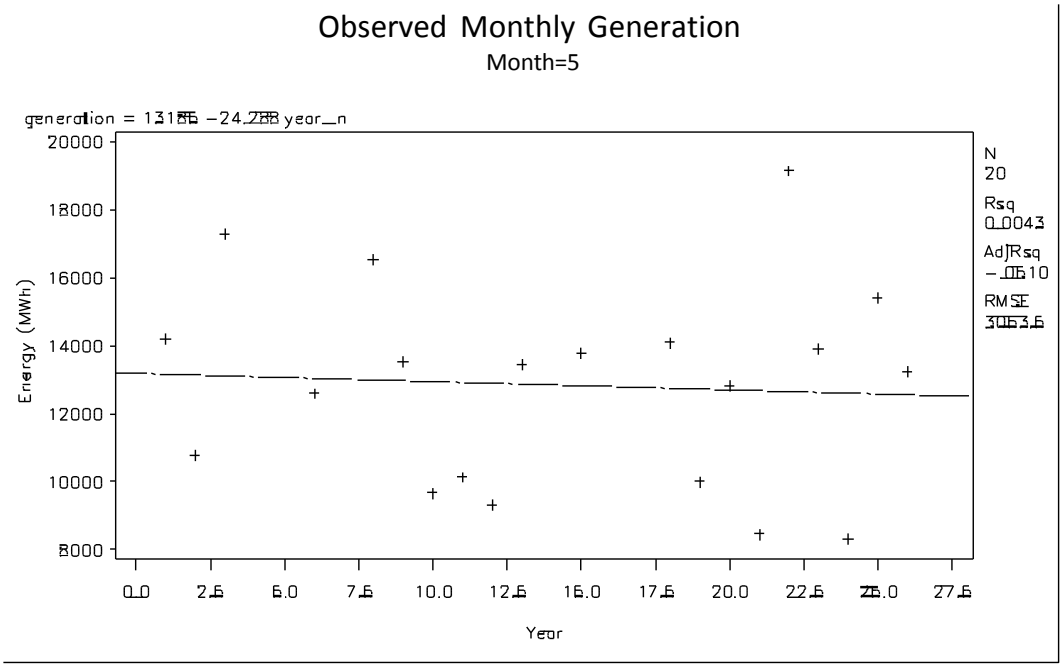


Figure 83. May Energy Generation Over Time

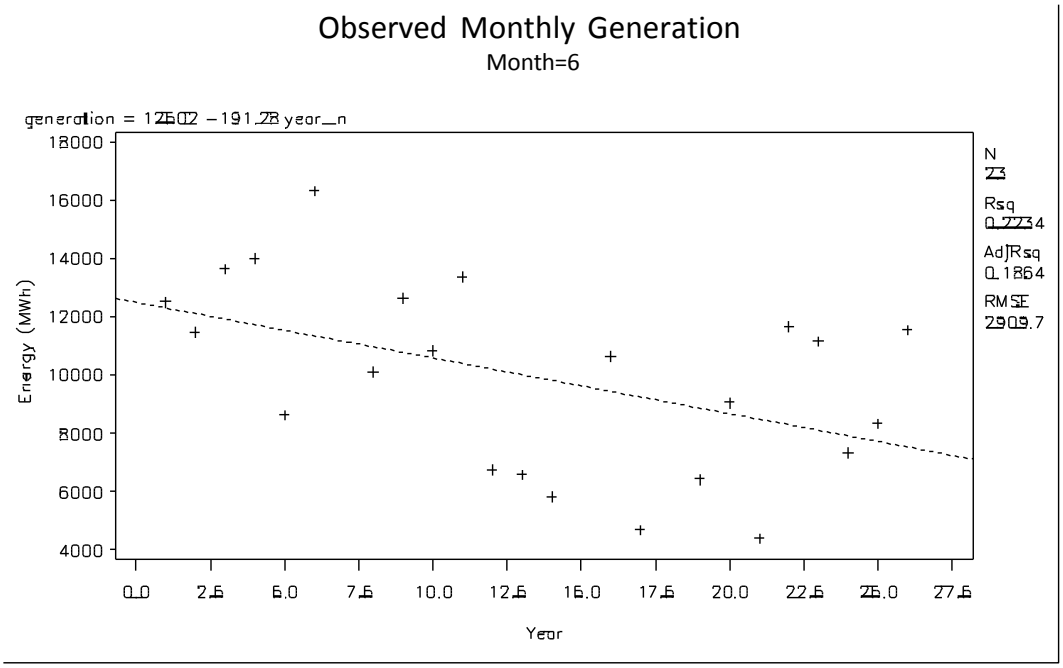


Figure 84. June Energy Generation Over Time

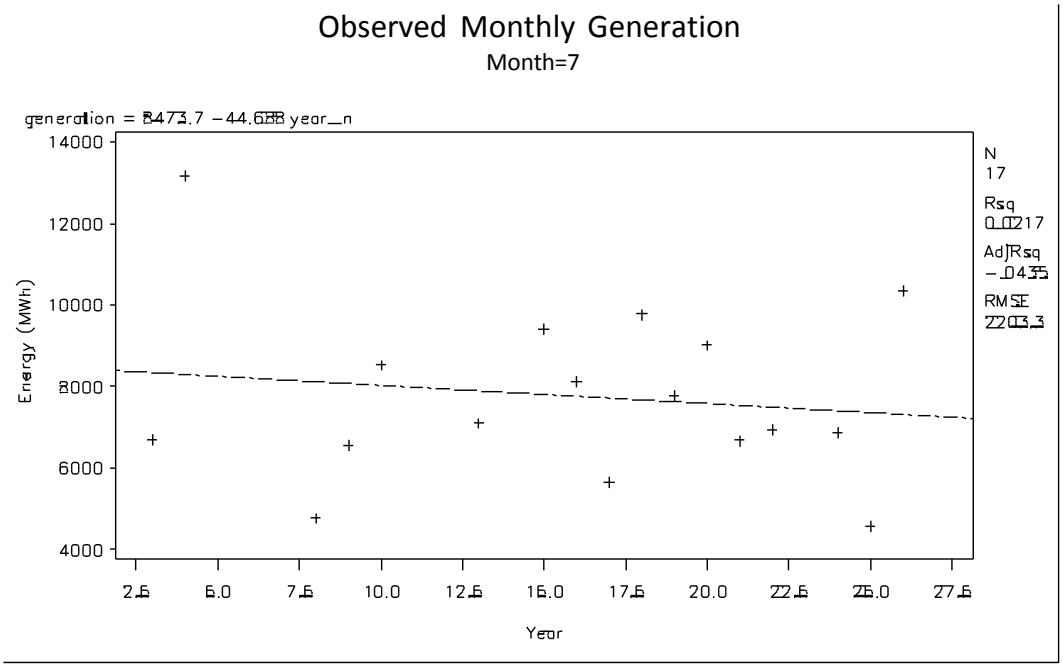


Figure 85. July Energy Generation Over Time

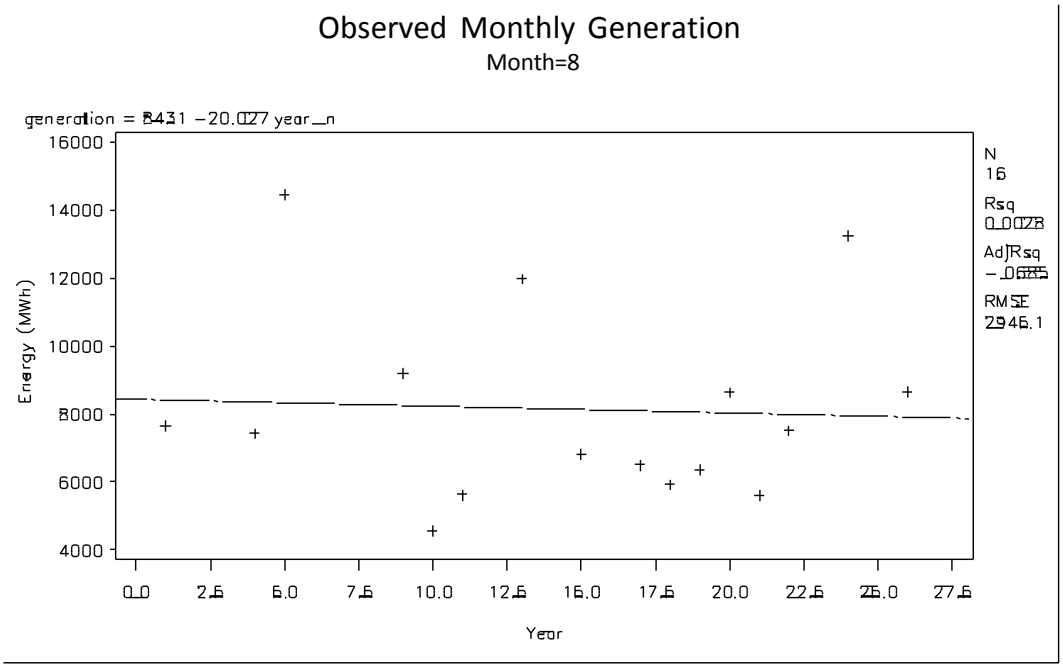


Figure 86. August Energy Generation Over Time

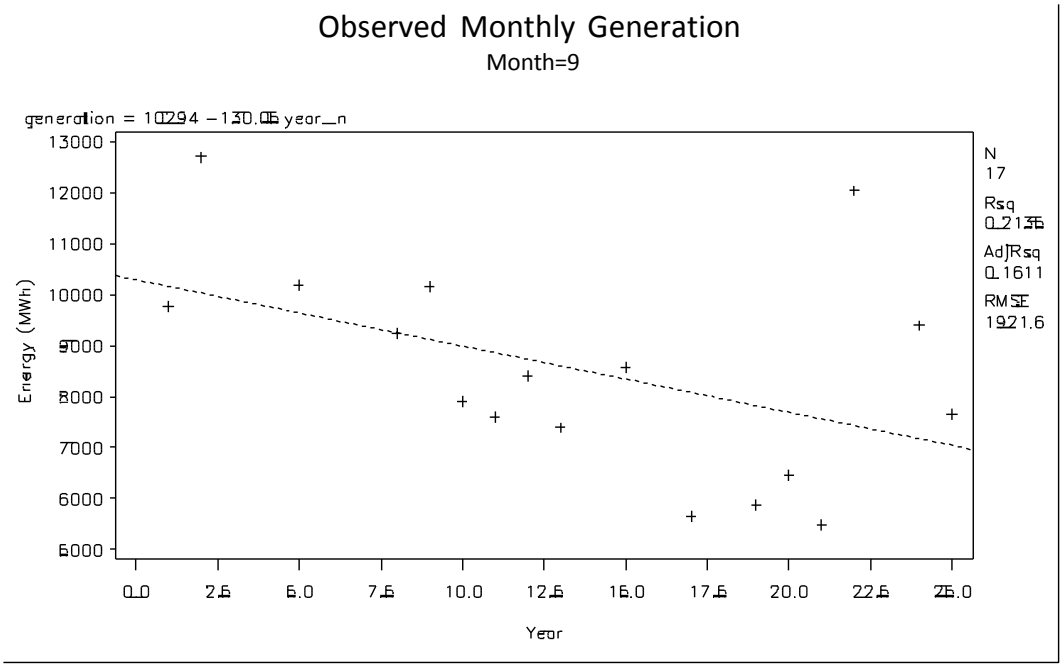


Figure 87. September Energy Generation Over Time

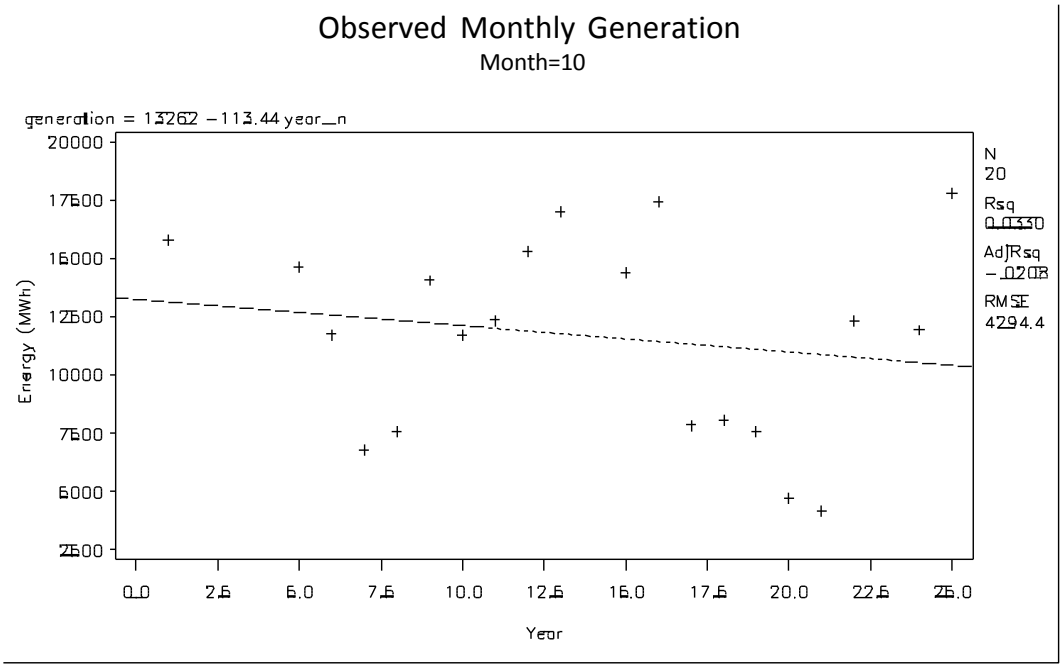


Figure 88. October Energy Generation Over Time

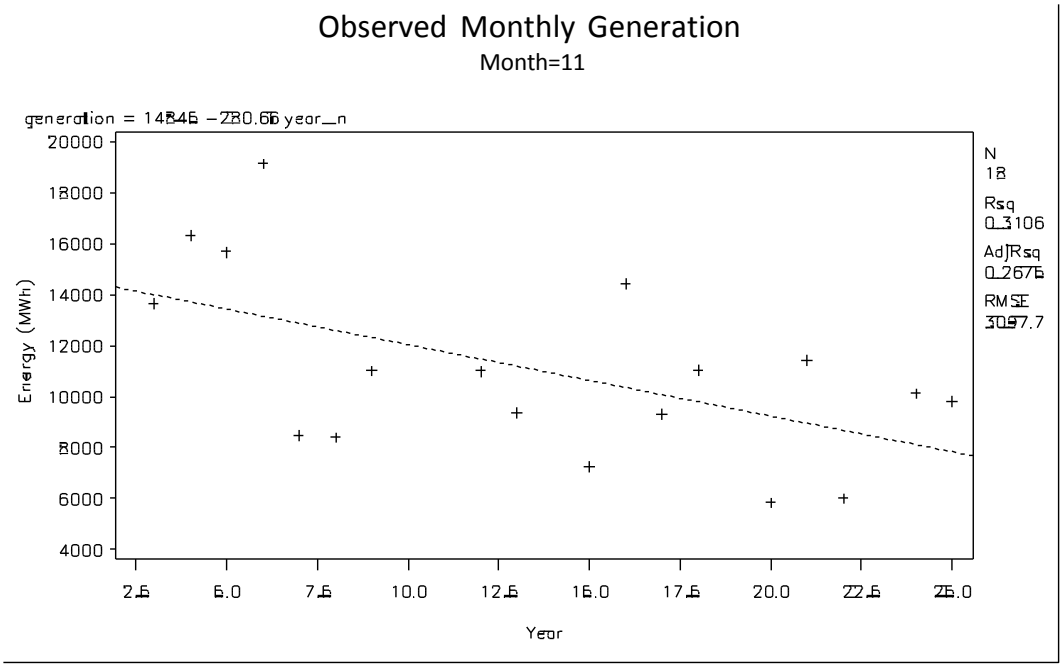


Figure 89. November Energy Generation Over Time

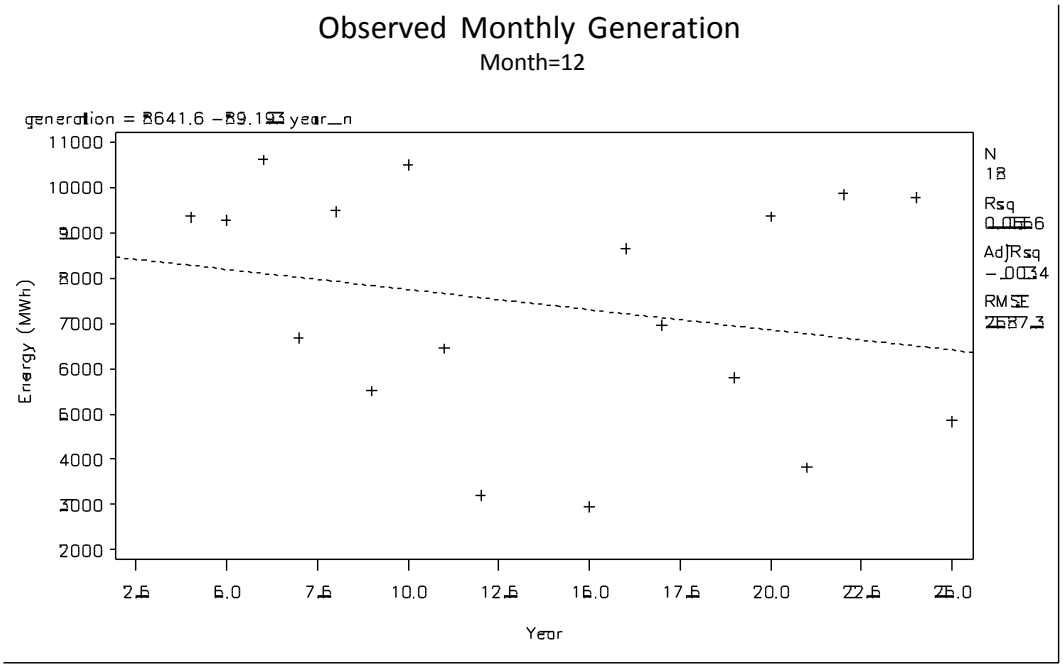


Figure 90. December Energy Generation Over Time

APPENDIX V. SAS PROGRAMMING

