7-2017

Benchmarking the Energy Intensity of Small Nebraska Wastewater Treatment Plants

Steven M. Hanna
University of Nebraska-Lincoln, shanna@huskers.unl.edu

BENCHMARKING THE ENERGY INTENSITY OF SMALL NEBRASKA WASTEWATER TREATMENT PLANTS

By

Steven M. Hanna

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Environmental Engineering

Under the Supervision of Professor Bruce I. Dvorak

Lincoln, Nebraska

July, 2017

BENCHMARKING THE ENERGY INTENSITY OF SMALL NEBRASKA WASTEWATER TREATMENT PLANTS

Steven M. Hanna, M.S.

University of Nebraska, 2017

Advisor: Bruce I. Dvorak

To help small communities improve the energy efficiency of their wastewater treatment plants, this study created energy benchmarking models for small wastewater plants serving populations of 10,000 or less and having average flows less than 1.5 million gallons per day (MGD). The purpose of these models is to allow comparisons among plants of similar type and size, identify what factors most significantly impact energy usage, and predict potential savings from changes in key plant characteristics.

Energy usage and plant data from 83 small, mechanical wastewater plants in Nebraska were collected and used to create energy benchmarking models. Data obtained from the Pennsylvania Department of Environmental Protection on 71 small Pennsylvania wastewater plants were also used for modeling and comparisons among the two states. The development of these benchmarking equations was modeled off the American Water Works Association Research Foundation and ENERGY STAR models for large wastewater treatment plants. Separate models were created by state with an overall model created for all plant types, as well as models based off the three most common plant types (extended aeration, oxidation ditch, sequencing batch reactor).

The models predict either intensity (MWh/MG) or usage (kWh/year) for both electric use only and total energy use. Key variables found in most models include extended aeration plant type, supplemental energy usage for sludge treatment, average flow, percent design flow, climate controlled floor area, effluent ammonia-nitrogen, and influent carbonaceous biochemical oxygen demand (CBOD). The resulting models

suggest that the variability of effluent NH_3-N limits may be a more important parameter in determining energy usage than influent and effluent CBOD for small plants. Like past studies, flow was found to explain much of the variation in energy use. Some variables that have not shown up as significant in previous studies may only be significant for small plants. These include climate controlled floor area, supplemental energy usage for sludge treatment, and presence of dewatering equipment. Some variables, such as automatic DO controls, thought to be significant, were found not to be significant. Differences between the Nebraska and Pennsylvania models suggest these types of models may be more region specific.

ACKNOWLEDGEMENTS

There are many people I would like to thank for their help with this study. First, and foremost, I would like to thank my advisor, Dr. Dvorak, for your amazing help and guidance in everything from making the transition to graduate school as smooth as possible to helping me completing my thesis. Your enthusiasm for teaching really shows through your patience, understanding, and ability to easily explain difficult subjects and abstract ideas. I grew academically, as well as professionally through your guidance.

I would also like to thank my graduate committee: Dr. Dahab, Dr. Hanford, and Dr. Williams for their help and willingness to review my work and provide constructive criticism.

I must also thank those who helped me collect all of the data for this study: Matt Thompson, Jack Micek, Ranil Philipose, David Hansen, and Mara Zelt. Your help was vital to obtaining the data I needed to create the best possible models. Your hard work was much appreciated. I want to especially thank Matt for helping me as a co-supervisor of the undergraduate interns and for always lending a hand for anything needed.

I would also like to thank Brianna Hitt from UNL's Stats Help desk who helped me learn how to use SAS better and created programs that allowed me to model the data.

I would like to thank my family, especially my girlfriend, Mary Cate, for supporting me the whole way through and for putting up with the long hours I put into this work. Your support has been what I needed when times were rough.

Lastly, I would like to thank my late father, to whom I dedicate this thesis to. Your hard work, perseverance, and values are what inspired me to be the man I am today.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

APPENDICES

Chapter 1: Introduction

1.1 Background

Energy consumption has become a greater concern in the wastewater industry in recent years (Mizuta and Shimada, 2010). The Electric Power Research Institute (EPRI, 2013) estimates wastewater treatment plants use approximately 30.2 billion kWh per year, or 0.8% of total annual U.S. electricity use, and expect energy consumption to continue to increase. According to the United States Environmental Protection Agency's (EPA) 2012 Clean Watersheds Needs Survey (2012), almost 80% of centralized wastewater treatment plants serve communities with populations of 10,000 or less. Small communities often have limited finances and therefore face challenges in maintaining efficient wastewater treatment plants (EPA, 2017).

Over 95% of wastewater treatment plants in Nebraska serve communities with a population less than 10,000, which are often considered "small" (EPA, 2012). In addition to a lack of funding for energy efficiency improvements, small communities in Nebraska often lack the expertise to conduct energy assessments to determine what areas of a plant can be improved. One of the steps to becoming energy efficient is to determine a plant's baseline energy usage and compare this usage to a best practice benchmark (Public Service Commission of Wisconsin, 2016, Carlson and Walburger, 2007).

Several studies have created energy benchmarks for wastewater treatment plants, but few have focused on small wastewater plants serving populations of 10,000 or less. The American Water Works Association Research Foundation (AwwaRF), now known as the Water Research Foundation, created a benchmarking model equation that scores wastewater treatment plants based on plant energy usage and several other plant characteristics (Carlson and Walburger, 2007). One of the limitations of AwwaRF's model is that it cannot be used by plants treating less than 0.6 million gallons per day (MGD) (Carlson and Walburger, 2007). The Pennsylvania Department of Environmental Protection (PDEP) also conducted an energy benchmarking study of wastewater treatment plants (PDEP, 2011). Their study included over 70 plants treating 1.5 MGD or less, however, their benchmarks are only based on the what treatment process is used and average flow (PDEP, 2011). As shown by AwwaRF's (2007) model, additional plant characteristics to treatment type and flow can have a significant effect on plant energy usage.

The study presented here creates energy benchmarking models for small Nebraska wastewater treatment plants using a similar methodology used in AwwaRF's model development. The method for creating these benchmarking equations involves collection of energy consumption and plant data from small wastewater treatment plants in Nebraska. Data from PDEP's (2011) study is also used to create separate Pennsylvania benchmarking models to provide a comparison of the models between the two states. Multiple linear regression is used to determine which factors affect energy usage the most at small plants and to create models for a comparison between Nebraska and Pennsylvania.

1.2 Objectives

The objectives of this study were to:

1. Assess the energy usage of wastewater treatment plants in Nebraska serving populations of 10,000 or less and treating average flows less than 1.5 million

gallons per day (MGD); and identifying factors that may affect energy usage at these small plants.

2. Create energy benchmarking models using multiple linear regression that predict the electric or energy intensity or usage of small wastewater treatment plants based on data collected from small Nebraska plants.
3. Based off the results of the models, identify which variables, or factors, affect energy usage the most for these small wastewater plants.
4. Create similar energy benchmarking models for small Pennsylvania plants based on data collected from a previous study and compare the two sets of benchmarking models in order to determine if these types of models are nationally relevant or if they may be more appropriate for specific regions or states.

1.3 Thesis Organization

This thesis is organized into five chapters. A literature review can be found in Chapter 2 that gives an overview of previous energy benchmarking studies, as well as an overview of energy usage at small wastewater treatment plants and factors that have been previously found to be significant in regards to energy usage at small plants. Chapter 3 describes the methods used for data collection and statistical analysis. Chapter 4 discusses the results of the statistical modeling including discussion of the final benchmarking models along with a step-by-step example of how to use one these benchmarking models. Chapter 5 lists a summary of the conclusions, as well as recommendations for future research. The references and appendices are attached and include supplemental information such as forms used for data collection, specific model output, and additional plant data.

Chapter 2: Literature Review

2.1 Introduction

In order to place this study into context, an examination of the available technical literature was performed. These topics include past benchmarking studies of wastewater treatment plants, comparisons of small and large wastewater plants, common process types used for small plants, energy and oxygen usage at wastewater treatment plants, and a background of multiple linear regression.

2.2 Past Benchmarking Studies of Wastewater Treatment Plants

Benchmarking the energy usage of wastewater treatment plants has become an increasing priority in recent years due to ever-increasing energy prices, as well as the need to curb the emission of greenhouse gases (Mizuta and Shimada, 2010). Not only are electric rates increasing, but effluent requirements are becoming more stringent, which requires plants to use more energy intensive processes (EPA, 2010). According to a joint report published by the Electric Power Research Institute and the Water Research Foundation, it is estimated that wastewater treatment accounts for 30.2 billion kWh per year or 0.8% of electricity use in the United States (EPRI, 2013). In addition, energy costs account for between 25% and 40% of a typical wastewater treatment plant's operating budget (NYSERDA, 2008). For small systems, a municipality's water and wastewater system can account for 20 to 40% of the municipality's electricity bill (NEO, 2016). A starting point for improving energy efficiency in any sector is to benchmark energy usage.

A review of literature on the study of energy usage and energy benchmarking of wastewater treatment plants reveals that most of the studies concern large plants. There are few studies on the benchmarking of small wastewater plants. In addition, there are two types of benchmarking studies. In one type of study a model equation is developed based on several key plant characteristics (e.g. flow, BOD load, etc.) to help benchmark energy usage. The other type of benchmarking study reports average or median values of either energy intensity (energy usage normalized by either volume of flow, mass of organic loading, or population) or energy usage by treatment process. In addition, some studies measured actual energy use, while others determined energy use based on theoretical calculations (Young and Koopman, 1991). Both types of studies identify energy efficiency strategies and provide general recommendations for what types of processes or equipment are most or least energy efficient.

One of the most well-known efforts to benchmark the energy usage of wastewater treatment plants was a study conducted by AwwaRF. This study created a benchmarking model that predicts energy usage based on several plant characteristics and was used to create the U.S. Environmental Protection Agency's (EPA) ENERGY STAR benchmarking model for wastewater treatment plants (ENERGY STAR, 2014). Several other studies that did not create model equations, but still analyzed plant energy usage to create benchmarks are discussed subsequently. These include studies in Japan (Mizuta and Shimada, 2010), China (Yang *et al*., 2010), Spain (Trapote *et al.*, 2013) Portugal (Silva and Rosa, 2015), New York (NYSERDA, 2008), Wisconsin (Focus on Energy, 2016), Florida (Young and Koopman, 1991), California (PG&E, 2006), and Pennsylvania (PDEP, 2011).

In addition to AwwaRF (Carlson and Walburger, 2007) and ENERGY STAR (2014), another important reference used in this study was the Pennsylvania Department of Environmental Protection's (PDEP) benchmarking study. PDEP's study focused on plants of all flow ranges, but included 82 plants with average flows less than 1.5 MGD (PDEP, 2011). PDEP shared data from their benchmarking study and the data were used to create Pennsylvania benchmarking models for comparison to Nebraska benchmarking models. The AwwaRF and ENERY STAR models are discussed in Section 2.2.1, while the PDEP and other past benchmarking studies are discussed in Section 2.2.2. Section 2.2.3 summarizes the range of energy intensities (energy usage normalized by flow) of wastewater treatment plants found in past benchmarking studies.

2.2.1 AwwaRF and ENERGY STAR Benchmarking Models

AwwaRF's benchmarking study followed similar methodology as the EPA's ENERGY STAR benchmarking score for buildings (Carlson and Walburger, 2007). Much of the approach and methodology of the effort to benchmark small Nebraska wastewater plants is based on AwwaRF's study.

One of AwwaRF's (Carlson and Walburger, 2007) project goals for their benchmarking study was to create a metric that allows for the comparison of wastewater treatment plant energy use among peers. Data were collected on 266 wastewater treatment plants across the United States with average flows greater than 0.6 MGD, average influent BOD greater than 30 and less than 1000 mg/L, treatment electricity use greater than 100,000 kWh per year, and effluent BOD greater than 0 mg/L. The data collected included level of treatment, plant processes, operating conditions, flow volumes, loading, and energy use (Carlson and Walburger, 2007).

AwwaRF set out to create a multi-parameter model that is able to capture the impacts of key plant characteristics on energy use (Carlson and Walburger, 2007). The model was developed using ordinary least squares multiple linear regression analysis with a forward stepwise variable selection approach. Variables were added one at a time to the model to test the impact each had on energy use. Both binary (yes or no) and continuous variables were used in the analysis. Variables were selected for each iteration in the model if their t-test values were above 2.0 (alternatively, with p-values below 0.05). After several iterations, six parameters were selected for inclusion in the final model and the R^2 was 0.89. These six parameters were average flow (MGD), average influent BOD (mg/L), average effluent BOD (mg/L), load factor ($\frac{\text{Average Daily Flow, MGD}}{\text{Average Design Daily Flow, MGD}}$), trickle filtration (yes or no), and nutrient removal (yes or no).

In addition to creating a model to predict annual energy use, AwwaRF (Carlson and Walburger, 2007) went a step further, following the methodology of EPA's ENERGY STAR score for buildings, and developed a score from 1 to 100 (100 = best, 1 = worst) that rates plants based on the ratio of their predicted and actual energy usage. The EPA developed a model using the data and information from AwwaRF's study to create the ENERGY STAR benchmarking model for wastewater treatment plants (ENERGY STAR, 2014). The ENERGY STAR score model for wastewater treatment plants differs slightly from AwwaRF's model. The ENERGY STAR (2014) model was developed using weighted least squares multiple linear regression and instead of predicting annual energy usage, it predicts energy usage per gallon of flow treated. The same six parameters used in AwwaRF's model were used in the ENERGY STAR model, but two extra variables for the effect of weather were added to the ENERGY STAR model: Heating Degree Days and Cooling Degree Days (ENERGY STAR, 2014). These same variables were investigated in AwwaRF's (Carlson and Walburger, 2007) study and found to be statistically significant, but it was decided to leave these weather parameters out of the final model because they are not as readily available to plant operators as other data needed for the model.

One of the limitations of both the ENERGY STAR and AwwaRF models is that they produce outlying results when using the models on smaller plants (average flows < 0.6 MGD) (Carlson and Walburger, 2007, ENERGY STAR, 2014). Therefore, the purpose of the Nebraska benchmarking models is to fill in the data gap and to create a tool for small wastewater plants to use to compare energy use among peers.

2.2.2 Other Past Benchmarking Studies

As mentioned previously, several past studies benchmarked the energy usage of wastewater treatment plants, both in the U.S. and elsewhere. These studies analyzed energy usage, as well as energy usage per unit flow, organic load, and population equivalent. While most of the benchmarking studies focused on large plants, many of the same findings and principles apply to small plants. A summary of key findings from these past studies is discussed subsequently.

Several past studies focused on large plants, but included some small plants in their analysis (Trapote *et al.*, 2013, NYSERDA, 2008, PG&E, 2006, ENERWATER, 2015, Silva and Rosa, 2015, and Mizuta and Shimada, 2010). Three studies by state agencies, Wisconsin's Focus On Energy (Public Service Commission of Wisconsin, 2016), Young and Koopman (1991), and PDEP (2011), focused mainly on plants treating less than 5 MGD. PDEP's study was especially influential for the Nebraska models.

The goal of PDEP's benchmarking study was to provide a tool for operators and managers of wastewater treatment plants in Pennsylvania that allows them to evaluate and reduce plant electric consumption (PDEP, 2011). Surveys were sent out to all plants in Pennsylvania requesting treatment type and unit processes, design and actual plant loading, electric consumption, electric costs, and effluent limits (PDEP, 2011). Full data from 117 plants were collected with 82 of the 117 treating average flows less than 1.5 MGD. Electric intensities (MWh of electricity consumed per million gallons of wastewater treated), as well as electricity consumed per pound of BOD treated, were calculated by treatment type (Extended Air, Conventional Activated Sludge, Sequencing Batch Reactor, Oxidation Ditch, and Trickling Filter) with benchmark values set as the lowest 10% intensity value for each plant type.

PDEP's (2011) study concluded that Extended Aeration plants are much more energy intensive than other plant types with trickling filters being the least energy intensive. They also concluded the closer a plant is to its design capacity, for both flow and BOD loading, the less energy intensive the plant is. Looking at the energy intensity of plants with fine bubble diffusers versus plants with coarse bubble diffusers, PDEP showed that plants with fine bubble diffusers were less energy intensive. Similar to many benchmarking studies, PDEP's (2011) report includes several common energy efficiency strategies for wastewater treatment plants.

Other past benchmarking studies concluded similar findings as to what was found in PDEP's report in addition to reporting baseline energy intensities for the plants analyzed in each study. These baseline energy intensities are presented in Table 2.2 of

Section 2.2.3. A summary of the findings from these past benchmarking studies is summarized in Tables 2.1 and 2.1B.

Table 2.1: A Summary of Important Findings from Past Benchmarking Studies of Wastewater Treatment Plants.

Benchmarking Study Conclusions	Source(s)
Energy costs can account for 25-40% of operating costs.	NYSERDA, 2008
Energy consumption varies by country, size, amount of pollutant loading, and treatment technology used.	Silva and Rosa, 2015, ENERWATER, 2015
Extended aeration plants are much less energy efficient than other plant types.	Silva and Rosa, 2015, ENERWATER, 2015
Fixed Film plants are the most energy efficient plant type due to the absence of aeration.	NYSERDA, 2008, ENERWATER, 2015, Young and Koopman, 1991
The greatest consumers of energy at wastewater treatment plants are aeration and pumping. • Other major consumers include anaerobic digester heating, dewatering equipment, and UV disinfection.	Young and Koopman, 1991, Silva and Rosa, 2015, Trapote *et al.*, 2013, ENERWATER, 2015, NYSERDA, 2008
Energy Intensity (MWh/MG) and energy consumption per unit organic load removed is inversely proportional to the amount of wastewater flow treated.	Silva and Rosa, 2015, ENERWATER, 2015, Mizuta and Shimada, 2010, Trapote *et al.*, 2013, Young and Koopman, 1991, NYSERDA, 2008
The relationship between flow/organic loading and intensity (MWh/MG or MWh/lb-BOD) is more variable than the relationship between flow/organic loading and consumption (MWh).	ENERWATER, 2015
Plants operating closer to their design capacity are more energy efficient.	Silva and Rosa, 2015, Young and Koopman, 1991
Larger plants are more efficient due to: • Economy of scale • Better and more efficient technology and controls • More stable flow and organic loading conditions • Use of biogas from anaerobic digesters allows for the production of energy and supplements energy use • Greater operator presence and training	ENERWATER, 2015, Trapote *et al.*, 2013, Young and Koopman, 1991

Table 2.1B: A Summary of Important Findings from Past Benchmarking Studies of Wastewater Treatment Plants (Continued).

Benchmarking Study Conclusions	Source(s)
Energy Conservation measures include: • Aeration improvements • Solids handling improvements • Waste heat recovery • Inflow and infiltration reduction • Flow equalization • Use of variable frequency drives • Stabilizing the demand by offsetting peak loads. • Building improvements (lighting, HVAC, etc.) • Optimizing process control settings • Utilizing anaerobic digester biogas for supplemental energy	NYSERDA, 2008, PG&E, 2006
Centralization of treatment allows for greater energy efficiency.	Trapote *et al.*, 2013, Mizuta and Shimada, 2010

2.2.3 Energy Intensity Values for Large Wastewater Treatment Plants

Several of the past benchmarking studies mentioned in the previous sections have found energy intensity values for large wastewater treatment plants. Some of these studies included small plants, but most of the plants analyzed in these studies focused on large plants. Energy intensities for studies that included small and large plants are summarized in Table 2.2, while intensities for studies that focused mainly on large plants are summarized in Table 2.3.

Table 2.2: Energy Intensities of Small and Large Wastewater Treatment Plant Studies.

Country	Study	Energy Intensity Range (MWh/MG)	Number of Plants	Approximate Average Flow Range (MGD)
United States	PDEP (2011)	0.4-46.0	117	0.01-95
	Wisconsin's Focus on Energy (2016)	2.3-7.3	85	0->5
	Young & Koopman (1991)	1.5-4.5	5	0.1-2.8
	NYSERDA (2008)	1.1-4.6	174	≤1.0 - ≥75
	PG&E (2006)	0.3-16.4	73	0->22
Spain	Trapote *et al.* (2013)	1.1-9.5	90	0.01-25
Japan	Mizuta & Shimada (2010)	1.1-14.2	985	0.03-130

Table 2.3: Energy Intensities of Studies Focused on Large Wastewater Treatment Plants.

Country	Study	Energy Intensity Range (MWh/MG)	Number of Plants	Approximate Average Flow Range (MGD)
United States	AwwaRF/ENERGY STAR (2007)	1.6-3.3	266	0.6-250
	EPRI (2013)[1]	0.7-3.0	-	-
China	Yang *et al.* (2010)	0.8-1.3	599	2.5-150
Portugal	Silva & Rosa (2015)[2]	1.2-5.1	17	-

[1]EPRI only provided estimates of energy intensity, not observed energy intensities. Therefore, no flow range is listed.
[2]Silva & Rosa did not list the flow range for the plants benchmarked.

2.3 Differences Between Small and Large Wastewater Treatment Plants

There are many differences between small and large wastewater treatment plants. The differences between the plants lead to different ways in which energy is used and how best to manage its usage. A comparison is made here to show the need for separate energy benchmarking models for large and small plants. The EPA considers small plants as treating less than 1 MGD or serving populations of 10,000 or less (EPA, 2012). While the majority of wastewater flows in the United States are treated by large capacity plants, over 80% of centralized wastewater treatment plants serve small communities (EPA, 2012). Small wastewater plants are generally much simpler than large plants, but different challenges are faced by each group (EPA, 2010). Differences in flow, organic loading, treatment processes used, and other characteristics drive the need for separate benchmarking models for large and small wastewater treatment plants.

One of the main differences between large and small wastewater plants is the variation in flow and loading. Smaller plants experience much more extreme variations in flow and loading than larger plants (Boller, 1997). The peaking factor (maximum flow or organic load divided by average flow or organic load) for large plans is normally between

1.5 and 2, but can be as large as 5 for small systems (Boller, 1997). Small plants must be designed to treat larger variations in flow so that unit processes are not upset (Crites and Tchobanoglous, 1998). Therefore, small plants are often overdesigned for the average flow and result in less energy efficient plants (Foladari *et al.*, 2015).

In addition to the differences in flow and organic loading, different treatment processes are often used for small plants. Small plants are often simpler and easier to operate (Tchobanoglous *et al.*, 2014). Plants used for small communities that are simpler than conventional plants and are better equipped to handle greater fluctuations in flow and organic loading include extended aeration plants, sequencing batch reactors, and oxidation ditches (EPA, 2000). As the name implies, extended aeration plants, as well as sequencing batch reactors and oxidation ditches, have much longer solids retention times (SRTs) on the order of 20 to 40 days, as opposed to conventional activated sludge plants that have SRTs between 3 and 15 days (Tchobanoglous *et al.*, 2014). Longer detention times allow these plants to better handle shock loadings (Tchobanoglous *et al.*, 2014).

Another difference between large and small plants is the difference in available capital or finances. A larger community will have more people to share the costs of investing in their wastewater treatment plant than a smaller community. In Nebraska, small communities are seeing declines in population as technological advances in large-scale farming have become more prevalent, leading to less jobs being available and causing young people to move out of rural areas and into larger metropolitan and urban areas in search of better employment opportunities (NDEQ, 2014). This decline in population is further exacerbating the financial needs of small wastewater treatment plants in Nebraska and emphasizes the differences in available financing for small and large communities.

Related to financing is the technology and operational differences in large and small plants. Small plants are unable to afford automated controls and newer, more efficient equipment. Due to limited finances, small plants often are designed with manual controls that do not allow for adjustment as the flow or organic loading changes. Therefore, equipment at small plants is operated to handle peak loads 24/7 even when the flow and organic loading are not at their peak, which therefore wastes energy (Young and Koopman, 1991). In a similar manner, because of a lack of financing, small plants are unable to hire or retain skilled operators (EPA, 2017). For many of the small plants in Nebraska, operators have other duties in town and are only at the wastewater plant a couple of hours a day. According to Boller (1997), the operator strongly influences treatment performance. If a community cannot hire a skilled operator, then their plant may not operate efficiently.

2.4 Oxygen Usage in Wastewater Treatment

Looking back at AwwaRF's benchmarking model for wastewater treatment plants, one can see the correlation between energy use and oxygen demand (Carlson and Walburger, 2007). Because aeration is one of the main users of electricity at wastewater treatment plants, it is important to understand why aeration or oxygen is needed. One of the main objectives of biological wastewater treatment is to transform, or oxidize, waste via microorganisms (Tchobanoglous *et al.*, 2014). In the aerobic treatment of wastewater, oxygen consumed by microorganisms drives the metabolic reactions that transform the waste into acceptable end products (Tchobanoglous *et al.*, 2014). In order to provide oxygen to the microorganisms, conventional treatment provides it via aeration.

Aeration of wastewater is the greatest consumer of energy at wastewater treatment plants (EPA, 2010). Young and Koopman (1991) determined aeration can account for 54-97% of a small wastewater treatment plant's total electricity consumption. Oxygen is needed for removal of carbonaceous matter, as well as nutrients such as nitrogen (Tchobanoglous *et al.*, 2014). Differing amounts of oxygen are needed for the removal of carbonaceous and nitrogenous material (Tchobanoglous *et al.*, 2014). Almost all of the wastewater treatment plants in Nebraska are required to meet effluent limits for ammonia. None of the small Nebraska plants analyzed in this study were required to meet effluent limits for Total Nitrogen or Total Phosphorous. The following discussion details the background behind the oxygen needed to oxidize both carbonaceous organic matter and ammonia.

The amount of oxygen needed for the removal of carbonaceous material, carbonaceous biochemical oxygen demand (CBOD), is the amount of oxygen needed for the oxidation of waste, the synthesis of new cells, and endogenous respiration (Tchobanoglous, *et al.*, 2014). Endogenous respiration is the term that describes when organic matter is used up and the new cells start consuming their own cell tissue to obtain energy for cell maintenance (Tchobanoglous *et al.*, 2014). These three processes can be shown with the following generalized chemical reactions [2.1], [2.2] and [2.3] with COHNS representing organic waste (composed of carbon, oxygen, hydrogen, nitrogen, and sulphur) and $C_5H_7NO_2$ representing cell tissue (Tchobanoglous *et al.*, 2014).

Energy Reaction:

$$\text{COHNS} + O_2 \rightarrow CO_2 + NH_3 + H_2O + \text{other end products} + \text{energy} \qquad [2.1]$$

Cell Synthesis Reaction:

$$\text{COHNS} + O_2 + \text{bacteria} + \text{energy} \rightarrow C_5H_7NO_2(\text{New Cells}) \qquad [2.2]$$

Endogenous Respiration:

$$C_5H_7NO_2 + 5O_2 \rightarrow 5CO_2 + NH_3 + 2H_2O \qquad [2.3]$$

According to Tchobanglous *et al.* (2014), the oxygen requirement for the removal of CBOD is between 0.90 and 1.3 pounds of O_2 per pound of CBOD.

The need to remove ammonia (NH_4-N) and nitrite (NO_2-N) in wastewater is due to concerns over the effect of ammonia on receiving water dissolved oxygen concentrations and fish toxicity, concerns over eutrophication, and to prevent groundwater contamination by ammonia (Tchobanoglous *et al.*, 2014). As with the removal of CBOD, ammonia and nitrite are removed via aeration of the wastewater (Tchobanglous *et al.*, 2014). The process for converting ammonia and nitrite into nitrate is termed nitrification. Nitrification is a two-step process where 1) ammonia is oxidized to form nitrite [2.4] and then 2) nitrite is oxidized to form nitrate [2.5] (Tchobanglous *et al.*, 2014). The following chemical equations, [2.4] and [2.5], show this two-step process.

$$2NH_4^+ + 3O_2 \rightarrow 2NO_2^- + 4H^+ + 2H_2O \qquad [2.4]$$

$$2NO_2^- + O_2 \rightarrow 2NO_3^- \qquad [2.5]$$

The total oxidation reaction is:

$$NH_4^+ + 2O_2 \rightarrow NO_3^- + 2H^+ + H_2O \qquad [2.6]$$

Based on this total oxidation reaction [2.6], the oxygen requirement for the oxidation of ammonia to nitrate, nitrogenous biochemical oxygen demand (NBOD), is 4.57 pounds of oxygen per pound of ammonia (Tchobanoglous *et al.*, 2014).

Nitrifying bacteria have a much slower reproduction rate than bacteria that breakdown carbonaceous material (Tchobanglous *et al.*, 2014). Because nitrifying bacteria take a long time to reproduce, wastewater must be kept in the aeration basin for a longer time in order for nitrifying bacteria to reach a significant population (Tchobanoglous *et al.*, 2014). This affects the design of wastewater treatment plants because aeration basins must be designed to be larger to increase the solids retention time (SRT) if nitrification is desired (Tchobanoglous *et al.*, 2014). Longer SRTs result in lower sludge yields (Crites and Tchobanoglous, 1998, EPA, 2000). In addition to longer SRTs, nitrifying bacteria require higher dissolved oxygen (DO) levels that are above 1.0 mg/L with the maximum nitrification rate occurring when the DO level is 3 to 4 mg/L (Tchobanoglous *et al.*, 2014).

The differences in the amount of oxygen required to remove CBOD and ammonia help explain the amount of aeration required and therefore the amount of energy consumed for aeration. While the influent concentration of ammonia is much less than CBOD in typical domestic wastewaters, it is important to keep in mind the amount of oxygen needed to remove both to understand energy usage in aeration systems.

2.5 Common Process Types for Small Wastewater Treatment Plants

There are three main plant types used in small communities. These include extended aeration plants, oxidation ditches, and sequencing batch reactors (EPA, 2000). Each type has unique characteristics about them and how they are operated. They are all biological aeration processes (EPA, 2000). The control of aeration for each plant type is somewhat unique, which leads to differing amounts of energy being used. All three plant types have long detention times and are equipped to handle large variations in flow rates (EPA, 2000). In addition to differing energy use due to different plant types used, sludge stabilization can also be a significant energy user. The following sections detail each main plant type, how each is operated, challenges operators face with each plant type, and overviews of aerobic and anaerobic digestion.

2.5.1 Extended Aeration

The extended aeration process is a modification of the conventional activated sludge process (EPA, 2000). They are typically designed to treat flow rates between 0.002 and 0.1 MGD (EPA, 2000). The treatment basin is aerated to provide oxygen to the microorganisms that break down suspended waste, but differing from conventional treatment, the wastewater is aerated for a more extended period of time than conventional treatment, which in turn uses more energy (EPA, 2000, Tchobanoglous *et al.*, 2014).

The treatment train for a typical extended aeration plant starts with screening or grinding of the influent to prevent large solids from harming equipment downstream. Sometimes flow is then sent to a primary clarifier where solids settle out, or it goes directly from screening/grinding to the aeration basin. The wastewater is then aerated using air diffusers that bubble air through the wastewater so that microorganisms can oxidize the suspended organic matter. The aerated water next flows to a final clarifier while the sludge from the aeration basin is sent to an aerobic digester. After the final clarifier, disinfection of the water occurs using either UV lights or chemical means (chlorine) and the disinfected water is sent to the receiving body of water. A process flow diagram for a typical extended aeration plant and an aerial view of a typical extended aeration plant are shown in Figures 2.1 and 2.2, respectively.

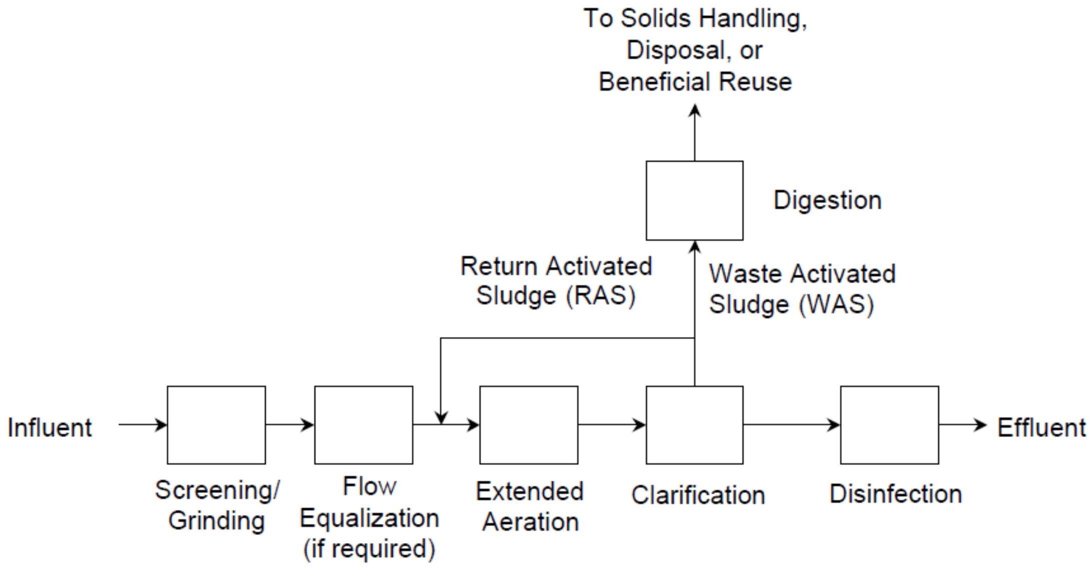


Figure 2.1: Process Flow Diagram for a Typical Extended Aeration Plant (EPA, 2000).

Figure 2.2: Aerial View of a Typical Extended Aeration Plant (Google Maps, 2017a).

Extended aeration plants can be modified to provide nutrient removal (EPA, 2000). Most plants built before the 1980s were generally not designed for nutrient removal, including extended aeration plants (Tchobanoglous *et al.*, 2014). Plants designed only for the removal of carbonaceous organic material have smaller treatment basins, leading to shorter solids retention times, and have lower oxygen requirements (Tchobanoglous *et al.*, 2014). Plants not designed for nitrification must provide a greater amount of aeration and increase their SRTs by wasting sludge less often in order to achieve complete nitrification (Li and Wu, 2014).

Extended aeration plants tend to use more energy than other common small plant types (PDEP, 2011, Silva and Rosa, 2015, ENERWATER, 2015). An energy consumption study of 746 wastewater treatment plants in China found extended aeration plants to have the highest energy intensity out of all plant types studied (Yang *et al.*, 2010). Yang *et al.* (2010) found extended aeration plants to have an average energy intensity of 1.29 MWh/MG. The flow range of the extended aeration plants studied was not mentioned, but it was implied that all flow ranges were analyzed, which explains the low value obtained in their study.

Extended aeration plants were also found to be the most energy intensive plant type in PDEP's (2011) study which analyzed 26 extended aeration plants in Pennsylvania with average flows less than 1.5 MGD. The median energy intensity for extended aeration plants in PDEP's (2011) study was 11.8 MWh/MG, a much larger value than what was found in China. Silva and Rosa (2015) reported an average energy intensity for extended aeration plants in Portugal somewhat in the middle of PDEP (2011) and Yang *et al.* (2010) at 2.7-5.1 MWh/MG. ENERWATER (2015), a study by the European Commission, gathered data on 118 plants serving populations less than 2,000 from several studies in the past in various countries. They reported a range of energy intensity for extended aeration plants as between 0.8 and 20.8 MWh/MG (ENERWATER, 2015). The wide range in intensity values may be explained by factors other than flow.

2.5.2 Oxidation Ditch

Oxidation ditches are easily identified by their "racetrack" shape. The oxidation ditch is a ring or oval-shaped channel usually equipped with mechanical surface aerators (Crites and Tchobanoglous, 1998). Oxidation ditches operate in complete-mix extended aeration mode (EPA, 2000). Figure 2.3 shows a general process flow diagram for an oxidation ditch, while Figure 2.4 shows an aerial view of a typical, small oxidation ditch plant.

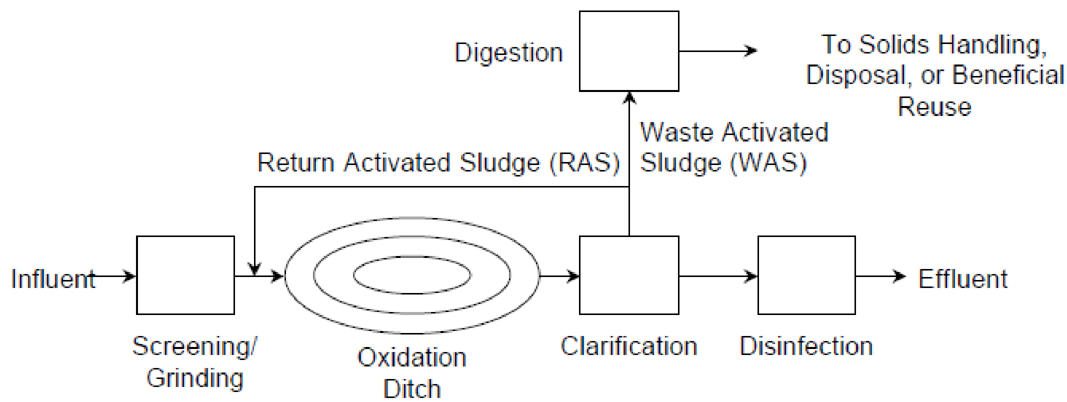


Figure 2.3: Typical Oxidation Ditch Process Flow Diagram (EPA, 2000).

Figure 2.4: Aerial View of a Small Oxidation Ditch Plant (Google Maps, 2017b).

Oxidation ditches are normally aerated using brush-type or surface type mechanical aerators (Tchobanoglous *et al.*, 2014). These aerators, or the water level in the ditch, can sometimes be raised or lowered in order to decrease or increase the dissolved oxygen levels. These surface aerators are much less efficient at transferring oxygen to the wastewater than diffused aeration (Tchobanoglous *et al.*, 2014).

Oxidation ditches can provide complete nutrient removal with slight process modifications (EPA, 2000). When oxidation ditches were first being built in the United States in the 1960s, however, most were not designed for nitrification or denitrification (EPA, 1992). It was determined later, however, that due to the fact that they were designed to operate as an extended aeration process with long SRTs, they can achieve significant nitrification if enough oxygen is supplied (EPA, 1992). Denitrification can be achieved in oxidation ditches by creating an anoxic zone, or a zone where there is no oxygen available and nitrate replaces oxygen as the electron acceptor (Tchobanoglous *et al.*, 2014).

Energy usage at oxidation ditches is mainly in the aerators and pumps. Several benchmarking studies have included oxidation ditches in their analyses. These include Mizuta and Shimada (2010), Yang *et al.* (2010), PDEP (2011), and Wisconsin's Focus on Energy (Public Service Commission of Wisconsin, 2016). Mizuta and Shimada (2010) and Yang *et al.* (2010) analyzed oxidation ditches of all flow ranges and reported average energy intensities between 1.1 and 7.8 MWh/MG. PDEP (2011) and Wisconsin's Focus on Energy (2016) studies included 7 and 19 oxidation ditches, respectively, with average flows less than 1.5 MGD. PDEP (2011) reported the average energy intensity for oxidation ditches as between 2.2 and 6.6 MWh/MG. Wisconsin's Focus on Energy

(2016) reported an average energy intensity for oxidation ditches as 3.9 MWh/MG. Larger plants have lower intensities, so it is not a surprise that PDEP and Wisconsin's Focus on Energy averages are slightly higher than the other two studies with the exception of a few outliers in Mizuta and Shimada (2010) and Yang *et al.* (2010).

2.5.3 Sequencing Batch Reactor

Sequencing batch reactors (SBRs) are a fill and draw batch process, with all of the treatment occurring in one basin (EPA, 2000, Tchobanoglous *et al.*, 2014). There are usually at least two treatment basins at small plants that act in parallel, allowing for continuous flow into the plant (EPA, 2000). The only treatment not occurring in the basins are sludge digestion and disinfection (EPA, 2000). A typical cycle for an SBR is 3 hours of fill, 2 hours of aeration, 0.5 hours of settling, and 0.5 hours of withdrawal of the supernatant (Tchobanoglous *et al.*, 2014). Cycle times can vary for each plant, as well as what occurs during each cycle. An example of a cycle is shown in Figure 2.5. An aerial view of an SBR plant is shown in Figure 2.6.

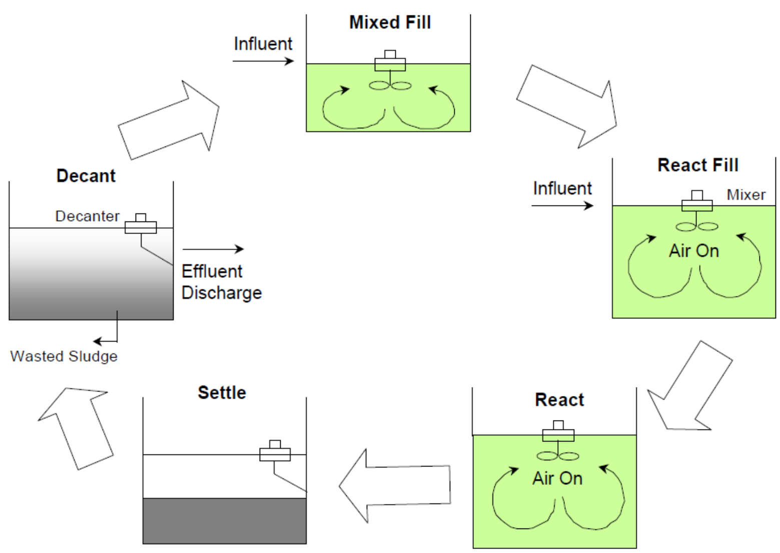


Figure 2.5: Process Flow Diagram for a Sequencing Batch Reactor (EPA, 2000).

Figure 2.6: Aerial View of a Typical SBR Plant (Google Maps, 2017c).

SBRs are typically designed for treating flow rates between 0.01 to 0.2 MGD, but larger SBR systems do exist (EPA, 2000). They typically use diffused aeration with mechanical mixers used to keep the solids suspended when the aeration is off (Tchobanoglous *et al.*, 2014). In regards to nutrient removal, because of the ability to change cycle times, the reactor can be set for aerobic, anoxic, and anaerobic conditions in order to achieve biological nutrient removal (NEIWPCC, 2005, EPA, 1999, EPA, 1985). The only process change needed to achieve nitrification is to ensure the aeration cycle time is sufficiently long enough and that the dissolved oxygen levels are high enough to allow for the completion of ammonia-nitrogen oxidation (EPA, 1985). Because of their ability to handle varying loads, SBRs are often used to treat both municipal and industrial wastewaters (NEIWPCC, 2005).

In a technology assessment performed by the EPA in 1985, when SBRs were first being introduced, it was estimated that due to the fill and draw nature of SBRs, they could be more energy efficient than both oxidation ditches and extended aeration plants of similar size (EPA, 1985). In practice, this is generally true and has been shown by several studies. Yang *et al.* (2010) found an average energy intensity for SBRs of 1.3 MWh/MG. In their study, this intensity was more efficient than extended aeration plants, but slightly more energy intensive than oxidation ditches (Yang *et al.*, 2010). PDEP (2011) also showed that SBRs were less energy intensive than extended aeration plants, but more intensive than oxidation ditches with a median energy intensity of 6.1 MWh/MG. ENERWATER (2015) and EPRI (2013) reported low energy intensities for SBRs in the range of 0.7 to 1.5 MWh/MG, however, both reports included SBRs of all flow ranges in their calculations and EPRI's value is an estimate based on theoretical values, not measured energy usage. This relatively large range in reported values of energy intensities may again imply that other factors besides flow affect energy usage.

2.5.4 Aerobic Sludge Digestion

Aerobic sludge digestion is typically used in plants treating less than 5 MGD (Tchobanoglous *et al.*, 2014). It is the most common sludge stabilization method for small wastewater plants in Nebraska. The goal of sludge (solids) stabilization is to reduce pathogens, eliminate odors, and prevent putrefication in the solids (Tchobanoglous *et al.*, 2014). Aerobic sludge digestion is similar to the activated sludge process. As the supply of food (waste) becomes depleted, the bacterial cells start consuming themselves for energy. When cell tissue is consumed, it is oxidized into carbon dioxide, water, and ammonia (Tchobanoglous *et al.*, 2014).

While aerobic digestion is less expensive, for both capital and operational costs, and requires fewer special skills than other stabilization methods for reliable operation, it is often more energy intensive (Nowak, 2006, Tchobanoglous *et al.*, 2014). Solids retention times for aerobic digesters range from 40 to 60 days depending on the outside temperature (Tchobanoglous *et al.*, 2014). According to a study published by the European Commission (ENERWATER, 2015) that studied the energy consumption of wastewater treatment plants and unit processes of plants around the world, aerobic digestion is the greatest consumer of energy out of all the stabilization methods. ENERWATER (2015) reported a median energy intensity for aerobic digestion for plants serving less than 2,000 people to be 2.0 MWh/MG.

2.5.5 Anaerobic Sludge Digestion

Anaerobic digestion is the biological conversion of organic matter by fermentation (absence of oxygen) in a heated reactor to produce methane gas and carbon dioxide (Tchobanoglous *et al.*, 2014). The anaerobic digestion process is much more complicated than aerobic digestion and requires skilled operators to run efficiently (Tchobanoglous *et al.*, 2014). Most anaerobic digesters are operated in the mesophilic range, or at temperatures between 85 and 100°F and are heated by either the combustion of biogas produced from the digester or natural gas (Tchobanoglous *et al.*, 2014). Generally in the past, anaerobic digestion was only thought to be economically feasible for plants serving populations of more than 50,000 (Nowak, 2006). However, recent studies have shown that it may be economically feasible for plants serving as few as 7,500 (Gretzschel *et al.*, 2014).

A study published by the European Commission (ENERWATER, 2015) reports an average energy intensity of the anaerobic digestion process as 0.02 MWh/MG for plants serving populations between 50,000 and 100,000. For larger plants, the production of biogas can meet much of the energy needs for plant operation (Tchobanoglous *et al.*, 2014). Due to seasonal fluctuations of biogas production, smaller plants with anaerobic digesters generally are only able to use biogas for heating their digesters and often must use supplemental natural gas when not enough biogas is produced (Wong and Law-Flood, 2011). The complexities of operation, as well as fluctuations in biogas production make the use of anaerobic digesters less attractive for small wastewater treatment plants.

2.6 Energy Usage at Wastewater Treatment Plants

Energy in wastewater treatment plants is used throughout the treatment process. The main users for all plant sizes, however, are pumping and aeration operations (EPA, 2010). Other treatment processes requiring significant amounts of energy include solids treatment and processing and disinfection (EPA, 2010). The following sections detail main energy users at small plants, as well as a breakdown of energy usage for each unit process employed at wastewater treatment plants.

2.6.1 Main Users of Energy at Small Wastewater Plants

Much like large wastewater treatment plants, the majority of energy usage at small wastewater plants can be attributed to two main pieces of equipment: aeration blowers and pumps (Young and Koopman, 1991 and Foladari *et al.*, 2015). Young and Koopman (1991) and Foladari *et al.* (2015) conducted energy consumption studies of small wastewater treatment plants. Each study measured the energy use of unit processes in five small wastewater treatment plants (Young and Koopman, 1991, Foladari *et al.*, 2015). Each study directly measured electric usage of equipment using electric meters. The average flow range for the plants studied by Young and Koopman (1991) was 0.1-2.8 MGD and Foladari *et al.* (2015) studied plants with average flows between 0.03 and 0.8 MGD. Plants types studied in these two papers included extended aeration plants, trickling filters, and contact stabilization plants (Young and Koopman, 1991, Foladari *et* *al.,* 2015). Young and Koopman (1991) found that aeration can account for 54-97% of plant electricity use. In large plants, aeration accounts for a smaller percentage of total energy use in the range of 25-60% of total energy usage (Silva and Rosa, 2015). Other equipment such as clarifier scrapers, grit-removal devices, and mechanical bar screens use very little energy at small wastewater plants in comparison to aeration equipment and pumps (Young and Koopman, 1991, Foladari *et al.,* 2015). Figure 2.7 shows the percent energy use of unit processes in a small wastewater plant from Young and Koopman's (1991) study that uses most of its energy for aeration.