This appendix contains the most critical SAS programs utilized during the data analysis. There are three main sections:

1. Empirical data analysis
    a. Contains the program used for the monthly observed energy data
2. Bootstrap data analysis
    a. Contains some basic data analysis programming along with the Survey Select procedure utilized for bootstrapping
3. Capacity data analysis
    a. Contains the program used for analyzing hourly wind speeds

Empirical Data Analysis

```
DATA observed_monthly_sums;
 INFILE 'G:\Thesis\!Final Data Sets\sas input_monthly sums_102009.csv'
    DLM= ',' DSD FIRSTOBS=2 MISSOVER;
 INPUT year_n     year     month  _FREQ_    days_per_month    Days_missing
    monthly_sums  generation;
 label month = 'Month';
 label generation = ' Energy (MWh) ';

DATA observed_energy;
    SET observed_monthly_sums;
    IF year = 2007 then delete;

PROC SORT data = observed_energy;
     BY month;

ods output ParameterEstimates = LognormalParmeters ;
ods output GoodnessofFit = LognormalPvalues ;
ods output FitQuantiles = LognormalQuantiles ;
```

```
options nobyline ;
PROC UNIVARIATE data = observed_energy ciquantdf (alpha = .05) normal plot ;
      *BY month;
       VAR       generation;
       HISTOGRAM   generation / lognormal (theta =est sigma = est zeta = est color =
black )
                                cframe = white
                          vscale    = Percent
                                 height = 4
                          cfill     = H1507050
                               font = calibri
                                    infont = calibri
                                    curvelegend = Lognormal;