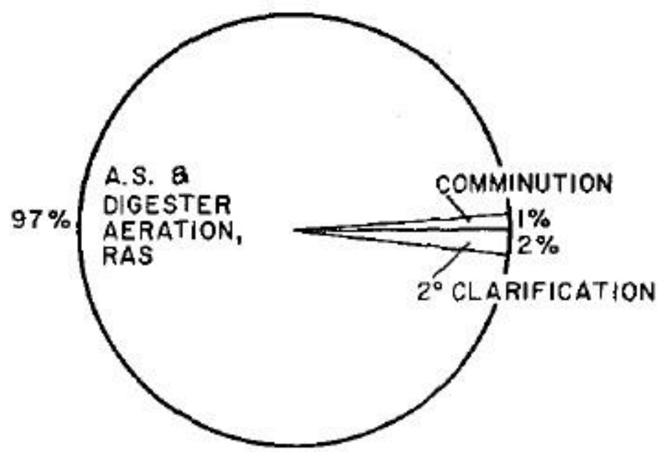


Figure 2.7: Distribution of Energy Use at an Extended Aeration Plant with an Average Flow of 0.15 MGD (Young and Koopman, 1991).

An area of treatment that is often ignored in regards to energy consumption at small plants is the aerobic digestion process (Foladari *et al.,* 2015). Even though related to aeration, aerobic digestion can use a significant amount of energy at small plants. ENERWATER (2015) found that the average energy intensity of aerobic digestion to be around 2.0 MWh/MG for communities serving 2,000 people or less. Dewatering of the sludge can also be an energy intensive process if mechanical means are used. Although not always present at small facilities, when dewatering equipment is present, the energy used for dewatering can be significant for facilities depending on the amount of time the equipment is operated (Foladari *et al.*, 2015).

2.6.2 Energy Use by Common Wastewater Treatment Processes

Electricity use and management in the water and wastewater industries was addressed in a study conducted by the Electric Power Research Institute (EPRI, 2013). One of the objectives of this study was to estimate unit process electric consumption for individual wastewater unit processes (EPRI, 2013). Although unit process energy use has been addressed several times in previous sections of this literature review, EPRI (2013) presents estimates for all possible wastewater unit processes used today in the industry. These estimates were based on theoretical calculations as well as data from AwwaRF's (Carlson and Walburger, 2007) benchmarking study. EPRI (2013) developed daily energy consumption estimates for unit processes for plants with average flows of 1, 5, 10, 20, 50, 100, and 250 MGD. The estimates for a plant with an average flow of 1 MGD are presented in Table 2.4 on the following page.

Table 2.4: Estimates of Electric Intensity (kWh/day) of Wastewater Treatment Unit Processes for a Plant with Average Flow of 1 MGD (EPRI, 2013).

Unit Process		Electric Usage (kWh/day)
Wastewater Pumping		220
Primary Treatment	Odor Control	150
	Grit Removal, Aerated	130
	Grit Removal. Forced vortex	160
	Primary Clarifiers	30
	Ballasted Sedimentation	75
Secondary Treatment	Trickling Filters	630
	Biological Nutrient Removal (BNR) mixing	110
	Aeration without nitrification	720
	Aeration with nitrification	1,080
	Secondary Clarifiers	85
	Sequencing Batch Reactors	1,090
	Membrane Bioreactors	2,700
Solids Handling, Treatment & Disposal	Aerobic Digestion	1,000
	Gravity Belt Thickener	30
	Centrifuge Thickening	80
	Screw Press	20
	Centrifuge Dewatering	260
	Thermal Drying	221
Filtration & Disinfection	UV Disinfection	225
	Depth Filtration	100
	Surface Filtration (e.g. cloth filters)	50
Nonprocess loads (buildings, lighting, computers, pneumatics, etc.)		300

Table 2.4 presents the differences in energy consumption between different wastewater treatment processes and helps illustrate the fact that different processes use differing amounts of energy. The amount of energy used at plants can be broken down even simpler by showing a pie chart of what common process types use the most energy at large wastewater treatment plants. Figure 2.8 shows a breakdown of typical energy

end-uses in large wastewater treatment plants. These percentages were developed based on numerous energy audits of wastewater treatment plants (EPRI, 2013). Understanding the relative energy use by unit processes for large plants can help one understand the relative energy use at small plants.

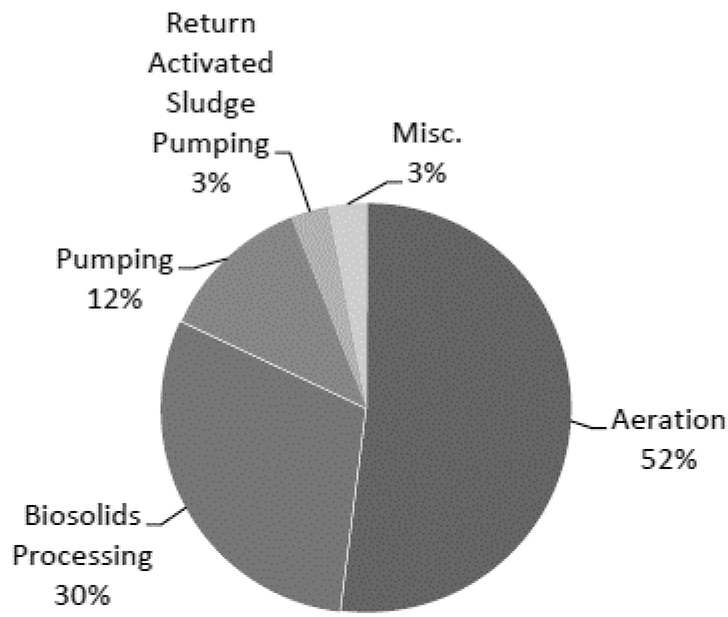


Figure 2.8: Typical Energy End-Uses in Municipal Wastewater Treatment (EPRI, 2013).

2.7 Statistical Analysis Background

The models created in this study were created using multiple linear regression (MLR) analysis. MLR is one of the most powerful and most used statistical techniques (Ngo, 2012, Sheather, 2009). It has been shown to be an appropriate measure to predict the energy usage or intensity of wastewater treatment plants (Carlson and Walburger, 2007, ENERGY STAR, 2014). MLR models the relationship between a dependent (response) variable, Y, and p independent (explanatory) variables x_i where i = 1, …, p (Sheather, 2009). The x_i variables are linearly related to Y by linear coefficients β_i where i = 1, …, p. The general form of a multiple linear regression equation is $Y = \beta_0 + \beta_1 x_1 +$

$\beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon$, where β_0 is the intercept and ε is the random error (Sheather, 2009).

The goal of regression analysis is to minimize the difference between the actual observed value, y_i, and the predicted value, $\widehat{y_i}$ (Sheather, 2009). The difference between y_i and $\widehat{y_i}$ is known as the residual, $\widehat{e_i}$. A typical way of choosing the model coefficients β_0, β_1, …, and β_p is to use the least squares method which chooses the model coefficients that minimize the sum of squared residuals (RSS) (Sheather, 2009). The RSS is expressed as $\sum_{i=1}^{n} \widehat{e_i^2} = \sum_{i=1}^{n} (y_i - \widehat{y_i})^2 = \sum_{i=1}^{n} (y_i - \widehat{\beta_0} + \widehat{\beta_1} x_{1i} + \widehat{\beta_2} x_{2i} + \cdots + \widehat{\beta_p} x_{pi})^2$ (Sheather, 2009). The estimates of β_0, β_1, …, and β_p ($\widehat{\beta_0}$, $\widehat{\beta_1}$, …, and $\widehat{\beta_p}$) are found by setting the derivative of the RSS with respect to each beta parameter equal to zero and solving the system of equations.

The steps involved in creating a multiple linear regression model include variable selection, evaluation of the model assumptions, and validation of the model (Ngo, 2012, Sheather, 2009). There are several criteria for choosing the "best" set of predictor variables, known as selection criteria (Sheather, 2009). One way of selecting the "best" model is to use R^2-Adjusted as the selection criterion. R^2-Adjusted is similar to R^2, but R^2-Adjusted takes into account the number of predictors, *p*, as well as the sample size, *n* (Sheather, 2009). The "best" model is the one that maximizes the R^2-Adjusted. While a general understanding of the variables is necessary for analysts, selection criteria help point in the right direction (Ngo, 2012).

Using the selection criteria, variable selection techniques include all-subsets selection and stepwise selection (Sheather, 2009). All-subsets selection calculates the selection criterion for each possible combination of predictor variables, while stepwise analyzes a sequential subset of predictor variables (Sheather, 2009). There are three types of stepwise selection methods, forward, backward, and alternating (Bilder and Loughin, 2015). Forward selection adds predictors one at a time and selects the predictor that maximizes the selection criterion after each step. Backward selection starts with all possible predictors, removes the predictor that increases the selection criterion the most, and moves on to the next step until the removal of predictors does not increase the selection criterion. Alternating selection is a hybrid of both forward and backward selection where terms can be added in one step, but removed in another (Bilder and Loughin, 2015).

Once variables are selected, model assumptions must be evaluated (Ngo, 2012). Model assumptions are that Y is linearly related to x, the errors are independent of each other, the errors have a common variance, and the errors are normally distributed (Sheather, 2009). One method of evaluating the model assumptions is to look at diagnostic plots (Ngo, 2012, Sheather, 2009). If the model assumptions are violated, the most common way to solve this is by transforming the response or explanatory variables (Ngo, 2012). Using the natural log transformation is one type of transformation used (Ngo, 2012).

Another type of problem that causes model assumptions to not hold is multicollinearity (Ngo, 2012, Sheather, 2009). Multicollinearity is when predictor variables are highly correlated to one another (Ngo, 2012). Multicollinearity can cause non-significant variables to show up as significant or even cause parameter estimates to have opposite signs from what is expected (Ngo, 2012). One way of checking for multicollinearity is to calculate the coefficient of correlation (r) between each pair of independent variables (Ngo, 2012). If the coefficient of correlation between any two variables is close to 1 or -1, then they are highly correlated. In order to prevent any problems in the model, a solution is to remove one of the correlated independent variables. Another way of checking for multicollinearity is to calculate the variance inflation factor (VIF) for each variable in the model (Ngo, 2012, Sheather, 2009). The VIF is $\frac{1}{1-R_j^2}$, where R_j^2 is the R^2 between x_j and the other x's (Sheather, 2009). A general rule for identifying multicollinearity is if the VIF is greater than five (Sheather, 2009), however, sometimes 10 is used as the cutoff (Ngo, 2012).

The final step in MLR is model validation (Ngo, 2012, Sheather, 2009). Model validation includes examining predicted values and applying the model to a new data set and examining the model fit (Ngo, 2012, Sheather, 2009). Applying the model to a new data set is a good way of testing the validity of the final model (Sheather, 2009).

2.8 Summary

Several benchmarking efforts have been made in the past to benchmark the energy usage or intensity of wastewater treatment plants, but most have focused on large plants. An example of the problems associated with using these benchmarks on small plants is the fact that using AwwaRF's or Energy Star's benchmarking model equations on small plants produces outlying results. Few studies have focused on benchmarking the energy usage of small wastewater treatment plants. In addition to large plant models producing outlying results for small plants, the differences in small and large plants may suggest a better way to model energy usage may be to create separate models for large and small plants.

The three most common small plant types found in Nebraska are conventional extended aeration plants, oxidation ditches, and sequencing batch reactors. The energy consumption and operation of each process type differs, including how they provide oxygen to remove both carbonaceous and nitrogenous organic material, and illustrates a potential need to consider plant type as a potentially influential factor in regards to energy usage. The energy usage in all three common small plant types is mainly required to provide aeration of the wastewater. Aeration is needed to remove both CBOD and ammonia, with differing amounts of aeration needed for each. In addition to aeration, the next most energy intensive processes employed at small wastewater treatment plants are aerobic digestion and pumping.

The most common way to model the relationship between a single dependent variable and a number of independent variables is to use multiple linear regression. Multiple linear regression has been shown to be an appropriate method of benchmarking the energy usage or intensity of wastewater treatment plants as seen by the AwwaRF and Energy Star benchmarking models.

Chapter 3: Methods

3.1 Introduction

In order to perform this study, data were collected and then analyzed. The key steps included determining what data to collect, and then collecting the data using cross-trained undergraduate interns and graduate students. Once the data were collected, data analysis began and consisted of organizing the data, creating visual representations of the data, building multiple statistical models, comparing these models to determine the best or most useful ones, and validating the models that were chosen. Final conclusions were then based off these models. The following sections detail the steps taken to create the final models in this study.

3.2 Wastewater Treatment Plants Studied

In order to identify potential plants to include in this study, a list was obtained from the Nebraska Department of Environmental Quality (NDEQ) containing 268 permitted minor Publicly Owned Treatment Works (POTWs) in Nebraska with average design flows of less than 1.0 Million Gallons per Day (MGD). The list contained permit information, as well as contact information for the cognizant official and head operator of each plant. It was decided not to include lagoon systems, primarily due to their very small size, low energy usage, and considerable differences from conventional, mechanical wastewater plants.

One of the goals was to gather data from as many small, mechanical plants in Nebraska as possible serving populations of 10,000 or less and with average flows, not average design flows, less than 1.5 MGD. Additional mechanical plants were added to the NDEQ list by going through a list of Nebraska towns by population and looking at their average flows on NDEQ's online public records (NDEQ, 2016). The total number of Nebraska plants meeting the above criteria was 109.

In addition to the plants in Nebraska, the Pennsylvania Department of Environmental Protection (PDEP) was contacted about obtaining data from a similar study that they had published in 2011 (PDEP 2011). Their list contained energy usage data and plant characteristics for 81 Pennsylvania plants with average flows less than 1.5 MGD. While the same data were not collected for the Pennsylvania plants, this data set was used for comparison to Nebraska plants.

3.3 Energy Usage Determination

Plant energy usage was determined by reviewing the utility bills for each plant. The three main fuel sources used at small Nebraska plants included electricity, natural gas, and propane. Towns were solicited over the phone to provide utility bills for their wastewater treatment plants. Utility bills were either obtained from town clerks or directly from a town's utility provider. A detailed description of how utility bills were obtained is provided in Appendix A under the title, "Pre-Assessment Guide." Multiple years of energy usage were obtained for most plants to determine if energy usage had changed drastically between the baseline year of 2015 and previous years. Table 3.1 provides the percentage of plants out of the original 109 that provided multiple years of utility bills. Information recorded from utility bills included usage, electric demand, meter read dates, and meter numbers.

Table 3.1: Multiple Years of Utility Bills

Number of years of utility bills obtained	**Percentage of Plants**
1 year	91%
> 1 year, but < 2 years	43%
> 2 years	18%

3.4 Wastewater Plant Characteristic Determination and Data Collection

Plant data collected included energy usage, plant characteristics and processes, influent and effluent water quality data, climatic information, information on how the plant is operated, and other plant information related to energy usage. The information collected was similar to what had been collected for previous energy benchmarking studies performed by ENERGY STAR and AwwaRF (ENERGY STAR, 2014, Carlson and Walburger 2007). Table 3.2 lists the data collected from Nebraska plants and the main source of the data. Data for each plant were recorded in the Wastewater Facility Energy Use Assessment Forms and Assessment Spreadsheets, examples of both are provided in Appendix B.

A detailed description of how data were collected and where it was found before visiting a plant is provided in Appendix A under the title, “Pre-Assessment Guide.” The Pre-Assessment Guide details the information that can be acquired before visiting the wastewater plants.

Table 3.2: Data Collected and Sources of Data

Data Category	Characteristic	Main Source of Data
Community Information	Population	Nebraska League of Municipalities
	Facility Address, AWIN Score	NDEQ[1]
	Contact Information	NDEQ[1] or Operator
Flow Data	Average Design Flow, MGD	NDEQ[1]
	Average Flow, MGD	ECHO[2]
Water Quality Parameters	Average Influent $CBOD_5$, mg/L Average Effluent $CBOD_5$, mg/L Average Influent TSS, mg/L Average Effluent TSS, mg/L Average Effluent NH3N, mg/L	ECHO[2]
	Lowest NH3N Discharge Limit, mg/L	NDEQ[1]
Climate and Building Information	Annual Sum of HDDs, Annual Sum of CDDs	NOAA[3]
	Climate Controlled Floor Area, sq. ft.	Google Earth
Energy Usage Data	2015 Electric Usage (kWh), 2015 Natural Gas Usage (therms), 2015 Propane Usage (gallons)	Utility Bills
Plant Type	Oxidation Ditch, Seq. Batch Reactor, Extended Aeration, Trickling Filter, Other	NDEQ[1]
Secondary Treatment Aeration Type	Fine Diffusers, Course Diffusers, Mechanical Aerators	NDEQ[1] or Operator
Industrial User Data	Total Industrial Flow (MGD), Total Industrial Loading (lbs-CBOD)	NDEQ[1]
Sludge Management	Aerobic Digestion, Anaerobic Digestion, Lime Stabilization, Supplemental Energy Use for Sludge, Dewatering Equipment Type, Sludge Disposal Type	NDEQ[1] or Operator
	Biogas Usage	Operator
Operator Information	Number of Operators, Years of Experience, Recent Staff Changes	Operator
Other Factors/Equipment	Presence of Automatic DO Controls, Presence of Variable Frequency Drives, Disinfection Type	NDEQ[1]/Operator/ Site Visit
	Types of Pumps and Blowers Used	Operator/Site Visit

1. NDEQ (NDEQ, 2016)- http://deq.ne.gov/NDEQProg.nsf/OnWeb/PRS
2. ECHO (EPA, 2016)- https://echo.epa.gov/
3. NOAA (NOAA, 2016)- http://sercc.com/nowdata.html

As a way to validate information from online resources, as well as to gather further information, site visits and operator interviews were conducted to better improve the quality of the data. Due to time constraints, only 89 plants were visited during the summer of 2016, but additional plants were solicited over the phone for any missing data. With these additional plants, 95 of the 109 plants originally targeted provided the requisite data to be included in model creation.

3.5 Data Compilation and Confirmation

After each site visit, plant data were compiled into a master spreadsheet. A site visit narrative was written for each plant that summarized the visit, gave additional information about the plant not noted on the assessment forms, and provided any energy efficiency opportunities the assessor observed during the visit. The master spreadsheet can be found in Appendix G. An example site visit narrative is provided in Appendix C. Each assessment form and site narrative was peer reviewed by the undergraduate and graduate students collecting the data for this study.

After compiling the data, the actual observed Energy Intensities (EI) were calculated for each plant for the baseline year of 2015. EI was calculated by taking the annual energy usage in megawatt-hours (MWh) for all fuel types consumed at the plant divided by the amount of wastewater treated in 2015 in millions of gallons (MG). The median EI's by plant type are listed in Table D.1 in Appendix D for both the Nebraska and Pennsylvania plants. Median Observed Electric Intensities (EI_e) were calculated in a similar manner as EI, but with electric use only instead of overall energy use. The EI_e's of both Nebraska and Pennsylvania plants are listed in Table D.2 in Appendix D. Note that only electric usage was obtained for Pennsylvania plants and that all data for the Pennsylvania plants were collected in 2008.

3.6 Modeling Background

Multiple linear regression (MLR) methods were used to model this data. MLR is one of the most widely used statistical procedures employed when modeling the relationship between one dependent variable, Y, and two or more predictor variables X_1, $X_2, \ldots, X_p$ (Sheather, 2009). The general model equation is of the form:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + \varepsilon_i$$

where i is the number of observations, ε_i is random error in Y_i, and the response variable Y is predicted from p predictor variables $X_1, X_2, \ldots, X_p$ with the relationship between Y and $X_1, X_2, \ldots, X_p$ being linear in the parameters $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ (Sheather, 2009).

For this study, MLR was utilized to model plant data with the intent of predicting the Electric Intensity (MWh/MG), Energy Intensity (MWh/MG), Annual Electric Usage (kWh/year), and Annual Energy Usage (kWh/year) of small Nebraska wastewater plants. ENERGY STAR and AwwaRF have shown that MLR is an appropriate method for modeling the energy/electric intensity and usage of wastewater treatment facilities (ENERGY STAR, 2014, Carlson and Walburger, 2007). The basic approach to MLR is to determine the combination of predictor variables that best predict the response variable, while at the same time, not violating model assumptions.

The first model assumption is that the response is linearly related to the predictor variables. The other assumptions are that the errors are independent, normally distributed, and have a constant or common variance (Sheather, 2009). Models are created and then assessed on whether or not they violate the model assumptions listed above. Model assumptions are evaluated by observing diagnostic plots as well as evaluating outliers and observations with high leverage. When model assumptions do not hold, transformation of the response or predictor variables can result in a valid model. Removal of outliers and highly influential observations can also result in a more valid model if these observations are truly different from other observations in the data set.

While there is little agreement in statistical literature on what defines a model as "best" (Sheather, 2009), a common method is to choose the model that maximizes R^2-Adjusted. The Adjusted-R^2 is just the R^2, the proportion of the total sample variability in the *Y*'s explained by the regression equation, but with an adjustment added in for the number of predictors in the model so that irrelevant predictor variables are not included in the final model (Sheather, 2009). The final step, once a model is chosen, is to test its predictive ability by fitting the model to a new or test data set.

3.6.1 Modeling Approach

Model formulation was performed using SAS 9.4 (SAS Institute Inc., 2014). The SAS code used for modeling can be found in Appendix E. The first part of model formulation is fitting all predictor variables in the model. Checking the correlations in the Nebraska data set of all predictors to both the response and all of the other predictors showed high correlations between many of the predictors. These high correlations revealed redundancy in some of the potential predictor variables. Because of these high correlations, many of the predictor variables were removed from the data analysis in order to reduce the level of multicollinearity. The results of the correlation analysis of the Nebraska data set were used to delete any redundant predictor variables in the Pennsylvania data set.

Removing many of the highly-correlated factors helped narrow down the number of potential predictors. Other potential predictors were deleted due to poor quality of data, small sample size, missing data, or subjective rankings. A list of the final set of potential predictor variables for the Nebraska data set, as well potential predictors that were removed from the analysis and explanations as to why each one was removed is provided in Appendix G (Tables G.1-G.14).

The final number of potential predictors in the Nebraska data set was narrowed down to 25 variables. In comparison, the final number of potential predictors for the Pennsylvania data set was eight. For the Nebraska data set, 11 of the 25 potential predictor variables were continuous variables and the rest were binary (e.g., 1 = Yes, 0 = No). Examining the distributions of the continuous variables, almost all of the variables were skewed right and therefore required transformation. When data are skewed right, Velleman and Hoaglin (1981) suggest using the log transformation in order to obtain normality. The log transformation was performed for all of the continuous variables except for HDDs and CDDs. Depending on the data set, the CDDs and HDDs were transformed using a power transformation of either two or three. After transformation, the distributions of the continuous variables were approximately normally distributed.

The full model was run once more, now with the final list of potential predictors. Not all factors were found to be significant (p-value < 0.05) in any of the full models and the R^2 values were quite low. In order to increase the R^2 and get a better fit, variable selection was run using a stepwise selection method. The stepwise selection option in SAS uses an approach that is a combination of forward and backward stepwise selection. Factors can be added and taken away in multiple steps of the model creation based on each variable's *F* statistic. The process ends when none of the variables outside of the model have an *F* statistic that is significant at the stated significance level and all of the variables in the model do have *F* statistics at the stated significance level (SAS Institute Inc., 2014). The default significance level for the stepwise selection method in SAS is 0.15 for both staying in the model and for entry into the model. The significance level was not changed in order to evaluate which variables were approaching significance. Evaluating the p-values of the remaining variables after stepwise selection, variables were removed from the model if their p-values were above 0.05. The remaining predictors were found to be significant at the 95% confidence level and diagnostic plots were examined. SAS output and diagnostic plots can be found in Appendix E for each respective model. Model comparisons and fits were evaluated by examining each model's R^2-Adjusted, as well as the Root Mean Squared Errors (RMSE).

Examining the diagnostic plots revealed several outliers and points of high leverage. Plants indicated as being outliers were further investigated. Some of the plants investigated were dropped from the model due to either questionable flow data or outlying energy usage due to plant type. Some of the smaller fixed film plants had extremely low energy usage and were deemed different enough from the rest of the plants in the data set to warrant removal from the analysis. Once the outliers had been removed, variable selection was rerun and models were finalized.

To test the predictive ability of the models, a test data set was fitted to the model. Since utility bills and other plant information that changes year to year was collected for multiple years, the test data set was the data for plants from the year 2014. Interpretation and discussion of the final models can be found in the Results and Discussion chapter.

3.6.2 Types of Models Created

There were 4 types of models created. The first group of models predict electric intensity (MWh/MG), which normalizes annual electricity use by volume of flow treated. The second group of models predict annual electric usage (kWh/year). The third group of models predicts energy intensity (MWh/MG) which has the same units as electric intensity, but takes into account energy usage from all fuel types at a plant (electricity, natural gas, and propane) and normalizes the annual energy usage by the amount of flow treated. The final group of models predict annual energy usage (kWh/yr). Annual energy usage considers all fuel types used at a plant, but is not normalized by flow.

Within the three groups of models, models were created by plant type (extended aeration, oxidation ditch, and sequencing batch reactor) and by state (Pennsylvania and Nebraska). Comparisons were made between the models and interpretations of the results are discussed in the following chapter.

Chapter 4: Results and Discussion

4.1 Introduction

To aid the understanding of energy use by small wastewater utilities, benchmarking models were developed. These models predict Electric Intensity, Electric Usage, Energy Intensity, and Energy Usage for small Nebraska Wastewater treatment plants and are presented in this chapter, along with model interpretations, general discussion of the results, and a comparison of these small Nebraska plants to a set of small Pennsylvania wastewater plants. The models help explain much of the variability in energy and electric intensity and usage among small Nebraska wastewater plants. This variability stems from plant characteristics such as flow, plant type, equipment used, and several other characteristics.

Data were collected from 95 wastewater treatment plants in Nebraska treating less than 1.5 MGD. Further analysis and investigation led to the removal of outliers and certain fixed film plants. Outliers were removed due to missing data (1 plant), flow data that were determined to be erroneous (3 plants), or incorrect utility bills (2 plants).

Outliers removed due to incorrect utility bills included one community that provided the electric bills for the community's drinking water wells and another community for which a drinking water well was connected to the same electric meter as its wastewater plant. The correct energy consumption for these two plants was not collected due to time constraints. After removing the five outlying plants, as well as seven fixed film plants, the final number of plants included in the Nebraska energy intensity models was 83.

4.2 Removal of Fixed Film Only Plants from the Models

Analyzing the data collected and creating many different models led to the decision that the seven solely fixed film secondary treatment plants should be left out of the overall models because they used significantly less energy than the other plants. The median Energy Intensity of these seven fixed film plants was 2.33 MWh/MG. Several sources have documented that fixed film plants use significantly less energy than plants with aeration systems for secondary treatment (Tchobanoglous et al., 2014, EPRI, 2013). Six fixed film plants had a combination of fixed film and suspended growth secondary treatment (pre-or-post aeration, aerated filters, etc.) and were left in the data set because their energy consumption was similar to the other plants. The median Energy Intensity of these six combination fixed film plants was 5.23 MWh/MG, similar to the overall median Energy Intensity of the other plant types, which was 5.47 MWh/MG.

4.3 Importance of Flow and Percent Design Capacity

Much of the variation in energy usage between plants may be explained by differences in flow, as well as where plants run in regards to percent design flow, ($\frac{\text{Average Daily Flow, MGD}}{\text{Average Design Daily Flow, MGD}}$). Figures 4.1(a) and 4.1(b) show the strong relationship between average flow (MGD) and annual energy usage (kWh/year) and average flow and annual electric usage (kWh/year), respectively. According to the R^2 value shown in Figure 4.1(a), 74% of the variation in annual energy usage among small Nebraska plants can be explained by differences in flow. The R^2 value seen here is similar to what was found in AwwaRF's Energy Index Development study (Carlson and Walburger, 2007) of large plants, which found the R^2 value between average flow and annual energy usage to be 0.82.

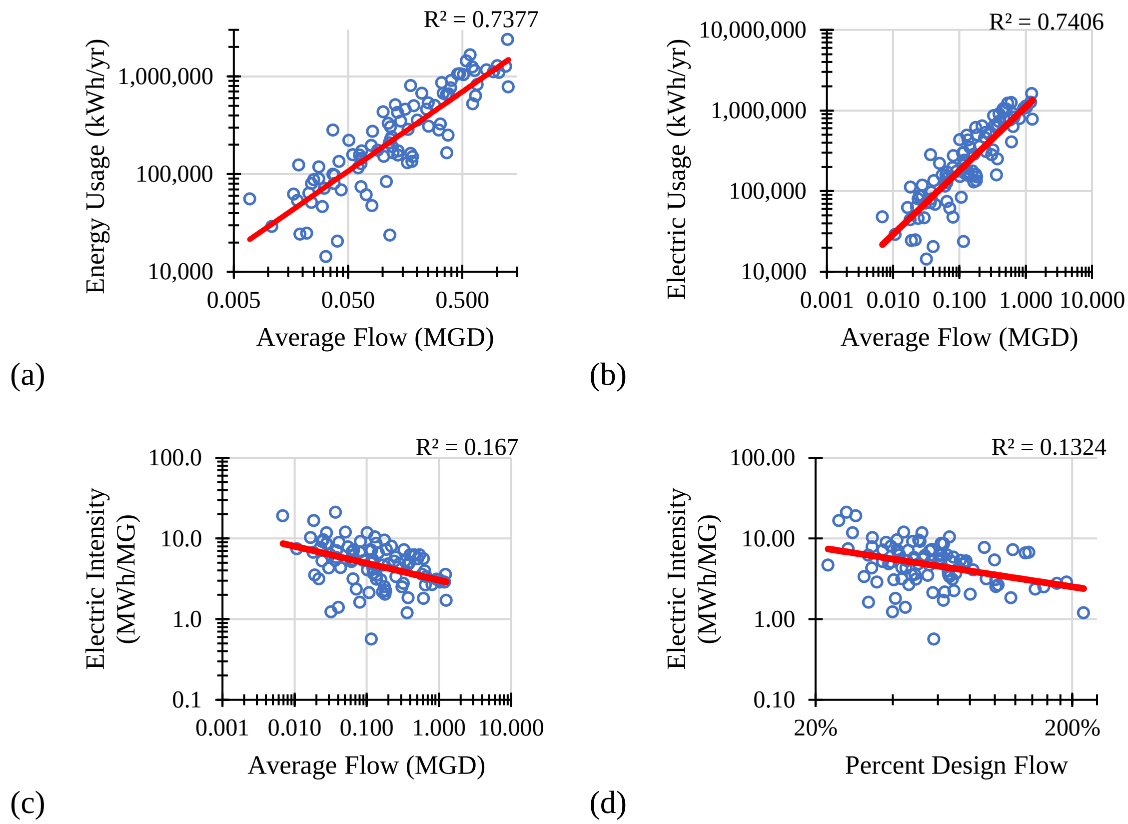


Figure 4.1: Flow and Percent Design Flow vs Electric Intensity and Energy/Electric Usage. (a) Flow *vs.* Annual Energy Usage, (b) Flow *vs.* Annual Electric Usage, (c) Flow *vs.* Electric Intensity, and (d) Percent Design Flow *vs.* Electric Intensity.

The relationship between average flow and Electric Intensity (MWh/MG) was not nearly as strong, but this is because the response variable is electric usage per million gallons of wastewater treated. Therefore, the explanatory power of flow is hidden in the response variable and not included in the R^2 value seen in Figure 4.1(c) (ENERGY STAR, 2014). There is, however, still a clear relationship between intensity and flow; as flow increases, intensity decreases.

Percent design flow is also an important factor in regards to energy consumption, especially at small plants. Figure 4.1(d) shows that as percent design flow increases, intensity decreases. This can be attributed to the fact that many of the plants in the study were underloaded in comparison to their design load. Young and Koopman (1991) observed that many of the motors at small wastewater plants are underloaded, which leads to low energy efficiencies. The percent design flow affects the efficiencies of all the equipment at the plant with lower percent design flows leading to higher energy intensities. Young and Koopman (1991) also observed that many small plants have equipment that runs at full bore regardless of what flow is coming into the plant. This was also seen for small Nebraska plants and may be attributed to the fact that the plants are not able to "turn down" or adjust the treatment processes as the flow or loading change because they lack either the equipment, such as automatic DO controls, or are not manually operated by frequently adjusting settings to an optimum level, which therefore leads to inefficiencies.

The trends observed in Figure 4.1 can be explained by the concept of the economy of scale. The Energy Star and AwwaRF models for larger plants both indicate an economy of scale that shows on average, larger plants are more energy efficient on a per unit volume of water treated basis (ENERGY STAR, 2014 and Carlson and Walburger, 2007). In addition, based on anecdotal evidence from plant operators and various city clerks, larger plants are able to invest more money in their wastewater plants and can therefore afford newer, more efficient equipment such as variable frequency drives (VFDs) and automatic DO controls. Although flow and percent design flow accounted for much of the variation in energy usage between plants, the goal of this study is to determine additional factors that might influence energy usage at small plants.

4.4 Electricity Models

In most cases, electricity is the main form of energy used by municipal wastewater treatment plants. Therefore, two sets of electricity models were created that predict two different metrics. One subset of the electricity models predicts a plant's electric intensity (MWh/MG), while the other subset predicts a plant's annual electric usage (kWh/year). For the Nebraska data set, 75% (62 of 83) of plants use electricity as their sole energy source. For half of the plants that use additional sources of energy (natural gas and propane), electricity accounts for 90% or greater of their total energy use.

Electricity models were investigated not only because most Nebraska plants only use electricity, but also because no natural gas or propane usage was recorded for the Pennsylvania data set. An appropriate comparison of plants from the two states could only be done if they were compared based on electric consumption. The following subsections detail the Electric Intensity and Electric Usage models, as well as comparisons between the Nebraska and Pennsylvania plants.

4.4.1 Electric Intensity Models

One way of considering electricity use is by dividing the annual electricity use by the annual volume of wastewater flow. This is called electric intensity (MWh/MG). Dividing by flow normalizes the electricity use. This was done because even though the focus of the study was on plants treating 1.5 MGD or less, there was still a large range in average flow (0.01-1.3 MGD). Using Electric Intensity (MWh/MG) as the response variable, models were created using a stepwise multiple linear regression approach.

The variables that were found to be significant at the 95% level (p-value ≤ 0.05) included both binary and continuous variables. When using the models, binary variables are assigned a 1 if the plant has a certain characteristic described by the respective binary variable or a 0 if the plant does not. The binary variables found to be significant for at least one of the electric intensity models (with their abbreviations in parentheses) included (1) Extended Aeration plant type (EA), (2) Supplemental Energy Usage for Sludge Treatment (SE Sldg), (3) Aerobic Digestion (Aer Digest), (4) Dewatering Equipment (DWE), (5) Variable Frequency Drives (VFDs), (6) UV Disinfection (UV), (7) Industrial Loadings (Ind. Load), and (8) a plant's state (Pennsylvania = 1, Nebraska = 0).

The continuous variables that were found to be significant for at least one of the electric intensity models (with their abbreviations and units in parentheses) included (9) Climate Controlled Floor Area (CCFA, ft^2), (10) Daily Average Flow (Ave. Flow, MGD), (11) Percent Daily Design Flow (% Design Flow, $\frac{\text{Average Daily Flow, MGD}}{\text{Average Design Daily Flow, MGD}}$), (12) Average Influent $CBOD_5$ or BOD_5 (mg/L), (13) Average Effluent NH_3-N (mg/L), and (14) Percent Design BOD Loading ($\frac{\text{Average Daily BOD Load,}\frac{\text{lbs}}{\text{d}}}{\text{Average Daily Design BOD Load,}\frac{\text{lbs}}{\text{d}}}$).

The regression models for Electric Intensity are listed in Table 4.1. The table lists the intercepts, coefficients of the explanatory variables, and regression statistics for each of the models created. Explanatory variables with positive coefficients predict an increase in electric intensity, while variables with negative coefficients predict a decrease in electric intensity. Most of the continuous variables, along with the response variable, Electric Intensity, were transformed using the natural log (ln) and are denoted by an asterisk.

Table 4.1: Electric Intensity Model Coefficients

Model Description	Intercept	Binary Variables (Yes = 1, No = 0): Plant Type	Binary: Sludge Management			Binary: Equip.	Binary: Disinfect	Binary: User Type	Binary: State	Continuous Variables: Bldgs	Continuous: Flow		Continuous: Water Quality			Regression Statistics		
		EA	SE Sldg	Aer Digest	DWE	VFDs	UV	Ind. Load	Penn	CCFA	Ave. Flow	% Design Flow	Inf. $CBOD_5$ or BOD_5	Eff. NH_3-N	% Design BOD Load	n	R^2-adj	RMSE
Overall NE	-2.06	0.257	0.264		0.272					0.165*	-0.323*	-0.268*	0.256*			71	0.52	0.36
NE EA	-1.52									0.320*	-0.444*					34	0.36	0.41
NE OD	-1.22		0.417			-0.404				0.233*	-0.420*			-0.167*		25	0.71	0.27
NE SBR^	0.425		1.24					1.83				-6.04*		-0.536*		13	0.94	0.53
Overall Penn	-2.38						0.222				-0.316*		0.632*		-0.310*	71	0.74	0.39
Penn & NE	-1.86	0.182		0.237					0.200		-0.226*	-0.328*	0.500*			141	0.64	0.40
Penn EA	-0.762						0.495				-0.211*	-1.77	0.585*			25	0.77	0.37
Penn SBR	-2.48										-1.57		0.872*			19	0.73	0.40

*Variables that were transformed using natural log (ln) transformation.

^Response Variable (Electric Intensity) for NE SBR Model not transformed. All other response variables for the other models transformed using the natural log (ln) transformation.

Abbreviations:

EA: Extended Aeration

SE Sldg: Supplemental Energy Usage for Sludge Treatment

DWE: Dewatering Equipment

Aer Digest: Aerobic Digestion (Only for data sets that include Pennsylvania data)

CCFA: Climate Controlled Floor Area

RMSE: Root Mean Squared Error

Overall models were created for both Nebraska and Pennsylvania, as well as plant type specific models for each state. For further comparison, a combination model was created that combined data from the two states.

The same data were not collected as part of the studies from Nebraska and Pennsylvania. The Nebraska data set included information on the presence of VFDs and the use of Supplemental Energy Usage for Sludge Treatment. The variable for Supplemental Energy Usage for Sludge Treatment denotes plants that use either aerobic digestion or heated anaerobic digestion that requires energy input in addition to the use of biogas from the digester. The Pennsylvania data set included an aerobic digestion variable because there was no heating information collected for the anaerobic Pennsylvania plants. Another difference in the data sets was that influent $CBOD_5$ was collected for Nebraska, while influent BOD_5 was collected for Pennsylvania.

These models provide insights to key factors influencing electricity use in small wastewater plants. The models varied from having 2 to 7 explanatory variables, with many of the same variables showing up in multiple models. Two of the variables seen in most of the models, average flow and percent design flow, have shown up in past models of larger wastewater plants (Carlson and Walburger, 2007, Mizuta and Shimada, 2010), and was illustrated in Figure 4.1.

Extended aeration plant type shows up for the overall Nebraska model with a positive coefficient. Out of the three main small plant types in Nebraska for this data set, extended aeration plants were the majority (34 of 83), and are more energy intensive than other plant types used for small communities (Tchobanoglous et al, 2014, EPA, 2000). One can also see in Table D.1 in Appendix D that the observed median electric intensities for extended aeration plants for both Nebraska and Pennsylvania are much higher than the other plant types.

Supplemental Energy for Sludge Treatment is included in most of the Nebraska models. This variable included both plants that use aerobic digestion and plants that use supplemental energy (both natural gas or electric) to heat anaerobic digesters. Aerobic digestion is a large energy user at wastewater plants (EPRI, 2013). Gretzschel *et al.* (2014) found that even under the best circumstances, anaerobic stabilization is only economically feasible for plants serving population equivalents greater than 7,500, in part since it is difficult to produce sufficient biogas to self-heat at smaller scales. Only 4 of the 83 plants in the study are above this population level. The SE Sldg variable did not show up as significant in the NE EA model, likely since most (68%) of the EA plants employ aerobic digestion for sludge treatment.

Dewatering equipment was another variable with a positive coefficient that only showed up in the Nebraska overall model. Only larger plants tended to have dewatering equipment, but a comparison of plants larger than 0.2 MGD with and without dewatering equipment showed a median difference in electric intensity of 2.26 MWh/MG. While dewatering equipment is not a major consumer of energy at large plants, it still has somewhat of an impact (EPRI, 2013). This impact may be greater for smaller plants leading to greater electric intensity.

Influent organic loading ($CBOD_5$ or BOD_5) has been shown to significantly impact energy usage in previous studies (Carlson and Walburger, 2007, PDEP, 2011). It shows up in the electric intensity models with a positive coefficient for Nebraska and

Pennsylvania plants. According to the models, a higher influent organic loading is predicted to increase electric intensity.

One of the variables that was found in this study to be somewhat unique for small plants was climate controlled floor area. CCFA was only collected for Nebraska plants and has a positive coefficient. One of the reasons why this variable is significant is because of the great variability in the data set. The minimum CCFA was zero, while the maximum was approximately over 21,000 ft^2. An explanation as to why there was such a large difference in CCFA between the plants is that several small communities in Nebraska had maintenance/storage garages on-site, connected to the one and only electric meter at the plant, that housed equipment for all the utilities in town. Other communities stored their utility equipment elsewhere. The energy used to heat these garages could not be differentiated easily, so CCFA was included to account for these differences.

Several variables in the Nebraska models were only found to be significant in one or two of the plant type models. These include VFDs, industrial loadings, and average effluent NH_3-N. VFDs showed up for the oxidation ditch model with a negative coefficient and have been shown to decrease energy use at oxidation plants when the VFDs are connected to the aeration blower or rotary surface aerators and adjusted appropriately (DOE, 2012). The industrial loadings variable shows up for the SBR model because SBRs are most often used for communities with large variations in flow or organic loading, such as the case when a community has a significant industrial user (EPA, 2000, Tchobanoglous *et al.*, 2014).

Another variable that only showed up in the plant type models was average effluent NH_3-N. In AwwaRF's Energy Benchmarking Development study (Carlson and

Walburger, 2007), a similar binary variable for the presence of nutrient removal was included in their model as well. The coefficient for the Nebraska electric intensity models is negative, meaning an increase in the effluent NH_3-N, leads to a decrease in the electric intensity. This is consistent with AwwaRF's model because the coefficient in their model was positive, meaning if a plant had nutrient removal, they saw an increase in energy usage/intensity. While the presence of a decrease in the amount of ammonia is not necessarily indicative of nutrient removal, the variable in Nebraska models quantifies the amount of electricity per million gallons of flow for a decrease or increase in the effluent NH_3-N.

4.4.2 Electric Usage Models

The electric usage models predict the annual electric usage of small wastewater plants. The electric usage models resulted in the same variables as the electric intensity models, with the same variables being significant for the same type of model (e.g., overall Nebraska, NE EA, etc.). The only difference is the model coefficients and R^2 values. The model coefficients and regression statistics for the electric usage models are listed in Table 4.2. The R^2 values for the electric usage models are generally much higher than their electric intensity model counterparts. This is due to the fact that the response variable is not normalized for flow and the explanatory power of flow is no longer hidden in the response variable. In addition, total electric use is roughly proportional to flow. Thus, by having a wide range of flow rates, and electric use values proportional to flow, a higher R^2 value is obtained.

Table 4.2: Electric Usage Model Coefficients

Model	Intercept	Binary Variables (Yes = 1, No = 0)							Continuous Variables						Regression Statistics		
		Plant Type	Sludge Management		Equip	Disinfect	State	Bldgs	Flow		Water Quality						
		EA	SE Sldg	DWE	VFDs	UV	Penn	CCFA	Ave Flow	% Design Flow	Eff. NH_3-N	Inf. $CBOD_5$ or BOD_5	% Design BOD Load	n	R^2-Adj	RMSE
NE Overall	10.8	0.257	0.264	0.273				0.164*	0.677*	-0.268*		0.256*		71	0.88	0.36
Penn Overall	10.4					0.222			0.684*			0.632*	-0.306*	71	0.85	0.39
NE & Penn	10.9	0.182	0.236				0.200		0.774*	-0.328*		0.500*		141	0.86	0.40
NE EA	11.3							0.320*	0.556*					34	0.85	0.41
NE OD	11.6		0.416		-0.400			0.231*	0.581*		-0.170*			25	0.89	0.27

*Variables transformed with the natural log (ln).

The specific plant type models for the two main Pennsylvania plants types, EA and SBR, as well as the NE SBR model, were not created for predicting electric usage because the distributions of the electric usage for these plant types was non-normal and required more complex transformations that are not easy to interpret.

4.4.3 Comparison to Small Pennsylvania Wastewater Treatment Plants

The Pennsylvania Department of Environmental Protection (PDEP) conducted a study similar to this one (PDEP, 2011). The study examined the electric intensities of 117 wastewater plants and collected much of the same types of data as the Nebraska study. However, the study did not develop any benchmarking models. PDEP shared the data which were modeled in a similar manner as the Nebraska data. Of the 117 plants in the Pennsylvania study, 71 met the same criteria as this study of having average flows less than 1.5 MGD, no fixed film plants, being mechanical, secondary treatment plants (no lagoon systems); and having no missing data. These 71 plants were used to create the Pennsylvania models listed in Table 4.1.

The Pennsylvania plants and models were generally similar to the Nebraska models, with a few notable differences. Average flow, percent design flow, and influent BOD_5 were highly significant in many of the Pennsylvania models. For both states, extended aeration plants were highest in electric intensity and fixed film plants were lowest in intensity. The Pennsylvania plants seemed to be slightly higher in electric intensity on average. This difference can be shown statistically in the combined Nebraska and Pennsylvania model in Table 4.1.

Looking at the variables found to be significant for the Pennsylvania models, one sees that UV disinfection is included only in the Pennsylvania models. This may be due to the fact that only 32% of the Pennsylvania plants use UV, while 81% of Nebraska plants do. The greater variety of disinfection types used may have allowed UV to stand out in the Pennsylvania data, but not in the Nebraska data.

Another difference between the two sets of models was that influent organic loading (e.g. BOD_5 concentration) was found to be significant in all of the Pennsylvania models, but it was only significant in the overall Nebraska model. The difference here may be due to differences in data quality. It was confirmed from PDEP officials who worked on the study that influent organic loading was, on average, sampled more frequently than the Nebraska influent organic loading. At the time of the study, Pennsylvania plants were generally required to sample influent BOD_5 once per month, while almost all of the Nebraska plants were required to sample influent $CBOD_5$ only once per year.

Another reason for this difference could be the fact that influent organic loading for Pennsylvania was measured as BOD_5, while influent organic loading for Nebraska was measured as $CBOD_5$. The difference between $CBOD_5$ and BOD_5 is that $CBOD_5$ does not take into account nitrogenous biochemical oxygen demand (NBOD) (Tchobanoglous *et al.*, 2014). This could explain why effluent ammonia-nitrogen was found to be significant in some of the Nebraska models, while influent $CBOD_5$ was not. The influent and effluent $CBOD_5$ was less variable than the effluent ammonia-nitrogen. Albertson (1995) suggests that $CBOD_5$ is an improper test for influent wastewater because it understates the true strength of the wastewater by 20-40%. With 65 of the 83 (78%) Nebraska plants having ammonia-nitrogen limits, even though most of the plants in the data set not being designed for nutrient removal, means that plants are increasing their aeration in order to meet their permit limits for ammonia-nitrogen and therefore using more energy. The ammonia-nitrogen being treated may be masking the relationship between influent $CBOD_5$ and energy usage.

The combined model used plants from both the Nebraska and Pennsylvania data sets and included a binary variable for what state the plant was in (Pennsylvania = 1, Nebraska = 0). This state variable was found to be significant with a positive coefficient, indicating that the Pennsylvania plants were significantly higher in electric intensity. While this does not determine why the plants were more electric intensive, it points out that there are some underlying differences. The differences between the models show that these types of benchmarking models may be more region or even state specific. A more detailed investigation may pinpoint the key differences more concisely, but this was beyond the scope of this study.

4.5 Overall Energy Models

The Energy models predict the overall energy intensity or usage of all fuel types used on site at a wastewater plant. As stated previously, the main energy source for plants in this study was electricity, with 75% (62 of 83) using strictly electricity. The other 21 plants, however, used natural gas or propane to heat buildings and/or anaerobic digesters. Some plants used a significant amount of natural gas or propane, with 4 plants having 20% or more of their total energy use coming from sources other than electricity. Taking into account energy use from all fuel sources allows one to make a fair comparison between plants that do and do not use other fuel sources besides electricity. The two metrics used for response variables for the energy models include Energy Intensity (MWh/MG) and Annual Energy Usage (kWh/year).

4.5.1 Energy Intensity Models

Following the same form as Tables 4.1 and 4.2, Table 4.3 on the following page lists the Energy Intensity model coefficients. The models include much of the same variables seen in the electric models. These models show a relatively good fit to the data, but not as good as the electric models. The models include one overall model that includes all plants in the study, while the other three models are for the three main plant types (EA, OD, and SBR). Models for Pennsylvania were not created using Energy Intensity because natural gas and propane usage was not collected for these plants.

One of the main differences in the Energy models versus the Electric models is that average influent $CBOD_5$ showed up as significant in the Electric models, but not for the energy models. The p-value for $CBOD_5$, when it was forced into the overall Nebraska energy intensity model, was only 0.13. This could be due to the fact that $CBOD_5$ is treated mainly by equipment that strictly uses electricity. The addition of natural gas and propane use into the model may have weakened the relationship between energy usage and influent organic loading. As stated previously, it could also be due to the fact that ammonia may be having a greater impact on energy usage than $CBOD_5$.

Table 4.3: Energy Intensity Model Coefficients

Model Description	Intercept	Binary Variables (Yes or No)					Continuous Variables*				Regression Statistics			
		Plant Type	Sludge Management		Equip	User Type	Bldgs	Flow		Water Quality				
		EA	SE Sldg	DWE	VFDs	Ind. Load	CCFA	Ave. Flow	% Design Flow	Eff. NH_3-N	n	R^2-adj	R^2	RMSE
NE Overall	-0.950	0.251	0.254	0.316			0.200	-0.327	-0.288		83	0.49	0.52	0.37
NE EA	-1.58						0.326	-0.462			34	0.38	0.41	0.41
NE OD	-1.16		0.426		-0.411		0.232	-0.397		-0.166	25	0.69	0.76	0.28
NE SBR	0.564		0.199			0.484			-1.35	-0.139	13	0.96	0.97	0.10

*All continuous variables transformed using the natural log (ln) transformation.
-All response variables transformed using the natural log (ln) transformation.
Abbreviations:

EA: Extended Aeration Plant
SE Sldg: Supplemental Energy Usage for Sludge Treatment
DWE: Dewatering Equipment
Ind. Load: Industrial Loadings
CCFA: Climate Controlled Floor Area
RMSE: Root Mean Squared Error

4.5.2 Energy Usage Models

The Energy Usage models include the same variables as the Energy Intensity models. Much like the differences between the Electric Intensity and Electric Usage models, the only differences between the Energy Intensity and Energy Usage models are the values of the coefficients and the R^2 values. The model coefficients and regression statistics for the Energy Usage models are listed in Appendix E. Just as before, the R^2 values are higher for the energy usage models than the energy intensity models due to the difference in the response variables used for each set of models.

4.6 Model Comparison Between Years

In order to confirm the sensitivity of the Nebraska model between different years of data, a comparison between years was made. Of the 95 plants data were collected for in this study, only 46 (48%) provided utility bills for both 2015 and 2014. The baseline year for the models mentioned in the previous sections was 2015. After strictly fixed film plants were removed, data from 39 plants remained for the analysis. Other data that changes from year to year, such as water quality data, was collected for the appropriate time periods through online resources such as ECHO and NDEQ discharge monitoring reports. The overall Nebraska energy intensity model (Table 4.3) was used for comparison because it did not include influent CBOD as an explanatory variable since the average influent CBOD used in the 2015 models was an average of at least the past 2 or 3 years due to the fact that sampling of influent data for most plants is once per year.

A model was created combining data from both 2014 and 2015. A binary variable for year (2014 = 1, 2015 = 0) was included along with the same variables as before and stepwise selection was employed to develop the model. The binary variable was not found to be significant (p-value > 0.05), indicating that the energy intensity between years did not change significantly. The average percent change in plant energy usage between 2014 and 2015 was only ±7%. However, the average change in energy intensity between 2014 and 2015 was ±21%. This means there were larger changes in the reported flow rate, potentially due to inflow and infiltration during wet years and/or imprecision in flow measurement.

For further analysis, the 2015 energy intensity model was next tested using plant data from 2014. Using the 2015 model variables and fitting the model to the 2014 data resulted in only 2 of the 6 variables showing up to be significant at the 95% level (p-values < 0.05). The 2 variables were Extended Aeration plant type (EA) and Dewatering Equipment (DWE). Two variables approaching significance included average flow (p-value = 0.09) and percent design flow (p-value = 0.14).

Testing the model with another year of data resulted in the conclusion that the model is less than ideal. While using a test data set is seen as a good way of evaluating the performance of regression models, small sample sizes of test data sets does not work well for evaluating the performance of regression models (Sheather, 2009). The small sample size of the 2014 data set is not representative of the overall population of small wastewater plants in Nebraska and is therefore not suited to evaluate the performance of the models developed in this study. A larger data set of 2014 data may better evaluate the performance of the models, but this is for future studies. In the absence of a firm evaluation of the models, the models, as they currently stand, are still a good starting point for benchmarking the energy use of small wastewater plants.

4.7 Potential Sources of Unexplained Model Variability

There are many difficult to model factors that contribute to variability in energy usage between communities. These may include poor quality data, quantity of data available, equipment or plant age, level of maintenance, and level of operator training. Small Nebraska plants typically only have one to three certified operators and most plants are typically primarily manually controlled or equipment settings are manually set. Often, many operators have multiple duties for the municipality and are at the plant for less than 40 hrs/wk. For a handful of the Nebraska plants in our study, there were step changes in energy usage that corresponded to a change in operators. Table 4.4 shows five plants in the data set where step changes in energy usage corresponded to a change in operators.

Table 4.4: Changes in Electric Intensity with a Change in Operator.

Community	Previous Electric Intensity (MWh/MG)	Current Electric Intensity (MWh/MG)	% Change
A	5.50	3.09	-44%
B	7.11	4.46	-37%
C	5.05	8.61	70%
D	7.18	6.80	-5%

The large changes in energy usage seen by making operational changes can explain some of the variability not captured by the models. A significant degree of variability can occur based on operational decisions, in these cases it can be 5 to 70% changes (up to 3.5 MWh/MG). This suggests the importance of operator training and decisions and its impact on the variability in energy usage between otherwise similar plants.

4.8 Model Uses

The resulting models created in this study have several different uses. They can be used as a guide for creating similar models in the future, provide a baseline for comparison of individual plants, and allow for the estimation of overall electricity usage and/or energy usage for new plants, or groups of plants, for which actual energy data is not available. The different response variables can be used to predict either intensity or usage, depending on what model is used.

The steps taken to create these energy benchmarking models have been described in this research, with the main source being the Methods chapter (Ch. 3). Researchers can follow similar steps for future energy benchmarking studies, regardless of what wastewater plants are being studied and where they are located. Everything from what data to collect and how to collect the data are detailed in this paper along with how to create multiple linear regression models using the collected data. Analyzing the steps for collecting the data and how to analyze the data may save future researchers time for additional analysis on top of the analysis discussed in this paper.

Individual plants can use these models as a comparison to other plants, or to get a general idea of where their plant should be in regards to energy usage/intensity. Operators or even state officials may use these models for energy management plans for individual plants. Using these models gives individual plants a starting point or goal in regards to energy usage. A benefit of having separate models for both total energy use and electricity use makes it easier for plant managers and operators to compare actual plant energy usage to the model estimated usage. This is due to the fact that natural gas and/or propane usage is not often readily available to plant operators, especially at small plants.

State officials/regulators may also use these models to identify the least and most energy efficient plants without needing to collect utility bills. Utility companies may use the energy usage models to determine how much energy a new plant will use. The benchmarking models in this study provide a good starting point for comparing plant electric and/or energy usage/intensity.

4.9 Using the Benchmarking Models: A Detailed Example

The following example shows how one of the models created in this study can be used. The Nebraska Overall model for Electric Intensity is used here for one of the small Nebraska wastewater treatment plants used in this study.

Step 1

- Gather 12 months of electric use information.
- Gather plant information/characteristics from the past 12 months such as average flow (MGD), climate controlled floor area (ft^2), average daily design flow (MGD), and average influent $CBOD_5$ (mg/L).

Table 4.5: Example Plant Characteristics

Electricity Use (kWh)	118,309
Average Flow (MGD)	0.028
Average Influent $CBOD_5$ (mg/L)	185
Climate Controlled Floor Area (ft^2)	600
Average Daily Design Flow (MGD)	0.10
Extended Aeration Plant	1 (Yes)
Supplemental Energy Usage for Sludge Treatment	1 (Yes)
Dewatering Equipment	0 (No)

Step 2

- Compute actual Electric Intensity (MWh/MG).
    - Divide the annual electricity use by the annual volume of flow treated.

$$\textbf{Actual Electric Intensity} = \frac{118{,}309\ \frac{\text{kWh}}{\text{year}} * (1\ \text{MWh}/1000\ \text{kWh})}{0.028\ \text{MGD} * (365\ \text{days}/\text{year})} = 11.6\ \text{MWh/MG}$$

Step 3

- Compute predicted Electric Intensity (MWh/MG) using the model equation.

- Enter in data gathered in Step 1.
- Transform continuous data (Climate Controlled Floor Area, Average Flow, Percent Design Flow, and Average Influent $CBOD_5$) using the natural log (Ln).
- Multiply the transformed plant values by the respective model coefficients and sum these values up to find the Ln(Predicted Electric Intensity).
- Retransform the final sum by taking the exponential of the sum.

Table 4.6: Computing the Predicted Electric Intensity Using the Model Equation

Variable	Actual Plant Value	Transformed Value	Model Coefficient	Coefficient * Transformed Plant Value
Model Intercept	-	-	-2.06	-2.06
Extended Aeration (1 if yes, 0 if no)	1	-	0.257	0.257
Sup. Energy for Sldg Trt (1 if yes, 0 if no)	1	-	0.264	0.264
Dewatering Equip. (1 if yes, 0 if no)	0	-	0.272	0
Climate Controlled Floor Area (ft^2)	600	6.40	0.165	1.06
Average Flow (MGD)	0.028	-3.58	-0.323	1.16
Percent Design Flow	0.28	-1.27	-0.268	0.340
Average Influent $CBOD_5$ (mg/L)	185	5.22	0.256	1.34
			Ln(Predicted Electric Intensity)	2.36
			$e^{\text{Ln(Predicted Electric Intensity)}}$	10.6

Predicted Electric Intensity = 10.6 MWh/MG

Step 4

- Compare the Actual Electric Intensity to the Predicted, as well as the Nebraska Median Electric Intensity for the plant type.

Table 4.7: Comparing the Actual and Model Predicted Electric Intensities (MWh/MG) and Median Electric Intensity (MWh/MG) for the Specific Plant Type.

Actual Electric Intensity	11.6
Predicted Electric Intensity	10.6
Median Nebraska Extended Aeration Electric Intensity	6.2

This comparison shows that while the plant's actual electric intensity is much higher than the median for its plant type, it is not as far away from the model predicted electric intensity. This is due to the fact that the model takes into account that the plant is a small extended aeration plant, with aerobic digestion, 600 ft^2 of climate controlled floor area, an average flow of 0.028 MGD, a percent design flow of 28%, and an average influent CBOD of 185 mg/L, while the median value does not. Because the actual intensity is higher than both the predicted and median intensities, this comparison shows that there are opportunities for improvement at the plant in regards to energy efficiency.

In addition to comparisons among plants, these model equations can provide estimates of the savings for changes in plant characteristics based on the statistical data. For example, using the plant data in the previous example, one could estimate the potential energy savings from switching to lime stabilization from aerobic digestion. If a recommendation of switching to lime stabilization is implemented, a statistical estimate of the annual energy savings can be estimated using the model equation as follows.

Assuming all other plant variables stay the same and using a zero instead of a one for the supplemental energy usage for sludge treatment, the model equation calculates the newly predicted electric intensity as 8.1 MWh/MG. The model estimates an overall average savings of 30% or 3.5 MWh/MG. Using the average volume of flow treated per year and assuming the price of electricity as $0.08/kWh, this equates to an average savings of about 3,600 kWh per year or $288 per year. This calculation of course does not take into account capital costs or costs for materials such as lime, but it is strictly an estimate of the annual energy savings. Using the same method, one could also determine the savings or, decrease in electric intensity, of increased capacity/flow or even determine the savings of a decrease in the influent $CBOD_5$ loading such as what occurs when an industrial loading is no longer present. Simple calculations as these can help operators justify implementing energy efficiency recommendations that are not easily quantifiable and determine their predicted average savings.

4.10 Summary

The overall results of the benchmarking models in this study bring about several conclusions. Several variables that showed up as significant in the Energy Star (2014) and AwwaRF (Carlson and Walburger, 2007) models also showed up in the models developed for small Nebraska wastewater plants. These include average flow, percent design flow, and average effluent ammonia-nitrogen (nutrient removal in the Energy Star and AwwaRF models). There were certain variables that may only be significant for small wastewater treatment plants. These include climate controlled floor area, presence of dewatering equipment, supplemental energy usage for sludge treatment, presence of industrial loadings, and the presence of VFDs. Some variables that were thought to be significant in regards to energy consumption were found not to be significant. An example of this is the presence of automatic DO controls not being statistically significant.

Another important factor for small wastewater plants is nutrient removal, or more specifically, nitrification. Although the total amount of oxygen needed to oxidize $CBOD_5$ is anticipated to be significantly greater than the amount needed to oxidize ammonia, nutrient removal can still have an important impact on plant energy use. Because many of the small plants in this study were not designed for nutrient removal, plants are forced to operate differently than they were designed, which leads to inefficiencies. The fact that effluent ammonia-nitrogen concentration was showing up as significant more often than $CBOD_5$ is a result of the added oxygen demand required to achieve nitrification.

Additionally, it may also be due to the fact that there was greater variability in the effluent ammonia-nitrogen than influent or effluent $CBOD_5$. Greater variability was observed in the Nebraska dataset for the regulatory effluent limits for ammonia-nitrogen (no limit or 0.4 to 32 mg/L) than for effluent $CBOD_5$. The effluent $CBOD_5$ limit for all the Nebraska plants was 25 mg/L. The estimated sample variance of the reported effluent ammonia-nitrogen concentrations was 23.4, while the estimated sample variance of the reported effluent $CBOD_5$ concentrations was only 11.1. This greater variability may be why the effluent ammonia-nitrogen concentration showed up as significant more often than $CBOD_5$ in the Nebraska models. Using an appropriate measure that takes into account both carbonaceous and nitrogenous biochemical oxygen demand (such as BOD_5) may lead to a better representation of the energy used for treatment.

In addition to what factors affect energy usage at small wastewater plants, the models also show that energy benchmarking models for small plants may be state or region specific, based on the differences between the Nebraska and Pennsylvania models. These differences stem from differences in technology, time period data was gathered, and regulatory requirements.

The resulting Adjusted-R^2 values are summarized for all of the models in Table 4.8. The electric models had higher Adjusted-R^2 values than the energy models. This may be due to the fact that most plants in the study use strictly electricity. The usage models had higher Adjusted-R^2 values than the intensity models because the predictive power of flow was not hidden in the response variable for the usage models. The Adjusted-R^2 values are good measures of fit, but the variability in plant energy usage/intensity cannot be modeled perfectly. These models, therefore, are a good tool for estimating the electric or energy usage/intensity of small wastewater treatment plants, but not perfect. Further research should be conducted to improve these models.

Table 4.8: Comparing Model Adjusted-R^2 values

Model	**Response Variable**			
	Electric Intensity (MWh/MG)	Annual Electric Usage (kWh/year)	Energy Intensity (MWh/MG)	Annual Energy Usage (kWh/year)
Overall NE	0.52	0.88	0.49	0.88
NE EA	0.47	0.86	0.38	0.84
NE OD	0.71	0.89	0.69	0.90
NE SBR	0.94	-*	0.96	-*

*Models for NE SBR plants for electric and energy usage could not be created due to non-normality in the distribution of electric and energy usage for the SBR data set.

Chapter 5: Conclusions and Recommendations

5.1 Conclusions

Energy usage, plant characteristics and processes, influent and effluent water quality data, climatic information, information on how the plant is operated, and other plant information related to energy usage from 83 small wastewater treatment plants in Nebraska and 71 small wastewater plants from Pennsylvania was collected and used in this study to create energy benchmarking model equations. The benchmarking models predict either electric or total energy (electricity + natural gas + propane usage) intensity (energy consumed per unit flow treated, MWh/MG) or annual consumption (kWh/year). From this research, the following conclusions were made:

- The data from both Nebraska and Pennsylvania fit the models well with model Adjusted-R^2 values ranging from 0.38 to 0.96. The best fit was found with the usage models (kWh/year) as opposed to the intensity models (MWh/MG) mainly because the response variable in the intensity models includes flow. In addition, the electric models tended to provide a better fit to the data than the total energy models. This may be explained by the fact that most plants in the data set use only electricity.
- There were some similarities between the Nebraska models and the AwwaRF/Energy Star models. Both sets of models included average daily flow, percent design flow ($\frac{\text{Average Daily Flow, MGD}}{\text{Average Design Daily Flow, MGD}}$), and a variable related to nutrient removal. The fact that these are significant in small and large plant models provides further confirmation that flow and nutrient removal are important factors for both large and small plants in regards to energy usage.

- Some variables that did not show up as significant in previous benchmarking studies may only be significant for small plants. These include climate controlled floor area, supplemental energy usage for sludge treatment, and presence of dewatering equipment.
- Some variables thought to be significant in regards to energy usage were found to be not as significant as originally thought. Examples of these variables include presence of automatic dissolved oxygen controls and influent organic loading (with the exception of the Nebraska overall models looking at electric intensity/usage only).
- Average effluent NH_3-N concentration was found to be a significant parameter in determining total energy use in more Nebraska models than influent carbonaceous oxygen demand (CBOD) concentration. Ammonia-nitrogen may have been showing up as significant due to the fact that effluent limits for ammonia-nitrogen varied much more than effluent CBOD.
- The comparison of models between Nebraska and Pennsylvania revealed underlying differences that are difficult to identify. The differences in small wastewater plant models between states may indicate that energy benchmarking models may be more state or region specific.

5.2 Recommendations for Future Research

Due to the fact that more data can always be collected and analyzed, suggestions for future research are presented that identify ways in which the models created in this thesis may be improved or modified. These recommendations suggest collecting better quality data, as well as more detailed data in general. The following are suggestions for future research:

- Average influent CBOD for Nebraska plants was only sampled once per year. While CBOD showed up as significant for only two of the small Nebraska models, it was shown to be highly significant in the AwwaRF and ENERGY STAR Models (Carlson and Walburger, 2007; ENERGY STAR, 2014). Potentially, CBOD was not as significant for the smaller plants in part due to many of them being operated without DO control and at a constant blower rate year-round regardless of temperature changes and changes in flow rate. Obtaining more samples throughout the year from plants may be more representative of actual CBOD loadings and could change the model results.
- The data collected for the plants in this study were not as detailed as what is collected for an energy audit. Young and Koopman (1991) and Foladari *et al.* (2015) carried out detailed energy audits on five small wastewater plants in each study. They also benchmarked the plants, both by overall energy consumption and by unit processes. Following the methodology of Young and Koopman (1991) and Foladari *et al.* (2015), one could audit 5 to 10 small plants, collecting more detailed information, and produce benchmarks for unit processes at small wastewater plants in order to confirm or challenge the findings of Young and Koopman (1991), Foladari *et al.* (2015), and other studies that determined unit process benchmarks for small wastewater plants.
- The economy of scale for this study and AwwaRF's (Carlson and Walburger, 2007) was quite influential, especially when looking at the influence flow rate had

on energy consumption. AwwaRF (Carlson and Walburger, 2007) and ENERGY STAR (2014) both removed plants treating less than 0.6 MGD because they were on the lower end of the flow spectrum. The flow range for the Nebraska study was from 0.01 to 1.5 MGD. The influence of flow at an even smaller scale, 0.01 to 0.5 MGD may produce different results. Therefore, it may be beneficial to create benchmarking models on an even smaller scale.

- Since AwwaRF believes that a nationwide model is able to accurately predict the energy usage of plants in different regions of the U.S., it might be worthwhile to gather more data on small plants from more states or regions, specifically in the Midwest (close to Nebraska) for comparison. Recreating the models with more plants from different states, but from the same region, may make more sense than a national model for small plants, but more data collection is needed. On the other hand, it may show that small plant models may be applicable on a national scale.
- AwwaRF (Carlson and Walburger, 2007) and ENERGY STAR (2014) both created scoring tables for the output of their benchmarking models on a scale of 1 to 100 with 100 being the most energy efficient and 1 being the least. This scoring method creates a simple score for plant managers, operators, and city officials to interpret. Creating a similar scoring table for Nebraska plants may be beneficial for helping plant managers, operators, and city officials better understand the benchmarking model results.
- Other factors shown to be significant in the AwwaRF and ENERGY STAR models, but not in the small system models were HDDs and CDDs. HDDs and CDDs may not have shown up as significant because only one state was included

to create the small system models. AwwaRF and ENERGY STAR used plants from states around the U.S. A recommendation for future research would be to collect data from small plants throughout the U.S. This may further show that climate plays an important factor in small plant energy use.

Chapter 6: References

Albertson, O.E. (1995), "Is $CBOD_5$ Test Viable for Raw and Settled Wastewater?" *Journal of Environmental Engineering*, 121(7), 515-520.

Bilder, C., Loughin, T. (2015), *Analysis of Categorical Data with R,* CRC Press, Boca Raton, FL.

Boller, M., (1997), "Small Wastewater Treatment Plants-A Challenge to Wastewater Engineers," *Water Sci. Tech.*, 35(6), 1-12, https://doi.org/10.1016/S0273-1223(97)00089-9.

Carlson, S.W., Walburger, A. (2007) *Energy Index Development for Benchmarking Water and Wastewater Utilities.* Awwa Research Foundation: Denver, CO.

Crites, R., Tchobanoglous, G., (1998), *Small and Decentralized Wastewater Management Systems,* WCB/McGraw-Hill, New York, NY.

Electric Power Research Institute and Water Research Foundation (EPRI) (2013) *Electricity Use and Management in the Municipal Water Supply and Wastewater Industries*, EPRI Report No. 3002001433.

ENERGY STAR (2014) *Score for Wastewater Treatment Plants in the United States.* ENERGY STAR Portfolio Manager. Retrieved 2.8.2017 from https://www.energystar.gov/sites/default/files/tools/Wastewater.pdf.

ENERWATER (2015) Deliverable 2.1 Study of published energy data. European Commission, Horizon 2020 ENERWATER Project. Retrieved January 1st, 2017 from http://www.enerwater.eu/wp-content/uploads/2015/10/ENERWATER_D2.1-Study-of-published-energy-data-Will-contain-data-of-at-least-500-WWTPs.pdf.

EPA. (EPA/600/S2-85/007, March 1985). Technology Assessment of Sequencing Batch Reactors. Retrieved June 5th, 2017 from https://nepis.epa.gov/Exe/ZyPDF.cgi/2000TO6Q.PDF?Dockey=2000TO6Q.PDF.

EPA. (EPA/832/R-92/003, September 1992). Evaluation of Oxidation Ditches for Nutrient Removal. Retrieved June 6th, 2017 from https://nepis.epa.gov/Exe/ZyPDF.cgi/200045V0.PDF?Dockey=200045V0.PDF.

EPA. (EPA/932/F-99/073, September 1999). Wastewater Technology Fact Sheet:

Sequencing Batch Reactors. Retrieved May 23rd, 2017 from https://www3.epa.gov/npdes/pubs/sbr_new.pdf.

EPA. (EPA/832/F-00/016, September 2000). Wastewater Technology Fact Sheet: Package Plants. Retrieved Jan. 4th, 2017 from https://www3.epa.gov/npdes/pubs/package_plant.pdf

EPA. (EPA/832/R-10/005, September 2010). Evaluation of Energy Conservation Measures for Wastewater Treatment Facilities. Retrieved May 5th, 2016 from https://nepis.epa.gov/Exe/ZyPDF.cgi/P1008SBM.PDF?Dockey=P1008SBM.PDF.

EPA. (EPA/830/R-15/005, January 2016). Clean Watersheds Needs Survey 2012. Retrieved November 29th, 2016 from https://www.epa.gov/sites/production/files/2015-12/documents/cwns_2012_report_to_congress-508-opt.pdf.

EPA. Enforcement and Compliance History Online (ECHO). Environmental Protection Agency. https://echo.epa.gov/ (2.8.2016).

EPA (2017), "Small and Rural Wastewater Systems: Challenges and Needs," https://www.epa.gov/small-and-rural-wastewater-systems/learn-about-small-wastewater-systems#wastewater.

Foladori, P., Vaccari, M., and Vitali, F. (2015), "Energy audit in small wastewater treatment plants: methodology, energy consumption indicators, and lessons learned," *Water Science & Technology*, 72(6), 1007-1015, http://dx.doi.org/10.2166/wst.2015.306.

Google Maps (2017a), *Dodge, NE Wastewater Treatment Plant*, Retrieved from https://www.google.com/maps/@41.7174401,-96.8798168,107m/data=!3m1!1e3.

Google Maps (2017b), *Eagle, NE Wastewater Treatment Plant*, Retrieved from https://www.google.com/maps/@40.8013465,-96.4276257,102m/data=!3m1!1e3.

Google Maps (2017c), *Cozad, NE Wastewater Treatment Plant*, Retrieved from https://www.google.com/maps/@40.8408536,-99.9771587,129m/data=!3m1!1e3.

Gretzschel, O., Schmitt, T., Hansen, J., Siekmann, K., and Jakob, J. (2014), "Sludge Digestion Instead of Aerobic Stabilization-A Cost Benefit Analysis Based on Experiences in Germany," *Water Science & Technology*, 69(2), 430-437, http://dx.doi.org/10.2166/wst.2013.711.

Li, B., and Wu, G., (2014), "Effects of Sludge Retention Times on Nutrient Removal and Nitrous Oxide Emission in Biological Nutrient Removal Processes," *International Journal of Environmental Research and Public Health*, 11, 3553-3569, http://dx.doi.org/10.3390/ijerph110403553.

Mizuta, K., and Shimada, M. (2010), "Benchmarking Energy Consumption in Municipal Wastewater Treatment Plants in Japan," *Water Science & Technology*, 62(10), 2256-2262, http://dx.doi.org/10.2166/wst.2010.510.

National Oceanic and Atmospheric Administration (NOAA). NOWData-NOAA Online Weather Data. Accessed 2.8.2016. http://sercc.com/nowdata.html.

NEO (Nebraska Energy Office) (2016), "Municipalities Saving Energy and Lowering Costs for Wastewater Treatment," http://www.neo.ne.gov/neq_online/dec2016/dec2016-08.html, (5.22.2017).

NDEQ (2014). Assessing Wastewater Infrastructure Needs Executive Summary (AWIN). Retrieved March 28th, 2016 from NDEQ staff.

NDEQ. Public Records Searches and Interactive Map Server. Nebraska Department of Environmental Quality. http://deq.ne.gov/NDEQProg.nsf/OnWeb/PRS (2.8.2016).

New England Interstate Water Pollution Control Commission (NEIWPCC) (2005). Sequencing Batch Reactor Design and Operational Considerations. Lowell, MA. Retrieved May 12th, 2016 from http://www.neiwpcc.org/neiwpcc_docs/sbr_manual.pdf.

New York State Energy Research and Development Authority. (NYSERDA/0817/November 2008). Statewide Assessment of Energy Use by The Municipal Water and Wastewater Sector. Retrieved December 20th, 2016 from https://www.nyserda.ny.gov/-/media/Files/EERP/Commercial/Sector/Municipal-Water-Wastewater-Facilities/nys-assessment-energy-use-wwtp-biogas-recovery.pdf.

Ngo, T.H.D. (2012), "The Steps to Follow in a Multiple Regression Analysis," *Proceedings of the SAS® Global Forum 2012 Conference*. Cary, NC: SAS Institute Inc.

Nowak, O. (2006), "Optimizing the Use of Sludge Treatment Facilities at Municipal

WWTPs," *Journal of Environmental Science and Health,* Part A, 41(9), 1807-1817, http://dx.doi.org/10.1080/10934520600778986.

Pacific Gas and Electric (PG&E) Company (2006) *Energy Baseline Study For Municipal Wastewater Treatment Plants.*

PDEP. *Electric Use at Pennsylvania Sewage Treatment Plants.* Pennsylvania Department of Environmental Protection. 2011.

Public Service Commission of Wisconsin (2016) Energy Best Practices: Water & Wastewater Industry. Focus on Energy program. Retrieved Oct. 24th, 2016 from https://focusonenergy.com/sites/default/files/WW-Best-Practices_web_1.pdf.

SAS Institute Inc. 2014. SAS/STAT® 13.2 User's Guide. Cary, NC: SAS Institute Inc.

Sheather, S.J. (2009), *A Modern Approach to Regression with R,* Springer Science+Business Media, LLC, New York, NY.

Silva, C., and Rosa, M.J., (2015), "Energy performance indicators of wastewater treatment: a field study with 17 Portuguese plants," *Water Science & Technology,* 72(4), 510-519, http://dx.doi.org/ 10.2166/wst.2015.189

Tchobanoglous, G., Burton, F., Stensel, H. D., Tsuchihashi, R., Abu-Orf, M., Bowden, G., and Pfrang, W. (2014). *Wastewater Engineering: Treatment and Resource Recovery* . McGraw-Hill, New York, NY.

Trapote, A., Albaladejo, A., Simon, P. (2013), "Energy Consumption in an urban wastewater treatment plants: the case of Murcia Region (Spain)," *Civil Engineering and Environmental Systems,* 31(4), 304-310, http://dx.doi.org/10.1080/10286608.2013.866106.

United States Department of Energy (DOE): National Renewable Energy Laboratory (2012) *Energy Efficiency Strategies for Municipal Wastewater Treatment Facilities.* NREL/TP-7A30-53341.

Velleman, P. F., and Hoaglin, D. C. (1981). *Applications, basics, and computing of exploratory data analysis*. Duxbury Press, Boston, MA.

Wong, S., and Law-Flood, A. (2011), Tapping the Energy Potential of Municipal Wastewater Treatment: Anaerobic Digestion and Combined Heat and Power in Massachusetts. Massachusetts Department of Environmental Protection.

Retrieved June 2nd, 2017 from http://www.mass.gov/eea/docs/dep/water/priorities/chp-11.pdf.

Yang, L., Zeng, S., Chen, J., He, M., and Yang, W. (2010), “Operational energy performance assessment system of municipal wastewater treatment plants,” *Water Science & Technology,* 62(6), 1361-1370, http://dx.doi.org/10.2166/wst.2010.394.

Young, D.F., and Koopman B. (1991), “Electricity Use in Small Wastewater Treatment Plants,” *Journal of Environmental Engineering,* 117(3), 300-307, http://dx.doi.org/10.1061/(ASCE)0733-9372(1991)117:3(300).

Appendices

Appendix A: Pre-Assessment Guide

Filling out the Assessment Survey Forms-Summer 2016

The following is a guide on how and where to find information about certain small wastewater plants in Nebraska before you go out and visit them. Having some of the information before visiting the plant makes the visit much simpler and allows you to focus on other aspects of the plant such as how exactly the plant is operated in regards to energy usage, rather than the plant characteristics and influent and effluent water quality. In addition, it helps validate information found online. The information that you will need to find is listed on both the Wastewater Facility Energy Use Assessment sheet and the Assessment Spreadsheet. The goal is to find as much information as possible before doing an on-site assessment. Much of the information that you will be gathering this summer is available online, but the information is not always 100% reliable. Because it is not always reliable, you will need to confirm it with the operators during your visits to these plants. The following steps take you through the Wastewater Facility Energy Use Assessment sheet and Assessment Spreadsheet and tell you what information can be found online and where you can find it. A lot of the information can be found in multiple places, but for the sake of consistency and accuracy, certain documents should be used over others. These "other" documents should only be used if the ones suggested to be used cannot be found or contain obviously erroneous information. You'll understand what I mean once you really start looking at these documents. This guide should help you find a lot of information about the plant, but if you have any questions, feel free to ask Steve or Dr. Dvorak.

<u>Wastewater Facility Energy Assessment Form</u>

1. **Contact Information**
   The Facility contact can be found on the List of Plants Excel file. More than likely, you will be meeting with the operator, so write down the operator's name and phone number. Later on, you can ask for their email, or if you meet with someone else, you can change the information. To find the Facility Address, you will want to look at an inspection form on the NDEQ website. The NDEQ has all public records online, but it can be a little difficult to find the information you are looking for unless you have that facility's IIS number. A facility's IIS number is basically the facility's ID number used by the NDEQ. Luckily, we were provided with these numbers and they can be found on the List of Plants Excel file. The next thing you'll want to do is go to the NDEQ's website. Once on the NDEQ home page, scroll to the bottom and click on "Public Records Search." This link will direct you to another page. On the new page click on the link towards the top of the page labeled, "Public Records Search." A new window will pop up titled, "Nebraska Enterprise Content Management Portal." In the search box labeled, "DEQ Facility Number," enter the IIS number for the plant of interest and click "Search." A list of documents pertaining to the facility of interest will show up.

To find the Facility Address, find the most recent Inspection document. The inspection document is labeled as, “DEQ Inspection,” under the Document Type column and is labeled as, “COMPLIANCE,” under the DEQ Description column. Also, make sure the inspection document is for a wastewater inspection, not an air pollution inspection. A wastewater inspection is labeled as “PCS” under the DEQ Program column. An air pollution inspection is labeled as “AIR” in the DEQ Program column. The Facility Address can be found in this document labeled as Facility Location. We want the Facility Location, not the mailing address. If the plant does not have an address, go to Google Earth, find the wastewater plant, and record the cross streets that are next to the plant.

2. Plant Characteristics

a. Design Flow

The Design Flow can be found on the NDEQ website in the inspection document. The design flow data can also be found in NDEQ documents labeled “Fact Sheet,” “NPDES Permit Application,” or “Discharge Monitoring Report.” It is preferred that the design flow listed in the inspection document be used, as this seems to be the most accurate source out of all the other documents. Most of the time, the data is the same for all the different documents for the plant, but sometimes the documents do not agree. It is always a good idea to cross-reference documents to confirm not only that the design flow is correct, but also other data found in these NDEQ documents is correct. Always make sure to cite where the data come from, just in case discrepancies arise.

b. Population

Population data can be found on the List of Plants Excel file. The population data were supplied to us by the Nebraska League of Municipalities.

c. Industrial Users

Information on industrial users can be found on the NPDES Permit Application document for the plant on the NDEQ website. The table containing the number of industrial users and their respective flows and loadings can be found on the fourth page of the NPDES Permit Application. Some plants do not have all the information for each industrial user, but record whatever information is listed in this table.

d. Type of Discharge and Frequency

The type of discharge can be found on the NDEQ website in the document titled “NPDES Municipal Wastewater” under the DEQ Description column. It is labeled as “DEQ Application” under the Document Type column. This document is the plant’s discharge permit application. The Type of Discharge can be found towards the top of the fourth page of the application. Make sure to use the most recent application. If the type of discharge is intermittent, make sure to also mark down the frequency of discharge. The frequency can either be in discharges/year or discharges/day.

e. Buildings, Floor Area, and Plant Diagram

Count the number of buildings on-site using Google Maps and use the measure tool to find the total floor area of all buildings on site. For the Plant Diagram, take a screenshot of an aerial view of the plant using Google Maps. Make sure to outline the buildings and label them describing what is inside of the buildings. During the site visit, don't forget to double check that you measured the right buildings, that the dimensions seem to be correct, and that you have adequately described what they have inside of them.

f. Treatment Processes used at the Facility and Other Information

The treatment processes used at the plant can be found by looking at the most recent inspection document on the NDEQ website. The inspection document should have most of what you need, but it is also important to check the Fact Sheet and NPDES Municipal Wastewater Application to check if any other information can be found. Another good source for treatment processes and any other plant information are Engineering Specs or Studies and Operation Guides. The Operations Guide is labeled as "DEQ Plan" under the Document Type column and as "Operations and Maintenance Manual" under the DEQ Description column. Engineering Specs or Studies are labeled under the Document Type column as "DEQ Plan" and labeled as "Facility Engineering Report" under the DEQ Description column. The Engineering Reports and Operation Manuals list all the processes at the plant and specific details about all the equipment. Not all plants have Engineering Reports or Operation Manuals listed on the NDEQ website, but if they do, take a look at them to find other information such as information on treatment processes, pumps and motors used at the plant, types of diffusers, sludge treatment information, and other information listed on the Wastewater Facility Energy Assessment Form.

Assessment Spreadsheet

1. Energy Use/Utility Bills

Before visiting the community, it is important to obtain the utility bills for the wastewater plant. It is advised to acquire the utility bills before the visit, just in case the bills are difficult to read or if there are any errors. Our goal is to get up to three years-worth of energy bills. When calling communities, make sure to call the town clerk first because they are usually the ones who have access to the plant's utility bills. When requesting the utility bills for the wastewater plant, it is important to clarify that we need bills for the wastewater plant, not the drinking water plant, from the past three years for all forms of energy used at the plant including electricity, natural gas, and propane. Also, make sure to request for monthly usage (kWh of electricity or therms of natural gas), monthly cost information, monthly demand (kW), demand charges, billing dates, and meter numbers. Once you have obtained the bills, fill out the columns for the monthly utility bills, labeling each meter number, starting with the oldest bill. Only record the energy used on-site. Do not record energy used by lift stations. If it is unclear which meters are on-site and which ones are for lift stations, record all meters and their usages and confirm with the meter numbers on-site when you visit the plant. If the plant uses natural gas, in addition to electricity, and the billing dates do not line up, list the natural gas billing dates by matching them up to the electric billing dates as best as you can. Most plants that use propane may not meter their propane usage, but they should be able to provide receipts for when the propane tank was filled up, how much propane was dispensed, and how much it cost. Much like the natural gas usage, try your best to line up when the tank was filled with the electric billing dates.

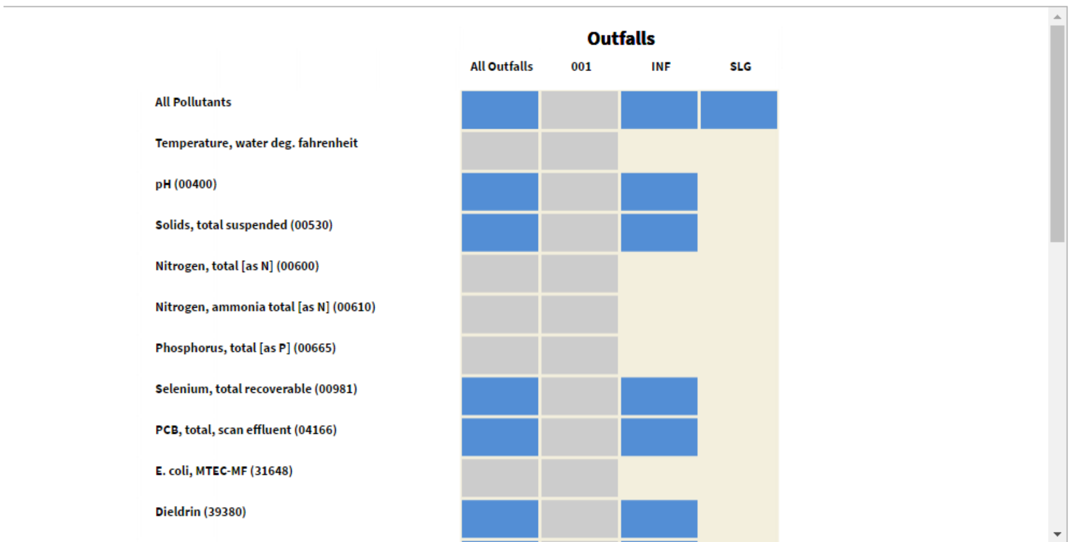
Many times, town clerks will be either very busy, or hesitant to participate in the study because of the amount of work it takes to put together utility bills for the wastewater plant from the past three years. Many of these town clerks have different responsibilities in their town other than town clerk and simply don't have time to look for utility bills. If they seem to be hesitant or say they do not have time to provide the bills, mention the NEO's Utility Release Form. The Utility Release From allows us to contact their energy suppliers directly about the energy usage at their plant. The only thing the town clerk needs to do is provide their wastewater facility address, the names of their energy suppliers, their account numbers, and their signature. Have them either email you the signed release form, or fax or mail it to the NEO directly. Once we have their signed release form, we can directly contact their suppliers about their energy usage. While the preferred method is to have the town clerks provide the bills directly, the Utility Release Form is a secondary option to be used.

2. Influent and Effluent Water Quality

a. Influent and Effluent Flow, CBOD, and TSS and Effluent NH3-N

The plant characteristics can be found in a variety of documents. However, the best and easiest place to get the influent and effluent flow, influent and effluent CBOD, influent and effluent TSS, and effluent NH3-N data are from ECHO. ECHO stands for Enforcement and Compliance History Online. It is an EPA website that contains information about any kind of facility that has an EPA permit and that discharges pollutants of any kind. ECHO compiles all of the water quality characteristics that a plant is required to report. The same information can be found in the documents listed for the plant on the NDEQ website and are titled "Discharge Monitoring Reports (DMR)." DMRs are monthly reports that document data such as effluent CBOD, effluent pH, effluent TSS, and several other influent and effluent characteristics. ECHO compiles the information on the DMRs and presents it in tables that can easily be copy and pasted onto the Assessment Spreadsheet.

Go to the following website: https://echo.epa.gov/. In the search box in the middle of the page, type the community in which the wastewater plant you are investigating is located. A list of facilities in the community with EPA permits will appear. Click on the link for the wastewater plant. Once on the page for the wastewater plant, click on the link that says, "CWA Effluent Charts." This will take you to the page containing the information that you need. The following chart will appear:



Click on any of the boxes to get information about influent and effluent flow, influent and effluent CBOD, influent and effluent TSS, and effluent NH3-N. Once you have clicked on the box corresponding to the data you are looking for, you can view the data in table form by clicking on the blue box labeled "Show/Hide Table." Almost all the plants you will be going to this summer are only required to report influent data once a year. Typically, ECHO only has influent data from the past 3 or 4 years. You will need to find influent data for flow, CBOD, and TSS.

Plants are typically required to report effluent data every month. Therefore, you will record the effluent values for CBOD, TSS, and NH3-N. As with influent data, be sure to mark the range of dates. For both influent and effluent CBOD, TSS, and NH3-N, make sure you are recording the Concentration (mg/L) and not Quantity (kg/d).

Record the influent and effluent data for as far back as we have electric bills for. When recording the influent and effluent data onto the Assessment Spreadsheet, make sure to line up the month and year the sample was recorded with the ending month of the electric bill. This is important because we want to be consistent and because not all plants will have electric bills beginning and ending in the same month.

b. Current Discharge Limit for NH_3-N (mg/L)

The Current Discharge Limit for NH_3-N can be found on the documents on the NDEQ's website. The limit can be found on the document labeled as "DEQ Issued Permit" under the Document Type column and as "NPDES Municipal wastewater" under the DEQ Description column. This is the plant's NPDES Permit that lists all the discharge limits for the plant. The NH3-N limit can be found on the Table labeled, "Seasonal Discharge Limits and Monitoring Requirements for Ammonia." Most plants will have discharge limits, but some are only required to monitor the effluent ammonia. For plants that are only required to monitor effluent ammonia, leave the Current Discharge Limit for NH3-N blank.

The limits are listed by season (Summer, Winter, and Spring). Make sure to record the concentration limits and that the Permit you are looking at is the most up to date. It is advised to double check with the operators about the ammonia limits during the site visits.

3. Heating Degree Days (HDDs) and Cooling Degree Days (CDDs)

HDDs and CDDs are measures of how much a facility needs to cool or heat their buildings throughout the year based on how many degrees the outside temperature deviates from the baseline of 65°F each day of the year. These can be found on NOAA's website called NOWData, which stands for NOAA Online Weather Data. Go to the following website: http://www.sercc.com/nowdata.html. Here, NOAA has NWS Offices listed by Region. Nebraska is divided into 3 offices, Omaha, Hastings, and North Platte. The Omaha office covers the Eastern portion of the state, Hastings the middle, and North Platte the Western portion including most of the panhandle. Most of the weather information for the towns to be visited this summer can be found under these three offices. However, there are several towns in the panhandle, northeast corner, and southwestern corner that are not covered by these three NWS Offices. You will need to use the Cheyenne, WY Office for parts of the panhandle, the Goodland, KS Office for part of the southwestern corner of the state, and the Sioux Falls, SD Office for part of the northeastern corner of the state.

Clicking on the appropriate Weather Office will bring you to the following screen:

Scroll through the list of locations to find the town you are assessing and click on it. Next, click on the Monthly Summarized Data option under Product. Under Options, specify the Year Range by how far back you have electric bills for the plant of interest. Change the variable option to HDD base 65 for HDDs or CDD base 65 for CDDs. Leave the Summary Option as sum and click Go. This will give you the HDDs or CDDs by month for the years you specified. Record the monthly HDD and CDD values on the Assessment Spreadsheet by matching the months up with the ending months for the electric bills. In addition to recording the values, make sure to record the station the values were recorded from. Label the station on the Assessment Spreadsheet exactly as it appears on the NOWData site. Take note that the HDD values are listed on the NOWData website by season (July-June), while CDD values are listed by year (January-December).

If the community you are looking for does not appear on the list of locations, click on the View Map option. Zoom in to where your community is located and click on the option saying, "Show more stations." This will show additional stations not on the location list. These additional stations sometimes have the same information as the main stations, but sometimes do not. It is important to check these additional stations to find your community in order to have the most accurate climate information.

Sometimes, stations, even the stations appearing on the main list, have missing data. Missing data are labeled as "M" under the month that has missing data. If a community is missing 3 or less months for the time frame in which you are looking, find the nearest station and fill in the data gaps, but list where that month's data came from. Also, check the months surrounding the missing month for both communities to determine if they are similar values. If a station has more than three missing months of HDD and CDD data, use the next closest station.

Additional Notes

In summary, this is the information that you can find online before your visit:

Wastewater Facility Assessment Form:

- Contact Information
- Facility Location/Address
- Average Design Flow
- Population
- Industrial User Information
- Type of Discharge and Frequency
- Number of buildings, total floor area, Plant Diagram, and Building Descriptions
- Treatment Processes used at the Facility
- Sources of Data
- Any other data that can be found in the documents online that is also on the Assessment Form.

Assessment Spreadsheet:

- Utility Bills (Include all forms of energy: electricity, natural gas, propane)
- Average Influent and Effluent Flow
- Average Influent and Effluent CBOD
- Average Influent and Effluent TSS
- Average Effluent NH_3-N
- Current Discharge Limits for NH3-N
- HDDs
- CDDs
- Source community for HDD and CDD values

All information found online must be confirmed with the operators during the plant visits. All other information on the Wastewater Facility Energy Assessment Form and Assessment Spreadsheet not found during the Pre-Assessment can only be filled out by visiting the wastewater plants and interviewing the operators.

Appendix B: Example Assessment Forms

Wastewater Facility Energy Assessment Form

Assessor and Reviewer Information

Assessor:____

Date and Time of Visit:______________________________________

Assessment Form Reviewer:___________________________________

Date of Review:____

Contact Information

Facility Name: ______________________________________

Facility Address: ______________________________________

____

Facility Contact: ______________________________________

Contact Phone:_________________ E-mail:______________________

Plant Characteristics

Design Flow:______________MGD

Has there been a large difference in flow over the past 3 years?
☐Yes ☐No

If yes, ask the operator why and explain.

____

____

____

____

Population served:_______________(Source: Nebraska League of Municipalities)

Plant Loading:

Has there been a large difference in influent loading (BOD, TSS, NH_3-N) over the past 3 years?☐Yes ☐No

If yes, ask the operator and explain.

____

____

____

____

Are there any industrial or commercial users that discharge to the plant? ☐Yes ☒No

If yes, please fill in the following table:

Industrial/Commercial Users			
Name of Business	**Description of Business**	**Average Flow Discharged per Day (MGD)**	**Average Loading (lbs-BOD/day)**

Type of Discharge: ☐Continuous ☐Intermittent

If intermittent discharge, what is the frequency of discharge? ________________________

____

Number of Buildings on-site:_________________________

Total floor area of buildings on-site:_________________________

Plant Diagram: Draw a plant diagram over a screen grab of an aerial photo of the plant labeling buildings and treatment areas. If there is not enough space on this page, attach the diagram as a separate page.