                               inset lognormal ( ksdpval= "P-Value" (5.4) sigma  (5.2) theta (5.2)
zeta (5.2)) / height = 3 cfill = ywh ;
                                        inset mean (5.2) std (5.2) n (5.0) / height = 3 cfill = ywh ;
       QQPLOT  generation /lognormal (theta =est sigma = est zeta = est color =
H1507050)
                                cframe = white
                          font = calibri
                                infont = calibri;

          OUTPUT OUT = observed_quantiles
                                mean = mean min = min max = max std = st_dev
                                P1= P99 P5 = P95 P10 =P90 Q1 = P75 Median = P50 Q3= P25
P90 = P10 P95 = P5 P99 = P1;

TITLE1 font = calibri 'Observed Monthly Generation';
TITLE2 height = 1.5 font = calibri ' #byline';
RUN;
OPTIONS byline;

DATA observed_lognormal_stats;
     MERGE LognormalParmeters LognormalPvalues LognormalQuantiles;
     BY Month;
PROC SORT data = observed_monthly_sums;
       BY year;
PROC MEANS DATA = observed_monthly_sums sum mean std min max lclm uclm
noprint;
             BY year;
             VAR generation ;
                      output out = mv_ci
sum = sum mean = mean min = min max = max std = St_Dev lclm = Lower_Limit uclm
= Upper_Limit;
RUN;
```

```
PROC SORT data = observed_energy;
      BY month;
PROC REG DATA = observed_energy;
      BY month;
     MODEL  generation = year_n;
    PLOT generation*year_n;
    PLOT RESIDUAL. * year_n;
        PLOT RESIDUAL. * generation;

PROC ANOVA data = observed_energy;
    CLASS month;
    MODEL generation = month;
    MEANS month / snk;
RUN;
```

Bootstrap Data Analysis

```
DATA superior1;
    SET superior;
    IF Wind_Speed = . THEN delete;
     Label Wind_Speed = 'Wind Speed (m/s)';
RUN;

/*Deleting holdout year*/
IF year = 2007 THEN DELETE;
RUN;
/*Converting wind speed to power(99 MW wind facility) with directional specific
losses*/
DATA power;
      SET superior2;
          IF Wind_Speed LT 3.5 THEN MW = 0;
          IF Wind_Speed GT 25 THEN MW = 0;
          IF Wind_Speed = . THEN delete;
          IF Wind_Speed GE 3.5 and Wind_Speed LE 25
        THEN MW = ((1500 / (1+ 2.718281828 ** (6.25976602 + -0.750823075
*(Wind_Speed))))* Loss_Factor * 66/1000);

RUN;
PROC SORT DATA = power;
      BY month;

/*Summing the hourly power output for each day of each month of each year*/
```

```
PROC SORT DATA = power;
      BY year month day;
PROC MEANS DATA = power sum noprint;
            BY year month day;
            VAR MW ;
            output out = daily_sums
      sum = sum_MW  ;
RUN;

/*Removing days with incomplete data*/

DATA all_daily_sums;
     SET daily_sums;
     IF _FREQ_ = 24;
     DROP _type_;
RUN;

PROC SORT DATA = all_daily_sums;
      BY month;

DATA days_permonth;
INPUT month _NSIZE_;
DATALINES;
1   31
2   28
3   31
4   30
5   31
6   30
7   31
8   31
9   30
10  31
11  30
12  31
;
/*Bootstrapping with PROC SURVEYSELECT – was run with seed = 0 first to get
random seed value, which was then used in program to get the same results every time
the program was run*/

PROC SURVEYSELECT data = all_daily_sums out=outboot
                        method=srs
                        sampsize = days_permonth
                        seed = 304560000
                        rep=2000;
              STRATA month;
```

```
RUN;
PROC SORT DATA = outboot;
    BY month replicate;

/*Summing months by replicate number creating 2000 bootstrap months for each month
of the year*/

PROC MEANS DATA=outboot sum noprint;
    BY month replicate;
    VAR sum_MW;
    OUTPUT OUT =outall sum = sum_MW;
    label sum_MW = 'MWh';
RUN;
PROC SORT DATA = outall;
    BY month;

/*Analyzing the bootstrapped months with PROC UNIVARIATE*/

ods output ParameterEstimates = LognormalParmeters ;
ods output GoodnessofFit = LognormalPvalues ;
ods output FitQuantiles = LognormalQuantiles ;
OPTIONS nobyline;
PROC UNIVARIATE data=outall normal;
        BY month;
      VAR      sum_MW;
      HISTOGRAM sum_MW /lognormal (theta =est sigma = est zeta = est color = black
)
                              cfill     = H1507050
                          font = calibri
                              infont = calibri
                              height = 4
                              curvelegend = Lognormal;
                              inset lognormal ( adpval= "P-Value" (5.4) sigma (5.2) theta
(5.0) zeta (5.2)) / height = 3 cfill = ywh ;
                              inset mean (5.2) std (5.2) n (5.0) / height = 3 cfill = ywh ;
                QQPLOT sum_MW / lognormal (theta =est sigma = est zeta = est);