Building Descriptions:

Treatment Processes used at the facility: Mark all that apply.

Type of Bioreactor-Suspended Growth

- SBR ☐
- Oxidation Ditch ☐
- Conventional Activated Sludge ☐
- Extended Aeration ☐
- Other* ☐

Type of Bioreactor-Fixed Film

- Rotating Biological Contactor ☐
- Trickling Filter ☐

Activated Sludge

- Mechanical Aerators ☐
- Coarse Bubble ☐
- Fine Bubble ☐
- Pure Oxygen ☐

Nutrient Removal

- Biological Nitrification ☐
- Biological Denitrification ☐
- Biological P Removal ☐

Disinfection

- Chemical ☐
- Ultraviolet (UV) ☐

Sludge Treatment

- Thickening ☐
- Dewatering ☐
- Pumped to lagoon ☐

Sludge Digestion

- Aerobic ☐
- Anaerobic ☐

Sludge Disposal

- Composting ☐
- Land Application ☐
- Incineration ☐
- Landfill ☐
- Hauled off-site ☐

If "Other Bioreactor" was chosen, please list what type is used and explain the process:__

____

____

How many lift stations does the community have for its wastewater?_________________

Pumps, Motors, and Aeration System

Types of pumps and blowers used at the plant. Mark all that apply.
☐Centrifugal ☐High speed turbo ☐Rotary-Lobe positive displacement

Do they have any pumps/motors that need to be replaced soon?
☐Yes ☐No ☐Unknown

If yes, which ones?____

____

____

Do any of the motors used at the plant have variable frequency drives (VFDs)?
☐Yes ☐No

If yes, how many and which ones?__________________________________

____

____

____

Is the secondary treatment aeration system controlled automatically via DO levels and/or pressure differentials?
☐Yes ☐No

If yes, describe how the system is operated (what level is the DO set to, etc.), and if no automated DO controls are used, how is the aeration system controlled?

____

____

____

____

Is the DO in any of the aeration basins >2.0 mg/L at any time? ☐Yes ☐No

If yes, at approximately what value does the DO level peak?_________________________

____

What was the DO level at the time of your visit?___________________________________

Has the plant ever checked/tested the efficiency of the pumps/blowers in the plant?
☐ Yes ☐ No ☐Unknown

If diffused aeration is used, how often are the diffusers cleaned?______________________

____

If diffused aeration is used, how often is the air filter(s) changed?____________________

____

Sludge Treatment and Digestion

If sludge is pumped to a storage lagoon, how often is the lagoon emptied/cleaned out?___

____

If aerobic digestion is used, how are the blowers controlled?
☐ Automated DO controls ☐ Operator judgement

If operator judgement is used, please explain the procedure that is used to determine when the digester blowers are run:

____

____

____

____

If dewatering of the sludge is used, what type of dewatering technology is used?
☐Centrifuge ☐Belt-Filter Press ☐Drying Beds☐Other:__________________

If anaerobic digestion is used, does the plant produce biogas?
☐Yes ☐No

Do they use the biogas for energy?
☐Yes ☐No

If yes, how much is produced per year?____

____

Disinfection

If UV disinfection is used and the plant discharges intermittently, do they turn off the UV system when the plant is not discharging? ☐Yes ☐No

If UV disinfection is used, are the lamps self-cleaning? ☐Yes ☐No

If no, how often are the lamps cleaned?____

Lighting, Heating, and Cooling

Mark the types of lights used at the facility:
☐Fluorescent ☐Incandescent ☐LED ☐Halogen

☐Other:_______________________________

Are any of the lights controlled by motion sensors? ☐Yes ☐No
☐Unknown

Do all of the lights have a switch that turns them on and off? ☐Yes ☐No
☐Unknown

Do the buildings on site have programmable thermostats? ☐Yes ☐No
☐Unknown

Are the outdoor lights controlled by photo cells? ☐Yes ☐No
☐Unknown

List the buildings that are heated during the winter, how they are heated (e.g. space heaters, forced air furnace, heat pump, infrared lamps, etc.), and why they are heated (e.g. water tanks need heating, cold sensitive equipment storage, operator comfort, etc.).

____

Which buildings are air conditioned and what type of cooling do they use? (e.g. central air, window units, other)

Energy Use

Please enter the total amount of energy used (electricity, natural gas, etc.) and total energy cost from the past 36 billing cycles or the past 36, consecutive billing cycles available. Please attach in the Example Spreadsheet posted on the Drive. We are mostly concerned with consumption rather than how much it cost.

For natural gas and electric use, list the number of meters, their locations, and what equipment they measure.

Did you confirm each meter number listed on the utility bills with the meters at the plant?

☐Yes ☐No

Take pictures of the meters with their meter numbers being readable and attach them to this Assessment Form. If a photo of a meter number cannot be taken (lift station), please make a note explaining why.

Were there any extraordinary events that occurred during the billing cycle above that effected plant energy use? (Examples being extremely cold winter, drought, malfunctioning equipment, etc.)

☐Yes ☐No ☐Unknown

If yes, please explain.

____

____

____

____

Operator Information

Is the operator/community looking to implement any energy efficiency measures (Wanting to replace a motor, install LEDs, install VFDs, etc.)
☐Yes ☐No

If yes, please explain.

____

____

____

____

Does the operator have any energy efficiency suggestions besides what they are already looking to implement?
☐Yes ☐No

If yes, please explain them below.

____

____

____

____

How many people work at the plant? (Number of operators)_________________________

How long has each person been working at the plant?______________________________

____

____

____

Have there been any staff changes at the plant in the past 5 years or so? If yes, explain.

☐Yes ☐No

____

____

____

____

Who should we send the final benchmarking results to?

Name:____

Position:____

Mailing Address:____

____

Phone:____

Additional Notes: In this section, also note any E2 suggestions you have as you walk through the plant.

Wastewater Energy Assessment Form Appendix

Design Flow Source:

☐ NDEQ Fact Sheet, Document Date:______________________________
☐ NPDES Permit Application, Date:_______________________________
☐ NDEQ Inspection Sheet, Date of Inspection:______________________
☐ NDEQ DMR, Dates of DMR's used:_____________to______________
☐ Other:____

Current Average Effluent Flow Source:

☐NDEQ Fact Sheet, Document Date:______________________________
☐ NPDES Permit Application, Date:_______________________________
☐ NDEQ Inspection Sheet, Date of Inspection:______________________
☐ NDEQ DMR, Dates of DMR's used:_____________to______________
☐ECHO, Dates used:___________________to_______________________
☐ Other:____

Source(s) of Plant Loading Data:

Influent BOD:
☐ ECHO, Dates used to calculate average:______________to___________
☐ NDEQ Inspection Sheet, Date:______________________
☐ NDEQ DMR, Dates of DMR's:_____________________to___________

Effluent BOD:
☐ ECHO, Dates used to calculate average:______________to___________
☐ NDEQ Inspection Sheet, Date:___________________________________
☐ NDEQ DMR, Dates of DMR's:_____________________to___________

Influent TSS:
☐ ECHO, Dates used to calculate average:______________to___________
☐ NDEQ Inspection Sheet, Date:___________________________________
☐ NDEQ DMR, Dates of DMR's:_____________________to___________

Effluent TSS:
☐ ECHO, Dates used to calculate average:______________to___________
☐ NDEQ Inspection Sheet, Date:___________________________________
☐ NDEQ DMR, Dates of DMR's:_____________________to___________

Effluent NH_3-N:
☐ ECHO, Dates used to calculate average:______________to___________
☐ NDEQ Inspection Sheet, Date:___________________________________
☐ NDEQ DMR, Dates of DMR's:_____________________to___________

Source of Building area:

☐ On-site measurement ☐Google Earth Estimate
☐ Local Staff Estimate ☐On-site Estimate

Table B.1: Example Assessment Spreadsheet

Energy Consumption														
Electric Bill Dates							Natural Gas Bill Dates							
Starting Date	Ending Date	Monthly Electricity Consumption (kWh)	Monthly Electric Cost (Cost for kWh usage only) ($)	Monthly Electric Demand (kW)	Monthly Demand Cost ($)	Total Monthly Electricity Cost (Demand cost+kWh cost) ($)	Starting Date	Ending Date	Monthly Natural Gas Consumption (therms)	Monthly Natural Gas Cost ($)	Date Propane Tank was Filled	How much Propane was Filled (gal)	Propane Cost ($)
		Meter #		Meter #					Meter #				

Water Quality								Climate			
Influent Flow (MGD)	Effluent Flow (MGD)	Influent CBOD (mg/L)	Influent TSS (mg/L)	Effluent CBOD (mg/L)	Effluent TSS (mg/L)	Effluent NH_3-N (mg/L)	Current Discharge Limit for NH_3-N (mg/L)	Monthly HDD	Community HDD was taken from	Monthly CDD	Community CDD was taken from

Appendix C: Example Site Visit Narrative

Example Site Visit Narrative-Minden

On Friday, May 20th, 2016, Steven Hanna, Matt Thompson, Ranil Philipose, Jackson Micek, and David Hansen visited the wastewater treatment plant in Minden, NE for an energy assessment. They met Ryan Hurst, the operator, at the plant who was able to show them around and answer questions about the processes, operation, and energy usage at the plant. The plant is a SBR type plant that has an average flow of 0.196 MGD. The plant was originally built as a conventional activated sludge plant, but was converted into a SBR plant in 1999. The wastewater first passes through a comminutor and is then diverted to one of the two SBR basins. The treatment sequence for one cycle begins with mix fill where wastewater enters the basin while the mixer is on. Next, the blower for the basin turns on while the basin continues to fill. This stage is called the react fill sequence. After the basin is full, the influent is diverted to the other basin and both the blower and the mixer continue to operate during what is called as the react sequence. After the react sequence, the blower and mixer turn off and the solids settle during the settling sequence. Once the solids are settled, the clear supernatant is decanted and discharged to a nearby stream during the decant sequence. Once the effluent has been discharged, sludge is wasted to one of the two sludge lagoons and the cycle starts over again.

The visit to Minden on May 20th started off with the investigators confirming with Ryan Hurst flow data and water quality characteristics found before the visit from NDEQ and EPA resources online. All of the data were confirmed to be correct and the investigators started to go through the missing data on the Wastewater Facility Energy Assessment survey form. Ryan was very helpful in providing all of the data needed for the survey form and keeps excellent records of both water quality characteristics and maintenance records. The investigators feel very confident in the data Ryan was able to provide. Ryan became the operator for Minden in May of 2015, but was previously an operator at the Seward, NE wastewater plant. He was very knowledgeable and is very active with young professionals in Nebraska in the water and wastewater industry. The previous operator had been there since the plant was converted into a SBR and was not as concerned with energy use as Ryan has been. The difference between Ryan and the previous operator can be seen in the electric bills from the past year. There is a large difference between the electric consumption of March and April of 2015 and March and April of 2016. One example of the difference between Ryan and the previous operator was that the previous operator left heaters on in rooms that did not need heating.

In regards to energy usage, the main user of energy is the blowers. One of the blowers was replaced in 2015 and performs much more efficiently than the other two. The other two blowers have had regular maintenance, but they are both the same blowers installed in the 1999 plant conversion. Ryan said he has been having trouble maintaining DO levels in the first basin because it is aerated by one of the older blowers. The plant has the capability to operate using DO levels, but Ryan has not been able to use this capability because of the inefficiency of the old blowers. This results in unnecessarily high DO levels at night. The higher levels of DO at night are an example of wasting energy. If a more efficient blower was installed, Ryan could operate the aeration by using automatic DO controls and therefore save energy, especially at night.

Another area for energy savings is the steel maintenance building located on-site. The steel maintenance building was built in 1972 when the original plant was built. It is lightly insulated, but there were some gaps in the insulation found. The building needs to be heated during the winter because vacuum and jetter trailers are stored in the building and must always have water in them and their diesel engines need to be kept warm in order to be ready for emergencies. Currently, the building is heated using two electric heaters set to run via thermostats. It is a rather large building and consumes a large portion of the heating costs during the winter. Ryan said they rarely use the building in the winter and the only reason it is heated is because of the vacuum and jetter trailers. A possible recommendation for saving on energy is to heat only the trailers and not the whole building. One could do this by using an engine block heater for the diesel engines and a water tank heater for the water tank. Another recommendation is to install better insulation.

In regards to lighting, all of the lighting on-site uses fluorescent bulbs. The plant could install LEDs to save on energy, but the payback might be too long to implement this recommendation. An area that might merit LEDs regardless of payback is the steel building. It currently uses fluorescent lighting, but it provides poor lighting and therefore poor working conditions. Ryan stated one of the reasons they do not spend a lot of time working in the maintenance building is because of the poor lighting. Installing LEDs in the steel maintenance building would provide more lighting and safer working conditions.

In conclusion, the main areas for improved energy efficiency in the future are the blowers and heating of the steel maintenance building. It is recommended to replace one of the old blowers with a newer, more efficient one. It is also recommended to heat only the trailers in the steel building instead of the whole building itself. It is believed that these are the main areas of concern in regards to energy usage and should be further investigated.

Appendix D: Median Energy and Electric Intensity by Plant Type

Table D.1: Median Energy Intensity by Plant Type for Nebraska and Pennsylvania WWTFs with Flows <1.5 MGD.

Plant Type	Nebraska Plants		Pennsylvania Plants
	Median Energy Intensity (MWh/MG) (# of plants)	Range (MWh/MG)	Median Energy Intensity (MWh/MG)* (# of plants)
Seq. Batch Reactor	5.6 (13)	2.0-12.5	6.1 (19)
Ox. Ditch	5.2 (26)	1.8-11.7	4.5 (7)
Ext. Aeration	6.5 (38)	1.2-22.1	11.8 (26)
Trickling Filter	2.8 (6)	1.2-8.0	0.8 (6)
Other	4.6 (12)	0.6-7.8	4.5 (23)
All Plants	5.4 (95)	0.6-22.1	5.6 (81)

*Only electricity usage was obtained for Pennsylvania plants, therefore Energy Intensity here denotes the electric usage per unit flow.

Table D.2: Median Electric Intensity by Plant Type for Nebraska and Pennsylvania WWTFs with Flows <1.5 MGD.

Plant Type	Nebraska Plants		Pennsylvania Plants
	Median Electric Intensity (MWh/MG) (# of plants)	Range (MWh/MG)	Median Electric Intensity (MWh/MG) (# of plants)
Seq. Batch Reactor	5.6 (13)	2.0-9.5	6.1 (19)
Ox. Ditch	5.1 (25)	1.8-11.9	4.5 (7)
Ext. Aeration	6.2 (38)	1.2-20.9	11.8 (26)
Fixed Film	3.1 (13)	0.6-6.2	0.9 (9)
Other	3.7 (5)	1.7-5.8	4.3 (20)
All Plants	5.2 (94)	0.6-20.9	5.6 (81)

Appendix E: SAS Code and Output

Example SAS Code with Annotations:

The data analysis for this research was done using SAS® 9.4. The PROC REG procedure was used to create the multiple linear regression models.

```
/* Read in Data */
proc import out=wastewater
datafile="C:\Users\Steven\Documents\Grad Research\Data Analysis\Data
Sets\Wastewater Data_Sldg.xlsx"
dbms=xlsx REPLACE;
getnames=YES;
run;
proc print; run;
```

This first block of code reads the data set into SAS from an Excel file. For the different models created, different files were imported. The only line that changes in this first block is the location and file name of the data set that is imported.

```
/* Kendall tau Correlations */
proc corr data=wastewater best=5 Kendall;
run;
```

The Correlations block of code determines which variables are highly correlated to each other. A value closer to 1 or -1 meant the two variables were very highly correlated to each other. Many of the original predictors were deleted or consolidated into fewer variables after looking at the correlations in order to avoid multicollinearity in the model.

```
/* Distributions of Continuous Variables */
proc kde data=wastewater;
univar Design_Flow / plots= (density);
univar Climate_Control_Floor_Area / plots= (density);
univar Avg_Eff_Flow / plots= (density);
univar Avg_Eff_CBOD / plots= (density);
univar Avg_Eff_TSS / plots= (density);
univar Avg_Eff_NH3N / plots= (density);
univar Annual_Sum_HDDs / plots= (density);
univar Annual_Sum_CDDs / plots= (density);
univar Percent_Design_Flow / plots= (density);
univar Annual_Electric_Usage / plots= (density);
univar Energy_Intensity_Flow / plots= (density);
univar Energy_Intensity_Elec / plots= (density);
run;
```

This next block of code checks the distributions of the continuous variables. The output gives Gaussian Kernal Density plots for each of the continuous variables. Transformation of the variables were made depending on if the distributions were non-normal.

```
/* Transformations */
data log_ww;
set wastewater;
lDF = log(Design_Flow);
```

```
lCC_Floor = log(Climate_Control_Floor_Area);
lAEF = log(Avg_Eff_Flow);
lAEC = log(Avg_Eff_CBOD);
lAET = log(Avg_Eff_TSS);
lAEN = log(Avg_Eff_NH3N);
lPDF = log(Percent_Design_Flow);
Log_Use = log(Total_2015_Energy_Usage);
Log_EI = log(Energy_Intensity_Flow);
Log_EIE = log(Energy_Intensity_Elec);
tHDD = Annual_Sum_HDDs**3;
drop Design_Flow  Climate_Control_Floor_Area
     Avg_Eff_Flow Avg_Eff_CBOD Avg_Eff_TSS
     Avg_Eff_NH3N Percent_Design_Flow
     Total_2015_Energy_Usage Energy_Intensity_Elec
Energy_Intensity_Flow
       Annual_Sum_HDDs Annual_Nat_gas;
run;
proc print; run;
```

The transformations block of code creates a new data set that transforms the continuous variables depending on how skewed their distributions were. Generally, the natural log transformation was used because much of the continuous data had skewed right distributions.

```
/* Stepwise Selection */
proc reg data=log_ww outest=betas covout
      plots(label) = (CooksD RStudentbyLeverage);
id Facility_Community;
model Log_EI = IND_Load--lPDF tHDD / selection=stepwise vif;
output out=pred p=phat;
run;
```

The Stepwise Selection block of code uses stepwise variable selection that chooses the best model according to the F statistics of the variables. Variables can be added and deleted in multiple steps throughout the selection. The output gives the final model, but further interpretation of this model is required. The output also provides diagnostic plots and labels any outliers and points of high leverage. The diagnostic plots are examined for model validity and some of the outliers are deleted after further investigation.

```
/* Remove Outliers */
data log_ww_noouts;
set log_ww;
IF Facility_Community= "Newcastle WWTF" then delete;
IF Facility_Community= "Greenwood WWTF" then delete;
IF Facility_Community= "Pender WWTF" then delete;
IF Facility_Community= "Petersburg WWTF" then delete;
IF Facility_Community= "Wood River WWTF" then delete;
run;
proc print; run;
```

Outliers are deleted in this block of code and a new data set is created for further analysis.

```
/* Re-run Stepwise Selection without outliers */
proc reg data=log_ww_noouts outest=betas covout plot(label)=(CooksD
RStudentbyleverage Diagnostics);
```

```
id Facility_Community;
model Log_EI = IND_Load--lPDF tHDD / selection=stepwise vif;
output out=pred p=phat;
run;
```

Stepwise selection of the data set without the outliers is conducted with this next block of code. The output is investigated and the diagnostic plots are examined once more.

```
/* Simplified Model */
proc reg data=log_ww_noouts outest=betas covout;
model Log_EI = EA_Bioreactor Sldg_Sup_Energy Dewater_Equip lCC_Floor
lAEF lPDF / vif;
output out=pred p=phat;
run;
```

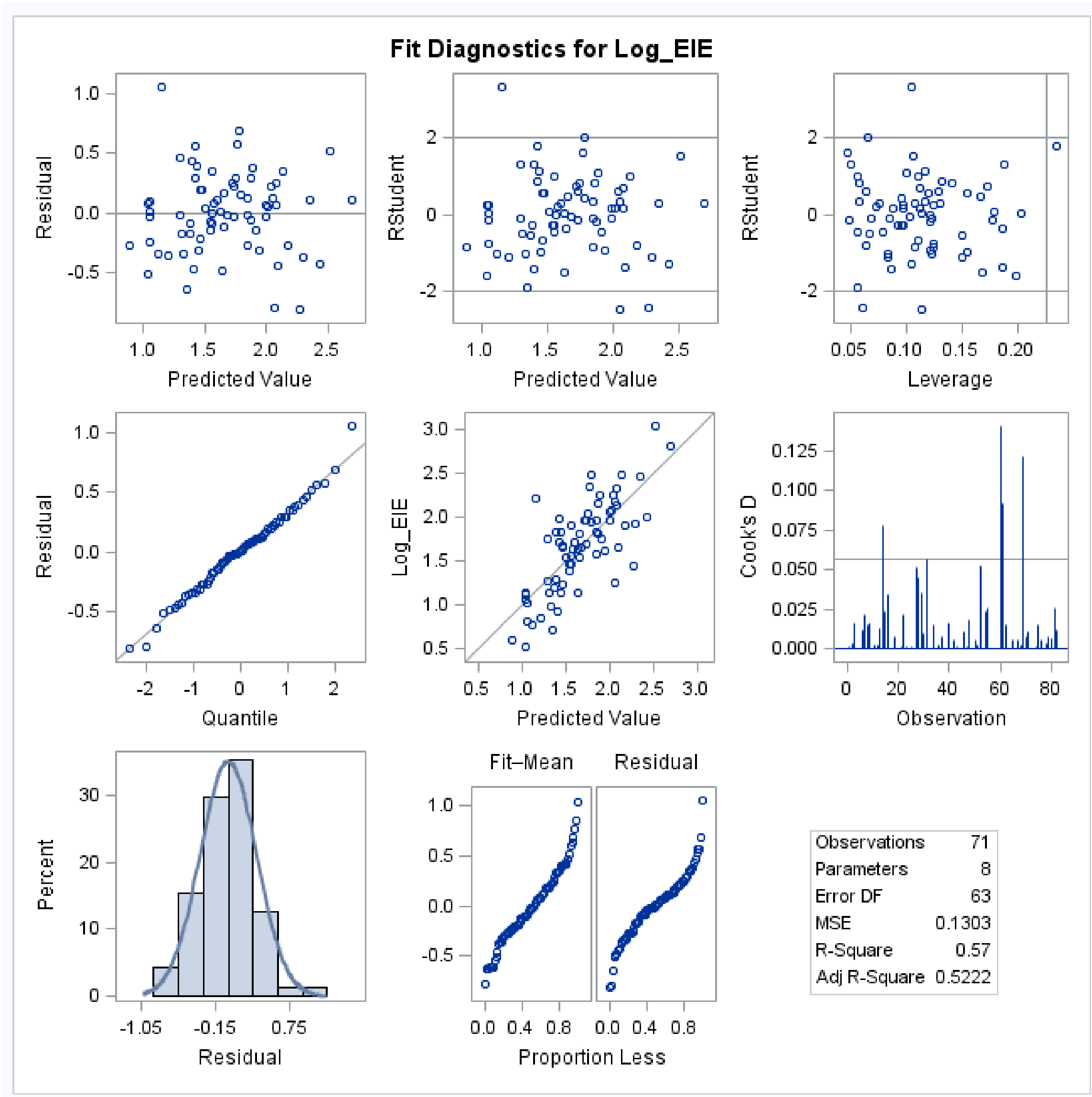
The final model variables are put into the model statement and the model is run without using the stepwise selection option.

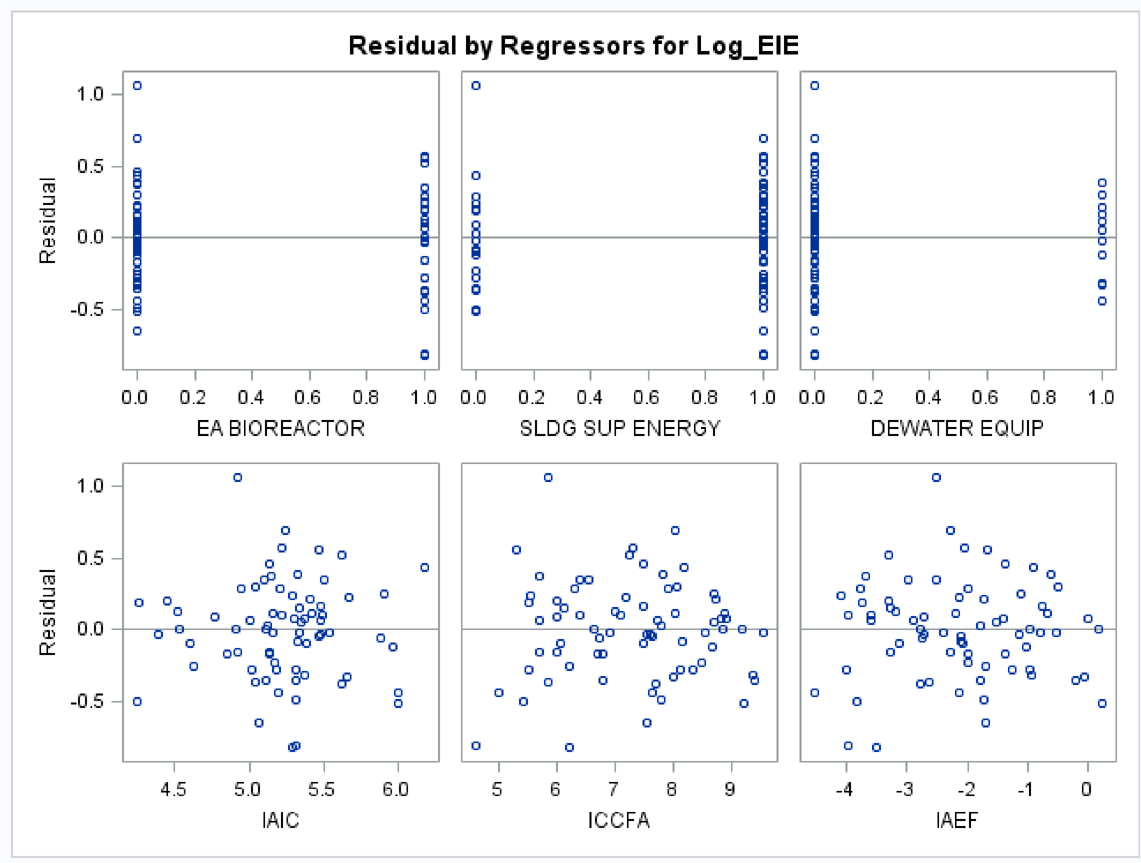
Models and Diagnostic Plots

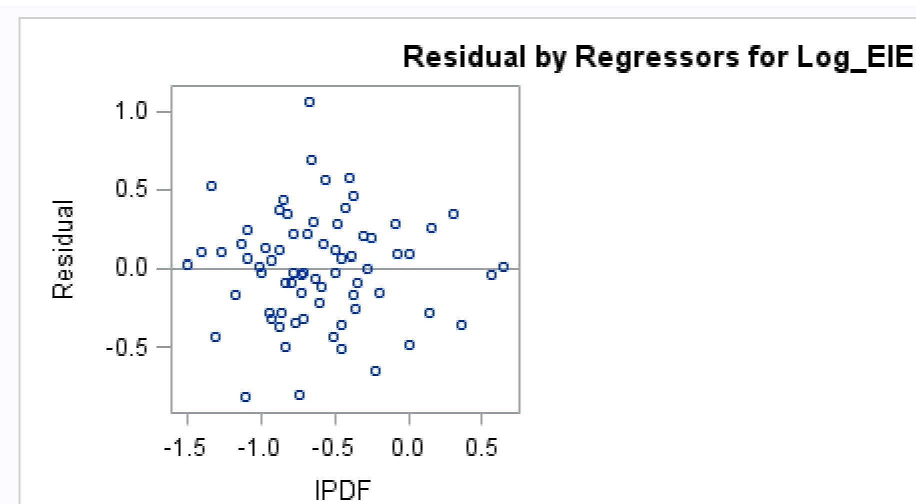
1. <u>NE Overall Plant Type Electric Intensity</u>
   - Response: Electric Intensity
   - Data Set: Wastewater Data_1_10_17
   - Deleted Greenwood, Pender, and Petersburg.
   - No strictly fixed film plants.

Variable	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	-2.06161	0.78361	-2.63	0.0107	0
EA_BIOREACTOR	0.25651	0.10673	2.40	0.0192	1.49957
SLDG_SUP_ENERGY	0.26408	0.10203	2.59	0.0120	1.07337
DEWATER_EQUIP	0.27243	0.13814	1.97	0.0530	1.36112
lAIC	0.25642	0.12191	2.10	0.0394	1.25006
lCCFA	0.16456	0.06239	2.64	0.0105	3.15278
lAEF	-0.32341	0.07290	-4.44	<.0001	3.99016
lPDF	-0.26770	0.11838	-2.26	0.0272	1.50273

Root MSE	0.36100	**R-Square**	0.5700
Dependent Mean	1.66614	**Adj R-Sq**	0.5222
Coeff Var	21.66699		


Fit Diagnostics for Log_EIE
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_EIE
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 71
Parameters 8
Error DF 63
MSE 0.1303
R-Square 0.57
Adj R-Square 0.5222

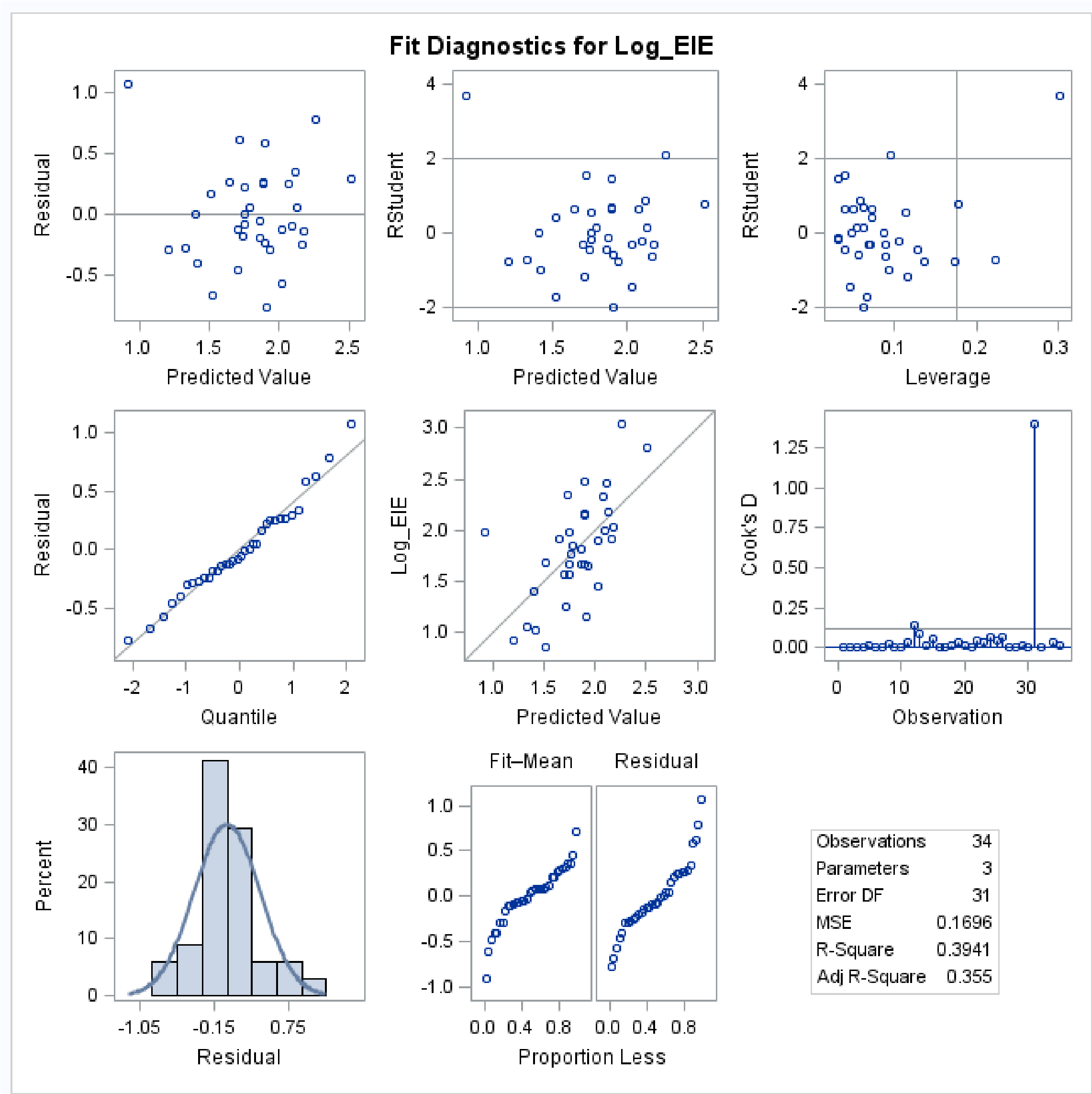

Residual by Regressors for Log_EIE
Residual
1.0
0.5
0.0
-0.5
0.0 0.2 0.4 0.6 0.8 1.0
EA BIOREACTOR
SLDG SUP ENERGY
DEWATER EQUIP
4.5 5.0 5.5 6.0
IAIC
5 6 7 8 9
ICCFA
-4 -3 -2 -1 0
IAEF

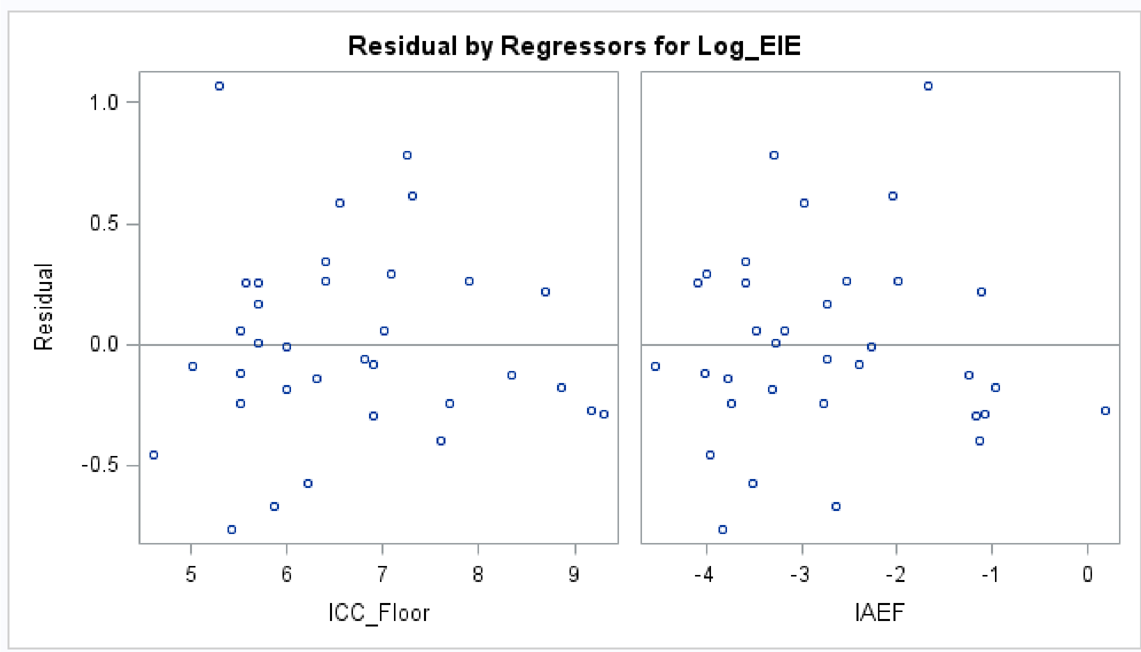

Residual by Regressors for Log_EIE
Residual
1.0
0.5
0.0
-0.5
-1.5 -1.0 -0.5 0.0 0.5
IPDF

2. <u>NE EA Electric Intensity</u>
    - Response: Electric Intensity
    - Data Set: Ext_Aeration
    - Deleted Greenwood, Pender, and Petersburg

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	-1.52355	0.84260	-1.81	0.0803	0
lCC_Floor	0.32009	0.09231	3.47	0.0016	2.51214
lAEF	-0.44383	0.09885	-4.49	<.0001	2.51214

Root MSE	0.41186	**R-Square**	0.3941
Dependent Mean	1.81141	**Adj R-Sq**	0.3550
Coeff Var	22.73692		


Fit Diagnostics for Log_EIE
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_EIE
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 34
Parameters 3
Error DF 31
MSE 0.1696
R-Square 0.3941
Adj R-Square 0.355

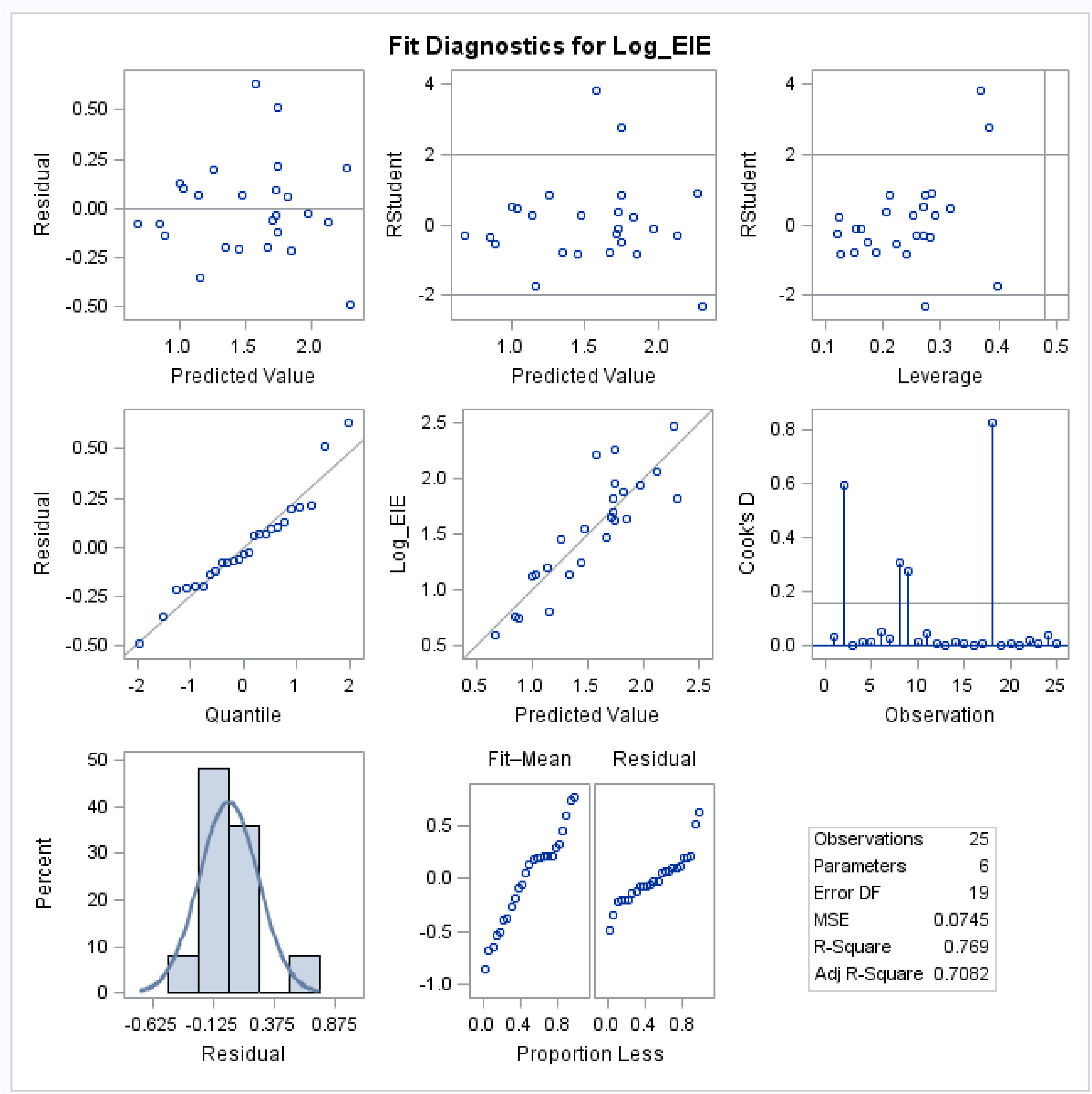


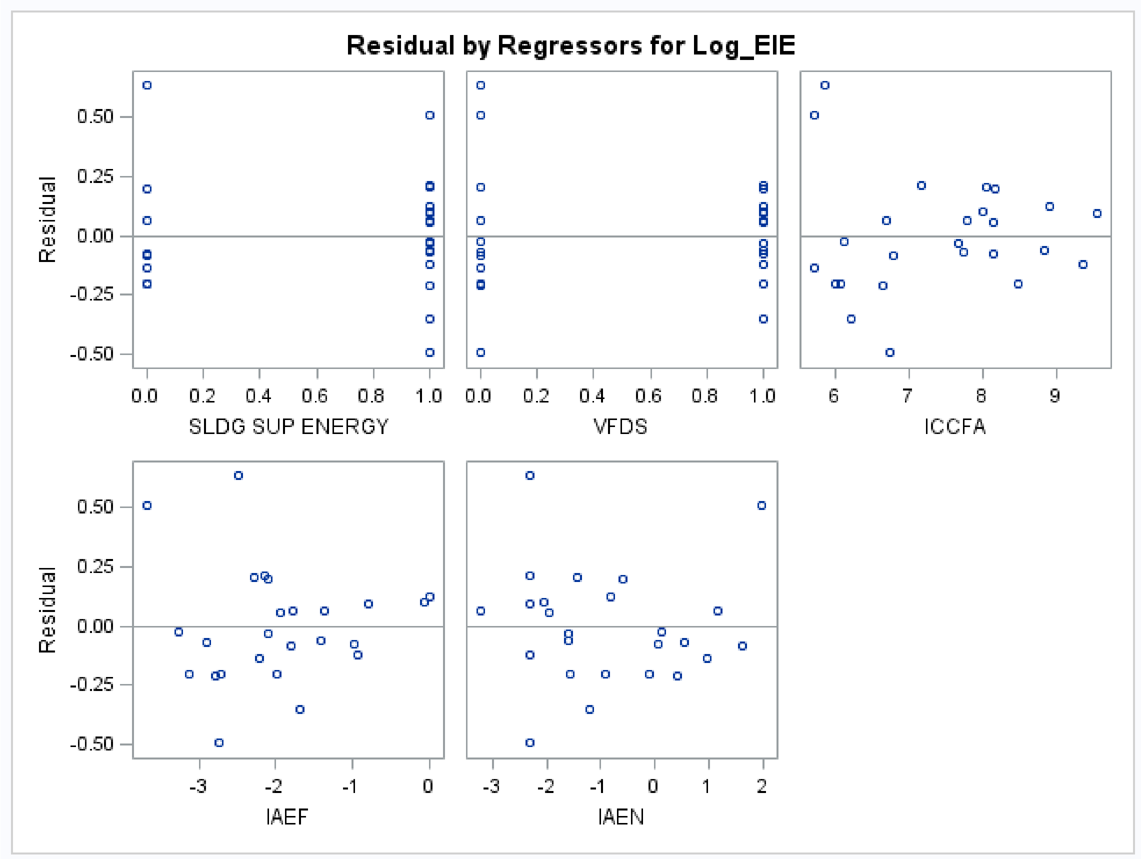
Residual by Regressors for Log_EIE
Residual
1.0
0.5
0.0
-0.5
5
6
7
8
9
ICC_Floor
-4
-3
-2
-1
0
IAEF

3. NE OD Electric Intensity Model
   - Data Set: OD Plants
   - Response: Electric Intensity
   - Could possibly delete Randolph and Arnold because they are large outliers, but I don't know why.

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	-1.21445	0.63155	-1.92	0.0696	0
SLDG_SUP_ENERGY	0.41690	0.11800	3.53	0.0022	1.07603
VFDS	-0.40380	0.18634	-2.17	0.0431	2.90664
lCCFA	0.23247	0.07744	3.00	0.0073	2.69708
lAEF	-0.41941	0.08906	-4.71	0.0002	2.24417
lAEN	-0.16649	0.04935	-3.37	0.0032	1.57560

Root MSE	0.27302	**R-Square**	0.7690
Dependent Mean	1.53004	**Adj R-Sq**	0.7082
Coeff Var	17.84406		


Fit Diagnostics for Log_EIE
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_EIE
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 25
Parameters 6
Error DF 19
MSE 0.0745
R-Square 0.769
Adj R-Square 0.7082

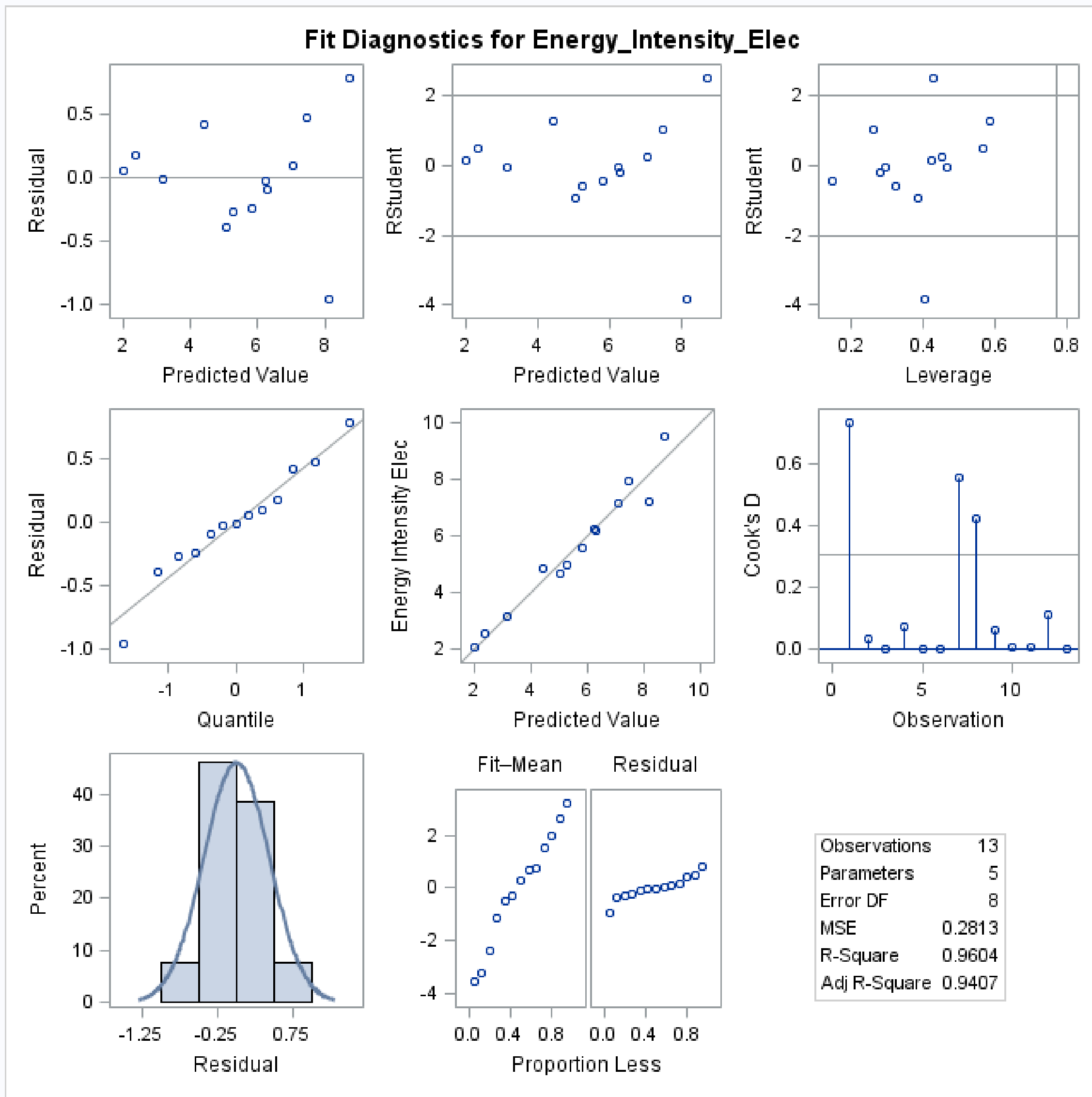


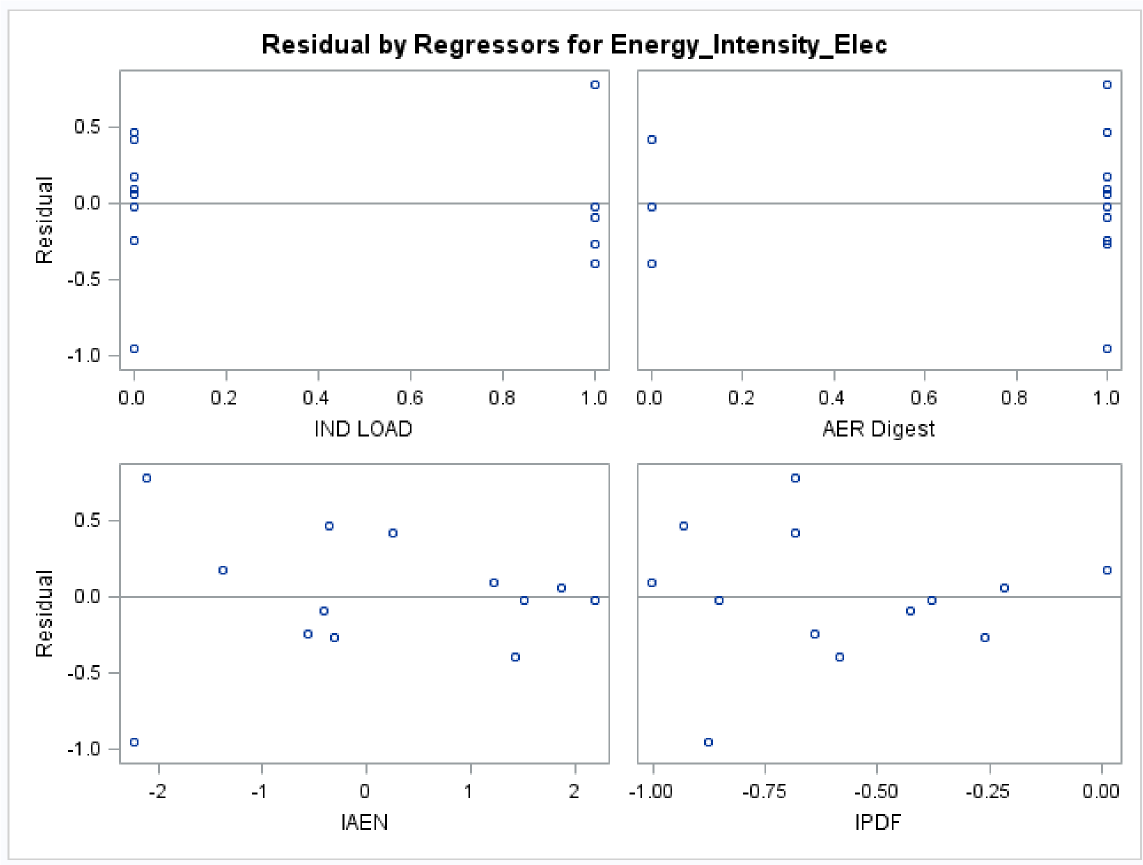
Residual by Regressors for Log_EIE
Residual
0.50
0.25
0.00
-0.25
-0.50
0.0
0.2
0.4
0.6
0.8
1.0
SLDG SUP ENERGY
VFDS
6
7
8
9
ICCFA
-3
-2
-1
0
IAEF
1
2
IAEN

4. <u>NE SBR Electric Intensity</u>
   - Data Set: SBR Plants
   - Response: Electric Intensity<= Actually EIE, not transformed
   - No outliers deleted
   - No inf data used.
   - Dropped UV because it had opposite sign. Dropped CDDs because while it was significant, it had a small effect on the overall R-sq.

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	0.42516	0.58879	0.72	0.4908	0
IND_LOAD	1.83436	0.32497	5.64	0.0005	1.15525
AER_Digest	1.23565	0.43766	2.82	0.0224	1.57155
lAEN	-0.53584	0.11905	-4.50	0.0020	1.31599
lPDF	-6.03809	0.52083	-11.59	<.0001	1.09496

Root MSE	0.53035	**R-Square**	0.9604
Dependent Mean	5.53375	**Adj R-Sq**	0.9407
Coeff Var	9.58386		

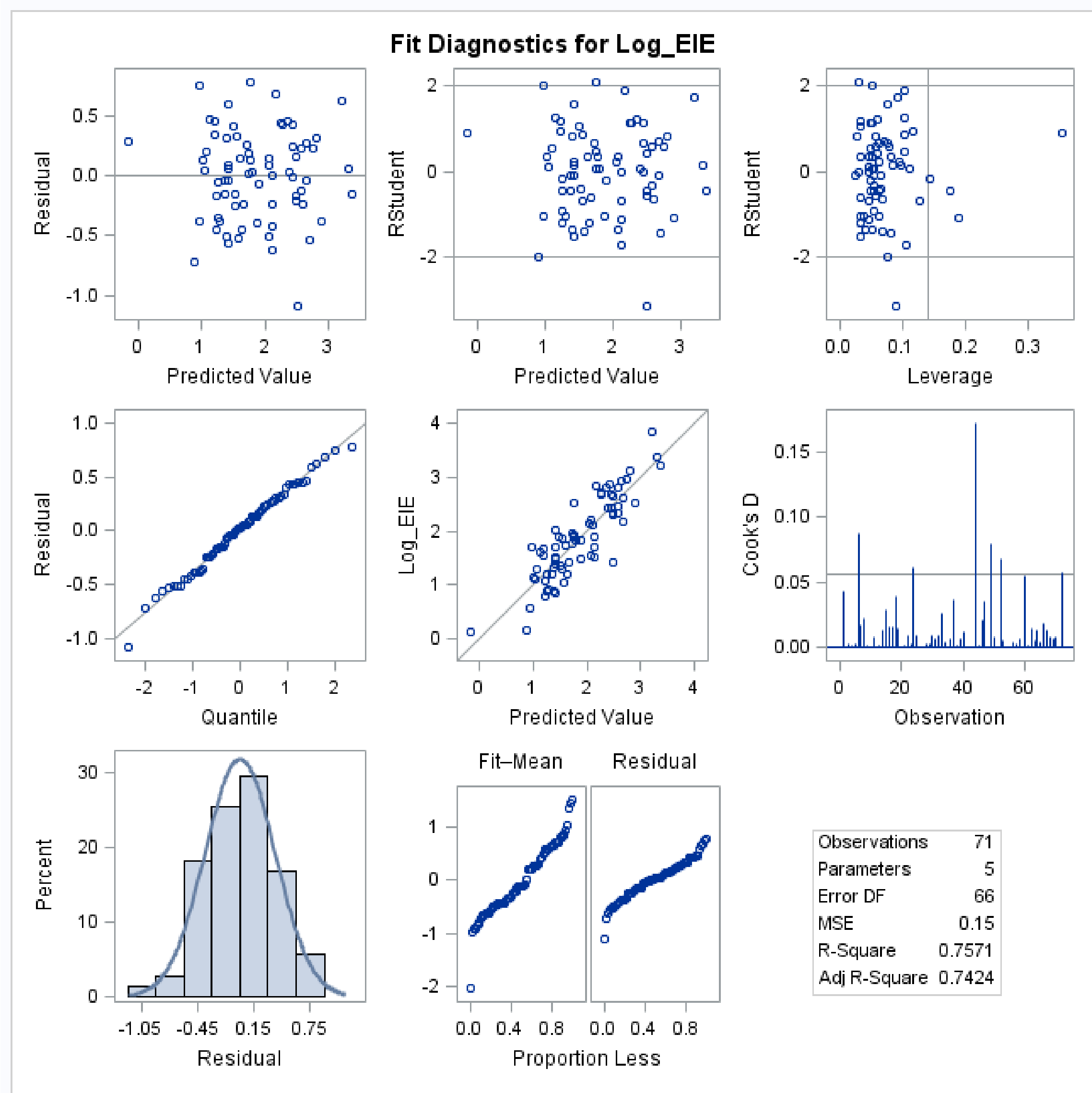

Fit Diagnostics for Energy_Intensity_Elec
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Energy Intensity Elec
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 13
Parameters 5
Error DF 8
MSE 0.2813
R-Square 0.9604
Adj R-Square 0.9407


Residual by Regressors for Energy_Intensity_Elec
Residual
0.5
0.0
-0.5
-1.0
0.0
0.2
0.4
0.6
0.8
1.0
IND LOAD
AER Digest
-2
-1
0
1
2
IAEN
-1.00
-0.75
-0.50
-0.25
0.00
IPDF

5. Overall Pennsylvania Electric Intensity Model

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	-2.37781	0.59993	-3.96	0.0002	0
UV	0.22182	0.10169	2.18	0.0327	1.07198
lAF	-0.31633	0.03434	-9.21	<.0001	1.17852
lBOD	0.63229	0.10847	5.83	<.0001	1.35070
lLC	-0.30630	0.12507	-2.45	0.0170	1.25229

Root MSE	0.38731	**R-Square**	0.7571
Dependent Mean	1.86745	**Adj R-Sq**	0.7424
Coeff Var	20.73986		

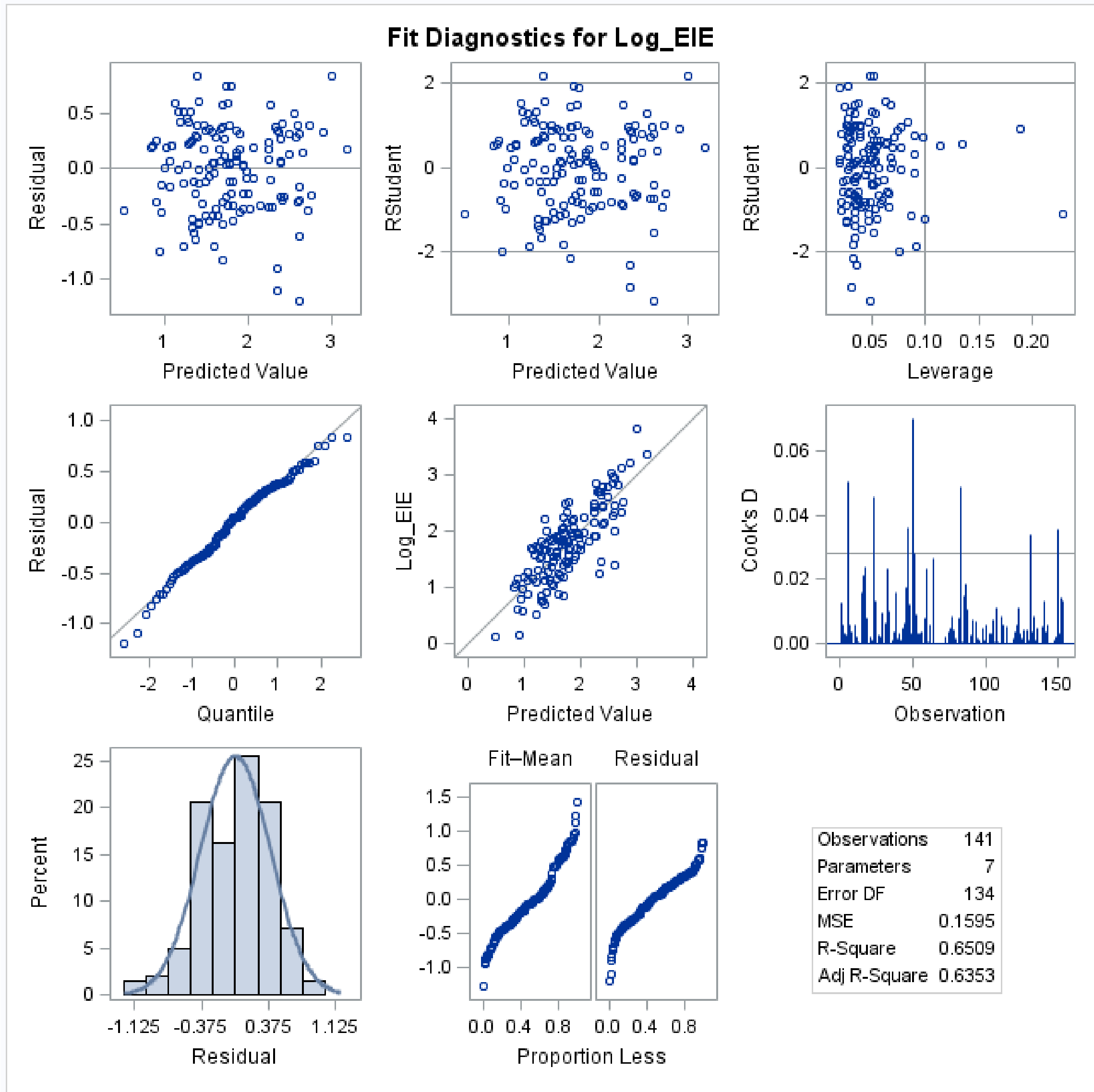

Fit Diagnostics for Log_EIE
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_EIE
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 71
Parameters 5
Error DF 66
MSE 0.15
R-Square 0.7571
Adj R-Square 0.7424

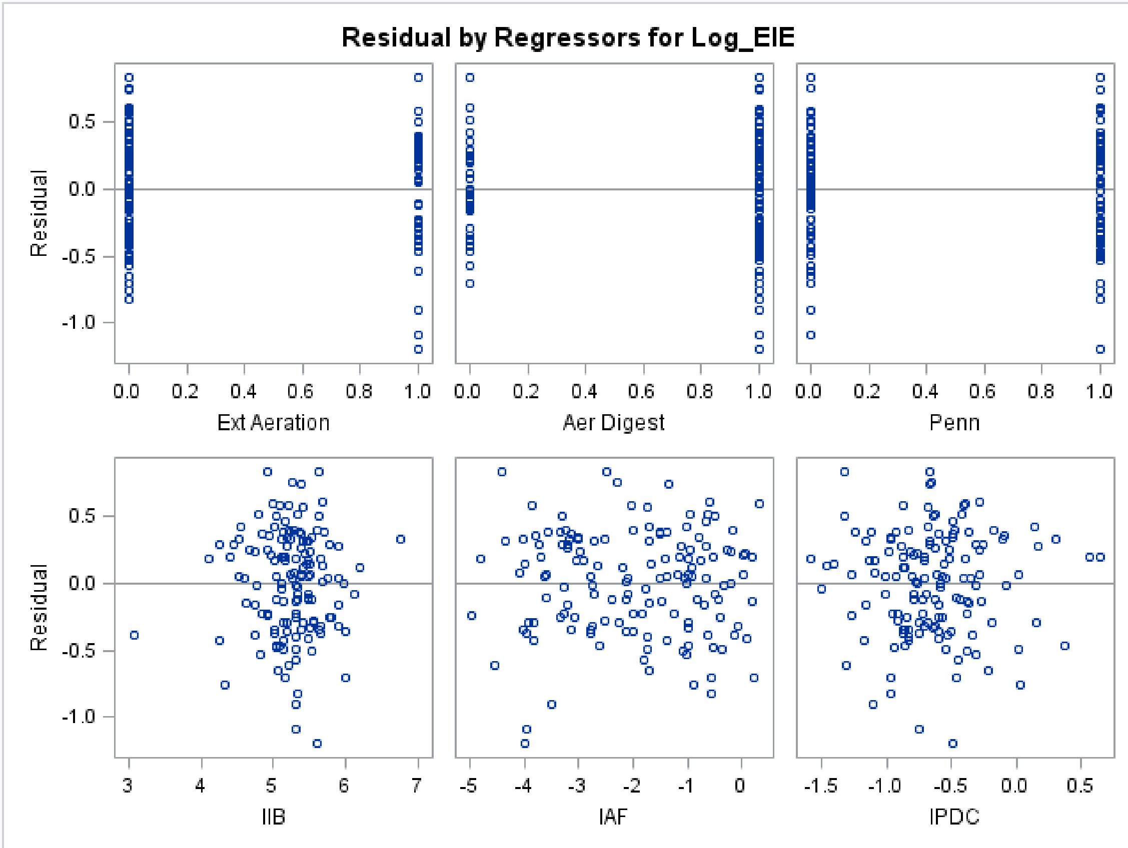

Residual by Regressors for Log_EIE
Residual
0.5
0.0
-0.5
-1.0
0.0
0.2
0.4
0.6
0.8
1.0
UV
-5
-4
-3
-2
-1
0
IAF
3
4
5
6
7
IBOD
-2.0
-1.5
-1.0
-0.5
0.0
ILC

6. NE and Penn Electric Intensity Combined Model
   - Response: Electric Intensity
   - Data Set: Penn and NE
   - Pennsylvania and Nebraska plants
   - No Strictly Fixed Film plants removed from NE plants. Fixed Film labeled plants removed from Penn data set.

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	-1.86477	0.40622	-4.59	<.0001	0
Ext_Aeration	0.18153	0.08727	2.08	0.0394	1.59115
Aer_Digest	0.23652	0.08452	2.80	0.0059	1.08338
Penn	0.20127	0.06948	2.90	0.0044	1.06679
lIB	0.49963	0.07907	6.32	<.0001	1.08598
lAF	-0.22578	0.03413	-6.62	<.0001	1.82076
lPDC	-0.32797	0.09568	-3.43	0.0008	1.23585

Root MSE	0.39936	**R-Square**	0.6509
Dependent Mean	1.76864	**Adj R-Sq**	0.6353
Coeff Var	22.57975		

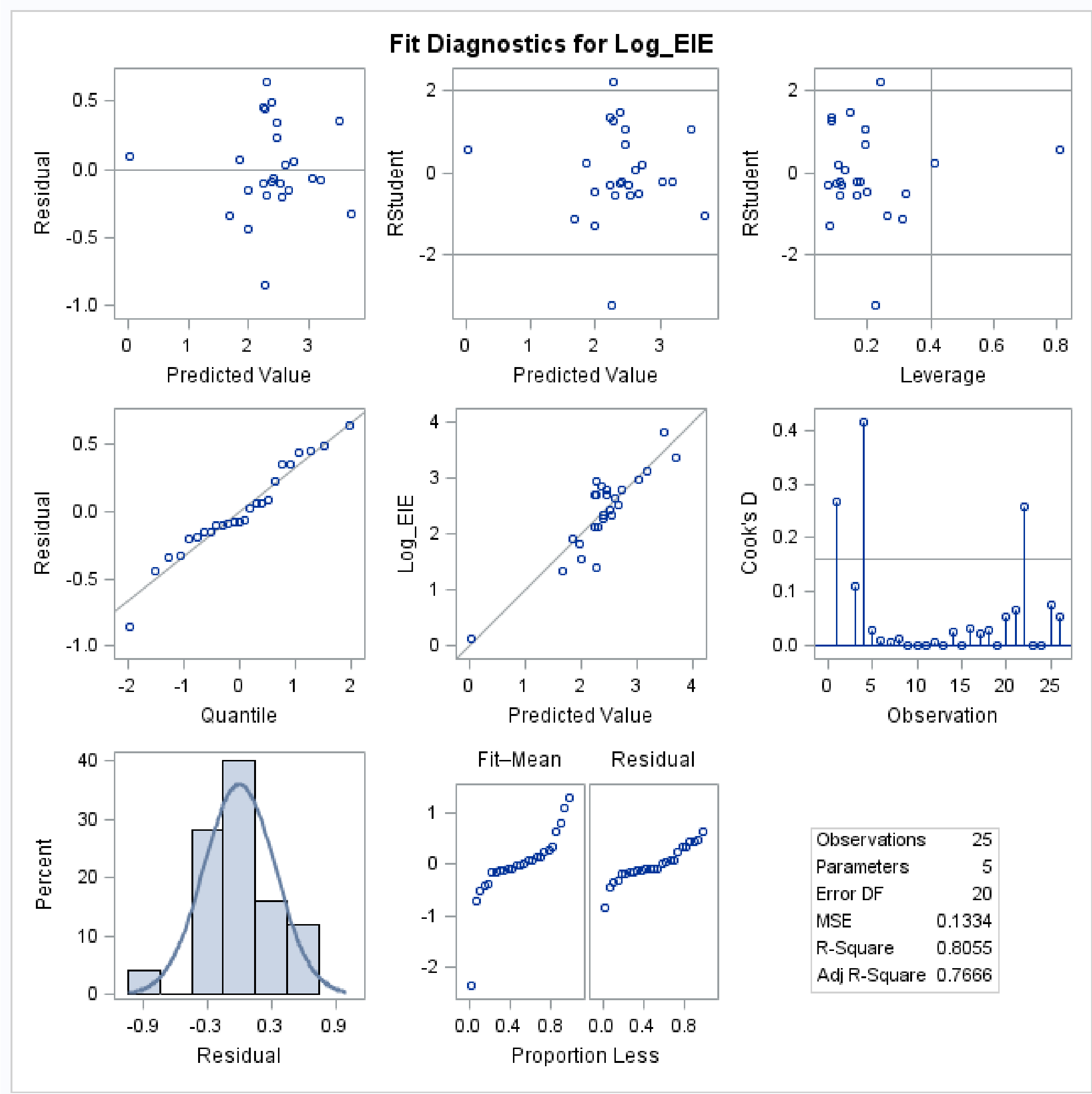

Fit Diagnostics for Log_EIE
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_EIE
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 141
Parameters 7
Error DF 134
MSE 0.1595
R-Square 0.6509
Adj R-Square 0.6353

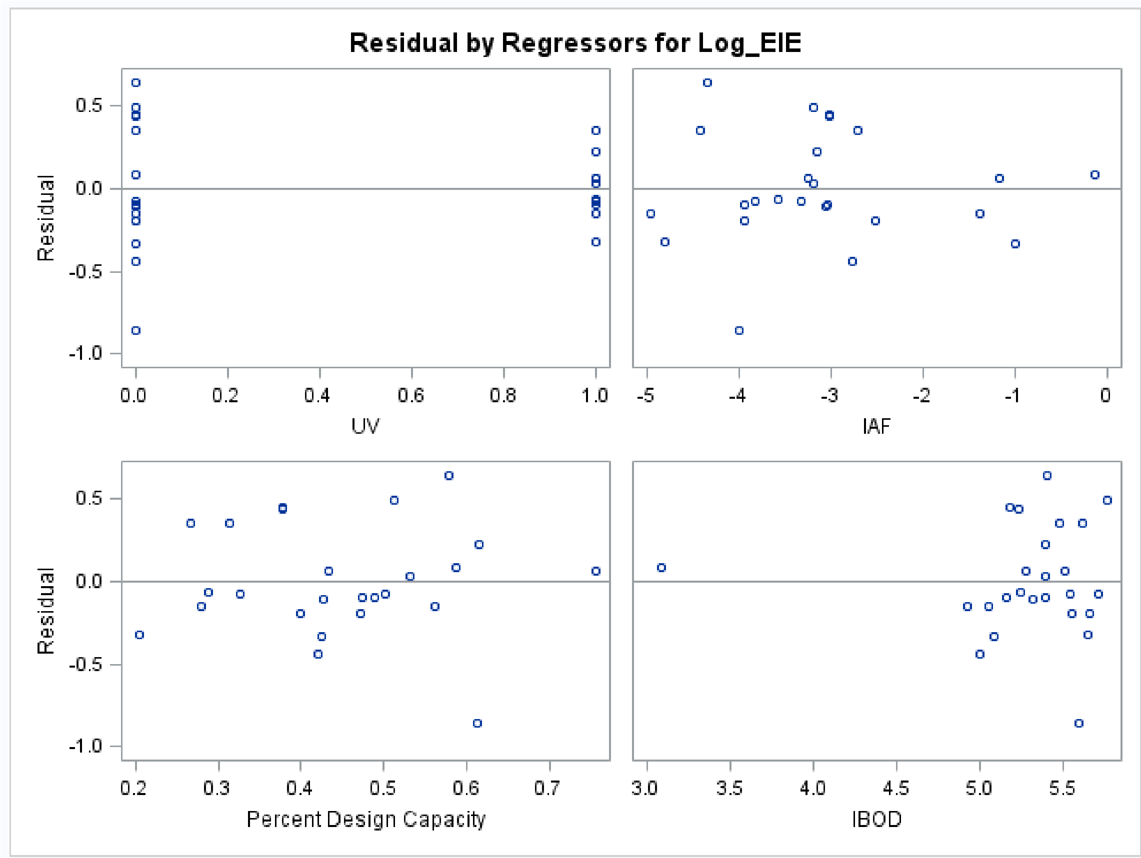

Residual by Regressors for Log_EIE
Residual
0.5
0.0
-0.5
-1.0
0.0 0.2 0.4 0.6 0.8 1.0
Ext Aeration
Aer Digest
Penn
3 4 5 6 7
IIB
-5 -4 -3 -2 -1 0
IAF
-1.5 -1.0 -0.5 0.0 0.5
IPDC

7. <u>Penn EA Model</u>
   - Response: Electric Intensity
   - Data Set: Penn EA
   - No outliers deleted
   - Dropped Eff_CBOD because 5 plants were missing this data.
   - Inf BOD in units of mg/L
   - Sludge Digestion not included because all use aerobic digestion.

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	-0.76168	0.87930	-0.87	0.3966	0
UV	0.49528	0.15262	3.25	0.0041	1.04790
lAF	-0.21137	0.08757	-2.41	0.0255	1.91419
Percent_Design_Capacity	-1.77527	0.65258	-2.72	0.0132	1.30200
lBOD	0.58450	0.18303	3.19	0.0046	1.59605

Root MSE	0.36520	**R-Square**	0.8055
Dependent Mean	2.38640	**Adj R-Sq**	0.7666
Coeff Var	15.30344		

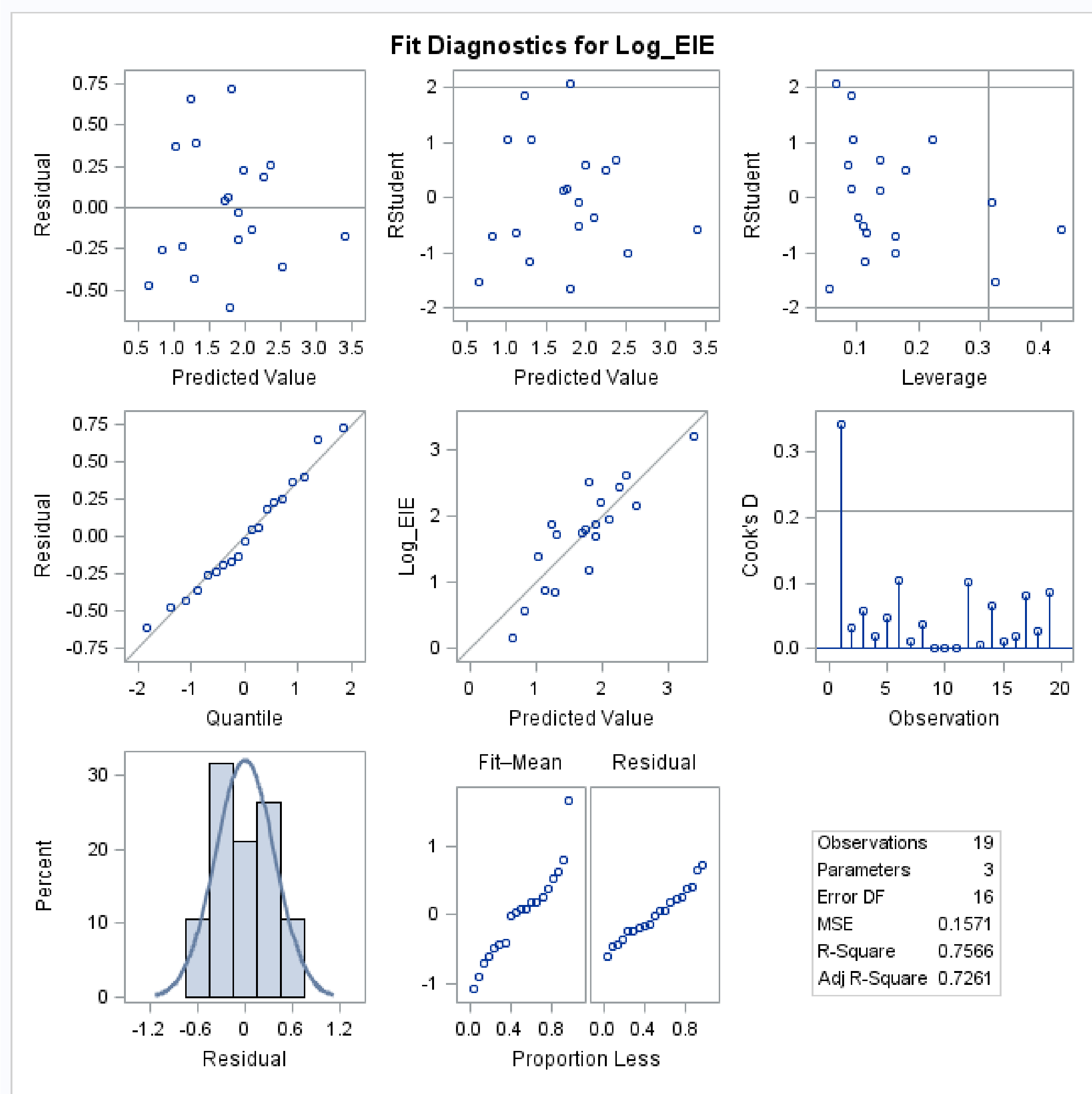

Fit Diagnostics for Log_EIE
Residual
Predicted Value
RStudent
Leverage
Quantile
Log_EIE
Cook's D
Observation
Percent
Fit–Mean
Residual
Proportion Less
Observations 25
Parameters 5
Error DF 20
MSE 0.1334
R-Square 0.8055
Adj R-Square 0.7666

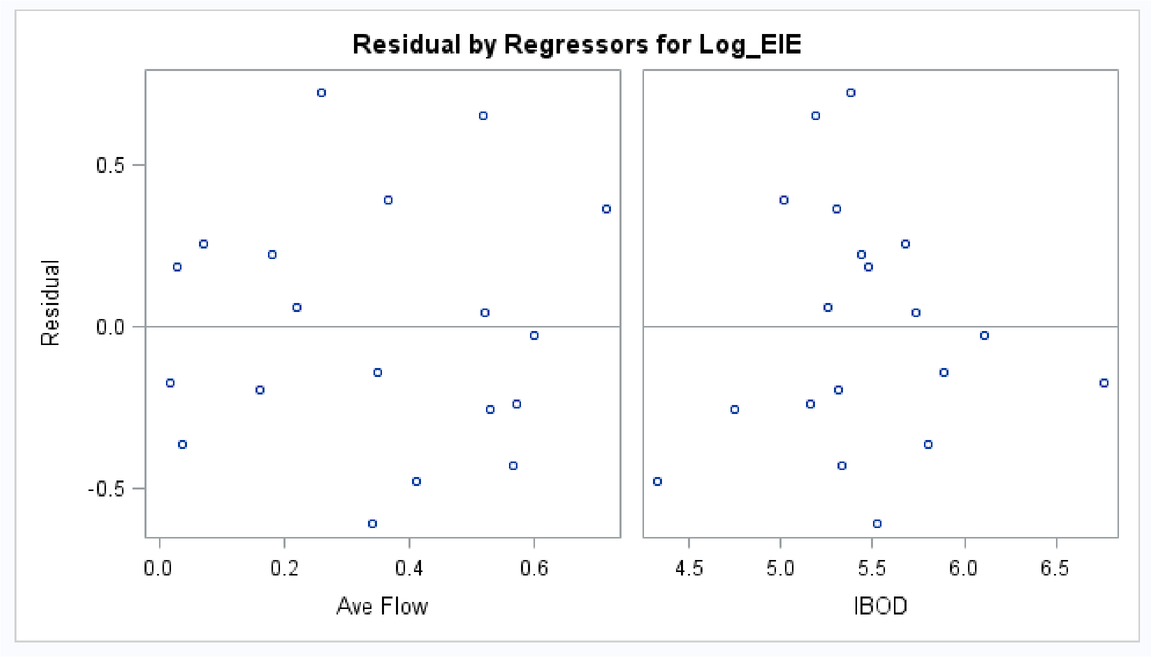

Residual by Regressors for Log_EIE
Residual
0.5
0.0
-0.5
-1.0
0.0
0.2
0.4
0.6
0.8
1.0
UV
-5
-4
-3
-2
-1
0
IAF
0.2
0.3
0.4
0.5
0.6
0.7
Percent Design Capacity
3.0
3.5
4.0
4.5
5.0
5.5
IBOD

8. <u>Penn SBR Model</u>
   - Response: Electric Intensity
   - Data Set: Penn SBR
   - Dropped Eff. CBOD because there were 2 plants with Eff. CBOD missing.
   - Inf. BOD in units of mg/L

Variable	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	-2.47961	1.11996	-2.21	0.0417	0
Ave_Flow	-1.57191	0.45677	-3.44	0.0034	1.14669
lBOD	0.87231	0.19313	4.52	0.0004	1.14669

Root MSE	0.39633	**R-Square**	0.7566
Dependent Mean	1.73499	**Adj R-Sq**	0.7261
Coeff Var	22.84326		

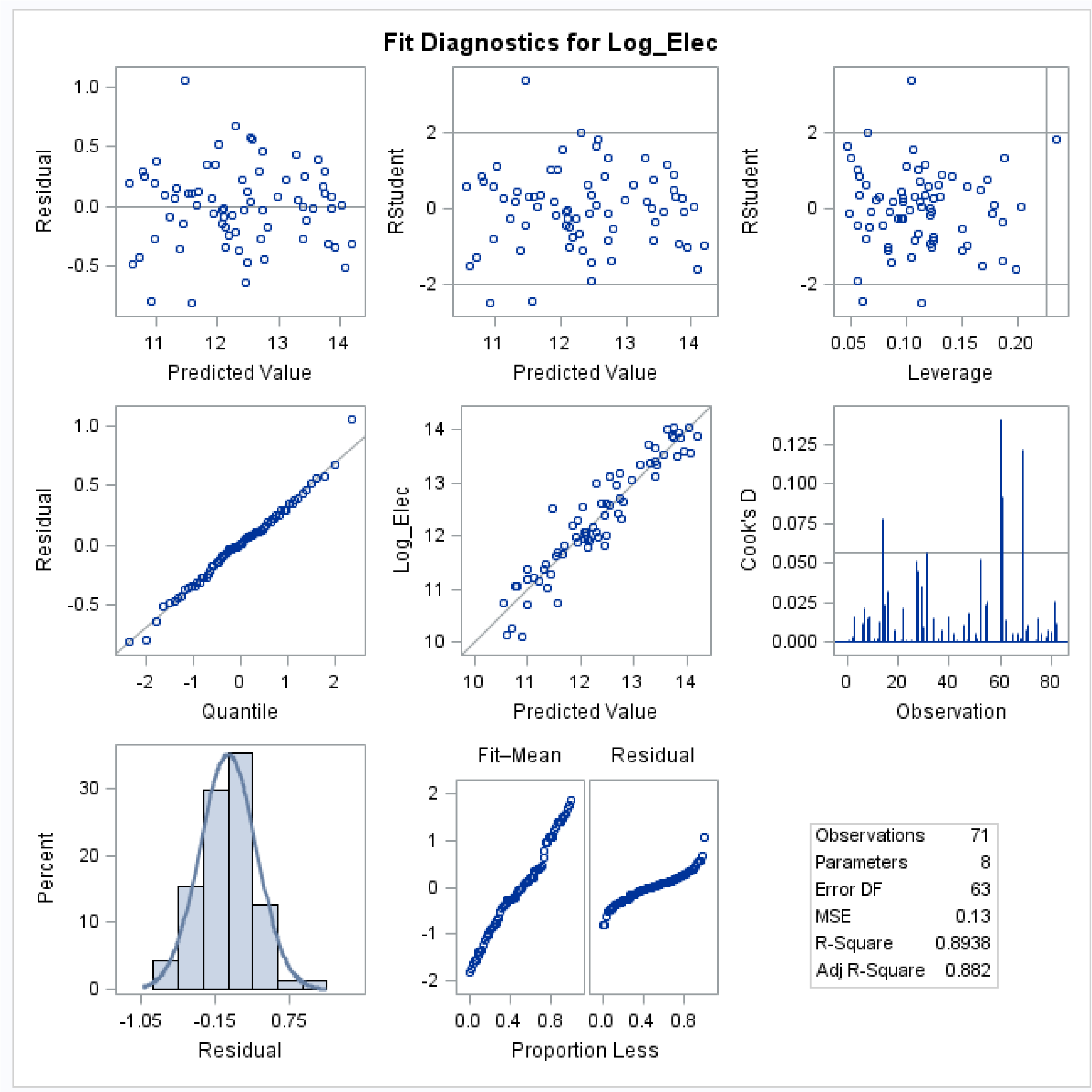

Fit Diagnostics for Log_EIE
Residual
Predicted Value
RStudent
Leverage
Quantile
Log_EIE
Cook's D
Observation
Percent
Residual
Fit–Mean
Proportion Less
Observations 19
Parameters 3
Error DF 16
MSE 0.1571
R-Square 0.7566
Adj R-Square 0.7261

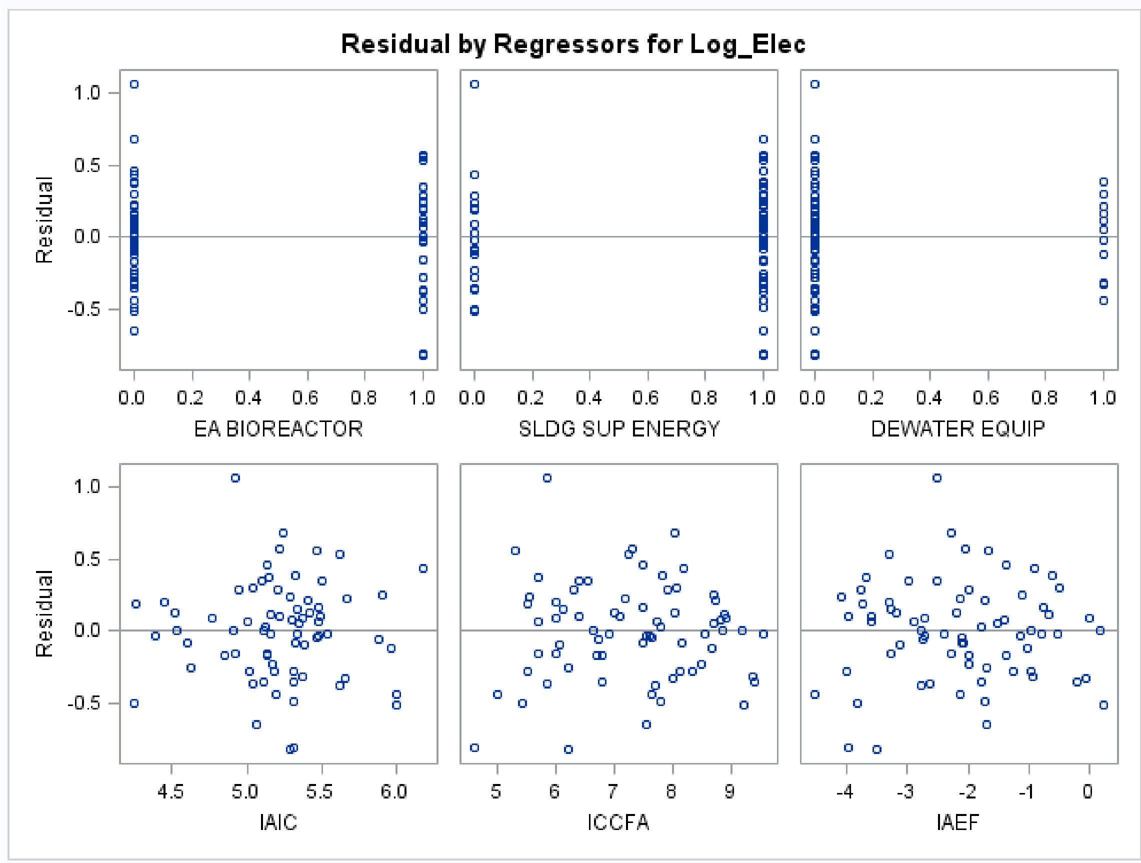

Residual by Regressors for Log_EIE
Residual
0.5
0.0
-0.5
0.0
0.2
0.4
0.6
Ave Flow
4.5
5.0
5.5
6.0
6.5
IBOD

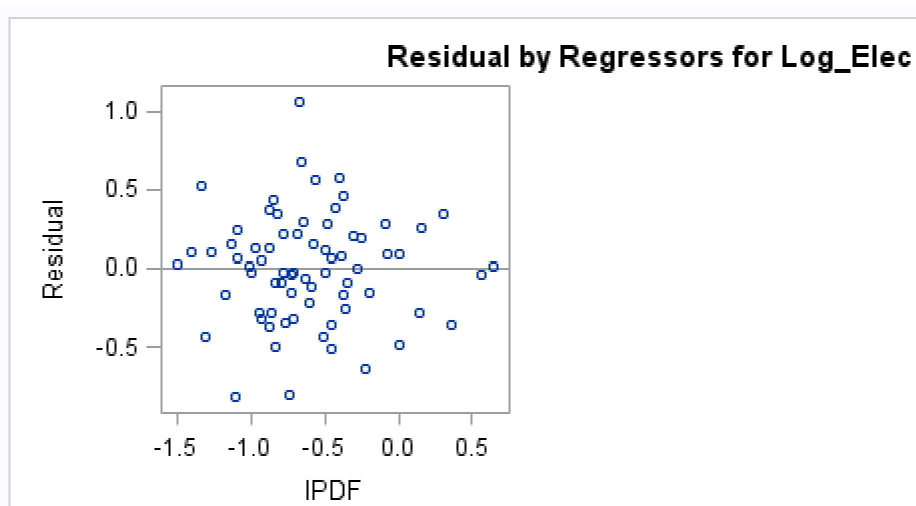
9. NE Electric Usage All Plant Types Model

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	10.74884	0.78270	13.73	<.0001	0
EA_BIOREACTOR	0.25709	0.10661	2.41	0.0188	1.49957
SLDG_SUP_ENERGY	0.26374	0.10191	2.59	0.0120	1.07337
DEWATER_EQUIP	0.27289	0.13798	1.98	0.0523	1.36112
lAIC	0.25646	0.12176	2.11	0.0392	1.25006
lCCFA	0.16421	0.06232	2.64	0.0106	3.15278
lAEF	0.67701	0.07282	9.30	<.0001	3.99016
lPDF	-0.26795	0.11824	-2.27	0.0269	1.50273

Root MSE	0.36058	**R-Square**	0.8938
Dependent Mean	12.35319	**Adj R-Sq**	0.8820
Coeff Var	2.91893		


Fit Diagnostics for Log_Elec
Residual
Predicted Value
RStudent
Leverage
Quantile
Log_Elec
Cook's D
Observation
Percent
Fit–Mean
Residual
Proportion Less
Observations 71
Parameters 8
Error DF 63
MSE 0.13
R-Square 0.8938
Adj R-Square 0.882

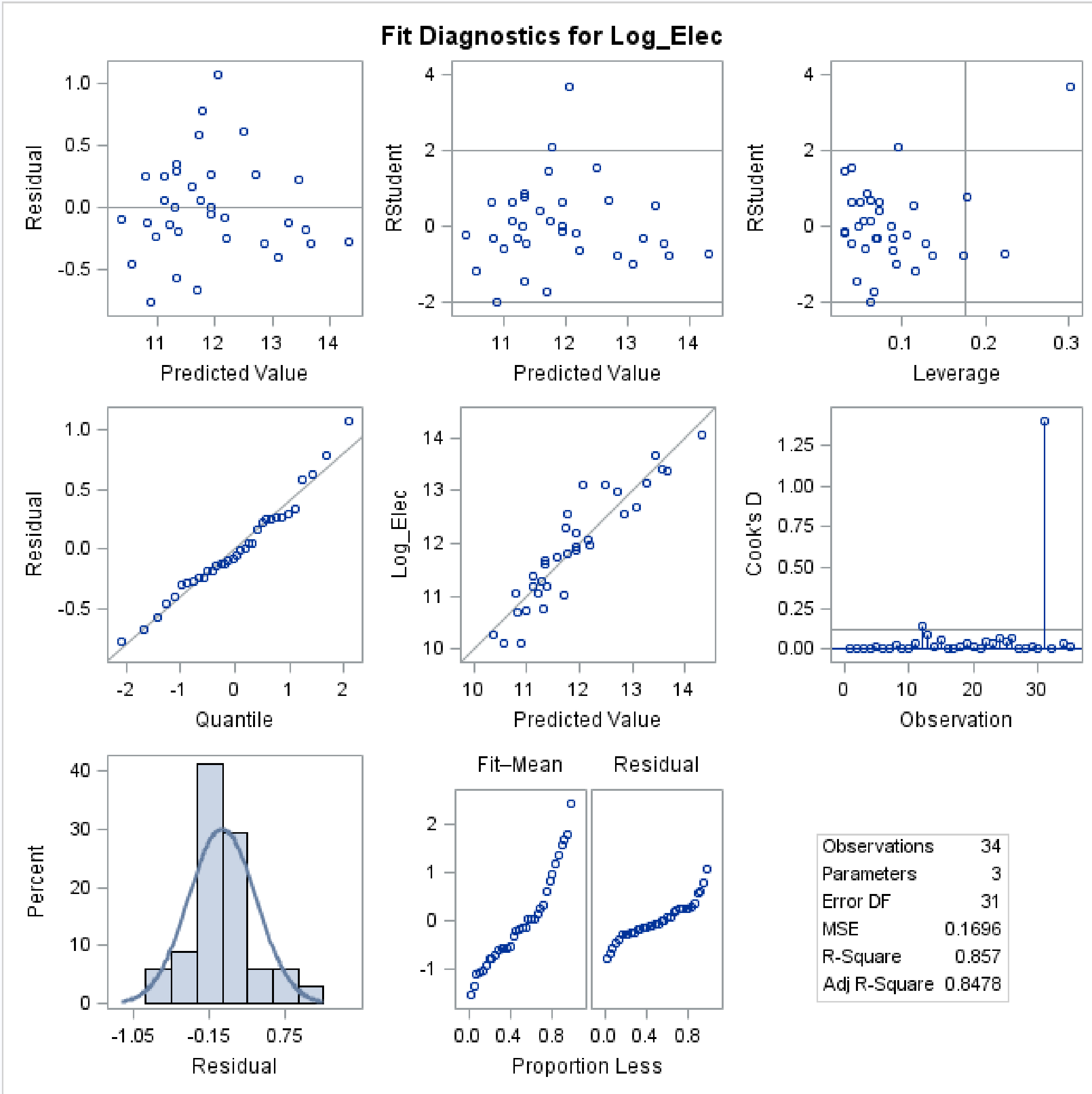

Residual by Regressors for Log_Elec
Residual
EA BIOREACTOR
SLDG SUP ENERGY
DEWATER EQUIP
IAIC
ICCFA
IAEF

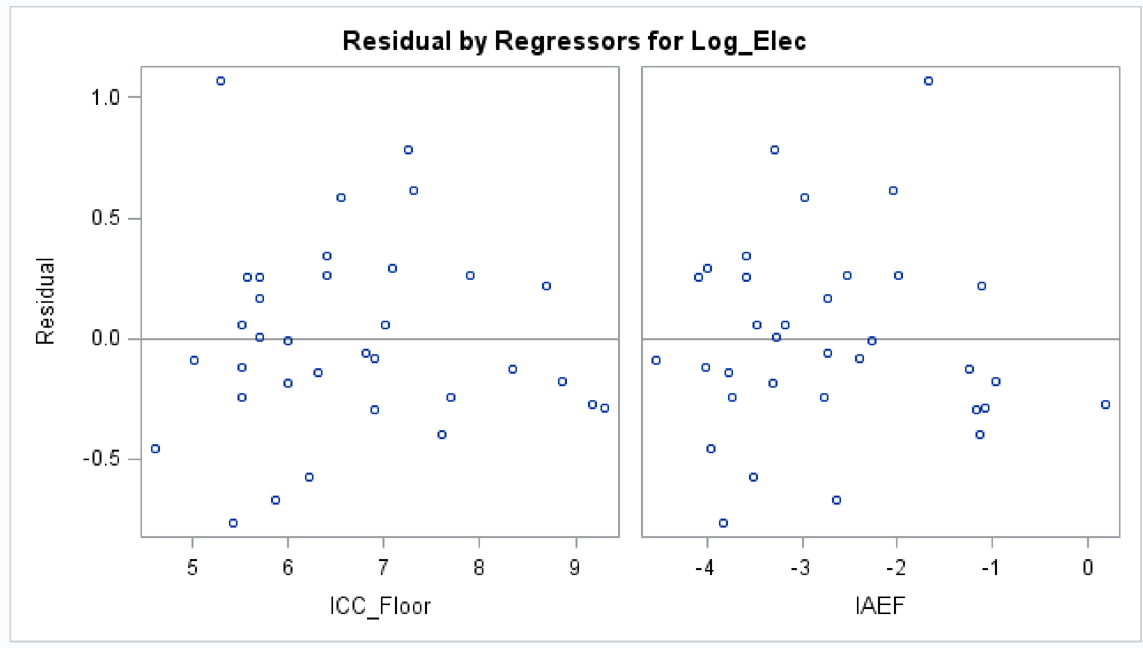

Residual by Regressors for Log_Elec
Residual
IPDF

10. NE EA Electric Usage Model

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	11.28411	0.84260	13.39	<.0001	0
lCC_Floor	0.32009	0.09231	3.47	0.0016	2.51214
lAEF	0.55617	0.09885	5.63	<.0001	2.51214