       OUTPUT OUT = final mean = mean min = min max = max std = st_dev  P1= P1
P5 = P5 P10 =P10 Q1 = P25 Median = P50 Q3= P75 P95 = P95 P99 = P99
pctlpts=2.5, 97.5 pctlpre=ci;
    TITLE1 font = calibri ' Simulated Monthly Energy Generation ';
    TITLE2 height = 1.5 font = calibri ' #byline';
RUN;
OPTIONS byline;
RUN;

DATA lognormal_stats;
    MERGE LognormalParmeters LognormalPvalues LognormalQuantiles;
```

```
    BY Month;

RUN;

DATA logtransform;
    SET outall;
    IF month = 1
    THEN ln_sum_MW = LOG ( sum_MW + 3012.73 );
    IF month = 2
    THEN ln_sum_MW = LOG ( sum_MW + 4425.8 );
    IF month = 3
    THEN ln_sum_MW = LOG ( sum_MW + 9420.37);
    IF month = 4
    THEN ln_sum_MW = LOG ( sum_MW + 26246.9);
    IF month = 5
    THEN ln_sum_MW = LOG ( sum_MW + 12980.6 );
    IF month = 6
    THEN ln_sum_MW = LOG ( sum_MW + 8688.12);
    IF month = 7
    THEN ln_sum_MW = LOG ( sum_MW + 5997.47);
    IF month = 8
    THEN ln_sum_MW = LOG ( sum_MW + 4980.59);
    IF month = 9
    THEN ln_sum_MW = LOG ( sum_MW + 5142.34);
    IF month = 10
    THEN ln_sum_MW = LOG ( sum_MW + 8152.15);
    IF month = 11
    THEN ln_sum_MW = LOG ( sum_MW + 7623.03);
    IF month = 12
    THEN ln_sum_MW = LOG ( sum_MW + 2443.58);
PROC SORT data = logtransform;
    BY month;
RUN;

PROC UNIVARIATE data = logtransform CIQUANTNORMAL (type = lower alpha =
.05);
    BY month;
    VAR ln_sum_MW;
        HISTOGRAM ln_sum_MW / normal (mu = est sigma = est)cfill      = H1507050
                        font = calibri
                            infont = calibri
                            curvelegend = normal;
                            inset normal ( ksdpval= "P-Value" mean std ) / cfill = ywh ;
        QQPLOT ln_sum_MW/ normal (mu = est sigma = est);

RUN;
```

```
PROC ANOVA data = logtransform;
     CLASS month;
     MODEL ln_sum_MW = month;
     MEANS month / SNK;

run;

PROC ANOVA data = power_nozero;
     CLASS month;
     MODEL Wind_Speed = month;
     MEANS month / SNK;

run;
```

Capacity Data Analysis

```
/*Analyzing hourly wind speed and power data*/
DATA power_nozero;
     SET power;
     If Wind_Speed GT 0;
PROC SORT DATA = power_nozero;
      BY month;
OPTIONS nobyline;
PROC UNIVARIATE data = power_nozero;
     BY month;
      VAR Wind_Speed;
      HISTOGRAM Wind_Speed / weibull (theta = 0 sigma = est c = est color = black )
cfill     = H1507050
                              font = calibri
                                   infont = calibri
                                   height = 4
                                   curvelegend = weibull;
                                   inset weibull ( adpval= "P-Value" (5.4) sigma (5.2) theta (5.0)
c (5.2)) / height = 3 cfill = ywh POS = NE  ;
                                   inset mean (5.2) std (5.2) n (5.0)/  height = 3 cfill = ywh POS =
SE ;
     TITLE1 font = calibri 'Monthly Wind Speed Analysis';
     TITLE2 height = 1.5 font = calibri '#byline';
RUN;
OPTIONS byline;
RUN;;
```