Root MSE	0.41186	**R-Square**	0.8570
Dependent Mean	11.90130	**Adj R-Sq**	0.8478
Coeff Var	3.46063		

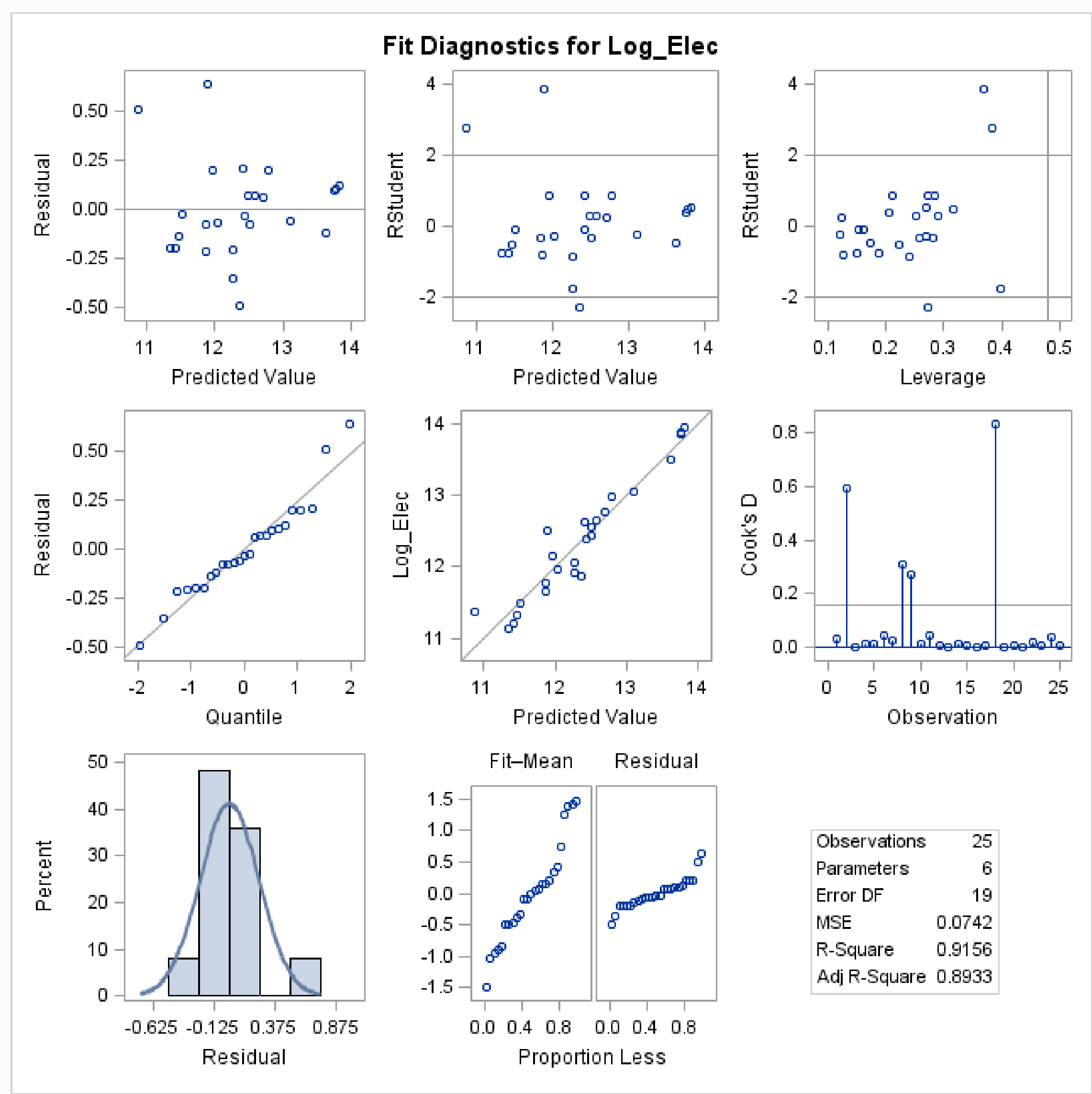

Fit Diagnostics for Log_Elec
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_Elec
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 34
Parameters 3
Error DF 31
MSE 0.1696
R-Square 0.857
Adj R-Square 0.8478

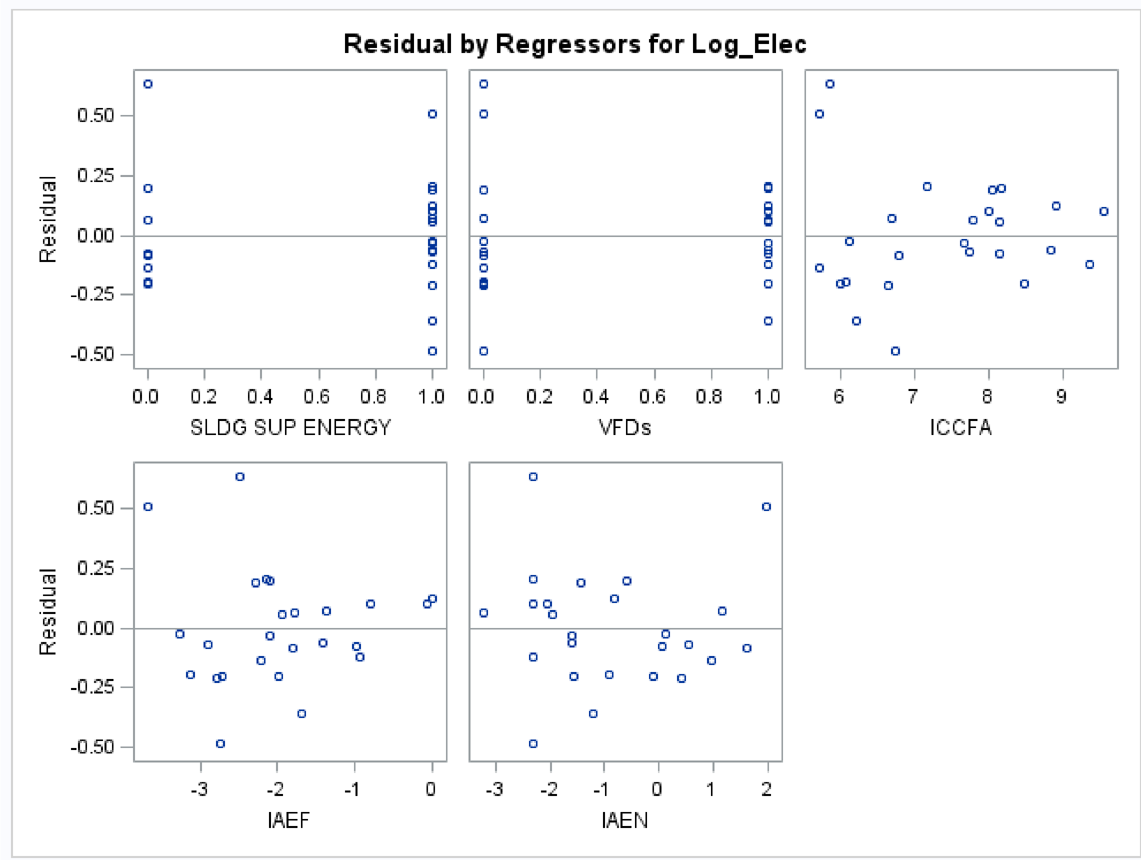

Residual by Regressors for Log_Elec
Residual
1.0
0.5
0.0
-0.5
5
6
7
8
9
ICC_Floor
-4
-3
-2
-1
0
IAEF

11. NE OD Electric Usage Model

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	11.60081	0.63031	18.41	<.0001	0
SLDG_SUP_ENERGY	0.41642	0.11777	3.54	0.0022	1.07603
VFDs	-0.39980	0.18597	-2.15	0.0447	2.90664
lCCFA	0.23122	0.07729	2.99	0.0075	2.69708
lAEF	0.58073	0.08888	6.53	<.0001	2.24417
lAEN	-0.16597	0.04926	-3.37	0.0032	1.57560

Root MSE	0.27248	**R-Square**	0.9156
Dependent Mean	12.36306	**Adj R-Sq**	0.8933
Coeff Var	2.20402		


Fit Diagnostics for Log_Elec
Residual
Predicted Value
RStudent
Leverage
Quantile
Log_Elec
Cook's D
Observation
Percent
Residual
Fit–Mean
Proportion Less
Observations 25
Parameters 6
Error DF 19
MSE 0.0742
R-Square 0.9156
Adj R-Square 0.8933

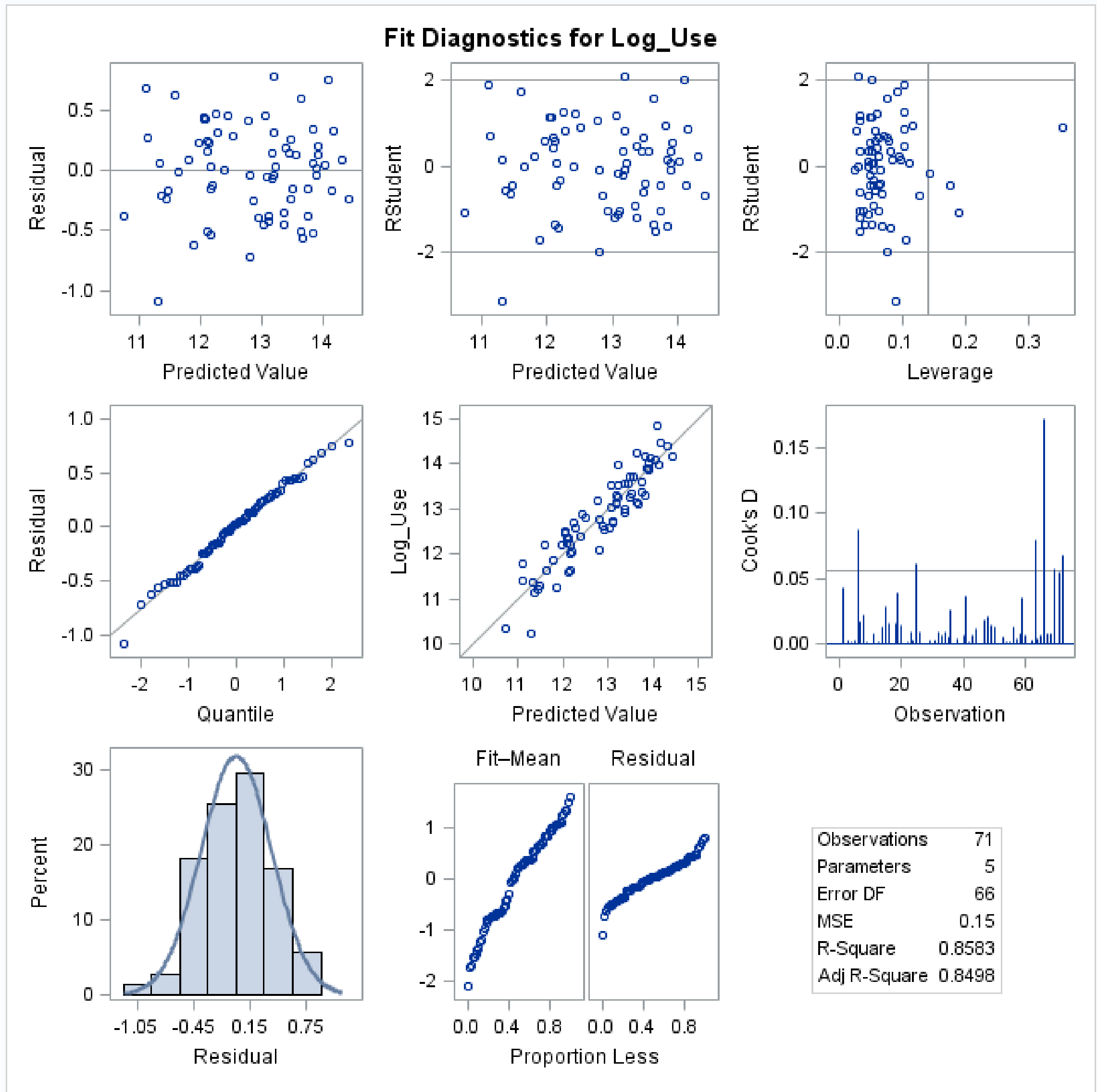


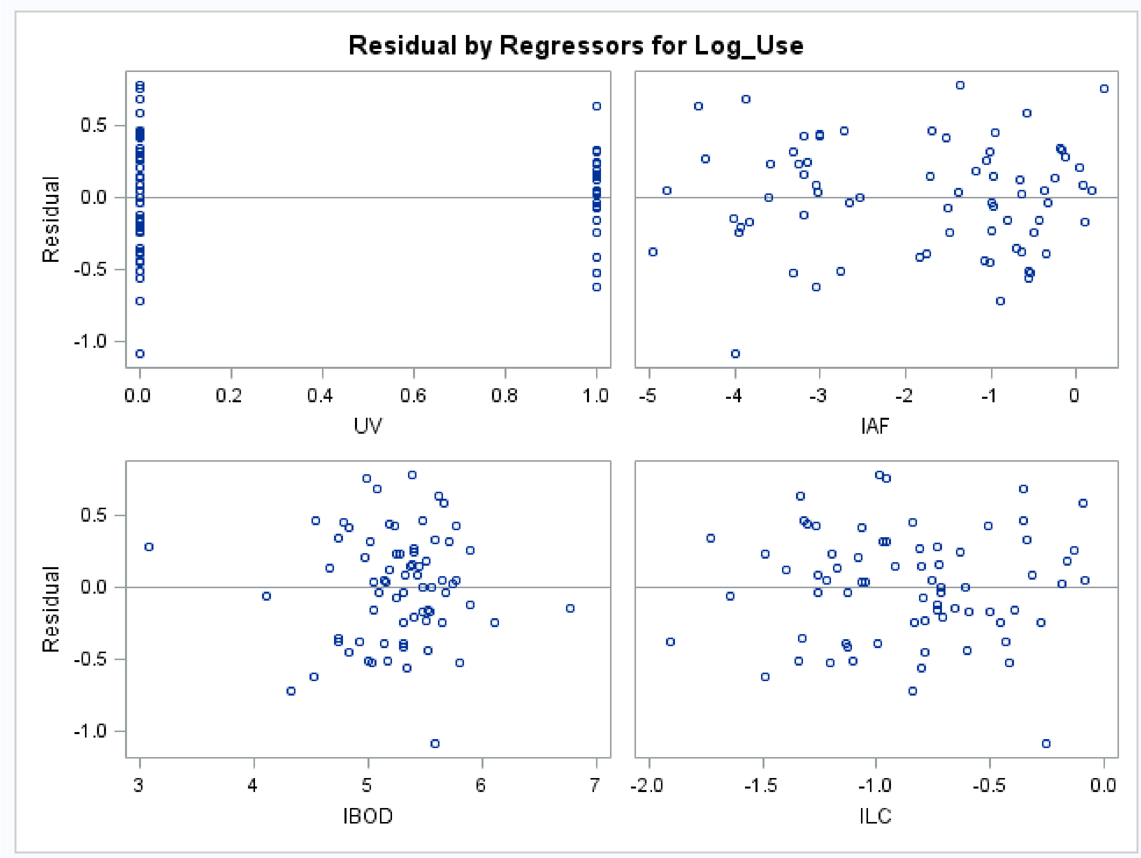
Residual by Regressors for Log_Elec
Residual
0.50
0.25
0.00
-0.25
-0.50
0.0
0.2
0.4
0.6
0.8
1.0
SLDG SUP ENERGY
VFDs
6
7
8
9
ICCFA
-3
-2
-1
0
1
2
IAEF
IAEN

12. Penn Electric Usage All Plant Types

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	10.42984	0.59993	17.39	<.0001	0
UV	0.22182	0.10169	2.18	0.0327	1.07198
lAF	0.68367	0.03434	19.91	<.0001	1.17852
lBOD	0.63229	0.10847	5.83	<.0001	1.35070
lLC	-0.30630	0.12507	-2.45	0.0170	1.25229

Root MSE	0.38731	**R-Square**	0.8583
Dependent Mean	12.83161	**Adj R-Sq**	0.8498
Coeff Var	3.01839		

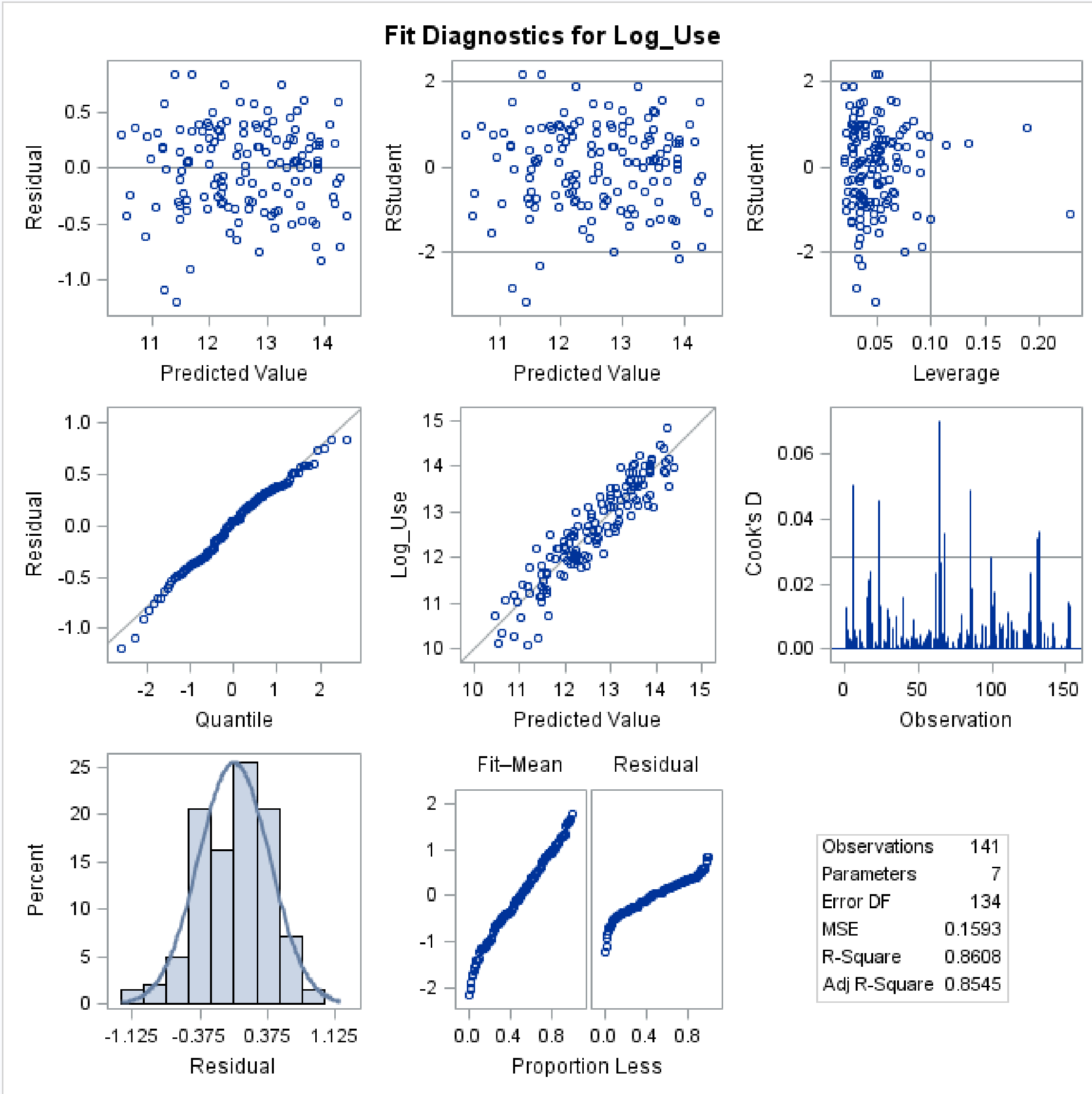

Fit Diagnostics for Log_Use
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_Use
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 71
Parameters 5
Error DF 66
MSE 0.15
R-Square 0.8583
Adj R-Square 0.8498

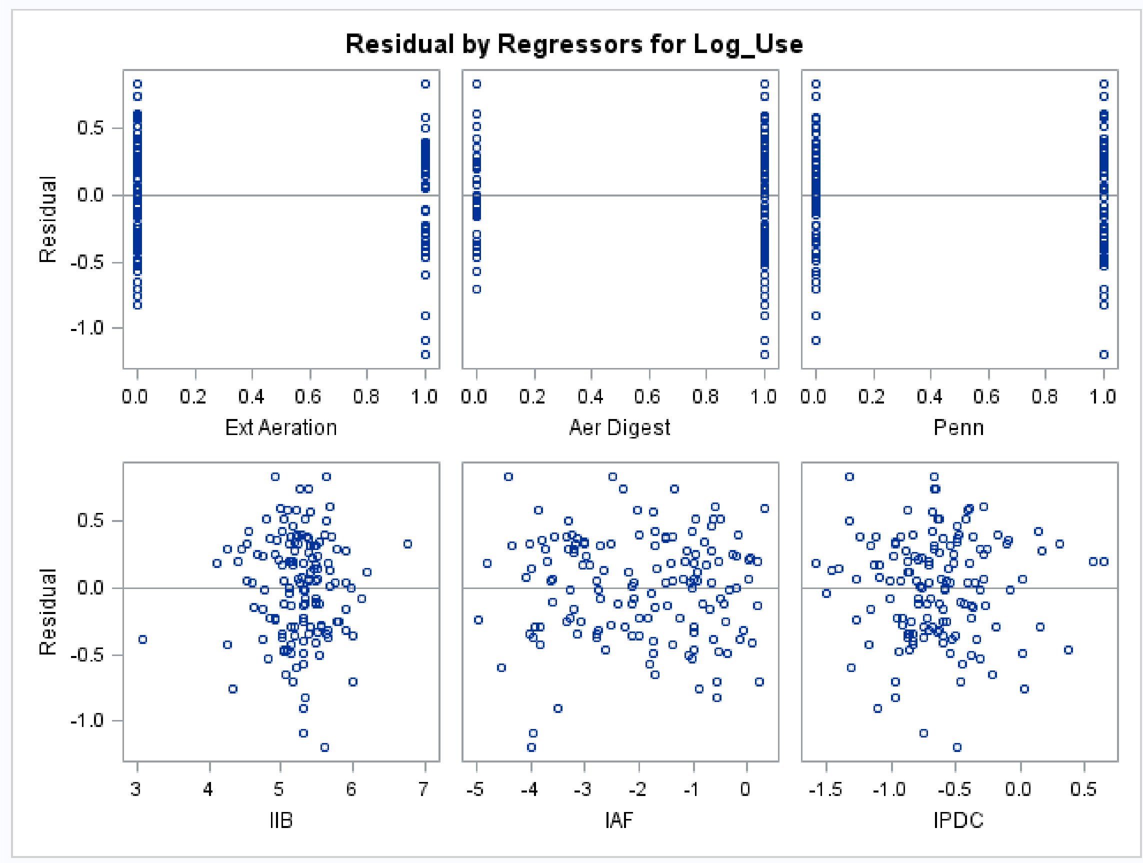

Residual by Regressors for Log_Use
Residual
0.5
0.0
-0.5
-1.0
0.0
0.2
0.4
0.6
0.8
1.0
UV
-5
-4
-3
-2
-1
0
IAF
3
4
5
6
7
IBOD
-2.0
-1.5
-1.0
-0.5
0.0
ILC

13. Penn and NE Electric Usage Model

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	10.94249	0.40602	26.95	<.0001	0
Ext_Aeration	0.18184	0.08723	2.08	0.0390	1.59115
Aer_Digest	0.23631	0.08448	2.80	0.0059	1.08338
Penn	0.20149	0.06944	2.90	0.0043	1.06679
lIB	0.49970	0.07903	6.32	<.0001	1.08598
lAF	0.77429	0.03411	22.70	<.0001	1.82076
lPDC	-0.32800	0.09563	-3.43	0.0008	1.23585

Root MSE	0.39916	**R-Square**	0.8608
Dependent Mean	12.60188	**Adj R-Sq**	0.8545
Coeff Var	3.16747		


Fit Diagnostics for Log_Use
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_Use
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 141
Parameters 7
Error DF 134
MSE 0.1593
R-Square 0.8608
Adj R-Square 0.8545

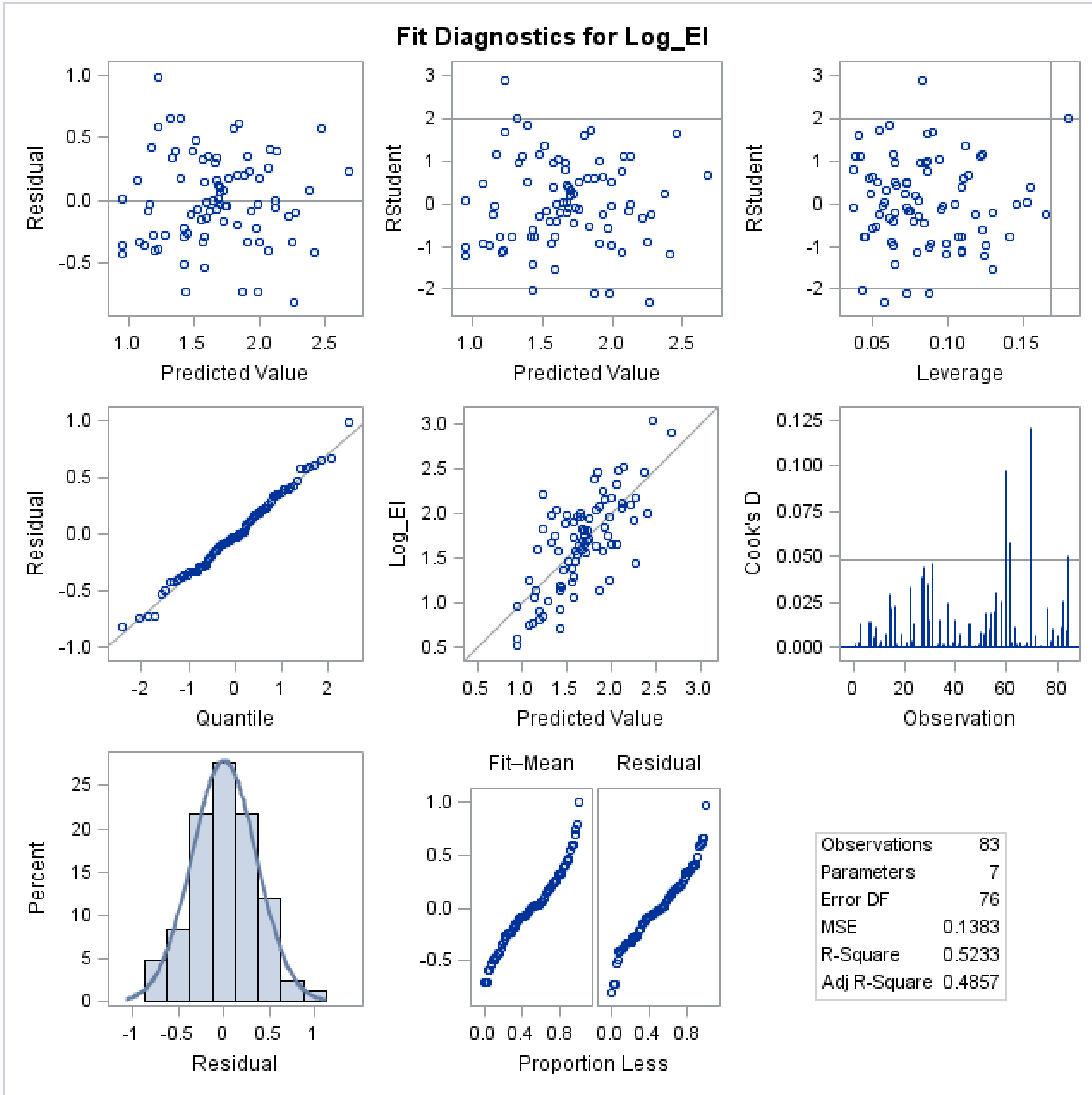

Residual by Regressors for Log_Use
Residual
0.5
0.0
-0.5
-1.0
0.0
0.2
0.4
0.6
0.8
1.0
Ext Aeration
Aer Digest
Penn
3
4
5
6
7
IIB
-5
-4
-3
-2
-1
0
IAF
-1.5
-1.0
-0.5
0.0
0.5
IPDC

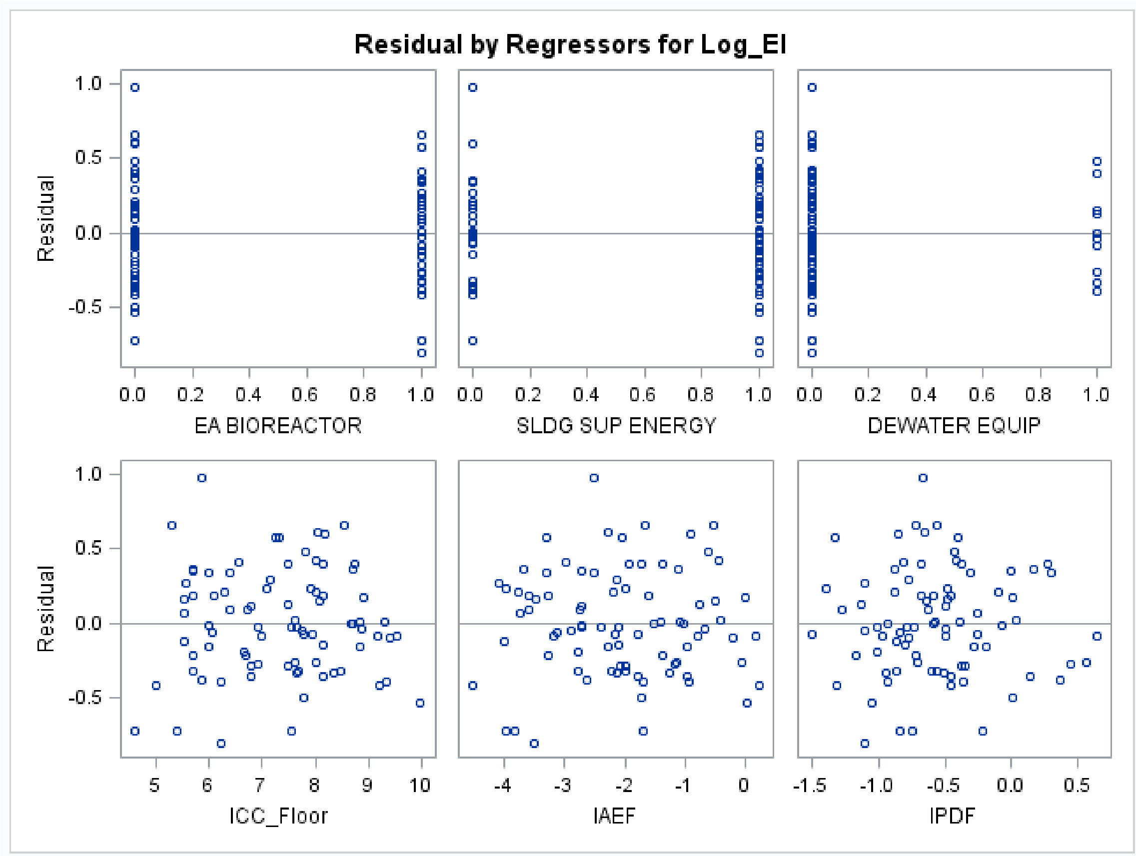
14. NE Overall Plant Type Energy Intensity

- Data Set: Wastewater Data_Sldg_noFF
- Response: Energy Intensity
- No inf. data used.
- Deleted Greenwood, Pender, and Petersburg.
- Removed Eff. CBOD and CDDs because they produced opposite signs in the model.
- Removed HDDs because the p-value was only 0.09. Removing HDDs did not change R-sq.
- No Strictly Fixed Film plants.

Variable	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	-0.94982	0.51192	-1.86	0.0674	0
EA_BIOREACTOR	0.25067	0.09971	2.51	0.0141	1.44272
SLDG_SUP_ENERGY	0.25349	0.09278	2.73	0.0078	1.06179
DEWATER_EQUIP	0.31584	0.13537	2.33	0.0223	1.26430
lCC_Floor	0.20025	0.05698	3.51	0.0007	2.99831
lAEF	-0.32646	0.06803	-4.80	<.0001	3.77268
lPDF	-0.28748	0.10623	-2.71	0.0084	1.37356

Root MSE	0.37191	**R-Square**	0.5233
Dependent Mean	1.66858	**Adj R-Sq**	0.4857
Coeff Var	22.28877		


Fit Diagnostics for Log_EI
Residual
RStudent
Predicted Value
Leverage
Quantile
Log_EI
Cook's D
Observation
Percent
Fit–Mean
Residual
Proportion Less
Observations 83
Parameters 7
Error DF 76
MSE 0.1383
R-Square 0.5233
Adj R-Square 0.4857

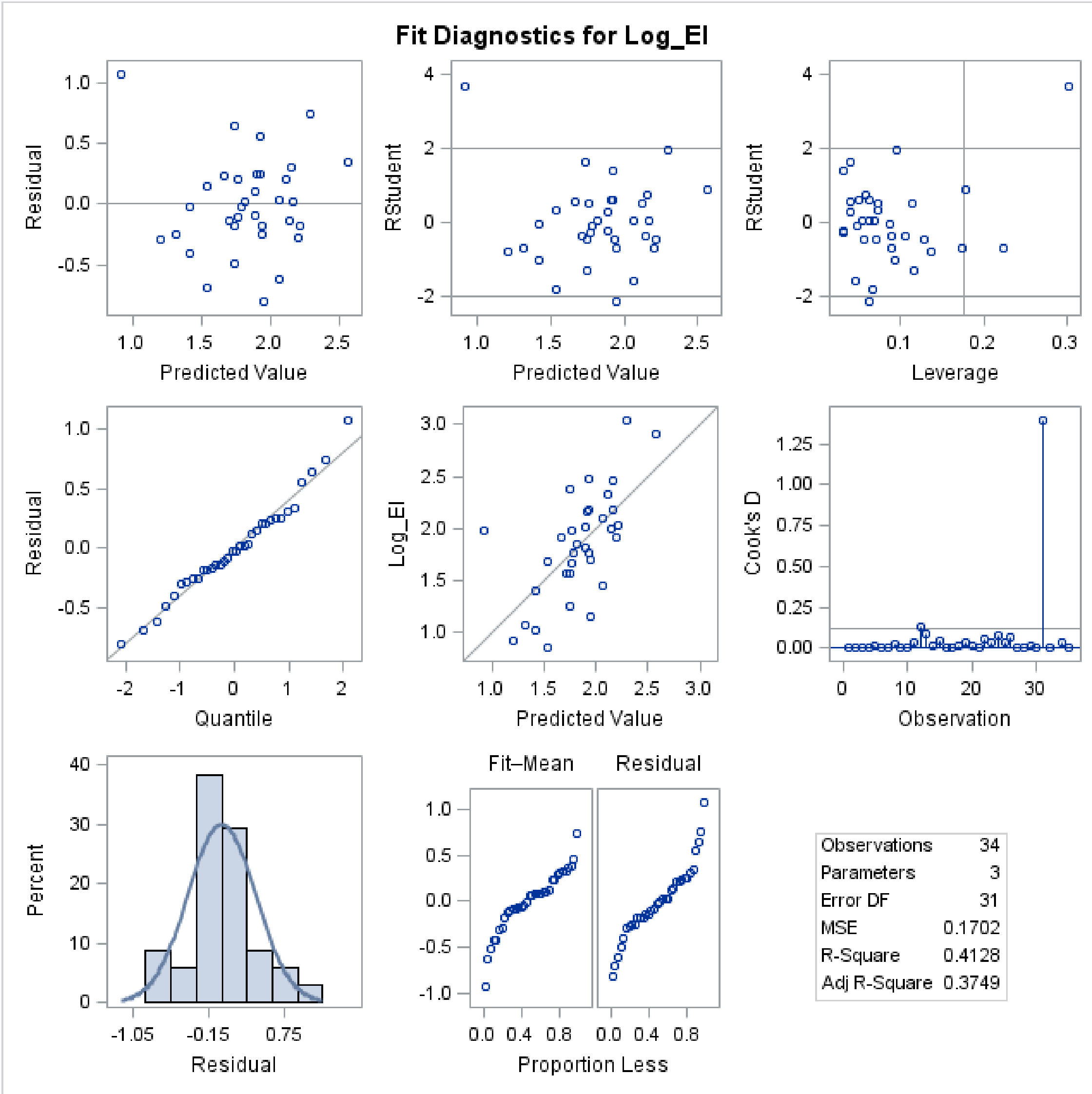

Residual by Regressors for Log_EI
Residual
EA BIOREACTOR
SLDG SUP ENERGY
DEWATER EQUIP
ICC_Floor
IAEF
IPDF
1.0
0.5
0.0
-0.5
0.0 0.2 0.4 0.6 0.8 1.0
5 6 7 8 9 10
-4 -3 -2 -1 0
-1.5 -1.0 -0.5 0.0 0.5

15. NE EA Energy Intensity

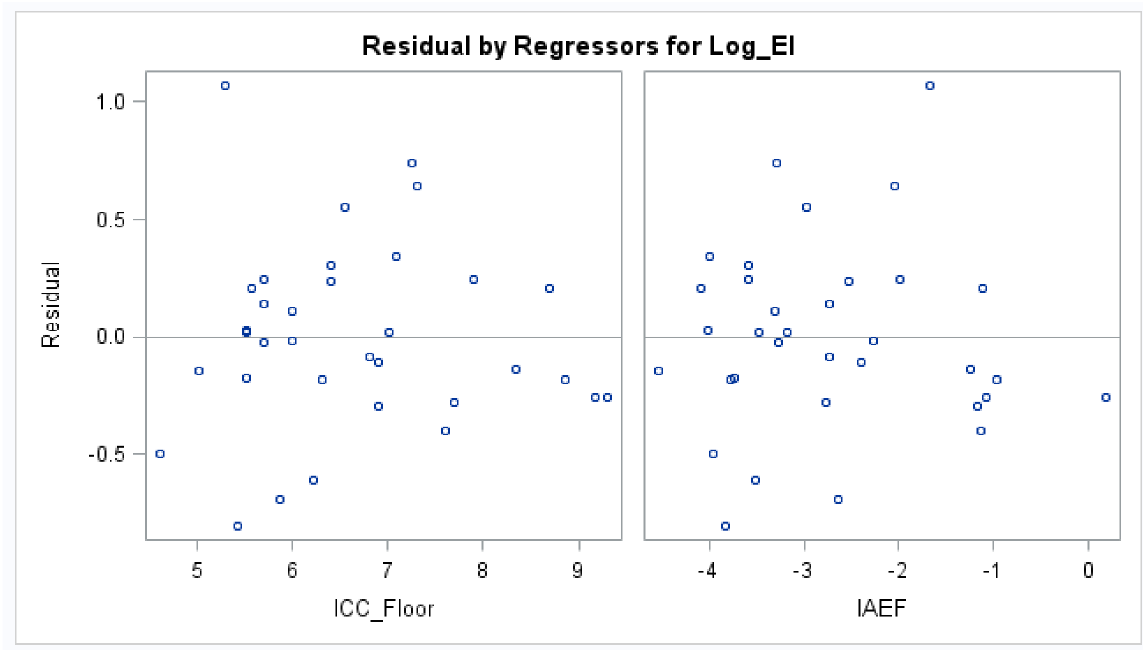
- Response: Energy Intensity
- Data Set: Ext_Aeration
- Deleted Greenwood, Pender, and Petersburg. Reasoning: Greenwood (bills), Petersburg and Pender (flow).
- Not deleting Petersburg improves R-sq adj to 0.404, but Petersburg has so much influence on the model (Cook's D) that I left them out of the model.
- Syracuse was also highly influential because of its small CCFA, but could not delete because there was nothing to justify its deletion.

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	-1.58381	0.84408	-1.88	0.0700	0
lCC_Floor	0.32551	0.09248	3.52	0.0014	2.51214
lAEF	-0.46199	0.09902	-4.67	<.0001	2.51214

Root MSE	0.41258	**R-Square**	0.4128
Dependent Mean	1.83654	**Adj R-Sq**	0.3749
Coeff Var	22.46520		



Fit Diagnostics for Log_EI
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_EI
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 34
Parameters 3
Error DF 31
MSE 0.1702
R-Square 0.4128
Adj R-Square 0.3749



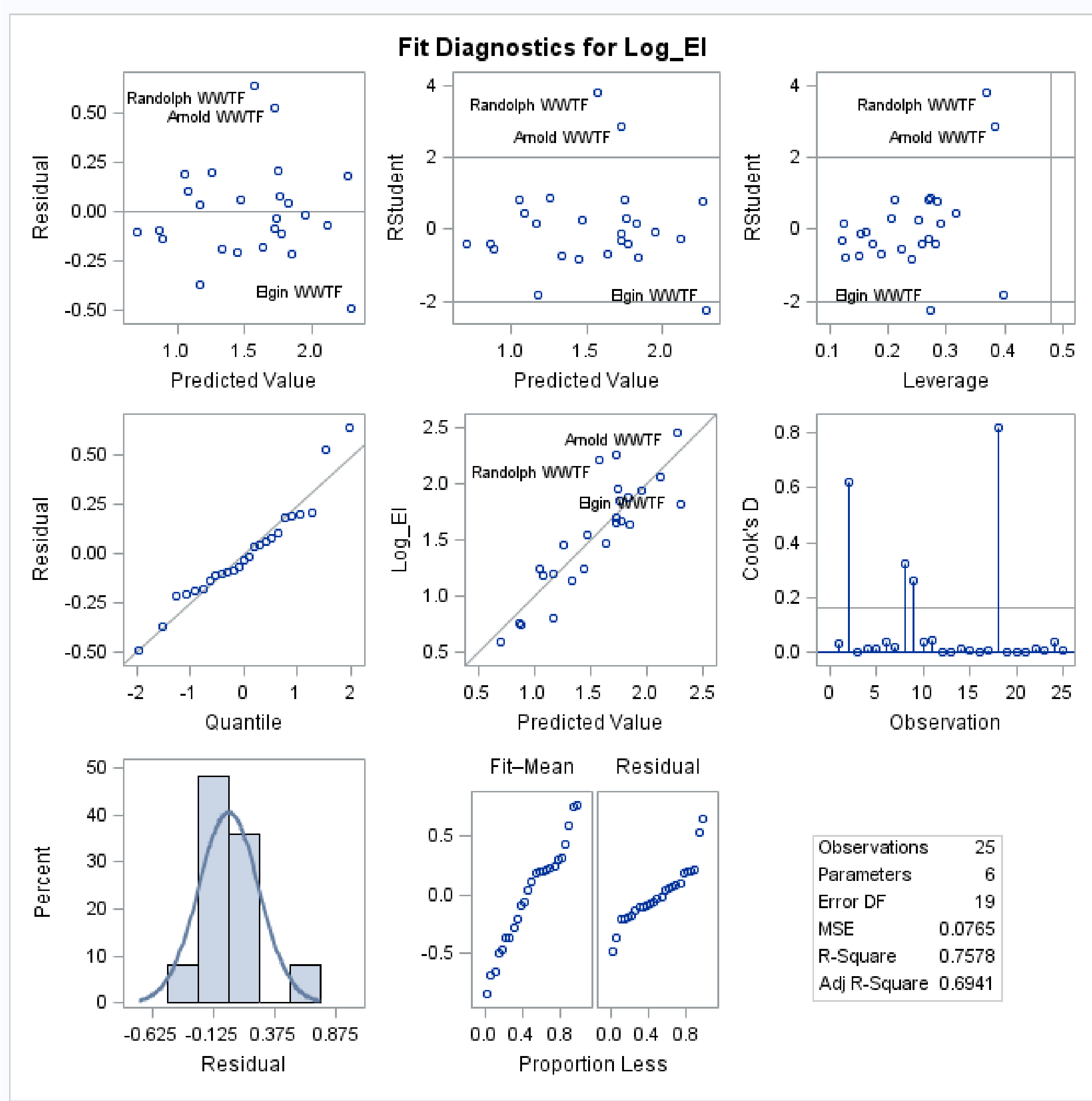
Residual by Regressors for Log_EI
Residual
1.0
0.5
0.0
-0.5
5
6
7
8
9
ICC_Floor
-4
-3
-2
-1
0
IAEF

16. <u>NE OD Energy Intensity</u>

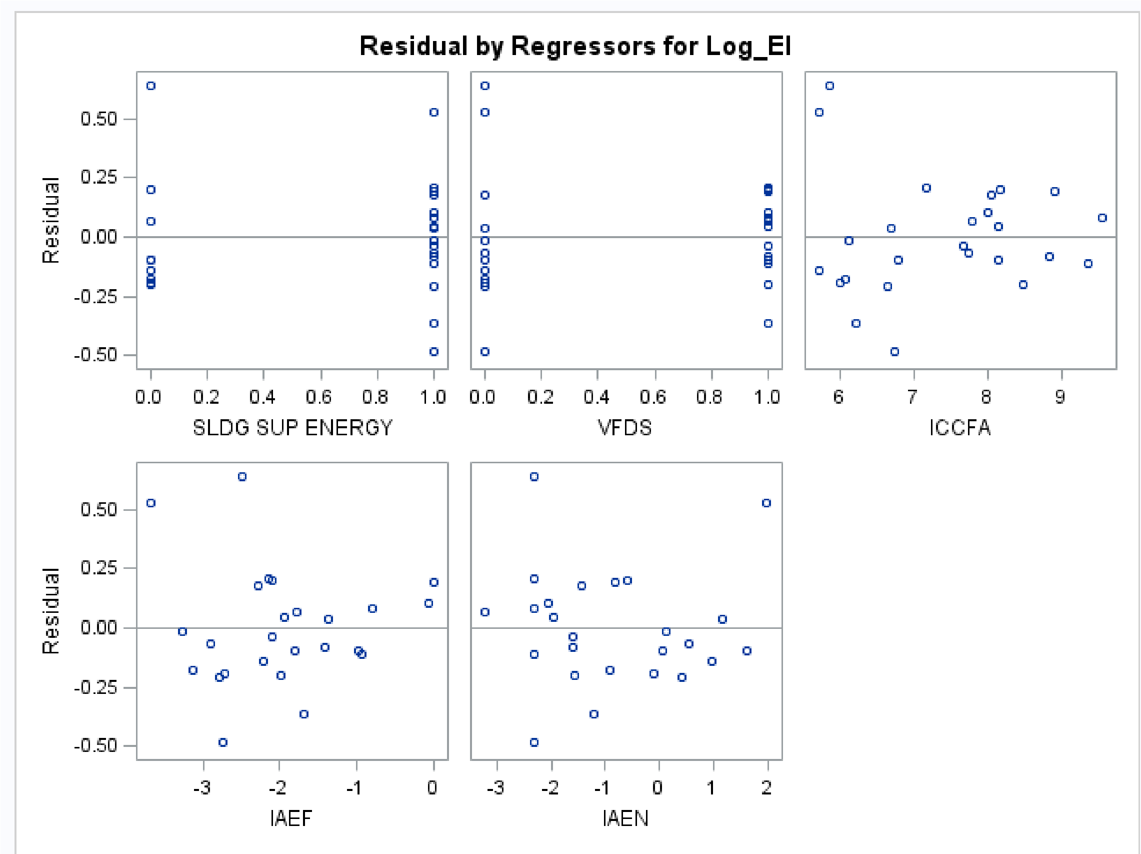
- Response: log(Energy Intensity), MWh/MG
- Only NE plants
- Data Set: OD Plants
- Small sample size: 25 plants
- Looks like the errors aren't quite normal, but that might be because of the sample size.

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	-1.15929	0.63979	-1.81	0.0858	0
SLDG_SUP_ENERGY	0.42625	0.11954	3.57	0.0021	1.07603
VFDS	-0.41101	0.18877	-2.18	0.0423	2.90664
lCCFA	0.23195	0.07845	2.96	0.0081	2.69708
lAEF	-0.39712	0.09022	-4.40	0.0003	2.24417
lAEN	-0.16547	0.05000	-3.31	0.0037	1.57560

Root MSE	0.27659	**R-Square**	0.7578
Dependent Mean	1.53874	**Adj R-Sq**	0.6941
Coeff Var	17.97475		



Fit Diagnostics for Log_EI
Randolph WWTF
Arnold WWTF
Elgin WWTF
Residual
RStudent
Predicted Value
Leverage
Quantile
Log_EI
Cook's D
Observation
Percent
Residual
Fit–Mean
Proportion Less
Observations 25
Parameters 6
Error DF 19
MSE 0.0765
R-Square 0.7578
Adj R-Square 0.6941

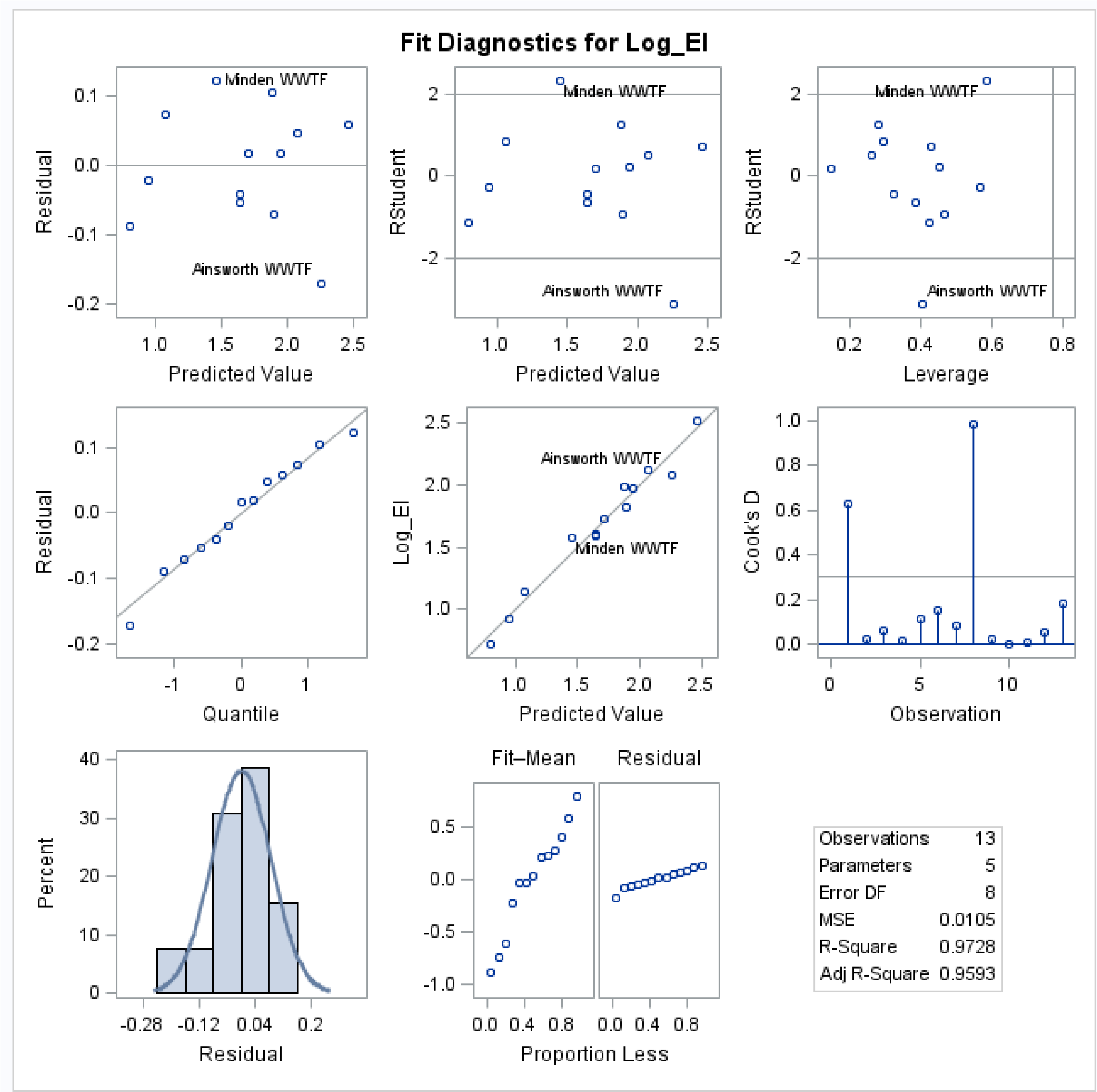

Residual by Regressors for Log_EI
Residual
0.50
0.25
0.00
-0.25
-0.50
0.0 0.2 0.4 0.6 0.8 1.0
SLDG SUP ENERGY
VFDS
6 7 8 9
ICCFA
-3 -2 -1 0
IAEF
-3 -2 -1 0 1 2
IAEN

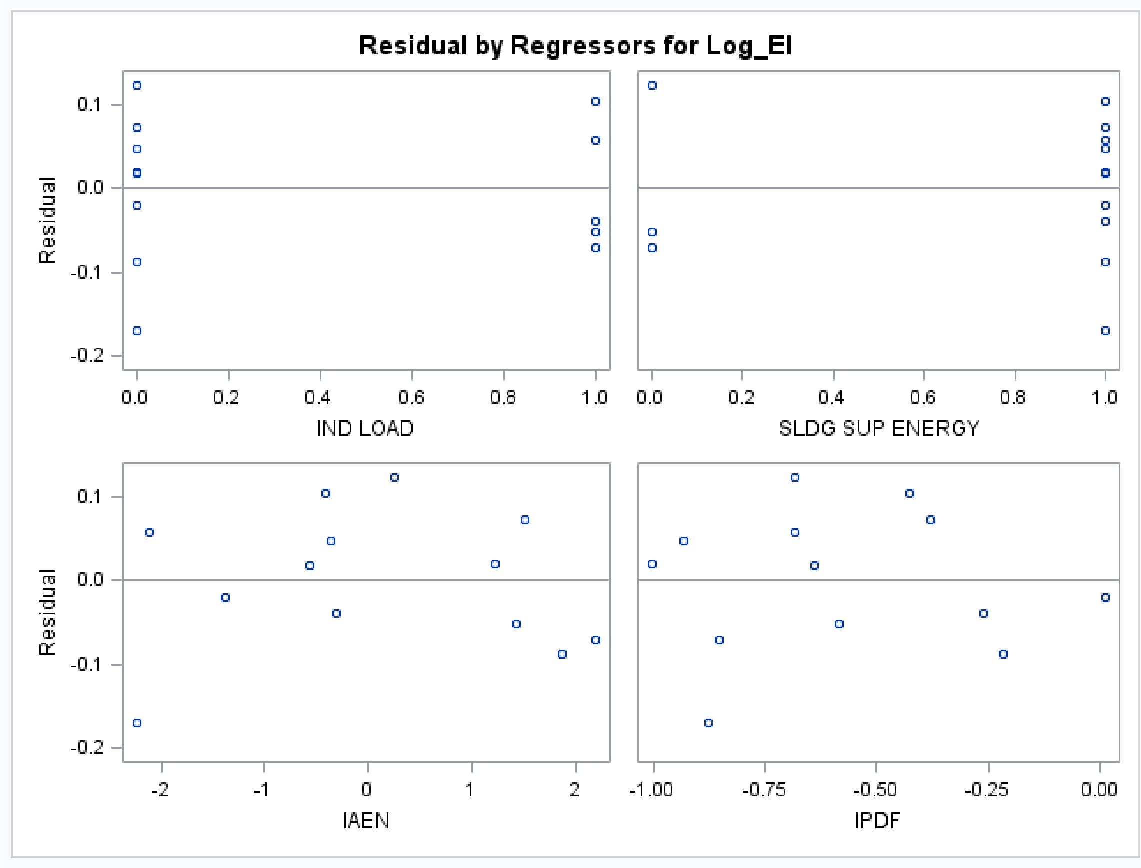
17. NE SBR Energy Intensity

- Response: log(Energy Intensity), MWh/MG
- Only NE plants
- Small sample size: 13 plants
- Data Set: SBR Plants

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	0.56370	0.11395	4.95	0.0011	0
IND_LOAD	0.48403	0.06289	7.70	<.0001	1.15525
SLDG_SUP_ENERGY	0.19898	0.08470	2.35	0.0467	1.57155
lAEN	-0.13916	0.02304	-6.04	0.0003	1.31599
lPDF	-1.35221	0.10079	-13.42	<.0001	1.09496

Root MSE	0.10264	**R-Square**	0.9728
Dependent Mean	1.67450	**Adj R-Sq**	0.9593
Coeff Var	6.12933		


Fit Diagnostics for Log_EI
Minden WWTF
Ainsworth WWTF
Residual
Predicted Value
RStudent
Leverage
Quantile
Log_EI
Cook's D
Observation
Percent
Fit–Mean
Residual
Proportion Less
Observations 13
Parameters 5
Error DF 8
MSE 0.0105
R-Square 0.9728
Adj R-Square 0.9593

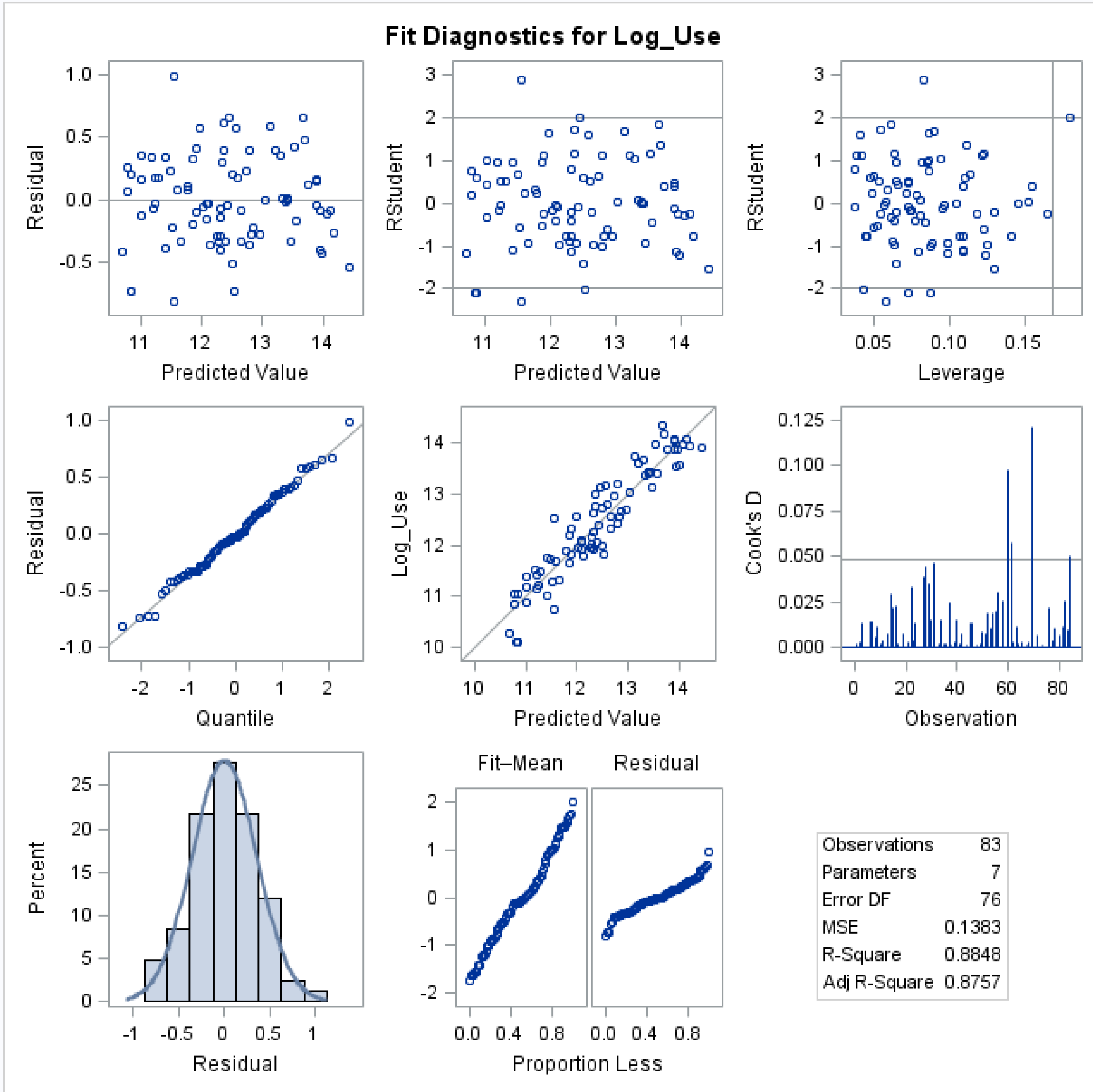

Residual by Regressors for Log_EI
Residual
0.1
0.0
-0.1
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
IND LOAD
SLDG SUP ENERGY
-2
-1
0
1
2
IAEN
-1.00
-0.75
-0.50
-0.25
0.00
IPDF

18. NE Overall Plant Type Energy Usage

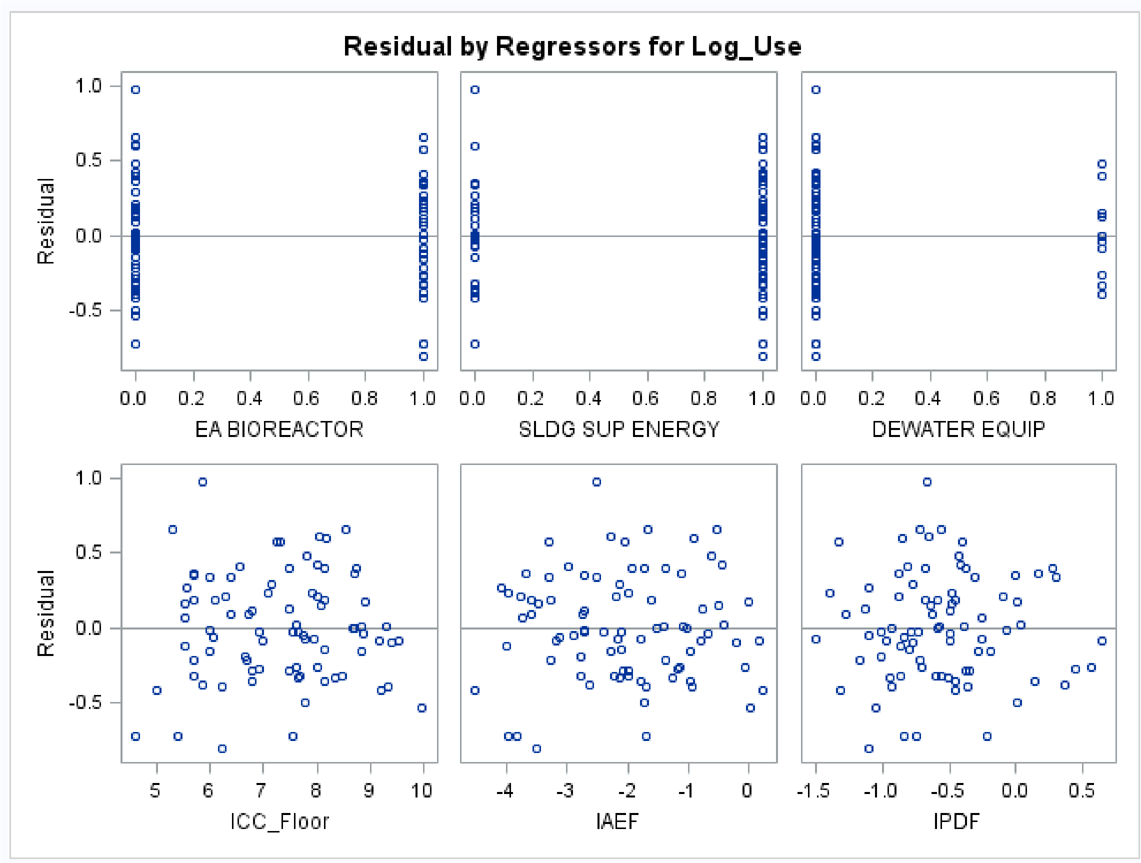
- Data Set: Wastewater Data_Sldg
- Response: Energy Usage
- Deleted Greenwood, Pender, and Petersburg.
- No Strictly Fixed Film plants
- No influent data.

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	11.85786	0.51193	23.16	<.0001	0
EA_BIOREACTOR	0.25067	0.09971	2.51	0.0141	1.44272
SLDG_SUP_ENERGY	0.25348	0.09278	2.73	0.0078	1.06179
DEWATER_EQUIP	0.31583	0.13538	2.33	0.0223	1.26430
lCC_Floor	0.20025	0.05698	3.51	0.0007	2.99831
lAEF	0.67355	0.06803	9.90	<.0001	3.77268
lPDF	-0.28749	0.10623	-2.71	0.0084	1.37356

Root MSE	0.37191	**R-Square**	0.8848
Dependent Mean	12.42225	**Adj R-Sq**	0.8757
Coeff Var	2.99392		



Fit Diagnostics for Log_Use
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_Use
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 83
Parameters 7
Error DF 76
MSE 0.1383
R-Square 0.8848
Adj R-Square 0.8757

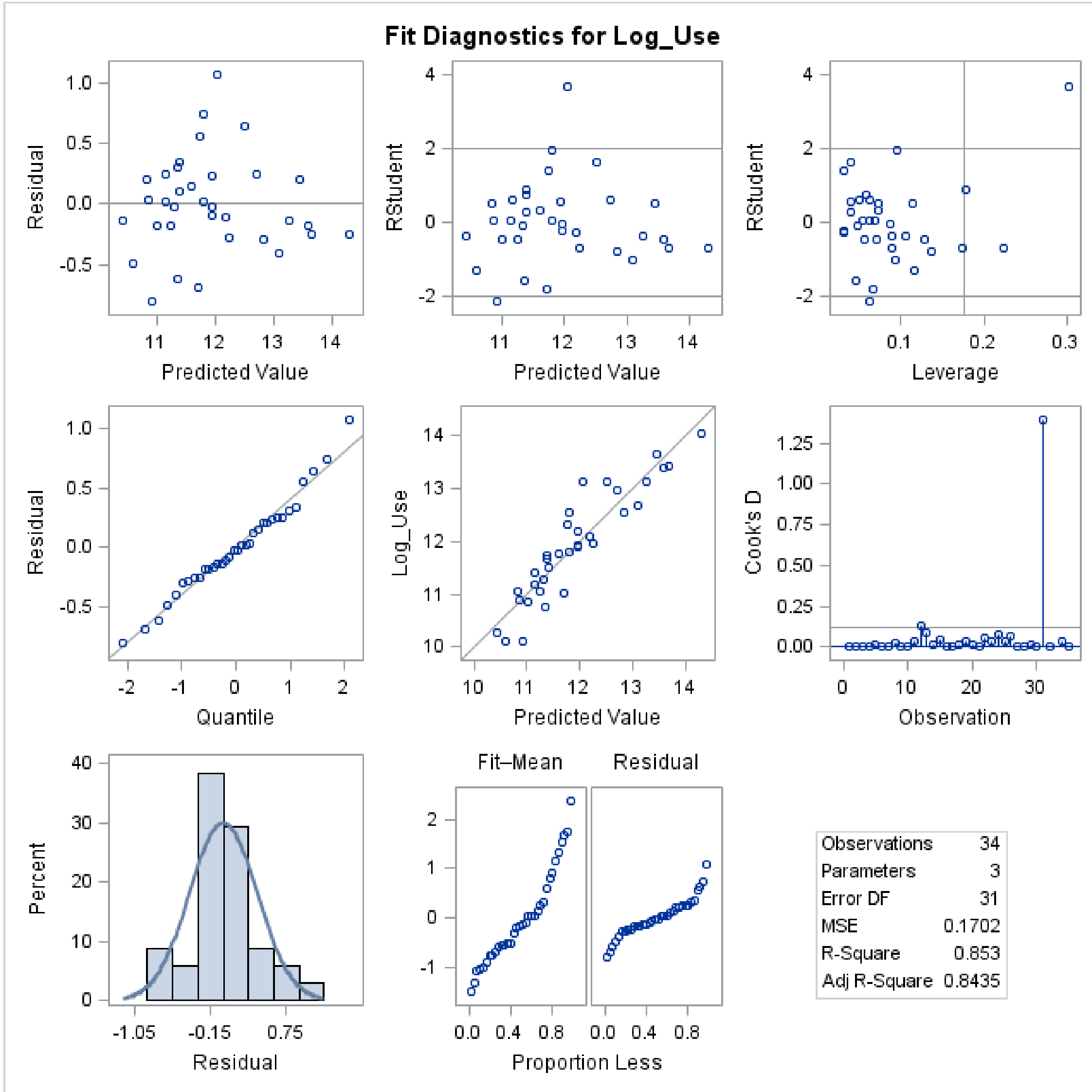

Residual by Regressors for Log_Use
Residual
1.0
0.5
0.0
-0.5
0.0 0.2 0.4 0.6 0.8 1.0
EA BIOREACTOR
SLDG SUP ENERGY
DEWATER EQUIP
5 6 7 8 9 10
ICC_Floor
-4 -3 -2 -1 0
IAEF
-1.5 -1.0 -0.5 0.0 0.5
IPDF

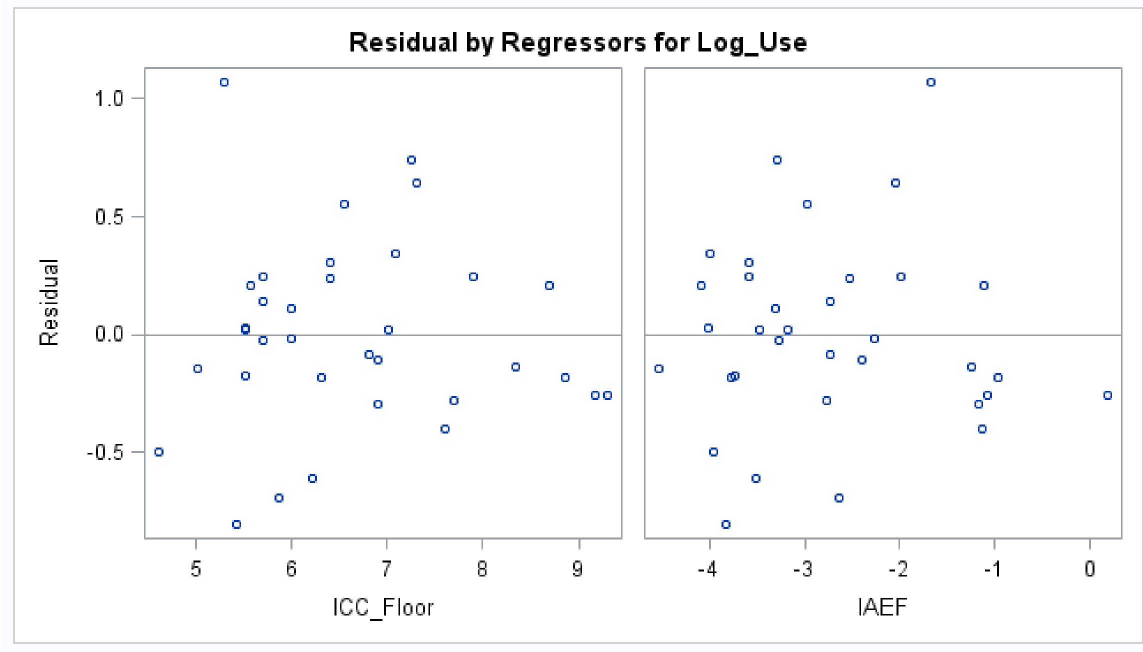
19. <u>NE EA Energy Usage Model</u>

- Data Set: Ext_Aeration
- Response: Energy Usage (kWh/yr)
- Deleted Greenwood, Pender, and Petersburg.
- Decided not to delete Syracuse from the data set, even though they have a small Climate Controlled Floor area in comparison to their plant size. Keeping Syracuse in the data set did not change which variables showed up as significant, but the R-sq went down by 0.04.

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	11.22384	0.84408	13.30	<.0001	0
lCC_Floor	0.32551	0.09248	3.52	0.0014	2.51214
lAEF	0.53801	0.09902	5.43	<.0001	2.51214

Root MSE	0.41258	**R-Square**	0.8530
Dependent Mean	11.92643	**Adj R-Sq**	0.8435
Coeff Var	3.45940		


Fit Diagnostics for Log_Use
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_Use
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 34
Parameters 3
Error DF 31
MSE 0.1702
R-Square 0.853
Adj R-Square 0.8435



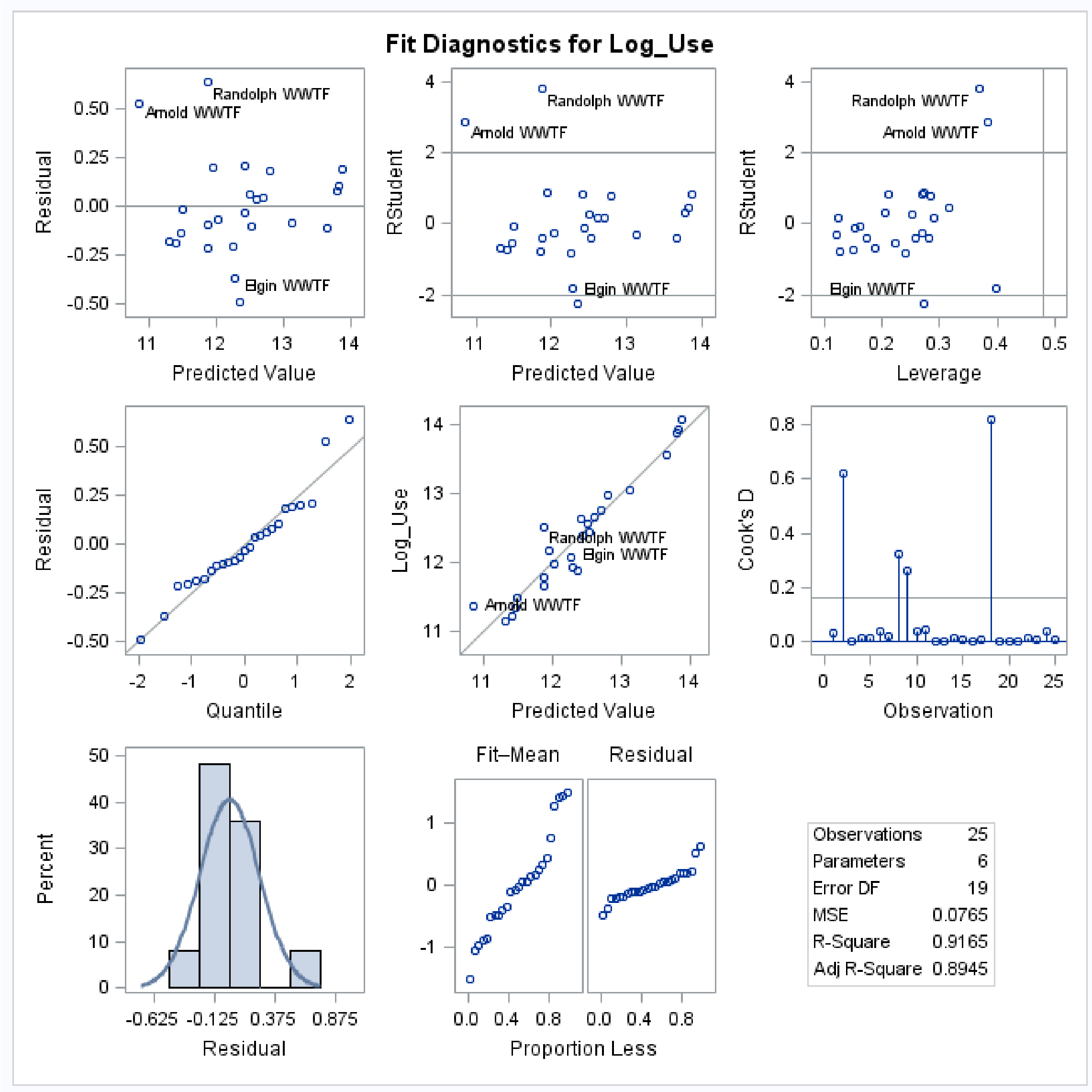
Residual by Regressors for Log_Use
Residual
1.0
0.5
0.0
-0.5
5
6
7
8
9
ICC_Floor
-4
-3
-2
-1
0
IAEF

20. <u>NE OD Model for Energy Usage</u>

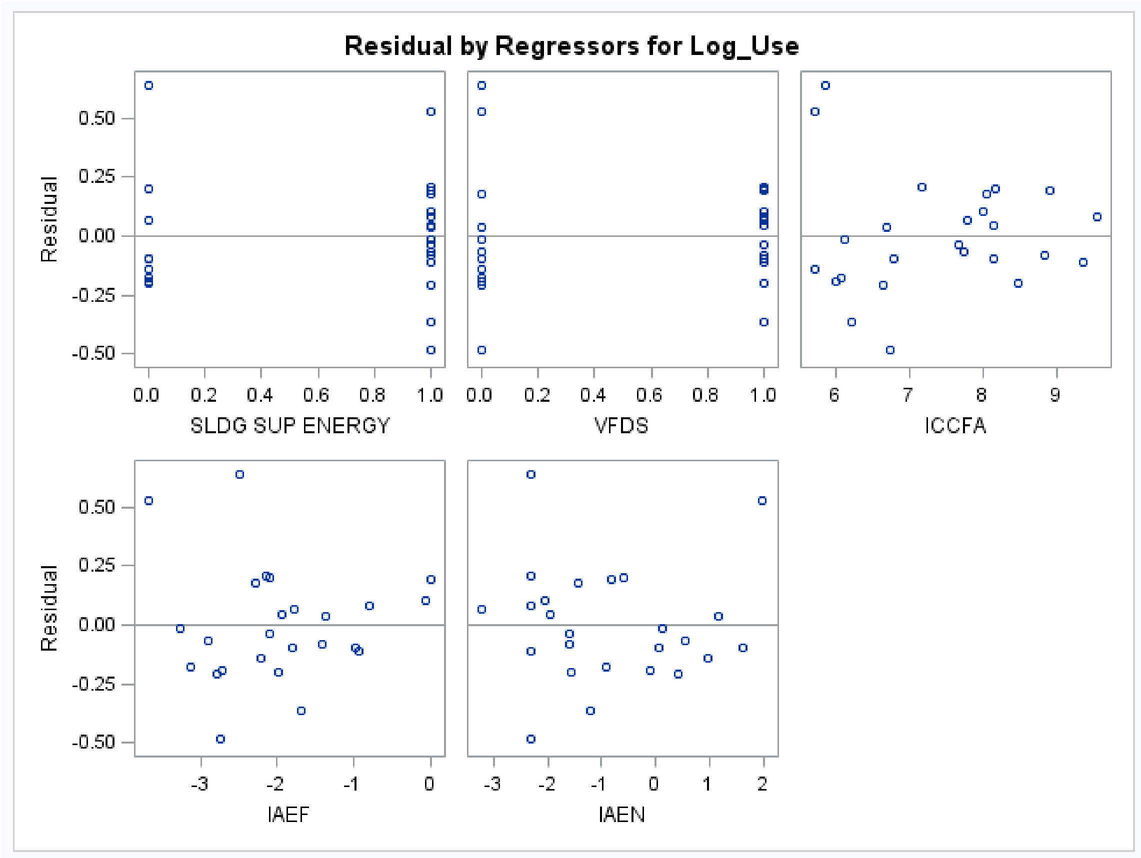
- Data Set: OD Plants
- Response: Energy Usage
- No outliers deleted.
- Same results as the OD Electric Intensity Model. I could also look into deleting Randolph and Arnold for this model as well.

Variable	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	11.64836	0.63979	18.21	<.0001	0
SLDG_SUP_ENERGY	0.42625	0.11954	3.57	0.0021	1.07603
VFDS	-0.41101	0.18877	-2.18	0.0423	2.90664
lCCFA	0.23195	0.07845	2.96	0.0081	2.69708
lAEF	0.60288	0.09022	6.68	<.0001	2.24417
lAEN	-0.16547	0.05000	-3.31	0.0037	1.57560

Root MSE	0.27659	**R-Square**	0.9165
Dependent Mean	12.37232	**Adj R-Sq**	0.8945
Coeff Var	2.23552		



Fit Diagnostics for Log_Use
Residual
Predicted Value
RStudent
Leverage
Randolph WWTF
Arnold WWTF
Elgin WWTF
Quantile
Log_Use
Cook's D
Observation
Percent
Residual
Fit–Mean
Proportion Less
Observations 25
Parameters 6
Error DF 19
MSE 0.0765
R-Square 0.9165
Adj R-Square 0.8945

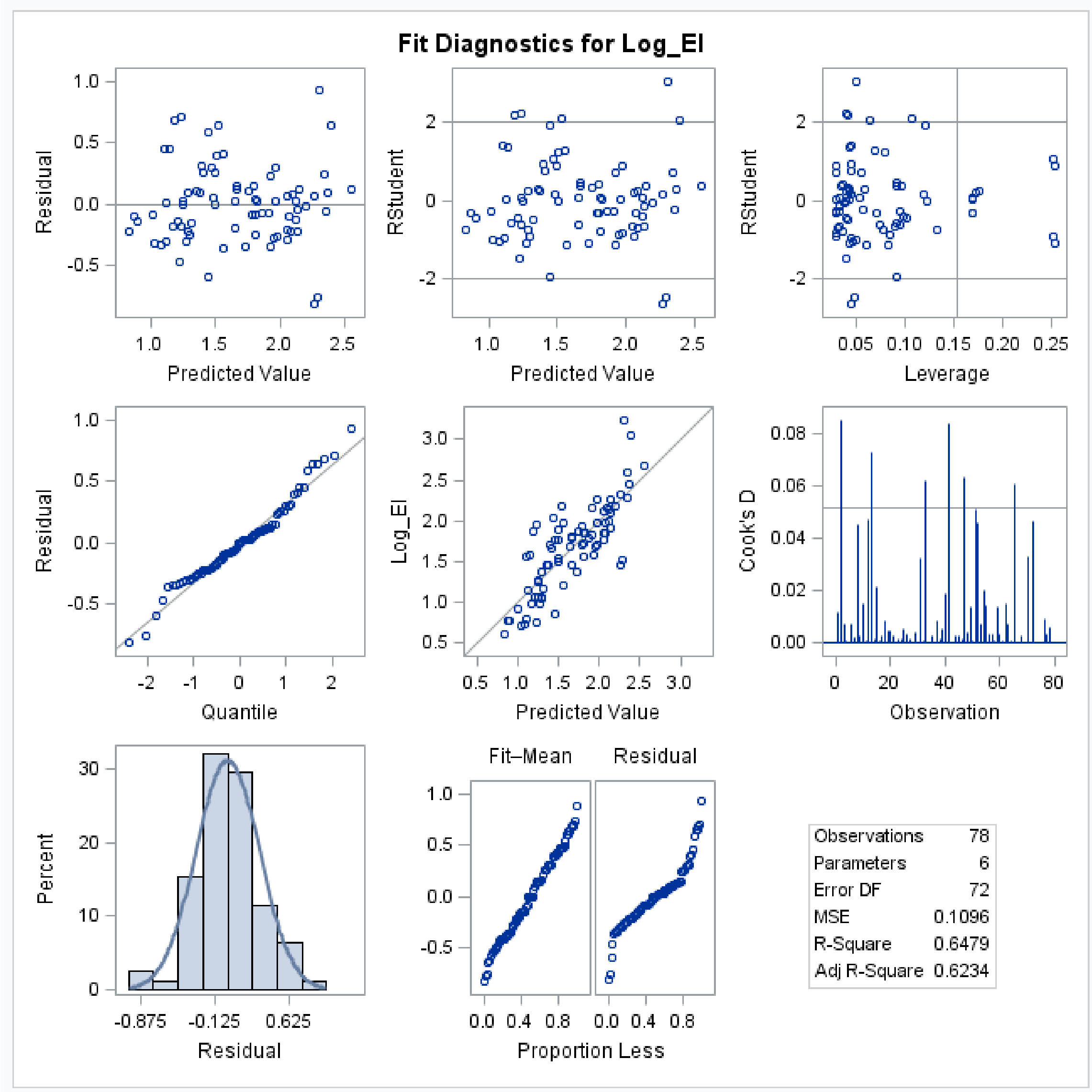

Residual by Regressors for Log_Use
Residual
0.50
0.25
0.00
-0.25
-0.50
0.0 0.2 0.4 0.6 0.8 1.0
SLDG SUP ENERGY
VFDS
6 7 8 9
ICCFA
-3 -2 -1 0
IAEF
-3 -2 -1 0 1 2
IAEN

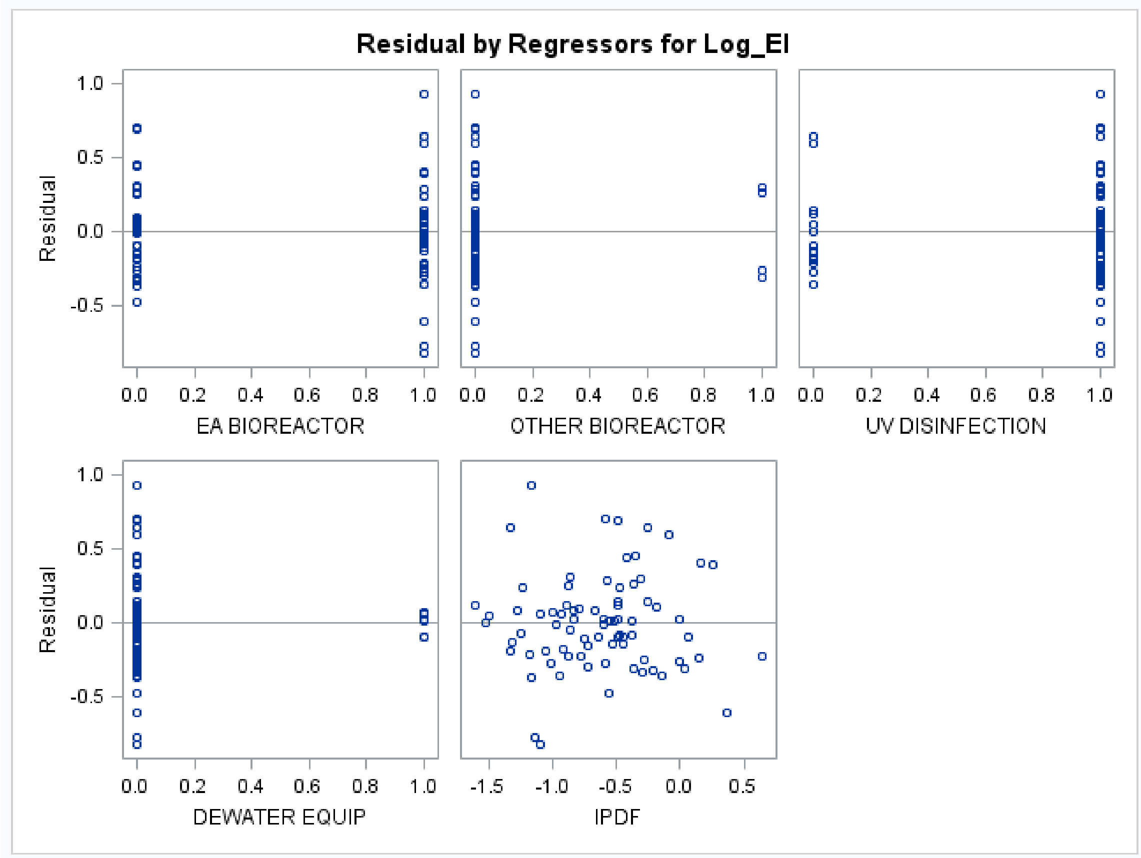
21. NE 2014 and 2015 Combined Energy Intensity Model

- Data Set: Wastewater Data_2014_2015
- Response: Energy Intensity
- No Fixed Film
- No Inf Data

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	0.65085	0.12366	5.26	<.0001	0
EA_BIOREACTOR	0.74023	0.08160	9.07	<.0001	1.18353
OTHER_BIOREACTOR	0.38211	0.18232	2.10	0.0396	1.15053
UV_DISINFECTION	0.26060	0.10141	2.57	0.0122	1.07756
DEWATER_EQUIP	0.62336	0.15050	4.14	<.0001	1.14425
lPDF	-0.55998	0.08558	-6.54	<.0001	1.10427

Root MSE	0.33112	**R-Square**	0.6479
Dependent Mean	1.66513	**Adj R-Sq**	0.6234
Coeff Var	19.88544		


Fit Diagnostics for Log_EI
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_EI
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 78
Parameters 6
Error DF 72
MSE 0.1096
R-Square 0.6479
Adj R-Square 0.6234

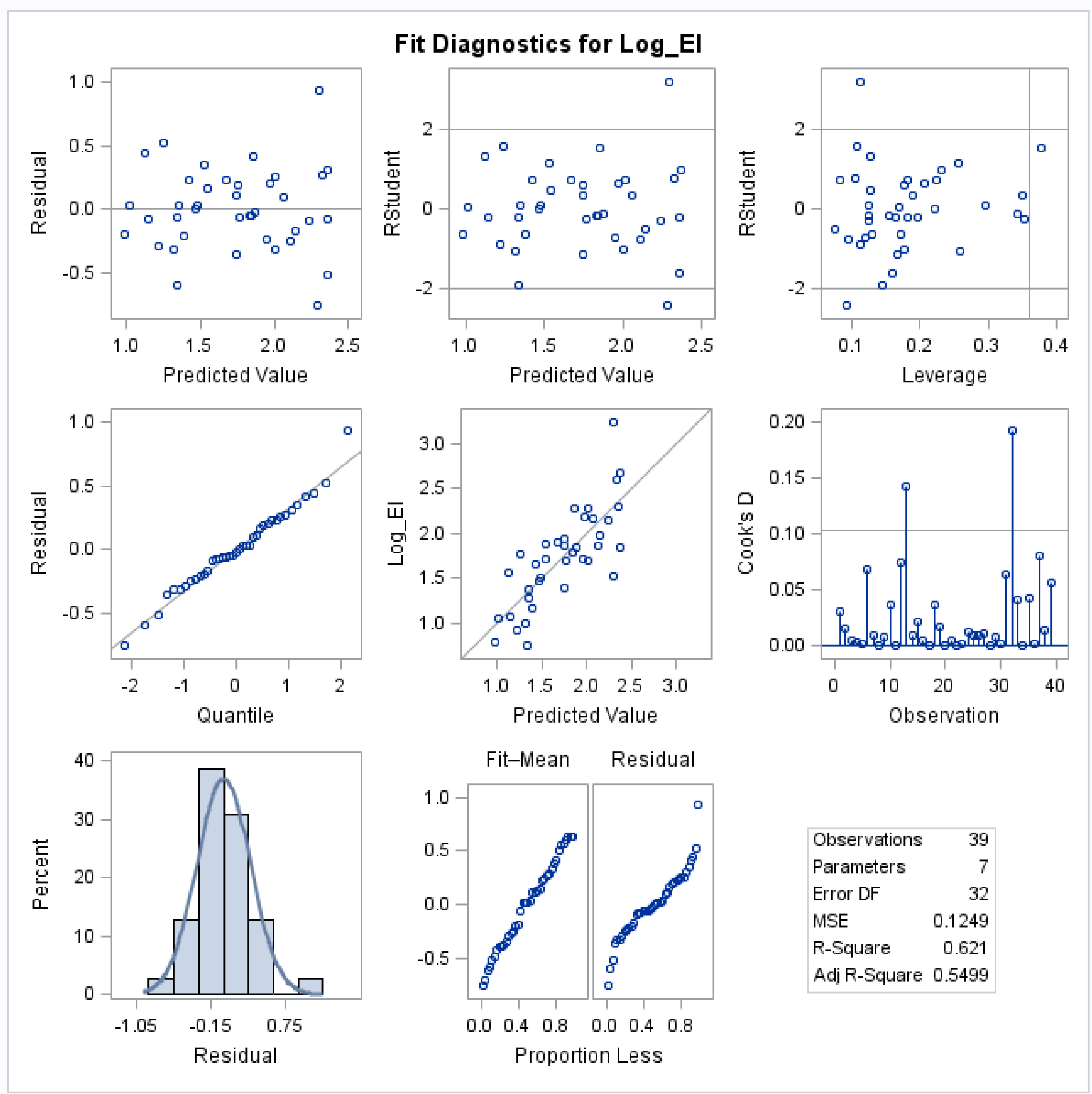


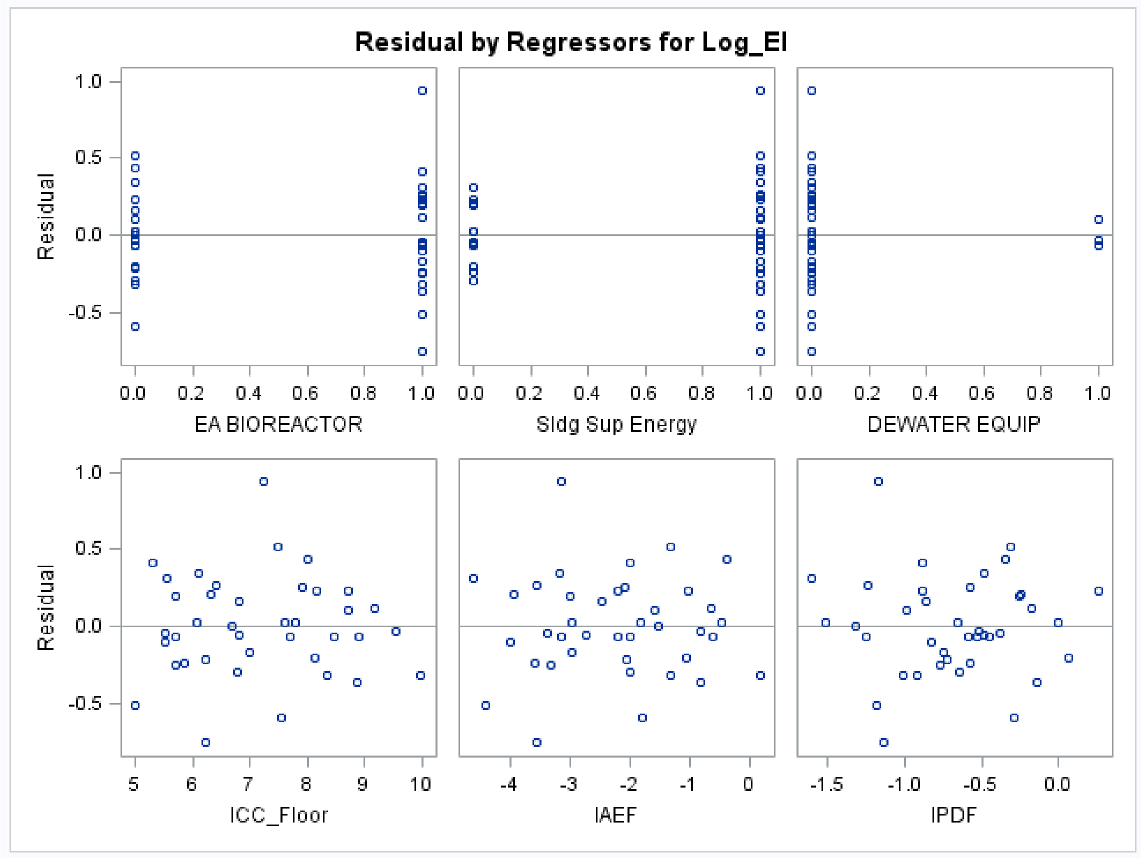
Residual by Regressors for Log_EI
Residual
1.0
0.5
0.0
-0.5
0.0
0.2
0.4
0.6
0.8
1.0
EA BIOREACTOR
OTHER BIOREACTOR
UV DISINFECTION
DEWATER EQUIP
-1.5
-1.0
-0.5
0.0
0.5
IPDF

22. Testing the NE Energy Intensity Model with 2014 Data

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	0.12239	0.75275	0.16	0.8719	0
EA_BIOREACTOR	0.52189	0.15255	3.42	0.0017	1.81552
Sldg_Sup_Energy	0.17945	0.13831	1.30	0.2037	1.32741
DEWATER_EQUIP	0.50448	0.24650	2.05	0.0490	1.34724
lCC_Floor	0.08254	0.08214	1.00	0.3225	3.55186
lAEF	-0.19051	0.10848	-1.76	0.0886	5.29402
lPDF	-0.24015	0.15744	-1.53	0.1370	1.43741

Root MSE	0.35340	**R-Square**	0.6210
Dependent Mean	1.72958	**Adj R-Sq**	0.5499
Coeff Var	20.43275		


Fit Diagnostics for Log_EI
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Log_EI
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 39
Parameters 7
Error DF 32
MSE 0.1249
R-Square 0.621
Adj R-Square 0.5499


Residual by Regressors for Log_EI
Residual
EA BIOREACTOR
Sldg Sup Energy
DEWATER EQUIP
ICC_Floor
IAEF
IPDF
1.0
0.5
0.0
-0.5
0.0 0.2 0.4 0.6 0.8 1.0
5 6 7 8 9 10
-4 -3 -2 -1 0
-1.5 -1.0 -0.5 0.0

Appendix F: Data Set Characteristics

Nebraska Data Set Characteristics

- Data set contained characteristics on 84 Nebraska plants.
- No solely fixed film plants included in the data
- One plant in this data set was not used in the Energy Intensity models due to missing data.
- Only 71 plants used for the electric intensity/usage models due to missing influent $CBOD_5$ data.
- Binary Variable Notation
    - Does the plant have this characteristic/equipment/process?
        - Yes = 1
        - No = 0
    - Abbreviations
        - IND LOAD: Industrial Loadings
        - OD: Oxidation Ditch
        - SBR: Sequencing Batch Reactor
        - EA: Extended Aeration
        - Other Plant: Other Plant Type
        - Fixed Film: Fixed Film Plant Type
        - Fine Diffs: Fine Bubble Diffusers (Main treatment basin)
        - Coarse Diffs: Coarse Bubble Diffusers (Main treatment basin)
        - Mech. Aerators: Mechanical Aerators (Main treatment basin)
        - UV: UV Disinfection
        - SE Sldg: Supplemental Energy Usage for Sludge Treatment (Aerobic Digestion or Heated Anaerobic Digestion)
        - VFDs: Variable Frequency Drives
        - Auto DO Control: Automatic Dissolved Oxygen Controls
        - DWE: Dewatering Equipment (Belt filter press, centrifuge, rotary screw press, and rotary drum centrifuge)
- Units for continuous variables
    - Flow/Design Flow: MGD
    - Climate Controlled Floor Area: ft^2
    - Influent and Effluent CBOD, TSS, and NH_3-N: mg/L
    - Energy Usage: kWh/year
    - Energy/Electric Intensity: MWh/MG

Table F.1: Continuous Variables for Nebraska Plants

Variable	n	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
Design Flow	84	0.409	0.198	0.025	0.050	1.000	3.000	0.519
Climate Controlled Floor Area	84	2928	1800	0	260	7200	21125	3682
Avg. Inf. CBOD	72	200.3	187.7	69.8	102.3	287.7	483.0	79.6
Avg. Flow	84	0.237	0.123	0.011	0.024	0.611	1.264	0.281
Avg. Eff. CBOD	84	4.83	3.83	1.55	2.50	7.23	24.71	3.32
Avg. Inf. TSS	77	229.5	211.2	54.8	105.0	355.0	628.5	113.4
Avg. Eff. TSS	84	8.17	7.40	2.67	4.50	12.75	45.30	5.16
Avg. Eff. NH_3-N	82	2.18	0.57	0.04	0.10	5.73	29.23	4.81
ANNUAL SUM HDDs	84	5725	5701	4166	5273	6297	6677	456
ANNUAL SUM CDDs	84	955	969	386	729	1215	1243	181
Percent Design Flow	84	0.64	0.55	0.22	0.33	1.03	1.90	0.33
Total 2015 Energy Usage	84	405243	223682	24283	63749	1093160	1664334	400788
Energy Intensity	84	6.09	5.51	1.69	2.51	9.56	20.92	3.33
Electric Intensity	84	5.86	5.30	1.69	2.51	9.52	20.92	3.20

Table F.2: Binary Variables (Yes = 1, No = 0) for Nebraska Plants

IND LOAD	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	61	AVG FLOW	0.179	0.102	0.011	0.023	0.322	1.264	0.263
		Energy Intensity	6.23	5.61	1.69	2.49	9.56	20.92	3.63
1	23	AVG FLOW	0.391	0.376	0.024	0.065	0.811	1.044	0.274
		Energy Intensity	5.71	5.42	1.82	3.13	7.29	12.44	2.35
OD	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	59	AVG FLOW	0.245	0.124	0.011	0.022	0.642	1.264	0.293
		Energy Intensity	6.45	5.82	1.69	2.65	10.79	20.92	3.57
1	25	AVG FLOW	0.219	0.122	0.025	0.044	0.457	1.012	0.254
		Energy Intensity	5.23	5.15	1.82	2.15	9.14	11.70	2.52
SBR	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	71	AVG FLOW	0.234	0.118	0.011	0.024	0.642	1.264	0.297
		Energy Intensity	6.11	5.47	1.69	2.65	9.56	20.92	3.43
1	13	AVG FLOW	0.254	0.181	0.066	0.112	0.541	0.611	0.172
		Energy Intensity	5.96	5.61	2.03	2.51	8.33	12.44	2.80
EA	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	49	AVG FLOW	0.312	0.177	0.025	0.061	0.811	1.264	0.299
		Energy Intensity	5.27	5.15	1.69	2.15	8.33	12.44	2.46
1	35	AVG FLOW	0.133	0.051	0.011	0.019	0.329	1.200	0.217
		Energy Intensity	7.23	6.72	2.33	2.89	11.66	20.92	4.02
Other Plant	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	79	AVG FLOW	0.222	0.121	0.011	0.024	0.587	1.200	0.260
		Energy Intensity	6.23	5.61	1.82	2.51	10.25	20.92	3.36
1	5	AVG FLOW	0.483	0.252	0.118	0.118	1.264	1.264	0.489
		Energy Intensity	3.81	3.66	1.69	1.69	5.82	5.82	1.73
FIXED FILM	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	78	AVG FLOW	0.203	0.118	0.011	0.024	0.457	1.264	0.256
		Energy Intensity	6.16	5.54	1.69	2.49	10.25	20.92	3.42
1	6	AVG FLOW	0.678	0.615	0.473	0.473	1.044	1.044	0.215
		Energy Intensity	5.20	5.23	2.87	2.87	7.77	7.77	1.72

Table F.2 (cont.)

FINE DIFFS	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	57	AVG FLOW	0.172	0.108	0.011	0.024	0.396	1.044	0.226
		Energy Intensity	6.07	5.47	1.82	2.33	10.25	20.92	3.39
1	27	AVG FLOW	0.375	0.286	0.018	0.019	0.811	1.264	0.335
		Energy Intensity	6.12	5.56	1.69	2.89	8.33	18.34	3.25
COARSE DIFFS	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	47	AVG FLOW	0.314	0.168	0.018	0.038	0.936	1.264	0.335
		Energy Intensity	5.77	5.29	1.69	2.24	9.14	18.34	3.05
1	37	AVG FLOW	0.139	0.103	0.011	0.023	0.340	0.655	0.145
		Energy Intensity	6.49	5.83	2.03	2.51	10.79	20.92	3.65
MECH AERATORS	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	63	AVG FLOW	0.245	0.136	0.011	0.023	0.611	1.264	0.285
		Energy Intensity	6.36	5.81	1.69	2.76	10.25	20.92	3.49
1	21	AVG FLOW	0.215	0.118	0.025	0.044	0.457	1.012	0.275
		Energy Intensity	5.28	5.15	1.82	2.15	9.14	11.70	2.71
UV	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	16	AVG FLOW	0.350	0.173	0.011	0.023	1.044	1.264	0.398
		Energy Intensity	5.45	5.06	1.69	2.15	9.14	12.44	2.77
1	68	AVG FLOW	0.211	0.120	0.017	0.024	0.541	1.200	0.242
		Energy Intensity	6.24	5.68	1.82	2.51	10.25	20.92	3.45
SE Sldg	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	24	AVG FLOW	0.179	0.087	0.017	0.023	0.376	1.264	0.261
		Energy Intensity	4.91	4.87	1.69	2.11	7.65	10.25	2.26
1	60	AVG FLOW	0.261	0.136	0.011	0.026	0.649	1.200	0.287
		Energy Intensity	6.56	5.97	2.03	2.82	11.23	20.92	3.58
VFDs	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	48	AVG FLOW	0.093	0.065	0.011	0.019	0.202	0.340	0.083
		Energy Intensity	6.36	6.11	2.03	2.49	10.25	18.34	3.20
1	36	AVG FLOW	0.430	0.370	0.028	0.118	1.012	1.264	0.333
		Energy Intensity	5.73	5.22	1.69	2.65	8.33	20.92	3.51

Table F.2 (cont.)

AUTO DO CONTROL	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	60	AVG FLOW	0.204	0.076	0.011	0.022	0.615	1.264	0.296
		Energy Intensity	6.28	5.82	1.69	2.41	10.52	20.92	3.70
1	24	AVG FLOW	0.320	0.249	0.112	0.118	0.611	0.936	0.225
		Energy Intensity	5.62	5.22	2.24	3.34	8.06	12.44	2.11
DWE	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	73	AVG FLOW	0.207	0.112	0.011	0.024	0.587	1.264	0.277
		Energy Intensity	6.04	5.42	1.69	2.49	9.56	20.92	3.46
1	11	AVG FLOW	0.437	0.457	0.118	0.177	0.611	0.936	0.228
		Energy Intensity	6.40	5.61	3.26	4.90	8.33	12.44	2.39

Pennsylvania Data Set Characteristics

- Data set contained characteristics on 72 Pennsylvania plants.
- No fixed film plants included in the data
- One plant in this data set was not used in the models due to missing data.
- Continuous Variables
  - Flow/Design Flow: MGD
  - Influent BOD: lbs-BOD/day
  - Effluent $CBOD_5$: mg/L
  - Electric Usage: MWh/year
  - Percent Design Capacity, based on flow
    - $\frac{\text{Average Flow, MGD}}{\text{Average Design Flow, MGD}}$
  - BOD Load Capacity
    - $\frac{\text{Average BOD Loading, lbs-BOD/day}}{\text{Average Design BOD Loading, lbs-BOD/day}}$
- Binary Variable Notation
  - Does the plant have this characteristic/equipment/process?
    - Yes = 1
    - No = 0
  - Abbreviations (Different from Nebraska data set)
    - Ext. Aeration: Extended Aeration
    - Aer Digest: Aerobic Digestion
    - Ana Digest: Anaerobic Digestion

Table F.3: Continuous Variables for Pennsylvania Plants

Variable	N	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
Design Flow	72	0.624	0.430	0.020	0.043	1.500	2.300	0.624
Ave Flow	72	0.340	0.255	0.007	0.020	0.824	1.378	0.337
Inf BOD	72	541.05	337.80	8.00	40.85	1351.00	2265.00	595.86
Eff CBOD	58	6.15	4.52	0.16	2.06	13.22	44.41	6.57
Elec Usage	72	571.0	402.9	27.4	88.6	1305.2	2789.6	525.2
Percent Design Capacity	72	0.53	0.52	0.21	0.33	0.73	1.15	0.17
BOD Load Capacity	72	0.46	0.45	0.15	0.26	0.73	0.92	0.19
Electric Intensity	72	8.58	5.96	1.12	2.47	17.16	46.05	7.46

Table F.4: Binary Variables (Yes=1, No=0) for Pennsylvania Plants

OD	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	65	Ave Flow	0.307	0.220	0.007	0.019	0.715	1.378	0.321
		Electric Intensity	9.032	6.419	1.121	2.469	17.439	46.046	7.702
1	7	Ave Flow	0.650	0.649	0.219	0.219	1.077	1.077	0.343
		Electric Intensity	4.389	4.530	2.176	2.176	6.641	6.641	1.505
SBR	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	53	Ave Flow	0.340	0.219	0.007	0.019	0.880	1.378	0.372
		Electric Intensity	9.037	5.734	1.121	2.902	17.439	46.046	8.022
1	19	Ave Flow	0.340	0.350	0.018	0.027	0.600	0.715	0.219
		Electric Intensity	7.307	6.136	1.188	1.771	13.749	24.985	5.569
Ext Aeration	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	46	Ave Flow	0.469	0.397	0.018	0.041	1.038	1.378	0.335
		Electric Intensity	5.933	4.611	1.188	2.430	11.597	24.985	4.403
1	26	Ave Flow	0.112	0.042	0.007	0.012	0.355	0.880	0.190
		Electric Intensity	13.264	11.758	1.121	3.789	22.496	46.046	9.340
Other	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	52	Ave Flow	0.268	0.170	0.007	0.018	0.649	1.077	0.288
		Electric Intensity	9.893	7.001	1.121	2.430	18.679	46.046	8.150
1	20	Ave Flow	0.529	0.471	0.021	0.044	1.150	1.378	0.387
		Electric Intensity	5.169	4.268	2.455	2.653	9.505	17.161	3.527
Aer Digest	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	11	Ave Flow	0.560	0.528	0.063	0.228	1.038	1.077	0.313
		Electric Intensity	4.128	3.605	1.771	2.455	5.539	7.413	1.585
1	59	Ave Flow	0.304	0.184	0.007	0.018	0.824	1.378	0.332
		Electric Intensity	9.569	6.641	1.121	2.469	18.679	46.046	7.879

Table F.4 (cont)

Ana Digest	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	68	Ave Flow	0.339	0.240	0.007	0.019	0.834	1.378	0.343
		Electric Intensity	8.825	6.166	1.121	2.469	17.439	46.046	7.589
1	2	Ave Flow	0.528	0.528	0.496	0.496	0.560	0.560	0.045
		Electric Intensity	4.934	4.934	2.455	2.455	7.413	7.413	3.506
Fine Diffs	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	43	Ave Flow	0.302	0.184	0.008	0.018	0.824	1.378	0.344
		Electric Intensity	9.593	6.136	1.121	3.306	18.679	46.046	8.652
1	28	Ave Flow	0.398	0.354	0.007	0.036	0.834	1.200	0.330
		Electric Intensity	7.214	5.867	1.771	2.430	14.909	19.457	5.011
Coarse Diffs	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	33	Ave Flow	0.449	0.445	0.018	0.049	0.834	1.100	0.319
		Electric Intensity	7.032	4.693	1.771	2.430	14.909	24.985	5.698
1	38	Ave Flow	0.246	0.056	0.007	0.013	0.715	1.378	0.332
		Electric Intensity	10.064	7.278	1.121	3.306	18.679	46.046	8.568
Mech Aerators	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	52	Ave Flow	0.310	0.172	0.007	0.019	0.775	1.378	0.347
		Electric Intensity	9.495	6.720	1.121	2.902	17.439	46.046	7.902
1	19	Ave Flow	0.422	0.378	0.018	0.021	1.038	1.077	0.312
		Electric Intensity	6.356	4.530	1.188	1.771	17.161	24.985	5.758
UV	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	48	Ave Flow	0.390	0.358	0.007	0.019	1.038	1.378	0.369
		Electric Intensity	7.192	4.836	1.121	2.354	16.438	24.985	5.447
1	23	Ave Flow	0.235	0.160	0.008	0.028	0.585	0.834	0.242
		Electric Intensity	11.707	6.858	2.837	3.881	22.496	46.046	10.015

Table F.4 (cont.)

DO controls	n	Variable	Mean	Median	Min	10th Pctl	90th Pctl	Max	Std Dev
0	58	Ave Flow	0.302	0.202	0.007	0.018	0.775	1.378	0.326
		Electric Intensity	9.051	6.584	1.121	2.430	18.679	46.046	7.879
1	14	Ave Flow	0.498	0.378	0.021	0.041	1.077	1.100	0.348
		Electric Intensity	6.633	4.611	2.469	2.837	17.161	17.439	5.129
Dewater Equip	**n**	**Variable**	**Mean**	**Median**	**Min**	**10th Pctl**	**90th Pctl**	**Max**	**Std Dev**
0	44	Ave Flow	0.211	0.065	0.007	0.018	0.649	0.880	0.252
		Electric Intensity	10.781	8.880	1.121	3.134	19.457	46.046	8.541
1	26	Ave Flow	0.570	0.524	0.049	0.184	1.100	1.378	0.354
		Electric Intensity	5.216	4.611	1.771	2.430	7.143	14.909	3.213

Appendix G: Data Collected

The following Tables G.2-G.13 contain data collected for each Nebraska plant in this study. These tables include only the potential predictor variables that were used in the final model creation. These tables do not contain all of the data originally collected. They contain all of the variables that were not filtered out due to redundant variables, poor quality data, small sample size, missing data, or subjective rankings. The list of predictor variables that were filtered out, as well as the reasons why they were left out of model creation, are listed in Table G.14. The final set of potential predictor variables listed in Tables G.2-G.13 contain 25 plant variables. The set consists of 14 binary variables (1 = the plant has this characteristic, 0 = the plant does not have the characteristic) and 11 continuous variables. Table G.1 lists all the final potential predictor variables and response variables, along with their abbreviations and units, if applicable.

Table G.1: Final Potential Predictor Variables

	Variable	**Abbreviation**	**Units**
Binary Variables	Conventional Extended Aeration	EA	-
	Oxidation Ditch	OD	-
	Sequencing Batch Reactor	SBR	-
	Fixed Film Plant	FF	-
	Other Plant Type	Other	-
	Supplemental Energy Usage For Sludge Treatment	SE Sldg	-
	Dewatering Equipment	DWE	-
	Fine Bubble Diffusers	Fine Diffs	-
	Coarse Bubble Diffusers	Coarse Diffs	-
	Mechanical Aerators	Mech. Aer.	-
	UV Disinfection	UV	-
	Industrial Loadings	Ind. Load	-
	Variable Frequency Drives	VFDs	-
	Automatic Dissolved Oxygen Controls	ADC	-
Continuous Variables	Average Flow	AF	MGD
	Average Design Flow	ADF	MGD
	Percent Design Flow	PDF	-
	Average Influent $CBOD_5$	AIC	mg/L
	Average Effluent $CBOD_5$	AEC	mg/L
	Average Influent TSS	AIT	mg/L
	Average Effluent TSS	AET	mg/L
	Average Effluent NH_3-N	AEN	mg/L
	Climate Controlled Floor Area	CCFA	ft^2
	Annual Sum of Heating Degree Days	HDDs	degree-days
	Annual Sum of Cooling Degree Days	CDDs	degree-days
Response Variables	Annual Energy Usage	-	kWh/year
	Annual Electric Usage	-	kWh/year
	Energy Intensity	-	MWh/MG
	Electric Intensity	-	MWh/MG

Table G.2: Ainsworth-Crete Plant Data

Facility	Plant Type					Sludge Treatment		Aeration Equipment			Disinfection	User Type
	EA	OD	SBR	FF	Other	SE Sldg	DWE	Fine Diff	Coarse Diff	Mech. Aer.	UV	Ind. Load
Ainsworth	0	0	1	0	0	1	0	1	0	0	1	0
Albion	0	1	0	0	0	1	0	0	0	1	1	0
Arnold	0	1	0	0	0	1	0	0	0	1	1	0
Ashland	0	1	0	0	0	1	0	0	1	1	1	0
Atkinson	0	0	1	0	0	1	0	0	1	0	1	1
Auburn	1	0	0	0	0	1	0	1	0	0	0	1
Aurora	0	1	0	0	0	1	1	0	0	1	0	1
Avoca	1	0	0	0	0	0	0	0	1	0	0	0
Bancroft	1	0	0	0	0	1	0	0	1	0	1	0
Bassett	0	1	0	0	0	1	0	0	0	1	1	0
Battle Creek	0	0	1	0	0	1	0	0	1	0	1	0
Beemer	1	0	0	0	0	1	0	0	1	0	1	0
Bennet	1	0	0	0	0	1	0	1	0	0	1	0
Blair	0	0	0	0	1	0	0	1	0	0	0	0
Brownville	1	0	0	0	0	1	0	0	1	0	0	0
Burwell	0	0	0	1	0	0	0	0	0	0	1	0
Cambridge	0	1	0	0	0	1	0	0	0	1	1	0
Cedar Rapids	1	0	0	0	0	1	0	0	1	0	1	0
Central City	0	0	1	0	0	1	1	1	0	0	1	0
Ceresco	1	0	0	0	0	1	0	1	0	0	1	0
Chadron	0	0	0	0	1	1	0	1	1	0	1	0
Coleridge	1	0	0	0	0	1	0	0	1	0	1	0
Cozad	0	0	1	0	0	1	1	1	0	0	1	1
Crawford	0	0	0	0	1	1	0	0	1	0	1	0
Crete	0	0	0	1	0	1	0	1	0	0	1	1

Table G3: Ainsworth-Crete (Continued)

Facility	Binary Var.		Continuous Variables								
	Other Equip.		Flow			Water Quality Data					Bldgs
	VFDs	ADC	AF	DF	PDF	AIC	AEC	AIT	AET	AEN	CCFA
Ainsworth	0	1	0.11	0.27	0.42	225	5.33	222.00	6.33	0.11	3,052
Albion	1	1	0.12	0.26	0.46	289	2.25	579.30	6.17	0.10	1,300
Arnold	0	0	0.03	0.06	0.42	172	5.58	112.15	7.58	7.15	300
Ashland	1	1	0.24	0.36	0.68	200	3.17	249.60	8.92	0.20	6,800
Atkinson	0	1	0.11	0.15	0.77		2.67		2.70	0.74	2,800
Auburn	1	1	0.29	0.74	0.39	203	2.72	628.50	5.17	0.33	4,200
Aurora	1	1	0.94	1.90	0.49	288	2.80	276.10	8.10	0.13	3,000
Avoca	0	0	0.02	0.03	0.91	141	5.25	241.40	5.25	0.10	550
Bancroft	0	0	0.08	0.06	1.36	164	3.92	285.75	7.25	0.54	600
Bassett	0	0	0.06	0.17	0.37	92	2.50	91.20	4.83	1.50	770
Battle Creek	0	0	0.14	0.20	0.68	127	3.41	162.00	5.64	4.52	900
Beemer	0	0	0.04	0.11	0.38	92	3.50	180.00	8.67	0.10	1,100
Bennet	0	0	0.06	0.15	0.42	277	2.83	405.50	5.33	1.98	2,200
Blair	1	0	1.26	2.00	0.63	402	12.06	216.20	11.54	29.23	10,000
Brownville	0	0	0.01	0.04	0.27	181	3.68	338.00	7.67	0.14	150
Burwell	0	0	0.08	0.25	0.32	317	20.08	314.00	15.83	22.16	225
Cambridge	0	0	0.10	0.20	0.52	190	3.60	147.75	5.80	0.24	3,100
Cedar Rapids	0	0	0.03	0.04	0.63	241	3.25	246.25	7.58		300
Central City	1	1	0.22	0.56	0.39	211	1.61	160.00	2.67	0.70	6,000
Ceresco	0	0	0.05	0.12	0.44	246	7.00	437.00	12.75	16.57	700
Chadron	1	0	0.66	0.64	1.03		5.49	237.50	13.24	1.21	2,000
Coleridge	0	0	0.04	0.08	0.49	138	6.42	168.00	10.92	5.73	300
Cozad	1	1	0.54	0.83	0.65	205	7.88	246.00	14.43	0.67	2,500
Crawford	1	1	0.12	0.18	0.71	99	7.66	105.00	6.67		1,800
Crete	1	0	0.64	0.98	0.66		6.50	210.10	9.33	4.55	3,000

Table G.4: Ainsworth-Crete (Continued)

Facility	Climate		Energy Usage			
	HDDs	CDDs	Annual Energy Usage	Annual Electric Usage	Energy Intensity	Electric Intensity
Ainsworth	5737	904	330,760	295,160	8.06	7.19
Albion	6167	760	303,200	303,200	7.07	7.07
Arnold	5752	689	87,260	87,260	9.56	9.56
Ashland	5674	993	461,376	461,376	5.19	5.19
Atkinson	6523	786	208,000	208,000	4.97	4.97
Auburn	5228	1177	503,400	503,400	4.82	4.82
Aurora	5502	1048	1,112,755	1,060,660	3.26	3.10
Avoca	5697	1087	63,749	63,749	7.65	7.65
Bancroft	6102	822	196,330	196,330	6.72	6.72
Bassett	4437	904	115,359	115,359	5.15	5.15
Battle Creek	6040	756	155,990	155,990	3.12	3.12
Beemer	6102	822	134,676	134,676	8.87	8.87
Bennet	5400	1243	156,600	156,600	6.83	6.83
Blair	4166	1215	781,560	781,560	1.69	1.69
Brownville	5228	1177	28,980	28,980	7.39	7.39
Burwell	6042	846	47,570	47,570	1.61	1.61
Cambridge	5276	1003	433,560	433,560	11.70	11.70
Cedar Rapids	5255	794	89,576	86,822	8.82	8.55
Central City	5502	1048	670,773	640,280	8.33	7.95
Ceresco	5352	1243	221,364	221,364	11.93	11.93
Chadron	6633	602	632,500	632,500	2.65	2.65
Coleridge	6476	735	80,204	80,204	5.81	5.81
Cozad	5855	897	1,439,383	1,225,200	7.29	6.20
Crawford	6423	682	165,000	165,000	3.66	3.66
Crete	5352	1243	1,150,040	921,500	4.91	3.93

Table G.5: Crofton-Minden Plant Data

Facility	Plant Type					Sludge Treatment		Aeration Equipment			Disinfection	User Type
	EA	OD	SBR	FF	Other	SE Sldg	DWE	Fine Diff	Coarse Diff	Mech. Aer.	UV	Ind. Load
Crofton	0	1	0	0	0	0	0	0	0	1	1	0
Dakota City	1	0	0	0	0	0	0	1	0	0	0	0
David City	0	0	1	0	0	0	0	1	0	0	0	1
Decatur	1	0	0	0	0	1	0	0	1	0	1	0
Dodge	1	0	0	0	0	1	0	0	1	0	1	0
Eagle	0	1	0	0	0	1	0	0	0	1	1	0
Edison	1	0	0	0	0	0	0	0	1	0	1	0
Elgin	0	1	0	0	0	1	0	0	0	1	1	0
Elkhorn	0	1	0	0	0	1	0	0	0	1	1	0
Fairbury	1	0	0	0	0	1	0	1	0	0	1	1
Falls City	1	0	0	0	0	1	0	1	0	0	1	0
Friend	1	0	0	0	0	1	0	0	1	0	1	0
Geneva	0	0	1	0	0	1	1	0	1	0	0	1
Genoa	1	0	0	0	0	0	0	1	0	0	0	0
Gothenburg	1	0	0	0	0	1	0	1	0	0	1	1
Greenwood	1	0	0	0	0	1	0	0	1	0	1	0
Hartington	1	0	0	0	0	1	0	0	1	0	1	1
Hebron	0	1	0	0	0	0	0	0	0	1	1	0
Hickman	0	1	0	0	0	1	0	1	0	0	1	0
Holdrege	0	0	0	1	0	1	1	1	0	0	1	1
Kimball	0	1	0	0	0	0	0	1	0	0	0	0
Laurel	0	1	0	0	0	0	0	0	0	1	1	0
Madison	0	1	0	0	0	0	0	0	0	1	1	1
Malcolm	1	0	0	0	0	1	0	0	1	0	1	0
McCook	0	0	0	1	0	1	0	1	0	0	1	1
Minden	0	0	1	0	0	0	0	1	0	0	0	0

Table G.6: Crofton-Minden (Continued)

Facility	Binary Var.		Continuous Variables								
	Other Equip.		Flow			Water Quality Data					Bldgs
	VFDs	ADC	AF	DF	PDF	AIC	AEC	AIT	AET	AEN	CCFA
Crofton	0	0	0.04	0.10	0.44	218	5.92	130.70	14.75	0.40	430
Dakota City	0	0	0.09	0.20	0.46	208	14.28	343.00	14.76	25.43	1,000
David City	1	1	0.40	0.94	0.43	483	15.80	382.25	45.30	8.92	3,600
Decatur	0	0	0.03	0.09	0.33	199	5.67	200.30	8.50	0.11	500
Dodge	1	0	0.04	0.14	0.26	277	9.57	194.25	9.75	0.58	1,400
Eagle	1	1	0.18	0.27	0.69	102	4.46	108.70	9.00	0.29	500
Edison	0	0	0.02	0.05	0.43	70	1.78	87.30	8.19	0.08	225
Elgin	0	0	0.06	0.12	0.53	361	2.92	336.70	10.50	0.10	830
Elkhorn	1	0	1.01	1.00	1.01	217	4.00	355.00	10.00	0.44	7,300
Fairbury	1	1	0.33	0.28	1.18	366	6.23	252.94	12.10	0.10	6,000
Falls City	1	0	1.20	0.63	1.90	167	2.94	256.00	6.00	1.19	9,600
Friend	0	0	0.10	0.13	0.82	170	5.16	111.00	7.99	0.27	400
Geneva	1	1	0.18	0.35	0.51	223	3.37	207.00	10.00	0.12	6,200
Genoa	0	0	0.07	0.11	0.61		3.83		8.17	1.67	900
Gothenburg	1	1	0.38	0.50	0.75	136	1.55	104.00	3.36	0.24	7,000
Greenwood	0	0	0.04	0.09	0.45	240	7.09		13.36	13.56	700
Hartington	0	0	0.13	0.20	0.67	185	4.17	224.00	6.25	0.66	1,500
Hebron	1	1	0.12	0.27	0.45	206	1.95	202.40	3.41	0.56	3,500
Hickman	0	1	0.25	0.82	0.31	170	8.00	150.50	8.08	3.19	800
Holdrege	1	1	0.51	0.84	0.61	174	3.58	204.40	4.23	0.74	7,200
Kimball	1	1	0.17	0.75	0.22	168	3.36	280.30	2.69	0.04	2,400
Laurel	0	0	0.11	0.19	0.57		3.75		5.67	2.64	300
Madison	1	0	0.38	0.33	1.16	178	4.85	109.60	3.02	1.04	3,400
Malcolm	1	0	0.03	0.10	0.28	185	3.00	188.30	7.58	0.20	600
McCook	1	1	0.81	1.75	0.46	165	4.32	239.00	4.30	0.26	12,000
Minden	0	0	0.20	0.40	0.51		3.69	156.00	5.96	1.28	3,420

Table G.7: Crofton-Minden (Continued)

Facility	Climate		Energy Usage			
	HDDs	CDDs	Annual Energy Usage	Annual Electric Usage	Energy Intensity	Electric Intensity
Crofton	6243	835	68,686	68,686	4.33	4.33
Dakota City	5935	889	175,818	175,818	5.29	5.29
David City	6052	855	909,172	909,172	6.21	6.21
Decatur	6102	822	46,436	46,436	4.26	4.26
Dodge	5690	932	282,460	282,460	20.92	20.92
Eagle	5697	1087	150,144	150,144	2.24	2.24
Edison	5273	1003	24,868	24,868	3.13	3.13
Elgin	6069	729	143,283	143,283	6.13	6.13
Elkhorn	5397	1215	1,284,501	1,134,528	3.48	3.07
Fairbury	5454	1092	864,040	864,040	7.19	7.19
Falls City	5228	1177	1,267,097	1,258,760	2.89	2.87
Friend	5677	969	152,000	152,000	4.04	4.04
Geneva	5099	1122	803,699	615,300	12.44	9.52
Genoa	5701	1008	144,520	144,520	6.09	6.09
Gothenburg	5733	897	661,600	661,600	4.77	4.77
Greenwood	5352	1243	20,480	20,480	1.39	1.39
Hartington	6297	863	512,161	492,120	10.79	10.37
Hebron	5174	1205	190,960	190,960	4.29	4.29
Hickman	5352	1243	309,400	309,400	3.34	3.34
Holdrege	5276	1003	1,041,240	1,041,240	5.56	5.56
Kimball	6677	386	286,320	286,320	4.67	4.67
Laurel	5745	917	83,360	83,360	2.11	2.11
Madison	5746	917	250,000	250,000	1.82	1.82
Malcolm	5352	1243	118,309	118,309	11.66	11.66
McCook	5412	1041	1,163,913	791,100	3.93	2.67
Minden	5794	985	356,543	356,543	4.84	4.84

Table G.8: Monroe-Superior Plant Data

Facility	Plant Type					Sludge Treatment		Aeration Equipment			Disinfection	User Type
	EA	OD	SBR	FF	Other	SE Sldg	DWE	Fine Diff	Coarse Diff	Mech. Aer.	UV	Ind. Load
Monroe	1	0	0	0	0	0	0	0	1	0	1	0
Nebraska City	0	0	0	1	0	0	1	0	0	0	0	1
Nebraska Crossing	1	0	0	0	0	1	0	1	0	0	1	0
Neligh	0	0	0	0	1	1	1	0	1	0	1	0
Newcastle	0	0	0	1	0	0	0	0	0	0	1	0
North Bend	1	0	0	0	0	1	0	0	1	0	1	0
Oakland	0	0	1	0	0	1	0	0	1	0	1	0
Odell	1	0	0	0	0	0	0	0	1	0	1	0
Ogallala	0	0	0	1	0	1	0	1	0	0	1	0
O'Neill	1	0	0	0	0	0	0	0	1	0	1	1
Pawnee City	0	1	0	0	0	1	0	0	1	0	1	1
Pender	1	0	0	0	0	1	0	0	1	0	1	1
Petersburg	1	0	0	0	0	0	0	0	1	0	1	0
Plattsmouth	0	0	0	1	0	1	0	0	0	0	0	0
Ponca	0	1	0	0	0	1	0	0	0	1	1	0
Randolph	0	1	0	0	0	0	0	0	0	1	0	0
Seward	0	0	0	1	0	1	0	0	0	0	1	0
Shubert	1	0	0	0	0	1	0	1	0	0	1	0
Sidney	0	0	1	0	0	1	1	1	0	0	0	0
Snyder	1	0	0	0	0	0	0	0	1	0	1	0
Spencer	1	0	0	0	0	0	0	0	1	0	1	0
Springfield	1	0	0	0	0	1	0	0	1	0	1	0
Stanton	0	0	1	0	0	1	0	0	1	0	1	0
Stella	1	0	0	0	0	1	0	1	0	0	1	0
Stuart	0	1	0	0	0	1	0	0	0	1	1	1
Superior	0	0	0	1	0	1	0	0	0	0	1	0

Table G.9: Monroe-Superior (Continued)

Facility	Binary Var.		Continuous Variables								
	Other Equip.		Flow			Water Quality Data					Bldgs
	VFDs	ADC	AF	DF	PDF	AIC	AEC	AIT	AET	AEN	CCFA
Monroe	0	0	0.04	0.05	0.74	86	4.08	54.80	5.92	0.12	400
Nebraska City	1	0	1.24	1.88	0.66	202	18.60	1496.20	16.10	1.05	20,000
Nebraska Crossing	0	0	0.02	0.08	0.25	243	5.75	282.50	5.25	5.87	1,200
Neligh	1	1	0.12	0.20	0.61	404	2.27	274.30	4.53	0.34	2,100
Newcastle	0	0	0.03	0.08	0.40	263	6.42	346.50	6.25	0.28	200
North Bend	0	0	0.32	0.18	1.75	81	6.00	83.00	12.17	0.10	2,000
Oakland	0	0	0.18	0.18	1.01	203	5.08	302.20	7.00	0.25	2,400
Odell	0	0	0.07	0.05	1.44	155	2.67	142.60	5.67	0.35	350
Ogallala	1	0	0.59	1.20	0.49	175	4.07	226.90	7.55	0.73	5,100
O'Neill	0	0	0.34	0.60	0.57		6.65		12.08	9.89	11,000
Pawnee City	1	1	0.15	0.11	1.32		7.00	341.30	8.67	0.14	3,400
Pender	0	0	0.37	0.17	2.22		3.00	151.80	12.79	5.74	1,600
Petersburg	0	0	0.01	0.02	0.29		3.75		8.50	0.94	300
Plattsmouth	1	0	0.62	1.50	0.41	106	11.65	190.10	7.98	10.36	6,000
Ponca	0	0	0.06	0.17	0.33	150	4.00	191.00	11.33	1.72	2,300
Randolph	0	0	0.08	0.16	0.51	138	3.50	134.30	6.33	0.10	350
Seward	1	0	0.67	1.01	0.67	332	14.40	549.90	19.38	0.95	11,200
Shubert	0	0	0.02	0.04	0.48	204	3.75	190.25	15.25	1.48	100
Sidney	1	1	0.61	1.16	0.53	155	3.97	208.10	5.46	0.57	3,200
Snyder	0	0	0.07	0.07	1.00		2.92		9.42	0.10	300
Spencer	0	0	0.03	0.05	0.62		3.08		4.75	2.02	250
Springfield	0	0	0.14	0.22	0.62	183	3.83	211.20	4.58	0.10	2,700
Stanton	0	0	0.07	0.18	0.37	239	3.31	246.50	5.06	3.41	1,900
Stella	0	0	0.02	0.04	0.42	150	3.83	127.70	9.83	0.13	250
Stuart	0	0	0.04	0.07	0.56	208	5.75	141.75	9.17	1.14	450
Superior	1	0	0.16	0.39	0.41	301	10.55	859.75	7.25	2.53	7,000

Table G.10: Monroe-Superior (Continued)

Facility	Climate		Energy Usage			
	HDDs	CDDs	Annual Energy Usage	Annual Electric Usage	Energy Intensity	Electric Intensity
Monroe	5701	1008	99,682	71,399	7.42	5.32
Nebraska City	5228	1177	2,393,823	1,611,630	5.28	3.56
Nebraska Crossing	5631	954	123,867	111,678	18.34	16.54
Neligh	6296	677	226,000	226,000	5.25	5.25
Newcastle	5980	880	14,296	14,296	1.22	1.22
North Bend	5701	980	324,600	324,600	2.76	2.76
Oakland	6102	822	162,100	162,100	2.51	2.51
Odell	5359	1092	61,352	61,352	2.33	2.33
Ogallala	5703	835	1,664,334	763,300	7.77	3.56
O'Neill	6466	786	672,858	645,480	5.42	5.20
Pawnee City	5482	1050	347,200	347,200	6.56	6.56
Pender	6102	822	164,799	158,880	1.23	1.19
Petersburg	6069	729	55,718	47,895	22.12	19.02
Plattsmouth	5314	1215	522,593	405,100	2.33	1.80
Ponca	5935	889	157,280	157,280	7.83	7.83
Randolph	6296	677	273,503	273,503	9.14	9.14
Seward	5400	1243	820,572	820,572	3.34	3.34
Shubert	5482	1050	24,283	24,283	3.50	3.50
Sidney	6171	701	1,251,284	1,247,280	5.61	5.59
Snyder	6105	822	127,298	127,298	5.37	5.37
Spencer	5806	1037	71,200	71,200	6.29	6.29
Springfield	5631	954	428,544	428,544	8.63	8.63
Stanton	6063	766	172,300	172,300	7.15	7.15
Stella	5482	1050	53,754	44,304	8.14	6.71
Stuart	6604	658	96,998	96,998	6.99	6.99
Superior	5474	1051	461,945	176,680	8.01	3.06

Table G.11: Syracuse-York Plant Data

Facility	Plant Type					Sludge Treatment		Aeration Equipment			Disinfection	User Type
	EA	OD	SBR	FF	Other	SE Sldg	DWE	Fine Diff	Coarse Diff	Mech. Aer.	UV	Ind. Load
Syracuse	1	0	0	0	0	1	0	0	1	0	1	0
Table Rock	1	0	0	0	0	0	0	0	1	0	1	1
Valentine	0	0	0	0	1	1	0	1	0	0	1	0
Waco	1	0	0	0	0	1	0	0	1	0	1	0
Wahoo	0	1	0	0	0	1	1	0	0	1	1	1
Waterloo	0	1	0	0	0	1	0	0	0	1	1	0
Wausa	0	1	0	0	0	0	0	0	0	1	0	1
Waverly	0	1	0	0	0	1	1	0	1	0	1	1
Wayne	0	0	0	1	0	1	1	0	1	0	1	1
Weeping Water	1	0	0	0	0	1	0	0	1	0	1	0
West Point	0	0	1	0	0	0	1	1	0	0	1	1
Western	1	0	0	0	0	0	0	0	1	0	1	0
Wilber	0	1	0	0	0	0	0	0	0	1	1	0
Wisner	0	0	1	0	0	1	0	0	1	0	1	0
Wood River	0	0	0	1	0	0	0	0	0	0	0	1
Wymore	0	1	0	0	0	0	0	0	0	1	0	0
York	0	0	0	1	0	1	0	0	0	0	0	1

Table G.12: Syracuse-York (Continued)

Facility	Binary Var.		Continuous Variables								
	Other Equip.		Flow			Water Quality Data					Bldgs
	VFDs	ADC	AF	DF	PDF	AIC	AEC	AIT	AET	AEN	CCFA
Syracuse	0	0	0.19	0.33	0.57	236	6.77	413.30	7.76	1.67	200
Table Rock	0	0	0.02	0.03	0.77	70	2.25	69.30	6.50	0.10	250
Valentine	1	0	0.25	0.37	0.69	171	3.25	171.00	5.00	0.58	1,800
Waco	0	0	0.02	0.05	0.48	177	3.28	136.00	5.13	0.89	-
Wahoo	1	0	0.46	0.75	0.61	253	2.67	218.50	5.75	0.10	14,000
Waterloo	1	0	0.12	0.25	0.48	237	6.50	250.75	6.75	0.20	2,100
Wausa	0	0	0.07	0.07	0.93	118	5.42	99.30	10.38	0.90	400
Waverly	1	1	0.40	1.00	0.40	215	3.36	283.25	9.17	0.10	11,500
Wayne	1	0	0.47	1.47	0.32	241	3.67	258.70	15.70	1.48	1,800
Weeping Water	0	0	0.31	0.20	1.56		4.00		4.58	0.10	1,000
West Point	1	1	0.36	0.65	0.56	388	3.92	583.25	6.50	4.16	5,800
Western	0	0	0.02	0.05	0.33	199	1.61	262.70	6.42	0.04	260
Wilber	1	0	0.14	0.25	0.55	177	3.50	171.60	4.50	0.21	4,800
Wisner	0	0	0.18	0.23	0.80	158	7.23	169.50	8.92	6.47	1,900
Wood River	0	0	0.12	0.20	0.58	270	7.00	157.50	6.75	1.35	1,900
Wymore	0	0	0.17	0.26	0.64	202	2.58	344.00	5.92	4.99	880
York	1	0	1.04	3.00	0.35		24.71	310.40	13.00	1.14	21,125

Table G.13: Syracuse-York (Continued)

Facility	Climate		Energy Usage			
	HDDs	**CDDs**	**Annual Energy Usage**	**Annual Electric Usage**	**Energy Intensity**	**Electric Intensity**
Syracuse	5611	1087	498,960	498,960	7.27	7.27
Table Rock	5482	1050	51,040	45,977	5.83	5.25
Valentine	6297	800	535,040	535,040	5.82	5.82
Waco	5700	979	80,212	80,212	9.16	9.16
Wahoo	5314	1215	1,053,395	1,035,200	6.32	6.21
Waterloo	5348	1079	241,440	241,440	5.47	5.47
Wausa	6243	835	74,343	74,343	3.13	3.13
Waverly	5352	1243	763,704	732,300	5.28	5.07
Wayne	5825	917	1,064,400	1,064,400	6.17	6.17
Weeping Water	5314	1215	282,360	282,360	2.49	2.49
West Point	6102	822	651,073	616,704	4.90	4.64
Western	5677	969	62,458	62,458	10.25	10.25
Wilber	5352	1243	172,960	172,960	3.46	3.46
Wisner	6102	822	134,280	134,280	2.03	2.03
Wood River	5497	980	23,741	23,741	0.56	0.56
Wymore	5359	1092	130,200	130,200	2.15	2.15
York	5764	979	1,093,160	1,093,160	2.87	2.87

Table G.14: Variables dropped from Model Creation

Dropped Variables	**Reason for dropping**
Population	Was highly correlated with Ave Design Flow.
Per Capita Flow, gal/cap-d	The compnents of per-capita flow are already included as variables (ave flow and population).
Per Capita CBOD Loading, lbs-CBOD/cap-d	The compnents of per-capita CBOD loading are already included as variables (ave flow, ave inf. CBOD, and population). Also, influent CBOD has poor data quality due to the frequency of sampling (once per year).
AWIN Score	Kept showing up as significant, but with wrong sign (model would show higher the AWIN Score, the more efficient)
Total Industrial Flow, MGD	Poor data quality. Recorded from NPDES permit applications filled out by operators. Many operators estimate total industrial flow.
Total Industrial Loading, lbs-CBOD	Poor data quality. Recorded from NPDES permit applications filled out by operators. Many operators estimate total industrial loading or don't record it at all.
Continuous Discharge?	All but maybe 1 or 2 of the plants discharge continuously.
Discharges per year	See note above for continuous discharge
Number of Buildings	Highly correlated with number of buildings heated during the winter. Consolidated into climate controlled floor area.
Total Floor Area, sq ft	Highly correlated with climate controlled floor area because most plants heat all of their buildings during the winter. Consolidated into climate controlled floor area.
Nutrient Removal?	Not characterized correctly. Plants that had effluent NH3N limits considered to have nutrient removal.
Disinfection?	Most plants (89%) have disinfection.
Chemical Disinfection?	Most plants with disinfection (89%) have UV disinfection. Only 9 use chemical disinfection.
Composting Sludge?	Only one plant composts sludge.

Table G.14: Variables dropped from Model Creation (Continued)

Dropped Variables	**Reason for dropping**
Hauling Sludge?	This was supposed to be for plants that haul sludge to another plant for treatment, but was confused by interns to mean haul away for land application. In addition, energy usage for hauling (diesel or gasoline) was not recorded.
Land Application of Sludge?	Majority of plants land apply sludge.
Lagoon disposal of sludge?	Hard to characterize. Some plants with lagoons were marked down as land apply because they eventually clean the lagoons out.
Number of Lift Stations	Not related to onsite energy usage. Was recorded for future use/assessments.
Motor Needs Replacing Soon	Marked down for future assessments. Hard to categorize. Subjective depending on who visited the plant and what the operator believes needs replacing.
Efficiency of Motors ever Tested?	Only 5 of the 94 plants have ever tested the efficiency of their motors.
Belt Filter Press?	Only 9 of the 94 plants have a belt filter press.
Centrifuge?	Only 1 plant has a centrifuge.
Drying Beds?	Not recorded consistently. Only 8 of 94 indicated as having drying beds.
Reed Beds?	Not recorded consistently. Only 1 of 94 indicated as having reed beds.
Rotary Drum Centrifuge?	Only 1 plant has a rotary drum centrifuge.
Rotary Screw Press?	Only 1 plant has a rotary screw press.
Self-cleaning UV lamps?	Only 7 of 94 have self-cleaning UV lamps. Not significant energy user.
Number of Buildings heated in the winter	Highly correlated with number of buildings because most plants heat all of their buildings during the winter. Consolidated into climate controlled floor area.
Number of buildings cooled in the summer	Usually only the lab/one room is air conditioned at a majority of these plants. Not a significant energy user.

Table G.14: Variables dropped from Model Creation (Continued)

Number of Operators	Inconsistent recording. Number of operators was not always clear. Did not determine how many certified operators they have, only the number of people that work at the plant.
Staff Changes in Recent Years?	Inconsistent recording.
Average Influent TSS, mg/L	Poor data quality. Influent samples only taken once a year.
Lowest Ammonia Discharge Limit, mg/L	Too many missing values. 22 of the 94 plants don't have NH3N limits.